

# AN ASYMPTOTIC EXPANSION FOR THE DISTRIBUTION OF THE LINEAR DISCRIMINANT FUNCTION<sup>1</sup>

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

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

where

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

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

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

Now it is well known that if the Mahalanobis distance  $D^2$  between two populations  $\Pi_1$  and  $\Pi_2$ , defined by

$$(1.4) \quad D^2 = (\mathbf{u}^{(1)} - \mathbf{u}^{(2)})' \mathbf{\Sigma}^{-1} (\mathbf{u}^{(1)} - \mathbf{u}^{(2)}),$$

is not zero, then as  $N_1$ ,  $N_2$  and  $n$  tend to infinity  $\bar{\mathbf{X}}^{(1)} \rightarrow \mathbf{u}^{(1)}$ ,  $\bar{\mathbf{X}}^{(2)} \rightarrow \mathbf{u}^{(2)}$  and  $\mathbf{S} \rightarrow \mathbf{\Sigma}$  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.

<sup>1</sup> 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  $\Pi_1$  or from  $\Pi_2$ ; that is, if for any real constant  $c$  we denote by

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

the probability that  $W < \frac{1}{2}D^2 + cD$  when  $\mathbf{X}_{N_1+N_2+1}$  comes from  $\Pi_1$  and similarly

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

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

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

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

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

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

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

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

$$(2.5) \quad \begin{aligned} L_1(d; D) &= (2N_1D^2)^{-1}[d^4 + p(d^2 + Dd)], \\ L_2(d; D) &= (2N_2D^2)^{-1}[(d^2 - Dd)^2 + p(d^2 - Dd)], \\ L_3(d; D) &= (4n)^{-1}[(2d^2 - Dd)^2 + 2(p + 1)(3d^2 - Dd)], \\ Q_{11}(d; D) &= (4N_1^2D^4)^{-1}[2d^4(d^2 + Dd) + p(d^2 + Dd)^2], \\ Q_{22}(d; D) &= (4N_2^2D^4)^{-1}[2(d^2 - Dd)^3 + p(d^2 - Dd)^2], \end{aligned}$$

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

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<sup>2</sup>:

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

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

PROOF. We have only to substitute the identities

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

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  $\Pi_1$  or  $\Pi_2$  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.

$$\begin{aligned}
(2.9) \quad P_r\{W < 0 \mid \Pi_1\} &= \Phi\left(-\frac{D}{2}\right) + \frac{a_1}{N_1} + \frac{a_2}{N_2} + \frac{a_3}{n} \\
&\quad + \frac{b_{11}}{N_1^2} + \frac{b_{22}}{N_2^2} + \frac{b_{12}}{N_1N_2} + \frac{b_{13}}{N_1n} + \frac{b_{23}}{N_2n} + \frac{b_{33}}{n^2} + O_3, \\
(2.10) \quad P_r\{W > 0 \mid \Pi_2\} &= \Phi\left(-\frac{D}{2}\right) + \frac{a_2}{N_1} + \frac{a_1}{N_2} + \frac{a_3}{n} \\
&\quad + \frac{b_{22}}{N_1^2} + \frac{b_{11}}{N_2^2} + \frac{b_{12}}{N_1N_2} + \frac{b_{23}}{N_1n} + \frac{b_{13}}{N_2n} + \frac{b_{33}}{n^2} + O_3,
\end{aligned}$$

<sup>2</sup> 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

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

and  $d_0^i$  are constants defined by

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

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

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

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

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

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

$$(2.14) \quad Dd_0^i = 2d_0^{i+1} + 2(i - 1)d_0^{i-1} \quad (i = 1, 2, \dots, 7).$$

