AN ASYMPTOTIC EXPANSION FOR THE DISTRIBUTION OF THE LINEAR DISCRIMINANT FUNCTION[1]

By Masashi Okamoto

Osaka University, Japan

0. Summary. The distribution of the linear discriminant function W, Anderson's classification statistic (1951), is investigated by several authors: Bowker (1960), Bowker and Sitgreaves (1961), Sitgreaves (1952, 1961), etc. Since the exact distribution is too complicated to be used numerically, as indicated by Sitgreaves (1961), we present here an asymptotic expansion of the distribution with respect to three numbers N_1, N_2 and n representing degrees of freedom. This is a generalization of the result of Bowker and Sitgreaves who deal with a special case where $N_1 = N_2 = N$ and $n = 2N - 2$.

1. Introduction. Let $\mathbf{X}_i (i = 1, \ldots, N_1 + N_2 + 1)$ and $\mathbf{S}$ be random p-vectors and a random $p \times p$-matrix, respectively, distributed independently, $\mathbf{X}_i (i = 1, \cdots, N_1)$ according to a normal distribution $\Pi_1 : N(\boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma})$; $\mathbf{X}_i (i = N_1 + 1, \cdots, N_1 + N_2)$ according to $\Pi_2 : N(\boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma})$; $\mathbf{X}_{N_1+N_2+1}$ according to either Π_1 or Π_2; and finally $n\mathbf{S}$ according to $W(n, \boldsymbol{\Sigma})$, a Wishart distribution with n degrees of freedom and variance matrix $\boldsymbol{\Sigma}$. When we put

$$W = [\mathbf{X}_{N_1+N_2+1} - \tfrac{1}{2}(\bar{\mathbf{X}}^{(1)} + \bar{\mathbf{X}}^{(2)})]'\mathbf{S}^{-1}(\bar{\mathbf{X}}^{(1)} - \bar{\mathbf{X}}^{(2)}) \tag{1.1}$$

where

$$\bar{\mathbf{X}}^{(1)} = (1/N_1) \sum_{i=1}^{N_1} \mathbf{X}_i \quad \text{and} \quad \bar{\mathbf{X}}^{(2)} = (1/N_2) \sum_{i=N_1+1}^{N_1+N_2} \mathbf{X}_i, \tag{1.2}$$

W is a generalization of the linear discriminant function W, Anderson's classification statistic, which corresponds to the special case

$$\mathbf{S} = \frac{1}{N_1 + N_2 - 2}\left\{\sum_{i=1}^{N_1} (\mathbf{X}_i - \bar{\mathbf{X}}^{(1)})(\mathbf{X}_i - \bar{\mathbf{X}}^{(1)})' + \sum_{i=N_1+1}^{N_1+N_2} (\mathbf{X}_i - \bar{\mathbf{X}}^{(2)})(\mathbf{X}_i - \bar{\mathbf{X}}^{(2)})'\right\}. \tag{1.3}$$

Now it is well known that if the Mahalonobis distance D^2 between two populations Π_1 and Π_2, defined by

$$D^2 = (\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)}), \tag{1.4}$$

is not zero, then as N_1, N_2 and n tend to infinity $\bar{\mathbf{X}}^{(1)} \to \boldsymbol{\mu}^{(1)}$, $\bar{\mathbf{X}}^{(2)} \to \boldsymbol{\mu}^{(2)}$ and $\mathbf{S} \to \boldsymbol{\Sigma}$ in probability and hence the limiting distribution of W is $N(\frac{1}{2}D^2, D^2)$ or

Received May 8, 1962; revised February 12, 1963.

[1] This investigation was supported in part by the U. S. Public Health Service (Grant RG-9469).

$N(-\frac{1}{2}D^2, D^2)$ according as $\mathbf{X}_{N_1+N_2+1}$ comes from Π_1 or from Π_2; that is, if for any real constant c we denote by

$$(1.5) \qquad P_1(c; D) = P_r\{W < \tfrac{1}{2}D^2 + cD \mid \Pi_1\}$$

the probability that $W < \frac{1}{2}D^2 + cD$ when $\mathbf{X}_{N_1+N_2+1}$ comes from Π_1 and similarly

$$(1.6) \qquad P_2(c; D) = P_r\{W < -\tfrac{1}{2}D^2 + cD \mid \Pi_2\},$$

then both $P_1(c; D)$ and $P_2(c; D)$ tend to $\Phi(c)$, cdf of $N(0, 1)$.

The purpose of this paper is to evaluate $P_1(c; D)$ and $P_2(c; D)$ in an asymptotic expansion with respect to N_1^{-1}, N_2^{-1} and n^{-1}. The formulae are derived by the "Studentization" method of Hartley (1938) and of Welch (1947), which has been shown by Ito (1956, 1960) and Siotani (1957) to be useful also for some multivariate problems. An analogous result is given by the present author (1961) for the quadratic discriminant function which appears in discriminating two normal populations with common mean vector and different variance matrices.

2. The main result. We shall first show that $P_2(c; D)$ can be derived from $P_1(c; D)$.

LEMMA 1. *Let $P_1^*(c; D)$ be the expression obtained by interchanging N_1 and N_2 in $P_1(c; D)$, then*

$$(2.1) \qquad P_2(c; D) = 1 - P_1^*(-c; D).$$

PROOF. Since we obtain $-W$ by interchanging $\bar{\mathbf{X}}^{(1)}$ and $\bar{\mathbf{X}}^{(2)}$ in W defined in (1.1), we have $P_r\{-W < \frac{1}{2}D^2 + cD \mid \Pi_2\} = P_1^*(c; D)$, which together with (1.6) implies (2.1).

Now the main theorem of this paper is

THEOREM. *If $D > 0$, then*

$$(2.2) \qquad P_1(c; D) = [1 + L(d; D) + Q(d; D)]\Phi(c) + O_3,$$

where d stands for the differential operator d/dc, $\Phi(c)$ for the cdf *of $N(0, 1)$ and*

$$(2.3) \qquad L(d; D) = \sum_{i=1}^{3} L_i(d; D),$$

$$(2.4) \qquad Q(d; D) = \tfrac{1}{2}[L(d; D)]^2 + \sum_{i \leqq j=1}^{3} Q_{ij}(d; D),$$

$$(2.5) \qquad \begin{aligned} L_1(d; D) &= (2N_1D^2)^{-1}[d^4 + p(d^2 + Dd)], \\ L_2(d; D) &= (2N_2D^2)^{-1}[(d^2 - Dd)^2 + p(d^2 - Dd)], \\ L_3(d; D) &= (4n)^{-1}[(2d^2 - Dd)^2 + 2(p+1)(3d^2 - Dd)], \end{aligned}$$

$$Q_{11}(d; D) = (4N_1^2D^4)^{-1}[2d^4(d^2 + Dd) + p(d^2 + Dd)^2],$$

$$Q_{22}(d; D) = (4N_2^2D^4)^{-1}[2(d^2 - Dd)^3 + p(d^2 - Dd)^2],$$

$$
\begin{aligned}
Q_{12}(d; D) &= (2N_1N_2D^4)^{-1}[2d^4(d^2 - Dd) + pd^4], \\
Q_{13}(d; D) &= (2N_1nD^2)^{-1}[4d^4(2d^2 - Dd) + 2(5p + 7)d^4 - D^2d^2 \\
&\qquad + (p^2 + p)(3d^2 + Dd)], \\
Q_{23}(d; D) &= (2N_2nD^2)^{-1}[2(d^2 - Dd)(2d^2 - Dd)^2 + 2(5p + 7)d^4 \\
&\qquad - 4(3p + 4)Dd^3 + (3p + 4)D^2d^2 \\
&\qquad + (p^2 + p)(3d^2 - Dd)], \\
Q_{33}(d; D) &= (12n^2)^{-1}[2(2d^2 - Dd)^2(7d^2 - 2Dd) + 9(15p + 13)d^4 \\
&\qquad - 24(4p + 3)Dd^3 + 3(5p + 3)D^2d^2 + 6(6p^2 + 13p \\
&\qquad + 9)d^2 - 6(p + 1)^2Dd]
\end{aligned}
\tag{2.6}
$$

and finally O_3 *stands for the term of the third order with respect to* $(N_1^{-1}, N_2^{-1}, n^{-1})$.

