Publ. Math. Debrecen
63/1-2 (2003), 51–66

Information geometry of random walk distribution

By NASSAR H. ABDEL-All (Assiut),
H. N. ABD-ELLAH (Assiut) and H. M. MOUSTAFA (Assiut)

Abstract. Information geometry (Geometry and Nature) has emerged from the study of invariant properties of the manifold of probability distributions. It is regarded as a mathematical discipline having rapidly developing areas of applications as well as giving new trends in geometrical and topological methods.

Information geometry has many applications belonging to many different fields, for instance, statistical inference, linear systems and time series, neural networks and nonlinear systems, linear programing, convex analysis and completely integrable dynamical systems, quantum information geometry and geometric modelling.

Here, we give a brief account of information geometry and the deep relationship between differential geometry and statistics [1], [4], [5], [10], [11]. The parameter space of a random walk distribution using its Fisher matrix is defined. The Riemannian and scalar curvatures of the parameter space are calculated. The differential equations of the geodesic are obtained and solved. The relations between the J-divergence and the geodesic distance in that space are found.

1. Geometrical and statistical preliminaries

The concept of a metric and a tensor is, in generality, well known, we turn now to statistical applications. Therefore, we give some formulas and definitions of the geometrical and statistical concepts which are necessary for later use.

Mathematics Subject Classification: 53B20, 62B10.
Key words and phrases: information geometry, statistical manifold and random walk distribution.

Definition 1.1 ([10]). A statistical model is a family S of probability distributions P of random variables which is smoothly parametrized by a finite number of real parameters θ^i as follows:

$$S = \{P_\theta : \theta \in R^n,\ \theta = (\theta^i)\}.$$

The statistical model S carries the structure of smooth Riemannian manifold, with respect to which $\theta = (\theta^i)$ plays the role of coordinates of a point $P_\theta \in S$, and whose metric $\langle\, ,\, \rangle$ is defined by the Fisher's information matrix $(g_{ij}(\theta))$, i.e.

$$dS^2 = \sum_{i,j=1}^{n} g_{ij}(\theta) d\theta^i d\theta^j, \tag{1.1}$$

where $g_{ij}(\theta)$ is given as:

$$g_{ij}(\theta) = \int \frac{\partial \ln f}{\partial \theta^i} \frac{\partial \ln f}{\partial \theta^j} f\ dx = -E\left[\frac{\partial^2 \ln f}{\partial \theta^i \partial \theta^j}\right], \tag{1.2}$$

where E denotes the expectation with respect to the distribution f [5]. Here $\left\{\frac{\partial \ln f}{\partial \theta^i}\right\}$ is a basis of a vector space of random variables, which can be identified with the tangent space $T_\theta\, S$. Thus, we can define the operator of covariant differentiation and some operator of different connection using the one-one correspondence between statistical model and Riemannian manifold. To illustrate this statement, let M an n-dimensional parameter space, $\theta = \left(\theta^1, \theta^2, \ldots, \theta^n\right)$ be an arbitrary parameter vector in M, x a random variable and $f(x;\theta)$ be the probability density function. Let (θ^i) and $(\theta^i + d\theta^i)$ be two neighbouring points. The infinitesimal distance dS between the points (θ^i) and $(\theta^i + d\theta^i)$ is defined by (1.1).

C. R. Rao [11] has proved that the Fisher's information matrix has the properties of a Riemannian metric and the parameter space M becomes a Riemannian space with a metric tensor $g_{ij}(\theta)$.

Definition 1.2 ([9]). Let $\theta^i = \theta^i(s)$ be the equations of a curve of class C^2 on the Riemannian manifold M. Then the curve is a geodesic (curve with minimal distance) if and only if

$$\frac{d^2\theta^i}{ds^2} + \sum_{j,k=1}^{n} \Gamma^i_{jk} \frac{d\theta^j}{ds} \frac{d\theta^k}{ds} = 0, \quad i = 1, \ldots, n \tag{1.3}$$

where s is the parameter of arc length and the quantities

$$\Gamma^i_{jk} = \frac{1}{2}\sum_l g^{il}\left(\frac{\partial g_{jl}}{\partial \theta^k} + \frac{\partial g_{kl}}{\partial \theta^j} - \frac{\partial g_{jk}}{\partial \theta^l}\right) \tag{1.4}$$

are called the Christoffel symbols of the 2nd kind and (g^{il}) is the contravariant metric tensor field of the covariant metric tensor field (g_{il}), i.e; $\sum_{k=1}^{2} g^{ik} g_{jk} = \delta^i_j$.

