Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 51 (5) (2022), 1472 – 1483
DOI : 10.15672/hujms.1094273

Research Article

Generating function for generalized Fisher information measure and its application to finite mixture models

Omid Kharazmi*[1], Narayanaswamy Balakrishnan[2]

[1] *Department of Statistics, Faculty of Mathematical Sciences, Vali-e-Asr University of Rafsanjan, Iran*
[2] *Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada*

Abstract

In this work, we consider generating function for generalized Fisher information measure and use it to develop some results for this measure. Next, we study generalized Fisher information for the mixing parameter vector of a finite mixture density function and develop some results for this model. Further, we propose a Jensen-type divergence measure, namely, Jensen-generalized Fisher information (JGFI), and establish some properties for this measure and its generating function. Finally, for illustrative purposes, we examine a real example from image processing and provide some numerical results in terms of JGFI measure.

Mathematics Subject Classification (2020). 94A15, 62B10

Keywords. Fisher information, generalized Fisher information, generating function, Jensen-Fisher information, finite mixture distributions

1. Introduction

In information theory, several measures have been introduced for capturing the information content of a probabilistic model. There are generally two main classes of information measures, namely, entropy-type and Fisher-type, in the literature. Among these classes, Shannon entropy and Fisher information are the most important information measures, and these have been used rather extensively. Shannon entropy originated from the pioneering work of [14], based on a study of the behavior of systems described by probability density (or mass) functions. Nearly two decades earlier, Fisher [6] had proposed another information measure for describing the interior properties of a probabilistic model, which has since become vital to likelihood-based inferential methods. Fisher information, as well as Shannon entropy, have found many key applications including in statistical inference, physics, thermodynamics and information theory. It is possible for complex systems to be completely described by means of their behavior (Shannon) and their architecture (Fisher) information measures. One may refer to [1] and [15] for some discussion in this regard.

*Corresponding Author.
Email addresses: omid.kharazmi@vru.ac.ir (O. Kharazmi), bala@mcmaster.ca (N. Balakrishnan)
Received: 28.03.2022; Accepted: 07.06.2022

In information theory, generating functions have been defined for probability densities to determine information quantities such as Shannon information and Kullback-Leibler divergence; see [7, 8]. Recently, through the use of information generating function, Kharazmi and Balakrishnan [9] have proposed Jensen-information generating (JIG) function and discussed in detail its connections with other well-known information measures such as Kullback-Leibler, Jensen-Shannon, Jensen-extropy and Jensen-Taneja information divergences.

Papaioannou et al. [11] proposed generating function for Fisher information of a probability density function $f_\theta(x)$ with $\theta \in \Theta \subseteq R$, whose second derivative, evaluated at 0, gives the well-known Fisher information measure about parameter θ. Let X be a continuous random variable on the support $\mathcal{X}$ with density function $f_\theta(x)$ and score function $\rho_\theta(x) = \frac{\partial \log f_\theta(x)}{\partial \theta}$. Then, the Fisher information generating (FIG) function of density $f_\theta(x)$ (or parameter θ), for any $\alpha > 0$, is defined as

$$G_{FI}(\theta,\alpha) = G_{FI}(f_\theta,\alpha) = \int_{\mathcal{X}} \mathrm{e}^{\alpha\rho_\theta(x)} f_\theta(x)dx, \tag{1.1}$$

provided the integral exists. For simplicity in notation, the integration domain will be omitted through out the paper unless there is a need to explicitly specify it. Papaioannou et al. [11] then showed the following properties of $G_{FI}(\theta,\alpha)$ in (1.1):

$$(i)\quad G_{FI}(\theta,0) = 1;$$
$$(ii)\quad \frac{\partial^2 G_{FI}(\theta,\alpha)}{\partial\alpha^2}\bigg|_{\alpha=0} = I(\theta),$$

where $I(\theta)$ is the Fisher information about parameter θ defined as

$$I(\theta) = \int \Big\{\rho_\theta(x)\Big\}^2 f_\theta(x)dx. \tag{1.2}$$

There is another kind of Fisher information, known as Fisher information of the density itself. Let X be a continuous random variable with density function f and $\rho(x) = \frac{\partial \log f(x)}{\partial x}$. Then, the Fisher information of density f itself is defined as

$$I(f) = \int \Big\{\rho(x)\Big\}^2 f(x)dx. \tag{1.3}$$

Incidentally, the Fisher information measures in (1.2) and (1.3) are identical when θ is a location parameter, or equivalently, when the density f belongs to the location family of distributions.

