The Annals of Mathematical Statistics
1969, Vol. 40, No. 3, 935–941

ON LIMITING DISTRIBUTIONS FOR SUMS OF A RANDOM NUMBER OF INDEPENDENT RANDOM VECTORS[1]

By Leon Jay Gleser

The Johns Hopkins University

1. Introduction. Consider a sequence of $p \times 1$ random (column) vectors $\{y_n\}$, $n = 1, 2, \cdots$. Suppose that there exists a sequence $\{B_n\}$ of real $p \times p$ nonsingular matrices and a proper p-variate distribution function $F(y)$ such that

$$\lim_{n\to\infty} \mathcal{L}(B_n^{-1} y_n) = \mathcal{L}(y^*), \tag{1.1}$$

where y^* is a $p \times 1$ random vector having the distribution function $F(y)$. (The notation $\mathcal{L}(y^*)$ denotes the law or distribution of y^*. $\text{Lim}_{n\to\infty} \mathcal{L}(Z_n) = \mathcal{L}(Z)$ means that Z_n converges in law (converges weakly) to Z. The notation $\mathcal{L}(\mathfrak{N}(0, \sigma^2 I))$ used later is short for the law of a multivariate normal random variable with mean vector 0 and covariance matrix $\sigma^2 I$.) Suppose further that we have an infinite sequence $\{\nu_n\}$, $n = 1, 2, \cdots$, of positive integer-valued random variables, and a sequence $\{k_n\}$ of positive integers such that

$$\lim_{n\to\infty} k_n = \infty, \qquad \text{plim}_{n\to\infty} k_n^{-1} \nu_n = 1. \tag{1.2}$$

We are interested in conditions under which

$$\lim_{n\to\infty} \mathcal{L}(B_{k_n}^{-1} y_{\nu_n}) = \mathcal{L}(y^*). \tag{1.3}$$

In the scalar case ($p = 1$), Anscombe [2] found a sufficient condition for (1.3) to hold. One extension of that theorem (Theorem 1 of [2]) to the vector case ($p > 1$) is the following.

Theorem 1.1. *If the sequences $\{y_n\}$, $\{B_n\}$, $\{\nu_n\}$, and $\{k_n\}$ satisfy* (1.1) *and* (1.2), *then for* (1.3) *to hold, it is sufficient that for given $\epsilon > 0$, $\eta > 0$, there exists a positive integer n_0 and a positive number c such that for all $n \geq n_0$,*

$$P\{\max_{n' : |n - n'| < cn} \|B_n^{-1}(y_n - y_{n'})\|_2 < \epsilon\} > 1 - \eta. \tag{1.4}$$

Here, for a $p \times 1$ vector $Z = (Z_1, Z_2, \cdots, Z_p)'$, the notation $\|Z\|_2$ represents the L_2 norm of Z, i.e., $\|Z\|_2 = (Z'Z)^{\frac{1}{2}}$. *The notation $\|Z\|_\infty$ is used to represent the L_∞ norm of Z,* i.e., $\|Z\|_\infty = \max_{1 \leq j \leq p} |Z_j|$.

Note. We note that nothing is supposed concerning the dependence of ν_n on the random vectors y_k.

Theorem 1.1 is proven in Section 2. The proof closely resembles that given by Anscombe [2] in the scalar case, and consequently is only briefly sketched.

Received 12 December 1968.

[1] This research was sponsored in part by the Office of Naval Research, under Contract NONR 4010(09) awarded to the Department of Statistics, The Johns Hopkins University. This paper, in whole or in part, may be reproduced for any purpose of the United States Government.

For the scalar case, Anscombe [2] conjectured that condition (1.4) is both necessary and sufficient (given the truth of (1.1) and (1.2)) for (1.3) to hold. When y_n is the sum $\sum_{i=1}^{n} x_i$ of independent (scalar) random variables x_i, Mogyoródi [5] found a condition equivalent to (1.4) (but more easily verified) which is both necessary and sufficient for (1.3). One extension of his result to the vector case is the following theorem, which is the major result of this paper.

