ANNALES UNIVERSITATIS MARIAE CURIE-SKŁODOWSKA
LUBLIN–POLONIA
VOL. XXXIV, 9 SECTIO A 1980

Statistics Departament
University of Adelaide, South Australia
Instytut Matematyki
Uniwersytet Marii Curie-Skłodowskiej

Kerwin W. MORRIS, Dominik SZYNAL

Convergence in Distribution of Multiply-indexed Arrays, with Applications in MANOVA

Zbieżność według rozkładu wielowskaźnikowych tablic z zastosowaniami w MANOVA

Сходимость по распределению мультииндексных таблиц с приложениями в MANOVA

1. Introduction. The importance of convergence in distribution in statistical inference arises as follows. The data $y_1, .., y_n$ arising from n performances of a given random process ϵ is used to calculate various quantities of interest, $\underset{\sim}{W}_n = (W_{1n}, .., W_{ln})'$ say, which are then used to construct significance, confidence intervals etc. These require the evaluation of probabilities of the form $P(\underset{\sim}{W}_n \in A)$, for given sets $A \in R^l$. If the distribution of $\underset{\sim}{W}_n$ is intractable, an approximation to $P(\underset{\sim}{W}_n \in A)$ is available when the sample size n is large in the case when the sequence $\{\underset{\sim}{W}_n\}$ converges in distribution to a variate $\underset{\sim}{W}$ with know distribution for then ([2], Theorem 2.1)

$$\lim_{n \to \infty} P(\underset{\sim}{W}_n \in A) = P(\underset{\sim}{W} \in A)$$

for all sets A of practical interest.

Consider now the situation in e.g. MANOVA. There are now several (k say) independent random processes $\epsilon_1, .., \epsilon_k$, the data arises from n_i performances of ϵ_i, $i = 1, .., k$, and leads to quantities of interest of the form $W_{n_1 \ldots n_k}$. An approximation to $P(\underset{\sim}{W}_{n_1 \ldots n_k} \in A)$ when all the sample sizes $n_1, .., n_k$ are large may then be important in practice for similar reasons. Such approximations are provided by the type of convergence in distribution of multiply-indexed arrays $\{\underset{\sim}{W}_{n_1 \ldots n_k}\}$ of random vectors that is defined below.

2. Multiply-indexed arrays. We discuss in some detail only the case $k = 2$, since the treatment when $k > 2$ presents no additional difficulties.

2. 1 Definition and notation. We call a set of real numbers $\{a_{n_1 n_2}; n_1 \geq 1, n_2 \geq 1\}$ a doubly-indexed array. It may be conveniently pictured in table form –

$n_1 \backslash n_2$	1	2	3	...
1	a_{11}	a_{12}	a_{13}	...
2	a_{21}	a_{22}	a_{23}	...
3	a_{31}	a_{32}	a_{33}	...
⋮	⋮	⋮	⋮	

In view of later use in MANOVA, it will be convenient to write $N = \text{diag}(n_1, n_2)$, $a_{n_1 n_2} = \{a_N\}$, and the array a_N.

Further, if $N_1 = \text{diag}(n_{11}, n_{12})$, $N_2 = \text{diag}(n_{21}, n_{22})$, we shall write $N_1 > N_2$ if $n_{11} > n_{21}$ and $n_{12} > n_{22}$, with a similar meaning for $N_1 \geq N_2$.

Finally by 'N is arbitrarily large' we shall mean that n_1 and n_2 are both arbitrarily large.

2.2 Limit points. We say that α (finite) is a limit point of $\{a_N\}$ if for arbitrary $\epsilon > 0$ and $n > 0$, $\exists N \geq nI$ such that $|a_N - \alpha| < \epsilon$.

There is the usual extension to infinite limit points. (There will be similar extension below, which in general will not be mentioned explicitly.)

We show now, by a standard argument, thet every array $\{a_N\}$ has a limit point.

In the case when $\{a_N\}$ is bounded, $a_N \in J_0 = [a, b] \; \forall N$ say, construct a sequence $\{J_n\}$ of closed intervals by repeated subdivision of J_0, viz., for $n = 1, 2, .., J_n$ is the left half of J_{n-1} if this half contains terms a_N with N arbitrarily large, and otherwise J_n is the right half. Then $\{J_n\}$ defines the point $\alpha = \bigcap_1^\infty J_n$. This point α is a limit point of $\{a_N\}$, since $\forall n \; J_n$ contains α and terms a_N with N arbitrarily large. Moreover, $\alpha = \liminf a_N$, since for arbitrary $\epsilon > 0$, $\exists n_0$ sueh that $a_N > \alpha - \epsilon \; \forall N \geq n_0 I$.

If $\{a_N\}$ is not bounded, a similar argument shows that a_N has a limit point, which now may be infinite.

2.3 Subarrays. Let S be a subset of diagonal matrices N that contains matrices that are arbitrarily large. We call $\{a_N, N \in S\}$ a subarray of $\{a_N\}$.

Limit points of subarrays are defined in the obvious way, and it follows, as in 2.2, that every subarray has a limit point.

2.4 Convergence. We say that $\{a_N\}$ converges to α (finite), and write $\lim_{N \to \infty} a_N = \alpha$, if for arbitrary $\epsilon > 0$, $\exists n_0$ such that $|a_N - \alpha| < \epsilon \; \forall N \geq n_0 I$.

