Jensen–Inaccuracy Information Measure

Availability:
This version is available at: 11583/2994645 since: 2024-11-21T09:46:05Z

Publisher:
MDPI

Published
DOI:10.3390/e25030483

Terms of use:

Publisher copyright

(Article begins on next page)

19 February 2025

Jensen–Inaccuracy Information Measure

Omid Kharazmi [1,*] , **Faezeh Shirazinia** [1] , **Francesco Buono** [2,†] **and Maria Longobardi** [3]

1 Department of Statistics, Faculty of Mathematical Sciences, Vali-e-Asr University of Rafsanjan, Rafsanjan 7718897111, Iran

2 Dipartimento di Matematica e Applicazioni "Renato Caccioppoli", Università degli Studi di Napoli Federico II, 80138 Naples, Italy

3 Dipartimento di Biologia, Università degli Studi di Napoli Federico II, 80138 Naples, Italy

* Correspondence: omid.kharazmi@vru.ac.ir

† Current address: Institute of Statistics, RWTH Aachen University, 52056 Aachen, Germany.

Abstract: The purpose of the paper is to introduce the Jensen–inaccuracy measure and examine its properties. Furthermore, some results on the connections between the inaccuracy and Jensen–inaccuracy measures and some other well-known information measures are provided. Moreover, in three different optimization problems, the arithmetic mixture distribution provides optimal information based on the inaccuracy information measure. Finally, two real examples from image processing are studied and some numerical results in terms of the inaccuracy and Jensen–inaccuracy information measures are obtained.

Keywords: Jensen–inaccuracy measure; inaccuracy measure; optimal information; Rényi entropy; average entropy

MSC: 62B10; 94A17

1. Introduction

In recent decades, several researchers have studied information theory and its applications in various fields such as statistics, physics, economics, and engineering. Shannon entropy [1] is a fundamental quantity in information theory and it is defined for a continuous random variable X having probability density function (PDF) f on support $\mathcal{X}$ as

$$H(X) = -\int_{\mathcal{X}} f(x) \log f(x) dx, \tag{1}$$

where log denotes the natural logarithm. Throughout the paper, the support will be omitted in all the integrals. Several extensions of Shannon entropy have been considered by many researchers (see, for instance, Rényi and Tsallis entropies [2,3]) by providing one-parameter families of entropy measures.

Moreover, several divergence measures based on entropy have been defined in order to measure similarity and dissimilarity between two density functions. Among them, chi-square, Kullback–Leibler, Rényi divergences, and their extensions have been introduced; for further details, see Nielsen and Nock [4], Di Crescenzo and Longobardi [5], and Van Erven and Harremos [6].

Let X and Y be two random variables with density functions f and g, respectively. Then, the Kullback–Leibler divergence [7] is defined as

$$KL(f,g) = \int f(x) \log \frac{f(x)}{g(x)} dx, \tag{2}$$

provided the integral exists. Because of some limitations of Shannon entropy, Kerridge [8] proposed a measure, known as the inaccuracy measure (or the Kerridge

measure). Consider f and g as two probability density functions. Then, the inaccuracy measure between f and g is given by

$$\mathcal{K}(f,g) = H(f) + KL(f,g) = -\int f(x)\log g(x)dx.$$

Several extensions of the inaccuracy measure have been developed, as well as Shannon entropy. For more details, see Kayal and Sunoj [9] and the references therein.

In the literature, the class of Jensen divergences has been studied in an extensive way as a general technique in developing information measures. Recently, other information measures such as the Jensen–Shannon, Jensen–Fisher, and Jensen–Gini measures have been studied as generalizations of well-known quantities. For further details, see Lin [10], Sánchez-Moreno et al. [11], and Mehrali et al. [12].

However, the link between the inaccuracy measure and the Jensen concept has remained unknown so far. Therefore, the main motivation in this paper is to present the Jensen–inaccuracy information (JII) measure and its properties. Let us remark that the introduction of JII measure is motivated by the fact that it can be expressed as mixture of well-known divergence measures, and it is close to the arithmetic–geometric divergence measure. Nevertheless, our new measure obtains better results when studying the similarity between elements in the field of image quality assessment. Furthermore, we establish some results associated with the connection between inaccuracy and Jensen–inaccuracy information measures and some other measures of discrimination such as Rényi entropy, average entropy, and Rényi divergence. Next, we show that the arithmetic mixture distribution provides optimal information under three different optimization problems based on the inaccuracy information measure. In the following, some well-known and useful information measures are recalled.

An extended version of the Shannon entropy measure for $\alpha > 0$ and $\alpha \neq 1$, is defined by Rényi [2] as

$$R_\alpha(f) = \frac{\log\left(\int_{\mathcal{X}} f^\alpha(x)dx\right)}{1-\alpha}. \tag{3}$$

Several applications of Rényi entropy have been discussed in the literature.

Furthermore, the Rényi divergence of order $\alpha > 0$ and $\alpha \neq 1$ between density functions f and g is defined by

$$D_\alpha(f,g) = \frac{\log\left(\int f^\alpha(x)g^{1-\alpha}(x)dx\right)}{\alpha-1}. \tag{4}$$

The information measures in (3) and (4) become Shannon entropy and Kullback–Leibler divergence measures, respectively, when α tends to 1.

Another important diversity measure between two continuous density functions f and g is the chi-square divergence, defined as

$$\chi^2(f,g) = \int \frac{(f(x)-g(x))^2}{f(x)}dx. \tag{5}$$

In a similar manner, we can define $\chi^2(g,f)$.

