THE INSTITUTE OF STATISTICS

THE CONSOLIDATED UNIVERSITY OF NORTH CAROLINA

USE OF MOMENTS IN DISTRIBUTION THEORY: A MULTIVARIATE CASE

N.A. Volodin Samuel Kotz N.L. Johnson

August 1991

Mimeo Series #2063

DEPARTMENT OF STATISTICS
Chapel Hill, North Carolina

Use of Moments in Distribution Theory: A Multivariate Case

N.A. Volodin
Tashkent
U.S.S.R.

Samuel Kotz
University of Maryland
College Park, MD 20742
U.S.A.

N.L. Johnson
University of North Carolina
Chapel Hill, NC 27599-3260
U.S.A.

ABSTRACT

In recent papers, Johnson and Kotz (1990a,b) have explored the utility of moment calculations as a simple way of establishing distributional forms. In particular a characterization theorem for beta distributions was proved. In this paper these methods are extended to multivariate problems, and a result established for Dirichlet distributions.

Key words: Characterization, Dirichlet distribution, Limit distribution, Moments, Random matrices.

INTRODUCTION

In Johnson and Kotz (1990a), elementary tools - the so-called 'moment methods' - were employed to obtain distributions and characterizations of distributions of random mixtures of form

$$Z = WX_1 + (1-W)X_2$$

(or, more generally

$$Z = \sum_{j=1}^{m} W_j X_j),$$

where the X's are mutually independent, and have a common beta distribution, and the W's are independent of the X's.

In Johnson and Kotz (1990b), the method was extended to the distributions of variables of type

$$Y = \sum_{j=0}^{\infty} (-1)^j \prod_{i=1}^{j} X_i$$

In particular the following result was obtained:

"If X_0, X_1,... are independent and identically distributed (i.i.d.) random variables, the distribution of Y is beta (a+b,b) if and only if the common distribution of the X's is beta (a,b)."

Similar, more general, results were recently obtained by Chamayou and Letac (1991) using different, more advanced methodology. See, in particular Example 7 of Chamayou and Letac (page 20). Devroye, Letac and Seshadri (1986) used a moment method for determination of distributions of random intervals - which was also the second topic discussed by Johnson and Kotz (1990a,b).

In the present paper, we extend this methodology to derive characterizations of multivariate distributions - in particular, of Dirichlet distributions. These distributions have been applied, with increasing frequency, in statistical modelling, distribution theory and Bayesian inference. See, for example, Aitchison (1986) and Fang, Kotz and Ng (1990).

The essential features of the moment method are:

(i) establishing the existence of a limit;

(ii) establishing a recurrence relation among variables having the initial and limit distributions;

(iii) using (ii), obtaining formulae linking moments of the original and limit distributions, indicating that the moments of either one are determined by the moments of the other;

(iv) for a given original distribution, demonstrating that the equations are satisfied by the moments of a distribution.

Provided the moments determine the distributions (as is certainly the case when the ranges of variation of the random variables are finite, which is so for Dirichlet distributions), it follows that this distribution is the limit distribution if and only if the original distribution is the one whose moments have been used in the calculations.

2. RESULTS

THEOREM: Let $\underset{\sim}{Y}_i^{(n)} = (Y_{i1}^{(n)},\ldots,Y_{ik}^{(n)})$ $(i=1,\ldots,k;\ n=1,2,\ldots)$ be i.i.d. (k-vector) random variables and

$$\underset{\sim}{Y}^{(n)} = \begin{bmatrix} \underset{\sim}{Y}_1^{(n)} \\ \vdots \\ \underset{\sim}{Y}_k^{(n)} \end{bmatrix} .$$

The the limit distribution of each row $\underset{\sim}{X}_1^{(n)},\ldots,\underset{\sim}{X}_k^{(n)}$ of

$$\underset{\sim}{X}^{(n)} = \underset{\sim}{Y}^{(n)}\ \underset{\sim}{Y}^{(n-1)} \ \cdots\ \underset{\sim}{Y}^{(1)}$$

is Dirichlet with parameters ka, ka, ..., ka (D(ka,ka,...,ka)) if and only if the distribution of each $\underset{\sim}{Y}_i^{(n)}$ is D(a,a,...a).

