SOME RESULTS ON THE DISTRIBUTION OF TWO RANDOM MATRICES USED IN CLASSIFICATION PROCEDURES

By D. G. Kabe

Wayne State University and Karnatak University

1. Introduction and summary. Wald [5] and Anderson [1] discuss a classification problem as follows. We have $N_1 + N_2 + 1$ independent p dimensional random vectors. We know that the first N_1 vectors are observations from a population Π_1, the following N_2 are observations from a population Π_2, and the last vector is either from Π_1 or Π_2. Let us assume that the probability distribution in both Π_1 and Π_2 is multivariate normal with the same covariance matrix Σ; the vector of expected values being μ_1 in Π_1 and μ_2 in Π_2. The values of μ_1, μ_2, and Σ are unknown. Let X denote the $p \cdot (N_1 + N_2 + 1)$ matrix of observations. On the basis of X we wish to classify the last observation, $X_{N_1+N_2+1}$ as coming from Π_1 or Π_2. For this purpose both Wald and Anderson propose classification statistics. Wald proposes the statistic

$$U = X'_{N_1+N_2+1} S^{-1}(\bar{X}_1 - \bar{X}_2), \tag{1.1}$$

where

$$\bar{X}_1 = (1/N_1) \sum_{t=1}^{N_1} X_t, \qquad \bar{X}_2 = (1/N_2) \sum_{t=N_1+1}^{N_1+N_2} X_t, \tag{1.2}$$

and

$$S = (1/N_1 + N_2 - 2) \cdot \left[\sum_{t=1}^{N_1} (X_t - \bar{X}_1)(X_t - \bar{X}_1)' + \sum_{t=N_1+1}^{N_1+N_2} (X_t - \bar{X}_2)(X_t - \bar{X}_2)' \right]. \tag{1.3}$$

Anderson considers the statistic

$$W = X'_{N_1+N_2+1} S^{-1}(\bar{X}_1 - \bar{X}_2) - \tfrac{1}{2}(\bar{X}_1 + \bar{X}_2)' S^{-1}(\bar{X}_1 - \bar{X}_2). \tag{1.4}$$

It is known [4] that the sampling distributions of U and W are contained as special cases in the sampling distribution of the statistic

$$V = Y_1' A^{-1} Y_2, \tag{1.5}$$

where Y_1 and Y_2 are p dimensional random normal vectors with mean vectors ξ_1 and ξ_2, say, respectively, and A is a $p \times p$ symmetric matrix having the Wishart distribution with n degrees of freedom; the three sets are independently distributed with the same covariance matrix Σ. Let M denote the matrix

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{pmatrix} = Y' B^{-1} Y, \tag{1.6}$$

Received December 5, 1961; revised July 17, 1962.

where

$$(1.7)\qquad Y = (Y_1, Y_2), \quad \text{and} \quad B = A + YY',$$

then the statistic V is given by the relation

$$(1.8)\qquad V = m_{12}((1 - m_{11})(1 - m_{22}) - m_{12}^2)^{-1}.$$

Extending earlier work of Wald [5], and Anderson [1], Sitgreaves [4] obtains the distributions of the matrices M and M^*, where

$$(1.9)\qquad M^* = Y'A^{-1}Y = M(I - M)^{-1}.$$

Both Anderson [1] and Sitgreaves [4] assume that the vectors ξ_1 and ξ_2 are proportional. In this paper we obtain these distributions without assuming the proportionality of ξ_1 and ξ_2. For this purpose we require the following results.

