A Sequence of Escort Distributions and Generalizations of Expectations on $q$-Exponential Family

Hiroshi Matsuzoe

Department of Computer Science and Engineering, Nagoya Institute of Technology, Nagoya 466-8555, Japan; matsuzoel@nitech.ac.jp; Tel.: +81-52-735-5143

Abstract: In the theory of complex systems, long tailed probability distributions are often discussed. For such a probability distribution, a deformed expectation with respect to an escort distribution is more useful than the standard expectation. In this paper, by generalizing such escort distributions, a sequence of escort distributions is introduced. As a consequence, it is shown that deformed expectations with respect to sequential escort distributions effectively work for anomalous statistics. In particular, it is shown that a Fisher metric on a $q$-exponential family can be obtained from the escort expectation with respect to the second escort distribution, and a cubic form (or an Amari–Chentsov tensor field, equivalently) is obtained from the escort expectation with respect to the third escort distribution.

Keywords: escort distribution; escort expectation; statistical manifold; deformed exponential family; Tsallis statistics; information geometry

MSC: 53A15; 53B50; 62F99; 94A14

1. Introduction

Long tailed probability distributions and their related probability distributions are important objects in anomalous statistical physics (cf. [1–3]). For such long tailed probability distributions, the standard expectation does not exist in general. Therefore, the notion of escort distribution has been introduced [4]. Since an escort distribution gives a suitable weight for tail probability, the escort expectation which is the expectation with respect to an escort distribution is more useful than the standard one.

In anomalous statistics, a deformed exponential function and a deformed logarithm function play essential roles. In fact, a deformed exponential family is an important statistical model in anomalous statistics. Such a statistical model is described by such a deformed exponential function. In particular, the set of all $q$-normal distributions (or Student’s $t$-distributions, equivalently) is a $q$-exponential family, which is described by a $q$-deformed exponential function [5] (see also [6,7]).

On the other hand, a generalized score function is defined from a deformed logarithm function. In the previous works, the author showed that a deformed score function is unbiased with respect to the escort expectation [8,9]. This implies that a deformed score function is regarded as an estimating function on a deformed exponential family. In addition, in information geometry, it is known that a deformed exponential family has a statistical manifold structure. Then a deformed score function is regarded as a tangent vector on this statistical manifold [6,10]. Therefore, properties of escort expectations are closely related to geometric structures on a deformed exponential family.

In this paper, we introduce a sequence of escort distributions, then we consider a sequential structure of escort expectations. It is known that a deformed exponential family naturally has at
least three kind of different statistical manifold structures [6,11]. Then we show that such statistical
manifold structures can be obtained from a sequential structure of escort expectations. In particular,
we show that a Fisher metric on a $q$-exponential family can be obtained from the deformed expectations
with respect to the second escort distribution, and a cubic form (or an Amari–Chentsov tensor field,
equivalently) is obtained from the deformed expectations with respect to the third escort distribution.

This paper is written based on the proceeding paper [7]. However, this paper focuses on
deformed expectations of a $q$-exponential family, whereas the previous paper focused on deformed
independences. We remark that several authors have been studying deformed expectations recently.
See [12,13], for example.

2. Deformed Exponential Families

In this paper, we assume that all objects are smooth for simplicity. Let us review preliminary facts
about deformed exponential functions and deformed exponential families. For more details, see [2,6],
for example. Historically, Tsallis [14] introduced the notion of $q$-exponential function and Naudts [5]
introduced the notion of $q$-exponential family together with a further generalization. Such a historical
note is provided in [2].

Let $R_{++}$ be the set of all positive real numbers, $R_{++} := \{ x \in R | x > 0 \}$. Let $\chi$ be a strictly
increasing function from $R_{++}$ to $R_{++} := \{ x \in R | x > 0 \}$. We define a $\chi$-logarithm function
or a deformed logarithm function by

$$\ln_{\chi} s := \int_{1}^{s} \frac{1}{\chi(t)} dt.$$ 

The inverse of $\ln_{\chi} s$ is called a $\chi$-exponential function or a deformed exponential function, which is
defined by

$$\exp_{\chi} t := 1 + \int_{0}^{t} u(s) ds,$$

where the function $u(s)$ is given by $u(\ln_{\chi} s) = \chi(s)$.

From now on, we suppose that $\chi$ is a power function, that is, $\chi(t) = t^{q}$. Then the deformed
logarithm and the deformed exponential are defined by

$$\ln_{q} s := \frac{s^{1-q} - 1}{1 - q}, \quad (s > 0),$$

$$\exp_{q} t := (1 + (1 - q)t)^{\frac{1}{1-q}}, \quad (1 + (1 - q)t > 0).$$

We say that $\ln_{q} s$ is a $q$-logarithm function and $\exp_{q} t$ is a $q$-exponential function. In this case,
the function $u(s)$ is given by

$$u(s) = (1 + (1 - q)s)^{\frac{q}{1-q}} = \{\exp_{q} s\}^{q}.$$ 

By taking a limit $q \to 1$, these functions coincide with the standard logarithm $\ln s$ and the standard
exponential $\exp t$, respectively.

