The Annals of Mathematical Statistics
1968, Vol. 39, No. 3, 855–858

ON THE DISTRIBUTIONS OF DIRECTION AND COLLINEARITY FACTORS IN DISCRIMINANT ANALYSIS

BY D. G. KABE

Dalhousie University

1. Introduction and summary. Let A and B be two symmetric positive definite matrices of order p, having the independent Wishart densities

$$g(A) = C_1 \exp\{-\tfrac{1}{2} \operatorname{tr} A\}|A|^{(n-q-p-1)/2}, \tag{1.1}$$

where

$$C_1^{-1} = 2^{(n-q)p/2} \pi^{p(p-1)/4} \textstyle\prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n-q+1-i]), \tag{1.2}$$

and

$$g(B) = C_2 \exp\{-\tfrac{1}{2} \operatorname{tr} B\}|B|^{(q-p-1)/2} \textstyle\sum_{\alpha=0}^{\infty} (\lambda^2 b_{11})^{\alpha}/ [2^{2\alpha}\alpha!\Gamma(\tfrac{1}{2}n + \alpha)], \tag{1.3}$$

where

$$C_2^{-1} = \exp\{\tfrac{1}{2}\lambda^2\}2^{qp/2}\pi^{p(p-1)/4} \textstyle\prod_{i=2}^{p} \Gamma(\tfrac{1}{2}[q + 1 - i]). \tag{1.4}$$

Anderson [1] calls the density (1.3) as noncentral linear Wishart density. If we define the matrix V by the relations

$$A = CVC', \qquad A + B = CC', \tag{1.5}$$

where C is a lower triangular matrix of order p. Then Kshirsagar [4] finds the noncentral multivariate linear beta density of V to be

$$g(V) = C_3|V|^{(n-q-p-1)/2}|I - V|^{(q-p-1)/2}\Phi(v_{11}), \tag{1.6}$$

where

$$\Phi(v_{11}) = {}_1F_1[\tfrac{1}{2}n, \tfrac{1}{2}q; \tfrac{1}{2}\lambda^2(1 - v_{11})], \tag{1.7}$$

and

$$C_3 = \exp\{-\tfrac{1}{2}\lambda^2\} \textstyle\prod_{i=2}^{p} \Gamma(\tfrac{1}{2}[n + 1 - i])\pi^{-p(p-1)/4} \cdot[\prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n - q + 1 - i]) \prod_{i=2}^{p} \Gamma(\tfrac{1}{2}[q + 1 - i])]^{-1}. \tag{1.8}$$

Kshirsagar [4] has used this distribution of V to derive the distribution of the test criterion for testing the adequacy of a single hypothetical discriminant function as defined by Williams [6]. Continuing his earlier work, Kshirsagar [5] now uses the distribution of V to obtain the distributions of the direction, and the collinearity factor of this single discriminant function. In case $\xi = \alpha' x$ denotes the discriminant function, then Bartlett [3] gives a factorization of $\Lambda = |V|$ as

$$\Lambda = \Lambda_1\Lambda_2\Lambda_3, \tag{1.9}$$

Received April 17, 1967.

where

$$\Lambda_1 = 1 - (\alpha' B\alpha)/(\alpha'(A + B)\alpha),$$

$$\text{(1.10)} \quad \Lambda_2 = [1 - \alpha' B(A + B)^{-1} B\alpha/\alpha' B\alpha]/\Lambda_1, \quad \text{direction factor},$$

$$\Lambda_3 = \Lambda/\Lambda_1\Lambda_2, \quad \text{the partial collinearity factor.}$$

Assuming $\alpha = (1, 0, 0, \cdots, 0)$ and factorizing the density of V in terms of rectangular coordinates T, where $V = TT'$ and T lower triangular, Kshirsagar [5] expresses Λ_1, Λ_2, and Λ_3 as function of the elements of T, and thus obtains the densities of Λ_1, Λ_2, and Λ_3. Bartlett [3] gives an alternative factorization of Λ as $\Lambda = \Lambda_1\Lambda_4\Lambda_5$, where Λ_4 is called the collinearity factor and Λ_5 the partial direction factor, where Λ_4 and Λ_5 are certain functions of the elements of the matrices A and B. Kshirsagar expresses Λ_4 and Λ_5 as functions of the elements of T and obtains their distributions.

