

QÜESTIÓ, vol. 26, 3, p. 429-441, 2002

IMPLEMENTACIÓN DEL CÁLCULO DE POLINOMIOS ZONALES Y APLICACIONES EN ANÁLISIS MULTIVARIANTE

J. RODRÍGUEZ-AVI
A. J. SÁEZ-CASTILLO
A. CONDE-SÁNCHEZ
Universidad de Jaén*

En este trabajo se describe la implementación de un algoritmo para el cálculo de polinomios zonales, así como dos aplicaciones explícitas de éstos en el ámbito del análisis multivariante. Concretamente, esta implementación permite obtener resultados de sumación aproximados para funciones hipergeométricas de argumento matricial que, a su vez, pueden utilizarse en la génesis de distribuciones multivariantes discretas con frecuencias simétricas. De igual forma, se pone en práctica un conocido resultado teórico que caracteriza la distribución de la menor raíz característica de una matriz aleatoria con distribución de Wishart.

Implementing calculus of zonal polynomials and applications in Multivariate Analysis

Palabras clave: Polinomios zonales, distribuciones discretas, distribución de Wishart, funciones hipergeométricas de argumento matricial

Clasificación AMS (MSC 2000): 60E05, 62E15

*Dep. de Estadística e I.O. Despacho 147, Ed. D-3, Paraje Las Lagunillas, Universidad de Jaén, 23071 Jaén, España. Telf: + 34 953 012.207; fax: + 34 953 012.222; email: jravi@ujaen.es.

–Recibido en diciembre de 2000.

–Aceptado en abril de 2002.

1. INTRODUCCIÓN

Los polinomios zonales son una extensión multivariante de las funciones potenciales. La familia de polinomios simétricos y homogéneos que constituyen ha sido ampliamente utilizada dentro de la Estadística Matemática clásica en la expresión de densidades y distribuciones de formas cuadráticas en poblaciones normales, para extender resultados conocidos de la estadística univariante a un ambiente multivariante (Muirhead, 1982). No obstante, no se conocen fórmulas generales para estos polinomios, ya que su definición, como vamos a ver a continuación, los caracteriza como autofunciones de un cierto operador diferencial.

En concreto, dada una matriz simétrica $X_{m \times m}$ con raíces características x_1, \dots, x_m y $\kappa = (k_1, \dots, k_m)$ una partición de k en no más de m partes, el polinomio zonal de X correspondiente a κ , denotado por $C_\kappa(X)$, es un polinomio simétrico homogéneo de grado k en las raíces características x_1, \dots, x_m que debe verificar las siguientes tres condiciones:

1. El término de mayor peso es, salvo constante, $x_1^{k_1} \dots x_m^{k_m}$.
2. Verifica la ecuación diferencial

$$(1) \quad \Delta_X C_\kappa(X) = \alpha_\kappa C_\kappa(X)$$

donde

$$\Delta_X = \sum_{i=1}^m x_i^2 \frac{\partial^2}{\partial x_i^2} + \sum_{i=1}^m \sum_{j=1, j \neq i}^m \frac{x_i^2}{x_i - x_j} \frac{\partial}{\partial x_i}$$

y

$$\alpha_\kappa = \sum_{i=1}^m k_i (k_i - i) + k(m - 1).$$

3. $(tr X)^k = (x_1 + \dots + x_m)^k = \sum_\kappa C_\kappa(X)$, donde por \sum_κ denotamos la suma en todas las particiones del grado k .

Alternativamente a esta última condición, que se utiliza para la normalización de los polinomios zonales, puede considerarse la evaluación de estos polinomios en un punto concreto. En este sentido

$$(2) \quad C_\kappa(I_m) = 2^{2k} k! \left(\frac{1}{2} \right)_\kappa \frac{\prod_{i < j}^p (2k_i - 2k_j - i + j)}{\prod_{i < j}^p (2k_i + p - i)!}$$

donde p es el número de partes no nulas de κ y, en general,

$$(3) \quad (a)_{\kappa} = (a)_{k_1} \left(a - \frac{1}{2}\right)_{k_2} \cdots \left(a - \frac{1}{2}(m-1)\right)_{k_m},$$

con $(a)_k = a(a+1)\cdots(a+k-1)$, $(a)_0 = 1$.