We shall defer the proof to the following sections and state now two corollaries, the first of which gives the linear terms in another form[2]:

COROLLARY 1. *Let* $\phi(c)$ *be the density of* $N(0, 1)$, *then*

$$
\begin{aligned}
P_1(c; D) &= \Phi(c) + \{(2N_1D^2)^{-1}[3c - c^3 + p(D - c)] \\
&\quad + (2N_2D^2)^{-1}[2D + 3c - c(D + c)^2 - p(D + c)] \\
&\quad + (4n)^{-1}[2(D + 3c) - c(D + 2c)^2 - 2p(D + 3c)]\}\phi(c) + O_2 .
\end{aligned}
\tag{2.7}
$$

PROOF. We have only to substitute the identities

$$
\begin{aligned}
&d\Phi(c) = \phi(c), \qquad d^2\Phi(c) = -c\phi(c), \\
&d^3\Phi(c) = (c^2 - 1)\phi(c), \qquad d^4\Phi(c) = (3c - c^3)\phi(c)
\end{aligned}
\tag{2.8}
$$

into the term $L(d; D)\Phi(c)$ of the theorem.

In many situations the discrimination is performed in the following way: We regard an observed value of $\mathbf{X}_{N_1+N_2+1}$ as coming from Π_1 or Π_2 according as the observed value of W is positive or negative. For this procedure the error probabilities of two kinds are given by

COROLLARY 2.

$$
\begin{aligned}
P_r\{W < 0 \mid \Pi_1\} &= \Phi\left(-\frac{D}{2}\right) + \frac{a_1}{N_1} + \frac{a_2}{N_2} + \frac{a_3}{n} \\
&\quad + \frac{b_{11}}{N_1^2} + \frac{b_{22}}{N_2^2} + \frac{b_{12}}{N_1N_2} + \frac{b_{13}}{N_1n} + \frac{b_{23}}{N_2n} + \frac{b_{33}}{n^2} + O_3,
\end{aligned}
\tag{2.9}
$$

$$
\begin{aligned}
P_r\{W > 0 \mid \Pi_2\} &= \Phi\left(-\frac{D}{2}\right) + \frac{a_2}{N_1} + \frac{a_1}{N_2} + \frac{a_3}{n} \\
&\quad + \frac{b_{22}}{N_1^2} + \frac{b_{11}}{N_2^2} + \frac{b_{12}}{N_1N_2} + \frac{b_{23}}{N_1n} + \frac{b_{13}}{N_2n} + \frac{b_{33}}{n^2} + O_3,
\end{aligned}
\tag{2.10}
$$

[2] When we put $N_1 = N_2 = N$ and $n = 2N - 2$, Formula (2.7) agrees with Theorem 3 of Bowker and Sitgreaves (1961) up to terms of order N^{-1} provided that the coefficients a_{3j} $(j = 1, 2)$ in the latter are changed in sign. The term of the second order is not yet checked.

where

$$a_1 = (2D^2)^{-1}(d_0^4 + 3p\,d_0^2), \qquad a_2 = (2D^2)^{-1}[d_0^4 - (p-4)d_0^2],$$
$$a_3 = \tfrac{1}{2}(p-1)d_0^2,$$
$$b_{11} = (8D^4)^{-1}[d_0^8 + 6(p+2)d_0^6 + (p+2)(9p+16)d_0^4 + 20p(p+2)d_0^2],$$
$$b_{22} = (8D^4)^{-1}[d_0^8 - 2(p-10)d_0^6 + (p-6)(p-16)d_0^4 + 4(p-4)(p-6)d_0^2],$$
$$b_{12} = (4D^4)^{-1}[d_0^8 + 2(p+8)d_0^6 - 3(p^2 - 10p - 16)d_0^4 - 12p(p-6)d_0^2],$$
$$b_{13} = (4D^2)^{-1}(p-1)[d_0^6 + 3(p+4)d_0^4 + 6(p+4)d_0^2],$$
$$b_{23} = (4D^2)^{-1}(p-1)[d_0^6 - (p-8)d_0^4 - 2(p-4)d_0^2],$$
$$b_{33} = \tfrac{1}{8}(p-1)[(p+1)d_0^4 + 4(p-12)d_0^2]$$

and d_0^i *are constants defined by*

$$d_0^i = (d^i/dc^i)\Phi(c)\mid_{c=-D/2} \qquad (i = 2, 4, 6, 8). \tag{2.11}$$

Proof. Since from (1.5), $P_r\{W < 0 \mid \Pi_1\}' = P_1(-D/2; D)$, the theorem and (2.9) imply

$$\begin{aligned} a_i &= N_i L_i(d; D)\Phi(c)\mid_{c=-D/2}, \\ b_{ij} &= N_i N_j[(1 - \tfrac{1}{2}\delta_{ij})L_i(d; D)L_j(d; D) + Q_{ij}(d; D)]\Phi(c)\mid_{c=-D/2}, \end{aligned} \tag{2.12}$$

where N_3 stands for n. Now we have a recurrence formula

$$d^{i+1}\Phi(c) = -(i-1)d^{i-1}\Phi(c) - cd^i\Phi(c) \tag{2.13}$$

by operating d^{i-1} on the second relation of (2.8); and hence, by putting $c = -D/2$,

$$Dd_0^i = 2d_0^{i+1} + 2(i-1)d_0^{i-1} \qquad (i = 1, 2, \cdots, 7). \tag{2.14}$$

Substituting (2.5) and (2.6) into the right-hand side of (2.12) and letting D, which appear in the coefficients there, vanish by repeated application of (2.14), we obtain the expressions of the a_i's and b_{ij}'s presented in the lemma, which proves (2.9). The dual equality (2.10) follows from (2.9) and Lemma 1.

At the end of this paper we present Table 1 giving values of coefficients a_i and b_{ij} appearing in (2.9) and (2.10) for selected values of p and D: $p = 1, 2, 3, 5, 7, 10, 20, 50$; $D = 1, 2, 3, 4, 6, 8$, as well as Table 2 giving the first three terms $\Phi(-D/2)$, $\sum a_i/N_i$ and $\sum b_{ij}/(N_iN_j)$ when $N_1 = N_2 = 100$ and $n = 198$. When $N_1 = N_2 = N$ and $n = 2N - 2$ Table 2 is applicable for other values of N than 100 by using the fact that the first and the second order terms are approximately proportionate to N^{-1} and N^{-2}, respectively. (The author wishes to thank Mr. N. Nakajima for his assistance in preparing the tables.)

