A GEOMETRIC APPROACH OF PROBABILITY DISTRIBUTIONS

IOANA ANTONIA BRANEA AND IOANA RĂDULESCU (LĂZĂRESCU)

Abstract. Over the past several decades, Information Geometry has had a significant impact, providing a powerful lens to understand and manipulate information that led to advancements in theory, algorithms, and practical applications in various fields, ranging from statistics and machine learning to optimization, quantum information theory, and physics. In this paper we provide a brief theoretical background of statistical models and we conduct an extensive differential geometric study on the set of exponential and Bernoulli distributions. Our results reveal that the statistical models given by the exponential distribution and the one given by the Bernoulli distribution are 1-type curves in $\mathbb{R}^2$.

Mathematics Subject Classification (2010): 53B12, 62B11
Key words: Probability density functions, Fisher metric, submanifolds of finite type.

Article history:
Received: June 5, 2023
Received in revised form: July 16, 2023
Accepted: July 17, 2023

1. Introduction

Geometric modeling is part of one of the most recent branches of mathematics, Information Geometry, where tools from Statistics and Differential Geometry are used to study information loss, statistical inference, and estimation. Information Geometry has applicability in various domains such as physics, signal processing, computer science, machine learning, neuroscience, and optimization in high-dimensional spaces.

In the pioneering work on Information Geometry of Amari and Nagaoka [1], we are introduced to the notion of a manifold of probability density functions. A representative example is the set of normal distributions with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in (0, \infty)$:

$$p(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbb{R},$$

that can be treated as a two-dimensional surface. By endowing it with a Riemannian metric (usually the Fisher information matrix due to the work of Rao [8] and Jeffreys [6]), this becomes a space of constant negative curvature (see [1], [3]).

In the present paper we introduce a differential geometric study of the exponential distribution space and the Bernoulli distribution space consisting of explicit computations of the Fisher metric, Christoffel symbols of the first and the second kind, the geodesics, and the Laplace-Beltrami operator. Using a specific immersion for each of the two models, we also prove that the set of exponential distributions and the set of Bernoulli distributions are 1-type curves in $\mathbb{R}^2$.

The paper is structured as follows. In Section 2 the theoretical background of Probability Theory and Statistics (Subsection 2.1), Differential Geometry (Subsection 2.2), and Statistical Manifolds (Subsection 2.3) is presented. The differential geometric study of the set on exponential distributions is introduced in

Subsection 3.1 and in Subsection 3.2 we present the geometric study on the set of Bernoulli distributions. Section 4 concludes the paper.

2. Preliminaries

In this section, we provide a brief overview of the background knowledge necessary to comprehend the topic, as well as the notations used throughout the paper.

2.1. **Probability Theory and Statistics.** The two interconnected fields provide a systematic framework for understanding, analyzing, and interpreting data.

Let $(\Omega, \mathcal{F}, P)$ be a *probability space*, where

- Ω is the set of all possible outcomes;
- σ-field $\mathcal{F}$ is a collection of subsets of Ω that is closed under complements and countable intersections;
- P is a probability function, i.e. a measure on $\mathcal{F}$ for which $P(\Omega) = 1$.

A *random variable* X on $(\Omega, \mathcal{F}, P)$ is a function $X : \Omega \to \mathbb{R}$ that satisfies

$$X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}, \quad \forall A \in \mathcal{B}, \tag{2.1}$$

where $\mathcal{B}$ is the Borel algebra on the set of real numbers. There are two classes of random variables:

1. *discrete random variables* $X : \Omega \to \chi = \{x_1, x_2, \dots\}$ for which the density function $p : \chi \to \mathbb{R}$ satisfies

$$p(x) = \begin{cases} P(X = x_i), & x = x_i, i = \{1, 2, \dots\}, \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad \sum_k p(x_k) = 1. \tag{2.2}$$

2. *continuous random variables* $X : \Omega \to \chi \in \mathbb{R}^n$ for which the density function $p : \chi \to \mathbb{R}$ satisfies

$$P(X \in D) = \int_D p(x)\,\mathrm{d}x \qquad \text{and} \qquad \int_\chi p(x)\,\mathrm{d}x = 1. \tag{2.3}$$

From the numerical characteristics of random variables, we recall the *expectation value*

$$\mathbf{E}(X) = \begin{cases} \displaystyle\sum_i x_i p(x_i), & \text{if } X \text{ is a discrete random variable,} \\ \displaystyle\int_\chi x p(x)\,\mathrm{d}x, & \text{if } X \text{ is a continuous random variable} \end{cases}. \tag{2.4}$$

For a more detailed introduction to Probability Theory and Statistics, we refer to [2].

