The Annals of Probability
1982, Vol. 10, No. 2, 374–381

ON THE RATE OF CONVERGENCE IN THE WEAK LAW OF LARGE NUMBERS

By Peter Hall

Australian National University

Let $\{Z_n\}$ be a sequence of random variables converging in probability to zero. If the convergence is also in L^2, it is common to measure the rate of convergence by the L^2 norm of Z_n. However, in many interesting cases the variables Z_n do not have finite variance, and then it seems appropriate to study the truncated L^2 norm, $\Delta_n = E[\min(1, Z_n^2)]$. We put Δ_n forward as a global measure of the rate of convergence. The paper concentrates on the case where Z_n is a normalised sum of independent and identically distributed random variables, and we derive very precise descriptions of the rate of convergence in this situation.

1. Introduction. Let $\{Z_n, n \geq 1\}$ be a sequence of random variables converging in distribution to the variable Z. When Z has a continuous distribution it is natural to measure the rate of convergence of Z_n to Z in the uniform norm, but when Z is discontinuous this approach is not meaningful. In the case $Z \equiv 0$ the rate of convergence may be described by the large deviations $P(|Z_n| > \varepsilon)$, for each $\varepsilon > 0$. However, these do not provide a global account of the rate of convergence, and there seems to be a need for an alternative approach.

In the theory of mathematical statistics it is common to measure the distance of an estimator, $\hat{\theta}_n$, from the true parameter value, θ, in the L^2 norm. If $Z_n = \hat{\theta}_n - \theta$ then a description of the rate of convergence is given by the behaviour of $E(Z_n^2)$. However, in many cases of more general interest the variable Z_n does not have finite variance, and so we should study instead the *truncated* L^2 norm of Z_n. That is, we direct attention to $\Delta_n \equiv E[\min(1, Z_n^2)] = E[A(Z_n)]$, where $A(x) = \min(1, x^2)$. Note that $Z_n \to_p 0$ if and only if $\Delta_n \to 0$, and that for any $\varepsilon > 0$, $\Delta_n \geq \min(1, \varepsilon^2)P(|Z_n| > \varepsilon)$. Therefore the probabilities of large deviations are bounded in a very simple way by the quantity Δ_n, and at the same time Δ_n gives information about the behaviour of Z_n near the origin. This paper is devoted to studying the asymptotic properties of Δ_n when $Z_n = (\sum_1^n X_i - a_n)/b_n$ is a normalised sum of independent and identically distributed random variables.

Suppose $b_n \to \infty$, and for some sequence $\{a_n\}$, $(\sum_1^n X_i - a_n)/b_n \to_p 0$. We shall determine those sequences $\{a_n\}$ which give the fastest rate of convergence to zero in the sense of the measure Δ_n. There are three convenient choices for the "best" sequence: (i) the median, $a_n = \text{med}(\sum_1^n X_i)$; (ii) the sum of truncated means, $a_n = nE[X_1 I(|X_1| \leq b_n)]$, where $I(E)$ denotes the indicator function of the event E; and (iii) the median of the truncated sum, $a_n = \text{med}(\sum_1^n X_i I(|X_i| \leq b_n))$. We shall show that for any given sequence $\{b_n\}$ the rates of convergence with any of these choices are the same, and that they cannot be beaten by any other choice of $\{a_n\}$.

To avoid trivialities we shall assume that the common distribution of the summands is not concentrated at a single point; apart from this, our results are completely general. In the special case where the mean is finite it is usual to set $a_n = nE(X_1)$, and we shall investigate the influence of this choice of a_n on the rate of convergence. In general, centring at the mean rather than the sum of truncated means gives a slightly inferior rate. We shall give examples in which the rate calculated for the optimal a_n is asymptotically negligible in comparison with that calculated for $a_n = nE(X_1)$.

Received August 1980; revised April 1981.

Key words and phrases. Independent and identically distributed random variables, rate of convergence, weak law of large numbers.

AMS 1980 *subject classification.* Primary 60F05, 60G50; secondary 60F25.