Similarly, we say that the subarray $\{a_N, N \in S\}$ converges to α (finite) if for arbitrary $\epsilon > 0$, $\exists n_0$ such that $|a_N - \alpha| < \epsilon \; \forall N \in S$ such that $N \geq n_0 I$.

The usual results then follow. As an example, we prove that if α is a limit point of $\{a_N\}$ then there exists a subarray that converges to α.

In the case when α is finite, let $\{\epsilon_i\}$ be a null sequence of positive terms, and construct a set $S = \{N_i\}$ as follows.

Choose $N_1 > I$ such that $| a_{N_1} - \alpha | < \epsilon_1$, then successively choose $N_{i+1} > N_i$ such that $| a_{N_{i+1}} - \alpha | < \epsilon_{i+1}, i = 1, 2, ..,$ (such N_i always exist, from 2.2).

For given $\epsilon > 0, \exists j$ such that $\epsilon_i < \epsilon \, \forall i \geqslant j$. Then $| a_{N_i} - \alpha | < \epsilon \, \forall i \geqslant j$.

Writing $n_0 = \min (n_{j1}, n_{j2})$, where $N_j = \operatorname{diag} (n_{j1}, n_{j2})$, then $S \cap \{ N, N \geqslant n_0 I \} = \{ N_i, i \geqslant j \}$, whence $| a_N - \alpha | < \epsilon \, \forall N \in S$ such that $N \geqslant n_0 I$ and the subarray $\{ a_N, N \in S \}$ converges to α.

There is a similar result for limit points of subarrays.

We mentione one further result, viz., that if $\{ a_N \}$ converges to α, then every subarray of $\{ a_N \}$ converges to α. And there is the corresponding result for convergent subarrays.

2.5 Lim inf and lim sup. The following treatment is parallel to Feller's treatment of lim inf and lim sup ([4], IV. 1), and uses his $\cap$, $\cup$ notation.

We first introduce a sequential ordering of the terms of $\{ a_N \}$ with $N \geqslant nI$, viz.

$$a_{nn}, a_{n+1\,n}, a_{nn+1}, a_{n+2\,n}, a_{n+1\,n+1}, a_{nn+2}, a_{n+3n}, \cdots .$$

Next, consider the sequence $\{ w_n \}$, where

$$w_n = a_{nn} \cap a_{n+1\,n} \cap a_{nn+1} \cap a_{n+2\,n} \cap \ldots = \bigcap_{N \geqslant nI} a_N .$$

Clearly $w_n \uparrow$, whence $\{ w_n \}$ convergence to a limit, α say. Thus, in the case when α is finite, for arbitrary

$$\epsilon > 0, \exists\, n_0 \text{ such that } \alpha - \epsilon < w_n \leqslant \alpha \, \forall n \geqslant n_0 \tag{1}$$

We now show that for arbitrary $\epsilon > 0$,

$$\exists\, n_1 \text{ such that } a_N > \alpha - \epsilon \, \forall N \geqslant n_1 I, \text{ and} \tag{i}$$

$$\exists\, N \text{ arbitrarily large such that } a_N < \alpha + \epsilon, \tag{ii}$$

from which it follows that $\alpha = \lim \inf a_N$. Firstly, since by definition $w_{n_0} \leqslant a_N \, \forall N \geqslant n_0 I$, then, from (1), (i) holds with $n_1 = n_0$. Next, suppose that (ii) does not hold. Then $\exists\, \epsilon_1 > 0$ and n_2 such that $a_N \geqslant \alpha + \epsilon_1 \, \forall N \geqslant n_2 I$. But then $w_n \geqslant w_{n_2} \geqslant \alpha + \epsilon_1 \, \forall n \geqslant n_2$, which contradicts (1).

There is a similar treatment for lim sup a_N.

2.6 Fatou's lemma and the dominated convergence theorem. We consider now an array $\{ f_N(\underset{\sim}{x}) \}$ of functions $f : R^l \to R^1$. Then, from 2.5, for each $\underset{\sim}{x}$,

$$w_n(\underset{\sim}{x}) = \bigcap_{N \geqslant nI} f_N(\underset{\sim}{x})$$

defines an increasing sequence $\{ w_n(x) \}$ that converges to lim inf $f_N(x)$.

Fatou's lemma. ([4], IV. 2) *Suppose that* $\{ f_N(x) \}$ *is an array of non-negative functions, and that* $F(x)$ *is a distribution function* (d.f.).

If f_N *is integroble for all* N, *i.e. if*

$$E[f_N] = \int_{R^Q} f_N(\underset{\sim}{x})\, dF(\underset{\sim}{x}) < \infty \;\forall N$$

then

$$E\,[\liminf f_N] \leq \liminf E\,[f_N]\,.$$

Proof. Define a sequence of functions $\{f_n\}$ as follows. For each $n \geq 1$, choose N_n such that $N_n \geq nI$, and define

$$f_n = f_{N_n} \tag{2}$$

By definition of w_n, $w_n \leq f_n \;\forall n$, whence $E(w_n) \leq E(f_n) \;\forall n$, and so

$$\liminf E(w_n) \leq \liminf E(f_n)\,. \tag{3}$$

Since $w_n \uparrow$, $\lim w_n = \liminf f_N$, and w_n is integrable for all n, then, by the monotone convergence theorem ([4], IV. 2), $\{E(w_n)\}$ converges, and $\lim E(w_n) = E(\lim w_n)$. It follows then, using (3), that

