DOCUMENT RESUME

ED 069 678 TM 002 129

AUTHOR Olkin, Ingram
TITLE Monotonicity Properties of Dirichlet Integrals with Applications to the Multinomial Distribution and the Anova Test; A Draft.
INSTITUTION Educational Testing Service, Princeton, N.J.
SPONS AGENCY National Science Foundation, Washington, D.C.
REPORT NO ETS-RB-72-36
PUB DATE Aug 72
NOTE 13p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Analysis of Variance; Evaluation Techniques; Mathematical Models; Measurement Techniques; *Statistical Analysis; Technical Reports; Testing

ABSTRACT

Bounds for the tails of Dirichlet integrals are established by showing that each integral as a function of the limits is a Schur function. In particular, it is shown how these bounds apply to the simultaneous analysis of variance test and to the multinomial distribution. (Author)

RESEARCH

BULLETIN

RB-72-36

MONOTONICITY PROPERTIES OF DIRICHLET INTEGRALS WITH APPLICATIONS TO THE MULTINOMIAL DISTRIBUTION AND THE ANOVA TEST

Ingram Olkin

Stanford University
and Educational Testing Service

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the author. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
August 1972

MONOTONICITY PROPERTIES OF DIRICHLET INTEGRALS WITH APPLICATIONS TO THE MULTINOMIAL DISTRIBUTION AND THE ANOVA TEST

Ingram Olkin
Stanford University and Educational Testing Service

SUMMARY

Bounds for the tails of Dirichlet integrals are established by showing that each integral as a function of the limits is a Schur function. In particular, it is shown how these bounds apply to the simultaneous analysis of variance test and to the multinomial distribution.

KEY WORDS: Dirichlet integrals, simultaneous analysis of variance test, multinomial distribution, inequalities, Schur function, majorization, least favorable distribution.

MONOTONICITY PROPERTIES OF DIRICHLET INTEGRALS WITH APPLICATIONS TO THE MULTINOMIAL DISTRIBUTION AND THE ANOVA TEST[1]

Ingram Olkin
Stanford University and Educational Testing Service

1. Introduction. The present paper is concerned with establishing bounds for the Dirichlet integrals:

$$L_1(a) = \int_0^{a_1} \cdots \int_0^{a_k} \prod_1^k t_i^{w-1}(1-\Sigma t_i)^d \prod_1^k dt_i \quad , \tag{1.1}$$

$$L_2(a) = \int_0^{a_1} \cdots \int_0^{a_k} \prod_1^k t_i^{w-1}(1-\Sigma t_i)^{-d} \prod_1^k dt_i \quad , \tag{1.2}$$

$$U(a) = \int_{a_1}^{\infty} \cdots \int_{a_k}^{\infty} \prod_1^k t_i^{w-1}(1+\Sigma t_i)^{-d} \prod_1^k dt_i \quad , \tag{1.3}$$

where $a = (a_1,\ldots,a_k)$, and d is such that the integrals exist.

The main result is that under certain conditions, $-L_1(a)$, $-L_2(a)$, and $-U(a)$ are Schur functions in $(a_1,\ldots,a_k)$, (see e.g., Berge (1963) or Marshall, Olkin and Proschan (1967)). As a consequence of this fact it follows that if

$$(a_1,\ldots,a_k) \succ (b_1,\ldots,b_k)$$

[1]Work supported in part by Educational Testing Service and by National Science Foundation Grant GP-32326X at Stanford University.

in the sense that, after possible reordering, with $a_1 \geq \cdots \geq a_k$, $b_1 \geq \cdots \geq b_k$, that

$$\sum_1^m a_i \geq \sum_1^m b_i \;, \quad m = 1,\ldots,k-1 \;, \quad \sum_1^k a_i = \sum_1^k b_i \tag{1.4}$$

holds, then

$$L_1(a) \leq L_1(b) \;, \; L_2(a) \leq L_2(b) \;, \; U(a) \leq U(b) \quad . \tag{1.5}$$

In this way we can generate many inequalities. Both the multinomial distribution and the simultaneous analysis of variance test are related to $L_2(a)$ and $U(a)$, respectively, and we provide several new results as a consequence of (1.4).

