

Article

Not peer-reviewed version

---

# Joint Probability Densities as Symmetric Covariant Tensor Densities on Riemannian Manifold

---

[Manouchehr Amiri](#) \*

Posted Date: 19 October 2023

doi: 10.20944/preprints202310.1154.v1

Keywords: Joint probability density, Covariant Tensor, Riemannian manifold, Symmetrized Covariant derivative, Taylor expansion



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

# Joint Probability Densities as Symmetric Covariant Tensor Densities on Riemannian Manifold

Manouchehr Amiri

Independent Researcher; manoamiri@gmail.com

**Abstract:** This paper presents the tensor properties of joint probability densities on a Riemannian manifold. Initially we develop a binary data matrix to record variables' values of a large number of particles confining in a closed system at a certain time in order to retrieve the joint probability densities of related variables. By introducing the particle oriented coordinate and generalized inner product as a multi-linear operation on the basis of this coordinate, we extract the set of joint probabilities and prove them to meet covariant tensor properties on a general Riemannian space of variables. Based on the Taylor expansion of scalar field in Riemannian manifolds, it has been shown that the symmetrized iterative covariant derivatives of cumulative probability function defined on Riemannian manifold also gives the set of related joint probability densities equivalent to the aforementioned multi-linear method. We show these covariant tensors reduce to classical ordinary partial derivatives in ordinary Euclidean space with Cartesian coordinates and gives the formal definition of joint probabilities by partial derivatives of cumulative distribution function. The equivalence between symmetrized covariant derivative and generalized inner product has been concluded. Some examples of well-known physical tensors clarify that many deterministic physical variables are presented as tensor densities with an interpretation similar to probability densities.

**Keywords:** joint probability density; covariant tensor; riemannian manifold; symmetrized covariant derivative; taylor expansion

## 1. Introduction

The current classical theory of probability distributions have been based on the presumptive Euclidean space of parameters and random variables. In this sense the joint probability densities being derived from sequential partial derivatives of corresponding cumulative distribution function [1–3]. The non-Euclidean manifolds endowed with Riemannian metrics in the context of probability theory, have been introduced by many authors for phase and parameter space [4–7]. Moreover using differential geometry, some variations of information theory having been devised on non-Euclidean (Riemannian) metric spaces by Fisher, Rao, Amari and others [5,8]. In these approaches, probability distributions of various models exhibited as points on some Riemannian manifolds. By this background, differential geometry technique could be applied to analyze the probability distribution manifolds. The most applicable metric tensor on these spaces that first introduced by Fisher and Rao, is the so-called Fisher information metric [9]. Applying the Riemannian metric to define the basic concepts in statistics such as mean and covariance matrix of random variable, has also been introduced in other works [6,7,10]. In this short article we introduce a new model based on binary data arrays and matrices to analyse the variables distributions of a system of particles on the Riemannian manifold. A cloud of a large number of particles at a certain time is considered in the space of variables without analysis of its dynamical evolution and quantum physics uncertainties restrictions. The coordinates of variables in this space, is divided to infinitesimal intervals while each interval labelled by an ordered integer number. The variables of each particle occupies just one infinitesimal interval on each variable coordinate  $x^{\nu}$  labelled by some integer number  $i$ . For any specified infinitesimal interval on each coordinates, there is a specific array with binary entries  $\{0, 1\}$  that determines the particles whose  $x^{\nu}$  variable restricted to this interval. Based on this background, in sec (2) we develop the concept of particle oriented coordinates which span a flat Euclidean space on which we embed Riemannian

manifolds of variable coordinates. Collection of all binary arrays of all variables, yield a binary matrix containing entire information of the particles system. By introducing a generalized inner product for a set of vectors, joint probability densities of variables, is calculated as inner product of the binary arrays that stands for some vectors in cotangent space at each point on manifold and consequently the tensor density properties of joint probability density is proved. In sec (3) based on tensor density property of the joint probabilities of variables at any point on manifold, we present a new definition for joint probability densities by symmetrized covariant derivative of cumulative distribution function. A new method to connect concepts in continuous and discrete probability theory and a novel interpretation of covariant derivative by generalized inner product has been proposed. In section (4) some examples that reveal the tensor density properties of famous physical entities are presented.

**Definition 1.** *The space (manifold)  $\mathcal{M}$  of variables spanned by coordinates  $x^\nu$  with  $1 \leq \nu \leq d$  where  $d$  is the dimension of manifold. So the number of involved independent variables (degree of freedom) is  $d$ . The variable space  $\mathcal{M}$  in present article, generally presumed to be a Riemannian manifold with local coordinates  $x^\nu$  and basis vectors  $e_\nu = \frac{\partial}{\partial x^\nu}$  at any point  $p \in \mathcal{M}$ . If we divide the coordinate  $x^\nu$  into  $m_\nu$  small intervals  $\Delta x^\nu(i)$ , while  $m_\nu$  is a large number, then integer  $i$  refers to the  $i$ th interval of this parameter. This means that  $i$  stands for the coordinate value of point  $p$  along  $x^\nu$  and ranges between 0 and a large integer  $m_\nu$ :*

$$1 \leq i \leq m_\nu \quad (1)$$

The whole manifold includes all coordinates  $x^\nu$  and associated basis vectors. The overall number of intervals reads as:

$$m = \sum_{\nu=1}^d m_\nu \quad (2)$$

In this setting the vector space  $\mathcal{M}$  is a lattice space where the coordinates of points on manifold specified by  $d$  digit numbers, therefor the related field of  $\mathcal{M}$  will be  $\mathbb{Z}$ . At the limit  $m_\nu \rightarrow \infty$  manifold  $\mathcal{M}$ , is a smooth one.

