

# Distributions of Ratios: From Random Variables to Random Matrices\*

Thu Pham-Gia<sup>#</sup>, Noyan Turkkan

Department of Mathematics and Statistics, Universite de Moncton, Canada

E-mail: <sup>#</sup>phamgit@umoncton.ca

Received March 25, 2011; revised April 17, 2011; accepted April 25, 2011

## Abstract

The ratio  $R$  of two random quantities is frequently encountered in probability and statistics. But while for unidimensional statistical variables the distribution of  $R$  can be computed relatively easily, for symmetric positive definite random matrices, this ratio can take various forms and its distribution, and even its definition, can offer many challenges. However, for the distribution of its determinant, Meijer  $\mathbf{G}$ -function often provides an effective analytic and computational tool, applicable at any division level, because of its reproductive property.

**Keywords:** Matrix Variate, Beta Distribution, Generalized-F Distribution, Ratios, Meijer  $\mathbf{G}$ -Function, Wishart Distribution, Ratio

## 1. Introduction

In statistical analysis several important concepts and methods rely on two types of ratios of two independent, or dependent, random quantities. In univariate statistics, for example, the F-test, well-utilized in Regression and Analysis of variance, relies on the ratio of two independent chi-square variables, which are special cases of the gamma distribution. In multivariate analysis, similar problems use the ratio of two random matrices, and also the ratio of their determinants, so that some inference using a statistic based on the latter ratio, can be carried out. Latent roots of these ratios, considered either individually, or collectively, are also used for this purpose.

But, although most problems related to the above ratios are usually well understood in univariate statistics, with the expressions of their densities often available in closed forms, there are still considerable gaps in multivariate statistics. Here, few of the concerned distributions are known, less computed, tabulated, or available on a computer software. Results abound in terms of approximations or asymptotic estimations, but, as pointed out by Pillai ([1] and [2]), more than thirty years ago, asymptotic methods do not effectively contribute to the practical use of the related methods. Computation for hypergeometric functions of matrix arguments, or for zonal polynomials, are only in a state of development [3], and,

at the present time, we still know little about their numerical values, to be able to effectively compute the power of some tests.

The use of special functions, especially Meijer  $\mathbf{G}$ -functions and Fox  $\mathbf{H}$ -functions [4] has helped a great deal in the study of the densities of the determinants of products and ratios of random variables, a domain not fully explored yet, computationally. A large number of common densities can be expressed as  $\mathbf{G}$ -functions and since products and ratios of  $\mathbf{G}$ -functions distributions are again  $\mathbf{G}$ -functions distributions, this process can be repeatedly applied. Several computer applicable forms of these functions have been presented by Springer [5]. But it is the recent availability of computer routines to deal with them, in some commercial softwares like **Maple** or **Mathematica**, that made their use quite convenient and effective [6]. We should also mention here the increasing role that the  $\mathbf{G}$ -function is taking in the above two softwares, in the numerical computation of integrals [7].

In Section 2 we recall the case of the gamma distribution and various results related to the ratio of two gammas in univariate statistics. In Section 3, going into matrix variate distributions, we consider first the classical case where both  $\mathbf{A}$  and  $\mathbf{B}$  are Wishart matrices, leading to the two types of matrix variate beta distributions for their ratios. The two associated determinant distributions are the two Wilks's lambdas, with density expressed in terms of  $\mathbf{G}$ -functions. It is of interest to note that the variety of

\*Research partially supported by NSERC grant A9249 (Canada).

special mathematical functions used in the previous section can all be expressed as **G**-functions, making the latter the only tool really required.

In Section 4, extensions of the results are made in several directions, to several Wishart matrices and to the matrix variate Gamma distribution. Several ratios have their determinant distribution established here, for example the ratio of a matrix variate beta distribution to a Wishart matrix, generalizing the ratio of a beta to a chi square. Two numerical examples are given and in Section 5 an example and an application are presented. Finally, although ratios are treated here, products, which are usually simpler to deal with, are also sometimes studied.

## 2. Ratio of Two Univariate Random Variables

We recall here some results related to the ratios of two independent r.v.'s so that the reader can have a comparative view with those related to random matrices in the next section.

First, for  $X \sim N(\mu_X, \sigma_X^2)$ , independent of  $Y \sim N(\mu_Y, \sigma_Y^2)$ ,  $R = X/Y$  has density established numerically by Springer ([5], p. 148), using Mellin transform. Pham-Gia, Turkkan and Marchand ([8]) established a closed form expression for the more general case of the bivariate normal  $(X, Y) \sim N(\mu_X, \mu_Y; \sigma_X^2, \sigma_Y^2; \rho)$ , using Kummer hypergeometric function  ${}_1F_1(\cdot)$ . Springer ([5], p.156) obtained another expression but only for the case  $\mu_X = \mu_Y = 0$ .

The gamma distribution,  $X \sim Ga(\gamma, \delta)$ , with density  $g(x; \gamma, \delta) = x^{\delta-1} \exp(-x/\gamma) / [\gamma^\delta \Gamma(\delta)]$ ,  $x > 0$  has as special case  $Ga(n/2, 2)$  called the Chi-square with  $n$  degrees of freedom  $\chi_n^2$ , with density :

$$g(x; n) = x^{\frac{n}{2}-1} \exp(-x/2) / [2^{n/2} \Gamma(n/2)], x > 0.$$

With independent Chi-square variables  $X_1 \sim \chi_{n_1}^2$  and  $X_2 \sim \chi_{n_2}^2$ , we can form the two ratios  $W_1 = X_1/(X_1 + X_2)$  and  $W_2 = X_1/X_2$  which have, respectively, the standard beta distribution on  $[0,1]$  (or beta of the first kind  $beta^I(\gamma_1, \gamma_2)$ , with density defined on  $[0,1]$  by:

$$f(x) = x^{\gamma_1-1} (1-x)^{\gamma_2-1} / B(\gamma_1, \gamma_2), \gamma_1, \gamma_2 > 0.$$

and standard betaprime distribution (beta of the second kind  $beta^{II}(\gamma_1, \gamma_2)$ ) defined on  $[0, \infty]$  by

$$f(y) = y^{\gamma_1-1} / [B(\gamma_1, \gamma_2)(1+y)^{\gamma_1+\gamma_2}].$$

The Fisher-Snedecor variable  $F_{\nu_1, \nu_2}$  is just a multiple of the univariate beta prime  $W_2$ .

For independent  $X_i \sim Ga(\gamma_i, \delta_i), i = 1, 2$ , we can form

the same ratios  $T_1 = X_1/(X_1 + X_2)$  and  $T_2 = X_1/X_2$ .  $T_2$  has the Generalized-F distribution,  $T_2 \sim GF(\alpha, \beta, \lambda)$ , with density :

$$f(t; \alpha, \beta, \lambda) = \frac{\lambda^\alpha t^{\alpha-1}}{B(\alpha, \beta)(1+\lambda t)^{\alpha+\beta}},$$

where  $\alpha = \delta_1, \beta = \delta_2, \lambda = \gamma_2/\gamma_1$  and we have  $F_{\delta_1, \delta_2} = T_2(\delta_1/2, \delta_2/2, \delta_1/\delta_2)$ .

Naturally, when  $\delta_1 = \delta_2$ ,  $T_1 \sim beta^I(\gamma_1, \gamma_2)$  and  $T_2 \sim beta^{II}(\gamma_1, \gamma_2)$ .

Starting now from the standard beta distribution defined on  $(0,1)$ , Pham-Gia and Turkkan [9] give the expression of the density of  $R$ , and also of  $W = X/(X + Y)$ , using Gauss hypergeometric function  ${}_2F_1(\cdot)$ . For the general beta defined on a finite interval  $(a,b)$ , Pham-Gia and Turkkan [10] gives the density of  $R$ , using Appell's function  $F_D(\cdot)$ , a generalization of  ${}_2F_1(\cdot)$ . In [11], several cases are considered for the ratio  $X/Y$ , with  $Y \sim Ga(\alpha, \beta)$ . In particular, the Hermite and Tricomi functions,  $H_\nu(\cdot)$  and  $\psi(\cdot)$ , are used. The Generalized-F variable being the ratio of two independent gamma variables, its ratio to an independent Gamma is also given there, using  $\psi(\cdot)$ . Finally, the ratio of two independent Generalized-F variables is given in [11], using Appell function  $F_D(\cdot)$  again. All these operations will be generalized to random matrices in the following section.

