Reprinted from

S A N K H Y Ā

THE INDIAN JOURNAL OF STATISTICS

SERIES A, VOLUME 47, PART 1, 1985

RAO'S DISTANCE FOR NEGATIVE MULTINOMIAL DISTRIBUTIONS*

By

JOSEP M. OLLER and CARLES M. CUADRAS

STATISTICAL PUBLISHING SOCIETY

CALCUTTA

* *(versión revisada)*

Sankhyā: The Indian Journal of Statistics
1985, Volume **47**, Series A, Pt. 1, pp. 75-83.

RAO'S DISTANCE FOR NEGATIVE MULTINOMIAL DISTRIBUTIONS

By JOSEP M. OLLER and CARLES M. CUADRAS
University of Barcelona, Spain

SUMMARY. Rao (1945) proposed a method, based on Fisher's information matrix, for measuring distance between distributions of a parametric family satisfying certain regularity conditions. In this paper, Rao's (1945) method is applied to obtain the distance between two negative multinomial distributions. Some other properties are discussed too.

1. Introduction

The question of introducing a distance between different statistical populations has been considered by various authors. If we assume that all the information for constructing such a distance is contained in the probability density function of a random vector X, supposedly existing and restricted to each population, it will not be generally satisfactory to characterize each populations by their mean value of the random vector X, since the latter does not determine uniquely the probability density function associated with each population.

A reasonable alternative would be to allow that the probability density function of the random vector X, in any of the populations studied, to belong to a certain parametric family, $p(\cdot|\theta)$. Thus, a population could be characterized by $\theta = (\theta^1, \ldots, \theta^n)$, element of a parametric space Ω.

It is also reasonable to require that the proposed distance on the parametric space Ω, possesses the property of being invariant under any admissible transformation of the parameters, since the latter does not affect the probability density function, $p(\cdot|\theta)$, of the random vector X. In addition, the distance has to be invariant for admissible transformations of the random vector X, since it must be independent of the method by which the measurements are attained.

The method proposed by Rao (1945) and studied later by Atkinson and Mitchell (1981) and Burbea and Rao (1982a, b), allow us to define a distance on the parametric space Ω with the above mentioned characteristics. If the

---

AMS (1980) subject classifications : 62H30, 94A17, 53A35.

Key words and phrases : Information metric, geodesic distance in probability, multinomial, negative distributions.

parametric family satisfy certain regularity conditions, the Fisher information matrix defines a covariant symmetric tensor field of the second order on the parametric space Ω,

$$g_{ij} = E\left(\frac{\partial \ln p(X|\theta)}{\partial \theta^i}\frac{\partial \ln p(X|\theta)}{\partial \theta^j}\right) \quad (i,j=1,\ldots,n) \qquad \ldots \quad (1.1)$$

and may also be taken as a metric tensor field on Ω, thereby rendering Ω as a Riemannian manifold. The Rao distance between two points θ_A and θ_B of Ω is then defined as their geodesic distance with respect to the metric induced by (1.1).

In this paper, Rao's (1945) method is applied to obtain a distance between two negative multinomial distributions. We also discuss some other differential geometric properties of these distributions.

2. Development

Let F be a holomorphic function in the unit disk $\Delta = \{z \in \mathbb{C} : |z| < 1\}$ with a power-series expansion

$$F(z) = \sum_{m=0}^{\infty} b_m z^m \quad (z \in \Delta) \qquad \ldots \quad (2.1)$$

with $b_m \geqslant 0 \ \ (m = 0,1,\ldots)$ and not all b_m are zero. For $\alpha = (\alpha_1,\ldots,\alpha_n) \in \mathbb{Z}_+^n$ and $\theta = (\theta^1,\ldots,\theta^n) \in \mathbb{C}^n$, using multinomial notation, we write

$$|\alpha| = \alpha_1 + \cdots + \alpha_n \quad , \quad \alpha! = \alpha_1! \cdots \alpha_n!$$
$$\theta^\alpha = (\theta^1)^{\alpha_1} \cdot \cdots \cdot (\theta^n)^{\alpha_n} \quad , \quad \Theta = \theta^1 + \cdots + \theta^n. \qquad \ldots \quad (2.2)$$

Define

$$\Omega = \left\{\theta = (\theta^1,\ldots,\theta^n) \in \mathbb{R}^n : \ \theta^j > 0 \ \ (j=1,\ldots,n), \Theta < 1\right\} \qquad \ldots \quad (2.3)$$

and let

$$p(\alpha|\theta) = \frac{1}{F(\Theta)}\frac{|\alpha|!}{\alpha!} b_{|\alpha|}\theta^\alpha \ \ (\alpha \in \mathbb{Z}_+^n, \theta \in \Omega), \qquad \ldots \quad (2.4)$$

then $p(\alpha|\theta)$ is a probability distribution defined on $\mathbb{Z}_+^n \times \Omega$, where $\mathbb{Z}_+^n$ is the sample space and Ω is the parameter space.