**Definition 2.** *Regarding the definition of cumulative distribution or joint distribution function (CDF) [1], [2]. We may define a function  $F(x^1, x^2, \dots, x^d)$  on any point  $P(x^1, x^2, \dots, x^d) \in \mathcal{M}$  as follows:*

$$\begin{aligned} 0 &\leq F \leq 1 \\ \lim_{x^i \rightarrow \infty} F &= 1, \quad 1 \leq i \leq d \\ \lim_{x^i \rightarrow -\infty} F &= 0, \quad 1 \leq i \leq d \end{aligned} \quad (3)$$

### Axiom

For a system consisting of a large number particles, there is a smooth and differentiable probability density function on the manifold  $\mathcal{M}$  that yields the density of particles at any volume element  $dV$  in manifold. This postulate is consistent with the accepted postulates of the kinetic theory of gases.

## 2. Introducing Particle Dependent Coordinates

Assume a system consisting of a large number  $N$  of particles confined in an interval of space-time. Taking into account of such a system of particles, brings us the advantage of choosing a sufficient huge number of particles, moreover we could substitute particles by any kind of systems defined by their arbitrary points in parametric space. Suppose a set of independent parameters labelled by  $\nu$  is to be considered in a small interval of time  $\Delta\tau$ . We label the particles by integer numbers up to  $N$ , and divide the possible range of each parameter into such small intervals that satisfy the accuracy of the experiment. These intervals defined as in definition 1, denoted by  $\Delta x^\nu(i)$  where  $i$  stands for the ordered location number of an intervals on the coordinate  $x^\nu$ :

$$x^v(i-1) < x^v(i) < x^v(i+1) \quad (4)$$

So the variable's value of each particle falls in just one of these intervals.

**Definition 3.** Let the basis vectors  $\varepsilon_1 = (1, 0, 0, \dots)$ ,  $\varepsilon_2 = (0, 1, 0, \dots)$ , ...,  $\varepsilon_N = (0, 0, 0, \dots, 1)$  span a vector space  $\mathcal{V}$  of the dimension  $N$  over the field  $\mathbb{Z}$ , where  $N$  is the number of particles.  $\mathcal{V}$  regarded a lattice Euclidean spaces endowed with Cartesian coordinates. the particles are labelled by ordered integer numbers, then  $\varepsilon_1$  is the basis for the first particle and  $\varepsilon_2$  is the basis for second particle and so on. Here any particle specifies an independent (basis) coordinate with two possible values 0 and 1. Obviously these basis are orthogonal. We call these set of basis as particle oriented coordinate that as a coordinate chart is homeomorphic to a sub-space of Euclidean lattice space  $\mathbb{Z}^N$ . In this case the dual basis presentation coincides the same original basis i.e.  $\varepsilon^{*i} = \varepsilon_i$ . The dual basis could also be represented as a binary array like ;  $\varepsilon^{*3} = (0, 0, 1, \dots, 0)$ . At any point on manifold  $\mathcal{M}$  the basis  $\left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^d}\right)$  span the tangent space  $T_P\mathcal{M}$ . The basis  $\varepsilon^{*\mu}$  acts on the basis  $\partial_\nu = \frac{\partial}{\partial x^\nu}$  of tangent space at the point  $P \in \mathcal{M}$  by the relation  $\varepsilon^{*\mu}(\partial_\nu) = \delta_\nu^\mu$  because  $\varepsilon^{*\mu}$  is specified to the variable  $x^\nu$  and acts on  $\partial_\nu$  to return the component "1" for the related particle, while its action on  $\partial_\mu$  which is independent from  $\partial_\nu$  returns "0". Consequently, the vectors  $\varepsilon^{*\mu}$  are living in cotangent space  $T_P^*\mathcal{M}$ .

For each interval  $\Delta x^v(i)$ , there are some particles that their variable  $x^v$  places in this interval. Let record the results of these outcomes in an array whose entries are 0 or 1 in such a way that for particles with parameter value in interval  $\Delta x^v(i)$ , the corresponding value in array reads 1 and otherwise 0. For instance for the first particle, if its value of parameter  $x^v$  falls in the range of  $\Delta x^v(i)$ , it returns 1, and otherwise 0. If this process iterates for all particles then we obtain an array of entries for this interval which could be arranged as a vector  $e_v(i)$ . This vector is a binary array which carries the information of this interval for system of particles e.g.