TABLE 1

Coefficients in the asymptotic expansion of the error probability

p	D	Linear coefficients			Quadratic coefficients					
		a_1	a_2	a_3	b_{11}	b_{22}	b_{12}	b_{13}	b_{23}	b_{33}
1	1	0.02200	0.02200	0	$-0.0^{2}378$	$-0.0^{2}378$	$-0.0^{2}756$	0	0	0
	2	0.03025	0.03025	0	$-0.0^{2}378$	$-0.0^{2}378$	$-0.0^{2}756$	0	0	0
	3	0.02428	0.02428	0	$-0.0^{2}114$	$-0.0^{2}114$	$-0.0^{2}228$	0	0	0
	4	0.01350	0.01350	0	$0.0^{3}844$	$0.0^{3}844$	$0.0^{2}169$	0	0	0
	6	$0.0^{2}1662$	$0.0^{2}1662$	0	$0.0^{3}623$	$0.0^{3}623$	$0.0^{2}125$	0	0	0
	8	$0.0^{4}6690$	$0.0^{4}6690$	0	$0.0^{4}544$	$0.0^{4}544$	$0.0^{3}109$	0	0	0
2	1	0.2861	-0.06601	0.08802	0.164	-0.0231	0.141	-0.0412	$0.0^{2}275$	-1.06
	2	0.1210	0	0.1210	0.0227	$-0.0^{2}756$	0.0151	0.0907	-0.0302	-1.39
	3	0.05666	0.01349	0.09714	0.0107	$-0.0^{2}380$	$0.0^{2}694$	0.108	-0.0159	-1.03
	4	0.02362	0.01012	0.05399	$0.0^{2}648$	$-0.0^{3}685$	$0.0^{2}580$	0.0759	$0.0^{2}169$	-0.499
	6	$0.0^{2}2216$	$0.0^{2}1477$	$0.0^{2}6648$	$0.0^{2}127$	$0.0^{3}426$	$0.0^{2}169$	0.0138	$0.0^{2}425$	-0.0366
	8	$0.0^{4}7944$	$0.0^{4}6272$	$0.0^{3}2676$	$0.0^{4}798$	$0.0^{4}462$	$0.0^{3}126$	$0.0^{3}797$	$0.0^{3}404$	$-0.0^{4}669$
3	1	0.5501	-0.1540	0.1760	0.123	0.0127	-0.0406	-0.281	0.0715	-2.07
	2	0.2117	-0.03025	0.2420	0.0567	$-0.0^{2}378$	$-0.0^{2}756$	0.181	-0.0605	-2.66
	3	0.08904	$0.0^{2}2698$	0.1943	0.0306	$-0.0^{2}451$	$0.0^{2}447$	0.257	-0.0452	-1.89
	4	0.03374	$0.0^{2}6749$	0.1080	0.0152	$-0.0^{2}169$	$0.0^{2}675$	0.182	$-0.0^{2}675$	-0.864
	6	$0.0^{2}2770$	$0.0^{2}1293$	0.01330	$0.0^{2}210$	$0.0^{3}254$	$0.0^{2}198$	0.0321	$0.0^{2}702$	-0.0399
	8	$0.0^{4}9199$	$0.0^{4}5854$	$0.0^{3}5352$	$0.0^{3}110$	$0.0^{4}387$	$0.0^{3}140$	$0.0^{2}178$	$0.0^{3}744$	$0.0^{2}214$
5	1	1.078	-0.3301	0.3521	-0.587	0.249	-1.39	-1.35	0.407	-3.92
	2	0.3932	-0.09074	0.4839	0.147	0.0265	-0.189	0.363	-0.121	-4.84
	3	0.1538	-0.01889	0.3886	0.0941	$-0.0^{4}890$	-0.0356	0.676	-0.144	-3.16
	4	0.05399	0	0.2160	0.0418	$-0.0^{2}211$	$-0.0^{3}844$	0.486	-0.0540	-1.19
	6	$0.0^{2}3878$	$0.0^{3}9233$	0.02659	$0.0^{2}434$	$-0.0^{4}128$	$0.0^{2}211$	0.0820	$0.0^{2}813$	0.0532
	8	$0.0^{3}1171$	$0.0^{4}5018$	$0.0^{2}1071$	$0.0^{3}183$	$0.0^{4}252$	$0.0^{3}158$	$0.0^{2}432$	$0.0^{2}124$	0.0134

7	1	1.606	−0.5061	0.5281	−2.13	0.706	−4.07	−3.22	1.01	−5.55
	2	0.5747	−0.1512	0.7259	0.268	0.0870	−0.552	0.544	−0.181	−6.53
	3	0.2186	−0.04047	0.5828	0.189	0.0121	−0.122	1.26	−0.297	−3.79
	4	0.07424	$-0.0^{2}6749$	0.3239	0.0806	$-0.0^{3}422$	−0.0211	0.911	−0.142	−0.972
	6	$0.0^{2}4986$	$0.0^{3}5540$	0.03989	$0.0^{2}733$	$-0.0^{3}177$	$0.0^{2}161$	0.150	$0.0^{2}332$	0.279
	8	$0.0^{3}1422$	$0.0^{4}4181$	$0.0^{2}1606$	$0.0^{3}274$	$0.0^{4}140$	$0.0^{3}163$	$0.0^{2}760$	$0.0^{2}148$	0.0337
10	1	2.398	−0.7701	0.7921	−6.02	1.80	−10.6	−7.50	2.40	−7.57
	2	0.8469	−0.2420	1.089	0.507	0.234	−1.44	0.817	−0.272	−8.17
	3	0.3157	−0.07285	0.8742	0.392	0.0451	−0.340	2.43	−0.628	−3.55
	4	0.1046	−0.01687	0.4859	0.162	$0.0^{2}606$	−0.0752	1.78	−0.349	0.364
	6	$0.0^{2}6648$	0	0.05983	0.0133	$-0.0^{3}231$	$-0.0^{3}277$	0.284	−0.0150	0.868
	8	$0.0^{3}1798$	$0.0^{4}2927$	$0.0^{2}2409$	$0.0^{3}445$	$0.0^{5}129$	$0.0^{3}145$	0.0139	$0.0^{2}137$	0.0813
20	1	5.039	−1.650	1.672	−32.6	9.03	−53.6	−34.6	11.3	−10.8
	2	1.754	−0.5444	2.299	1.79	1.22	−7.33	1.72	−0.575	−5.75
	3	0.6395	−0.1808	1.846	1.58	0.282	−1.83	8.98	−2.61	7.50
	4	0.2058	−0.05062	1.026	0.631	0.0620	−0.461	6.64	−1.70	13.6
	6	0.01219	$-0.0^{2}1847$	0.1263	0.0453	$0.0^{2}126$	−0.0166	1.02	−0.172	4.99
	8	$0.0^{3}3053$	$-0.0^{4}1255$	$0.0^{2}5084$	$0.0^{2}130$	$-0.0^{5}492$	$-0.0^{3}132$	0.0473	$-0.0^{2}306$	0.388
50	1	12.96	−4.291	4.313	−238	63.7	−381	−235	77.8	12.7
	2	4.476	−1.452	5.928	10.2	8.71	−52.3	4.45	−1.48	74.1
	3	1.611	−0.5046	4.760	9.92	2.16	−13.3	52.9	−16.6	135
	4	0.5095	−0.1518	2.646	3.87	0.546	−3.52	39.4	−11.8	134
	6	0.02881	$-0.0^{2}7386$	0.3257	0.255	0.0211	−0.158	5.89	−1.53	37.3
	8	$0.0^{3}6815$	$-0.0^{3}1380$	0.01311	$0.0^{2}657$	$0.0^{3}310$	$-0.0^{2}296$	0.260	−0.0540	2.67

TABLE 2

First three terms of the error probability ($N_1 = N_2 = 100$, $n = 198$)