2. ALGORITMO DE CÁLCULO

Como se ha comentado, de la definición no puede deducirse una fórmula explícita, aunque sí un algoritmo de cálculo (James, 1968), que describimos seguidamente.

Sea $\kappa = (k_1, \dots, k_m)$ una partición del entero k . Entonces, el monomio simétrico de una matriz $X_{m \times m}$ con raíces características x_1, \dots, x_m correspondiente a κ se define como

$$M_{\kappa}(X) = x_1^{k_1} \cdots x_m^{k_m} + \text{términos simétricos.}$$

El algoritmo de James establece que los polinomios zonales son combinaciones lineales de monomios simétricos. Concretamente:

$$(4) \quad C_{\kappa}(X) = \sum_{\lambda \leq \kappa} c_{\kappa, \lambda} M_{\lambda}(X),$$

donde la sumatoria se hace sobre todas las particiones λ de k tales que $\lambda \leq \kappa$ y

$$(5) \quad c_{\kappa, \lambda} = \sum_{\lambda < \mu \leq \kappa} \frac{[(l_i + t) - (l_j - t)]}{\rho_{\kappa} - \rho_{\lambda}} c_{\kappa, \mu}.$$

En esta expresión,

$$(6) \quad \rho_{\kappa} = \sum_{i=1}^m k_i (k_i - i);$$

además, $\lambda = (l_1, \dots, l_m)$ y $\mu = (l_1, \dots, l_i + t, \dots, l_j - t, \dots, l_m)$, para $t = 1, \dots, l_j$, son tales que cuando las partes de la partición μ se organizan en orden descendente, μ es mayor que λ y menor o igual que κ en orden lexicográfico. La sumatoria en (5) es sobre todas las posibles μ , y una suma vacía se toma como cero.

Mediante este algoritmo se determinan todos los coeficientes $c_{\kappa, \lambda}$ salvo el de mayor peso; es decir, se determina el polinomio zonal salvo normalización. Desde el punto de vista computacional, el algoritmo puede inicializarse considerando $c_{\kappa, \kappa} = 1$ y calculando los coeficientes $c_{\kappa, \lambda}$ salvo constante multiplicativa, que se obtiene mediante (2).

La implementación del algoritmo se ha realizado empleando el programa MATLAB. Los correspondientes archivos fuente y las instrucciones para su ejecución no aparecen aquí por brevedad, aunque pueden encontrarse en la dirección

<http://www.ujaen.es/dep/estinv/pdi/ajsaez/software.html>.

3. APLICACIONES

En este apartado se presentan sendas aplicaciones en el ámbito de la estadística multivariante que pueden llevarse a cabo gracias al conocimiento explícito de un buen número de polinomios zonales calculados mediante la metodología expuesta en la sección anterior.

3.1. Aproximación de funciones hipergeométricas de argumento matricial y génesis de distribuciones multivariantes

Se define la función hipergeométrica de argumento matricial de parámetros a_1, \dots, a_p y b_1, \dots, b_q asociada a la matriz simétrica $X_{m \times m}$ como

$$(7) \quad {}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; X) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} \cdots (a_p)_{\kappa} C_{\kappa}(X)}{(b_1)_{\kappa} \cdots (b_q)_{\kappa} k!},$$

donde \sum_{κ} denota la suma sobre todas las particiones $\kappa = (k_1, \dots, k_m)$ de k y $C_{\kappa}(X)$ es el polinomio zonal de X correspondiente a κ . Ningún parámetro b_j del denominador puede ser cero o un medio de un entero menor o igual que $\frac{1}{2}(m-1)$, ya que, en caso contrario, alguno de los denominadores de la serie se anularía. De igual forma, si algún parámetro del numerador es un entero negativo, esto es, $a_i = -n$, entonces la función es un polinomio de grado mn .

La serie converge para todo X si $p \leq q$; converge para $\|X\| < 1$ si $p = q + 1$, donde por $\|X\|$ notamos el máximo de los valores absolutos de las raíces características de X ; y salvo que la serie tenga un número finito de términos no nulos, diverge para todo $X \neq 0$ si $p > q + 1$.