## 3. Distribution of the Ratio of Two Matrix Variates

### 3.1. Three Types of Distributions

Under their full generality, rectangular  $(p \times q)$  random matrices can be considered, but to avoid several difficulties in matrix operations and definitions, we consider only symmetric positive definite matrices. Also, here, we will be concerned only with the non-singular matrix, with its null or central distribution and the exact, non-asymptotic expression of the latter.

For a random  $(p \times q)$  symmetric matrix,  $p \geq 1$ , there are three associated distributions. We essentially distinguish between:

1) the distribution of its elements (*i.e.* its  $p(p+1)/2$  independent elements  $e_{ij}$  in the case of a symmetric matrix), called here its *elements distribution* (with matrix input), for convenience. This is a mathematical expression relating the components of the matrix, but usually, it is too complex to be expressed as an equation (or several equations) on the elements  $e_{ij}$  themselves and hence, most often, it is expressed as an equation on its determinant.

2) the univariate distribution of its determinant (de-

noted here by  $|X|$  instead of  $\det(X)$  (with positive real input), called *determinant distribution*, and

3) the distribution of its latent roots (with  $p$ -vector input), called *latent roots distribution*.

These three distributions evidently become a single one for a positive univariate variable, then called its density. The literature is mostly concerned with the first distribution [12] than with the other two. The focal point of this article is in the determinant distribution.

### 3.2. General Approach

When dealing with a random matrix, the distribution of the elements within that random matrix constitutes the first step in its study, together with the computation of its moments, characteristic function and other properties.

Let  $X$  and  $Y$  be two symmetric, positive definite independent ( $p \times q$ ) random matrices with densities (or elements distributions)  $f_X(X)$  and  $f_Y(Y)$ . As in univariate statistics, we first define two types of ratios, type I of the form  $X/X+Y$ , and type II, of the form:  $X/Y$ . However, for matrices, there are several ratios of type II that can be formed:  $Z_1 = XY^{-1}$ ,  $Z_2 = Y^{-1}X$  and  $Z_3 = Y^{-1/2}X(Y^{-1/2})^T$ , where  $Y^{1/2}$  is the symmetric square root of  $Y$ , i.e.  $Y^{1/2}Y^{1/2} = Y$ , beside the two formed with the Cholesky decomposition of  $Y$ ,  $Z_4 = (U^T)^{-1}X(U)^{-1}$  and  $Z_5 = (V^T)^{-1}X(V)^{-1}$ , where  $U$  is upper triangular, with  $UU^T = Y$ , and  $V$ , lower triangular, with  $VV^T = Y$ .

1) For elements distribution, we only consider  $Z_3$ , which is positive and symmetric, but there are applications of  $Z_4$  and  $Z_5$  in the statistical literature. We can determine the matrix variate distribution of  $Z_3$  from those of  $X$  and  $Y$ . In general, for two matrix variates  $A$  and  $B$ , with joint density  $f_{A,B}(A,B)$  the density of  $G = A^{-1/2}BA^{-1/2}$  is obtained by a change of variables:  $h(G) = \int A f_{A,G}(A,G) |A|^{-(p+1)/2} dA$ , with  $dA$  associated with all elements of  $A$ . When  $A$  and  $B$  are independent, we have  $f_{G,A}(G,A) = f_B(GA) f_A(A)$ .

Gupta and Kabe [13], for example, compute the density of  $G$  from the joint density  $g(A,B)$ , using the approach adopted by Phillips [14].

$$g(G) = K \int \left[ \left| \frac{\partial}{\partial F_{11}} \right|^p \phi(F_{11}, F_{22}) \right]_{F_{11} = -G^2 F_{22} G^2} dF_{22},$$

where in the block division of  $A$ ,  $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ , we

have:  $A_{11}(p \times p)$ ,  $A_{22}(q \times q)$  while  $\phi(F_{21}, F_{22})$  is the joint characteristic function of  $A_{21}$  et  $A_{22}$ .

$$\text{For } g(A,B) = K \exp \left\{ -\frac{\text{tr}(A+B)}{2} \right\} |A|^{\frac{n-p-1}{2}} |B|^{\frac{q-p-1}{2}},$$

they obtained  $g(G) = K |G|^{\frac{q-p-1}{2}} |I+G|^{\frac{n+q}{2}}$ , i.e.

$$G \sim \text{Beta}^{\text{II}} \left( \frac{q}{2}, \frac{n}{2} \right).$$

Similarly, for Type I ratio, we can also have 5 types of ratios, and we will consider  $U_s = (X+Y)^{-1/2} X(X+Y)^{-1/2T}$ , the symmetric form of the ratio  $X(X+Y)^{-1}$  for elements distribution.

2) In general, the density of the determinant of a random matrix  $|A|$ , can be obtained from its elements distribution  $f(\cdot)$  of the previous section, in some simple cases, by using following relation for differentials:  $d|A| = |A| \text{tr}(A^{-1}(dA))$  ([15], p. 150). In practice, frequently, we have to manipulate  $|A|$  directly, often using an orthogonal transformation, to arrive at a product of independent diagonal and off-diagonal elements.

The determinants of the above different ratios  $Z_i, i = 1, \dots, 5$ , have univariate distributions that are identical, however, and they are of much interest since they will determine the null distribution useful in some statistical inference procedures.

3) Latent roots distributions for ratios, and the distributions of some associated statistics, remain very complicated, and in this article, we just mention some of their basic properties. For the density of the latent roots  $\{l_1, \dots, l_p\}$ , we have, using the elements density  $f(\cdot)$ :

$$h(l_1, \dots, l_p) = \frac{\pi^{\frac{p^2}{2}}}{\Gamma_p(p/2)} \prod_{i < j} (l_i - l_j) \int_{O(p)} f(HLH^T) (dH),$$

$$(l_1 > \dots > l_p)$$

where  $O(p)$  is the orthogonal group,  $H$  is an orthogonal ( $p \times p$ ) matrix,  $(dH)$  is the Haar invariant measure on  $O(p)$ , and  $L = \text{diag}(l_1, \dots, l_p)$  ([16], p. 105).

### 3.3. Two Kinds of Beta, of Wilks's Statistic, and of Latent Roots Distributions

We examine here the elements, determinant and latent roots distributions of a random matrix called the beta matrix variate, the homologous of the standard univariate beta.

First, the Wishart distribution, the matrix generalization of the chi-square distribution, played a critical part in the development of multivariate statistics.  $V \sim W_p(n, C)$ , is called a Wishart Matrix with parameters  $n$  and  $C$  if its density is:

$$f(V) = \text{etr}(-C^{-1}V/2) |V|^{\frac{n-p-1}{2}} \left/ \left[ 2^{np/2} |C|^{n/2} \Gamma_p(n/2) \right] \right. \quad (1)$$

$V > 0, n \geq p$ . It is a particular case of the Gamma matrix variate  $W \sim \text{Gap}(\alpha, C)$ , with matrix density:

$$f(\mathbf{W}) = \text{etr}(-\mathbf{C}^{-1}\mathbf{W}) |\mathbf{W}|^{\alpha - \frac{p+1}{2}} / \left[ |\mathbf{C}|^\alpha \Gamma_p(\alpha) \right] \quad (2)$$

with  $\mathbf{W} > 0, \mathbf{C}(p \times p) > 0$  and  $\alpha > (p-1)/2$

Hence, If  $\mathbf{S} \sim \mathbf{W}_p(n, \Sigma)$ , then  $\mathbf{S} \sim \mathbf{Ga}_p(n/2, \Sigma/2)$ .

**THEOREM 1:** Let  $\mathbf{A} \sim \mathbf{W}_p(m_A, \Sigma)$  et  $\mathbf{B} \sim \mathbf{W}_p(m_B, \Sigma)$ , with  $\mathbf{A}$  and  $\mathbf{B}$  independent ( $p \times p$ ) positive definite symmetric matrices, and  $m_A$  and  $m_B$  integers, with  $m_A, m_B > p$ . Then, for the three above-mentioned distributions, we have:

### 3.3.1. Elements Distribution

1) The ratio  $\mathbf{U} = (\mathbf{A} + \mathbf{B})^{-1/2} \mathbf{A} (\mathbf{A} + \mathbf{B})^{-1/2}$  has the matrix beta distribution  $\text{beta}_p^I(m_A/2, m_B/2)$ , with density given by (3).