A statistical model $S_q$ is called a $q$-exponential family if

$$S_q := \left\{ p(x, \theta) \Big| p(x; \theta) = \exp_{q} \left[ \sum_{i=1}^{n} \theta^{i} F_{i}(x) - \psi(\theta) \right], \theta \in \Theta \subset \mathbb{R}^{n} \right\},$$

where $F_{1}(x), \ldots, F_{n}(x)$ are functions on a sample space $\Omega$, $\theta = (\theta^{1}, \ldots, \theta^{n})$ is a parameter, and $\psi(\theta)$
is the normalization with respect to the parameter $\theta$. Under suitable conditions, $S_q$ is regarded as a
manifold with a local coordinate system $\{ \theta^{i} \}$. In this case, we call $\{ \theta^{i} \}$ a natural coordinate system.
In this paper, we focus on the \( q \)-exponential case. However, many results for the \( q \)-exponential family can be generalized for the \( \chi \)-exponential family (cf. [6,8]). We remark that a \( q \)-exponential family and a \( \chi \)-exponential family have further generalizations. See [15], for example.

**Example 1** (Student’s \( t \)-distribution (cf. [2,6,7])). Fix a number \( q \) \((1 < q < 1 + 2/d, \ d \in \mathbb{N})\), and set \( \nu = -d - 2/(1 - q) \). We define a \( d \)-dimensional Student’s \( t \)-distribution with degree of freedom \( \nu \) or a \( q \)-Gaussian distribution by

\[
p_{q}(x; \mu, \Sigma) := \frac{\Gamma \left( \frac{1}{q} \right)}{(\pi \nu)^{\frac{d}{2}} \Gamma \left( \frac{d}{2} \right) \sqrt{\det(\Sigma)}} \left[ 1 + \frac{1}{\nu} (x - \mu)^{T}(x - \mu) \right]^{-\frac{d}{2q} - \frac{1}{2}},
\]

where \( X = (X_1, \ldots, X_d) \) is a random vector on \( \mathbb{R}^d \), \( \mu = (\mu_1, \ldots, \mu_d) \) is a location vector on \( \mathbb{R}^d \) and \( \Sigma \) is a scale matrix on \( \text{Sym}^+ (d) \). For simplicity, we assume that \( \Sigma \) is invertible. Otherwise, we should choose a suitable basis \( \{ v^a \} \) on \( \text{Sym}^+ (d) \) such that \( \Sigma = \sum_a w_a v^a \). Then, the set of all Student’s \( t \)-distributions is a \( q \)-exponential family. In fact, setting parameters by

\[
z_q = \frac{(\pi \nu)^{\frac{d}{2}} \Gamma \left( \frac{d}{2} \right) \sqrt{\det(\Sigma)}}{\Gamma \left( \frac{1}{q} \right)}, \quad \tilde{\nu} = \frac{z_q^{-1}}{(1 - q)d + 2}, \quad \text{and} \quad \theta = 2\tilde{\nu} \mu,
\]

we have

\[
p_{q}(x; \mu, \Sigma) = \frac{1}{z_q} \left[ 1 + \frac{1}{\nu} (x - \mu)^{T}(x - \mu) \right]^{-\frac{d}{2q} - \frac{1}{2}} = \exp_q \left[ -\frac{1}{\nu} (x - \mu)^{T}(x - \mu) + \ln_q \frac{1}{z_q} \right] = \exp_q \left[ \sum_{i=1}^{d} \theta' x_i - \sum_{i=1}^{d} \tilde{\nu}_i x_i^2 - 2 \sum_{i<j} \tilde{\nu}_{ij} x_i x_j - \frac{1}{4} \theta \tilde{\nu}^{-1} \theta + \ln_q \frac{1}{z_q} \right].
\]

Since \( \theta \in \mathbb{R}^d \) and \( \tilde{\nu} \in \text{Sym}^+ (d) \), the set of all Student’s \( t \)-distributions is a \( d(d + 3)/2 \)-dimensional \( q \)-exponential family. The normalization \( \psi(\theta) \) is given by

\[
\psi(\theta) = \frac{1}{4} \theta \tilde{\nu}^{-1} \theta - \ln_q \frac{1}{z_q}.
\]

A univariate Student’s \( t \)-distribution is a well-known probability distribution in elementary statistics. We denote it by

\[
t_{\nu}(x; \mu, \sigma) := \frac{1}{Z_{q}} \exp_q \left[ -\frac{(x - \mu)^2}{(3 - q)\sigma^2} \right], \tag{3}
\]

where \( \mu \in \mathbb{R} \) is a location parameter, \( \sigma \in \mathbb{R}_+ \) is a scale parameter, and \( Z_{q} \) is the normalization defined by

\[
Z_{q} = \sqrt{\frac{3-q}{q-1}} \text{Beta} \left( \frac{3-q}{2(q-1)}, \frac{1}{2} \right) \sigma.
\]

In this case, the degree of freedom is \( \nu = (3 - q)/(q - 1) \). Conversely, the parameter \( q \) is give by

\[
q = \frac{\nu + 3}{\nu + 1}. \tag{4}
\]
3. Escort Distributions and Generalizations of Expectations

In anomalous statistics, a generalized expectation, called an escort expectation, is often discussed since the standard expectation does not exist in general (cf. [2,5,6]). In this section, we recall generalizations of expectations and introduce a sequential structure of escort distributions.