THEOREM 1.2. *If* $\{y_n\}$, $\{B_n\}$ $\{\nu_n\}$, $\{k_n\}$ *satisfy* (1.1) *and* (1.2), *if* $y_n = \sum_{i=1}^{n} x_i$, $n = 1, 2, \cdots$, *is the sum of the independent random vectors* x_i, $i = 1, 2, \cdots$, *and if the random variables* $\|B_n^{-1}x_k\|_2$ *are infinitesimal,* $k = 1, 2, \cdots, n$ (c.f. [4]), *then* (1.3) *holds if and only if*

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \lambda_{k_n}\lambda_{[k_n(1+\delta)]}^{-1} = 1, \tag{1.5}$$

where $[a]$ *denotes the integer part of the real number* a *and*

$$\lambda_n \equiv \min \textit{root of } (B_nB_n')^{\frac{1}{2}}. \tag{1.6}$$

Mogyoródi [5] remarks that (1.5) is satisfied when $\lambda_n = n^{\alpha}L(n)$, $\alpha > 0$, where for all $c > 0$,

$$\lim_{n\to\infty} L([cn])/L(n) = 1.$$

In Section 3, we present the proof of Theorem 1.2. This vector extension of Mogyoródi's theorem [5] provides a convenient and less restrictive proof of a recent result of A. Albert [1]. This application of Theorem 1.2 appears in Section 4.

2. Proof of Theorem 1.1. Given $\epsilon > 0$, $\eta > 0$, let n_0 and c satisfy (1.4), and let n_0 also be large enough so that for any r for which $k_r > n_0$,

$$P\{|\nu_r - k_r| < ck_r\} > 1 - \eta. \tag{2.1}$$

Such a choice is clearly possible since $\{k_n\}$ and $\{\nu_n\}$ satisfy (1.2). For fixed r satisfying $k_r > n_0$, let E denote the event

$$E: |\nu_r - k_r| < ck_r \quad \text{and} \quad \|B_{k_r}^{-1}(y_{k_r} - y_{\nu_r})\|_\infty < \epsilon$$

and let $S(n)$, $S'(n)$, and T denote the events

$$S(n): \max_{n':|n'-n|<cn} \|B_n^{-1}(y_{n'} - y_n)\|_2 < \epsilon,$$

$$S'(n): \max_{n':|n'-n|<cn} \|B_n^{-1}(y_{n'} - y_n)\|_\infty < \epsilon,$$

$$T: |\nu_r - k_r| < ck_r.$$

Then $S \equiv S(k_r) \subset S'(k_r) \equiv S'$, and

$$P(E) \geqq P(S' \cap T) = P(S') - P(S' \cap T^c)$$
$$\geqq P(S') - P(T^c) \geqq P(S) - P(T^c),$$

where T^c denotes the complement of T. Consequently, we conclude from (1.4) and (2.1) that

$$P(E) > 1 - 2\eta. \tag{2.2}$$

Define the events

$$D = \{B_{k_r}^{-1} y_{\nu_r} \text{ componentwise} \leqq y\},$$
$$F = \{\|B_{k_r}^{-1}(y_{k_r} - y_{\nu_r})\|_\infty < \epsilon\},$$
$$R^+ = \{B_{k_r}^{-1} y_{k_r} \text{ componentwise} \leqq y + \epsilon\mathbf{1}\},$$
$$R^- = \{B_{k_r}^{-1} y_{k_r} \text{ componentwise} \leqq y - \epsilon\mathbf{1}\},$$

where $\mathbf{1}' = (1, 1, \cdots, 1)$. Noting that $E = T \cap F$ and that D is the union of $D \cap E$ and $D \cap E^c$, we find that

$$P(D) \leqq P(D \cap E) + P(E^c) \leqq P(D \cap F) + P(E^c) \leqq P(R^+) + P(E^c),$$

and

$$P(D) \geqq P(D \cap F) \geqq P(R^- \cap F) \geqq P(R^-) - P(F^c) \geqq P(R^-) - P(E^c).$$

