GENERALIZATION OF SVERDRUP'S LEMMA AND ITS APPLICATIONS TO MULTIVARIATE DISTRIBUTION THEORY

By D. G. Kabe

Wayne State University and Karnatak University

1. Introduction and summary. The multivariate sampling distribution theory underlying a multivariate nornal law has a significant role in Multivariate Statistical Analysis. Several methods are available for the derivation of the usual sampling distributions, see, e.g., [2], [7], [8], [9], and [10]. This is yet another attempt in the same direction. However, the method which we present in this paper is elegant and straightforward, as the desired distributions are obtained in a rather unified way, either directly from the probability law of the sample or from the Wishart distribution. Our method is based on a generalization of Sverdrup's lemma [12] which we give below. The generalized Sverdrup's lemma is indeed implicit in many derivations in multivariate statistical literature without its explicit statement, see, e.g., Anderson's derivation of the integral representation of the noncentral Wishart distribution ([1], p. 417). Our purpose in this paper is to make an explicit statement of this implied lemma, and point out that the lemma stated here may be used as a powerful tool in multivariate distribution theory. Since our method is easily understood, we give only two applications. Several other applications follow on similar lines, see, Kabe [3], [4]. We have used fairly standard notation in this paper.

2. Generalization of Sverdrup's lemma. Sverdrup's lemma [12] may be stated as follows. Let y be a N component column vector, D a given $q \times N$ matrix of rank $q(<N)$, then

$$\int_{y'y=u, Dy=v} f(y'y, Dy)\, dy = \frac{1}{2} C(N-q)|DD'|^{-\frac{1}{2}} f(u, v)[u - v'(DD')^{-1}v]^{\frac{1}{2}(N-q-2)}. \tag{2.1}$$

Here v is a q component column vector, dy, as usual, denotes the product of the differentials of the elements of y, and $C(N)$ represents the surface area of a unit N dimensional sphere. The integral is considered as a part of the volumn integral over the appropriate range of the variables of integration, i.e., the range is $-\infty < y < \infty$, $u \leqq y'y \leqq u + du$, $v \leqq Dy \leqq v + dv$. In case the integrand f is a suitable density function, then the right hand side of (2.1) obviously represents the joint density of the variables u and v. A similar lemma follows when the rank of D is less than q.

As a generalization of (2.1) we have the following lemma.

Lemma. *Let Y be a $p \times N$ matrix all of whose components range over the entire N dimensional cartesian Euclidean space, D a given $q \times N$ matrix of rank $q(<N)$,*

Received 6 July 1961; revised 28 September 1964.

$N \geqq p + q$. Then

$$(2.2)\quad \begin{aligned} &\int_{YY'=G,DY'=V'} f(YY', DY')\, dY \\ &= 2^{-p} \prod_{i=1}^{p} C(N - p - q + i)|DD'|^{-\frac{1}{2}p} f(G, V')|G - V(DD')^{-1}V'|^{\frac{1}{2}(N-p-q-1)}. \end{aligned}$$

Here G is a $p \times p$ positive definite symmetric matrix, V is a $p \times q$ matrix, and the integral is considered as a part of the volume integral over the range $-\infty < Y < \infty$, $G \leqq YY' \leqq G + dG$, $V \leqq YD' \leqq V + dV$. In case the integrand f is a suitable density function, then the right hand side of (2.2) obviously represents the joint density of the matrices G and V. A similar lemma holds when the rank of D is less than q.

We proceed to prove (2.2). Now with $DY' = V'$, we have a $(N - q) \times N$ matrix C orthogonal to D and satisfying the condition $CC' = I$, where I is the identity matrix of order $N - q$. Setting $CY' = U$, where U is a $p \times (N - q)$ matrix, we find that

$$(2.3)\quad \begin{pmatrix} D \\ C \end{pmatrix} Y' = \begin{pmatrix} V' \\ U' \end{pmatrix}.$$

The Jacobian J of the transformation from Y to $(V \quad U)$ is easily found to be

$$(2.4)\quad J = \left| \begin{pmatrix} D \\ C \end{pmatrix} \begin{pmatrix} D \\ C \end{pmatrix}' \right|^{-\frac{1}{2}p} = |DD'|^{-\frac{1}{2}p}.$$

We also observe that

$$(2.5)\quad YY' = (VU) \left[\begin{pmatrix} D \\ C \end{pmatrix} \begin{pmatrix} D \\ C \end{pmatrix}' \right]^{-1} (VU)' = V(DD')^{-1}V' + UU'.$$

It follows that the integral (2.2) is now equal to the integral

$$(2.6)\quad |DD'|^{-\frac{1}{2}p} \int_{UU'=G-V(DD')^{-1}V'} f(V(DD')^{-1}V' + UU', V')\, dU,$$

which is easily evaluated by using the known integral ([2], p. 319, Lemma 13.3.1)

$$(2.7)\quad \int_{YY'=G} f(YY')\, dY = 2^{-p}\prod_{i=1}^{p} C(N - p + i) f(G)|G|^{\frac{1}{2}(N-p-1)},$$

and gives us the right hand side of the equation (2.2).