En el caso $m = 1$, la serie (7) se reduce la función hipergeométrica generalizada clásica

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; x) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_p)_k x^k}{(b_1)_k \cdots (b_q)_k k!}.$$

Podemos subrayar dos casos especiales de (7): los dados por la función ${}_0F_0(X) = \text{etr}(X)$, generalización de la serie exponencial, y la función ${}_1F_0(a; X) = \det(I_m - X)^{-a}$,

generalización de la serie binomial $(1-x)^{-a}$. Más allá de estos dos casos no existen expresiones generales de las funciones hipergeométricas de argumento matricial. Es por ello que para obtener resultados de sumación se suele acudir a aproximaciones numéricas mediante técnicas de análisis que proporcionan el valor de las funciones en algunos casos concretos.

Nosotros proponemos desarrollar estas funciones como series de monomios simétricos, utilizando la implementación del algoritmo de James. De esta forma, pueden obtenerse resultados aproximados de sumación de dichas funciones para un amplio rango de valores de los parámetros. Asimismo, este desarrollo en serie es el adecuado para utilizarlas como generatrices de distribuciones multivariantes discretas de probabilidad, distribuciones que cuentan con una interesante propiedad de simetría en sus frecuencias.

Tabla 1. Desarrollo truncado de la función ${}_2F_1(a, b; c; I_2)$.

a	b	c	${}_2F_1(a, b; c; I_2)$	$\sum_{k=0}^{20} \sum_{\rho} f_{\rho} M_{\rho}(I_2)$	Diferencia
2	2	10	3.0569	3.0547	0.0022
3	3	10	20.0909	19.0750	1.0159
5	5	25	13.8359	13.8299	0.006
4	5	20	15.7342	15.7046	0.0296
8	10	50	55.1096	55.0856	0.024
10	10	50	172.7901	172.2222	0.5679
12	12	65	260.1934	259.5012	0.6922
15	15	85	699.2147	695.7047	3.51

Concretamente, si en el desarrollo de la serie hipergeométrica dada en (7), expresamos los polinomios zonales en términos de los monomios simétricos, se tiene que

$$\begin{aligned}
 (8) \quad {}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; X) &= \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} \cdots (a_p)_{\kappa}}{(b_1)_{\kappa} \cdots (b_q)_{\kappa}} \frac{C_{\kappa}(X)}{k!} \\
 &= \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} \cdots (a_p)_{\kappa}}{(b_1)_{\kappa} \cdots (b_q)_{\kappa}} \frac{\sum_{\lambda \leq \kappa} c_{\kappa, \lambda} M_{\lambda}(X)}{k!} \\
 &= \sum_{k=0}^{\infty} \sum_{\kappa} \sum_{\lambda \leq \kappa} \frac{1}{k!} \frac{(a_1)_{\kappa} \cdots (a_p)_{\kappa}}{(b_1)_{\kappa} \cdots (b_q)_{\kappa}} c_{\kappa, \lambda} M_{\lambda}(X) \\
 &= \sum_{k=0}^{\infty} \sum_{\rho} f_{\rho} M_{\rho}(X),
 \end{aligned}$$

donde de nuevo la suma \sum_{ρ} es en todas las particiones ρ de cada entero k . Una vez calculados los coeficientes f_{ρ} , la serie puede aproximarse mediante sus sumas parciales.

En orden a considerar la bondad de estas aproximaciones, hay que tener en cuenta que la cola de la serie truncada en el grado r , contiene tan sólo términos en las raíces características de X de grado superior a r .

