REVSTAT – Statistical Journal
Volume 17, Number 3, July 2019, 297–320

ON WEIGHTED KULLBACK–LEIBLER DIVERGENCE FOR DOUBLY TRUNCATED RANDOM VARIABLES

Authors: Rajesh Moharana
– Department of Mathematics, National Institute of Technology Rourkela, Rourkela-769008, India
rajeshmoharana31@gmail.com

Suchandan Kayal
– Department of Mathematics, National Institute of Technology Rourkela, Rourkela-769008, India
suchandan.kayal@gmail.com , kayals@nitrkl.ac.in

Received: July 2016 Revised: January 2017 Accepted: March 2017

Abstract:

- In this communication, we study doubly truncated weighted Kullback–Leibler divergence (KLD) between two nonnegative random variables. The proposed measure is a generalization of the dynamic weighted KLD introduced by Yasaei Sekeh *et al.* (2013). In reliability theory and survival analysis, it plays a significant role to study several aspects of a system when lifetimes fall in a time interval. It is showed that under some conditions, the proposed measure determines the distribution function uniquely. Further, characterization theorems for various lifetime distributions are proved. The effect of the monotone transformation on the proposed measure is studied. Some inequalities and bounds in terms of useful measures are obtained and finally, few applications are provided.

Key-Words:

- *weighted Kullback–Leibler divergence; generalized failure rate; weighted geometric vitality function; proportional (reversed) hazard model; likelihood ratio order.*

AMS Subject Classification:

- 94A17, 62E10, 62N05, 60E15.

1. INTRODUCTION

Kullback–Leibler divergence (see Kullback and Leibler, 1951) is an important measure in information theory, which has proven to be useful in reliability analysis and other related fields. It measures similarity (closeness) between two statistical distributions. To be specific, let X and Y be two nonnegative absolutely continuous random variables associated with probability density functions (pdf) f and g, and cumulative distribution functions (cdf) F and G, respectively. Then the KLD between f and g is given by

$$D_{KL}(X||Y) = \int_0^{+\infty} f(x) \ln\Big(\frac{f(x)}{g(x)}\Big) dx = E_f\Big(\ln\Big(\frac{f(X)}{g(X)}\Big)\Big), \tag{1.1}$$

where "ln" stands for the natural logarithm. We remark that $D_{KL}(X||Y)$ is nonnegative, not symmetric in f and g, zero if the distributions match exactly. It is scale invariant, that is, for two nonnegative random variables $Z_1 = aX$ and $Z_2 = aY$ with $a > 0$, we have $D_{KL}(X||Y) = D_{KL}(Z_1||Z_2)$. Note that $D_{KL}(X||Y)$ given by (1.1), which is a special case of Csiszar's ϕ-divergence measure can be viewed as a measure of the information loss in the fitted model relative to that in the reference model. For some recent development on KLD, we refer to Kasza and Solomon (2015) and Sankaran *et al.* (2016).

In recent past, there have been considerable interest to enlarge the concept of uncertainty by introducing nonnegative weight function. Belis and Guiasu (1968) first proposed the notion of (discrete) weighted entropy. It takes two kind of uncertainty into consideration. One of them is related to objective probability and other is related to utility. In analogy to Belis and Guiasu (1968), Di Crescenzo and Longobardi (2006) considered the weighted differential entropy for a nonnegative absolutely continuous random variable X as $S^w(X) = -\int_0^{+\infty} x f(x) \ln f(x) dx$. It is shift dependent, though the differential entropy $S(X) = -\int_0^{+\infty} f(x) \ln f(x) dx$ is not. Besides weighted differential entropy, several authors introduced and studied some other weighted information measures. In this direction, we refer to Suhov and Yasaei Sekeh (2015), Mirali *et al.* (2017), Nourbakhsh *et al.* (2016), and Rajesh *et al.* (2017).

Recently, based on the concept of weighted differential entropy, Yasaei Sekeh *et al.* (2013) considered weighted KLD as

$$D^w_{KL}(X||Y) = \int_0^{+\infty} x f(x) \ln\Big(\frac{f(x)}{g(x)}\Big) dx = E_f\Big(X \ln\Big(\frac{f(X)}{g(X)}\Big)\Big), \tag{1.2}$$

which takes into account the qualitative characteristic related to utility. To illustrate the importance of the weighted KLD, we consider the following example.

