J. Statist. Reserch
Vol. 20, 1986

NEW REPRESENTATIONS FOR THE DISTRIBUTION OF A CLASS OF LIKELIHOOD RATIO CRITERIA

by

Panagis G. Moschopoulos
The University of Texas at Dallas
P.O. Box 830688
Richardson, Texas 75083-0688

ABSTRACT

An asymptotic expansion of the distribution function of a class of likelihood ratio criteria is obtained by a straightforward inversion of the corresponding characteristic function. In particular, computational alternatives to Box's (1949) series are obtained and illustrated.

Key Words: Likelihood ratio, moments, generalized Bernoulli polynomials.

1. INTRODUCTION

A number of likelihood ratio criteria statistics, say U, have moments of the form

$$E(U)^h = \int_0^1 u^h f(u)du = C \prod_{j=1}^{p} \frac{\Gamma[z+a_j+h]}{\Gamma[z+b_j+h]} \qquad (1.1)$$

For $0 < u < 1$ and hence the distribution of U is completely determined by the moments. Two important examples are as follows:

1) Linear hypothesis. Let $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_N$ be a set of N p-variate observations, $\underset{\sim}{x}_\alpha$ drawn from a p-variate normal distribution $N_p(B\underset{\sim}{Z}_\alpha, \Sigma)$; the matrix $\Sigma(p\text{x}p)$ is the common unknown covariance matrix and $B(p\text{x}q)$ is the matrix of the unknown regression coefficients. Let λ be the likelihood ratio criterion for testing the hypothesis $H_1: B_1(p\text{x}q_1) = B_1^*$, i.e., the hypothesis that a submatrix B_1 of B is equal to a given matrix. Then, when H_1 is assumed true, the h-th moment of $U = \lambda^{2/N}$ is of the form (1.1) with $z = (n-q)/2$, $a_j = (1-j)/2$, $b_j = q/2+a_j$, see Anderson (1958), p. 188.

2) Sphericity. Let λ_1 be the likelihood ratio criterion for testing the hypothesis H_2: $\Sigma = \sigma^2 I$, where Σ is the covariance matrix of a p-variate normal distribution $N_p(\underset{\sim}{\mu}, \Sigma)$. When H_2 is assumed true, then the h-th moment of $U = \lambda_1^{\frac{N-1}{N}}$ is of the form (1.1) where there are p-1 gamma ratios instead of p, $z = (N-1)/2$, $a_j = -j/2$, $b_j = p^{-1}j$, see Khatri and Srivastava (1979), p. 209.

Exact series representations for the distribution of are most commonly obtained through the use of the inverse Mellin integral transform, e.g. see Consul (1966), Mathai and Rathie (1971); diredt integration methods, Wilks

(1935); successive convolutions, Schatzoff (1966). Pertinent among the series used for computational purposes is Box's (1949) series, see Anderson (1958) p. 203. The series is used widely for approximations and exact computations, e.g., see Pillai and Gupta (1969), Nagarsenker and Pillai (1973).

Box's series is obtained in Moschopoulos (1983) by asymptotically expanding the gamma ratios in (1.1) and taking the inverse Mellin tranform. This leads to alternative analytic expressions for all the coefficients. In this paper: 1) An asymptotic expansion of the distribution of $-\log U$ is derived by inverting the characteristic function; 2) simple computational formulae are given for computing the asymptotic expansion to any desired degree of accuracy.

2. THE RESULT

Consider a random variable U with support on (0,1) having moments of the form (1.1). Then, for $i^2 = -1$, the characteristic function M(h) of $-\log U$ is:

$$M(h) = E(e^{-ih\log U}) = E(U^{-ih}) = C\prod_{j=1}^{p}\frac{\Gamma[z+a_j-ih]}{\Gamma[z+b_j-ih]}. \quad (2.1)$$

where C is such that M(0) = 1.

We seek an expansion of the characteristic function above in a series in powers of m^{-1} where $m = z-\alpha$ for α fixed and $z\to\infty$; the constant α is later determined so that the convergence of the series is accelerated. Upon substituting $m = z-\alpha$ in (2.1) and letting

$$U_* = U^m,\ \alpha_j = a_j + \alpha\ ,\ \beta_j = b_j + \alpha,\ j = 1,\ldots,p,$$

the expression (2.1) is modified to the following

$$M'(h) = E(e^{-ih\log U_*}) = C\prod_{j=1}^{p}\frac{\Gamma[m(1-ih)+\alpha_j]}{\Gamma[m(1-ih)+\beta_j]}. \quad (2.2)$$