Consequently, from (2.2) and the definitions of R^+ and R^-:

$$P\{B_{k_r}^{-1} y_{k_r} \text{ componentwise} \leqq y - \epsilon\mathbf{1}\} - 2\eta \leqq P(D)$$
$$\leqq P\{B_{k_r}^{-1} y_{k_r} \text{ componentwise} \leqq y + \epsilon\mathbf{1}\} + 2\eta.$$

Thus, if y is a continuity point of $F(y)$, convergence at y (and thus the result (1.3)) follows from the fact that p-variate cumulative distribution functions are componentwise monotonic and almost everywhere continuous. □

3. Proof of Theorem 1.2. The necessity of (1.1) and (1.5) for (1.3) can be shown by setting $B_n = \lambda_n I$ and then mimicking the proof [5] of Mogyoródi for the scalar case. To show the sufficiency of (1.1) and (1.5) for (1.3), we only need show that (1.5) implies (1.4), for then our result is a corollary of Theorem 1.1.

Let $y_n = \sum_{i=1}^n x_i$. Now

$$\max_{(n':|n'-n|<cn)} \|B_n^{-1}(y_{n'} - y_n)\|_2$$
$$\leqq \max_{(n':|n'-n|<cn)} \|y_{n'} - y_n\|_2 \lambda_n^{-1} \leqq \max_{(n':|n'-n|<cn)} \max_{1\leqq i \leqq p} p^{\frac{1}{2}} |y_{n'i} - y_{ni}| \lambda_n^{-1}$$
$$\leqq \max_{1\leqq i\leqq p} \max_{(n':|n'-n|<cn)} p^{\frac{1}{2}} \left|\sum_{k=1}^{n'} x_{ki} - \sum_{k=1}^{n} x_{ki}\right| \lambda_n^{-1},$$

where $y_{n'i}$, y_{ni}, x_{ki} denote the ith component of $y_{n'}$, y_n, and x_k, respectively. Let $S \equiv S(n)$ denote the set where $\max_{n':|n'-n|<cn} \|B_n^{-1}(y_{n'} - y_n)\|_2 < \epsilon$, and let

$$S_i \equiv S_i(n) = \{\max_{n':|n'-n|<cn} p^{\frac{1}{2}} \left|\sum_{k=1}^{n'} x_{ki} - \sum_{k=1}^{n} x_{ki}\right| \lambda_n^{-1} < \epsilon\}.$$

Then $P(S) \geqq P(\bigcap_{i=1}^p S_i) \geqq 1 - \sum_{i=1}^p P(S_i^c)$. In proving the scalar version of Theorem 1.2, Mogyoródi [5] showed that (1.5) was sufficient for showing that given $\epsilon > 0$ and $\eta' = \eta/p$, there is a c_i and a n_{0i} such that for any $n \geqq n_{0i}$,

$$P(S_i^*) > 1 - \eta'$$

where

$$S_i^* \equiv S_i^*(n) = \{\max_{n':|n'-n|<c_i n} p^{\frac{1}{2}} \left|\sum_{k=1}^{n'} x_{ki} - \sum_{k=1}^{n} x_{ki}\right| \lambda_n^{-1} < \epsilon\}.$$

Set $c = \min_{1 \leqq i \leqq p} c_i > 0$ and $n_0 = \max_{1 \leqq i \leqq p} n_{0i}$, and it follows that

$$P(S_i) \geqq P(S_i^*) > 1 - \eta' = 1 - \eta p^{-1}.$$

Thus for $n \geqq n_0$, $c = \min_{1 \leqq i \leqq p} c_i$,

$$P(S) \geqq 1 - \sum_{i=1}^{p} P(S_i^c) \geqq 1 - p(\eta p^{-1}) = 1 - \eta,$$

so that (1.2) holds. This proves the theorem.