The rest of this paper is organized as follows. In Section 2, we first introduce the Jensen–inaccuracy information (JII) measure. Then, we show that JII can be expressed as a mixture of the Kullback–Leibler divergence measures. We show that the Jensen–inaccuracy information measure has a close connection with the arithmetic–geometric divergence measure. Furthermore, we present an upper bound for the JII measure in terms of chi-square divergence measures. The (w,α)-Jensen–inaccuracy measure is also introduced in this section as an extended version of JII. We study the inaccuracy information measure for the escort and generalized escort distributions in Section 3. In Section 4, we consider the

average entropy and define the average inaccuracy measure. Furthermore, some results are given in this regard. In Section 5, we show that the arithmetic mixture distribution involves optimal information under three different optimization problems in terms of the inaccuracy measure. Then, in Section 6, a real example is presented in order to study a problem in image processing, and some numerical results are presented in terms of the inaccuracy and Jensen–inaccuracy information measures. More precisely, our measure is useful for detection of similarity between images. Finally, in Section 7, concluding remarks are provided.

2. The Jensen–Inaccuracy Measure

In this section, we introduce the Jensen–inaccuracy measure and then provide a representation for this information measure in terms of Kullback–Leibler divergence measure. Furthermore, we explore the possible connection between Jensen–inaccuracy and arithmetic–geometric divergence measures. We also provide an upper bound for the Jensen–inaccuracy measure based on chi-square divergence measure. At the end of this section, we introduce (w, α)-Jensen–inaccuracy and establish a result for this extended measure.

Definition 1. *Let f, f_0, and f_1 be three density functions. Then, the Jensen–inaccuracy measure between f_0 and f_1 with respect to f is defined by*

$$\mathcal{JK}(f, f_0, f_1) = \frac{1}{2}\mathcal{K}(f, f_0) + \frac{1}{2}\mathcal{K}(f, f_1) - \mathcal{K}\left(f, \frac{f_0 + f_1}{2}\right). \tag{6}$$

Theorem 1. *The $\mathcal{JK}(f, f_0, f_1)$ inaccuracy measure in (6) is non-negative.*

Proof. From the convexity properties of $-\log x, x > 0$, function we have

$$-\log\left(\frac{f_0(x) + f_1(x)}{2}\right) \leq -\frac{1}{2}\log f_0(x) - \frac{1}{2}\log f_1(x). \tag{7}$$

Now, by multiplying both sides of (7) by $f(x)$ and then integrating with respect to x, we obtain

$$-\int f(x)\log\left(\frac{f_0(x) + f_1(x)}{2}\right)dx \leq -\int f(x)\frac{1}{2}\log f_0(x)dx - \int f(x)\frac{1}{2}\log f_1(x)dx,$$

as required. □

2.1. The Jensen–Inaccuracy Measure and its Connection to Kullback–Leibler Divergence

Here, we provide a representation for the Jensen–inaccuracy measure in terms of Kullback–Leibler divergence measure.

Theorem 2. *A representation for the Jensen–inaccuracy measure in (6) based on mixture of Kullback–Liebler divergence measures is given by*

$$\mathcal{JK}(f, f_0, f_1) = \frac{1}{2}KL(f, f_0) + \frac{1}{2}KL(f, f_1) - KL\left(f, \frac{f_0 + f_1}{2}\right). \tag{8}$$

Proof. According to the definition of the Jensen–inaccuracy measure and the relations

$$\mathcal{K}(f, f_0) = KL(f, f_0) + H(f),$$

$$\mathcal{K}(f, f_1) = KL(f, f_1) + H(f),$$

$$\mathcal{K}\left(f, \frac{f_0 + f_1}{2}\right) = KL\left(f, \frac{f_0 + f_1}{2}\right) + H(f),$$

we have

$$
\begin{aligned}
\mathcal{JK}(f, f_0, f_1) &= \frac{1}{2}\mathcal{K}(f, f_0) + \frac{1}{2}\mathcal{K}(f, f_1) - \mathcal{K}\left(f, \frac{f_0 + f_1}{2}\right) \\
&= \frac{KL(f, f_0) + H(f)}{2} + \frac{KL(f, f_1) + H(f)}{2} - KL\left(f, \frac{f_0 + f_1}{2}\right) - H(f) \\
&= \frac{1}{2}KL(f, f_0) + \frac{1}{2}KL(f, f_1) - KL\left(f, \frac{f_0 + f_1}{2}\right),
\end{aligned}
$$

as required. □

Next, we extend the definition of the Jensen–inaccuracy measure based on $n + 1$ density functions.

Definition 2. *Let $X_1, \ldots, X_n$, and Y be random variables with density functions $f_1, \ldots, f_n$, and f, respectively, and $\alpha_1, \ldots, \alpha_n$ be non-negative real numbers such that $\sum_{i=1}^n \alpha_i = 1$. Then, the Jensen–inaccuracy measure is defined as*

$$\mathcal{JK}_{\boldsymbol{\alpha}}(f, f_1, \ldots, f_n) = \sum_{i=1}^n \alpha_i \mathcal{K}(f, f_i) - \mathcal{K}\left(f, \sum_{i=1}^n \alpha_i f_i\right). \tag{9}$$

Theorem 3. *The $\mathcal{JK}_{\boldsymbol{\alpha}}(f, f_1, \ldots, f_n)$ information measure in (9) can be written in terms of Kullback–Leibler divergence as*

$$\mathcal{JK}_{\boldsymbol{\alpha}}(f, f_1, \ldots, f_n) = \sum_{i=1}^n \alpha_i KL(f, f_i) - KL\left(f, \sum_{i=1}^n \alpha_i f_i\right). \tag{10}$$