To obtain the expansion we employ the following well known result on the expansion of the ratio of two gamma functions, see Luke Vol I (1969), p. 33. For $|\arg(t+a)|\leq\pi-\varepsilon$, $\varepsilon>0$ and $(a)_n = a(a+1)\ldots(a+n-1)$, we have

$$\frac{\Gamma(t+\alpha_j)}{\Gamma(t+\beta_j)} = t^{\alpha_j-\beta_j}\sum_{r=0}^{k} G(j,r)t^{-r}+O(t^{-k-1}) \tag{2.3}$$

where

$$G(j,r) = (-1)^r B_r(\alpha_j,\alpha_j-\beta_j+1)\ (\beta_j-\alpha_j) r/r!$$

and $B_r(x,a)$ is the generalized Bernoulli polynomial of order r. For given x and a, the values of these polynomials are easily derived on the computer for all r by using $B_0(x,a)=1$ and the recursive formula, Luke (1969) p. 20

$$aB_r(x,a+1) = (a-r)B_r(x,a) + r(x-a)B_{r-1}(x,a). \tag{2.4}$$

Put now $t=m(1-ih)$ in (2.3) and multiply the p series for $j=1,\ldots,p$. This leads to the following asymptotic expansion for the product of gamma ratios in (2.2)

$$\prod_{j=1}^{p}\frac{\Gamma[m(1-ih)+\alpha_j]}{\Gamma[m(1-ih)+\beta_j]} = [m(1-ih)]^{-\rho}\sum_{r=0}^{k}\gamma_r^{(p)}[m(1-ih)]^{-r}+O(m^{-k-1}), \tag{2.5}$$

where $\rho = \sum_{j=1}^{p}(\beta_j-\alpha_j) > 0$ and the coefficients $\gamma_r^{(p)}$, $r=0,1,\ldots$ are easily derived on the computer recursively. We have

$$\gamma_r^{(1)} = G(1,r),$$

and

$$\gamma_r^{(s)} = \sum_{\ell=0}^{r}\gamma_r^{(s-1)}\ G(s,r-\ell),\ s=2,3,\ldots,p,$$

for $r=0,1,\ldots,k$. Similarly we obtain an asymptotic expansion for the product of the gamma ratios denoted by C. This time (2.3) is applied with t=m and α_j and β_j interchanged. The result is

$$C = m^{\rho} \sum_{r=0}^{k} \delta_r^{(p)} m^{-r} + O(m^{-k-1}) \tag{2.6}$$

where

$$\delta_r^{(1)} = G'(1,r),$$

$$\delta_r^{(s)} = \sum_{\ell=0}^{r} \delta_r^{(s-1)} G'(s, r-\ell), \quad s=2,3,\ldots,p$$

and G' is G with α_j and β_j interchanged.

Thus, the characteristic function M'(h) in (2.2), using (2.5) and (2.6) is expressed as

$$M(h) = \sum_{r=0}^{k} c_r m^{-r} (1-ih)^{-(r+\rho)} + O(m^{-k-1}) \tag{2.7}$$

where

$$c_r = \sum_{\ell=0}^{r} \gamma_\ell^{(p)} \delta_{r-\ell}^{(p)}.$$

The density and distribution function of $-\log \nu$ are obtained by term-by-term inversion of (2.7). On putting 2h for h in (2.7), since the term $(1-2ih)^{-(r+\rho)}$ is the characteristic function of a chi-square distribution with $2(r+\rho)$ degrees of freedom, the density of $y=-2 \log \nu$ * is obtained as

$$g(y) = \sum_{r=0}^{k} c_r m^{-r} \frac{y^{r+\rho-1} e^{-\frac{y}{2}}}{\Gamma(r+\rho) \cdot 2^{r+\rho}} + O(m^{-k-1}), \quad y>0. \tag{2.8}$$

The distribution function $F(x) = Pr(-2m \log u \le x)$ is

$$F(x) = \sum_{r=0}^{k} c_r m^{-r} P_r\{\chi^2_{2(r+\rho)} \le x\} + O(m^{-k-1}). \tag{2.9}$$

Finally, the constant α is determined so that the $O(m^{-1})$ term in (2.9) vanishes. For this we require

$$\gamma_1^{(p)} = \sum_{j=1}^{p} (\beta_j - \alpha_j) B_1(\alpha_j, \alpha_j - \beta_j + 1)$$

$$= \frac{1}{2} \sum_{j=1}^{p} (b_j - a_j)(a_j + b_j - 1 - 2\alpha) = 0. \tag{2.10}$$

3. NUMERICAL ILLUSTRATIONS

We now illustrate the method in the case of the linear hypothesis. As noted earlier, in this case $z = \frac{n}{2} = \frac{(N-q)}{2}$, $a_j = (1-j)/2$ and $b_j = a_j + q/2$. For these values of a_j and b_j we have $\rho = pq/2$, and from (2.10) for $\gamma_1^{(p)} = 0$ we obtain $\alpha = (q-p-1)/4$ and hence $m = (1/2)(n + (q-p-1)/2)$. It follows that (2.9) becomes an alternative form of Box's series; note however the simplicity in the computation of the coefficients c_r.