$$e_v(i) = (1, 0, 0, 1, 1, 0, 1, \dots) \quad (5)$$

Each of these 0 and 1 connected to a specific particle in the system and the total number of entries equals the total number of particles  $N$ . In this example the first entry 1 means that the value of variable  $x^v$  for particle with label 1 is in the  $i$ th interval. therefor these vectors are members of a vector space of dimension  $N$ . As is defined in definition 3, One may attribute to any particle, an independent basis  $\varepsilon_\nu$  for  $i$  th particle as

$$\varepsilon_\nu = (0, 0, 0, \dots, 1, \dots, 0) \quad (6)$$

where only the  $i$  th entry takes the value 1. The result of projection of a set of  $\varepsilon_\nu$  on the coordinate  $x^v$  turns out the array  $e_\nu(i)$  in equation (5). This means that the sum of basis  $\varepsilon_\nu$  of all particles with the common value of  $x^v$  yields the vector  $e_\nu(i)$ . For example the vector in equation (5), is the sum  $e_\nu(i) = \varepsilon_1 + \varepsilon_4 + \varepsilon_5 + \dots$

The vectors  $e_\nu(i)$  at each point on  $x^v$  belongs to a sub-space  $\mathcal{N}$  of  $\mathcal{V}$  defined in definition 3. Obviously the dimension of  $\mathcal{N}$  is  $d$ . At any point  $P$  on  $\mathcal{M}$ , If the tangent space spanned by  $\left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^d}\right)$  denoted by  $T_P\mathcal{M}$  and tangent space at the same point on  $\mathcal{N}$  denoted by  $T_P\mathcal{N}$ , then there is a one to one mapping between these spaces.

**Remark 1.** Respect to definition 3, the vectors  $e^{*v}(i)$  as the sum over  $\varepsilon^{*i}$  at any point  $P(x^1, x^2, \dots, x^d) \in \mathcal{M}$  is a functional on  $P$  with local coordinates  $\left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^d}\right)$ , thus belongs to the dual space of  $\mathcal{M}$  i.e.  $e_\nu(i) \in T_P^*\mathcal{M}$

**Lemma 1.** The set of basis  $e_\nu(i)$  for fixed  $\nu$  are mutually orthogonal.

**Proof.** Any particle takes just one value and consequently one interval on  $x^v$  coordinate. So as was shown in definition 4, columns of  $\mathfrak{D}_{m_v \times N}$  carries just one entry 1 while other entries takes 0. If the  $n$  th component of  $e_v(i)$  be denoted by  $[e_v(i)]_n$ , then the inner product  $\langle e_v(i), e_v(j) \rangle$  can be read as a sum over all particles:

$$\langle e_v(i), e_v(j) \rangle = \sum_{n=1}^N [e_v(i)]_n [e_v(j)]_n \quad (7)$$

The components  $[e_v(i)]_n$  and  $[e_v(j)]_n$  do not take the 1 value simultaneously, because a particle can not take two values on  $x^v$ , thus it is easy to conclude that inner product  $\langle e_v(i), e_v(j) \rangle$  vanishes for  $i \neq j$

$$\langle e_v(i), e_v(j) \rangle = \delta_{ij} \quad (8)$$

and orthogonality of these bases is proved.  $\square$

**Definition 4.** For a fixed  $v$ , we define the matrix  $\mathfrak{D}_{m_v \times N}$  with  $e_v(i)$  as  $i$ th row.  $m_v$  as described in definition 1, is the number of infinitesimal intervals on  $x^v$ . Obviously each column of this matrix contains just one entry 1 and other entries are 0, because each column belongs to a particle which occupies just one value (interval) at  $\Delta x^v(i)$ .

Because of orthogonality, the set of  $\{e_v(i) | 1 \leq i \leq m_v\}$  for a fixed  $v$  span a tangent space  $T_p^v \mathcal{N}$  which is in one to one mapping with tangent  $T_p^v \mathcal{M}$  with coordinates  $x^v(i)$ .

$$\text{Span} \{e_v\} = T_p^v \mathcal{N} \quad (9)$$

$$T_p^v \mathcal{N} \subset T_p \mathcal{N} \quad (10)$$

Manifold  $\mathcal{N}$  is an lattice Euclidean space but the manifold  $\mathcal{M}$  as a lattice space, is not necessarily flat and may be endowed by a general Riemannian metric.

The probability density of particles within interval  $dx^v(i)$  which will be denoted by  $f_v(i)$ , is proportional to the ratio of total number of entry "1" in the  $e_v(i)$ , denoted by  $n_v(i)$ , to total particles number  $N$ :

$$f_v(i) dx^v(i) = \frac{n_v(i)}{N} \quad (11)$$

It is straight forward to conclude that  $n_v(i)$  equals the inner product  $\langle e_v(i), e_v(i) \rangle$ . In an infinitesimal limit of  $\Delta x^v(i)$  the basis  $e^{*v}(i)$  as the dual vector of  $e_v(i)$  approaches to  $dx^v(i)$  which belongs to the cotangent (dual) space basis of  $T_p^{*v} \mathcal{M}$  while carries the information about that interval in considered system. Respect to coordinate transformations  $x^v \rightarrow \bar{x}^v$  in variable space we apply the rule of  $dx^v(i)$  transformation with summation on  $v$ :

$$d\bar{x}^\mu = \frac{\partial \bar{x}^\mu}{\partial x^v} dx^v \quad (12)$$

In this equation and all tensor equations in next sections, we use Einstein's summation convention on repeated covariant and contravariant indices.