D		1	2	3	4	6	8
Principal term $\Phi\left(-\frac{D}{2}\right)$		.30854	.15866	.06681	.02275	$.0^{2}1350$	$.0^{4}3167$
Term of the first order	$p =$ 1	$.0^{3}4401$	$.0^{3}6049$	$.0^{3}4857$	$.0^{3}2700$	$.0^{4}3324$	$.0^{5}1338$
	2	$.0^{2}2645$	$.0^{2}1821$	$.0^{2}1192$	$.0^{3}6101$	$.0^{4}7051$	$.0^{5}2774$
	3	$.0^{2}4850$	$.0^{2}3037$	$.0^{2}1899$	$.0^{3}9503$	$.0^{3}1078$	$.0^{5}4208$
	5	$.0^{2}9260$	$.0^{2}5469$	$.0^{2}3312$	$.0^{2}1631$	$.0^{3}1823$	$.0^{5}7079$
	7	.01367	$.0^{2}7901$	$.0^{2}4724$	$.0^{2}2311$	$.0^{3}2568$	$.0^{5}9949$
	10	.02028	.01155	$.0^{2}6844$	$.0^{2}3331$	$.0^{3}3686$	$.0^{4}1426$
	20	.04233	.02371	.01391	$.0^{2}6733$	$.0^{3}7413$	$.0^{4}2861$
	50	.1085	.06019	.03510	.01694	$.0^{2}1859$	$.0^{4}7166$
Term of the second order	$p =$ 1	$-.0^{5}151$	$-.0^{5}151$	$-.0^{6}455$	$.0^{6}337$	$.0^{6}249$	$.0^{7}217$
	2	$-.0^{5}083$	$-.0^{4}294$	$-.0^{4}201$	$-.0^{5}766$	$.0^{6}320$	$.0^{7}841$
	3	$-.0^{4}538$	$-.0^{4}572$	$-.0^{4}346$	$-.0^{4}111$	$.0^{5}139$	$.0^{6}211$
	5	$-.0^{3}321$	$-.0^{3}113$	$-.0^{4}478$	$-.0^{5}460$	$.0^{5}655$	$.0^{6}659$
	7	$-.0^{3}803$	$-.0^{3}168$	$-.0^{4}403$	$.0^{4}200$	$.0^{4}157$	$.0^{5}136$
	10	$-.0^{2}193$	$-.0^{3}251$	$.0^{4}101$	$.0^{4}906$	$.0^{4}370$	$.0^{5}291$
	20	$-.0^{2}917$	$-.0^{3}521$	$.0^{3}517$	$.0^{3}619$	$.0^{3}173$	$.0^{4}122$
	50	$-.0631$	$-.0^{2}130$	$.0^{2}516$	$.0^{2}491$	$.0^{2}118$	$.0^{4}790$

3. Derivation of the asymptotic expansion. To obtain the asymptotic expansion of $P_1(c; D)$ we first consider the characteristic function

$$\psi(t) = E\{\exp[itD^{-1}(W - D^2/2)] \mid \Pi_1\} \tag{3.1}$$

of the random variable $D^{-1}(W - D^2/2)$ when $\mathbf{X}_{N_1+N_2+1}$ comes from Π_1. From a well-known property of the conditional probability we have

$$\psi(t) = E^{\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S}}\{E[\exp[itD^{-1}(W - D^2/2)] \mid \bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S}; \Pi_1]\}. \tag{3.2}$$

Since the distribution of W is invariant under any linear transformation performed on $\mathbf{X}_i (i = 1, 2, \cdots, N_1 + N_2 + 1)$ and $\mathbf{S}$, we may suppose

$$\boldsymbol{\mu}^{(1)} = \mathbf{0}, \qquad \boldsymbol{\mu}^{(2)} = \boldsymbol{\mu}_0, \qquad \boldsymbol{\Sigma} = \mathbf{I} \text{ (identity)} \tag{3.3}$$

without any loss of generality, where $\boldsymbol{\mu}_0$ denotes a p-vector with the first component D and the others 0. From the definition of W its conditional distribution given $(\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S})$ is (one-dimensional) $N(\hat{\mu}, \hat{\sigma}^2)$, where

$$\begin{aligned} \hat{\mu} &= \hat{\mu}(\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S}) = -\tfrac{1}{2}(\bar{\mathbf{X}}^{(1)} + \bar{\mathbf{X}}^{(2)})'\mathbf{S}^{-1}(\bar{\mathbf{X}}^{(1)} - \bar{\mathbf{X}}^{(2)}), \\ \hat{\sigma}^2 &= \hat{\sigma}^2(\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S}) = (\bar{\mathbf{X}}^{(1)} - \bar{\mathbf{X}}^{(2)})'\mathbf{S}^{-2}(\bar{\mathbf{X}}^{(1)} - \bar{\mathbf{X}}^{(2)}). \end{aligned} \tag{3.4}$$

Therefore, if we put

$$\Psi(\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S}) = \exp\,[itD^{-1}(\hat{\mu} - \tfrac{1}{2}D^2) - \tfrac{1}{2}t^2D^{-2}\hat{\sigma}^2], \tag{3.5}$$

then

$$\psi(t) = E^{\bar{\mathbf{X}}^{(1)},\bar{\mathbf{X}}^{(2)},\mathbf{S}}\Psi(\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S}). \tag{3.6}$$

Since the function Ψ is analytic about the point $(\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S}) = (\mathbf{0}, \boldsymbol{\mu}_0, \mathbf{I})$ specified by (3.3) it holds that

$$\begin{aligned}\Psi(\bar{\mathbf{X}}^{(1)},\bar{\mathbf{X}}^{(2)}, \mathbf{S}) = \exp\Bigg[\sum_{i=1}^{p} \bar{X}_i^{(1)} \frac{\partial}{\partial\mu_i^{(1)}} + \sum_{i=1}^{p} (\bar{X}_i^{(2)} - \mu_{0i}) \frac{\partial}{\partial\mu_i^{(2)}} \\ + \sum_{i\leqq j=1}^{p} (S_{ij} - \sigma_{ij}) \frac{\partial}{\partial\sigma_{ij}}\Bigg]\Psi(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma})\,|_0\,,\end{aligned} \tag{3.7}$$

where $\bar{X}_i^{(1)}$, $\bar{X}_i^{(2)}$, $\mu_i^{(1)}$, $\mu_i^{(2)}$ and μ_{0i} are components of the vectors $\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \boldsymbol{\mu}^{(1)}$, $\boldsymbol{\mu}^{(2)}$ and $\boldsymbol{\mu}_0$, respectively, σ_{ij} are elements of the symmetric matrix $\boldsymbol{\Sigma}$ which is regarded as a function of $\sigma_{ij} (i \leqq j = 1, \cdots, p)$, δ is Kronecker's symbol and the notation $|_0$ means the value at the point (3.3). If we put

$$\begin{aligned}\boldsymbol{\partial}^{(k)} &= (\partial_i^{(k)})_{i=1,\cdots,p}\,, k = 1, 2, \partial_i^{(k)} = \frac{\partial}{\partial\mu_i^{(k)}}\,, \\ \boldsymbol{\partial} &= (\partial_{ij})_{i,j=1,\cdots,p}\,, \partial_{ij} = \partial_{ji} = \tfrac{1}{2}(1 + \delta_{ij}) \frac{\partial}{\partial\sigma_{ij}} \ (i \leqq j),\end{aligned} \tag{3.8}$$

then (3.7) is written in a matrix form as

$$\begin{aligned}\Psi(\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S}) = \exp\,[\bar{\mathbf{X}}^{(1)\prime}\boldsymbol{\partial}^{(1)} + (\bar{\mathbf{X}}^{(2)} - \boldsymbol{\mu}_0)'\boldsymbol{\partial}^{(2)} \\ + \operatorname{tr}\,(\mathbf{S} - \mathbf{I})\boldsymbol{\partial}]\Psi(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma})\,|_0\,.\end{aligned} \tag{3.9}$$

Substituting (3.9) into (3.6) and using the associative law

$$E(\Delta\Psi) = (E\Delta)\Psi, \tag{3.10}$$

where Δ stands for the factor involving exp in (3.9), which may be justified by extending the theory of asymptotic expansions (see, for instance, Jeffreys (1962), pp. 14–16) to the multi-dimensional case, we obtain

$$\psi(t) = \Theta\Psi(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma})\,|_0\,, \tag{3.11}$$

where Θ is a differential operator defined formally by

$$\Theta = E^{\bar{\mathbf{X}}^{(1)},\bar{\mathbf{X}}^{(2)},\mathbf{S}} \exp\,[\bar{\mathbf{X}}^{(1)\prime}\boldsymbol{\partial}^{(1)} + (\bar{\mathbf{X}}^{(2)} - \boldsymbol{\mu}_0)'\boldsymbol{\partial}^{(2)} + \operatorname{tr}\,(\mathbf{S} - \mathbf{I})\boldsymbol{\partial}]. \tag{3.12}$$

