The Annals of Statistics
1975, Vol. 3, No. 2, 467-472

CENTRAL AND NONCENTRAL DISTRIBUTIONS OF WILKS' STATISTIC IN MANOVA AS MIXTURES OF INCOMPLETE BETA FUNCTIONS[1]

BY MARIETTA J. TRETTER AND G. WILLIAM WALSTER

University of Wisconsin

In this paper it is shown that the distribution of the product of independently distributed beta random variables can be represented as a mixture of beta distributions. This result is then applied to derive the distributions of Wilks' statistic in the central and the noncentral linear cases. In the noncentral linear case the results of computations agree with those of Gupta (1971).

1. Introduction. Let $\mathbf{X}_1(f_1 \times p)$ and $\mathbf{X}_2(f_2 \times p)$ have the following joint density function:

$$f(\mathbf{X}_1, \mathbf{X}_2) = [(2\pi)^p|\mathbf{\Sigma}|]^{-(f_1+f_2)/2} \operatorname{etr}[-\mathbf{\Sigma}^{-1}\{\mathbf{X}_2'\mathbf{X}_2 + (\mathbf{X}_1 - \boldsymbol{\mu})'(\mathbf{X}_1 - \boldsymbol{\mu})\}/2]\,. \tag{1.1}$$

Without loss of generality (because all MANOVA hypotheses of equality of mean vectors can be reduced to this form) the hypothesis to be tested is H_0: $\boldsymbol{\mu}(f_1 \times p) = \mathbf{0}$ against H_1: $\boldsymbol{\mu} \neq \mathbf{0}$. The likelihood ratio criterion for this test is given by $\Lambda = |\mathbf{E}|/|\mathbf{H} + \mathbf{E}|$; where $\mathbf{H} = \mathbf{X}_1'\mathbf{X}_1$ and $\mathbf{E} = \mathbf{X}_2'\mathbf{X}_2$. In terms of the nonzero roots, ϕ_i, of the equation $|\mathbf{H} - \phi\mathbf{E}| = 0$, $\Lambda = \prod_{i=1}^r 1/(1 + \phi_i)$; where $r = \min(p, f_1)$, f_1, f_2 are the degrees of freedom, and $0 < \phi_1 \leqq \phi_2 \leqq \cdots \leqq \phi_r < \infty$. Wilks (1932) first presented the statistic Λ and included a solution (1.3) to his Type B integral equation (1.2). Later, Wilks (1935) used equation (1.3) to obtain the exact distribution of Λ for a few special cases, one of which was found to be incorrect by Consul (1966). In 1941, Wald and Brookner used equation (1.3) to derive an approximation to the null distribution of Λ. In what follows, equation (1.3) is the basis for the mixture representation of the distribution of Λ in both central and noncentral linear cases.

The Type B integral equation as given by Wilks is:

$$\int_0^B w^h g(w)\,dw = CB^h \prod_{i=1}^p \frac{\Gamma[b(i) + h]}{\Gamma[c(i) + h]}\,; \tag{1.2}$$

where $C = \prod_{i=1}^p \Gamma[c(i)]/\Gamma[b(i)]$, B and $g(w)$ are independent of h, and $b(i)$, $c(i)$ are real and positive such that $b(i) < c(i)$, for $i = 1, \cdots, p$.

Received June 1973; revised February 1974.

[1] This research was supported by the National Science Foundation, Grant No. GS-30822X and the Russell Sage Foundation.

AMS 1970 *subject classifications.* Primary—Statistics; Secondary—Distribution Theory, Multivariate Analysis, Analysis of Variance.

Key words and phrases. Incomplete beta function, mixtures, Wilks' likelihood ratio criterion, Wilks' LRC distributions, multivariate analysis of variance.

