Filomat 37:19 (2023), 6525–6536
https://doi.org/10.2298/FIL2319525K

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Some new results for residual Fisher information distance

Omid Kharazmi[a,*], **Sanku Dey**[b]

[a]*Department of Statistics, Faculty of Mathematical Sciences, Vali-e-Asr University of Rafsanjan, Iran*
[b]*Department of Statistics, St. Anthony's College, Shillong, Meghalaya, India*

Abstract. Fisher information plays a pivotal role throughout statistical inference especially in optimal and large sample studies in estimation theory. It also plays a key role in physics, thermodynamic, information theory and other applications. In this paper, we establish some new results on residual Fisher information distance (RFID) between residual density functions of two systems. Further, some results on RFID and their relations to other reliability measures are investigated along with some comparison of systems based on stochastic ordering. A lower bound for RFID measure is provided based on quadratic form of hazards functions. In addition, RFID measure for equilibrium distributions are studied. Finally, we establish some results associated with residual Fisher information (RFI) and RFID measures of escort and generalized escort distributions.

1. Introduction

The Fisher information (FI) is not only integral to statistical inference but also is considered fundamental in statistics, information theory, physics, and allied disciplines (see, for example, Fisher (1929), Cover and Thomas (1991), Frieden (2004) and Shao (2003)). Besides, Fisher information has been applied in areas such as biology, social science, geophysics, and encryption of covert information, etc. FI can be used as a measure to quantify the amount of information that the acquired data contains about an unknown parameter or vector of parameters that one wishes to estimate from the data of a statistical model. If the data carries greater the amount of information about the unknown parameter or vector of parameters, the higher is the accuracy (i.e., the smaller the standard deviation) of estimation of parameter or vector of parameters. The amount of information is stochastic in nature and is determined by considering how the likelihood function of observed data changes with the value of unknown parameter or vector of parameters of interest. If the likelihood function of the observed data does not change significantly with the change in value of the unknown parameter or vector of parameters of interest, it indicates that the observed data contains relatively little information about the parameter or vector of parameters. Whereas, if the likelihood of the data is sensitive to changes in the value of the parameters of interest, it indicates that the observed data carries a relatively greater amount of information about the parameter of interest and can be estimated with relatively high accuracy. Let us consider a random variable X (continuous or discrete) with a distribution function F_θ having a probability density function f_θ, where $\theta \in \Theta \subseteq R$. We assume throughout the paper

2020 *Mathematics Subject Classification*. 62B10.
Keywords. Equilibrium distribution; Escort distribution; Fisher information distance; Residual density function; Stochastic ordering.
Received: 11 March 2021; Accepted: 15 June 2021
Communicated by Biljana Popović
* Correspondng author: Omid Kharazmi
Email addresses: omid.kharazmi@vru.ac.ir (Omid Kharazmi), sankud66@gmail.com (Sanku Dey)

that $f_\theta(x)$ is differentiable with respect to both θ and x. The Fisher information of random variable X (or distribution F_θ) about the parameter θ , based on an observation x of X, is defined as

$$\mathcal{I}(\theta) \quad = \quad E\Big[\frac{\partial \log f_\theta(X)}{\partial\theta}\Big]^2. \tag{1}$$

If f_θ is sharply peaked with respect to changes in θ, it indicates the "correct" value of θ from the data, or equivalently, the data X provides a lot of information about the parameter θ. If the likelihood function is flat and spread-out, then it would take many samples of X to estimate the actual "true" value of θ that would be obtained using the entire population being sampled. Moreover, the Fisher information $I(\theta)$ measures how high the peaks are in the log-likelihood function. It plays crucial role, in both classical and Bayesian statistical inference, in derivation of optimal estimators particularly in large sample studies when one is interested to study asymptotic properties of estimators; see, for example Lehmann and Casella (1998) and Shaked and Shanthikumar (2007). Fisher information is widely used in optimal experimental design.

If a random variable X has density $f(x)$, under the condition that the derivative of f exists for all values on its support, the Fisher information of the density is defined as

$$\mathcal{I}(f) = E\Big[\rho^2(X)\Big], \tag{2}$$

where $\rho(x) = \frac{f'(x)}{f(x)}$ is called the score function corresponding to f.

The definition of the Fisher information distance (relative Fisher information) was first propounded by Otto and Villani (2000), in context of translationally-invariant case. Thereafter, Fisher information distance has been used in many applications in different problems and fields; Venkatesan and Plastino (2014), Yáñez et al. (2008), Yamano (2013), Bobkov et al. (2014) and Johnson (2004). It is to be noted that the first general analysis of the relative Fisher information was presented by Zegres (2002).

Given two random variables X and Y with absolutely continuous density functions f and g, respectively, the Fisher information distance (FID) between X and Y (or f and g) is defined by

$$\mathcal{D}(f,g) = E_f\Big[\Big(\rho_f(X) - \rho_g(X)\Big)^2\Big], \tag{3}$$

where $\rho_f(x)$ and $\rho_g(x)$ are the score functions corresponding to f and g, respectively. Recently, time-dependent version of (2) and (3) have been proposed by Kharazmi and Asadi (2018). These information measures so-called RFI and RFID.

