PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 124, Number 3, March 1996

RANDOMLY SAMPLED RIEMANN SUMS AND COMPLETE CONVERGENCE IN THE LAW OF LARGE NUMBERS FOR A CASE WITHOUT IDENTICAL DISTRIBUTION

ALEXANDER R. PRUSS

(Communicated by Rick Durrett)

Abstract. Let the points $\{x_{nk}\}$ be independently and uniformly randomly chosen in the intervals $\left[\frac{k-1}{n}, \frac{k}{n}\right]$, where $k = 1, 2, ..., n$. We show that for a finite-valued measurable function f on $[0,1]$, the randomly sampled Riemann sums $\frac{1}{n}\sum_{k=1}^{n} f(x_{nk})$ converge almost surely to a finite number as $n \to \infty$ if and only if $f \in L^2[0,1]$, in which case the limit must agree with the Lebesgue integral. One direction of the proof uses Bikelis' (1966) non-uniform estimate of the rate of convergence in the central limit theorem. We also generalize the notion of sums of i.i.d. random variables, subsuming the randomly sampled Riemann sums above, and we show that a result of Hsu, Robbins and Erdős (1947, 1949) on complete convergence in the law of large numbers continues to hold. In the Appendix, we note that a theorem due to Baum and Katz (1965) on the rate of convergence in the law of large numbers also generalizes to our case.

1. Introduction

We fix a set I with a probability measure μ, and a sequence of partitions $\mathcal{P}_n$ of I into μ-measurable subsets $\{I_{n1}, ..., I_{nn}\}$ with $\mu(I_{nk}) = 1/n$. We suppose that we have a large underlying probability space (Ω, P) and that for each fixed n we have a collection of independent random variables $\{x_{n1}, ..., x_{nn}\}$ with x_{nk} being μ-uniformly distributed over I_{nk}, i.e., $P(x_{nk} \in U) = \mu(U)/\mu(I_{nk}) = n\mu(U)$ for $U \subseteq I_{nk}$. Throughout, f shall denote a finite-valued μ-measurable function on I.

Theorem 1. *If $f \in L^2(\mu)$, then, with probability one,*

$$\frac{1}{n}\sum_{k=1}^{n} f(x_{nk}) \to \int_I f\,d\mu \quad \text{as} \quad n \to \infty.$$

Conversely, if we additionally assume that **all** *the $\{x_{nk}\}$ are independent and f is a μ-measurable function not in $L^2(\mu)$, then, with probability one, $\frac{1}{n}\sum_{k=1}^{n} f(x_{nk})$ is unbounded and in particular divergent.*

Note. In the first part of the theorem there are no assumptions on the correlations, if any, between the $\{x_{nk}\}_{k=1}^{n}$ and the $\{x_{n'k'}\}_{k'=1}^{n'}$ for $n \neq n'$; in the converse part

Received by the editors May 9, 1994 and, in revised form, September 15, 1994.
1991 *Mathematics Subject Classification.* Primary 60F15, 26A42; Secondary 60F10.
Key words and phrases. Riemann sums, complete convergence, Lebesgue integral, law of large numbers, central limit theorem.

they are assumed to be independent. We refer to $\frac{1}{n}\sum_{k=1}^{n} f(x_{nk})$ as a "randomly sampled Riemann sum".

To get a more vivid view of our result, take μ to be Lebesgue measure on $I = [0, 1]$ and set $I_{nk} = \left[\frac{k-1}{n}, \frac{k}{n}\right]$. Then the theorem describes when the Riemann sums with sampling points x_{nk} chosen randomly with uniform distribution on each interval I_{nk} of the regular partition $\mathcal{P}_n$ converge almost surely to the Lebesgue integral as $n \to \infty$. It should be noted that in this case it is not difficult to prove the theorem in an easy elementary way for non-negative monotone functions f (see [10, Appendix B]), but the proof does not seem to generalize to other functions.

Recently, C. S. Kahane [6] proved that if $f \in L^p(\mu)$ for $p > 2$, then the requisite almost sure convergence follows. However the question of what happens for $p \leq 2$ was left open. As Kieffer and Stanojević [7] note, it is not very difficult to see that it suffices to have $f \in L^1(\mu)$ in order to obtain L^1-convergence of the randomly sampled Riemann sums to the Lebesgue integral. One proof proceeds first by showing that if $f \in L^2(\mu)$, then in fact we have L^2-convergence (this is a consequence of [6, inequality (2)]), and then by approximating a function $f \in L^1(\mu)$ by bounded, hence L^2, functions. (See [10] for details.) Thus, the interesting question is the almost sure convergence. Kieffer and Stanojević [7] did prove a positive result about almost sure convergence for $f \in L^1$, but they were working with partitions $\mathcal{P}_n$ such that $\mathcal{P}_{n+1}$ is a refinement of $\mathcal{P}_n$ and with carefully chosen correlations between the $\{x_{nk}\}$ and the $\{x_{n+1,k'}\}$, and thus our work, like that of Kahane, is independent of theirs.