Following the work of [6], considerable attention has been paid to providing extensions of the Fisher information in (1.2) and the Fisher information of density function in (1.3). Jensen-Fisher (JF) information divergence and generalized Fisher information are two such important extensions of Fisher information; see [3, 4, 13]. In particular, Bercher [3] introduced a k-generalized Fisher information (generalized Fisher information) as an extension of (1.2), for $k > 0$, in the form

$$I_k(\theta) = \int |\rho_\theta(x)|^k f_\theta(x)dx. \tag{1.4}$$

Bercher [3] then showed that this measure fits well in the context of non-extensive thermostatistics and provided some results in this connection. Recently, Bobkov [4] considered the density version of $I_k(f_\theta)$ in (1.4) of the form

$$I_k(f) = \int |\rho(x)|^k f(x)dx, \tag{1.5}$$

where $\rho(x) = \frac{f'(x)}{f(x)}$ is the score function corresponding to the density f. This is also a generalization of the Fisher information in (1.3). Bobkov [4] then presented some interesting properties of this generalized Fisher information measure $I_k(f)$, and specifically showed that this measure is convex with respect to the density function f.

The purpose of the present work is first to define generating functions for the two types of generalized Fisher information and then establish some interesting properties of these functions, and secondly to study generalized Fisher information for the mixing parameter of a general finite-mixture density function.

Let $f_1, ..., f_n$ be n continuous density functions. Then, a finite mixture density, with mixing parameter vector $\boldsymbol{\theta} = (\theta_1, ..., \theta_{n-1})$, for $n \geq 2$, is given by

$$f_{\boldsymbol{\theta}}(x) = \frac{1}{n-1}\sum_{j=1}^{n-1}\theta_j f_j(x) + \left(1 - \frac{\sum_{j=1}^{n-1}\theta_j}{n-1}\right) f_n(x), \tag{1.6}$$

where $0 \leq \theta_i \leq 1,\ i = 1, ..., n-1$. In the special cases of $n = 2$ and $n = 3$ in (1.6), the 2-component and 3-component mixture densities are deduced with corresponding densities as

$$f_\theta(x) = \theta f_1(x) + (1-\theta) f_2(x),\ 0 \leq \theta \leq 1 \tag{1.7}$$

and

$$f_\theta(x) = \frac{\theta_1}{2} f_1(x) + \frac{\theta_2}{2} f_2(x) + \left(1 - \frac{\theta_1 + \theta_2}{2}\right) f_3(x),\ 0 \leq \theta_i \leq 1,\ i = 1, 2, \tag{1.8}$$

respectively. In the later part of this work, we first propose Jensen-generalized Fisher information based on $I_k(f)$ in (1.5) and then present generating function of this information measure. With regard to these Jensen measures and connections between them, some results are also established.

The rest of this paper is organized as follows. In Section 2, we introduce generating function of generalized Fisher information about a parameter θ and establish some new properties for it. We show that this can be expressed in terms of different orders of generalized Fisher information measures. In Section 3, we consider the finite mixture density function in (1.6) and establish some results for the generalized Fisher information measure of the mixing parameter vector. We specifically show that the Fisher information of the mixing parameter vector is connected to the higher order chi-square divergence, Pearson-Vajda χ^k divergence. Next, in Section 4, Jensen-generalized Fisher information measure and its generating function are discussed. We show that the latter generates different orders of the Jensen-generalized Fisher information measure. In Section 5, we consider a real example on image processing and present some numerical results in terms of the JGFI measure. Finally, we present some concluding remarks in Section 6.

2. Generating function of generalized Fisher information

Inspired by the work of [11], we now introduce the generating function for the generalized Fisher information measure in (1.4), which assists in defining different orders of generalized Fisher information. This is achieved through repeated derivatives of the generating function.

Definition 2.1. Let X be a continuous random variable on the support $\mathcal{X}$ with density function $f_\theta(x)$ and score function $\rho_\theta(x) = \frac{\partial \log f_\theta(x)}{\partial \theta}$. Then, the generating function of generalized Fisher information (GFGFI) of density $f_\theta(x)$ (or parameter θ), for any $\alpha > 0$, is defined as

$$G_{GFI}(\theta, \alpha) \equiv G_{GFI}(f_\theta, \alpha) = \int_{\mathcal{X}} e^{\alpha|\rho_\theta(x)|} f_\theta(x) dx, \tag{2.1}$$

provided the integral exists.

From Definition 2.1, we obtain the following lemma.

Lemma 2.2. *Suppose the random variable X has density function $f_\theta(x)$. Then, a series representation for the GFGFI in (2.1) is given by*

$$G_{GFI}(\theta,\alpha)=\sum_{k=0}^{\infty}\frac{\alpha^k}{k!}I_k(\theta), \tag{2.2}$$

where $I_k(\theta)$ is the generalized Fisher information of order k presented in (1.4).

Proof. From the definition of GFGFI measure in (2.1) and by making use of Maclaurin expansion and Fubini's theorem, we get

$$\begin{aligned} G_{GFI}(\theta,\alpha) &= E\left[\mathrm{e}^{\alpha|\rho_\theta(X)|}\right] \\ &= \sum_{k=0}^{\infty}\frac{\alpha^k}{k!}\int\left|\rho_\theta(x)\right|^k f_\theta(x)dx \\ &= \sum_{k=0}^{\infty}\frac{\alpha^k}{k!}I_k(\theta), \end{aligned}$$

as required. □

Now, let N be a random variable having Poisson distribution with mean α. Then, from Lemma 2.2, an alternative representation for the GFGFI is given by

$$G_{GFI}(\theta,\alpha)=\sum_{k=0}^{\infty}\frac{\alpha^k}{k!}I_k(\theta)=\mathrm{e}^{\alpha}\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\alpha}\alpha^k}{k!}I_k(\theta)=\mathrm{e}^{\alpha}E_N\left[I_N(\theta)\right],$$

where $I_N(\theta)$ is the generalized Fisher information of order N as given in (1.4).

