ITERATED LIMITS AND THE CENTRAL LIMIT THEOREM FOR DEPENDENT VARIABLES

by

GEORGE MARSAGLIA

Special report to the Office of Naval Research of work at Chapel Hill under Project NR 042 031 for research in probability and statistics.

Institute of Statistics
Mimeograph Series No. 93
February 1954

ITERATED LIMITS AND THE CENTRAL LIMIT THEOREM

FOR DEPENDENT VARIABLES.[1]

by
George Marsaglia
University of North Carolina

1. Introduction. Section 2 of this paper gives some results on iterated limits which may be considered generalizations of well-known results [1, p. 254]. Section 3 applies these results to give easy proofs of some central limit theorems for m-dependent variables.

2. Iterated Probability Limits. Here, we use the strong sense of an iterated limit: for constants a_{ij}, $i, j = 1, 2, \ldots$, $\lim_j \lim_i a_{ij} = a$ means

$$\lim_{j \to \infty} \left(\overline{\lim_{i \to \infty}} \; |a_{ij} - a| \right) = 0. \tag{1}$$

We note that (1) holds if, and only if, for each $\epsilon > 0$ there exist integers $M, N_1, N_2, \ldots$ such that if the pair (i,j) satisfies $j > M$, $i > N_j$, then $|a_{ij} - a| < \epsilon$.

DEFINITION 1. Let f, f_{ij}, $i, j = 1, 2, \ldots$ be random variables. Then

$$\operatorname{plim}_j \operatorname{plim}_i f_{ij} = f$$

means, for every $\epsilon > 0$,

$$\lim_j \lim_i P(|f_{ij} - f| > \epsilon) = 0$$

1. Work sponsored by the Office of Naval Research under Contract No. N7-onr-28402.

THEOREM 1. *Let* h_{ij}, g_{ij}, $i, j = 1, 2, \ldots$ *be random variables. Let* G *be a function such that at each of its continuity points* x

$$\lim_j \lim_i P(g_{ij} \leq x) = G(x),$$

and suppose

$$\operatorname{plim}_j \operatorname{plim}_i h_{ij} = 0.$$

Then

$$\lim_j \lim_i P(g_{ij} + h_{ij} \leq x) = G(x).$$

Let $\epsilon = 8\partial > 0$ and a continuity point x of G be given. We shall exhibit integers M, N_1, N_2, ... such that

$$|P(g_{ij} + h_{ij} \leq x) - G(x)| < \epsilon \tag{2}$$

if $j > M$ and $i > N_j$. First, choose β so that G is continuous at $x + \beta$, at $x - \beta$, and so that

$$|G(x + \beta) - G(x - \beta)| < \partial. \tag{3}$$

Then choose M, N_1, N_2, ... so that, simultaneously,

$$P(|h_{ij}| > \beta) < \partial \tag{4}$$

$$|P(g_{ij} \leq x) - G(x)| < \partial \tag{5}$$

$$|P(g_{ij} \leq x - \beta) - G(x - \beta)| < \partial \tag{6}$$

$$|P(g_{ij} \leq x + \beta) - G(x + \beta)| < \partial \tag{7}$$

whenever (i,j) satisfies $j > M$ and $i > N_j$. Then for such a pair (i,j), let $F(x) = P(g_{ij} + h_{ij} \leq x)$, $H(x,\beta) = P(g_{ij} + h_{ij} \leq x, |h_{ij}| \leq \beta)$, $L(x,\beta) = P(g_{ij} \leq x, |h_{ij}| \leq \beta)$ and $Q(x) = P(g_{ij} \leq x)$. We have

$$|F(x)-G(x)| \leq |F(x)-H(x,\beta)| + |H(x,\beta)-L(x,\beta)| + |L(x,\beta)-Q(x)| + |Q(x)-G(x)|.$$

Now by (4) and (5), each of the terms on the right except the second is bounded by δ, and since $L(x-\beta,\beta) \leq H(x,\beta) \leq L(x+\beta,\beta)$ and $L(x-\beta,\beta) \leq L(x,\beta) \leq L(x+\beta,\beta)$,

$$|H(x,\beta) - L(x,\beta)| \leq |L(x+\beta,\beta) - L(x-\beta,\beta)| < 5\delta$$

by (4), (7), (3) and (6), since

$$|L(x+\beta,\beta)-L(x-\beta,\beta)| \leq |L(x+\beta,\beta)-Q(x+\beta)| + |Q(x+\beta)-G(x+\beta)| + |G(x+\beta)-G(x-\beta)|$$
$$+ |G(x-\beta)-Q(x-\beta)| + |Q(x-\beta)-L(x-\beta,\beta)|$$

Hence $|F(x)-G(x)| < 8\delta = \epsilon$, which is condition (2).

THEOREM 2. <u>Under the conditions of Theorem</u> 1, <u>if there exist constants</u> a_{ij} <u>such that</u> $\lim_j \lim_i a_{ij} = a > 0$, <u>and if</u> G <u>is continuous</u> at x/a then $\lim_j \lim_i P(a_{ij}g_{ij} \leq x) = G(x/a)$.

