



Department of Economics
University of Southampton
Southampton SO17 1BJ
UK

**Discussion Papers in
Economics and Econometrics**

1999

This paper is available on our website
<http://www.soton.ac.uk/~econweb/dp/dp99.html>

The Density of a Quadratic Form in a Vector Uniformly Distributed on the n-Sphere*

by

Grant Hillier
email: ghh@soton.ac.uk

Department of Economics
University of Southampton
Southampton SO17 1BJ
U.K.

Discussion Paper No. 9902
Revised, July 1999

Abstract

There are many instances in the statistical literature in which inference is based on a normalized quadratic form in a standard normal vector, normalized by the squared length of that vector. Examples include both test statistics (the Durbin-Watson statistic), and estimators (serial correlation coefficients).

Although much studied, no general closed-form expression for the density function of such a statistic is known. This paper gives general formulae for the density in each open interval between the characteristic roots of the matrix involved. Results are given for the case of distinct roots, which need not be assumed positive, and when the roots occur with multiplicities greater than one. Starting from a representation of the density as a surface integral over an $(n - 2)$ -dimensional hyperplane, the density is expressed in terms of top-order zonal polynomials involving difference-quotients of the characteristic roots of the matrix in the numerator quadratic form

* I thank the Editor, Peter Phillips, the three referees, and Karim Abadir, Giovanni Forchini, and Bent Nielsen, for helpful comments that greatly improved the presentation of the paper.

1 Introduction

There are numerous problems in the statistical literature in which inference is based on a statistic that has the form of a normalised quadratic form in standard normal variates, i.e., a statistic of the form:

$$Q = y^0 B y = y^0 y; \quad (1)$$

with $y \gg N(0; I_n)$, and B an $n \times n$ real symmetric matrix that may or may not, depending on the context, be positive definite. Prominent examples of test statistics that have the form (1) are the Durbin-Watson statistic, and other (scale-invariant) diagnostic test statistics for the linear model, while examples of estimators that have this form are serial correlation coefficients - see Chapter 6 of Anderson (1971), for instance. The maximum likelihood estimator (MLE) for the coefficient of the lagged variable in a first-order autoregressive (AR(1)) model is also closely related to, but not quite identical to, Q in (1).

Despite its apparently simple form, and notwithstanding a huge literature on various special cases of Q that arise - in particular, the Durbin-Watson statistic and first-order serial correlation coefficient - the density function of Q remains unknown. For practical purposes this has not been a serious problem, because the moments of Q are relatively easy to derive (and calculate), as indeed is its characteristic function (see below), and either of these can be used to construct excellent approximations to the true density, or to tail area probabilities in the case of test statistics (see, for example, Henshaw (1966) for the Durbin-Watson case).

In fact, it is fair to say that two of the earliest contributors to the literature on the problem, von Neumann (1941) and Koopmans (1942), provided what are still today virtually all of the known key features of the exact density (as distinct from approximations to it, which have been developed considerably recently). In particular, the results that (a) Q is independent of its denominator, (b) the density is non-analytic at the characteristic roots of the matrix B , but analytic in the open intervals between those roots (see also Mulholland (1965), (1970), and Saldanha and Tomei (1996) for further analysis of this phenomenon), and (c) the form of the density is different in each of these open intervals, all appear in either or both of von Neumann (1941) and Koopmans (1942). In addition, von Neumann gives (recursive) expressions for the moments of Q , and evaluates some of the low order moments, and Anderson (1971), Corollary 6.7.3, gives an explicit formula for the density in the case where the roots of B each occur with multiplicity 2.

Nevertheless, no general closed-form expression for the density has so far been obtained, and it remains a puzzle that the problem should be so difficult for such an innocuous statistic. One of the standard routes to the density function - direct

inversion of the characteristic function - has been perfectly adequate in delivering approximations (usually asymptotic) to the density function, at least in special cases, but seems incapable of yielding a closed-form expression. Thus, it has become clear that a new approach is needed.

In this paper we derive a closed-form expression for the density of Q , for each of the $n_j - 1$ open intervals between the characteristic roots of B . We do not need to assume that B is positive definite, but will assume to begin with that its roots, λ_i , are distinct. The case in which roots of B occur with multiplicity greater than one will be discussed briefly in Section 5. Since it is known that the density is non-analytic at the roots, we exclude these points from consideration throughout. Indeed, it will become apparent that there are several steps in the derivations to follow that fail if Q is at a root λ_i , so these points must be excluded for the formulae that follow to be valid. It is plausible that some technical device could be invoked to produce, from the results to be given below, a single expression for the density valid over the entire domain of definition, but we do not explore that possibility here.

As indicated above, our starting point is the key, and is quite different from that adopted in virtually all previous work on the problem, although it has close connections with the approach taken by Anderson (1971) who focused on the distribution function of Q . Our point of departure will be a representation of the density as a surface integral over (essentially) the level set of Q . This representation of the density, by itself, produces many of its qualitative features as immediate consequences, and also shows precisely why the problem is difficult. The plan of the paper is as follows. In Section 2 we continue this preliminary discussion, giving some elementary properties of Q , and introducing some notation needed for the development of the main results. Section 3 gives, first, a brief introduction to the differential-geometric argument that leads to the surface integral representation of the density, then the integral representation itself, and finally some elementary but important properties of the density that follow directly from it. The main results for Q appear in Section 4, and in Section 5 we extend these to the case where the roots of B occur with multiplicity greater than one. All proofs and technical details are given in the Appendix.

The distribution assumptions on y declared in (1) mean that we are confining our attention to the "null" distributions of the various statistics that have the form (1). In the case of test statistics this is usually the density under the null hypothesis, and is, of course, of direct interest, but in the case of estimators, like serial correlation coefficients, it will yield the density only for specific (null) values of the model parameters. Thus, further work is needed to obtain the density of, for instance, the MLE in the AR(1) model, in full generality, but the results given in the present paper are a key step in that direction. Results for such "non-null" densities will be reported in a separate paper. Finally, it may be worth pointing out that, although our focus here is on exact distribution results, the results clearly also apply to statistics that, asymptotically, have the representation (1). That is, if $y \sim y_T$ is a function of T (the sample size, say), $y_T \xrightarrow{d} N(0; I_n)$; and $B_T \xrightarrow{p} B$; as $T \rightarrow \infty$; then the results to

follow yield the asymptotic distribution of $Q_T = y_T^0 B_T y_T = y_T^0 y_T$.

2 Some Preliminary Results And Notation

Because the normal density is spherically symmetric, i.e., the density is invariant under $y \rightarrow Hy$, for $H \in O(n)$, the group of $n \times n$ orthogonal matrices, it is clear that B in (1) may be assumed diagonal, with characteristic roots $\lambda_1, \dots, \lambda_n$ on the diagonal. Also, since the density is invariant under permutations of the elements of y , we may always assume that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.

Transforming y to $v = y(y^0 y)^{-1/2}$ and $r^2 = y^0 y$, the volume element (dy) (Lebesgue measure on R^n) transforms to: $(dy) = 2^{i-1} (r^2)^{\frac{n}{2}i-1} (dr^2) (v^0 dv)$, where $(v^0 dv)$ denotes (unnormalized) Haar measure on the surface of the unit sphere in R^n ; S_n (see Muirhead (1982), Chapter 2). It is straightforward to check that v and r^2 are independent, that $r^2 \sim \tilde{\Lambda}^2(n)$ and v is uniformly distributed on S_n , and that

$$Q = v^0 B v; \tag{2}$$

It follows from this that Q is independent of its denominator, r^2 , and also that these properties, and the density of Q itself, continue to hold under any spherically symmetric density for y , because any member of that family induces a uniform density on S_n for v . Hence, the results that follow hold more generally than under the Gaussian assumption $y \sim N(0; I_n)$.

Since $v \in S_n$ may be regarded as the first column of a random matrix $H \in O(n)$, uniformly distributed on $O(n)$, standard results on integration over the orthogonal group (see Muirhead (1982), Chapter 7; in particular Theorem 7.2.5, p. 243), together with the representation (2) for Q , yield the characteristic function of Q :

$$\begin{aligned} \tilde{A}_Q(t) &= E[\exp(itv^0 B v)] \\ &= {}_1F_1(1; 2; n; 2; itB); \end{aligned} \tag{3}$$

Here and in what follows we use the notation and definitions of hypergeometric functions of matrix argument first defined by Herz (1955) and explained in James (1964) and Muirhead (1982), Chapter 7. In general, a matrix-argument hypergeometric function can be expressed as an infinite series whose k_j th term is a linear combination of the zonal polynomials $C_\lambda(B)$ of the matrix B indexed by the (ordered) partitions, λ , of k with n or fewer parts. However, in (3), and in all of the results to be presented below, only top-order zonal polynomials (corresponding to the partition $\lambda = (k; 0; \dots; 0)$), which we denote by $C_{[k]}(B)$, are involved in the expansions, essentially because Q is one-dimensional. Thus, the results to be given below are very much simpler than the general hypergeometric functions that are prominent in much multivariate distribution theory. In what follows we shall need to exploit some of the properties of the top-order zonal polynomials. A brief introduction to these

polynomials is given in the Appendix. In particular, Lemma A.1 in the Appendix gives an explicit formula for the polynomials that is essential to our results.

The series expansion for the ${}_1F_1$ (confluent) hypergeometric function in (3) converges for all real t , and $A_Q(t)$ is infinitely differentiable at the origin, so that the moments of all orders of Q exist and are given by:

$$E[Q^k] = (1)_k C_{[k]}(B) / (n)_k \quad (4)$$

Here and throughout, $(c)_k$ denotes the Pochhammer symbol (forward factorial):

$$(c)_k = c(c+1)\cdots(c+k-1); (c)_0 = 1 \quad (5)$$

The computations on pp. 373–378 of von Neumann (1941) essentially give recursive formulae for the top-order zonal polynomials $C_{[k]}(\mathbb{C})$, well before their introduction, formal definition, and generalisation, by James (1961a); (1961b).

The density of a positive definite quadratic form in standard normal variables, $W = y^0 B y$, with $y \gg N(0; I_n)$, i.e., the numerator of Q , may also be expressed in terms of a confluent hypergeometric function with matrix argument, and was first given by James (1964) (equation (133)), in slightly different notation, in the form:

$$\text{pdf}_W(w) = [2^{n-2} \Gamma(n-2) |B|^{-\frac{1}{2}}]^{-1} {}_1F_1(1-2; n-2; \frac{1}{2} w B^{-1}) \quad (6)$$

Note that here and later we denote the random variable of interest by an upper-case letter, and a particular value of that variable by its lower-case counterpart.