Now, observe that for any measurable function G we have

$$(1.1)\qquad \sum_{k=1}^{n} E[G(f(x_{nk}))] = \sum_{k=1}^{n} n \int_{I_{nk}} G(f(t))\, d\mu(t) = n \int_{I} G(f(t))\, d\mu(t),$$

provided these expressions make sense. This simple identity encodes the central property of the $\{f(x_{nk})\}_{k=1}^{n}$ and prompts us to make the following generalization:

Definition. Let $\xi_1, ..., \xi_n$ be random variables, and let Ξ be a random variable possibly defined on a different probability space. Then, $\xi_1, ..., \xi_n$ are said to be a **regular cover** of (the distribution of) Ξ provided we have

$$(1.2)\qquad E[G(\Xi)] = \frac{1}{n}\sum_{k=1}^{n} E[G(\xi_k)],$$

for any measurable function G for which both sides make sense. If $\xi_1, ..., \xi_n$ are in addition independent, then we say they form an **independent regular cover** of Ξ.

From (1.1) we see that $f(x_{n1}), ..., f(x_{nn})$ form a regular cover of f, where f is considered a random variable on the probability space (I, μ). Note that condition (1.2) is equivalent to saying that the average of the distribution (respectively, characteristic) functions of the $\{\xi_k\}_{k=1}^{n}$ is equal to the distribution (respectively, characteristic) function of Ξ. It is also to be noted that if $\xi_1, ..., \xi_n$ are identically distributed, then clearly they are a regular cover of $\Xi = \xi_1$. The main part of this paper will then be devoted to the proof of the following result.

Theorem 2. *Let $\xi_{n1}, ..., \xi_{nn}$ be an independent regular cover of a random variable Ξ. Let $S_n = \xi_{n1} + ... + \xi_{nn}$. Let ε be a positive number, and let A_n be a sequence of real numbers. Suppose that*

$$E[\Xi^2] < \infty \qquad \text{and} \qquad \limsup_{n\to\infty} |E[\Xi] - A_n| < \varepsilon. \tag{1.3}$$

Then

$$\sum_{n=1}^{\infty} P(|S_n - nA_n| \geq \varepsilon n) < \infty. \tag{1.4}$$

Conversely, if (1.4) *holds, then*

$$E[\Xi^2] < \infty \qquad \text{and} \qquad \limsup_{n\to\infty} |E[\Xi] - A_n| \leq \varepsilon. \tag{1.5}$$

Remarks. Hsu and Robbins [5] showed that if the ξ_{nk} are i.i.d., then (1.3) implies (1.4) (more precisely, they proved the corresponding implication in the corollary below, from which the implication in Theorem 2 can easily be made to follow), and Erdős [4] proved all of Theorem 2, still under the assumption that the ξ_{nk} are i.i.d. Erdős' elementary proof of the i.i.d. analog of condition (1.3) implying (1.4) can be adapted to our more general case using a symmetrization argument; the methods of Duncan and Szynal [3] can also be used for this purpose. However, in this paper we choose to use a simple method based on a theorem of A. Bikelis [2]. Some alternate proofs of Theorem 2 may be found in [10]. Note also that by the covering identity (1.2), we have $E[S_n] = \sum_{k=1}^{n} E[\xi_{nk}] = nE[\Xi]$. We shall refer to (1.3) or (1.5) as the "second moment condition".

With general A_n, we can easily find examples, even with the ξ_{nk} i.i.d., where (1.4) holds while $\limsup |E[\Xi] - A_n| = \varepsilon$ so that (1.3) does not hold. On the other hand, if A_n is constant, the author cautiously conjectures that (1.4) implies (1.3). This conjecture is equivalent to saying that whenever $E[\Xi^2] < \infty$ then $\sum_{n=1}^{\infty} P(S_n \geq nE[\Xi]) = \infty$. With the ξ_{nk} i.i.d., the conjecture is true, as shown by Erdős [4] who notes that the finite second moment and the identical distribution allow one to apply the central limit theorem to conclude that $\lim_{n\to\infty} P(S_n \geq nE[\Xi]) \geq 1/2$ (with equality if and only if the variables are not almost surely constant). However, the simple example of randomly sampled Riemann sums for $f(x) = \operatorname{sgn}(x - \frac{1}{2})$ on $I = [0, 1]$ shows that the central limit theorem is not available in our more general case.

In any case, letting $A_n \equiv L$ be constant in Theorem 2 immediately yields the following result which characterizes complete convergence in the law of large numbers for the case of regular covering.

Corollary. *If Ξ and S_n are as in Theorem 2, then the following two statements are equivalent for any real number L:*

(a) $\sum_{n=1}^{\infty} P(|S_n/n - L| \geq \varepsilon) < \infty$ *for every* $\varepsilon > 0$.