Example 1.1. Let X_1 and Y_1 be two nonnegative absolutely continuous random variables with pdfs $f_1(x) = 2x$, $0 < x < 1$ and $g_1(x) = 2(1-x)$, $0 < x < 1$, respectively. We consider another random variables X_2 and Y_2 with pdfs $f_2(x) = x/2,\ 0<x<2$ and $g_2(x) = (2-x)/2,\ 0<x<2$, respectively. From (1.1) we obtain $D_{KL}(X_1||Y_1)=1$ and $D_{KL}(X_2||Y_2)=1$. Also, from (1.2) we get $D^w_{KL}(X_1||Y_1)=1$ and $D^w_{KL}(X_2||Y_2) = 2$. Thus, from the objective probability point of view, KLD measures are same. But, when we take the qualitative characteristics into consideration, they differ. Here, $D^w_{KL}(X_1||Y_1) < D^w_{KL}(X_2||Y_2)$.

Note that when the weight function "x" equals to 1, $D^w_{KL}(X||Y)$ coincides with the standard KLD given by (1.1). Yasaei Sekeh *et al.* (2013) considered weighted KLD between two residual (truncated from left) lifetime distributions and two past (truncated from right) lifetime distributions. But there exist several situations in real life, where statistical data are not only truncated from left or right side, but also from both sides. When data are truncated from left and right sides, we call it as doubly truncated. Doubly truncated data play a central role in various statistical analysis of survival data. Doubly truncated failure time occurs if the failure of an individual occurs within a certain interval. In medical science, the induction time data in AIDS are doubly truncated, since HIV was unknown to us before the year 1982. Also, during a survival experiment, sometimes it is required to collect data after an engineering system starts operating and before it fails. Let X denote the lifetime of a system. Then the conditional random variable $(X|x < X < y)$ is known as doubly truncated lifetime. That is, event time of an individual lies within a certain time interval (x, y) is only observed. Therefore, an individual is not observed if it's event time does not fall in this predefined interval. Hence, information on the subject outside this interval is not available to the investigator. Misagh and Yari (2012) considered doubly truncated (truncated from both sides) KLD as

$$D_{KL}(X||Y;t_1,t_2) = \int_{t_1}^{t_2} \frac{f(x)}{\Delta F} \ln\Big(\frac{f(x)/\Delta F}{g(x)/\Delta G}\Big)dx, \tag{1.3}$$

where $\Delta F = F(t_2)-F(t_1)$, $\Delta G = G(t_2)-G(t_1)$ and $(t_1,t_2) \in \mathbb{D} = \{(x,y)|F(x) < F(y) \text{ and } G(x) < G(y)\}$. In this paper, we consider doubly truncated weighted KLD between f and g, which is given by

$$D^w_{KL}(X||Y;t_1,t_2) = \int_{t_1}^{t_2} x\frac{f(x)}{\Delta F} \ln\Big(\frac{f(x)/\Delta F}{g(x)/\Delta G}\Big)dx. \tag{1.4}$$

Note that $D^w_{KL}(X||Y;t_1,t_2)$ is a generalization of the measures considered by Yasaei Sekeh *et al.* (2013) in the sense that it reduces to the weighted KLD between two residual lives and two past lives, when t_2 tends to $+\infty$ and t_1 tends to 0, respectively. Mathematically,

$$\lim_{t_2\to+\infty} D^w_{KL}(X||Y;t_1,t_2) = \int_{t_1}^{+\infty} x\frac{f(x)}{\bar F(t_1)} \ln\Big(\frac{f(x)/\bar F(t_1)}{g(x)/\bar G(t_1)}\Big)dx$$

and

$$\lim_{t_1\to 0} D^w_{KL}(X||Y;t_1,t_2) = \int_{0}^{t_2} x\frac{f(x)}{F(t_2)} \ln\Big(\frac{f(x)/F(t_2)}{g(x)/G(t_2)}\Big)dx.$$

The doubly truncated weighted KLD given by (1.4) may take negative values, though the doubly truncated KLD always take nonnegative values. The expression given by (1.4) measures the weighted discrepancy between two systems with lifetimes X and Y, which have survived up to time t_1 and have seen to be down at time t_2. In fact, (1.4) can be used to measure divergence between two distributions having different supports. We consider the following example to illustrate this.

