

NECESSARY AND SUFFICIENT CONDITIONS FOR COMPLETE CONVERGENCE IN THE LAW OF LARGE NUMBERS

BY SØREN ASMUSSEN AND THOMAS G. KURTZ

University of Copenhagen and University of Wisconsin-Madison

Relationships between the growth of a sequence N_k and conditions on the tail of the distribution of a sequence X_l of i.i.d. mean zero random variables are given that are necessary and sufficient for

$$\sum_{k=1}^{\infty} P \left\{ \left| \frac{1}{N_k} \sum_{l=1}^{N_k} X_l \right| > \epsilon \right\} < \infty.$$

The results are significant for distributions satisfying $E(|X_l|) < \infty$ but $E(|X_l|^\beta) = \infty$ for some $\beta > 1$. Necessary and sufficient conditions for the finiteness of sums of the form

$$\sum_{n=1}^{\infty} \gamma(n) P \left\{ \left| \frac{1}{n} \sum_{l=1}^n X_l \right| > \epsilon \right\}$$

are obtained as a corollary.

1. Introduction. Throughout this paper X_1, X_2, X_3, \dots will be a sequence of independent identically distributed random variables with

$$(1.1) \quad E|X_n| < \infty, EX_n = 0$$

and we will define $F(t) = P(|X_n| \geq t)$, $p(n, \epsilon) = P(|X_1 + \dots + X_n|/n > \epsilon)$. We are interested in relationships between the growth of a sequence N_k and conditions on the tail F of the X_l that imply

$$(1.2) \quad \sum_{k=1}^{\infty} p(N_k, \epsilon) < \infty \text{ for all } \epsilon > 0.$$

Finiteness of the sum in (1.2) is an indication of the rate of convergence (termed "complete convergence" by Hsu and Robbins (1947)) of the averages to zero and is the necessary and sufficient condition (by Borel-Cantelli) for the almost sure convergence of a triangular array of independent random variables. The sum can also be interpreted as the expected number of times $|\sum_{l=1}^{N_k} X_l| > \epsilon N_k$. Questions of this type arise in the study of sums of independent random variables indexed by lattices and more general partially ordered sets, for example, Smythe (1974), as well as in the study of supercritical branching processes, for example, Asmussen (1978) and Athreya and Kaplan (1978). In that setting, the sequence is replaced by a triangular array and N_1, N_2, \dots are the generation sizes of the branching process (of course, here N_1, N_2, \dots are random variables and the $p(N_k, \epsilon)$ are conditional probabilities).

In Theorems A and B $\psi(x)$ will be a strictly increasing function on $[0, \infty)$ satisfying $\sup_{x \geq 1} \psi(x+1)/\psi(x) < \infty$; $\psi^{-1}(x)$ will be absolutely continuous with

Received December 5, 1977; revised November 17, 1978.

AMS 1970 subject classifications. Primary 60F10; secondary 60F05, 60F15.

Key words and phrases. Law of large numbers, triangular array, complete convergence, error estimates.

derivative $\gamma(x)$; and for some θ and some $C > 0$ $\gamma(x)$ will satisfy $\gamma(yx) \leq Cy^\theta\gamma(x)$ for $x \geq 0$ and $y \geq 1$. Note that for some C_1, C_2 and some $\alpha, \beta \leq 0$, $C_1x^\alpha \leq \psi(x) \leq C_2e^{\beta x}$. The requirement on γ is automatic if ψ is convex ($\theta = 0$) and also is satisfied if $\psi(x) = x^\alpha, \alpha > 0$ ($\theta = (1 - \alpha)/\alpha$). If γ is increasing then this condition is equivalent to dominated variation as defined in Feller (1969).

THEOREM A. *Let t_k, N_k satisfy*

$$(1.3) \quad \lim_{k \rightarrow \infty} N_k / \psi(t_k) = 1.$$

If $\inf_k t_k - t_{k-l} \equiv a > 0$ for some $l > 0$ and

$$(1.4) \quad \int_0^\infty t\gamma(t)F(t) dt = Eg(|X_l|) < \infty$$

where $g(t) = \int_0^t s\gamma(s) ds$, then (1.2) holds. Conversely, if $\sup_k t_k - t_{k-1} \equiv A < \infty$ and (1.2) holds, then (1.4) holds.