2. The main results. We next prove the main results concerning $L_1(a)$, $L_2(a)$ and $U(a)$. Theorem 1 is stated more generally and encompasses both $L_1(a)$ and $L_2(a)$.

To show that a function $F(a) \equiv F(a_1,\ldots,a_k)$ is a Schur function, we must show that

$$\left(\frac{\partial F(a)}{\partial a_i} - \frac{\partial F(a)}{\partial a_j} \right) (a_i - a_j) \geq 0 \quad , \tag{2.1}$$

for all i and j.

Theorem 1. If $f(x)$ is a non-negative monotone decreasing function,

$$L(a;w) = \int_0^{a_1} \cdots \int_0^{a_k} f(\Sigma t) \prod_1^k t_i^{w_i - 1} dt_i \quad , \tag{2.2}$$

where $a_1 \geq \cdots \geq a_k \geq 0$, $0 \leq w_1 \leq \cdots \leq w_k$, then $-L(a;w)$ is a Schur function in $(a_1, \ldots, a_k)$, whenever the integral is finite.

Proof. Because of symmetry, we need only consider $(i,j) = (1,2)$ in showing that (2.1) holds for $-L(a;w)$. It is immediate from (2.2) -- noting the condition on $f(x)$ -- that

$$\frac{\partial L(a;w)}{\partial a_1} = \int_0^{a_2} \cdots \int_0^{a_k} f(a_1 + \sum_2^k t_i)\, a_1^{w_1 - 1} \prod_2^k t_i^{w_i - 1} dt_i \tag{2.3}$$

$$= \prod_1^k a_i^{w_i} \int_0^1 \cdots \int_0^1 a_1^{-1} f(a_1 + a_2 z + \sum_3^k a_i y_i)\, z^{w_2 - 1} \prod_3^k y_i^{w_i - 1} dz \prod_3^k dy_i \quad .$$

Condition (2.1) for $-L(a;w)$ will be satisfied if

$$\int_0^1 \cdots \int_0^1 [a_1^{-1} f(a_1 + a_2 z + \sum_3^k a_i y_i) z^{w_2 - 1} - a_2^{-1} f(a_1 z + a_2 + \sum_3^k a_i y_i) z^{w_1 - 1}] \tag{2.4}$$

$$dz \prod_3^k dy_i \leq 0 \quad .$$

A sufficient condition for (2.4) to hold is that, pointwise,

$$a_2 f(a_1 + a_2 z + Q) z^{w_2 - 1} \leq a_1 f(a_1 z + a_2 + Q) z^{w_1 - 1} \quad . \tag{2.5}$$

For $0 \le z \le 1$, $(a_1+a_2z+Q) \ge (a_1z+a_2+Q)$, so that $f(a_1+a_2z+Q) \le f(a_1z+a_2+Q)$. Since $0 \le w_1 \le w_2$, $z^{w_2} \le z^{w_1}$, and (2.4) holds.||

The result for $U(a)$ is more delicate and a pointwise argument does not carry through.

Theorem 2. If $a_1 \ge \ldots a_k \ge 0$, $0 \le w_1 \le \ldots \le w_k$, and

$$U(a;w) = \int_{a_1}^{\infty} \cdots \int_{a_k}^{\infty} \frac{\prod_1^k t_i^{w_i-1} dt_i}{(1+\Sigma t_i)^d} , \tag{2.6}$$

where $d \ge \Sigma w_i$, then $-U(a;w)$ is a Schur function.