The Riemann curvature tensor is defined by [6]

$$R^l_{ijk} = \Gamma^l_{ik,j} - \Gamma^l_{ij,k} + \sum_{t=1}^{2}\left(\Gamma^l_{jt}\Gamma^t_{ik} - \Gamma^l_{kt}\Gamma^t_{ij}\right) \tag{1.5}$$

where the comma denotes the partial derivative. The Riemann tensor can be written using its covariant components as the following

$$R_{iklm} = \sum_{j=1}^{2} g_{ij} R^j_{klm}. \tag{1.6}$$

Thus we can obtain a tensor of rank two as

$$\left.\begin{aligned} R_{ij} &= \sum_{k,m} g^{km} R_{mijk} = \sum_k R^k_{ijk} \\ &= \sum_k \left(\Gamma^k_{ik,j} - \Gamma^k_{ij,k}\right) + \sum_{k,t=1}^{2}\left(\Gamma^k_{jt}\Gamma^t_{ik} - \Gamma^k_{kt}\Gamma^t_{ij}\right) \end{aligned}\right\}. \tag{1.7}$$

This tensor is called the Ricci tensor.

The scalar curvature R is defined by the contraction performed twice on the Ricci tensor [6], i.e.,

$$R = \sum_{i,j=1}^{2} g^{ij} R_{ij}. \tag{1.8}$$

In order to express the relation between the geodesic distance and the J-divergence, let us remind that the geodesic distance, the distance along a geodesic between two points p and q, is defined by [11]

$$S(p,q) = \int_p^q \sqrt{g_{ij}(\theta) d\theta^i d\theta^j}, \quad \theta^i = \theta^i(s). \tag{1.9}$$

On the other hand, the J-divergence between two extremely close probability density functions $p = f(x;\theta)$ and $q = f(x;\theta + d\theta)$ is given by [7], [8]

$$J(p,q) = \int (p-q)\ln\frac{p}{q}dx. \tag{1.10}$$

Explicitly, we have locally (Taylor expansion)

$$J(p,q) = \left[\int \frac{\partial \ln f}{\partial \theta^i}\frac{\partial \ln f}{\partial \theta^j} f\, dx\right] d\theta^i d\theta^j.$$

Taking (1.2) into account, we can represent the J-divergence in the form

$$J(p,q) = \sum_{i,j} g_{ij}(\theta) d\theta^i d\theta^j,$$

which, because of (1.1), means that the J-divergence between the functions $f(x;\theta)$ and $f(x;\theta + d\theta)$ coincides locally (for small displacements $d\theta$) with the square of the geodesic distance between the points (θ^i) and $(\theta^i + d\theta^i)$. So, locally, the J-divergence and the geodesic distance are closely connected by the relation

$$J(p,q) = dS^2. \tag{1.11}$$

The divergence is a (nonsymmetrical) generalization of the square of the Riemannian distance based on the Fisher information matrix.

2. Geometry of the random walk statistical model

Suppose that a particle, moving along a line tends to move with a uniform velocity ν. Suppose also that the particle is subject to linear Brownian motion, which causes it to take a variable amount of time to cover a fixed distance d. It can be shown that the time Y required to cover the distance is a random variable with probability density function

$$f_Y(y) = \frac{1}{\sqrt{2\pi\beta y^3}}\, d\, e^{-(d-\nu y)^2}/2\beta y, \quad y \succ 0, \tag{2.1}$$

where β is the diffusion constant. On substituting $\nu = d/\mu$ and $\beta = d^2/\lambda$ into (2.1), we obtain

$$f_Y(y;\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi y^3}}\, e^{-\lambda(y-\mu)^2}/2\mu^2 y, \quad y,\lambda,\mu \succ 0. \tag{2.2}$$

These distributions are called first passage time distributions of a Brownian motion with positive drift [13]. Because of the inverse relationship between the cumulant generating function of the first passage time distribution and that of the random distribution, Tweedie [12] proposed the name Inverse Gaussian (IG) for the first passage time distribution. For some purposes, it is convenient to work with the reciprocal of the Inverse Gaussian variate Y, which will be denoted by X. In terms of the moving particle, this corresponds to average speed. The probability density function of X is

$$f_X(x;\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi x}} \exp\left\{-\frac{\lambda x}{2} + \frac{\lambda}{\mu} - \frac{\lambda}{2\mu^2 x}\right\}, \quad x,\lambda,\mu \succ 0. \tag{2.3}$$

This distribution is called the random walk distribution [3], [13], where the mean $E(X)$ and the variance $V(X)$ are given as

$$E(X) = \frac{1}{\mu} + \frac{1}{\lambda} \quad \text{and} \quad V(X) = \frac{2}{\lambda^2} + \frac{1}{\mu\lambda}. \tag{2.4}$$