En la Tabla 1 se muestran resultados aproximados de sumación referidos a la función bivalente ${}_2F_1$, para distintos valores de los parámetros cuando $X = I_2$. Todos ellos se comparan con el valor exacto de la suma, conocido gracias al Teorema de Sumación de Gauss en su versión multivariante. En general, podemos establecer una cota del error, ϵ , cometido mediante la suma de los $M + 1$ primeros términos cuando $a_1, a_2, b_1 > 0, b_1 > a_1 + a_2$ y $\|X\| \leq 1$, dada por

$$\begin{aligned}
 \sum_{k=M+1}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} (a_2)_{\kappa}}{(b_1)_{\kappa}} \frac{C_{\kappa}(X)}{k!} &\leq \sum_{k=M+1}^{\infty} \sum_{\kappa} \frac{C_{\kappa}(X)}{k!} \\
 &= \sum_{k=M+1}^{\infty} \frac{\text{traza}(X)^k}{k!} \\
 &\leq \sum_{k=M+1}^{\infty} \frac{m^k}{k!} \\
 &= e^m - \sum_{k=0}^M \frac{m^k}{k!}
 \end{aligned}
 \tag{9}$$

Para una discusión más detallada de la metodología de este método de sumación puede verse Gutiérrez *et al.* (2000).

Por otra parte, las funciones hipergeométricas clásicas de argumento escalar, fundamentalmente la función de Gauss ${}_2F_1$ y sus extensiones y particularizaciones, pueden emplearse para generar las distribuciones discretas más usuales (Johnson, Kotz y Kemp, 1992; Gutiérrez y Rodríguez, 1997; Rodríguez *et al.*, 2001a). También se han introducido extensiones bivariantes de algunas funciones hipergeométricas de argumento escalar, como la F_1 o la F_3 , para generar distribuciones bivariantes discretas.

En este sentido, las funciones hipergeométricas de argumento matricial también pueden utilizarse para generar distribuciones discretas, en este caso multivariantes, que generalicen sus homólogas univariantes. Para ello tan sólo es necesario considerar el desarrollo en serie dado en (8), de manera que las probabilidades de estas distribuciones vienen dadas por las constantes f_{ρ} convenientemente normalizadas. La principal característica de estas distribuciones es que sus frecuencias son invariantes frente a permutaciones de las variables que la forman; es decir, si $X_{m \times 1}$ es un vector aleatorio generado por una función de este tipo, se verifica:

$$P[X = (i_1, \dots, i_m)] = P[X = \sigma(i_1, \dots, i_m)],
 \tag{10}$$

donde (i_1, \dots, i_m) es cualquier vector posible para X y σ es cualquier permutación. Esta propiedad hace que estas distribuciones sean modelos adecuados para fenómenos

multivariantes discretos cuyas componentes no son necesariamente independientes pero sí idénticamente distribuidas, de manera que las probabilidades de ocurrencia no dependan del orden de las variables. Así, se han utilizado con éxito, por ejemplo, en la modelización del número de ingresos diarios en dos boxes del Servicio de Urgencias del Hospital de San Agustín (Linares, España) (Rodríguez *et al*, 2001*b*). Otros ámbitos adecuados de aplicación aparecen, por ejemplo, en la modelización del número de fallos en varias máquinas indistinguibles en paralelo o el número de llamadas a varios servidores idénticos dispuestos también en paralelo.

En la Figura 1 se muestran ejemplos de poligonales de frecuencias generadas mediante funciones ${}_1F_1$ y ${}_2F_1$ bivariantes.

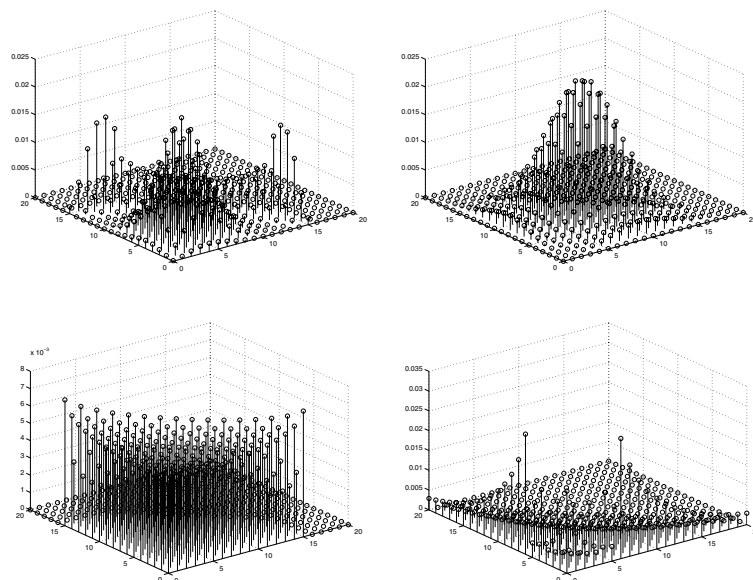


Figura 1. Poligonal de frecuencias para distribuciones generadas por funciones ${}_1F_1$ (a. y b.) y ${}_2F_1$ (c. y d.).