As we shall see, the density of Q may also be expressed in terms of hypergeometric functions of matrix argument, and closely related functions, but the functions involved turn out to be related to the Gaussian ${}_2F_1$ function, rather than the confluent ${}_1F_1$ function.

Notes on notation: (1) we work throughout with vectors, x , say, whose elements are positive, and indicate this simply by $x > 0$; (2) we repeatedly transform variables of integration (usually by simple scale changes), but to avoid a plethora of new symbols (and because these are just variables of integration), we often retain the same symbols for the new variables as were used for the old.

3 Surface Integral Representation of the Density

The results given in the previous section rely on the representation (2) for Q as a quadratic form in a vector, v , that is uniformly distributed on the unit sphere S_n , and the characteristic function (3) is obtained by integrating over the surface $S_n = \{v; v \in \mathbb{R}^n; v^0 v = 1\}$, an $(n-1)$ -dimensional manifold embedded in \mathbb{R}^n . Our point of departure will be to express the density of Q as an integral over an $(n-2)$ -dimensional manifold, so we begin this section with a brief introduction to this differential-geometric representation of a density. Full technical details, and the

statement of the main result that we rely on, may be found in Tjur (1980), especially Chapter 8. Before doing so we introduce a different representation for Q that will be the basis of our approach.

In the normal case, and hence also in the case of any spherically symmetric density for y , it is clear that Q can be written in the form:

$$Q = (S_{i=1}^n x_i) = (S_{i=1}^n x_i); \quad (7)$$

where $x_i \sim \sqrt{y_i^2}$; $i = 1, \dots, n$ are independent $\hat{A}^2(1)$ random variables, and $\lambda_1, \lambda_2, \dots, \lambda_n$ are the (ordered) characteristic roots of B . We do not insist that the roots λ_i are positive, and will state the main results of the paper under the assumption that these roots are distinct. In Section 5 we briefly indicate the modifications needed when the roots of B occur with multiplicities greater than one. Since, from (7), Q is a convex combination of the λ_i , it is clear that $\lambda_i > 0$.

Now, in general, let the underlying vector of random variables be $x \in X \subset \mathbb{R}^n$, and consider a (possibly vector-valued) statistic $T : X \rightarrow Y \subset \mathbb{R}^p$, with $p \leq n$. Assume that T is continuously differentiable, and that the $p \times n$ matrix $DT(x) = \partial T_i(x) / \partial x_j$; $i = 1, \dots, p$; $j = 1, \dots, n$, has full rank p whenever $T(x) = t$, where t is a particular value of T . Then, the t_j level set of T ; $M(t) = \{x \in X; T(x) = t\}$, is an $(n - p)$ -dimensional (differentiable) manifold embedded in \mathbb{R}^n . This means that, in the neighbourhood of each point $x \in M(t)$, there is a one-to-one differentiable function $f : W \rightarrow \mathbb{R}^n$, with W an open subset of \mathbb{R}^{n-p} , such that $f(W) = M(t) \cap \mathbb{R}^n$, so that points $y \in W$ provide local coordinates for points in $M(t)$ near x (and f a local coordinate chart for $M(t)$ near x). Because $M(t)$ can be covered by a system of overlapping local coordinate charts, one can define, in a canonical way, a volume element, $(dM(t))$, everywhere on $M(t)$, and, accordingly, integrate functions defined on $M(t)$. Given a local coordinate chart f , and defining $Df(y)$ as $DT(x)$ has been defined above, we have:

$$(dM(t)) = |J(Df(y))|^{1/2} (dy); \quad (8)$$

where (dy) denotes ordinary Lebesgue measure on \mathbb{R}^{n-p} , and $|J|$ denotes the determinant of the indicated matrix. It is straightforward to confirm that the volume element thus defined does not depend on the system of coordinate charts used for $M(t)$. Using these ideas, Tjur (1980), Proposition 8.1.2, shows that the density of T at the point t is given by the surface integral:

$$\text{pdf}_T(t) = \int_{M(t)} |J(DT(x))|^{1/2} \text{pdf}(x) (dM(t)); \quad (9)$$

where $\text{pdf}(x)$ denotes the density of the underlying vector x .

In principle this expression for the density could be applied directly to Q in (7), with $x = (x_1, \dots, x_n)$ defined on the non-negative orthant in \mathbb{R}^n . However, it turns out that it is slightly simpler to work first with a bivariate statistic consisting of the

numerator and denominator of Q in (7) separately, and this is the route we shall take. Thus, consider first the statistic

$$T(x) = (S_{i=1}^n x_i; S_{i=1}^n x_i)^0 = (T_1; T_2)^0; \quad (10)$$

so that $Q = T_1/T_2$. The level set of T ,

$$M = M(t_1; t_2) = \{x; x > 0; T_1 = t_1; T_2 = t_2\};$$

where t_1 and t_2 are fixed values of T_1 and T_2 (satisfying $t_1 > 0, t_2 > 0$), is evidently the intersection of an $(n-2)$ -dimensional hyperplane with the non-negative orthant in R^n , and it is straightforward to check that $DT(x)$ is of rank 2 for all x . The level set of $T; M$, is therefore an $(n-2)$ -dimensional manifold. The joint density of $(T_1; T_2)$ at $(t_1; t_2)$ is given by equation (9) above, with

$$j(DT(x))(DT(x))^0 j = [n S_{i=1}^{n-2} (S_{i=1}^n)^2]; \quad (11)$$

and

$$\text{pdf}(x) = (2^i)^{-\frac{n}{2}} [\prod_{i=1}^n x_i]^{-\frac{1}{2}} \exp\{-\sum_{i=1}^n x_i\}; \quad (12)$$

so that (9) becomes, for this case:

$$\begin{aligned} \text{pdf}_T(t_1; t_2) &= (2^i)^{-\frac{n}{2}} [n S_{i=1}^{n-2} (S_{i=1}^n)^2]^{-\frac{1}{2}} \exp\{-\sum_{i=1}^n t_i\} \\ &\int_M [\prod_{i=1}^n x_i]^{-\frac{1}{2}} (dM); \end{aligned} \quad (13)$$

The results that follow can easily be modified to yield the joint density of $(T_1; T_2)$ directly from (13), but since Q and its denominator are independent, it is simpler to first obtain the density of Q , then multiply by $\text{pdf}_{T_2}(t_2)$ and transform $(Q; T_2)$ to $(T_1 = QT_2; T_2)$ to obtain the joint density. Thus, we shall focus here on the density of Q .

Transforming variables in the integral in (13) to $x_i = t_2 q_i; i = 1; \dots; n$, the manifold M is mapped into a new manifold M' , say, and it is easy to see (using (8)) that $(dM) = t_2^{n-2} (dM')$, so that (13) becomes

$$\begin{aligned} \text{pdf}_T(t_1; t_2) &= (2^i)^{-\frac{n}{2}} [n S_{i=1}^{n-2} (S_{i=1}^n)^2]^{-\frac{1}{2}} \exp\{-\sum_{i=1}^n t_2 q_i\} t_2^{\frac{n}{2}-2} \\ &\int_{M'} [\prod_{i=1}^n x_i]^{-\frac{1}{2}} (dM'); \end{aligned} \quad (14)$$

where M' is now the manifold $\{x; x > 0; \sum_{i=1}^n x_i = t_1/t_2; \sum_{i=1}^n x_i = 1\}$. But, transforming now from $(T_1; T_2)$ to $(Q; T_2)$, the integral in (14) no longer depends on t_2 , so we see again that Q and T_2 are independent and $T_2 \gg \hat{A}^2(n)$. Integrating out t_2 , we obtain a surface integral formula for the density of Q at the point $Q = q$:

$$\text{pdf}_Q(q) = \frac{1}{2} [n S_{i=1}^{n-2} (S_{i=1}^n)^2]^{-\frac{1}{2}}$$

$$\int_S \left[\sum_{i=1}^n x_i \right]^{i-1} dS; \quad \sum_{i=1}^n x_i < q < \sum_{i=1}^n 1; \quad (15)$$

with S now the manifold $S = \{x; x_i > 0; \sum_{i=1}^n x_i = q; \sum_{i=1}^n 1 = 1\}$. Note that the density vanishes outside $[\sum_{i=1}^n x_i]$ because, for q outside this interval, the hyperplanes $\sum_{i=1}^n x_i = q$ and $\sum_{i=1}^n 1 = 1$ do not intersect in the non-negative orthant.

In view of (15), the problem has been reduced to the evaluation of the surface integral in the last line of that equation. The first step in this process is to choose coordinates for S , and since S is flat, this can be done globally (rather than only locally as in the general case discussed above). For this purpose it is convenient to use $x_2; \dots; x_{n-1}$ as coordinates, setting

$$x_1 = (\sum_{i=2}^n x_i)^{i-1} [q - \sum_{i=2}^n x_i];$$

$x_i = x_i$ for $i = 2; \dots; n-1$; and

$$x_n = (\sum_{i=2}^n x_i)^{i-1} [\sum_{i=2}^n x_i];$$

Here and henceforth we use the notation S_r to indicate a sum $\sum_{i=r}^n 1$, with a similar abbreviation for products. With these coordinates we find from (8) that:

$$dS = (\sum_{i=2}^n x_i)^{i-1} [n \sum_{i=2}^n x_i^2 - (\sum_{i=2}^n x_i)^2]^{1/2} dx_2 \dots dx_{n-1}; \quad (16)$$

and $x_2; \dots; x_{n-1}$ are (because of the non-negativity of x_1 and x_n) constrained to lie in the region

$$R = \{x; x_i > 0; \sum_{i=2}^n x_i < q - \sum_{i=2}^n x_i; \sum_{i=2}^n x_i < \sum_{i=2}^n 1\};$$

Define the new variable

$$F = (Q - \sum_{i=2}^n x_i) = (\sum_{i=2}^n 1 - F); \quad F > 0; \quad (17)$$

and the sequence of constants

$$\tilde{A}_i = (\sum_{i=2}^n 1) = (\sum_{i=2}^n 1); \quad i = 2; \dots; n-1; \quad (18)$$

so that $\tilde{A}_2 > \dots > \tilde{A}_{n-1} > 0$: Using the coordinates $(x_2; \dots; x_{n-1})$ for S , and transforming $Q \rightarrow F$; (the Jacobian is $(\sum_{i=2}^n 1)(1+F)^{i-1}$); (15) becomes

$$\text{pdf}_F(f) = \frac{1}{2} f^{i-1} (n-2)(1+f)^{i-1}$$

\int_R

$$\int_R (f - (1+f) \sum_{i=2}^n x_i)^{i-1} (1 - (1+f) \sum_{i=2}^n x_i)^{i-1} \left[\sum_{i=2}^n x_i \right]^{i-1} dx_2 \dots dx_{n-1}; \quad (19)$$

where we have temporarily put $\tilde{A}_i = (\sum_{i=2}^n 1) = (\sum_{i=2}^n 1); i = 2; \dots; n-1$; and the region of integration is defined by

$$R = \{x; x_i > 0; \sum_{i=2}^n x_i < f(1+f); \sum_{i=2}^n (1 - \tilde{A}_i) x_i < 1 = (1+f)g\};$$

