Statistics & Probability Letters 2 (1984) 165-168
North-Holland

May 1984

ON THE DISTRIBUTION OF THE PRODUCT OF INDEPENDENT BETA RANDOM VARIABLES

Jen TANG

Marquette University, Milwaukee, WI 53233, USA

A.K. GUPTA

Bowling Green State University, Bowling Green, OH 43403, USA

Received September 1983

Abstract: In this paper, exact distribution of the product of independent beta random variables has been derived and its structural form is given together with recurrence relations for the coefficients of this representation. These recurrence relations yield a direct computational algorithm for computing the percentage points of many test criteria in multivariate statistical analysis.

AMS 1979 Subject Classification: Primary 62H10; Secondary 62H15.

Keywords and phrases: exact density, coefficients, recurrence relations, testing multivariate statistical analysis.

1. Introduction

Wilks (1932) has considered the following type-B integral equation:

$$\int_0^1 w^h g_p(w)\,\mathrm{d}w = \prod_{j=1}^{p} \frac{\Gamma[c(j)]\Gamma[b(j)+h]}{\Gamma[b(j)]\Gamma[c(j)+h]} \quad (h=0, 1, 2, \ldots), \tag{1.1}$$

where $g_p(w)$ is unknown and independent of h, $b(j)$ and $c(j)$ are real and positive such that $c(j) > b(j)$, for $j = 1, 2, \ldots, p$, and $\Gamma[\cdot]$ is the gamma function. Evidently, $g_p(w)$ is the density function of $W_p \stackrel{\text{def}}{=} \prod_{j=1}^{p} t_j$ say, where $\mathscr{L}(t_j) = \text{Beta}[b(j), c(j)-b(j)]$, $j = 1, \ldots, p$, are independent. Here we define the Beta$[m, n]$ density as $B^{-1}[m, n]\, u^{m-1}(1-u)^{n-1}$, $0 < u < 1$.

The distribution of W_p, known as Wilks' type-B distribution is very useful in multivariate statistical hypothesis testing. In fact, many likelihood ratio test criteria of multivariate hypotheses can be expressed as such products; for example, (i) general linear hypothesis (see Anderson, 1958, p. 187), (ii) equality of several mean vectors with same covariance matrix (see Anderson, 1958, p. 211); that is, the MANOVA hypothesis, (iii) independence between every pair of sets of variables (Anderson, 1958, p. 237), (iv) the sphericity hypothesis (Anderson, 1958, p. 259), (v) equality of covariance matrices (Anderson, 1958, p. 247).

Wilks (1932) has expressed $g_p(w)$ in the form of a $(p-1)$-fold multiple integral, which he was able to evaluate for some special cases only. Using the Fourier inversion theorem, Wald and Brookner (1941) have obtained the structural form for $g_p(w)$ as a mixture of beta functions. Recently, Gupta (1977), Walster and Trotter (1980) and Nandi (1980) have obtained some formulas as $(p-1)$-fold series for computing the coefficients of the mixture representation. Schatzoff (1966), Pillai and Gupta (1969), and Gupta (1971) have applied convolution technique to obtain expressions for the MANOVA model for small p, which Khatri and Bhargava (1981) have modified to obtain $g_p(w)$ in the form of a p-fold multiple series. Consul (1968) has used the inverse Mellin transform and some special properties of hypergeometric functions to derive the exact distributions of W_p for a limited number of cases in various problems. Using contour integration, Mathai (1971) has obtained the pdf of W_p in series form. Mathai (1973) has written a review of different methods of obtaining exact distributions of multivariate test criteria and pointed out shortcomings of these methods.

In this paper, we obtain the exact distribution of W_p by deriving both the structural form for $g_p(w)$ and *simple recurrence relations* for the coefficients of this representation. Recurrence relations yield a direct computational algorithm for computing the percentage points for many test criteria, hence our results are very useful from theoretical as well as practical points of view.

2. Distribution of product of independent beta random variables

We state the main result of this paper as follows.

Theorem 2.1. *The exact density function of W_p is given by*

$$g_p(w) = K_p w^{b(p)-1}(1-w)^{f(p)-1} \sum_{r=0}^{\infty} \sigma^{(p)}(1-w)^r, \quad 0<w<1, \tag{2.1}$$

where K_p, f(p) are defined by

$$K_m = \prod_{j=1}^{m} \{\Gamma[c(j)]/\Gamma[b(j)]\}, \qquad f(m) = \sum_{j=1}^{m} [c(j)-b(j)],$$

for $1 \leqslant m \leqslant p$, and $\sigma_r^{(p)}$ can be obtained from the following recurrence relation:

$$\sigma_r^{(k)} = \frac{\Gamma[f(k-1)+r]}{\Gamma[f(k)+r]} \sum_{s=0}^{r} \left[(c(k)-b(k-1))_s/s!\right] \sigma_{r-s}^{(k-1)} \quad \text{for } r=0,1,2,\ldots,\ k=2,3,\ldots,p, \tag{2.2}$$

with initial values

$$\sigma_0^{(1)} = \Gamma^{-1}[f(1)], \qquad \sigma_r^{(1)} = 0 \quad \text{for } r=1,2,\ \ldots .$$

Proof. For $2 \leqslant k \leqslant p$, p fixed, write

$$W_{k,p} = \prod_{j=1}^{k} t_j = \left(\prod_{j=1}^{k-1} t_j\right) t_k = W_{k-1,p} t_k$$

say, where $W_{p,p} = W_p$ and we assume that the density of $W_{k-1,p}$ can be written as

$$g_{k-1,p}(w) = K_{k-1} w^{b(k-1)-1}(1-w)^{f(k-1)-1} \sum_{r=0}^{\infty} \sigma_r^{(k-1)}(1-w)^r.$$