Remark. It is worth noting the fact that (in the context of Theorem 1.2) not only do (1.1) and (1.5) together imply (1.4), but also (1.1) and (1.4) imply (1.5). The proof of the latter assertion closely parallels that of Mogyoródi [5] for the scalar case.

4. An application of Theorem 1.2. Consider a sequence $\{u_n\}$ of independent random variables (scalars) satisfying

$$u_i = \beta' g^{(i)} + e_i,$$

β a fixed $p \times 1$ vector, $g^{(i)}$ a known $p \times 1$ vector, e_i a random observation obeying distribution function $H(e)$ with mean 0 and finite variance σ^2, $0 < \sigma^2 < \infty$. Let

$$G_n = (g^{(1)}, g^{(2)}, \cdots, g^{(n)}): p \times n,$$

$n = p, p + 1, \cdots$, and assume that rank $G_n = p$. Define $T_n^2 = G_n G_n'$ and let $\lambda_n = (\text{min root } T_n^2)^{\frac{1}{2}}$. Finally, let

$$L_n = T_n^{-1} G_n = (l_n^{(1)}, l_n^{(2)}, \cdots, l_n^{(n)})$$

where $l_n^{(i)}: p \times 1$ is the ith column of L_n. The customary estimator of β based on n observations on the u_i is

$$\hat{\beta}(n) \equiv (G_n G_n')^{-1} G_n u(n) = T_n^{-1} L_n u(n)$$

where $u(n) = (u_1, u_2, \cdots, u_n)'$ is the observed sample. Gleser [3] has proved the following corollary to a theorem in [4], p. 103.

Theorem 4.1. *If*

$$\lim_{n\to\infty} \max_{1 \leqq i \leqq n, 1 \leqq j \leqq p} |l_{n \cdot j}^{(i)}| = 0 \tag{4.1}$$

where $l_{n \cdot j}^{(i)}$ *is the* jth *component of* $l_n^{(i)}$, *then*

$$\lim_{n\to\infty} \mathcal{L}(T_n(\hat{\beta}(n) - \beta) = \mathcal{L}(\mathfrak{N}(0, \sigma^2 I)). \tag{4.2}$$

Remark. It may be of practical use to point out that (4.1) is equivalent to

$$\lim_{n\to\infty} \max_{1 \leqq i \leqq n} g^{(i)\prime} (G_n G_n')^{-1} g^{(i)} = 0. \tag{4.1$'$}$$

This can be seen from the fact that

$$p^{-1} l_n^{(i)\prime} l_n^{(i)} \leqq (\max_{1 \leqq j \leqq p} |l_{n \cdot j}^{(i)}|)^2 \leqq l_n^{(i)\prime} l_n^{(i)},$$

and $l_n^{(i)\prime} l_n^{(i)} = g^{(i)\prime} (G_n G_n')^{-1} g^{(i)}$.

Now, consider a sequence of positive integer-valued random variables $\{\nu_n\}$,

$n = 1, 2, \cdots$, such that $\text{plim}_{n\to\infty} n^{-1}\nu_n = 1$. From Theorem 1.2, we can prove the following result.

THEOREM 4.2. *If* (4.1) *holds, then*

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \lambda_n \lambda_{[n(1+\delta)]}^{-1} = 1, \tag{4.3}$$

is necessary and sufficient for

$$\lim_{n\to\infty} \mathcal{L}(T_n(\hat{\beta}(\nu_n) - \beta)) = \mathcal{L}(\mathfrak{N}(0, \sigma^2 I)). \tag{4.4}$$

PROOF. Set $x_k = e_k g^{(k)}$, $B_n = T_n$, $\nu_n = \nu_n$, and $k_n = n$ in Theorem 1.2. Then

$$T_n(\hat{\beta}(n) - \beta) = T_n^{-1} \sum_{i=1}^n x_i,$$

so if $y_n = \sum_{i=1}^n x_i$, we have

$$T_n(\hat{\beta}(n) - \beta) = T_n^{-1} y_n .$$

Our problem is now in the form of Theorem 1.2, and (4.4) follows as a direct consequence of (4.3), Theorem 4.1, and Theorem 1.2.