The solution of this integral equation is in the form:

$$g(w) = (Kw^{b(p)-1}(1 - w/B)^{\nu(p)-\beta(p)-1}B^{-b(p)}) \times \int_0^1 \cdots \int_0^1 \prod_{i=1}^{p-1} v_i^{c(i)-b(i)-1}(1 - v_i)^{\nu(p-i)-\beta(p-i)-1} \times [1 - \xi_i(1 - w/B)]^{b(i)-c(i+1)}\, dv_i\,; \tag{1.3}$$

where $\xi_1 = v_1$, $\xi_i = \{v_1 + v_2(1 - v_1) + \cdots + v_i(1 - v_1) \cdots (1 - v_{i-1})\}$, for $i = 2, \cdots, p-1$; $\nu(i) = \sum_{j=0}^{i-1} c(p-j)$, and $\beta(i) = \sum_{j=0}^{i-1} b(p-j)$, for $i = 1, \cdots, p$; and $K = \prod_{i=1}^{p} \Gamma[c(i)]/\Gamma[b(i)]\Gamma[c(i) - b(i)]$. Then, $g(w)$ is the distribution of $w = B\prod_{i=1}^{p} t_i$; where t_i is a beta random variable with parameters $2b(i)$ and $2[c(i) - b(i)]$. (If X is a beta random variable with parameters a and b, denote $\mathscr{L}(x) = B(a, b\colon x)$ with pdf: $f(x) = B(a/2, b/2)^{-1}x^{a/2-1}(1 - x)^{b/2-1}$.)

Relating the Type B equation to Λ, it is well known in the central case that Λ can be expressed as a product of independent beta random variables, Anderson (1958). In the noncentral linear case, Kshirsagar (1961) has shown that Λ', say, can be written as a product of independent beta random variables where one beta is noncentral. Using Kshirsagar's relationship between the beta random variables and Λ, these results can be formalized as follows. Write $\mathbf{E} = \mathbf{CLC}'$ where $\mathbf{C}$ is lower triangular such that $\mathbf{E} + \mathbf{H} = \mathbf{CC}'$. Further, write $\mathbf{L} = \mathbf{TT}'$ where $\mathbf{T}$ is a lower triangular matrix $[t_{ij}]$ of order p. Then, $\Lambda = \prod_{i=1}^{p} t_{ii}^2$; where t_{ii} is the ith diagonal element of T (hereafter denote t_{ii}^2 by t_i).

THEOREM 1.1. *In the central case Λ (p variates; f_1, f_2 degrees of freedom) is distributed as $\prod_{i=1}^{p} t_i$; where*

$$\mathscr{L}(t_i) = B(f_2 - i + 1, f_1\colon t_i)\,. \tag{1.4}$$

THEOREM 1.2. *In the noncentral linear case Λ' is distributed as $\prod_{i=1}^{p} t_i$; where*

$$\mathscr{L}(t_1) = e^{-\lambda} \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} B(f_2, f_1 + 2l\colon t_1)\,, \tag{1.5}$$

and $t_i (i = 2, \cdots, p)$ is distributed as in (1.4). *The noncentrality parameter, λ, is: $\lambda = \delta^2/2$; where δ^2 is the single nonzero root of the equation $|\boldsymbol{\mu}'\boldsymbol{\mu} - \delta^2\boldsymbol{\Sigma}| = 0$, Anderson* (1946).

These results allow application of the integral equation (1.3), with $B = 1$, to represent the pdf of Λ and Λ' in terms of $g(w)$. To evaluate $g(w)$ it is necessary to extract and collect factors of $(1 - w)^j$ from the definite integrals in (1.3). Both Wilks (1932) and Wald and Brookner (1941) indicate that expanding the terms in a McClaurin series results in a uniformly converging series that is term wise integrable.

Expanding each of the ξ_i in $g(w)$ results in a product of $(p - 1)$ series. Each term in the product of series (also uniformly convergent) can be integrated term by term yielding coefficients σ_j, say, of $(1 - w)^j$. Thus, factors of $(1 - w)^j$ can be extracted with the resulting density function, $g(w)$, in a form leading to a mixture:

$$g(w) = (Kw^{b(p)-1}(1 - w)^{\nu(p)-\beta(p)-1}) \sum_{j=0}^{\infty} \sigma_j(1 - w)^j\,. \tag{1.6}$$

While the properties of mixtures are well known, it is surprising that these properties have not been explored more widely. The method of mixtures was formally introduced by Robbins (1948) and subsequently applied to quadratic forms of normal variates by Robbins and Pitman (1949) and Ruben (1962).