To simplify the notation in evaluating the integral in (19), we now transform to new variables of integration $z_i = (1 + f)^{-i} x_{i+1} = f^{-i} x_{i+1}$; $i = 1; \dots; n-1$; and also define a new sequence of constants (setting, from now on, $m = n-1$)

$$\begin{aligned} a_i &= f^{-i} x_{i+1} = f^{-i} \tilde{A}_{i+1} \\ &= f^{-i} x_{i+1} = f^{-i} \tilde{A}_{i+1}; \quad i = 1; \dots; m; \end{aligned} \quad (20)$$

Note that $0 < a_1 < a_2 < \dots < a_m$, and that:

$$a_i > 1 \text{ or } a_i < 1 \text{ as } f > \tilde{A}_{i+1} \text{ or } f < \tilde{A}_{i+1}; \quad (21)$$

With this notation the density (19) becomes

$$\begin{aligned} \text{pdf}_F(f) &= \frac{1}{2} f^{\frac{n}{2}-1} (n-2)! \left[\prod_{i=1}^m (\tilde{A}_{i+1} (1 + \tilde{A}_i)) \right]^{-1} f^{\frac{m-1}{2}} (1 + f)^{-\frac{m+2}{2}} \\ &\int_R \prod_{i=1}^m (1 - S_1 z_i)^{-\frac{1}{2}} (1 - S_1 a_i z_i)^{-\frac{1}{2}} \left[\prod_{i=1}^m z_i \right]^{-\frac{1}{2}} (dz_1 \dots dz_m); \end{aligned} \quad (22)$$

with $R = \{z > 0; S_1 z_i < 1; S_1 a_i z_i < 1\}$ (the sums running from $i = 1$ to m): Thus the region of integration becomes a polyhedral region in R^m bounded by the coordinate axes and the two hyperplanes $S_1 z_i = 1$ and $S_1 a_i z_i = 1$:

Some elementary properties of the density $\text{pdf}_F(f)$ can be deduced directly from the form of the integral in (22). First, it is clear that the integral is a symmetric function of $a_1; \dots; a_m$ (that is, is invariant under permutations of the a_i 's), because on permuting the a_i 's we can then apply the inverse permutation to the z_i 's, leaving the integral unchanged. More importantly, the 'shape' of the region of integration R depends on the number of the a_i that are greater than one, and hence on where f lies in relation to the sequence of constants \tilde{A}_i . For instance, if $m = 2$ ($n = 4$), the lines $z_1 + z_2 = 1$ and $a_1 z_1 + a_2 z_2 = 1$ (bounding R in (22)) do not cross in the non-negative orthant if either a_1 and a_2 are both less than, or both greater than, one, but do cross there if $a_1 < 1$ and $a_2 > 1$:

Thus, the fact that the density of Q has a different form in each interval between the roots of B follows directly from (22). Before considering the evaluation of the integral in (22), we first establish an important simple property of the integral (and hence the density of F) - a kind of symmetry that implies that only half of the relevant intervals need be dealt with separately (cf. also the discussion on pp. 389-390 in von Neumann (1941)).

For each $r = 0; \dots; m$; define

$$I_r(a_1; \dots; a_m) = \int_R \prod_{i=1}^m (1 - S_1 z_i)^{-\frac{1}{2}} (1 - S_1 a_i z_i)^{-\frac{1}{2}} \left[\prod_{i=1}^m z_i \right]^{-\frac{1}{2}} (dz_1 \dots dz_m); \quad (23)$$

where $R = \{z > 0; S_1 z_i < 1; S_1 a_i z_i < 1\}$; and the subscript r on I_r indicates that $0 < a_i < 1$ for $i = 1; \dots; r$; $a_i > 1$ for $i = r+1; \dots; m$: We have:

Proposition 1 : For $r = 0; \dots; m$,

$$I_r(a_1; \dots; a_m) = [\prod_{i=1}^m a_i]^{i-\frac{1}{2}} I_{m_i-r}(a_1^{-1}; \dots; a_m^{-1}); \quad (24)$$

Since I_r yields the density of F when $\tilde{A}_{r+2} < f < \tilde{A}_{r+1}$, and I_{m_i-r} the density when $\tilde{A}_{m_i-r+2} < f < \tilde{A}_{m_i-r+1}$; (24) means that we need only evaluate the integral in (22) for $r = m-2$ if m is even, and $r = (m-1)-2$ if m is odd. The proof of Proposition 1 is given in the Appendix; the result is a simple consequence of the observation that the transformation $z_i \rightarrow 1/z_i = a_i z_i$ leaves the form of the integral in (23) unchanged, but replaces each a_i by a_i^{-1} :

4 Main Results

4.1 The extreme intervals $f < \tilde{A}_{n_i-1}$ and $f > \tilde{A}_2$.

We now evaluate I_m ; i.e., the integral in (23) when all m of the a_i are less than one. For the case $m = 1$ the result is standard:

$$\int_0^1 z^{i-\frac{1}{2}} (1-z)^{i-\frac{1}{2}} (1-az)^{i-\frac{1}{2}} dz = \frac{1}{4} {}_2F_1(1=2; 1=2; 1; a); \quad (25)$$

where

$${}_2F_1(a; b; c; x) = \sum_{k=0}^{\infty} \frac{[(a)_k (b)_k = (c)_k] [x^k = k!]}{j \times j < 1}; \quad (26)$$

denotes the Gaussian hypergeometric function. We now give a generalisation of this result to the m_i dimensional case, as in (23):

Lemma 2 : Assume that $0 < a_i < 1; i = 1; \dots; m$, and let $A = \text{diag}(a_1; \dots; a_m)$. Then:

$$I_m(a_1; \dots; a_m) = [\prod_{i=1}^m a_i]^{\frac{m+1}{2}} {}_2F_1(1=2; 1=2; (m+1)=2; A); \quad (27)$$

where $I_m(a_1; \dots; a_m)$ denotes the integral in equation (23).

Because the numerator parameters are $1/2$, only the top-order zonal polynomials $C_{[k]}(A)$ appear in the series expansion of the hypergeometric function. Thus, the expansion for the ${}_2F_1$ function in (27) is, in this case:

$${}_2F_1(1=2; 1=2; (m+1)=2; A) = \sum_{k=0}^{\infty} \frac{[(1=2)_k (1=2)_k = k! ((m+1)=2)_k] C_{[k]}(A)}{k!}; \quad (28)$$

The detailed proof of Lemma 2 appears in the Appendix, but it will be helpful here to indicate the method of proof, to prepare for the case where some of the a_i are greater than one.

Outline Proof of Lemma 1:

To integrate over the region $fz > 0; \sum_{i=1}^m z_i < 1$ it is natural to first transform from z_1 to $b_1 = z_1/(1 - \sum_{i=2}^m z_i)$, so that $0 < b_1 < 1$; then from z_2 to $b_2 = z_2/(1 - \sum_{i=3}^m z_i)$, so that $0 < b_2 < 1$; and so on sequentially. This sequence of transformations defines a (non-linear) transformation $z \rightarrow b = (b_1; \dots; b_m)$ that maps the region $\sum_{i=1}^m z_i < 1$ onto the open unit m -cube:

$$C_m = \{b_i; i = 1; \dots; m; 0 < b_i < 1\}$$

Now, when $a_i < 1$ for all $i = 1; \dots; m$, it is easy to see that C_m is a proper subset of the region corresponding to $\sum_{i=1}^m a_i z_i < 1$ in the coordinates $(b_1; \dots; b_m)$, so that the integral I_m in (23) is over all of C_m . The result in Lemma 2 then follows by (tediously) expanding the integrand in (23) (expressed in terms of the b_i) and integrating over C_m , followed by the use of Lemma A.1 in the Appendix.

¥

In the case where some of the a_i are greater than one, C_m is no longer a subset of the region corresponding to $\sum_{i=1}^m a_i z_i < 1$; and the restrictions that this second inequality imposes on $(b_1; \dots; b_m)$ need to be taken account of.

In view of Lemma 2, equation (22), and Proposition 1 we can state:

Theorem 3 : Let $D = \text{diag} \tilde{A}_2^{i_1-1}; \dots; \tilde{A}_{n_i-1}^{i_1-1} g$:

For $0 < f < \tilde{A}_{n_i-1}$,

$$\text{pdf}_F(f) = [B(1=2; (m+1)=2)]^{i_1-1} [{}_{i_1-2}(\tilde{A}_i=(1+\tilde{A}_i))]^{i_1-1} f^{\frac{m_i-1}{2}} (1+f)^{i_1 \frac{m+2}{2}}$$

$$\in {}_2F_1(1=2; 1=2; (m+1)=2; fD); \quad (29)$$

For $f > \tilde{A}_2$,

$$\text{pdf}_F(f) = [B(1=2; (m+1)=2)]^{i_1-1} [{}_{i_1-2}(1+\tilde{A}_i)]^{\frac{1}{2}} f^{i_1 \frac{1}{2}} (1+f)^{i_1 \frac{m+2}{2}}$$

$$\in {}_2F_1(1=2; 1=2; (m+1)=2; (fD)^{i_1-1}); \quad (30)$$

Here, as usual, $B(a; c)$ denotes the Beta coefficient $\frac{\Gamma(a)\Gamma(c)}{\Gamma(a+c)}$.

Remarks:

1. The series in (29) and (30) are power series in f and f^{i_1-1} respectively, with coefficients that are symmetric functions of the \tilde{A}_i , as anticipated.
2. The densities of F in these two intervals are obviously relatives of the $F(m+1; 1)$ and $F(1; m+1)$ densities, respectively. The density of F on the interval $f > \tilde{A}_2$ is obtained from that of F on the interval $f < \tilde{A}_{n_i-1}$ by transforming to F^{i_1-1} , and replacing $(\tilde{A}_2; \dots; \tilde{A}_{n_i-1})$ by $(\tilde{A}_2^{i_1-1}; \dots; \tilde{A}_{n_i-1}^{i_1-1})$. This property of the density will shortly be seen to hold generally (Corollary 1 below).
3. The series in (29) and (30) converge for all f in the indicated interval, but diverge otherwise. Thus, in particular, these expressions are undefined for F at \tilde{A}_{n_i-1} or \tilde{A}_2 .