Our approach is to construct a sequence of positive constants $\{\delta_n\}$ which depend on the distribution of X_1 in a very simple way, and which have the property that

$$0 < \liminf_{n\to\infty} \Delta_n/\delta_n \le \limsup_{n\to\infty} \Delta_n/\delta_n < \infty.$$

We write this as $\Delta_n \asymp \delta_n$. Thus, all the necessary information about the asymptotic behaviour of Δ_n is contained in the simpler sequence $\{\delta_n\}$. This makes it very easy to obtain characterizations of the rate of convergence; see Corollaries 1, 2 and 3.

Let $X_1, X_2, \cdots$ be independent and identically distributed symmetric random variables whose tails are regularly varying with exponent α, $0 < \alpha < 2$. Let $S_n = \sum_1^n X_i$ and $b_n = n^r$. Then $Z_n = S_n/b_n \to_p 0$ if $r > 1/\alpha$, and in this case for any $\varepsilon > 0$ there exist constants C_1 and C_2 such that $P(|S_n/b_n| > \varepsilon) \le C_1\Delta_n \le C_2 P(|S_n/b_n| > \varepsilon)$. The second inequality follows from the facts that for large n, $P(|S_n/b_n| > \varepsilon) > C_3 nP(|X| > b_n)$, with use of the techniques of Erdös (1949, page 290), and $\Delta_n \le C_4 nP(|X| > b_n)$, with use of our Theorem 1 together with the results of Feller (1971, page 281). Therefore in this case, $E[A(Z_n)]$ has the same rate of convergence as probabilities of large deviations.

In stating and proving our results we shall use the symbol C, with or without subscripts, to denote positive generic constants which depend on the underlying distribution of the summands but do not depend on n or other parameters.

2. The results. Let $X, X_1, X_2, \cdots$ be independent and identically distributed random variables with common distribution function F, and set $S_n = \sum_1^n X_i$. For any two norming constants a and b ($\neq 0$), define

$$\Delta(a, b) = E[\min\{1, ((S_n - a)/b)^2\}].$$

Let $\{b_n, n \ge 1\}$ be a sequence of positive constants diverging to infinity and with the property that for some sequence $\{a_n\}$, $(S_n - a_n)/b_n \to_p 0$. It follows from Corollary 1, page 245 of Loève (1963) that a suitable choice for a_n is $a_n = \text{med}(S_n)$, and from Theorem 1, page 258 of Petrov (1975) that another choice is $a_n = nE[XI(|X| \le b_n)]$. Our first result shows that either of these versions of a_n gives an optimal rate of convergence. Define

$$S_n' = \textstyle\sum_1^n X_i(|X_i| \le b_n) \qquad \text{and} \qquad \delta_{n1} = n\{b_n^{-2}E[X^2I(|X| \le b_n)] + P(|X| > b_n)\}.$$

Note that $(S_n - a_n)/b_n \to_p 0$ for some sequence $\{a_n\}$ if and ony if $\delta_{n1} \to 0$; see Gnedenko and Kolmogorov (1954, page 105).

THEOREM 1. *Suppose $(S_n - a_n)/b_n \to_p 0$ for some sequence $\{a_n\}$, and the distribution of X is not concentrated at a point. Then*

$$\inf_a \Delta(a, b_n) \ge C_1\delta_{n1}. \tag{1}$$

With a_n equal to any one of $\text{med}(S_n)$, $\text{med}(S_n')$ *or* $nE[XI(|X| \le b_n)] = E(S_n')$, *we have*

$$\Delta(a_n, b_n) \le C_2\delta_{n1}.$$

In particular with this choice of a_n, $\Delta_n(a_n, b_n) \asymp \delta_{n1}$.

Theorem 1 demonstrates clearly the advantage of centring at the median or the sum of truncated means. However in many real situations it is difficult to determine either of these quantities, and one must resort to somewhat cruder location constants. When the mean is finite it is common to centre at the mean of S_n, and in this case there is no loss of generality in assuming that $E(X) = 0$ and $a_n = 0$. In many cases when $E|X| = \infty$ it is also usual to take $a_n = 0$, for very often b_n increases at a faster rate than $nE[XI(|X| \le b_n)]$, and so the centring constants are asymptotically negligible:

$$nb_n^{-1}E[XI(|X| \le b_n)] \to 0.$$

This is the case in several of the characterizations of rates of convergence given by Katz (1962), Baum and Katz (1963, 1965), Heyde and Rohatgi (1967) and Chow and Lai (1975).