2.2. **Differential Geometry.** By investigating curves, surfaces, manifolds, and studying concepts like tangent vectors, curvature, and metrics, Differential Geometry reveals the intrinsic properties of geometric objects.

Let $(M, \mathcal{A})$ be a *differentiable manifold,* where M is a topological space and $\mathcal{A} = \{(U_i, h_i) : i \in I\}$ is the *atlas*, i.e. a collection of *charts* which are bijective mappings between open subsets of M and open subsets of $\mathbb{R}^m$. An *immersion* is a mapping $x : M \to \mathbb{R}^m$ that has rank $n = \dim M$.

Remember that a *Riemannian metric* g on a differentiable manifold is a symmetric, positive definite bilinear form on the tangent space. The pair (M, g) is called a *Riemannian manifold.* The Riemannian metric allows for the definition of various geometric quantities such as the *Christoffel symbols of the first kind*

$$\Gamma_{ij,k} = \frac{1}{2}\left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right), \tag{2.5}$$

and the *Christoffel symbols of the second kind*

$$\Gamma^k_{ij} = \frac{1}{2}\sum_{l=1}^{n} g^{kl}\left(\frac{\partial g_{jl}}{\partial x^i} + \frac{\partial g_{il}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l}\right). \tag{2.6}$$

With the Riemannian metric, one can also define *geodesics*, which are the paths that locally minimize distance

$$\frac{d^2x^k}{dt^2} + \sum_{i,j=1}^{n} \Gamma^k_{ij} \cdot \frac{dx^i}{dt} \cdot \frac{dx^j}{dt} = 0, \qquad k = \overline{1,n}. \tag{2.7}$$

The *Laplace-Beltrami operator* is defined by

$$\Delta f := -\frac{1}{\sqrt{\det g}} \sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(g^{ij} \cdot \sqrt{\det g} \cdot \frac{\partial f}{\partial x_j}\right), \tag{2.8}$$

where g^{ij} is the inverse of the Fisher metric g_{ij}. A function satisfying $\Delta f = 0$ is called *harmonic*.

Recall that a *submanifold* is a subset of a manifold that itself possesses the structure of a manifold. It is well-known (see [4]) that an isometric immersion $x : M \to \mathbb{R}^m$, $x = (x^1, \ldots, x^m)$, $x^i \in C^\infty(M), i = \overline{1,m}$ satisfies

$$x^i = x^i_0 + \sum_{t=p_i}^{q_i} x^i_t, \quad i = \overline{1,m}, \tag{2.9}$$

with $x^i_0 \in \mathbb{R}$ and x^i_t eigenfunctions of the Laplace-Beltrami operator.

Chen [4] defines *submanifolds of finite type* by denoting

$$p = \min\{p_i : i = \overline{1,m}\} \in \mathbb{N}^* \quad \text{and} \quad q = \min\{q_i : i = \overline{1,m}\} \in \mathbb{N}^* \cup \{\infty\} \tag{2.10}$$

as follows.

Definition 2.1. [4] *A compact submanifold M in $\mathbb{R}^m$ is said to be of finite type if q from* (2.10) *is finite. Otherwise, M is of infinite type.*

If the set $\{t \in \{p, p+1, \ldots, q\} : x_t \neq 0\}$ has exactly k elements, then M is said to be of k-type.

Finally, we present the following characterization for the submanifolds of finite type.

Theorem 2.2. [4] *Let $x : M \to \mathbb{R}^m$ be an isometric immersion of a compact, n-dimensional Riemannian manifold M. Then M is of finite type if and only if there is a non-trivial polynomial P such as*

$$P(\Delta)H = \Delta^k H + c_1\Delta^{k-1}H + \cdots + c_{k-1}\Delta H + c_k H = 0, \quad c_i \in \mathbb{R}, \quad i = \overline{1,k}, \tag{2.11}$$

where H is the ***mean curvature vector*** *defined by*

$$H = \frac{1}{n}\sum_{i=1}^{n} h(e_i, e_i), \tag{2.12}$$

for any orthonormal frame $e_1, ..., e_n$.

For more details concerning the concepts presented in this subsection we refer to [5], [7], and [4].

2.3. **Statistical Manifolds.** The notions presented in the above subsections intertwine, resulting the theory behind statistical (or parametric) models. Explicitly, a family of probability distributions which depends on a finite number of parameters can be considered a parameterized surface.