REMARK (1): If the joint distribution of $Y_{i1}^{(n)},\ldots,Y_{ik}^{(n)}$ is D(a,a,...a), then the joint probability density function is

$$f_{\underset{\sim}{Y}_i^{(n)}}(\underset{\sim}{y}) = \frac{\Gamma(ka)}{\{\Gamma(a)\}^k} \prod_{i=1}^{k} y_i^{a-1} \quad \left(0 \le y_i;\ \sum_{i=1}^{k} y_i = 1\right) \qquad (1)$$

The range of variation of each $Y_{ij}^{(n)}$ is finite (in fact [0,1]).

REMARK (2): $\sum_{j=1}^{k} Y_{ij}^{(n)} = 1 = \sum_{j=1}^{k} X_{ij}^{(n)}$, so that $\underset{\sim}{X}^{(n)}$ and $\underset{\sim}{Y}^{(n)}$ are stochastic matrices.

REMARK (3): When k = 2, D(a,a) is a beta (a,a) distribution and D(2a,2a) is a beta (2a,2a) distribution.

PROOF: The proof will be given in four stages, corresponding to (i) - (iv) of the Introduction

(i) <u>Existence of the limit</u>

Denoting the (i,j)-th element of $\underset{\sim}{X}^{(n)}$ by $X_{ij}^{(n)}$,

$$X_{ij}^{(1)} = Y_{ij}^{(1)}$$

and

$$X_{ij}^{(n+1)} = \sum_{u=1}^{k} Y_{iu}^{(n)} X_{uj}^{(n)} \tag{2}$$

so that $X_{ij}^{(n+1)}$ is a weighted mean of $X_{1j}^{(n)},\dots,X_{kj}^{(n)}$, if

$$\sum_{u=1}^{k} Y_{iu}^{(n)} = 1 \text{ and } Y_{iu}^{(n)} \geq 0 \quad (u=1,\dots,k)$$

as in (1). Hence

$$\min_u(X_{uj}^{(n)}) \leq X_{ij}^{(n)} \leq \max_u(X_{uj}^{(n)})$$

and further

$$\min_u(X_{uj}^{(1)}) \leq \min_u(X_{uj}^{(2)}) \leq \dots \leq \min_u(X_{uj}^{(n)}) \leq X_{ij}^{(n)} \leq \max_u(X_{uj}^{(n)})$$
$$\leq \dots \leq \max_u(X_{nj}^{(1)}) \quad (u,i,j = 1,\dots,k) \tag{3}$$

As $n \to \infty$, $\min_u(X_{uj}^{(n)})$ and $\max_u(X_{uj}^{(n)})$ must each tend to a limit. We now show that these limits coincide with probability 1, and hence this must also be

$$\lim_{n\to\infty} X_{ij}^{(n)} = X_{ij}, \text{ say } \quad (\text{for all } i) \tag{4}$$

provided $\Pr[\max_u(Y_{iu}^{(1)}) < 1] > 0$.

If $\min_u(X_{uj}^{(n)}) = \max_u(X_{uj}^{(n)})$, then all $\{X_{uj}^{(n)}\}$ $(u=1,\ldots,k)$ are equal (to this common value), and this is also their common limit value $X_{1j} = X_{2j} = \ldots = X_{kj}$. Otherwise, defining

$$D_j^{(n)} = \max_u(X_{uj}^{(n)}) - \min_u(X_{uj}^{(n)})$$

we have, from (4)

$$\Pr[D_j^{(n+1)} \leq (1-\eta)D_j^{(n)} | \{X_{uj}^{(n)}\},\ j = 1,\ldots k] > \epsilon > 0$$

for some fixed $\eta > 0$ and $\epsilon > 0$, for all $\{X_{uj}^{(n)}\}$.