4.2 The Intermediate Intervals $\tilde{A}_{r+2} < f < \tilde{A}_{r+1}$

Consider now the m_i-1 cases where, for some $r \in \{1, \dots, m_i-1\}$; $0 < a_i < 1$ for $i \leq r$; $a_i > 1$ for $i \geq r+1$; so that $\tilde{A}_{r+2} < f < \tilde{A}_{r+1}$ (or $\tilde{q}_{r+2} < q < \tilde{q}_{r+1}$). As we have seen, in each of these cases the hyperplanes $\sum z_i = 1$ and $\sum a_i z_i = 1$ that (together with the coordinate axes) form the boundary of the region of integration in (22), intersect, and the region changes shape depending on the value of r . For any fixed r it is possible, in principle, to evaluate the required integral in equation (23) by using Fourier-Motzkin elimination to decompose the region into disjoint components, and then integrate over each component separately (see Schechter (1998)). This approach, however - even in the case, as here, where there are only two bounding hyperplanes - does not yield analytically tractable expressions. Thus, instead, we work with the coordinates introduced in the outline proof of Lemma 1 above.

The generalisation of Lemma 1 to the case where some of the a_i are less than, and some greater than, one, is provided by:

Lemma 4 : Let

$$I_r(a_1; \dots; a_m) = \int_R (1 - \sum_{i=1}^r z_i)^{i-1} (1 - \sum_{i=1}^r a_i z_i)^{i-1} \prod_{i=1}^r [z_i]^{i-1} (dz_1 \dots dz_m); \quad (31)$$

with $R = \{z; z_i > 0; \sum_{i=1}^r z_i < 1; \sum_{i=1}^r a_i z_i < 1\}$, and assume that $0 < a_i < 1$ for $i \leq r$; $a_i > 1$ for $i > r+1$: Let $A_r = \text{diag}\{a_1; \dots; a_r\}$, and let $A_{r+1} = \text{diag}\{a_{r+1}; \dots; a_m\}$. Then:

$$I_r = \int_{\mathcal{C}_m} \prod_{i=1}^m (1 - \sum_{j=1}^i z_j)^{i-1} \prod_{i=1}^m [z_i]^{i-1} \prod_{j:k=0}^{\infty} c_r(j; k) [(1-2)_j (1-2)_{k-j} / k!] C_{[j]}(A_r) C_{[k]}(A_{r+1}^{-1}); \quad (32)$$

Whenever non-vanishing, the coefficients $c_r(j; k)$ are given by:

$$\begin{aligned} c_r(j; k) &= 1 \quad \text{if } j = k; \\ &= \binom{p}{k-j} \binom{q}{j} \binom{q-1}{k-j} \quad \text{if } j < k; \\ &= \binom{q}{j-k} \binom{p}{j} \binom{p-1}{j-k} \quad \text{if } j > k; \end{aligned} \quad (33)$$

where $p = (r-1)$ and $q = (m-r)$: If either $q = p$ is an integer, or $p = q$ is an integer, or both,

$$c_r(j; k) = 0 \text{ if } j > k + q \text{ and/or } j < k - p. \quad (34)$$

The proof of Lemma 2, including the detailed properties of the coefficients $c_r(j; k)$, is given in the Appendix. As before, we merely outline the method of proof here.

Outline proof of Lemma 2:

In the coordinates $(b_1; \dots; b_m)$ introduced in the outline proof of Lemma 1, the region of interest $R = \{z; z_i > 0; \sum_{i=1}^r z_i < 1; \sum_{i=1}^r a_i z_i < 1\}$ is, when some of the a_i are greater than one, a subset of C_m . To see this, consider the recursive expression for $d_m = \sum_{i=1}^r a_i z_i$ given in equation (A.14) in the Appendix:

$$d_m = a_m b_m + (1 - b_m) d_{m-1}; \quad (35)$$

Since this expresses d_m as a convex combination of a_m , which we now assume is greater than one, and d_{m-1} ; a necessary condition for $d_m < 1$ is clearly that $d_{m-1} < 1$; and this, together with the condition:

$$b_m < u_m = (1 - d_{m-1}) / (a_m - d_{m-1}) < 1; \quad (36)$$

are necessary and sufficient to ensure that $d_m < 1$: That is, in terms of the coordinates $(b_1; \dots; b_m)$, the region R becomes:

$$R = C_m \setminus \{0 < b_m < u_m; d_{m-1} < 1\}; \quad (37)$$

If $a_i < 1$ for $i = m - j - 1$ (i.e., $r = m - j - 1$), the condition $d_{m-j-1} < 1$ is automatically satisfied, but if not this argument can be repeated for b_{m-j-1} , and so on sequentially for $b_{m-j-2}; \dots; b_{r+1}$, until the remaining condition, $d_r < 1$; is automatically satisfied. By a suitable sequential set of transformations of the variables $(b_m; b_{m-1}; \dots; b_{r+1})$ we can therefore, for each choice of r , map the region of integration in the integral in (23) onto $C_m = C_{m-j-r} \in C_r$, and the integral is then relatively easily evaluated. The full details of this process are relegated to the Appendix. \neq

The various cases relevant to the vanishing of the coefficients $c_r(j; k)$ in Lemma 2 are summarised in Table 1.

TABLE 1

Properties of the numerical coefficients in Lemma 2

	$\textcircled{+} = (r - j - 1) = 2$	$\textcircled{-} = (m - j - r - j - 1) = 2$
m, r both even	non-integer	non-integer
m even, r odd	integer = p	integer = q
m odd, r even	non-integer	integer = q
m, r both odd	integer = p	non-integer

$$c_r(j; k) = 0 \text{ if } j > k + q \text{ or } j < k - p:$$

Applying Lemma 2 in equation (22) for the density of f we obtain:

Theorem 5 : For each $r = 1; \dots; m - j - 1$; for f in the interval $\tilde{A}_{r+2} < f < \tilde{A}_{r+1}$,

$$\text{pdf}_F(f) = [B((r - j - 1) = 2; (m - j - r - j - 1) = 2)]^{-1} \left[\prod_{i=2}^{r+1} \tilde{A}_i \right]^{i-1} \left[\prod_{i=2}^{m+1} (1 + \tilde{A}_i) \right]^{-\frac{1}{2}} \\ \int_0^f f^{\frac{r-1}{2}} (1+f)^{i-\frac{m+2}{2}} \sum_{j,k=0}^{\infty} [(1=2)_j (1=2)_k = j! k!] c_r(j; k) f^{j+k} C_{[j]}(D_r) C_{[k]}(D_{r+1}^{-1}); \quad (38)$$

where $D_r = \text{diag} \tilde{A}_2^{-1}; \dots; \tilde{A}_{r+1}^{-1} \mathbf{g}$; $D_{r+1} = \text{diag} \tilde{A}_{r+2}^{-1}; \dots; \tilde{A}_{m+1}^{-1} \mathbf{g}$, and the numerical coefficients $c_r(j; k)$ are as described in Lemma 2.

Remarks

1. It is straightforward to check that in fact Theorem 2 holds for $r = 0$ and $r = m$: in each of these cases one of the summations in (38) terminates at zero, and the numerical coefficients in (33) reduce to those given in Theorem 1 (with $(\textcircled{+}; \textcircled{-}) = (j - 1 = 2; (m - j - 1) = 2)$ in the first case, $(\textcircled{+}; \textcircled{-}) = ((m - j - 1) = 2; j - 1 = 2)$ in the second). Thus, Theorem 2 subsumes Theorem 1.
2. When $\textcircled{+}$ and $\textcircled{-}$ are both integers (i.e., m is even and r is odd), the coefficients vanish outside the strip of $(j; k)$ values $k - j \leq k + q$, so that the power of f in

the series expansion in (38), $j \leq k$, remains in the interval $[j, p; q]$. This agrees with the results in von Neumann (1941).

3. For fixed r the density of F is evidently a relative of the $F(r+1; m; r+1)$ density, varying, therefore, with r . Moreover, the generalisation of the "inverse symmetry" mentioned in remark 2 following Theorem 1 is clear from Theorem 2:

Corollary 6 : For $r = 0; \dots; m$, let $\text{pdf}_F^{(r)}(f; \tilde{A}_2; \dots; \tilde{A}_{n_i-1})$ denote the density of F in the interval $\tilde{A}_{r+2} < f < \tilde{A}_{r+1}$, let $\tilde{F} = F^{-1}$ (and $\tilde{f} = f^{-1}$). Then:

$$\text{pdf}_F^{(r)}(f; \tilde{A}_2; \dots; \tilde{A}_{n_i-1}) = \text{pdf}_{\tilde{F}}^{(m-r+1)}(\tilde{f}; \tilde{A}_2^{-1}; \dots; \tilde{A}_{n_i-1}^{-1}); \quad (39)$$

the latter being defined on the interval $\tilde{A}_{m-r+1}^{-1} < \tilde{f} < \tilde{A}_{m-r+2}^{-1}$:

Corollary 1 is implied by Proposition 1 and equation (22), but is manifest more explicitly in Theorem 2. Evidently, the density of F possesses a kind of symmetry analogous to the "obvious" property of the $F(\alpha_1; \alpha_2)$ density, that $F^{-1}(\alpha_1; \alpha_2) \sim F(\alpha_2; \alpha_1)$, except that here the role of the parameters \tilde{A}_i is also "inverted".

To obtain the density of Q itself we simply need to transform $F \rightarrow Q$, (Jacobian $(\frac{\partial x_1 \dots \partial x_n}{\partial q}) = (\frac{\partial x_1 \dots \partial x_n}{\partial q})^2$), and set $f = (q - x_n) = (\frac{\partial x_1 \dots \partial x_n}{\partial q})$ in Theorem 2. For completeness we state this result as:

Corollary 7 : With the assumptions and notation given in Theorem 2, for each $r = 0; \dots; m$; and q in the interval $x_{r+2} < q < x_{r+1}$:

$$\begin{aligned} \text{pdf}_Q(q) &= [B((r+1)=2; (m-r+1)=2)]^{i-1} [\prod_{i=2}^{r+1} \tilde{A}_i]^{i-\frac{1}{2}} [\prod_{i=2}^{m+1} (1 + \tilde{A}_i)]^{\frac{1}{2}} \\ &\quad \times \int_{j;k=0}^{\infty} [(1=2)_j (1=2)_{k=j+k}] c_r(j; k) [(q - x_n) = (\frac{\partial x_1 \dots \partial x_n}{\partial q})]^{j+k} C_{[j]}(D_r) C_{[k]}(D_{r+1}^{-1}); \quad (40) \end{aligned}$$

As indicated in Section 2, given the density of Q , it is straightforward to obtain the joint density of T_1 and T_2 from Theorem 2, because T_2 is independent of Q .

