WEAK LAWS OF LARGE NUMBERS FOR DEPENDENT SUMMANDS

by Z. GOVINDARAJULU, *Department of Statistics, University of Kentucky, Lexington, Ky* 40506, *U.S.A.*

(Communicated by S. M. Shah, F.N.A.)

(*Received* 3 *February* 1973)

Using the method of operators, the author obtains certain weak laws of large numbers for dependent random variables. These results constitute generalizations of an earlier paper by the author (Govindarajulu 1970).

1. INTRODUCTION

In an earlier paper, Govindarajulu (1970) has proved certain weak laws of large numbers using techniques analogous to those of Trotter (1959). Since the summands are assumed to be independent, weak convergence implies convergence almost everywhere (see Loeve 1960, p. 249). In this paper we shall extend the results of the earlier paper (Govindarajulu 1970) for dependent summands.

2. NOTATION

Let $(X_{1n}, .., X_{nn})$, $n = 1, 2, \ldots$ be a sequence of random variables such that $F^{(n)}(x_1, \ldots, x_n)$ denote the joint distribution function of $X_{1n}, \quad X_{nn}$ having marginals $F_{in}(x)$ $(i = 1, \ldots, n)$. Let

$$\alpha_{in} = \int_{|x| \leqslant b_n} x dF_{in}(x), \quad i = 1, \ldots, n$$

where $b_n \to \infty$ such that $b_n/n < K$, where K denotes a finite generic constant.

Also, let $S_n = \sum_{i=1}^{n} Z_{in}$, $Z_{in} = (X_{in} - \alpha_{in})/b_n$. Further $\{Y_i\}$ be a sequence of independent random variables that are degenerate at zero. Then we have the results given in § 3.

3. Main Results

Theorem 3.1—With the above notation, S_n converges to zero in probability if

$$\text{(i)} \quad \sum_{i=1}^{n} \int_{|x| > b_n} dF_{in}(x) \to 0 \text{ and}$$

$$\text{(ii)} \quad b_n^{-1} \sum_{i=1}^{n} \int_{|x| \leqslant b_n} |x - \alpha_{in}| \, dF_{in}(x) \to 0$$

as $n \to \infty$.

Proof : It suffices to show that $\| T_{S_n} h - T_Y h \| \to 0$ for every $h \in C_2$ where $(T_X h)(y) = Eh(X+y)$ for every real y and $\| h \| = \sup_x | h(x) |$. Now, let τ_i be the operator associated with Y_i $(i = 1, \ldots, n)$ and let

$$(T_i^* h)(y) = E\{h(Z_{in} + y) \mid X_{1n}, \ldots, X_{i-1,n}\} \quad i = 1, \ldots, n.$$

Then one can write

$$T_{S_n} h - T_Y h = \sum_{i=1}^{n} T_1^* \ldots T_{i-1}^* (T_i^* - \tau_i) \tau_{i+1} \ldots \tau_n h$$

$$= \sum_{i=1}^{n} \tau_{i+1} \ldots \tau_n T_1^* T_2^* \ldots T_{i-1}^* (T_i^* - \tau_i) h$$

since the Y_i are mutually independent and without loss of generality can be chosen to be independent of $X_{1n}, \ldots, X_{nn}$. So consider

$$T_1^* T_2^* \ldots T_{i-1}^* (T_i^* - \tau_i) h$$

$$= \int_{-\infty}^{\infty} \cdots \int \Big[\int_{-\infty}^{\infty} \Big\{ h\Big(\sum_{j=1}^{i} Z_{jn} + y \Big) - h\Big(\sum_{j=1}^{i-1} Z_{jn} + y \Big) \Big\}$$

$$\times dF_i(x_i | X_{1n}, \ldots, X_{i-1,n}) \Big] dF^{(i-1)}(x_1, \ldots, x_{i-1})$$

$$= \int_{-\infty}^{\infty} \Big[\int \overset{\infty}{\underset{-\infty}{\cdots}} \int \Big\{ h\Big(\sum_{j=1}^{i} Z_{jn} + y\Big) - h\Big(\sum_{j=1}^{i-1} Z_{jn} + y\Big)\Big\}$$

$$\times dF(x_1, \ldots, x_{i-1} X_{in} \mid x_i)\Big] \, dF_i(x_i)$$

$$= \int_{|x_i| \leqslant b_n} [\quad] \, dF_i(x_i) + \int_{|x_i| > b_n} [\quad] \, dF_i(x_i).$$