Kshirsagar's [5] derivations of these distributions, although elegant, are lengthy and involved, as he uses several lower triangular matrix transformations of the rectangular coordinates in his derivations. It might perhaps be of pedagogical interest to express Λ_1, Λ_2, Λ_3, Λ_4, and Λ_5 as functions of the elements of V itself and thus derive their distributions. The present derivations are shorter and neater as they avoid most of the transformations used by Kshirsagar [5].

Since the distributions of Λ's are well known beta distributions, they have been derived without the normalizing constants. We assume that all the integrals occurring in this paper are evaluated over appropriate ranges of the variable of integration.

2. Distribution theory. Using Anderson's notation ([2], p. 106) we partition V and $(I - V)$ as

$$\text{(2.1)} \qquad V = \begin{pmatrix} v_{11} & v'_{(1)} \\ v_{(1)} & V_{22} \end{pmatrix}, \qquad I - V = \begin{pmatrix} 1 - v_{11} & -v'_{(1)} \\ -v_{(1)} & I - V_{22} \end{pmatrix},$$

and write the density of V as

$$\text{(2.2)} \quad g(v_{11}, v_{(1)}, V_{22}) \propto v_{11}^{(n-q-p-1)/2}(1 - v_{11})^{(q-p-1)/2}\Phi(v_{11})$$
$$\cdot |V_{22} - v_{(1)}v'_{(1)}/v_{11}|^{(n-q-p-1)/2}|I - V_{22} - v_{(1)}v'_{(1)}/(1 - v_{11})|^{(q-p-1)/2}.$$

Further setting

$$\text{(2.3)} \quad Z = V_{22} - v_{(1)}v'_{(1)}/v_{11},$$
$$R = (I - v_{(1)}v'_{(1)}/[v_{11}(1 - v_{11})])^{-\frac{1}{2}}Z(I - v_{(1)}v'_{(1)}/[v_{11}(1 - v_{11})])^{-\frac{1}{2}},$$

we find the joint density of v_{11}, $v_{(1)}$, and R to be

$$g(v_{11}, v_{(1)}, R)$$
$$\text{(2.4)} \quad \propto v_{11}^{(n-q-p-1)/2}(1 - v_{11})^{(q-p-1)/2}\Phi(v_{11})(1 - v'_{(1)}v_{(1)}/v_{11}(1 - v_{11}))^{(n-p-1)/2}$$
$$\cdot |R|^{(n-q-p-1)/2}|I - R|^{(q-p-1)/2}.$$

It may be noted that

$$\Lambda_1 = V_{11}, \quad \Lambda_2 = (1 - v'_{(1)}v_{(1)}/v_{11}(1 - v_{11})), \quad \Lambda_3 = |R|. \tag{2.5}$$

Now making the transformation $v_{(1)} = (v_{11}(1 - v_{11}))^{\frac{1}{2}}\delta$, and integrating over the range $\delta'\delta = (1 - \Lambda_2)$, we find that the densities of v_{11}, Λ_2, and Λ_3 are mutually independent. v_{11} has a noncentral beta density, Λ_2 has a central beta density, and $\Lambda_3 = |R|$ has a density that is identical with that of a product of $p - 1$ independent beta variates, the ith variate has a beta distribution with parameters $\frac{1}{2}(n - q - i)$ and $\frac{1}{2}(q - 1)$, $i = 1, 2, \cdots, p - 1$. This is the result derived by Kshirsagar ([5], p. 222).

Now we proceed to obtain the distributions of Λ_4 and Λ_5. It may be easily verified that

$$\Lambda_5 = (1 + v'_{(1)}Z^{-1}v_{(1)}/(v_{11}(1 - v_{11})))^{-1}. \tag{2.6}$$

Since Z is positive definite symmetric of order $p - 1$, there exists a nonsingular $(p - 1) \times (p - 1)$ matrix P such that $PP' = Z$. The joint density of v_{11}, P, and $v_{(1)}$ may be obtained by using the result (2.5), and we find that

$$\begin{aligned} g(v_{11}, p, v_{(1)}) \propto\ & v_{11}^{(n-q-p-1)/2}(1 - v_{11})^{(q-p-1)/2}\Phi(v_{11})|PP'|^{(n-q-p)/2} \\ & \cdot|I - PP' - v_{(1)}v'_{(1)}/v_{11}(1 - v_{11})|^{(q-p-1)/2}. \end{aligned} \tag{2.7}$$