COROLLARY I. *(1.4) holds if and only if (1.2) holds for some [and then every] sequence N_k satisfying $\lim_{k \rightarrow \infty} N_k / \psi(hk) = 1$ for some $h > 0$. Also, (1.4) is equivalent to the existence of sequences t_k, N_k satisfying (1.2), (1.3) and $\lim_{k \rightarrow \infty} t_{k+1} - t_k = 0$.*

COROLLARY II. *(1.4) holds if and only if*

$$(1.5) \quad \sum_{n=0}^\infty (\psi^{-1}(n+1) - \psi^{-1}(n))p(n, \epsilon) < \infty \quad \text{for all } \epsilon > 0.$$

THEOREM B. *Define $H(t) = \sup_{s \geq t} sF(s)$. Then (1.2) holds for every sequence N_k satisfying*

$$(1.6) \quad \inf_k N_k / \psi(hk) > 0 \quad \text{for some } h > 0$$

if and only if

$$(1.7) \quad \int_0^\infty \gamma(t)H(t) dt < \infty.$$

REMARK. One might expect that if (1.2) holds and $M_k > N_k$, then $\sum p(M_k, \epsilon) < \infty$ for all $\epsilon > 0$. A comparison of Theorems A and B shows that this is not the case in general. However, roughly speaking, moment assumptions only slightly stronger than (1.1) will ensure that (1.4) and (1.7) are the same condition and in fact are both equivalent to

$$(1.8) \quad \int_0^\infty \psi^{-1}(t)F(t) dt < \infty.$$

More precisely, always (1.8) \Rightarrow (1.7) \Rightarrow (1.4) and (1.4) \Leftrightarrow (1.8) for any concave ψ and also, e.g., for $\psi(x) = x^\beta, \beta > 0$. We verify these implications below. For an example where (1.4) holds but not (1.7), let $\gamma(t) = 1/t$ and let the X_l be concentrated on $\{\pm 2, \pm 2^4, \dots, \pm 2^{j^2}, \dots\}$ with $P(X_l = 2^{j^2}) = P(X_l = -2^{j^2}) = c/j^2 2^{j^2}$.

EXAMPLE. In connection with the branching process example above, we note the following particular cases. If $N_k / wm^k \rightarrow 1$ with $m > 1, 0 < w < \infty$, then (1.2) holds without further assumptions than (1.1) (from a generalization of Corollary I). This may fail if it is only known that $N_k \geq w_1 m_1^k$ with $0 < w_1 < \infty, m_1 > 1$, but

here at least the condition $E|X_l| \log^+ |X_l| < \infty$ is sufficient for (1.2) (Theorem B and the remark). However, if $\liminf N_{k+1}/N_k \geq m_2 > 1$, then again (1.1) is sufficient for (1.2) (Theorem A with $\psi(x) = e^x$, $t_k = \log N_k$).

There are many results in the literature related to those given here (see the review paper by Hanson (1970) and Heathcote (1967)). In particular, Katz (1963) proved that for $\lambda \geq 1$

$$(1.9) \quad \sum_{n=1}^{\infty} n^{\lambda-2} p(n, \epsilon) < \infty \quad \text{for every } \epsilon > 0$$

if and only if $E|X_l|^\lambda < \infty$, thereby generalizing earlier results of Spitzer (1956), Erdős (1949) and Hsu and Robbins (1947). This result is the special case of Corollary II with $\psi(x) = x^{1/(\lambda-1)}$ $\lambda > 1$ and $\psi(x) = e^x$ for $\lambda = 1$. More recently Smythe (1974) has given a result similar to the first part of Theorem A but with more restrictions on ψ and on the relationship between ψ and N_k . In connection with the idea of replacing $tF(t)$ by $H(t)$ in Theorem B, see also Franck and Hanson (1966), and in connection with the last part of Corollary I, Dvoretzky (1949).