Let $S_q$ be a $q$-exponential family. For a given $p(x;\theta) \in S_q$ we define the $q$-escort distribution $P_q(x;\theta)$ of $p(x;\theta)$ and the normalized $q$-escort distribution $P_q^{\text{esc}}(x;\theta)$ by

$$P_q(x;\theta) := P_{q,1}(x;\theta) := \{p(x;\theta)\}^q,$$

$$P_q^{\text{esc}}(x;\theta) := \frac{1}{Z_q(p)}\{p(x;\theta)\}^q, \quad \text{where} \quad Z_q(p) = \int_\Omega \{p(x;\theta)\}^q dx,$$

respectively. For a $q$-exponential family $S_q = \{p_q(x;\theta)\}$, the set of normalized escort distributions $S_{q'} = \{P_q^{\text{esc}}(x;\theta)\}$ is a $q'$-exponential family with $q' = (2q - 1)/q$.

Example 2. Let $t_v(x;\mu,\sigma)$ be a univariate Student’s $t$-distribution with degree of freedom $v$. Then its normalized escort distribution is also a univariate Student’s $t$-distribution with degree of freedom $v + 2$. In fact, from Equation (4), a direct calculation shows that

$$q' = \frac{2q - 1}{q} = \frac{v + 5}{v + 3}.$$  

This implies that the degree of freedom $v' = v + 2$. Therefore, we obtain a sequence of escort distributions from a given Student’s $t$-distribution $t_v$:

$$t_v \rightarrow t_{v+2} \rightarrow t_{v+4} \rightarrow \cdots.$$  

This sequence is called a $\tau$-sequence, and the procedure to obtain from a given $t$-distribution to another $t$-distribution through an escort distribution is called the $\tau$-transformation [16].

For a given $p_q(x;\theta) \in S_q$, we can define the escort of an escort distribution

$$\bar{P}_q(x;\theta) := P_{q,2}(x;\theta) := q\{P_q(x;\theta)\}^{q'} = q\{p_q(x;\theta)\}^{2q-1}.$$  

We call $\bar{P}_q(x;\theta)$ the second escort distribution of $p_q(x;\theta)$. The coefficient $q$ before $\{p_q(x;\theta)\}^{2q-1}$ comes from considerations of $U$-information geometry [17]. We will discuss in the latter part of Section 5.

Similarly, we can define the $n$-th escort distribution $P_{q,(n)}(x;\theta)$ from the sequence of escort distributions:

$$P_{q,(1)}(x;\theta) := \{P_{q,2}(x;\theta)\} = \{q\{P_q(x;\theta)\}^{q'}\} = \{q\{p_q(x;\theta)\}^{2q-1}\}.$$  

Let $f(x)$ be a function on $\Omega$. The $q$-expectation $E_{q,p}[f(x)]$ and the normalized $q$-expectation $E_{q,p}^{\text{esc}}[f(x)]$ with respect to $p(x;\theta) \in S_q$ are defined by

$$E_{q,p}[f(x)] := \int_\Omega f(x)P_q(x;\theta)dx,$$

$$E_{q,p}^{\text{esc}}[f(x)] := \int_\Omega f(x)P_q^{\text{esc}}(x;\theta)dx,$$

respectively. We denote by $E_{q,p}[f(x)]$ the expectation with respect to the second escort distribution $\bar{P}_q(x;\theta)$, that is,

$$E_{q,p}[f(x)] := \int_\Omega f(x)\bar{P}_q(x;\theta)dx = q\int_\Omega f(x)\{p_q(x;\theta)\}^{2q-1}dx.$$
Since a differential of a power function is also a power function, we can give a characterization for escort distributions.

**Proposition 1.** Suppose that \( S_q \) is a \( q \)-exponential family defined by (1). Then the \( n \)-th escort distribution is given by the \( n \)-th differential of \( q \)-exponential function. That is, by setting \( u(t) = (\exp_q t)' \), we have the following formula:

\[
p_q(x; \theta) = \exp_q \left( \sum_{i=1}^{n} \theta^i F_i(x) - \psi(\theta) \right),
\]

\[
P_q(x; \theta) = P_q(x; \theta) = \exp_q \left( \sum_{i=1}^{n} \theta^i F_i(x) - \psi(\theta) \right),
\]

\[
\tilde{P}_q(x; \theta) = P_q(x; \theta) = \exp_q \left( \sum_{i=1}^{n} \theta^i F_i(x) - \psi(\theta) \right),
\]

\[
\vdots
\]

\[
\vdots
\]

\[
\vdots
\]

\[
P_{q,(n)}(x; \theta) = u^{(n-1)} \left( \sum_{i=1}^{n} \theta^i F_i(x) - \psi(\theta) \right),
\]

**Proof.** Since a \( q \)-exponential function is \( \exp_q(x) = (1 + (1 - q)x)^{1/(1-q)} \), its differential is given by

\[
u(x) = \frac{1-q}{1-q}(1 + (1 - q)x)^{\frac{1}{1-q} - 1} = (1 + (1 - q)x) \left( \frac{d}{dt} \exp_q x \right) = \{\exp_q x\}'^q.
\]

Therefore, we obtain \( P_q(x; \theta) = u \left( \sum_{i=1}^{n} \theta^i F_i(x) - \psi(\theta) \right) \).

By induction, the \( n \)-th differential of \( u(x) \) coincides with the \( n \)-th escort distribution \( P_{q,(n)} \), which is given by Equation (5).

4. **Statistical Manifolds and Their Generalized Conformal Structures**

In this section, we review the geometry of statistical manifolds. For more details about the geometry of statistical manifolds, see [18,19].