Proof. Because of symmetry, we need only consider $(i,j) = (1,2)$ in showing that (2.1) holds for $-U(a;w)$. It is immediate that

$$\frac{\partial U(a;w)}{\partial a_1} = -\int_{a_2}^{\infty} \cdots \int_{a_k}^{\infty} \frac{a_1^{w_1-1} \prod_2^k t_i^{w_i-1} dt_i}{(1+a_1+t_2+\ldots+t_k)^d} . \tag{2.7}$$

Let $t_2-a_2 = z$ and interchange order of integration; then (2.7) becomes

$$\frac{\partial U(a;w)}{\partial a_1} = -\int_{a_3}^{\infty} \cdots \int_{a_k}^{\infty} (\prod_3^k t_i^{w_i-1} dt_i) \int_0^{\infty} \frac{a_1^{w_1-1} (z+a_2)^{w_2-1} dz}{(1+a_1+a_2+z+t_3+\ldots+t_k)^d} . \tag{2.8}$$

We now use a pointwise argument on the inner integral with $\sum_3^k t_i$ fixed. Let $z = (1+a_1+a_2+\sum_3^k t_i)v \equiv sv$ then the inner integral becomes

$$\int_0^\infty \frac{a_1^{w_1-1}(sv+a_2)^{w_2-1}}{s^{d-1}(1+v)^d}\,dv \quad . \tag{2.9}$$

Consequently, if $a_1 \geq a_2$, $\left(\frac{\partial U(a;w)}{\partial a_1} - \frac{\partial U(a;w)}{\partial a_2}\right) \leq 0$ provided

$$\int_0^\infty \frac{[a_1^{w_1-1}(sv+a_2)^{w_2-1} - a_2^{w_2-1}(sv+a_1)^{w_1-1}]}{(1+v)^d}\,dv \geq 0 \quad . \tag{2.10}$$

The ordering $0 \leq w_1 \leq w_2$ guarantees that the integrand be non-negative, so that (2.10) holds.‖

3. An application to the multinomial distribution. Let $X = (X_1,\ldots,X_k)$ have the multinomial distribution

$$P\{X=x\} = \binom{n}{x_1,\ldots,x_k} \prod_1^k \theta_i^{x_i} \quad , \tag{3.1}$$

where $x = (x_1,\ldots,x_k)$, $\Sigma x_i = n$, $\theta_1 \geq \cdots \geq \theta_k \geq 0$, $\Sigma\theta_i = 1$, and consider the tail probability $P\{X_1 \geq r,\ldots,X_k \geq r|\theta_1,\ldots,\theta_k\}$, with $r \leq n/k$.

Alam (1970) obtains lower and upper bounds for $P\{X_1 \geq r,\ldots,X_k \geq r|$ $\theta_1,\ldots,\theta_k\}$ by averaging some of the θ's, namely

$$P\{X_1 \geq r,\ldots,X_k \geq r \mid \theta^*,\theta_k,\ldots,\theta_k\} \leq P\{X_1 \geq r,\ldots,X_k \geq r \mid \theta_1,\ldots,\theta_k\} \quad (3.2)$$

$$\leq P\{X_1 \geq r,\ldots,X_k \geq r \mid \bar{\theta},\ldots,\bar{\theta}\} \quad ,$$

where $\theta^* = 1-(k-1)\theta_k$ and $\bar{\theta} = \Sigma\theta_i/k$. It has been shown by Olkin and Sobel (1965) that $P\{X_1 \geq r,\ldots,X_k \geq r \mid \theta_1,\ldots,\theta_k\}$ has a representation in terms of the Dirichlet integral

$$P\{X_1 \geq r,\ldots,X_k \geq r \mid \theta_1,\ldots,\theta_k\} = \kappa(n,r,k) \int_0^{\theta_1} \cdots \int_0^{\theta_k} (1-\sum_1^k t_i)^d \prod_1^k t_i^{r-1} dt_i \quad , \quad (3.3)$$

where $d=n-kr \geq 0$, and $\kappa(n,r,k) = \Gamma(n+1)/[\{\Gamma(r)\}^k \Gamma(n-kr+1)]$. We may now make use of Theorem 1 with $w_1 = \ldots = w_k = r$. Thus, if $0 \leq p_i \leq 1$, $0 \leq q_i \leq 1$, and $(p_1,\ldots,p_k) \succ (q_1,\ldots,q_k)$, then

$$P\{X_1 \geq r,\ldots,X_k \geq r \mid p_1,\ldots,p_k\} \leq P\{X_1 \geq r,\ldots,X_k \geq r \mid q_1,\ldots,q_k\} \quad . \quad (3.4)$$

The results of Alam are special cases of (3.4) since

$$(\theta^*,\theta_k,\ldots,\theta_k) \succ (\theta_1,\ldots,\theta_k) \succ (\bar{\theta},\ldots,\bar{\theta}) \quad .$$

Clearly, many other intermediate bounds can now be obtained.