Proof. From the definition of $\mathcal{JK}_{\boldsymbol{\alpha}}(f, f_1, \ldots, f_n)$ in (9), we have

$$
\begin{aligned}
\mathcal{JK}_{\boldsymbol{\alpha}}(f, f_1, \ldots, f_n) &= \sum_{i=1}^n \alpha_i \mathcal{K}(f, f_i) - \mathcal{K}\left(f, \sum_{i=1}^n \alpha_i f_i\right) \\
&= \sum_{i=1}^n \alpha_i \left(KL(f, f_i) + H(f)\right) - KL\left(f, \sum_{i=1}^n \alpha_i f_i\right) - H(f) \\
&= \sum_{i=1}^n \alpha_i KL(f, f_i) - KL\left(f, \sum_{i=1}^n \alpha_i f_i\right),
\end{aligned}
$$

as required. □

2.2. Connection between Jensen–Inaccuracy and Arithmetic–Geometric Divergence Measures

Now, we explore the connection between the Jensen–inaccuracy and arithmetic–geometric divergence measures. Then, we provide an upper bound for the Jensen–inaccuracy measure based on chi-square divergence measure.

Let f_0 and f_1 be two density functions. Then, the arithmetic–geometric divergence measure is defined as

$$T(f_0, f_1) = \int \left(\frac{f_0(x) + f_1(x)}{2}\right) \log\left(\frac{f_0(x) + f_1(x)}{2\sqrt{f_0(x)f_1(x)}}\right) dx. \tag{11}$$

For more details, see Taneja [13].

In the following definition, we provide an extension of the arithmetic–geometric divergence measure to n-density functions.

Definition 3. *Let $X_1, \ldots, X_n$ be random variables with density functions $f_1, \ldots, f_n$, respectively, and $\alpha_1, \ldots, \alpha_n$ be non-negative real numbers such that $\sum_{i=1}^{n} \alpha_i = 1$. Then, the extended arithmetic–geometric divergence measure is defined as*

$$T(f_1, \ldots, f_n; \boldsymbol{\alpha}) = \int f_T(x) \log \left(\frac{f_T(x)}{\prod_{i=1}^{n} f_i^{\alpha_i}(x)} \right) dx, \tag{12}$$

where $f_T(x) = \sum_{i=1}^{n} \alpha_i f_i(x)$ and $\boldsymbol{\alpha}$ is a brief notation to denote $\alpha_1, \ldots, \alpha_n$.

In the following, we explore the connection between the Jensen–inaccuracy measure in (9) and the arithmetic–geometric divergence measure in (12).

Theorem 4. *If $f = f_T$, then, we have*

$$\mathcal{JK}_{\boldsymbol{\alpha}}(f, f_1, \ldots, f_n) = T(f_1, \ldots, f_n; \boldsymbol{\alpha}), \tag{13}$$

where $T(f_1, \ldots, f_n; \boldsymbol{\alpha})$ is the extended arithmetic–geometric divergence measure defined in (12) and $\boldsymbol{\alpha}$ is a brief notation to denote $\alpha_1, \ldots, \alpha_n$.

Proof. From the assumption $f = f_T$ and Theorem 3, we have

$$
\begin{aligned}
\mathcal{JK}_{\boldsymbol{\alpha}}(f_T, f_1, \ldots, f_n) &= \sum_{i=1}^{n} \alpha_i KL(f_T, f_i) \\
&= \sum_{i=1}^{n} \alpha_i \int f_T(x) \log \left(\frac{f_T(x)}{f_i(x)} \right) dx \\
&= \int f_T(x) \log \left[\prod_{i=1}^{n} \left(\frac{f_T(x)}{f_i(x)} \right)^{\alpha_i} \right] dx \\
&= \int f_T(x) \log \left(\frac{f_T(x)}{\prod_{i=1}^{n} f_i^{\alpha_i}(x)} \right) dx \\
&= T(f_1, \ldots, f_n; \boldsymbol{\alpha}).
\end{aligned}
$$

as required. □

Remark 1. *From the inequality $\log x \leq x - 1$ for $x > 0$, it is easy to obtain an upper bound for $\mathcal{JK}_{\boldsymbol{\alpha}}(f_T, f_1, \ldots, f_n)$ based on the chi-square divergence measure:*

$$\mathcal{JK}_{\boldsymbol{\alpha}}(f_T, f_1, \ldots, f_n) = \sum_{i=1}^{n} \alpha_i KL(f_T, f_i) \leq \sum_{i=1}^{n} \alpha_i \chi^2(f_i, f_T).$$

2.3. The (w, α)-Jensen–Inaccuracy Measure

Here, (p, w)−Jensen–inaccuracy is defined. Moreover, we establish a result for this extended measure.

Definition 4. *Let f, f_0 and f_1 be three density functions. Then, the (w, α)-Jensen–inaccuracy measure between f_0 and f_1 with respect to f is defined by*

$$\mathcal{JK}_{w,\alpha}(f, f_0, f_1) = w\mathcal{K}(f, (1-p)f_0 + pf_1) + (1-w)\mathcal{K}(f, pf_0 + (1-p)f_1) - \mathcal{K}\left(f, (1-\bar{p})f_0 + \bar{p}f_1 \right),$$

where $\bar{p} = wp + (1-w)(1-p)$.