It is therefore important to estimate the rate of convergence when we take $a_n = 0$. Our next theorem solves this problem. Define

$$\delta_{n2} = \delta_{n1} + |nb_n^{-1}E[XI(|X| \leq b_n)]|^2.$$

Note that $S_n/b_n \to_p 0$ if and only if $\delta_{n2} \to 0$.

THEOREM 2. *Suppose that $S_n/b_n \to_p 0$, and the distribution of X is not concentrated at a point. Then $\Delta(0, b_n) \stackrel{\cup}{\cap} \delta_{n2}$.*

From Theorems 1 and 2 we see that centring at the median rather than the origin (in the case of zero mean or, on occasion, infinite mean) provides a significant improvement in the rate of convergence if and only if δ_{n1} is asymptotically negligible in comparison with $|nb_n^{-1}E[XI(|X| \leq b_n)]|^2$. To compare the asymptotic behaviour of these quantities we shall consider the special case of a distribution with regularly varying tails.

Suppose $P(|X| > x) = x^{-\alpha}L(x)$ where $\alpha \geq 0$ and L is slowly varying at infinity. Integrating by parts we find that

$$x^{-2}E[X^2I(|X| \leq x)] + P(|X| > x) = 2x^{-2}\int_0^x uP(|X| > u)\,du$$
$$\sim 2(2-\alpha)^{-1}x^{-\alpha}L(x) \qquad \text{if} \quad 0 \leq \alpha < 2.$$

(See Feller (1971, page 281) for the requisite theory of regular variation.) Therefore if $0 \leq \alpha < 2$ the quantities $x^{-2}E[X^2I(|X| \leq x)]$ and $P(|X| > x)$ exhibit the same asymptotic behaviour, and $\delta_{n1} \sim Cnb_n^{-\alpha}L(b_n)$. If $\alpha \geq 2$ and $E(X^2) < \infty$ then $\delta_{n1} \sim nb_n^{-2}E(X^2)$. If $\alpha = 2$ and $E(X^2) = \infty$ then

$$x^{-2}E[X^2I(|X| \leq x)] \sim 2x^{-2}\int_1^x u^{-1}L(u)\,du,$$

which dominates the behaviour of $P(|X| > x)$. Consequently

$$\delta_{n1} \sim 2nb_n^{-2}\int_1^{b_n} u^{-1}L(u)\,du.$$

To elucidate the behaviour of δ_{n2} we consider first the case $0 < \alpha < 1$. Then

$$|E[XI(|X| \leq x)]| \leq E[|X|I(|X| \leq x)] \leq \int_0^x P(|X| > u)\,du$$
$$\sim (1-\alpha)^{-1}x^{1-\alpha}L(x),$$

and it follows that $\delta_{n1} \stackrel{\cup}{\cap} \delta_{n2}$. If $\alpha > 1$ and $E(X) = 0$ then for $x > 1$,

$$|E[XI(|X| \leq x)]| \leq E[|X|I(|X| \geq x)] \leq CxP(|X| > x),$$

and again $\delta_{n1} \stackrel{\cup}{\cap} \delta_{n2}$. The case $\alpha = 1$ is the exception. If $\alpha = 1$, $E|X| = \infty$ and

(2) $$P(X > x) - P(X < -x) \sim Cx^{-1}L(x)$$

for $0 < C \leq 1$, then for large x,

$$|E[XI(|X| \leq x)]| \geq \tfrac{1}{2}C\int_1^x u^{-1}L(u)\,du \gg L(x).$$

In the case $L(x) = (\log x)^{-1}$ and $b_n = n \log n$ we have $\delta_{n1} \sim C(\log n)^{-2}$ while $\delta_{n2} \sim C(\log\log n/\log n)^2$. If $\alpha = 1$, $E|X| < \infty$, $E(X) = 0$ and (2) holds then for large x,

$$|E[XI(|X| \leq x)]| = |E[XI(|X| > x)]| \geq \tfrac{1}{2}C\int_x^\infty u^{-1}L(u)\,du \gg L(x).$$

In the case $L(x) = (\log x)^{-2}$ and $b_n = n$ we have again $\delta_{n1} \sim C(\log n)^{-2}$ and $\delta_{n2} \sim C(\log\log n/\log n)^2$.