To illustrate the computations, let w be the $100(1 - \alpha)\%$ point so that $Pr(-2\ m \log U \geq w) = \alpha$. These percentiles w, for $\alpha = .01$ or $.05$ and given n, p, q, are available in Table 47 of Pearson and Hartley (1972). Table 1 below illustrates the convergence of $F(x)$ to the true probability for selected percentiles. The numbers in parentheses denote the number of terms from (2.9) needed to achieve the indicated probability. Except in some rather extreme cases (e.g., n=8, p=7, q=10) in which the number of terms is up to fifteen, less than ten terms are generally required; for $n \geq 10$ three or four terms were needed for these results. The CPU time was absolutely negligible due to the efficiency of the formulae. Computations were done in single precision, using the IMSL routine MDGAM to evaluate the chi-square probabilities.

This research was partially sponsored by an award from the State of Texas Organized Research Fund. Thanks are due to Dalya Scheibal for carrying out the computations, and to A. M. Mathai and an anonymous referee for helpful suggestions for improving the presentation.

Table 1. THE CONVERGENCE OF F(z) IN (2.9)

α	p	n	q=2	q=6	q=10
.05	3	4	.0499(4)	.0498(8)	.0499(10)
		8	.0499(3)	.0498(4)	.0498(6)
		12	.0500(3)	.0500(3)	.0499(4)
		17	.0500(2)	.0499(3)	.0499(4)
		32	.0499(2)	.0499(3)	.0498(4)
		62	.0500(1)	.0500(2)	.0501(2)
.05	7	8	.0499(9)	.0501(10)	.0500(14)
		12	.0499(4)	.0501(5)	.0502(6)
		16	.0499(4)	.0501(4)	.0498(7)
		21	.0499(4)	.0501(3)	.0498(4)
		36	.0498(3)	.0501(3)	.0500(3)
		66	.0500(2)	.0499(2)	.0500(3)
.01	3	4	.0100(5)	.0099(8)	.0100(11)
		8	.0099(2)	.0100(4)	.0099(6)
		12	.0099(2)	.0100(3)	.0099(4)
		17	.0100(2)	.0100(3)	.0099(4)
		32	.0100(2)	.0100(2)	.0100(3)
		62	.0100(1)	.0100(2)	.0099(2)
,01	7	8	.0100(8)	.0099(12)	.0099(15)
		12	.0099(4)	.0099(6)	.0100(8)
		16	.0100(4)	.0099(4)	.0100(6)
		21	.0099(3)	.0099(3)	.0099(4)
		36	.0100(2)	.0099(2)	.0099(3)
		66	.0100(2)	.0099(2)	.0100(2)

(   ) = required number of terms from (2.9).

REFERENCES

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

Box, G. E. P. (1949). A general distribution theory for a class of likelihood ratio criteria. Biometrika 36, 317-346.

Consul, P. C. (1966). Exact distributions of the likelihood ratio. Ann. Math. Statist. 37, 1319-1330.

Lee, Y. S. (1972b). Some results on the distribution of Wilks' likelihood ratio criterion. Biometrika 59, 649-664.

Luke, Y. L. (1969). The Special Functions and Their Approximations. Vol. I, Acad. Press. New York.

Mathai, A. M., and Rathie, P. N. (1971). The exact distribution of Wilks' criterion. Ann. Math. Statist. 42, 3, 1010-1019.

Khatri C. G., Srivastava M. S. (1979). An Introduction to Multivariate Statistics. North Holland: New York.

Moschopoulos, P. G. (1983). A general procedure for deriving distributions. Commun. Statist. A, 12(17), 2005-2015.

Nagarsenker, B. N., and Pillai K. C. S. (1973). The distribution of the sphericity test criterion. J. Mult. Anal. 3, 226-235.

Pearson, E. S., Hartley, H. O. (1972). Biometrika Tables for Statisticians. Cambridge University Press.

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

Schatzoff, M. (1966). Exact distribution of Wilks' likelihood ratio criterion. Biometrika 53, 347-58.

Wilks, S. S. (1935). On the independence of k sets of normally distributed variables. Econometrica 3, 309-326.