### 2.1. Binary Data Matrix

Combining all the row vectors  $e_v(i)$  for all  $d$  variables of  $N$  particles, gives a matrix  $\mathfrak{D}_{m \times N}$  with  $m$  rows defined in equation (2) and  $N$  as the total number of particles. This matrix itself is also the conjoint of  $d$  matrices each for a specific variable. As an example if for a variable with coordinate  $x^v$ ,

there are  $m_v$  intervals, then a matrix  $\mathfrak{D}_{m_v \times N}$  with  $m_v$  rows and  $N$  columns could be identified as a part of  $\mathfrak{D}_{m \times N}$  where  $e^{*v}(i)$  constitute the set of dual basis for coordinates  $x^v$  at any location  $i$ . The columns of  $\mathfrak{D}_{m_v \times N}$  carries just one entry 1, because each column corresponds to one particle that its variable's value falls in just one of intervals  $\Delta x^v$ , namely  $\Delta x^v(i)$ . Obviously combining all  $\mathfrak{D}_{m_v \times N}$  result in the data matrix  $\mathfrak{D}_{m \times N}$ . As we showed in equation (8) the set of  $\{e_v(i)\}$  span a tangent space  $T_P \mathcal{N}$  on a point  $P \in \mathcal{N}$ .

**Remark 2.** Joint probability density of two variable  $x^v(i)$  and  $x^\mu(j)$  is proportional to the number of particles that simultaneously have the same parameter values of  $x^v(i)$  and  $x^\mu(j)$  or equivalently are confined in  $\Delta x^v(i)$  and  $\Delta x^\mu(j)$  where  $i$  and  $j$  indicate the values of the coordinates  $x^v$  and  $x^\mu$  (i.e. coordinate values). If this number presented by  $n^{\mu v}$  the exact form of joint probability density reads as:

$$f_{\mu v} dV = \frac{n^{\mu v}}{N} \quad \mu \neq v \quad (13)$$

$dV = dx^\mu(i) dx^v(j)$  is the volume element.

It is noteworthy to remind that for  $\mu = v$  the joint density  $f_{\mu v}$  and  $n_{\mu v}$  reduce to  $f_\mu$  and  $n_\mu$  respectively as will be shown in Lemma 3.

**Definition 5.** We define the generalized inner (scalar) product  $\odot$  for vectors  $U, V, W, \dots$  in a vector space  $\mathbb{V}$  with orthogonal local coordinates as a multi-linear map:

$$\Phi : \mathbb{V}^m \rightarrow \mathbb{R} \quad ; \quad U \odot V \odot W \dots = \sum_{n=1}^m u_n v_n w_n \dots \quad (14)$$

Where  $u_n, v_n, \dots$  are  $n$ -th components of  $U, V, \dots$  respectively. This is simply a generalization of inner products in usual definition. For inner product of two basis vector in Cartesian coordinate on Euclidean tangent or cotangent space, results in metric tensor. As an example the inner product for  $e_i$  and  $e_j \in T_P \mathcal{M}$  results in the metric tensor  $g_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle$

**Lemma 2.** Generalized scalar operation  $\odot$  is linear and commutative.

$$\begin{aligned} U \odot V &= V \odot U \\ \lambda_1 U \odot \lambda_2 V &= \lambda_1 \lambda_2 V \odot U \\ (U + V) \odot W &= U \odot W + V \odot W \end{aligned} \quad (15)$$

**Proof.** Due to the definition 5, it is straightforward to derive the equations of linear and commutative properties.  $\square$

If the vector  $U$  is a binary vector, the idempotent property also holds true:

**Lemma 3.** For inner product  $\odot$  of  $k$  vectors when  $k \geq 3$ , the repeated vectors reduces to one vector

$$U \odot V \odot V \odot W = U \odot V \odot W \quad (16)$$

**Proof.** Applying definition in equation (14) and commutative property of  $\odot$  proves the equation (15).  $\square$

## 2.2. Joint probability densities as tensor densities

**Theorem 1.** The joint probability density  $f_{\mu v \xi \dots}(i, j, k, \dots)$  for particles with common coordinate values  $x^\mu(i)$ ,  $x^v(j)$ ,  $x^\xi(k) \dots$  can be given by generalized inner product of the basis  $\{e_\mu(i), e_v(j), e_\xi(k), \dots\}$  via the equation:

$$f_{\mu\nu\zeta..}(i, j, k, ..) dV(i, j, k, ..) = \lim_{N \rightarrow \infty} \frac{1}{N} e_{\mu}(i) \odot e_{\nu}(j) \odot e_{\zeta}(k) .. \quad (17)$$

**Proof.** By using the equation (13) for basis  $\{e_{\mu}(i), e_{\nu}(j), e_{\zeta}(k), ..\}$  and by omitting the location indices  $i, j, k, ..$  we have:

$$e_{\mu} \odot e_{\nu} .. = \sum_{n=1}^m (e_{\mu})_n (e_{\nu})_n .. \quad (18)$$