The simplicity of the rates of convergence given in theorems 1 and 2 makes it almost a trivial matter to obtain characterizations of rates of convergence of the type given by Hsu and Robbins (1947), Erdös (1949), Katz (1962), Baum and Katz (1963, 1965), Heyde and Rohatgi (1967), Chow and Lai (1975), Lai and Lan (1976) and Gut (1978). We present below only three sample results.

COROLLARY 1. *Suppose* $\Delta(\text{med } S_n, n^\gamma) \to 0$ *where* $\gamma > \frac{1}{2}$, *and* $0 \le \beta < 2\gamma - 1$. *Then*

$$\sum_1^\infty n^{-1+\beta}\Delta(\text{med } S_n, n^\gamma) < \infty$$

if and only if $E[|X|^{(\beta+1)/\gamma}] < \infty$, *and if* $\Delta(0, n^\gamma) \to 0$ *then*

$$\sum_1^\infty n^{-1+\beta}\Delta(0, n^\gamma) < \infty$$

if and only if both $E[|X|^{(\beta+1)/\gamma}] < \infty$ *and*

$$\int_1^\infty x^{-3+(\beta+2)/\gamma}|E[XI(|X| \le x)]|^2\,dx < \infty.$$

COROLLARY 2. *Suppose* $\gamma > \frac{1}{2}$ *and* $0 < \beta < 2\gamma - 1$. *Then*

$$\Delta(\text{med } S_n, n^\gamma) = O(n^{-\beta})$$

if and only if $P(|X| > x) = O(x^{-(\beta+1)/\gamma})$, *and*

$$\Delta(0, n^\gamma) = O(n^{-\beta})$$

if and only if both $P(|X| > x) = O(x^{-(\beta+1)/\gamma})$ *and*

$$|E[XI(|X| \le x)]|^2 = O(x^{2-(\beta+2)/\gamma}).$$

COROLLARY 3. *Suppose* $\gamma > \frac{1}{2}$, *and set* $\beta = 2\gamma - 1$. *The following three conditions are equivalent*:

$$\Delta(\text{med } S_n, n^\gamma) = O(n^{-\beta});$$

$$E|X| < \infty \quad \text{and} \quad \Delta(nEX, n^\gamma) = O(n^{-\beta});$$

$$E(X^2) < \infty.$$

Interestingly, Corollary 3 provides a rate of convergence which is equivalent to the existence of finite variance. For similar results in the case of rates of convergence in the central limit theorem, see Egorov (1973) and Heyde (1973).

For certain specific values of β and γ some of the conditions may be simplified. If $\beta + 1 > \max(\gamma, 2\gamma - 1)$ and $E[|X|^{(\beta+1)/\gamma}] < \infty$ then

$$\int_1^\infty x^{-3+(\beta+2)/\gamma}|E[XI(|X| \le x)]|^2\,dx < \infty$$

if and only if $E(X) = 0$, and if $\beta + 1 > \max(\gamma, 2\gamma - 1)$ and $P(|X| > x) = O(x^{-(\beta+1)/\gamma})$ then

$$|E[XI(|X| \le x)]|^2 = O(x^{2-2(\beta+1)/\gamma})$$

if and only if $E(X) = 0$.

We provide below a sketch of the proofs of Corollaries 1, 2 and 3.

(i) *Corollary* 1. Using integral approximations to series, it may be proved that either of the series

$$\Sigma n^{-1+\beta}\cdot nP(|X| > n^\gamma) \quad \text{or} \quad \Sigma n^{-1+\beta}\cdot n^{1-2\gamma}E[X^2I(|X| \le n^\gamma)]$$

converges if and only if $E[|X|^{(\beta+1)/\gamma}] < \infty$, and that when this condition holds the series and integral

$$\Sigma n^{-1+\beta} | n^{1-\gamma} E[XI(|X| \le n^{\gamma})] |^2 \quad \text{and} \quad \int_1^{\infty} x^{1+\beta-2\gamma} | E[XI(|X| \le x^{\gamma})] |^2 \, dx$$

converge or diverge together.