For the lower tail of the multinomial distribution (3.1), we have the representation

$$C(\theta;r) \equiv P\{X_1 < r,\ldots,X_m < r \mid \theta_1,\ldots,\theta_k\}/\kappa(n,r,m) \tag{3.5}$$

$$= \int_{\theta_1}^{1-\sum_2^m \theta_i} \int_{\theta_2}^{1-\sum_3^m \theta_i - t_1} \cdots \int_{\theta_m}^{1-\sum_1^{m-1} t_i} (1-\Sigma t_i)^d \prod_1^m t_i^{r-1}\, dt_i \ ,$$

where $m \leq \min(k-1,n)$ and $d = n-kr \geq 0$. As in Theorem 1, a direct differentiation of (3.5) with respect to θ_1, followed by the change of variables $t_2-\theta_2 = v$, $t_j-\theta_j = z_j$, $j=3,\ldots,m$, yields

$$\frac{\partial C(\theta;r)}{\partial\theta_1} = -\int_0^{\theta_0}\int_0^{\theta_0-v} \cdots \int_0^{\theta_0-v-\sum_3^{m-1} z_j} \{\theta_1(v+\theta_2)\prod_3^m(z_j+\theta_j)\}^{r-1}(\theta_0-v-\sum_3^m z_j)^d\, dr \prod_3^m dz_j \ ,$$

where $\theta_0 = 1-\sum_1^m \theta_i$. By symmetry,

$$\frac{\partial C(\theta;r)}{\partial\theta_2} = -\int_0^{\theta_0}\int_0^{\theta_0-v} \cdots \int_0^{\theta_0-v-\sum_3^{m-1} z_j} \{\theta_2(v+\theta_1)\prod_3^m(z_j+\theta_j)\}^{r-1}(\theta_0-v-\sum_3^m z_j)^d\, dv \prod_3^m dz_j \ .$$

That $-C(\theta;r)$ is a Schur function follows from the fact that $(\theta_1-\theta_2)\{(\theta_1 v+\theta_1\theta_2)^{r-1} - (\theta_2 v+\theta_1\theta_2)^{r-1}\} \geq 0$ for all v.

As a consequence, we have the

<u>Corollary</u>. If $(p_1,\ldots,p_k) \succ (q_1,\ldots,q_k)$, then

$$P\{X_1 < r,\ldots,X_m < r \mid q_1,\ldots,q_k\} \geq P\{X_1 < r,\ldots,X_m < r \mid p_1,\ldots,p_k\} \ .$$

4. An application to the simultaneous analysis of variance model. Suppose two hypotheses are tested using the same error variance for each test, so that we have

$$F_1 = \frac{q_1/n_1}{q_0/n_0}, \quad F_2 = \frac{q_2/n_1}{q_0/n_0},$$

where q_0, q_1, and q_2 are independently distributed as χ^2 variates with n_0, n_1, and n_2 d.f. respectively. Kimball (1951) obtained the inequality

$$P\{F_1 \leq F_{1\alpha}, F_2 \leq F_{2\alpha}\} \geq P\{F_1 \leq F_{1\alpha}\} P\{F_2 \leq F_{2\alpha}\}, \tag{4.1}$$

where $F_{1\alpha}$ and $F_{2\alpha}$ are the 100 α percent points of the distributions of F_1 and F_2. This inequality is of interest in that it provides a bound for the probability of making no errors of the first kind. We may use Theorem 2 to obtain a bound for $P\{F_1 \geq F_{1\alpha}, F_2 \geq F_{2\alpha}\}$. Suppose that $n_1=n_2=n$, then

