JOURNAL OF MULTIVARIATE ANALYSIS 39, 202–208 (1991)

Entropy Inequalities for Some Multivariate Distributions

SHYAMAL DAS PEDDADA* AND DONALD ST. P. RICHARDS†

University of Virginia

Communicated by C. R. Rao

In this paper, we derive some monotonicity properties of generalized entropy functionals of various multivariate distributions. These include the distributions of random eigenvalues arising in many hypothesis testing problems in multivariate analysis; the multivariate Liouville distributions; and the noncentral Wishart distributions. © 1991 Academic Press, Inc.

1. INTRODUCTION

Suppose that p is the probability density function of a continuous random vector (or matrix) X. Then the entropy of X,

$$H(p) = -\int p(x) \log p(x)\, dx \tag{1.1}$$

is well known to play an important role in probability and statistics. We refer to Karlin and Rinott [5, 6], Marshall and Olkin [9], Rao [12], and Rényi [13] for applications of the entropy function (1.1) to probability and statistics.

Rényi [13] and Karlin and Rinott [5, 6] have studied the *generalized entropy functional*

$$H_\alpha(p) = \frac{1}{1-\alpha} \log\left(\int [p(x)]^\alpha\, dx\right), \qquad \alpha > 0. \tag{1.2}$$

Received December 14, 1990; revised March 6, 1991.

AMS 1985 subject classifications: primary 60E99, 62E15, 62H10.

Key words and phrases: entropy functionals, Liouville distributions, multivariate beta distributions, Schur-concavity, Schur-convexity, Selberg's integrals, noncentral Wishart distributions.

* Research performed while on leave at the Department of Mathematics, Cornell University.

† Partially supported by NSF Grant DMS-8802929 and by the Center for Advanced Studies, University of Virginia.

0047-259X/91 $3.00

Note that $H_\alpha(p) \to H(p)$ as $\alpha \to 1$. In some instances, it is simpler to compute $H_\alpha(p)$ and then recover $H(p)$ by taking limits as $\alpha \to 1$. In treating $H_\alpha(p)$, it is enough to study the properties of the functional

$$G_\alpha(p) = \int [p(x)]^\alpha \, dx. \tag{1.3}$$

Karlin and Rinott [5, 6] derived monotonicity properties of the functional $G_\alpha(p)$ for a large class of univariate and multivariate density functions.

In this paper we study monotonicity properties of $G_\alpha(p)$ when X is a vector of eigenvalues of a multivariate beta matrix (Muirhead [10]); and when X has a multivariate Liouville distribution (Gupta and Richards [3], Marshall and Olkin [9]). For these distributions, we evaluate the functional $G_\alpha(p)$ and use the techniques of [5, 6] to determine its Schur-concavity and Schur-convexity properties. These results are given in Sections 2 and 3, respectively.

When X is a central Wishart matrix, Karlin and Rinott [6] obtained the concavity properties of $G_\alpha(p)$ in terms of the matrix $\Sigma = E(X)$. This raises the question of whether similar results are valid for noncentral Wishart matrices. In general, this appears to be a difficult problem because the complicated nature of the noncentral Wishart density function (cf., Muirhead [10]) prevents exact evaluation of the functional $G_\alpha(p)$. In Section 4 we evaluate $G_2(p)$ in one case, and show that in this case $G_2(p)$ is Schur-concave in the eigenvalues of Σ.

2. Eigenvalue Distributions

Suppose that the parameter $\theta = (\theta_1, \theta_2) \in R^2_+$, and $X = (X_1, ..., X_n)$ is a vector of random eigenvalues with density function

$$p(x; \theta) = c_n(\theta_1, \theta_2, \gamma) \prod_{j=1}^{n} x_j^{\theta_1 - 1}(1 - x_j)^{\theta_2 - 1} \prod_{1 \leqslant i < j \leqslant n} |x_i - x_j|^{2\gamma}, \tag{2.1}$$

where $0 < x_1 < x_2 < \cdots < x_n < 1$, γ is a positive real constant, and $c_n(\theta_1, \theta_2, \gamma)$ is the normalizing constant. The functions (2.1) arise as the densities of eigenvalues of multivariate beta matrices (Muirhead [10]), and in many hypothesis testing problems in multivariate analysis (Andersson, Brøns, and Jensen [1], Andersson and Perlman [2]). In these problems, the constant γ is $\frac{1}{2}$, 1, or 2, and the value of the normalizing constant is [1]

$$c_n(\theta_1, \theta_2, \gamma) = n! \prod_{i=1}^{n} \frac{\Gamma(\theta_1 + \theta_2 + (2n - i - 1)\gamma)\, \Gamma(\gamma + 1)}{\Gamma(\theta_1 + (n - i)\gamma)\, \Gamma(\theta_2 + (n - i)\gamma)\, \Gamma(i\gamma + 1)}. \tag{2.2}$$

Note that (2.1) remains a density function for all positive θ_1, θ_2, and γ.