Denote the set of probability distributions on χ that depends on n parameters $\xi = \left(\xi^1, \dots, \xi^n\right)$ by

$$\mathcal{S} = \{p_\xi = p(x;\xi)\}. \tag{2.13}$$

$\mathcal{S}$ is a subset of $\mathcal{P}(\chi) = \left\{f : \chi \to \mathbb{R} : f \geq 0, \int_\chi f dx = 1\right\}$. If the mapping $\xi \to p_\xi$ is an immersion, then the set $\mathcal{S}$ is a *statistical model* of dimension n.

In our computations, we will make use of the *log-likelihood function* given by

$$\ell_x(\xi) = \ell(p_\xi)(x) = \ln p_\xi(x). \tag{2.14}$$

Recall that the *Fisher information matrix* is given by

$$g_{ij}(\xi) = \mathbf{E}\left[\frac{\partial \ell_x(\xi)}{\partial \xi^i} \cdot \frac{\partial \ell_x(\xi)}{\partial \xi^j}\right], \quad \forall i, j \in \{1, \dots, n\}, \tag{2.15}$$

where $\xi = (\xi^1, \dots, \xi^n) \in \mathbb{R}$. It is easy to prove (see e.g. [3, Proposition 1.6.2]) that for any statistical model, the Fisher information matrix is a Riemannian metric. As a consequence, the pair $(\mathcal{S}, g)$ can be organized as a manifold.

For a more detailed presentation of statistical models, we refer to [1] or [3].

3. Main Results

In the bellow paragraphs, we present a differential geometric study for the exponential distribution and the Bernoulli distribution. For the rest of this paper, let $(\Omega, \mathcal{F}, P)$ be a probability space.

3.1. **Exponential Distribution.** We denote the family of exponential distribution by

$$\mathcal{S} = \{p_\xi(x) = \xi e^{-\xi x} : \xi > 0, x \geq 0\}. \tag{3.1}$$

3.1.1. *The Fisher Metric.* We start our study with the computation of the Fisher information matrix, that will be used in the sequel as a Riemannian metric for the corresponding manifold.

Proposition 3.1. *The Fisher information matrix of $\mathcal{S}$ is given by*

$$g_{11}(\xi) = \frac{1}{\xi^2}. \tag{3.2}$$

Proof. It is known (see e.g. [3]) that the Fisher information matrix can be written as:

$$g_{ij}(\xi) = -\mathbf{E}\left[\frac{\partial^2 \ell_x(\xi)}{\partial \xi^i \partial \xi^j}\right]. \tag{3.3}$$

For $\mathcal{S}$, the log-likelihood function is

$$\ell_x(\xi) = \ln\left(\xi e^{-\xi x}\right) = \ln \xi - \xi x,$$

hence

$$\frac{\partial \ell_x(\xi)}{\partial \xi} = \frac{1}{\xi} - x \implies \frac{\partial^2 \ell_x(\xi)}{\partial \xi^2} = -\frac{1}{\xi^2}.$$

Finally, formulas (3.3) and (2.3) provide the Fisher information matrix

$$g_{11}(\xi) = -\mathbf{E}\left[-\frac{1}{\xi^2}\right] = \int_0^\infty \frac{1}{\xi^2} p_\xi(x)\,\mathrm{d}x = \frac{1}{\xi^2}\int_0^\infty p_\xi(x)\,\mathrm{d}x = \frac{1}{\xi^2}.$$

□

3.1.2. *The Christoffel Symbols.* We now use the Fisher information matrix to compute the Christoffel Symbols for the manifold $(\mathcal{S}, g)$.

Proposition 3.2. *The Christoffel symbols of the first and the second kind of $(\mathcal{S}, g)$ are given by:*

$$\Gamma_{11,1} = -\frac{1}{\xi^3} \qquad \text{and} \qquad \Gamma^1_{11} = -\frac{1}{\xi}. \tag{3.4}$$

Proof. Applying (2.5) and (2.6), we obtain:

$$\Gamma_{11,1} = \frac{1}{2}\left(\frac{\partial g_{11}}{\partial \xi} + \frac{\partial g_{11}}{\partial \xi} - \frac{\partial g_{11}}{\partial \xi}\right) = -\frac{1}{2} \cdot \frac{2}{\xi^3} = -\frac{1}{\xi^3},$$

respectively

$$\Gamma^1_{11} = \frac{1}{2} g^{11}\left(\frac{\partial g_{11}}{\partial \xi} + \frac{\partial g_{11}}{\partial \xi} - \frac{\partial g_{11}}{\partial \xi}\right) = -\frac{1}{2} \cdot \xi^2 \cdot \frac{2}{\xi^3} = -\frac{1}{\xi}.$$

□

3.1.3. *The Geodesics.* Using the Christoffel symbols in (2.7), we can compute the geodesics.