The purpose of this work is to establish some new properties of RFID measure between two residual lifetime distributions. Let X be a nonnegative random variable denoting a duration such as a lifetime where we assume that it has the distribution function F, and the probability density function f. The random variable of interest is the *residual random variable*, $X_t = X|X > t$, on the set

$$\mathcal{S}_t = \{x : x > t\} \quad t \in (0, b) \;\; b \le \infty.$$

Hence, the distribution of interest for computing information is the residual distribution with survival function

$$\bar{F}_t(x) = \begin{cases} \frac{\bar{F}(x)}{\bar{F}(t)} & x \in \mathcal{S}_t \\ 1 & \text{otherwise,} \end{cases} \tag{4}$$

provided that $\bar{F}(t) < \infty$, where $\bar{F} = 1 - F$ denotes the survival function of X; see Barlow and Proschan (1975) for pertinent details.

In this work, our main objective is to establish some new results for RFID measure. We examine the relationship between RFID and some of reliability quantities. Moreover, we study RFID measure for equilibrium and escort distributions. The rest of this paper is organized as follows. In Section 2, we first establish some new results for RFID measure and provide results associated with stochastic ordering in order to compare the system lifetimes. We then obtain a lower bound for the RFID measure based on quadratic form of hazards functions. In Section 3, we study RFID measure for equilibrium distributions. It is shown that RFID measure between two equilibrium distributions is connected with hazard and mean residual functions of underlying variable. In Section 4, we establish some results associated with RFI and RFID measures of escort and generalized escort distributions. Finally, some concluding remarks are made in Section 5.

2. New results for Residual Fisher information distance (RFID)

Assuming that $f_t(x)$ and $g_t(x)$ denote the density functions corresponding to residual lifetimes variables X and Y, respectively. We now introduce RFI and RFID measures that will be used in the sequel.

Definition 2.1. Let X_t be a residual random variable with an absolutely continuous density function $f_t(x)$. The residual Fisher information of $f_t(x)$ is defined as

$$\mathcal{I}(f;t) \;=\; E\Big(\rho^2(X)|X>t\Big), \tag{5}$$

where $t>0$ and $b\leq\infty$ the right extremity of the support of X, i.e., $F(b)=1$.

Definition 2.2. *The RFID between f_t and g_t is defined as*

$$\mathcal{D}(f,g;t) \;=\; E\Big[\Big(\rho_f(X)-\rho_g(X)\Big)^2|X>t\Big]. \tag{6}$$

Clearly for a non-negative random variable X, RFID reduce to FID, when $t\to 0$.

For more details, see Kharazmi and Asadi (2018). Next, we establish some new results associated with RFID measure.

Theorem 2.3. *Given two random variables X and Y with RFID measure $\mathcal{D}(f,g;t)$. Then under the condition of Lemma 2.1 from Kharazmi and Asadi (2018), we have*

$$\mathcal{D}(f,g;t)\geq\Big(r_f(t)+E\Big(\rho_g(X)|X>t\Big)\Big)^2, \tag{7}$$

where $r_f(t)=\frac{f(t)}{\bar{F}(t)}$ is hazard function of variable X.

Proof : From Lemma 2.1 of [20], with $g(x)\equiv 1$, gives $E\Big[\rho_f(X)|X>t\Big]=-r_f(t)$. Hence we get

$$\begin{aligned} Var\Big[\Big(\rho_f(X)-\rho_g(X)\Big)|X>t\Big] &= E\Big[\Big(\rho_f(X)-\rho_g(X)\Big)^2|X>t\Big] \\ &\quad -E^2\Big[\Big(\rho_f(X)-\rho_g(X)\Big)|X>t\Big] \\ &= \mathcal{D}(f,g;t)-\Big(r_f(t)+E(\rho_g(X)|X>t)\Big)^2 \\ &\geq 0. \end{aligned}$$

Given two random variables X and Y with absolutely continuous density functions f and g, respectively. The variable X is said to be less than Y in hazard rate order, $X\leq_{hr}Y$, if $r_f(x)\geq r_g(x)$, for all x in the union of supports of X and Y, where $r_f(x)(r_g(x))$ is the hazard rate of $X(Y)$. The following theorem provides an interesting lower bound for RFID measure based on the hazard functions.

Theorem 2.4. *Given two random variables X and Y with hazard functions $r_f(x)$ and $r_g(x)$, respectively. If $\rho_g(x)$ is decreasing and $X\leq_{hr}Y$, then*