(b) $E[\Xi^2] < \infty$ *and* $L = E[\Xi]$.

In the terminology of Hsu and Robbins [5], condition (a) of the corollary is asserting the complete convergence of S_n/n to L.

Assuming Theorem 2, and letting $\xi_{nk} = f(x_{nk})$ and $\Xi = f$, we obtain Theorem 1. For, suppose $f \in L^2$. Then, condition (b) of the corollary is satisfied when we put $L = \int_I f\,d\mu$, and so, for any $\varepsilon > 0$, condition (a) implies that $\sum_{n=1}^\infty P(|S_n/n - L| \geq \varepsilon)$ converges, so that by Borel-Cantelli, $S_n/n = \frac{1}{n}\sum_{k=1}^n f(x_{nk})$ must converge almost surely to L, as desired. Conversely, suppose $f \notin L^2$ and *all* the $\{x_{nk}\}$ are independent. Then, condition (1.5) in Theorem 2 must fail for each $\varepsilon > 0$ and for every real sequence A_n, so that $\sum_{n=1}^\infty P(|S_n/n| \geq \varepsilon)$ must diverge. Thus, since the $\{S_n\}_{n=1}^\infty$ are independent, the Borel-Cantelli lemma implies that, with probability one, $|S_n/n| \geq \varepsilon$ for infinitely many n, and so $\frac{1}{n}\sum_{k=1}^n f(x_{nk})$ must be unbounded with probability one, as desired.

2. The case of $E[|\Xi|^p] < \infty$ for $p > 2$

C. S. Kahane [6] proceeds to prove the convergence of the randomly-sampled Riemann sums for $f \in L^p(\mu)$ where $p = 2 + \delta > 2$ by showing that

$$P\left(\left|\frac{1}{n}\sum_{k=1}^n f(x_{nk}) - \int_I f d\mu\right| \geq \varepsilon\right) \leq C_\varepsilon n^{-(1+\delta/2)} \|f\|_{2+\delta}^{2+\delta}, \tag{2.1}$$

for some constant C_ε depending only on ε. The almost sure convergence then follows by the Borel-Cantelli lemma. We would like to prove our result for a general square integrable f by truncation, but, unfortunately, (2.1) is too weak to help there. In our work we shall suppose $0 < \delta \leq 1$ and use A. Bikelis' [2] non-uniform estimate for convergence in the central limit theorem instead of (2.1).

We now state the relevant result of Bikelis. Let $\xi_1, ..., \xi_n$ be independent random variables. Define $B_{pn} = \sum_{k=1}^n E[|\xi_k|^p]$ and $\Delta_n^2 = \sum_{k=1}^n \mathrm{Var}[\xi_k]$. Let $S_n = \sum_{k=1}^n \xi_k$, and let $F_n = P(S_n \leq y)$ be its distribution function. Let $G_n = P(S_n < y)$. We denote the distribution function of a normal random variable of mean zero and variance one by Φ. Using characteristic function arguments, A. Bikelis proved the following result.

Proposition 1 (A. Bikelis [2]). *There is an absolute constant $K < \infty$ such that if $0 < \delta \leq 1$ and each of $\xi_1, ..., \xi_n$ is assumed to have a finite moment of order $2 + \delta$, then*

$$|F_n(x + E[S_n]) - \Phi(x/\Delta_n)| \leq K \frac{B_{2+\delta,n}}{\Delta_n^{2+\delta} + |x|^{2+\delta}}.$$

Remark. Taking left limits, we see that we may also put G_n in place of F_n in the above inequality.

The Proposition and the above remark almost automatically give us the desired improvement over (2.1), besides supplying us with a concise proof of Kahane's result [6] in our apparently more general regular covering case. For, suppose that $\xi_1, ..., \xi_n$ form an independent regular cover of Ξ, and that $E[|\Xi|^p] < \infty$ for some $p = 2 + \delta > 2$. Then, by the covering identity (1.2), for any $r > 0$ we have $B_{rn} = nE[|\Xi|^r]$. Also, $E[S_n] = nE[\Xi]$. Thus, under the assumptions that $\Delta_n \neq 0$ and $x > 0$, we have by Proposition 1 and the remark immediately following it,

$$P(|S_n - nE[\Xi]| \geq x) \leq 2(1 - \Phi(x/\Delta_n)) + 2K \frac{nE[|\Xi|^{2+\delta}]}{x^{2+\delta}}, \tag{2.2}$$

where we have used the symmetry of the normal distribution Φ. Now, noting that

$$\Delta_n = \left(\sum_{k=1}^{n}(E[\xi_{nk}^2] - (E[\xi_{nk}])^2)\right)^{1/2} \leq \left(\sum_{k=1}^{n} E[\xi_{nk}^2]\right)^{1/2} = B_{2,n}^{1/2} = (nE[\Xi^2])^{1/2},$$