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  $\sum b_{ij}/(N_i N_j)$  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 <sup>3</sup> 78	-0.0 <sup>3</sup> 78	-0.0 <sup>7</sup> 56	0	0	0
	2	0.03025	0.03025	0	-0.0 <sup>3</sup> 78	-0.0 <sup>3</sup> 78	-0.0 <sup>7</sup> 56	0	0	0
	3	0.02428	0.02428	0	-0.0 <sup>3</sup> 14	-0.0 <sup>3</sup> 14	-0.0 <sup>7</sup> 228	0	0	0
	4	0.01350	0.01350	0	0.0 <sup>8</sup> 44	0.0 <sup>8</sup> 44	0.0 <sup>1</sup> 69	0	0	0
	6	0.0 <sup>1</sup> 662	0.0 <sup>1</sup> 662	0	0.0 <sup>6</sup> 23	0.0 <sup>6</sup> 23	0.0 <sup>1</sup> 25	0	0	0
	8	0.0 <sup>4</sup> 690	0.0 <sup>4</sup> 690	0	0.0 <sup>5</sup> 44	0.0 <sup>5</sup> 44	0.0 <sup>1</sup> 09	0	0	0
	1	0.2861	-0.06601	0.08802	0.164	-0.0231	0.141	-0.0412	0.0 <sup>2</sup> 75	-1.06
	2	0.1210	0	0.1210	0.0227	-0.0 <sup>7</sup> 56	0.0151	0.0907	-0.0302	-1.39
2	3	0.05666	0.01349	0.09714	0.0107	-0.0 <sup>3</sup> 80	0.0 <sup>6</sup> 94	0.108	-0.0159	-1.03
	4	0.02362	0.01012	0.05399	0.0 <sup>6</sup> 48	-0.0 <sup>4</sup> 85	0.0 <sup>5</sup> 80	0.0759	0.0 <sup>1</sup> 69	-0.499
	6	0.0 <sup>2</sup> 216	0.0 <sup>1</sup> 477	0.0 <sup>6</sup> 648	0.0 <sup>1</sup> 27	0.0 <sup>4</sup> 26	0.0 <sup>1</sup> 69	0.0138	0.0 <sup>2</sup> 425	-0.0366
	8	0.0 <sup>7</sup> 944	0.0 <sup>4</sup> 6272	0.0 <sup>2</sup> 676	0.0 <sup>7</sup> 98	0.0 <sup>4</sup> 62	0.0 <sup>1</sup> 26	0.0 <sup>7</sup> 97	0.0 <sup>3</sup> 404	-0.0 <sup>4</sup> 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 <sup>3</sup> 78	-0.0 <sup>7</sup> 56	0.181	-0.0605	-2.66
	3	0.08904	0.0 <sup>2</sup> 698	0.1943	0.0306	-0.0 <sup>4</sup> 51	0.0 <sup>4</sup> 47	0.257	-0.0452	-1.89
	4	0.03374	0.0 <sup>6</sup> 749	0.1080	0.0152	-0.0 <sup>1</sup> 69	0.0 <sup>6</sup> 75	0.182	-0.0 <sup>6</sup> 75	-0.864
	6	0.0 <sup>2</sup> 770	0.0 <sup>1</sup> 293	0.01330	0.0 <sup>2</sup> 10	0.0 <sup>3</sup> 254	0.0 <sup>1</sup> 98	0.0321	0.0 <sup>7</sup> 02	-0.0399
	8	0.0 <sup>4</sup> 9199	0.0 <sup>4</sup> 5854	0.0 <sup>5</sup> 352	0.0 <sup>1</sup> 10	0.0 <sup>3</sup> 87	0.0 <sup>1</sup> 40	0.0 <sup>1</sup> 78	0.0 <sup>7</sup> 44	0.0 <sup>2</sup> 14
	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
5	3	0.1538	-0.01889	0.3886	0.0941	-0.0 <sup>4</sup> 890	-0.0356	0.676	-0.144	-3.16
	4	0.05399	0	0.2160	0.0418	-0.0 <sup>2</sup> 11	-0.0 <sup>8</sup> 44	0.486	-0.0540	-1.19
	6	0.0 <sup>3</sup> 878	0.0 <sup>4</sup> 233	0.02659	0.0 <sup>4</sup> 34	-0.0 <sup>1</sup> 28	0.0 <sup>2</sup> 11	0.0820	0.0 <sup>8</sup> 13	0.0532
	8	0.0 <sup>1</sup> 171	0.0 <sup>4</sup> 5018	0.0 <sup>1</sup> 071	0.0 <sup>1</sup> 83	0.0 <sup>2</sup> 52	0.0 <sup>1</sup> 58	0.0 <sup>4</sup> 32	0.0 <sup>1</sup> 24	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*6749	0.3239	0.0806	-0.0*422	-0.0211	0.911	-0.142	-0.972