Summation carried out on all particles. Since components of  $(e_{\mu})_n$  take two values 0 or 1; the non-vanishing terms are those with components that simultaneously take 1, and therefore right side sum of equation (17) reduces to the number of particles whose parameter values simultaneously are located in the intervals  $i, j, k, ..$  on corresponding  $x^{\mu}, x^{\nu}, x^{\zeta}, ..$  coordinates. This number is denoted as  $n_{\mu\nu\zeta..}$ . Normalization of  $n_{\mu\nu\zeta..}$  as defined in equation (12) by  $1/N$  yields the ratio of this number to total number of particles and consequently gives the joint probability  $f_{\mu\nu\zeta..}$  of particles with common coordinate values  $x^{\mu}(i), x^{\nu}(j) ..$

$$f_{\mu\nu\zeta..} dV(i, j, k, ..) = \frac{n_{\mu\nu\zeta..}}{N} = \frac{1}{N} e_{\mu}(i) \odot e_{\nu}(j) \odot e_{\zeta}(k) .. \quad (19)$$

□

The joint probability  $f_{\mu\nu\zeta..} dV(i, j, k, ..)$  and  $n_{\mu\nu\zeta..}$  are smooth functions at the limit  $N \rightarrow \infty$  and  $dV \rightarrow 0$ . The equation (18) reveals a specific configuration or state of related system for which there is a specific set of joint probabilities  $f_{\mu\nu\zeta..}$  that represents the exact state of it. Any configuration (state) of this system can be represented by such a specific set  $\{e_{\mu}(i), e_{\nu}(j), e_{\zeta}(k) ..\}$  for all points on manifold  $\mathcal{M}$  and vice versa. Evidently the order of indices  $\mu, \nu, \zeta ..$ , does not affect on the related joint probability density. The example of this symmetry could be found in symmetric properties of metric tensors  $g_{ij} = \langle e_i, e_j \rangle$  or  $g^{ij} = \langle e^{*i}, e^{*j} \rangle$  respect to their lower and upper indices respectively. We show the tensor properties of  $f_{\mu\nu\zeta..}$  in the next theorem.

**Lemma 4.**  $\frac{n_{\mu\nu\zeta..}}{N}$  are covariant tensors.

**Proof.** Let  $V_{\nu} = T_p^{\nu} \mathcal{N} \subset T_p \mathcal{N}$  and  $V^{*\nu} = T_p^{*\nu} \mathcal{N} \subset T_p^* \mathcal{N}$ . From equation (18) we have

$$\frac{n_{\mu\nu\zeta..}}{N} = \frac{1}{N} e_{\mu}(i) \odot e_{\nu}(j) \odot e_{\zeta}(k) .. \quad (20)$$

Since  $e_{\mu} \in V_{\mu}$ , the equation (19) is a map  $\phi$  from  $V_{\mu} \times V_{\nu} \times V_{\zeta} ..$  to  $\mathbb{R}$  or in a brief notation:

$$\phi : V^d \rightarrow \mathbb{R} \quad (21)$$

Respect to linear properties of  $\odot$  in equations (14)  $\phi$  is a multi-linear map with the following property:

$$\phi(e_1, .., ae_i + be_i, .., e_d) = e_1 \odot .. \odot ae_i \odot .. \odot e_d + e_1 \odot .. \odot be_i \odot .. \odot e_d \quad (22)$$

$$\phi(e_1, .., ae_i + be_i, .., e_d) = a\phi(e_1, .., e_i, .., e_d) + b\phi(e_1, .., e_i, .., e_d) \quad (23)$$

There is a one to one map between local tangent vectors  $\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, .., \frac{\partial}{\partial x^d}$  on  $\mathcal{M}$  and tangent vectors  $e_1, e_2, .., e_d$  on  $\mathcal{N}$  at any point  $P$ . Since  $\mathbb{Z}$  is the common field of both vector spaces, therefore the tangent spaces  $T_p \mathcal{N}$  and tangent space  $T_p \mathcal{M}$  are isomorphic. The term  $\frac{n_{\mu\nu\zeta..}}{N}$ , due to the axiom in section (1) and equation (12) is smooth and differentiable function and as multi-linear map defined in (19) is

covariant tensor at any point  $p \in \mathcal{M}$ , with the rank  $\leq d$ . Thus under the coordinate transformation  $x^\nu \rightarrow \bar{x}^\nu$  we have:

$$\frac{\bar{n}_{\mu\nu\zeta\dots}}{N} = \left( \frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\beta}{\partial \bar{x}^\nu} \frac{\partial x^\gamma}{\partial \bar{x}^\zeta} \dots \right) \frac{n_{\alpha\beta\gamma\dots}}{N} \quad (24)$$

□

**Theorem 2.** Joint probability density  $f_{\mu\nu\zeta\dots}$  defined in (18) is a covariant tensor density of weight -1.

**Proof.** With the coordinate transformation  $x^\nu \rightarrow \bar{x}^\nu$  equations (19), (20) and (24) gives:

$$\bar{f}_{\mu\nu\zeta\dots} d\bar{V} = \frac{\bar{n}_{\mu\nu\zeta\dots}}{N} = \left( \frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\beta}{\partial \bar{x}^\nu} \frac{\partial x^\gamma}{\partial \bar{x}^\zeta} \dots \right) \frac{n_{\alpha\beta\gamma\dots}}{N} \quad (25)$$

Respect to equation (19) we have:

$$\bar{f}_{\mu\nu\zeta\dots} d\bar{V} = \left( \frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\beta}{\partial \bar{x}^\nu} \frac{\partial x^\gamma}{\partial \bar{x}^\zeta} \dots \right) f_{\alpha\beta\gamma\dots} dV \quad (26)$$

Due to the definition of volume elements  $dV$  and  $d\bar{V}$  we obtain:

$$\bar{f}_{\mu\nu\zeta\dots} = |J|^{-1} \left( \frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\beta}{\partial \bar{x}^\nu} \frac{\partial x^\gamma}{\partial \bar{x}^\zeta} \dots \right) f_{\alpha\beta\gamma\dots} \quad (27)$$

□

Where  $|J|$  denoted as determinant of Jacobian. Respect to the definition of tensor density,  $\bar{f}_{\mu\nu\zeta\dots}$  in equation (26) is a tensor density of weight -1.