In this case, a use of (1.1) and (2.4) gives

$$g_{ij}(\theta) = f(\Theta)\left(\frac{1}{\theta^i}\delta_{ij} + \frac{f'(\Theta)}{f(\Theta)}\right) \quad (i,j=1,\ldots,m) \qquad \ldots \quad (2.5)$$

where

$$f(z) = \frac{F'(z)}{F(z)} \quad (z \in \Delta). \qquad \ldots \quad (2.6)$$

If we choose

$$F(z) = (1-z)^{-r} = \sum_{m=0}^{\infty} \frac{1}{m!} \frac{\Gamma(m+r)}{\Gamma(r)} z^m \qquad (r > 0, \quad z \in \Delta) \qquad \ldots \ (2.7)$$

we obtain the family of negative multinomial distributions $\{P_r(\cdot|\cdot)\}$, $r > 0$,

$$P_r(\alpha|\theta) = \frac{\Gamma(|\alpha|+r)}{\alpha!\,\Gamma(r)} \theta^{\alpha}(1-\Theta)^r \qquad \ldots \ (2.8)$$

In this case

$$f(z) = \frac{F'(z)}{F(z)} = (\log F(z))' = r(1-z)^{-1} \qquad \frac{f'(z)}{f(z)} = (\log f(z))' = (1-z)^{-1}, \qquad \ldots \ (2.9)$$

therefore,

$$g_{ij}(\theta) = \frac{r}{\theta^{n+1}} \left(\frac{1}{\theta^i}\delta_{ij} + \frac{1}{\theta^{n+1}} \right) \qquad (i, j = 1, \ldots, n) \qquad \ldots \ (2.10)$$

where $\theta^{n+1} = 1 - \Theta$ and δ_{ij} is a Kronecker delta. It follows that the tensor in (2.10) is positive definite on the parametric space Ω, and thus Ω is a Riemannian manifold (see also Hicks, 1965).

To analyze whether the Riemannian manifold Ω is Euclidean or not, that is, whether or not there exists an admissible transformation of the coordinates (parameters) which reduces the metric tensor field to a constant tensor field, we proceed to calculate the Riemann-Christoffel tensor of the first kind (covariant curvature tensor) R_{hijk} of the metric (2.10). This gives

$$R_{hijk} = \frac{r}{4\theta^i\theta^h(\theta^{n+1})^2} \left(\delta_{ij}\delta_{kh} - \delta_{ik}\delta_{jh} + \frac{\theta^i(\delta_{kh} - \delta_{jh}) + \theta^h(\delta_{ij} - \delta_{ik})}{\theta^{n+1}} \right)$$

$$(h, i, j, k = 1, \ldots, n). \qquad \ldots \ (2.11)$$

Evidently, $R_{hijk} \neq 0$ if and only if $n > 1$, in which case the manifold Ω is not Euclidean. An alternative expression for R_{hijk} is obtained by using (2.10) and (2.11), namely

$$R_{hijk} = -\frac{1}{4r}(g_{hj}g_{ik} - g_{hk}g_{ij}) \qquad (h, i, j, k = 1, \ldots, n). \qquad \ldots \ (2.12)$$

The Riemannian curvature K is therefore

$$K = -\frac{1}{4r}. \qquad \ldots \ (2.13)$$

It follows that the space Ω, is isotropic, and has a constant negative curvature. Therefore, it is locally isometric to the Poincaré hyperbolic space. An interesting property is deduced from this : given $\theta_a, \theta_B \in \Omega$ there is only geodesic line that joins both points (see Hicks, 1965 and Spivak, 1979).