letting $x = \varepsilon n$, and supposing that Γ is a finite positive constant such that

$$\sqrt{E[\Xi^2]} \leq \Gamma \tag{2.3a}$$

we obtain from (2.2),

$$P(|S_n - nE[\Xi]| \geq \varepsilon n) \leq Ce^{-\sqrt{n}\cdot\varepsilon/\Gamma} + 2Kn^{-(1+\delta)}\varepsilon^{-(2+\delta)}E[|\Xi|^{2+\delta}], \tag{2.3b}$$

where C is an absolute constant chosen so that $2(1-\Phi(y)) \leq Ce^{-y}$ for $y \geq 0$. Note that (2.3b) was derived under the assumption that $\Delta_n \neq 0$, but if $\Delta_n = 0$, then in fact the left-hand side of (2.3b) vanishes, and the inequality trivially continues to hold. Thus, (2.3b) is the desired replacement for (2.1).

3. Some lemmas

Write $y^+ = \max(y, 0)$ and $y^- = \max(-y, 0)$. We recall the following elementary result.

Lemma 1. *Let T be any finite number and fix $\varepsilon > 0$. Then if Ξ is a random variable, we have $E[(\Xi^+)^2] < \infty$ if and only if $\sum_{n=1}^{\infty} nP(\Xi > \varepsilon n + T) < \infty$.*

The following lemma will be used to show that we may truncate Ξ.

Lemma 2. *In the setup of Theorem 2, $\sum_{n=1}^{\infty}\sum_{k=1}^{n} P(|\xi_{nk}| \geq n)$ is finite if and only if $E[\Xi^2] < \infty$.*

Proof. By the covering identity (1.2) applied to $G(y) = 1_{\{|y|\geq n\}}$, we have

$$\sum_{n=1}^{\infty}\sum_{k=1}^{n} P(|\xi_{nk}| \geq n) = \sum_{n=1}^{\infty} nP(|\Xi| \geq n).$$

By the preceding lemma, this is finite if and only if $E[\Xi^2]$ is finite. □

The following modification of a standard lemma is proved just as in the standard case $t = 0$.

Lemma 3. *Suppose a_n is a sequence of non-increasing non-negative numbers. Then $\sum_{n=1}^{\infty} n^t a_n$ converges if and only if $\sum_{m=0}^{\infty} 2^{m(t+1)} a_{2^m}$ converges.*

As a corollary, we obtain the following somewhat curious criterion for a random variable to have a finite second moment.

Lemma 4. *Let Ξ be a random variable. Let ε and α be any positive constants. Suppose τ_n is a non-decreasing sequence of real numbers such that $P(\Xi \geq \tau_n) \geq \alpha/n$. Then, $E[(\Xi^+)^2] < \infty$ if and only if $\sum_{n=1}^{\infty} nP(\Xi > \varepsilon n + \tau_n) < \infty$.*

Proof. If $E[(\Xi^+)^2] < \infty$, then since $\tau_n \geq \tau_1$, we have

$$\sum_{n=1}^{\infty} nP(\Xi > \varepsilon n + \tau_n) \leq \sum_{n=1}^{\infty} nP(\Xi > \varepsilon n + \tau_1) < \infty$$

by Lemma 1. Conversely, suppose that $\sum_{n=1}^{\infty} nP(\Xi > \varepsilon n + \tau_n)$ converges. Since τ_n is non-decreasing, we may apply Lemma 3 to conclude that

$$\sum_{m=0}^{\infty} 2^{2m} P(\Xi > 2^m \varepsilon + \tau_{2^m}) < \infty. \tag{3.1}$$

I claim that for m sufficiently large we have $\tau_{2^{m+1}} \leq 2^m\varepsilon + \tau_{2^m}$. For, if m is such that $\tau_{2^{m+1}} > 2^m\varepsilon + \tau_{2^m}$, then $2^{2m}P(\Xi > 2^m\varepsilon + \tau_{2^m}) \geq 2^{2m}P(\Xi \geq \tau_{2^{m+1}}) \geq 2^{2m}(\alpha/2^{m+1}) = 2^{m-1}\alpha$, which cannot happen for infinitely many m if (3.1) is to hold. Thus, choose a natural M such that if $m \geq M$, then $\tau_{2^{m+1}} \leq 2^m\varepsilon + \tau_{2^m}$. Then, iterating this inequality for $m \geq M$ we find that $\tau_{2^m} < 2^m\varepsilon + \tau_{2^M}$. Adding $2^m\varepsilon$ to both sides, we find that whenever $m \geq M$ then $2^{m+1}\varepsilon + \tau_{2^M} > 2^m\varepsilon + \tau_{2^m}$, and, hence, $P(\Xi > 2^{m+1}\varepsilon + \tau_{2^M}) \leq P(\Xi > 2^m\varepsilon + \tau_{2^m})$. By (3.1) we now conclude that $\sum_{m=0}^{\infty} 2^{2m}P(\Xi > 2^{m+1}\varepsilon + \tau_{2^M})$ must converge, and so by another application of Lemma 3 we find that $\sum_{n=1}^{\infty} nP(\Xi > 2n\varepsilon + \tau_{2^M})$ converges. Since τ_{2^M} is a finite constant, Lemma 1 now shows that $E[(\Xi^+)^2] < \infty$. □