(ii) *Corollary* 2. If $nP(|X| > n^{\gamma}) = O(n^{-\beta})$ then $n^{1-2\gamma}E[X^2 I(|X| \le n^{\gamma})] = O(n^{-\beta})$, and Corollary 2 is then easily proved.

(iii) *Corollary* 3. Here $n^{1-2\gamma}E[X^2 I(|X| \le n^{\gamma})] = O(n^{-\beta})$ if and only if $E(X^2) < \infty$, and this in turn implies that $nP(|X| > n^{\gamma}) = O(n^{-\beta})$, by Chebychev's inequality.

3. The proofs.

PROOF OF THEOREM 1. From the symmetrization inequalities (see Loève (1963, page 245)) we see that

$$\begin{aligned}\delta_{n1}^s &\equiv b_n^{-2}E[(X_1 - X_2)^2 I(|X_1 - X_2| \le b_n)] + P(|X_1 - X_2| > b_n) \\ &= 2b_n^{-2} \int_0^{b_n} uP(|X_1 - X_2| > u)du \le 4b_n^{-2} \int_0^{b_n} uP(|X| > \tfrac{1}{2}u)\, du \le 16\delta_{n1},\end{aligned}$$

and if $m = \text{med}(X)$ then for large n,

$$\delta_{n1}^s \ge \tfrac{1}{2}b_n^{-2} \int_0^{b_n} uP(|X - m| > u)\, du \ge Cb_n^{-2} \int_0^{b_n} uP(|X| > 2u)\, du \ge \tfrac{1}{4}C\delta_{n1}.$$

To obtain the middle inequality, observe that for $b_n > |m|$,

$$\int_{|m|}^{b_n} uP(|X - m| > u)\, du \ge \int_{|m|}^{b_n} uP(X > 2u)\, du.$$

It follows that $\delta_{n1} \le C_1\delta_{n1}^s \le C_2\delta_{n1}$. Next let S_n^* be an independent copy of S_n, and set $S_n^s = S_n - S_n^*$ and $B(x) = 1 - e^{-x^2/2}$. There exists a constant $C > 0$ such that $B(x) \le A(x) \equiv \min(1, x^2) \le CB(x)$ for all x, and for any $c > 0$ and any a,

$$\begin{aligned}E[B(S_n^s/cb_n)] &= \int_0^{\infty} P(|S_n^s| > cb_n x)xe^{-x^2/2}\, dx \\ &\le 2\int_0^{\infty} P(|S_n - a| > \tfrac{1}{2}cb_n x)xe^{-x^2/2}\, dx \\ &\le 2E[A((S_n - a)/\tfrac{1}{2}cb_n)],\end{aligned}$$

using the weak symmetrization inequalities. In view of the estimates above, the result (1) will follow if we prove that in the case of a *symmetric* distribution, and for any $c > 0$,

$$E[B(S_n/cb_n)] \ge C\delta_{n1}. \tag{3}$$

The case of a general c is treated exactly as $c = 1$, and so we shall make this assumption.

Let ϕ_n denote the (real valued) characteristic function of $XI(|X| \le b_n)$, and F_0 the distribution function giving unit mass to the origin. Then

$$\begin{aligned}&[1 - \phi_n^n(t/b_n)]/it = \int_{-\infty}^{\infty} [P(S_n'/b_n \le x) - F_0(x)]e^{itx}\, dx \quad \text{and} \\ &ite^{-t^2/2} = (2\pi)^{-1/2} \int_{-\infty}^{\infty} xe^{-x^2/2}e^{itx}\, dx,\end{aligned} \tag{4}$$

and so by Parseval's identity,

$$(5)\qquad -\int_{-\infty}^{\infty}[P(S_n'/b_n\le x)-F_0(x)]xe^{-x^2/2}\,dx=(2\pi)^{-1/2}\int_{-\infty}^{\infty}[1-\phi_n^n(t/b_n)]e^{-t^2/2}\,dt.$$