Proposition 3.3. *The geodesics of $(\mathcal{S}, g)$ are given by*

$$\xi(t) = e^{c_1 t + c_2}, \tag{3.5}$$

where c_1, c_2 are constants.

Proof. Applying (2.7), we have

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} + \Gamma^1_{11} \frac{\mathrm{d}\xi}{\mathrm{d}t} \cdot \frac{\mathrm{d}\xi}{\mathrm{d}t} = 0 \iff \frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} - \frac{1}{\xi} \cdot \left(\frac{\mathrm{d}\xi}{\mathrm{d}t}\right)^2 = 0.$$

We obtained the homogeneous differential equation

$$\xi'' - \frac{(\xi')^2}{\xi} = 0.$$

We divide the above equation by ξ ($\xi \neq 0$)

$$\frac{\xi''}{\xi} - \frac{(\xi')^2}{\xi^2} = 0 \iff \frac{\xi'' \cdot \xi - \xi' \cdot \xi'}{\xi^2} = 0 \iff \left(\frac{\xi'}{\xi}\right)' = 0,$$

and by integration, we have

$$\frac{\xi'}{\xi} = c_1 \iff \frac{\mathrm{d}\xi}{\mathrm{d}t} = c_1 \xi \iff \frac{\mathrm{d}\xi}{\xi} = c_1 \,\mathrm{d}t \iff \int \frac{\mathrm{d}\xi}{\xi} = \int c_1 \,\mathrm{d}t$$

$$\iff \ln(\xi) = c_1 t + c_2 \implies \xi(t) = e^{c_1 t + c_2}, \quad c_1, c_2 \text{ constants},$$

concluding the proof. □

3.1.4. *The Laplace-Beltrami Operator.* Using (2.8), we will compute Δf with respect to g. In the end of this subsection, we will find those functions f that are harmonic.

Proposition 3.4. *The Laplace-Beltrami operator operates on differentiable functions $f : \mathcal{S} \to \mathbb{R}$ via*

$$\Delta f = -\xi\left(\frac{\partial f}{\partial \xi} + \xi \frac{\partial^2 f}{\partial \xi^2}\right). \tag{3.6}$$

Proof. It is easy to see that $\det g = \frac{1}{\xi^2}$ and $g^{11} = \xi^2$. Then

$$\Delta f = -\frac{1}{\sqrt{\frac{1}{\xi^2}}} \frac{\partial}{\partial \xi}\left(\xi^2 \cdot \frac{1}{\xi} \cdot \frac{\partial f}{\partial \xi}\right) = -\xi \frac{\partial}{\partial \xi}\left(\xi \frac{\partial f}{\partial \xi}\right) = -\xi\left(\frac{\partial f}{\partial \xi} + \xi \frac{\partial^2 f}{\partial \xi^2}\right).$$

□

Next we consider the case $\Delta f = 0$. We have the homogeneous differential equation:

$$f' + \xi f'' = 0.$$

We denote by $u(\xi) = f'(\xi)$. Then the above relation becomes

$$u + \xi u' = 0 \iff \xi \frac{\mathrm{d}u}{\mathrm{d}\xi} = -u \iff \frac{\mathrm{d}u}{u} = -\frac{\mathrm{d}\xi}{\xi}.$$

By integration, we have

$$\ln u = -\ln \xi + \ln c_1 \iff \ln u = \ln\left(\frac{c_1}{\xi}\right) \iff u = \frac{c_1}{\xi}$$

$$\iff \frac{\mathrm{d}f}{\mathrm{d}\xi} = \frac{c_1}{\xi} \iff \int \mathrm{d}f = \int \frac{c_1}{\xi}\,\mathrm{d}\xi \iff f(\xi) = c_1 \ln \xi + c_2, \quad c_1, c_2 \text{ constants.}$$

3.1.5. *Submanifold of finite type.* In this subsection, we will use the framework provided by Chen in [4] to study the family of exponential distributions as a curve in $\mathbb{R}^2$.

Theorem 3.5. *The set of exponential distributions is a 1-type curve in* $\mathbb{R}^2$*.*

Proof. We consider the immersion $x : \mathcal{S} \to \mathbb{R}^2$ defined by:

$$x(\xi) = (\cos(\ln \xi), \sin(\ln \xi))\,. \tag{3.7}$$

Indeed, we have

$$\frac{\partial x}{\partial \xi} = \left(-\sin(\ln \xi) \cdot \frac{1}{\xi}, \cos(\ln \xi) \cdot \frac{1}{\xi}\right)$$

and

$$g_{11} = \left\langle \frac{\partial x}{\partial \xi}, \frac{\partial x}{\partial \xi} \right\rangle = \sin^2(\ln \xi) \cdot \frac{1}{\xi^2} + \cos^2(\ln \xi) \cdot \frac{1}{\xi^2} = \frac{1}{\xi^2},$$

where $\langle \cdot, \cdot \rangle$ is the inner product.