$$E\,(\liminf f_n) \leq \liminf E(f_N)\,. \tag{4}$$

Since (4) holds for all sequences $\{f_n\}$ satisfying (2), it is enough to show that there exists such a sequence for which $\liminf E(f_n) = \liminf E(f_N)$. To show this, consider the array $\{E(f_N)\}$, and write $\alpha = \liminf E(f_N)$. By 2.4, there exists a subarray $\{E(f_N), N \in S\}$ that converges to α. Consider now the corresponding subarray $\{f_N, N \in S\}$, and construct from it a sequence $\{f_n\}$ as follows. Choose any element N_1 of S and define $f_1 = f_{N_1}$; then for $n = 1, 2, ..$, choose an element N_{n+1} of S such that $N_{n+1} > N_n$ and define $f_{n+1} = f_{N_{n+1}}$. Then

(i) $N_n \geq nI \;\forall n$, so that $\{f_n\}$ satisfies (2).

(ii) Since $\{E(f_n)\} = \{E(f_N), N \in S_1\}$, where $S_1 = \{N_i\} \subset S$,

and, by 2.4, the subarray $\{E(f_N), N \in S_1\}$ converges to α, then $\lim_{n\to\infty} E(f_n) = \alpha$ and $\liminf E(f_n) = \liminf E(f_N)$, as required.

The following theorem then follows from Fatou's lemma in the standard way (see e.g. [4], IV. 2).

Dominated convergence theorem. *If $\{f_N(\underset{\sim}{x})\}$ is an array such that f_N is integrable $\forall N$, and that $\lim_{N\to\infty} f_N(\underset{\sim}{x}) = f(\underset{\sim}{x})$ pointwise, and that there exists an integrable function u such that $|f_N(\underset{\sim}{x})| < u(\underset{\sim}{x}) \;\forall \underset{\sim}{x}$, then*

$$\lim_{N\to\infty} E(f_N) = E(f)\,.$$

2.7 Helly's theorem. Helly's theorem ([4], VIII. 6) may be generalized to arrays $\{F_N(\underset{\sim}{x})\}$ of d.f. The proof is essentially the same as the proof in Feller, and depends on the following lemma.

Lemma. *If $\{f_N(x)\}$ is a given array of bounded functions ($R^{\ell} \to R^l$) and $\{a_i\}$ is a given sequence of points in R^{ℓ}, then there exists a subarray $\{f_N, N \in S\}$ that converges at all points a_i.*

Proof. (c.f. [4], VIII. 6). By 2.2, the bounded array $\{f_N(a_1)\}$ has a limit point, and hence contains a convergent subarray $\{f_N(a_1), N \in S_1\}$. Proceeding in this way, the bounded subarray $\{f_N(a_2), N \in S_1\}$ has a limit point, and hence contains a convergent subarray $\{f_N(a_2), N \in S_2\}$. Continuing this procedure, we generate a sequence of sets $S_1 \supset S_2 \supset \ldots \supset S_n \supset \ldots$ such that for each $i = 1, 2, \ldots, \{f_N(a_i), N \in S_i\}$ is a convergent subarray.

For each $n \geqslant 1$ we now choose an element $N_n \in S_n$ such that $N_n \geqslant nI$, and define $S = \{N_n\}$. Then the subarray $f_N, N \in S$ has the required property. For consider $\{f_N(a_i), N \in S\}$. Since $N_n \in S_n \subset S_i \;\forall\, n \geqslant i$, then apart from a finite number of terms, $\{f_N(a_i), N \in S\}$ is a subarray of $\{f_N(a_i), N \in S_i\}$ which we know converges.

Thus $\{f_N(a_i), N \in S\}$ converges for $i = 1, 2, \ldots$

The generalizations of these results when $k > 2$ are now used to develop a theory of convergence in distribution for multiply-indexed arrays.

3. Convergence in distribution for multiply-indexed arrays. Let $\{W_n\}$ be a k-fold multiply-indexed array of $\ell \times l$ vector variates and W an $\ell \times l$ vector variate. We denote the corresponding d.f. by $\{F_N(x)\}$ and $F(x)$, the corre-characteristic functions (c.f.) by $\{\zeta_N(t)\}$ and $\zeta(t)$, and write

$$E_N(f) = \int_{R^{\ell}} f(x)\, dF_N(x)$$

and

$$E(f) = \int_{R^{\ell}} f(x)\, dF(x)\,.$$

Definition. We say that $\{W_N\}$ converges in distribution to W, and write $W_N \xrightarrow{D} W$, iff $\lim_{N\to\infty} F_N(x) = F(x) \;\forall$ continuity points x of F.

Theorem 1. *$W_N \xrightarrow{D} W$ if and only if either*

(i) $\lim_{N\to\infty} P(W_N \in I) = P(W \in I)$ *for all bounded open 'rectangles' I such that* $P(W \in \partial I) = 0$,

or (ii) $\lim_{N\to\infty} P(W_N \in A) = P(W \in A)$ *for all Borel sets A such that* $P(W \in \partial A) = 0$,

or (iii) $\lim_{N\to\infty} E_N(f) = E(f)$ *for all bounded and continuous functions* $f: R^l \to R^1$,

or (iv) $\lim_{N\to\infty} \zeta_N(t) = \zeta(t) \;\forall\, t$.