	6	0.0*4986	0.0*5540	0.03989	0.0*733	-0.0*177	0.0*161	0.150	0.0*332	0.279
10	8	0.0*1422	0.0*4181	0.0*1606	0.0*274	0.0*140	0.0*163	0.0*760	0.0*148	0.0337
	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*606	-0.0752	1.78	-0.349	0.364
20	6	0.0*6648	0	0.05983	0.0133	-0.0*231	-0.0*277	0.284	-0.0150	0.868
	8	0.0*1798	0.0*2927	0.0*2409	0.0*445	0.0*129	0.0*145	0.0139	0.0*137	0.0813
	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
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*1847	0.1263	0.0453	0.0*126	-0.0166	1.02	-0.172	4.99
	8	0.0*3053	-0.0*1255	0.0*5084	0.0*130	-0.0*492	-0.0*132	0.0473	-0.0*306	0.388
	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*7386	0.3257	0.255	0.0211	-0.158	5.89	-1.53	37.3
	8	0.0*6815	-0.0*1380	0.01311	0.0*657	0.0*310	-0.0*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 <sup>2</sup> 1350	.0 <sup>4</sup> 3167
Term of the first order	$p = 1$	.0 <sup>3</sup> 4401	.0 <sup>6</sup> 049	.0 <sup>3</sup> 4857	.0 <sup>2</sup> 2700	.0 <sup>4</sup> 3324	.0 <sup>5</sup> 1338
	2	.0 <sup>2</sup> 2645	.0 <sup>2</sup> 1821	.0 <sup>2</sup> 1192	.0 <sup>6</sup> 101	.0 <sup>4</sup> 7051	.0 <sup>5</sup> 2774
	3	.0 <sup>2</sup> 4850	.0 <sup>3</sup> 037	.0 <sup>2</sup> 1899	.0 <sup>9</sup> 503	.0 <sup>8</sup> 1078	.0 <sup>5</sup> 4208
	5	.0 <sup>9</sup> 260	.0 <sup>5</sup> 469	.0 <sup>3</sup> 312	.0 <sup>2</sup> 1631	.0 <sup>8</sup> 1823	.0 <sup>5</sup> 7079
	7	.01367	.0 <sup>2</sup> 7901	.0 <sup>4</sup> 724	.0 <sup>2</sup> 2311	.0 <sup>2</sup> 2568	.0 <sup>5</sup> 9949
	10	.02028	.01155	.0 <sup>6</sup> 844	.0 <sup>2</sup> 3331	.0 <sup>3</sup> 686	.0 <sup>4</sup> 1426
	20	.04233	.02371	.01391	.0 <sup>6</sup> 733	.0 <sup>7</sup> 413	.0 <sup>4</sup> 2861
	50	.1085	.06019	.03510	.01694	.0 <sup>2</sup> 1859	.0 <sup>4</sup> 7166
Term of the sec- ond or- der	$p = 1$	-.0 <sup>5</sup> 151	-.0 <sup>5</sup> 151	-.0 <sup>6</sup> 455	.0 <sup>6</sup> 337	.0 <sup>6</sup> 249	.0 <sup>7</sup> 217
	2	-.0 <sup>6</sup> 083	-.0 <sup>4</sup> 294	-.0 <sup>4</sup> 201	-.0 <sup>6</sup> 766	.0 <sup>6</sup> 320	.0 <sup>7</sup> 841
	3	-.0 <sup>5</sup> 538	-.0 <sup>4</sup> 572	-.0 <sup>4</sup> 346	-.0 <sup>4</sup> 111	.0 <sup>6</sup> 139	.0 <sup>6</sup> 211
	5	-.0 <sup>3</sup> 321	-.0 <sup>3</sup> 113	-.0 <sup>4</sup> 478	-.0 <sup>5</sup> 460	.0 <sup>6</sup> 655	.0 <sup>6</sup> 659
	7	-.0 <sup>8</sup> 803	-.0 <sup>3</sup> 168	-.0 <sup>4</sup> 403	.0 <sup>4</sup> 200	.0 <sup>4</sup> 157	.0 <sup>5</sup> 136
	10	-.0 <sup>2</sup> 193	-.0 <sup>3</sup> 251	.0 <sup>4</sup> 101	.0 <sup>4</sup> 906	.0 <sup>4</sup> 370	.0 <sup>5</sup> 291
	20	-.0 <sup>9</sup> 17	-.0 <sup>3</sup> 521	.0 <sup>3</sup> 517	.0 <sup>6</sup> 19	.0 <sup>3</sup> 173	.0 <sup>4</sup> 122
	50	-.0631	-.0 <sup>2</sup> 130	.0 <sup>2</sup> 516	.0 <sup>4</sup> 491	.0 <sup>2</sup> 118	.0 <sup>4</sup> 790