$$\mathcal{D}(f,g;t)\geq\Big(r_f(t)-r_g(t)\Big)^2. \tag{8}$$

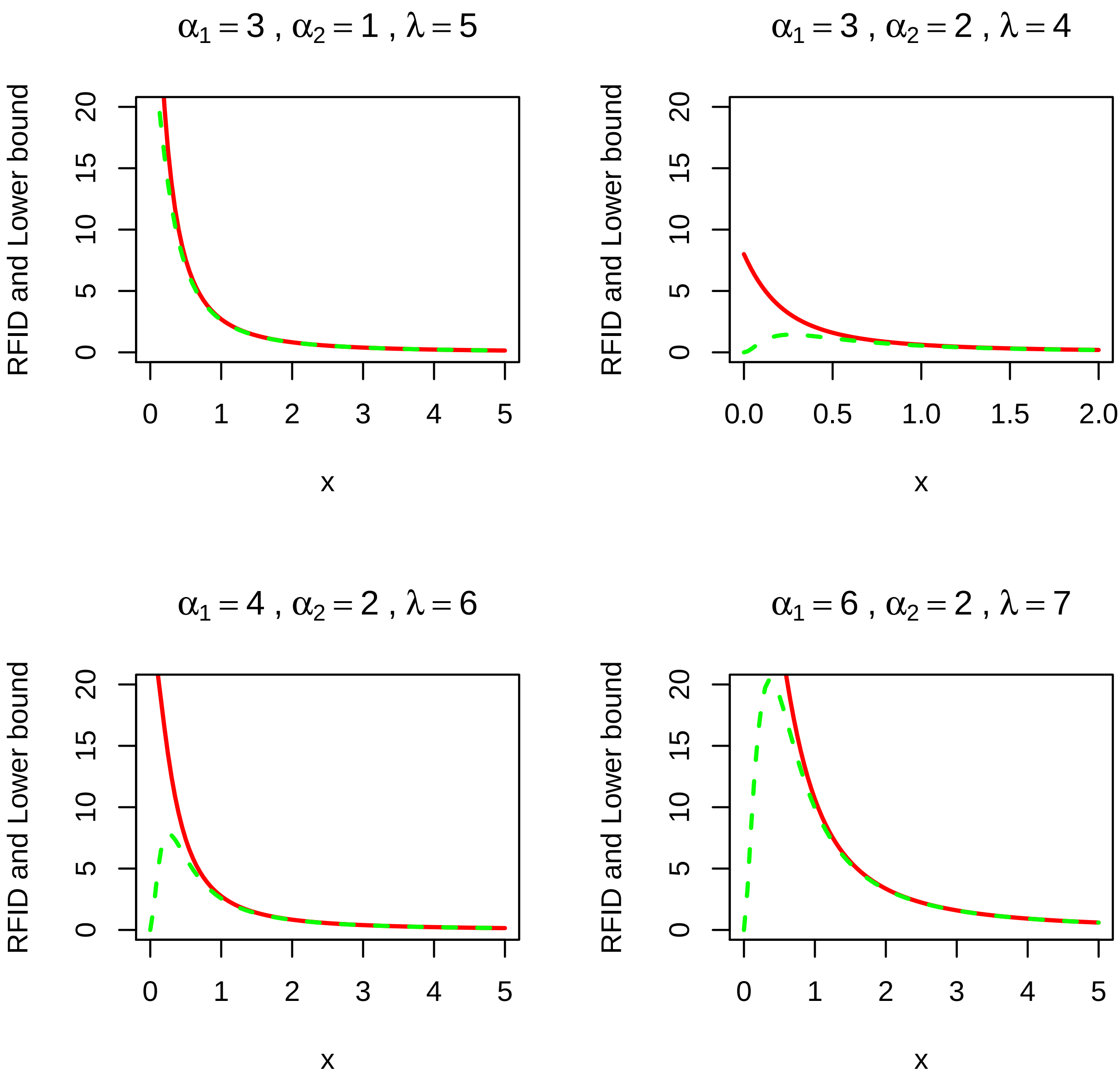

Figure 1: Plots of $D(f_1, f_2; t)$ measure (solid line) and it's corresponding lower bound (dashed line) for some selected values of α_1, α_2 and λ.

Proof : Since $E\Big[\rho_g(Y)|Y > t\Big] = -r_g(t)$ and using the Theorem 2.3 and assumption $X \leq_{hr} Y$, we have

$$\begin{aligned}\mathcal{D}(f,g;t) &\geq \Big(r_f(t) + E(\rho_g(X)|X > t)\Big)^2 \\ &\geq \Big(r_f(t) + E(\rho_g(Y)|Y > t)\Big)^2 \\ &\geq \Big(r_f(t) - r_g(t)\Big)^2,\end{aligned}$$

where the second inequality follows from the fact that $X \leq_{hr} Y$ implies $E(\phi(X)|X > t) \leq E(\phi(Y)|Y > t)$ for any increasing function $\phi(x)$ (see Shaked and Shanthikumar (2007)).

In a similar way, we can show that if $\rho_f(x)$ is decreasing and $Y \leq_{hr} X$, then

$$\mathcal{D}(g,f;t) \geq \Big(r_f(t) - r_g(t)\Big)^2.$$

From Theorem 2.4, we have

$$\Big(r_f(t) - r_g(t)\Big)^2 \leq \frac{\mathcal{D}(f,g;t) + \mathcal{D}(g,f;t)}{2}.$$

Example 2.5. *Let X_i, $i = 1,2$, be distributed as gamma distribution with density*

$$f_i(x) = \frac{\lambda^{\alpha_i}}{\Gamma(\alpha_i)} x^{\alpha_i - 1} e^{-\lambda x}, \quad x > 0,\ \alpha_i, \lambda > 0,\ i = 1,2.$$