Moreover, if $W_N \xrightarrow{D} W$, then the convergence in (iv) *is uniform for all t in any bounded domain of R^l.*

Proof. The proof of (ii) and (iii) depends only on the content of 2.1 – 2.4 and follows step for step the corresponding proof of Bilingsley ([2], § 2). The fact that (i) ⇒ (iii) similarly follows the proof of the theorem in [4], VIII. 1. The fourth part (a continuity theorem for c.f.) depends also on 2.6 – 2.7, and can be proved in the same way as the corresponding 'ordinary' theorem, as e.g. in [4], XV. 3 or in [3], Chapter 11.

In the case when $\underset{\sim}{W}$ has a singular distribution concentrated at the single point $\underset{\sim}{\alpha}$, we say that $\{\underset{\sim}{W}_N\}$ converges in probability to $\underset{\sim}{\alpha}$, and we shall write $\underset{\sim}{W}_N \xrightarrow{D} \underset{\sim}{\alpha}$, as well as the standard $\underset{\sim}{W}_N \xrightarrow{P} \underset{\sim}{\alpha}$.

All the standard results for convergence in distribution of sequences of vector variates have their obvious counterparts in the theory of convergence in distribution of multiply-indexed arrays. We recall in particular two results. The first states that, if $\underset{\sim}{W}_{1N} \xrightarrow{D} \underset{\sim}{W}_1$ and $\underset{\sim}{W}_{2N} \xrightarrow{D} \underset{\sim}{\alpha}_2$, then writing $\underset{\sim}{W}_N = (\underset{\sim}{W}'_{1N}, \underset{\sim}{W}'_{2N})'$, $\underset{\sim}{W}_N \xrightarrow{D} \underset{\sim}{W} = (\underset{\sim}{W}'_1, \underset{\sim}{W}'_2)'$, where $P(\underset{\sim}{W}_2 = \underset{\sim}{\alpha}_2) = 1$, i.e. the limiting joint distribution is singular, and concentrated on the hyperplane $\underset{\sim}{W}_2 = \underset{\sim}{\alpha}_2$. In such a case, we shall write

$$\begin{pmatrix} \underset{\sim}{W}_{1N} \\ \underset{\sim}{W}_{2N} \end{pmatrix} \xrightarrow{D} \begin{matrix} \underset{\sim}{W}_1 \\ \underset{\sim}{\alpha}_2 \end{matrix}$$

The second result, which has widespread application, we state as a theorem.

Theorem 2.

$$\underset{\sim}{W}_N \xrightarrow{D} \underset{\sim}{W} \Rightarrow \phi(\underset{\sim}{W}_N) \xrightarrow{D} \phi(\underset{\sim}{W})$$

for every Borel-measurable function $\phi : R^{\ell} \to R^q$ such that $P(\underset{\sim}{W} \in D_\phi) = 0$, where

$$D_\phi = \{\underset{\sim}{x}; \phi(\underset{\sim}{x}) \text{ is discontinuous}\}.$$

The proof again follows step for step the corresponding proof in Billingsley ([2], Corollary 3 of theorem 3.3).

4. Some asymptotic results in MANOVA. As an application of § 3, we now derive some asymptotic results in MANOVA, on the assumption of a common non-singular covariance matrix Σ.

4.1 MANOVA notation. We suppose that the data is obtained from n_i performances of the random process $\mathcal{E}_i$, $i = 1, .., k$, where $\mathcal{E}_1, .., \mathcal{E}_k$ are independent processes. For $\mathcal{E}_i$, we denote the $p \times 1$ variate by $\underset{\sim}{y}_i$ and its mean by $\underset{\sim}{\mu}_i$. We denote the corresponding $n_i \times p$ data matrix by Y_i, and the sample mean and covariance matrix by $\bar{\underset{\sim}{y}}_i$ and $S_{(i)}$. We write

$$\sum_1^k n_i = n, \qquad N = \operatorname{diag}(\underset{\sim}{n}_1, .., \underset{\sim}{n}_k),$$

$$\underset{k \times p}{M} = \begin{pmatrix} \underset{\sim}{\mu}'_1 \\ \vdots \\ \underset{\sim}{\mu}'_k \end{pmatrix} = (\mu_{ij}), \qquad \underset{k \times p}{\bar{Y}_N} = \begin{pmatrix} \bar{\underset{\sim}{y}}'_1 \\ \vdots \\ \bar{\underset{\sim}{y}}'_k \end{pmatrix} = (y_{ij}),$$

$$\underset{n_i \times p}{Y_i} = \begin{pmatrix} \underset{\sim}{y}'_{i1} \\ \vdots \\ \underset{\sim}{y}'_{in_i} \end{pmatrix}, \qquad \underset{n \times p}{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_k \end{pmatrix},$$

and $\underset{p\times p}{S} = \frac{1}{n-k} \sum_i (n_i - 1)\, S_{(i)} = \frac{1}{n-k} \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\underset{\sim}{y}_{ij} - \bar{\underset{\sim}{y}}_i)(\underset{\sim}{y}_{ij} - \bar{\underset{\sim}{y}}_i)'$.

