جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة بابل/ كلية التربية

تطبيق مبرهنة الغاية المركزية على توزيعات احتمالية باستخدام الهندسة التفاضلية

رسالة مقدمة الى
قسم الرياضيات– كلية التربية- جامعة بابل
وهي جزء من متطلبات نيل درجة الماجستير في علوم الرياضيات

من قبل

وسن عباس جاسم الحسناوي

بإشراف

أ. م. حسين احمد علي

2006 م **1427هـ**

In this chapter, we introduce some basic ideas and important concepts of differential geometry such as tensor, first fundamental form, second fundamental form, Gaussian curvature, geodesics, etc.

We shall denote the familiar three dimensional Euclidean space (traditionally denoted R^3) as E^3. In studying the geometry of a surface in E^3 we find that some of its most important geometric properties belong to the surface itself and not the surrounding Euclidean space.

The property of a surface which depends only on the metric form is an intrinsic. For example, Gaussian curvature is an intrinsic property of a surface.

1.1 Tensor [6]

An nth- rank tensor in m- dimensional space is a mathematical object that has n indices and m^n components and obeys certain transformation rules. Each index of a tensor ranges over the number of dimensions of space. However, the dimension of the space is largely irrelevant in most tensor equations (with the notable exception of the contracted kronecker delta). Tensors are generalizations of scalars (that have no indices), vectors (that have exactly one index), and matrices (that have exactly two indices) to an arbitrary number of indices.

Tensors provide a natural and concise mathematical framework for formulating and solving problems in areas of physics such as elasticity, fluid mechanics, and general relativity.

The notation for a tensor is similar to that of a matrix (i.e., $A=(a_{ij})$), except that a tensor $a_{ijk\ldots}, a^{ijk\cdots}, a_i^{jk}$, etc., may have an arbitrary number of indices where i, j, k...=1, 2, ..., m. In addition, a tensor with rank r+s may be of mixed type (r, s), consisting of r so-called “contravariant” (upper) indices and s “covariant” (lower) indices. Note that the positions of the slots in which contravariant and covariant indices are placed are significant so, for example, $a_{«¬}^{\text{a}}$ is distinct from $a_{«}^{¬\text{a}}$.

While the distinction between covariant and contravariant indices must be made for general tensor, the two are equivalent for tensors in three-dimensional Euclidean space, and such tensors are known as cartesian tensors.

Objects that transform like zeroth-rank tensors are called scalars, those that transform like first- rank tensors are called vectors, and those that transform like second-rank tensors are called matrices. In tensor notation, a vector v would be written v_i, where i= 1, 2, …, *m*, and matrix is a tensor of type (1, 1), which would be written a_i^j in tensor notation.

Tensors may be operated on by other tensors (such as metric tensors, the permutation tensor, or the kronecker delta) or by tensor operators (such as the covariant or semicolon derivatives). The manipulation of tensor indices to produce identities or to simplify expressions is known as index gymnastics, which includes index lowering and index raising as special cases. These can be achieved through multiplication by a so-called metric tensor g_{ij}, g^{ij}, g_i^j, etc., e.g.,

$$g^{ij} A_j = A^i \qquad i, j = 1, 2, \ldots, m \qquad \text{…. (1.1)}$$

$$g_{ij} A^j = A_i \qquad \text{…. (1.2)}$$

The metric tensor is a tensor of rank 2 that is used to measure distance between any two points in a given space.

Tensor notation can provide a very concise way of writing vector and more general identities. For example, in tensor notation, the dot product *u.v* is simply written

$$u.v = u_i v^i, \qquad \text{…. (1.3)}$$

where repeated indices are summed over (Einstein summation).

Similarly, the cross product can be concisely written as

$$(u \times v)_i = \in_{ijk} u^j v^k , \qquad \text{…. (1.4)}$$

where $\in_{ijk}$ is the permutation tensor

Contravariant second-rank tensors are objects which transform as

$$A'^{ij} = \frac{\partial x'_i}{\partial x_k} \frac{\partial x'_j}{\partial x_l} A^{kl} \qquad i, j, k, l=1, 2, \dots, m \qquad \text{…. (1.5)}$$

Covariant Second- rank tensors are objects which transform as

$$C'_{ij} = \frac{\partial x_k}{\partial x'_i} \frac{\partial x_l}{\partial x'_j} C_{kl} \quad i, j, k, l = 1,2,...,m \qquad \text{…. (1.6)}$$

Mixed Second- rank tensors are objects which transform as

$$B'^i_j = \frac{\partial x'_i}{\partial x_k} \frac{\partial x_l}{\partial x'_j} B^k_l \quad i, j, k, l = 1,2,...,m \qquad \text{…. (1.7)}$$

where $x'_i = x'_i(x_1, x_2, ..., x_m)$ is the coordinate transformation, and $x_i = x_i(x'_1, x'_2, ..., x'_m)$ is its inverse.

Definition (1.1.1) [7]

A *second-tensor rank symmetric tensor* is defined as a tensor *A* for which

$$A^{mn} = A^{nm}$$

Definition (1.1.2) [8]

An *antisymmetric* (also called *alternating*) tensor is a tensor which changes sign when two indices are switched. For example, a tensor $A^{x_1,...,x_n}$ such that

$$A^{x_1 ,...,x_i ,...,x_j,...,x_n} = -A^{x_n,...,x_i,...,x_j,...,x_1}$$

where $x_1,...,x_n$ are indices.

The simplest nontrivial antisymmetric tensor is therefore an antisymmetric rank-2 tensor, which satisfies $A^{mn} = -A^{nm}$.

1.2 The First Fundamental Form [9, 10]

Suppose M is a surface determined by $\vec{X}(u,v) \subset E^3$ and suppose $\vec{\sim}(t)$ is a curve on M, the variable t is called the parameter of the curve, $t \in [a,b]$ for $a,b \in R$. Then we can write $\vec{\sim}(t) = \vec{X}(u(t),v(t))$ (($u(t)$, $v(t)$)) is a curve in R^2 whose image under $\vec{X}$ is $\vec{\sim}$). Then

$$\vec{\sim}'(t) = \frac{\partial \vec{X}}{\partial u}\frac{du}{dt} + \frac{\partial \vec{X}}{\partial v}\frac{dv}{dt} = u'\vec{X}_1 + v'\vec{X}_2 \qquad \text{…. (1.8)}$$

If $s(t)$ represents the arc length along $\vec{\sim}$ (with $s(a) = 0$) then

$$s(t) = \int_a^t \left\|\vec{\sim}'(r)\right\| dr \qquad \text{…. (1.9)}$$

where r is a real variable.

And

$$\frac{ds}{dt} = \left\|\vec{\sim}'(t)\right\| \qquad \text{…. (1.10)}$$

so

$$\left(\frac{ds}{dt}\right)^2 = \left\|\vec{\sim}'(t)\right\|^2 = \vec{\sim}'.\vec{\sim}' = \left(u'\vec{X}_1 + v'\vec{X}_2\right).\left(u'\vec{X}_1 + v'\vec{X}_2\right)$$

$$= u'^2\left(\vec{X}_1.\vec{X}_1\right) + 2u'v'\left(\vec{X}_1.\vec{X}_2\right) + v'^2\left(\vec{X}_2.\vec{X}_2\right).$$

Following Gauss' notation (briefly) we denote

$$E = \vec{X}_1.\vec{X}_1, \qquad F = \vec{X}_1.\vec{X}_2, \qquad G = \vec{X}_2.\vec{X}_2 \qquad \text{…. (1.11)}$$

and have

$$\left(\frac{ds}{dt}\right)^2 = E\left(\frac{du}{dt}\right)^2 + 2F\left(\frac{du}{dt}\frac{dv}{dt}\right) + G\left(\frac{dv}{dt}\right)^2 \qquad \text{.... (1.12)}$$

or in differential notation

$$ds^2 = E(du)^2 + 2F(dudv) + G(dv)^2 \qquad \text{…. (1.13)}$$

Definition (1.2.1) [9]

Let M be a surface determined by $\vec{X}(u,v)$. The *first fundamental form* (or more commonly *metric form*) of M is $\left(\frac{ds}{dt}\right)^2$ or $(ds)^2$ as defined in formulas (1.12) and (1.13).

Definition (1.2.2) [10]

The *matrix of the first fundamental form* of a surface M determined by $\vec{X}(u,v)$ is

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix} \equiv \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$$

where E, F, G are as defined in formula (1.11). This matrix determines dot products of tangent vectors.

If $\vec{v} = a\vec{X}_1 + b\vec{X}_2$ and $\vec{\mu} = c\vec{X}_1 + d\vec{X}_2$ are vectors tangent to a surface M at a given point, then

$$\begin{aligned} \vec{v}.\vec{\mu} &= \left(a\vec{X}_1 + b\vec{X}_2\right)\left(c\vec{X}_1 + d\vec{X}_2\right) \\ &= Eac + F(ad + bc) + Gbd \\ &= \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} \end{aligned}$$

Notation [9]

We now replace the parameters u and v with u^1 and u^2 in formula (1.13).

We then have

$$ds^2 = g_{11}\left(du^1\right)^2 + 2g_{12}du^1du^2 + g_{22}\left(du^2\right)^2 = \sum_{i,j} g_{ij}du^i du^j \quad \text{…. (1.14)}$$

where the summation is taken over the set {1, 2}. If $\vec{v}$ is a vector tangent to M at a point $\vec{p}$ and $\vec{v} = \left(v^1, v^2\right)$ in the basis $\left\{\vec{X}_1, \vec{X}_2\right\}$ for the tangent plane at $\vec{p}$, then we have $\vec{v} = \sum_i v^i \vec{X}_i$.

If $\vec{\alpha}(t)$ is a curve on M where $\vec{\alpha}$ is represented by $\vec{X}\left(u^1(t), u^2(t)\right)$ then

$$\vec{\alpha}'(t) = u^{1'}(t)\vec{X}_1 + u^{2'}(t)\vec{X}_2 = \sum_i u^{i'}\vec{X}_i \, .$$

1.3 The Second Fundamental Form [11]

We have treated a path $\vec{\alpha}(t)$ along a surface M as if it were the trajectory of a particle in E^3. We then interprete $\vec{\alpha}''(t)$ as the acceleration of the particle. Well, a particle can accelerate in two ways:
(1) it can accelerate in the direction of travel, and (2) it can accelerate by changing its direction of travel. We can therefore decompose $\vec{\alpha}''$ into two components, $\vec{\alpha}''_{\vec{T}}$ (representing acceleration in the direction of travel) and $\vec{\alpha}''_{\vec{N}}$ (representing acceleration that changes the direction of travel). We may have dealt with this by taking $\vec{\alpha}''_{\vec{T}}$ as the component of $\vec{\alpha}''$ in the direction of $\vec{\alpha}'$ computed as $\vec{\alpha}''_{\vec{T}} = \left(\vec{\alpha}''.\frac{\vec{\alpha}'}{\|\vec{\alpha}'\|}\right)\frac{\vec{\alpha}'}{\|\vec{\alpha}'\|}$

and $\vec{\alpha}''_{\vec{N}}$ as the "remaining component" of $\vec{\alpha}$ (that is, $\vec{\alpha}''_{\vec{N}} = \vec{\alpha}'' - \vec{\alpha}''_{\vec{T}}$).

The unit tangent vector $\vec{T}(s) = \vec{\alpha}'(s) = u^{i'}\vec{X}_i$ with $\vec{\alpha}$ parameterized in terms of arc length s, $\vec{\alpha} = \vec{\alpha}(s) = \vec{X}\left(u^1(s), u^2(s)\right)$. We can see that $\vec{\alpha}''(s) = \vec{T}'(s)$ is a vector normal to $\vec{\alpha}'$ where $\vec{T}' = k\vec{N}$. We again decompose $\vec{\alpha}''$ into two orthogonal components, but this time we make explicit use of the surface M. We wish to write:

$$\vec{\alpha}'' = \vec{\alpha}''_{\tan} + \vec{\alpha}''_{nor}$$

where $\vec{\alpha}''_{\tan}$ is the component of $\vec{\alpha}''$ tangent to M and $\vec{\alpha}''_{nor}$ is the component of $\vec{\alpha}''$ normal to M. Notice that $\vec{\alpha}''_{\tan}$ will be a linear combination of $\vec{X}_1$ and $\vec{X}_2$ (they are a basis for the tangent plane) and

$\vec{\sim}''_{nor}$ will be a multiple of the unit normal vector to *M*, $\vec{U}\left(\text{calculated as}\,\vec{U}=\dfrac{\vec{X}_1\times\vec{X}_2}{\left\|\vec{X}_1\times\vec{X}_2\right\|}\right)$.

Since $\vec{\sim}(s)=\vec{X}\left(u^1(s),u^2(s)\right)$ and $\vec{\sim}'=u^{i'}\vec{X}_i$ (here, 'means *d*/*ds*), then

$$\vec{\sim}''=u^{i''}\vec{X}_i+u^{i'}\vec{X}_i'=u^{i''}\vec{X}_i+u^{i'}\frac{d\vec{X}_i}{ds}$$

Now $u^{i''}\vec{X}_i$ is part of $\vec{\sim}''_{\tan}$, but $u^{i'}\vec{X}_i'$ may also have a component in the tangent plane. Well,

$$\frac{d\vec{X}_i}{ds}=\frac{d}{ds}\left[\vec{X}_i\left(u^1(s),u^2(s)\right)\right]=\frac{\partial\vec{X}_i}{\partial u^1}\frac{du^1}{ds}+\frac{\partial\vec{X}_i}{\partial u^2}\frac{du^2}{ds}$$

$$=\frac{\partial\vec{X}_i}{\partial u^1}u^{1'}+\frac{\partial\vec{X}_i}{\partial u^2}u^{2'}=\frac{\partial\vec{X}_i}{\partial u^j}u^{j'}.$$

If we denote $\dfrac{\partial^2\vec{X}}{\partial u^i\partial u^j}=\vec{X}_{ij}$ (we have assumed continuous second partials, so the order of differentiation doesn't matter) then we have $\dfrac{d\vec{X}_i}{ds}=\vec{X}_{ij}u^{j'}$.

So acceleration becomes

$$\vec{\sim}''=u^{r''}\vec{X}_r+u^{i'}u^{j'}\vec{X}_{ij}.$$

We now need only to write $\vec{X}_{ij}$ in terms of a component in the tangent plane (and so in terms of $\vec{X}_1$ and $\vec{X}_2$) and a component normal to the tangent plane (which will be a multiple of $\vec{U}$).

Definition (1.3.1) [11]

With the notation above, we define the *formula of Gauss* as

$$\vec{X}_{ij}=\Gamma_{ij}^{r}\vec{X}_r+L_{ij}\vec{U}. \qquad \text{…. (1.15)}$$

That is we define L_{ij} as the projection of $\vec{X}_{ij}$ in the direction $\vec{U}$. Notice, however, that Γ_{ij}^{r} may not be the projection of $\vec{X}_{ij}$ onto $\vec{X}_r$ since the $\vec{X}_r$'s are not orthonormal.

Note [11]

Since projections are computed from dot products, we immediately have that

$$L_{ij} = \vec{X}_{ij}.\vec{U} = \vec{X}_{ij}.\frac{\vec{X}_1 \times \vec{X}_2}{\left\| \vec{X}_1 \times \vec{X}_2 \right\|} \quad \text{.... (1.16)}$$

We therefore have

$$\vec{\alpha}'' = \vec{\alpha}''_{\tan} + \vec{\alpha}''_{nor} = (u^{r''} + \Gamma_{ij}^{r} u^{i'} u^{j'})\vec{X}_r + (L_{ij} u^{i'} u^{j'})\vec{U} \quad \text{.... (1.17)}$$

Definition (1.3.2) [12]

The *second fundamental form* of surface *M* is the matrix $\begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}$ where the determinate of this matrix is *L*. The projections L_{ij} are defined in formula (1.16).

1.4 Gaussian Curvature [13]

If $f(x, y, z)$ is a (scalar valued) function, then for c a constant, $f(x, y, z)=c$ determines a *surface* (we assume all second partials of *f* are continuous and so the surface is *smooth*). The *gradient* of *f* is $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$.

If $\vec{v}_\circ$ is a vector tangent to the surface $f(x,y,z)=c$ at point $\vec{p}_\circ = (x_\circ, y_\circ, z_\circ)$, then $\nabla f(x_\circ, y_\circ, z_\circ)$ is orthogonal to $\vec{v}_\circ$ (and so ∇f is orthogonal to the surface). The equation of a plane tangent to the surface can be calculated using ∇f as the normal vector for the plane.

Definition (1.4.1) [13, 14]

Let $\vec{v}$ be a unit vector tangent to a smooth surface $M \subset E^3$ at a point $\vec{p}$ (again making no distinction between a vector and a point). Let $\vec{U}$ be a unit vector normal (perpendicular) to M at point $\vec{p}$. The plane through point $\vec{p}$ which contains vectors $\vec{v}$ and $\vec{U}$ intersects the surface in a curve $\vec{\sim}_{\vec{v}}$ called the *normal section* of M at $\vec{p}$ *in the direction* $\vec{v}$.