5 Multiplicities of the Roots

We now consider (briefly) the case in which the roots of B are not distinct. Suppose that there are s distinct roots $x_1 > x_2 > \dots > x_s$, and that x_i occurs with multiplicity n_i ; so that $\sum_{i=1}^s n_i = n$. As in (7), we can write:

$$Q = (S_{i=1}^s x_i) = (S_{i=1}^s x_i); \quad (41)$$

but the x_i are now independent $\hat{A}^2(n_i)$ variates, $i = 1; \dots; s$: This problem is of interest both in the context of our original problem, and because the numerator of Q in (41)

is simply a linear combination of independent $\tilde{A}^2(n_i)$ random variables, a form that also arises frequently.

The argument in Section 3 leading to equation (22) is readily adapted to this case, giving the following analogue of that equation:

$$\text{pdf}_F(f) = \prod_{i=1}^m (n_i=2) [2^{\frac{n_i}{2}} \prod_{j=1}^{n_i} (n_j=2)]^{-1} \left[\prod_{i=2}^m (\tilde{A}_i = (1 + \tilde{A}_i))^{n_i} \right] f^{\frac{n_1 - n_{s_1} - 2}{2}} (1 + f)^{n_1/2} \int_{\mathbf{R}} \prod_{i=1}^m (1 - \sum_{j=1}^m z_j)^{\frac{n_i}{2} - 1} (1 - \sum_{j=1}^m a_j z_j)^{\frac{n_{s_i}}{2} - 1} \left[\prod_{i=1}^m z_i^{\frac{n_i}{2} - 1} \right] (dz_1 \dots dz_m); \quad (42)$$

with, now, $m = s_j - 2$; $\tilde{A}_i = (\dots, i, \dots) = (\dots, 1, \dots)$; $i = 2, \dots, s_j - 1$; $a_i = f = \tilde{A}_{i+1}$; $i = 1, \dots, m$, and $\mathbf{R} = \{z; z > 0; \sum_{i=1}^m z_i < 1; \sum_{i=1}^m a_i z_i < 1\}$.

The integral in the second line of (42) can be evaluated by exactly the same procedures as we have applied to the integral in (23). For example, the generalised version of Lemma 1 becomes:

Lemma 8 : Let $m = s_j - 2$; $\sum_{i=1}^s n_i = n$; $\mathbf{A} = \text{diag}\{a_1 I_{n_2}, \dots, a_m I_{n_{s_j-1}}, \mathbf{g}\}$; with $a_i < 1$ $\forall i = 1, \dots, m$;

$$\mathbf{R} = \{z; z > 0; \sum_{i=1}^m z_i < 1\};$$

and

$$I_m(a_1, \dots, a_m) = \int_{\mathbf{R}} \prod_{i=1}^m (1 - \sum_{j=1}^m z_j)^{\frac{n_i}{2} - 1} (1 - \sum_{j=1}^m a_j z_j)^{\frac{n_{s_i}}{2} - 1} \left[\prod_{i=1}^m z_i^{\frac{n_i}{2} - 1} \right] (dz_1 \dots dz_m); \quad (43)$$

Then:

$$I_m = \left[\prod_{i=1}^{s_j-1} (n_i=2) \right]^{-1} ((n_j - n_s)=2) {}_2F_1(1=2; (n_s - j - 2)=2; (n_j - n_s)=2; \mathbf{A}); \quad (44)$$

The proof of Lemma 3 is omitted, but it relies essentially just on Lemma A.3 in the Appendix, which provides an identity for the top-order zonal polynomials of matrices having the structure of the matrix \mathbf{A} given in the Lemma. The analogue of Proposition 1 obviously provides the corresponding version of Lemma 3 when the a_i are all greater than one. For the case where some of the a_i are greater than, and some less than, one, the following generalisation of Lemma 2 can be deduced by following the procedure given in the Appendix for the case of distinct roots:

Lemma 9 : Let $m = s_j - 2$; $\sum_{i=1}^s n_i = n$;

$$\mathbf{A}_r = \text{diag}\{a_1 I_{n_2}, \dots, a_r I_{n_{r+1}}, \mathbf{g}\};$$

$$\mathbf{A}_{r+1} = \text{diag}\{a_{r+1} I_{n_{r+2}}, \dots, a_m I_{n_{m+1}}, \mathbf{g}\};$$

where $0 < a_i < 1$ for $i \leq r$; and $a_i > 1$ for $i \geq r + 1$: Let $p_r = \sum_{i=1}^{r+1} n_i$; and let

$$I_r(a_1, \dots, a_m) = \int_{\mathbf{R}} \prod_{i=1}^m (1 - \sum_{j=1}^m z_j)^{\frac{n_i}{2} - 1} (1 - \sum_{j=1}^m a_j z_j)^{\frac{n_{s_i}}{2} - 1} \left[\prod_{i=1}^m z_i^{\frac{n_i}{2} - 1} \right] (dz_1 \dots dz_m);$$

where

$$R = \{z; z > 0; \sum_{i=1}^m z_i < 1; \sum_{i=1}^m a_i z_i < 1\}$$

Then:

$$I_r = \int_{\sum_{i=r+1}^m a_i z_i^{\frac{n_i+1}{2}}} \int_{\sum_{i=1}^r z_i (n_i=2)=[i (p_r=2)_i ((n_i - p_r)=2)]} \prod_{j:k=0}^{\infty} [(1=2)_j (1=2)_k=j!k!] c_r(j; k) C_{[j]}(A_r) C_{[k]}(A_{r+1}^{-1}); \quad (45)$$

where the coefficients $c_r(j; k)$ are as described in Lemma 2 except that $\alpha = p_r=2_i - 1$ and $\beta = (n_i - p_r)=2_i - 1$:

These results, when combined with equation (42), yield results exactly analogous to Theorems 1 and 2 for the density of F : We omit the detailed statements of these results.

Anderson (1971), Section 6.7, gives the density of Q when $n = 2N$ and the roots of B each occur with multiplicity 2. In this special case the integrand in (42) is evidently a constant, and I_r in Lemma 4 is simply the volume of the polyhedral region R (and α and β in Lemma 4 are both integers). It is straightforward, though tedious, to confirm that in this special case Lemma 4 yields Anderson's (1971) Corollary 6.7.3.

6 Discussion

The results given above completely characterise the density function of Q , and the representation of the density as a surface integral given in Section 3 provides, by itself, considerable insight into the nature of the density. The analysis also explains why the problem has presented such difficulties for more conventional approaches: integration over a polyhedral region in R^m is no trivial matter. Nevertheless, it is unlikely that the exact expressions for the density given here will replace existing methods for evaluating p_j values, or critical values, for test statistics of the form (1). Although we do intend to explore the usefulness of Theorem 2 for these purposes in future work, the value of both the methods used, and the results themselves, is more likely to reside in their contribution to analytic work on distribution theory.

Implicit in the results given in the paper are various integral formulae that may well have applications elsewhere. For example, the usual expression for a density as the inverse Fourier transform of its characteristic function, when applied to (3), must yield (40), implying a Fourier inversion formula that can doubtless be generalised to cover other (similar) cases. Also, equation (25) is an example of an elliptic integral (of the first kind) - see, for instance, Abramowitz and Stegun (1964), Chapter 17, and the references therein. Lemmas 1 - 4 thus provide explicit multivariate generalisations of some known formulae for elliptic integrals, and show that, as in the univariate case, such integrals are related to the Gaussian hypergeometric function (in the multivariate case, with matrix argument). The argument in the paper also shows that these

integrals can be represented as surface integrals over high-dimensional hyperplanes, and that property can also doubtless be generalised. Development of these matters is obviously well beyond the scope of the present paper.

Of more statistical interest would be to generalise the results to the matrix-variate case, i.e., to statistics of the form $Q = g((Y^0 Y)^{-1} Y^0 B Y)$; where Y is, say, $n \times p$ ($n \geq p$) with independent $N(0, 1)$ elements, and $g(\cdot)$ is a scalar-valued function of its $p \times p$ matrix argument (e.g., $g = \text{trace}$, or $g = \text{det}$). Or, particularly in the case of test statistics, to generalise to the case where $y \gg N(1; S)$; with $1 \notin 0$ and/or $S \notin I_n$. Again, these matters are certainly of interest but beyond our present scope.

Finally, it would be of interest to relate the exact results given here to known asymptotic (as $n \rightarrow \infty$) results for Q . Phillips (1986), analysing the Wald statistic, obtained an exact expression (in an operator form) for the (non-null) density of, essentially, a statistic of the form $W = y^0 B y$; with B also random but independent of y . In that case the asymptotic analysis is straightforward, producing in an elegant manner all known asymptotic results for such statistics. In the case of W , though, the density is analytic over its entire domain, and it seems clear that the case dealt with here may require a more delicate treatment. In addition, the asymptotics evidently depend on assumptions made about the behaviour of B as $n \rightarrow \infty$; and hence on the context. We shall not pursue this further here.

REFERENCES

- Abramowitz, M., and I. Stegun (1964), *Handbook of Mathematical Functions*, Dover Publications, New York.
- Anderson, T.W., (1971), *The Statistical Analysis of Time Series*, John Wiley and Sons, New York.
- Henshaw, R.C. (1966), "Testing Single-Equation Least Squares Regression Models for Autocorrelated Disturbances", *Econometrica*, 34, 646-660.
- Herz, C.S., (1955), "Bessel functions of matrix argument", *Annals of Mathematics*, 61; 3; 474 j 523:
- James, A.T., (1964), "Distributions of matrix variates and latent roots derived from normal samples", *Ann. Math. Statist.*, 35; 475 j 501:
- James, A.T., (1961a), "The distribution of noncentral means with known covariance", *Ann. Math. Statist.*, 32; 874 j 882:
- James, A.T., (1961b), "Zonal polynomials of the real positive definite symmetric matrices", *Ann. Math.*, 74; 456 j 469:
- Koopmans, T., (1942), "Serial correlation and quadratic forms in normal variables", *Ann. Math. Statist.*, 12; 14 j 33:
- Muirhead, R.J., (1982), *Aspects of multivariate statistical theory*, John Wiley and Sons, New York.
- Mulholland, H.P., (1965), "On the degree of smoothness and on singularities in distributions of statistical functions", *Proc. Camb. Phil. Soc.*, 61; 721 j 739:
- Mulholland, H.P., (1970), "On singularities of sampling distributions, in particular for ratios of quadratic forms", *Biometrika*, 57; 155 j 174:
- Phillips, P.C.B., (1986), "The Exact Distribution of the Wald Statistic", *Econometrica*, 54, 881-895.
- Rainville, E.D., (1960), *Special Functions*, Macmillan, New York.
- Saldanha, N.C. and C. Tomei, (1996), "The accumulated distribution of quadratic forms on the sphere", *Linear Algebra and its Applications*, 245; 335 j 351:
- Schechter, M., (1998), "Integration over a polyhedron: an application of the Fourier-Motzkin elimination method", *Amer. Math. Monthly*, 105; 246 j 251:
- Tjur, Tue, (1980), *Probability based on Radon measures*, Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, New York.
- von Neumann, John, (1941), "Distribution of the ratio of the mean square successive difference to the variance", *Ann. Math. Statist.*, 12; 367 j 395:

APPENDIX

Preliminary Remark

In the following proofs we repeatedly integrate multiple power series term-by-term. To avoid tedious repetition, except where the issue is in doubt we do so without explicitly mentioning the justification for this process, namely, that the series produced by it actually converges. We also use a variety of standard devices to manipulate multiple power series, in particular, that of "summing by diagonals". We do not elaborate on these methods here; the reader is referred to Rainville (1960) (especially pp.56 ; 58) for background on these and other methods that will be used below. We use the notation $A < I$ for A diagonal to indicate that the diagonal elements of A are all less than unity.