2.1. The metric tensors and the Fisher information matrix. The likelihood function of a random sample consisting of n observations $x_1,\dots,x_n$ from a random walk distribution is

$$L(x_i,\mu,\lambda) = \prod_{i=1}^{n} \sqrt{\frac{\lambda}{2\pi}} (x_i)^{\frac{-1}{2}} \exp\left\{-\frac{\lambda x_i}{2} + \frac{\lambda}{\mu} - \frac{\lambda}{2\mu^2 x_i}\right\}.$$

Thus, we have

$$\ln L(\mu,\lambda) = \frac{n}{2}\ln\frac{\lambda}{2\pi} - \frac{1}{2}\sum_{i=1}^{n}\ln x_i - \frac{\lambda}{2}\sum_{i=1}^{n} x_i + \frac{n\lambda}{\mu} - \frac{\lambda}{2\mu^2}\sum_{i=1}^{n}\frac{1}{x_i}. \tag{2.5}$$

The asymptotic variances and covariances of the maximum likelihood estimates $\hat{\lambda}$ and $\hat{\mu}$ can be obtained by inverting the Fisher information matrix (1.2), in which the elements are the negatives of the expected values of the second partials of the log likelihood function.

Using (2.5), one can see that the Fisher information matrix (1.2) (the components of the covariant metric tensor) is given by

$$(g_{ij}) = \operatorname{diag}\left(\frac{n\lambda}{\mu^3}, \frac{n}{2\lambda^2}\right), \tag{2.6}$$

and the variance covariance matrix (g^{ij}) (the components of the contravariant metric tensor) is the inverse of the Fisher information matrix and is given by

$$(g^{ij}) = \operatorname{diag}\left(\frac{2\lambda^2}{n}, \frac{\mu^3}{n\lambda}\right). \tag{2.7}$$

2.2. The Riemannian and scalar curvatures of the random walk manifold. Consider the statistical model (2-dimensional Riemannian manifold) corresponding to a family of random walk probability distributions.

From (1.4), (2.6) and (2.7) one can see that the Christoffel symbols are given as follows:

$$\left.\begin{array}{lll} \Gamma^1_{11} = 0, & \Gamma^2_{11} = \dfrac{\mu^3}{2\lambda^4}, & \Gamma^1_{12} = \Gamma^1_{21} = \dfrac{-1}{\lambda}, \\ \Gamma^1_{22} = \dfrac{3\lambda^3}{\mu^4}, & \Gamma^2_{22} = \dfrac{1}{2\lambda}, & \Gamma^2_{12} = \Gamma^2_{21} = \dfrac{-3}{2\mu}. \end{array}\right\} \tag{2.8}$$

Thus, using (1.5), it is easy to see that the Riemannian curvature tensor is given by

$$R^2_{121} = R^2_{122} = -R^2_{112} = -\frac{5(\mu^5 + 3\lambda^5)}{4\mu^2\lambda^5},$$

$$R^1_{212} = -R^1_{221} = -\frac{5(\mu^5 + 3\lambda^5)}{2\mu^5\lambda^2}$$

and the other components are equal to zero, and from (1.6) we have

$$R_{1212} = R_{1222} = -R_{1122} = -R_{2211} = -\frac{5n(\mu^5 + 3\lambda^5)}{4\mu^5\lambda^4}$$

and the other components are equal to zero. Thus, using (1.7) we find that the Ricci curvature tensor is given by

$$R_{11} = \frac{5(\mu^5 + 3\lambda^5)}{4\mu^2\lambda^5}, \quad R_{22} = \frac{5(\mu^5 + 3\lambda^5)}{2\mu^5\lambda^2}, \quad R_{12} = R_{21} = 0.$$

Thus, using (1.8) and (2.3), we have the following theorem [6]:

Theorem 1. *The scalar curvature R and the Gaussian curvature G of the 2-dimensional parameter space, constituted by the random walk distributions, are*

$$R = \frac{5}{n}\left[\left(\frac{\mu}{\lambda}\right)^3 + 3\left(\frac{\lambda}{\mu}\right)^2\right], \quad G = \frac{1}{2}R. \tag{2.9}$$

2.3. Geodesics on the random walk manifold. We next introduce the notion of straightness or geodesic in random walk distributions. From the Riemannian point of view, a curve is a geodesic if it is a minimum-length line connecting two points. In general, in order to define a straight line (geodesic), we need the notion of the covariant derivative by which the curvature of a curve is defined.