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

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

of the random variable  $D^{-1}(W - D^2/2)$  when  $\mathbf{X}_{N_1+N_2+1}$  comes from  $\Pi_1$ . From a well-known property of the conditional probability we have

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

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

$$(3.3) \quad \mathbf{u}^{(1)} = \mathbf{0}, \quad \mathbf{u}^{(2)} = \mathbf{u}_0, \quad \Sigma = \mathbf{I} \text{ (identity)}$$

without any loss of generality, where  $\mathbf{u}_0$  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

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

Therefore, if we put

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

then

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

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

$$(3.7) \quad \begin{aligned} \Psi(\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S}) = \exp \left[ \sum_{i=1}^p \bar{X}_i^{(1)} \frac{\partial}{\partial \mu_i^{(1)}} + \sum_{i=1}^p (\bar{X}_i^{(2)} - \mu_{0i}) \frac{\partial}{\partial \mu_i^{(2)}} \right. \\ \left. + \sum_{i \leq j=1}^p (S_{ij} - \sigma_{ij}) \frac{\partial}{\partial \sigma_{ij}} \right] \Psi(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \mathbf{\Sigma})|_0, \end{aligned}$$

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

$$(3.8) \quad \begin{aligned} \mathfrak{a}^{(k)} &= (\partial_i^{(k)})_{i=1, \dots, p}, k = 1, 2, \partial_i^{(k)} = \frac{\partial}{\partial \mu_i^{(k)}}, \\ \mathfrak{a} &= (\partial_{ij})_{i, j=1, \dots, p}, \partial_{ij} = \partial_{ji} = \frac{1}{2}(1 + \delta_{ij}) \frac{\partial}{\partial \sigma_{ij}} (i \leq j), \end{aligned}$$

then (3.7) is written in a matrix form as

$$(3.9) \quad \begin{aligned} \Psi(\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S}) = \exp [\bar{\mathbf{X}}^{(1)'} \mathfrak{a}^{(1)} + (\bar{\mathbf{X}}^{(2)} - \mathbf{u}_0)' \mathfrak{a}^{(2)} \\ + \text{tr} (\mathbf{S} - \mathbf{I}) \mathfrak{a}] \Psi(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \mathbf{\Sigma})|_0. \end{aligned}$$

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

$$(3.10) \quad E(\Delta \Psi) = (E\Delta) \Psi,$$

where  $\Delta$  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

$$(3.11) \quad \psi(t) = \Theta \Psi(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \mathbf{\Sigma})|_0,$$

where  $\Theta$  is a differential operator defined formally by

$$(3.12) \quad \Theta = E^{\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \mathbf{S}} \exp [\bar{\mathbf{X}}^{(1)'} \mathfrak{a}^{(1)} + (\bar{\mathbf{X}}^{(2)} - \mathbf{u}_0)' \mathfrak{a}^{(2)} + \text{tr} (\mathbf{S} - \mathbf{I}) \mathfrak{a}].$$