Using the artifice, for suitable i, j, γ,

$$|P(g_{ij} \leq \frac{x}{a_{ij}}) - G(\frac{x}{a})| \leq |P(g_{ij} \leq \frac{x}{a_{ij}}) - P(g_{ij} \leq \frac{x}{a})| + |P(g_{ij} \leq \frac{x}{a}) - G(\frac{x}{a})|$$

$$|P(g_{ij} \leq \frac{x}{a_{ij}}) - P(g_{ij} \leq \frac{x}{a})| \leq |P(g_{ij} \leq \frac{x}{a-\gamma}) - P(g_{ij} \leq \frac{x}{a+\gamma})| ,$$

the proof is routine. The details are omitted.

3. Applications to partitioned sequences of m-dependent random variables. Let $x_1, x_2, \ldots$ be an m-dependent sequence of random variables with zero means. For each pair (n, k) with $2m < k \leq n$, define

$$y_i = x_{ik-k+1} + \ldots + x_{ik-m} \qquad i = 1, 2, \ldots$$

$$g_{nk} = \sum_{1}^{\left[\frac{n}{k}\right]} y_i, \qquad t_{nk}^2 = E(g_{nk}^2)$$

$$s_n^2 = E(x_1 + \ldots + x_n)^2, \quad h_{nk} = \frac{1}{s_n}\left(\sum_{1}^{n} x_i - g_{nk}\right) \quad .$$

Since we shall be dealing with $\lim_k \lim_n$ relations, g_{nk}, $n < k$, may be defined indifferently.

According to Theorems 1 and 2, if

(8) $$\lim_k \lim_n \frac{t_{nk}}{s_n} = 1$$

(9) $$\operatorname{plim}_k \operatorname{plim}_n h_{nk} = 0$$

(10) $$\lim \lim P\left(\frac{g_{nk}}{t_{nk}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt$$

then

(11) $$\lim_n P\left(\frac{x_1 + \ldots x_n}{s_n} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt.$$

The following theorems give conditions which imply (8), (9), and (10).

THEOREM 3. <u>If there exist constants</u> $\alpha > 2$, $B > 0$ <u>such that</u>

(12) $$\frac{n}{s_n^2} < B, \qquad n = 1, 2, \ldots$$

(13) $$E(x_n^2) < B, \qquad n = 1, 2, \ldots$$

(14) $$\lim_n \frac{\left(\sum_1^n E(|x_i|^{\alpha})\right)^{1/\alpha}}{s_n} = 0$$

<u>then condition</u> (11) <u>holds.</u>

We first establish (8) and (9). One readily finds, for $2m < k \leq n$,

(15) $$|s_n^2 - t_{nk}^2| < \left(\left[\frac{n}{k}\right] + k^2\right)8m^2B \quad .$$

and

(16) $$E(h_{nk}^2) < \frac{1}{s_n^2}\left(\left[\frac{n}{k}\right] + k^2\right)8m^2B \quad .$$

But, using (12),

(17) $$\lim_k \lim_n \frac{\left[\frac{n}{k}\right] + k^2}{s_n^2} = \lim_k \lim_n \frac{\left[\frac{n}{k}\right]}{s_n^2} = \lim\left(\frac{1}{k}\,\overline{\lim}\,\frac{n}{s_n^2}\right) = 0.$$

Relations (15), (16) and (17) imply (8) and (9).

Condition (10) will be true, by Liapounoff's Theorem [4, p. 284] if, for large k,

$$\lim_{n} \frac{\left(\sum_{i=1}^{\left[\frac{k}{n}\right]} E(|y_i|^\alpha) \right)^{1/\alpha}}{t_{nk}} = 0.$$

Now $E(|y_i|^\alpha) \leq k^\alpha \sum_{j=ik-k+1}^{ik-m} E(|x_j|^\alpha)$, so that

$$\lim_{n} \frac{\left(\sum_{i=1}^{\left[\frac{k}{n}\right]} E(|y_i|^\alpha) \right)^{1/\alpha}}{t_{nk}} \leq \lim_{n} \frac{k \left(\sum_{i=1}^{n} E(|x_i|^\alpha) \right)^{1/\alpha}}{s_n} \cdot \frac{s_n}{t_{nk}} = 0.$$

by (14) and (8), if k is large.

THEOREM 4. _If_ $x_1, x_2, \ldots$ _is a stationary m-dependent sequence with zero means, then_ (11) _holds_.

For in that case, (12) holds, and, since the variances are bounded, (8) and (9) are established as above. (10) holds, since, for each $k > 2m$, the sequence $y_1, y_2, \ldots$ is stationary and independent.

REFERENCES

1. Cramer, H., Mathematical Methods of Statistics, Princeton, 1946.

2. Diananda, P. H., "Some Probability Limits with Statistical Applications," Proc. Camb. Phil. Soc. (1953), pp. 239-246.

3. Hoeffding, W. and Robbins, H. E., "The Central Limit Theorem for Dependent Random Variables," Duke Math. Jr., (1948), pp. 773-780.

4. Uspensky, J. U., Introduction to Mathematical Probability, McGraw Hill, 1937.