Now the function Ψ in (3.11) is determined by

$$\Psi(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma}) = \exp\,[itD^{-1}(\mu - \tfrac{1}{2}D^2) - \tfrac{1}{2}t^2D^{-2}\sigma^2], \tag{3.13}$$

where

$$\text{(3.14)}\quad \begin{aligned} \mu &= \hat{\mu}(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma}) = -\tfrac{1}{2}(\boldsymbol{\mu}^{(1)} + \boldsymbol{\mu}^{(2)})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)}), \\ \sigma^2 &= \hat{\sigma}^2(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma}) = (\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})'\boldsymbol{\Sigma}^{-2}(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)}) \end{aligned}$$

which follow from (3.5) and (3.4), respectively. On the other hand $\bar{\mathbf{X}}^{(1)}$, $\bar{\mathbf{X}}^{(2)}$ and $n\mathbf{S}$ follow independently $N(\mathbf{0}, N_1^{-1}\mathbf{I})$, $N(\boldsymbol{\mu}_0, N_2^{-1}\mathbf{I})$ and $W(n, \mathbf{I})$, respectively; and hence, using the formula of the moment generating function of each distribution, we have

$$\text{(3.15)}\quad \Theta = \exp\left[\frac{1}{2N_1}\boldsymbol{\partial}^{(1)\prime}\boldsymbol{\partial}^{(1)} + \frac{1}{2N_2}\boldsymbol{\partial}^{(2)\prime}\boldsymbol{\partial}^{(2)} - \operatorname{tr}(\boldsymbol{\partial}) - \frac{n}{2}\log\left|\mathbf{I} - \frac{2}{n}\boldsymbol{\partial}\right|\right].$$

Substituting the expansion $-\log|\mathbf{I} - \mathbf{A}| = \operatorname{tr}(\mathbf{A}) + \frac{1}{2}\operatorname{tr}(\mathbf{A}^2) + \frac{1}{3}\operatorname{tr}(\mathbf{A}^3) + \cdots$, we have

$$\begin{aligned} \Theta &= \exp\left[\frac{1}{2N_1}\boldsymbol{\partial}^{(1)\prime}\boldsymbol{\partial}^{(1)} + \frac{1}{2N_2}\boldsymbol{\partial}^{(2)\prime}\boldsymbol{\partial}^{(2)} + \frac{1}{n}\operatorname{tr}(\boldsymbol{\partial}^2) + \frac{4}{3n^2}\operatorname{tr}(\boldsymbol{\partial}^3) + \cdots\right] \\ &= 1 + \frac{1}{2N_1}\boldsymbol{\partial}^{(1)\prime}\boldsymbol{\partial}^{(1)} + \frac{1}{2N_2}\boldsymbol{\partial}^{(2)\prime}\boldsymbol{\partial}^{(2)} + \frac{1}{n}\operatorname{tr}(\boldsymbol{\partial}^2) \\ &\quad + \frac{1}{8N_1^2}[\boldsymbol{\partial}^{(1)\prime}\boldsymbol{\partial}^{(1)}]^2 + \frac{1}{8N_2^2}[\boldsymbol{\partial}^{(2)\prime}\boldsymbol{\partial}^{(2)}]^2 + \frac{1}{4N_1N_2}\boldsymbol{\partial}^{(1)\prime}\boldsymbol{\partial}^{(1)}\boldsymbol{\partial}^{(2)\prime}\boldsymbol{\partial}^{(2)} \quad \text{(3.16)} \\ &\quad + \frac{1}{2N_1 n}\boldsymbol{\partial}^{(1)\prime}\boldsymbol{\partial}^{(1)}\operatorname{tr}(\boldsymbol{\partial}^2) + \frac{1}{2N_2 n}\boldsymbol{\partial}^{(2)\prime}\boldsymbol{\partial}^{(2)}\operatorname{tr}(\boldsymbol{\partial}^2) \\ &\quad + \frac{1}{n^2}\{\tfrac{1}{2}[\operatorname{tr}(\boldsymbol{\partial}^2)]^2 + \tfrac{4}{3}\operatorname{tr}(\boldsymbol{\partial}^3)\} + O_3 \end{aligned}$$

or

$$\begin{aligned} \Theta &= 1 + \frac{1}{2N_1}\sum[\partial_r^{(1)}]^2 + \frac{1}{2N_2}\sum[\partial_r^{(2)}]^2 + \frac{1}{n}\sum\partial_{rs}^2 \\ &\quad + \frac{1}{8N_1^2}\sum[\partial_r^{(1)}\partial_s^{(1)}]^2 + \frac{1}{8N_2^2}\sum[\partial_r^{(2)}\partial_s^{(2)}]^2 + \frac{1}{4N_1N_2}\sum[\partial_r^{(1)}\partial_s^{(2)}]^2 \quad \text{(3.17)} \\ &\quad + \frac{1}{2N_1 n}\sum[\partial_r^{(1)}\partial_{st}]^2 + \frac{1}{2N_2 n}\sum[\partial_s^{(2)}\partial_{st}]^2 \\ &\quad + \frac{1}{n^2}[\tfrac{1}{2}\sum\partial_{rs}^2\partial_{tu}^2 + \tfrac{4}{3}\sum\partial_{rs}\partial_{st}\partial_{tr}] + O_3, \end{aligned}$$

each subscript r, s, t and u running over the range $1, 2, \cdots, p$.

Now we have to turn from the c.f. $\psi(t)$ to the cdf $P_1(c; D)$. It is shown in Section 5 that from (3.11), (3.13) and (3.17) we have the expression

$$\text{(3.18)}\quad \psi(t) = \sum_{\nu=0}^{8} a_\nu(-it)^\nu e^{-t^2/2} + O_3.$$

The Fourier transform

$$(3.19) \qquad \int_{-\infty}^{\infty} e^{itx}\, d\Phi_{(x)}^{(\nu)} = (-it)^{\nu} e^{-t^2/2}$$

(cf., for example, Cramér (1946), p. 225) then gives

$$(3.20) \qquad P_1(c; D) = \sum_{\nu} a_\nu \Phi^{(\nu)}(c) + O_3 = [\sum_{\nu} a_\nu d^\nu] \Phi(c) + O_3 .$$

By comparing (3.18) and (3.20) it is seen that by using $\theta = -it$ instead of t the expressions become simpler; i.e., to each term $\theta^\nu e^{\theta^2/2}$ of the c.f. corresponds a term $d^\nu \Phi(c)$ of the cdf. We put

$$(3.21) \qquad A = A(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \mathbf{\Sigma}) = -D^{-1}\theta\mu + \tfrac{1}{2}D^{-2}\theta^2\sigma^2 + \tfrac{1}{2}D\theta$$

and rewrite (3.11) as

$$(3.22) \qquad \psi(t) = \Theta e^{A} \,|_0 .$$

In Sections 5 we shall calculate the terms of (3.22), and hence of $P_1(c; D)$, arising from the operation of each term of (3.17) on e^A.

4. Lemmas.

LEMMA 2. *Let* $\mathbf{\Sigma} = (\sigma_{ij})$ *be a non-singular symmetric matrix. We regard two matrices* $\mathbf{\Sigma}^{-1} = (\sigma^{ij})$ *and* $\mathbf{\Sigma}^{-2} = (\sigma_2^{ij})$ *as functions of* σ_{rs} $(r \leqq s)$ *and denote by subscript rs the result obtained by operating* $\partial_{rs} = \frac{1}{2}(1 + \delta_{rs})\partial/\partial\sigma_{rs}$, *then*

$$(4.1) \qquad \begin{aligned} (\sigma^{ij})_{rs} &= -\tfrac{1}{2}(\sigma^{ir}\sigma^{sj} + \sigma^{is}\sigma^{rj}), \\ (\sigma_2^{ij})_{rs} &= -\tfrac{1}{2}(\sigma^{ir}\sigma_2^{sj} + \sigma^{is}\sigma_2^{rj} + \sigma_2^{ir}\sigma^{sj} + \sigma_2^{is}\sigma^{rj}). \end{aligned}$$