Now the function  $\Psi$  in (3.11) is determined by

$$(3.13) \quad \Psi(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \mathbf{\Sigma}) = \exp [itD^{-1}(\mu - \frac{1}{2}D^2) - \frac{1}{2}t^2D^{-2}\sigma^2],$$



where

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

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  $N(\mathbf{0}, N_1^{-1}\mathbf{I})$ ,  $N(\mathbf{u}_0, N_2^{-1}\mathbf{I})$  and  $W(n, \mathbf{I})$ , respectively; and hence, using the formula of the moment generating function of each distribution, we have

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

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

$$(3.16) \quad \begin{aligned} \Theta &= \exp \left[ \frac{1}{2N_1} \mathfrak{d}^{(1)'} \mathfrak{d}^{(1)} + \frac{1}{2N_2} \mathfrak{d}^{(2)'} \mathfrak{d}^{(2)} + \frac{1}{n} \text{tr}(\mathfrak{d}^2) + \frac{4}{3n^2} \text{tr}(\mathfrak{d}^3) + \dots \right] \\ &= 1 + \frac{1}{2N_1} \mathfrak{d}^{(1)'} \mathfrak{d}^{(1)} + \frac{1}{2N_2} \mathfrak{d}^{(2)'} \mathfrak{d}^{(2)} + \frac{1}{n} \text{tr}(\mathfrak{d}^2) \\ &\quad + \frac{1}{8N_1^2} [\mathfrak{d}^{(1)'} \mathfrak{d}^{(1)}]^2 + \frac{1}{8N_2^2} [\mathfrak{d}^{(2)'} \mathfrak{d}^{(2)}]^2 + \frac{1}{4N_1 N_2} \mathfrak{d}^{(1)'} \mathfrak{d}^{(1)} \mathfrak{d}^{(2)'} \mathfrak{d}^{(2)} \\ &\quad + \frac{1}{2N_1 n} \mathfrak{d}^{(1)'} \mathfrak{d}^{(1)} \text{tr}(\mathfrak{d}^2) + \frac{1}{2N_2 n} \mathfrak{d}^{(2)'} \mathfrak{d}^{(2)} \text{tr}(\mathfrak{d}^2) \\ &\quad + \frac{1}{n^2} \left\{ \frac{1}{2} [\text{tr}(\mathfrak{d}^2)]^2 + \frac{4}{3} \text{tr}(\mathfrak{d}^3) \right\} + O_3 \end{aligned}$$

or

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

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

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

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

The Fourier transform

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

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

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

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

$$(3.21) \quad A = A(\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \Sigma) = -D^{-1} \theta \mu + \frac{1}{2} D^{-2} \theta^2 \sigma^2 + \frac{1}{2} D \theta$$

and rewrite (3.11) as

$$(3.22) \quad \psi(t) = \Theta e^A|_0.$$

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

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

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

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

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

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

Let  $\mathbf{E}_{rs}$  be the matrix with all elements 0 except the  $(r, s)$  element which is 1, then  $(\Sigma)_{rs} = (\mathbf{E}_{rs} + \mathbf{E}_{sr})/2$ . 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  $\partial_{rs}^2$ , then

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

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

$$(4.5) \quad \begin{aligned} (\sigma^{ij})_{rs,rs} \mid 0 &= \frac{1}{2} (\delta_{ijr} + \delta_{ijs} + 2\delta_{ijrs}), \\ (\sigma_2^{ij})_{rs,rs} \mid 0 &= \frac{3}{2} (\delta_{ijr} + \delta_{ijs} + 2\delta_{ijrs}), \end{aligned}$$

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

LEMMA 4.