We put $\theta = (\theta^i) = (\theta^1, \theta^2) = (\mu, \lambda)$, $j, k = 1, 2$ in (1.3), and using (2.8), we have the geodesic differential equations

$$\left.\begin{aligned} \ddot{\mu} - \tfrac{2}{\lambda}\dot{\mu}\dot{\lambda} + \frac{3\lambda^3}{\mu^4}\dot{\lambda}^2 = 0, \qquad & \cdot = \frac{d}{ds} \\ \ddot{\lambda} - \tfrac{3}{\mu}\dot{\mu}\dot{\lambda} + \frac{\mu^3}{2\lambda^4}\dot{\mu}^2 + \frac{1}{2\lambda}\,\dot{\lambda}^2 = 0. \qquad & \end{aligned}\right\} \tag{2.10}$$

These equations can be put in the following equation

$$2\mu^5\lambda^4\ddot{\lambda} - 3\mu^4\lambda^5\ddot{\mu} + \mu^8\dot{\mu}^2 + \lambda^3(\mu^5 - 9\lambda^5)\dot{\lambda}^2 = 0. \tag{2.11}$$

In the following, the theoretical and numerical solutions of the geodesic differential equations (2.11) are obtained.

The theoretical solutions:

(i) For $\mu = \text{Const.}$ in (2.11), we have

$$\Gamma\left[\frac{3}{10}, \lambda^5\right] = a_1 s + b_1 \tag{2.12}$$

where $\Gamma[a, z]$ is the Incomplete Gamma function given by [2]

$$\Gamma[a, z] \simeq z^{a-1}\, e^{-z}\left[1 + \frac{a-1}{z} + \frac{(a-1)(a-2)}{z^2} + \cdots\right]. \tag{2.13}$$

Using approximation of the first order, we have

$$\lambda^{-\frac{7}{2}} e^{-\lambda^5} = a_1 s + b_1, \quad a_1, b_1 \in R \tag{2.14}$$

(see Figure 1).

(ii) At $\lambda = \text{Const.}$ in (2.11), we have

$$\Gamma\left[\frac{1}{5}, \mu^5\right] = a_2 s + b_2. \tag{2.15}$$

As in (2.14), we have

$$\mu^{-4} e^{-\mu^5} = a_2 s + b_2, \quad a_2, b_2 \in R \tag{2.16}$$

(see Figure 2).

From (2.14) and (2.16) we get an implicit relation between μ and λ as

$$a_1 \mu^{-4} e^{-\mu^5} - a_2 \lambda^{-\frac{7}{2}} e^{-\lambda^5} = a_1 b_2 - a_2 b_1 \tag{2.17}$$

(see Figure 3).

The numerical solutions of the geodesic differential equations (2.11) corresponding to the theoretical solutions (2.11), (2.16) and (2.17) are translated to figures as (4), (5) and (6), respectively.

3. The relation between the J-divergence and the geodesic distance

Although, globally no relation between the J-divergence and the geodesic distance has been found yet, locally we can obtain a relation given by formula (1.11). In geometrical language the geodesic distance is invariant but the J-divergence is not invariant. That is why the relation between them is nonlinear and generally much more complicated. We have investigated this relation in the case of random walk distributions.

3.1. The J-divergence in a random walk statistical model.

Theorem 2. *The J-divergence of the random walk probability density functions $p(x;\theta_p)$ and $q(x;\theta_q)$ satisfies the following equation:*

$$\begin{aligned} J(p,q) &= \frac{(\lambda_q - \lambda_p)^2}{2\lambda_q \lambda_p} + \frac{(\lambda_q - \lambda_p)(\mu_q - \mu_p)}{2\mu_q \mu_p} \\ &\quad + (\mu_p - \mu_q)\left(\frac{\lambda_q}{2\mu_q^2} - \frac{\lambda_p}{2\mu_p^2}\right). \end{aligned} \tag{3.1}$$

PROOF. From (2.3), the functions $p(x;\theta_p)$ and $q(x;\theta_q)$ are defined by the following equations:

$$\left.\begin{aligned} p(x;\theta_p) &= \sqrt{\frac{\lambda_p}{2\pi x}} \exp\left\{-\frac{\lambda_p x}{2} + \frac{\lambda_p}{\mu_p} - \frac{\lambda_p}{2\mu_p^2 x}\right\}, \\ q(x;\theta_q) &= \sqrt{\frac{\lambda_q}{2\pi x}} \exp\left\{-\frac{\lambda_q x}{2} + \frac{\lambda_q}{\mu_q} - \frac{\lambda q}{2\mu_q^2 x}\right\}, \end{aligned}\right\} \tag{3.2}$$

where $\theta_p = (\lambda_p, \mu_p)$, $\theta_q = (\lambda_q, \mu_q)$.