2) Similarly, the ratio  $\mathbf{V} = \mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2}$  has a  $\text{beta}_p^{II}(m_A/2, m_B/2)$  distribution if  $\Sigma = \mathbf{I}_p$  and  $\mathbf{B}^{1/2}$  is symmetric. Its density is given by (4).

### 3.3.2. Determinant Distribution

1) For  $\mathbf{U} \sim \text{beta}_p^I(m_A, m_B)$ , its determinant  $|\mathbf{U}|$  has Wilks's distribution of the first type, denoted by  $|\mathbf{U}| \sim \Lambda^I(2(m_A + m_B), p, 2m_B)$  (to follow the notation in Kshirsagar [17], expressed as a product of independent  $\text{beta}^I$ , and its density, is given by (5).

2) For  $\mathbf{V} \sim \text{beta}_p^{II}(m_A, m_B)$ , its determinant  $|\mathbf{V}|$  has Wilks's distribution of the second type, denoted by  $|\mathbf{V}| \sim \Lambda^{II}(2(m_A + m_B), p, 2m_B)$ , expressed as a product of  $p$  independent univariate beta primes  $\text{beta}^{II}$ , and its density is given by (6).

### 3.3.3. Latent Roots Distribution

1) The latent roots of  $\mathbf{U}$ : The null-density of the latent roots  $\{f_1, \dots, f_p\}$  of  $\mathbf{U} = (\mathbf{A} + \mathbf{B})^{-1/2} \mathbf{A} (\mathbf{A} + \mathbf{B})^{-1/2}$  or  $\mathbf{A} (\mathbf{A} + \mathbf{B})^{-1}$  is:

$$f(l_i) = \frac{\pi^{p/2}}{\Gamma_p(p/2)} \frac{C(p, n_1) C(p, n_2)}{C(p, n_1 + n_2)} \cdot \left[ \prod_{i=1}^p f_i^{\frac{n_1 - p - 1}{2}} (1 - f_i)^{\frac{n_2 - p - 1}{2}} \right] \prod_{i < j} (f_i - f_j)$$

defined in the sector  $f_1 > f_2 > \dots > f_p > 0$ , with

$$C(p, ni) = \left[ 2^{\frac{pni}{2}} \Gamma_p \left( \frac{n_i}{2} \right) \right]^{-1},$$

where  $\Gamma_p(v) = \pi^p p(p-1)/4 \prod_{j=1}^p \Gamma \left( v - \frac{j-1}{2} \right)$ , (If  $n_1 < p$

but  $n_2 \geq p$ , we can make the changes

$(p, n_1, n_2) \rightarrow (n_1, p, n_1 + n_2 - p)$  to obtain the right expression).

2) The latent roots of  $\mathbf{V}$ : For  $\mathbf{G} = \mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2}$  or  $\mathbf{A} \mathbf{B}^{-1}$ , the roots  $\{l_1, \dots, l_p\}$  have density:

$$f(l_i) = \frac{\pi^{p/2}}{\Gamma_p(p/2)} \frac{C(p, n_1) C(p, n_2)}{C(p, n_1 + n_2)} \cdot \left[ \prod_{i=1}^p l_i^{p+1-n_1} (1+l_i)^{-\frac{n_1+n_2}{2}} \right] \prod_{i < j} (l_i - l_j)$$

We have also:  $l_i = f_i / (1 - f_i)$ .

**PROOF:** The proofs of part 1) and part 3) are found in most textbooks in multivariate analysis. For part 2, see [6]. **QED.**

The following explicative notes provide more details on the above results.

### 3.4. Explicative Notes

**1) On Elements Distributions:** The positive definite symmetric random matrix  $\mathbf{U}$  has a beta distribution of the first kind,  $\mathbf{U} \sim \text{beta}_p^I(a, b)$  if its density is of the form:

$$f(\mathbf{U}) = \frac{|\mathbf{U}|^{a-\frac{1}{2}(p+1)} |\mathbf{I}_p - \mathbf{U}|^{b-\frac{1}{2}(p+1)}}{\beta_p(a, b)}, \quad 0 < \mathbf{U} < \mathbf{I}_p, \quad (3)$$

It has a beta of the second kind distribution, denoted by  $\mathbf{V} \sim \text{beta}_p^{II}(a, b)$  if its density is of the form:

$$f(\mathbf{V}) = \frac{|\mathbf{V}|^{\frac{n-p+1}{2}}}{\beta_p(a, b) |\mathbf{I}_p + \mathbf{V}|^{a+b}}, \quad 0 < \mathbf{V} \quad (4)$$

where  $a, b > (p+1)/2$ , are positive real numbers, and  $\beta_p(a, b)$  is the beta function in  $R^p$ , i.e.

$$\beta_p(a, b) = \frac{\Gamma_p(a) \Gamma_p(b)}{\Gamma_p(a+b)}, \quad \text{with}$$

$$\Gamma_p(a) = \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^p \Gamma \left( a - \frac{i-1}{2} \right).$$

The transformations from  $\mathbf{U} \sim \text{beta}_p^I(a, b)$  to  $\mathbf{V} \sim \text{beta}_p^{II}(a, b)$  and vice-versa are simple ones. Also, similarly to the univariate case, where the beta prime is also called the gamma-gamma distribution, frequently encountered in Bayesian Statistics [11],  $\mathbf{V} \sim \text{beta}_p^{II}(a, b)$  can also be obtained as the continuous mixture of two Wishart densities, in the sense of  $\mathbf{V} \sim \mathbf{W}_p(n, \Omega)$ , with  $\Omega \sim \mathbf{W}_p(n_0, \Omega_0)$ . Also, for  $\mathbf{U}$  and  $\mathbf{V}$  above,  $\mathbf{I}_p - \mathbf{U} \sim \text{beta}_p^I(b, a)$  and  $\mathbf{V}^{-1} \sim \text{beta}_p^{II}(b, a)$ .

2) For the general case where  $\Sigma \neq \alpha \mathbf{I}_p$ ,  $\mathbf{V}$  is not necessarily  $\text{beta}_p^{II}$ , as pointed out by Olkin and Rubin [18]. Several other reasons, such as its dependency on  $\Sigma$  and on  $(\mathbf{A} + \mathbf{B})$ , makes  $\mathbf{V}$  difficult to use, and Perlman [19] suggested using

$\mathbf{V}^* = (\mathbf{A} + \mathbf{B})^{-1/2} \mathbf{B}^{-1} \mathbf{A} (\mathbf{A} + \mathbf{B})^{1/2} \sim \text{beta}_p^{II}(m_A/2, m_B/2)$  which does not have these weaknesses. We will use this definition as the matrix ratio type II of  $\mathbf{A}$  and  $\mathbf{B}$  when considering its elements distribution.

**2) On Determinant Distributions:** a) For

$U \sim \text{beta}_p^I(m_A, m_B)$  its determinant  $|U|$  has Wilks's distribution of the first type, denoted by  $|U| \sim \Lambda^I(2(m_A + m_B), p, 2m_B)$  (to follow the notation in [17]), expressed as a product of  $p$  univariate betas of the first kind, and this expression has been treated fully in [6]. This result is obtained via a transformation and by considering the elements on the diagonal.