2. Central case. From Theorem 1.1 it is clear that the conditions of the Type B integral equation are satisfied in the central case; where Λ corresponds to w, $B = 1$, $b(i) = (f_2 - i + 1)/2$, and $c(i) = (f_1 + f_2 - i + 1)/2$. For convenience w is substituted for Λ in the results to follow.

To liberate $(1 - w)^j$ in the definite integrals of $g(w)$, expand each of the $(p - 1)$ terms containing ξ_i in a McClaurin series, i.e.

$$\{1 - \xi_i(1 - w)\}^{-(c(i+1)-b(i))} = \sum_{k(i)=0}^{\infty} B_{k(i)}\xi_i^{k(i)}(1 - w)^{k(i)}\,; \tag{2.1}$$

where $B_{k(i)} = \Gamma[a + k(i)]/\Gamma[a]\Gamma[k(i) + 1]$, and $a = c(i + 1) - b(i) = (f_1 - 1)/2$.

Taking the product of the $(p - 1)$ series, from (2.1) we have:

$$\prod_{i=1}^{p-1} \sum_{k(i)=0}^{\infty} B_{k(i)}\xi_i^{k(i)}(1 - w)^{k(i)} = \sum_{j=0}^{\infty} (1 - w)^j \sum_{[k(i)]} \prod_{i=1}^{p-1} B_{k(i)}\xi_i^{k(i)}\,; \tag{2.2}$$

where $[k(i)]$ denotes the sum over all $k(i)$ such that $\sum_{i=1}^{p-1} k(i) = j$. Since the product series is uniformly convergent, integration and summation is interchanged with the result:

$$\begin{aligned} g(w) &= K \sum_{j=0}^{\infty} w^{b(p)-1}(1 - w)^{\nu(p)-\beta(p)-1+j} \\ &\quad\times \sum_{[k(i)]} \int_0^1 \cdots \int_0^1 \prod_{i=1}^{p-1} B_{k(i)} v_i^{c(i)-b(i)-1} \\ &\quad\times (1 - v_i)^{\nu(p-i)-\beta(p-i)-1}\xi_i^{k(i)}\, dv_i\,. \end{aligned} \tag{2.3}$$

Although Wilks, and Wald and Brookner, indicate that the integration over v_i can be performed, the literature contains no evidence of the recursion relation in the form:

$$\xi_i = v_i + (1 - v_i)\xi_{i-1}\,, \qquad i = 2, \cdots, p - 1\,. \tag{2.4}$$

This is the relation that allows development of a general recursive algorithm for evaluating the definite integrals and thus the coefficients of $(1 - w)^j$. Essentially, the recursion relation allows one to extract each v_i in the product of the ξ_i and integrate.

Before performing the integration it simplifies notation to observe that $c(i) - b(i) = f_1/2 = b$, say, and $\nu(p - i) - \beta(p - i) = (f_1/2)(p - i) = bp - bi$.

Beginning the integration with the $(p - 1)$ integral, consider:

$$\begin{aligned} \xi_{p-1}^{k(p-1)} &= \{v_{p-1} + \xi_{p-2}(1 - v_{p-1})\}^{k(p-1)} \\ &= \sum_{r(p-1)=0}^{k(p-1)} \binom{k(p-1)}{r(p-1)} v_{p-1}^{k(p-1)-r(p-1)}(1 - v_{p-1})^{r(p-1)}\xi_{p-2}^{r(p-1)}\,. \end{aligned} \tag{2.5}$$