Example 1.2. Let X and Y be two nonnegative absolutely continuous random variables with pdfs $f(x) = \frac{3}{4}(2x + x^2),\ 0 < x < 1$ and $g(x) = \frac{1}{8}(1 + 3y)$, $0 < y < 2$, respectively. As supports of the distributions are different, therefore, the expression given by (1.1) can not be used to compute the divergence between f and g. In this situation, one may use (1.4) for finding divergence. In particular, $D^w_{KL}(X||Y; 0.1, 0.5) = 0.033427$ and $D^w_{KL}(X||Y; 0.3, 0.9) = 0.014291$.

Again, the doubly truncated weighted KLD given by (1.4) can be expressed in terms of the doubly truncated weighted Shannon entropy and the doubly truncated weighted inaccuracy as

$$D^w_{KL}(X||Y; t_1, t_2) = -S^w(X; t_1, t_2) + I^w(X||Y; t_1, t_2), \tag{1.5}$$

where $S^w(X; t_1, t_2)$ and $I^w(X||Y; t_1, t_2)$ are defined in the next section.

The paper is arranged as follows. First, in Section 2, we recall some preliminary definitions. Few characterization results are proved in Section 3. Further, various lifetime distributions are characterized from the relationships among reliability measures. In Section 4, we analysis the effect of the affine transformations on the doubly truncated weighted KLD. Then, few inequalities and bounds are obtained in Section 5. Section 6 contains few examples in support of the results obtained in Section 5. Finally, some concluding remarks have been added in Section 7.

Throughout the paper, the random variables are taken to be nonnegative and absolutely continuous. The terms increasing and decreasing are used in nonstrict sense. The differentiation, integration and expectation wherever used are assumed to exist.

2. PRELIMINARY RESULTS

In this section, we present some preliminary definitions and results which are useful for the rest of the paper. Let X and Y be two nonnegative absolutely continuous random variables with pdfs f and g, and cdfs F and G, respectively.

Then the generalized failure rate (GFR) functions of $(X|t_1 < X < t_2)$ and $(Y|t_1 < Y < t_2)$ are given by (see Navarro and Ruiz, 1996)

$$h_1^X(t_1, t_2) = \frac{f(t_1)}{\Delta F}, \qquad h_2^X(t_1, t_2) = \frac{f(t_2)}{\Delta F} \tag{2.1}$$

and

$$h_1^Y(t_1, t_2) = \frac{g(t_1)}{\Delta G}, \qquad h_2^Y(t_1, t_2) = \frac{g(t_2)}{\Delta G}, \tag{2.2}$$

respectively for $(t_1, t_2) \in \mathbb{D}$. Note that when t_2 tends to $+\infty$, $h_1^X(t_1, t_2)$ reduces to the failure rate of X, and when t_1 tends to zero, $h_2^X(t_1, t_2)$ reduces to reversed failure rate of X. Similarly for the random variable Y. Navarro and Ruiz (1996) showed that the distribution function can be uniquely determined by GFR functions.

Definition 2.1. Let X be a nonnegative random variable with pdf f and cdf F. Then the generalized conditional mean (GCM) of a doubly truncated random variable $(X|t_1 < X < t_2)$ is given by

$$\mu_X(t_1, t_2) = E(X|t_1 < X < t_2) = \int_{t_1}^{t_2} \frac{xf(x)}{\Delta F} dx. \tag{2.3}$$

For some characterizations based on (2.3), one may refer to Ruiz and Navarro (1996).

Definition 2.2. Let X be a nonnegative random variable with pdf f and cdf F. Then the geometric vitality function for $(X|t_1 < X < t_2)$ is defined as

$$G_X(t_1, t_2) = E(\ln X|t_1 < X < t_2) = \int_{t_1}^{t_2} \frac{\ln x f(x)}{\Delta F} dx. \tag{2.4}$$

Note that $G_X(t_1, t_2)$ gives the geometric mean life of X truncated at two points t_1 and t_2. Nair and Rajesh (2000) gave some applications of geometric vitality function. Sunoj *et al.* (2009) discussed few properties of this measure and showed that it determines the distribution function uniquely. The weighted version of the measure given by (2.4) is defined as follows.