The left hand side equals $E[B(S_n'/b_n)]$, while the right hand side dominates $C\int_0^1[1-\phi_n^n(t/b_n)]\,dt$. Now, $\log\phi_n(t/b_n)\le-[1-\phi_n(t/b_n)]$ and so

$$1-\phi_n^n(t/b_n)\ge1-\exp\{-n[1-\phi_n(t/b_n)]\}.$$

If $nb_n^{-2}E[X^2I(|X|\le b_n)]\le1$ then

$$n[1-\phi_n(t/b_n)]=nE\{[1-\cos(tX/b_n)]I(|X|\le b_n)\}\le 1/2$$

for $t\in(0,1)$, and consequently $1-\phi_n^n(t/b_n)\ge Cn[1-\phi_n(t/b_n)]$. Therefore

$$\begin{aligned}E[B(S_n'/b_n)]&\ge C_1n\int_0^1[1-\phi_n(t/b_n)]\,dt\\&=C_1n\int_0^{b_n}dP(|X|\le x)\int_0^1[1-\cos(tx/b_n)]\,dt\\&\ge C_2n\int_0^{b_n}(x/b_n)^2\,dP(|X|\le x)=C_2nb_n^{-2}E[X^2I(|X|\le b_n)].\end{aligned}$$

Using techniques of Erdös (1949, page 290) we may prove that $P(|S_n|>b_n)\ge CnP(|X|>b_n)$ for large n, and consequently $\Delta(0,b_n)\ge C_1P(|S_n|>b_n)\ge C_2nP(|X|>b_n)$. (Note that we are considering a symmetric distribution.) It follows that for any x,

$$\begin{aligned}P(|S_n'|>b_nx)&\le nP(|X|>b_n)+P(|S_n|>b_nx)\\&\le CE[B(S_n/b_n)]+P(|S_n|>b_nx),\end{aligned}$$

and so

$$E[B(S_n'/b_n)]=\int_0^{\infty}P(|S_n'|>b_nx)xe^{-x^2/2}\,dx\le CE[B(S_n/b_n)].$$

Combining these estimates we deduce that (3) (with $c=1$) holds for δ_{n1} sufficiently small. For any $\varepsilon>0$, $\delta_{n1}>\varepsilon$ for only a finite number of values of n, and so the constant C may be chosen so that (3) holds for all n.

The next step is to prove that with $d_n=nE[XI(|X|\le b_n)]$ and for any $c>0$,

$$(6)\qquad E[B((S_n'-d_n)/cb_n)]\le C\delta_{n1}.$$

We no longer restrict our attention to the symmetric case. From (6) and the weak symmetrization inequalities it follows that with $e_n=\operatorname{med}(S_n')$ and for any $c>0$, $E[B((S_n'-e_n)/cb_n)]\le C\delta_{n1}$. Since

$$P(|(S_n-a)/b_n|>x)\le\delta_{n1}+P(|(S_n'-a)/b_n|>x)$$

for any a and x, we may conclude immediately that $\Delta(d_n,cb_n)+\Delta(e_n,cb_n)\le C\delta_{n1}$. Using the weak symmetrization inequalities again we may now deduce that $\Delta(\operatorname{med}S_n,b_n)\le C\delta_{n1}$.

The proof of (6) for arbitrary c is identical to the proof for $c=1$, and so we shall make this assumption. Using the techniques leading to (5) we may obtain

$$B_n\equiv E[B((S_n'-d_n)/b_n)]=(2/\pi)^{1/2}\int_0^{\infty}\mathscr{R}\ell[1-\phi_n^n(t/b_n)e^{-itd_n/b_n}]e^{-t^2/2}\,dt.$$

Write $n \log \phi_n(t/b_n) - itd_n/b_n = \alpha_n(t) + i\beta_n(t)$ for real valued functions α_n and β_n. Necessarily $\alpha_n \le 0$, and consequently

$$\mathscr{R}\ell[1 - \phi_n^n(t/b_n)e^{-itd_n/b_n}] = 1 - [\cos \beta_n(t)]e^{\alpha_n(t)} \le |\alpha_n(t)| + |\beta_n(t)|.$$