Since this is a finite series it may be integrated term by term. Noticing that,

$$\begin{aligned} &\int_0^1 v_{p-1}^{b+k(p-1)-r(p-1)-1}(1 - v_{p-1})^{b+r(p-1)-1}\, dv_{p-1} \\ &\qquad = B[b + k(p - 1) - r(p - 1), b + r(p - 1)]\,, \end{aligned} \tag{2.6}$$

the first integration over v_{p-1} leaves the definite integral part of the solution:

$$\begin{aligned}(2.7)\qquad &\textstyle\sum_{r(p-1)=0}^{k(p-1)} C[p-1, k(p-1), r(p-1)] \\ &\quad \times \int_0^1 \cdots \int_0^1 (\prod_{i=1}^{p-2} v_i^{b-1}(1-v_i)^{bp-bi-1})\xi_1^{k(1)} \cdots \\ &\quad \xi_{p-3}^{k(p-3)}\xi_{p-2}^{k(p-2)+r(p-1)} \prod_{i=1}^{p-2} dv_i\,;\end{aligned}$$

where $C[i, k(i), r(i)] = \binom{k(i)}{r(i)} B[b + k(i) - r(i), bp - bi + r(i)]$. The next term, $\xi_{p-2}^{k(p-1)+r(p-1)}$, is expanded in a finite series and integrated with respect to v_{p-2} and the procedure is continued for all $(p-1)$ integrals giving (1.6). The coefficients σ_j are:

$$\begin{aligned}(2.8)\qquad \sigma_j = &\textstyle\sum_{[k(i)]} (\prod_{i=1}^{p-1} B_{k(i)}) \sum_{r(p-1)=0}^{k(p-1)} C[p-1, k(p-1), r(p-1)] \\ &\times \textstyle\sum_{r(p-2)=0}^{k(p-2)+r(p-1)} C[p-2, k(p-2)+r(p-1), r(p-2)] \cdots \\ &\textstyle\sum_{r(2)=0}^{k(2)+r(3)} C[2, k(2)+r(3), r(2)]B[b+k(1)+r(2), bp-b]\,.\end{aligned}$$

Hence the coefficients of $(1-w)^j$ are in the form of a finite sum of products of beta functions with all positive arguments. The σ_j are all positive and the series converges, since $0 \leqq w \leqq 1$. Thus the cdf obtained by interchanging integration and summation:

$$(2.9)\qquad G(w) = K \textstyle\sum_{j=0}^{\infty} \sigma_j \int_0^w u^{b(p)-1}(1-u)^{bp+j-1}\,du\,.$$

Multiplying and dividing by $B[b(p), bp+j]$, and adopting the notation $I(a, b: x) = B(a, b)^{-1} \int_0^x u^{a-1}(1-u)^{b-1}\,du$ for the incomplete beta function, yields:

$$(2.10)\qquad G(w) = K \textstyle\sum_{j=0}^{\infty} \sigma_j B[b(p), bp+j]I[b(p), bp+j: w]\,,$$

which is the central cdf of Wilks' Λ written as a mixture of incomplete beta functions.

3. Noncentral case. From Theorem 1.2, Λ' is also expressed as a product of t_i; where the joint distribution is now an infinite series of products of beta random variables:

$$(3.1)\qquad \mathscr{L}(\textstyle\prod_{i=1}^{p} t_i) = e^{-\lambda} \sum_{l=0}^{\infty} \dfrac{\lambda^l}{l!} B(f_2, f_1 + 2l: t_1) \prod_{i=2}^{p} B(f_2 - i + 1, f_1: t_i)\,.$$

Wilks' solution can be applied to each term in the series. Replacing Λ' by w' and $c(1)$ by ${}_lc(1) = c(1) + l$,

$$(3.2)\qquad g_l(w') = {}_lK \textstyle\sum_{j=0}^{\infty} {}_l\sigma_j w'^{b(p)-1}(1-w')^{bp+j+l-1}\,;$$

where ${}_lK$ and ${}_l\sigma_j$ are the same as K and σ_j except ${}_lc(1)$ replaces $c(1)$.