Let \( (S, g) \) be a Riemannian manifold and \( \nabla \) be a torsion-free affine connection on \( S \). We say that the triplet \( (S, \nabla, g) \) is a statistical manifold if \( \nabla g \) is totally symmetric. In this case, we can define a totally symmetric \((0,3)\)-tensor field by

\[
C(X, Y, Z) := (\nabla_X g)(Y, Z) = X g(Y, Z) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z),
\]

where \( X, Y \) and \( Z \) are arbitrary vector fields on \( S \). The tensor field \( C \) is called a cubic form or an Amari–Chentsov tensor field.

The notion of statistical manifold was introduced by Lauritzen [20]. He called the triplet \( (S, g, C) \) a statistical manifold. In this paper, the definition is followed to Kurose [18]. Though these two definitions are different, the other statistical manifold structure can be obtained from a given one. However, the motivation for the notion of conformal equivalence using \( (S, g, C) \) is different from that one using \( (S, \nabla, g) \), which we will discuss in the latter part of this section.

For a given statistical manifold \( (S, \nabla, g) \), we can define another torsion-free affine connection \( \nabla^* \) on \( S \) by

\[
Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla^*_X Z).
\]
The connection $\nabla^*$ is called the dual connection of $\nabla$ with respect to $g$. The triplet $(S, \nabla^*, g)$ is also a statistical manifold, which is called the dual statistical manifold of $(S, \nabla, g)$. The cubic form is given by the difference of two affine connections $\nabla^*$ and $\nabla$:

$$C(X,Y,Z) = g(\nabla^*_X Y - \nabla_X Y, Z).$$

We define generalized conformal structures for statistical manifolds followed to Kurose [18]. Two statistical manifolds $(S, \nabla, g)$ and $(S, \nabla, \bar{g})$ are said to be 1-conformally equivalent if there exists a function $\lambda : S \to \mathbb{R}_{++}$ such that

$$\bar{g}(X,Y) = \lambda g(X,Y),$$

$$\nabla_X Y = \nabla^*_X Y - g(X,Y)\text{grad}_g(\ln \lambda),$$

where grad$_g(\ln \lambda)$ is the gradient vector field of $\ln \lambda$ with respect to $g$, that is, $g(X, \ln \lambda) = X(\ln \lambda)$. We say that $(S, \nabla, g)$ is 1-conformally flat if $(S, \nabla, g)$ is locally 1-conformally equivalent to a flat statistical manifold.

Two statistical manifolds $(S, \nabla, g)$ and $(S, \nabla, \bar{g})$ are said to be $(−1)$-conformally equivalent if there exists a function $\lambda : S \to \mathbb{R}_{++}$ such that

$$\bar{g}(X,Y) = \lambda g(X,Y),$$

$$\nabla_X Y = \nabla^*_X Y + d(\ln \lambda)(Y)X + d(\ln \lambda)(X)Y,$$

where $d(\ln \lambda)(X) = X(\ln \lambda)$. If two statistical manifolds $(S, \nabla, g)$ and $(S, \nabla, \bar{g})$ are 1-conformally equivalent, then their dual statistical manifolds $(S, \nabla^*, g)$ and $(S, \nabla^*, \bar{g})$ are $(−1)$-conformally equivalent.

**Proposition 2.** If two statistical manifolds $(S, \nabla, g)$ and $(S, \nabla, \bar{g})$ are 1-conformally equivalent, then their cubic forms have the following relation:

$$\frac{1}{\lambda} \bar{C}(X,Y,Z) = C(X,Y,Z) + g(Y,Z)d(\ln \lambda)(X) + g(Z,X)d(\ln \lambda)(Y) + g(X,Y)d(\ln \lambda)(Z).$$

**Proof.** From Equations (7) and (8), we obtain

$$\nabla_X Y = \nabla^*_X Y + d(\ln \lambda)(Y)X + d(\ln \lambda)(X)Y + g(X,Y)\text{grad}_g(\ln \lambda).$$

By taking an inner product with respect to $g$, we obtain the result. ☐

5. **Statistical Manifold Structures on q-Exponential Families**

In this section, we consider statistical manifold structures on a $q$-exponential family. It is known that a $q$-exponential family naturally has at least three kinds of statistical manifold structures (cf. [6,8]). We reformulate these structures from the viewpoint of the sequence of escort distributions. In this paper, we omit the details about information geometry. See [21,22] for further details.

Firstly, we review basic facts about $q$-exponential family. Let $S_q$ be a $q$-exponential family. The normalization $\psi(\theta)$ on $S_q$ is convex, but may not be strictly convex. In fact, we obtain the following proposition.
Proposition 3. Let \( S_q = \{ p(x; \theta) \} \) be a \( q \)-exponential family. Then the normalization function \( \psi(\theta) \) is convex.