Now we proceed to consider the applications of the Lemma (2.2) to distribution theory. However, we also require following useful results.

3. Some useful results. In case D is a row vector d' of N components, then from (2.2) it follows that

$$(3.1)\quad \begin{aligned} &\int_{YAY'=G} f(YAY', d'Y')\, dY \\ &= 2^{-p}\prod_{i=1}^{p} C(N - p - 1 + i)|A|^{-\frac{1}{2}p} \int f[G, z'(d'A^{-1}d)^{\frac{1}{2}}]|G - zz'|^{\frac{1}{2}(N-p-2)}\, dz. \end{aligned}$$

The range of the integration on the right hand side of (3.1) is determined by the condition that $(G - zz')$ is positive semidefinite.

Using a known integral ([6], p. 268) it may be shown that

$$(3.2)\quad \int_{y'Ay=u} f(y'Ay) \exp\{d'y\}\, dy = \tfrac{1}{2}C(N-1)|A|^{-\frac{1}{2}}B(\tfrac{1}{2}, \tfrac{1}{2}(N-1))f(u)u^{\frac{1}{2}(N-2)} \cdot \sum_{r=0}^{\infty} \Gamma(\tfrac{1}{2}N)u^r(d'A^{-1}d)^r/\Gamma(\tfrac{1}{2}N + r)2^{2r}r!.$$

In equations (3.1) and (3.2) A is a positive definite $N \times N$ symmetric matrix. We use the result ([2], p. 176, Example 6) to obtain that

$$(3.3)\quad \begin{aligned} &\int \exp\{-\tfrac{1}{2}\operatorname{tr} \Sigma^{-1}G\}|G|^{\frac{1}{2}(N-p-1)}(\delta'G\delta)^r\, dG \\ &= (d/d\lambda)^r_{\lambda=0} \int \exp\{-\tfrac{1}{2}\operatorname{tr}(\Sigma^{-1} - \lambda\delta\delta')G\}|G|^{\frac{1}{2}(N-p-1)}\, dG \\ &= (2\pi)^{\frac{1}{2}pN}|\Sigma|^{\frac{1}{2}N}2^{p+r}\Gamma(\tfrac{1}{2}N + r)(\delta'\Sigma^{-1}\delta)^r/\textstyle\prod_{i=1}^{p} C(N-p+i)\Gamma(\tfrac{1}{2}N). \end{aligned}$$

Here dG denotes the product of the differentials of the distinct elements of the matrix G, whose diagonal elements range over 0 to ∞ and nondiagonal elements range over $-\infty$ to ∞; δ is a p component column vector. The interchange of differentiation and integration in the formula (3.1) is easily justified. We now proceed with the distribution theory.

We assume that all the integrals occurring in this paper are evaluated over appropriate ranges of the variables of integration.

4. The distribution of the matrix of regression coefficients. We consider a multivariate normal regression model, $Y = BX + E$, where Y is a $p \times N$ matrix of observations, B is the $p \times q$ matrix of the population regression coefficients, the $q \times N$ matrix X of fixed variates is of rank $q(< N)$, and the usual pN component error vector that is obtained from the elements of E has a pN variate normal distribution with mean vector zero and covariance matrix $\Sigma \otimes I$. Here Σ is $p \times p$ positive definite symmetric matrix. The maximum likelihood estimate $\hat{B}$ of B is $\hat{B}' = (XX')^{-1}XY'$. Obviously the joint density of the matrices $YY' = G$, and $\hat{B}$ is

$$(4.1)\quad g(G, \hat{B}) = \int_{YY'=G,(XX')^{-1}XY'=\hat{B}'} (2\pi)^{-\frac{1}{2}pN}|\Sigma|^{-\frac{1}{2}N} \cdot \exp\left\{-\frac{1}{2}\operatorname{tr} \Sigma^{-1}[YY' - 2BXY' + BXX'B']\right\} dY.$$