Example 2.3. Let X be an exponential variable with probability density function (PDF) $f(x)=\lambda\mathrm{e}^{-\lambda x},\ \lambda>0,\ x>0$. From (2.1), for $\alpha\leq\lambda$, the GFGFI does not exist, but for $\lambda>\alpha$, we have

$$G_{GFI}(\lambda,\alpha)=\frac{\lambda\mathrm{e}^{\frac{\alpha}{\lambda}}}{\alpha+\lambda}(1-\mathrm{e}^{-\frac{\alpha+\lambda}{\lambda}})+\frac{\lambda\mathrm{e}^{-\frac{\alpha}{\lambda}}}{\lambda-\alpha}\mathrm{e}^{-\frac{\lambda-\alpha}{\lambda}}.$$

From this expression, we find

$$\frac{\partial}{\partial\alpha}G_{GFI}(\lambda,\alpha) = \frac{\alpha e^{\frac{\alpha}{\lambda}}+\frac{\lambda}{e}}{(\alpha+\lambda)^2}+\frac{\lambda}{e(\lambda-\alpha)^2},$$

$$\frac{\partial^2}{\partial\alpha^2}G_{GFI}(\lambda,\alpha) = e^{\frac{\alpha}{\lambda}}\frac{\lambda^2+\alpha^2}{\lambda(\lambda+\alpha)^3}+4\lambda\alpha\frac{3\lambda^2+\alpha^2}{e(\lambda^2-\alpha^2)^3}.$$

Thence, we have

$$\frac{\partial^2}{\partial\alpha^2}G_{GFI}(\lambda,\alpha)\big|_{\alpha=0}=\frac{1}{\lambda^2}=I(\lambda).$$

A 3D-plot of this $G_{GFI}(\lambda,\alpha)$ measure, for $\lambda\in(2,3.5)$ and $\alpha\in(0,1.5)$, is presented in Figure 1.

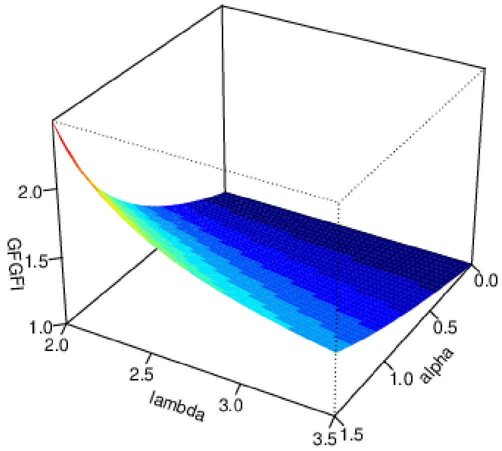

Figure 1. 3D-plot of the GFGFI in Example 2.3.

It is seen from Figure 1 that the GFGFI is increasing with respect to both parameters α and λ.

3. Generalized Fisher information measure for finite mixture distributions

In this section, we study the generalized Fisher information for parameter $\boldsymbol{\theta}$ of a finite mixture distribution. We then focus on 2-component and 3-component mixture densities as particular cases. For this purpose, we first consider the definition of a higher-order chi-square divergence. Let f and g be two density functions on the common support. Then, the well-known Pearson-Vajda χ^k divergence between f and g is defined as

$$\chi^k(f,g) = \int \frac{|f(x) - g(x)|^k}{f^{k-1}(x)} dx.$$

In an analogous manner, we can define $\chi^k(g,f)$. For more details, see [10]. Then, the following theorem gives a representation for the generalized Fisher information in (1.4), denoted by $I_k(\theta_i), i = 1, ..., n-1$, for the n-component finite mixture model in (1.6).

Theorem 3.1. *The generalized Fisher information measure of the mixture PDF in (1.6) about parameter $\theta_i, i = 1, ..., n-1$, is given by*

$$I_k(\theta_i) = \frac{1}{|\theta_i - (n-1)|^k} \chi^k(f_{\boldsymbol{\theta}}, f_{\boldsymbol{\theta}_{-i}}), \;\; i = 1, ..., n-1, \tag{3.1}$$

where $\chi^k(f_{\boldsymbol{\theta}}, f_{\boldsymbol{\theta}_{-i}})$ is the Pearson-Vajda divergence between density functions $f_{\boldsymbol{\theta}}$ and $f_{\boldsymbol{\theta}_{-i}}$, with

$$f_{\boldsymbol{\theta}_{-i}}(x) = \frac{n-2}{n-1} f_i(x) + \frac{1}{n-1} \sum_{j=1, j\neq i}^{n-1} \theta_j f_j(x) + \frac{1}{n-1}\Big(1 - \sum_{j=1, j\neq i}^{n-1} \theta_j\Big) f_n(x),$$

and $\boldsymbol{\theta}_{-i} = (\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_{n-1})$ is the vector obtained by removing the i-th component of $\boldsymbol{\theta}$.