For $\alpha_1 > 2$, we have

$$D(f_1, f_2; t) = \lambda^2 \frac{(\alpha_2 - \alpha_1)^2 \Gamma(\alpha_1 - 2, \lambda t)}{\Gamma(\alpha_1, \lambda t)},$$

where $\Gamma(\alpha_1, t) = \int_t^\infty x^{\alpha_1 - 1} e^{-x} dx$, is incomplete gamma function. On other hand, the hazard rate function is

$$r_{f_i}(t) = \frac{\lambda^{\alpha_i} t^{\alpha_i - 1} e^{-\lambda t}}{\Gamma(\alpha_1, \lambda t)},\ \ i = 1,2.$$

Now using the Theorem 2.4, a lower bound for $D(f_1, f_2; t)$ is given as

$$D(f_1, f_2; t) \geq \Big(\frac{\lambda^{\alpha_1} t^{\alpha_1 - 1} e^{-\lambda t}}{\Gamma(\alpha_1, \lambda t)} - \frac{\lambda^{\alpha_2} t^{\alpha_2 - 1} e^{-\lambda t}}{\Gamma(\alpha_2, \lambda t)}\Big)^2.$$

The plots of $D(f_1, f_2; t)$ measure and it's corresponding lower bound have been plotted in Figure 1.

We now give some theorems corresponding to RFID measure between densities of two transformed variables.

Theorem 2.6. *Given two random variables X and Y with absolutely continuous density functions f and g, respectively, and ϕ be a nonnegative increasing, twice differentiable and invertible function. Then*

$$\mathcal{D}(f_\phi, g_\phi; t) = E\Big[\frac{1}{\phi'(X)^2}\left[\frac{f'(X)}{f(X)} - \frac{g'(X)}{g(X)}\right]^2 \Big| X > \phi^{-1}(t)\Big]. \tag{9}$$

Proof: *Using the definition of dynamic Fisher information distance and transformed variables $\phi(X)$ and $\phi(Y)$,* *we have*

$$\begin{aligned}
\mathcal{D}(f_\phi, g_\phi; t) &= \frac{1}{\bar{F}(\phi^{-1}(t))}\int_t^\infty \left[\frac{(\frac{f(\phi^{-1}(x))}{\phi'(\phi^{-1}(x))})'}{\frac{f(\phi^{-1}(x))}{\phi'(\phi^{-1}(x))}} - \frac{(\frac{g(\phi^{-1}(x))}{\phi'(\phi^{-1}(x))})'}{\frac{g(\phi^{-1}(x))}{\phi'(\phi^{-1}(x))}}\right]^2 \frac{f(\phi^{-1}(x))}{\phi'(\phi^{-1}(x))}dx \\
&= \frac{1}{\bar{F}(\phi^{-1}(t))}\int_{\phi^{-1}(t)}^\infty \frac{1}{\phi'(x)^2}\left[\frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}\right]^2 f(x)dx \\
&= = E\Big[\frac{1}{\phi'(X)^2}\left[\frac{f'(X)}{f(X)} - \frac{g'(X)}{g(X)}\right]^2 \Big|X > \phi^{-1}(t)\Big].
\end{aligned}$$

As a special case, when ϕ is linear, that is $\phi(x) = ax + b$, we get

$$\begin{aligned}
\mathcal{D}(f_\phi, g_\phi; t) &= \frac{1}{\bar{F}((t-b)/a)}\int_{(t-b)/a}^\infty \frac{1}{a^2}[\frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}]^2 f(x)dx \\
&= \frac{1}{a^2}\mathcal{D}(f, g; \frac{t-b}{a}).
\end{aligned}$$

Theorem 2.7. *Given two random variables X and Y with absolutely continuous density functions f and g, respectively, and ϕ be a differentiable and invertible function. Then:*

(i) If ϕ is convex and increasing, $\mathcal{D}(f_\phi, g_\phi; t) \leq \frac{1}{(\phi'(\phi^{-1}(t)))^2}\mathcal{D}(f, g; \phi^{-1}(t))$;

(ii) If ϕ is concave and decreasing, $\mathcal{D}(f_\phi, g_\phi; t) \geq \frac{1}{(\phi'(\phi^{-1}(t)))^2}\mathcal{D}(f, g; \phi^{-1}(t))$.

Proof: By using the assumptions and Theorem 2.6, we have

$$\begin{aligned}
\mathcal{D}(f_\phi, g_\phi; t) &= \frac{1}{\bar{F}(\phi^{-1}(t))}\int_{\phi^{-1}(t)}^\infty \frac{1}{\phi'(x)^2}\left[\frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}\right]^2 f(x)dx \\
&\leq \frac{1}{\left(\phi'(\phi^{-1}(t))\right)^2}\frac{1}{\bar{F}(\phi^{-1}(t))}\int_{\phi^{-1}(t)}^\infty \left[\frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}\right]^2 f(x)dx \\
&= \frac{1}{\left(\phi'(\phi^{-1}(t))\right)^2}\mathcal{D}(f, g; \phi^{-1}(t)),
\end{aligned}$$

as required for Part (i). Part (ii) can be proved similarly.