Example (1.4.2) [11]

Let $M : x^2 + y^2 = 1$ (an infinitely tall right circular cylinder of radius 1). To find the normal section of M at the point $\vec{p}$=(1,0,0) in the direction $\vec{v}$=(0,1,0). We can use a normal vector to M at $\vec{p}$ as

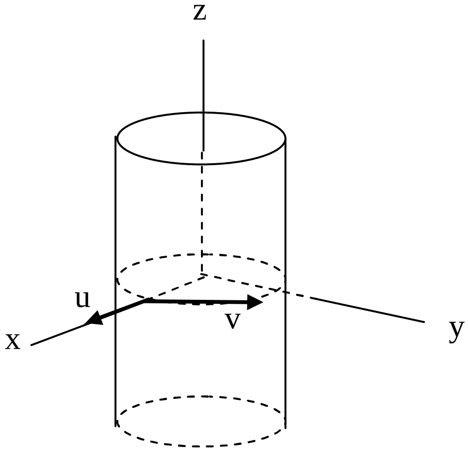
$$\nabla\left(x^2 + y^2\right) = \left(2x, 2y, 0\right)\Big|_{(1,0,0)} = (2,0,0)$$

Therefore, we take $\vec{U}$=(1,0,0). The plane containing $\vec{U}$ and $\vec{v}$ has as a normal vector

$$\vec{U} \times \vec{v} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = (0,0,1).$$

Therefore the equation of this plane is $0(x-1) + 0(y-1) + 1(z-0) = 0$ or z=0. The intersection of this plane and the surface is

$$\vec{\sim}_{\vec{v}} = \left\{(x, y, z) \middle| x^2 + y^2 = 1, \quad z = 0\right\}.$$



Definition (1.4.3) [11]

Let $\vec{\sim}_{\vec{v}}$ be a normal section to a smooth surface M at point $\vec{p}$ in the direction $\vec{v}$. Let $\vec{U}$ be a unit normal to M at $\vec{p}$ (-$\vec{U}$ is also a unit normal to M at $\vec{p}$). The *normal curvature* of M at $\vec{p}$ *in the* $\vec{v}$ *direction with respect to* $\vec{U}$, denoted $k_{n,\vec{U}}(\vec{v})$, is

$$k_{n,\vec{U}}(\vec{v}) = \frac{\vec{U}.\vec{N}}{R(\vec{v})} \qquad \text{…. (1.18)}$$

where $\vec{N}$ is the principal normal vector of $\vec{\sim}_{\vec{v}}$ at $\vec{p}$ and $R(\vec{v})$ is the radius of the osculating circle to $\vec{\sim}_{\vec{v}}$ at $\vec{p}$. If $\vec{\sim}_{\vec{v}}$ has zero curvature at $\vec{p}$, we take $k_{n,\vec{U}}(\vec{v})$ =0.

Note [12, 13]

We have defined the normal curvature of a surface at a point $\vec{p}$ in the direction $\vec{v} : k_n(\vec{v})$. Therefore, for a given point on a surface, there are an infinite number of (not necessarily distinct) curvatures (one for each "direction"). We can think of $k_n(\vec{v})$ as a function mapping the vector space $T_{\vec{p}}(M)$ (the plane tangent to surface M at point $\vec{p}$) into R. That is $k_n : T_{\vec{p}}(M) \to R$. We need $\vec{v}$ to be a unit vector, so the domain of k_n is $\{\vec{v} \in T_{\vec{p}}(M) \,|\, \|\vec{v}\| = 1\}$. Therefore, k_n is a continuous function on a compact set and by the extreme value theorem (for metric spaces), k_n assumes a maximum and a minimum value.

Definition (1.4.4) [11]

Let M be a surface and $\vec{p}$ a point on the surface. Define $k_1 = \max k_n(\vec{v})$ and $k_2 = \min k_n(\vec{v})$ where the maximum and minimum are taken over the domain of k_n. k_1 and k_2 are called the *principal curvatures* of *M* at $\vec{p}$, and the corresponding directions are called *principal directions*. The product $K = K(P) = k_1 k_2$ is the *Gaussian curvature* of *M* at $\vec{p}$.

Theorem (1.4.5) [11]

The Gaussian curvature at any point $\vec{p}$ of a surface *M* is $K(\vec{p}) = L/g$ where $L = \det(L_{ij})$ and $g = \det(g_{ij})$.

Proof

First, if $\vec{v} = v^i \vec{X}_i$ then

$$\|\vec{v}\|^2 = \left(v^1 \vec{X}_1 + v^2 \vec{X}_2\right).\left(v^1 \vec{X}_1 + v^2 \vec{X}_2\right)$$

$$= (v^1)^2 \vec{X}_1.\vec{X}_1 + 2(v^1)(v^2)\vec{X}_1.\vec{X}_2 + (v^2)^2 \vec{X}_2.\vec{X}_2$$

$$= g_{mn} v^m v^n \left(g_{mn} = \vec{X}_m.\vec{X}_n\right).$$

Therefore finding extrema of $k_n(\vec{v})$ for $\|\vec{v}\| = 1$equivalent to finding extrema of

$$k = k_n(\vec{v}) = \frac{L_{ij} v^i v^j}{g_{mn} v^m v^n} \qquad \text{…. (1.19)}$$

for $\vec{v} \in T_{\vec{p}}(M)$ and $\vec{v} \neq \vec{0}$. If $k_n(\vec{v})$ is an extreme value of *k*, where $\vec{v} = v^i \vec{X}_i$, then $\frac{\partial k}{\partial v^1} = \frac{\partial k}{\partial v^2} = 0$ at $\vec{v}$ (that is, the gradient of *k* is 0 however, this gradient

is computed in a (v^1, v^2) coordinate system, not (*x*, *y*)). Now

$$\frac{\partial k}{\partial v^r} = \frac{\left[2L_{rj} v^j\right]\left(g_{mn} v^m v^n\right) - \left(L_{ij} v^i v^j\right)\left[2 g_{rn} v^n\right]}{\left(g_{mn} v^m v^n\right)^2}$$

for *r*= 1, 2. Now $k = \frac{L_{ij} v^i v^j}{g_{mn} v^m v^n}$, so replacing $L_{ij} v^i v^j$ with $k g_{mn} v^m v^n$ gives

$$\frac{\partial k}{\partial v^r} = \frac{2L_{rj}v^j\left(g_{mn}v^m v^n\right) - \left(kg_{mn}v^m v^n\right)2g_{rn}v^n}{\left(g_{mn}v^m v^n\right)^2}$$

$$= \frac{2L_{rj}v^j - 2kg_{rn}v^n}{g_{mn}v^m v^n} = \frac{2L_{rj}v^j - 2kg_{rj}v^j}{g_{mn}v^m v^n}$$

$$= \frac{2\left(L_{rj} - kg_{rj}\right)v^j}{g_{mn}v^m v^n},$$

for r= 1, 2. So at an extreme value, $\left(L_{ij} - kg_{ij}\right)v^j$ =0 for i = 1, 2. This is two linear equations in two unknowns (v^1 and v^2). Since $\vec{v}$ is nonzero, the only way this system can have a solution is for $\det\left(L_{ij} - kg_{ij}\right)$=0. That is:

$$\det\begin{bmatrix} L_{11} - kg_{11} & L_{12} - kg_{12} \\ L_{21} - kg_{21} & L_{22} - kg_{22} \end{bmatrix} = 0$$

or $\left(L_{11} - kg_{11}\right)\left(L_{22} - kg_{22}\right) - \left(L_{21} - kg_{21}\right)\left(L_{12} - kg_{12}\right) = 0$

or $L_{11}L_{22} - kL_{11}g_{22} - kL_{22}g_{11} + k^2 g_{11}g_{22} - L_{21}L_{12}$

$+ kL_{21}g_{12} + kL_{12}g_{21} - k^2 g_{12}g_{21} = 0$

or $k^2\left(g_{11}g_{22} - g_{12}g_{21}\right) - k\left(g_{11}L_{22} + g_{22}L_{11} - g_{12}L_{12} - g_{21}L_{21}\right)$

$+ \left(L_{11}L_{22} - L_{21}L_{12}\right) = 0$

or $k^2 g - k\left(g_{11}L_{22} + g_{22}L_{11} - 2g_{12}L_{12}\right) + L = 0$

since $L_{12} = L_{21}, L = \det\left(L_{ij}\right),$ and $g = \det\left(g_{ij}\right)$.

So for extrema of k we need

$$k^2 - k\left(\frac{g_{11}L_{22} + g_{22}L_{11} - 2g_{12}L_{12}}{g}\right) + \frac{L}{g} = 0$$

Since k_1 and k_2 are known to be roots of this equation, this equation factors as

$\left(k - k_1\right)\left(k - k_2\right) = k^2 - \left(k_1 + k_2\right)k + k_1k_2 = 0$. Therefore, the Gaussian curvature is $k_1k_2 = L/g$.■

Theorem (1.4.6) [11]

If $\vec{v}$ and $\vec{\mu}$ are principal directions for surface M at point $\vec{p}$ corresponding to k_1 (maximum normal curvature at $\vec{p}$) and k_2 (minimum normal curvature at $\vec{p}$) respectively, then if $k_1 \neq k_2$ we have $\vec{v}$ and $\vec{\mu}$ are orthogonal.

Definition (1.4.7) [9]

For a surface determined by $\vec{X}\left(u^1, u^2\right)$, the equations $\vec{U}_j = -L_j^i \vec{X}_i$ for j=1, 2 are the *equations of Weingarten*, where $L_j^i = L_j^i\left(u^1, u^2\right) = L_{jk} g^{ki}$ for i, j= 1, 2 and $\vec{U}_j$ is tangent to M. Therefore $\vec{U}_j$ is a linear combination of $\vec{X}_1$ and $\vec{X}_2$.

1.5 Geodesics [15]

A curve $\vec{\sim}(s)$ on a surface M can curve in two different ways. First, $\vec{\sim}$ can bend *along with* surface M (the "normal curvature" discussed above). Second, $\vec{\sim}$ can bend *within* the surface M (the "geodesic curvature" to be defined).

For curve $\vec{\sim}$ on surface M, $\vec{\sim}''$ can be written as components tangent and normal to M as $\vec{\sim}'' = \vec{\sim}''_{\tan} + \vec{\sim}''_{nor}$ where

$$\vec{\sim}''_{\tan} = \left(u^{r''} + \Gamma_{ij}^r u^{i'} u^{j'}\right)\vec{X}_r$$

$$\vec{\sim}''_{nor} = \left(L_{ij} u^{i'} u^{j'}\right)\vec{U}$$

and the parameters on the right hand side are defined in section (1.3). $\vec{\sim}''_{nor}$ reflects the curvature of $\vec{\sim}$ due to the bending of M and $\vec{\sim}''_{\tan}$ reflects the curvature of $\vec{\sim}$ within M. Now

$$\vec{\sim}''_{\tan}.\vec{U} = \left\{\left(u^{r''} + \Gamma_{ij}^r u^{i'} u^{j'}\right)\vec{X}_r\right\}.\vec{U} = 0$$

$$\left(\text{we know that } \vec{U} = \frac{\vec{X}_1 \times \vec{X}_2}{\left\|\vec{X}_1 \times \vec{X}_2\right\|}\right) \text{ and}$$

$$\vec{\sim}''_{\tan}.\vec{\sim}' = \vec{\sim}''_{\tan}.\vec{\sim}' + 0 = \vec{\sim}''_{\tan}.\vec{\sim}' + \vec{\sim}''_{nor}.\vec{\sim}'$$

$$(\vec{\sim}' = u^{i'}\vec{X}_i \text{ and } \vec{X}_i.\vec{U} = 0)$$

$$= \left(\vec{\sim}''_{\tan} + \vec{\sim}''_{nor}\right).\vec{\sim}' = \vec{\sim}''.\vec{\sim}' = 0$$

$$\left(\left\|\vec{\sim}'\right\| = \left\|\vec{\sim}'(s)\right\| = 1 \text{ and } ' = d/ds\right).$$

Therefore $\vec{\sim}''_{\tan}$ is orthogonal to both $\vec{U}$ and $\vec{\sim}'$. If we define $\vec{\mu}$ as the unit vector $\vec{\mu} = \vec{U} \times \vec{\sim}'$, then $\vec{\sim}''_{\tan}$ is a multiple of $\vec{\mu}$ (and $\vec{\mu}$ is a vector tangent to M).

1.6 Some Properties of Geodesic and Christoffel Symbols [11]

There are some important properties which need it in this thesis

(1) Let $\vec{\sim}(s)$ be a curve on M where s is arc length. The *geodesic curvature* of $\vec{\sim}$ at $\vec{\sim}(s)$ is the function $k_g = k_g(s)$ defined by

$$\vec{\sim}''_{\tan} = k_g\vec{\mu} = k_g(\vec{U} \times \vec{\sim}'). \qquad \text{…. (1.20)}$$

(2) The geodesic curvature k_g of curve $\vec{\sim}$ in surface M can be calculated as

$$k_g = \vec{U}.\vec{\sim}' \times \vec{\sim}''. \qquad \text{…. (1.21)}$$

(3) Let $\vec{\sim} = \vec{\sim}(s)$ be a curve on a surface M. Then $\vec{\sim}$ is a *geodesic* if $\vec{\sim}''_{\tan} = \vec{0}$ (or equivalently, if $\vec{\sim}'' = \vec{\sim}''_{nor}$) at every point of $\vec{\sim}$.

(4) A geodesic on a surface is, in a sense, as "straight" as a curve can be on the surface. That is, $\vec{\sim}$ has no curvature within the surface. For example, on a sphere the geodesics are great circles.

(5) If $\vec{\sim}$ is a geodesic on M then

$$u^{r''} + \Gamma^r_{ij}u^{i'}u^{j'} = 0$$

for r= 1,2 and $\vec{U}.\vec{\sim}' \times \vec{\sim}'' = 0$.

(6) Let $\vec{X}(u^1,u^2)$ be a surface and let g_{ij} and Γ_{ij}^r be as defined in sections (1.2) and (1.3). The *Christoffel symbols of the first kind* are

$$\Gamma_{ijk} = \Gamma_{ij}^r g_{rk} \text{ for } i, j, k = 1, 2. \qquad \text{…. (1.22)}$$

(7) The Γ_{ij}^r defined in section (1.3) are the *Christoffel symbols of the second kind*. Since $\Gamma_{ij}^r = \Gamma_{ji}^r$ then $\Gamma_{ijk} = \Gamma_{jik}$. Also, since $(g_{ij})^{-1} = (g^{ij})$, we have

$$\Gamma_{ij}^m = \Gamma_{ijk} g^{km}. \qquad \text{…. (1.23)}$$

(8) Let $\vec{X}(u^1,u^2)$ be a surface and let g_{ij} and Γ_{ij}^r be as defined in sections (1.2) and (1.3). Then

$$\Gamma_{ijk} = \vec{X}_{ij}.\vec{X}_k \qquad \text{…. (1.24)}$$

$$\Gamma_{ijk} = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial u^j} + \frac{\partial g_{jk}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^k}\right) \qquad \text{…. (1.25)}$$

and

$$\Gamma_{ij}^r = \frac{1}{2} g^{kr}\left(\frac{\partial g_{ik}}{\partial u^j} + \frac{\partial g_{jk}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^k}\right) \qquad \text{…. (1.26)}$$

(9) Let $\vec{X}(u^1,u^2)$ be a surface. Then the coordinates $\vec{X}_1$ and $\vec{X}_2$ are *orthogonal* if $g_{12} = g_{21} = 0$. (This makes sense since $g_{ij} = \vec{X}_i.\vec{X}_j$).

(10) Let $\vec{X}(u^1,u^2)$ be a surface and let g_{ij} and Γ_{ij}^r be as defined in sections (1.2) and (1.3). If $\vec{X}_1$ and $\vec{X}_2$ are orthogonal coordinates, then

$$\Gamma_{ij}^r = \frac{1}{2g_{rr}}\left(\frac{\partial g_{ir}}{\partial u^j} + \frac{\partial g_{jr}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^r}\right) \qquad \text{…. (1.27)}$$

(no sums over any of *i*, *j*, *r*).

(11) With the hypotheses of the previous property (with *i*, *j*, *r*= 1,2), when *j*=r

$$\Gamma_{ir}^{r} = \frac{1}{2g_{rr}} \frac{\partial g_{rr}}{\partial u^i} = \frac{1}{2} \frac{\partial}{\partial u^i} [\ln g_{rr}] \quad \text{.... (1.28)}$$

and when *i*=*j*≠*r*

$$\Gamma_{ii}^{r} = \frac{1}{2g_{rr}} \left(-\frac{\partial g_{ii}}{\partial u^r} \right) \quad \text{.... (1.29)}$$

(12) In dimensions 3 and greater, if coordinates are mutually orthogonal, then for *i*, *j*, *r* all distinct, $\Gamma_{ij}^{r} = 0$. (in the event that one or more of *i*, *j*, *r* are equal, the two properties (10) and (11) apply).

(13) In the case of orthogonal coordinates, if we return to Gauss' notation:

$$g_{11} = E, \quad g_{12} = g_{21} = F = 0, \quad g_{22} = G$$

we have the first fundamental form (or metric form) $ds^2 = Edu^2 + Gdv^2$ on surface $\vec{X}(u,v)$. In this notation, the Christoffel symbols are then

$$\Gamma_{11}^{1} = \frac{E_u}{2E} \qquad \Gamma_{22}^{2} = \frac{G_v}{2G}$$

$$\Gamma_{12}^{1} = \Gamma_{21}^{1} = \frac{E_v}{2E} \qquad \Gamma_{21}^{2} = \Gamma_{12}^{2} = \frac{G_u}{2G} \quad \text{.... (1.30)}$$

$$\Gamma_{22}^{1} = -\frac{G_u}{2E} \qquad \Gamma_{11}^{2} = -\frac{E_v}{2G}$$

1.7 The Curvature Tensor and the Theorema Egregium [16]

A property of a surface which depends only on the metric form is an intrinsic property. We have shown (Theorem 1.4.5) that the Gaussian curvature at a point $\vec{p}$ is $K(\vec{p}) = L/g$ where *L* is the second fundamental form and *g* is the determinate of the matrix of the first fundamental form (or metric form). Therefore, to show that curvature is an intrinsic property of a surface, we need to show that *L* is a function of the g_{ij} (and their derivatives) which make up the metric form.

Lemma (1.7.1) [9]

The coefficients of the second fundamental form and the Christoffel symbols are related as follows (for h=1,2):

$$\frac{\partial \Gamma_{ik}^{h}}{\partial u^{j}} - \frac{\partial \Gamma_{ij}^{h}}{\partial u^{k}} + \Gamma_{ik}^{r}\Gamma_{rj}^{h} - \Gamma_{ij}^{r}\Gamma_{rk}^{h} = L_{ik}L_{j}^{h} - L_{ij}L_{k}^{h} \quad \text{…. (1.31)}$$

Definition (1.7.2) [3]

For a surface M with Christoffel symbols as above, define

$$R_{ijk}^{h} = \frac{\partial \Gamma_{ik}^{h}}{\partial u^{j}} - \frac{\partial \Gamma_{ij}^{h}}{\partial u^{k}} + \Gamma_{ik}^{r}\Gamma_{rj}^{h} - \Gamma_{ij}^{r}\Gamma_{rk}^{h} \quad \text{…. (1.32)}$$

These make up the *Riemann-Christoffel curvature tensor* (with h=1, 2).