If $\varepsilon > 0$ is chosen sufficiently small then

$$|1 - \phi_n(t)| \le E(1 - \cos tX) + |E \sin tX| + P(|X| > b_n) < ½$$

for large n and $|t| < \varepsilon$. Therefore for $|t| < \varepsilon b_n$,

$$n \log \phi_n(t/b_n) - itd_n/b_n = -n[1 - \phi_n(t/b_n)] - itd_n/b_n + r_n(t)$$

where $|r_n(t)| \le Cn|1 - \phi_n(t/b_n)|^2$, and so for $0 < t < \varepsilon b_n$,

$$|\alpha_n(t)| + |\beta_n(t)| \le C(nE\{[1 - \cos(tX/b_n)]I(|X| \le b_n)\} + nE\{|(tX/b_n) - \sin(tX/b_n)|I(|X| \le b_n)\} + n|1 - \phi_n(t/b_n)|^2).$$

For $t > 0$ we have the following estimates:

$$\varepsilon_{n1} \equiv nE\{[1 - \cos(tX/b_n)]I(|X| \le b_n)\} \le ½t^2\delta_{n1};$$

$$\varepsilon_{n2} \equiv nE\{|(tX/b_n) - \sin(tX/b_n)|I(|X| \le b_n)\} \le ⅙t^3\delta_{n1};$$

$$n|1 - \phi_n(t/b_n)| \le \varepsilon_{n1} + \varepsilon_{n2} + n(t/b_n)|E[XI(|X| \le b_n)]|; \qquad \text{and}$$

$$\{b_n^{-1}|E[XI(|X| \le b_n]|\}^2 \le b_n^{-2}E[X^2I(|X| \le b_n)].$$

It follows that $|\alpha_n(t)| + |\beta_n(t)| \le C(1 + t^6)\delta_{n1}$, and therefore

$$B_n \le C\delta_{n1}\int_0^{\varepsilon b_n} (1 + t^6)e^{-t^2/2}\,dt + C\int_{\varepsilon b_n}^{\infty} e^{-t^2/2}\,dt.$$

Plainly $b_n^{-2} \le C\delta_{n1}$, and so (6) is true.

Proof of Theorem 2. Let $\{Z_{nm}, m \ge 1\}$ be a sequence of random variables with finite means and converging to S_n/b_n, and let F_{nm} and ψ_{nm} denote respectively the distribution function and characteristic function of Z_{nm}. In place of (4) we may prove that

$$[1 - \psi_{nm}(t)]/it = \int_{-\infty}^{\infty} [F_{nm}(x) - F_0(x)]e^{itx}\,dx,$$

which leads to an analogue of (5). Letting $m \to \infty$ in this result, we deduce that

$$E[B(S_n/b_n)] = -\int_{-\infty}^{\infty} [P(S_n/b_n \le x) - F_0(x)]xe^{-x^2/2}\,dx$$

$$= (2\pi)^{-1/2}\int_{-\infty}^{\infty} [1 - \phi^n(t/b_n)]e^{-t^2/2}\,dt.$$

The imaginary part of the right hand side is zero, and so writing $n \log \phi_n(t/b_n) = \alpha_n(t) + i\beta_n(t)$ for real valued functions α_n and β_n, we find that

(8) $$\int_0^1 [1 - \cos \beta_n(t)e^{\alpha_n(t)}]\,dt \le C_1E[B(S_n/b_n)] = C_1\int_0^{\infty} [1 - \cos \beta_n(t)e^{\alpha_n(t)}]e^{-t^2/2}\,dt \le C_2\int_0^{\infty} [|\alpha_n(t)| + |\beta_n(t)|^2]e^{-t^2/2}\,dt.$$

Using arguments very similar to those in the proof of Theorem 1, we deduce that for some