Top-order zonal polynomials

The expansion of the elementary hypergeometric function

$$\begin{aligned} {}_1F_0(1=2; tB) &= \prod_{i=1}^n (1 - t_{s,i})^{-\frac{1}{2}} \\ &= \sum_{k=0}^{\infty} [t^k (1=2)_{k=1} C_{[k]}(B)]; \end{aligned} \quad (A.1)$$

where $B = \text{diag}(s_1, \dots, s_n)$, may be regarded as a generating function for the $C_{[k]}(B)$ (cf. James, (1964), von Neumann (1941), p. 374). From it we can obtain an explicit expression for the polynomials, as follows. Let $C(k; n)$ denote the set of compositions of k with n parts, i.e., the set of length- n sequences $\alpha = (\alpha_1; \alpha_2; \dots; \alpha_n)$ of non-negative integers satisfying $\sum_{i=1}^n \alpha_i = k$. For $\alpha \in C(k; n)$, let $s^\alpha = \prod_{i=1}^n (s_i)^{\alpha_i}$, let $z^\alpha = \prod_{i=1}^n \alpha_i!$, and let

$$(1=2)^\alpha = \prod_{i=1}^n (1=2)_{\alpha_i}; \quad (A.2)$$

By expanding each term in the product in the first line of (A.1) separately and comparing coefficients of t^k we obtain:

Lemma A.1:

$$C_{[k]}(B) = [k! = (1=2)_k] \sum_{\alpha \in C(k; n)} (1=2)^\alpha s^\alpha = z^\alpha; \quad (A.3)$$

Since the coefficients in the sum here are symmetric (i.e., invariant under permutations of the elements of α); (A.3) yields an expansion for $C_{[k]}(B)$ in terms of the monomial symmetric functions of $(s_1; \dots; s_n)$, see Muirhead (1982), p. 234. Also, it is easy to see from (A.1) that $C_{[k]}(I_n) = (n=2)_{k=1} = (1=2)_k$, and that, if E_1 denotes the $n \times n$ matrix with (1,1) element unity, all other elements zero, $C_{[k]}(E_1) = 1$ for all

k. These two results, together with Muirhead (1982), Theorem 7.2.5, yield the characteristic function in equation (3) in the text, and also the density of W in equation (6).

In what follows we shall also need two further identities for these polynomials. The next two Lemmas give these identities for later use.

Lemma A.2: For $B \in \mathbb{R}^p$ and $B^i < I_p$:

$$C_{[k]}([B \quad I_p]^i) = \prod_{j=0}^{i-1} [I_p - B^j]^{-1} \sum_{k+j=j!} [(1=2)_{k+j} = j! (1=2)_k] C_{[k+j]}(B^i): \quad (\text{A.4})$$

Proof: From the generating function (A.1) we have

$$\prod_{j=0}^{i-1} [I_p - t(B \quad I_p)^j]^{-1} = \sum_{k=0}^{\infty} [(1=2)_{k=k!}] t^k C_{[k]}([B \quad I_p]^i):$$

On the other hand, the left-hand side is also equal to:

$$\prod_{j=0}^{i-1} [I_p - B^j]^{-1} \prod_{j=0}^{i-1} [I_p - (1+t)B^j]^{-1}:$$

Provided $B^i < I_p$, the second factor here can be expanded using (A.1), and we may then expand the term $(1+t)^j$ that occurs binomially, sum by diagonals, and equate the coefficients of powers of t to obtain (A.4). ¥

Lemma A.3: Let $B = \text{diag}\{I_{n_1}, \dots, I_{n_m}\}$; with $\sum_{i=1}^m n_i = n$: Then:

$$C_{[k]}(B) = \sum_{2C(k;m)} [k! = (1=2)_k] \prod_{i=1}^m (n_i = 2)_{k_i} [I_{n_i} = 0]: \quad (\text{A.5})$$

Proof: This follows directly from the generating function (A.1) on noting that:

$$\prod_{i=1}^m [I_{n_i} - tB^i]^{-1} = \prod_{i=1}^m (1 - t_{n_i})^{-1};$$

expanding each term in the product separately, and identifying the coefficients of powers of t . ¥

Proof of Proposition 1

In the case $a_i < 1$; it is clear that $\sum_{i=1}^m a_i z_i < 1$ implies $\sum_{i=1}^m a_i z_i < 1$ (but not conversely), so that the region $\{z; z > 0; \sum_{i=1}^m a_i z_i < 1\}$ is a proper subset of the region $\{z; z > 0; \sum_{i=1}^m a_i z_i < 1\}$, and the integral in (23) is over the region $\{z; z > 0; \sum_{i=1}^m a_i z_i < 1\}$. Likewise, if $a_i > 1$, the situation is reversed and the integral in (23) is over the region $\{z; z > 0; \sum_{i=1}^m a_i z_i < 1\}$. We shall first show that the result for each of

these cases can be obtained from that for the other, and then immediately generalise this observation to obtain Proposition 1.

Let us denote the integral in (23) for the case $a_i < 1 \forall i$ by $I_m(a_1; \dots; a_m)$, the subscript indicating that all m of the a_i 's are less than one:

$$I_m(a_1; \dots; a_m) = \int_{\substack{\mathbf{z} \\ S_1 z_i < 1}} (1 - S_1 z_i)^{i-1} (1 - S_1 a_i z_i)^{i-1} \prod_{i=1}^m [1 - z_i]^{i-1} dz_1 \dots dz_m: \quad (\text{A.6})$$

For the case $a_i > 1 \forall i$; the integrand is as in (A.6), but the region of integration becomes $\mathbf{z}; z > 0; S_1 a_i z_i < 1 \forall i$, so that, for this case, the integral in (23) becomes:

$$I_0(a_1; \dots; a_m) = \int_{\substack{\mathbf{z} \\ S_1 a_i z_i < 1}} (1 - S_1 z_i)^{i-1} (1 - S_1 a_i z_i)^{i-1} \prod_{i=1}^m [1 - z_i]^{i-1} dz_1 \dots dz_m: \quad (\text{A.7})$$

Now put $z_i = a_i z_i; i = 1; \dots; m$, in (A.7). The region of integration becomes $\mathbf{z}; z > 0; S_1 z_i < 1 \forall i$, and the integrand becomes (taking account of the Jacobian of the transformation):

$$\prod_{i=1}^m [1 - a_i]^{i-1} (1 - S_1 z_i)^{i-1} (1 - S_1 a_i^{-1} z_i)^{i-1} \prod_{i=1}^m [1 - z_i]^{i-1} dz_1 \dots dz_m;$$

where $a_i^{-1} < 1$ for $i = 1; \dots; m$. Hence, we see at once that, for the case $a_i > 1 \forall i$;

$$I_0(a_1; \dots; a_m) = \prod_{i=1}^m [1 - a_i]^{i-1} I_m(a_1^{-1}; \dots; a_m^{-1}): \quad (\text{A.8})$$

But, since it is just the number of a_i 's that are less than one that determines the shape of the region R in (23), and the transformation $z_i \rightarrow z_i = a_i z_i$ simply multiplies the integral in (23) by $\prod_{i=1}^m [1 - a_i]^{i-1}$ and replaces the a_i by the a_i^{-1} , Proposition 1 follows by the same argument.

Substitution of Inequalities

As mentioned in the text, to integrate an expression involving non-negative variables $z_1; \dots; z_m$, say, over the region $\mathbf{z}; z > 0; S_1 z_i < 1 \forall i$, it is natural to successively transform variables to, ...rst, $b_1 = z_1 = (1 - S_2 z_1)$, so that $0 < b_1 < 1$; then, in the resulting expression, to $b_2 = z_2 = (1 - S_3 z_2)$, so that $0 < b_2 < 1$; and so on, to ...nally obtain a function of $b_1; \dots; b_m$ to be integrated over $0 < b_i < 1; i = 1; \dots; m$. This process defines a transformation from $(z_1; \dots; z_m) \in \mathbf{z}; z > 0; S_{i=1}^m z_i < 1 \forall i$ to new variables $(b_1; \dots; b_m)$ that maps the region $S_{i=1}^m z_i < 1$ onto the (open) unit m -cube, $C_m = \{b_i; 0 < b_i < 1; i = 1; \dots; m\}$. It is straightforward to check that:

$$z_s = b_s \prod_{i=s+1}^m (1 - b_i); s = 1; \dots; m-1; z_m = b_m; \quad (\text{A.9})$$

so that

$$1 - S_{i=1}^m z_i = \prod_{i=1}^m (1 - b_i); \quad (\text{A.10})$$

and

$$\prod_{i=1}^m z_i = \left[\prod_{i=1}^m b_i \right] \left[\prod_{i=1}^m (1 - b_i)^{i-1} \right]; \quad (\text{A.11})$$

and that the Jacobian of the transformation is:

$$J = \prod_{i=1}^m (1 - b_i)^{i-1}; \quad (\text{A.12})$$

The region of integration, R , in (23), also requires that $\sum_{i=1}^m a_i z_i < 1$: Let $d_m = \sum_{i=1}^m a_i z_i$. In terms of the b_i ; $d_m = \sum_{j=1}^m a_j b_j [\prod_{i=j+1}^m (1 - b_i)]$. Define

$$d_s = \sum_{j=1}^s a_j b_j [\prod_{i=j+1}^s (1 - b_i)]; \quad s = 2; \dots; m; \quad (\text{A.13})$$

and $d_1 = a_1 b_1$, and notice that the d_s satisfy the recursion:

$$d_s = b_s a_s + (1 - b_s) d_{s-1}; \quad d_1 = a_1 b_1; \quad (\text{A.14})$$

with d_{s-1} a function only of $b_1; \dots; b_{s-1}$:

It is clear from (A.9) and (A.13) that the condition $d_m < 1$ is automatically satisfied if $a_i < 1/b_i$. That is, if $a_i < 1/b_i$, the integral in (23) can be evaluated by making the substitutions (A.9) in the integrand and integrating over all of C_m . We deal with this case next.