However,

$$\Big\| \int_{|x_i| > b_n} [\quad] \, dF_i(x_i) \Big\| \leqslant 2\|h\| P(|X_{in}| > b_n).$$

Consider

$$\int_{|x_i| \leqslant b_n} \int \overset{\infty}{\underset{-\infty}{\cdots}} \int \Big\{ h\Big(\sum_{j=1}^{i} Z_{jn} + y\Big) - h\Big(\sum_{j=1}^{i-1} Z_{jn} + y\Big)\Big\} dF(x_1, \ldots, x_i)$$

$$= b_n^{-1} \int_{|x_i| \leqslant b_n} \int \overset{\infty}{\underset{-\infty}{\cdots}} \int h' \Big(\sum_{j=1}^{i-1} Z_{jn} + y\Big) (x_i - \alpha_{in}) \, dF(x_1, \ldots, x_i)$$

$$+ (2b_n^2)^{-1} \int_{|x_i| \leqslant b_n} \int \overset{\infty}{\underset{-\infty}{\cdots}} \int h''(\xi) \, (x_i - \alpha_{in})^2 \, dF(x_1, \ldots, x_i)$$

where ξ lies between $\sum_{j=1}^{i-1} Z_{jn} + y$ and $\sum_{j=1}^{i} Z_{jn} + y$, after expanding

$h\Big(\sum_{j=1}^{i} Z_{jn} + y\Big)$ in Taylor series about $\sum_{j=1}^{i-1} Z_{jn} + y$. Thus,

$$\Big\| \int_{|x_i| \leqslant b_n} \int \overset{\infty}{\underset{-\infty}{\cdots}} \int [\quad] \, dF(x_1, \ldots, x_i) \Big\| \leqslant \frac{\|h'\|}{b_n}$$

$$\int_{|x| \leqslant b_n} |x_i - \alpha_{in}| dF_i(x) + \frac{\|h''\|}{2b_n^2} \int_{|x| \leqslant b_n} (x - \alpha_{in})^2 \, dF_i(x).$$

Now,

$$\|T_{S_n}h - T_Y h\| \leqslant \sum_{i=1}^{n} \|T_1^* \ldots T_{i-1}^*(T_i^* - \tau_i)h\|$$

$$< 2\|h\| \sum_{i=1}^{n} P(|X_{in}| > b_n)$$

$$+ \frac{\|h'\|}{b_n} \sum_{i=1}^{n} \int_{|x| \leqslant b_n} |x - \alpha_{in}| dF_i(x)$$

$$+ \frac{\|h''\|}{2b^2_n} \int_{|x| \leqslant b_n} (x - \alpha_{in})^2 \, dF_i(x).$$

However

$$b_n^{-1} \int_{|x| \leqslant b_n} (x - \alpha_{in})^2 dF_i(x) \leqslant 2 \int_{|x| \leqslant b_n} |x - \alpha_{in}| dF_i(x).$$

Consequently

$$\|T_{S_n} h - T_Y h\| \leqslant K \sum_{i=1}^{n} P(|X_{in}| > b_n) + \frac{K}{b_n} \sum_{i=1}^{n} \int_{|x| \leqslant b_n} |x - \alpha_{in}| dF_i(x) \to 0$$

due to the hypothesis of the theorem.

Next, let

$$\alpha_{in}^* = \int_{|x| \leqslant b_n} x dF_i(x | X_{1n}, \ldots, X_{i-1,n}), \; i = 1, \ldots, n$$

$$A_i(t) = t[1 - F_i(t) + F_i(-t)] \text{ and } B_n(t) = n^{-1} \sum_{i=1}^{n} A_i(t).$$

Also, let

$$S_n^* = \sum_{i=1}^{n} Z_{in}^*, \quad Z_{in}^* = (X_{in} - \alpha_{in}^*)/b_n.$$

Then, we have the following theorem.

Theorem 3.2—With the above notation, S_n^* converges to zero in probability if $B_n(t) \to 0$ as $n, t \to \infty$.