Definition 2.3. The weighted geometric vitality function of a nonnegative random variable X with pdf f and cdf F is given by

$$G_X^w(t_1, t_2) = E(X \ln X|t_1 < X < t_2) = \int_{t_1}^{t_2} \frac{x \ln x f(x)}{\Delta F} dx. \tag{2.5}$$

Definition 2.4. For a nonnegative random variable X with pdf f and cdf F, the doubly truncated weighted Shannon entropy is given by

$$S^w(X;t_1,t_2) = -\int_{t_1}^{t_2} x\frac{f(x)}{\Delta F}\ln\left(\frac{f(x)}{\Delta F}\right)dx. \tag{2.6}$$

In the following we consider the definition of weighted inaccuracy measure between two doubly truncated random variables.

Definition 2.5. The doubly truncated weighted inaccuracy measure between two nonnegative random variables X and Y is given by

$$I^w(X||Y;t_1,t_2) = -\int_{t_1}^{t_2} x\frac{f(x)}{\Delta F}\ln\left(\frac{g(x)}{\Delta G}\right)dx. \tag{2.7}$$

Next we recall the following important definition from Shaked and Shanthikumar (2007).

Definition 2.6. Let X and Y be two random variables with pdfs f and g, and cdfs F and G, respectively. We say that X is larger than Y in likelihood ratio order, denoted by $X \geq^{lr} Y$ if $f(x)/g(x)$ is increasing in x.

Log-sum inequality: Let m be a sigma finite measure. If f and g are positive and m integrable, then

$$\int f\log\left(\frac{f}{g}\right)dm \geq \left[\int fdm\right]\log\left[\frac{\int fdm}{\int gdm}\right]. \tag{2.8}$$

3. CHARACTERIZATIONS

In this section, we obtain some characterization results which may be used to describe probability distributions. The general characterization problem is to determine when the doubly truncated weighted KLD uniquely determines the distribution function. Yasaei Sekeh *et al.* (2013) showed that under some conditions, the weighted KLD for two residual and past lifetime distributions characterizes the distribution function uniquely. In the following theorem, we show that using relationship between doubly truncated weighted KLD and GCM, and under the condition on GFR functions, one can characterize one of the distributions when other is known.

Theorem 3.1. *Let X and Y be two absolutely continuous nonnegative random variables with pdfs f and g and cdfs F and G, respectively, such that*

$h_i^Y(t_1,t_2) \le h_i^X(t_1,t_2)$, $i=1,2$. *Then the doubly truncated weighted KLD given by* (1.4) *characterizes the distribution function* G *(or* F*), when* F *(or* G*) is known, provided* $D_{KL}^w(X||Y;t_1,t_2)=\mu_X(t_1,t_2)$.

Proof: Differentiating (1.4) partially with respect to t_1 (for any fixed t_2) and t_2 (for any fixed t_1), we get after simplification

$$\begin{aligned}\frac{\partial D_{KL}^w(X||Y;t_1,t_2)}{\partial t_1} &= h_1^X(t_1,t_2)\Big[D_{KL}^w(X||Y;t_1,t_2)+\mu_X(t_1,t_2)\\ &\quad +t_1\ln\Big(\frac{h_1^Y(t_1,t_2)}{h_1^X(t_1,t_2)}\Big)\Big]-h_1^Y(t_1,t_2)\mu_X(t_1,t_2)\end{aligned} \tag{3.1}$$

and

$$\begin{aligned}\frac{\partial D_{KL}^w(X||Y;t_1,t_2)}{\partial t_2} &= -h_2^X(t_1,t_2)\Big[D_{KL}^w(X||Y;t_1,t_2)+\mu_X(t_1,t_2)\\ &\quad +t_2\ln\Big(\frac{h_2^Y(t_1,t_2)}{h_2^X(t_1,t_2)}\Big)\Big]+h_2^Y(t_1,t_2)\mu_X(t_1,t_2).\end{aligned} \tag{3.2}$$