Theorem (1.7.3) [11]

(Gauss' Theorema Egregium)

The Gaussian curvature of a surface is an intrinsic property. That is, the Gaussian curvature of a surface is a function of the coefficients of the metric form and their derivatives.

Proof

From the lemma and definition of R_{ijk}^{h} we have

$$R_{ijk}^{h} = L_{ik}L_{j}^{h} - L_{ij}L_{k}^{h} \quad \text{…. (1.33)}$$

Now define $R_{mijk} = g_{mh}R_{ijk}^{h} = g_{mr}R_{ijk}^{r}$ …. (1.34)

Then

$$R_{ijk}^{r} = g^{mr}R_{mijk}\,.$$

Now the Riemann-Christoffel curvature symbols R_{ijk}^{h} are intrinsic and therefore R_{mijk} are also intrinsic. Multiplying (1.33) by g_{mh} gives (summing over h=1, 2)

$$g_{mh}R^h_{ijk} = g_{mh}L_{ik}L^h_j - g_{mh}L_{ij}L^h_k = g_{hm}L_{ik}L^h_j - g_{hm}L_{ij}L^h_k$$

or

$$R_{mijk} = L_{ik}L_{jm} - L_{ij}L_{km} \text{ since } g_{im}L^i_j = L_{jm}.$$

In particular, with *m*, *j*=1 and *i*, *k*=2

$$R_{1212} = L_{22}L_{11} - L_{21}L_{21}$$

$$= L_{11}L_{22} - L_{12}L_{12}$$

(since $L_{ij} = L_{ji}$)

$$= \det(L_{ij}) = L$$

Therefore, since R_{mijk} are intrinsic, then L is intrinsic and $K = L/g = R_{1212}/g$ is intrinsic.■

Corollary (1.7.4) [11]

For a surface M determined by $\vec{X}(u,v) = \vec{X}(u^1,u^2)$ the curvature is given by

$$K = \frac{1}{g}\left[F_{uv} - \frac{1}{2}E_{vv} - \frac{1}{2}G_{uu} + \left(\Gamma^h_{12}\Gamma^r_{12} - \Gamma^h_{22}\Gamma^r_{11}\right)g_{rh}\right] \quad \text{…. (1.35)}$$

where

$$g_{11} = \vec{X}_1.\vec{X}_1 = E$$

$$g_{12} = \vec{X}_1.\vec{X}_2 = F = g_{21}$$

$$g_{22} = \vec{X}_2.\vec{X}_2 = G$$

$$g = \det(g_{ij})$$

and

$$\Gamma^r_{ij} = \frac{1}{2}g^{kr}\left(\frac{\partial g_{ik}}{\partial u^j} + \frac{\partial g_{jk}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^k}\right)$$

where $(g_{ij})^{-1} = (g^{ij})$.

Corollary (1.7.5) [11]

For a surface M determined by $\vec{X}(u,v)$ with orthogonal coordinates $\left(\vec{X}_1.\vec{X}_2 = F = 0\right)$ the curvature is

$$K = \frac{-1}{2\sqrt{EG}}\left(\frac{\partial}{\partial u}\left[\frac{G_u}{\sqrt{EG}}\right] + \frac{\partial}{\partial v}\left[\frac{E_v}{\sqrt{EG}}\right]\right) \qquad \text{.... (1.36)}$$

Proof

With F=0 and equation (1.30) (which gives the Christoffel symbols in an orthogonal coordinate system in terms of E and G) we have

$$K = \frac{1}{EG}\left\{-\frac{1}{2}E_{vv} - \frac{1}{2}G_{uu} + \left(\left(\Gamma_{12}^1\right)^2 - \Gamma_{22}^1\Gamma_{11}^1\right)g_{11} + \left(\left(\Gamma_{12}^2\right)^2 - \Gamma_{22}^2\Gamma_{11}^2\right)g_{22}\right\} \text{(1.37)}$$

(since $g_{12} = g_{21} = 0$ and $\det\left(g_{ij}\right) = g_{11}g_{22} = EG$)

$$= \frac{1}{EG}\left\{-\frac{1}{2}E_{vv} - \frac{1}{2}G_{uu} + \left(\left(\frac{E_v}{2E}\right)^2 - \left(-\frac{G_u}{2E}\right)\left(\frac{E_u}{2E}\right)\right)E + \left(\left(\frac{G_u}{2G}\right)^2 - \left(\frac{G_v}{2G}\right)\left(-\frac{E_v}{2G}\right)\right)G\right\}$$

$$= \frac{1}{EG}\left\{-\frac{1}{2}E_{vv} - \frac{1}{2}G_{uu} + \left(\frac{E_v^2}{4E^2} + \frac{E_uG_u}{4E^2}\right)E + \left(\frac{G_u^2}{4G^2} + \frac{E_vG_v}{4G^2}\right)G\right\}$$

$$= -\frac{1}{EG}\left\{\frac{1}{2}E_{vv} + \frac{1}{2}G_{uu} - \frac{EE_v^2 + EE_uG_u}{4E^2} - \frac{GG_u^2 + E_vGG_v}{4G^2}\right\}$$

$$= \frac{-1}{2EG\sqrt{EG}}\left\{\sqrt{EG}\,E_{vv} + \sqrt{EG}\,G_{uu} - \sqrt{EG}\left(\frac{E_v^2 + E_uG_u}{2E}\right) - \sqrt{EG}\left(\frac{G_u^2 + E_vG_v}{2G}\right)\right\}$$

$$= \frac{-1}{2EG\sqrt{EG}}\left\{\sqrt{EG}\,E_{vv} + \sqrt{EG}\,G_{uu} - \frac{GE_v^2 + E_uGG_u}{2\sqrt{EG}} - \frac{EG_u^2 + EE_vG_v}{2\sqrt{EG}}\right\}$$

$$= \frac{-1}{2EG\sqrt{EG}}\left\{\sqrt{EG}\,E_{vv} + \sqrt{EG}\,G_{uu} - \frac{G_u\left(EG_u + E_uG\right)}{2\sqrt{EG}} - \frac{E_v\left(EG_v + E_vG\right)}{2\sqrt{EG}}\right\}$$

$$= \frac{-1}{2\sqrt{EG}}\left\{\frac{\sqrt{EG}\,G_{uu} - \frac{G_u\left(EG_u + E_uG\right)}{2\sqrt{EG}}}{EG} + \frac{\sqrt{EG}\,E_{vv} - \frac{E_v\left(EG_v + E_vG\right)}{2\sqrt{EG}}}{EG}\right\}$$

$$= \frac{-1}{2\sqrt{EG}}\left\{\frac{\partial}{\partial u}\left[\frac{G_u}{\sqrt{EG}}\right] + \frac{\partial}{\partial v}\left[\frac{E_v}{\sqrt{EG}}\right]\right\}. \blacksquare$$

In this chapter, we introduce some interesting connections between statistics and differential geometry which includes Rao distance based on Riemannian metric whose elements are the entries in the Fisher information matrix. The set of distributions form a manifold.

The distance between two distributions is the distance along the geodesic that connects the two distributions. Also, we study Riemannian metric and Gaussian curvature for some distributions.

The goal of this chapter is to compute Gaussian curvature of some well known distributions such as normal, Gamma, Cauchy and t distribution by using different formulas. If the Gaussian curvature of any distribution approaches to the Gaussian curvature of the normal distribution, we say that the distribution converges to normal distribution.

3.1 Riemannian Metrics Based on Fisher Information [3, 31]

We define the coefficients of the expected Fisher information matrix as equal to the coefficients of the first fundamental form (Riemannian metric) on the space of probabilities, known as Fisher information metric, which it is a metric tensor for a statistical differential manifold. It can be used to calculate the informational difference between measurement. It takes the following form where $f(x,{}^{-})$ be a class of probability densities.

$$g_{ij} = \int \frac{\partial \log f(x,{}^{-})}{\partial {}^{-}_{i}} \frac{\partial \log f(x,{}^{-})}{\partial {}^{-}_{j}} f(x,{}^{-})dx \qquad \text{.... (3.1)}$$

which can be thought of intuitively as:

“The distance between two points on a statistical differential manifold is the amount of information between them, i.e. the informational difference between them”.

An equivalent form of the above equation is:

$$g_{ij} = -\int \frac{\partial^2 \log f(x,\theta)}{\partial\theta_i \, \partial\theta_j} f(x,\theta)dx = -E\left[\frac{\partial^2 \log f(x,\theta)}{\partial\theta_i \, \partial\theta_j}\right] \quad \text{.... (3.2)}$$

We define the coefficients of the first fundamental form as:

$$g_{11} = E = -E\left[\frac{\partial^2 \log f(x,\theta)}{\partial\theta_1^2}\right]$$

$$g_{12} = g_{21} = F = -E\left[\frac{\partial^2 \log f(x,\theta)}{\partial\theta_1 \partial\theta_2}\right]$$

$$g_{22} = G = -E\left[\frac{\partial^2 \log f(x,\theta)}{\partial\theta_2^2}\right]$$

where (θ_1, θ_2) are two parameters of the probability density functions. It is clear that E, F and G are functions of the parameters θ_1 and θ_2. The expectations apply to the whole sample space where the random variables are defined.

3.2 Riemannian Metrics and Geodesics [4, 32]

We know that $f(x,\theta)$ be a class of probability densities, e.g. normal, binomial. Suppose that Θ be the set of all of the values of the parameter θ (Θ is a subset of R^n).

Assume that $\theta = (\theta_1, \theta_2, \ldots, \theta_n) \in \Theta$. Consider a quadratic differential metric in the form

$$ds^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}(\theta)d\theta_i d\theta_j \quad \text{.... (3.3)}$$

Let $\theta = \theta(t)$, $t_1 \le t \le t_2$ represent a curve that joins the points in Θ $p = \theta(t_1)$, $Q = \theta(t_2)$. Let C represent the set of all such curves. The geodesic distance between p and Q will be defined by

$$d(p,Q) = \min_{c}\left|\int_{t_1}^{t_2}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}(\theta)\theta_i \theta_j\right\}^{1/2} dt\right| \quad \text{.... (3.4)}$$

A curve for which this minimum is assumed will be called a geodesic curve. This curve may be found using the calculus of variations as a solution to the Euler Lagrange equations. For this problem they take the form of a system of ordinary differential equations

$$\sum_{i=1}^{n} g_{ij}\,\bar{\ }_i + \sum_{i=1}^{n}\sum_{j=1}^{n}\Gamma_{ijk}\,\bar{\ }_i\,\bar{\ }_j = 0, \quad k = 1, \ldots, n \qquad \ldots.\ (3.5)$$

with boundary conditions of the form $\bar{\ }(t_1) = \sim\ ,\ \bar{\ }(t_2) =$ where

$$\Gamma_{ijk} = \frac{1}{2}\left[\frac{\partial g_{jk}}{\partial \bar{\ }_i} + \frac{\partial g_{ki}}{\partial \bar{\ }_j} - \frac{\partial g_{ij}}{\partial \bar{\ }_k}\right] \qquad \ldots.\ (3.6)$$

3.3 Fisher Information Rao Distances between Probability Distributions[32]

We defined the elements of the Fisher information matrix as in form (3.2).

By substitution of the probability density into (3.2) the elements of the Fisher information (3.2) matrix and the Christoffel symbols (3.6) may be obtained. The system of equations (3.5) may be solved to get the geodesic curve and the geodesic distance may be obtained using (3.4). The geodesic distance may now be computed between two pdfs.

3.4 The Rao Distance for the Normal Distribution [32]

Three cases of distances to be considered for both the univariable case include:

1- distributions with the same standard deviations but different means.

2- distributions with the same means but different standard deviations.

3- distances between normal distributions where both the mean and standard deviation are different.

In the univariable distribution case we can see the following:

Consider two normal distributions with means $«_1$ and $«_2$ and common standard deviation $\pm$.

The Rao distance is

$$d = \frac{|«_1 - «_2|}{\pm} \qquad \text{…. (3.7)}$$

For two normal distributions with common mean μ and with standard deviations $\pm_1$ and $\pm_2$. The Rao distance for this case is

$$d = \left|\sqrt{2}\log\left(\frac{\pm_1}{\pm_2}\right)\right| = \left|\frac{1}{\sqrt{2}}\log\left(\frac{\pm_1^2}{\pm_2^2}\right)\right| \qquad \text{…. (3.8)}$$

For two normal distributions with respective means $«_1$ and $«_2$ and standard deviations $\pm_1$ and $\pm_2$

$$d = \sqrt{2}\left|\log\frac{1+¢(1,2)}{1-¢(1,2)}\right| \qquad \text{…. (3.9)}$$

Observe that in (3.10) $\delta<1$

where

$$¢(1,2) = \sqrt{\frac{(«_1 - «_2)^2 + 2(\pm_1 - \pm_2)^2}{(«_1 - «_2)^2 + 2(\pm_1 + \pm_2)^2}} \qquad \text{…. (3.10)}$$

Furthermore, the derivative of d with respect to δ is

$$d' = \sqrt{2}\left[\frac{2}{1-¢^2}\right] > 0 \ \text{ for } |\delta|<1 \qquad \text{…. (3.11)}$$

Thus, the Rao distance is an increasing function of δ. As the difference between the standard deviations remaining constant δ and hence the Rao distance increases. For a constant difference between $«_1$ and $«_2$ the Rao distance becomes larger for increasing standard deviations when

$$|«_1 - «_2| \le \sqrt{\frac{2(\pm_1 - \pm_2)(\pm_2 + (1-\pm_2)(\pm_1 - \pm_2))}{\pm_2}} \qquad \text{…. (3.12)}$$

For the quantity under the square root sign in (3.12) to be positive either $\pm_1 > 1$ and $\pm_1 > \pm_2$ or $\pm_1 < 1$ and $\pm_1 < \pm_2$.

The following theorem explains this case:

Theorem (3.4.1) [32]

Assume that C is a positive number on the unit interval where

$$\left|\frac{\pm_1 - \pm_2}{\pm_1 + \pm_2}\right| < C < 1 \text{ .}$$

Then a necessary and sufficient condition that

1. $¢ \le C$
2. $d \le \sqrt{2}\log\frac{1+C}{1-C}$

is that

$$|«_1 - «_2| \le \sqrt{\frac{2C^2(\pm_1 + \pm_2)^2 - 2(\pm_1 - \pm_2)^2}{1-C^2}} \qquad \text{…. (3.13)}$$

The condition (3.12) guarantees that the expression under the square root sign in (3.13) is positive.

3.5 Riemannian Metric For Some Distributions

(a) Normal Distribution [3]

Let Ω_1 be a location scale manifold of density that has the following general form:

$$\Omega_1 = \left\{ f(x) = \frac{1}{\sqrt{2®v^2}}\exp\left(-\frac{(x-u)^2}{2v^2}\right) \middle| (u,v) \in R \times R_+ \right\}$$

where *u* is the location parameter and *v* is the scale parameter. We also assume that the regular conditions of the information metric are satisfied. The first and second partial derivative, with respect to parametric lines *u* and *v*, are given as:

$$\frac{\partial^2 \ln f}{\partial u^2} = -\frac{1}{v^2}$$

$$\frac{\partial^2 \ln f}{\partial v^2} = \frac{1}{v^2} - \frac{3(x-u)^2}{v^4} \qquad \text{…. (3.14)}$$

$$\frac{\partial^2 \ln f}{\partial v \partial u} = -\frac{2(x-u)}{v^3}$$

It is commonly known that the expected value and variance of the random variable X are u and v^2, respectively.