Proof. From the definition of the generalized Fisher information in (1.4), for $i = 1, ..., n-1$, we have

$$\begin{aligned} I_k(\theta_i) &= E\left|\frac{\partial \log f_{\boldsymbol{\theta}}(X)}{\partial \theta_i}\right|^k \\ &= \frac{1}{(n-1)^k}\int \frac{\left|f_i(x) - f_n(x)\right|^k}{f_{\boldsymbol{\theta}}^k(x)} f_{\boldsymbol{\theta}}(x)dx \\ &= \frac{1}{\left|\theta_i - (n-2)\right|^k}\int \frac{\left|f_{\boldsymbol{\theta}_{-i}}(x) - f_{\boldsymbol{\theta}}(x)\right|^k}{f_{\boldsymbol{\theta}}^{k-1}(x)}dx \\ &= \frac{1}{\left|\theta_i - (n-2)\right|^k}\chi^k(f_{\boldsymbol{\theta}}, f_{\boldsymbol{\theta}_{-i}}), \end{aligned} \tag{3.2}$$

where the third equality follows from the fact that, for $i = 1, ..., n-1$,

$$f_i(x) - f_n(x) = \frac{n-1}{\theta_i - (n-2)}(f_{\boldsymbol{\theta}}(x) - f_{\boldsymbol{\theta}_{-i}}(x)). \tag{3.3}$$

□

From Theorem 3.1, we can see that the GFGFI measure is related to the Pearson-Vajda divergence measure as a higher-order chi-square divergence. Let us now consider the 2-component and 3-component mixture distributions as particular cases of the general mixture distribution in (1.6). Upon setting $n = 2$ and $n = 3$, the GFI measures of these sub-models are obtained as

$$I_k(\theta) = \frac{1}{|\theta - 1|^k}\chi^k(f_\theta, f_1)$$

and

$$I_k(\theta_i) = \frac{1}{\left|\theta_i - 2\right|^k}\chi^k(f_{\boldsymbol{\theta}}, f_{\boldsymbol{\theta}_{-i}}),\ i = 1, 2,$$

respectively, where

$$f_{\boldsymbol{\theta}_{-i}}(x) = \frac{1}{2}f_i(x) + \frac{1}{2}\sum_{j=1, j\neq i}^{2}\theta_j f_j(x) + \frac{1}{2}\Big(1 - \sum_{j=1, j\neq i}^{2}\theta_j\Big)f_3(x),\ i = 1, 2,$$

and

$$\boldsymbol{\theta}_{-i} = \begin{cases} \theta_2, & i = 1, \\ \theta_1, & i = 2. \end{cases}$$

Theorem 3.2. *The generating function of the generalized Fisher information measure of the mixture PDF in (1.6) about parameter $\theta_i, i = 1, ..., n-1$, is given by*

$$G_{GFI}(\theta_i, \alpha) = \sum_{k=0}^{\infty}\frac{\alpha^k}{k!\left|\theta_i - (n-1)\right|^k}\chi^k(f_{\boldsymbol{\theta}}, f_{\boldsymbol{\theta}_{-i}}),$$

where $\chi^k(\cdot, \cdot)$ is the Pearson-Vajda divergence measure of order k.

Proof. From Lemma 2.2 and Theorem 3.1, we obtain

$$\begin{aligned} G_{GFI}(\theta_i, \alpha) &= \sum_{k=0}^{\infty}\frac{\alpha^k}{k!}I_k(\theta_i) \\ &= \sum_{k=0}^{\infty}\frac{\alpha^k}{k!}\frac{1}{\left|\theta_i - (n-1)\right|^k}\chi^k(f_{\boldsymbol{\theta}}, f_{\boldsymbol{\theta}_{-i}}), \end{aligned}$$

as required. □

4. Jensen-generalized Fisher information measure and its generating function

We first propose the Jensen-generalized Fisher information based on the generalized Fisher information measure of density itself, and then introduce generating function of this measure. The connection between these two new information measures is then discussed.

4.1. Jensen-generalized Fisher information

Let $X_1, ..., X_n$ be random variables with density functions $f_1, ..., f_n$, respectively, and $\theta_1, ..., \theta_n$ be non-negative real numbers such that $\sum_{i=1}^{n} \theta_i = 1$. Then, the Jensen-generalized Fisher information (JGFI) measure of order k is defined as

$$JGFI_k(f_1, ..., f_n, \boldsymbol{\theta}) = \sum_{i=1}^{n} \theta_i I_k(f_i) - I_k\Big(\sum_{i=1}^{n} \theta_i f_i\Big). \tag{4.1}$$

Due to the convexity property of $I_k(f)$, it is easy to show that the JGFI measure in (4.1) is non-negative; see [4] about the convexity of $I_k(f)$ measure.