$$(4.6) \quad (\sigma^{11})_{rs,tu} \mid 0 = \frac{1}{3} (\sigma_2^{11})_{rs,tu} \mid 0 = \frac{1}{2} (\delta_{1rt} \delta_{su} + \delta_{1ru} \delta_{st} + \delta_{1st} \delta_{ru} + \delta_{1su} \delta_{rt}),$$

$$(4.7) \quad \begin{aligned} (\sigma^{11})_{rs,rs,tu} \mid 0 &= \frac{1}{4} (\sigma_2^{11})_{rs,rs,tu} \mid 0 \\ &= -\frac{1}{2} [\delta_{1r} \delta_{st} + \delta_{1s} \delta_{rt} + 2(\delta_{1rtu} + \delta_{1stu}) + 6\delta_{1rst}], \end{aligned}$$

$$(4.8) \quad \begin{aligned} (\sigma^{11})_{rs,st,tr} \mid 0 &= \frac{1}{4} (\sigma_2^{11})_{rs,st,tr} \mid 0 \\ &= -\frac{1}{4} [\delta_{1r} + \delta_{1s} + \delta_{1t} + (\delta_{1r} \delta_{st} + \delta_{1s} \delta_{tr} + \delta_{1t} \delta_{rs}) \\ &\quad + 2(\delta_{1rs} + \delta_{1st} + \delta_{1tr}) + 12\delta_{1rst}], \end{aligned}$$

$$(4.9) \quad \begin{aligned} (\sigma^{11})_{rs,rs,tu,tu} \mid 0 &= \frac{1}{5} (\sigma_2^{11})_{rs,rs,tu,tu} \mid 0 \\ &= \frac{1}{4} [\delta_{1r} (\delta_{st} + \delta_{su}) + \delta_{1s} (\delta_{rt} + \delta_{ru}) + \delta_{1t} (\delta_{ru} + \delta_{su}) \\ &\quad + \delta_{1u} (\delta_{rt} + \delta_{st}) + 2(\delta_{1rt} + \delta_{1ru} + \delta_{1st} + \delta_{1su}) \\ &\quad + 2(\delta_{1r} \delta_{st} + \delta_{1s} \delta_{rt} + \delta_{1t} \delta_{ru} + \delta_{1u} \delta_{rs}) \\ &\quad + 2(\delta_{1rt} \delta_{su} + \delta_{1ru} \delta_{st} + \delta_{1st} \delta_{ru} + \delta_{1su} \delta_{rt}) \\ &\quad + 6(\delta_{1rst} + \delta_{1rsu} + \delta_{1rtu} + \delta_{1stu}) + 40\delta_{1rstu}]. \end{aligned}$$

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

$$e^A|_0 = e^{\theta^2/2},$$

since (3.14) gives

$$(5.2) \quad \begin{aligned} \mu &= \hat{\mu}(\mathbf{0}, \mathbf{u}_0, \mathbf{I}) = \frac{1}{2} \mathbf{u}_0' \mathbf{u}_0 = \frac{1}{2} D^2, \\ \sigma^2 &= \hat{\sigma}^2(\mathbf{0}, \mathbf{u}_0, \mathbf{I}) = \mathbf{u}_0' \mathbf{u}_0 = D^2, \end{aligned}$$

and hence  $A|_0 = \theta^2/2$  from (3.21).

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

$$(5.3) \quad (2N_1)^{-1} \sum_{r=1}^p [\partial_r^{(1)}]^2 e^A|_0.$$

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

$$(5.4) \quad \begin{aligned} \mu &= \hat{\mu}(\mathbf{u}^{(1)}, \mathbf{u}_0, \mathbf{I}) = -\frac{1}{2} (\mathbf{u}^{(1)} + \mathbf{u}_0)' (\mathbf{u}^{(1)} - \mathbf{u}_0), \\ \sigma^2 &= \hat{\sigma}^2(\mathbf{u}^{(1)}, \mathbf{u}_0, \mathbf{I}) = (\mathbf{u}^{(1)} - \mathbf{u}_0)' (\mathbf{u}^{(1)} - \mathbf{u}_0). \end{aligned}$$