2. PROOFS. The results are based on the following theorems from Kurtz (1972).

THEOREM C. *Let X_1, X_2, \dots be independent random variables with mean zero and let $S_m = \sum_{k=1}^m a_k X_k$ for some sequence $a_k \geq 0$. Let $F(t) = \sup_k P\{|X_k| \geq t\}$ and define $\eta = \sup_k a_k \int_{1/a_k}^{\infty} F(t) dt$ and $e = \sum_k a_k \int_{1/a_k}^{\infty} F(t) dt$.*

(a) *If $\eta < 2\delta$ and $0 \leq \alpha < 1$, then*

$$P\{\sup_m |S_m| > \delta + e\} \leq (\alpha + 1) \left[\frac{4}{(2\delta - \eta)^{\alpha+1}} + 1 \right] \sum_k \int_0^{1/a_k} u^\alpha F(u/a_k) du.$$

(b) *Let $\epsilon > 0$ and $L > 1$. If*

$$\alpha \leq 1 + \epsilon(L - 1), \eta < \delta/2^L, \sum_k \int_0^{1/a_k} u^{1-\epsilon} F(u/a_k) du \leq M < \infty,$$

then

$$P\{\sup_m |S_m| > \delta + e\} \leq \left(\sum_k \int_0^{1/a_k} u^\alpha F(u/a_k) du \right) (\alpha + 1) \times \left[1 + \sum_{l=1}^L \frac{(\alpha(\alpha + 1)M)^{l-1}}{\prod_{m=0}^{l-1} (\delta/2^{m+1} - \eta)^{\alpha+1}} + \frac{\alpha(\alpha(\alpha + 1)M)^{L-1}}{\prod_{m=0}^{L-1} (\delta/2^{m+1} - \eta)^{\alpha+1}} \right].$$

Note that

$$\sum_k \int_0^{1/a_k} u^{1-\epsilon} F(u/a_k) du = \sum_k a_k \int_0^{1/a_k} (a_k v)^{1-\epsilon} F(v) dv \leq \sum_k a_k \int_0^\infty F(v) dv.$$

Consequently when $\sum_k a_k = 1$, for example, we may take $\epsilon = 1$ and replace M by $\int_0^\infty F(v) dv$.

COROLLARY III. *Let $a_k = 1/N$, $k = 1, \dots, N$ and $a_k = 0$ for $k > N$. Let $\eta = 1/N \int_N^\infty F(t) dt$, $e = \int_N^\infty F(t) dt$ and $M = \int_0^\infty F(t) dt$.*

(a) *If $\eta < 2\delta$ and $\alpha \leq 1$, then*

$$P\{\sup_m |S_m| > \delta + e\} \leq CN \int_0^1 u^\alpha F(uN) du$$

where C depends on α, δ and η , and is increasing in η .

(b) Let $L > 1, \alpha \leq 1 + (L - 1)$ and $\eta < \delta/2^L$. Then

$$P\{\sup_m |S_m| > \delta + \epsilon\} \leq CN \int_0^1 u^\alpha F(uN) du$$

where C depends on α, δ, M and η and is increasing in η .

We start by proving the first part of Theorem A. Taking $\alpha = 2 + \theta$ in Corollary III, (1.2) is implied by

$$(2.1) \quad \sum N_k \int_0^1 u^{2+\theta} F(N_k u) du < \infty.$$

Let $M = \sup N_k/\psi(t_k), m = \inf N_k/\psi(t_k)$. Since we can break the sum in (1.2) into l parts, without loss of generality we can assume $l = 1$. Furthermore we can assume $t_1 = 2$ and hence

$$\sup_k \sup_{t_{k-1} < t \leq t_k} \frac{\psi(t_k)}{\psi(t)} \equiv K < \infty.$$

Then (2.1) is bounded by

$$\begin{aligned} M \sum_k \psi(t_k) \int_0^1 u^{2+\theta} F(mu\psi(t_k)) du &= M \int_0^1 u^{2+\theta} \sum_k \psi(t_k) F(mu\psi(t_k)) du \\ &\leq M \int_0^1 u^{2+\theta} \sum_k \frac{1}{a} \int_{t_{k-1} \vee (t_k-1)}^{t_k} \psi(t_k) F(mu\psi(t)) dt du \\ (2.2) \quad &\leq \frac{KM}{a} \int_0^1 u^{2+\theta} \sum_k \int_{t_{k-1} \vee (t_k-1)}^{t_k} \psi(t) F(mu\psi(t)) dt du \\ &\leq \frac{KM}{a} \int_0^1 u^{2+\theta} \int_0^\infty \psi(t) F(mu\psi(t)) dt du. \end{aligned}$$