2. Some useful results. Anderson and Girshick [3] essentially prove that

$$(2.1)\qquad \begin{aligned}\int_{XX'=D} & f(XX') \exp\{\operatorname{tr} \Sigma^{-1}\mu X'\}\, dX \\ &= \pi^{p(2n-2t-p+1)/4}\left[\prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n - t + 1 - i])\right]^{-1} f(D) \\ &\quad \cdot \int \exp\{\lambda_1 v_{11} + \lambda_2 v_{22} + \cdots + \lambda_t v_{tt}\} |D - VV'|^{(n-p-t-1)/2}\, dV.\end{aligned}$$

Here X is a $p \times n$ matrix, μ a given $p \times n$ matrix of rank $t(\leqq p)$, as usual dX denotes the product of the differentials of the elements of the matrix X, D is a $p \times p$ symmetric matrix of rank p, and $\lambda_1^2, \cdots, \lambda_t^2$, are the nonzero roots of the equation

$$(2.2)\qquad |\Sigma^{-1}\mu\mu'\Sigma^{-1} - \lambda I| = 0,$$

the elements of the matrix V being assumed to be arranged as

$$(2.3)\qquad V = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1t} \\ v_{21} & v_{22} & \cdots & v_{2t} \\ \cdot & \cdot & \cdots & \cdot \\ v_{p1} & v_{p2} & \cdots & v_{pt} \end{pmatrix},$$

and the range of integration with respect to V is determined by the condition that the matrix $(B - VV')$ is positive semidefinite.

It is obvious ([2], p. 417) that the noncentral planar p dimensional Wishart density is a particular case of the equation (2.1); however, when $n = p$, this density may also be written as ([2], p. 419)

$$(2.4)\qquad \begin{aligned} W(D, &\Sigma, \omega_1, \omega_2; p, p, 2) \\ &= C_1 \exp\{-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1} D\} |D|^{-\frac{1}{2}} \int \exp\{\omega_1 z_{11} + \omega_2 z_{22}\} |I - ZZ'|^{(p-5)/2}\, dZ_1\, dZ_2,\end{aligned}$$

where Z_1 and Z_2 are column vectors of the matrix

$$Z = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix}, \tag{2.5}$$

and ω_1^2 and ω_2^2 are the nonzero roots of the equation

$$|\Sigma^{-1}\mu\mu'\Sigma^{-1} - \lambda D^{-1}| = 0, \tag{2.6}$$

as the rank of μ is now assumed to be 2. The range of integration with respect to Z is determined by the condition that the matrix $(I - ZZ')$ is positive semi-definite. The constant C_1 is given by the expression

$$\begin{aligned} C_1^{-1} &= \exp\{\tfrac{1}{2}(k_1^2 + k_2^2)\}2^{p^2/2}\Pi^{p+p(p-1)/4}|\Sigma|^{p/2} \\ &\quad \cdot \prod_{i=1}^{2}\Gamma(\tfrac{1}{2}[p-1-i])\prod_{i=1}^{p-2}\Gamma(\tfrac{1}{2}[p-1-i]), \end{aligned} \tag{2.7}$$

where k_1^2 and k_2^2 are the nonzero roots of the equation

$$|\Sigma^{-1}\mu\mu'\Sigma^{-1} - \lambda\Sigma^{-1}| = 0. \tag{2.8}$$

Now we consider the integral

$$\begin{aligned} I = \int \exp\{-\tfrac{1}{2}\operatorname{tr}\Sigma^{-1}D + \omega_1 z_{11} + \omega_2 z_{22}\}|D|^{(n-p+1)/2} \\ \cdot|I - ZZ'|^{(p-5)/2}\,dZ_1\,dZ_2\,dD' \end{aligned} \tag{2.9}$$

which is easily evaluated by using Theorem 5 of Anderson's ([2], p. 424); and we find that

$$\begin{aligned} I &= |\Sigma|^{(n+2)/2}2^{p(n+2)/2}\pi^{p+p(p-1)/4} \\ &\quad \cdot\prod_{i=1}^{2}\Gamma(\tfrac{1}{2}[p-1-i])\prod_{i=3}^{p}\Gamma(\tfrac{1}{2}[n+3-i])G(k_1^2, k_2^2), \end{aligned} \tag{2.10}$$