**Remark 3.** Considering the commutative property of  $\odot$  from equations (15), the tensor density  $f_{\mu\nu\zeta\dots}$  is symmetric respect to covariant indices  $\mu, \nu, \zeta \dots$  as expected for a joint probability. On the other hand repeated covariant indices reduces to not-repeated indices as we showed in Lemma 3:

$$f_{\alpha\beta\beta\gamma\dots} = f_{\alpha\beta\gamma\dots} \quad (28)$$

### 3. Joint Probability Densities as Symmetrized Covariant Derivatives of Cumulative Distribution Function

In the context of probability theory, the joint probability density of multiple (random) variables which are defined on a flat Euclidean space equipped with Cartesian coordinate, presented as a sequence of partial derivative of cumulative distribution function [1,2] and [3]:

$$f_{ijk\dots} = \frac{\partial^n F}{\partial x^i \partial x^j \partial x^k \dots} = (\partial_i \partial_j \partial_k \dots) F \quad (29)$$

Where  $F$  stands for cumulative distribution function (CDF) and  $\partial_i = \frac{\partial}{\partial x^i}$ . In this sequence of partial derivatives, the repetition of indices is not allowed. Obviously  $f_{ijk\dots}$  is not a tensor and does not meet the transformation requirement of tensors. Therefore in a flat Euclidean space with Cartesian coordinates  $(x^1, x^2, \dots)$ , joint probability density under coordinate transformation  $x^\nu \rightarrow \bar{x}^\nu$  will transform as:

$$\bar{f}(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^d) = |J|^{-1} f(x^1, x^2, \dots, x^d) \quad (30)$$

### 3.1. Taylor Expansion

$f_{\mu\nu\zeta\dots}$  as a covariant tensor is full symmetric respect to covariant indices and fulfils this property of joint probability densities. Joint probability densities for cumulative distribution function on Euclidean flat manifolds, are derived by regular partial derivatives respect to the contravariant coordinates  $x^i$ .

**Lemma 5.** *If a differentiable smooth scalar function  $\phi$  is defined on  $d$ -dimensional Euclidean space with Cartesian coordinates, then the Taylor expansion of  $\phi$  around a point  $(x_0^1, x_0^2, \dots, x_0^d)$  when for all  $\nu$ ,  $(x^\nu - x_0^\nu) \rightarrow 0$  will read as:*

$$\phi(x^1, x^2, \dots, x^d) = \phi(x_0^1, x_0^2, \dots, x_0^d) + \frac{\partial\phi}{\partial x^1}(x^1 - x_0^1) + \frac{\partial\phi}{\partial x^2}(x^2 - x_0^2) + \dots + \frac{\partial^d\phi}{\partial x^1\partial x^2\dots\partial x^d}(x^1 - x_0^1) \dots (x^d - x_0^d) \quad (31)$$

**Proof.** For a function  $\phi(x^1, x^2, \dots, x^d)$  defined on a  $d$ -dimensional space, at the limit  $(x^\nu - x_0^\nu) \rightarrow 0$  if the order of  $(x^1 - x_0^1) \dots (x^\nu - x_0^\nu)$  is higher than  $d$ , it will be negligible. Therefore the expansion at this limit will close at the  $d$ th order of the partial derivatives.  $\square$

Let  $\phi = F$  be the cumulative distribution function. Because  $(x^1 - x_0^1) \dots (x^d - x_0^d)$  in Cartesian coordinates is the volume element, its coefficient in the last term is joint probability density for  $d$  variables at  $x^\nu \rightarrow x_0^\nu$  as was shown in equation (29).

The Taylor expansion could be generalized to Taylor expansion on Riemannian manifold by means of *symmetrized covariant derivatives*. Symmetrization of multiple covariant derivative (symmetrized covariant derivative) could be accomplished by routine symmetrized form [11,12]:

$$T_{\mu\nu\dots\kappa} = \frac{1}{k!} \nabla_{(\mu} \nabla_{\nu} \dots \nabla_{\kappa)} F \quad (32)$$

This notation with bracket around indices is the abbreviation for the sum of all permutations over  $1 \leq k \leq d$  indices. In this sequence of covariant derivatives, the repetition of indices is not allowed. Obviously  $T_{\mu\nu\dots\kappa}$  is a tensor. In the Taylor expansion for  $F$  on a Riemannian manifold, the last symmetrized consecutive covariant derivatives is the coefficients of  $(x^1 - x_0^1) \dots (x^d - x_0^d)$  at the limit  $x^\nu \rightarrow x_0^\nu$  where the coordinates  $x^i$  are the local coordinates of Riemannian manifold  $\mathcal{M}$ . At this limit the last term reads as:

$$T_{\mu\nu\dots\kappa} dx^1 dx^2 \dots dx^d = \frac{1}{d!} \nabla_{(\mu} \nabla_{\nu} \dots \nabla_{\kappa)} F dx^1 dx^2 \dots dx^d \quad (33)$$