4. Sufficiency of the second moment condition in Theorem 2

We shall show that (1.3) implies (1.4) in Theorem 2. Rescaling if necessary, assume that $\varepsilon = 1$. Now, suppose that $E[\Xi^2] < \infty$ and $\limsup_{n\to\infty} |A_n - E[\Xi]| < 1$. We truncate by setting $\Xi'_n = \Xi \cdot 1_{\{|\Xi|<n\}}$. Let $\xi'_{nk} = \xi_{nk} \cdot 1_{\{|\xi_{nk}|<n\}}$ and put $S'_n = \sum_{k=1}^{n} \xi'_{nk}$. We easily see that $\xi'_{n1}, ..., \xi'_{nn}$ form a regular cover of Ξ'_n. Let $\Gamma = \sqrt{E[\Xi^2]}$. Then, clearly, $\sqrt{E[(\Xi'_n)^2]} \leq \Gamma$, and so (2.3a) is satisfied with Ξ'_n in place of Ξ. We shall apply (2.3b) with $\delta = 1$ to the primed variables. (Note that no information would be gained by using another $\delta \in (0, 1)$.)

Now, clearly

$$P(S_n \neq S'_n) \leq P\left(\bigcup_{k=1}^{n} \{|\xi_{nk}| \geq n\}\right) \leq \sum_{k=1}^{n} P(|\xi_{nk}| \geq n).$$

But by Lemma 2, the last term is summable over n. Then,

$$P(|S_n - nA_n| \geq n) \leq P(|S'_n - nA_n| \geq n) + P(S_n \neq S'_n),$$

and so it suffices for us to show that $P(|S'_n - nA_n| \geq n)$ is summable. Note also that since $E[|\Xi - \Xi'_n|] \to 0$, and $\limsup |E[\Xi] - A_n| < 1$, then for some $\gamma > 0$ and n sufficiently large, we will have $|E[\Xi'_n] - A_n| \leq 1 - \gamma$. For such n, we then have

$$\begin{aligned} P(|S'_n - nA_n| \geq n) &\leq P(|S'_n - nE[\Xi'_n]| + n(1-\gamma) \geq n) \\ &= P(|S'_n - nE[\Xi'_n]| \geq \gamma n). \end{aligned}$$

Now, apply (2.3b) with $\Gamma = \sqrt{E[\Xi^2]}$ and $\delta = 1$, to conclude that for such sufficiently large n we have

$$P(|S'_n - nA_n| \geq n) \leq Ce^{-\sqrt{n}\cdot\gamma/\Gamma} + 2Kn^{-2}\gamma^{-3}E[|\Xi'_n|^3].$$

We have seen that it suffices for us to show that the left-hand side of this is summable, and clearly $\sum_{n=1}^{\infty} \exp -(\sqrt{n}\cdot\gamma/\Gamma) < \infty$, so that it suffices to show that

$$\sum_{n=1}^{\infty} n^{-2}E[|\Xi'_n|^3] < \infty. \tag{4.1}$$

But, using Fubini's theorem,

$$\sum_{n=1}^{\infty} n^{-2} E[|\Xi_n'|^3] = \sum_{n=1}^{\infty} E[n^{-2} \cdot |\Xi|^3 \cdot 1_{\{|\Xi|<n\}}] = E\left[\sum_{n=\lfloor|\Xi|+1\rfloor}^{\infty} n^{-2}|\Xi|^3\right]$$
$$\leq E\left[2\lfloor|\Xi|+1\rfloor^{-1}|\Xi|^3\right] \leq 2E[\Xi^2] < \infty,$$

where we have used the elementary estimate that $\sum_{n=m}^{\infty} n^{-2} \leq 2/m$. Thus, (4.1) holds and the proof that condition (1.3) implies (1.4) is complete. □

5. Necessity of the second moment condition in Theorem 2

We shall show that (1.4) implies (1.5) in Theorem 2. (A longer but more intuitive proof based on an association inequality is given in [10, §5.2]; however, it does not work for Theorem 3 in the Appendix, below.) Assume that (1.4) holds. Without loss of generality put $\varepsilon = 1$. We thus have $\sum_{n=1}^{\infty} P(|S_n - nA_n| \geq n) < \infty$. We plan to apply Lemma 4 to Ξ in order to conclude that $E[(\Xi^+)^2] < \infty$.