Then by an integral formula of Selberg [14] (cf., Karlin and Studden [7, Chap. IV, Section 6]), (2.2) again gives the value of the normalizing constant $c_n(\theta_1, \theta_2, \gamma)$.

Now we can state our result concerning the monotonicity of the functional $G_\alpha(p)$ for the density function (2.1).

THEOREM 2.1. *For the density function* (2.1), *the entropy functional* $G_\alpha(p)$ *satisfies*:

(a) *If* $0<\alpha<1$ *then* $G_\alpha(p)$ *is Schur-concave in* (θ_1, θ_2);

(b) *If* $\alpha>1$ *and* $\theta_j>1-\alpha^{-1}-n\gamma$, $j=1,2$, *then* $G_\alpha(p)$ *is Schur-convex in* (θ_1, θ_2).

In the course of proving Theorem 2.1, we will need the following result of Karlin and Rinott [6].

LEMMA 2.2 [6]. *For* $0<\alpha<1$, *the function* $g_\alpha(x)=\Gamma(\alpha x-\alpha+1)/[\Gamma(x)]^\alpha$ *is log-concave in* $x>0$. *If* $\alpha>1$ *then* $g_\alpha(x)$ *is log-convex over the domain* $x>(\alpha-1)/\alpha$.

Proof of Theorem 2.1. By (1.3), (2.2), and Selberg's integral, we have

$$G_\alpha(p)=\frac{[c_n(\theta_1, \theta_2, \gamma)]^\alpha}{[n!]^{\alpha-1}\, c_n(\alpha\theta_1-\alpha+1, \alpha\theta_2-\alpha+1, \alpha\gamma)}. \tag{2.3}$$

Define

$$h_\alpha(x)=\prod_{i=1}^{n}\frac{\Gamma(\alpha x+(n-i)\alpha\gamma-\alpha+1)}{[\Gamma(x+(n-i)\gamma)]^\alpha}, \qquad x>0;$$

then we can rewrite (2.3) as

$$G_\alpha(p)=h_\alpha(\theta_1)\, h_\alpha(\theta_2)\, k_\alpha(\theta_1+\theta_2),$$

where the explicit form of the function k_α will not be required.

If $0<\alpha<1$ then, by Lemma 2.1, the function h_α is log-concave in $x>0$; and if $\alpha>1$ then h_α is log-convex in $x>1-\alpha^{-1}-n\gamma$. Therefore $h_\alpha(\theta_1)\, h_\alpha(\theta_2)$, and in turn $G_\alpha(p)$, is Schur-concave in $(\theta_1, \theta_2)\in R_+^2$ if $0<\alpha<1$. Similarly, if $\alpha>1$ then $G_\alpha(p)$ is Schur-convex in (θ_1, θ_2) for $\theta_j>1-\alpha^{-1}-n\gamma$, $j=1,2$. ∎

Similar to Theorem 2.1 we can also obtain results on the Schur-concavity and Schur-convexity of entropy functionals for the distributions of eigenvalues of $n\times n$ Wishart matrices $W_n(d, I_n)$, i.e., with d degrees of freedom and expectation the identity matrix; as well as for the distributions of eigenvalues of central F matrices. In the case of the Wishart matrix distribution, similar results were obtained by Karlin and Rinott [6].

3. Multivariate Liouville Distributions

Let the continuous random vector $X=(X_1, X_2, ..., X_n)$ have a multivariate Liouville distribution [3, 9]. Then the joint density function of X is of the form

$$p(x; \theta)=c_n(\theta)\, f\left(\sum_{i=1}^{n} x_i\right) \prod_{i=1}^{n} x_i^{\theta_i-1}, \tag{3.1}$$

where $x_i>0$, $i=1, ..., n$; $f: R_+ \to R_+$ is continuous and the parameter $\theta=(\theta_1, \theta_2, ..., \theta_n) \in R_+^n$.

It may be shown [3, 9] that the normalizing constant $c_n(\theta)$ is given by

$$[c_n(\theta)]^{-1}=\frac{\prod_{i=1}^{n} \Gamma(\theta_i)}{\Gamma(\sum_{i=1}^{n} \theta_i)} \int_0^\infty t^{\sum_{i=1}^{n}\theta_i-1} f(t)\, dt. \tag{3.2}$$

The analog of Theorem 2.1 for the Liouville distributions is the following.