In order to obtain the distance between two points of the parametric space Ω, where each point equivalent to a negative multinomial distribution, we have to calculate the geodesic equations. Taking into account that the inverse of the tensor-metric g_{ij} is given by

$$g^{ij} = \frac{\theta^{n+1}}{r}\left(\delta^{ij}\theta^i - \theta^i\theta^j\right) \qquad (i,j = 1,\ldots,n) \qquad \ldots \quad (2.14)$$

and defining $\delta_{ijk} = \delta_{ij}\delta_{ik}$, the Christoffel symbols of the second kind are

$$\Gamma^k_{ij} = -\frac{\delta_{ijk}}{2\theta^i} + \frac{\delta_{ik}+\delta_{jk}}{2\theta^{n+1}} \qquad (i,j,k = 1,\ldots,n). \qquad \ldots \quad (2.15)$$

Therefore, the differential equations of the geodesics may be written as

$$\frac{d^2\theta^k}{ds^2} - \frac{1}{2\theta^k}\left(\frac{d\theta^k}{ds}\right)^2 + \frac{1}{\theta^{n+1}}\frac{d\theta^k}{ds}\sum_{i=1}^{n}\frac{d\theta^i}{ds} = 0 \qquad (k = 1,\ldots,n). \qquad \ldots \quad (2.16)$$

Summing up the above n equations gives

$$\frac{d^2\theta^{n+1}}{ds^2} + \sum_{k=1}^{n}\frac{1}{2\theta^k}\left(\frac{d\theta^k}{ds}\right)^2 - \frac{1}{\theta^{n+1}}\left(\frac{d\theta^{n+1}}{ds}\right) = 0, \qquad \ldots \quad (2.17)$$

and taking into account that the vector $\left(\frac{d\theta^1}{ds},\ldots,\frac{d\theta^n}{ds}\right)$ has unit length, that is

$$\frac{r}{\theta^{n+1}}\left(\sum_{k=1}^{n}\frac{1}{\theta^k}\left(\frac{d\theta^k}{ds}\right)^2 + \frac{1}{\theta^{n+1}}\left(\frac{d\theta^{n+1}}{ds}\right)^2\right) = 1, \qquad \ldots \quad (2.18)$$

we obtain from (2.17) and (2.18) that

$$\frac{1}{\theta^{n+1}}\frac{d^2\theta^{n+1}}{ds^2} - \frac{3}{2}\left(\frac{1}{\theta^{n+1}}\frac{d\theta^{n+1}}{ds}\right)^2 + \frac{1}{2r} = 0. \qquad \ldots \quad (2.19)$$

The latter equation may be resolved via the transformation

$$u = \frac{1}{\theta^{n+1}}\frac{d\theta^{n+1}}{ds}, \qquad \ldots \quad (2.20)$$

yielding

$$u = -\frac{1}{\sqrt{r}}\tanh\left(\frac{s}{2\sqrt{r}} + C\right) \qquad \ldots \quad (2.21)$$

where C is an integration constant.

From (2.21) and (2.16) we obtain

$$\frac{1}{\theta^k}\frac{d^2\theta^k}{ds^2} - \frac{1}{2}\left(\frac{1}{\theta^k}\frac{d\theta^k}{ds}\right)^2 + \frac{1}{\theta^k}\frac{d\theta^k}{ds}\frac{1}{\sqrt{r}}\tanh\left(\frac{s}{2\sqrt{r}} + C\right) = 0 \qquad (k = 1,\ldots,n). \qquad \ldots \quad (2.22)$$

Equations (2.22) have the particular solution

$$\theta^k = A_k = \text{const.} \quad (k = 1, \ldots, n). \qquad \ldots \quad (2.23)$$

More general solutions can be obtained through the transformation

$$z^k = \theta^k \left(\frac{d\theta^k}{ds} \right)^{-1} \quad (k = 1, \ldots, n), \qquad \ldots \quad (2.24)$$

which reduces (2.22) to an independent set of linear differential equations of the form

$$\frac{dz^k}{ds} - \frac{1}{\sqrt{r}} \tanh\left(\frac{s}{2\sqrt{r}} + C \right) z^k = \frac{1}{2} \quad (k = 1, \ldots, n). \qquad \ldots \quad (2.25)$$

The general solution (2.25) is

$$z^k = \cosh^2\left(\frac{s}{2\sqrt{r}} + C \right) \left(\sqrt{r} \tanh\left(\frac{s}{2\sqrt{r}} + C \right) + B_k \right) \quad (k = 1, \ldots, n), \quad \ldots \quad (2.26)$$

where the B_k are constants of integration.