**PROPOSITION 1:** For integers  $n, p$  and  $q$ , with  $n \geq q \geq p$ ,

a) The density of the ratio  $|U| \sim \Lambda^I(n, p, q)$ , is: (see (5)).

$$\text{where } K = \prod_{i=1}^p \frac{\Gamma\left(\frac{n+1-i}{2}\right)}{\Gamma\left(\frac{n-q+1-i}{2}\right)}.$$

b) For  $V \sim \text{beta}_p^{II}(m_A, m_B)$ , the latent roots of  $AB^{-1}$ ,  $B^{-1}A$ , and  $B^{-1/2}AB^{-1/2}$  are the same, and so are the three determinants, and  $|V|$  can be expressed as a product of  $p$  independent univariate beta primes, which, in turn, can be expressed as Meijer's  $G$ -functions, i.e.  $|V| \sim \Lambda^{II}(2(m_A + m_B), p, 2m_B)$ , or the Wilks's distribution of the second type. Hence, for the above ratio  $V$

$$|V| = \prod_{j=1}^p T_j, 0 < |V| < \infty, \text{ with}$$

$$T_j \sim \text{beta}^{II}\left(\frac{m_A - (j-1)}{2}, \frac{m_B - (j-1)}{2}\right),$$

where  $X \sim \text{beta}^{II}(\alpha, \beta)$  if it has as univariate density

$$f(x) = \frac{1}{B(\alpha, \beta)} \frac{x^{\alpha-1}}{(1+x)^{\alpha+\beta}} = \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \mathbf{H}_{1,1}^{1,1} \left[ x \middle| \begin{matrix} (-\beta, 1) \\ (\alpha-1, 1) \end{matrix} \right] \quad (6)$$

**PROPOSITION 2:** Wilks's statistic of the second kind,  $\Lambda^{II}(n, p, q)$ , has as density (see (7)).

$$\text{with } A = \prod_{j=1}^p \frac{1}{\Gamma\left(\frac{n-q}{2} - \frac{(j-1)}{2}\right) \Gamma\left(\frac{q}{2} - \frac{(j-1)}{2}\right)}.$$

In the general case  $n, p, q$  can have non-integral values and we can also have the case  $q < p$ .

**PROOF:** See (6).

**REMARK:** The cdf of  $Y$  is expressible in closed form, using the hypergeometric function  ${}_2F_1$  of matrix argument. ([12], p. 166) and the moments are

$$E(|Y|^h) = \frac{\Gamma_p(a+h)\Gamma_p(b-h)}{\Gamma_p(a)\Gamma_p(b)}, \text{ for}$$

$$-a + (p-1)/2 \leq h \leq b - (p-1)/2$$

**3) On Latent Roots distributions:**

The distribution of the latent roots  $\{f_1, f_2, \dots, f_p\}$  of  $X \sim \text{beta}_p^I(a, b)$  was made almost at the same time by five distinguished statisticians, as we all know. It is sometimes referred to as the generalized beta, but the marginal distributions of some roots might not be univariate beta. Although they are difficult to handle, partially due to their domain of definition as a sector in  $R^p$ , their associations with the Selberg integral, as presented in [20], has permitted to derive several important results. A similar expression applies for the latent roots of  $Y \sim \text{beta}_p^{II}(a, b)$ .

### 4. Extensions and Generalizations

From the basic results above, extensions can be made into several directions and various applications can be found. We will consider here elements and determinant distributions only. Let us recall that for univariate distributions there are several relations between the beta and the Dirichlet, and these relations can also be established for the matrix variate distribution.

$$h(u) = \begin{cases} K \cdot \mathbf{G}_{p,p}^{p,p} \left[ u \middle| \begin{matrix} \frac{n}{2} - 1, \frac{n-1}{2} - 1, \dots, \frac{n-(p-1)}{2} - 1 \\ \frac{n-q}{2} - 1, \frac{n-q-1}{2} - 1, \dots, \frac{n-q-(p-1)}{2} - 1 \end{matrix} \right], u > 0 \\ 0, u < 0 \end{cases} \quad (5)$$

$$f(x) = A \cdot \mathbf{G}_{p,p}^{p,p} \left[ x \middle| \begin{matrix} -\frac{q}{2}, -\left(\frac{q}{2} - \frac{1}{2}\right), \dots, -\left(\frac{q}{2} - \frac{p-1}{2}\right) \\ \frac{n-q}{2} - 1, \frac{n-q}{2} - \frac{3}{2}, \dots, \frac{n-q}{2} - \frac{p+1}{2} \end{matrix} \right] \quad (7)$$

4.1. Extensions

4.1.1. To Several Matrices

There is a wealth of relationships between the matrix variate Dirichlet distributions and its components [12], but because of space limitation, only a few can be presented here. The proof of the following results can be found in [18] and [19], where the question of independence between individual ratios and the sum of all matrices is discussed.

1) In Matrix-variate statistics, if  $\{S_0, S_1, \dots, S_k\}$  are  $(k + 1)$  independent  $(p \times p)$  random Wishart matrices,  $S_i \sim W_p(n_i, \Sigma)$ , then, the matrix variates  $\{U_i\}_{i=1}^k$ , defined “cumulatively” by:

$$\begin{aligned} U_1 &= (S_0 + S_1)^{-1/2} S_1 (S_0 + S_1)^{-1/2}, \dots, \\ U_j &= (S_0 + \dots + S_j)^{-1/2} S_j (S_0 + \dots + S_j)^{-1/2}, \dots, \\ U_k &= (S_0 + \dots + S_k)^{-1/2} S_k (S_0 + \dots + S_k)^{-1/2}, \\ U_0 &= S_0 + S_1 + \dots + S_k \end{aligned}$$

are mutually independent, with  $U_j, j = 1, \dots, k$ , having a matrix variate beta distrib.  $beta_p^I(n_j, n_0 + \dots + n_{j-1})$ .

2) Similarly, as suggested by Perlman [19], defining,

$$V_j = (S_0 + \dots + S_j)^{-1/2} [S_j (S_0 + \dots + S_{j-1})^{-1}] (S_0 + \dots + S_j)^{1/2}$$

then  $V_j \sim beta_p^{II}(n_j, n_0 + \dots + n_{j-1})$  and are mutually independent. Concerning their determinants, we have:

**THEOREM 2:** Let  $\{S_0, S_1, \dots, S_k\}$  be  $(k + 1)$  indep  $(p \times p)$  random Wishart matrices and  $U_j$  and  $V_j$

defined as above. Then the two products  $\Omega_j = \prod_{i=1}^j U_i$

and  $\Omega_j = \prod_{i=1}^j V_i, 1 \leq j \leq k$ , as well as any ratio

$U_j/U_k, j \neq k$  and  $V_j/V_k, j \neq k$ , have the densities of their determinants expressed in closed forms in terms of Meijer **G**-functions.

**PROOF:** We have  $|U_j| \sim \Lambda^I\left(2 \sum_{k=0}^j n_k, p, 2 \sum_{k=0}^{j-1} n_k\right)$ ,

$1 \leq j \leq k, |V_j| \sim \Lambda^{II}\left(2 \sum_{k=0}^j n_k, p, 2 \sum_{k=0}^{j-1} n_k\right)$ , and the result is

immediate from Section 2 of Theorem 1 and the expression of the **G**-function for products and ratios of independent **G**-function distributions given in the Appendix. Exact expressions are not given here to save space but are available upon request.

**QED.**

Also, considering the sum  $T = \sum_{j=0}^k S_j$ , if we take:

$$W_0 = T^{-1/2} S_0 T^{-1/2}, W_j = T^{-1/2} S_j T^{-1/2} \text{ and}$$

$T_j = W_0^{-1/2} W_j W_0^{-1/2}, 1 \leq j \leq k$ , then

1) the matrix vector  $(W_1, \dots, W_k)$  has the matrix variate Dirichlet of type I distribution, i.e.

$(W_1, \dots, W_k) \sim Dir^I(\{n_j, n_0\})$ , with density:

$$f(W_1, \dots, W_k) = c \prod_{j=1}^k |W_j|^{\frac{n_j-p-1}{2}} \left| I - \sum_{j=1}^k W_j \right|^{\frac{n_0-p-1}{2}}, W_j > 0,$$

$$I - \sum_{j=1}^k W_j > 0, \text{ where } c = \frac{\prod_{i=0}^k C(p, n_i)}{C(p, n)}, \text{ with } n = \sum_{j=0}^k n_j$$

$$\text{and } C(p, n) = \left[ 2^{\frac{pn}{2}} \pi^{p(p-1)/2} \prod_{j=1}^p \Gamma\left(\frac{n-j+1}{2}\right) \right]^{-1}.$$

and  $W_j \sim beta_p^I(n_j/2, (n-n_j)/2)$ .

