The Annals of Statistics
1978, Vol. 6, No. 1, 235-238

POSITIVE DEPENDENCE OF THE ROOTS OF A WISHART MATRIX

BY RICHARD L. DYKSTRA AND JOHN E. HEWETT

University of Missouri—Columbia

It is shown that the characteristic roots of a Wishart matrix (identity covariance matrix) and the roots of $S_1 S_2^{-1}$ and $S_1(S_1 + S_2)^{-1}$ where S_1, S_2 are independent $p \times p$ Wishart matrices with the same covariance matrix, satisfy certain types of dependency relationships. That is, it is shown that these roots are (a) positive orthant dependent, (b) associated, (c) stochastically increasing in sequence, and (d) positively likelihood ratio dependent. An example of how this may be used in obtaining simultaneous confidence intervals is also included.

1. Introduction. The joint probability density function of the ordered roots of a $p \times p$ Wishart matrix with covariance matrix I and m degrees of freedom is well known. That is, if $\infty > \lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_p \geqq 0$ denote these roots, their joint density is given by

$$f_1(\lambda_1, \cdots, \lambda_n) = \frac{\pi^{p/2} \prod_{i=1}^{p} \lambda_i^{\frac{1}{2}(m-p-1)} \exp -\frac{1}{2} \sum_{i=1}^{p} \lambda_i \prod_{i<j} (\lambda_i - \lambda_j)^+}{2^{pm/2} \prod_{i=1}^{p} \{\Gamma[\frac{1}{2}(m+1-i)] \cdot \Gamma[\frac{1}{2}(p+1-i)]\}}, \qquad \lambda_p \geqq 0$$

where $(\lambda_i - \lambda_j)^+ = \max\{(\lambda_i - \lambda_j), 0\}$. Anderson (1958) among others gives a derivation of this density. Since the roots themselves are important statistics and since the distribution of a given subset of these roots is quite difficult and intractable (see, for example, Waikar (1975)), any pattern to the behavior of these roots is of interest. It is the purpose of this note to point out that the characteristic roots of a Wishart matrix satisfy certain types of positive dependence relationships.

2. Types of positive dependence. Lehmann (1966) has defined several types of dependence relationships for the bivariate case which have been extended to the multivariate situation. The following types of positive dependence will be of interest to us in this paper.

(a) *Positive orthant dependence.* We shall say that random variables $X_1, X_2, \cdots, X_p$ are positively orthant dependent if

$$P(X_i \leqq x_i; i = 1, \cdots, p) \geqq \prod_{i=1}^{p} P(X_i \leqq x_i)$$

for every choice of $x_1, x_2, \cdots, x_p$. This type of dependence is discussed in several sources including Esary, Proschan and Walkup (1967) and Dykstra, Hewett and Thompson (1973).

Received October 1976; revised March 1977.

AMS 1970 *subject classification.* Primary 62H99.

Key words and phrases. Wishart distribution, characteristic roots, positive dependence, totally positive of order two.

(b) *Associativity.* The random variables $X_1, \cdots, X_p$ are associated if $\operatorname{Cov}(g(X_1, \cdots, X_p), h(X_1, \cdots, X_p)) \geqq 0$ for all functions g and h which are nondecreasing in each argument and for which the covariance is defined. This concept has been extensively discussed in Esary et al. (1967).

(c) *Stochastically increasing in sequence.* Esary and Proschan (1968) define random variables $X_1, \cdots, X_p$ to be stochastically increasing in sequence if the conditional distribution function of X_i given $X_{i+1} = x_{i+1}, \cdots, X_p = x_p$, i.e., $F(x_i \mid x_{i+1}, \cdots, x_p)$, is nonincreasing in $x_{i+1}, \cdots, x_p$ for $i = 1, \cdots, p-1$.

(d) *Positively likelihood ratio dependent.* This concept is defined in Dykstra et al. (1973). The random variables $X_1, \cdots, X_p$ have this property if for $x_j' \geqq x_j$, $j = 1, \cdots, p$,

$$f_i(x_i, x_{i+1}, \cdots, x_p) f_i(x_i', x_{i+1}', x_p') \geqq f_i(x_i', x_{i+1}, \cdots, x_p) f_i(x_i, x_{i+1}', \cdots, x_p'), \quad i = 1, 2, \cdots, p-1,$$

where f_i denotes the joint density of $X_i, X_{i+1}, \cdots, X_p$.