From this, we could easily derive the coefficient of the first fundamental form

$$E = \frac{1}{v^2} \quad , \quad F = 0, \quad G = \frac{2}{v^2}$$

$$\text{Metric Tensor} = \begin{pmatrix} \frac{1}{v^2} & 0 \\ 0 & \frac{2}{v^2} \end{pmatrix}$$

or

$$ds^2 = \frac{1}{v^2}(du)^2 + \frac{2}{v^2}(dv)^2 \qquad \text{…. (3.15)}$$

(b) Cauchy Distribution [3]

Let Ω_2 be the location scale manifold of density which has the following general form:

$$\Omega_2 = \left\{ f(x) = \frac{v}{®} \frac{1}{v^2 + (x-u)^2} \middle| x \in R, (u,v) \in R \times R_+ \right\}$$

where u is the location parameter and v is the scale parameter. The logarithm of the likelihood function of Cauchy density with one observation can be written as

$$\ln f = \ln \frac{v}{®} - \ln\left(v^2 + (x-u)^2\right). \qquad \text{…. (3.16)}$$

As before, we can derive the first two partial derivatives with respect to the parametric lines *u* and *v*.

$$\frac{\partial^2 \ln f}{\partial u^2} = \frac{2\left((x-u)^2 - v^2\right)}{\left((x-u)^2 + v^2\right)^2},$$

$$\frac{\partial^2 \ln f}{\partial v^2} = \frac{-1}{v^2} + \frac{2\left(v^2 - (x-u)^2\right)}{\left(v^2 + (x-u)^2\right)^2}, \qquad \text{…. (3.17)}$$

$$\frac{\partial^2 \ln f}{\partial v \partial u} = \frac{-4v(x-u)}{\left(v^2 + (x-u)^2\right)^2}.$$

Taking the expected values of equations (3.17), we finally get the following results:

$$E = -E\left(\frac{\partial^2 \ln f}{\partial u^2}\right) = \frac{1}{2v^2}, \quad F = -E\left(\frac{\partial^2 \ln f}{\partial u \partial v}\right) = 0, \text{ and}$$

$$G = -E\left(\frac{\partial^2 \ln f}{\partial v^2}\right) = \frac{1}{2v^2}.$$

$$\text{Metric Tensor} = \begin{pmatrix} \frac{1}{2v^2} & 0 \\ 0 & \frac{1}{2v^2} \end{pmatrix}$$

or

$$ds^2 = \frac{1}{2v^2}(du)^2 + \frac{1}{2v^2}(dv)^2 \qquad \text{…. (3.18)}$$

(c) Student t Distribution [3]

Let Ω_3 be the location scale manifold of density that has the student *t* distribution and generally has the form:

$$\Omega_3 = \left\{ f(x) = \frac{1}{v} \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\circledR r}\,\Gamma\left(\frac{r}{2}\right)} \left(1 + \frac{1}{r}\left(\frac{x-u}{v}\right)^2\right)^{-\frac{r+1}{2}} \middle|\, x \in R, (u,v) \in \mathrm{R} \times \mathrm{R}_+ \right\}$$

where u is location parameter and v is scale parameter. Let us define the following variables to simplify the notation:

$$a = \frac{1}{r}, \quad b = \frac{r+1}{2} \quad , \qquad C_r = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\circledR r}\,\Gamma\left(\frac{r}{2}\right)} .$$

Then the logarithm of likelihood function of family t, can be written as follows:

$$\ln f(x) = \ln C_r - b \ln\left(1 + a\left(\frac{x-u}{v}\right)^2\right) - \ln v \qquad \text{…. (3.19)}$$

From equation (3.19), we can derive the first and second partial derivatives:

$$\frac{\partial^2 \ln f}{\partial u^2} = \frac{2ab\left(a(x-u)^2 - v^2\right)}{\left(a(x-u)^2 + v^2\right)^2}$$

$$\frac{\partial^2 \ln f}{\partial v^2} = v^{-2} \times$$

$$\left[1 + \frac{-6ab(x-u)^2 v^{-2}\left(1 + a(x-u)^2 v^{-2}\right) + 4a^2 b(x-u)^4 v^{-4}}{\left(1 + a(x-u)^2 v^{-2}\right)^2}\right] \text{….(3.20)}$$

$$\frac{\partial^2 \ln f}{\partial v \partial u} = \frac{4abv(x-u)}{\left(a(x-u)^2 + v^2\right)^2}$$

We can now take the expected values (3.20), and have the following results:

$$E = -E\left(\frac{\partial^2 \ln f}{\partial u^2}\right) = -2ab\left(\frac{-r}{v^2(r+3)}\right) = \frac{r+1}{v^2(r+3)},$$

$$F = -E\left(\frac{\partial^2 \ln f}{\partial v \partial u}\right) = 0,$$

$$G = -E\left(\frac{\partial^2 \ln f}{\partial v^2}\right) = \frac{-1}{v^2}\left(1 - 3\frac{r+1}{r+3}\right).$$

$$\text{Metric Tensor} = \begin{pmatrix} \frac{r+1}{v^2(r+3)} & 0 \\ 0 & \frac{-1}{v^2}\left(1 - 3\frac{r+1}{r+3}\right) \end{pmatrix}$$

or

$$ds^2 = \frac{r+1}{v^2(r+3)}(du)^2 - \frac{1}{v^2}\left(1 - 3\frac{r+1}{r+3}\right)(dv)^2 \qquad \ldots.\ (3.21)$$

(d) Gamma Distribution [4]

We know that

$$p(x|\sim,\ ) = \frac{\sim}{\Gamma(\sim)} x^{\sim-1}\ e^{-\ x},\ x > 0$$

be probability density function of Gamma distribution.

The parameter space $\Theta = (\sim,\ )$ may be viewed as a manifold. Consider a new set of coordinates (α, η) where η=log (α/β). The distributions then take the form

$$f(x|\sim,¦) = \frac{\sim^{\sim} x^{\sim-1} e^{-¦\sim-\sim x e^{-¦}}}{\Gamma(\sim)},\ x>0$$

$$\ln f = \sim\ln\sim + (\sim-1)\ln x - ¦\sim - \sim x e^{-¦} - \ln\Gamma(\sim) \qquad \ldots.\ (3.22)$$

From equation (3.22), we can derive the first and second partial derivatives:

$$\frac{\partial^2 \ln f}{\partial \sim^2} = \frac{1}{\sim} - \left(\frac{\Gamma'(\sim)}{\Gamma(\sim)}\right)', \frac{\partial^2 \ln f}{\partial ¦ \partial \sim} = 0, \frac{\partial^2 \ln f}{\partial ¦^2} = -\sim x e^{-¦} \qquad \text{.... (3.23)}$$

For the Gamma family of distributions, the coefficients of the first fundamental form:

$$E = -E\left(\frac{\partial^2 \ln f}{\partial \sim^2}\right) = \left(\frac{\Gamma'(\sim)}{\Gamma(\sim)}\right)' - \frac{1}{\sim} = \Psi'(\sim) - \frac{1}{\sim}$$

where $\Psi(\sim) = \frac{\Gamma'(\sim)}{\Gamma(\sim)}$ is the digamma function.

$$F = -E\left(\frac{\partial^2 \ln f}{\partial ¦ \partial \sim}\right) = 0$$

$$G = -E\left(\frac{\partial^2 \ln f}{\partial ¦^2}\right) = \sim$$

Metric Tensor= $\begin{pmatrix} ¥(\sim) & 0 \\ 0 & \sim \end{pmatrix}$

or

$$ds^2 = ¥(\sim)(d\sim)^2 + \sim(d¦)^2 \qquad \text{…. (3.24)}$$

where $¥(\sim) = \left(\frac{\Gamma'(\sim)}{\Gamma(\sim)}\right)' - \frac{1}{\sim}$

(e) Pareto Distribution

$$f(x) = \frac{\sim}{x^{\quad +1}} \qquad \sim < x < \infty, \qquad \sim, \quad > 0$$

Then the logarithm of likelihood function of family Pareto, can be written as follows:

$$\ln f = \ln \quad + \quad \ln \sim - (\quad + 1)\ln x \qquad \text{…. (3.25)}$$

As before, we can derive the first two partial derivatives with respect to the parametric lines α and β.

$$\frac{\partial^2 \ln f}{\partial \sim^2} = -\frac{}{\sim^2}, \quad \frac{\partial^2 \ln f}{\partial \ \partial \sim} = \frac{1}{\sim}, \frac{\partial^2 \ln f}{\partial \ ^2} = -\frac{1}{\ ^2} \qquad \text{…. (3.26)}$$

We can now take the expected values of (3.26), and have the following results:

$$E = -E\left(\frac{\partial^2 \ln f}{\partial \sim^2}\right) = \frac{}{\sim^2}$$

$$F = -E\left(\frac{\partial^2 \ln f}{\partial \ \partial \sim}\right) = -\frac{1}{\sim}$$

$$G = -E\left(\frac{\partial^2 \ln f}{\partial \ ^2}\right) = \frac{1}{\ ^2}$$

$$\text{Metric Tensor=}\begin{pmatrix} \frac{}{\sim^2} & -\frac{1}{\sim} \\ -\frac{1}{\sim} & \frac{1}{\ ^2} \end{pmatrix}$$

or

$$ds^2 = \frac{}{\sim^2}(d\sim)^2 - \frac{2}{\sim} d\sim d \ + \frac{1}{\ ^2}(d \ )^2 \qquad \text{…. (3.27)}$$

(f) Wald Distribution

$$f(x;\text{«},\text{ª}) = \sqrt{\frac{\text{ª}}{2\text{®}x^3}} \exp\left(-\frac{\text{ª}(x-\text{«})^2}{2\text{«}^2 x}\right), \quad x > 0, \quad \text{«},\text{ª} > 0$$

The logarithm of the likelihood function of wald density can be written as

$$\ln f = \frac{1}{2}\ln \text{ª} - \frac{\text{ª}(x-\text{«})^2}{2\text{«}^2 x} - \frac{1}{2}\ln 2\text{®}x^3 \qquad \text{…. (3.28)}$$

From equation (3.28), we can derive the first and second partial derivatives with respect to the parametric lines μ and λ

$$\frac{\partial^2 \ln f}{\partial \text{«}^2} = \frac{-\text{ª}(3x - 2\text{«})}{\text{«}^4}$$

$$\frac{\partial^2 \ln f}{\partial \text{ª} \partial \text{«}} = \frac{-(x - \text{«})}{\text{«}^3} \qquad \text{…. (3.29)}$$

$$\frac{\partial^2 \ln f}{\partial \text{ª}^2} = -\frac{1}{2\text{ª}^2}$$

We can now take the expected values of (3.29), and have the following results:

$$E = -E\left(\frac{\partial^2 \ln f}{\partial \text{«}^2}\right) = \frac{\text{ª}}{\text{«}^3}$$

$$F = -E\left(\frac{\partial^2 \ln f}{\partial \text{ª} \partial \text{«}}\right) = 0$$

$$G = -E\left(\frac{\partial^2 \ln f}{\partial \text{ª}^2}\right) = \frac{1}{2\text{ª}^2}$$

$$\text{Metric Tensor} = \begin{pmatrix} \frac{\text{ª}}{\text{«}^3} & 0 \\ 0 & \frac{1}{2\text{ª}^2} \end{pmatrix}$$

or

$$ds^2 = \frac{\text{ª}}{\text{«}^3}(d\text{«})^2 + \frac{1}{2\text{ª}^2}(d\text{ª})^2 \qquad \text{…. (3.30)}$$

3.6 The Gaussian Curvature of the Probability Distribution [3]

First, we define the six well known Christoffel symbols as:

$$\Gamma_{11}^1 = \frac{GE_u - 2FF_u + FE_v}{2\left(EG - F^2\right)}, \qquad \Gamma_{12}^2 = \frac{EG_u - FE_v}{2\left(EG - F^2\right)},$$

$$\Gamma_{11}^2 = \frac{2EF_u - EE_v - FE_u}{2\left(EG - F^2\right)}, \qquad \Gamma_{22}^1 = \frac{2GF_v - GG_u - FG_v}{2\left(EG - F^2\right)}, \qquad \text{…. (3.31)}$$

$$\Gamma_{12}^1 = \frac{GE_v - FG_u}{2\left(EG - F^2\right)}, \qquad \Gamma_{22}^2 = \frac{EG_v - 2FF_v + FG_u}{2\left(EG - F^2\right)}.$$

Since E, F and G are functions of parameters (u, v) and are continuously twice differentiable, E_u, E_v, F_u, F_v, G_u and G_v all exists and are all well defined. Additionally, no assumption is made regarding F=0, and so the parametric lines are not necessarily orthogonal. However, if F=0, the six Christoffel symbols can be greatly simplified.

Now, we select four formulas that can be used to compute the Gaussian curvature of the distributions:

(A) $$K = \frac{1}{\sqrt{EG}}\left(\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\right)\right)$$

(B)

$$K = -\frac{1}{2\sqrt{EG-F^2}}\left[\frac{\partial}{\partial u}\frac{G_u - F_v}{\sqrt{EG-F^2}} - \frac{\partial}{\partial v}\frac{F_u - E_v}{\sqrt{EG-F^2}}\right] - \frac{1}{4\left(EG-F^2\right)^2}\begin{vmatrix} E & F & G \\ E_u & F_u & G_u \\ E_v & F_v & G_v \end{vmatrix}$$

(C) $$K = \frac{1}{D}\left[\frac{\partial}{\partial v}\left(\frac{D}{E}\Gamma_{11}^2\right) - \frac{\partial}{\partial u}\left(\frac{D}{E}\Gamma_{12}^2\right)\right] = \frac{1}{D}\left[\frac{\partial}{\partial u}\left(\frac{D}{G}\Gamma_{22}^1\right) - \frac{\partial}{\partial v}\left(\frac{D}{G}\Gamma_{12}^1\right)\right]$$

where $D^2 = EG - F^2$

(D) $$K = \frac{R_{1212}}{EG-F^2} = \frac{(12,12)}{EG-F^2},$$

where (12, 12)=$R_{1212} = \sum_{m=1}^{2} R_{121}^m g_{m2}$,

$$R_{ijk}^1 = \frac{\partial}{\partial u_j}\Gamma_{ik}^1 - \frac{\partial}{\partial u_i}\Gamma_{jk}^1 + \Gamma_{ik}^m\Gamma_{mj}^1 - \Gamma_{jk}^m\Gamma_{mi}^1, \text{ sum on } m,$$

where the quantities of R_{ijk}^1 are components of a tensor of the fourth order. This tensor is called the mixed Riemann curvature tensor. Notice that g_{11}, g_{12} and g_{22} are simply tensor notation for E, F and G.

Clearly formula (A) is a special case that is valid only when the parametric lines are orthogonal. Formula (D) is a general form represented in Riemann symbols of the first and second kind, respectively. In formula

(D), R_{1212}, the inner product of the mixed Riemann curvature tensor and the metric tensor, is called the covariant Riemann curvature tensor; it is a covariant tensor of the fourth order. The components R^1_{ijk} and R_{1212} are also known as Riemann symbols of the first and second kind, respectively. Notice that Riemann symbols of the second kind will satisfy the relation $R_{1212} = -R_{1221} = -R_{2112} = R_{2121}$, the well-known property of skew-symmetry with respect to the last two indices.

It is useful to be aware of the fact that the Christoffel symbols depend only on the coefficients of the first fundamental form and their derivatives. The same holds true for the mixed Riemann curvature tensor. From this point of view, as long as we can find the coefficients of the first fundamental form of a given distribution and their first and second derivatives, we can uniquely define the corresponding Christoffel symbols and hence mixed Riemann curvature tensors. Thus, the process of computing the covariant Riemann curvature tensor and Gaussian curvature is simplified.

From a different perspective, we know that the mixed Riemann curvature tensor will link with the coefficient of the second fundamental form, namely *e*, *f*, and *g*, by

$R^n_{121} = g^{n2}(eg - f^2)$, where

$$g^{11} = \frac{G}{EG - F^2}, \quad g^{12} = \frac{-F}{EG - F^2}, \quad g^{22} = \frac{E}{EG - F^2}.$$

The above relation can then be easily used to derive $R_{1212} = eg - f^2$ and the result will be the original fundamental definition of Gaussian curvature.

3.7 Some Examples

In this section, we give the details of some examples which will deal with the location-scale family of densities and methods of finding those with negative Gaussian curvature.

Chen, W.W.S. used the formula (D) to compute our Gaussian curvature with the p.d.f. of normal, Cauchy distribution and *t* family distribution but, Gruber, M. H. J. used the formula (B) for the normal and Gamma distribution.

Example (1) [3]

The metric tensor of the normal distribution as calculated in section (3.5)-a where the coefficients of the metric tensor are $E=\frac{1}{v^2}$, F=0, $G=\frac{2}{v^2}$, as well as their corresponding derivatives with respect to the parametric lines u and v:

$$E_u=0,\quad E_v=\frac{-2}{v^3},\quad G_u=0,\quad G_v=\frac{-4}{v^3},\quad EG=\frac{2}{v^4}\ .$$

Substituting the listed results into formula (D) to compute the Gaussian curvature. Chen, W.W.S. found that $\Gamma_{11}^{1}=\Gamma_{12}^{2}=\Gamma_{22}^{1}=0$.

$$R_{121}^{2}=\frac{\partial}{\partial v}\Gamma_{11}^{2}-\Gamma_{21}^{1}\Gamma_{11}^{2}+\Gamma_{11}^{2}\Gamma_{22}^{2}=\frac{-1}{2v^2}$$

$$R_{1212}=R_{121}^{2}G=\frac{-1}{v^4}$$

$$K=\frac{R_{1212}}{EG}=\frac{-1}{v^4}\frac{v^4}{2}=-\frac{1}{2}.$$

Also, Gruber M. H. J. found the same result when he used formula (B).

Example (2) [3]

The metric tensor of the Cauchy distribution as calculated in section (3.5)-b. The derivatives of the coefficients of the metric tensor and six Christoffel symbols are all straightforward computations. Due to the fact that the Cauchy distribution is the same as the normal distribution, that is, $\Gamma_{11}^{1} = \Gamma_{12}^{2} = \Gamma_{22}^{1} = 0$, Chen, W.W.S. used formula (D) to derive the Gaussian curvature.

$$R_{121}^{2} = \frac{\partial}{\partial v}\Gamma_{11}^{2} - \Gamma_{21}^{1}\Gamma_{11}^{2} + \Gamma_{11}^{2}\Gamma_{22}^{2} = \frac{-1}{v^2}$$

$$EG = \frac{1}{4v^4}$$

$$R_{1212} = \sum_{m=1}^{2} R_{121}^{m} g_{m2} = R_{121}^{2} G = \frac{-1}{2v^4}$$

$$K = \frac{R_{1212}}{EG} = 4v^4\left(\frac{-1}{2v^4}\right) = -2\,.$$

This result shows that Cauchy distribution does not converge to the normal distribution.

Example (3) [3]

The metric tensor of the t distribution as calculated in section (3.5)-c. The procedures to compute the derivative of the coefficient of the metric tensor and six Christoffel symbols became routine procedures.

Chen, W.W.S. computed the Riemann symbols of the first and second kind, respectively. Thus, the Gaussian curvature is calculated as:

$$R_{121}^{2} = \frac{\partial}{\partial v}\Gamma_{11}^{2} - \Gamma_{21}^{1}\Gamma_{11}^{2} + \Gamma_{11}^{2}\Gamma_{22}^{2} = \frac{-(r+1)}{2rv^2}$$

$$R_{1212} = \sum_{m=1}^{2} R_{121}^{m} g_{m2} = R_{121}^{2} G = \frac{-(r+1)}{v^4(r+3)}$$

$$K = \frac{R_{1212}}{EG} = \frac{-(r+3)}{2r}$$

This result shows that t distribution converges to the normal distribution at certain value of r.

Example (4) [4]

The metric tensor of the Gamma distribution as calculated in section (3.5)-d. The Gaussian curvature of the Gamma family of distributions depends only on $\sim \in (0,\infty)$.

Gruber M.H.J. used the formula (B) to compute the Gaussian curvature. To find $K(\sim)$ he used the formula $K(\sim) = -\frac{1}{4}\frac{\left(\sim¥(\sim)\right)'}{\left(\sim¥(\sim)\right)^2}$, which obtained by applying Gauss Remarkable Theorem. It can also be shown that for normal distributions with parameters μ and σ that the Gaussian curvature is always $-\frac{1}{2}$. Thus,

$$\lim_{\alpha\to\infty} K(\alpha) = -\frac{1}{2}, \quad \lim_{\alpha\to 0} K(\alpha) = -\frac{1}{4}.$$

As $\sim \to \infty$ the curvature of the Gamma family of distributions tends toward $-\frac{1}{2}$ the curvature of the normal family of distributions.