Moreover, differentiating (2.3) with respect to t_1 and t_2, we obtain

$$\frac{\partial \mu_X(t_1,t_2)}{\partial t_1}=h_1^X(t_1,t_2)\Big[\mu_X(t_1,t_2)-t_1\Big] \tag{3.3}$$

and

$$\frac{\partial \mu_X(t_1,t_2)}{\partial t_2}=-h_2^X(t_1,t_2)\Big[\mu_X(t_1,t_2)-t_2\Big], \tag{3.4}$$

respectively. Again, differentiating $D_{KL}^w(X||Y;t_1,t_2)=\mu_X(t_1,t_2)$ with respect to t_i (for fixed $t_j,\ j\neq i$), $i,j=1,2$ and using (3.1), (3.2), (3.3) and (3.4) we get

$$t_ih_i^X(t_1,t_2)\Big[1+\ln\Big(\frac{h_i^Y(t_1,t_2)}{h_i^X(t_1,t_2)}\Big)\Big]+[h_i^X(t_1,t_2)-h_i^Y(t_1,t_2)]\mu_X(t_1,t_2)=0. \tag{3.5}$$

The above equation given by (3.5) can be further written as

$$g_i(x)=t_i[1+\ln x]+(1-x)\mu_X(t_1,t_2)=0,\ \ i=1,2, \tag{3.6}$$

where $x=h_i^Y(t_1,t_2)/h_i^X(t_1,t_2)$ and $0<x<1$. Thus, for any fixed t_2 and arbitrary t_1, $h_1^Y(t_1,t_2)/h_1^X(t_1,t_2)$ is a positive solution of the equation $g_1(x)=0$. Also, for any fixed t_1 and arbitrary t_2, $h_2^Y(t_1,t_2)/h_2^X(t_1,t_2)$ is a positive solution of the equation $g_2(x)=0$. After some simple calculations, it is easy to show that both the solutions of $g_1(x)=0$ and $g_2(x)=0$ are unique. Hence, using the result that GFR functions uniquely determine the distribution function (see Navarro and Ruiz, 1996), the proof follows. This completes the proof of the theorem. □

Theorem 3.2. *Let X and Y be two absolutely continuous nonnegative random variables with pdfs f and g and cdfs F and G, respectively. Assume $h_1^Y(t_1,t_2) = \theta h_1^X(t_1,t_2)$ and $D_{KL}^w(X||Y;t_1,t_2) > (<)\,(\theta-1)\mu_X(t_1,t_2) - t_1\log\theta$, where $\theta > 0$. If $D_{KL}^w(X||Y;t_1,t_2)$ is strictly increasing (decreasing) in t_1 for fixed t_2, then $D_{KL}^w(X||Y;t_1,t_2)$ characterizes the distribution function uniquely.*

Proof: Under the given hypothesis, (3.1) is reduced to

$$\begin{aligned}\frac{\partial D_{KL}^w(X||Y;t_1,t_2)}{\partial t_1} &= h_1^X(t_1,t_2)\Big[D_{KL}^w(X||Y;t_1,t_2)\\ &\quad +(1-\theta)\mu_X(t_1,t_2)+t_1\ln\theta\Big].\end{aligned} \tag{3.7}$$

The expression in (3.7) can be written further as

$$\begin{aligned}u(x) &= \frac{\partial D_{KL}^w(X||Y;t_1,t_2)}{\partial t_1} - x\Big[D_{KL}^w(X||Y;t_1,t_2)+(1-\theta)\mu_X(t_1,t_2)+t_1\ln\theta\Big]\\ &= 0,\end{aligned} \tag{3.8}$$

where $x = h_1^X(t_1,t_2)$ is a positive solution of $u(x) = 0$. Since $D_{KL}^w(X||Y;t_1,t_2) > (<)(\theta-1)\mu_X(t_1,t_2) - t_1\log\theta$, therefore from (3.8) it is easy to show that

$$\lim_{x\to+\infty} u(x) = -\infty\ (+\infty). \tag{3.9}$$

Again, $D_{KL}^w(X||Y;t_1,t_2)$ is strictly increasing (decreasing) in t_1 for fixed t_2. Therefore,

$$\lim_{x\to 0} u(x) = \frac{\partial D_{KL}^w(X||Y;t_1,t_2)}{\partial t_1} > (<)0. \tag{3.10}$$