3.2. Distribución de la menor raíz característica de una matriz de Wishart en términos de polinomios zonales

Dentro de la Estadística Matemática, los polinomios zonales y las funciones hipergeométricas de argumento matricial han sido utilizados, entre otras aplicaciones, en la expresión de la distribución de formas cuadráticas asociadas a poblaciones normales. Así, por ejemplo, expresiones que involucran funciones hipergeométricas aparecen en

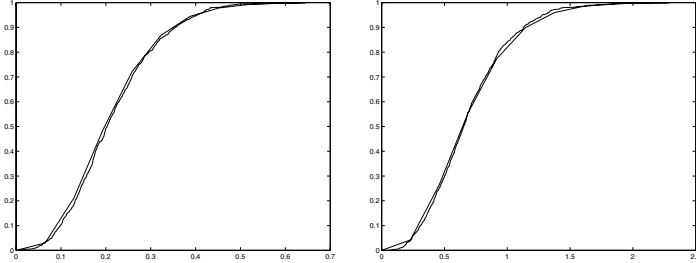


Figura 2. Funciones de distribución teórica y empírica para la menor raíz característica de $A^1 \rightarrow W_3(10, \Sigma_1)$ (a.) y $A^2 \rightarrow W_3(10, \Sigma_2)$ (b.).

la distribución χ^2 (Pearson, 1900), de Wishart central y no central (Wishart, 1928), de las raíces características de la matriz de covarianza de una distribución gaussiana (James, 1964), en la ratio de la varianza F (Fisher, 1924), de la F no central con p y n grados de libertad (Fisher, 1924), etc. En este apartado describimos, precisamente, una aplicación de la implementación realizada que nos permite evaluar la distribución de probabilidad de la menor raíz característica de una matriz de Wishart.

Concretamente (Muirhead, 1982), si l_m es la menor raíz característica de S , donde $A = nS$ es $W_m(n, \Sigma)$, y $r = \frac{1}{2}(n - m - 1)$ es un entero positivo, entonces

$$(11) \quad P_{\Sigma}(l_m > x) = \text{etr} \left(-\frac{1}{2}nx\Sigma^{-1} \right) \sum_{k=0}^{mr} \sum_{\kappa}^* \frac{C_{\kappa}(\frac{1}{2}nx\Sigma^{-1})}{k!},$$

donde \sum_{κ}^* denota la suma sobre todas las particiones $\kappa = (k_1, \dots, k_m)$ de k tales que $k_1 \leq r$.

La puesta en práctica de este teorema tan sólo implica el desarrollo de una serie finita en términos de polinomios zonales, así que puede realizarse de manera exacta hasta distintos valores de m y n , aunque no muy elevados.

Nosotros hemos considerado una aplicación en donde $m = 3$ y $n = 10$. Se han simulado mediante el programa matemático MATLAB 1.000 valores de matrices aleatorias $X_{n \times m} \rightarrow N(0, I_n \otimes \Sigma)$. A partir de estas 1.000 matrices, $\mathbf{x}_{n \times m}^1, \dots, \mathbf{x}_{n \times m}^{1,000}$, se calculan 1.000 valores simulados de sendas distribuciones de Wishart, mediante la expresión $\mathbf{A}^i = \mathbf{x}^{i'} \cdot \mathbf{x}^i$, $i = 1, \dots, 1,000$, de modo que la distribución teórica de estas matrices es la de $A \rightarrow W_3(10, \Sigma)$. Posteriormente se han dividido estas matrices por $n = 10$ y a cada una de estas nuevas matrices se les ha calculado su menor raíz característica. Esto genera una muestra aleatoria simple, $l_1, \dots, l_{1,000}$, de 1.000 valores de la variable aleatoria l_m , menor raíz característica de A/n .