If we denote the value at $\mathbf{\Sigma} = \mathbf{I}$ or $\sigma_{ij} = \delta_{ij}$ *by the symbol* $|_0$, *then*

$$(4.2) \qquad \begin{aligned} (\sigma^{ij})_{rs}\,|_0 &= -\tfrac{1}{2}(\delta_{ir}\delta_{sj} + \delta_{is}\delta_{rj}), \\ (\sigma_2^{ij})_{rs}\,|_0 &= -(\delta_{ir}\delta_{sj} + \delta_{is}\delta_{rj}). \end{aligned}$$

PROOF. As in Anderson (1958), p. 348, by differentiating the identity $\mathbf{\Sigma}\mathbf{\Sigma}^{-1} = \mathbf{I}$ we have $(\mathbf{\Sigma})_{rs}\mathbf{\Sigma}^{-1} + \mathbf{\Sigma}(\mathbf{\Sigma}^{-1})_{rs} = \mathbf{0}$; and hence

$$(4.3) \qquad (\mathbf{\Sigma}^{-1})_{rs} = -\mathbf{\Sigma}^{-1}(\mathbf{\Sigma})_{rs}\mathbf{\Sigma}^{-1}.$$

Let $\mathbf{E}_{rs}$ be the matrix with all elements 0 except the (r, s) element which is 1, then $(\mathbf{\Sigma})_{rs} = (\mathbf{E}_{rs} + \mathbf{E}_{sr})/2$. Substituting this into (4.3) we have the first equation of (4.1). Substituting it into

$$(\sigma_2^{ij})_{rs} = (\sum_k \sigma^{ik}\sigma^{kj})_{rs} = \sum_k [(\sigma^{ik})_{rs}\sigma^{kj} + \sigma^{ik}(\sigma^{kj})_{rs}],$$

we have the second equation of (4.1). (4.2) follows from (4.1) at once.

LEMMA 3. *Let subscript rs, rs denote the differential coefficient by* ∂_{rs}^2, *then*

$$(\sigma^{ij})_{rs,rs} = \frac{1}{4}\left[(\sigma^{ir}\sigma^{sr} + \sigma^{is}\sigma^{rr})\sigma^{sj} + \binom{i,\ j}{r,\ s} + (r,\ s)\right],$$

$$(4.4)\quad (\sigma_2^{ij})_{rs,rs} = \frac{1}{4}\left[(\sigma^{ir}\sigma^{sr} + \sigma^{is}\sigma^{rr})\sigma_2^{sj} + \sigma^{ir}(\sigma^{sr}\sigma_2^{sj} + \sigma^{ss}\sigma_2^{rj} + \sigma_2^{sr}\sigma^{sj} + \sigma_2^{ss}\sigma^{rj}) + \binom{i,\ j}{r,\ s} + (r,\ s)\right],$$

where the symbol $\binom{i,\ j}{r,\ s}$ *represents an expression obtained by interchanging i and j together with r and s in the whole expression (in square bracket) appearing in the left of the symbol and (r, s) analogously. Moreover*

$$(4.5)\quad \begin{aligned}(\sigma^{ij})_{rs,rs}\mid_0 &= \tfrac{1}{2}(\delta_{ijr} + \delta_{ijs} + 2\delta_{ijrs}),\\ (\sigma_2^{ij})_{rs,rs}\mid_0 &= \tfrac{3}{2}(\delta_{ijr} + \delta_{ijs} + 2\delta_{ijrs}),\end{aligned}$$

where $\delta_{rs\cdots v}$ *in the sequel is* 1 *when* $r = s = \cdots = v$ *and* 0 *otherwise.*

LEMMA 4.

$$(4.6)\quad (\sigma^{11})_{rs,tu}\mid_0 = \tfrac{1}{3}(\sigma_2^{11})_{rs,tu}\mid_0 = \tfrac{1}{2}(\delta_{1rt}\delta_{su} + \delta_{1ru}\delta_{st} + \delta_{1st}\delta_{ru} + \delta_{1su}\delta_{rt}),$$

$$(4.7)\quad \begin{aligned}(\sigma^{11})_{rs,rs,tu}\mid_0 &= \tfrac{1}{4}(\sigma_2^{11})_{rs,rs,tu}\mid_0\\ &= -\tfrac{1}{2}[\delta_{1r}\delta_{stu} + \delta_{1s}\delta_{rtu} + 2(\delta_{1rtu} + \delta_{1stu}) + 6\delta_{1rstu}],\end{aligned}$$

$$(4.8)\quad \begin{aligned}(\sigma^{11})_{rs,st,tr}\mid_0 &= \tfrac{1}{4}(\sigma_2^{11})_{rs,st,tr}\mid_0\\ &= -\tfrac{1}{4}[\delta_{1r} + \delta_{1s} + \delta_{1t} + (\delta_{1r}\delta_{st} + \delta_{1s}\delta_{tr} + \delta_{1t}\delta_{rs})\\ &\quad + 2(\delta_{1rs} + \delta_{1st} + \delta_{1tr}) + 12\delta_{1rst}],\end{aligned}$$

$$(4.9)\quad \begin{aligned}(\sigma^{11})_{rs,rs,tu,tu}\mid_0 &= \tfrac{1}{5}(\sigma_2^{11})_{rs,rs,tu,tu}\mid_0\\ &= \tfrac{1}{4}[\delta_{1r}(\delta_{st} + \delta_{su}) + \delta_{1s}(\delta_{rt} + \delta_{ru}) + \delta_{1t}(\delta_{ru} + \delta_{su})\\ &\quad + \delta_{1u}(\delta_{rt} + \delta_{st}) + 2(\delta_{1rt} + \delta_{1ru} + \delta_{1st} + \delta_{1su})\\ &\quad + 2(\delta_{1r}\delta_{stu} + \delta_{1s}\delta_{rtu} + \delta_{1t}\delta_{rsu} + \delta_{1u}\delta_{rst})\\ &\quad + 2(\delta_{1rt}\delta_{su} + \delta_{1ru}\delta_{st} + \delta_{1st}\delta_{ru} + \delta_{1su}\delta_{rt})\\ &\quad + 6(\delta_{1rst} + \delta_{1rsu} + \delta_{1rtu} + \delta_{1stu}) + 40\delta_{1rstu}].\end{aligned}$$

The proofs of Lemmas 3 and 4 are omitted. Details are given in the author's seminar report (1962).

5. Calculation of each term in the asymptotic expansion. If it is shown that

$$(5.1)\quad \psi(t) = [1 + L(\theta; D) + Q(\theta; D)]e^{\theta^2/2} + O_3,$$

where $\theta = -it$ and the functional forms of L and Q are given by (2.3) through

(2.6), then we have (2.2) in view of the correspondence of c.f. and cdf described in the last paragraph of Section 3. Since the calculation is straightforward we shall only sketch it, leaving details again to [10].

5.1. *The principal term.* From (3.17) and (3.22) the principal term in $\psi(t)$ is

$$e^A \mid_0 = e^{\theta^2/2},$$

since (3.14) gives

$$\mu = \hat{\mu}(\mathbf{0}, \boldsymbol{\mu}_0, \mathbf{I}) = \tfrac{1}{2}\boldsymbol{\mu}_0'\boldsymbol{\mu}_0 = \tfrac{1}{2}D^2, \qquad (5.2)$$
$$\sigma^2 = \hat{\sigma}^2(\mathbf{0}, \boldsymbol{\mu}_0, \mathbf{I}) = \boldsymbol{\mu}_0'\boldsymbol{\mu}_0 = D^2,$$

and hence $A \mid_0 = \theta^2/2$ from (3.21).