From (1.10), the J-divergence can be represented by the sum

$$J(p,q) = I(p,q) + I(q,p), \tag{3.3}$$

where

$$I(p,q) = \int p(x;\theta_p) \ln \frac{p(x;\theta_p)}{q(x;\theta_q)} dx. \tag{3.4}$$

Using the well-known properties of a probability density function

$$\int p(x;\theta)dx = 1, \quad \int q(x;\theta)dx = 1, \quad \int xp(x;\theta)dx = E(x), \tag{3.5}$$

and (3.2), (3.4) and (3.5), we can express the I-divergence of $p(x;\theta_p)$ and $q(x;\theta_q)$ in the following way:

$$\begin{aligned} I(p,q) &= \int p(x,\theta_p) \left\{\ln \frac{\lambda_p}{\lambda_q} + \frac{(\lambda_q - \lambda_p)}{2} x + \left(\frac{\lambda_p}{\mu_p} - \frac{\lambda_q}{\mu_q}\right) \right. \\ &\quad \left. + \left(\frac{\lambda_q}{2\mu_q^2} - \frac{\lambda_p}{2\mu_p^2}\right) \frac{1}{x}\right\} dx \\ &= \ln \frac{\lambda_p}{\lambda_q} + \frac{(\lambda_q - \lambda_p)}{2} A_p + \left(\frac{\lambda_p}{\mu_p} - \frac{\lambda_q}{\mu_q}\right) + \left(\frac{\lambda_q}{2\mu_q^2} - \frac{\lambda_p}{2\mu_p^2}\right) \mu_p, \end{aligned} \tag{3.6}$$

where $\mu_p = E\left(\frac{1}{X}\right) = E(Y)$ and Y is the inverse Gaussian random variable.

Similary, we have

$$I(q,p) = \ln \frac{\lambda_q}{\lambda_p} + \frac{(\lambda_p - \lambda_q)}{2} A_q + \left(\frac{\lambda_q}{\mu_q} - \frac{\lambda_p}{\mu_p}\right) + \left(\frac{\lambda_p}{2\mu_p^2} - \frac{\lambda_q}{2\mu_q^2}\right) \mu_q, \tag{3.7}$$

where A_p and A_q are the mean values of the random walk distributions of $p(x;\theta_p)$ and $q(x;\theta_q)$ respectively, which are given by (2.4).

Using (3.3), (3.6) and (3.7) we get

$$J(p,q) = \frac{1}{2}(\lambda_q - \lambda_p)(A_p - A_q) + (\mu_p - \mu_q)\left(\frac{\lambda_q}{2\mu_q^2} - \frac{\lambda_p}{2\mu_p^2}\right). \tag{3.8}$$

Thus, using (2.4), we have the proof. □

Special cases:

(i) If $\mu_p = \mu_q$, then the equation (3.1) reduces to

$$J_\mu(p,q) = \frac{(\lambda_q - \lambda_p)^2}{2\lambda_q\lambda_p}; \tag{3.9}$$

(ii) If $\lambda_q = \lambda_p = \lambda$, then the equation (3.1) takes the form

$$J_\lambda(p,q) = \frac{\lambda}{2}(\mu_p + \mu_q)\left(\frac{\mu_p - \mu_q}{\mu_p\mu_q}\right)^2. \tag{3.10}$$

3.2. The geodesic distance in a random walk manifold. Here we describe the geodesic distance in one and two-dimensional parameter spaces. Let $p(x;\theta_p)$ and $q(x;\theta_q)$ be two points in a random walk manifold correspondence to a statistical model of random walk probability density functions.

3.2.1. *One-dimensional parameter space* (space curve). From (1.9), the geodesic distance between two points $(\theta_p) = (\lambda_p, \mu_p)$, $(\theta_q) = (\lambda_q, \mu_q)$ on a space curve immersed in the random walk manifold is given as

$$S(p,q) = \Big|\int_{\theta_p}^{\theta_q} \sqrt{g_{11}(\theta^1)}d\theta^1\Big|. \tag{3.11}$$

From (2.3), we consider the following two cases:

(i) The λ is fixed, $\mu = \theta^1$ is variable (μ-parametric curve). In this case, we have

$$f(x;\theta^1) = \sqrt{\frac{\lambda}{2\pi x}}\exp\left\{-\frac{\lambda x}{2} + \frac{\lambda}{\theta^1} - \frac{\lambda}{2(\theta^1)^2 x}\right\}.$$

Thus, we get

$$g_{11}(\theta^1) = -E\left[\frac{\partial^2 \ln f}{\partial(\theta^1)^2}\right] = \frac{\lambda}{(\theta^1)^3} \tag{3.12}$$

and, because of (3.11) and (3.12), the geodesic distance is

$$S_\lambda = 2\left[\sqrt{\frac{\lambda}{\mu_p}} - \sqrt{\frac{\lambda}{\mu_q}}\right]. \tag{3.13}$$