Theorem 4.2 was implicitly proved by Albert [1] under the conditions:

$$\lim_{n\to\infty} \text{tr } (G_n G_n')^{-1} = 0, \tag{4.5}$$

$$\lim_{n\to\infty} \max_{1 \leq i \leq n} (g^{(i)\prime} g^{(i)}) \text{ tr } (G_n G_n')^{-1} = 0, \tag{4.6}$$

$$\limsup_{n\to\infty} (\text{max root } G_n G_n')\lambda_n^{-2} < \infty, \tag{4.7}$$

$$\lim_{\delta\to 0} \limsup_{n\to\infty} |\text{tr } (G_{[n(1+\delta)]} G'_{[n(1+\delta)]})[\text{tr } (G_n G_n')]^{-1} - 1| = 0. \tag{4.8}$$

Since

$$g^{(i)\prime}(G_n G_n')^{-1} g^{(i)} \leq g^{(i)\prime} g^{(i)} \text{ tr } (G_n G_n')^{-1},$$

it follows that (4.6) implies (4.1)$'$, and thus (4.6) implies (4.1). Now, Albert's proof [1] of (4.4) takes no cognizance of the special nature of his stopping rule ν_n. Thus, since given (4.1), (4.3) is necessary and sufficient for (4.4), we can expect that Albert's conditions [(4.5) through (4.8)] imply (4.3). To demonstrate that this assertion is correct, we first prove the following lemma.

LEMMA 4.3. *Assume that* $A: s \times s$ *is positive semi-definite and that* $B: s \times s$ *is either positive semi-definite or negative semi-definite* ($A \geq 0$, $B \geq 0$ *or* $B \leq 0$). *For any matrix* C, *let* $\lambda_{\max}(C) \equiv$ maximum root C *and* $\lambda_{\min}(C) =$ minimum root C. *Then*

$$\lambda_{\min}(A + B) \leq \lambda_{\min}(A) + |\text{tr } B|. \tag{4.9}$$

PROOF. Let $w_A : s \times 1$ be such that $\|w_A\|_2 = 1$ and $w_A' A w_A = \lambda_{\min}(A)$. Then

$$\begin{aligned} \lambda_{\min}(A + B) &= \min_{\|w\|_2 = 1} w'(A + B)w \\ &\leq w_A'(A + B)w_A = \lambda_{\min}(A) + w_A' B w_A \\ &\leq \lambda_{\min}(A) + |w_A' B w_A|. \end{aligned} \tag{4.10}$$

Since all of the roots of B have the same sign, and since $(w_A'w_A)^{\frac{1}{2}} = \|w_A\|_2 = 1$, it follows that $|w_A'Bw_A| \leqq |\text{tr } B|$. This fact, with (4.10), proves (4.9).

Let $n^* = [n(1 + \delta)]$. Applying Lemma 4.3 with $A = G_{n^*}G_{n^*}'$, $B = G_nG_n' - G_{n^*}G_{n^*}'$ ($G_nG_n' - G_{n^*}G_{n^*}'$ is $\geqq 0$ if $n^* \leqq n$, and is $\leqq 0$ if $n^* \geqq n$), we find that

$$\lambda_n^2 \leqq \lambda_{n^*}^2 + |\text{tr } (G_nG_n' - G_{n^*}G_{n^*}')| \tag{4.11}$$

and thus

$$\lambda_n^2\lambda_{n^*}^{-2} \leqq 1 + |(\text{tr } G_nG_n')(\text{tr } G_{n^*}G_{n^*}')^{-1} - 1|(\text{tr } G_{n^*}G_{n^*}')\lambda_{n^*}^{-2}. \tag{4.12}$$

From (4.12), it is clear that (4.7) and (4.8) imply (4.3).