Hence

$$\Pr[D_j^{(n+1)} \leq (1-\eta)^m] < \Pr[Z \geq m] \tag{5}$$

(since $D_j^{(1)} \leq 1$) where Z has a binomial distribution with parameters (n,ϵ). So, for any $\delta > 0$ and any $\epsilon^* > 0$

$$\Pr[D_j^{(n)} < \delta \quad \text{for all } n > n_0\ (\delta,\epsilon,\epsilon^*)] > 1 - \epsilon^* \tag{6}$$

(Note that $D_j^{(n)}$ cannot decrease with n, so if $D_j^{(n)} < \delta$, then $D_j^{(r)} < \delta$ for all $r > n$. Choose m in (5) greater than $(\log \delta)/\{\log(1-\eta)\}$ and n_0 large enough to make

$$\Pr[Z \geq m] > 1 - \epsilon^*.)$$

Note that each $X_{ij}^{(n)}$ $(i=1,\ldots,k)$ tends to the same limiting value (see (4)) as $n \to \infty$ (though not, in general, the same for all j). *A fortiori*, the limit distribution of $\underset{\sim}{X}_i^{(n)}$ is the same for all i. Further, there is no need to investigate the joint limit distribution of $\underset{\sim}{X}_1^{(n)},\ldots,X_k^{(n)}$ since given any set of $\underset{\sim}{X}_i$ values, the values of each of the other rows $\{\underset{\sim}{X}_j\}$ $(j \neq i)$ will be the same as those of $\underset{\sim}{X}_i$.

(ii) <u>Recurrence relation among variables with original and limit distribution</u>

If $\underset{\sim}{X}_i^{(n)}$ has a limit distribution, this will also be the limit distribution (as $n \to \infty$) of the i-th row of $\underset{\sim}{Y}^{(n)}\, \underset{\sim}{Y}^{(n-1)} \ldots \underset{\sim}{Y}^{(2)}$. Hence we have

$$\underset{\sim}{X}^{*} = \underset{\sim}{X}\underset{\sim}{Y} \tag{7}$$

where $\underset{\sim}{X}^{*}$ and $\underset{\sim}{X}$ each have the limit distribution, $\underset{\sim}{Y}$ has the $\underset{\sim}{Y}^{(1)}$ distribution, and $\underset{\sim}{X}$ and Y are mutually independent.

(iii) <u>Formulae linking moments of the original and limit distributions</u>

From (7)

$$X_{ij}^{*} = \sum_{u=1}^{k} X_{iu}\, Y_{uj} \tag{8}$$

Since each $\underset{\sim}{Y}_u$ has the same distribution we see that each row $\underset{\sim}{X}_i^{*} = (X_{i1},\ldots,X_{ik})$ has the same distribution.

We denote

$$E\left[\prod_{j=1}^{k} X_{ij}^{s_j}\right] = E\left[\prod_{j=1}^{k} X_{ij}^{*s_j}\right] \text{ by } \mu_{s_1\ldots s_k} \quad \text{(for all i)}$$

and

$$E\left[\prod_{j=1}^{k} Y_{uj}^{t_j}\right] \text{ by } \nu_{t_1\ldots t_k} \quad \text{(for all u).}$$

From (8)

$$\mu_{s_1\ldots s_k} = E\left[\prod_{j=1}^{k}\left\{\sum_{u=1}^{k} X_{iu}\, Y_{uj}\right\}^{s_j}\right] = E\left[\prod_{j=1}^{k}\left\{\sum_{\underset{\sim}{h}_j}\begin{bmatrix} & s_j & \\ h_{j1} & \cdots & h_{jk}\end{bmatrix}\prod_{u=1}^{k}(X_{iu}\, Y_{uj})^{h_{ju}}\right\}\right] \tag{9}$$

where

$$\begin{bmatrix} & s_j & \\ h_{j1} & \cdots & h_{jk}\end{bmatrix} = (s_j!)\left\{\prod_{u=1}^{k}(h_{ju}!)\right\}^{-1}$$

and $\sum_{\underset{\sim}{h}_j}$ denotes summation over all nonnegative integers $\underset{\sim}{h}_j = (h_{j1},\ldots,h_{jk})$