PROOF : One can write

$$T_1^* \cdot T_{i-1}^* (T_i^* - \tau_i)h = \int_{-\infty}^{\infty} \cdots \int \left[\int_{|x_i| \leqslant b_n} \left\{ h\left(\sum_{j=1}^{i} Z_{jn} + y \right) - h\left(\sum_{j=1}^{i-1} Z_{jn} + y \right) \right\} dF_i(x_i | X_{1n}, \ldots, x_{i-1,n}) \right] dF(x_1, \ldots, x_{i-1})$$

$$+ \int_{-\infty}^{\infty} \cdots \int \left[\int_{|x| > b_n} \{ \ \} \right] = \int_{-\infty}^{\infty} \cdots \int \left[\int_{|x_i| \leqslant b_n} \left\{ h'\left(\sum_{j=1}^{i-1} Z_{jn} + y \right) Z_{in} + \frac{Z^2_{in}}{2} h''(\xi) \right\} dF_i(x_i | X_{1n}, \ldots, X_{i-1,n}) \right] dF(x_1, \ldots, x_{i-1})$$

$$+ \int_{-\infty}^{\infty} \cdots \int \left[\int_{|x_i| > b_n} \{ \ \} \right].$$

Consider

$$\left\| \int_{-\infty}^{\infty} \cdots \int \int_{|x_i| > b_n} \{ \qquad \} dF_i(x_i | \ldots) \, dF(x_1, \ldots, x_{i-1}) \right\|$$

$$\leqslant 2 \|h\| \int_{|x_i| > b_n} \int_{-\infty}^{\infty} \cdots \int dF_i(x_1, \ldots, x_i) = 2KP[|X_{in}| > b_n].$$

$$b^{-1}_n \int_{-\infty}^{\infty} \int \left[\int_{|x_i| \leqslant b_n} h' \left\{ \sum_{j=1}^{i-1} Z_{jn} + y \right\} (x_i - \alpha_{in}^*) dF_i(x_i | \ldots) dF(x_1, \ldots, x_{i-1}) \right.$$

$$= -b_n^{-1} \int_{-\infty}^{\infty} \cdots \int h' \left(\sum_{j=1}^{i-1} Z_{jn} + y \right) \alpha_{in}^* \left[\int_{|x_i| > b_n} dF_i(x_i | \ldots) \right] dF(x_1, \ldots, x_{i-1})$$

$$\|LHS\| \leqslant \|h'\| \int_{|x_i|>b_n} \int_{-\infty}^{\infty} \int dF(x_1, \ldots, x_i)$$

$$= \|h'\| P[|X_{in}| > b_n],$$

and $\quad P[|X_{in}| > b_n] \leqslant b_n^{-1} A_i(b_n).$

Next, consider

$$\| \int_{-\infty}^{\infty} \cdots \int \int_{|x_i| \leqslant b_n} \frac{Z_{in}^2}{2} h''(\xi) dF(x_1, \ldots, x_i)\| \leqslant \frac{\|h''\|}{2b_n^2}$$

$$\int_{-\infty}^{\infty} \cdots \int \int_{|x_i| \leqslant b_n} (x_i - \alpha_{in}^*)^2 \, dF_i(x_i)| \ldots) dF(x_1, \ldots, x_{i-1})$$

$$\leqslant \frac{\|h''\|}{b_n^2} \int_{|x| \leqslant b_n} x^2 dF_i(x),$$

since $\alpha_{in}^{*2} \leqslant \int_{|x_i \leqslant |b_n} x^2_i dF_i(x_i | X_{1n}, \ldots, X_{i-1,n}).$

Also

$$\int_{|x| \leqslant b_n} x^2 dF_i(x) < b_n A_i(b_n) + 2 \int_0^{b_n} A_i(x) \, dx$$

after integrating by parts once. Now

$$\int_0^{b_n} A_i(x) dx = \int_0^{t_0} A_i(x) dx + \int_{t_0}^{b_n} A_i(x) \, dx.$$

Thus

$$\|T_{S_n} h - T_Y h\| < K \sum_{i=1}^{n} P[|X_{in}| > b_n]$$

$$+ b_n^{-1} \sum_{i=1}^{n} A_i(b_n) + t_0^2 \frac{n}{b_n^2} + \frac{2n}{b_n^2} \int_{t_0}^{b_n} B_n(t) \, dt$$

$$\leqslant KB_n(b_n) + \frac{K}{b_n} + K(b_n - t_0)\delta / b_n$$

$$\to 0 \text{ as } n, t \to \infty.$$

This completes the proof of Theorem 3.2.