Proof. Let \( u(x) = (\exp_q x)\) and \( \partial_i = \partial / \partial \theta^i \). Then we have

\[
\begin{align*}
\partial_i p(x; \theta) &= u \left( \sum \theta^k F_k(x) - \psi(\theta) \right) (F_i(x) - \partial_i \psi(\theta)), \\
\partial_i \partial_j p(x; \theta) &= u' \left( \sum \theta^k F_k(x) - \psi(\theta) \right) (F_i(x) - \partial_i \psi(\theta))(F_j(x) - \partial_j \psi(\theta)) \\
&\quad - u \left( \sum \theta^k F_k(x) - \psi(\theta) \right) \partial_i \partial_j \psi(\theta).
\end{align*}
\]

(9)

Since \( \partial_i \int_{\Omega} p(x; \theta) dx = \int_{\Omega} \partial_i p(x; \theta) dx = 0 \) and \( \int_{\Omega} \partial_i \partial_j p(x; \theta) dx = 0 \), we have

\[
\begin{align*}
Z_q(p) &= \int_{\Omega} \left( \int_{\Omega} \left( \sum \theta^k F_k(x) - \psi(\theta) \right) dx \right) u(x) dx, \\
\partial_i \partial_j \psi(\theta) &= \frac{1}{Z_q(p)} \int_{\Omega} u' \left( \sum \theta^k F_k(x) - \psi(\theta) \right) (F_i(x) - \partial_i \psi(\theta))(F_j(x) - \partial_j \psi(\theta)) dx.
\end{align*}
\]

(10)

For an arbitrary vector \( c = (c_1, c_2, \ldots, c^n) \in \mathbb{R}^n \), since \( Z_q(p) > 0 \) and \( u''(x) > 0 \), we have

\[
\sum_{i,j=1}^n c^i c^j (\partial_i \partial_j \psi(\theta)) = \frac{1}{Z_q(p)} \int_{\Omega} u'' \left( \sum \theta^k F_k(x) - \psi(\theta) \right) \left( \sum_{i=1}^n c^i (F_i(x) - \partial_i \psi(\theta)) \right)^2 dx \geq 0.
\]

This implies that the Hessian matrix \( (\partial_i \partial_j \psi(\theta)) \) is semi-positive definite. \( \Box \)

We assume that \( \psi \) is strictly convex in this paper. Under this assumption, we can induce many geometric structures for a \( q \)-exponential family.

Since \( \psi \) is strictly convex, we can define a Riemannian metric and a cubic form by

\[
\begin{align*}
\alpha^q_i(\theta) &= \partial_i \partial_j \psi(\theta), \\
\beta^q_{ijk}(\theta) &= \partial_i \partial_j \partial_k \psi(\theta).
\end{align*}
\]

We call \( \alpha^q \) and \( \beta^q \) a \( q \)-Fisher metric and a \( q \)-cubic form, respectively [23,24]. Since \( \alpha^q \) is a Hessian of a function \( \psi \), \( \alpha^q \) is a Hessian metric, and \( \psi \) is the potential of \( \alpha^q \) with respect to the natural coordinate \( \{ \theta^i \} \) [25].

For a fixed real number \( \alpha \), set

\[
\alpha^q \left( \nabla^q_{X} Y, Z \right) := \alpha^q \left( \nabla^q_0 Y, Z \right) - \frac{\alpha}{2} \beta^q \left( X, Y, Z \right),
\]

(11)

where \( \nabla^q_0 \) is the Levi-Civita connection with respect to \( \alpha^q \). Since \( \alpha^q \) is a Hessian metric, from standard arguments in Hessian geometry [25], \( \nabla^q_{(c)} := \nabla^q_{(1)} \) and \( \nabla^q_{(m)} := \nabla^q_{(-1)} \) are flat affine connections and mutually dual with respect to \( \alpha^q \). Therefore, the triplets \( (S_q, \nabla^q_{(c)}, \alpha^q) \) and \( (S_q, \nabla^q_{(m)}, \alpha^q) \) are flat statistical manifolds, and the quadruplet \( (S_q, \alpha^q, \nabla^q_{(c)}, \nabla^q_{(m)}) \) is a dually flat space.

Under \( q \)-expectations, we have the following proposition (cf. [10]).

Proposition 4. For \( S_q \) a \( q \)-exponential family, (1) Set \( \eta_i = E_q^{\exp_q}[F_i(x)] \). Then \( \{ \eta_i \} \) is a \( \nabla^q_{(m)} \)-affine coordinate system such that

\[
\alpha^q \left( \frac{\partial}{\partial \theta^i}, \frac{\partial}{\partial \eta_j} \right) = \delta^i_j.
\]

(2) Set \( \phi(\eta) = E_q^{\exp_q}[\log_q p(x; \theta)] \), then \( \phi(\eta) \) is the potential of \( \alpha^q \) with respect to \( \{ \eta_i \} \).
Next, let us consider the standard Fisher metric and the standard cubic form. Suppose that $S := \{p(x; \theta)\}$ is a statistical model. Set $p_\theta := p(x; \theta)$, for simplicity. We define the (standard) Fisher metric $g^F$ on $S_\theta$ by

$$g^F_{ij}(\theta) := \int_{\Omega} (\partial_i \ln p_\theta)(\partial_j \ln p_\theta)p_\theta dx,$$

and the (standard) cubic form or the Amari–Chentsov vector field $C^F$ by

$$C^F_{ijk} (\theta) := \int_{\Omega} (\partial_i \ln p_\theta)(\partial_j \ln p_\theta)(\partial_k \ln p_\theta)p_\theta dx.$$

From similar arguments of (11), we can define an $\alpha$-connection $\nabla^{(\alpha)}$ on $S_\theta$, and we can obtain a statistical manifold structure $(S_\theta, \nabla^{(\alpha)}, g^F)$. In this case, $(S_\theta, \nabla^{(\alpha)}, g^F)$ is called an invariant statistical manifold [21,22].