Remark 4.1. From (4.1), in the special case when $k = 2$, we obtain Jensen-Fisher information proposed originally by [13].

4.2. Generating function of Jensen-generalized Fisher information measure

We now introduce the generating function of the generalized Fisher information, which is based on the density function. Then, we present the generating function of Jensen-generalized Fisher information measure and discuss a connection between them.

Definition 4.2. Let X be a continuous random variable with density function $f(x)$ and score function $\rho(x) = \frac{\partial \log f(x)}{\partial x}$. Then, the generating function of the generalized Fisher information (GFGFI) of density $f(x)$ itself, for any $\alpha > 0$, is defined as

$$G_{GFI}(f, \alpha) = \int \mathrm{e}^{\alpha|\rho(x)|} f(x) dx, \tag{4.2}$$

provided the integral exists.

As in Lemma 2.2, the above generating function of the generalized Fisher information measure can be expressed as

$$G_{GFI}(f, \alpha) = \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} I_k(f),$$

where $I_k(f)$ is the generalized Fisher information of order k in (1.5).

As mentioned earlier, the quantity $I_k(f)$ provides information about the density f itself; for pertinent details, see Bobkov [4]. It is worthwhile to note that when θ is a location parameter, that is, $f_\theta(x) = f(x - \theta)$, under some regularity conditions, we have $\partial f(x-\theta)/\partial\theta = -\partial f(x-\theta)/\partial x$, and so the generalized Fisher information $I_k(\theta)$ in (1.4) and the generalized Fisher information of the density function $I_k(f)$ in (1.5) become identical.

Analogous to Definition 4.2, we now propose the generating function of the Jensen-generalized Fisher information measure.

Definition 4.3. Let the variables $X_1, ..., X_n$ have density functions $f_1, ..., f_n$, respectively. Then, the generating function of Jensen-generalized Fisher information (GFJGFI) is defined as

$$G_{JGFI,\alpha}(f_1, ..., f_n; \boldsymbol{\theta}) = \sum_{i=1}^{n} \theta_i G_{GFI}(f_i, \alpha) - G_{GFI}\Big(\sum_{i=1}^{n} \theta_i f_i, \alpha\Big),$$

where $\boldsymbol{\theta} = (\theta_1, ..., \theta_n)$, with $\theta_i \geq 0$ and $\sum_{i=1}^{n} \theta_i = 1$.

In Corollary 4.5 below, we show that GFJGFI is non-negative, and before that, we present the following theorem.

Theorem 4.4. *Let the variables $X_1, ..., X_n$ have density functions $f_1, ..., f_n$, respectively. Then, the GFJGFI is an infinite mixture of scaled $JGFI_k(f_1, ..., f_n, \boldsymbol{\theta})$ measures in (4.1), given by*

$$G_{JGFI,\alpha}(f_1, ..., f_n; \boldsymbol{\theta}) = \sum_{k=0}^{\infty} p_\alpha(k) D_k^\alpha(f_1, ..., f_n; \boldsymbol{\theta}),$$

where $D_k^\alpha(f_1, ..., f_n; \boldsymbol{\theta}) = \frac{JGFI_k(f_1,...,f_n;\boldsymbol{\theta})}{e^{-\alpha}}$, $p_\alpha(k) = \frac{e^{-\alpha}\alpha^k}{k!}$, and $\boldsymbol{\theta} = (\theta_1, ..., \theta_n)$, with $\theta_i \geq 0$, $\sum_{i=1}^n \theta_i = 1$.

Proof. Let $D_k^\alpha(f_1, ..., f_n; \boldsymbol{\theta}) = \frac{JGFI_k(f_1,...,f_n;\boldsymbol{\theta})}{e^{-\alpha}}$. Then, by the definition of $G_{JGFI,\alpha}(f_1, ..., f_n; \boldsymbol{\theta})$ and by using Lemma 2.2, we have

$$\begin{aligned} G_{JGFI,\alpha}(f_1, ..., f_n; \boldsymbol{\theta}) &= \sum_{i=1}^n \theta_i G_{GFI}(f_i, \alpha) - G_{GFI}\Big(\sum_{i=1}^n \theta_i f_i, \alpha\Big) \\ &= \sum_{i=1}^n \theta_i \Big\{\sum_{k=0}^{\infty} \frac{\alpha^k}{k!} I_k(f_i)\Big\} - \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} I_k\Big\{\sum_{i=1}^n \theta_i f_i\Big\} \\ &= \sum_{k=0}^{\infty} \frac{\alpha^k}{k!}\Big\{\sum_{i=1}^n \theta_i I_k(f_i) - I_k\Big(\sum_{i=1}^n \theta_i f_i\Big)\Big\} \\ &= \sum_{k=0}^{\infty} p_\alpha(k) D_k^\alpha(f_1, ..., f_n; \boldsymbol{\theta}), \end{aligned}$$

as required. □

Corollary 4.5. *The GFJGFI measure is non-negative.*

Proof. By Theorem 4.4, the result follows readily. □

Theorem 4.6. *Let the variables $X_1, ..., X_n$ have density functions $f_1, ..., f_n$, respectively, and $G_{JGFI,\alpha}(f_1, ..., f_n; \boldsymbol{\theta})$ be the corresponding generating function of the Jensen-generalized Fisher measure. Then,*

$$\frac{\partial^k}{\partial \alpha^k} G_{JGFI,\alpha}(f_1, ..., f_n; \boldsymbol{\theta})\Big|_{\alpha=0} = JGFI_k(f_1, ..., f_n; \boldsymbol{\theta}),$$

where $JGFI_k(f_1, ..., f_n; \boldsymbol{\theta})$ is the Jensen-generalized Fisher information of order k in (4.1).