where

$$\begin{aligned} G(k_1^2, k_2^2) &= \sum_{\alpha,\beta_1,\beta_2=0}^{\infty}\frac{(k_1^2)^{\alpha+\beta_1}(k_2^2)^{\alpha+\beta_2}\Gamma(\frac{1}{2}[n+2]+\alpha+\beta_1)}{2^{2\alpha+\beta_1+\beta_2}\alpha!\beta_1!\beta_2!\Gamma(\frac{1}{2}[p-1]+\alpha)} \\ &\quad \cdot\frac{\Gamma(\frac{1}{2}[n+2]+\alpha+\beta_2)\Gamma(\frac{1}{2}[n+1]+\alpha)}{\Gamma(\frac{1}{2}p+2\alpha+\beta_1+\beta_2)\Gamma(\frac{1}{2}[n+2]+\alpha)}. \end{aligned} \tag{2.11}$$

Here we note that the process of integration carries the quantities ω_1, ω_2, to the quantities k_1, k_2; i.e., the process of integration with respect to D replaces the matrix D in (2.6) by the matrix Σ. We now proceed to obtain the distributions of the matrices M and M^*.

3. The distribution of the matrix M. It is easily seen [4] that the joint density of the matrices B and Y is given by the expression

$$f(B, Y) = C_2|B - YY'|^{(n-p-1)/2}\exp\{-\tfrac{1}{2}\operatorname{tr}\Sigma^{-1}B + \operatorname{tr}\Sigma^{-1}\xi Y'\}, \tag{3.1}$$

where

$$(3.2) \qquad \xi = (\xi_1, \xi_2),$$

and

$$(3.3) \qquad C_2^{-1} = |\Sigma|^{(n-2)/2} \exp \{\tfrac{1}{2} \operatorname{tr} \Sigma^{-1}\xi\xi'\} 2^{p(n+2)/2} \pi^{p+p(p-1)/4} \prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n+1-i]).$$

Now consider the integral

$$(3.4) \qquad I = \int_{Y'B^{-1}Y=M} |I - Y'B^{-1}Y|^{(n-p-1)/2} \exp \{\operatorname{tr} \Sigma^{-1}\xi Y'\}\, dY,$$

which is easily evaluated by using (2.1); and we have that

$$(3.5) \qquad \begin{aligned} I = {} & C_3 |B|\, |I - M|^{(n-p-1)/2} \\ & \cdot \int |M - ZZ'|^{(p-5)/2} \exp \{\gamma_1 z_{11} + \gamma_2 z_{22}\}\, dZ_1\, dZ_2, \end{aligned}$$

where

$$(3.6) \qquad C_3^{-1} = \pi^{(5-2p)/2} \Gamma(\tfrac{1}{2}[p-2]) \Gamma(\tfrac{1}{2}[p-3]),$$

and γ_1^2 and γ_2^2 are the nonzero roots of the equation

$$(3.7) \qquad |\Sigma^{-1}\xi\xi'\Sigma^{-1} - \lambda B^{-1}| = 0,$$

i.e., of the equation

$$(3.8) \qquad |\xi'\Sigma^{-1}B\Sigma^{-1}\xi - \lambda I| = 0.$$

It may be noted that in (3.5) the range of integration is determined by the condition that the matrix $(M - ZZ')$ is positive semidefinite, while in (2.4) the range is given by the condition that the matrix $(I - ZZ')$ is positive semidefinite. However, by a suitable transformation (3.5) may be written as

$$(3.9) \qquad \begin{aligned} I = {} & C_3 |B|\, |I - M|^{(n-p-1)/2} |M|^{(p-3)/2} \\ & \cdot \int \exp \{\delta_1 t_{11} + \delta_2 t_{22}\} |I - TT'|^{(p-5)/2}\, dT_1\, dT_2, \end{aligned}$$