Let μ_n be a median of S_n. Let X^s denote the symmetrization of a random variable X. Inspired by [8], note that

$$P(S_n - \mu_n \geq 2n) \leq P(|S_n - \mu_n| \geq 2n) \leq 2P(|S_n^s| \geq 2n) \leq 4P(|S_n - nA_n| \geq n),$$

by standard symmetrization inequalities [9, §17.1A]. Thus we have

$$\sum_{n=1}^{\infty} P(S_n - \mu_n \geq 2n) < \infty. \tag{5.1}$$

Since the choice of medians in all the preceding work was arbitrary, define μ_n to be the smallest median of S_n, i.e., let $\mu_n = \inf\{\mu : P(S_n \leq \mu) \geq 1/2\}$.

Now, set $\alpha = 1/4$ and $\beta = 1/3$, and note that $(1-\alpha)(1-\beta) = 1/2$. For each n, choose a finite τ_n so that we have $\tau_n \geq \tau_{n-1}$ for $n > 1$, as well as

$$P(\Xi \geq \tau_n) \geq \alpha/n \qquad \text{and} \qquad P(\Xi \leq \tau_n) \geq 1 - \alpha/n. \tag{5.2}$$

One may, e.g., let $\tau_n = \sup\{\tau : P(\Xi \geq \tau) \geq \alpha/n\}$. Note that

$$\begin{aligned} P(\xi_{nk} > \tau_n) &\leq P(\xi_{n1} > \tau_n) + ... + P(\xi_{nn} > \tau_n) \\ &= nP(\Xi > \tau_n) = n(1 - P(\Xi \leq \tau_n)) \leq \alpha, \end{aligned} \tag{5.3}$$

where we have used the covering identity (1.2) as well as (5.2). Furthermore, for every $k \in \{1, ..., n\}$, choose a finite number ρ_{nk} such that

$$P(S_n - \xi_{nk} \geq \rho_{nk}) \geq \beta \qquad \text{and} \qquad P(S_n - \xi_{nk} \leq \rho_{nk}) \geq 1 - \beta. \tag{5.4}$$

One may, e.g., let $\rho_{nk} = \sup\{\rho : P(S_n - \xi_{nk} \geq \rho) \geq \beta\}$.

Then,

$$\begin{aligned} P(S_n \leq \tau_n + \rho_{nk}) &\geq P(\xi_{nk} \leq \tau_n \text{ and } S_n - \xi_{nk} \leq \rho_{nk}) \\ &= P(\xi_{nk} \leq \tau_n)P(S_n - \xi_{nk} \leq \rho_{nk}) \\ &= (1 - P(\xi_{nk} > \tau_n))P(S_n - \xi_{nk} \leq \rho_{nk}) \\ &\geq (1-\alpha)(1-\beta) = \frac{1}{2}, \end{aligned}$$

where we have used the independence of $S_n - \xi_{nk}$ and ξ_{nk}, as well as (5.3) and (5.4). Thus, we see that either $\tau_n + \rho_{nk}$ is a median of S_n or else it is greater than all the medians of S_n, and in either case $\tau_n + \rho_{nk} \geq \mu_n$ since we have chosen μ_n to be the smallest median of S_n.

Let $T_{nk} = \{\xi_{nk} > 2n + \tau_n\}$ and $R_{nk} = \{S_n - \xi_{nk} \geq \rho_{nk}\}$. We may now apply the argument from [4] as follows:

$$
\begin{aligned}
(5.5)\qquad P(S_n \geq 2n + \mu_n) &\geq P(S_n > 2n + \tau_n + \rho_{nk}) \geq P\left[\bigcup_{k=1}^{n} (T_{nk} \cap R_{nk})\right] \\
&= \sum_{k=1}^{n} P[(T_{n1} \cap R_{n1})^c \cap ... \cap (T_{n,k-1} \cap R_{n,k-1})^c \cap (T_{nk} \cap R_{nk})] \\
&\geq \sum_{k=1}^{n} P[T_{n1}{}^c \cap ... \cap T_{n,k-1}{}^c \cap T_{nk} \cap R_{nk}] \\
&\geq \sum_{k=1}^{n} [P(T_{nk} \cap R_{nk}) - P((T_{n1} \cup ... \cup T_{n,k-1}) \cap T_{nk})] \\
&= \sum_{k=1}^{n} [P(T_{nk})P(R_{nk}) - P(T_{n1} \cup ... \cup T_{n,k-1})P(T_{nk})] \\
&\geq \sum_{k=1}^{n} P(T_{nk}) \left(P(R_{nk}) - \sum_{k=1}^{n} P(T_{nk}) \right),
\end{aligned}
$$

where we have used the independence of T_{nk} and R_{nk}, as well as the independence of the $\{T_{nk}\}_{k=1}^n$. But, $P(R_{nk}) \geq \beta$ by (5.4). Moreover,