It should be noted that the concept of positive likelihood ratio dependence is closely connected with the concept of TP_2 (totally positive of order 2) densities. In particular a nonnegative function $h(\lambda_1, \lambda_2)$ is said to be TP_2 if

$$\begin{vmatrix} h(\lambda_1, \lambda_2) & h(\lambda_1', \lambda_2) \\ h(\lambda_1, \lambda_2') & h(\lambda_1', \lambda_2') \end{vmatrix} \geqq 0$$

for all $\lambda_i' \geqq \lambda_i$, $i = 1, 2$ (see Karlin (1968) for an extensive discussion on total positivity). Then it is straightforward to show that positive likelihood ratio dependence is equivalent to $f_j(x_j, x_{j+1}, \cdots, x_p)$ being TP_2 as a function of λ_j and λ_k (holding the other variables fixed) for $j = 1, 2, \cdots, p-1$ and $k > j$.

It is established in Theorem 2 of Dykstra et al. (1973) that these types of dependence are successively stronger. That is, each type of positive dependence also implies the types listed before it. Thus if we can show that the characteristic roots satisfy the fourth type of dependence, they must also satisfy the other three. We are now ready to state our result.

3. Main result.

THEOREM. *The characteristic roots of a Wishart matrix with identity covariance matrix satisfy the four types of positive dependence defined in Section* 2.

PROOF. From previous statements it will suffice to show that $f_j(\lambda_j, \lambda_{j+1}, \cdots, \lambda_p)$ is TP_2 as a function of λ_j and λ_k for $j = 1, 2, \cdots, p-1$ and $k > j$. First consider the case $j = 1$. Since the joint density of all p roots is of the form

$$\prod_{i=1}^{p} g_i(\lambda_i) \cdot \prod_{i<j} (\lambda_i - \lambda_j)^+ ,$$

$f_1(\lambda_1, \cdots, \lambda_p)$ will be TP_2 in each pair of variables, say λ_k and λ_m, $(k < m)$ if

$$\prod_{i<j} (\lambda_i - \lambda_j)^+$$

is TP_2 in λ_k and λ_m. However, by writing out the TP_2 criterion and cancelling like terms, it is straightforward to show this is the case.

To verify that $f_2(\lambda_2, \cdots, \lambda_p)$ is TP_2 in each pair of variables, we wish to use Theorem 5.1 (page 123) of Karlin (1968). This theorem states if a strictly positive function $f(x_1, x_2, x_3)$ is TP_2 in *each* pair of variables while the third is held fixed, and if the function $g(x_1, x_2)$ is TP_2, then for a σ-finite measure μ

$$h(x_2, x_3) = \int f(x_1, x_2, x_3) g(x_1, x_2)\, d\mu(x_1)$$

is TP_2 (the variables being defined over linear sets).

If we let $g(\lambda_1, \lambda_2) = (\lambda_1 - \lambda_2)^+$ and $f(\lambda_1, \lambda_2, \cdots, \lambda_p) = f_1(\lambda_1, \cdots, \lambda_p)/(\lambda_1 - \lambda_2)^+$, and integrate with respect to variable λ_1, then Theorem 5.1 of Karlin guarantees that $f_2(\lambda_2, \cdots, \lambda_p)$ is TP_2 in λ_i and λ_j, $j > i \geqq 2$.

By using this theorem again, this time with $g(\lambda_2, \lambda_3) = (\lambda_2 - \lambda_3)^+$ and $f(\lambda_2, \cdots, \lambda_p) = f_2(\lambda_2, \cdots, \lambda_p)/(\lambda_2 - \lambda_3)^+$, and integrating with respect to λ_2, we have that $f_3(\lambda_3, \cdots, \lambda_p)$ is TP_2 in each pair of variables.

Repeating this procedure as often as necessary gives the desired result.