Theorem 2.8. *Let X, Y and Z have densities f, g and h, and distribution functions F, G and H, respectively. Then:*

(i) If h is increasing, $\mathcal{D}(f_H, g_H; H(t)) \leq \frac{1}{h(t)^2}\mathcal{D}(f, g; t)$;

(ii) If h is decreasing, $\mathcal{D}(f_H, g_H; H(t)) \geq \frac{1}{h(t)^2}\mathcal{D}(f, g; t)$.

Proof: From Theorem 2.7, and the assumption that g is increasing, we have

$$\begin{aligned}
\mathcal{D}(f_H, g_H; H(t)) &= \frac{1}{\bar{F}(t)}\int_t^\infty \frac{1}{h(x)^2}[\frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}]^2 f(x)dx \\
&\leq \frac{1}{\bar{F}(t)}\int_t^\infty \frac{1}{h(t)^2}[\frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}]^2 f(x)dx \\
&= \frac{1}{h(t)^2}\mathcal{D}(f, g; t),
\end{aligned}$$

as required for Part (i). Using the assumption that g is decreasing, Part (ii) can be proved in similar manner.

Theorem 2.9. *Let X and Y be lifetime variables with the densities f and g, the score functions $\rho_f(x)$ and $\rho_g(x)$, and cumulative hazard functions $R_f(x)$ and $R_g(x)$, respectively.Then, for $k > 0$, we have*

$$E[L_\rho^k(X)R_f(X)] = \int_0^\infty f(u)E[L_\rho^k(X)|X > u]du,$$

where $L_\rho(x) = |\rho_f(x) - \rho_g(x)|$ and $R_f(t) = \int_0^t r_f(x)dx$.

Proof: We have

$$\begin{aligned} E\Big[L_\rho^k(X)R_f(X)\Big] &= \int_0^\infty L_\rho^k(x)R_f(x)f(x)dx \\ &= \int_0^\infty \Big\{\int_0^x r_f(u)du\Big\}L_\rho^k(x)f(x)dx \\ &= \int_0^\infty r_f(u)\Big\{\int_u^\infty L_\rho^k(x)f(x)dx\Big\}du \\ &= \int_0^\infty f(u)E\Big[L_\rho^k(X)|X > u\Big]du. \end{aligned}$$

The theorem has the following special cases.

- When $k = 1$, from $E[\rho_f(X)|X > u] = -r_f(u)$, we get

$$E[L_\rho(X)R_f(X)] = \begin{cases} -E[r_f(X)] - \int_0^\infty f(u)E[\rho_g(X)|X > u]du, & \rho_g(x) \leq \rho_f(x), \\ E[r_f(X)] + \int_0^\infty f(u)E[\rho_g(X)|X > u]du, & \rho_f(x) \leq \rho_g(x). \end{cases}$$

- For $k = 2$, based on the definition of $\mathcal{D}(f, g; t)$, we can obtain

$$\begin{aligned} E[L_\rho^2(X)R_f(X)] &= \int_0^\infty f(u)E[L_\rho^2(X)|X > u]du \\ &= E[\mathcal{D}(f, g; X)]. \end{aligned}$$

Let X and Y be continuous variables with density functions f and g, respectively. Variable X is said to be less than variable Y in likelihood ratio order, $X \leq_{lr} Y$, if $\frac{g(x)}{f(x)}$ is increasing in x for all x in the union of supports of X and Y.

Definition 2.10. Let X and Y be two random variables with residual Fisher information distance $\mathcal{D}(f, g; t)$ and $\mathcal{D}(g, f; t)$, respectively. X is said to be less than Y in residual Fisher information distance, denoted by $X \leq_{RFID} Y$, if $\mathcal{D}(g, f; t) \leq \mathcal{D}(g, f; t)$, for all t.

Theorem 2.11. *Let X and Y have densities f and g, respectively.*

(i) Assume that $\frac{f}{g}$ is log-concave. If $X \leq_{lr} Y$, then $X \leq_{RFID} Y$,

(ii) Assume that $\frac{g}{f}$ is log-convex. If $Y \leq_{lr} X$, then $X \leq_{RFID} Y$.

Proof: $X \leq_{lr} Y$ implies that $\frac{f'(x)}{f(x)} \leq \frac{g'(x)}{g(x)}$ for all x. Since $\frac{f}{g}$ is log-concave, therefore, $\Big(\rho_f(x) - \rho_g(x)\Big)^2$ is increasing. Hence from the fact that $X \leq_{lr} Y \Rightarrow X \leq_{hr} Y$, we get

$$\frac{\int_t^b \Big(\rho_f(x) - \rho_g(x)\Big)^2 f(x)dx}{\bar{F}(t)} \leq \frac{\int_t^b \Big(\rho_f(x) - \rho_g(x)\Big)^2 g(x)dx}{\bar{G}(t)}, \tag{10}$$

where the inequality follows from the fact that $X \leq_{hr} Y$ implies $E(\phi(X)|X > t) \leq E(\phi(Y)|Y > t)$ for any increasing function $\phi(x)$ (see, Shaked, and Shanthikumar (2007)). From (10), Part (i) is proved. Similarly, Part (ii) follows part (i) from the assumption that $\frac{g}{f}$ is log-convex.