$$
\begin{aligned}
\sum_{k=1}^{n} P(T_{nk}) = \sum_{k=1}^{n} P(\xi_{nk} > 2n + \tau_n) &= nP(\Xi > 2n + \tau_n) \\
&= n(1 - P(\Xi \leq 2n + \tau_n)) \leq \alpha,
\end{aligned}
$$

where we have used the covering identity (1.2) as well as (5.2). Thus, (5.5) implies that

$$
P(S_n \geq 2n + \mu_n) \geq (\beta - \alpha) \sum_{k=1}^{n} P(T_{nk}) = \frac{1}{12} nP(\Xi > 2n + \tau_n).
$$

Now, applying (5.1), we conclude that

$$
\sum_{n=1}^{\infty} nP(\Xi_n > 2n + \tau_n)
$$

converges, so that $E[(\Xi^+)^2] < \infty$ by Lemma 4, as τ_n is non-decreasing and satisfies (5.2). In exactly the same way we see that $E[(\Xi^-)^2] < \infty$ and so Ξ must have a finite second moment.

Now, it remains to show that $\limsup |E[\Xi] - A_n| \leq 1$. To obtain a contradiction, suppose that $\limsup |E[\Xi] - A_n| > 1$. Then, there is a $\gamma > 0$ such that $|E[\Xi] - A_n| \geq 1 + \gamma$ for infinitely many n. Now, for any such n, we either have $E[\Xi] - A_n \geq 1 + \gamma$

or $E[\Xi] - A_n \le -1-\gamma$. Assume that we have $E[\Xi] - A_n \ge 1+\gamma$ for infinitely many n, since the proof in the other case is analogous. For any such n we have $A_n + 1 \le E[\Xi] - \gamma$ so that

$$P(|S_n/n - A_n| \ge 1) \ge P(S_n/n \ge A_n + 1) \ge P(S_n/n \ge E[\Xi] - \gamma). \tag{5.4}$$

But, we have already showed that (1.3) implies (1.4) in Theorem 2 from which by Borel-Cantelli it follows that $S_n/n \to E[\Xi]$ almost surely, and hence also in probability. Thus, the right-hand side of (5.4) must converge to unity as $n \to \infty$. Since, by our assumption, the inequalities in (5.4) hold for infinitely many n, the left-hand side of (5.4) cannot be summable, and so (1.4) cannot occur, a contradiction. Thus, indeed we must have $\limsup |E[\Xi] - A_n| < 1$ as desired, and as we have seen that $E[\Xi^2] < \infty$, it follows that (1.5) is satisfied. □

Appendix
Rate of convergence in probability

A proof of the convergence in L^1 of the randomly sampled Riemann sums for $f \in L^1$ has been outlined in the Introduction. The same proof can be made to show that if Ξ has a finite expectation and $\xi_{n1}, ..., \xi_{nn}$ form an independent regular cover of Ξ, then $S_n/n \to E[\Xi]$ in L^1 and hence also in probability, where $S_n = \sum_{k=1}^n \xi_{nk}$. If $p = 1$, the following theorem can be interpreted as giving an estimate on the rate of this convergence in probability.

Theorem 3. *Let $\xi_{n1}, ..., \xi_{nn}$ be an independent regular cover of a random variable Ξ and set $S_n = \xi_{n1} + ... + \xi_{nn}$. Let $0 < p < 2$. Then, for any real number L the following two statements are equivalent:*
(a) $\sum_{n=1}^\infty n^{-1} P(|S_n - nL| \ge \varepsilon n^{1/p}) < \infty$ *for every* $\varepsilon > 0$.
(b) $E[|\Xi|^p] < \infty$ *and either* $p < 1$ *or* $L = E[\Xi]$.

Remarks. The theorem is due to Baum and Katz [1] in the i.i.d. case. Note also that condition (b) need not imply condition (a) if $p \ge 2$, even in the i.i.d. case. For, if the ξ_{nk} are i.i.d. and have a finite second moment, then, as long as they are not almost surely constant, the central limit theorem implies that $(S_n - nE[\Xi])/n^{1/2}$ converges in distribution to a normal random variable, and thus we see that, in such a case, (a) cannot hold for any $p \ge 2$.