This result shows that Gamma distribution converges to the normal distribution.

In this chapter, we explain some theorems of statistics about continuous random variable, some continuous distributions and Fisher information. Also, we show the theorems of law of the large numbers and central limit theorem.

2.1 Random Variable [17, 18]

Given a random experiment with sample space S. A *Random Variable* (r.v.)X is a function which assigns to each element $\mu \in S$ a real number $X(\mu)$ in the set E.

There are two types of random variables, discrete and continuous. A random variable has either probability mass function (discrete random variable) or probability density function (continuous random variable).

A random variable X is called *continuous* if E is the set of the real numbers R or any subset of R. (or a continuous random variable is one which takes an infinite number of possible values. Continuous random variables are usually measurements. Examples include height, weight, the amount of sugar in an orange, the time required to run a mile.

2.2 Probability Density Function [19, 20]

The *probability density function* (p.d.f), $f(x)$, of a continuous random variable X, is a function from R to $[o,\infty)$.

The p.d.f. must satisfy the following two conditions:

a. $f(x) \geq 0$ for all x.

b. $\int_{-\infty}^{\infty} f(x)dx = 1.$

2.3 Distribution Function of a Continuous Random Variable

If X is a random variable defined on the sample space S, the *cumulative distribution function* (c.d.f), or simply the *distribution function*

F of the random variable X, is the function from R to [0, 1], and is defined for all $x \in R$, by

$$F(x) = p[X \leq x] = \int_{-\infty}^{x} f(u)du \qquad \text{.... (2.1)}$$

In other word, $F(x)$ denotes the probability that the random variable X takes on a value that is less than or equal to x.

For each random variable X there is one and only one distribution function. [20]

We can show that the continuous distribution function have the following properties: [17, 21]

1- $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$

2- $F(\infty) = \lim_{x \to \infty} F(x) = 1$

3- F is a non decreasing function, that is, if $a \leq b$, then $F(a) \leq F(b)$

4- $f(x) = \frac{d}{dx} F(x)$

5- $p(a \leq X \leq b) = p(X \leq b) - p(X \leq a)$

$$= F(b) - F(a)$$

$$= \int_{a}^{b} f(x)dx \text{ for } a < b$$

2.4 Mathematical Expectation

The *expected value* (or *population mean*) of a random variable indicates its average or central value.

The expected value of a random variable X is symbolized by $E(X)$ or μ.

If X is a continuous random variable with probability density function $f(x)$, then the expected value of X is defined by

$$« = E(X) = \int_{-\infty}^{\infty} x f(x) dx \qquad \text{…. (2.2)}$$

also,

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx \qquad \text{…. (2.3)}$$

where $g(X)$ is a function of a continuous r.v. [18, 21]

The properties of the expectation are: [22]

1- $E(a)=a$ for a constant a

2- $E(a_1 X + a_2) = a_1 E(X) + a_2$ for constants a_1, a_2

3- $E\left(\frac{X}{Y}\right) = E(X).E\left(\frac{1}{Y}\right)$ if X and Y independent r.v's.

4- $E[ag(X)] = aE[g(X)]$.

5- $E[a_1 g_1(X) \pm a_2 g_2(X)] = a_1 E[g_1(X)] \pm a_2 E[g_2(X)]$.

6- $E[g_1(X)] \le E[g_2(X)]$ if $g_1(x) \le g_2(x)$ for all x.

2.5 Variance of Random Variable [18, 21]

The (*population*) *variance* of a random variable is a non- negative number which gives an idea of how widely spread the values of the random variable are likely to be; the larger the variance, the more scattered the observations on average.

The variance of a r.v.X is symbolized by var(X) or v(X) or $\pm^2$ and is defined by

$$\text{var}(X) = \pm^2 = E[X - E(X)]^2 = E(X^2) - [E(X)]^2 \qquad \text{…. (2.4)}$$

where $E(X)$ is the expected value of X.

If X is a continuous r.v. then

$$\text{var}(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - «^2$$

The *standard deviation* of *X* is symbolized by σ, and is defined as

$$\pm = \sqrt{\mathrm{var}(X)}$$

2.6 Some Special Continuous Distributions

(a) Gamma Distribution [23]

A continuous random variable *X*, taking all real values in the range (0, ∞) is said to have a Gamma distribution with parameters α and β if its probability density function is given by

$$f(x) = \frac{{}^{\sim} x^{\sim -1} e^{- \ x}}{\Gamma(\sim)} \qquad \begin{matrix} x > 0 \\ \sim, \quad > 0 \end{matrix} \qquad \text{.... (2.5)}$$

where α, β are constants. We write *X*~ G (α, β).

The gamma function $\Gamma(\sim)$ is defined as

$$\Gamma(\sim) = \int_0^{\infty} x^{\sim -1} e^{-x} dx$$ for $\sim > 0$ and has the following properties:

1- If α is a positive integer, then $\Gamma(\sim) = (\sim -1)!$

2- $\Gamma(\sim) = (\sim -1)\Gamma(\sim -1)$ for any positive real number α

3- $\Gamma\left(\frac{1}{2}\right) = \sqrt{®}$

The mean of this distribution is $\frac{\sim}{\ }$, and the variance is $\frac{\sim}{\ ^{2}}$.

The distribution function is

$$F(x) = 1 - \sum_{j=0}^{\sim -1} \frac{e^{- \ x} (\ x)^{j}}{j!}$$

(b) Normal Distribution [17]

A continuous random variable *X*, taking all real values in the range (-∞, ∞) is said to have a normal distribution with parameters μ and σ if its probability density function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad \begin{matrix} -\infty < x < \infty \\ -\infty < \mu < \infty, \sigma > 0 \end{matrix} \qquad \text{.... (2.6)}$$

where μ, σ are real constants. We write X~ N (μ, σ^2).

The mean of this distribution is μ and the variance is σ^2.

The distribution function is

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du$$

The values of $F(x)$ can be taken from a special table.

Notes [17]

1. If X~ N (μ, σ^2), then $Z = \frac{X - \mu}{\sigma}$ is a standard normal variable with mean 0 and variance 1.
2. The p.d.f. of the standard normal variable Z is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad -\infty < z < \infty \qquad \text{.... (2.7)}$$

and the corresponding distribution function

$\Phi(z) = P[Z \leq z]$ is given by

$$\Phi(z) = \int_{-\infty}^{z} f(u)du = \int_{-\infty}^{z} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du$$

The values of $\Phi(z)$ can be taken from a special table.

(c) Cauchy Distribution [22]

A continuous random variable X, taking all real values in the range (-∞, ∞) is said to have a Cauchy distribution with parameters α and β if its probability density function is given by

$$f(x) = \frac{1}{\pi\beta\left[1+\left(\frac{x-\alpha}{\beta}\right)^2\right]} \qquad \begin{matrix} -\infty < x < \infty \\ -\infty < \alpha < \infty, \quad \beta > 0 \end{matrix} \qquad \text{…. (2.8)}$$

where α, β are constants. We write X~ Cauchy (α, β).

The mean and the variance of this distribution do not exist.

The distribution function F is given by

$$F(x) = \frac{1}{\pi}\tan^{-1}\left(\frac{x-\alpha}{\beta}\right) + \frac{1}{2}$$

(d) Exponential Distribution [23]

A continuous random variable X, taking all real values in the range $[0, \infty)$ is said to have an exponential distribution with parameter λ if its probability density function is given by

$$f(x) = \lambda e^{-\lambda x} \qquad \begin{matrix} x \geq 0 \\ \lambda > 0 \end{matrix} \qquad \text{…. (2.9)}$$

where λ is a constant. We write $X \sim \text{Exp}(\lambda)$.

The mean of this distribution is $\frac{1}{\lambda}$ and the variance is $\frac{1}{\lambda^2}$.

The distribution function F is given by

$$F(x) = \int_0^x f(u)du = \int_0^x \lambda e^{-\lambda u} du = 1 - e^{-\lambda x}.$$

(e) Beta Distribution [22]

A continuous random variable X, taking all real values in the range $(0, 1)$ is said to have a beta distribution with parameters a and b if its probability density function is given by

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} \qquad \begin{matrix} 0 < x < 1 \\ a, b > 0 \end{matrix} \qquad \text{…. (2.10)}$$

where a, b are constants. We write X~ Beta (a, b).

The Beta function $B(a, b)$ is defined as

$$B(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}dx \qquad \text{for } a>0 \text{ , } b>0.$$

The Beta function is related to the gamma function according to the following formula:

$$B(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

The mean of this distribution is $\frac{a}{a+b}$ and the variance is

$$\frac{ab}{(a+b+1)(a+b)^2}$$

The distribution function F is given by

$$F(x)=\frac{1}{B(a,b)}\int_0^x u^{a-1}(1-u)^{b-1}du$$

The values of $F(x)$ can be taken from a special table.

(f) Double Exponential (Laplace) Distribution [22]

A continuous random variable X, taking all real values in the range $(-\infty,\infty)$ is said to have a double exponential distribution with parameters α and β if its probability density function is given by

$$f(x)=\frac{1}{2}e^{-\frac{|x-\sim|}{\ }} \qquad \begin{array}{c} -\infty<x<\infty \\ -\infty<\sim<\infty, \quad\ >0 \end{array} \qquad \text{…. (2.11)}$$

where α, β are constants. We write X~ Lap (α, β).

The mean of this distribution is $\sim$ and the variance is $2\ ^2$

The distribution function F is given by

$$F(x)=\frac{1}{2}e^{-\frac{\sim-x}{\ }} \qquad x<\sim$$

$$=1-\frac{1}{2}e^{-\left(\frac{x-\sim}{\ }\right)} \qquad x\geq\sim$$

(g) Pareto Distribution [22]

A continuous random variable *X*, taking all real values in the range (*α*, ∞) is said to have a Pareto distribution with parameters α and β if its probability density function is given by

$$f(x)=\frac{\beta\alpha^{\beta}}{x^{\beta+1}} \qquad \alpha<x<\infty,\quad \alpha,\beta>0 \qquad \text{.... (2.12)}$$

where α and β are constants. We write *X*~ Par (α, β).

The mean of this distribution is $\frac{\alpha\beta}{\beta-1}$ for $\beta>1$ and the variance is

$$\frac{\beta\alpha^{2}}{(\beta-1)^{2}(\beta-2)} \text{ for } \beta>2.$$

The distribution function *F* is given by

$$F(x)=\int_{\alpha}^{x}\frac{\beta\alpha^{\beta}}{u^{\beta+1}}du=\alpha^{\beta}\left[\alpha^{-\beta}-x^{-\beta}\right].$$

(h) Wald Distribution [24]

A continuous random variable *X*, taking all real values in the range (0, ∞) is said to have a Wald distribution with parameters μ and λ if its probability density function is given by

$$f(x)=\sqrt{\frac{\lambda}{2\pi x^{3}}}\,e^{-\frac{\lambda(x-\mu)^{2}}{2\mu^{2}x}} \qquad x>0,\quad \mu,\lambda>0 \qquad \text{.... (2.13)}$$

where *μ* and λ are constants. We write *X*~ Wald (*μ*, λ).

The mean of this distribution is *μ* and the variance is $\frac{\mu^{3}}{\lambda}$

The distribution function *F* is given by

$$F(x)=F^{*}\left\{(x-1)\sqrt{\frac{\lambda}{x\mu}}\right\}+e^{\frac{2\lambda}{\mu}}.F^{*}\left\{-(x+1)\sqrt{\frac{\lambda}{x\mu}}\right\}$$

since $F^*(.)$ represents distribution function in the standard normal distribution.

We can use the following approximation:

$F(x) \approx 1 - e^{-\frac{\text{ª}(x-2)}{2\text{«}}} .\ln x$ when X is relatively large.

(i) Weibull Distribution [23]

A continuous random variable *X*, taking all real values in the range (0, ∞) is said to have a Weibull distribution with parameters α and β if its probability density function is given by

$$f(x) = \sim\ x^{\ -1} e^{-\sim x} \qquad \text{for} \qquad \begin{matrix} x > 0 \\ \sim > 0, \quad > 0 \end{matrix} \qquad \text{…. (2.14)}$$

where α, β are constants. We write *X*~ Weibull (α, β).

The mean of this distribution is $\sim^{-1/}\ \Gamma\left(1+\frac{1}{\ }\right)$ and the variance is

$$\sim^{-2/}\ \left[\Gamma\left(1+\frac{2}{\ }\right) - \left[\Gamma\left(1+\frac{1}{\ }\right)\right]^2\right]$$

The distribution function is

$$F(x) = 1 - e^{-\sim x} \qquad \text{for } x > 0$$

(j) Student's t Distribution [22]

A continuous random variable *X*, taking all real values in the range (-∞, ∞) is said to have a t distribution with *r* degrees of freedom if its probability density function is given by

$$f(x) = \frac{\Gamma[(r+1)/2]}{\Gamma(r/2)} \frac{1}{\sqrt{r\text{®}}} \frac{1}{\left(1+\frac{x^2}{r}\right)^{(r+1)/2}} \qquad \begin{matrix} -\infty < x < \infty \\ r > 0 \end{matrix} \qquad \text{…. (2.15)}$$

where *r* is the parameter. We write $X \sim t_r$

The mean of this distribution is $\mu=0$ for $r > 1$, and the variance is $\frac{r}{r-2}$ for $r>2$

The distribution function F is given by

$$F(x)=\frac{\Gamma[(r+1)/2]}{\Gamma(r/2)}\frac{1}{\sqrt{r\pi}}\int_{-\infty}^{x}\frac{1}{\left(1+\frac{u^2}{r}\right)^{(r+1)/2}}du$$

The values of $F(x)$ can be taken from a special table.

2.7 Fisher Information

In statistics and information theory, the Fisher information (denoted $I(\theta)$) is the variance of the score. Fisher information is thought of as the amount of information that an observable random variable carries about an unobservable parameter θ upon which the probability distribution of X depends. Since the expectation of the score is zero, the variance is also the second moment of the score and so the Fisher information can be written as

$$I(\theta)=E\left[\left[\frac{\partial}{\partial\theta}\log f(X;\theta)\right]^2\right] \quad \text{…. (2.16)}$$

where f is the probability density function of random variable X and, consequently, $0 \le I(\theta) < \infty$. Then it is the expectation of the square of the score. A random variable carrying high Fisher information implies that the absolute value of the score is frequently high.

Note that the information as defined above is not a function of a particular observation, as the random variable X has been averaged out [25].

Under certain regularity conditions:

$$I(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2}\log f(X;\theta)\right] \qquad \text{…. (2.17)}$$

We can prove that as follows:

$$\frac{\partial^2}{\partial \theta^2}\log f(x;\theta) = \frac{\partial}{\partial \theta}\left[\frac{1}{f(x;\theta)}\frac{\partial f(x;\theta)}{\partial \theta}\right]$$

$$= -\frac{1}{f^2(x;\theta)}\left[\frac{\partial f(x;\theta)}{\partial \theta}\right]^2 + \frac{1}{f(x;\theta)}\frac{\partial^2 f(x;\theta)}{\partial \theta^2}$$

$$= -\left[\frac{\partial \log f(x;\theta)}{\partial \theta}\right]^2 + \frac{1}{f(x;\theta)}\frac{\partial^2 f(x;\theta)}{\partial \theta^2}$$

Assuming that the regularity conditions allow us to take the differentiation outside the integration sign:

$$E\left[\frac{1}{f(\mathrm{X};\theta)}\frac{\partial^2 f(\mathrm{X};\theta)}{\partial \theta^2}\right] = \int_{\mathrm{X}} \frac{\partial^2 f(x;\theta)}{\partial \theta^2}dx$$

$$= \frac{\partial^2}{\partial \theta^2}\int_{\mathrm{X}} f(x;\theta)dx$$

$$= \frac{\partial^2}{\partial \theta^2}1 = 0$$

Thus $E\left[-\frac{\partial^2}{\partial \theta^2}\log f(X;\theta)\right] = E\left[\left(\frac{\partial}{\partial \theta}\log f(X;\theta)\right)^2\right] = I(\theta)$. [26]■

In the case when there are d parameters, thus making θ a vector of length d, then the Fisher information matrix (FIM) is defined as having the (i, j) element as

$$\left(I(\theta)\right)_{i,j} = E\left[\frac{\partial}{\partial \theta_i}\log f(X;\theta)\frac{\partial}{\partial \theta_j}\log f(X;\theta)\right] \qquad \text{…. (2.18)}$$

or

$$\left(I(\theta)\right)_{i,j} = -E\left[\frac{\partial^2}{\partial \theta_i \partial \theta_j}\log f(X;\theta)\right] \qquad \text{…. (2.19)}$$

The FIM is a $d \times d$ and symmetric matrix. [25]

2.8 The Characteristic Function [17]

The characteristic function (c.f.) of a r.v.X is denoted by $\phi_X(t)$, and is defined as

$$\phi_X(t) = E\left(e^{itX}\right) \qquad \text{…. (2.20)}$$

$= \int_{-\infty}^{\infty} e^{itx} f(x)dx$ if X is a continuous r.v., where $i = \sqrt{-1}$, and t is any real number.

2.9 Convergence of Random Variables [27]

The convergence of sequence of random variables to some limiting random variable is an important concept in probability theory, and its applications to statistics and stochastic processes. Throughout the following, we assume that $\{X_n\}$ is a sequence of random variables, and X is a random variable, and all of them are defined on the same probability space. There exist several kinds of convergence of the sequence $\{X_n\}$ of r.v's, which are:

1- Convergence in Probability (or Stochastic Convergence). [17, 27]

A sequence of random variables $\{X_n\}$ $(n \geq 1)$ is said to be convergent in probability to a r.v. X (or weakly convergent), written as $X_n \xrightarrow{p} X$, if for every ε >0,

$$\lim_{n\to\infty} p\{|X_n - X| \geq \varepsilon\} = 0 \qquad \text{…. (2.21)}$$

or equivalently, $p\{|X_n - X| < \varepsilon\} \to 1$ as $n \to \infty$

2- Convergence Almost Surely [17, 27]

A sequence of r.v.'s $\{X_n\}$ is said to be convergent almost surely (or strongly) to a r.v.X, written as $X_n \xrightarrow{a.s.} X$ if

$$p\left\{\lim_{n\to\infty} X_n = X\right\} = 1 \qquad \text{…. (2.22)}$$

In this case, we write $\lim_{n\to\infty} X_n = X$ with probability one (simply $w.p.1$)

3- Convergence in Distribution [17, 27]

The sequence $\{X_n\}$ of r.v.'s is said to be convergent in distribution (or in law) to the r.v.X if the distribution function $F_n(x)$ of X_n converges to the distribution function $F(x)$ of X at every continuity point x of F. It is written as $X_n \xrightarrow{d} X$.