5.2. *The linear terms.* First the term of order N_1^{-1} in $\psi(t)$ is

$$(2N_1)^{-1}\sum_{r=1}^{p} [\partial_r^{(1)}]^2 e^A \mid_0 . \qquad (5.3)$$

Since differentiation is concerned with only $\boldsymbol{\mu}^{(1)}$ it makes the calculation simpler to put $\boldsymbol{\mu}^{(2)} = \boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma} = \mathbf{I}$ before differentiation and put $\boldsymbol{\mu}^{(1)} = \mathbf{0}$ after. Then (3.14) is written as

$$\mu = \hat{\mu}(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}_0, \mathbf{I}) = -\tfrac{1}{2}(\boldsymbol{\mu}^{(1)} + \boldsymbol{\mu}_0)'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}_0), \qquad (5.4)$$
$$\sigma^2 = \hat{\sigma}^2(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}_0, \mathbf{I}) = (\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}_0)'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}_0).$$

If we denote the result of $\partial_r^{(1)}$ and $[\partial_r^{(1)}]^2$ by subscript r and r, r, respectively, then

$$[\partial_r^{(1)}]^2 e^A = (A_r^2 + A_{r,r})e^A. \qquad (5.5)$$

Therefore (5.3) is written as

$$(2N_1)^{-1}\left[\sum_{r=1}^{p} A_r^2 \mid_0 + \sum_{r=1}^{p} A_{r,r} \mid_0\right] e^{\theta^2/2} = L_1(\theta; D)e^{\theta^2/2} \quad \text{(say)}. \qquad (5.6)$$

Differentiation of (3.21) yields

$$A_r = -D^{-1}\theta(\mu)_r + \tfrac{1}{2}D^{-2}\theta^2(\sigma^2)_r = D^{-1}\theta\mu_r^{(1)} + D^{-2}\theta^2(\mu_r^{(1)} - \mu_{0r}), \qquad (5.7)$$
$$A_{r,r} = D^{-1}\theta + D^{-2}\theta^2,$$

and hence

$$A_r \mid_0 = -D^{-2}\theta^2\mu_{0r} = -D^{-1}\theta^2\delta_{1r}, \qquad A_{r,r} \mid_0 = D^{-2}(\theta^2 + D\theta). \qquad (5.8)$$

Thus we have

$$L_1(\theta; D) = (2N_1D^2)^{-1}[\theta^4 + p(\theta^2 + D\theta)]. \qquad (5.9)$$

Similarly the term of order N_2^{-1} is $L_2(\theta; D)e^{\theta^2/2}$ where

$$L_2(\theta; D) = (2N_2D^2)^{-1}[(\theta^2 - D\theta)^2 + p(\theta^2 - D\theta)]. \qquad (5.10)$$

For the later reference we state the result

$$(5.11)\qquad A_r\,|_0 = D^{-1}(\theta^2 - D\theta)\delta_{1r}\,, \qquad A_{r,r}\,|_0 = D^{-2}(\theta^2 - D\theta),$$

which corresponds to (5.8).

The last linear term in $\psi(t)$ is of order n^{-1}; that is,

$$(5.12)\qquad n^{-1}\sum_{r,s=1}^{p} \partial_{rs}^2 e^A\,|_0 = L_3(\theta; D)e^{\theta^2/2} \quad \text{(say)}.$$

This time we put $\mathbf{\mu}^{(1)} = \mathbf{0}$ and $\mathbf{\mu}^{(2)} = \mathbf{\mu}_0$ before differentiation and put $\mathbf{\Sigma} = \mathbf{I}$ after. Then (3.14) becomes

$$(5.13)\qquad \begin{aligned} \mu &= \hat{\mu}(\mathbf{0}, \mathbf{\mu}_0, \mathbf{\Sigma}) = \tfrac{1}{2}\mathbf{\mu}_0'\mathbf{\Sigma}^{-1}\mathbf{\mu}_0 = \tfrac{1}{2}D^2\sigma^{11}, \\ \sigma^2 &= \hat{\sigma}^2(\mathbf{0}, \mathbf{\mu}_0, \mathbf{\Sigma}) = \mathbf{\mu}_0'\mathbf{\Sigma}^{-2}\mathbf{\mu}_0 = D^2\sigma_2^{11}, \end{aligned}$$

where we put $\mathbf{\Sigma}^{-1} = (\sigma^{ij})$ and $\mathbf{\Sigma}^{-2} = (\sigma_2^{ij})$. Using the notation in Section 4, we have

$$(5.14)\qquad L_3(\theta; D) = n^{-1}\left[\sum_{r,s=1}^{p} A_{rs}^2\,|_0 + \sum_{r,s=1}^{p} A_{rs,rs}\,|_0\right].$$

From (3.21) and (5.13) follows

$$(5.15)\qquad \begin{aligned} A_{rs} &= -\tfrac{1}{2}D\theta(\sigma^{11})_{rs} + \tfrac{1}{2}\theta^2(\sigma_2^{11})_{rs}\,, \\ A_{rs,rs} &= -\tfrac{1}{2}D\theta(\sigma^{11})_{rs,rs} + \tfrac{1}{2}\theta^2(\sigma_2^{11})_{rs,rs}\,, \end{aligned}$$

which implies in view of (4.2) and (4.5) that

$$(5.16)\qquad \begin{aligned} A_{rs}\,|_0 &= -\tfrac{1}{2}(2\theta^2 - D\theta)\delta_{1rs}\,, \\ A_{rs,rs}\,|_0 &= \tfrac{1}{4}(3\theta^2 - D\theta)(\delta_{1r} + \delta_{1s} + 2\delta_{1rs}). \end{aligned}$$

Substituting this into (5.14), we obtain

$$(5.17)\qquad L_3(\theta; D) = (4n)^{-1}[(2\theta^2 - D\theta)^2 + 2(p+1)(3\theta^2 - D\theta)].$$

Combining these results, we have the linear terms of $\psi(t)$ in the form $[\sum_{i=1}^{3} L_i(\theta; D)]e^{\theta^2/2}$.

5.3. *The quadratic terms.* Having settled the linear terms, we shall now turn to the second order terms. First the term of order N_1^{-2} is

$$(5.18)\qquad (8N_1^2)^{-1}\sum_{r,s=1}^{p}[\partial_r^{(1)}\partial_s^{(1)}]^2 e^A\,|_0 = Q_{11}'(\theta; D)e^{\theta^2/2},$$

where we have to determine Q_{11}'. Further differentiation of (5.5) gives

$$(5.19)\qquad \begin{aligned} [\partial_r^{(1)}\partial_s^{(1)}]^2 e^A = [&(A_r^2 + A_{r,r})(A_s^2 + A_{s,s}) + 4A_rA_sA_{r,s} \\ &+ 2A_{r,s}^2 + 2(A_rA_{r,s,s} + A_sA_{r,r,s}) + A_{r,r,s,s}]e^A. \end{aligned}$$

Since (5.7) implies that $A_{r,s}$ for $r \neq s$ as well as every derivative of A of more than second order are zero, it holds that

$$\sum_{r,s=1}^{p} [\partial_r^{(1)}\partial_s^{(1)}]^2 e^A = \left\{\left[\sum_{r=1}^{p} (A_r^2 + A_{r,r})\right]^2 + \sum_{r=1}^{p} (4A_r^2A_{r,r} + 2A_{r,r}^2)\right\} e^A.$$

If we put

$$(5.20)\quad \begin{aligned} Q_{11}(\theta; D) &= (8N_1^2)^{-1}\sum_{r=1}^{p} (4A_r^2A_{r,r} + 2A_{r,r}^2) \mid_0 \\ &= (4N_1^2D^4)^{-1}[2\theta^4(\theta^2 + D\theta) + p(\theta^2 + D\theta)^2] \end{aligned}$$

which follows from (5.8), then we have

$$(5.21)\quad Q_{11}'(\theta; D) = \tfrac{1}{2}[L_1(\theta; D)]^2 + Q_{11}(\theta; D)$$

with L_1 defined in (5.9). Thus the term of order N_1^{-2} is written as

$$(5.22)\quad \{\tfrac{1}{2}[L_1(\theta; D)]^2 + Q_{11}(\theta; D)\}e^{\theta^2/2}.$$

Similarly the term of order N_2^{-2} is

$$(5.23)\quad \{\tfrac{1}{2}[L_2(\theta; D)]^2 + Q_{22}(\theta; D)\}e^{\theta^2/2}$$

with L_2 and Q_{22} defined in (5.10) and (2.6), respectively.