(ii) The μ is fixed, $\lambda = \theta^1$ is variable (λ-parametric curve). Then the probability density function is

$$f(x;\theta^1) = \sqrt{\frac{\theta^1}{2\pi x}}\exp\left\{-\frac{\theta^1 x}{2} + \frac{\theta^1}{\mu} - \frac{\theta^1}{2\mu^2 x}\right\},$$

which means that

$$g_{11}(\theta^1) = \frac{1}{2(\theta^1)^2}. \tag{3.14}$$

Using (3.11) and (3.14), we obtain the geodesic distance as

$$S_\mu = \frac{1}{\sqrt{2}}\ln\frac{\lambda_q}{\lambda_p}. \tag{3.15}$$

The J-divergence and the geodesic distance in a random walk manifold are represented by (3.1) and (3.11), respectively, but to obtain a direct relation between these two general forms is hardly possible. That is why we consider the following

Theorem 3. *Let $p(x;\theta_p)$ and $q(x;\theta_q)$ be random walk probability density functions with parameters (λ_p, μ_p) and (λ_q, μ_q) respectively. Then the J-divergence and the geodesic distance are connected in the following ways:*

(i) *If μ is fixed:*

$$J_\mu = Cosh\,(\sqrt{2}S_\mu) - 1. \tag{3.16}$$

(ii) *If λ is fixed:*

$$J_\lambda = \frac{(\mu_q + \mu_p)(\sqrt{\mu_q} + \sqrt{\mu_p}\,)^2}{8\mu_q\mu_p}S_\lambda^2. \tag{3.17}$$

Proof. (i) In case of μ fixed, the J-divergence is given by the equation (3.9) and the geodesic distance satisfies the equation (3.15). From this we obtain

$$\frac{\lambda_q}{\lambda_p} = e^{\sqrt{2}S_\mu}. \tag{3.18}$$

Substituting (3.18) in (3.9), we obtain (3.16).

(ii) When λ is fixed, the J-divergence and the geodesic distance satisfy (3.10) and (3.13), respectively. Now, it is easy to see the connection (3.17) and the proof is finished. □

3.2.2. *Two-dimensional parameter space.* The geodesic distance between two points θ_p, θ_q in a random walk manifold is given, using (1.9), by the following two formulas:

$$S(p,q) = \int_{\theta_p}^{\theta_q} \sqrt{g_{11} + g_{22}\left(\frac{d\lambda}{d\mu}\right)^2}\, d\mu, \tag{3.19}$$

and

$$S(p,q) = \int_{\theta_p}^{\theta_q} \sqrt{g_{11}\left(\frac{d\mu}{d\lambda}\right)^2 + g_{22}}\, d\lambda. \tag{3.20}$$

From (2.12), (2.15) and using

$$\frac{\partial\Gamma[a,x]}{\partial x} = x^{a-1}\ e^{-x},$$

we have

$$\frac{d\lambda}{d\mu} = \frac{a_1}{a_2\sqrt{\lambda}} e^{\lambda^5-\mu^5}.$$

Therefore, and from (2.6), (3.19) and (3.20), we have the relation between the geodesic distances $S = S(\lambda)$, $\mu =$ const. and $S = S(\mu)$, $\lambda =$ const. which are given by the figures (7) and (8), respectively.

From the Fisher information matrix (2.6), it follows that the parameter space of the random walk distribution is represented by a two-dimensional manifold given by a system of partial differential equations

$$g_{11} = \langle \xi_\mu, \xi_\mu \rangle = \frac{\lambda}{\mu^3}, \quad g_{12} = \langle \xi_\mu, \xi_\lambda \rangle = 0, \quad g_{22} = \langle \xi_\lambda, \xi_\lambda \rangle = \frac{1}{2\lambda^2}, \tag{3.21}$$

where $\xi = \xi(\theta) = \xi(\mu, \lambda)$ is a 3-dimensional vector-valued function of two parameters μ, λ, $\langle , \rangle$ is the induced inner product (1.1), $\xi_\mu = \frac{\partial \xi}{\partial \mu}$ and $\xi_\lambda = \frac{\partial \xi}{\partial \lambda}$ are tangent vectors to a 2-dimensional (surface) random walk manifold immersed in a 3-dimensional space. The solution of the system (3.21) yields two figures 9 and 10.

Geometrical and statistical interpretations of the parameters:

From (3.21), $g_{12} = 0$ means that the random walk manifold is patched by an orthogonal net, i.e. at each point $\theta = (\lambda, \mu)$, the λ-parametric curve cuts the μ-parametric curve perpendicularly as shown in Figures 9 and 10.

From (2.9), it is easy to see that $R = \text{Const.}$ (zero) if $\mu = c\lambda$ ($\mu = -\lambda$), where c is a real number. This means that along the curve $\mu = c\lambda$, the parameter space has constant Gaussian curvature. If $c = -1$, the parameter space is flat ($R = 0$), i.e. plane.