2) the matrix vector  $(T_1, \dots, T_k)$  has the matrix variate Dirichlet of type II distribution, i.e.

$(T_1, \dots, T_k) \sim Dir^{II}(\{n_j, n_0\})$ , with density:

$$f(T_1, \dots, T_k) = c \frac{\prod_{j=1}^k |T_j|^{\frac{n_j-p-1}{2}}}{\prod_{j=1}^k \left| I + \sum_{j=1}^k T_j \right|^{\frac{n}{2}}}, T_j > 0$$

and  $T_i \sim beta_p^{II}(n_i/2, (n-n_i)/2)$ .

**THEOREM 3:** Under the same hypothesis as Theorem 2 the densities of the determinants of  $R = |W_i/W_j|, i \neq j$  and of  $R^* = |T_i/T_j|$ , can be obtained in closed form in terms of **G**-functions.

**PROOF:** We have  $|W_j| \sim \Lambda^I(n_j + n_0, p, n_0)$  and  $|T_j| \sim \Lambda^{II}(n_j + n_0, p, n_0), 1 \leq j \leq k$  and the conclusion is immediate from Section 2 of Theorem 1.

**QED.**

4.1.2. To the Matrix Variate Gamma Distribution

In the preceding sections we started with the Wishart distribution. However, it is a more general to consider the Gamma matrix variate distribution,  $W \sim Ga(\alpha, C)$ .

Here, we know that  $\frac{|W|}{|C|} \sim \prod_{i=1}^p X_i$ , with independent

$X_i \sim Ga\left(\alpha - \frac{i-1}{2}, 1\right)$ . Hence, the density of  $|W|$  is :

$$f(w) = |C| K \cdot G_{0,p}^p \left( w \left| \alpha - 1, \alpha - \frac{3}{2}, \dots, \alpha - \frac{p+1}{2} \right. \right), w > 0 \tag{8}$$

$$\text{with } K = \left[ \prod_{i=1}^p \Gamma\left(\alpha - \frac{i-1}{2}\right) \right]^{-1}.$$

Although for two independent Matrix Gammas,  $W_1$  and  $W_2$ , their symmetric ratio of the second type,  $W_2^{-1/2} W_1 W_2^{-1/2}$ , suffers from the same definition diffi-

culty as with the Wishart, the same recommendations made by Perlman [19] can be implemented to obtain the well-defined elements distribution of the Generalized  $F$  Matrix variate  $W \sim GF_p(\alpha_1, \alpha_2)$ . The matrix variate  $GF_p(a, b)$ -distribution, a scaled form of the  $beta_p^{\text{II}}(\alpha, \beta)$ , is encountered in multivariate regression, just like its univariate counterpart. But we have, for the ratio of the two determinants

$$|W| = |W_1|/|W_2| \cong \frac{\prod_{i=1}^p X_i}{\prod_{i=1}^p Y_i} = \prod_{i=1}^p T_i, \text{ with independent} \\ T_i \sim beta^{\text{II}}\left(\alpha_1 - \frac{i-1}{2}, \alpha_2 - \frac{i-1}{2}\right) \quad (9)$$

Hence, the distribution of  $|W|$ , the determinant of  $W \sim GF_p(\alpha_1, \alpha_2)$  has density:

$$f(z) = A \cdot \mathbf{G}_{p,p}^p \left[ x \left| \begin{array}{c} -\alpha_2, -\left(\alpha_2 - \frac{1}{2}\right), \dots, -\left(\alpha_2 - \frac{p-1}{2}\right) \\ \alpha_1 - 1, \alpha_1 - \frac{3}{2}, \dots, \alpha_1 - \frac{p+1}{2} \end{array} \right. \right] \quad (10)$$

with  $A = \left[ \prod_{j=1}^p \Gamma\left(\alpha_1 - \frac{j-1}{2}\right) \Gamma\left(\alpha_2 - \frac{j-1}{2}\right) \right]^{-1}$ , which is of the same form as (4).

The determinants of the product and ratio of the two  $GF_p$  variates, can now be computed, using the method given in the Appendix, extending the univariate case established for the generalized-F [11].

### 4.2. Further Matrix Ratios

There are various types of ratios encountered in the statistical literature, extending the univariate results of Pham-Gia and Turkkan [21] on divisions by the univariate gamma variable. We consider the following four matrix variates, which include all cases considered previously ( $\theta_1$  is a special case of  $\theta_4$ ):

Let  $\theta_1 \sim Wishart(W_p(n_i, \Sigma))$ , with integer  $n$ ,  $\theta_2 \sim Beta$  type I ( $beta_p^{\text{I}}(a, b)$ ),  $\theta_3 \sim Beta$  type II or  $GF_p(a, b)$  ( $Beta^{\text{II}}(a, b)$  or  $GF_p(a, b)$ ),  $a, b > (p+1)/2$ ,  $\theta_4 \sim Gamma(\alpha, C)$ , ( $Ga_p(\alpha, C), \alpha > 0, C > 0$ ).

The elements distributions of various products  $\theta_j^{1/2} \theta_i \theta_j^{1/2}$ , and ratios  $\theta_j^{-1/2} \theta_i \theta_j^{1/2}$  and  $(\theta_i + \theta_j)^{-1/2} \theta_i (\theta_i + \theta_j)^{-1/2}$ , for independent  $\theta_i, \theta_j$ ,  $1 \leq i, j \leq 4$ , can be carried out, but will usually lead to quite complex results. Some results when both  $\theta_i$  and  $\theta_j$  are  $\theta_2$ -matrices are obtained by Bekker, Roux and Pham-Gia [22].

However, for their determinants, we have:

$$|\theta_1| = |\Sigma| \sum_{i=1}^p \chi_{n-i+1}^2,$$

$$|\theta_2| \sim \Lambda^{\text{I}}(2(a+b), p, 2b) \sim \prod beta^{\text{I}}\left(a - \frac{j-1}{2}, b - \frac{j-1}{2}\right),$$

$$|\theta_3| \sim \Lambda^{\text{II}}(2(a+b), p, 2b) \sim \prod beta^{\text{II}}\left(a - \frac{j-1}{2}, b - \frac{j-1}{2}\right),$$

$$|\theta_4| \sim C \prod_{i=1}^p ga\left(\alpha - \frac{i-1}{2}, 1\right).$$

The expressions of the densities of  $|\theta_i|, 1 \leq i \leq 4$ , in terms of  $\mathbf{G}$ -functions are given in previous sections, except for  $|\theta_1|$  the density of which is given by (12) below.  $\Lambda^{\text{I}}$  and  $\Lambda^{\text{II}}$  here are generalized Wilks's variables, with  $a$  and  $b$  positive constants, instead of being integers.

**THEOREM 4:** For any of the above matrix variates

$\theta_j$   
a) The density of the determinant of any product  $\psi_{i,j}^{(1)} = |\theta_i \theta_j|$ , and ratio  $\psi_{i,j}^{(2)} = |\theta_i / \theta_j|$ ,  $0 \leq i, j \leq 4$ , in the above list, can be expressed in terms of Meijer  $\mathbf{G}$ -functions.

b) Furthermore, subject to the independence of all the factors involved, any product and ratio of different  $|\theta_i|, 1 \leq i \leq 4$ , and of different  $\psi_{i,j}^{(1)}$  and  $\psi_{i,j}^{(2)}$ ,  $0 \leq i, j \leq 4$ , can also have their densities expressed in terms of  $\mathbf{G}$ -functions.

**PROOF:** The proof is again based on the reproductive property of the Meijer  $\mathbf{G}$ -functions when product and ratio operations are performed, with the complex expressions for these operations presented in [6], and reproduced in the Appendix. Computation details can be provided by the authors upon request.

**QED.**

**REMARK:** Mathai [23] considered several types of integral equations associated with Wilks' work and provide solutions to these equations in a general theoretical context. The method presented here can be used to give a  $\mathbf{G}$ -function or  $\mathbf{H}$ -function form to these solutions, that can then be used for exact numerical computations.

## 5. Example and Application

We provide here an example using Theorem 4, with two graphs and also an engineering application.