Proof of Lemma 1

On transforming from z to $(b_1; \dots; b_m)$ as described above, and using (A.9); (A.12), (23) becomes:

$$I_m = \int_{C_m} [\prod_{i=1}^m b_i]^{i-1} [\prod_{i=1}^m (1 - b_i)^{\frac{i-2}{2}}] (1 - d_m)^{i-1} (db_1 db_2 \dots db_m); \quad (\text{A.15})$$

with d_m as given in (A.13) (setting $s = m$). Expanding the last term in the integrand as a power series in d_m , then expanding the powers of d_m multinomially, we have:

$$\sum_{k=0}^{\infty} (1-d_m)^k \sum_{2C(k;m)} [a_i^{c_i} \dots] g(b_1; \dots; b_m); \quad (\text{A.16})$$

with

$$g(b_1; \dots; b_m) = b [\prod_{s=2}^m (1 - b_s)^{\sum_{i=1}^{s-1} k_i}]; \quad (\text{A.17})$$

Term-by-term integration of the series is justified by its uniform convergence on the region of integration, C_m , and it is straightforward to obtain the result given in Lemma 1 on using the expansion for the top-order zonal polynomials given in Lemma A.1, together with the fact that, for $c > 0$; $(c)_k = \prod_{j=1}^k (c + j - 1)$.

Proof of Lemma 2

Assume now that $a_i < 1/b_i$ for $i \leq r$, and $a_i > 1/b_i$ for $i \geq r+1$: In this case the condition $b_i \leq 1/a_i$ no longer ensures that $d_m < 1$; the condition $d_m < 1$ imposes restrictions on $(b_{r+1}; \dots; b_m)$. In fact, from the recursion (A.14) we see that, for each $s > r$, we must have $d_s < 1$ and

$$0 < b_s < (1 - d_{s-1}) / a_s = u_s; \quad \text{say.}$$

Our strategy will be to make a sequence of transformations of $b_m; b_{m-1}; \dots; b_{r+1}$ that converts the region of integration for the new variables back to $C_{m;r}$. The remaining terms involve only d_r , which is a function only of $b_1; \dots; b_r$, and, since $a_i < 1$ for $i \leq r$, this can be integrated over all of C_r , as in the proof of Lemma 1.

For $s = m; m-1; \dots; r+1$, define

$$u_s = (1 - d_{s-1}) = (a_s - d_{s-1}) < 1; \quad (\text{A.18})$$

At each stage of the sequential process described above we transform b_s to:

$$t_s = b_s(1 - u_s) = [u_s(1 - b_s)]; \quad (\text{A.19})$$

so that $0 < t_s < 1$ on the interval $0 < b_s < u_s$. We have:

$$\begin{aligned} b_s &= u_s t_s = [1 - u_s(1 - t_s)] \\ 1 - b_s &= (1 - u_s) = [1 - u_s(1 - t_s)] \end{aligned} \quad (\text{A.20})$$

and the Jacobian of the transformation is:

$$J_s = u_s(1 - u_s) = [1 - u_s(1 - t_s)]^2; \quad (\text{A.21})$$

To simplify the notation we also introduce some further (temporary) definitions at this point. For each $s = m; m-1; \dots; r+1$, define

$$c_s = a_s - 1 > 0; \quad (\text{A.22})$$

and the sequence of functions

$$\begin{aligned} h_s &= a_s - d_{s-1} - (1 - t_s)(1 - d_{s-1}) \\ &= c_s + t_s(1 - d_{s-1}); \end{aligned} \quad (\text{A.23})$$

a function only of $b_1; \dots; b_{s-1}$ and t_s . Rewriting (A.20) and (A.21) in this notation we have:

$$\begin{aligned} b_s &= t_s(1 - d_{s-1}) = h_s \\ 1 - b_s &= c_s = h_s; \end{aligned} \quad (\text{A.24})$$

and

$$J_s = c_s(1 - d_{s-1}) = h_s^2; \quad (\text{A.25})$$

Finally,

$$1 - d_s = c_s(1 - t_s)(1 - d_{s-1}) = h_s; \quad (\text{A.26})$$

Now, using (A.24) and (A.26) to transform from b_m to t_m in the integrand of (23) gives:

$$c_m^{\frac{m-1}{2}} t_m^{\frac{1}{2}} (1 - t_m)^{\frac{1}{2}} h_m^{\frac{m}{2}}; \quad (\text{A.27})$$

Only the last term, h_m , involves the remaining b_i^0 s. Transforming $b_{m_i-1} ! t_{m_i-1}$ in the same way (assuming that $a_{m_i-1} > 1$), the terms in b_{m_i-1} become:

$$c_{m_i-1}^{m_i-1} t_{m_i-1}^{i-\frac{1}{2}} (1 - d_{m_i-2})^{\frac{1}{2}} [c_m h_{m_i-1} + c_{m_i-1} t_m (1 - t_{m_i-1}) (1 - d_{m_i-2})]^i \frac{m}{2}; \quad (\text{A.28})$$

and the last term here (the only term apart from d_{m_i-2} involving the remaining b_i^0 s again) simplifies to:

$$[c_m c_{m_i-1} + (c_m t_{m_i-1} + c_{m_i-1} t_m (1 - t_{m_i-1})) (1 - d_{m_i-2})]^i \frac{m}{2}; \quad (\text{A.29})$$

Iterating this process of transformation, the recursion should be clear: at each step (transforming $b_{m_i s}$, say) (i) a term $c_{m_i s}^{m_i-1}$ is introduced, (ii) the term $(1 - d_{m_i s_i-1})$ appears with a power increased by 1/2 from the previous step (beginning at $-1=2$), and (iii) if the term raised to the power $m=2$ (the only term involving the remaining b_i^0 s) at the previous step was, say,

$$p_{m_i s+1} + q_{m_i s+1} (1 - d_{m_i s});$$

that term becomes

$$p_{m_i s} + q_{m_i s} (1 - d_{m_i s_i-1});$$

with

$$p_{m_i s} = c_{m_i s} p_{m_i s+1} \quad (\text{A.30})$$

and

$$q_{m_i s} = p_{m_i s+1} t_{m_i s} + q_{m_i s+1} c_{m_i s} (1 - t_{m_i s}); \quad (\text{A.31})$$

These relations define the $p_{m_i s}$; $q_{m_i s}$ recursively, beginning with $p_m = c_m$; $q_m = t_m$, and we continue the process of transformation until the term remaining is $(1 - d_r)$, i.e., until $m_i s = r + 1$: At this point the integrand has become (ignoring, for the moment, terms other than d_r that involve only $b_1; \dots; b_r$):

$$[c_{r+1}^{m_i-1} c_i]^{\frac{m_i-1}{2}} [c_{r+1} t_i]^i \frac{1}{2} [c_{r+1} (1 - t_i)^{\frac{m_i-1}{2}}] (1 - d_r)^{\frac{m_i-1}{2}} \int [p_{r+1} + q_{r+1} (1 - d_r)]^i \frac{m}{2}; \quad (\text{A.32})$$

Now, from (A.30) it is clear that $p_{r+1} = c_{r+1}^{m_i-1} c_i$. Removing this factor from the last term in (A.32), and using (A.31), we find that:

$$q_{r+1} = q_{r+1} p_{r+1} = c_{r+1}^{m_i-1} t_{r+1} + \sum_{s=r+2}^{m_i} c_i^{s-1} t_i [c_{s=r+1}^{i-1} (1 - t_s)]; \quad (\text{A.33})$$

The first and last terms in (A.32) are thereby modified to $[c_{r+1}^{m_i-1} c_i]^i \frac{1}{2}$ and

$$[1 + q_{r+1} (1 - d_r)]^i \frac{m}{2} \quad (\text{A.34})$$

respectively.

We now seek to integrate out $t_{r+1}; \dots; t_m$ over the unit $(m - r)_j$ cube $C_{m_i r}$, and $b_1; \dots; b_r$ over C_r . Unfortunately, this cannot be done by direct expansion of (A.34), because $q_{r+1}(1 - d_r)$ is not bounded below one over this region. Thus, we must take a slightly less direct route. We first express the term (A.34) that now occurs in the integrand as a Laplace transform:

$$[1 + q_{r+1}(1 - d_r)]^{\frac{m}{2}} = [j(m=2)]^{i-1} \int_{w>0} \exp\{j w[1 + q_{r+1}(1 - d_r)]\} g w^{\frac{m}{2}i-1} dw: \quad (\text{A.35})$$

This integral converges for all $(b_1; \dots; b_r; t_{r+1}; \dots; t_m) \in C_r \times C_{m_i r}$, and we may interchange the order of integration with respect to w and the b 's and t 's.

Next, in (A.35), change the variable of integration to $w = w(1 - d_r)$. The last two terms in the integrand (A.32) then become:

$$(1 - d_r)^{i \frac{r+1}{2}} [j(m=2)]^{i-1} \int_{w>0} \exp\{j w(1 - d_r)\} w^{\frac{m}{2}i-1} dw: \quad (\text{A.36})$$

Notice that this step has the effect of separating terms involving $(1 - d_r)$ from that involving q_{r+1} , so that, to begin with, we can deal with the integral over C_r separately from that over $C_{m_i r}$. Thus, consider first the terms in $(1 - d_r)$.