Note that

$$(1-\bar{p})f_0(x) + \bar{p}f_1(x) = w\left((1-p)f_0(x) + pf_1(x)\right) + (1-w)\left(pf_0(x) + (1-p)f_1(x)\right).$$

Theorem 5. *A representation for (w, α)-Jensen–inaccuracy measure based on Kullback–Leibler divergence is given by*

$$\mathcal{JK}_{w,\alpha}(f, f_0, f_1) = wKL(f, (1-p)f_0 + pf_1) + (1-w)KL(f, pf_0 + (1-p)f_1)$$
$$- KL\left(f, (1-\bar{p})f_0 + \bar{p}f_1\right),$$

Proof. It can be proven in the same manner as Theorem 2. □

3. Inaccuracy Information Measure of the Escort and Generalized Escort Distributions

The escort distribution is a baseline definition in non-extensive statistical mechanics and coding theory; it is closely related to Tsallis and Rényi entropies. For more details, see Bercher [14]. We show that the inaccuracy measure between an arbitrary density and its corresponding escort density can be expressed as a mixture of Shannon and Rényi entropies. Furthermore, another finding associated with the inaccuracy measure between a generalized escort distribution and each of its components reveals some interesting connections in terms of Kullback–Leibler and Rényi divergences.

Let f be a density function. Then, the escort density with order $\alpha > 0$ associated with f is defined as

$$f_\alpha(x) = \frac{f^\alpha(x)}{\int f^\alpha(x)dx}. \tag{14}$$

Theorem 6. *Let f be a density function and f_α be an escort density corresponding to f. Then, for $\alpha > 0$, we obtain:*
(i) $\quad \mathcal{K}(f, f_\alpha) = \alpha H(f) + (1-\alpha)R_\alpha(f);$
(ii) $\quad \mathcal{K}(f_\alpha, f) = \frac{1}{\alpha}H(f_\alpha) + \frac{\alpha-1}{\alpha}R_\alpha(f),$
where $R_\alpha(f)$ is Rényi entropy in (3).

Proof. From the definition of inaccuracy measure between f and f_α, we have

$$\mathcal{K}(f, f_\alpha) = -\int f(x)\log f_\alpha(x)dx = -\int f(x)\log\left(\frac{f^\alpha(x)}{\int f^\alpha(x)dx}\right)dx$$

$$= -\alpha\int f(x)\log f(x)dx + \int f(x)\log\left(\int f^\alpha(x)dx\right)dx$$

$$= \alpha H(f) + \log\left(\int f^\alpha(x)dx\right)$$

$$= \alpha H(f) + (1-\alpha)R_\alpha(f),$$

which proves (i). Next

$$\mathcal{K}(f_\alpha, f) = -\int f_\alpha(x)\log f(x)dx = -\int\left(\frac{f^\alpha(x)}{\int f^\alpha(x)dx}\right)\log f(x)dx$$

$$= -\frac{1}{\alpha}\int\left(\frac{f^\alpha(x)}{\int f^\alpha(x)dx}\right)\log f^\alpha(x)dx$$

$$= -\frac{1}{\alpha}\int\left(\frac{f^\alpha(x)}{\int f^\alpha(x)dx}\right)\log\left(\frac{f^\alpha(x)}{\int f^\alpha(x)dx}\right)dx$$

$$\quad -\frac{1}{\alpha}\int\left(\frac{f^\alpha(x)}{\int f^\alpha(x)dx}\right)\log\left(\int f^\alpha(x)dx\right)dx$$

$$= \frac{1}{\alpha}H(f_\alpha) + \frac{\alpha-1}{\alpha}R_\alpha(f),$$

which proves (ii). □

Let f and g be two probability density functions. Then, the generalized escort density for $\alpha \in (0,1)$ is defined as

$$h_\alpha(x) = \frac{f^\alpha(x)g^{1-\alpha}(x)}{\int f^\alpha(x)g^{1-\alpha}(x)dx}. \tag{15}$$

Theorem 7. *The inaccuracy information measure between f and the generalized escort density h_α is given by*

$$\mathcal{K}(f,h_\alpha) = (\alpha-1)D_\alpha(f,g) - \alpha KL(f,g) - \mathcal{K}(f,g), \tag{16}$$

where $D_\alpha(f,g)$ is the relative Rényi entropy defined by

$$D_\alpha(f,g) = \frac{\log\left(\int f^\alpha(x)g^{1-\alpha}(x)dx\right)}{\alpha-1}. \tag{17}$$

Proof. From the definition of $\mathcal{K}(f,h_\alpha)$, we derive

$$
\begin{aligned}
\mathcal{K}(f,h_\alpha) &= -\int f(x)\log h_\alpha(x)dx = -\int f(x)\log\left(\frac{f^\alpha(x)g^{1-\alpha}(x)}{\int f^\alpha(x)g^{1-\alpha}(x)dx}\right)dx\\
&= -\int f(x)\log\left(\frac{f(x)}{g(x)}\right)^\alpha dx + \int f(x)\log\left(\int f^\alpha(x)g^{1-\alpha}(x)dx\right)dx - \mathcal{K}(f,g)\\
&= -\alpha\int f(x)\log\frac{f(x)}{g(x)}dx + \log\left(\int f^\alpha(x)g^{1-\alpha}(x)dx\right) - \mathcal{K}(f,g)\\
&= (\alpha-1)D_\alpha(f,g) - \alpha KL(f,g) - \mathcal{K}(f,g),
\end{aligned}
$$

as required. $\square$

Theorem 8. *Let f_0 and f_1 be two density functions and consider the arithmetic and geometric mixture densities, respectively, as $f_a(x) = pf_0(x) + (1-p)f_1(x)$ and $f_g(x) = \frac{f_0^p(x)f_1^{1-p}(x)}{\int f_0^p(x)f_1^{1-p}(x)dx}$. Then, a lower bound for $\mathcal{K}(f_a,f_g)$ is given by*

$$\mathcal{K}(f_a,f_g) \geq H(f_a) + (1-p)D_p(f_0,f_1). \tag{18}$$

Proof. From the definition $\mathcal{K}(f_a,f_g)$, by using the arithmetic mean–geometric mean inequality, we have