### 5.1. Example

Let  $\theta_i \sim beta_p^{\text{II}}\left(\frac{n_i}{2}, p, \frac{n_i}{2}\right), i=1, 2$ . The densities of  $Y =$

$|\theta_1| |\theta_2|$  and  $R = \frac{|\theta_1|}{|\theta_2|}$  can be obtained in closed form as follows:

$$f(y) = A \cdot \mathbf{G}_{p_1 p_2}^{p_1 p_2} y \left[ \begin{array}{l} -\frac{q_1}{2}, -\left(\frac{q_1}{2} - \frac{1}{2}\right), \dots, -\left(\frac{q_1}{2} - \frac{p_1 - 1}{2}\right) \quad -\frac{q_2}{2}, -\left(\frac{q_2}{2} - \frac{1}{2}\right), \dots, -\left(\frac{q_2}{2} - \frac{p_2 - 1}{2}\right) \\ \frac{n_1 - q_1}{2} - 1, \frac{n_1 - q_1}{2} - \frac{3}{2}, \dots, \frac{n_1 - q_1}{2} - \frac{p_1 + 1}{2} \quad \frac{n_2 - q_2}{2} - 1, \frac{n_2 - q_2}{2} - \frac{3}{2}, \dots, \frac{n_2 - q_2}{2} - \frac{p_2 + 1}{2} \end{array} \right] \quad (11)$$

for  $y > 0$  with  $p = p_1 + p_2$ ,  $A = A_1 A_2$  and  $A_i = \prod_{j=1}^{p_i} \frac{1}{\Gamma\left(\frac{n_i - q_i}{2} - \frac{(j-1)}{2}\right) \Gamma\left(\frac{q_i}{2} - \frac{(j-1)}{2}\right)}$ ,  $i = 1, 2$ .

Similarly, the ratio  $R$  has density:

$$f(r) = A \cdot \mathbf{G}_{p_1 p_2}^{p_1 p_2} r \left[ \begin{array}{l} -\frac{q_1}{2}, -\left(\frac{q_1}{2} - \frac{1}{2}\right), \dots, -\left(\frac{q_1}{2} - \frac{p_1 - 1}{2}\right) \quad -\frac{n_2 - q_2}{2}, -\frac{n_2 - q_2}{2} + \frac{1}{2}, \dots, -\frac{n_2 - q_2}{2} + \frac{p_2 - 1}{2} \\ \frac{n_1 - q_1}{2} - 1, \frac{n_1 - q_1}{2} - \frac{3}{2}, \dots, \frac{n_1 - q_1}{2} - \frac{p_1 + 1}{2} \quad \left(\frac{q_2}{2} - 1\right), \left(\frac{q_2}{2} - \frac{3}{2}\right), \dots, \left(\frac{q_2}{2} - \frac{p_2 + 1}{2}\right) \end{array} \right] \quad (12)$$

for  $r > 0$ .

**PROOF:** Using (7) above, we can derive (11) and (12) by applying the approach presented in the Appendix.

**QED.**

**Figure 1** and **Figure 2** give, respectively, these two densities, for  $p_1 = p_2 = 4, q_1 = 9, q_2 = 12, n_1 = 13$  and  $n_2 = 18$ , using the MAPLE software.

### 5.2. Application to Multivariate Process Control

1) **Ratios of Variances:** The generalized variance is treated in [24]. We consider  $X \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  and a random sample  $\{X_1, \dots, X_N\}$  of  $X$ . We define

$A = \sum_{\alpha=1}^N (X_\alpha - \bar{X})(X_\alpha - \bar{X})^T$ , the sample sum of squares

and products matrix, and the sample covariance matrix  $S = A/(n-1)$ . We have:

$$|A|/|\boldsymbol{\Sigma}| \cong \chi_n^2 \cdots \chi_{n-p+1}^2 \quad (13)$$

where the  $\chi_j^2$  variables, with  $j$  degrees of freedom,  $n-p+1 \leq j \leq n$  are independent.

This is also the distribution of  $n^p \frac{|S|}{|\boldsymbol{\Sigma}|}$ .

Now let  $S_1$  and  $S_2$  be two independent sample covariance matrices of sizes  $n_1$  and  $n_2$  respectively. From (1), we have  $S_i \sim W_p(n_i - 1, \boldsymbol{\Sigma}_i / (n_i - 1))$ .

a) The sample generalized variance

$Y_i = (n_i - 1)^{p_i} |S_i| / |\boldsymbol{\Sigma}_i|$  has density:

$$h(y_i) = \left\{ \begin{array}{l} \frac{1}{2^p} \left( \prod_{j=0}^{p-1} \frac{1}{\Gamma\left(\frac{n_i - j}{2}\right)} \right) \mathbf{G}_{0 p}^{p 0} \left[ \frac{y_i}{2^p} \left| \frac{(n-2)}{2}, \frac{(n-3)}{2}, \dots, \frac{(n-(p+1))}{2} \right|, y_i > 0 \right. \\ \left. 0, \quad y_i \leq 0, i = 1, 2 \right] \end{array} \right\} \quad (14)$$

2) Now, the two types of ratios  $V = S_2^{-1/2} S_1 S_2^{-1/2}$  and  $U = (S_1 + S_2)^{-1/2} S_1 (S_1 + S_2)^{-1/2}$ , can be determined.

We have  $U \sim \text{Beta}^I((n_1 - 1)/2, (n_2 - 1)/2)$  and  $|U| \sim \Lambda^I(n_1 + n_2 - 2, p, n_2 - 1)$ .  $V$  has a density which is not necessarily  $\text{Beta}^{II}((n_1 - 1)/2, (n_2 - 1)/2)$ , but we have:  $|V| \sim \Lambda^{II}(n_1 + n_2 - 2, p, n_2 - 1)$ .

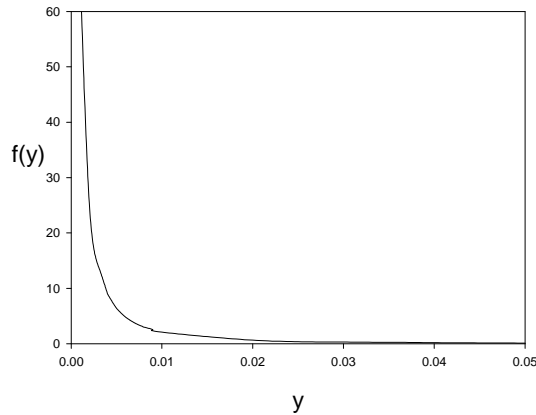
2) **Application:** In an industrial environment we wish to monitor the variations of a normal process

$X \sim N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ , using the variations of the ratio of two random sample covariance matrices taken from that environment, against the fluctuations of its control environment  $Y \sim N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ , represented by a similar ratio.

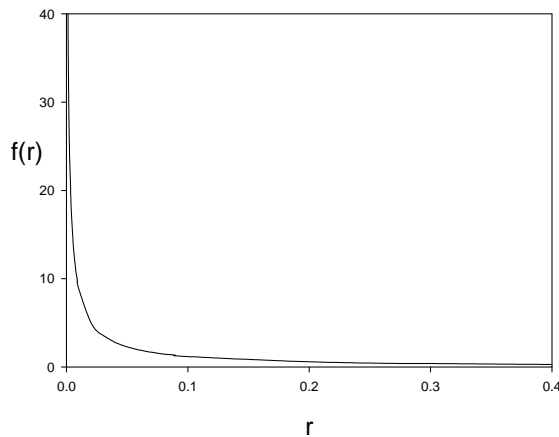
For clarity, we will proceed in several steps:

a) Let  $\theta_x = |S_1|/|S_2|$ , with  $|S_1|$  and  $|S_2|$  being two random samples of sizes  $n_1$  and  $n_2$  from  $N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$  and, similarly,  $\theta_y = |S_3|/|S_4|$ , with  $S_3$  and  $S_4$  being two random samples of sizes  $n_3$  and  $n_4$





**Figure 1.** Density of  $X_1/X_2$ , with  $X_1 \sim \Lambda^{\text{II}}(13,4,9), X_2 \sim \Lambda^{\text{II}}(18,4,12)$ .