A Fisher metric and a cubic form have the following representation using a sequence of escort distributions,

**Theorem 1.** Let $S_\theta$ be a $q$-exponential family. For $p(x; \theta) \in S_\theta$, suppose that $P_{q,(2)}(x; \theta)$ and $P_{q,(3)}(x; \theta)$ are the second and the third escort distribution of $p(x; \theta)$, respectively. Then the Fisher metric $g^F$ and the cubic form $C^F$ are given as follows:

$$\begin{align*}
g^F_{ij}(\theta) &= \frac{1}{q} \int_{\Omega} (\partial_i \ln p_\theta)(\partial_j \ln p_\theta) P_{q,(2)}(x; \theta) dx, \quad (12) \\
C^F_{ijk}(\theta) &= \frac{1}{q(2q - 1)} \int_{\Omega} (\partial_i \ln p_\theta)(\partial_j \ln p_\theta)(\partial_k \ln p_\theta) P_{q,(3)}(x; \theta) dx. \quad (13)
\end{align*}$$

**Proof.** Differentiating the $q$-logarithm, we have

$$\partial_i \ln p_\theta = \partial_i \left( p_\theta^{1-q} - 1 \right) = p_\theta^{-q} \partial_i p(\theta) = p_\theta^{-q} \partial_i \ln p(\theta).$$

Therefore, we obtain

$$\frac{1}{q} \int_{\Omega} (\partial_i \ln p_\theta)(\partial_j \ln p_\theta) P_{q,(2)}(x; \theta) dx = \int_{\Omega} p_\theta^{1-q}(\partial_i \ln p_\theta)(\partial_j \ln p_\theta) p_\theta^{2q-1}(x; \theta) dx$$

$$= \int_{\Omega} (\partial_i \ln p_\theta)(\partial_j \ln p_\theta) p_\theta(x; \theta) dx$$

$$= g^F_{ij}(\theta).$$

By a similar argument, we obtain the representation for $C^F$. $\square$

We define an $\alpha$-divergence $D^{(\alpha)}$ with $\alpha = 1 - 2q$ and a $q$-relative entropy (or a normalized Tsallis relative entropy) $D^T_q$ by

$$\begin{align*}
D^{(1-2q)}(p(x), r(x)) &= \frac{1}{q} E_{q,p} [\ln_q p(x) - \ln_q r(x)] \\
&= \frac{1 - \int_{\Omega} p(x)^q r(x)^{1-q} dx}{q(1-q)}, \quad (14)
\\
D^T_q(p(x), r(x)) &= E_{q,p}^\text{esc} [\ln_q p(x) - \ln_q r(x)] \\
&= \frac{1 - \int_{\Omega} p(x)^q r(x)^{1-q} dx}{(1-q)Z_q(p)}, \quad (15)
\end{align*}$$

respectively. It is known that the $\alpha$-divergence $D^{(1-2q)}(r, p)$ induces a statistical manifold structure $(S_\theta, g^F, \nabla^{(2q-1)})$, where $g^F$ is the Fisher metric on $S_\theta$ and $\nabla^{(2q-1)}$ is the $\alpha$-connection with $\alpha = 2q - 1$, and the $q$-relative entropy $D^T_q(r, p)$ induces $(S_\theta, g, \nabla^{q(e)})$. 
Theorem 2 (cf. [10, 24]). For a $q$-exponential family $S_q$, two statistical manifolds $(S_q, g^F, \nabla^{(2q-1)})$ and $(S_q, g, \nabla^{q(c)})$ are 1-conformally equivalent. In particular, an invariant statistical manifold $(S_q, g^F, \nabla^{(2q-1)})$ is 1-conformally flat. Riemannian metrics and cubic forms have the following relations:

$$
S^q_{ij}(\theta) = \frac{q}{Z_q(p)} g^q_{ij}(\theta),
$$

$$
C^q_{ijk}(\theta) = \frac{q}{Z_q(p)} ((2q-1)C^F_{ijk}(\theta) - \frac{q}{Z_q(p)} \{ g^F_{ij}(\theta) \partial_i \ln Z_q(p) + g^F_{jk}(\theta) \partial_j \ln Z_q(p) + g^F_{ki}(\theta) \partial_k \ln Z_q(p) \}).
$$

**Proof.** The results were essentially obtained in [10]. However, we give a simpler proof for Equations (16) and (17). The key idea is a sequence of escort distributions and the escort representations of $g^F$ and $C^F$ in Theorem 1.

From Equation (10), we directly obtain the conformal equivalence relation (16) using the escort representation of $g^F$ in (12).