Theorem 2.12. *Let X, Y and Z have densities f, g and h, respectively. If $Z \leq_{lr} Y \leq_{lr} X$, then*

$$\mathcal{D}(f,g;t) \leq \mathcal{D}(f,h;t).$$

Proof: $Z \leq_{lr} Y \leq_{lr} X$ implies that $\frac{f'(x)}{f(x)} \leq \frac{g'(x)}{g(x)} \leq \frac{h'(x)}{h(x)}$ for all x. From this, we have

$$\left(\rho_f(x) - \rho_g(x)\right)^2 \leq \left(\rho_f(x) - \rho_h(x)\right)^2.$$

Hence, we have

$$\frac{\int_t^b \left(\rho_f(x) - \rho_g(x)\right)^2 f(x)dx}{\bar{F}(t)} \leq \frac{\int_t^b \left(\rho_f(x) - \rho_h(x)\right)^2 f(x)dx}{\bar{F}(t)},$$

as required.

3. RFID measure for equilibrium distributions

Assume that $\bar{F}(x)$ is the survival function of a nonnegative continuous random variable X with finite mean μ. The random variable X_e is said to be the equilibrium random variable corresponding to the random variable X, if the density function of X_e is given by

$$f_e(x) = \frac{\bar{F}(x)}{\mu}, \quad x > 0. \tag{11}$$

The equilibrium distributions arise in renewal theory as the asymptotic distributions of the waiting time till the next event and the time since the last event at time t. Before presenting the theorem, we recall that the mean residual lifetime (MRL) of continuous random variable X with survival function $\bar{F}$ is defined at time t as

$$m(t) = E(X - t|X > t) = \frac{\int_t^b \bar{F}(x)dx}{\bar{F}(t)},$$

provided that $\bar{F}(t) > 0$. Note that $m(0) = \mu$ is the mean of X.

Let X and Y be two continuous random variables with density functions f and g and corresponding equilibrium densities f_e and g_e, respectively.

Theorem 3.1. *The RFID between two equilibrium distributions f_e and g_e can be represented as*

$$\mathcal{D}(f_e, g_e; t) = \frac{E\left[\frac{\left(r_f(X) - r_g(X)\right)^2}{r_f(X)}|X > t\right]}{m_f(t)},$$

where $m_f(t)$ denote the MRL of variable X.

Proof: From the definition of RFID measure, we have

$$\begin{aligned}\mathcal{D}(f_e, g_e; t) &= \frac{\int_t^b \left\{\frac{f_e'(x)}{f_e(x)} - \frac{g_e'(x)}{g_e(x)}\right\}^2 \bar{F}(x)dx}{\int_t^b \bar{F}(x)dx} \\ &= \frac{\int_t^b \left\{r_f(x) - r_g(x)\right\}^2 \bar{F}(x)dx}{\int_t^b \bar{F}(x)dx} \\ &= \frac{\int_t^b \frac{\left\{r_f(X)-r_g(X)\right\}^2}{r_f(X)} f(x)dx}{\int_t^b \bar{F}(x)dx} \\ &= \frac{E\left[\frac{\left(r_f(X)-r_g(X)\right)^2}{r_f(X)} | X > t\right]}{m_f(t)}. \end{aligned} \tag{12}$$

4. RFI and RFID measures for escort and generalized escort distributions

Let X be a variable with density function f. Then, the variable X_e is said to be the escort random variable corresponding to X, if for any positive real number η, the density function of X_e is given by

$$f_\eta(x) = \frac{f^\eta(x)}{\int f^\eta(x)dx},$$

provided $\int f^\eta(x)dx < \infty$. It should be noted that this is a special case of weighted distributions in which the weight function has been chosen to be $w(x) = f^{\eta-1}(x)$. An extension of escort distribution is known as generalized escort distribution. Assume that variables X and Y have density functions f and g, respectively. Then, the corresponding generalized escort density is given by

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

Note that this again is a special case of weighted distributions in which the weight function has been chosen to be $w(x) = \left\{\frac{f(x)}{g(x)}\right\}^{\eta-1}$. Escort distributions have found many applications in different areas. The significance of this distribution in chaos theory and thermodynamics has been emphasized by Beck and Schögl (1995). Further, this distribution finds application in coding theory wherein it can be used to modify the weights of the words with low probabilities (see, Bercher (2012)). Moreover, it was proved by Bercher (2012) that escort distributions can be used as one of the solutions for a minimum Kullback-Liebler discrimination related to state transition model framework. For more details, see Asadi et al. (2018). The following theorems present RFI and RFID measures associated with the escort and generalized escort distributions.

Definition 4.1. Let X, Y and Z be three random variables with score functions $\rho_f(x)$, $\rho_g(x)$ and $\rho_h(x)$ respectively. It is said that Z has additive proportional score function if $\rho_h(x) = \eta\rho_f(x) + (1-\eta)\rho_g(x)$, where η is positive constant.