By multiplying and dividing to  $\sqrt{g}$  and taking into account the property  $\nabla_\mu g_{\nu\zeta} = 0$ , we get:

$$T_{\mu\nu\dots\kappa} dx^1 dx^2 \dots dx^d = \frac{1}{\sqrt{g}d!} \nabla_{(\mu} \nabla_{\nu} \dots \nabla_{\kappa)} F \sqrt{g} dx^1 dx^2 \dots dx^d \quad (34)$$

In this equation the last term  $F \sqrt{g} dx^1 dx^2 \dots dx^d$  is a scalar density which is invariant under integration on the possible domain and the term  $\sqrt{g} dx^1 dx^2 \dots dx^d$  is invariant volume element. Since  $\sqrt{g}$  is a tensor density of weight 1, the term :

$$\frac{1}{\sqrt{g}d!} \nabla_{(\mu} \nabla_{\nu} \dots \nabla_{\kappa)} F \quad (35)$$

is a tensor density of weight -1. Therefore respect to equation (27), and the fact that this term is the coefficient of invariant volume element, it equals to the joint probability density:

$$f_{\mu\nu\zeta\dots} = \frac{1}{\sqrt{g}d!} \nabla_{(\mu} \nabla_{\nu} \dots \nabla_{\kappa)} F \quad (36)$$

$f_{\mu\nu\dots\kappa}$  is symmetric respect to indices as expected for a joint probability density. Actually these covariant derivatives of any order remain as tensors and preserve the tensor properties of  $f_{\mu\nu\zeta\dots}$ . The

equation (34) after contraction by metric tensors  $g^{\alpha\mu}$  could be presented in the contravariant form. We define  $g^{\alpha\mu}\nabla_\mu = \nabla^\alpha$  as "contravariant derivative" [13]. Because  $\nabla_\xi g^{\alpha\mu} = 0$ , the covariant derivative  $\nabla_\xi$  and  $g^{\alpha\mu}$  are commutative, thus after multiplying both side of equation (36) by  $g^{\alpha\mu}g^{\beta\nu}..g^{\kappa\delta}$  we obtain:

$$f^{\alpha\beta.. \delta} = \frac{1}{\sqrt{g}d!} \nabla^{(\alpha} \nabla^{\beta} .. \nabla^{\delta)} F \quad (37)$$

This is the joint probability density defined in dual space of variables.

Regarding the definition of covariant derivative of the scalar  $F$ , for the first order covariant derivative we have:

$$\nabla_\nu F = \frac{\partial}{\partial x^\nu} F \quad (38)$$

Adding more symmetrized covariant derivatives leads to the terms that contains Christoffel symbols [14]:

$$\nabla_{(\mu} \nabla_{\nu)} F = \frac{1}{2} (\nabla_\mu \nabla_\nu + \nabla_\nu \nabla_\mu) F = \nabla_\mu \nabla_\nu F = \partial_\mu \partial_\nu F - \Gamma_{\mu\nu}^\xi \partial_\xi F \quad (39)$$

Where the identity  $\nabla_\mu \nabla_\nu F = \nabla_\nu \nabla_\mu F$  is used in equation (39).

**Lemma 6.** In a flat Euclidean manifold with Cartesian coordinates equation (34) reduces to (28).

**Proof.** In an Euclidean space with Cartesian coordinates all terms of  $\Gamma_{\mu\nu}^\xi$  vanish and  $\nabla_\mu$  reduces to  $\partial_\nu$ :

$$\nabla_{(\mu} \nabla_{\nu} .. \nabla_{\kappa)} = d! \partial_\mu \partial_\nu .. \partial_\kappa \quad (40)$$

Taking into account the identity  $g = 1$  for this case, the equation (33) becomes:

$$f_{\mu\nu.. \kappa} = \partial_\mu \partial_\nu .. \partial_\kappa F \quad (41)$$

□

Therefor the joint probability density in equation (28) is the special case of equation (33).

### 3.2. Equivalence of Symmetrized Covariant Derivative and Generalized Inner Product

In previous sections the joint probability densities have been derived by two distinct method 1) Symmetrized covariant derivatives on Riemannian manifold 2) Generalized inner product defined on "Particle Oriented coordinates". This conveys the idea of equivalence of these two operation on basis vectors of connected and discrete spaces respectively. Based on equations (18) and (34) this equivalence can be shown by these notations:

$$\begin{aligned} \nabla_\mu F &\sim e_\mu \\ \nabla_{(\mu} \nabla_{\nu)} F &\sim e_\mu \odot e_\nu \\ \nabla_{(\mu} \nabla_\nu \nabla_{\xi)} F &\sim e_\mu \odot e_\nu \odot e_\xi \end{aligned} \quad (42)$$

These relations suggests a new method to connect the concepts in discrete and continuous probability theory and a novel interpretation of covariant derivatives in differential geometry.