We have

$$\frac{\partial^2 x}{\partial \xi^2} = \left(\frac{-\cos(\ln \xi)\frac{1}{\xi}\xi + \sin(\ln \xi)}{\xi^2}, \frac{-\sin(\ln \xi)\frac{1}{\xi}\xi - \cos(\ln \xi)}{\xi^2}\right)$$
$$= \left(\frac{-\cos(\ln \xi) + \sin(\ln \xi)}{\xi^2}, \frac{-\sin(\ln \xi) - \cos(\ln \xi)}{\xi^2}\right).$$

Applying (3.6), we obtain:

$$\Delta x = -\xi \left(\frac{\partial x}{\partial \xi} + \xi \frac{\partial^2 x}{\partial \xi^2}\right)$$
$$= -\xi \left[\left(\frac{-\sin(\ln \xi)}{\xi}, \frac{\cos(\ln \xi)}{\xi}\right) + \xi \left(\frac{-\cos(\ln \xi) + \sin(\ln \xi)}{\xi^2}, \frac{-\sin(\ln \xi) - \cos(\ln \xi)}{\xi^2}\right)\right]$$
$$= -\xi \left(\frac{-\cos(\ln \xi)}{\xi}, \frac{-\sin(\ln \xi)}{\xi}\right) \implies \Delta x = (\cos(\ln \xi), \sin(\ln \xi))\,.$$

It is known (see [4]) that H satisfies

$$\Delta x = -nH, \tag{3.8}$$

where n is the dimension of the submanifold and H is the mean curvature vector. Then

$$H = -\Delta x \implies H = (-\cos(\ln \xi), -\sin(\ln \xi))\,.$$

The first and second order partial derivatives are

$$\frac{\partial H}{\partial \xi} = \left(\frac{\sin(\ln \xi)}{\xi}, \frac{-\cos(\ln \xi)}{\xi}\right),$$

$$\frac{\partial^2 H}{\partial \xi^2} = \left(\frac{\cos(\ln \xi) - \sin(\ln \xi)}{\xi^2}, \frac{\sin(\ln \xi) + \cos(\ln \xi)}{\xi^2} \right).$$

Hence

$$\Delta H = -\xi \left[\left(\frac{\sin(\ln \xi))}{\xi}, \frac{-\cos(\ln \xi)}{\xi} \right) + \xi \left(\frac{\cos(\ln \xi) - \sin(\ln \xi)}{\xi^2}, \frac{\sin(\ln \xi) + \cos(\ln \xi)}{\xi^2} \right) \right]$$

$$\Longrightarrow \Delta H = -\left(\cos(\ln \xi), \sin(\ln \xi)\right).$$

We showed that the mean curvature vector satisfies the following relation

$$\Delta H - H = 0,$$

so, by Theorem 2.2 we conclude that $\mathcal{S}$ is a 1-type curve in $\mathbb{R}^2$. □

Corollary 3.6. *An eigenvalue of the Laplace-Beltrami operator is 1.*

3.2. **Bernoulli Distribution.** In this section we study the manifold of Bernoulli distributions for a single experiment with 2 possible outcomes. The family set of Bernoulli distributions is given by

$$\mathcal{S} = \{p(\xi; k) = \xi^k (1-\xi)^{1-k} : \ 0 < \xi < 1, k \in \{0, 1\}\}. \tag{3.9}$$

3.2.1. *The Fisher Metric.*

Proposition 3.7. *The Fisher information matrix has one element given by*

$$g_{11}(\xi) = \frac{1}{\xi(1-\xi)}. \tag{3.10}$$

Proof. The log-likelihood for the Bernoulli probability density function is given by

$$\ell_k(\xi) = \ln p(\xi; k) = \ln(\xi^k (1-\xi)^{1-k}) = k \ln \xi + (1-k) \ln(1-\xi). \tag{3.11}$$

The first and second derivatives of the log-likelihood with respect to the parameter ξ are given by

$$\frac{\partial \ell_k(\xi)}{\partial \xi} = \frac{k - \xi}{\xi(1-\xi)}. \tag{3.12}$$

$$\frac{\partial^2 \ell_k(\xi)}{\partial \xi^2} = -\frac{k}{\xi^2} - \frac{1-k}{(1-\xi)^2}. \tag{3.13}$$