From Figure 7, it follows that for large (small) values of λ the geodesic distance or divergence approaches infinity (a finite value). From Figure 8, it is easy to see that if the parameter μ tends to infinity (zero) the geodesic distance or divergence tends to zero (a finite value).

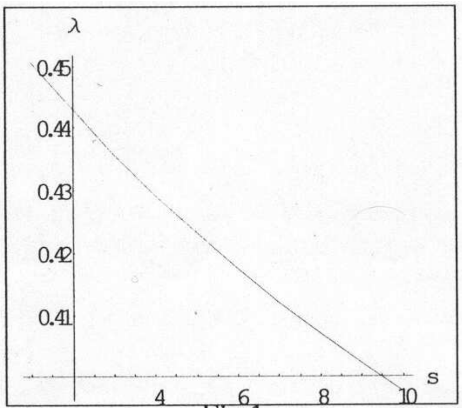

Fig.1

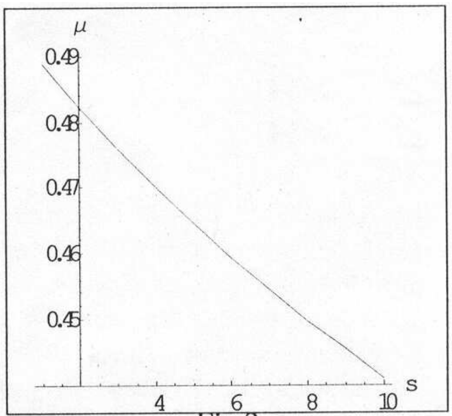

Fig.2

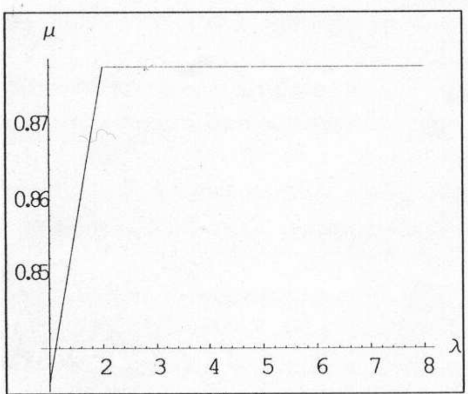

Fig.3

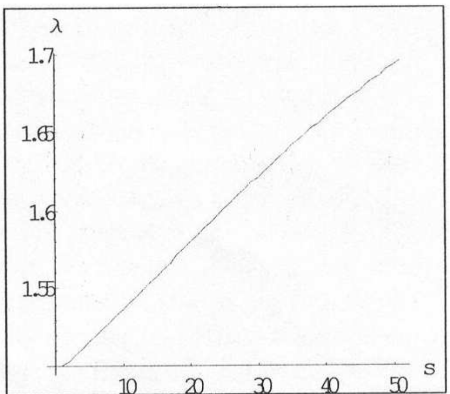

Fig.4

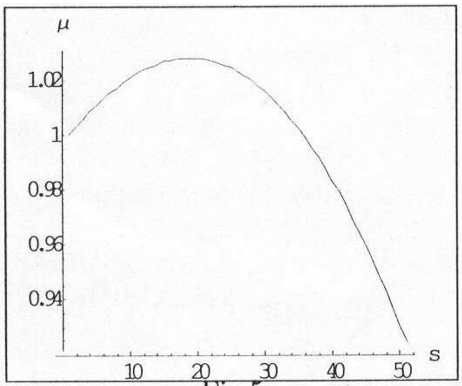

Fig.5

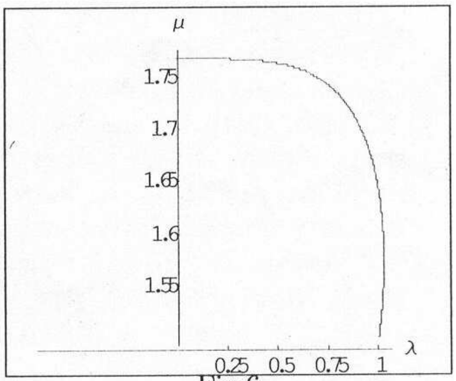

Fig.6

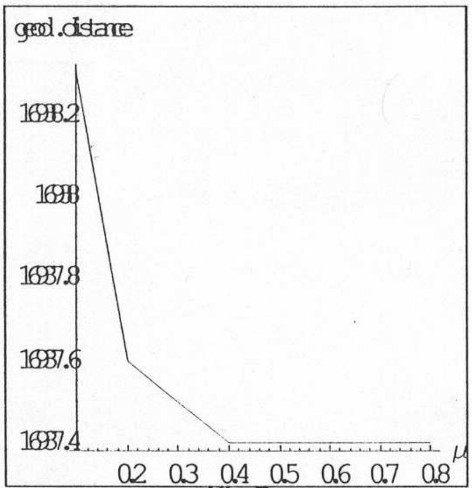

Fig.7

Fig.8

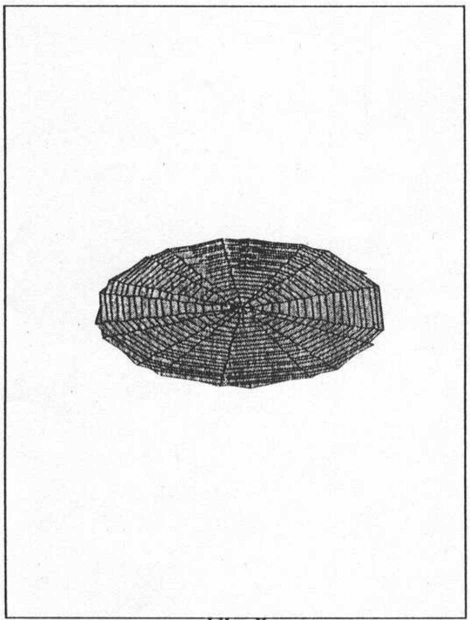

Fig.9

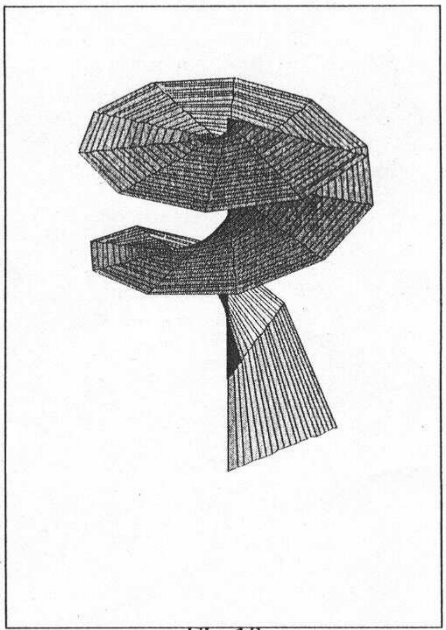

Fig.10

References

[1] N. H. Abdel-All, Information geometry and its applications, presented in Inter. Conf. on differential geometry and its applications, Cairo Univ., Fac. of Sci., Maths Dept., 19–26 June, Egypt, 2001.

[2] M. Abramowitz and I. A. Stegun (eds.), Handbook of Mathematical functions, *New York, Dover*, 1970.

[3] K. E. Ahmad, Random walk density function with unknown origin, *Mathl. Comput. Modelling* **18**, no. 8 (1993), 83–92.

[4] S. Amari, Differential geometrical methods in statistics, *Springer lecture Notes in Statistics* **28** (1985).

[5] S. Amari, Differential geometry of a parametric family of invertible linear systems Riemannian metric, dual affine connections and divergence, *Mathematical systems theory* **20** (1987b), 53–82.