En la Figura 2 aparecen superpuestas la función de distribución empírica, calculada a partir de sendas muestras aleatorias simples, y la función de distribución teórica, dada

por (11), para matrices

$$\Sigma_1 = \begin{pmatrix} 1 & 0,58 & 0,61 \\ 0,58 & 1 & 0,58 \\ 0,61 & 0,58 & 1 \end{pmatrix}, \Sigma_2 = \begin{pmatrix} 6,83 & 6,40 & 2,84 \\ 6,40 & 8,02 & 2,79 \\ 2,84 & 2,79 & 3,98 \end{pmatrix}$$

Como puede comprobarse, son prácticamente idénticas, confirmando que los datos empíricos quedan perfectamente ajustados mediante la distribución teórica.

4. DISCUSIÓN

La metodología descrita en este trabajo, que involucra el cálculo explícito de los polinomios zonales, proporciona la evaluación, exacta en algunos casos y aproximada en otros, de cualquier serie hipergeométrica de argumento matricial. Esto permite la puesta en práctica de numerosos resultados que involucran a estas series y que hasta ahora, o no se habían aplicado a casos reales, o si se habían aplicado, tenían que limitarse a casos donde técnicas de análisis numérico permitieran algún resultado concreto. A su vez, también permite abrir nuevas líneas de investigación, como la de la génesis de distribuciones multivariantes de probabilidad mediante funciones hipergeométricas de argumento matricial.

Hay que decir, no obstante, que el éxito de esta metodología está supeditado a las limitaciones computacionales en el cálculo de los polinomios zonales y, sobre todo, de los monomios simétricos en términos de los cuales vienen definidos.

5. AGRADECIMIENTOS

Deseamos agradecer a los referees sus comentarios y sugerencias, los cuales han sido de gran ayuda a la hora de elaborar este manuscrito.

REFERENCIAS

- Gutiérrez, R. and Rodríguez, J. (1997). «Family of Pearson Discrete Distributions Generated by the Univariate Hypergeometric function ${}_3F_2(\alpha_1, \alpha_2, \alpha_3; \gamma_1, \gamma_2; \lambda)$ ». *Appl. Stochastics Models and Data Anal*, 13, 115-125.
- Gutiérrez R., Rodríguez J. and Sáez A. J. (2000). «Approximation of hypergeometric functions of matricial argument through their development in series of zonal polynomials». *Electronic Transactions in Numeric Analysis*, 11, 121-130.

- James, A. T. (1961). «Zonal polynomials of the real positive definite symmetric matrices». *Annals of Mathematics*, 74, 456-469.
- James, A. T. (1968). «Calculation of zonal polynomial coefficient by use of the Laplace-Beltrami operator». *Ann. Math. Stat.*, 39, 1711-1718.
- Johnson, N. L., Kotz, S. and Kemp, A. W. (1992). *Univariate Discrete Distributions*. Wiley, New York. Second Edition.
- Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. Wiley. New York.
- Rodríguez, J., Conde, A. and Sáez, A. (2001a). «A new class of discrete distributions with complex parameters». *Accepted in Statistical Papers*.
- Rodríguez, J., Conde, A., Olmo, M. J. and Sáez A. J. (2001b). «Discrete multivariate distributions symmetric in frequencies generated by the Gauss function ${}_2F_1$ of matrixial argument». (*Submitted to Test*).
- Takemura, A. (1984). «Zonal Polynomials». *Institute of Mathematical Statistics*. Hayward, California.

ENGLISH SUMMARY

IMPLEMENTING CALCULUS OF ZONAL POLYNOMIALS AND APPLICATIONS IN MULTIVARIATE ANALYSIS

J. RODRÍGUEZ-AVI
A. J. SÁEZ-CASTILLO
A. CONDE-SÁNCHEZ
Universidad de Jaén*

The implementation of an algorithm that permits the calculation of zonal polynomials and two applications in multivariate analysis are described in this paper. This implementation provides summation results of hypergeometric functions of matricial argument that may be used in generating discrete multivariate distributions whose frequencies are symmetrical. Development of hypergeometric series of matricial argument also allows a well-known result that characterizes the distribution of the lowest eigenvalue of a random matrix with Wishart distribution to be put into practice.