Proof. Letting $\rho_i(x) = \frac{\partial}{\partial x} \log f_i(x)$ and $\rho_T(x) = \frac{\partial}{\partial x} \log\left(\sum_{i=1}^n \theta_i f_i(x)\right)$, and then taking the kth derivative of $G_{JGFI,\alpha}(f_1, ..., f_n; \boldsymbol{\theta})$ with respect to α, we obtain

$$\begin{aligned} \frac{\partial^k}{\partial \alpha^k} G_{JGFI,\alpha}(f_1, ..., f_n; \boldsymbol{\theta})|_{\alpha=0} &= \frac{\partial^k}{\partial \alpha^k}\Big\{\sum_{i=1}^n \theta_i G_{GFI}(f_i, \alpha) - G_{GFI}\Big(\sum_{i=1}^n \theta_i f_i, \alpha\Big)\Big\}\Big|_{\alpha=0} \\ &= \sum_{i=1}^n \theta_i \int \frac{\partial^k}{\partial \alpha^k} e^{\alpha|\rho_i(x)|} f_i(x) dx\Big|_{\alpha=0} \\ &\quad - \int \frac{\partial^k}{\partial \alpha^k} e^{\alpha|\rho_T(x)|} \sum_{i=1}^n \theta_i f_i(x) dx\Big|_{\alpha=0} \\ &= \sum_{i=1}^n \theta_i \int |\rho_i(x)|^k f_i(x) dx - \int |\rho_T(x)|^k \sum_{i=1}^n \theta_i f_i(x) dx \\ &= \sum_{i=1}^n \theta_i I_k(f_i) - I_k\Big(\sum_{i=1}^n \theta_i f_i\Big) \\ &= JGFI_k(f_1, f_2, ..., f_n; \boldsymbol{\theta}), \end{aligned}$$

as required. $\square$

5. Application of JGFI measure

We now demonstrate an application of the JGFI measure in (4.1) to image processing. Let $X_1, \ldots, X_n$ be a random sample from density f. Then, the kernel estimate of density f, based on kernel function K with bandwidth $h > 0$ at a given point x, is given by

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\Big(\frac{x - X_i}{h}\Big). \tag{5.1}$$

Further, the non-parametric estimate of the first derivative of density f, at a given point x, is given by

$$\begin{aligned} \hat{f}^{(1)}(x) &= \frac{d}{dx}\frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{x-X_i}{h}\Big) = \frac{1}{nh^2}\sum_{i=1}^{n}\frac{d}{dx}K\Big(\frac{x-X_i}{h}\Big) \\ &= \frac{1}{nh^2}\sum_{i=1}^{n} K^{(1)}\Big(\frac{x-X_i}{h}\Big). \end{aligned} \tag{5.2}$$

For more details, see [5] and [2]. Upon making use of (5.1) and (5.2), the integrated non-parametric estimate of the generalized Fisher information associated with variable X with density f, is given by

$$\begin{aligned} \widehat{I_k}(f) &= \int \left|\frac{\hat{f}^{(1)}(x)}{\hat{f}(x)}\right|^k \hat{f}(x)dx \\ &= \frac{1}{nh^{k+1}} \int \left|\frac{\sum_{i=1}^n K^{(1)}(\frac{x-X_i}{h})}{\sum_{i=1}^n K(\frac{x-X_i}{h})}\right|^k \sum_{i=1}^{n} K\Big(\frac{x-X_i}{h}\Big)dx. \end{aligned} \tag{5.3}$$

From (5.3) and with the use of Gaussian kernel $K(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}$, we have

$$\begin{aligned} \widehat{I_k}(f) &= \int \left|\frac{\hat{f}^{(1)}(x)}{\hat{f}(x)}\right|^k \hat{f}(x)dx \\ &= \frac{1}{n(2\pi)^{\frac{1}{2}}h^{k+1}} \int \left|\frac{\sum_{i=1}^n (\frac{x-X_i}{h})e^{\frac{-1}{2}(\frac{x-X_i}{h})^2}}{\sum_{i=1}^n e^{\frac{-1}{2}(\frac{x-X_i}{h})^2}}\right|^k \sum_{i=1}^{n} e^{\frac{-1}{2}(\frac{x-X_i}{h})^2}dx. \end{aligned} \tag{5.4}$$