The above assumptions, which we shall call the model G, can be summed up as follows:

$$G: \text{The rows of } Y \text{ are independent vector variates, and}$$
$$E(Y) = XM,\ \mathrm{Var}(Y) = \Sigma \otimes I_n, \text{ where} \tag{5}$$
$$\underset{n\times k}{X} = \begin{bmatrix} \underset{\sim}{1}_1 & \underset{\sim}{0} & \dots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{1}_2 & \dots & \underset{\sim}{0} \\ \vdots & \vdots & & \vdots \\ \underset{\sim}{0} & \underset{\sim}{0} & \dots & \underset{\sim}{1}_k \end{bmatrix}$$

Note that

$$X'X = N,\ X'Y = N\bar{Y}_N \text{ and } r(X) = k \tag{6}$$

and further, that each column of $E(Y) \subset \mathcal{R}(X) \subset R^n$, that $P = X(X'X)^{-1}X' = XN^{-1}X'$ is the orthogonal projector matrix (o.p. matrix) onto $\mathcal{R}(X)$, and that $(n-k)S = Y'(I-P)Y$.

We now consider the usual kind of MANOVA hypothesis H, viz.

$$H: M = X_1 B_1$$
$$\text{where } X_1 \text{ is a known } k \times r \text{ matrix of rank } r. \tag{7}$$

When H is true, $E(Y) = XX_1B_1 = X_0B_1$ where $X_0 = XX_1$ has rank r, each column of $E(Y) \subset \mathcal{R}(X_0) \subset \mathcal{R}(X)$, the o.p. matrix onto $\mathcal{R}(X_0)$ is

$$P_0 = XX_1(X_1'NX_1)^{-1}X_1'X'$$

and

$$S_0 = \frac{l}{n-r} Y'(I-P_0)Y$$

is an unbiassed estimate of Σ.

The MANOVA table for testing H is then

Source	SSP	DF	MSSP
H vs. G	$Y'(P-P_0)Y = S_1$	$k-r$	S_1
Within class	$Y'(I-P)Y = S$	$n-k$	S
Total	$Y'(I-P_0)Y = S_0$	$n-r$	S_0

4.2 A central limit theorem.

Theorem 3. *On G,* $N^{1/2}(\bar{Y}_N - M) \xrightarrow{D} W \sim N(0, \Sigma \otimes I_k)$.

Proof. Writing

$$\underset{k \times p}{T} = \begin{pmatrix} \underset{\sim}{t}_1' \\ \vdots \\ \underset{\sim}{t}_k' \end{pmatrix}$$

the c.f. $\zeta_N(T)$ of $W_N = N^{1/2}(\bar{Y}_N - M)$ is

$$\zeta_N(T) = E[\exp(i\,Tr(T'W_N))] = E[\prod_{j=1}^{k} (\exp(i\,\underset{\sim}{t}_j'(\bar{\underset{\sim}{y}}_j - \underset{\sim}{\mu}_j)\sqrt{n_j}))] =$$
$$= \prod_{j=1}^{k} E[\exp(i\,\underset{\sim}{t}_j'(\bar{\underset{\sim}{y}}_j - \underset{\sim}{\mu}_j)\sqrt{n_j})] = \prod_{j=1}^{k} \phi_{n_j}^{(j)}(\underset{\sim}{t}_j)$$

since $\bar{\underset{\sim}{y}}_1, .., \bar{\underset{\sim}{y}}_k$ are independent where

$$\phi_{n_j}^{(j)}(\underset{\sim}{t}) = E[\exp(i\,\underset{\sim}{t}'(\bar{\underset{\sim}{y}}_j - \underset{\sim}{\mu}_j)\sqrt{n_j}\,], \quad j = 1, .., k.$$

But it is known from the ordinary multivariate central limit theorem that, for given $\underset{\sim}{t}_j$,

$$\lim_{n_j \to \infty} \phi_{n_j}^{(j)}(\underset{\sim}{t}_j) = \exp(-0.5\,\underset{\sim}{t}_j'\Sigma \underset{\sim}{t}_j), \quad j = 1, .., k$$

Thus, for given T,

$$\lim_{N \to \infty} \zeta_N(T) = \prod_{j=1}^{k} \exp(-0.5\,\underset{\sim}{t}_j'\Sigma \underset{\sim}{t}_j) = \exp(-0.5\,Tr(T\Sigma T')) = E[\exp(iTr(T'W))]\,,$$

and the theorem follows from theorem 1.

4.3 The asymptotic distribution of S_1. Suppose that A is a $p \times p$ symmetric matrix. By $\underset{\sim}{A}$ we shall mean the $p(p+1)/2 \times 1$ vector

$$\underset{\sim}{A} = (a_{11}, .., a_{1p}, a_{22}, .., a_{2p}, .., a_{pp})\,.$$

Theorem 4. *When H is true, $\underset{\sim}{S}_1 \xrightarrow{D} \underset{\sim}{V}$, where $V = U'U$ and $\underset{(k-r) \times p}{U} \sim N(0, \Sigma \otimes I_{k-r})$.*

Proof. When H is true, the columns of XM lie in $\mathcal{R}(X_0) \subset \mathcal{R}(X)$, so that $(P - P_0)XM =$ $= 0$. Thus

$$S_1 = Y'(P - P_0)Y = (Y - XM)'(P - P_0)(Y - XM) = W_N'(I - P_N)W_N,$$

where, from (6) and the definitions of P and P_0,

$$\underset{k \times k}{P_N} = N^{1/2}X_1(X_1'NX_1)^{-1}X_1'N^{1/2},$$

the o.p. matrix onto the r-dimensional subspace Ω_N of R^k, where $\Omega_N = \mathcal{R}(N^{1/2}X_1)$.