Further transforming $v_{(1)}$ to η by the relation

$$v_{(1)} = P\eta, \tag{2.8}$$

we find the joint density of v_{11}, P, and η to be

$$\begin{aligned} & g(v_{11}, P, \eta) \\ & \propto v_{11}^{(n-q-p-1)/2}(1 - v_{11})^{(q-p-1)/2}\Phi(v_{11})(1 + \eta'\eta/(v_{11}(1 - v_{11})))^{-(n-q-p+1)/2} \\ & \cdot|P(I + \eta'\eta/v_{11}(1 - v_{11}))P'|^{(n-q-p+1)/2} \\ & \cdot|I - P(I + \eta'\eta/v_{11}(1 - v_{11}))P'|^{(q-p-1)/2}. \end{aligned} \tag{2.9}$$

Now we set

$$P(\mathrm{I} + \eta'\eta/v_{11}(1 - v_{11}))P' = W, \tag{2.10}$$

and find that

$$\begin{aligned} & g(v_{11}, W, \eta) \\ & \propto v_{11}^{(n-q-p-1)/2}(1 - v_{11})^{(q-p-1)/2}\Phi(v_{11})(1 + \eta'\eta/(v_{11}(1 - v_{11})))^{-(n-q)/2} \\ & \cdot|W|^{(n-q-p)/2}|I - W|^{(q-p-1)/2}. \end{aligned} \tag{2.11}$$

It is easily observed that

$$\Lambda_5 = (1 + \eta'\eta/v_{11}(1 - v_{11}))^{-1}. \tag{2.12}$$

Using the transformation $\gamma'\gamma = (1 - \Lambda_5)/\Lambda_5$, where $\eta = \gamma(v_{11}(1 - v_{11}))^{\frac{1}{2}}$, we find the density of v_{11}, Λ_5, and W to be

$$g(v_{11}, W, \Lambda_5) \tag{2.13}$$
$$\propto v_{11}^{(n-q-2)/2}(1 - v_{11})^{(q-2)/2}\Phi(v_{11})\Lambda_5^{(n-q-p-1)/2}(1 - \Lambda_5)^{(p-3)/2} \cdot |W|^{(n-q-p)/2}|I - W|^{(q-p-1)/2}.$$

Incidentally we note that $\Lambda_4 = |W|$, and hence we observe that the densities of $v_{11} = \Lambda_1$, Λ_4, and Λ_5 are mutually independent. Λ_5 has a beta density, and the density of Λ_4 is identical with that of a product of $p - 1$ independent beta variables.

3. Concluding remarks and acknowledgment. As pointed out by Kshirsagar ([5], pp. 217–218) a single discriminant function is adequate when all the population canonical correlations except the first is zero and the canonical variate corresponding to this nonzero canonical correlation coefficient is the required discriminant function. Obviously when two or more population canonical correlations are nonzero a single discriminant function is inadequate. No reference in statistical literature is available as what is to be done in this case. Wilks ([7], pp. 576–587) gives $\binom{q}{2}$ discriminant functions, under very restricted assumption about the means of the groups, to discriminate amongst q groups. However, Wilks considers a different aspect than the one considered by Kshirsagar [5]. Further the distributions derived in this paper assume that $\alpha = (1, 0, 0, \cdots, 0)$, it might be interesting to derive the distributions for a general α. There are several unexplored aspects of discriminant analysis and we hope research workers in Multivariate Analysis will soon deeply explore this branch of Multivariate Analysis.

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

REFERENCES

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

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

[3] Bartlett, M. S. (1951). The goodness of fit of a single hypothetical discriminant function in the case of several groups. *Ann. Eugen.* **16** 199–214.

[4] Kshirsagar, A. M. (1961). The noncentral multivariate beta distribution. *Ann. Math. Statist.* **32** 104–111.

[5] Kshirsagar, A. M. (1964). Distributions of the direction and collinearity factors in discriminant analysis. *Proc. Camb. Phil. Soc.* **60** 217–225.

[6] Williams, E. J. (1952). Some exact tests in multivariate analysis. *Biometrika* **39** 17–31.

[7] Wilks, S. S. (1962). *Mathematical Statistics.* Wiley, New York.