$$
\begin{aligned}
\mathcal{K}(f_a,f_g) &= -\int f_a(x)\log f_g(x)dx = -\int (pf_0(x)+(1-p)f_1(x))\log\left(\frac{f_0^p(x)f_1^{1-p}(x)}{\int f_0^p(x)f_1^{1-p}(x)dx}\right)dx\\
&= -\int (pf_0(x)+(1-p)f_1(x))\log\left(f_0^p(x)f_1^{1-p}(x)\right)dx - \log\left(\int f_0^p(x)f_1^{1-p}(x)dx\right)\\
&\geq -\int (pf_0(x)+(1-p)f_1(x))\log\left(pf_0(x)+(1-p)f_1(x)\right)dx - \log\left(\int f_0^p(x)f_1^{1-p}(x)dx\right)\\
&= H(pf_0+(1-p)f_1) + (1-p)D_p(f_0,f_1)\\
&= H(f_a) + (1-p)D_p(f_0,f_1),
\end{aligned}
$$

as required. $\square$

4. Inaccuracy Measure Based on Average Entropy

Let X be a random variable with PDF f. Then, the average entropy associated with f is defined as

$$AE(f) = -\int f(x)\log\left(\frac{f(x)}{\int f^2(x)dx}\right)dx \tag{19}$$

For pertinent details, see Kittaneh et al. [15].

Theorem 9. *Let f be a density function. Then, an upper bound for the Shannon entropy based on inaccuracy measure is given by*

$$H(f) \leq \mathcal{K}(f, f_2), \tag{20}$$

where f_2 is the corresponding escort distribution of the density f with order 2.

Proof. From the definition of average entropy, we have

$$
\begin{aligned}
AE(f) &= -\int f(x) \log \left(\frac{f(x)}{\int f^2(x) dx} \right) dx \\
&= -\int f(x) \log \left(\frac{f^2(x)}{\int f^2(x) dx} \right) dx + \int f(x) \log f(x) dx \\
&= \mathcal{K}(f, f_2) - H(f).
\end{aligned}
$$

Now, because $AE(f)$ is non-negative (see Theorem 1 of Kittaneh et al. [15]), we have

$$\mathcal{K}(f, f_2) \geq H(f), \tag{21}$$

as required. $\square$

Definition 5. *Let X be a random variable with density f. Then, the average inaccuracy measure is defined as*

$$\mathcal{AK}(f, g) = -\int f(x) \log \left(\frac{g(x)}{\int g^2(x) dx} \right) dx. \tag{22}$$

Theorem 10. *Let f and g be two PDFs. Then, the average inaccuracy measure between f and g, $\mathcal{AK}(f, g)$, can be expressed as*

$$\mathcal{AK}(f, g) = \mathcal{K}(f, g) - R_2(g). \tag{23}$$

Proof. From the definition of the average inaccuracy measure, we have

$$
\begin{aligned}
\mathcal{AK}(f, g) &= -\int f(x) \log \left(\frac{g(x)}{\int g^2(x) dx} \right) dx \\
&= -\int f(x) \log g(x) dx + \log \left(\int g^2(x) dx \right) \\
&= \mathcal{K}(f, g) - R_2(g),
\end{aligned}
$$

as required. $\square$

5. Optimal Information Model under Inaccuracy Information Measure

In this section, we prove that the arithmetic mixture distribution provides optimal information under three different optimization problems associated with inaccuracy information measures. For more details on optimal information properties of some statistical distributions, one may refer to Kharazmi et al. [16] and the references therein.

Theorem 11. *Let f, f_0, and f_1 be three density functions. Then, the solution to the optimization problem*

$$\min_f \mathcal{K}(f_0, f) \text{ subject to } \mathcal{K}(f_1, f) = \eta, \quad \int f(x) dx = 1, \tag{24}$$

is the arithmetic mixture density with mixing parameter $p = \frac{1}{1+\lambda_0}$, $\lambda_0 > 0$ is the Lagrangian multiplier, and η is a constraint associated with the optimization problem.

Proof. We use the Lagrangian multiplier technique in order to solve the optimization problem in (24). Thus, we have

$$\mathcal{L}(f, \lambda_0, \lambda_1) = - \int f_0(x) \log f(x) dx - \lambda_0 \int f_1(x) \log f(x) dx + \lambda_1 \int f(x) dx.$$

Now, differentiating with respect to f, we obtain

$$\frac{\partial}{\partial f} \mathcal{L}(f, \lambda_0, \lambda_1) = - \frac{f_0(x)}{f(x)} - \lambda_0 \frac{f_1(x)}{f(x)} + \lambda_1. \tag{25}$$

By setting (25) to zero, we derive the optimal density function as

$$f(x) = p f_0(x) + (1 - p) f_1(x),$$

where $p = \frac{1}{1+\lambda_0}$, as required. In fact, we obtain the solution based on $f(x)$ as

$$f(x) = \frac{f_0(x) + \lambda_0 f_1(x)}{\lambda_1}.$$

From the normalization condition, we have

$$\int f(x) dx = \int \left(\frac{f_0(x) + \lambda_0 f_1(x)}{\lambda_1} \right) dx = 1,$$

and then $\lambda_1 = 1 + \lambda_0$. □

Theorem 12. *Let f, f_0, and f_1 be three density functions. Then, the solution to the optimization problem,*

$$\min_f \{ w\mathcal{K}(f_0, f) + (1 - w)\mathcal{K}(f_1, f) \} \text{ subject to } \int f(x) dx = 1, \quad 0 \le w \le 1, \tag{26}$$

is the arithmetic mixture density with mixing parameter $p = w$.