Theorem (2.9.1) [17]

Let ¤$_n(t)$ be the characteristic function of X_n. If $X_n \xrightarrow{d} X$ then ¤$_n(t) \to$ ¤(t), where ¤(t) is the c.f. of X. If ¤$_n(t) \to$ ¤(t) and the limit function is continuous at t=0, then $X_n \xrightarrow{d} X$.

2.10 The Law of Large Numbers [28]

In probability theory, several laws of large numbers say that the average of a sequence of random variables with a common distribution converges to their common expectation, in the limit as the size of the sequence goes to infinity. The phrase "law of large numbers" is also sometimes used to refer to the principal that the probability of any possible event (even an unlikely one). Occurring at least once in a series increases with the number of events in the series. For example, the odds that we will win the lottery are very low, however, the odds that someone will win the lottery are quite good, provided that a large enough number of people purchased lottery tickets.

Theorem (2.10.1) [17, 28]

(The Weak Law of Large Numbers)

Let $X_1, X_2, ...,$ be a sequence of independent and identically distributed random variables (i.i.d. r.v's), each having a finite mean $\mu = E(X_i)$, i=1, 2, … . Then, for any ε>0 and $\text{var}(X_i) = \sigma^2 < \infty$,

$$p\left\{\left|\frac{X_1 + X_2 + ... + X_n}{n} - \mu\right| \geq \varepsilon\right\} \to 0 \text{ as } n \to \infty \qquad \text{…. (2.23)}$$

If we write $\overline{X}_n = \dfrac{X_1 + X_2 + ... + X_n}{n}$

then, from (2.23), it follows

$\overline{X}_n \xrightarrow{p} \mu$ as $n \to \infty$, where $\mu = E(\overline{X}_n)$.

Theorem (2. 10.2) [17, 28]

(The Strong Law of Large Numbers)

Let X_1, X_2, … be independent and identically distributed with a finite mean $\mu = E(X_i)$ and finite fourth central moment $\mu_4 = E(X_i - \mu)^4$ for i= 1, 2, … . Then $\overline{X}_n \xrightarrow{a.s.} \mu$.

That is $P\left\{\lim\limits_{n \to \infty} \overline{X}_n = \mu\right\} = 1$ …. (2.24)

2.11 Central limit Theorem [29]

The central limit theorem is one of the most remarkable results of the theory of probability. In its simplest form, the theorem states that the sum of a large number of independent observations from the same distribution has, under certain general conditions, an approximate normal distribution. Moreover, the approximation steadily improves as the number of observations increases. The theorem is considered the basis of probability theory, although a better name would be normal convergence theorem.

Theorem (2.11.1)

(The Central Limit Theorem)

Let X_1, X_2, X_3, … be a sequence of random variables which are defined on the same probability space, share the same probability distribution D and are independent. Assume that both the expected value ***μ*** and the standard deviation σ of D exist and are finite. Consider the sum $S_n = X_1 + ... + X_n$. Then the expected value of S_n is $n\boldsymbol{\mu}$ and its standard deviation is $\sigma\sqrt{n}$. Furthermore, informally speaking, the distribution of S_n approaches to the normal distribution $N(n\mu, n\sigma^2)$ as n approaches to ∞.

In order to clarify the word “approaches” in the last sentence, we standardize S_n by setting

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}.$$

Then the distribution of Z_n converges towards the standard normal distribution $N(0, 1)$ as n approaches to ∞ (this is convergence in distribution). This means: if $\Phi(z)$ is the cumulative distribution function of $N(0, 1)$, then for every real number z, we have

$$\lim_{n\to\infty} p\left(Z_n \le z\right) = \Phi(z),$$

or equivalently,

$$\lim_{n\to\infty} p\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \le z\right) = \Phi(z)$$

where

$\overline{X}_n = S_n / n = \left(X_1 + ... + X_n\right)/n$ is the “sample mean”. [30]

Proof

From theorem (2.9.1), it is sufficient to show that $¤_{Z_n}(t) \to ¤(t)$ as $n \to \infty$, where $¤_{Z_n}(t)$ and $¤(t)$ are the characteristic functions of Z_n and Z (standard normal distribution), respectively. We have

$$¤_{S_n}(t) = ¤_{X_1+\ldots+X_n}(t) = \prod_{i=1}^{n} ¤_{X_i}(t)$$

$$= \left[¤_X(t)\right]^n \text{ because } X_i \text{ are i.i.d.r.v's.}$$

Therefore,

$$¤_{Z_n}(t) = e^{-itn« / \sqrt{n}±} . ¤_{S_n}\left(\frac{t}{\sqrt{n}±}\right)$$

$$= e^{-it\sqrt{n}« / ±} . \left[¤_X\left(\frac{t}{\sqrt{n}±}\right)\right]^n .$$

By the Taylor series expansion of $¤_X(t)$, we get

$$¤_X(t) = 1 + itE(X) - \frac{t^2}{2!}E(X^2) + 0(t^2) .$$

Therefore,

$$¤_X\left(\frac{t}{\sqrt{n}±}\right) = 1 + \frac{it}{\sqrt{n}±}E(X) - \frac{t^2}{2n±^2}E(X^2) + 0\left(\frac{t^2}{n±^2}\right),$$

and

$$¤_{Z_n}(t) = e^{-\frac{it\sqrt{n}«}{±}}\left[1 + \frac{itE(X)}{\sqrt{n}±} - \frac{t^2}{2n±^2}E(X^2) + 0\left(\frac{t^2}{n±^2}\right)\right]^n$$

$$= e^{-\frac{it\sqrt{n}«}{±}}\left[1 + \frac{it«}{\sqrt{n}±} - \frac{t^2}{2n±^2}\left(±^2 + «^2\right) + 0\left(\frac{t^2}{n±^2}\right)\right]^n$$

$$= e^{-\frac{it\sqrt{n}«}{±}}\left[1 + \frac{it«}{\sqrt{n}±} - \frac{t^2}{2n} + 0\left(\frac{t^2}{n±^2}\right)\right]^n$$

We have

$$\left[1+\frac{it\text{«}\sqrt{n}/\pm\ -t^2/2}{n}+0\left(\frac{t^2}{n\pm^2}\right)\right]^n \longrightarrow e^{\frac{it\text{«}\sqrt{n}}{\pm}-t^2/2} \ as\ n\to\infty$$

Therefore

$$\lim_{n\to\infty}\text{¤}_{Z_n}(t)=e^{-\frac{it\sqrt{n}\text{«}}{\pm}}.e^{it\text{«}\sqrt{n}/\pm-t^2/2}=e^{-t^2/2}$$

$$=\text{¤}_Z(t),$$

where $\text{¤}_Z(t)=e^{-t^2/2}$ is the c.f. of the standard normal distribution.

By theorem (2.9.1), the distribution of Z_n approaches to the standard normal distribution. [17]

In this chapter, we try to use the different formulas which explained in chapter (3) to compute the Gaussian curvature for normal distribution, Cauchy, t distribution and Gamma distribution by using the formulas which do not use by others. Also, we find the Gaussian curvature of other distributions such as Pareto distribution and Wald distribution by using different formulas and show that if they are convergent with respect to normal distribution by comparing the value of Gaussian curvature for each distribution with the value of Gaussian curvature for the normal distribution.

4.1 Normal Distribution

(1) We use the formula (A) to compute the Gaussian curvature of the normal distribution.

$$K = \frac{-1}{\sqrt{EG}}\left(\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\right)\right)$$

We can find that

$$\sqrt{EG} = \frac{\sqrt{2}}{v^2} \text{ and } \frac{1}{\sqrt{EG}} = \frac{v^2}{\sqrt{2}}$$

$$K = \frac{-v^2}{\sqrt{2}}\left(\frac{\partial}{\partial v}\left(\frac{v}{\sqrt{2}}\frac{\partial}{\partial v}\frac{1}{v}\right)\right)$$

$$= \frac{v^2}{\sqrt{2}}\frac{\partial}{\partial v}\left(\frac{1}{v\sqrt{2}}\right) = \frac{v^2}{2}\frac{-1}{v^2} = -\frac{1}{2}$$

(2) We use the formula (C) to compute the Gaussian curvature of this distribution. We take the left hand side of formula (C) which has the form

$$K = \frac{1}{D}\left[\frac{\partial}{\partial v}\left(\frac{D}{E}\Gamma_{11}^{2}\right) - \frac{\partial}{\partial u}\left(\frac{D}{E}\Gamma_{12}^{2}\right)\right]$$

where $D^2 = EG - F^2$, $D = \sqrt{EG - F^2}$

We can find that

$$\Gamma_{12}^2 = 0$$

$$\Gamma_{11}^2 = \frac{-E_v}{2G} = \frac{2}{v^3} . \frac{v^2}{4} = \frac{1}{2v}$$

$$K = \frac{v^2}{\sqrt{2}} \left[\frac{\partial}{\partial v} \left(\frac{\sqrt{2}}{v^2} . v^2 . \frac{1}{2v} \right) \right]$$

$$= \frac{v^2}{\sqrt{2}} \left[\frac{\partial}{\partial v} \left(\frac{1}{\sqrt{2}v} \right) \right] = \frac{v^2}{2} \left(\frac{-1}{v^2} \right) = \frac{-1}{2}$$

Now, we take the right hand side of formula (C) which has the form

$$K = \frac{1}{D} \left[\frac{\partial}{\partial u} \left(\frac{D}{G} \Gamma_{22}^1 \right) - \frac{\partial}{\partial v} \left(\frac{D}{G} \Gamma_{12}^1 \right) \right]$$

We can find that

$$\Gamma_{22}^1 = 0$$

$$\Gamma_{12}^1 = \frac{E_v}{2E} = \frac{-2}{v^3} . \frac{v^2}{2} = \frac{-1}{v}$$

$$K = \frac{v^2}{\sqrt{2}} \left[\frac{\partial}{\partial v} \left(\frac{1}{\sqrt{2}v} \right) \right] = \frac{v^2}{2} \left(\frac{-1}{v^2} \right) = -\frac{1}{2}$$

4.2 Cauchy Distribution

(1) We use the formula (A) to compute the Gaussian curvature of the Cauchy distribution.

We can find that

$$EG = \frac{1}{4v^4}, \ \sqrt{EG} = \frac{1}{2v^2} \text{ and } \frac{1}{\sqrt{EG}} = 2v^2$$

$$K = -2v^2 \left(\frac{\partial}{\partial v} \left(\sqrt{2}v \frac{-1}{\sqrt{2}v^2} \right) \right)$$

$$= 2v^2 \left(\frac{-1}{v^2} \right) = -2$$

(2) We use the formula (B) to compute the Gaussian curvature of this distribution.

$$K=\frac{-1}{2\sqrt{EG-F^2}}\left[\frac{\partial}{\partial u}\frac{G_u-F_v}{\sqrt{EG-F^2}}-\frac{\partial}{\partial v}\frac{F_u-E_v}{\sqrt{EG-F^2}}\right]-\frac{1}{4\left(EG-F^2\right)^2}\begin{vmatrix}E & F & G\\ E_u & F_u & G_u\\ E_v & F_v & G_v\end{vmatrix}$$

We can find that

$$E_v=\frac{-1}{v^3},\ G_u=0,\ F=F_u=F_v=0$$

Then the determinant equal to 0 for F=0

$$K=\frac{-1}{2\sqrt{EG}}\left[\frac{\partial}{\partial v}\frac{E_v}{\sqrt{EG}}\right]$$

$$=-v^2\left(\frac{\partial}{\partial v}\frac{-2}{v}\right)=2v^2\frac{-1}{v^2}=-2$$

(3) We use the formula (C) to compute the Gaussian curvature of this distribution. We can find that

$$\Gamma_{11}^1=\Gamma_{12}^2=\Gamma_{22}^1=0$$

$$\Gamma_{21}^1=\frac{E_v}{2E}=\frac{-1}{v^3}.v^2=\frac{-1}{v}=\Gamma_{12}^1$$

$$\Gamma_{11}^2=\frac{-E_v}{2G}=\frac{1}{v}$$

$$\Gamma_{22}^2=\frac{G_v}{2G}=\frac{-1}{v^3}.v^2=\frac{-1}{v}$$

We take the left hand side of formula (C)

$$K=\frac{1}{\sqrt{EG}}\left[\frac{\partial}{\partial v}\left(\frac{\sqrt{EG}}{E}\Gamma_{11}^2\right)\right]$$

$$=2v^2\left[\frac{\partial}{\partial v}\frac{1}{v}\right]=2v^2\frac{-1}{v^2}=-2$$

Now, we take the right hand side of formula (C)

$$K = \frac{1}{\sqrt{EG}}\left[-\frac{\partial}{\partial v}\left(\frac{\sqrt{EG}}{G}\Gamma_{12}^{1}\right)\right]$$

$$= 2v^2\left(\frac{-1}{v^2}\right) = -2$$

This result shows that Cauchy distribution does not converge to the normal distribution.

4.3 t Distribution

(1) we use the formula (A) to compute the Gaussian curvature of the *t* distribution. We can find that

$$EG = \frac{2r(r+1)}{v^4(r+3)^2}, \qquad \sqrt{EG} = \frac{\sqrt{2r(r+1)}}{v^2(r+3)},$$

$$\frac{1}{\sqrt{EG}} = \frac{v^2(r+3)}{\sqrt{2r(r+1)}}$$

$$K = -\frac{v^2(r+3)}{\sqrt{2r(r+1)}}\left(\frac{\partial}{\partial v}\left(\frac{1}{\sqrt{-\frac{1}{v^2}\left(1-3\frac{r+1}{r+3}\right)}}\frac{\partial}{\partial v}\left(\frac{\sqrt{r+1}}{v\sqrt{r+3}}\right)\right)\right)$$

$$= -\frac{v^2(r+3)}{\sqrt{2r(r+1)}}\frac{\partial}{\partial v}\left(\frac{-\sqrt{r+1}}{v\sqrt{r+3}}.\frac{1}{\sqrt{-\left(1-3\frac{r+1}{r+3}\right)}}\right)$$

$$= \frac{-(r+3)}{2r}.$$

(2) we use the formula (B) to compute the Gaussian curvature of this distribution. We can find that

$$E_u = 0 \qquad , E_v = \frac{-2(r+1)}{(r+3)v^3}$$

$$G_u = 0 \qquad , \; G_v = \frac{2}{v^3}\left(1 - 3\frac{r+1}{r+3}\right)$$

For F=0 we can say that

$$K = -\frac{v^2(r+3)}{2\sqrt{2r}\sqrt{r+1}}\frac{\partial}{\partial v}\left(-\frac{\sqrt{2(r+1)}}{v}\frac{1}{\sqrt{r}}\right)$$

$$= -\frac{r+3}{2r}$$

(3) we use the formula (C) to compute the Gaussian curvature of this distribution. We can find

$$\Gamma_{11}^1 = \Gamma_{12}^2 = \Gamma_{22}^1 = 0$$

$$\Gamma_{11}^2 = \frac{r+1}{2rv} \qquad \Gamma_{12}^1 = -\frac{1}{v} \qquad \Gamma_{22}^2 = -\frac{1}{v}$$

We take the left hand side of formula (C)

$$K = \frac{1}{\sqrt{EG}}\left[\frac{\partial}{\partial v}\left(\frac{\sqrt{EG}}{E}\Gamma_{11}^2\right)\right]$$

$$= \frac{v^2(r+3)}{\sqrt{2r(r+1)}}\left(\frac{\partial}{\partial v}\frac{\sqrt{2r(r+1)}}{v^2(r+3)}.\frac{v^2(r+3)}{r+1}.\frac{r+1}{2rv}\right)$$

$$= \frac{v^2(r+3)}{2r}.\frac{-1}{v^2} = \frac{-(r+3)}{2r}.$$

Now, we take the right hand side of formula (C)

$$K = \frac{1}{\sqrt{EG}}\left[-\frac{\partial}{\partial v}\left(\frac{\sqrt{EG}}{G}\Gamma_{12}^1\right)\right]$$

$$= \frac{v^2(r+3)}{\sqrt{2r(r+1)}}\left[-\frac{\partial}{\partial v}\frac{\sqrt{2r(r+1)}}{v^2(r+3)}.\frac{v^2}{-\left(1-3\frac{r+1}{r+3}\right)}.\frac{-1}{v}\right]$$

$$= \frac{-(r+3)}{2r}$$

When r is very large the result of K converges to $-\frac{1}{2}$ that means t distribution converges to the normal distribution.

4.4 Gamma Distribution

(1) we use the formula (A) to compute the Gaussian curvature of Gamma distribution.

$$K = -\frac{1}{\sqrt{EG}}\left(\frac{\partial}{\partial \sim}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial \sim}\right) + \frac{\partial}{\partial \text{¦}}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial \text{¦}}\right)\right)$$

We can find

$$E = \Psi'(\sim) - \frac{1}{\sim} \qquad , \; F = 0 \qquad , \; G = \sim$$

$$EG = \sim\left(\Psi'(\sim) - \frac{1}{\sim}\right) = \sim\Psi'(\sim) - 1$$

$$\sqrt{EG} = \sqrt{\sim\Psi'(\sim) - 1}$$

$$K = -\frac{1}{\sqrt{\sim\Psi'(\sim)-1}}\left(\frac{\partial}{\partial\sim}\left(\frac{1}{\sqrt{\Psi'(\sim) - \frac{1}{\sim}}}\frac{\partial}{\partial\sim}\sqrt{\sim}\right)\right)$$

$$= \frac{\left(\Psi'(\sim) - \frac{1}{\sim}\right) + \sim\left(\Psi''(\sim) + \frac{1}{\sim^2}\right)}{4\left(\sim\Psi'(\sim) - 1\right)^2}$$

$$= \frac{\Psi'(\sim) + \sim\Psi''(\sim)}{4\left(\sim\Psi'(\sim) - 1\right)^2}$$

where $\Psi(\sim) = \dfrac{\Gamma'(\sim)}{\Gamma(\sim)}$ is the digamma function

$$\lim_{\sim\to\infty} K(\sim) = \lim_{\sim\to\infty}\frac{-\left(1/2\sim^2 + 0\left(1/\sim^3\right)\right)}{4\left(\left(1/4\sim^2\right) + 0\left(1/\sim^3\right)\right)} = -\frac{1}{2}$$

$$\lim_{\sim\to 0} K(\sim) = -\frac{1}{4}$$

(2) we use the formula (C) to compute the Gaussian curvature of this distribution.