## 4. Examples

In this section some examples of the application of joint probability densities in physics and engineering are presented. Actually the joint probability densities  $f^{\mu\nu\tilde{\xi}..}$  and  $f_{\mu\nu\tilde{\xi}..}$  as tensor densities, remind us the physical tensors such as stress-energy tensor  $T^{\mu\nu}$  which stands for definitions of energy density, momentum density or energy flow density in Riemannian curved space-time manifold.

Therefore these type of tensors could exemplify a physical realization of joint probability densities of two variables on the Riemannian space. the following example reveals the physical interpretation of joint probability densities as tensor density in various field of physics.

#### 4.1. Example 1

Four-current density has many applications in engineering and its transformation under  $x^\nu \rightarrow \bar{x}^\nu$  on a manifold with metric tensor  $g_{\mu\nu}$  is represented as a four-vector density by [15]:

$$\bar{j}^\mu \equiv |J|^{-1} \frac{\partial \bar{x}^\mu}{\partial x^\alpha} j^\alpha \quad (43)$$

where  $j^\mu$  is a contravariant tensor density of rank -1.

#### 4.2. Example 2

The stress-energy tensor  $T^{ij}$ , with  $i, j \in \{0, 1, 2, 3\}$  that has been introduced in the context of general relativity, as a contravariant tensor of rank 2 [16],  $\mu, \nu \in \{1, 2, 3\}$ , comprises of five physically different components;  $T^{00}$  as energy density,  $T^{\mu 0}$  as energy flux density,  $T^{0\mu}$  as momentum density,  $T^{\mu\mu}$  as pressure and  $T^{\mu\nu}$  with  $\mu \neq \nu$  as shear stress density. Thus the Stress-energy tensor  $T^{\mu\nu}$  is a physical example of joint probability density tensor in a deterministic case of physics such as general relativity. This fact conveys the idea that many deterministic physical variables have fundamentally a probabilistic origin.

### 5. Conclusion

In this article a general form of joint probability density compatible with Riemannian manifold of variables has been introduced. It is shown that the joint probability densities on Riemannian manifold transforms as tensor densities of weight -1. Approach to these results facilitated by introducing a binary data matrix that collects the variables information of a system of particles on a lattice Euclidean space embedded by the particle oriented coordinates where the joint probability densities identified as a new definition of generalized (multi-linear) inner products of basis vectors. By this method the tensor density properties of joint probability is proved. Based on Taylor expansion of scalar field in a Riemannian manifolds, it has been shown that the symmetrized iterative covariant derivatives of cumulative probability function defined on Riemannian manifold, also gives the set of related joint probability densities equivalent to the generalized inner products method. As an outcome the equivalence of symmetrized iterative covariant derivatives and multi-linear inner product is proved. It has been shown the reduction of the generalized joint probability density to the usual form of iterative partial derivative of cumulative function in Euclidean space of variables with Cartesian coordinates. A new method to connect concepts in continuous and discrete probability theory and a novel interpretation of covariant derivative by generalized inner product has been proposed. Some examples of well-known physical tensors convey us that many deterministic physical variables may have fundamentally a probabilistic origin that by the future works on this subject will be more clarified.

**Author Contributions:** The author confirms sole responsibility for the study conception and manuscript preparation.

**Funding:** No funding was received for conducting this study.

**Conflicts of Interest:** The author declares that he has no conflict of interest.

### References

1. DasGupta, A. *Fundamentals of probability: A first course*; Springer Science & Business Media, 2010.
2. Robinson, E.A.; others. *Probability theory and applications*; Springer, 1985.
3. Srinivasan, S.K.; Mehata, K. *Probability and random processes. (No Title)* 1978.

4. Amari, S.i.; Armstrong, J. Curvature of Hessian manifolds. *Differential Geometry and its Applications* **2014**, *33*, 1–12.
5. Amari, S.i.; Nagaoka, H. *Methods of information geometry*; Vol. 191, American Mathematical Soc., 2000.
6. Hendriks, H. Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions. *The Annals of Statistics* **1990**, pp. 832–849.
7. Pennec, X. Probabilities and statistics on riemannian manifolds: A geometric approach. PhD thesis, INRIA, 2004.
8. Burbea, J.; Radhakrishna Rao, C. Differential metrics in probability spaces. *Probability and mathematical statistics* **1984**, *3*, 241–258.
9. Rao, C.R. On the use and interpretation of distance functions in statistics. *Bulletin of the International Statistical Institute* **1954**, *34*, 90–100.
10. Bhattacharya, A.; Bhattacharya, R. Statistics on Riemannian manifolds: asymptotic distribution and curvature. *Proceedings of the American Mathematical Society* **2008**, *136*, 2959–2967.
11. Sathiapalan, B. Gauge invariance and equations of motion for closed string modes. *Nuclear Physics B* **2014**, *889*, 261–298.
12. Tsirulev, A.N. Analytic continuation of tensor fields along geodesics by covariant Taylor series. *Theoretical and Mathematical Physics* **1995**, *102*, 245–250.
13. Synge, J.L.; Schild, A. *Tensor calculus*; Vol. 5, Courier Corporation, 1978.
14. Viaclovsky, J.A. Math 865, Topics in Riemannian Geometry, 2007.
15. Cabral, F.; Lobo, F.S. Electrodynamics and spacetime geometry: foundations. *Foundations of Physics* **2017**, *47*, 208–228.
16. Wald, R.M. *General relativity*; University of Chicago press, 2010.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.