**Figure 2.** Density of  $X_1/X_2, X_1 \sim \Lambda^{\text{II}}(13,4,9), X_2 \sim \Lambda^{\text{II}}(18,4,12)$ .

from  $N_p(\mu, \Sigma)$ . Hence,

$$\eta = \frac{\theta_y}{\theta_x} = \frac{|S_3|/|S_4|}{|S_1|/|S_2|} = \frac{\Lambda^{\text{II}}(n_3 + n_4 - 2, p, n_4 - 1)}{\Lambda^{\text{II}}(n_1 + n_2 - 2, p, n_2 - 1)}$$

$$= \frac{\prod_{i=1}^p \text{beta}^{\text{II}}\left(\frac{n_3 - 1}{2} - \frac{i - 1}{2}, \frac{n_4 - 1}{2} - \frac{i - 1}{2}\right)}{\prod_{i=1}^p \text{beta}^{\text{II}}\left(\frac{n_1 - 1}{2} - \frac{i - 1}{2}, \frac{n_2 - 1}{2} - \frac{i - 1}{2}\right)}$$

b) For the denominator, its density is given by (7):

$$g(x) = \prod_{j=1}^p \frac{1}{\Gamma\left(\frac{n_1 - j}{2}\right)\Gamma\left(\frac{n_2 - j}{2}\right)}$$

$$\cdot \mathbf{G}_{p \ p}^p \left( r \left| \begin{matrix} -\frac{n_2 - 1}{2}, -\frac{(n_2 - 2)}{2}, \dots, -\frac{(n_2 - p)}{2} \\ \frac{n_1 - 3}{2}, \frac{n_1 - 4}{2}, \dots, \frac{n_1 - (p + 2)}{2} \end{matrix} \right. \right)$$

and similarly for the numerator.

c) Applying the ratio rule in the Appendix for these two independent **G**-function distributions, we have the density of  $\eta$ .

**QED.**

### 6. Conclusions

As shown in this article, for several types of random matrices, the moments of which can be expressed in terms of gamma functions, Meijer G-function provides a powerful tool to derive, and numerically compute, the densities of the determinants of products and ratios of these matrices. Multivariate hypothesis testing based on determinants can now be accurately carried out since the expressions of null distributions are, now, not based on asymptotic considerations.

### 7. References

- [1] K. C. S. Pillai, "Distributions of Characteristic Roots in Multivariate Analysis, Part I: Null Distributions," *Communications in Statistics—Theory and Methods*, Vol. 4, No. 2, 1976, pp. 157-184.
- [2] K. C. S. Pillai, "Distributions of Characteristic Roots in Multivariate Analysis, Part II: Non-Null Distributions," *Communications in Statistics—Theory and Methods*, Vol. 5, No. 21, 1977, pp. 1-62.
- [3] P. Koev and A. Edelman, "The Effective Evaluation of the Hypergeometric Function of a Matrix Argument," *Mathematics of Computation*, Vol. 75, 2006, pp. 833-846. [doi:10.1090/S0025-5718-06-01824-2](https://doi.org/10.1090/S0025-5718-06-01824-2)
- [4] A. M. Mathai and R. K. Saxena, "Generalized Hypergeometric Functions with Applications in Statistics and Physical Sciences," *Lecture Notes in Mathematics*, Vol. 348, Springer-Verlag, New York, 1973.
- [5] M. Springer, "The Algebra of Random Variables," Wiley, New York, 1984.
- [6] T. Pham-Gia, "Exact Distribution of the Generalized Wilks's Distribution and Applications," *Journal of Multivariate Analysis*, 2008, 1999, pp. 1698-1716.
- [7] V. Adamchik, "The Evaluation of Integrals of Bessel Functions via G-Function Identities," *Journal of Computational and Applied Mathematics*, Vol. 64, No. 3, 1995, pp. 283-290. [doi:10.1016/0377-0427\(95\)00153-0](https://doi.org/10.1016/0377-0427(95)00153-0)
- [8] T. Pham-Gia, T. N. Turkkan and E. Marchand, "Distribution of the Ratio of Normal Variables," *Communications in Statistics—Theory and Methods*, Vol. 35, 2006, pp. 1569-1591.
- [9] T. Pham-Gia and N. Turkkan, "Distributions of the Ratios of Independent Beta Variables and Applications," *Communications in Statistics—Theory and Methods*, Vol. 29, No. 12, 2000, pp. 2693-2715. [doi:10.1080/03610920008832632](https://doi.org/10.1080/03610920008832632)
- [10] T. Pham-Gia and N. Turkkan, "The Product and Quotient

- of General Beta Distributions,” *Statistical Papers*, Vol. 43, No. 4, 2002, pp. 537-550.  
[doi:10.1007/s00362-002-0122-y](https://doi.org/10.1007/s00362-002-0122-y)
- [11] T. Pham-Gia and N. Turkkan, “Operations on the Generalized F-Variables, and Applications,” *Statistics*, Vol. 36, No. 3, 2002, pp. 195-209.  
[doi:10.1080/02331880212855](https://doi.org/10.1080/02331880212855)
- [12] A. K. Gupta and D. K. Nagar, “Matrix Variate Distributions,” Chapman and Hall/CRC, Boca Raton, 2000.
- [13] A. K. Gupta and D. G. Kabe, “The Distribution of Symmetric Matrix Quotients,” *Journal of Multivariate Analysis*, Vol. 87, No. 2, 2003, pp. 413-417.  
[doi:10.1016/S0047-259X\(03\)00046-0](https://doi.org/10.1016/S0047-259X(03)00046-0)
- [14] P. C. B. Phillips, “The Distribution of Matrix Quotients,” *Journal of Multivariate Analysis*, Vol. 16, No. 1, 1985, pp. 157-161.  
[doi:10.1016/0047-259X\(85\)90056-9](https://doi.org/10.1016/0047-259X(85)90056-9)
- [15] A. M. Mathai, “Jacobians of Matrix Transformations and Functions of Matrix Argument,” World Scientific, Singapore, 1997.
- [16] R. J. Muirhead, “Aspects of Multivariate Statistical Theory,” Wiley, New York, 1982.  
[doi:10.1002/9780470316559](https://doi.org/10.1002/9780470316559)
- [17] A. Kshirsagar, “Multivariate Analysis,” Marcel Dekker, New York, 1972.
- [18] I. Olkin and H. Rubin, “Multivariate Beta Distributions and Independence Properties of the Wishart Distribution,” *Annals Mathematical Statistics*, Vol. 35, No.1, 1964, pp. 261-269. [doi:10.1214/aoms/1177703748](https://doi.org/10.1214/aoms/1177703748)
- [19] M. D. Perlman, “A Note on the Matrix-Variate F Distribution,” *Sankhya, Series A*, Vol. 39, 1977, pp. 290-298.
- [20] T. Pham-Gia, “The Multivariate Selberg Beta Distribution and Applications,” *Statistics*, Vol. 43, No. 1, 2009, pp. 65-79. [doi:10.1080/02331880802185372](https://doi.org/10.1080/02331880802185372)
- [21] T. Pham-Gia and N. Turkkan, “Distributions of Ratios of Random Variables from the Power-Quadratic Exponential family and Applications,” *Statistics*, Vol. 39, No. 4, 2005, pp. 355-372.
- [22] A. Bekker, J. J. J. Roux and T. Pham-Gia, “Operations on the Matrix Variate Beta Type I Variables and Applications,” Unpublished Manuscript, University of Pretoria, Pretoria, 2005.
- [23] A. M. Mathai, “Extensions of Wilks’ Integral Equations and Distributions of Test Statistics,” *Annals of the Institute of Statistical Mathematics*, Vol. 36, No. 2, 1984, pp. 271-288.
- [24] T. Pham-Gia and N. Turkkan, “Exact Expression of the Sample Generalized Variance and Applications,” *Statistical Papers*, Vol. 51, No. 4, 2010, pp. 931-945.