Now let $\underset{k \times (k-r)}{H_N}$ be a matrix whose columns are an orthonormal basis of $\Omega_N^{\perp}$. Then $H_N'H_N = I_{k-r}$, $H_NH_N' = I - P_N$, and

$$S_1 = U'_N U_N \tag{8}$$

where $U_N = H'_N W_N$.

We now show that $U_N \xrightarrow{D} U \sim N(0, \Sigma \otimes I_{k-r})$, from which the theorem follows by a simple application of theorem 2. From Theorems 3 and 1

$$\zeta_N(T) = \exp(-0.5\,Tr(T\Sigma T')) + f_N(T) \tag{9}$$

where $\lim_{N\to\infty} f_N(T) = 0$ uniformly for T in any bounded domain $A \subset R^{kp}$.

Consider now the c.f. $\phi_N(T_1)$ of U_N, where T_1 is $(k-r) \times p$.

$$\phi_N(T_1) = E[\exp(i\,Tr(T'_1 U_N))] = \zeta_N(H_N T_1), \quad \text{using (8)}$$

$$= \exp(-0.5\,Tr(H_N T_1 \Sigma T'_1 H'_N)) + f_N(H_N T_1) = \exp(-0.5\,Tr(T_1 \Sigma T'_1)) + f_N(H_N T_1)$$

since $H'_N H_N = I_{k-r}$.

For fixed T_1, choose in (9) $A = \{T; Tr(T'T) \leq Tr(T'_1 T_1)\}$. Since $Tr((H_N T_1)'(H_N T_1)) = Tr(T'_1 T_1)\ \forall N$, then, from (9) $\lim_{N\to\infty} f_N(H_N T_1) = 0$, and the result follows by an application of Theorem 1.

Theorem 5. *On G, $\underset{\sim}{S} \xrightarrow{D} \underset{\sim}{\Sigma}$.*

Proof. We write $\nu_j = n_j - 1$, $j = 1, \ldots, k$ and $\nu = n - k$, so that

$$\underset{\sim}{S} = \sum_{j=1}^{k} (\nu_j/\nu)\, \underset{\sim}{S}_{(j)}\,.$$

It is well-known that for each j $\underset{\sim}{S}_{(j)} \xrightarrow{D} \underset{\sim}{\Sigma}$ as $n_j \to \infty$. Thus, writing

$$\phi^{(j)}_{n_j}(\underset{\sim}{t}_j) = E[\exp(i\,\underset{\sim}{t}'_j \underset{\sim}{S}_{(j)})] = \exp(i\,\underset{\sim}{t}'_j \underset{\sim}{\Sigma}) + f^{(j)}_{n_j}(\underset{\sim}{t}_j)\,,$$

then $\lim_{n_j\to\infty} f^{(j)}_{n_j}(\underset{\sim}{t}_j) = 0$ uniformly for $\underset{\sim}{t}_j$ in any bounded domain.

Now write $g^{(j)}_{n_j}(\underset{\sim}{t}) = \exp(-i\,\underset{\sim}{t}'\underset{\sim}{\Sigma}) f^{(j)}_{n_j}(\underset{\sim}{t})$, $j = 1, \ldots, k$.

Then

$$\phi^{(j)}_{n_j}(\underset{\sim}{t}) = (1 + g^{(j)}_{n_j}(\underset{\sim}{t}))\exp(i\,\underset{\sim}{t}'\underset{\sim}{\Sigma})$$

and, since $|\exp(i\,\underset{\sim}{t}'\underset{\sim}{\Sigma})| = 1$, then, given $\epsilon_1 > 0$, $k > 0$, $\exists\, n_{0j}$ such that $|g^{(j)}_{n_j}(\underset{\sim}{t})| < \epsilon_1$ $\forall\, n_j \geq n_{0j}$ and

$$\forall\, \underset{\sim}{t} \in A = \{\underset{\sim}{t};\, \underset{\sim}{t}'\underset{\sim}{t} \leq k\}\,. \tag{10}$$

Consider now the c.f. $\underset{\sim}{\zeta}_N(\underset{\sim}{t})$ of S, viz.

$$\zeta_N(\underset{\sim}{t}) = E[\prod_{j=1}^{k} \exp(i\underset{\sim}{t}'(\nu_j/\nu)\underset{\sim}{S}_{(j)})] = \prod_{1}^{k} \phi_{nj}^{(j)}(\nu_j\underset{\sim}{t}/\nu),$$

since $S_1, \ldots, S_k$ are independent. Thus (all logs being principal-valued)

$$\log \zeta_N(\underset{\sim}{t}) = (i\underset{\sim}{t}'\Sigma)\sum_j (\nu_j/\nu) + \sum_j \log(l + g_{nj}^{(j)}(\nu_j\underset{\sim}{t}/\nu)) + (2C_N\pi)\, i\,,$$

where C_N is integer depending on N.