We can find

$$\Gamma_{11}^{1}=\frac{E_{\sim}}{2E}=\frac{\sim^{2}\Psi''(\sim)+1}{2\sim\left(\sim\Psi'(\sim)-1\right)}$$

$$\Gamma_{12}^{2}=\frac{G_{\sim}}{2G}=\frac{1}{2\sim}$$

$$\Gamma_{22}^{1}=\frac{-G_{\sim}}{2G}=\frac{-1}{2\left(\Psi'(\sim)-\frac{1}{\sim}\right)}$$

$$\Gamma_{11}^{2}=\Gamma_{12}^{1}=\Gamma_{22}^{2}=0$$

We take the left hand side of formula (C)

$$K=\frac{1}{D}\left[\frac{\partial}{\partial \text{¦}}\left(\frac{D}{E}\Gamma_{11}^{2}\right)-\frac{\partial}{\partial\sim}\left(\frac{D}{E}\Gamma_{12}^{2}\right)\right]$$

$$=\frac{1}{\sqrt{EG}}\left[-\frac{\partial}{\partial\sim}\left(\frac{\sqrt{EG}}{\Psi'(\sim)-\frac{1}{\sim}}\frac{1}{2\sim}\right)\right]$$

$$=\frac{-1}{2\sqrt{\sim\Psi'(\sim)-1}}\left[\frac{\partial}{\partial\sim}\frac{\sqrt{\sim\Psi'(\sim)-1}}{\sim\Psi'(\sim)-1}\right]$$

$$=\frac{\sim\Psi''(\sim)+\Psi'(\sim)}{4\left(\sim\Psi'(\sim)-1\right)^{2}}$$

$$\lim_{\sim\to\infty}K(\sim)=-\frac{1}{2}$$

$$\lim_{\sim\to 0}K(\sim)=-\frac{1}{4}$$

Now, we take the right hand side of formula (C)

$$K=\frac{1}{D}\left[\frac{\partial}{\partial\sim}\left(\frac{D}{G}\Gamma_{22}^{1}\right)-\frac{\partial}{\partial \text{¦}}\left(\frac{D}{G}\Gamma_{12}^{1}\right)\right]$$

$$= \frac{1}{\sqrt{EG}}\left(\frac{\partial}{\partial\sim}\left(\frac{\sqrt{EG}}{G}\Gamma_{22}^{1}\right)\right)$$

$$= \frac{-1}{2\sqrt{\sim\Psi'(\sim)-1}}\left(\frac{\partial}{\partial\sim}\frac{1}{\sqrt{\sim\Psi'(\sim)-1}}\right)$$

$$= \frac{\sim\Psi''(\sim)+\Psi'(\sim)}{4\left(\sim\Psi'(\sim)-1\right)^{2}}$$

$$\lim_{\sim\to\infty} K(\sim) = -\frac{1}{2} \quad , \quad \lim_{\sim\to 0} K(\sim) = -\frac{1}{4}$$

(3) we use the formula (D) to compute the Gaussian curvature of this distribution.

$$K = \frac{R_{1212}}{EG-F^{2}} = \frac{(12{,}12)}{EG-F^{2}}$$

where $(12{,}12) = R_{1212} = \sum_{m=1}^{2} R_{121}^{m} g_{m2}$

$$R_{121}^{1} = \frac{\partial}{\partial\text{¦}}\Gamma_{11}^{1} - \frac{\partial}{\partial\sim}\Gamma_{21}^{1} + \Gamma_{11}^{m}\Gamma_{m2}^{1} - \Gamma_{21}^{m}\Gamma_{m1}^{1} = 0$$

$$R_{121}^{2} = \frac{\partial}{\partial\text{¦}}\Gamma_{11}^{2} - \frac{\partial}{\partial\sim}\Gamma_{21}^{2} + \Gamma_{11}^{m}\Gamma_{m2}^{2} - \Gamma_{21}^{m}\Gamma_{m1}^{2}$$

$$= -\frac{\partial}{\partial\sim}\Gamma_{21}^{2} + \Gamma_{11}^{1}\Gamma_{12}^{2} - \Gamma_{21}^{2}\Gamma_{21}^{2}$$

$$= \frac{1}{2\sim^{2}} + \frac{\sim^{2}\Psi''(\sim)+1}{4\sim^{2}\left(\sim\Psi'(\sim)-1\right)} - \frac{1}{4\sim^{2}}$$

$$= \frac{\Psi'(\sim)+\sim\Psi''(\sim)}{4\sim^{2}\left(\sim\Psi'(\sim)-1\right)}$$

$$K = \frac{R_{121}^{2}}{E} = \frac{\Psi'(\sim)+\sim\Psi''(\sim)}{4\left(\sim\Psi'(\sim)-1\right)^{2}}$$

$$\lim_{\sim\to\infty} K(\sim) = -\frac{1}{2} \quad , \quad \lim_{\sim\to 0} K(\sim) = -\frac{1}{4}$$

This result shows that Gamma distribution converges to the normal distribution.

4.5 Pareto Distribution

(1) we use the formula (B) to compute the Gaussian curvature of this distribution.

$$K=-\frac{1}{2\sqrt{EG-F^2}}\left[\frac{\partial}{\partial\sim}\frac{G_{\sim}-F}{\sqrt{EG-F^2}}-\frac{\partial}{\partial}\frac{F_{\sim}-E}{\sqrt{EG-F^2}}\right]-\frac{1}{4\left(EG-F^2\right)^2}\begin{vmatrix}E & F & G\\ E_{\sim} & F_{\sim} & G_{\sim}\\ E & F & G\end{vmatrix}$$

where

$$\begin{vmatrix}E & F & G\\ E_{\sim} & F_{\sim} & G_{\sim}\\ E & F & G\end{vmatrix}=\begin{vmatrix}\frac{}{\sim^2} & -\frac{1}{\sim} & \frac{1}{\ \ ^2}\\ -\frac{2}{\sim^3} & \frac{1}{\sim^2} & 0\\ \frac{1}{\sim^2} & 0 & -\frac{2}{\ \ ^3}\end{vmatrix}=\frac{1}{\sim^4\ \ ^2}$$

Now, we can say that

$$K=-\frac{1}{4\left(EG-F^2\right)^2}\begin{vmatrix}E & F & G\\ E_{\sim} & F_{\sim} & G_{\sim}\\ E & F & G\end{vmatrix}$$

$$=-\frac{1}{4(1-\ \ )^2}$$

$$\lim_{\to 0}K(\ \ )=-\frac{1}{4} \quad , \quad \lim_{\to\infty}K(\ \ )=0$$

(2) we use the formula (C) to compute the Gaussian curvature of this distribution. We can find

$$E_{\sim}=-\frac{2}{\sim^3} \quad , E\ \ =\frac{1}{\sim^2} \quad , F_{\sim}=\frac{1}{\sim^2} \quad ,F\ \ =0$$

$$G_{\sim}=0 \quad , G\ \ =\frac{-2}{\ \ ^3}$$

$$\Gamma_{11}^1=-\frac{-2+}{2\sim(1-\ \ )} \quad , \Gamma_{12}^2=\frac{}{2\sim(1-\ \ )}$$

$$\Gamma_{11}^{2} = \frac{- \quad ^{2}}{2\sim^{2}(1- \quad)} \quad , \quad \Gamma_{22}^{1} = -\frac{\sim}{^{2}(1- \quad)}$$

$$\Gamma_{12}^{1} = \frac{1}{2 \quad (1- \quad)} \quad , \quad \Gamma_{22}^{2} = -\frac{1}{(1- \quad)}$$

We take the left hand side of formula (C)

$$K = \frac{1}{D}\left[\frac{\partial}{\partial}\left(\frac{D}{E}\Gamma_{11}^{2}\right) - \frac{\partial}{\partial\sim}\left(\frac{D}{E}\Gamma_{12}^{2}\right)\right]$$

$$= \frac{1}{\sqrt{EG-F^{2}}}\left[\frac{\partial}{\partial}\left(\frac{\sqrt{EG-F^{2}}}{E}\Gamma_{11}^{2}\right)\right]$$

$$= \frac{\sqrt{\quad}\sim}{\sqrt{1-}}\left[\frac{\partial}{\partial}\frac{-\sqrt{\quad}}{2\sim\sqrt{1-}}\right] = -\frac{1}{4(1- \quad)^{2}}$$

$$\lim_{\to 0} K(\quad) = -\frac{1}{4} \quad , \quad \lim_{\to\infty} K(\quad) = 0$$

Now, we take the right hand side of formula (C)

$$K = \frac{1}{D}\left[\frac{\partial}{\partial\sim}\left(\frac{D}{G}\Gamma_{22}^{1}\right) - \frac{\partial}{\partial}\left(\frac{D}{G}\Gamma_{12}^{1}\right)\right]$$

$$= \frac{1}{\sqrt{EG-F^{2}}}\left[-\frac{\partial}{\partial}\left(\frac{\sqrt{EG-F^{2}}}{G}\Gamma_{12}^{1}\right)\right]$$

$$= \frac{-\sqrt{\quad}\sim}{\sqrt{1-}}\left[\frac{\partial}{\partial}\frac{\sqrt{\quad}}{2\sim\sqrt{1-}}\right] = -\frac{1}{4(1- \quad)^{2}}$$

$$\lim_{\to 0} K(\quad) = -\frac{1}{4} \quad , \quad \lim_{\to\infty} K(\quad) = 0$$

(3) we use the formula (D) to compute the Gaussian curvature of this distribution.

$$K = \frac{R_{1212}}{EG-F^{2}}$$

$$R_{1212} = R_{121}^{1} g_{12} + R_{121}^{2} g_{22}$$

where

$$R^{1}_{121} = \frac{\partial}{\partial}\Gamma^{1}_{11} - \frac{\partial}{\partial\sim}\Gamma^{1}_{21} + \Gamma^{m}_{11}\Gamma^{1}_{m2} - \Gamma^{m}_{21}\Gamma^{1}_{m1}$$

$$= -\Gamma^{2}_{21}\Gamma^{1}_{21} = -\frac{1}{4\sim(1- \quad)^2}$$

$$R^{1}_{121}g_{12} = \frac{1}{4\sim^{2}(1- \quad)^2}$$

$$R^{2}_{121} = \frac{\partial}{\partial}\Gamma^{2}_{11} - \frac{\partial}{\partial\sim}\Gamma^{2}_{21} + \Gamma^{m}_{11}\Gamma^{2}_{m2} - \Gamma^{m}_{21}\Gamma^{2}_{m1}$$

$$= -\frac{2 \quad - \quad ^{2}}{2\sim^{2}(1- \quad)^2} + \frac{}{2\sim^{2}(1- \quad)} + \frac{-2 \quad + \quad ^{2}}{4\sim^{2}(1- \quad)^2}$$

$$+ \frac{}{2\sim^{2}(1- \quad)^2} + \frac{}{4\sim^{2}(1- \quad)^2} - \frac{^{2}}{4\sim^{2}(1- \quad)^2}$$

$$= \frac{-4 \quad + 3}{4\sim^{2}(1- \quad)^2} = -\frac{}{4\sim^{2}(1- \quad)^2}$$

$$R^{2}_{121}g_{22} = -\frac{1}{4 \quad \sim^{2}(1- \quad)^2}$$

$$R_{1212} = \frac{-1}{4 \quad \sim^{2}(1- \quad)^2}$$

$$K(\quad) = -\frac{1}{4(1- \quad)^2}$$

$$\lim_{\to 0} K(\quad) = -\frac{1}{4} \qquad , \qquad \lim_{\to\infty} K(\quad) = 0$$

This result shows that Pareto distribution does not converge to the normal distribution. But when the value of β lies between 0.29 and 0.3, $K(\quad) \to -\frac{1}{2}$.

4.6 Wald Distribution

(1) we use the formula (A) to compute the Gaussian curvature of this distribution. We know that

$$E=\frac{a}{«^{3}} \quad , F=0 \quad , G=\frac{1}{2a^{2}}$$

$$\sqrt{EG}=\frac{1}{\sqrt{2«^{3}a}}$$

Then

$$K=-\frac{1}{\sqrt{EG}}\left(\frac{\partial}{\partial «}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial «}\right)+\frac{\partial}{\partial a}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial a}\right)\right)$$

$$=-\sqrt{2«^{3}a}\,\frac{1}{2\sqrt{2a«^{3}}}=-\frac{1}{2}$$

(2) we use the formula (B) to compute the Gaussian curvature of this distribution.

We can find

$$E_{«}=\frac{-3a}{«^{4}} \quad , E_{a}=\frac{1}{«^{3}} \quad , F_{«}=0 \quad , F_{a}=0$$

$$G_{«}=0 \quad , G_{a}=-\frac{1}{a^{3}}$$

Then

$$K=-\frac{1}{2\sqrt{EG}}\left[\frac{\partial}{\partial «}\frac{G_{«}}{\sqrt{EG}}-\frac{\partial}{\partial a}\frac{-E_{a}}{\sqrt{EG}}\right]-\frac{1}{4(EG)^{2}}\begin{vmatrix}E & 0 & G\\ E_{«} & 0 & G_{«}\\ E_{a} & 0 & G_{a}\end{vmatrix}$$

$$=-\sqrt{a}\left(\frac{\partial}{\partial a}\sqrt{a}\right)=-\sqrt{a}\left(\frac{1}{2}\frac{1}{\sqrt{a}}\right)=-\frac{1}{2}$$

(3) we use the formula (C) to compute the Gaussian curvature of this distribution. We can find

$$\Gamma_{11}^{1}=\frac{E_{«}}{2E}=\frac{-3}{2«}$$

$$\Gamma_{11}^{2} = \frac{-E_{\text{a}}}{2G} = -\frac{\text{a}^{2}}{\text{«}^{3}}$$

$$\Gamma_{12}^{1} = \frac{E_{\text{a}}}{2E} = \frac{1}{2\text{a}}$$

$$\Gamma_{22}^{2} = \frac{G_{\text{a}}}{2G} = -\frac{1}{\text{a}}$$

$$\Gamma_{12}^{2} = \Gamma_{22}^{1} = 0$$

We take the left hand side of formula (C)

$$K = \frac{1}{D}\left[\frac{\partial}{\partial \text{a}}\left(\frac{D}{E}\Gamma_{11}^{2}\right) - \frac{\partial}{\partial \text{«}}\left(\frac{D}{E}\Gamma_{12}^{2}\right)\right]$$

$$= \sqrt{2\text{«}^{3}\text{a}}\left[\frac{\partial}{\partial \text{a}}\left(\frac{1}{\sqrt{2\text{«}^{3}\text{a}}}.\frac{\text{«}^{3}}{\text{a}}.\frac{-\text{a}^{2}}{\text{«}^{3}}\right)\right]$$

$$= -\sqrt{\text{a}}\left(\frac{\partial}{\partial \text{a}}\sqrt{\text{a}}\right) = -\sqrt{\text{a}}\left(\frac{1}{2\sqrt{\text{a}}}\right) = -\frac{1}{2}$$

We take the right hand side of formula (C)

$$K = \frac{1}{D}\left[\frac{\partial}{\partial \text{«}}\left(\frac{D}{G}\Gamma_{22}^{1}\right) - \frac{\partial}{\partial \text{a}}\left(\frac{D}{G}\Gamma_{12}^{1}\right)\right]$$

$$= \sqrt{2\text{«}^{3}\text{a}}\left[-\frac{\partial}{\partial \text{a}}\left(\frac{1}{\sqrt{2\text{«}^{3}\text{a}}}.2\text{a}^{2}.\frac{1}{2\text{a}}\right)\right]$$

$$= -\sqrt{\text{a}}\left(\frac{\partial}{\partial \text{a}}\sqrt{\text{a}}\right) = -\frac{1}{2}$$

(4) we use the formula (D) to compute the Gaussian curvature of this distribution.

$$K = \frac{R_{1212}}{EG - F^{2}}$$

$$R_{1212} = R_{121}^{1} g_{12} + R_{121}^{2} g_{22}$$

where

$$R_{121}^{1} = \frac{\partial}{\partial \text{a}}\Gamma_{11}^{1} - \frac{\partial}{\partial «}\Gamma_{21}^{1} + \Gamma_{11}^{m}\Gamma_{m2}^{1} - \Gamma_{21}^{m}\Gamma_{m1}^{1} = 0$$

$$R_{121}^{2} = \frac{\partial}{\partial \text{a}}\Gamma_{11}^{2} - \frac{\partial}{\partial «}\Gamma_{21}^{2} + \Gamma_{11}^{m}\Gamma_{m2}^{2} - \Gamma_{21}^{m}\Gamma_{m1}^{2}$$

$$= \frac{-2\text{a}}{«^{3}} + \frac{\text{a}}{«^{3}} + \frac{\text{a}}{2«^{3}} = \frac{-\text{a}}{2«^{3}}.$$

$$R_{1212} = R_{121}^{2} g_{22} = -\frac{1}{4«^{3}\text{a}}$$

$$K = -\frac{1}{2}$$

This result shows that Wald distribution converges to the normal distribution.

REFERENCES

[1] Barndorff- Nielsen, O.E., Cox, D. R. and Reid, N., "The Role of Differential Geometry in Statistical Theory", International Statistical Review, 54, 1, PP. 83-96, 1986.