[6] B. Y. Chen, Geometry of submanifolds, *Marcel Dekker*, *New York*, 1973.

[7] E. Hayama, R. Ivanova and T. Kawaguchi, A Riemannian approach to the J-divergence, *Tensor, N. S.* **57** (1996), 282–293.

[8] R. Ivanova and T. Kawaguchi, An interpretation of the Riemannian curvature in a statistical parameter space, *Tensor, N. S.* **57** (1996), 300–305.

[9] S. Kobayashi and K. Nomizu, Foundations of differential geometry, Vol. I, *Inter-sciencePublishers, a division of John wiley & Sons*, *New York, London*, 1963.

[10] M. K. Murray and J. W. Rice, Differential geometry and statistics, *Chapman and Hall*, 1993.

[11] C. R. Rao, Information and accuracy attainable in the estimation of statistical parameters, *Bull. Calcutta. Math. Soc.* **37** (1945), 81–91.

[12] M. C. K. Tweedie, Some statistical properties of the inverse Gaussian distribution, *Virginia Journal of Science* **7** (1956), 160–165.

[13] M. T. Wasan, First passage time distribution of Brownian motion, Monograph, *Dept. of Math., Queen's Univ.*, *Kingston, Ontario, Canada*, 1969.

NASSAR H. ABDEL-ALL, H. N. ABD-ELLAH, H. M. MOUSTAFA
DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE
ASSIUT UNIVERSITY
ASSIUT, 71516
EGYPT

E-mail: nhabdeal2002@yahoo.com
hamdy_n2000@yahoo.com

(Received July 9, 2001; revised January 11, 2002)