Substituting $x = mu\psi(t)$ the inner integral becomes

$$(2.3) \quad \int_0^\infty \frac{x}{m^2 u^2} \gamma\left(\frac{x}{mu}\right) F(x) dx \leq C \int_0^\infty \frac{x}{m^{2+\theta} u^{2+\theta}} \gamma(x) F(x) dx$$

and finiteness of (2.1) follows.

To prove the sufficiency of (1.7) in Theorem B we bound (2.1) by

$$(2.4) \quad \sum_k \int_0^1 u^{\theta+1} H(N_k u) du$$

and then using the monotonicity of H approximate (2.4) by

$$\begin{aligned} \sum_k \int_0^1 u^{\theta+1} H(\psi(k)u) du &\approx \int_0^1 u^{\theta+1} \int_0^\infty H(\psi(x)u) dx du \\ (2.5) \quad &= \int_0^1 u^{\theta+1} \int_0^\infty \frac{1}{u} \gamma\left(\frac{t}{u}\right) H(t) dt du \leq C \int_0^1 \int_0^\infty \gamma(t) H(t) dt du < \infty. \end{aligned}$$

The second part of Theorem A and the necessity of (1.7) is based upon

LEMMA. *If $0 < \epsilon < \delta$, then $p(N, \epsilon) \geq NF(N\delta)(1 - o(1))$.*

PROOF. Define

$$S_N = X_1 + \cdots + X_N, S_N^{(k)} = S_N - X_k \quad k = 1, \cdots, N,$$

$$M_N = |X_1| \vee \cdots \vee |X_N|.$$

Then

$$P(M_N \geq N\delta) = P(\cup_{k=1}^N \{|X_k| \geq N\delta\}) \geq NF(N\delta) - \binom{N}{2} F(N\delta)^2,$$

$$P(|S_N| < N\epsilon, M_N \geq N\delta) \leq \sum_{k=1}^N P(|S_N^{(k)}| \geq N(\delta - \epsilon), |X_k| \geq N\delta)$$

$$= NP(|S_{N-1}| \geq N(\delta - \epsilon))F(N\delta) = o(NF(N\delta))$$

and the assertion follows since

$$p(N, \epsilon) \geq P(|S_N| > N\epsilon, M_N \geq N\delta) = P(M_N \geq N\delta) - P(|S_N| \leq N\epsilon, M_N \geq N\delta).$$

To obtain the second part of Theorem A let $\kappa = \sup_{t > t_1} \sup_{0 < s < A} (\psi(t + S))/\psi(t)$. Then

$$\sum_k N_k F(N_k/M) \geq m \sum_k \psi(t_k) F(\psi(t_k)) \geq \frac{m}{A} \sum_k \int_{t_k}^{t_k + 1} \psi(t_k) F(\psi(t)) dt$$

$$\geq \frac{m}{A\kappa} \int_{t_1}^{\infty} \psi(t) F(\psi(t)) dt = \frac{m}{A\kappa} \int_{\psi^{-1}(t_1)}^{\infty} x\gamma(x) F(x) dx$$

and if (1.2) holds, then the left-hand side is finite by the lemma.