Outline of proof of Theorem 3. The proof in [1] that (b) implies (a) easily adapts to the more general regular covering case. On the other hand, assuming (a), we may prove that $E[|\Xi|^p] < \infty$ by easily adapting the proofs in §5 of the present paper. We now wish to prove that if (a) holds and $p \ge 1$, then $L = E[\Xi]$. In §5 we perform the analogous proof by noting that since (1.3) was already proved to imply (1.4), then we must have $S_n/n \to E[\Xi]$ in probability. However, the corresponding implication in Theorem 3 does not seem to immediately give convergence in probability and while we could carefully use the general conditions for the weak law of large numbers, we choose to bypass the convergence in probability by the following argument. By assumption, we have $\sum_{n=1}^\infty n^{-1} P(|S_n - nL| \ge \varepsilon n^{1/p})$ converging for every $\varepsilon > 0$. Furthermore, since we have already argued that (b) implies (a), we must also have $\sum_{n=1}^\infty n^{-1} P(|S_n - nE[\Xi]| \ge \varepsilon n^{1/p}) < \infty$. To obtain a contradiction, suppose that $E[\Xi] \ne L$. Fix $0 < \varepsilon < |E[\Xi] - L|/2$. Then, using the triangle inequality and the fact that $n^{1/p} \le n$ for $p \ge 1$, we see that whenever we have $|S_n - nE[\Xi]| < \varepsilon n^{1/p}$ then we must also have $|S_n - nL| \geq \varepsilon n^{1/p}$. Hence, $P(|S_n - nL| \geq \varepsilon n^{1/p}) \geq P(|S_n - nE[\Xi]| < \varepsilon n^{1/p}) = 1 - P(|S_n - nE[\Xi]| \geq \varepsilon n^{1/p})$. Thus,

$$\sum_{n=1}^{\infty} n^{-1} P(|S_n - nL| \geq \varepsilon n^{1/p}) \geq \sum_{n=1}^{\infty} [n^{-1} - n^{-1} P(|S_n - nE[\Xi]| \geq \varepsilon n^{1/p})].$$

Since $n^{-1}P(|S_n - nE[\Xi]| \geq \varepsilon n^{1/p})$ is summable, it follows that the right-hand side of the above inequality must be infinite, and hence we have a contradiction since we have assumed the left side to be finite. Hence, L must equal $E[\Xi]$ if $p \geq 1$. □

Acknowledgments

The author would like to thank Professors J. J. F. Fournier and J. B. Walsh for stimulating discussions and suggestions. The author is grateful to Professor M. T. Barlow for suggesting the term "regular cover". Finally, the author would like to express his gratitude to the referee for a careful reading and helpful suggestions.

Note added in proof

After this paper was accepted for publication, the author was kindly informed by Professor Allan Gut that the implications (a) ⇒ (b) in our Corollary and in Theorem 3 are special cases of his Theorems 2.1 and 5.1 [A. Gut, *Complete convergence for arrays*, Periodica Math. Hungarica **25** (1992), 51–75]. Since the first implication in our Theorem 1 is a direct consequence of the implication (a) ⇒ (b) in the Corollary, it follows that the first implication in our Theorem 1 is also a direct consequence of the work of Gut. Moreover Professor Gut has pointed out to the author that for i.i.d. random variables, the case $p = 1$ of Theorem 3 is due to F. L. Spitzer [*A combinatorial lemma and its application*, Trans. Amer. Math. Soc. **82** (1956), 323–339].

References

1. L. E. Baum and M. Katz, *Convergence rates in the law of large numbers*, Trans. Amer. Math. Soc. **120** (1965), 108–123. MR **33:**6679
2. A. Bikelis [= A. Bikyalis], *On estimates of the remainder term in the central limit theorem*, Litovskiĭ Mat. Sb. **6** (1966), 323–346. (Russian) MR **35:**1067
3. Richard Duncan and Dominik Szynal, *A note on the weak and Hsu-Robbins law of large numbers*, Bull. Polish Acad. Sci. Math. **32** (1984), 729–735. MR **86h:**60036
4. Paul Erdős, *On a theorem of Hsu and Robbins*, J. Math. Statist. **20** (1949), 286–291. MR **11:**40f
5. P. L. Hsu and H. Robbins, *Complete convergence and the law of large numbers*, Proc. Nat. Acad. Sci. U.S.A. **33** (1947), 25–31. MR **8:**470e
6. C. S. Kahane, *Evaluating Lebesgue integrals as limits of Riemann sums*, Math. Japonica **38** (1993), 1073–1076. MR **95b:**28002
7. J. C. Kieffer and Č. V. Stanojević, *The Lebesgue integral as the almost sure limit of random Riemann sums*, Proc. Amer. Math. Soc. **85** (1982), 389–392. MR **83h:**26015
8. O. I. Klesov, *Convergence of series of probabilities of large deviations of sums of independent identically distributed random variables*, Ukrain. Mat. Zh. **45** (1993), 770–784. (Russian) CMP 95:03

9. Michel Loève, *Probability theory*, Van Nostrand, New York, 1955. MR **16:**598f
10. A. R. Pruss, *Several proofs of a result concerning randomly sampled Riemann sums and complete convergence in the law of large numbers for a case without identical distribution*, preprint. ftp `math.ubc.ca:/pub/pruss/RandSumsBig.tex` (1995).

Department of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z2
E-mail address: `pruss@math.ubc.ca`