It follows, since

$$\ln \theta^k = \int \frac{1}{z_k} ds \quad (1, \ldots, n), \qquad \ldots \quad (2.27)$$

that

$$\theta^k = A_k \left(B_k + \tanh\left(\frac{s}{2\sqrt{r}} + C \right) \right)^2 \quad (k = 1, \ldots, n). \qquad \ldots \quad (2.28)$$

Equations (2.23) and (2.28) describe the geodesic lines, with the A_k as integration constants.

The integration constants will be chosen in such a way that they accomplish (2.18), and satisfy the initial conditions, that is, for $s = 0$ the geodesic departs from a point A, with coordinates $\left(a^1, \ldots, a^n\right)$ and for a positive value $s = d$, reaches the point B, $\left(b^1, \ldots, b^n\right)$. After some laborious calculations we obtain that the geodesic distance between A and B is given by

$$d = 2\sqrt{r} \cosh^{-1} \left(\frac{1 - \sum_{j=1}^{n} \sqrt{a^j b^j}}{\sqrt{a^{n+1} b^{n+1}}} \right) \qquad \ldots \quad (2.29)$$

Alternatively, (see also Oller and Cuadras, 1982),

$$d = 2\sqrt{r} \ln \left(\frac{1 - \sum_{j=1}^{n} \sqrt{a^j b^j} + \sqrt{\left(1 - \sum_{j=1}^{n} \sqrt{a^j b^j} \right)^2 - a^{n+1} b^{n+1}}}{\sqrt{a^{n+1} b^{n+1}}} \right) \qquad \ldots \quad (2.30)$$

3. Other results

(A). From (2.29) it follows that the distance between two negative multinomial distributions is not bounded. However, if we fix a^{n+1} and b^{n+1}, is easy to prove that

$$d \leqslant 2\sqrt{r}\cosh^{-1}\left(\frac{1}{\sqrt{a^{n+1}b^{n+1}}}\right) < \sqrt{r}\ln\left(\frac{4}{a^{n+1}b^{n+1}}\right). \qquad \ldots \ (3.1)$$

(B). It is also possible to establish a relationship between (2.29) and Bhattacharyya's distance (1946). This last distance can be obtained applying Rao's (1945) method to multinomial distribution (see Rao, 1949).

If we consider the multinomial distribution's parametric family, with known index N and parameters $\theta^1,\ldots,\theta^{n+1}$, where $\sum_{j=1}^{n+1}\theta^j = 1$, the Bhattacharyya's distance, d_B, between two points of the parametric space, $(a^1,\ldots,a^n)$ $(b^1,\ldots,b^n)$ is given by

$$d_B = 2\sqrt{N}\cos^{-1}\left(\sum_{j=1}^{n+1}\sqrt{a^j b^j}\right) \qquad \ldots \ (3.2)$$

Hence, if we fix a^{n+1} and b^{n+1}, both distances in (2.29) and (3.2), are decreasing functions of $\sum_{j=1}^{n}\sqrt{a^j b^j}$, and, following Shepard (1962), we conclude that the preorders associated with (2.29) and (3.2) are equal.

(C). When d is very small, we have

$$\cosh\left(\frac{d}{2\sqrt{r}}\right) \simeq 1 + \frac{d^2}{8r}, \qquad \ldots \ (3.3)$$

and thus

$$d \simeq 2\sqrt{2r\left(\frac{1-\sum_{j=1}^{n+1}\sqrt{a^j b^j}}{\sqrt{a^{n+1}b^{n+1}}}\right)}. \qquad \ldots \ (3.4)$$

It follows from (3.2) and (3.4) that

$$d \simeq 4\sqrt{r}\frac{\sin\left(\dfrac{d_B}{4\sqrt{N}}\right)}{\sqrt[4]{a^{n+1}b^{n+1}}} \simeq \sqrt{\frac{r}{N}}\frac{d_B}{\sqrt[4]{a^{n+1}b^{n+1}}} \qquad \ldots \ (3.5)$$

Thus, when a^{n+1} and b^{n+1} are fixed, d and d_B are nearly proportional, for small values of d.