Expanding the exponential $\exp\{j w(1 - d_r)\}$ in (A.36), multiplying by $(1 - d_r)^{i \frac{r+1}{2}}$, and expanding the resulting term $(1 - d_r)^{i(k + \frac{r+1}{2})}$ in a power series gives:

$$\sum_{j;k=0} [(k + (r + 1)=2)_j = j!k!] (j - w)^k d_r^j: \quad (\text{A.37})$$

Multinomial expansion of the powers d_r^j , using (A.13), followed by integration with respect to $b_1; \dots; b_r$ over C_r , exactly as in the proof of Lemma 1, gives:

$$[j \frac{r+1}{2} = j ((r + 1)=2)] \sum_{j;k=0} [(k + (r + 1)=2)_j (1=2)_j = j!k! ((r + 1)=2)_j] (j - w)^k C_{[j]}(A_r); \quad (\text{A.38})$$

where $A_r = \text{diag}\{a_1; \dots; a_r\}$. Using properties of the Pochhammer symbols it is easy to see that the sum on k in (A.38) can be reduced to:

$${}_1F_1(j + (r + 1)=2; (r + 1)=2; j - w);$$

which, on using the Kummer formula

$${}_1F_1(a; c; x) = \exp\{x\} {}_1F_1(c - a; c; -x);$$

becomes

$$\exp\{j w\} {}_1F_1(j; j; (r + 1)=2; w); \quad (\text{A.39})$$

the series here terminating at the j_i th term. Hence, (A.38) becomes:

$$[\frac{r+1}{2}]_i ((r+1)=2) \exp f_i w g \sum_{j=0}^{j_i} [(1=2)_j = j!] {}_1F_1(j; (r+1)=2; w) C_{[j]}(A_r); \quad (A.40)$$

It is straightforward to check that the series converges for all finite $w > 0$; essentially because $A_r < I_r$.

Consider next the term in (A.36) involving q_{r+1} . An exactly analogous process of expansion and integration produces:

$$[\frac{m_i r+1}{2}]_i ((m_i r+1)=2) \sum_{k=0}^{j_i} [(1=2)_k = k! ((m_i r+1)=2)_k] \int (i w)^k C_{[k]}((A_{r+1} i I_{m_i r})^{i-1}); \quad (A.41)$$

where $A_{r+1} = \text{diag } a_{r+1}; \dots; a_m g$, and the series (a confluent hypergeometric function) converges for all finite $w > 0$: But, using Lemma A.2, the series in (A.41) may be written as:

$$I_{m_i r} i A_{r+1}^{i-1} \sum_{k;s=0}^{j_i} [(1=2)_{k+s} = k! s! ((m_i r+1)=2)_k] (i w)^k C_{[k+s]}(A_{r+1}^{i-1});$$

Summing the double series by diagonals, and noting that

$$\sum_{k=0}^{j_i} \sum_{s=0}^{j_i-k} (i w)^k ((m_i r+1)=2)_k = {}_1F_1(j_i; (m_i r+1)=2; w);$$

(A.41) becomes:

$$[\frac{m_i r+1}{2}]_i ((m_i r+1)=2) I_{m_i r} i A_{r+1}^{i-1} \sum_{k=0}^{j_i} [(1=2)_k = k!] {}_1F_1(j_i; k; (m_i r+1)=2; w) C_{[k]}(A_{r+1}^{i-1}); \quad (A.42)$$

the series again converging for all $w > 0$ because $A_{r+1}^{i-1} < I_{m_i r}$.

Finally, we need now to integrate out w . Proceeding formally, we multiply (A.40) by (A.42) (and the term $[j_i (m=2)]_i^{-1} w^{m=2i-1}$ from (A.35)), and integrate the resulting double series term-by-term with respect to w . On multiplying by $[i]_{r+1} c_i i^{-\frac{1}{2}}$ we obtain:

$$[\frac{m+2}{2}]_i ((r+1)=2)_i ((m_i r+1)=2) [i]_{r+1} a_i i^{\frac{1}{2}} \sum_{j;k=0}^{j_i} [(1=2)_j (1=2)_k = j! k!] C_{[j]}(A_r) C_{[k]}(A_{r+1}^{i-1}) C_r(j; k); \quad (A.43)$$

with numerical coefficients $c_r(j; k)$ given by:

$$c_r(j; k) = \int_0^1 [j(m=2)]^{i-1} \exp\{j w g w^{\frac{m}{2}i-1} {}_1F_1(j; j; (r+1)=2; w) {}_1F_1(j; k; (m-j-r+1)=2; w) dw: \quad (\text{A.44})$$

This proves Lemma 2, apart from the properties of the numerical coefficients $c_r(j; k)$, to which we now turn.

Properties of the coefficients $c_r(j; k)$.

The hypergeometric functions that appear in (A.44) are proportional to the (generalised) Laguerre polynomials $L_j^{(\otimes)}(w)$ and $L_k^{(\ominus)}(w)$, defined by:

$$L_j^{(\otimes)}(w) = [(\otimes + 1)_j = j!] {}_1F_1(j; j; \otimes + 1; w); \quad (\text{A.45})$$

see Rainville (1960), Chapter 12. Setting $\otimes = (r-j-1)=2$; and $\ominus = (m-j-r-j-1)=2$; so that $\otimes + \ominus = m=2-j-1$, (A.44) becomes:

$$\begin{aligned} c_r(j; k) &= \int_0^1 [j!k! = (\otimes + 1)_j (\ominus + 1)_k] [j(m=2)]^{i-1} \exp\{j w g w^{\otimes+\ominus} L_j^{(\otimes)}(w) L_k^{(\ominus)}(w) dw: \\ &= [j!k! = (\otimes + 1)_j (\ominus + 1)_k] c_r^{\otimes}(j; k); \text{ say:} \end{aligned} \quad (\text{A.46})$$

Now, the generating function for the Laguerre polynomials,

$$g_{\otimes}(z; w) = \sum_{j=0}^{\infty} z^j L_j^{(\otimes)}(w) = (1-jz)^{-(\otimes+1)} \exp\{j wz = (1-jz)g\}; \quad (\text{A.47})$$

readily provides a generating function for the $c_r^{\otimes}(j; k)$ in (A.46):

$$\begin{aligned} g_{\otimes, \ominus}(z_1; z_2) &= \sum_{j,k=0}^{\infty} z_1^j z_2^k c_r^{\otimes}(j; k) \\ &= \int_0^1 [j(m=2)]^{i-1} \exp\{j w g w^{\otimes+\ominus} g_{\otimes}(z_1; w) g_{\ominus}(z_2; w) dw \\ &= (1-jz_1)^{-(\otimes)} (1-jz_2)^{-(\ominus)} (1-jz_1 z_2)^{i(\otimes+\ominus+1)}; \quad 0 < z_1; z_2 < 1: \end{aligned} \quad (\text{A.48})$$

That is, $c_r^{\otimes}(j; k)$ in (A.46) is the coefficient of $z_1^j z_2^k$ in the expansion of $g_{\otimes, \ominus}(z_1; z_2)$ in the last line of (A.48).

Expanding the three components of $g_{\otimes, \ominus}(z_1; z_2)$ we have:

$$\begin{aligned} g_{\otimes, \ominus}(z_1; z_2) &= (1-jz_1)^{-(\otimes)} (1-jz_2)^{-(\ominus)} (1-jz_1 z_2)^{i(\otimes+\ominus+1)} \\ &= \sum_{l,m,n=0}^{\infty} z_1^{l+m} z_2^{l+n} [(\otimes + \ominus + 1)_l (j^{-})_m (j^{\otimes})_n = l!m!n!]; \end{aligned} \quad (\text{A.49})$$

Suppose first that $k = j + t$, with $t \geq 0$, so that $k \geq j$: In (A.49), the term $z_1^j z_2^{j+t}$ occurs when $l = j + m$, $n = m + t$, and all $0 \leq m \leq j$. Hence, the coefficient of $z_1^j z_2^{j+t}$ is:

$$\begin{aligned} c_r^\alpha(j; j+t) &= \sum_{m=0}^j [(\alpha + \beta + 1)_{j-m} (i^-)_m (i^\alpha)_{m+t}] = [m!(j-m)!(m+t)!] \\ &= [(\alpha + \beta + 1)_j (i^\alpha)_{t=j}! t!] {}_3F_2(i^-; j; i^-; i^\alpha + t; t+1; i^\alpha + \beta + j; 1); \end{aligned} \quad (\text{A.50})$$

The ${}_3F_2$ function here is Saalschützian, and Saalschütz' Theorem (Rainville (1960), p. 87) applies, giving, for $t \geq 0$,

$$\begin{aligned} c_r^\alpha(j; j+t) &= [(\alpha + 1)_j (i^\alpha)_t (\beta + 1 + t)_j] = [j! t! (t+1)_j] \\ &= [(\alpha + 1)_j (\beta + 1)_k (i^\alpha)_t] = [j! k! (\beta + 1)_t]; \end{aligned}$$

so that, for $k \geq j$:

$$c_r(j; k) = (i^\alpha)_{k-j} (\beta + 1)_{k-j}; \quad (\text{A.51})$$

An exactly analogous argument for the case $j = k + t$ yields, for the case $j \geq k$:

$$c_r(j; k) = (i^-)_{j-k} (\alpha + 1)_{j-k}; \quad (\text{A.52})$$

If $j = k$, $c_r(j; k) = 1$ in both (A.51) and (A.52). If $\alpha = (r - 1) = 2$ is an integer, say $\alpha = p$; $c_r(j; k) = 0$ for $k - j > p$; or $j < k - p$; and if $\beta = (m - r - 1) = 2$ is an integer, say $\beta = q$, $c_r(j; k) = 0$ for $j - k > q$, or $j > k + q$: Hence, (A.51) and (A.52) yield the properties of the $c_r(j; k)$ given in Lemma 2.

It remains to prove that the double series in (A.43) converges, thus justifying term by term integration with respect to w . But, from Lemma A.1 it is clear that, for $0 < A_r < I_r$; and $0 < A_{r+1}^i < I_{m_i r}$; $0 < C_{[j]}(A_r) < a_r^j C_{[j]}(I_r) = a_r^j (r=2)_j = (1=2)_j$; and $0 < C_{[k]}(A_{r+1}^i) < a_{r+1}^{i k} C_{[k]}(I_{m_i r}) = a_{r+1}^{i k} ((m_i r)=2)_k = (1=2)_k$: Also, for all $j; k \geq 0$; $(r=2)_j < (\alpha + 1)_j$ and $((m_i r)=2)_k < (\beta + 1)_k$: Hence, the series in (A.43) is dominated termwise by the series

$$\sum_{j,k=0}^{\infty} z_1^j z_2^k [(\alpha + 1)_j (\beta + 1)_k j! k!] c_r(j; k);$$

with $z_1 = a_r$ and $z_2 = a_{r+1}^i$. But, in view of (A.46) and (A.48), this is the series expansion for the generating function $g_{\alpha, \beta}^-(z_1; z_2)$; which certainly converges (to the expression in the last line of (A.48)) for $0 < a_r < 1$; $0 < a_{r+1}^i < 1$: Hence, the series in (A.43) converges.

Outline Proof of Lemmas 3 and 4

We do not give the complete details for Lemmas 3 and 4, because they are exact analogues of those above for the case of n distinct roots. The procedure described above in the proof of Lemma 2 produces, when applied to the integral in equation (42) in the text, an expression proportional to that on the right in (A.5). The results stated in Lemmas 3 and 4 in the text then follow immediately from Lemma A.3.