For computing the Fisher metric coefficients we use the formula from [3, Proposition 1.6.3]

$$g_{ij}(\xi) = -\mathbf{E}\left[\frac{\partial^2 \ell_x(\xi)}{\partial \xi^i \, \partial \xi^j} \right].$$

In our case:

$$g_{11}(\xi) = \mathbf{E}\left[\frac{k}{\xi^2} + \frac{1-k}{(1-\xi)^2} \right]. \tag{3.14}$$

From the definition (2.4) of the expectation we have that

$$\mathbf{E}[k] = \xi$$
$$\mathbf{E}[1-k] = 1 - \xi.$$

So we obtain:

$$g_{11}(\xi) = \frac{1}{\xi} + \frac{1}{1-\xi} = \frac{1}{\xi(1-\xi)}.$$

□

3.2.2. *Christoffel Symbols.* Using the Fisher information matrix, we can compute the Christoffel symbols for the manifold $(\mathcal{S}, g)$.

Proposition 3.8. *The Christoffel symbols of first and second kind are given by*

$$\Gamma_{11,1} = \frac{2\xi - 1}{2\xi^2(1-\xi)^2} \tag{3.15}$$

and

$$\Gamma^1_{11} = \frac{2\xi - 1}{2\xi(1-\xi)}. \tag{3.16}$$

Proof. By applying formulas (2.5) and (2.6) we obtain

$$\Gamma_{11,1} = \frac{1}{2}\left(\frac{\partial g_{11}}{\partial \xi} + \frac{\partial g_{11}}{\partial \xi} - \frac{\partial g_{11}}{\partial \xi}\right) = \frac{1}{2}\left(-\frac{1-2\xi}{\xi^2(1-\xi)^2}\right) = \frac{2\xi - 1}{2\xi^2(1-\xi)^2}.$$

$$\Gamma^1_{11} = \frac{1}{2} g^{11}\left(\frac{\partial g_{11}}{\partial \xi} + \frac{\partial g_{11}}{\partial \xi} - \frac{\partial g_{11}}{\partial \xi}\right),$$

where g^{11} is the inverse of the Fisher matrix, in our case $g^{11} = \xi(1-\xi)$.

$$\Gamma^1_{11} = \frac{1}{2}\xi(1-\xi)\frac{2\xi - 1}{\xi^2(1-\xi)^2} = \frac{2\xi - 1}{2\xi(1-\xi)}.$$

□

3.2.3. *Geodesics.* By replacing the formulas obtained in (3.16) for the Christoffel symbols in (2.7), we can compute the geodesics equations.

Proposition 3.9. *The geodesics for the Bernoulli distribution model are given by*

$$\xi(t) = \frac{1}{2}(1 + \sin(c_1 t + c_2)), \quad c_1, c_2 \in \mathbb{R}. \tag{3.17}$$

Proof. From (2.7) we have

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} + \Gamma^1_{11}\frac{\mathrm{d}\xi}{\mathrm{d}t}\cdot\frac{\mathrm{d}\xi}{\mathrm{d}t} = 0 \iff \frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} + \frac{2\xi - 1}{2\xi(1-\xi)}\frac{\mathrm{d}\xi}{\mathrm{d}t}\cdot\frac{\mathrm{d}\xi}{\mathrm{d}t} = 0.$$

We make the substitution $\frac{\mathrm{d}\xi}{\mathrm{d}t} = u \implies \frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} = \frac{\mathrm{d}u}{\mathrm{d}\xi}\frac{\mathrm{d}\xi}{\mathrm{d}t} \iff \xi'' = u\frac{\mathrm{d}u}{\mathrm{d}\xi}$. By replacing this in the equation above, we get

$$u\frac{\mathrm{d}u}{\mathrm{d}\xi} + \frac{2\xi - 1}{2\xi(1-\xi)}u^2 = 0. \tag{3.18}$$

We distinguish 2 cases:

(1) $u = 0 \iff \frac{\mathrm{d}\xi}{\mathrm{d}t} = 0 \iff \xi(t) = c$, $c \in \mathbb{R}$;
(2) $u \neq 0$. We divide (3.18) by u, which leads to

$$\frac{\mathrm{d}u}{\mathrm{d}\xi} = -\frac{2\xi - 1}{2\xi(1-\xi)}u \iff \frac{\mathrm{d}u}{u} = \frac{1-2\xi}{2\xi(1-\xi)}\,\mathrm{d}\xi.$$