Clearly (1.4) is necessary in Theorem B as well. Therefore to verify the necessity of (1.7) we may assume (1.4) holds but that

$$\int_0^{\infty} \gamma(t) H(t) dt = \infty.$$

Let $T_1 = \{t : H(t) \neq tF(t)\}$ and $T_2 = \{t : H(t) = tF(t)\}$. Then

$$\int_0^{\infty} \gamma(t) H(t) dt = \int_{T_1} \gamma(t) H(t) dt + \int_{T_2} \gamma(t) tF(t) dt = \infty.$$

Since the second term on the right is finite the first is infinite. For every $t \in T_1$ there is a $s_t > t$ such that $H(u) = s_t F(s_t)$ for $t \leq u \leq s_t$. Since $H(t)$ is left continuous T_1 is a union of intervals $(t_n, s_n) \equiv (t_n, s_n)$ on which $H(t)$ is a constant (i.e., $s_n F(s_n)$). Therefore

$$\int_{T_1} \gamma(t) H(t) dt = \sum s_n F(s_n) \int_{t_n}^{s_n} \gamma(t) dt$$

$$= \sum s_n F(s_n) (\psi^{-1}(s_n) - \psi^{-1}(t_n)) = \infty.$$

We cannot assume $t_n \geq s_{n-1}$, but we can select a subset of the intervals $(t_n, s_n) \equiv (a_i, b_i)$ such that $a_i \geq b_{i-1}$ and

$$\sum b_i F(b_i) (\psi^{-1}(b_i) - \psi^{-1}(a_i)) = \infty.$$

Furthermore since

$$(\psi^{-1}(b_i) - \psi^{-1}(a_i)) b_i F(b_i) \leq \frac{b_i}{a_i} \int_{a_i}^{b_i} t \gamma(t) F(t) dt$$

we are able to select the intervals so that $\lim_{i \rightarrow \infty} b_i/a_i = \infty$. We obtain the desired

sequence by defining $N_k = [b_i]$ for

$$\psi^{-1}([b_i]) \geq k > \psi^{-1}([b_{i-1}])$$

and observe that

$$\begin{aligned} \Sigma N_k F(N_k) &\geq \Sigma(\psi^{-1}([b_i]) - \psi^{-1}([b_{i-1}]) - 1) \cdot [b_i] F([b_i]) \\ &\geq \Sigma(\psi^{-1}([b_i]) - \psi^{-1}(a_i) - 1) \cdot [b_i] F([b_i]) = \infty. \end{aligned}$$

The first part of Corollary I follows immediately from Theorem A. To obtain the second part, note that if t_k, N_k satisfy (1.2), (1.3) and $\lim_{k \rightarrow \infty} t_{k+1} - t_k = 0$ then (1.4) follows by the converse of Theorem A.

If (1.4) holds, then

$$(2.6) \quad \Sigma_{i=1}^{\infty} p([\psi(l/m)], 1/m) < \infty$$

for every m and hence we may select M_m such that $M_m/m \uparrow \infty$ as $m \uparrow \infty$ and

$$\Sigma_{i=M_m}^{\infty} p([\psi(l/m)], 1/m) < m^{-2}.$$

Let $t_1 < t_2 < t_3 < \dots$ be the ordering of

$$\cup_m \{l/m : M_m/m \leq l/m < M_{m+1}/(m+1)\}$$

and set $N_k = [\psi(t_k)]$. If $t_k \geq M_m/m$ then $t_{k+1} - t_k \leq m^{-1}$ and if $\epsilon > 1/m$ then

$$\Sigma_k p(N_k, \epsilon) \leq \Sigma_{t_k < M_m/m} p(N_k, \epsilon) + \Sigma_{l=m}^{\infty} l^{-2} < \infty.$$

For Corollary II, define $\Gamma_k = \{n : \psi(k) - 1 \leq n < \psi(k+1)\}$. Let $N_k, N'_k \in \Gamma_k$ satisfy

$$p(N'_k, \epsilon) \leq p(n, \epsilon) \leq p(N_k, \epsilon)$$

for all $n \in \Gamma_k$. Let $a_k^n = (\psi^{-1}(n+1) \wedge (k+1)) \vee k - (\psi^{-1}(n) \wedge (k+1)) \vee k$. Note that

$$\Gamma_k = \{n : a_k^n > 0\}, \quad \Sigma_k a_k^n = \psi^{-1}(n+1) - \psi^{-1}(n), \quad \Sigma_n a_k^n = 1.$$