**Appendix**

The **H**-function  $H(x)$  is defined as follows:

$$\mathbf{H}_{p,q}^{m,r} \left[ x \left| \begin{matrix} (a_1, \alpha_1), \dots, (a_p, \alpha_p) \\ (b_1, \beta_1), \dots, (b_q, \beta_q) \end{matrix} \right. \right]$$

$$= \frac{1}{2\pi i} \int_L \frac{\prod_{j=1}^m \Gamma(b_j - \beta_j s) \prod_{j=1}^r \Gamma(1 - a_j + \alpha_j s)}{\prod_{j=m+1}^q \Gamma(1 - b_j + \beta_j s) \prod_{j=r+1}^p \Gamma(a_j - \alpha_j s)} x^s ds$$

*i.e.* it is

the integral along the complex contour  $L$  of a ratio of products of gamma functions. **H**-function has a relationship with the generalized hypergeometric function:  ${}_pF_{q-1}(\cdot)$ . An equivalent form is obtained by replacing  $s$  by  $(-s)$ , which shows that the integrand is, in fact, a Mellin-transform. Numerically, **H**-functions can be computed by using the residue theorem in complex analysis, together with Jordan's lemma, but because of the complexity of this operation, it is only quite recently that it is available on softwares, although, in the past, several authors have suggested their own versions in their published works.

In this Appendix, the densities of the product and ratio of two independent random variables, whose densities are expressed in **H**-functions, is obtained. For **G**-functions, we have  $\alpha_i = \beta_j = 1, \forall i, j$ , and since these values are not affected by the operations the following results remain valid.

**1) Product:** Let  $X_j, j=1, \dots, s$ , be  $s$  independent **H**-function random variables, each with pdf:

$$\left( (a_{1,1}, \alpha_{1,1}), \dots, (a_{1,\eta}, \alpha_{1,\eta}); (a_{2,1}, \alpha_{2,1}), \dots, (a_{2,r_2}, \alpha_{2,r_2}); \dots; (a_{s,1}, \alpha_{s,1}), \dots, (a_{s,r_s}, \alpha_{s,r_s}) \right) ::$$

$$\left( (a_{1,(\eta+1)}, \alpha_{1,(\eta+1)}), \dots, (a_{1,p_1}, \alpha_{1,p_1}); (a_{2,(r_2+1)}, \alpha_{2,(r_2+1)}), \dots, (a_{2,p_2}, \alpha_{2,p_2}); \dots; (a_{s,(r_s+1)}, \alpha_{s,(r_s+1)}), \dots, (a_{s,p_s}, \alpha_{s,p_s}) \right)$$

(8)

b) Similarly, the lower parameter sequence (l.s.), of total length  $Q$ , consists of  $s$  consecutive subsequences of the type  $(b_{1,1}, \beta_{1,1}), \dots, (b_{1,m_1}, \beta_{1,m_1})$ , followed by  $s$  consecutive

$$\left( (b_{1,1}, \beta_{1,1}), \dots, (b_{1,m_1}, \beta_{1,m_1}); (b_{2,1}, \beta_{2,1}), \dots, (b_{2,m_2}, \beta_{2,m_2}); \dots; (b_{s,1}, \beta_{s,1}), \dots, (b_{s,m_s}, \beta_{s,m_s}) \right) ::$$

$$\left( (b_{1,(m_1+1)}, \beta_{1,(m_1+1)}), \dots, (b_{1,q_1}, \beta_{1,q_1}); (b_{2,(m_2+1)}, \beta_{2,(m_2+1)}), \dots, (b_{2,q_2}, \beta_{2,q_2}); \dots; (b_{s,(m_s+1)}, \beta_{s,(m_s+1)}), \dots, (b_{s,q_s}, \beta_{s,q_s}) \right)$$

(9)

**2) Ratio:** For the ratio  $W = X_1/X_2$ , its density is:

$$f_W(w) = A \mathbf{H}_{p_1+q_2, q_1+p_2}^{m_1+r_2, m_2+r_1} \left[ \frac{c_1}{c_2} w \left| \begin{matrix} (u.s.) (a_{1,1}, \alpha_{1,1}), \dots, \dots, (1 - b_{2,q_2} - 2\beta_{2,q_2}, \beta_{2,q_2}) \\ (l.s.) (b_{1,1}, \beta_{1,1}), \dots, \dots, (1 - a_{2,q_2} - 2\alpha_{2,q_2}, \alpha_{2,q_2}) \end{matrix} \right. \right], w > 0$$

(10)

$$f_j(x_j) = k_j \mathbf{H}_{p_j, q_j}^{m_j, r_j} \left[ c_j x_j \left| \begin{matrix} (a_{j,1}, \alpha_{j,1}), \dots, (a_{j,p_j}, \alpha_{j,p_j}) \\ (b_{j,1}, \beta_{j,1}), \dots, (b_{j,q_j}, \beta_{j,q_j}) \end{matrix} \right. \right], x_j > 0$$

(6)

The product of these variables,  $P = \prod_{j=1}^s X_j$ , is also a **H**-Function Random variable, with density

$$f_p(y) = \left( \prod_{j=1}^s k_j \right) \mathbf{H}_{P, Q}^{M, R} \left[ \left( \prod_{j=1}^s c_j \right) y \left| \begin{matrix} (u.s.) (a_{1,1}, \alpha_{1,1}), \dots, (a_{s,p_s}, \alpha_{s,p_s}) \\ (l.s.) (b_{1,1}, \beta_{1,1}), \dots, (b_{s,q_s}, \beta_{s,q_s}) \end{matrix} \right. \right], y > 0$$

(7)

where  $M = \sum_{j=1}^s m_j, R = \sum_{j=1}^s r_j, P = \sum_{j=1}^s p_j, Q = \sum_{j=1}^s q_j$ , and

the two parameter sequences, (u.s.) and (l.s.), are as follows:

a) The upper sequence of parameters (u.s.), of total length  $P$ , consists of  $s$  consecutive subsequences of the type  $(a_{1,1}, \alpha_{1,1}) \dots (a_{1,\eta}, \alpha_{1,\eta})$ , followed by  $s$  consecutive subsequences of the type  $(a_{1,(\eta+1)}, \alpha_{1,(\eta+1)}) \dots (a_{1,p_1}, \alpha_{1,p_1})$ , *i.e.* we have:

secutive subsequences of the type:

$(b_{1,(m_1+1)}, \beta_{1,(m_1+1)}), \dots, (b_{1,q_1}, \beta_{1,q_1})$ , *i.e.* we have:

where  $A = \left( \frac{k_1 k_2}{c_2^2} \right)$ .

a) The upper sequence of parameters (u.s.), of total length  $p_1 + q_2$ , consists of the four consecutive subsequences:

$(a_{11}, \alpha_{11}), \dots, (a_{1r_1}, \alpha_{1r_1})$  of length  $r_1$   
 $(1 - b_{2,1} - 2\beta_{2,1}, \beta_{2,1}), \dots, (1 - b_{2,m_2} - 2\beta_{2,m_2}, \beta_{2,m_2})$  of length  $m_2$  (11)

$(a_{1,r_1+1}, \alpha_{1,r_1+1}), \dots, (a_{1,p_1}, \alpha_{1,p_1})$  of length  $p_1 - r_1$ , and  
 $(1 - b_{2,m_2+1} - 2\beta_{2,m_2+1}, \beta_{2,m_2+1}), \dots, (1 - b_{2,q_2} - 2\beta_{2,q_2}, \beta_{2,q_2})$  of length  $q_2 - m_2$ .

b) The lower sequence (l.s.), of total length  $p_2 + q_1$ , also has 4 subsequences of respective lengths

$m_1, r_2, q_1 - m_1$  and  $p_2 - r_2$ :

$$(b_{1,1}, \beta_{1,1}), \dots, (b_{1,m_1}, \beta_{1,m_1})$$

$$(1 - a_{2,1} - 2\alpha_{2,1}, \alpha_{2,1}), \dots, (1 - a_{2,r_2} - 2\alpha_{2,r_2}, \alpha_{2,r_2}), \quad (12)$$

$(b_{1,m_1+1}, \beta_{1,m_1+1}), \dots, (b_{1,q_1}, \beta_{1,q_1})$ , and

$$(1 - a_{2,r_2+1} - 2\alpha_{2,r_2+1}, \alpha_{2,r_2+1}), \dots, (1 - a_{2,p_2} - 2\alpha_{2,p_2}, \alpha_{2,p_2}).$$

**PROOF:** The proofs of the above results, based on the Mellin transform of a function  $f(x)$  defined on  $R^+$ , and its inverse Mellin transform, as defined previously, are quite involved, but can be found in Springer (1984, p. 214).

**QED.**