The following example (due to M. S. Srivastava) demonstrates that (4.5) through (4.8) are not equivalent to (4.1) and (4.3). Consider

$$G_n = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 3 & & n \end{pmatrix}.$$

Then

$$G_nG_n' = n \begin{pmatrix} 1 & \frac{1}{2}(n+1) \\ \frac{1}{2}(n+1) & \frac{1}{6}(n+1)(2n+1) \end{pmatrix}$$

and

$$\begin{aligned}(\lambda_{\max}(G_nG_n'))\lambda_n^{-2} &= (\lambda_{\max}^2(G_nG_n'))|G_nG_n'|^{-1} \\ &\geqq (\text{tr } (G_nG_n'))^2\tfrac{1}{4}|G_nG_n'|^{-1} = (2n^2 + 3n + 7)^2\tfrac{1}{12}(n^2 - 1)^{-1},\end{aligned}$$

so that (4.7) is violated. On the other hand, it is not hard to show that G_n satisfies (4.1), and since

$$12\lambda_n^2 = n\{2n^2 + 3n + 7 - [(2n^2 + 3n + 7)^2 - 12n^2 + 12]^{\frac{1}{2}}\},$$

it is straightforward to show that λ_n satisfies (4.3).

It may be of interest to note that (4.1) and (4.3) are satisfied for general polynomial regression over the integers (i.e., where $u_s = \sum_{j=0}^{p-1} s^j\beta_j + e_s$), and even for such extreme cases as when

$$u_s = \beta_1 + \beta_2 t^s, \qquad s = 1, 2, \cdots,$$

for a known constant $t > 1$.

Conditions (4.1) and (4.3) are the most general conditions under which (4.4) can hold as long as no attention is paid to any special features of the stopping rule ν_n used. Since the stopping rules considered in Albert's paper (and also in Gleser [3]) are of a particular form, it is possible that in such a special context, condition (4.3) can be further relaxed.

As a final remark, we note that if

$$u^{(i)} = \boldsymbol{\beta} g^{(i)} + E^{(i)},$$

$\boldsymbol{\beta}$ a fixed $q \times p$ matrix, $g^{(i)}$ a known $p \times 1$ vector, $\{E^{(i)}\}$ a sequence of iid random

$q \times 1$ vectors drawn from a distribution $H(E)$ with mean vector 0 and covariance matrix Σ, $0 < \text{min root } \Sigma \leqq \text{max root } \Sigma < \infty$, the following theorem can be proved concerning the classical estimator

$$\hat{\boldsymbol{\beta}}(n) = U_n G_n' (G_n G_n')^{-1}$$

where $G_n = (g^{(1)}, g^{(2)}, \cdots, g^{(n)})$ is as before, and $U_n = (u^{(1)}, u^{(2)}, \cdots, u^{(n)})$: $q \times n$.

THEOREM 4.4. *With* $\{\nu_n\}$ *and* λ_n *defined as before, then if* (4.1) *and* (4.3) *hold,*

$$\lim_{n \to \infty} \mathcal{L}((G_n G_n')^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}(n) - \boldsymbol{\beta})' = \mathcal{L}(W) \tag{4.13}$$

where the p columns of W are independent, identically distributed as q-variate $\mathfrak{N}(0, \Sigma)$.

REFERENCES

[1] ALBERT, A. (1966). Fixed size confidence ellipsoids for linear regression parameters. *Ann. Math. Statist.* **37** 1602–1630.

[2] ANSCOMBE, F. J. (1952). Large sample theory of sequential estimation. *Proc. Cambridge Phil. Soc.* **48** 600–608.

[3] GLESER, L. J. (1965). On the asymptotic theory of fixed-size sequential confidence bounds for linear regression parameters. *Ann. Math. Statist.* **36** 463–467 (See also Correction Note, *Ann. Math. Statist.* **37** 1053–1055).

[4] GNEDENKO, B. V. and KOLMOGOROV, A. N. (1954). *Limit Distributions For Sums of Independent Random Variables.* Addison-Wesley, Cambridge.

[5] MOGYORÓDI, J. (1961). On limiting distributions for sums of a random number of independent random variables. *Publ. Math. Inst. Hungar. Acad. Sci.* **6** 365–371.