From the properties of Wilks' integral equation, $\int_0^{w'} g_l(u')\,du' = G_l(w')$ is a unique distribution function represented as a mixture, and $\sum_{l=0}^{\infty} e^{-\lambda}\lambda^l/l! = 1$. Thus the properties of a mixture are satisfied and $G(w')$ can be written as a mixture of the $G_l(w')(l = 0, \cdots, \infty)$:

$$(3.3)\qquad G(w') = e^{-\lambda} \textstyle\sum_{l=0}^{\infty} \dfrac{\lambda^l}{l!} \int_0^{w'} {}_lK \sum_{j=0}^{\infty} {}_l\sigma_j u'^{b(p)-1}(1-u')^{bp+j+l-1}\,du'\,.$$

Analogous to the central distribution, integration and summation can again be interchanged to give the cdf of Λ':

$$G(w') = e^{-\lambda} \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} {}_lK \sum_{j=0}^{\infty} {}_l\sigma_j B[b(p), bp + j + l] \times I[b(p), bp + j + l : w'] ; \tag{3.4}$$

which is again a mixture of incomplete beta functions.

4. Conclusion. The mixture representations of the cdf's of Λ and Λ' have two distinct computational advantages over previous representations. First, the properties of a mixture lead directly to an absolute bound on the error from truncating a mixture, Robbins and Pitman (1949) and Ruben (1962). Second, in the present case, the computation of σ_j, ${}_l\sigma_j$, and the incomplete beta functions can be performed using numerically stable recursion relations. Numerical instability severely limits the utility of the representations proposed by Schatzoff (1966) and Gupta (1971) for the cdf's of Λ and Λ', respectively. As Schatzoff points out,

> An unfortunate characteristic of the exact formulae is that the number of significant digits which must be retained in intermediate calculations increases as p, g and n increase.

Finally, the mixture representation for the cdf of Λ' has been programmed, Tretter (1973), and produced results that agree with those presented by Gupta (1971).

REFERENCES

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

[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. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

[4] Consul, P. C. (1966). On the exact distribution of the likelihood ratio criteria for testing linear hypotheses about regression coefficients. *Ann. Math. Statist.* **37** 1319–1330.

[5] Gupta, A. K. (1971). Noncentral distribution of Wilks' statistic in MANOVA. *Ann. Math. Statist.* **42** 1254–1261.

[6] Kshirsagar, A. M. (1961). The non-central multivariate beta distribution. *Ann. Math. Statist.* **32** 104–111.

[7] Pillai, K. C. S. and Gupta, A. K. (1969). On the exact distribution of Wilks' criterion. *Biometrika* **56** 109–118.

[8] Robbins, H. (1948). Mixtures of distributions. *Ann. Math. Statist.* **19** 360–369.

[9] Robbins, H. and Pitman, E. J. G. (1949). Application of the method of mixtures to quadratic forms in normal variates. *Ann. Math. Statist.* **20** 552–560.

[10] Roy, S. N. (1957). *Some Aspects of Multivariate Analysis.* Wiley, New York.

[11] Ruben, H. (1962). Probability content of regions under spherical normal distributions, IV: The distribution of homogeneous and non-homogeneous quadratic functions of normal variables. *Ann. Math. Statist.* **33** 542–570.

[12] SCHATZOFF, M. (1966). Exact distributions of Wilks's likelihood ratio criterion. *Biometrika* **53** 347–358.

[13] TRETTER, M. (1973). A mixture representation of the cumulative distribution of Wilks' likelihood ratio criterion in the null and non-null linear cases. Ph. D. dissertation, Univ. of Wisconsin.

[14] WALD, A. and BROOKNER, R. J. (1941). On the distribution of Wilks' statistic for testing independence of several groups of variates. *Ann. Math. Statist.* **12** 137–152.

[15] WILKS, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika* **24** 471–494.

[16] WILKS, S. S. (1935). On the independence of K sets of normally distributed statistical variables. *Econometrica* **3** 309–326.

DEPARTMENT OF SOCIOLOGY
UNIVERSITY OF WISCONSIN
SOCIAL SCIENCE BUILDING
MADISON, WISCONSIN 53706