The cross term of order $(N_1N_2)^{-1}$ can be written as

$$(5.24)\quad (4N_1N_2)^{-1}\sum_{r,s=1}^{p} [\partial_r^{(1)}\partial_s^{(2)}]^2 e^A \mid_0 = [L_1(\theta; D)L_2(\theta; D) + Q_{12}(\theta; D)]e^{\theta^2/2}.$$

It remains to determine Q_{12}. As with (5.20) it holds that

$$(5.25)\quad \begin{aligned} Q_{12}(\theta; D) &= (4N_1N_2)^{-1}\sum_{r,s=1}^{p} (4A_rA_sA_{r,s} + 2A_{r,s}^2) \mid_0 \\ &= (2N_1N_2D^4)^{-1}[2\theta^4(\theta^2 - D\theta) + p\theta^4] \end{aligned}$$

by substituting $A_r \mid_0$ and $A_s \mid_0$ given in (5.8) and (5.11), respectively, and also $A_{r,s} \mid_0 = -D^{-2}\theta^2\delta_{rs}$ which is easily verified.

Similarly the term of order $(N_1n)^{-1}$ is

$$(5.26)\quad (2N_1n)^{-1}\sum_{r,s,t=1}^{p} [\partial_r^{(1)}\partial_{st}]^2 e^A \mid_0 = [L_1(\theta; D)L_3(\theta; D) + Q_{13}(\theta; D)]e^{\theta^2/2}.$$

We shall determine Q_{13}. Adopting the notation which we have already used repeatedly, we have

$$(5.27)\quad \begin{aligned} [\partial_r^{(1)}\partial_{st}]^2 e^A &= [(A_r^2 + A_{r,r})(A_{st}^2 + A_{st,st}) + 4A_rA_{st}A_{r,st} \\ &\quad + 2A_{r,st}^2 + 2(A_rA_{r,st,st} + A_{st}A_{r,r,st}) + A_{r,r,st,st}]e^A, \end{aligned}$$

so that

$$(5.28)\quad \begin{aligned} Q_{13}(\theta; D) &= (2N_1n)^{-1}\sum_{r,s,t=1}^{p} [4A_rA_{st}A_{r,st} + 2A_{r,st}^2 \\ &\quad + 2A_rA_{r,st,st} + 2A_{st}A_{r,r,st} + A_{r,r,st,st}] \mid_0 . \end{aligned}$$

Differentiating (3.21) and using (4.2) and (4.5), we obtain the values of four quantities $A_{r,st}$, $A_{r,st,st}$, $A_{r,r,st}$ and $A_{r,r,st,st}$ at the point specified in (3.3), which determine the functional form of Q_{13} as shown in (2.6).

Analogous method gives

$$[L_2(\theta; D)L_3(\theta; D) + Q_{23}(\theta; D)]e^{\theta^2/2} \tag{5.29}$$

as the term of order $(N_2 n)^{-1}$, where L_2, L_3 and Q_{23} are given in (5.10), (5.14) and (2.6), respectively.

Last, the term of order n^{-2} is

$$\begin{aligned} n^{-2}\left[\frac{1}{2}\sum_{r,s,t,u=1}^{p} \partial_{rs}^2\partial_{tu}^2 + \frac{4}{3}\sum_{r,s,t=1}^{p} \partial_{rs}\partial_{st}\partial_{tr}\right]e^A \,\Big|_0 \\ = \{\tfrac{1}{2}[L_3(\theta; D)]^2 + Q_{33}(\theta; D)\}e^{\theta^2/2}. \end{aligned} \tag{5.30}$$

Q_{33} is obtained as follows. From two relations

$$\begin{aligned} \partial_{rs}^2\partial_{tu}^2 e^A = [(A_{rs}^2 + A_{rs,rs})(A_{tu}^2 + A_{tu,tu}) + 4A_{rs}A_{tu}A_{rs,tu} \\ + 2A_{rs,tu}^2 + 2(A_{rs}A_{rs,tu,tu} + A_{tu}A_{rs,rs,tu}) + A_{rs,rs,tu,tu}]e^A, \end{aligned}$$

and

$$\partial_{rs}\partial_{st}\partial_{tr}e^A = (A_{rs}A_{st}A_{tr} + A_{rs}A_{st,tr} + A_{st}A_{tr,rs} + A_{tr}A_{rs,st} + A_{rs,st,tr})e^A$$

it follows that

$$\begin{aligned} Q_{33}(\theta; D) = (2n^2)^{-1}\sum_{r,s,t,u=1}^{p} (4A_{rs}A_{tu}A_{rs,tu} + 2A_{rs,tu}^2 \\ + 4A_{tu}A_{rs,rs,tu} + A_{rs,rs,tu,tu}) \,\Big|_0 \\ + 4(3n^2)^{-1}\sum_{r,s,t=1}^{p} (A_{rs}A_{st}A_{tr} + 3A_{tr}A_{rs,st} + A_{rs,st,tr}) \,\Big|_0 \end{aligned} \tag{5.31}$$

Since A_{rs} as well as $A_{rs,rs}$ have already been computed and since the values of four quantities $A_{rs,tu}$, $A_{rs,rs,tu}$, $A_{rs,st,tr}$ and $A_{rs,rs,tu,tu}$ at the point (3.3) are given by Lemma 4, we can write Q_{33} in the form shown in (2.6).

Combining these results, we can represent the second order terms in the formula

$$\left\{\frac{1}{2}\left[\sum_{i=1}^{3} L_i(\theta; D)\right]^2 + \sum_{i \leqq j=1}^{3} Q_{ij}(\theta; D)\right\} e^{\theta^2/2},$$

which completes the calculation.

REFERENCES

[1] Anderson, T. W. (1951). Classification by multivariate analysis. *Psychometrika* **16** 31–50.

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

[3] Bowker, Albert H. (1960). A representation of Hotelling's T^2 and Anderson's classi-

fication statistic W in terms of simple statistics. Chapter 12 in Olkin, I. (ed.). *Contributions to Probability and Statistics*, Stanford Univ. Press.

[4] Bowker, Albert H. and Sitgreaves, Rosedith (1961). An asymptotic expansion for the distribution function of the W-classification statistic. Chapter 19 in Solomon, H. (ed.). *Studies in Item Analysis and Prediction*, Stanford Univ. Press.

[5] Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton Univ. Press.

[6] Hartley, H. O. (1938). Studentization and large-sample theory. *J. Roy. Statist. Soc. Suppl.* **5** 80–88.

[7] Ito, Koichi (1956), (1960). Asymptotic formulae for the distribution of Hotelling's generalized T_0^2 statistic. I. *Ann. Math. Statist.* **27** 1091–1105; II. *Ann. Math. Statist.* **31** 1148–1153.

[8] Jeffreys, Harold (1962). *Asymptotic Approximations*. Oxford Univ. Press.

[9] Okamoto, Masashi (1961). Discrimination for variance matrices. *Osaka Math. J.* **13** 1–39.

[10] Okamoto, Masashi (1962). Asymptotic expansion for the distribution of the linear discriminant function. (in Japanese). *Sem. Rep. Osaka Statist. Assoc.* **6** 97–117.

[11] Siotani, Minoru (1957). On the distributions of the Hotelling's T^2-statistics. *Ann. Inst. Statist. Math. Tokyo* **8** 1–14.

[12] Sitgreaves, Rosedith (1952). On the distribution of two random matrices used in classification procedures. *Ann. Math. Statist.* **23** 263–270.

[13] Sitgreaves, Rosedith (1961). Some results of the distribution of the W-classification statistic. Chapter 15 in Solomon, H. (ed.). *Studies in Item Analysis and Prediction*, Stanford Univ. Press.

[14] Welch, B. L. (1947). The generalization of "Student's" problem when several different population variances are involved. *Biometrika* **34** 28–35.