Theorem 4.2. *Let X, Y and Z be three non-negative absolutely continuous random variables with score functions $\rho_f(x)$,$\rho_g(x)$ and $\rho_h(x)$ respectively. If X, Y and Z have additive proportional score functions, then*

$$h(x) = \frac{f^\eta(x)g^{1-\eta}(x)}{\int_0^\infty f^\eta(x)g^{1-\eta}(x)dx}.$$

Proof: Since $\rho_h(x) = \eta\rho_f(x) + (1-\eta)\rho_g(x)$, then $\frac{h'(x)}{h(x)} = \eta\frac{f'(x)}{f(x)} + (1-\eta)\frac{g'(x)}{g(x)}$ by integrating both sides of this differential equation we obtain

$$\log\Big(h(x)\Big) = \eta\log(f(x)) + (1-\eta)\log(g(x)) + c,$$

so we have $h(x) = Mf^k(x)g^{1-k}(x)$. Since h is density function, we get $M = \frac{1}{\int_0^\infty f^\eta(x)g^{1-\eta}(x)dx}$.

Theorem 4.3. *Let X, Y and Z be lifetime variables with densities f, g, h and score functions $\rho_f(x)$, $\rho_g(x)$ and $\rho_h(x)$, respectively. If $\rho_h(x) = \alpha\rho_f(x) + (1-\alpha)\rho_g(x)$, $0 < \alpha < 1$, then*

(i) $\mathcal{D}(h, f; t) \leq \frac{(1-\alpha)^2}{M_\alpha^t}\Big\{\alpha\mathcal{D}(f, g; t) + (1-\alpha)\mathcal{D}(g, f; t)\Big\}$;

(ii) $\mathcal{D}(h, f; t) = (1-\frac{1}{\alpha})^2\mathcal{D}(h, g; t)$,

where $M_\alpha^t = \int_t^\infty f^\alpha(x)g^{1-\alpha}(x)dx$.

Proof: By using Theorem 4.2, it is evident that $h(x)$ is the generalized escort distribution with order α. Let us define $M_\alpha^t = \int_t^\infty f^\alpha(x)g^{1-\alpha}(x)dx$, so we have

$$\begin{aligned}
\mathcal{D}(h, f; t) &= \int_t^\infty \Big\{\alpha\rho_f(x) + (1-\alpha)\rho_g(x) - \rho_f(x)\Big\}^2 \frac{f^\alpha(x)g^{1-\alpha}(x)}{M_\alpha^t}dx \\
&= (1-\alpha)^2\int_t^\infty \Big\{\rho_f(x) - \rho_g(x)\Big\}^2 \frac{f^\alpha(x)g^{1-\alpha}(x)}{M_\alpha^t}dx \\
&\leq \frac{(1-\alpha)^2}{M_\alpha^t}\int_t^\infty (\Big\{\rho_f(x) - \rho_g(x)\Big\}^2\Big\{\alpha f(x) + (1-\alpha)g(x)\Big\}dx \\
&= \frac{(1-\alpha)^2}{M_\alpha^t}\Big\{\alpha\bar{F}(t)\mathcal{D}(f, g; t) + (1-\alpha)\bar{G}(t)\mathcal{D}(g, f; t)\Big\} \\
&\leq \frac{(1-\alpha)^2}{M_\alpha^t}\Big\{\alpha\mathcal{D}(f, g; t) + (1-\alpha)\mathcal{D}(g, f; t)\Big\},
\end{aligned}$$

where first inequality follows from the inequality of arithmetic and geometric means. This completes proof of Part (i). Since

$$\mathcal{D}(h, f; t) = (1-\alpha)^2\int_t^\infty (\rho_f(x) - \rho_g(x))^2\frac{f^\alpha(x)g^{1-\alpha}(x)}{M_\alpha^t}dx$$

and

$$\mathcal{D}(h, g; t) = \alpha^2\int_t^\infty (\rho_f(x) - \rho_g(x))^2\frac{f^\alpha(x)g^{1-\alpha}(x)}{M_\alpha^t}dx,$$

so, we have $\mathcal{D}(h, f; t) = (1-\frac{1}{\alpha})^2\mathcal{D}(h, g; t)$, as required for Part (ii).

Theorem 4.4. *Let X, Y and Z be lifetime variables with densities f, g, h, distributions F, G, H and score functions $\rho_f(x)$, $\rho_g(x)$ and $\rho_h(x)$, respectively.*

(i) If $\rho_h(x) = 3\rho_f(x)$, *then* $\mathcal{I}\Big(F(Z), F(t)\Big) = \frac{4\bar{F}(t)}{M_1^t}\mathcal{I}(f, t)$;

(ii) If $\rho_h(x) = 3\rho_f(x) + (1-3)\rho_g(x)$, *then* $\mathcal{I}\Big(F(Z), F(t)\Big) = \frac{4\bar{F}(t)}{M_2^t}\mathcal{I}(G(X), t)$,

where $M_1^t = \int_t^\infty f^3(x)(x)dx$ *and* $M_2^t = \int_t^\infty f^3(x)g^{1-3}(x)dx$.