By integrating both sides, we get

$$\ln u = \ln\left(c\xi(1-\xi)\right)^{\frac{1}{2}}, \quad c \in \mathbb{R}.$$

We obtain

$$u = (c\xi(1-\xi))^{\frac{1}{2}} \iff \frac{\mathrm{d}\xi}{\mathrm{d}t} = (c\xi(1-\xi))^{\frac{1}{2}} \iff \frac{\mathrm{d}\xi}{\sqrt{\xi(1-\xi)}} = c_1\,\mathrm{d}t, \quad c_1 \in \mathbb{R}.$$

By integrating both sides, we get

$$\arcsin(2\xi - 1) = c_1 t + c_2, \iff$$
$$\xi(t) = \frac{1}{2}(1 + \sin(c_1 t + c_2)), \quad c_1, c_2 \in \mathbb{R}.$$

This ends the proof. □

3.2.4. *Laplace-Beltrami operator and harmonic functions.*

Proposition 3.10. *The formula for the Laplace-Beltrami operator acting on smooth functions $f : \mathcal{S} \to \mathbb{R}$ is:*

$$\Delta f = \frac{2\xi - 1}{2}\frac{\partial f}{\partial \xi} - \xi(1-\xi)\frac{\partial^2 f}{\partial \xi^2}. \tag{3.19}$$

Proof. Using formula (3.8) for the Bernoulli distribution model, this becomes

$$\Delta f = -\sqrt{\xi(1-\xi)}\frac{\partial}{\partial \xi}\left(\xi(1-\xi)\sqrt{\frac{1}{\xi(1-\xi)}}\frac{\partial f}{\partial \xi}\right) \iff$$
$$\Delta f = -\sqrt{\xi(1-\xi)}\left(\frac{\partial}{\partial \xi}(\sqrt{\xi(1-\xi)}) \cdot \frac{\partial f}{\partial \xi} + \sqrt{\xi(1-\xi)} \cdot \frac{\partial^2 f}{\partial \xi^2}\right) \iff$$
$$\Delta f = -\sqrt{\xi(1-\xi)}\left(\frac{1-2\xi}{2\sqrt{\xi(1-\xi)}} \cdot \frac{\partial f}{\partial \xi} + \sqrt{\xi(1-\xi)} \cdot \frac{\partial^2 f}{\partial \xi^2}\right) \iff$$
$$\Delta f = \frac{2\xi - 1}{2}\frac{\partial f}{\partial \xi} - \xi(1-\xi)\frac{\partial^2 f}{\partial \xi^2}.$$

□

Proposition 3.11. *Harmonic functions have the following expression*

$$f(\xi) = c_1 \arcsin(2\xi - 1) + c_2, \qquad c_1, c_2 \in \mathbb{R}. \tag{3.20}$$

Proof. Harmonic functions are those that satisfy $\Delta f = 0$

$$\frac{2\xi - 1}{2}\frac{\mathrm{d}f}{\mathrm{d}\xi} - \xi(1-\xi)\frac{\mathrm{d}^2 f}{\mathrm{d}\xi^2} = 0.$$

We make the substitution $\frac{\mathrm{d}f}{\mathrm{d}\xi} = u \iff f'' = \frac{\mathrm{d}u}{\mathrm{d}\xi}$ and we obtain, by rearranging the terms and also taking into account that $0 < \xi < 1$

$$\frac{2\xi - 1}{2}u - \xi(1-\xi)\frac{\mathrm{d}u}{\mathrm{d}\xi} = 0 \iff$$
$$\frac{\mathrm{d}u}{u} = \frac{2\xi - 1}{2\xi(1-\xi)}\,\mathrm{d}\xi.$$

By integrating both sides, we get

$$\ln u = \ln\frac{c_1}{\sqrt{\xi(1-\xi)}} \iff u = \frac{c_1}{\sqrt{\xi(1-\xi)}} \iff \frac{\mathrm{d}f}{\mathrm{d}\xi} = \frac{c_1}{\sqrt{\xi(1-\xi)}} \iff$$
$$\mathrm{d}f = c_1\frac{1}{\sqrt{\xi(1-\xi)}}\,\mathrm{d}\xi \iff f(\xi) = c_1 \arcsin(2\xi - 1) + c_2, \qquad c_1, c_2 \in \mathbb{R}.$$

□

3.2.5. *Submanifold of finite type.* We use again the framework provided by Chen in [4, Chapter 6] to study the type of the family of Bernoulli distributions.

Theorem 3.12. *The family of Bernoulli distributions is* 1*-type curve in* $\mathbb{R}^2$.