4. Example. An easy but useful example of our result pertains to simultaneous confidence intervals. Several authors (for example, Hanumara and Thompson (1968), and Roy and Bose (1953)) develop methodology for simultaneous confidence intervals based upon the values l and u which are to be chosen so that

$$1 - 2\alpha = P(l \leqq \lambda_p < \lambda_1 \leqq u)\,. \tag{4.1}$$

Because of the difficulty in determining exact values for l and u, Hanumara and Thompson recommend that the conservative values (conservative because of the Bonferroni inequality) l_1 and u_1 be used where

$$\alpha = P(\lambda_p < l_1) = P(\lambda_1 > u_1)\,. \tag{4.2}$$

An idea of how conservative these values are may be obtained from the positive orthant dependence of the characteristic roots. That is,

$$\begin{aligned} P(l_1 \leqq \lambda_p < \lambda_1 \leqq u_1) &= 1 - P(\lambda_p < l_1) - P(\lambda_1 > u_1) + P(\lambda_p < l_1, \lambda_1 > u_1) \\ &\leqq 1 - 2\alpha + P(\lambda_p < l_1)P(\lambda_1 > u_1)\,. \end{aligned}$$

Thus

$$1 - 2\alpha < P(l_1 \leqq \lambda_p < \lambda_1 \leqq u_1) \leqq 1 - 2\alpha + \alpha^2\,,$$

so that if $\alpha = .025$, we may be assured that the true confidence coefficient is between .95 and .950625. Investigation seems to indicate that if $p = 2$, the true probability is near the middle of this interval. For larger values of p, the true probability tends to move towards the upper endpoint of the interval.

A similar type of argument is useful in trying to construct a critical region for tests involving the covariance matrix. Specifically, suppose one wishes to test $H_0 : \Sigma = \Sigma_0$ vs. $H_a : \Sigma \neq \Sigma_0$ where Σ is the covariance matrix of a p-variate normal distribution. Then if $n^{-1}S$ is the sample covariance matrix based on a sample of size $n + 1$, under H_0, $S\Sigma_0^{-1}$ has a Wishart distribution with $\Sigma = I$ and n df. For a 2α size test, Roy (1957) proposes as an acceptance region $l \leqq \lambda_p < \lambda_1 \leqq u$ where $\lambda_p < \cdots < \lambda_1$ are the roots of $S\Sigma_0^{-1}$ and l and u satisfy (4.1). Of course l and u are hard to obtain. If, however, we use l_1 and u_1 defined by (4.2), we have narrow bounds on the actual size of the test. (We could actually find even narrower bounds by putting unequal amounts of probability in each tail.)

5. Generalizations. Of course the exact same argument suffices for any set of random variables with a joint density of the form

$$\prod_{i=1}^{p} g_i(x_i) \cdot \prod_{i<j} (x_i - x_j)^+ .$$

In particular, it is worth noting that if $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_p \geqq 0$ denote the characteristic roots of either

$$S_1 S_2^{-1} \quad \text{or} \quad S_1(S_1 + S_2)^{-1}$$

where S_1, S_2 are independent $p \times p$ Wishart matrices with the same covariance matrix, the positive dependence relationships in Section 2 still hold.

Other interesting types of positive dependence are discussed by Shaked (1977). However, these depend upon the property of exchangeability which does not hold for the Wishart roots.

Acknowledgment. The authors wish to thank the associate editor for pointing out an error in the original proof.

REFERENCES

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

[2] DYKSTRA, R. L., HEWETT, JOHN E. and THOMPSON, W. A. (1973). Almost independent events. *Ann. Statist.* **1** 674–681.

[3] ESARY, J. D. and PROSCHAN, F. (1968). Generating associated random variables. *Boeing Scientific Research Labs* D1-82-0696.

[4] ESARY, J. D., PROSCHAN, F. and WALKUP, D. W. (1967). Association of random variables with applications. *Ann. Math. Statist.* **38** 1466–1474.

[5] HANUMARA, R. C. and THOMPSON, W. A. (1968). Percentage points of the extreme roots of a Wishart matrix. *Biometrika* **55** 505–512.

[6] KARLIN, SAMUEL (1968). *Total Positivity.* Stanford Univ. Press.

[7] LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1153.

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

[9] ROY, S. N. and BOSE, R. C. (1953). Simultaneous confidence interval estimation. *Ann. Math. Statist.* **24** 513–536.

[10] SHAKED, MOSHE (1977). A concept of positive dependence for exchangeable random variables. *Ann. Statist.* **5** 505–515.

[11] WAIKAR, V. B. (1975). Distributions of characteristic roots of random matrices. In *Statistical Distributions in Scientific Work* (G. P. Patil, S. Kotz and J. K. Ord, eds.). Reidel, Dordrecht.

DEPARTMENT OF STATISTICS
222 MATH. SCIENCES BUILDING
UNIVERSITY OF MISSOURI
COLUMBIA, MISSOURI