(D). Let E sample space and consider two finite partitions of E, $\mathcal{P}_1 = \{A_1,\ldots,A_{n+1}\}$ and $\mathcal{P}_2 = \{B_1,\ldots,B_{m+1}\}$. Let $\mathcal{P}_2$ be a refinement of

$\mathcal{P}_1$, and thus $B_{m+1} \subset A_{n+1}$, and consider the parametric family of the negative multinomial distributions (2.8), with parameters $\theta^i = \Pr(A_i)$ and $\theta^i = \Pr(B_i)$ respectively (notice that there are n independent parameters in the first case and m in the second one). Let A and B two statistical populations with

$$\begin{array}{lll} \Pr(A_i) = a^i \ (i = 1, \ldots, n) & \Pr(B_j) = p^j \ (j = 1, \ldots, m) & \text{(population } A) \\ \Pr(A_i) = b^i \ (i = 1, \ldots, n) & \Pr(B_j) = q^j \ (j = 1, \ldots, m) & \text{(population } B) \end{array} \quad \ldots \ (3.6)$$

then

$$d_1 = 2\sqrt{r}\cosh^{-1}\left(\frac{1 - \sum_{j=1}^{n} \sqrt{a^j b^j}}{\sqrt{a^{n+1} b^{n+1}}}\right) \leqslant d_2 = 2\sqrt{r}\cosh^{-1}\left(\frac{1 - \sum_{j=1}^{m} \sqrt{p^j q^j}}{\sqrt{p^{m+1} q^{m+1}}}\right) \quad \ldots \ (3.7)$$

Actually, it is sufficient to consider first the case

$$A_i = B_{i_1} \bigcup \cdots \bigcup B_{i_{s_i}} \qquad A_i \in \mathcal{P}_1,\ B_{i_j} \in \mathcal{P}_2 \quad (i = 1, \ldots, n+1) \quad \ldots \ (3.8)$$

and

$$a^i = p^{i_1} + \cdots + p^{i_{s_i}} \qquad b^i = q^{i_1} + \cdots + q^{i_{s_i}} \quad (i = 1, \ldots, n+1) \quad \ldots \ (3.9)$$

since $\mathcal{P}_2$ is a refinement of $\mathcal{P}_1$

$$\sqrt{a^i b^i} \geqslant \sum_{j=1}^{s_i} \sqrt{p^{i_j} q^{i_j}} \qquad (i = 1, \ldots, n+1) \quad \ldots \ (3.10)$$

because
$$\left(\sum_{j=1}^{s_i} \sqrt{p^{i_j} q^{i_j}}\right)^2 = a^i b^i - \sum_{j<k} \left(p^{i_j} q^{i_k} + p^{i_k} q^{i_j}\right) + 2\sum_{j<k} \sqrt{p^{i_j} q^{i_j} p^{i_k} q^{i_k}} \leqslant a^i b^i \ \ (i = 1, \ldots, n+1). \cdots \ (3.11)$$

On the other hand, if

$$A_{n+1} = B_{j_1} \bigcup \cdots \bigcup B_{j_h} \bigcup B_{m+1} \qquad A_{n+1} \in \mathcal{P}_1,\ B_s \in \mathcal{P}_2 \quad \ldots \ (3.12)$$

then

$$\sqrt{a^{n+1} b^{n+1}} \geqslant \sum_{k=1}^{h} \sqrt{p^{j_k} q^{j_k}} + \sqrt{p^{m+1} q^{m+1}}, \quad \ldots \ (3.13)$$

but since

$$1 - \sum_{j=1}^{n+1} \sqrt{a^j b^j} \geqslant 0 \quad \ldots \ (3.14)$$

we find that

$$\frac{1 - \sum_{j=1}^{n} \sqrt{a^j b^j}}{\sqrt{a^{n+1} b^{n+1}}} \leqslant \frac{1 - \sum_{j=1}^{n} \sqrt{a^j b^j} - \sum_{k=1}^{h} \sqrt{p^{j_k} q^{j_k}}}{\sqrt{p^{m+1} q^{m+1}}}. \quad \ldots \ (3.15)$$

However, since

$$\sum_{j=1}^{n}\sqrt{a^{j_k}b^{j_k}}+\sum_{k=1}^{h}\sqrt{p^{j_k}q^{j_k}}\geqslant\sum_{j=1}^{m}\sqrt{p^{j}q^{j}} \qquad \ldots \quad (3.16)$$

(3.7) follows at once, that is, $d_1 \leqslant d_2$: refining the partition associated with negative multinomial distribution, the distance (2.29) increases.