Now, let X and Y be two continuous random variables with density functions f_1 and f_2, respectively. Based on the random samples $\boldsymbol{X} = (X_1, ..., X_n)$ and $\boldsymbol{Y} = (Y_1, ..., Y_n)$ from densities f_1 and f_2, respectively, the integrated non-parametric estimate of the Jensen-generalized Fisher information $JGFI_k(f_1, f_2; \alpha = \frac{1}{2})$ can be obtained as

$$\begin{aligned} \widehat{JGFI}_k(f_1,f_2) &= \frac{1}{2}\widehat{I}_k(f_1) + \frac{1}{2}\widehat{I}_k(f_2) - \widehat{I}_k\Big(\frac{f_1+f_2}{2}\Big) \\ &= \frac{1}{2nh_1^{k+1}} \int \left|\frac{\sum_{i=1}^n K^{(1)}(\frac{x-X_i}{h_1})}{\sum_{i=1}^n K(\frac{x-X_i}{h_1})}\right|^k \sum_{i=1}^{n} K\Big(\frac{x-X_i}{h_1}\Big)dx \\ &+ \frac{1}{2nh_2^{k+1}} \int \left|\frac{\sum_{i=1}^n K^{(1)}(\frac{x-Y_i}{h_2})}{\sum_{i=1}^n K(\frac{x-Y_i}{h_2})}\right|^k \sum_{i=1}^{n} K\Big(\frac{x-Y_i}{h_2}\Big)dx \\ &- \frac{1}{2(h_1h_2)^k} \int \left\{\left|\frac{h_2^2\sum_{i=1}^n K^{(1)}(\frac{x-X_i}{h_1}) + h_1^2\sum_{i=1}^n K^{(1)}(\frac{x-Y_i}{h_2})}{h_2\sum_{i=1}^n K(\frac{x-X_i}{h_1}) + h_1\sum_{i=1}^n K(\frac{x-Y_i}{h_2})}\right|^k \left\{\frac{1}{nh_1}\sum_{i=1}^{n} K\Big(\frac{x-X_i}{h_1}\Big)\right.\right. \\ &\left.\left. + \frac{1}{nh_2}\sum_{i=1}^{n} K\Big(\frac{x-Y_i}{h_2}\Big)\right\}\right\}dx, \end{aligned} \tag{5.5}$$

where the bandwidths h_1 and h_2 are determined based on samples $\boldsymbol{X} = (X_1, ..., X_n)$ and $\boldsymbol{Y} = (Y_1, ..., Y_n)$, respectively. Now, using the Gaussian kernel $K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$ and its corresponding first derivative, the integral non-parametric estimate of $\widehat{JGFI}_k(f_1, f_2)$ in (5.5) is given as well as the estimate of $\widehat{I}_k(f)$ in (5.4). Finally, using the Cavalieri-Simpson rule for numerical integration, the empirical estimate of the JGFI measure in (5.5) can be obtained.

We now present an example of image processing and compute the JGFI measure between the original picture and each of its adjusted versions. Figure 2 shows a sample picture of two parrots (original picture) labeled as X and three adjusted versions of the original picture labeled as Y (increasing brightness), Z (increasing contrast), and W (gamma corrected). The available data of the main picture are 768×512 cells and the gray level of each cell has a value between 0 (black) and 1 (white). We apply JGFI measure in order to examine the amount of the dissimilarity between the original picture and each of its noisy versions. For this purpose, we consider three cases of the original image by creating noise and interference as $Y(= X + 0.3)$, $Z(= 2X)$, and $W(= \sqrt{X})$. For pertinent details, see ***EBImage*** package in ***R*** software [12].

Figure 2. Sample picture of two parrots with its adjusted versions.

We have plotted in Figure 3 the extracted histograms, with the corresponding empirical densities for pictures X, Y, Z, and W. As we can see from Figures 2 and 3, the highest degree of similarity is first related to W and then to Y, whereas Z has the highest degree of divergence with respect to the original picture X. We have presented the JGFI measure (for selected values of $k = 2$ and 2.5) for all four pictures in Table 1. It is easily seen that JGFI measure gets increased when the similarity gets decreased with respect to the original picture. Therefore, the JGFI measure can be considered as an efficient criteria for comparing the similarity between an original picture and its noisy versions.

Table 1. The JGFI measure between the original picture and each of its noisy versions.