By differentiating (9) and taking an integration, we obtain

$$
0 = \int_\Omega u'' \left( \sum \theta^i F_i(x) - \psi(\theta) \right) (F_i(x) - \partial_i \psi(\theta))(F_j(x) - \partial_j \psi(\theta))(F_k(x) - \partial_k \psi(\theta))dx - \int_\Omega u' \left( \sum \theta^i F_i(x) - \psi(\theta) \right) (F_k(x) - \partial_k \psi(\theta)) \partial_i \partial_j \psi(\theta)dx \\
- \int_\Omega u' \left( \sum \theta^i F_i(x) - \psi(\theta) \right) (F_i(x) - \partial_i \psi(\theta)) \partial_j \partial_k \psi(\theta)dx \\
- \int_\Omega u' \left( \sum \theta^i F_i(x) - \psi(\theta) \right) (F_j(x) - \partial_j \psi(\theta)) \partial_i \partial_k \psi(\theta)dx \\
- \int_\Omega u \left( \sum \theta^i F_i(x) - \psi(\theta) \right) \partial_i \partial_j \partial_k \psi(\theta)dx.
$$

Since $Z_q(p) = \int_\Omega P_q(x; \theta)dx$, we have

$$
\partial_i Z_q(p) = \partial_i \int_\Omega P_q(x; \theta)dx = \int_\Omega \partial_i P_q(x; \theta)dx = \int_\Omega \partial_i \ln Z_q(p)(F_i(x) - \partial_i \psi(\theta))dx.
$$

From the escort representation of $C^F$ in (13), and Proposition 1, we obtain Equation (17) since $g^q_{ij}(\theta) = \partial_i \partial_j \psi(\theta)$ and $C^q_{ijk}(\theta) = \partial_i \partial_j \partial_k \psi(\theta)$. 

We remark that the cubic form of $(S_q, g^F, \nabla^{(2q-1)})$ is not $C^F$ but $(2q-1)C^F$.

The difference of a $q$-divergence and a $q$-relative entropy is only the normalization $q/Z_q(p)$. This implies that a normalization for probability density imposes a generalized conformal change for a statistical model.

In the next part of this section, let us consider another statistical manifold on $S_q$ (cf. [6, 17, 26]).

Recall that a Fisher metric $g^F$ has the following representation:

$$
\bar{g}^F_{ij}(\theta) = \int_\Omega (\partial_i \ln p_\theta)(\partial_j p_\theta)dx.
$$

In information geometry, $\partial_i \ln p_\theta$ is called an $e$-representation (exponential representation) of $p_\theta$, and $\partial_i p_\theta$ is called a $m$-representation (mixture representation). Intuitively, $\partial_i \ln p_\theta$ and $\partial_i p_\theta$ are regarded as tangent vectors on a statistical model. Hence a Fisher metric is regarded as a $L^2$-inner product of $e$- and $m$-representations.

Let us generalize $e$- and $m$-representations for a $q$-exponential family. For $p_\theta \in S_q$, we call $\partial_i \ln q p_\theta$ a $q$-score function. Then we define a Riemannian metric $g^M$ by

$$
\bar{g}^M_{ij}(\theta) = \int_\Omega (\partial_i \ln q p_\theta)(\partial_j p_\theta)dx.
$$
By differentiating the above equation, we can define mutually dual torsion-free affine connections \( \nabla^{M(c)} \) and \( \nabla^{M(m)} \):

\[
\Gamma^{M(c)}_{ij,k}(\theta) := \int_{\Omega} (\partial_i \partial_j \ln q \theta)(\partial_k p \theta) dx, \\
\Gamma^{M(m)}_{ij,k}(\theta) := \int_{\Omega} (\partial_k \ln q \theta)(\partial_i \partial_j p \theta) dx,
\]

where \( \Gamma^{M(c)}_{ij,k} \) and \( \Gamma^{M(m)}_{ij,k} \) are the Christoffel symbols of \( \nabla^{M(c)} \) and \( \nabla^{M(m)} \) of the first kind, respectively. It is known that \( g^M \) is a Hessian metric, and the quadruplet \((S_q, g^M, \nabla^{M(c)}, \nabla^{M(m)})\) is a dually flat space. In addition, a natural parameter \( \{\theta^q\} \) is a \( \nabla^{M(c)} \)-affine coordinate system. Therefore, the cubic form for \((S_q, \nabla^{M(c)}, g^M)\) is

\[
C_{ijk}^M(\theta) = \Gamma^{M(m)}_{ij,k}(\theta). \tag{19}
\]

We remark that the statistical manifold structure \((S_q, \nabla^{M(c)}, g^M)\) is induced from a \( \beta \)-divergence \([17, 26]\) (or a density power divergence \([27]\)):

\[
D_{1-q}(p, r) := \int_{\Omega} \left\{ p(x)^{1-q} - r(x)^{1-q} - \frac{p(x)^{2-q} - r(x)^{2-q}}{2-q} \right\} dx. \tag{20}
\]

**Theorem 3.** For the statistical manifold structure \((S_q, \nabla^{M(c)}, g^M)\), the escort representations of the Riemannian metric \( g^M \) and the cubic form \( C^M \) are given as follows:

\[
g^M_{ij}(\theta) = \int_{\Omega} (\partial_i \ln q \theta)(\partial_j \ln q \theta) P_q(x; \theta) dx, \tag{21}
\]

\[
C_{ijk}^M(\theta) = \int_{\Omega} (\partial_i \ln q \theta)(\partial_j \ln q \theta)(\partial_k \ln q \theta) \tilde{P}_q(x; \theta) dx. \tag{22}
\]

**Proof.** For the Riemannian metric \( g^M \), since \( \partial_i p \theta = (\partial_i \ln q \theta) P_q(x; \theta) \), we immediately obtain Equation (21) from the definition of \( g^M \).