(E). Let $X_1,\ldots,X_m$ be an independent set of random vectors. Each X_i is distributed as (2.8) with $r=r_i$, $n=n_i$ and $\theta^i=\theta^i_j$. Then, the parametric space is defined by

$$\Omega=\Omega_1\times\cdots\times\Omega_m \qquad \ldots \quad (3.17)$$

where

$$\Omega_i=\left\{\left(\theta_i^1,\ldots,\theta_i^{n_i}\right)\in\mathbb{R}^{n_i}\,|\,\theta_i^j\geqslant 0,\ \ j=1,\ldots,\sum_{j=1}^{n_i}\theta_i^j<1\right\}\ \ (i=1,\ldots,m). \qquad \ldots \quad (3.18)$$

Let A and B two points of Ω, with $A=\left(a_1^1,\ldots,a_1^{n_1},\ldots,a_m^1,\ldots,a_m^{n_m}\right)$ and $B=\left(b_1^1,\ldots,b_1^{n_1},\ldots,b_m^1,\ldots,b_m^{n_m}\right)$. It is not difficult to prove that the geodesic distance between A and B is given by

$$d=2\sqrt{\sum_{j=1}^{m}r_j\left(\cosh^{-1}\left(\frac{1-\sum_{k=1}^{n_j}\sqrt{a_j^k b_j^k}}{\sqrt{a_j^{n_j+1}b_j^{n_j+1}}}\right)\right)^2} \qquad \ldots \quad (3.19)$$

because the metric tensor's matrix of the parametric space Ω is of the form

$$G=\operatorname{diag}\left(G_1,\ldots,G_m\right) \qquad \ldots \quad (3.20)$$

where $G_1,\ldots,G_m$ are the metric tensor's matrix associated with the parametric spaces $\Omega_1,\ldots,\Omega_m$ given by (3.18).

(F). We consider two particular cases :

(i) Negative binomial distribution. This is given by (2.8) by taking $n=1$. The obtained distance is

$$d=2\sqrt{r}\cosh^{-1}\left(\frac{1-\sqrt{a^1b^1}}{\sqrt{(1-a^1)(1-b^1)}}\right). \qquad \ldots \quad (3.21)$$

(ii) Geometric distribution. This is given by (2.8) by taking $n=1$, $r=1$. The distance is now

$$d=2\cosh^{-1}\left(\frac{1-\sqrt{a^1b^1}}{\sqrt{(1-a^1)(1-b^1)}}\right). \qquad \ldots \quad (3.22)$$

In both cases, the parametric space is Euclidean.

4. Remarks

In practice, each negative multinomial is characterized by maximum likelihood estimates of the θ^i. Substituting the latter (2.29), the maximum likelihood estimate of d is obtained. If we have m negative multinomial distributions, $N_1, \ldots, N_m$, we can estimate the distances $d^*(N_i, N_j)$ among them. Then, we would obtain the interdistances-matrix $\Delta = (d^*(N_i, N_j))$, and thus we can apply MDS (Multidimensional Scaling) techniques (Cuadras, 1981) to obtain a representation of the distributions as points in a low dimensional Euclidean space, generally a plane.

It is also possible to obtain a hierarchic classification of the distributions and its graphic output, the dendrogram, by using numerical taxonomy methods.

References

ATKINSON, C. and MITCHELL, A. F. S. (1981): Rao's distance measure. *Sankhyā*, **43**, A, 345-365.

BHATTACHARYYA, A. (1946): On a measure of divergence between two multinomial populations. *Sankhyā*, **7**, 401-406.

BURBEA, J. and RAO, C. R. (1982 a): Entropy differential metric, distance and divergence measures in probability spaces: a unified approach. *Jour. Multivariate Analysis*, **12**, 575-596.

——— (1982 b): Differential metrics in probability spaces. *Probability Math. Statist.*, **12**, 115-132.

CUADRAS, C. M. (1981): *Métodos de Análisis Multivariante*, Eunibar, Barcelona.

HICKS, N. J. (1965): *Notes on Differential Geometry*, Van Nostran, Princeton.

OLLER, J. M. and CUADRAS, C. M. On a defined distance for negative multinomial distribution. Abstracts of *XIth Inter. Biometric Conf.*, Toulouse, 69.

RAO, C. R. (1945): Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.*, **27**, 81-91.

——— (1949): On the distance between two populations. *Sankhyā*, **9**, 246-248.

SHEPARD, R. N. (1962): The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, **27**, 219-246.

SPIVAK, M. (1979): *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, Inc. Berkeley.

Paper received: August, 1983.

Revised: December, 1983.

Printed by Haradhan Chakrabarti at EKA PRESS, 204/1, Barrackpore Trunk Road,
Calcutta-35 and published by K. B. Goswami from Statistical Publishing Society,
204/1, Barrackpore Trunk Road, Calcutta 700 035.