subject to

$$\sum_{u=1}^{k} h_{ju} = s_j \quad (j = 1,\dots,k).$$

Equation (9) can be rearranged as

$$\mu_{s_1 \dots s_k} = \sum_{\underset{\sim}{h}_1} \dots \sum_{\underset{\sim}{h}_k} E\left[\prod_{u=1}^{k} X_{iu}^{h_{\cdot u}}\right] \prod_{j=1}^{k} \binom{s_j}{h_{j1} \dots h_{ju}} E\left[\prod_{u=1}^{k} \prod_{j=1}^{k} Y_{uj}^{h_{ju}}\right] \tag{10}$$

where $h_{\cdot u} = \sum_{j=1}^{k} h_{ju}$.

Noting that

$$E\left[\prod_{u=1}^{k} \prod_{j=1}^{k} Y_{uj}^{h_{ju}}\right] = \prod_{u=1}^{k} E\left[\prod_{j=1}^{k} Y_{uj}^{h_{ju}}\right] = \prod_{u=1}^{k} \nu_{h_{1u},\dots,h_{ku}}$$

$$= \prod_{j=1}^{k} \nu_{h_{1j},\dots,h_{kj}}$$

we have (for all $(s_1,\dots,s_k)$)

$$\mu_{s_1,\dots,s_k} = \sum_{\underset{\sim}{h}_1} \dots \sum_{\underset{\sim}{h}_k} \mu_{h_{\cdot 1},\dots,h_{\cdot k}} \prod_{j=1}^{k} \binom{s_j}{h_{j1},\dots,h_{jk}} \nu_{h_{1j},\dots,h_{kj}} \tag{11}$$

From the equations (11), the μ's can be obtained from the ν's, and conversely. (There are, of course, relationships among the μ's, and among the ν's, arising from Remark (2) following the statement of the theorem.)

(iv) Checking that the moment equations are satisfied

If $\underset{\sim}{Y}_i^{(n)}$ has the symmetrical Dirichlet distribution D(a,a,...a) for all i and all n, then

$$\nu_{h_{1j},\dots,h_{kj}} = \frac{\Gamma(ka)}{\{\Gamma(a)\}^k} \frac{\prod_{u=1}^{k} \Gamma(a + h_{uj})}{\Gamma(a + h_{\cdot j})} \tag{12}$$

Equations (11) now become

$$\mu_{s_1,\ldots,s_k} = \frac{\{\Gamma(ka)\}^k}{\{\Gamma(a)\}^{2k}} \sum_{\underset{\sim}{h}_1} \cdots \sum_{\underset{\sim}{h}_k} \mu_{h_{\cdot 1},\ldots,h_{\cdot k}} \prod_{j=1}^{k} \left\{ \binom{s_j}{h_{j1},\ldots,h_{jk}} \frac{\prod_{u=1}^{k} \Gamma(a+h_{uj})}{\Gamma(a+h_{\cdot j})} \right\} \tag{13}$$

We now show that equations (11) are satisfied if $\underset{\sim}{X}_i$ has the D(ka,ka,...,ka) distribution, with

$$\mu_{s_1,\ldots,s_k} = \frac{\Gamma(k^2a)}{\{\Gamma(ka)\}^k} \frac{\prod_{j=1}^{k} (ka+s_j)}{\Gamma(k^2a+s)} \tag{14.1}$$

where $s = \sum_{j=1}^{k} s_j$, and

$$\mu_{h_{\cdot 1},\ldots,h_{\cdot k}} = \frac{\Gamma(k^2a)}{\{\Gamma(ka)\}^k} \frac{\prod_{j=1}^{k} \Gamma(ka+h_{\cdot j})}{\Gamma(k^2a+s)} \tag{14.2}$$

(since $\sum_{j=1}^{k} h_{\cdot j} = \sum_{u=1}^{k} \sum_{j=1}^{k} h_{uj} = \sum_{u=1}^{k} s_u = s$).