where δ_1^2 and δ_2^2 are the roots of the equation

$$(3.10) \qquad |\xi'\Sigma^{-1}B\Sigma^{-1}\xi - \lambda M^{-1}| = 0,$$

and T_1 and T_2 are column vectors of the matrix

$$(3.11) \qquad T = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}.$$

Using (3.1), (3.4), and (3.9), the density of the matrix M is found to be

$$(3.12) \qquad \begin{aligned} f(M) = {} & C_2 C_3 |I - M|^{(n-p-1)/2} |M|^{(p-3)/2} \\ & \cdot \int \exp \{-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1}B + \delta_1 t_{11} + \delta_2 t_{22}\} \\ & \cdot |B|^{(n-p-1)/2} |I - TT'|^{(p-5)/2}\, dT_1\, dT_2\, dB, \end{aligned}$$

and by using (2.9) we have for the density of the matrix M the expression

$$f(M) = C_4 |I - M|^{(n-p-1)/2} |M|^{(p-3)/2} G(\phi_1^2, \phi_2^2), \tag{3.13}$$

where ϕ_1^2 and ϕ_2^2 are the roots of the equation

$$|\xi' \Sigma^{-1} \xi - \lambda M^{-1}| = 0, \tag{3.14}$$

and

$$C_4^{-1} = \exp\{\tfrac{1}{2} \operatorname{tr} \xi\xi'\} \Gamma(\tfrac{1}{2}) \Gamma(\tfrac{1}{2}[n + 1 - p]) \Gamma(\tfrac{1}{2}[n + 2 - p]). \tag{3.15}$$

We note that when ξ_1 and ξ_2 are proportional, say, $\theta_1 \xi^*$ and $\theta_2 \xi^*$, respectively, then (3.14) has only one root, viz.,

$$\phi_1^2 = (\xi^{*\prime} \Sigma^{-1} \xi^*)(\theta_1^2 m_{11} + 2\theta_1\theta_2 m_{12} + \theta_2^2 m_{22}). \tag{3.16}$$

Thus Sitgreaves' result ([4], p. 269, Equation 21) follows from (3.13) by setting

$$\alpha = 0, \quad \beta_2 = 0, \quad \phi_2 = 0, \quad \text{and} \quad \phi_1 = \phi_1 \text{ of (3.16)}. \tag{3.17}$$

4. The distribution of the matrix M^*. In (3.13) we set

$$M = M^*(I + M^*)^{-1}. \tag{4.1}$$

The Jacobian J of the transformation from M to M^* is easily found to be

$$J = |I + M^*|^{-3}. \tag{4.2}$$

Hence the density of the matrix M^* is given by

$$f(M^*) = C_4 |M^*|^{(p-3)/2} |I + M^*|^{-(n+2)/2} G(\chi_1^2, \chi_2^2), \tag{4.3}$$

where χ_1^2 and χ_2^2 are the roots of

$$|\xi' \Sigma^{-1} \xi - \lambda(I + M^*) M^{*-1}| = 0. \tag{4.4}$$

In a future communication, it is planned to generalize the results of this paper to the case of 3 populations, by using a result of Weibull's [6].

Acknowledgment. The author wishes to thank Dr. A. G. Laurent for some helpful suggestions.

REFERENCES

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

[2] Anderson, T. W. (1946). The noncentral Wishart distribution and certain problems of multivariate statistics. *Ann. Math. Statist.* **17** 409–431.

[3] Anderson, T. W. and Girshick, M. A. (1944). Some extensions of the Wishart distribution. *Ann. Math. Statist.* **15** 345–357.

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

[5] Wald, A. (1944). On a statistical problem arising in the classification of an individual into one of two groups. *Ann. Math. Statist.* **15** 145–162.

[6] Weibull, Martin (1953). The distributions of t- and F-statistics and of correlation and regression coefficients in stratified samples from normal populations with different means. *Skand. Aktuarietidskr.* **36** 9–106.