Remark 3.1 : If $\mu_i = \int x \, dF_i(x)$ (or $\mu_i^* = \int x dF_i(x | X_{1n}, \quad , X_{i-1,n})$) exist for $k = 1, \ldots, n$, then in Theorem 3.1 (or Theorem 3.2) α_{in} (or α_{in}^*) can be replaced by μ_i (or μ_i^*) provided

$$\sum_{i=1}^{n} \frac{\mu_i - \alpha_{in}}{b_n} = \sum_{i=1}^{n} \int_{|x| > b_n} x dF_i(x) \left(\text{or} \sum_{i=1}^{n} \int_{|x| > b_n} x dF_i(x | X_{1n}, \ldots, X_{i-1,n}) \right)$$

tends to zero as $n \to \infty$, which is satisfied when $F_{1n} = \ldots = F_{nn}$.

Remark 3.2: Using the same expansions as in the proof of Theorem 3.2, we see that S^*_n converges to zero in probability if

$$(i) \quad \sum_{i=1}^{n} \int_{|x| > b_n} dF_i \to 0$$

and

$$(ii) \quad \sum_{i=1}^{n} b_n^{-2} \int_{|x| \leqslant b_n} (x - \alpha_{in}^*)^2 dF_i(x) \to 0$$

as $n \to \infty$.

4. Case of Random Sample Size

Random sums based on a random number of variables commonly arise in statistics. Towards this we have the following result.

Theorem 4.1—Let $\{X_{1n}, X_{2n}, \ldots, X_{nn}\}$, $n = 1,2,\ldots$ be a sequence of random vectors such that $F^{(n)}(x_1, \ldots, x_n)$ denotes the joint distribution function of $X_{1n}, \ldots, X_{nn}$ having marginals $F_{in}(x)$ $(i = 1, \ldots, n)$. Let $N_1, N_2, \ldots$ be a sequence of positive integer-valued random variables. Define $S_{N_n} = \sum_{1}^{N_n} Z_{i,n}$ where $Z_{i,n} = (X_{in} - \alpha_{in})/b_n [X_{in} - \alpha_{in}^*)/b_n]$ and $\alpha_{in}[\alpha_{i,n}^*]$ and b_n are as defined in section 2 [Theorem 3.2]. Further, suppose that for any $\delta > 0$ there exists $n_0(\delta)$ and con stants $0 < a(\delta) < b(\delta) < \infty$ such that

$P(a < N_n/n < b) \geqslant 1-\delta$ for $n > n_0(\delta)$. Then S_{N_n} converges to zero in probability if either conditions (*i*) and (*ii*) of Theorem 3.1 hold [or $B_n(t)$ defined in Theorem 3.2 tends to zero as $n, t \to \infty$].

PROOF : We shall provide a slightly different proof than the one given by Govindarajulu (1970). It suffices to show

$$\|T_{S_{N_n}} h - T_Y h\| \to 0 \text{ as } n \to \infty.$$

Let $c_{k,n} = P(N_n = k)$, $k = 1, 2, \ldots$

Then,

$$T_{S_{N_n}} h - T_Y h = \Sigma c_{k,n} [T_{S_k} h - T_Y h]$$

Hence

$$\|T_{S_{N_n}} h - T_Y h\| = \sum_{k=[n a]}^{[n b]} c_{k,n} \|T_{S_k} h - T_Y h\| + 2 \|h\| \delta.$$

Now for every δ there exists a k_0 such that

$$\| T_{S_k} h - T_Y h\| \leqslant K \delta$$

provided the hypothesis of either Theorem 3.1 or Theorem 3.2 holds with n replaced by k.

Now choose $n_0' = [k_0/a] + 1$ and $n_0^* = \max (n_0', n_0)$. Then $n > n_0^*$ implies that $k > k_0$ which in turn implies that

$$\|T_{S_{N_n}} h - T_Y h\| \leqslant K\delta + 2 \| h \| \delta = K\delta.$$

5. THE MULTIVARIATE CASE

If the random variables are vector-valued the arguments employed in section 3 can be repeated verbatum in order to prove the multivariate versions of Theorems 3.1, 3.2 and 4.1, which are easy to formulate (for instance, see Govindarajulu 1970). Further, the multivariate case can be reduced to the univariate case by considering arbitrary linear combinations of the components.