Using the usual transformation technique for products and Euler's formula for the hypergeometric function ${}_2F_1$ (see Luke, 1969, p. 57), we can obtain the density of $W_{k,p}$ as follows:

$$g_{k,p}(w) = K_k w^{b(k)-1}(1-w)^{f(k)-1} \sum_{r=0}^{\infty} \sigma_r^{(k-1)} \Gamma[f(k-1)+r]$$
$$\times \Gamma^{-1}[f(k)+r](1-w)^r {}_2F_1[f(k-1)+r, c(k)-b(k-1); f(k)+r; 1-w]. \tag{2.3}$$

Since ${}_2F_1$ is a power series in $1-w$, we can rewrite (2.3) by combining powers of $1-w$ as

$$g_{k,p}(w) = K_k w^{b(k)-1}(1-w)^{f(k)-1} \sum_{r=0}^{\infty} \sigma_r^{(k)}(1-w)^r, \tag{2.4}$$

where $g_{p,p} = g_p$ and $\sigma_r^{(k)}$, $r = 0, 1, 2, \ldots$, clearly satisfy (2.2) for $2 \leqslant k \leqslant p$.

Note that both (2.1) and (2.3) with $k = p$ are representations for the density of W_p. It is easily seen from (2.2) that $\sigma_r^{(p)}$, $r = 0, 1, \ldots, m$, can be computed from $\sigma_r^{(k)}$, $r = 0, 1, 2, \ldots, m$, $k = 1, 2, \ldots, p-1$. Since $b(j)$ and $c(j)$ may depend on p, $g_{k,p}(w)$ given in (2.3) may not be the density of W_k unless $b(j)$ and $c(j)$ are independent of p. (2.2) provides formulas for computing distributions of W_k, $k = 1, 2, 3, \ldots$, successively and its pdf is given in (2.4).

The following corollary, which gives the exact density of the product of two independent beta variables, is an immediate consequence of Theorem 2.1.

Corollary 2.1. *The exact density of W_2 is given by*

$$g_2(w) = K_2 \Gamma^{-1}[f(2)]\, w^{b(2)-1}(1-w)^{f(2)-1} {}_2F_1[f(1), c(2)-b(1); f(2); 1-w] \quad \text{for } 0 < w < 1. \tag{2.5}$$

Note that many of Consul's results can easily be obtained from this corollary and simple transformations with the help of duplication formula for gamma functions (see Srivastava and Khatri, 1979, Chapter 7).

In many applications, it is possible to evaluate $\sigma_r^{(p)}$ using the fact that $\int_0^1 g_p(w)\, dw = 1$. From (2.1) we obtain

$$K_p^{-1} = \sum_{r=0}^{\infty} \sigma_r^{(p)} \Gamma[b(p)] \Gamma[f(p)+r] / \Gamma[b(p)+f(p)+r]. \tag{2.6}$$

Both sides of (2.6) can, in general, be expressed as power series in $n^{-1} = (\text{sample size})^{-1}$. Upon equating powers of n^{-1}, we can obtain different recurrence relations for $\sigma_r^{(p)}$, $r = 0, 1, 2, \ldots$ (see Tang, 1983).

3. Concluding remarks

The cumulative distribution function of W_p, obtained by integrating (2.1) term by term, is a mixture of incomplete beta functions which can also be evaluated recursively. The exact null pdf's of test criteria mentioned in Section 1 can be obtained by simple substitution of the parameters. Hence our results are useful for obtaining exact distributions and exact percentage points of many test criteria in multivariate hypothesis testing.

References

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

Consul, P.C. (1969), The exact distributions of likelihood ratio criteria for different hypotheses, in: Krishnaiah, ed., *Multivariate Analysis II* (Proc. 2nd Internat. Symp.) pp. 171–181.

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

Gupta, A.K. (1977), On the distribution of sphericity test criterion in the multivariate Gaussian distribution, *Aust. J. Statist.* **19**, 202–205.

Khatri, C.G. and R.P. Bhargava (1981), The distribution of product of independent beta random variables with application to multivariate analysis, *Ann. Inst. Statist. Math.* **33**, 281–293.

Luke, Y.L. (1969), *The Special Functions and Their Approximations, Vol. 1,* Academic Press, New York.

Mathai, A.M. (1971), An expansion of Meijer's *G*-function and the distribution of products of independent beta variates, *Statist. J.* **5**, 71–90.

Mathai, A.M. (1973), A review of the various methods of obtaining the exact distributions of multivariate test criteria, *Sankhya A* **35**, 39–60.

Nandi, S.B. (1980), On the exact distribution of a normalized ratio of the weighted geometric mean to the unweighted arithmetic mean in samples from gamma distributions, *J. Amer. Statist. Assoc.* **75**, 217–220.

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

Schatzoff, M. (1966), Exact distribution of Wilks' likelihood ratio criterion, *Biometrika* **53**, 347–358.

Srivastava, M.S. and C.G. Khatri (1979), *An Introduction to Multivariate Statistical Analysis,* North-Holland, Amsterdam, New York.

Tang, J. (1983), A note on the central and noncentral distributions of Wilks' statistics in MANOVA (unpublished).

Wold, A. and R.J. Brookner (1941), On the distribution of Wilks' statistic for testing independence of several group of variates, *Ann. Math. Statist.* **12**, 137–152.

Wolster, G.W. and M.J. Tretter (1980), Exact noncentral distribution of Wilks' and Wilks–Lawley criteria as mixtures of incomplete beta functions: For three tests, *Ann. Statist.* **8**, 1388–1390.

Wilks, S.S. (1932), Certain generalizations in the analysis of variance, *Biometrika* **24**, 471–494.