Proof. Making use of the Lagrangian multiplier technique as in Theorem 11, the result follows. □

Theorem 13. *Let f, f_0, and f_1 be three density functions and $T_\alpha(X) = \frac{f_0(X)}{f^2(X)}$. Then, the solution to the optimization problem,*

$$\min_f \mathcal{K}(f_0, f) \text{ subject to } E_f(T_\alpha(X)) = \eta, \quad \int f(x) dx = 1, \tag{27}$$

is the arithmetic mixture density with mixing parameter $p = \frac{1}{1+\lambda_0}$, $\lambda_0 > 0$ is the Lagrangian multiplier, $E_f(\cdot)$ is the expectation with respect to f and η is a constraint associated with the optimization problem.

Proof. The result follows analogously with the proof of Theorem 11. □

6. Application

In this section, we first consider the definition of histogram for a given image in the context of image quality assessment. Then, we illustrate two applications by using the inaccuracy and Jensen–inaccuracy measures.

6.1. Image and Histogram

A digital image is defined as a discrete set of small surface elements (pixel). One such digital image is a grayscale image in which each pixel only contains one value (its intensity).

These values are in the set $\{0, \ldots, L-1\}$, where L represents the number of intensity values. Suppose that n_k is the number of times in which the kth intensity appears in the image. Furthermore, the corresponding histogram of a digital image refers to a histogram of the pixel intensity values in the set $\{0, \ldots, L-1\}$ (one may refer to Gonzalez [17]).

6.2. Non-Parametric Jensen–Inaccuracy Estimation

We now show an application of the inaccuracy and Jensen–inaccuracy measures defined in (6) to image processing. Let $X_1, \ldots, X_n$ be a random sample with probability density function f. Then, the kernel estimate of density f based on kernel function K with bandwidth $h_X > 0$ at a fixed point x is given by

$$\hat{f}(x) = \frac{1}{nh_X} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_X}\right). \tag{28}$$

Similarly, the non-parametric estimate of density g with bandwidth $h_Y > 0$, based on random sample $Y_1, \ldots, Y_n$, is expressed as

$$\hat{g}(x) = \frac{1}{nh_Y} \sum_{i=1}^{n} K\left(\frac{x - Y_i}{h_Y}\right). \tag{29}$$

For more details, see Duong et al. [18]. Upon making use of (28) and (29), the integrated non-parametric estimate of the inaccuracy and Jensen–inaccuracy measures are given, respectively, by

$$
\begin{aligned}
\widehat{\mathcal{K}}(f,g) &= -\int \hat{f}(x) \log \hat{g}(x) dx \\
&= -\int \frac{1}{nh_X} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_X}\right) \log\left(\frac{1}{nh_Y} \sum_{i=1}^{n} K\left(\frac{x - Y_i}{h_Y}\right)\right) dx
\end{aligned}
$$

and

$$
\begin{aligned}
\widehat{\mathcal{JK}}(f, f_0, f_1) &= \frac{1}{2}\mathcal{K}(\hat{f}, \hat{f_0}) + \frac{1}{2}\mathcal{K}(\hat{f}, \hat{f_1}) - \mathcal{K}\left(\hat{f}, \frac{\hat{f_0} + \hat{f_1}}{2}\right) \\
&= -\int \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \log\left(\frac{1}{nh_0} \sum_{i=1}^{n} K\left(\frac{x - Y_i}{h_0}\right)\right) dx \\
&\quad -\int \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \log\left(\frac{1}{nh_1} \sum_{i=1}^{n} K\left(\frac{x - Y_i}{h_1}\right)\right) dx \\
&\quad +\int \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \log\left(\frac{\frac{1}{nh_0} \sum_{i=1}^{n} K\left(\frac{x - Y_i}{h_0}\right) + \frac{1}{nh_1} \sum_{i=1}^{n} K\left(\frac{x - Y_i}{h_1}\right)}{2}\right) dx,
\end{aligned}
$$

where h_0, h_1 and h are the corresponding bandwidths for the kernel estimations of the densities f_0, f_1 and f, respectively. Here we use Gaussian kernel $K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$.

Next, we present two examples of image processing (two reference images including grayscale cameraman and lake images) and compute the inaccuracy and Jensen–inaccuracy information measures between the original picture and each of its adjusted versions for both cases.

Cameraman image

Figure 1 shows the original cameraman picture denoted by X and three adjusted versions of this original picture considered as $Y(= X + 0.3)$ (increasing brightness), $Z(= 2 \times X)$ (increasing contrast), and $W(= 0.5 \times X + 0.5)$ (increasing brightness and decreasing contrast). The cameraman image includes 512×512 cells and the gray level of each cell has a value between 0 (black) and 1 (white).

Lake image

Figure 2 shows the original lake image that includes 512×512 cells and the level of the color gray of each cell assumes a value in the interval $[0, 1]$ (0 for black and 1 for white). This image labeled as X and three adjusted versions of it labeled as $Y(= X + 0.3)$ (increasing brightness), $Z(= 2 \times X)$ (increasing contrast) and $W(= \sqrt{X})$ (gamma corrected).