Proof. Consider the immersion $x : \mathcal{S} \to \mathbb{R}^2$,

$$x(\xi) = (2\xi^{\frac{1}{2}}, -2(1-\xi)^{\frac{1}{2}}), \qquad 0 < \xi < 1. \tag{3.21}$$

We have

$$\frac{\partial x}{\partial \xi} = (\xi^{-\frac{1}{2}}, (1-\xi)^{-\frac{1}{2}}).$$

$$g_{11} = \left\langle \frac{\partial x}{\partial \xi}, \frac{\partial x}{\partial \xi} \right\rangle = \frac{1}{\xi} + \frac{1}{1-\xi} = \frac{1}{\xi(1-\xi)},$$

the same as the Fisher matrix coefficient. The second order derivative is

$$\frac{\partial^2 x}{\partial \xi^2} = \left(-\frac{1}{2}\xi^{-\frac{3}{2}}, \frac{1}{2}(1-\xi)^{-\frac{3}{2}} \right).$$

By using (3.19) we obtain

$$\Delta x = \frac{2\xi - 1}{2}\left(\xi^{-\frac{1}{2}}, (1-\xi)^{-\frac{1}{2}}\right) - \xi(1-\xi)\left(-\frac{1}{2}\xi^{-\frac{3}{2}}, \frac{1}{2}(1-\xi)^{-\frac{3}{2}} \right).$$

After doing the computations we obtain

$$\Delta x = \left(\frac{1}{2}\sqrt{\xi}, -\frac{1}{2}\sqrt{1-\xi} \right).$$

We know (see e.g. [4]) that the mean curvature vector H satisfies

$$\Delta x = -nH, \tag{3.22}$$

where n is the dimension of the submanifold; in our case $n = 1$. This implies $\Delta x = -H \implies H = -\Delta x$, which means

$$H = \left(-\frac{1}{2}\sqrt{\xi}, \frac{1}{2}\sqrt{1-\xi} \right). \tag{3.23}$$

We compute the first and second order derivatives of H

$$\frac{\partial H}{\partial \xi} = \left(-\frac{1}{4\sqrt{\xi}}, -\frac{1}{4\sqrt{1-\xi}} \right). \tag{3.24}$$

$$\frac{\partial^2 H}{\partial \xi^2} = \left(\frac{1}{8\xi^{\frac{3}{2}}}, -\frac{1}{8(1-\xi)^{\frac{3}{2}}} \right). \tag{3.25}$$

By using the formula (3.19) for the Laplace-Beltrami operator, (3.24), and (3.25) we obtain

$$\Delta H = \frac{2\xi - 1}{2}\frac{\partial H}{\partial \xi} - \xi(1-\xi)\frac{\partial^2 H}{\partial \xi^2} \iff$$

$$\Delta H = \left(-\frac{1}{8}\sqrt{\xi}, \frac{1}{8}\sqrt{1-\xi}. \right) \tag{3.26}$$

From (3.23) and (3.26) we see that the following is true

$$-4\Delta H + H = 0. \tag{3.27}$$

By the characterization Theorem 2.2 for submanifolds of finite type, we conclude that $\mathcal{S}$ is a 1-type curve in $\mathbb{R}^2$. □

Corollary 3.13. *An eigenvalue of the Laplace-Beltrami operator is* $\frac{1}{4}$.

4. Conclusions

Inspired by the geometric study of the set of normal distributions presented by the authors in [3], we have conducted a similar study for the family of exponential distributions and for the family of Bernoulli distributions. In addition to the mentioned work, we have found appropriate immersions in order to apply Theorem 2.1 from [4] and conclude that the exponential distributions set and the Bernoulli distributions set are both 1-type curves in $\mathbb{R}^2$.

References

[1] S. Amari and H. Nagaoka, *Methods of Information Geometry*, American Mathematical Soc., Oxford Univ. Press **191**, 2000.
[2] G. Blom, *Probability and Statistics: Theory and Applications*, Springer, 1989.
[3] O. Calin and C. Udrişte, *Geometric Modeling in Probability and Statistics*, Springer, 2014.
[4] B.-Y. Chen, *Total Mean Curvature and Submanifolds of Finite Type*, World Scientific Publishing Company, 1984.
[5] M.P. do Carmo, *Riemannian Geometry*, Springer, 1992.
[6] H. Jeffreys, *An invariant form for the prior probability in estimation problems*, Proceedings of the Royal Society of London, Series A, **186(1007)** (1946), 453–461.
[7] J. Jost, *Riemannian Geometry and Geometric Analysis*, Springer, 2011.
[8] C.R. Rao, *Information and accuracy attainable in estimation of statistical parameters*, Bulletin of the Calcutta Math. Soc. **37** (1945), 81–91.

Interdisciplinary Doctoral School, Faculty of Mathematics and Computer Science, Transilvania University of Braşov, Romania

Email address: ioana.taca@unitbv.ro, ioana.radulescu@unitbv.ro