Since $\sum_j (\nu_j/\nu) = 1$, the theorem will follow by showing that, for fixed $\underset{\sim}{t}$,

$$\lim_{N\to\infty} \sum_j \log(l + g_{nj}^{(j)}(\nu_j\underset{\sim}{t}/\nu)) = 0\,.$$

Using the fact that

$$|\log(l+z)| < 2|z| \text{ if } |z| < 0.5\,,$$

then

$$|\sum_j \log(l + g_{nj}^{(j)}(\nu_j\underset{\sim}{t}/\nu))| < 2\sum_j |g_{nj}^{(j)}(\nu_j\underset{\sim}{t}/\nu)|$$

provided that

$$|g_{nj}^{(j)}(\nu_j\underset{\sim}{t}/\nu)| < 0.5\,, \quad j = l, \ldots, k.$$

For arbitrary $\epsilon > 0$, choose now $\epsilon_1 = \epsilon/2k$ and $k = \underset{\sim}{t}'\underset{\sim}{t}$ in (10), and write $n_0 = \max(n_{0l}, \ldots, n_{0k})$. Since $\nu_j\underset{\sim}{t}/\nu \in A\ \forall N$, it then follows from (10) that

$$2\sum_{l}^{k} |g_{nj}^{(j)}(\nu_j\underset{\sim}{t}/\nu)| < \epsilon\ \forall N \geqslant n_0 l$$

and hence that

$$\lim_{N\to\infty} \sum_j \log(l + g_{nj}^{(j)}(\nu_j\underset{\sim}{t}/\nu)) = 0\,.$$

4.4 The eigen-values of S_1S^{-1}. We now consider the asymptotic distribution of the e. values of S_1S^{-1} when H is true. Since Theorem 5 $\Rightarrow |S| \xrightarrow{D} |\Sigma| > 0$, it follows that the possible lack of definition of S^{-1} has no effect on the asymptotic distribution. Furthermore, since

$$r(S_1) \leqslant \rho = \min(p, k-r)\ \forall N,$$

with equality almost always when n is large, only the ρ largest e. values $\ell_1 \geqslant \ell_2 \geqslant \ldots \geqslant \ell_\rho$ are of interest.

Theorem 6. *When H is true, $\underset{\sim}{\ell}_N \xrightarrow{D} L$, where $\underset{\sim}{\ell}_N = (\ell_1, \ldots, \ell_\rho)'$, $\underset{\sim}{L} = (L_1, \ldots, L_\rho)'$, $L_1 \geqslant L_2 \geqslant \ldots \geqslant L_\rho$ are the largest e. values of $Z'Z$, and $\underset{(k-r)\times p}{Z} \sim N(0, I_{p(k-r)})$.*

Proof. From theorem 4 and 5

$$\begin{bmatrix} \underset{\sim}{S}_1 \\ \underset{\sim}{S} \end{bmatrix} \xrightarrow{D} \begin{bmatrix} \underset{\sim}{V} \\ \underset{\sim}{\Sigma} \end{bmatrix} \text{ when } H \text{ is true.}$$

Since $\ell_N = \phi(S_1, S)$, where ϕ is Borel-measurable and continuous when $S = \Sigma$, it follows from Theorem 2 that $\ell_N \xrightarrow{D} \phi(V, \Sigma)$, i.e. the vector of the ρ largest e. values of $V\Sigma^{-1} = U'U\Sigma^{-1}$, where from Theorem 1.4,

$$U \sim N(0, \Sigma \otimes I_{k-r}) .$$

Write now $\Sigma^{-1} = A^2$, where A is symmetric. Since $V\Sigma^{-1} = A^{-1}(AU'UA)A = A^{-1}(Z'Z)A$, where $Z = UA \sim N(0, I_{p(k-r)})$, then $V\Sigma^{-1}$ and $Z'Z$ have the same e. values, and the result follows.

This theorem allows us to write down the asymptotic distribution when H is true of some statistics commonly used in practice for testing H, viz. Hotelling's T_0^2, Pillai's $V^{(p)}$, and the statistic U, which is essentially the Normal theory likelihood-ratio statistic, where

$$T_0^2 = Tr(S_1 S^{-1}) = \sum_l^{\rho} \ell_i, \quad V^{(p)} = \frac{n-k}{n-r}\, Tr(S_1 S_0^{-1}) = \sum_l^{\rho} \frac{\ell_i}{l + (\ell_i/n-k)}$$

$$\text{and } U^{-1} = \prod_1^{\rho} \left(l + \frac{\ell_i}{n-k}\right)$$

see e.g. [1], Ch. 8).

Theorem 7. *When H is true, T_0^2, $V^{(p)}$ and $(n-k)(U^{-1} - I)$ each converges in distribution to*

$$Tr(Z'Z) \sim \chi^2_{p(k-r)} .$$

Proof. It follows immediately from Theorem 2 that

$$T_0^2 = \sum_1^{\rho} \ell_i \xrightarrow{D} \sum_1^{\rho} L_i = Tr(Z'Z) \sim \chi^2_{p(k-r)}$$

and similarly that

$$V^{(p)} \xrightarrow{D} \sum_1^{\rho} L_i/(l + 0.L_i) = Tr(Z'Z) .$$

Finally,

$$(n-k)(U^{-1} - l) = \sum_1^{\rho} \ell_i + (n-k)^{-1} \sum_{i \neq j} \ell_i \ell_j + \dots + \frac{1}{(n-k)^{\rho - l}} \prod_l \ell_i$$

$$\xrightarrow{D} \sum_1^{\rho} L_i + 0. \sum_{i \neq j} L_i L_j + \dots + 0.\Pi L_i = Tr(Z'Z).$$