Differentiating (3.8) with respect to x we have $u'(x) = -[D_{KL}^w(X||Y;t_1,t_2) + (1-\theta)\mu_X(t_1,t_2) + t_1\ln\theta] < (>)\ 0$ implies $u(x)$ is a decreasing (increasing) function in $x > 0$. Hence, $x = h_1^X(t_1,t_2)$ is the only solution of $u(x) = 0$. This completes the proof of the theorem. □

Theorem 3.3. *Let X and Y be two absolutely continuous nonnegative random variables with pdfs f and g and cdfs F and G, respectively. Assume that for $\theta > 0$, $h_2^Y(t_1,t_2) = \theta h_2^X(t_1,t_2)$ and $D_{KL}^w(X||Y;t_1,t_2) < (>)(\theta-1)\mu_X(t_1,t_2) - t_2\ln\theta$. If $D_{KL}^w(X||Y;t_1,t_2)$ is strictly increasing (decreasing) in t_2 for fixed t_1, then $D_{KL}^w(X||Y;t_1,t_2)$ characterizes the distribution function uniquely.*

Proof: Proof follows along the similar arguments of that of the Theorem 3.2. Hence it is omitted. □

It is noted that the conditions used in the above theorems are sufficient. Henceforth, we present characterization theorems for some useful continuous distributions. Let X and Y be two nonnegative absolutely continuous random variables with cdfs F and G, pdfs f and g, hazard rate functions λ_F and λ_G and reversed hazard rate functions r_F and r_G, respectively. Then X and Y are said to satisfy the proportional hazard rate model (PHRM) and proportional reversed hazard rate model (PRHRM) if for some $\theta > 0$,

$$\bar{G}(x) = [\bar{F}(x)]^{\theta} \text{ and } G(x) = [F(x)]^{\theta}, \tag{3.11}$$

respectively, where $\bar{F} = 1 - F$ and $\bar{G} = 1 - G$. The constant θ is known as proportionality constant. Several researchers used PHRM for survival data analysis. See, for instant, Cox (1972), Ebrahimi and Kirmani (1996) and Nair and Gupta (2007). On the other hand, for various results on PRHRM, we refer to Gupta and Gupta (2007) and Sankaran and Gleeja (2008). In the following consecutive theorems, we present characterizations of the first and second kind Pareto distributions.

Theorem 3.4. *Let X and Y be two absolutely continuous nonnegative random variables with pdfs f and g and cdfs F and G, respectively. Assume that F and G satisfies PHRM with proportionality constant $\theta > 0$. Then for $i = 1, 2$, the following relationship of the form*

$$\begin{aligned} D^w_{KL}(X||Y; t_1, t_2) \;+\; \mu_X(t_1, t_2)\left[\ln f(t_i) + (1 + \alpha\theta)\ln t_i + \ln\left(\frac{h^Y_i(t_1, t_2)}{h^X_i(t_1, t_2)}\right)\right] \\ = (1 + \alpha\theta)G^w_X(t_1, t_2) + E(X \ln f(X)|t_1 < X < t_2), \end{aligned} \tag{3.12}$$

holds if and only if X follows Pareto-I distribution with cdf $F(x) = 1 - (\beta/x)^{\alpha}$, $x > \beta > 0,\ \alpha > 0$.

Proof: The "if part" can be proved easily. To prove the "only if part", we assume that (3.12) holds. Using (1.4) and after simplification, we get from (3.12)

$$\begin{aligned} \int_{t_1}^{t_2} x f(x) \ln\left(\frac{f(x)/\Delta F}{g(x)/\Delta G}\right)dx \;+\; \left[\ln f(t_i) + (1+\alpha\theta)\ln t_i + \ln\left(\frac{g(t_i)}{\Delta G}\right) - \ln\left(\frac{f(t_i)}{\Delta F}\right)\right] \\ \times \int_{t_1}^{t_2} x f(x)dx \;=\; (1 + \alpha\theta)\int_{t_1}^{t_2} x f(x)\ln x dx \;+\; \int_{t_1}^{t_2} x f(x)\ln f(x)dx. \end{aligned} \tag{3.13}$$