$\varepsilon > 0$ and all $0 < t < b_n$,

$$|\alpha_n(t)| \le C(1+t^2)\delta_{n1}, \tag{9}$$

$$|\beta_n(t)| \le C(1+t^6)\{\delta_{n1} + nb_n^{-1}|E[XI(|X| \le b_n)]|\} \quad \text{and} \tag{10}$$

$$\begin{aligned} |\beta_n(t)| &\ge |nE[\sin(tX/b_n)]| - Cn|1-\phi(t/b_n)|^2 \\ &\ge tnb_n^{-1}|E[XI(|X| \le b_n)]| - nP(|X| > b_n) - Cn|1-\phi(t/b_n)|^2 \\ &\quad - nE\{|(tX/b_n) - \sin(tX/b_n)|I(|X| \le b_n)\} \\ &\ge tnb_n^{-1}|E[XI(|X| \le b_n)]| - C(1+t^6)\delta_{n1}. \end{aligned} \tag{11}$$

From (8), (9) and (10) we see that $E[B(S_n/b_n)] \le C\delta_{n2}$, while from (8), (9), (10) and (11) it follows that

$$\begin{aligned} E[B(S_n/b_n)] &\ge C_1 \int_0^1 [1 - \cos\beta_n(t)]\,dt - C_2 \int_0^1 |\alpha_n(t)|\,dt \\ &\ge C_3 \int_0^1 |\beta_n(t)|^2\,dt - C_4 \int_0^1 [|\alpha_n(t)| + |\beta_n(t)|^4]\,dt \\ &\ge C_5 |nb_n^{-1}E[XI(|X| \le b_n)]|^2 - C_6\delta_{n1}, \end{aligned}$$

using the fact that $nb_n^{-1}E[XI(|X| \le b_n)] \to 0$. Consequently

$$|nb_n^{-1}E[XI(|X| \le b_n)]|^2 \le C_1\{E[B(S_n/b_n)] + \delta_{n1}\} \le C_2E[A(S_n/b_n)],$$

using Theorem 1. Theorem 2 now follows from Theorem 1.

REFERENCES

BAUM, L. E. and KATZ, M. (1963). Convergence rates in the law of large numbers. *Bull. Amer. Math. Soc.* **69** 771–772.

BAUM, L. E. and KATZ, M. (1965). Convergence rates in the law of large numbers. *Trans. Amer. Math. Soc.* **120** 108–123.

CHOW, Y. S. and LAI, T. L. (1975). Some one-sided theorems on the tail distribution of sample sums with applications to the last time and largest excess of boundary crossings. *Trans. Amer. Math. Soc.* **208** 51–72.

EGOROV, V. A. (1973). On the rate of convergence to the normal law which is equivalent to the existence of a second moment. *Theor. Probability Appl.* **18** 175–180.

ERDÖS, P. (1949). On a theorem of Hsu and Robbins. *Ann. Math. Statist.* **20** 286–291.

FELLER, W. (1971). *Introduction to Probability Theory and its Applications.* Wiley, New York.

GNEDENKO, B. V. and KOLMOGOROV, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables.* Addison-Wesley, Reading, Mass.

GUT, A. (1978). Marcinkiewicz laws and convergence rates in the law of large numbers for random variables with multidimensional indices. *Ann. Probability* **6** 469–482.

HEYDE, C. C. (1973). On the uniform metric in the context of convergence to normality. *Z. Wahrsch. verw. Gebiete* **16** 151–156.

HEYDE, C. C. and ROHATGI, V. K. (1967). A pair of complementary theorems on convergence rates in the law of large numbers. *Proc. Cambridge Phil. Soc.* **63** 73–82.

HSU, P. L. and ROBBINS, H. (1947). Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. U.S.A.* **33** 25–31.

KATZ, M. (1963). The probability in the tail of a distribution. *Ann. Math. Statist.* **34** 312–318.

LAI, T. L. and LAN, K. K. (1976). On the last time and the number of boundary crossings related to the strong law of large numbers and the law of the iterated logarithm. *Z. Wahrsch. verw. Gebeite* **34** 59–71.

LOEVE, M. (1963). *Probability Theory.* 3rd ed. Van Nostrand, Princeton.

PETROV, V. V. (1975). *Sums of Independent Random Variables.* Springer, Berlin.

THE AUSTRALIAN NATIONAL UNIVERSITY
DEPARTMENT OF STATISTICS
FACULTY OF ECONOMICS
BOX 4, PO
CANBERRA, A.C.T. 2600 AUSTRALIA