Theorem 3.1. *For the density function* (3.1), *the entropy functional* $G_\alpha(p)$ *satisfies*:

(a) *If* $0<\alpha<1$ *then* $G_\alpha(p)$ *is Schur-concave in* $(\theta_1, \theta_2, ..., \theta_n)$;

(b) *If* $\alpha>1$ *and* $\theta_j>1-\alpha^{-1}, j=1, ..., n$, *then* $G_\alpha(p)$ *is Schur-convex in* $(\theta_1, \theta_2, ..., \theta_n)$.

Proof. The proof of this result is similar to that of Theorem 2.1. Let us adopt the notation

$$I(\beta; g):=\int_0^\infty t^\beta g(t)\, dt \tag{3.3}$$

for any function $g: R_+ \to R$ such that the integral (3.3) exists. Then applying (3.2), we obtain

$$G_\alpha(p)=\frac{I(\alpha\Sigma_{i=1}^{n}\theta_i-n\alpha+n; f^\alpha)[\Gamma(\sum_{i=1}^{n}\theta_i)]^\alpha \prod_{i=1}^{n} \Gamma(\alpha\theta_i-\alpha+1)}{I(\Sigma_{i=1}^{n}\theta_i-1; f)\, \Gamma(\alpha\sum_{i=1}^{n}\theta_i-n\alpha+n) \prod_{i=1}^{n} [\Gamma(\theta_i)]^\alpha}.$$

Letting

$$h_\alpha(x)=\Gamma(\alpha x-\alpha+1)/\Gamma^\alpha(x),$$

then we can rewrite $G_\alpha(p)$ as

$$G_\alpha(p) = k_\alpha\left(\sum_{i=1}^{n} \theta_i\right) \prod_{i=1}^{n} h_\alpha(\theta_i),$$

where the explicit form of the function k_α is not needed. By Lemma 2.2, $h_\alpha(x)$ is log-concave in $x > 0$ when $0 < \alpha < 1$; and $h_\alpha(x)$ is log-convex in $x > 1 - \alpha^{-1}$ when $\alpha > 1$. Then the statement of the theorem follows immediately. ∎

In the case when X has a Dirichlet or inverted Dirichlet distribution, Theorem 3.1 is due to Karlin and Rinott [6].

4. Noncentral Wishart Distributions

Suppose that the $n \times n$ matrix X has a noncentral Wishart distribution, $W_n(d, \Sigma, \Omega)$, with d degrees of freedom, $\Sigma = E(X)$, and noncentrality parameter Ω. The density function, p, of X exists for $d \geqslant n$. Relative to Lebesgue measure on the space of positive definite $n \times n$ matrices, we have [10, p. 442]

$$\begin{aligned} p(X) = {} & c_n (\det \Sigma)^{d/2} \exp(\tfrac{1}{2} \operatorname{tr} \Omega) \exp(-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1} X) \\ & \times (\det X)^{(d-n-1)/2} \, {}_0F_1(n/2; \tfrac{1}{4} \Omega \Sigma^{-1} X), \end{aligned} \tag{4.1}$$

where ${}_0F_1$ is a hypergeometric function of matrix argument [4, 10]; and throughout this section, c_n is generic notation for a constant depending only on n.

As noted above, Karlin and Rinott [6] obtained the monotonicity properties of $G_\alpha(p)$ in the central case ($\Omega = 0$). In the noncentral case, it appears to be difficult to evaluate $G_\alpha(p)$ in general. Despite the formidable nature of the density (4.1), we will evaluate $G_2(p)$ when $d = n + \frac{1}{2}$. This leads to the following result.

Proposition 4.1. *For $d = n + \frac{1}{2}$, the entropy functional $G_2(p)$ for the density* (4.1) *satisfies the following*:

(a) *$G_2(p)$ is log-concave in Σ*;

(b) *$G_2(p)$ is Schur-concave in $(\lambda_1, \ldots, \lambda_n)$, the vector of eigenvalues of Σ*;

(c) *If $\omega_1, \ldots, \omega_n$ are the eigenvalues of Ω, then $G_2(p)$ is strictly increasing in each ω_i, $i = 1, \ldots, n$*;