Keywords: Zonal polynomials, discrete distributions, Wishart distribution, hypergeometric functions of matricial argument

AMS Classification (MSC 2000): 60E05, 62E15

*Dep. de Estadística e I.O. Despacho 147, Ed. D-3, Paraje Las Lagunillas, Universidad de Jaén, 23071 Jaén, España. Telf: + 34 953 012.207; fax: + 34 953 012.222; email: jravi@ujaen.es.

–Received December 2000.

–Accepted April 2002.

Zonal polynomials constitute a family of symmetrical and homogeneous polynomials. They may be considered to be a multivariate extension of powers of a single variable and are defined as eigenfunctions of the differential operator

$$\Delta_X = \sum_{i=1}^m x_i^2 \frac{\partial^2}{\partial x_i^2} + \sum_{i=1}^m \sum_{j=1, j \neq i}^m \frac{x_i^2}{x_i - x_j} \frac{\partial}{\partial x_i},$$

where

$$\alpha_\kappa = \sum_{i=1}^m k_i(k_i - i) + k(m - 1),$$

in such a way that there is no explicit general formula for their evaluation.

Alternatively, James (1968) establishes an algorithm that permits zonal polynomials to be expressed as linear combinations of symmetric monomials, which are a common base of symmetrical and homogeneous polynomials. This algorithm has been implemented employing MATLAB. Source files may be obtained from

<http://www.ujaen.es/dep/estinv/pdi/ajsaez/software.html>

Moreover, hypergeometric functions of matricial argument,

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; X) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} \cdots (a_p)_{\kappa}}{(b_1)_{\kappa} \cdots (b_q)_{\kappa}} \frac{C_{\kappa}(X)}{k!},$$

where \sum_{κ} denotes the sum over all partitions $\kappa = (k_1, \dots, k_m)$ of k and $C_{\kappa}(X)$ is the zonal polynomial of X corresponding to κ , are multivariate extension of classical hypergeometric functions, where Pochhammer symbols are replaced by generalized hypergeometric coefficients and powers of single variables by zonal polynomials. As neither have an exact formula, we propose a development of the hypergeometric series in terms of symmetric monomials using the James algorithm for zonal polynomials: in this manner we obtain approximate values of the series.

In addition, hypergeometric functions of matricial argument may be used to generate probability functions of multivariate discrete distributions; these distributions extend the family of well-known univariate distributions generated by classical hypergeometric functions (Johnson, Kotz and Kemp, 1992). In this context, the development of hypergeometric functions of matricial argument considered in this paper is suitable for easily obtaining frequencies of the distributions as coefficients of the series in terms of symmetric monomials. The main characteristic of these distributions is that frequencies are invariable when permutations of the variables in the random vector are considered: this property is due to the symmetry of zonal polynomials.

In the field of Mathematical Statistics, hypergeometric series of matricial argument have been mainly used in expressing multivariate distributions of quadratic forms in normal

samples that extend analogous results in a univariate environment. Many of these results involve hypergeometric series without explicit expression, so they cannot be put into practice. In this sense, we consider an example where the distribution of the lowest eigenvalue of a random matrix with Wishart distribution is expressed as a finite sum of zonal polynomials. Specifically, if l_m is the lowest eigenvalue of S , where $A = nS$ is $W_m(n, \Sigma)$, and $r = \frac{1}{2}(n - m - 1)$ is a positive integer, then (Muirhead, 1982)

$$P_{\Sigma}(l_m > x) = \text{etr} \left(-\frac{1}{2}nx\Sigma^{-1} \right) \sum_{k=0}^{mr} \sum_{\kappa}^* \frac{C_{\kappa} \left(\frac{1}{2}nx\Sigma^{-1} \right)}{k!},$$

where \sum_{κ}^* denotes sum over all partitions $\kappa = (k_1, \dots, k_m)$ of k such that $k_1 \leq r$. We have expressed this result in terms of symmetric monomials in such a way that we can calculate exact probabilities of the distribution. The paper concludes with a simulation to show that the theoretical distribution function is close to the empirical distribution functions when we consider sample data for a Wishart sample.