[2] Kass, R. E. and Vos, P.W., "Geometrical Foundations of Asymptotic Inference", John Wiley & Sons, Inc, 1997.

[3] Chen, W. W. S., "On Computing Gaussian Curvature of Some Well Known Distributions", 1999.
http://www.irs.gov/pub/irs-soi/gauscurv.pdf

[4] Gruber, M. H. J., "Some Interesting Connections between Statistics and Differential Geometry", 2003.
http://www.rit.edu/~mjgsma/maatalk/geometrystatistics.html

[5] Arwini, K., Del Riego, L. and Dodson, C. T. J., "Universal Connection and Curvature for Statistical Manifold Geometry", June, 2005.
http://www.ma.umist.ac.uk/kd/PREPRINTS/univconexpf.pdf

[6] Weisstein, E. W., "Tensor-from MathWorld", Wolfram Research, Inc., 2005.
http://mathworld.wolfram.com/Tensor.html

[7] Weisstein, E. W., "Symmetric Tensor-from MathWorld", Wolfram Research, Inc., 2005.
http://mathworld.wolfram.com/SymmetricTensor.html

[8] Weisstein, E.W., "Antisymmetric Tensor- from MathWorld", Wolfram Research, Inc., 2005.

http://mathworld.wolfram.com/AntisymmetricTensor.html

[9] Do Carmo, M. P., "Differential Geometry of Curves and Surfaces", Prentice-Hall, Inc., New Jersey, 1976.

[10] "Chapter 14 Basics of the Differential Geometry of Surfaces", Nov. 2005.
http://www.seas.upenn.edu/~cis70005/cis700s|6pdf.pdf

[11] Gardner, B., "Differential Geometry (and Relativity) Classnotes", summer 2000.
http://www.etsu.edu/math/gardner/5310/notes.htm-lok-

[12] Fajstrup, L. and Raussen, M., "Visual Interactive Differential Geometry", August, 2004.
http://www.math.aau.dk/~fajstrup/UNDERVISNING/MAT3/GEOMETRI/NOTER/fundamentalformer.pdf

[13] O' Neill, B., "Elementary Differential Geometry", Academic Press, Inc., Orlando, 1966.

[14] Minru, Ph.D. "Differential Geometry Review: Part Two",
http://math.uh.edu/~minru/4350/tworeview.pdf

[15] Royster, D., "Chapter 5 Geodesics, Metrics, and Isometries", 1998.
http://www.math.uncc.edu/~droyster/courses/fall98/math4080/classnotes/geodesic.pdf

[16] Alekseevskij, D.V., Lychagin, V.V. and Vinogradov, "Geometry I", Encyclopedia of Mathematical Sciences, Vol.28, Springer-Verlag, Berlin, 1991.

[17] Fahady, K. S. and Shamoon, P. J., "Probability", Ministry of Higher Education and Scientific Research, 1990.

[18] Easton, V. J. and McColl, J.H., "Statistics Glossary-Random Variables and Probability Distributions", Steps. http://www.stats.gla.ac.uk/steps/glossary/probability_distributions.html-25k-

[19] Croarkin, C. and Tobias, P., "1.3.6. Probability Distributions", 2005. http://www.itl.nist.gov/div898/handbook/eda/section3/eda36.htm-5k-

[20] Gohari, P. "Chapter 3: Random Variables and Probability Distributions", Fall 2002. http://www.ece.concordia.ca/~gohari/engr371/3.pdf

[21] "Chapter 4. Continuous Random Variables and Probability Distributions", 2001. http://www.am.qub.ac.uk/users/j.condon/SOR101/sor101_lectures_2up_issue_chapter4.pdf

[22] Mood, A. M., Graybill, F. A. and Boes, D.C., "Introduction to the Theory of Statistics", 3d ed., McGraw- Hill, Inc., New York, 1974.

[23] Jensen, P.A. and Bard, J.F., "Probability Models. S3 Continuous Random Variables", 2001. http://www.me.utexas.edu/~jensen/ORMM/supplements/models/probmodel/continuous.pdf

[24] "Basics: Working with Arbitrary Distributions". http://www.addlink.es/pdf/AGDWeb250.pdf

[25] Wikipedia, The Free Encyclopedia, "Fisher Information", 2005.

http://en.wikipedia.org/wiki/Fisher_information.

[26] Hamza, A. and Hussein, D. "Statistical Inference", Ministry of Higher Education and Scientific Research, 1988.

[27] Wikipedia, The Free Encyclopedia, "Convergence of Random Variables", 2005.
http://en.wikipedia.org/wiki/convergence_of_random_variables

[28] Wikipedia, The Free Encyclopedia, "Law of Large Numbers", 2005.
http://en.wikipedia.org/wiki/Law_of_large_numbers

[29] "Central Limit Theorem in Statistics"
http://www.stattucino.com/berrie/clt.html-9k-

[30] Wikipedia, The Free Encyclopedia, "Central Limit Theorem", 2005.
http://en.wikipedia.org/wiki/central_limit_theorem.

[31] Wikipedia, The Free Encyclopedia, "Fisher Information Metric", 2005.
http://en.wikipedia.org/wiki/Fisher_information_metric-12k-

[32] Gruber, M.H.J., "Using Differential Geometry to Study Robustness in Bayesian Statistics for Shrinkage Estimators",2003.
http://www.rit.edu/~mjgsma/slides.pdf

Abstract

The using of mathematics to get a solution for many problems, is known from the past. According to truth we study the using of differential geometry to show which continuous distribution converges to normal distribution by connecting between differential geometry and statistics.

In particular, we illustrate the connection between Pareto distribution and Wald distribution with normal distribution by using some statistical theorems and differential geometry. The thesis consists of five chapters:

The first chapter introduces some definitions and concepts from differential geometry like tensor, the first and second fundamental form, Gaussian curvature, etc. From these concepts we get different formulas to calculate the value of Gaussian curvature based on the relation between Riemannian metrics and Fisher information.

The second chapter explains some definitions and concepts from probability and statistics needed in later chapters such as probability density function, continuous distribution function, some special continuous distributions, Fisher information, convergence of random variable, the law of large numbers and the central limit theorem.

The third chapter presents some connections between statistics and differential geometry, such as the definition of the coefficients of the expected Fisher information matrix as they equal to the coefficients of the first fundamental form (Riemannian metrics) given by:

$$g_{ij} = -\int \frac{\partial^2 \log f(x,\overline{\ })}{\partial\overline{\ }_i \partial\overline{\ }_j} f(x,\overline{\ })dx = -E\left[\frac{\partial^2 \log f(x,\overline{\ })}{\partial\overline{\ }_i \partial\overline{\ }_j}\right],$$

the relation between the Riemannian metrics and geodesics, Fisher information Rao distances between probability distributions, Riemannian metrics for some distributions, the Gaussian curvature of the probability

distributions, and the Christoffel symbols. Some examples are also given to compute the Gaussian curvature using different formulas.

In chapter four we use some methods to calculate the Gaussian curvature for some distributions. Also, we apply these methods to calculate the Gaussian curvature for Pareto distribution and Wald distribution.

Chapter five contains some conclusions and recommendations.

Acknowledgement

Praise is to Allah for helping me to complete this thesis.

I would like to express my gratitude to my supervisor Assist. Prof. Hussein Ahmed Ali for his advice, guidance and encouragement throughout the preparation of this work.

Also, I'm indebted to the dean of the college of education, Babylon university, and to all the staff of the department of Mathematics, especially the head of the department of Mathematics, for their cooperation.

Also I thank all those helped me very much in the fulfillment of this work.

My thanks are due to my family for their spiritual support during the period of writing this thesis.

5.1 Conclusions

(1) The gamma distribution and t distribution converge to normal distribution by using four formulas.

(2) The Cauchy distribution does not converge to normal distribution by using four formulas.

(3) By using three formulas, we find the Gaussian curvature for Pareto distribution equal to $\left(-\frac{1}{4}\right)$ when $\rightarrow 0$ and equal to (o) when $\rightarrow \infty$ that means this distribution does not converge with normal distribution except in the interval (0.29, 0.3).

(4) By using four formulas, we find the Gaussian curvature for Wald distribution equal to $\left(-\frac{1}{2}\right)$ that means this distribution converges to normal distribution.

(5) Implicit conclusion that whether one uses statistical or geometric mean curvature, the t_3 may be considered half way in between a normal and Cauchy distribution.

(6) The Gaussian curvature for the distributions (normal, gamma, Wald and t_3) lies between the Gaussian curvature of Pareto distribution and Cauchy distribution.

5.2 Recommendations

(1) Study which discrete distribution will have the convergence with normal distribution by using advanced mathematical formulas need to studies and researches in future with some conditions.

(2) Investigate the convergence of other continuous distributions using the different methods.

(3) The distributions of multivariate need to study by using the same formulas in future.

Examining Committee Certification

We certify that we have read this thesis entitled "**Application of Central Limit Theorem on Probability Distributions by Using Differential Geometry**" and, as an Examining Committee, we examined the student in its content, and what is related to it, and that in our opinion it is adequate with standing as a thesis for the degree of Master of Science in Mathematics.

Chairman
Signature:
Name:
Title:
Date: / / 2006

Member	**Member**
Signature:	**Signature:**
Name:	**Name:**
Title:	**Title:**
Date: / / 2006	**Date: / / 2006**

Member (Supervisor)
Signature:
Name:
Title:
Date: / / 2006

Approved by the Dean of the College of Education

Signature:
Name:
Title:
Date: / / 2006

Introduction

The studies of applying the differential geometry in statistical subject are very little and do not cover each the sides, also the references under this title are limited. The knowledge of student about the differential geometry is very important for the M. Sc. study.

This thesis is divided into five chapters:

Chapter one contains some important concepts of differential geometry, like tensor, Riemannian metric, second fundamental form, Gaussian curvature, geodesics and curvature tensor.

Chapter two presents some important concepts and theorems of mathematical statistics, such as continuous random variable, some continuous distributions and Fisher information. Also, we show the theorems of law of the large numbers and central limit theorem.

Chapter three gives some interesting connections between statistics and differential geometry.

Chapter four contains the results of computing the Gaussian curvature (K) of some continuous distributions, like normal, Cauchy, t, gamma, Pareto and wald.

Chapter five contains some conclusions and the recommendations.

There are some researchers who worked in this field in end of twenty century and beginning of twenty one century:

In 1986, Barndorff- Nielsen, O. E.[1] used the concept that the coefficients of the expected Fisher information matrix as equal to the coefficients of the first fundamental form.

In 1997, Kass, R.E. [2] used the same concept of Barndorff- Nielsen, O.E. using the following formula

$$-\frac{1}{\sqrt{EG}}\left(\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\right)+\frac{\partial}{\partial v}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\right)\right)$$

to compute the Gaussian curvature (K) of trinomial and t families. He gave the general form of a location-scale manifold of density:

$$\left\{p(x)=\frac{1}{v}f\left(\frac{x-u}{v}\right)\middle|(u,v)\in R\times R_{+}\right\}$$

For some density function *f*.

In 1999, Chen W.W.S. [3] provided a deeper and broader understanding of the meaning of Gaussian curvature, using some more general alternative computational methods. He used the formula

$$\frac{R_{1212}}{EG-F^2}=\frac{(12,12)}{EG-F^2}\quad\text{where } (12,12)=R_{1212}=\sum_{m=1}^{2}R_{121}^{m}g_{m2}$$

$$R_{ijk}^{1}=\frac{\partial}{\partial u_j}\Gamma_{ik}^{1}-\frac{\partial}{\partial u_i}\Gamma_{jk}^{1}+\Gamma_{ik}^{m}\Gamma_{mj}^{1}-\Gamma_{jk}^{m}\Gamma_{mi}^{1}\text{, sum on } m,$$

to compute the Gaussian curvature (K) for the distributions (normal, Cauchy and t family). He showed that in normal distribution Gaussian curvature $K=-\frac{1}{2}$, and in Cauchy distribution K=-2, while in t family distribution with r degrees of freedom, he get $K=-\frac{r+3}{2r}$. In other words, the Gaussian curvature of t_3 distribution is the geometric mean of the curvatures for the Cauchy and normal distribution.

In 2003, Gruber M.H.J. [4] used the following formula

$$-\frac{1}{2\sqrt{EG-F^2}}\left[\frac{\partial}{\partial u}\frac{G_u-F_v}{\sqrt{EG-F^2}}-\frac{\partial}{\partial v}\frac{F_u-E_v}{\sqrt{EG-F^2}}\right]-\frac{1}{4\left(EG-F^2\right)^2}\begin{vmatrix}E & F & G\\ E_u & F_u & G_u\\ E_v & F_v & G_v\end{vmatrix}$$

to compute the Gaussian curvature of gamma family of distributions and normal distribution. He illustrated some connections between the behaviour of Gaussian curvature of the gamma family of distributions and the central limit theorem as follows:

The random variable that has a Gamma distribution with $\alpha = n$ is the sum of exponential random variables. By the central limit theorem as $n \to \infty$ this random variable tends towards that of a normal distribution. As $n \to \infty$ the curvature of the gamma family of distributions tends towards $-\frac{1}{2}$, the curvature of the normal family of distributions.

In 2004, Arwini K., Del Riego L. and Dodson C. T. J. [5] provided formulae for universal connections and curvatures on exponential families and gave an explicit example for the manifold of gamma distributions.

CONTENTS

SUPERVISOR CERTIFICATION

I certify that this thesis was prepared under my supervision at the Department of Mathematics, College of Education, University of Babylon, as a partial fulfillment of the requirements for the degree of Master of Science in Mathematics.

Signature:

Name: ***Assist. Prof. Hussein Ahmed Ali.***

Date: / / 2006

RECOMMENDATION OF THE HEAD OF THE DEPARTMENT

In view of the available recommendations, I forward this research for debate by the Examining Committee.

Signature:

Name: Dr. Iftichar Mudhar Talib

Date: / / 2006

CHAPTER ONE

Some Definitions and Concepts from Differential Geometry

CHAPTER TWO

Some Definitions and Concepts from Statistics

CHAPTER THREE

Connections between Statistics and Differential Geometry

CHAPTER FOUR

Gaussian Curvature for Distributions

CHAPTER FIVE

Conclusions and Recommendations

REFERENCES

الاهـــــداء

الى ينبوع المحبة والتضحية والعطاء الدائم

والدي

الى نبع الحنان وشلال العطاء
الى القلب الذي يمدني بالدفء

امي الحبيبة

الى احباء قلبي..

إخواني وأخواتي

وسـن

بسم الله الرحمن الرحيم

﴿ وَأَنْ لَيسَ لِلإِنسَانِ إِلاَ مَا سَعَى* وَأَنَ سَعْيَهُ سَوْفَ يُرَى* ثُمَّ يُجزاهُ الجَزَاءَ الأَوفَى* وَأَنَّ إلى رَبِّكَ المُنْتَهى □

صدق الله العلي العظيم

سورة النجم آية (39-42)

الخـلاصة

إن استخدام الرياضيات لوضع الحل لمسائل عديدة، معروف من اماضي. ووفقـا لهـذه الحقيقة تم دراسة استخدام الهندسة التفاضلية لتبيان اي من التوزيعات المسـتمرة يقتـرب مـن التوزيع الطبيعي وذلك باستخدام الربط بين الهندسة التفاضلية والاحصاء. وبالاخص توضـيح الربط بين توزيع باريتو وتوزيع والد مع التوزيع الطبيعي باستخدام بعض النظريات الاحصائية والهندسة التفاضلية. وتتضمن الرسالة خمسة فصول:

الفصل الأول يقدم بعض التعاريف والمفاهيم من الهندسة التفاضـلية مثـل Tensor، الصيغة الأساسية الأولى والثانية، تقوس كاوس (Gaussian Curvature)، من هذه المفـاهيم تم الحصول على صيغ مختلفة لحساب قيمة تقوس كـاوس (Gaussian Curvature) علـى اساس العلاقة بين Riemannian Metric و Fisher Information.

الفصل الثاني يوضح بعض التعاريف والمفاهيم من الاحتمالية والاحصاء التي تحتـاج اليها الرسالة مثل دالة الكثافة الاحتمالية، دالة التوزيع المستمرة، بعض التوزيعـات المسـتمرة الخاصة، Fisher Information، تقارب المتغير العشوائي، قانون الاعداد الكبيرة ومبرهنـة الغاية المركزية.

الفصل الثالث يقدم بعض العلاقات بين الإحصاء والهندسة التفاضـلية، مثـل تعريـف المعاملات لـ Expected Fisher Information Matrix على أنها تساوي معاملات الصيغة الأساسية الأولى أو (Riemannian Metric) معطاة من قبل:

$$g_{ij} = -\int \frac{\partial^2 \log f(x,\theta)}{\partial \theta_i\, \partial \theta_j} f(x,\theta)dx = -E\left[\frac{\partial^2 \log f(x,\theta)}{\partial \theta_i\, \partial \theta_j}\right],$$

العلاقـة بـين Riemannian Metrics و Geodesics، Fisher Information Rao Distances بين التوزيعات الاحتمالية، Riemannian Metrics لبعض التوزيعات. تقـوس كاوس (Gaussian Curvature) للتوزيعات الاحتمالية، وتم تعريف Christoffel symbols مع بعض الأمثلة لحساب تقوس كاوس باستخدام صيغ مختلفة.

في الفصل الرابع تم استخدام بعض الطرق لحساب Gaussian Curvature لـبعض التوزيعات. كذلك تم تطبيق هذه الطرق لحساب Gaussian Curvature لتوزيع باريتو وتوزيع والد.

وتضمن الفصل الخامس الاستنتاجات والتوصيات.

REPUBLIC OF IRAQ
MINISTRY OF HIGHER EDUCATION
AND SCIENTIFIC RESEARCH
BABYLON UNIVERSITY
COLLEGE OF EDUCATION

Application of Central Limit Theorem on Probability Distributions by Using Differential Geometry

A THESIS
SUBMITTED TO THE DEPARTMENT OF MATHEMATICS- COLLEGE OF EDUCATION- UNIVERSITY OF BABYLON IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN MATHEMATICS

By
Wasan Abbas Jasim Al-Hisnawi

Supervised by
Assist. Prof. Hussein Ahmed Ali

2006