The integral (4.1) is evaluated by using (2.2), and we find that

$$(4.2)\quad g(G, \hat{B}) = 2^{-p}(2\pi)^{-\frac{1}{2}pN}|\Sigma|^{-\frac{1}{2}N}\prod_{i=1}^{p} C(N-p-q+i)|XX'|^{\frac{1}{2}p} \exp\{-\tfrac{1}{2}\operatorname{tr}\Sigma^{-1}[G - 2\hat{B}XX'B' + BXX'B']\}|G - \hat{B}XX'\hat{B}'|^{\frac{1}{2}(N-p-q-1)}.$$

Further, setting

$$(4.3)\quad G - \hat{B}XX'\hat{B}' = \hat{\Sigma},$$

we find that the densities of $\hat{\Sigma}$ and $\hat{B}$ are independent. $\hat{\Sigma}$ has Wishart density with $N - q$ degrees of freedom, and the pq component vector $\hat{B}$ has a pq variate normal density with mean vector B and covariance matrix $\Sigma \otimes (XX')^{-1}$.

Obviously the joint distribution of the sample mean vector and the sample covariance matrix may be obtained by using the method of this section.

In case the rank of the matrix X is $r(< q)$, then the model may be reparametrized as

$$Y = B\Delta\Delta'X + E = (\beta \ \beta_1) \begin{pmatrix} \Phi & 0 \\ 0 & 0 \end{pmatrix} H + E = \beta Z + E, \tag{4.4}$$

where

$$\Delta'XX'\Delta = \begin{pmatrix} \Phi^2 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \Delta'X = \begin{pmatrix} \Phi & 0 \\ 0 & 0 \end{pmatrix} H = \begin{pmatrix} Z \\ 0 \end{pmatrix}. \tag{4.5}$$

Here Δ is a $q \times q$ orthogonal matrix, H is a $N \times N$ orthogonal matrix, and Φ^2 is the $r \times r$ diagonal matrix of the nonzero roots of XX'. $Z = (\Phi \quad 0)H$ is $r \times N$ and of rank r, and β is $p \times r$ matrix of the estimable linear functions of the components of B. Using the transformation (4.5) we observe that the normal equations $XY' = XX'\hat{B}'$ reduce to the normal equations $ZY' = ZZ'\hat{\beta}'$. It follows that

$$\begin{aligned} g(G, \hat{\beta}) &= (2\pi)^{-\frac{1}{2}pN} |\Sigma|^{-\frac{1}{2}N} \int_{YY'=G, XY'=XX'\hat{B}'} \\ &\quad \cdot \exp\left\{-\frac{1}{2} \operatorname{tr} \Sigma^{-1}[YY' - 2BXY' + BXX'B']\right\} dY \\ &= (2\pi)^{-\frac{1}{2}pN} |\Sigma|^{-\frac{1}{2}N} \int_{YY'=G, ZY'=ZZ'\hat{\beta}'} \\ &\quad \cdot \exp\left\{-\frac{1}{2} \operatorname{tr} \Sigma^{-1}[YY' - 2\beta ZY' + \beta ZZ'\beta']\right\} dY. \end{aligned} \tag{4.6}$$

The second integral on the right hand side of the expression (4.6) is evaluated by using (2.2), and we find that $\hat{\Sigma} = G - \hat{\beta}ZZ'\hat{\beta}' = G - \hat{B}XX'\hat{B}'$ has now a Wishart distribution with $N - r$ degrees of freedom, and this is true for any particular solution $\hat{B}$ of the normal equations $XY' = XX'\hat{B}'$.

A careful observation will show that several distribution problems of multivariate analysis of variance and covariance theory involve manipulations of the integrals of the type (4.6) and its generalizations.

For further applications of generalized Sverdrup's lemma to distribution problems of regression theory, see, Kabe [3], [4].

5. Distribution of a random matrix used in classification theory. Following Sitgreaves [11] we write the joint density of a $p \times p$ positive definite symmetric matrix G, and a $p \times 2$ matrix Y as

$$g(G, Y) = C_1 \exp\{-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1}G + \operatorname{tr} k\delta'\Sigma^{-1}Y\} |G - YY'|^{\frac{1}{2}(N-p-1)}, \tag{5.1}$$

where

$$(5.2)\quad C_1^{-1} = \pi^{\frac{1}{4}p(p+1)} 2^{\frac{1}{2}p(N+2)} |\Sigma|^{\frac{1}{2}(N+2)} \prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[N - p + i]) \exp\{\tfrac{1}{2}(\delta'\Sigma^{-1}\delta)(k'k)\},$$

where k is a column vector of two components. Using (3.1), the joint density of the matrices G and $M = Y'G^{-1}Y$ may be easily found to be