Let us consider the expression for cubic form (22). The q-score function \( \partial_i \ln q \theta \) is unbiased under the q-expectation. In fact,

\[
E_\theta[p_q[\partial_i \ln q \theta]] = \int_{\Omega} (\partial_i \ln q \theta) P_q(x; \theta) dx = \int_{\Omega} \partial_i p \theta dx = 0.
\]

From Equation (19), we obtain

\[
C_{ijk}^M(\theta) = \Gamma^{M(m)}_{ij,k}(x; \theta),
\]

\[
= \int_{\Omega} (\partial_k \ln q \theta)(\partial_i \partial_j p \theta) dx
\]

\[
= \int_{\Omega} (\partial_k \ln q \theta)(\partial_i \partial_j P_q(x; \theta)) dx
\]

\[
= -\partial_i \psi(\theta) \int_{\Omega} (\partial_k \ln q \theta) P_q(x; \theta) dx + \int_{\Omega} (\partial_k \ln q \theta)(\partial_j \ln q \theta) (\partial_i P_q(x; \theta)) dx
\]

\[
= \int_{\Omega} (\partial_k \ln q \theta)(\partial_j \ln q \theta)(\partial_i \ln q \theta) \tilde{P}_q(x; \theta) dx.
\]

\(\square\)

We remark that Naudts [5] gave another generalization of Fisher metric \( g^N \), which is defined by

\[
g^N_{ij}(\theta) := \int_{\Omega} \frac{1}{P_{q \theta}(x; \theta)} (\partial_i p \theta)(\partial_j p \theta) dx,
\]
The metric \( g^N \) is conformally equivalent to \( g^M \) with conformal factor \( Z_q(p_\theta) = \int_\Omega \{ p(x; \theta) \}^q dx \). That is, \( g^N(\theta) = Z_q(p_\theta) g^M(\theta) \). (See also [6]). Naudts gave a further generalization of Fisher metric and he showed a Cramér–Rao type bound theorem [5].

6. Concluding Remarks

In this paper, we introduced a sequence of escort distributions. Then we gave representations of Riemannian metrics and cubic forms from a viewpoint of the sequence of escort distributions.

In particular, we can define the following \((0,2)\)-tensor fields on a \(q\)-exponential family. For \( p_\theta \in S_q \), set \( \eta_i = \partial_i \psi(\theta) \).

1. From the standard expectation, we obtain

\[
\hat{g}^{(0)}_{ij}(\theta) := g_{ij}(\theta) := \int_\Omega (\partial_i \ln q_p)(\partial_j \ln q_p) p_\theta dx = E_p[(F_i(x) - \eta_i)(F_j(x) - \eta_j)].
\]

The tensor \( G \) is a covariant tensor. However, it may not be important in anomalous statistics.

2. From the \(q\)-expectation, we obtain

\[
\hat{g}^{(1)}_{ij}(\theta) := \sigma^M_{ij}(\theta) = \int_\Omega (\partial_i \ln q_p)(\partial_j \ln q_p) \{p_\theta\}^q dx
= \frac{1}{q} \int_\Omega (\partial_i \ln q_p)(\partial_j \ln q_p) P_q(x; \theta) dx
= \frac{1}{q} E_{q,p}[(F_i(x) - \eta_i)(F_j(x) - \eta_j)].
\]

The Riemannian metric \( g^M \) is a Hessian metric, and it is induced from the \(\beta\)-divergence (20).

3. From the expectation with respect to the second escort distribution, we obtain

\[
\hat{g}^{(2)}_{ij}(\theta) := \hat{g}^F_{ij}(\theta) = \int_\Omega (\partial_i \ln q_p)(\partial_j \ln q_p) \{p_\theta\}^{2q-1} dx
= \frac{1}{q} \int_\Omega (\partial_i \ln q_p)(\partial_j \ln q_p) P_q(x; \theta) dx
= \frac{1}{q} E_{q,2,p}[(F_i(x) - \eta_i)(F_j(x) - \eta_j)].
\]

The Riemannian metric \( g^F \) is a Fisher metric. Hence \( g^F \) is invariant to the choice of reference measure on \( \Omega \), but it is not a Hessian metric. In addition, \( g^F \) is induced from the \(\alpha\)-divergence (14). The conformal Riemannian metric \( g^\theta \) is a \(q\)-Fisher metric. It is a Hessian metric, and it is induced from a normalized Tsallis relative entropy (15).

We may define a Riemannian metric and a cubic form from higher order escort expectations:

\[
\hat{g}^{(n)}_{ij}(\theta) := \int_\Omega (\partial_i \ln q_p)(\partial_j \ln q_p) p^{(n)}_{q,\theta}(x; \theta) dx,
\]

\[
\hat{C}^{(n)}_{ij}(\theta) := \int_\Omega (\partial_i \ln q_p)(\partial_j \ln q_p)(\partial_k \ln q_p) p^{(n+1)}_{q,\theta}(x; \theta) dx.
\]

Then we obtain a sequence of statistical manifold structures.

\((S_q, \hat{g}^{(1)}, C^{(1)}) \rightarrow (S_q, \hat{g}^{(2)}, C^{(2)}) \rightarrow \cdots \rightarrow (S_q, \hat{g}^{(n)}, C^{(n)}) \rightarrow \cdots\)

However, the geometric meaning of this sequence is not clear at this moment. Elucidating geometric properties of this sequence is a future problem.
Acknowledgments: This research was partially supported by JSPS (Japan Society for the Promotion of Science), KAKENHI (Grants-in-Aid for Scientific Research) Grant Numbers JP26108003 and JP15K04842.

Conflicts of Interest: The author declares no conflict of interest.

References