(It can also be shown somewhat similarly that

$$-n \log U \xrightarrow{D} \chi^2_{p(k-r)} .)$$

4.5 Estimation of B_1. If H is not rejected, the estimation of B_1 will often be of importance. We consider the asymptotic distribution of $\hat{B}_N$ when H is true, where $\hat{B}_N = (X_0' X_0)^{-1} X_0' Y$ is the matrix of minimum variance unbiassed linear estimates of B_1 when H is true. Since $E(\hat{B}_N) = (X_0' X_0)^{-1} X_0' X_0 B_1 = B_1$ and

$$\text{Var}(\hat{B}_N) = (I_p \otimes (X_0'X_0)^{-1} X_0')(\Sigma \otimes I_n)(I_p \otimes X_0 (X_0'X_0)^{-1}) = \Sigma \otimes (X_1'NX_1)^{-1},$$

it might be expected that $\hat{B}_N$ is asymptotically $N(B_1, \Sigma \otimes (X_1'NX_1)^{-1})$, in the sense that, if C_N and A are respectively $r \times r$ and $p \times p$ symmetric matrices such that

$$C_N^2 = X_1'NX_1, \; A^2 = \Sigma^{-1} \tag{11}$$

then $C_N(\hat{B}_N - B_1) A \xrightarrow{D} Z_1 \sim N(0, I_{pr})$.

To prove this, note first that when H is true $E(Y) = XM = X_0 B_1$, so that, using (6) and the notation of Theorem 3,

$$\hat{B}_N - B_1 = (X_0'X_0)^{-1} X_0' (Y - XM) = (X_1'NX_1)^{-1} X_0'N^{1/2} W_N$$

and $\quad C_N(\hat{B}_N - B_1) A = C_N (X_1'NX_1)^{-1} X_1'N^{1/2} W_N A = D_N W_N A$,

where $\quad D_N = C_N (X_1'NX_1)^{-1} X_1'N^{1/2}$.

The c.f. $\quad \phi_N(T_1)$ of $C_N(\hat{B}_N - B_1) A$ is

$$\phi_N(T_1) = E[\exp(i\, Tr(T_1'D_N W_N A))] = \zeta_N(D_N'T_1 A) =$$
$$= \exp(-0.5\, Tr(D_N'T_1 A \Sigma A T_1' D_N)) + f_N(D_N'T_1 A),$$

from (9), $\quad = \exp(-0.5\, Tr(T_1 T_1')) + f_N(D_N'T_1 A)$,

since, from (11) $A\Sigma A = I_p$ and $D_N D_N' = C_N (X_1'NX_1)^{-1} C_N = I_r$.

To show that $\lim_{N \to \infty} f_N(D_N'T_1 A) = 0$ for fixed T_1, note first that

$$Tr((D_N'T_1 A)'(D_N'T_1 A)) = Tr(T_1'T_1 \Sigma^{-1}) \; \forall N.$$

The result then follows from (9) by choosing

$$A = \{T; Tr(T'T) \leqslant Tr(T_1'T_1 \Sigma^{-1})\}.$$

Of more interest in practice is the result obtained by replacing Σ by S (or S_0, which is readily seen to converge in probability to Σ when H is true). If we write $A = \phi_2(\Sigma)$ and define

$$A_N = \phi_2(S), \tag{12}$$

it follows immediately from Theorem 2 that $A_N \xrightarrow{D} A$ and that

$$C_N(\hat{B}_N - B_1) A_N = (C_N(\hat{B}_N - B_1) A) A^{-1} A_N \xrightarrow{D} Z_1 A^{-1} A = Z_1,$$

which proves the following result.

Theorem 8. *When H is true*

$$\hat{B}_N \overset{\mathcal{L}}{\sim} N(B_1, S \otimes (X_1'NX_1)^{-1}),$$

in the sense that

$$C_N(\hat{B}_N - B_1)A_N \xrightarrow{D} N(0, I_{pr}),$$

where C_N *and* A_N *are defined in* (11) *and* (12).

REFERENCES

[1] Anderson, T. W. (1958), *An Introduction to Multivariate Statistical Analysis*, John Wiley, New York.

[2] Billingsley, P. (1971), *Weak Convergence of Measures:* Applications in Probability, Regional Conference Series in Applied Mathematics, 5. SIAM, Philadelphia.

[3] Breiman, L. (1968), *Probability*, Addison-Wesley, Reading, Mass.

[4] Feller, W. (1966), *An Introduction to Probability Theory and its Applications*, vol. II, John Wiley, New York.

STRESZCZENIE

W pracy prezentuje się uogólnienie zbieżności według rozkładu na wielowskaźnikowe tablice wektorów losowych. Rozważania te wykorzystuje się w analizie zbieżności według rozkładu statystyki T_0^2-Hotellinga i innych statystyk (w przypadku rozkładu różnego od normalnego) wykorzystywanych w MANOVA.

РЕЗЮМЕ

В работе представляются обобщение сходимости по распределению на мультииндексные таблицы случайных векторов. Эти исследования используются в анализе сходимости по распределению статистики T_0^2-Хотеллинга и других статистик (в случае распределения разного от нормального) использованных в MANOVA.