$$(5.3)\quad g(G, M) = C_1C_2 \exp\{-\tfrac{1}{2}\operatorname{tr}\Sigma^{-1}G\}|G|^{\frac{1}{2}(N-p+1)}|I - M|^{\frac{1}{2}(N-p-1)} \cdot \int \exp\{k'z(\delta'\Sigma^{-1}G\Sigma^{-1}\delta)^{\frac{1}{2}}\}|M - zz'|^{-(p-4)/2}\,dz,$$

where

$$(5.4)\quad C_2 = \tfrac{1}{4}\prod_{i=1}^{2} C(p - 3 + i),$$

and z is a column vector of two components. Writing $|M - zz'| = |M| \cdot [1 - z'M^{-1}z]$, and using (3.2) we first find the joint density of the variates G, M, aed $z'M^{-1}z = u$, and then integrate out this density with respect to u; we thus find that

$$(5.5)\quad g(G, M) = C_1C_2\pi\Gamma(\tfrac{1}{2}[p - 2])|I - M|^{\frac{1}{2}(N-p-1)}|M|^{\frac{1}{2}(p-3)} \exp\{-\tfrac{1}{2}\operatorname{tr}\Sigma^{-1}G\} \cdot |G|^{\frac{1}{2}(N-p+1)} \sum_{r=0}^{\infty} (\delta'\Sigma^{-1}G\Sigma^{-1}\delta)^r(k'Mk)^r/\Gamma(\tfrac{1}{2}p + r)2^{2r}r!$$

Now we use the result (3.3) to integrate out the result (5.5) with respect to G, and we have that

$$(5.6)\quad \begin{aligned} g(M) &= \Gamma(\tfrac{1}{2}[N + 1])[\Gamma(\tfrac{1}{2}[N - p + 2])\Gamma(\tfrac{1}{2}[N - p + 1])\Gamma(\tfrac{1}{2}[p - 1])\Gamma(\tfrac{1}{2})]^{-1} \\ &\cdot \exp\{-\tfrac{1}{2}(\delta'\Sigma^{-1}\delta)(k'k)\}|M|^{\frac{1}{2}(p-3)}|I - M|^{\frac{1}{2}(N-p-1)} \\ &\cdot \sum_{r=0}^{\infty} \Gamma(\tfrac{1}{2}[N + 2] + r)(\delta'\Sigma^{-1}\delta)^r(k'Mk)^r/\Gamma(\tfrac{1}{2}p + r)2^r r!, \end{aligned}$$

a result which agrees with that of Sitgreaves ([11], p. 269, Equation 21). The density (5.6) is called as the noncentral linear multivariate beta density. The analogy of the result (5.6) with a similar result given by Kshirsagar [5] may be noted.

Taking k to be a vector of one component $N^{\frac{1}{2}}$ and Y to be a $p \times 1$ vector $N^{\frac{1}{2}}\bar{y}$, we see that the matrix $Y'G^{-1}Y$ reduces to the quadratic form $N\bar{y}'G^{-1}\bar{y}$, and setting $N\bar{y}'G^{-1}\bar{y} = (T^2/(N - 1)/1 + T^2/(N - 1))$ we may obtain the noncentral distribution of Hotelling's T^2 by using the method of this section.

6. Acknowledgment. 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] Kabe, D. G. (1963). Multivariate linear hypothesis with linear restrictions. *J. Roy. Statist. Soc. Ser. B* **25** 348–351.

[4] Kabe, D. G. (1963). Estimation of a set of fixed variates for observed values of dependent variates with normal multivariate regression models subjected to linear restrictions. *Ann. Inst. Statist. Math.* **15** 51–59.

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

[6] MacRobert, T. M. (1950). *Functions of a Complex Variable*, MacMillan, London.

[7] Mahalanobis, P. C., Bose, R. C., and Roy, S. N. (1937). Normalization of variates and the use of rectangular coordinates in the theory of sampling distributions. *Sankhyā* **3** 1–40.

[8] Narain, R. D. (1948). A new approach to sampling distributions of the multivariate normal theory I. *J. Indian Soc. Agric. Statist.* **1** 59–69.

[9] Narain, R. D. (1948). A new approach to sampling distributions of the multivariate normal theory II. *J. Indian Soc. Agric. Statist.* **1** 137–146.

[10] Olkin, Ingram and Roy, S. N. (1954). On multivariate distribution theory. *Ann. Math. Statist.* **25** 329–339.

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

[12] Sverdrup, Erling (1947). Derivation of the Wishart distribution of the second order sample moments by straightforward integration of a multiple integral. *Skand. Aktuarietidskr.* **30** 151–166.