	JGFI($k = 2$)	JGFI($k = 2.5$)
$X \leftrightarrow Y$	0.00204	0.00227
$X \leftrightarrow Z$	0.00236	0.00243
$X \leftrightarrow W$	0.00096	0.00055

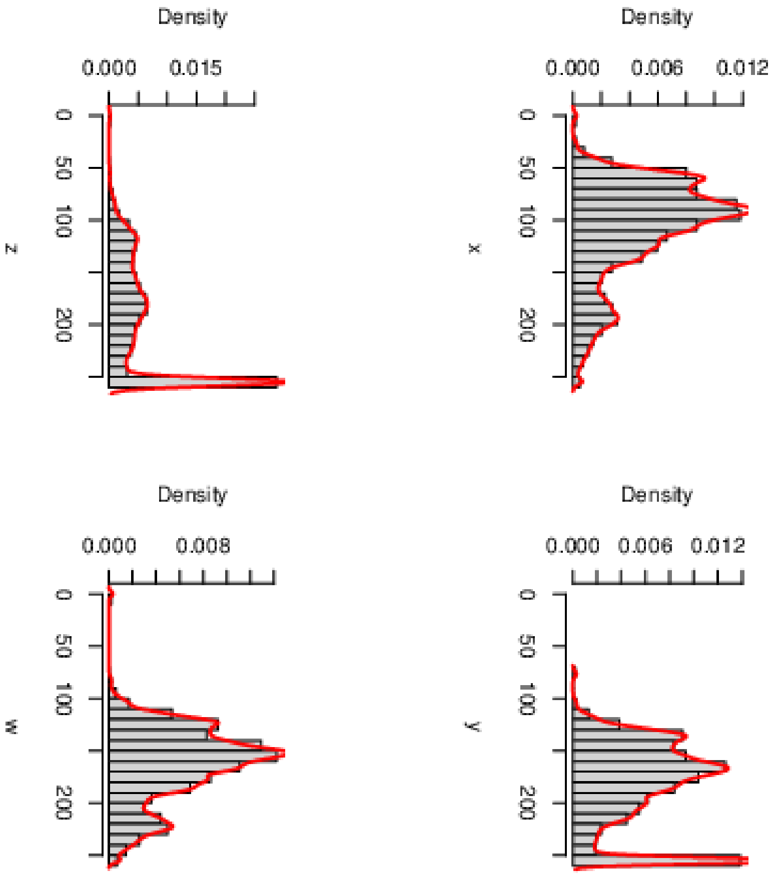

Figure 3. The histograms and the corresponding empirical densities for pictures X, Y, Z, and W.

6. Concluding remarks

In this work, we have considered generating function for the generalized Fisher information measure and established some properties of it. We have also considered a finite mixture density function and derived the generalized Fisher information for the mixing parameter vector. We have shown that the generalized Fisher information for the mixing parameter is connected to the Pearson-Vajda divergence. Further, the Jensen-generalized Fisher information and its generating function have been discussed. We have shown specifically that this generating function can be expressed as an infinite mixture of scaled-versions of Jensen-generalized Fisher information divergence measures. Finally, we have described an application of the JGFI measure by considering an example in image processing.

Acknowledgment. The authors express their sincere thanks to the Editor and the anonymous reviewers for their many useful comments and suggestions on an earlier version of this manuscript which let to this improvement version.

References

[1] N. Balakrishnan and A. Stepanov, *On the Fisher information in record data*, Statist. Probab. Lett. **76** (5), 537-545, 2006.

[2] P.K. Bhattacharya, *Estimation of a probability density function and its derivatives*, Sankhya A **29** (4), 373-382, 1967.

[3] J.F. Bercher, *Some properties of generalized Fisher information in the context of nonextensive thermostatistics*, Phys. A **392** (15), 3140-3154, 2013.

[4] S.G. Bobkov, *Moments of the scores*, IEEE Trans. Inf. Theory **65** (9), 5294-5301, 2019.

[5] T. Duong, M. Wand, J. Chacon and A. Gramacki, Package "ks", R package version: 1.13.5, 2022.

[6] R.A. Fisher, *Tests of significance in harmonic analysis*, Proc. R. Soc. A **125** (796), 54-59, 1929.

[7] S. Golomb, *The information generating function of a probability distribution (corresp.)*, IEEE Trans. Inf. Theory **12** (1), 75-77, 1966.

[8] S. Guiasu and C. Reischer, *The relative information generating function*, Inform. Sci. **35** (3), 235-241, 1985.

[9] O. Kharazmi and N. Balakrishnan, *Jensen-information generating function and its connections to some well-known information measures*, Statist. Probab. Lett. **170**, 1-10, 2021.

[10] F. Nielsen and R. Nock, *On the chi square and higher-order chi distances for approximating f-divergences*, IEEE Signal Process. Lett. **21** (1), 10-13, 2014.

[11] T. Papaioannou, K. Ferentinos and C. Tsairidis, *Some information theoretic ideas useful in statistical inference*, Methodol. Comput. Appl. Probab. **9** (2), 307-323, 2007.

[12] G. Pau, F. Fuchs, O. Sklyar, M. Boutros and W. Huber, *EBImage-an R package for image processing with applications to cellular phenotypes*, Bioinformatics **26** (7), 979-981, 2010.

[13] P. Sánchez-Moreno, A. Zarzo and J.S. Dehesa, *Jensen divergence based on Fisher's information*, J. Phys. A Math. **45** (12), 125305, 2012.

[14] C.E. Shannon, *A mathematical theory of communication*, Bell Labs Tech. J. **27** (3), 379-423, 1948.

[15] P. Zegers, *Fisher information properties*, Entropy **17** (7), 4918-4939, 2015.