Figure 1. The cameraman image with its three adjusted versions. First row (left panel) original X; first row (right panel) Y (increasing brightness); second row (left panel) Z (increasing contrast); second row (right panel) W (increasing brightness and decreasing contrast).
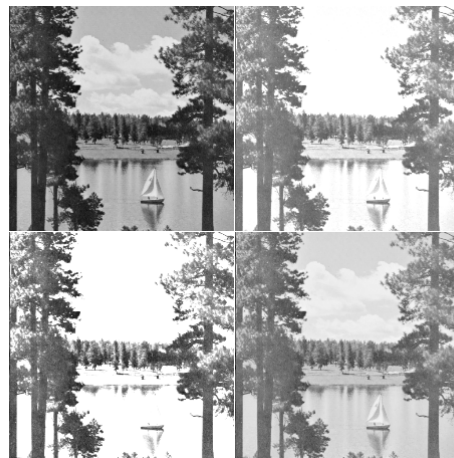
Figure 2. The lake image with its three adjusted versions. First row (left panel) original X; first row (right panel) Y (increasing brightness); second row (left panel) Z (increasing contrast); second row (right panel) W (gamma corrected).

Now, we first compute the inaccuracy between the original image and each of its adjusted versions. Then, we obtain the amount of the dissimilarity between each pair of the interference images with respect to original image based on the Jensen–inaccuracy information measure. For both images, we consider three interferences of the original images, as described above. For more details, see the *EBImage* package in *R* software [19].

The extracted histograms are plotted in Figures 3 and 4, with the corresponding empirical densities for pictures X, Y, Z, and W for the cameraman and lake images, respectively.

We can see from Figures 1 and 3 that the similarity has the highest degree related to W and then to Y, whereas Z has a divergence of the highest degree with respect to X that is the original picture. Moreover, from Figures 2 and 4, the same observation is also found for the lake image and its three adjusted versions. We have presented the inaccuracy and Jensen–inaccuracy information measures (for both cameraman and lake images) for all

pictures in Table 1. Therefore, the inaccuracy and Jensen–inaccuracy information measures can be considered as efficient criteria for comparing the similarity between an original picture and its adjusted versions.
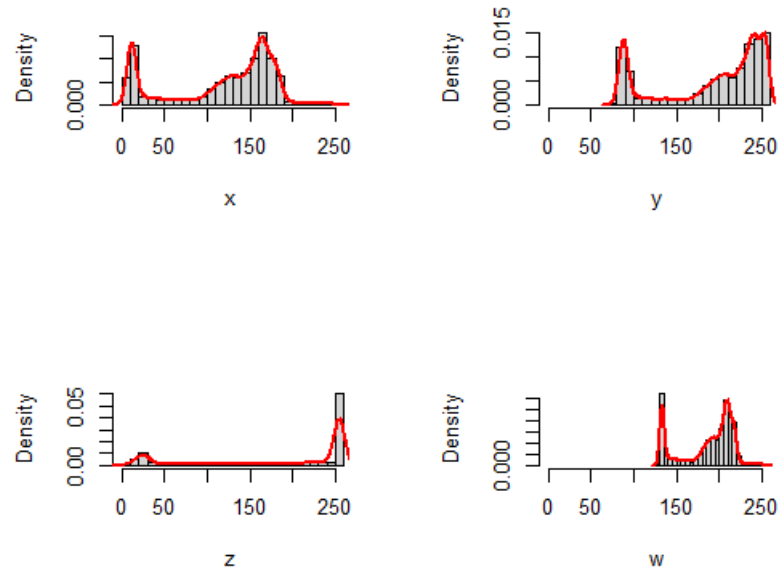

Figure 3. The histograms and the corresponding empirical densities for cameraman image and its three adjusted versions.
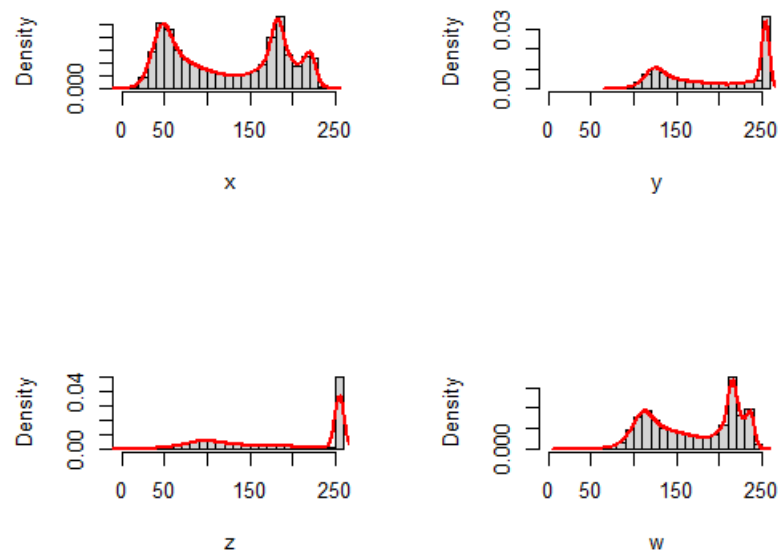

Figure 4. The histograms and the corresponding empirical densities for lake image and its three adjusted versions.

Table 1. Inaccuracy measures for the cameraman and lake images.

Cameraman Image		Lake Image	
Inaccuracy	**Jensen–Inaccuracy**	**Inaccuracy**	**Jensen–Inaccuracy**
$X \leftrightarrow Y$ 8.6142	$(Y, Z \parallel X)$ 0.9674	$X \leftrightarrow Y$ 10.7143	$(Y, Z \parallel X)$ 1.9235
$X \leftrightarrow Z$ 7.3654	$(Y, W \parallel X)$ 0.5707	$X \leftrightarrow Z$ 6.5744	$(Y, W \parallel X)$ 1.2558
$X \leftrightarrow W$ 9.2097	$(Z, W \parallel X)$ 1.4315	$X \leftrightarrow W$ 7.4086	$(Z, W \parallel X)$ 0.5131

According to Fan et al. [20], we have tried to observantly follow axioms 1 and 2 of axiomatic design theory, which has been proposed in recent decades; axiom 1 is about

verification of the validity of designs, and according to axiom 2, one must choose the best design among several options.