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

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

Therefore (5.3) is written as

$$(5.6) \quad (2N_1)^{-1} \left[ \sum_{r=1}^p A_r^2|_0 + \sum_{r=1}^p A_{r,r}|_0 \right] e^{\theta^2/2} = L_1(\theta; D) e^{\theta^2/2} \quad (\text{say}).$$

Differentiation of (3.21) yields

$$(5.7) \quad \begin{aligned} A_r &= -D^{-1} \theta(\mu)_r + \frac{1}{2} D^{-2} \theta^2(\sigma^2)_r = D^{-1} \theta \mu_r^{(1)} + D^{-2} \theta^2(\mu_r^{(1)} - \mu_{0r}), \\ A_{r,r} &= D^{-1} \theta + D^{-2} \theta^2, \end{aligned}$$

and hence

$$(5.8) \quad A_r|_0 = -D^{-2} \theta^2 \mu_{0r} = -D^{-1} \theta^2 \delta_{1r}, \quad A_{r,r}|_0 = D^{-2} (\theta^2 + D\theta).$$

Thus we have

$$(5.9) \quad L_1(\theta; D) = (2N_1 D^2)^{-1} [\theta^4 + p(\theta^2 + D\theta)].$$

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

$$(5.10) \quad L_2(\theta; D) = (2N_2 D^2)^{-1} [(\theta^2 - D\theta)^2 + p(\theta^2 - D\theta)].$$

For the later reference we state the result

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

which corresponds to (5.8).

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

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

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

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

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

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

From (3.21) and (5.13) follows

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

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

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

Substituting this into (5.14), we obtain

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

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

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

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

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

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

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

If we put

$$(5.20) \quad \begin{aligned} Q_{11}(\theta; D) &= (8N_1^2)^{-1} \sum_{r=1}^p (4A_r^2 A_{r,r} + 2A_{r,r}^2) |_0 \\ &= (4N_1^2 D^4)^{-1} [2\theta^4 (\theta^2 + D\theta) + p(\theta^2 + D\theta)^2] \end{aligned}$$

which follows from (5.8), then we have

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

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

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

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

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

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

The cross term of order  $(N_1 N_2)^{-1}$  can be written as

$$(5.24) \quad (4N_1 N_2)^{-1} \sum_{r,s=1}^p [\partial_r^{(1)} \partial_s^{(2)}]^2 e^A |_0 = [L_1(\theta; D) L_2(\theta; D) + Q_{12}(\theta; D)] e^{\theta^2/2}.$$

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

$$(5.25) \quad \begin{aligned} Q_{12}(\theta; D) &= (4N_1 N_2)^{-1} \sum_{r,s=1}^p (4A_r A_s A_{r,s} + 2A_{r,s}^2) |_0 \\ &= (2N_1 N_2 D^4)^{-1} [2\theta^4 (\theta^2 - D\theta) + p\theta^4] \end{aligned}$$

by substituting  $A_r |_0$  and  $A_s |_0$  given in (5.8) and (5.11), respectively, and also  $A_{r,s} |_0 = -D^{-2} \theta^2 \delta_{rs}$  which is easily verified.

Similarly the term of order  $(N_1 n)^{-1}$  is

$$(5.26) \quad (2N_1 n)^{-1} \sum_{r,s,t=1}^p [\partial_r^{(1)} \partial_{st}]^2 e^A |_0 = [L_1(\theta; D) L_3(\theta; D) + Q_{13}(\theta; D)] e^{\theta^2/2}.$$

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

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

so that

$$(5.28) \quad \begin{aligned} Q_{13}(\theta; D) &= (2N_1 n)^{-1} \sum_{r,s,t=1}^p [4A_r A_{st} A_{r,st} + 2A_{r,st}^2 \\ &\quad + 2A_r A_{r,st,st} + 2A_{st} A_{r,r,st} + A_{r,r,st,st}] |_0. \end{aligned}$$

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

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

as the term of order  $(N_2n)^{-1}$ , 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

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

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

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

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

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

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

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

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.