6. AN EXAMPLE

Let $Y_1, Y_2, \ldots$ be an independent sequence of random variables having $G_k(x)$ for their marginals where

$$G_k(x) = \begin{cases} 0 & \text{if } x < -(k+1) \\ c_k\left\{\dfrac{x^{-1}}{\log 2|x|} - \dfrac{(k+1)^{-1}}{\log 2(k+1)}\right\} & \text{if } -(k+1) \leqslant x < -1 \\ 1/2 & \text{if } -1 \leqslant x < 1 \\ \frac{1}{2} + c_k\left\{\dfrac{1}{\log 2} - \dfrac{x^{-1}}{\log 2x}\right\} & \text{if } 1 \leqslant x < k+1 \\ 1 & \text{if } x \geqslant k+1 \end{cases}$$

where $c_k = (1/2)\left\{\dfrac{1}{\log 2} - \dfrac{(k+1)^{-1}}{\log 2(k+1)}\right\}, k = 1, 2, \ldots$

Let $X_k = Y_k - Y_{k-1}, k = 1, \ldots, n$ with $Y_0 \equiv 0$. Also, let F_k denote the marginal distribution of X_k $(k = 1, \ldots, n)$. Since

$$x[1 - G_k(x)] \leqslant \frac{\log 2}{2}\left\{\frac{x(k+1)^{-1}}{\log 2(k+1)} - \frac{1}{\log 2x}\right\}$$

we have

$$B_n(x) = 2n^{-1}\sum_{k=1}^{n} x[1 - G_k(x)]$$

$$\leqslant \frac{x}{n}\log 2[\log\ \log\ 2(n+1) - \log\ \log 2] - \frac{\log 2}{\log 2x}$$

which tends to zero as $n, x \to \infty$. Now consider

$$B_n(x) = \frac{1}{n}\sum_{k=1}^{n} x[1 - F_k(x) + F_k(-x)] \leqslant \frac{1}{n}\sum_{k=1}^{n} x\,P[\,|Y_i - Y_{i-1}| \geqslant x]$$

$$\leqslant \frac{1}{n}\sum x\{P[\,|Y_i| \geqslant x/2] + P[\,|Y_{i-1}| \geqslant x/2]\}$$

$$\leqslant \frac{2}{n}\sum_{k=1}^{n} x\,P[\,|Y_i| \geqslant x/2]$$

$$= 4\,B_n(x) \to 0 \text{ as } n, x \to \infty.$$

Thus the condition of Theorem 3.2 is satisfied. Hence $\sum_{k=1}^{n} (X_k - \alpha_{kn}^*)/b_n$ converges to zero in probability where we take $b_n = n + 1$. In order to compute the $\alpha_{k,n}^*$ we need the conditional distributions. Consider

$$\begin{aligned} F_k^*(x) &= P[X_k \leqslant x \mid X_1 = x_1, \ldots, X_{k-1} = x_{k-1}] \\ &= P[Y_k - Y_{k-1} \leqslant x \mid Y_{k-1} = y], \; y = x_1 + \ldots + x_{k-1} \\ &= P[Y_k \leqslant x + y] = G_k(x + y), \; k = 1, \ldots, n. \end{aligned}$$

Hence

$$\alpha_{k,n}^* = \int_{|x|<n+1} x dF_k^*(x) = \int_{|x|<n+1} x dG_k(x+y)$$

$$= \int_{y-n-1}^{y+n+1} (z-y)\, dG_k(z)$$

$$= \int_{y-n-1}^{n+1+y} z\, dG_k(z) - y\,[G_k(y+n+1) - G_k(y-n-1)].$$

Now, it is straight forward to compute $\alpha_{k,n}^*$ when either (i) $-(k+1) \leqslant Y_{k-1} < -1$ or (ii) $1 \leqslant Y_{k-1} < k+1$.

REFERENCES

Govindarajulu, Z. (1970). On weak laws of large numbers. *Proc. Indian Acad. Sci.*, **70**, 266-74.

Loeve, M. (1960). Probability Theory. Van Nostrand Co., New York, p. 249.

Trotter, H. F. (1959). An elementary proof of the central limit theorem. *Archiv. Math.*, **10**, 226-34.