7. Conclusions

In this paper, by considering the inaccuracy measure, we have proposed Jensen–inaccuracy and (w, α)-Jensen–inaccuracy information measures. We have specifically shown that Jensen–inaccuracy is connected to the arithmetic–geometric divergence measure. Then, we have studied the inaccuracy measure between the escort distribution and its underling density. Furthermore, we have examined the inaccuracy measure between the generalized escort distribution and its components. It has been shown that these inaccuracy measures are closely connected with Rényi entropy, average entropy, and Rényi divergence. Interestingly, we have shown that the arithmetic mixture distribution provides optimal information under three different optimization problems associated with the inaccuracy measure. Finally, we have described two applications of the inaccuracy and Jensen–inaccuracy measures to image processing. We have considered three adjusted versions of the original cameraman and lake images and then have examined the dissimilarity between the original image and each of its adjusted versions for both cases.

Abbreviations

The following abbreviations are used in this manuscript:

$KL(f, g)$	Kullback–Leibler divergence between f and g
$\mathcal{K}(f, g)$	inaccuracy measure between f and g
PDF	probability density function
JII	Jensen–inaccuracy information measure
$\mathcal{JK}(f, f_0, f_1)$	Jensen–inaccuracy measure between f_0 and f_1 with respect to f
$\mathcal{JK}_{\boldsymbol{\alpha}}(f, f_1, \ldots, f_n)$	extended Jensen–inaccuracy measure based on $n + 1$ density functions
$T(f_0, f_1)$	arithmetic–geometric divergence measure
$T(f_1, \ldots, f_n; \boldsymbol{\alpha})$	extended arithmetic–geometric divergence measure
$\mathcal{JK}_{w,\alpha}(f, f_0, f_1)$	(w, α)-Jensen–inaccuracy measure between f_0 and f_1 with respect to f
$R_\alpha(f)$	Rényi entropy of order α
$D_\alpha(f, g)$	relative Rényi entropy of order α
$AE(f)$	average entropy associated with f
$\mathcal{AK}(f, g)$	average inaccuracy measure

References

1. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
2. Rényi, A. On measures of information and entropy. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Oakland, CA, USA, 1 January 1961; pp. 547–561.
3. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [CrossRef]
4. Nielsen, F.; Nock, R. On the chi square and higher-order chi distances for approximating f-divergences. *IEEE Signal Process. Lett.* **2013**, *21*, 10–13. [CrossRef]
5. Di Crescenzo, A.; Longobardi, M. Some properties and applications of cumulative Kullback-Leibler information. *Appl. Stoch. Model. Bus. Ind.* **2015**, *31*, 875–891. [CrossRef]
6. Van Erven, T.; Harremos, P. Rényi divergence and Kullback-Leibler divergence. *IEEE Trans. Inf. Theory* **2014**, *60*, 3797–3820. [CrossRef]
7. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
8. Kerridge, D.F. Inaccuracy and inference. *J. R. Stat. Soc. Ser. (Methodol.)* **1961**, *23*, 184–194. [CrossRef]
9. Kayal, S.; Sunoj, S.M. Generalized Kerridge's inaccuracy measure for conditionally specified models. *Commun.-Stat.-Theory Methods* **2017**, *46*, 8257–8268. [CrossRef]
10. Lin, J. Divergence measures based on the Shannon entropy. *IEEE Trans. Inf. Theory* **1991**, *37*, 145–151. [CrossRef]
11. Sánchez-Moreno, P.; Zarzo, A.; Dehesa, J.S. Jensen divergence based on Fisher's information. *J. Phys. A Math. Theor.* **2012**, *45*, 125305. [CrossRef]
12. Mehrali, Y.; Asadi, M.; Kharazmi, O. A Jensen-Gini measure of divergence with application in parameter estimation. *Metron* **2018**, *76*, 115–131. [CrossRef]
13. Taneja, I.J. Generalized symmetric divergence measures and the probability of error. *Commun.-Stat.-Theory Methods* **2013**, *42*, 1654–1672. [CrossRef]
14. Bercher, J.F. Source coding with escort distributions and Rényi entropy bounds. *Phys. Lett. A* **2009**, *373*, 3235–3238. [CrossRef]
15. Kittaneh, O.A.; Khan, M.A.; Akbar, M.; Bayoud, H.A. Average entropy: A new uncertainty measure with application to image segmentation. *Am. Stat.* **2016**, *70*, 18–24. [CrossRef]
16. Kharazmi, O.; Contreras-Reyes, J.E.; Balakrishnan, N. Optimal information, Jensen-RIG function and α-Onicescu's correlation coefficient in terms of information generating functions. *Phys. A Stat. Mech. Its Appl.* **2023**, *609*, 128362. [CrossRef]
17. Gonzalez, R.C. *Digital Image Processing*; Prentice Hall: New York, NY, USA, 2009.
18. Duong, T.; Duong, M.T.; Suggests, M.A.S.S. Package 'ks'. R Package Version, 1(5). 2022. Available online: https://cran.r-project.org/web/packages/ks/index.html (accessed on 15 January 2023).
19. Pau, G.; Fuchs, F.; Sklyar, O.; Boutros, M.; Huber, W. EBImage—An R package for image processing with applications to cellular phenotypes. *Bioinformatics* **2010**, *26*, 979–981. [CrossRef] [PubMed]
20. Fan, L.X.; Cai, M.Y.; Lin, Y.; Zhang, W.J. Axiomatic design theory: Further notes and its guideline to applications. *Int. J. Mater. Prod. Technol.* **2015**, *51*, 359–374. [CrossRef]