Proof: From Theorem 2.1 of [20], we have

$$\mathcal{I}(\phi(X), t) = \frac{1}{\bar{F}(\phi^{-1}(t))} \int_{\phi^{-1}(t)}^{\infty} \frac{1}{\phi'(x)^2} [\frac{f'(x)}{f(x)} - \frac{\phi''(x)}{\phi'(x)}]^2 f(x) dx. \tag{13}$$

Using the assumption, it is seen that $h(x)$ is escort distribution of $f(x)$ with order 3.
From $M_1^t = \int_t^\infty f^3(x)(x)dx$, we have

$$\begin{aligned}
\mathcal{I}(F(Z), F(t)) &= \int_t^\infty \frac{1}{f^2(x)} (3\rho_f(x) - \rho_f(x))^2 \frac{f^3(x)}{M_1^t} dx \\
&= 4 \int_t^\infty \rho_f^2(x) \frac{f(x)}{M_1^t} dx \\
&= \frac{4}{M_1^t} \int_t^\infty \rho_f^2(x) f(x) dx \\
&= \frac{4\bar{F}(t)}{M_1^t} \mathcal{I}(f, t),
\end{aligned}$$

as required for Part (i).

From assumption of Part (ii), $h(x)$ is generalized escort distribution of $f(x)$ and $g(x)$ with order 3. Since $M_2^t = \int_t^\infty f^3(x) g^{1-3}(x) dx$, we have

$$\begin{aligned}
\mathcal{I}\Big(F(Z), F(t)\Big) &= \int_t^\infty \frac{1}{f^2(x)} (3\rho_f(x) - 2\rho_g(x) - \rho_f(x))^2 \frac{f^3(x) g^{1-3}(x)}{M_2^t} dx \\
&= 4 \int_t^\infty \frac{(\rho_f(x) - \rho_g(x))^2}{g^2(x)} \frac{f(x)}{M_2^t} dx \\
&= \frac{4}{M_2^t} \int_t^\infty \frac{(\rho_f(x) - \rho_g(x))^2}{g^2(x)} f(x) dx \\
&= \frac{4\bar{F}(t)}{M_2^t} \mathcal{I}(G(X), t),
\end{aligned}$$

which is proves Part (ii).

5. Conclusion

In this paper, we have considered RFID measure and established some new properties associated with stochastic ordering in order to compare the lifetimes of two systems. We have shown that a lower bound for RFID measure can be expressed based on quadratic form of the corresponding hazard functions. In addition, we have provided some results of RFID measure in context of equilibrium and escort distributions. Besides, we have shown that RFID measure between two equilibrium distributions is connected with hazard and mean residual functions of underlying variable. Finally, we established some results associated with RFI and RFID measures of escort and generalized escort distributions.

References

[1] Asadi, M., Ebrahimi, N., Kharazmi, O., & Soofi, E. S. (2018). Mixture models, Bayes Fisher information, and divergence measures. IEEE Transactions on Information Theory, 65(4), 2316-2321.

[2] Barlow, R. E., & Proschan, F. (1975). Statistical theory of reliability and life testing: probability models. Florida State Univ Tallahassee.

[3] Beck, C., & Schögl, F. (1995). Thermodynamics of chaotic systems: an introduction (No. 4). Cambridge University Press.
[4] Bercher, J. F. (2012). A simple probabilistic construction yielding generalized entropies and divergences, escort distributions and q-Gaussians. Physica A: Statistical Mechanics and its Applications, 391(19), 4460-4469.
[5] S. G. Bobkov, G. P. Chistyakov, F. Götze, *Fisher information and the central limit theorem,* Probability theory and related fields, **159** (2014), 1-59.
[6] Cover, T. and Thomas, J. Elements of Information Theory, John Wiley & Sons, New York, NY, 1991.
[7] R. A. Fisher, *Tests of significance in harmonic analysis*, Proc R Soc Lond A Math Phys Sci, **125** (1929), 54-59.
[8] Frieden, B. R. Science from Fisher Information –A Unification, Cambridge University Press, Cambridge, 2004.
[9] Frieden, B. R. (2004). Science from Fisher Information: A Unification. Cambridge: Cambridge University Press. MR2069674
[10] Johnson, Oliver. Information theory and the central limit theorem. Vol. 8. London: Imperial College Press, 2004.
[11] Kharazmi, O., & Asadi, M. (2018). On the time-dependent Fisher information of a density function. Brazilian Journal of Probability and Statistics, 32(4), 795-814.
[12] Lehmann, E. L. and Casella, G. (1998). Information Theory and the Central Limit Theorem. New York: Springer.
[13] Shaked, M., & Shanthikumar, J. G. (2007). *Stochastic orders*. Springer Science & Business Media.
[14] F. Otto, C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal., **173** (2000),361-400.
[15] Shaked, M. and Shanthikumar, J. G. (2007). Stochastic Orders. New York: Springer Science and Business Media.
[16] Shao, J. (2003). Mathematical Statistics. New York: Springer. MR2002723.
[17] R. C. Venkatesan, A. Plastino, *Legendre transform structure and extremal properties of the relative Fisher information*, Phys. Lett. A, **20** (2014), 1341-1345.
[18] R. J. Yáñez, P. Sánchez-Moreno, A. Zarzo, J. S. Dehesa, *Fisher information of special functions and second-order differential equations*, J. Math. Phys., **49** (2008), 082104.
[19] T. Yamano, *Phase space gradient of dissipated work and information: A role of relative Fisher information*, J. Math. Phys., **54** (2013): 113301.
[20] P. Zebers, Some new results on the architecture, training process, and estimation error bounds for learning machines. 2002. PhD Thesis. The University of Arizona.