$$P\{F_1 \geq F_\alpha, F_2 \geq F_\alpha\} = k \int_c^\infty \int_c^\infty \frac{x^{\frac{1}{2}n-1} y^{\frac{1}{2}n-1}}{(1+x+y)^{\frac{1}{2}n_0+n}} \, dx\, dy, \tag{4.2}$$

where $c = nF_\alpha/n_0$ and k is a normalizing constant. Since $(2c,0) \succ (c,c)$, we obtain from Theorem 2 that

$$P\{F_1 \geq F_\alpha, F_2 \geq F_\alpha\} \geq P\{F_1 \geq 2F_\alpha, F_2 \geq 0\} = P\{F_1 \geq 2F_\alpha\}. \tag{4.3}$$

Since

$$P\{F_1 \le F_{1\alpha}, F_2 \le F_{2\alpha}\} = P\{F_1 \le F_{1\alpha}\} + P\{F_2 \le F_{2\alpha}\} + P\{F_1 \ge F_{1\alpha}, F_2 \ge F_{2\alpha}\} - 1 , \quad (4.4)$$

we obtain an alternative inequality to that of (4.1), namely,

$$P\{F_1 \le F_\alpha, \; F_2 \le F_\alpha\} \ge 2P\{F_1 \le F_\alpha\} - P\{F_1 \le 2F_\alpha\} \; . \quad (4.5)$$

This is to be compared with

$$P\{F_1 \le F_\alpha, \; F_2 \le F_\alpha\} \ge [P\{F_1 \le F_\alpha\}]^2 \; . \quad (4.6)$$

We wish the larger bound, so that we need to determine the sign of

$$[P\{F_1 \le F_\alpha\}]^2 - 2P\{F_1 \le F_\alpha\} + 1 - P\{F_1 \ge 2F_\alpha\} = [P\{F_1 \ge F_\alpha\}]^2 - P\{F_1 \ge 2F_\alpha\} \; . \quad (4.7)$$

It turns out that the difference is not always of one sign. When $n=2$,

$$[P\{F \ge c\}]^2 = (1 + \frac{2c}{n_0})^{-n_0} < (1 + \frac{4c}{n_0})^{\frac{-n_0}{2}} = P\{F \ge 2c\} \quad ,$$

so that (4.5) yields a better bound than (4.6). With $n_0 \to \infty$, the difference becomes

$$[P\{X_n^2 \ge c\}]^2 - P\{X_n^2 \ge 2c\} \; . \quad (4.8)$$

When $n=2m$, $P\{X_n^2 \geq c\} = \sum_0^{m-1} c^j/j!$, and a straightforward analysis shows that (4.8) is nonnegative.

For small values of c (less than $c_0(n,n_0)$), $[P\{F \geq c\}]^2$ is larger than $P\{F \geq 2c\}$, whereas for $c > c_0(n,n_0)$, $P\{F \geq 2c\}$ is larger than $[P\{F \geq c\}]^2$, where $c_0(n,n_0)$ depends on n and n_0. As either n or n_0 increase, the constant $c_0(n,n_0)$ tends to increase. Since c will, in general, be of moderate size, it appears that (4.5) is the better bound in practise.

Acknowledgement. The author is indebted to M. Sobel and V.R.R. Uppuluri for their comments, and for pointing out an application which led to Theorem 2; and to R. Miller for his suggestions concerning Section 4.

REFERENCES

Alam, K. (1970). Monotonicity properties of the multinomial distribution. Ann. Math. Statist. 41, 315-317.

Berge, C. (1963). Topological spaces including a treatment of multi-valued functions, vector spaces, and convexity. (Trans. by E. M. Patterson). London: Oliver and Boyd.

Kimball, A. W. (1951). On dependent tests of significance in the analysis of variance. Ann. Math. Statist. 22, 600-602.

Marshall, A. W., Olkin, I. and Proschan, F. (1967). Monotonicity of ratios of means and other applications of majorization. Inequalities-I, ed. by O. Shisha, Academic Press, New York.

Olkin, I. and Sobel, M. (1965). Integral expressions for tail probabilities of the multinomial and the negative multinomial distributions. Biometrika 52, 167-179.