The right-hand side of (13) can be written as

$$\frac{\Gamma(k^2\ a)}{\Gamma(k^2a+s)} \sum_{\underset{\sim}{h}_1} \cdots \sum_{\underset{\sim}{h}_k} \prod_{j=1}^{k} \left[\binom{s_j}{h_{j1},\ldots,h_{jk}} \prod_{u=1}^{k} \left\{ \frac{\Gamma(a+h_{uj})}{\Gamma(a)} \right\} \right]$$

$$= \frac{\Gamma(k^2a)}{\Gamma(k^2a+s)} \sum_{\underset{\sim}{h}_1} \cdots \sum_{\underset{\sim}{h}_k} \prod_{j=1}^{k} \left[\binom{s_j}{h_{j1},\ldots,h_{jk}} \prod_{u=1}^{k} \left\{ \frac{\Gamma(a+h_{ju})}{\Gamma(a)} \right\} \right]$$

(because $\prod_{j=1}^{k} \prod_{u=1}^{k} \Gamma(a+h_{uj}) = \prod_{j=1}^{k} \prod_{u=1}^{k} \Gamma(a+h_{ju})$)

$$= \frac{\Gamma(k^2a)}{\Gamma(k^2a+s)} \prod_{j=1}^{k} \sum_{\underset{\sim}{h}_j} \binom{s_j}{h_{j1},\ldots,h_{jk}} \frac{\Gamma(ka+s_j)}{\Gamma(ka)} E\left[\prod_{u=1}^{k} Y_u^{h_{ju}} \right] \Bigg\}$$

where $(Y_1,\ldots,Y_k)$ has a Dirichlet D(a,a,...a) distribution (See (1).)

Thus the right-hand side of (13) is equal to

$$\frac{\Gamma(k^2a)}{\{\Gamma(ka)\}^k\,\Gamma(k^2a + s)} \prod_{j=1}^{k} \Gamma(ka + s_j) = E\left[\left[\sum_{u=1}^{k} Y_u\right]^{s_j}\right]$$

$$\frac{\Gamma(k^2a)}{\{\Gamma(ka)\}^k\,\Gamma(k^2a + s)} \prod_{j=1}^{k} \Gamma(ka + s_j) = \mu_{s_1,\ldots,s_k}$$

(because $\sum_{u=1}^{k} Y_u = 1$).

The statement of the theorem now follows from the last paragraph of the Introduction.

REMARK (4): The distributions $D(a,a,\ldots,a)$ and $D(ka,ka,\ldots,ka)$ have the same marginal expected values $-k^{-1}$ - and the same pairwise correlation coefficients - $-(k-1)^{-1}$. However, the marginal variances of $D(a,a,\ldots,a)$ are $k^{-2}(ka + 1)^{-1}(k-1)$, while those of $D(ka,ka,\ldots,ka)$ are $k^{-2}(k^2a + 1)^{-1}(k-1)$.

3. CONCLUSING COMMENTS

The result obtained in the theorem is remarkable in that the relationship between original and limit distributions is so simple in the case of symmetrical Dirichlet distributions.

The methodology can be extended to the case of general exponential families, providing simplified and more 'transparent' proofs of the results of Chamayou and Letac (1991), mentioned earlier.

REFERENCES

Aitchison, J. (1986) The Statistical Analysis of Compositional Data, London: Chapman and Hall.

Chamayou, J.F. and Letac, G. (1991) Explicit stationary distributions for compositions of random functions and products of random matrices, *J. Theor. Prob.*, **4**, 3-36.

Devroye, L., Letac, G. and Seshadri, V. (1989) The limit behavior of an interval splitting scheme, *Statist. Prob. Letters*, **4**, 183-186.

Fang, K.-T., Kotz, S. and Ng, K.W. (1990) Symmetrical Multivariate and Related Distributions, New York: Chapman and Hall.

Johnson, N.L. and Kotz, S. (1990a) Randomly weighted averages: Some results and extensions, *Amer. Statistician*, **44**, 245-249.

Johnson, N.L. and Kotz, S. (1990b) Use of moments in deriving distributions and some characterizations, *Math. Scientist*, **15**, 42-52.