(d) *$G_2(p)$ is strictly log-convex in $(\omega_1, \ldots, \omega_n)$.*

Proof. To evaluate $G_2(p)$ when $d=n+\frac{1}{2}$, we need to evaluate the integral

$$\int_{X>0} (\det X)^{-1/2} \exp(-\operatorname{tr} \Sigma^{-1}X)[{}_0F_1(n/2; \tfrac{1}{4}\Omega\Sigma^{-1}X)]^2 \, dX, \qquad (4.2)$$

where dX denotes Lebesgue measure on the space $\{X>0\}$ of positive definite $n \times n$ matrices. By Herz's generalization [4, Eq. (5.8)] of Weber's second exponential integral, we see that (4.2) equals

$$c_n \exp(\tfrac{1}{2} \operatorname{tr} \Omega)(\det \Sigma)^{n/2} \, {}_0F_1(n/2; \tfrac{1}{16}\Omega^2).$$

Therefore

$$G_2(p) = c_n \exp(\operatorname{tr} \Omega)(\det \Sigma)^{n+1/4} \, {}_0F_1(n/2; \tfrac{1}{16}\Omega^2). \qquad (4.3)$$

Since $(\det \Sigma)^{n+1/4}$ is log-concave in Σ, then (a) follows. Next, (b) follows by the equivalence of log-concavity in Σ and Schur-concavity in $(\lambda_1, ..., \lambda_n)$ [6, Sections 7–8].

To prove (c), it suffices to show that ${}_0F_1(n/2; \Omega^2)$ is strictly increasing in each ω_i. This result follows from the zonal polynomial expansion [10, Section 7.3]

$${}_0F_1(n/2; \Omega^2) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{C_\kappa(\Omega^2)}{k!(n/2)_\kappa}$$

for the ${}_0F_1$ function. Here, the zonal polynomial $C_\kappa(\Omega)$ is a homogeneous, symmetric polynomial in $\omega_1, ..., \omega_n$. It is known (cf. [11, p. 1334]) that the zonal polynomials are strictly increasing in each ω_i; hence (c) follows.

To prove (d) note that by (4.3),

$$\left(\frac{\partial^2}{\partial\omega_i \, \partial\omega_j} \log G_2(p)\right) = \left(\frac{\partial^2}{\partial\omega_i \, \partial\omega_j} \log {}_0F_1\left(n/2; \frac{1}{16}\Omega^2\right)\right). \qquad (4.4)$$

By Khatri and Mardia [8, Eq. (2.11)], the right-hand side of (4.4) is a positive definite matrix. Then by Marshall and Olkin [9, p. 448, B.3.d], $\log G_2(p)$ is strictly convex in $(\omega_1, ..., \omega_n)$; hence (d) is proved. ∎

References

[1] Andersson, S. A., Brøns, H. K., and Jensen, S. T. (1983). Distribution of eigenvalues in multivariate statistical analysis. *Ann. Statist.* **11** 392–415.

[2] Andersson, S., and Perlman, M. D. (1984). Two testing problems relating the real and complex multivariate normal distributions. *J. Multivariate Anal.* **15** 21–51.

[3] Gupta, R. D., and Richards, D. St. P. (1987). Multivariate Liouville distributions. *J. Multivariate Anal.* **23** 233–256.

[4] Herz, C. S. (1955). Bessel functions of matrix argument. *Ann. Math.* **61** 474–523.

[5] Karlin, S., and Rinott, Y. (1981). Entropy inequalities for classes of probability distributions. I. The univariate case. *Adv. in Appl. Probab.* **13** 93–112.

[6] Karlin, S., and Rinott, Y. (1981). Entropy inequalities for classes of probability distributions. II. The multivariate case. *Adv. in Appl. Probab.* **13** 325–351.

[7] Karlin, S., and Studden, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics.* Interscience–Wiley, New York.

[8] Khatri, C. G., and Mardia, K. V. (1977). The von Mises–Fisher matrix distribution in orientation statistics. *J. Roy. Statist. Soc. Ser. B* **39** 95–106.

[9] Marshall, A. W., and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.

[10] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* Wiley, New York.

[11] Perlman, M. D., and Olkin, I. (1980). Unbiasedness of invariant tests for MANOVA and other multivariate problems. *Ann. Statist.* **8** 1326–1342.

[12] Rao, C. R. (1973). *Linear Statistical Inference and Its Applications.* Wiley, New York.

[13] Rényi, A. (1961). On measures of entropy and information. In *Proceedings, 4th Berkeley Sympos. Math. Statist. Probab., Vol. 1*, pp. 541–561.

[14] Selberg, A. (1944). Bemerkninger om et multipelt integral. *Nordisk. Mat. Tidskr.* **26** 71–78.