From this we have

$$(2.7) \quad \begin{aligned} \Sigma_k p(N'_k, \epsilon) &\leq \Sigma_{k,n} a_k^n p(n, \epsilon) \\ &= \Sigma_n (\psi^{-1}(n+1) - \psi^{-1}(n)) p(n, \epsilon) \leq \Sigma_k p(N_k, \epsilon). \end{aligned}$$

Since

$$\psi(k) - 1 \leq N_k \leq \psi(k+1)$$

there exist $k \leq t_k \leq k+1$ such that $\lim_{k \rightarrow \infty} N_k/\psi(t_k) = 1$. Since $\inf_k t_k - t_{k-2} \geq 1$ and $\sup_k t_k - t_{k-1} \leq 2$, Theorem A applies to the right hand side of (2.7) and similarly to the left.

It only remains to verify the claims in the remark following Theorem B. Note first that

$$\begin{aligned} \int_0^{\infty} (\psi^{-1}(t) - \psi^{-1}(0)) F(t) dt &= \int_0^{\infty} \int_0^t \gamma(s) ds F(t) dt \\ &= \int_0^{\infty} \gamma(t) \int_t^{\infty} F(s) ds dt \\ &= 2^{-1} \int_0^{\infty} \gamma(t/2) \int_{t/2}^{\infty} F(s) ds dt. \end{aligned}$$

Let $s_t \geq t$, $s_t F(s_t) = H(t)$. Then

$$\begin{aligned} C 2^\theta \int_0^\infty \gamma(t/2) \int_{t/2}^\infty F(s) ds dt &\geq \int_0^\infty \gamma(t) (s_t - t/2) F(s_t) dt \\ &\geq 2^{-1} \int_0^\infty \gamma(t) H(t) dt \end{aligned}$$

and it follows that (1.8) \Rightarrow (1.7). Finally if ψ is concave then (1.4) \Rightarrow (1.8) since

$$t\gamma(t) \geq \int_0^t \gamma(s) ds = \psi^{-1}(t) - \psi^{-1}(0),$$

and if $\psi(x) = x^\beta$ then $t\gamma(t) = \psi^{-1}(t)/\beta$.

REFERENCES

- [1] ASMUSSEN, SØREN (1978). Some martingale methods in the limit theory of supercritical branching processes. In *Advances in Probability and Related Topics*. (P. Ney ed.). Marcel Dekker, New York.
- [2] ATHREYA, K. B. and KAPLAN, N. (1978). Additive property and its applications in branching processes. In *Advances in Probability and Related Topics*. (P. Ney ed.). Marcel Dekker, New York.
- [3] DVORETZKY, ARYEH (1949). On the strong stability of a sequence of events. *Ann. Math. Statist.* **20** 296–299.
- [4] ERDŐS, PAUL (1949). On a theorem of Hsu and Robbins. *Ann. Math. Statist.* **20** 286–291.
- [5] FELLER, WILLIAM (1969). One sided analogs of Karamata's regular variation. *Enseignement math.* **15** 107–121.
- [6] FRANCK, W. E. and D. L. HANSON (1966). Some results giving rates of convergence in the law of large numbers. *Trans. Amer. Math. Soc.* **124** 347–359.
- [7] HANSON, D. L. (1970). Some results on convergence rates for weighted averages. In *Contributions to Ergodic Theory and Probability, Lecture Notes 160*. Springer-Verlag, Berlin, Heidelberg, New York.
- [8] HEATHCOTE, C. R. (1967). Complete exponential convergence and related topics. *J. Appl. Probability* **4** 217–256.
- [9] HSU, P. L. and H. ROBBINS (1947). Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. U.S.A.* **33** 25–31.
- [10] KATZ, MELVIN L. (1963). The probability in the tail of a distribution. *Ann. Math. Statist.* **34** 312–318.
- [11] KURTZ, THOMAS G. (1972). Inequalities for the law of large numbers. *Ann. Math. Statist.* **43** 1874–1883.
- [12] SMYTHE, R. T. (1974). Sums of independent random variables on partially ordered sets. *Ann. Probability* **2** 906–917.
- [13] SPITZER, FRANK (1956). A combinatorial lemma and its applications to probability theory. *Trans. Amer. Math. Soc.* **82** 323–339.

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN
COPENHAGEN, DENMARK

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706