Differentiating (3.13) with respect to t_i and then further calculations lead to

$$g(t_i) = k t_i^{-(\alpha\theta+1)},\ i = 1, 2 \text{ and } k > 0.$$

Hence the required result follows. This completes the proof. □

Theorem 3.5. *Let X and Y be two absolutely continuous random variables as described in Theorem 3.4 and satisfying PHRM with proportionality constant $\theta > 0$. Then for $i = 1, 2$, the following relationship of the form*

$$\begin{aligned} D^w_{KL}(X\|Y;t_1,t_2) + \ \mu_X(t_1,t_2)\left[\ln f(t_i) + (1+\alpha\theta)\ln(t_i-\gamma+\beta) + \ln\Big(\frac{h_i^Y(t_1,t_2)}{h_i^X(t_1,t_2)}\Big)\right] \\ = (1+\alpha\theta)G^w_Z(t_1,t_2) + E(X\ln f(X)|t_1 < X < t_2), \end{aligned} \tag{3.14}$$

where $G^w_Z(t_1,t_2) = E(X\ln(X-\gamma+\beta)|t_1 < X < t_2)$ holds if and only if X follows Pareto-II distribution with cdf $F(x) = 1 - [1 + (\frac{x-\gamma}{\beta})]^{-\alpha},\ x > \gamma > 0,\ \alpha, \beta > 0.$

Proof: The "if part" is straightforward and hence omitted. To prove the "only if part", let us assume that (3.14) holds. Then from (3.14) and (1.4) we obtain

$$\begin{aligned} \int_{t_1}^{t_2} xf(x)\ln\Big(\frac{f(x)/\Delta F}{g(x)/\Delta G}\Big)dx + \Big[\ln f(t_i) \ + \ (1+\alpha\theta)\ln(t_i-\gamma+\beta) + \ln\Big(\frac{g(t_i)}{\Delta G}\Big) \\ -\ln\Big(\frac{f(t_i)}{\Delta F}\Big)\Big]\int_{t_1}^{t_2} xf(x)dx \ = \ (1+\alpha\theta)\int_{t_1}^{t_2} xf(x)\ln(x-\gamma+\beta)dx \\ + \int_{t_1}^{t_2} xf(x)\ln f(x)dx. \end{aligned} \tag{3.15}$$

Differentiating (3.15) with respect to $t_i, i = 1, 2$ and after some algebraic calculations, we get

$$g(t_i) = k(t_i - \gamma + \beta)^{-(1+\alpha\theta)},\ i = 1, 2 \text{ and } k > 0.$$

Hence the result follows. This completes the proof of the theorem. □

Here, below we present a characterization theorem for Weibull distribution.

Theorem 3.6. *Let X and Y be two absolutely continuous nonnegative random variables as mentioned in Theorem 3.4. Also, assume that they satisfy PHRM with proportionality constant $\theta > 0$. Then the following relationship of the form*

$$\begin{aligned} D^w_{KL}(X\|Y;t_1,t_2) \ + \ \mu_X(t_1,t_2)\Big[\lambda\theta t_i^p + \ln f(t_i) + (1-p)\ln t_i + \ln\Big(\frac{h_i^Y(t_1,t_2)}{h_i^X(t_1,t_2)}\Big)\Big] \\ +\lambda\theta\mu_{X^{p+1}}(t_1,t_2) \ = \ (1-p)G^w_{Z^*}(t_1,t_2) + E(X\ln f(X)|t_1 < X < t_2), \quad i = 1, 2, \end{aligned} \tag{3.16}$$

where $\mu_{X^{p+1}}(t_1,t_2) = E(X^{p+1}|t_1 < X < t_2)$ and $G^w_{Z^*}(t_1,t_2) = E(X\ln(X-\alpha)|t_1 < X < t_2)$ holds if and only if X follows Weibull distribution with cdf $F(x) = 1 - \exp(-\lambda x^p),\ x > 0, p > 0, \lambda > 0.$

Proof: The "if part" is straightforward. To prove the "only if part", we first assume that (3.16) holds. Then using (1.4) in (3.16), we get

$$\begin{aligned}\int_{t_1}^{t_2} xf(x)\ln\Big(\frac{f(x)/\Delta F}{g(x)/\Delta G}\Big)dx \;+\; \Big[\lambda\theta t_i^p + \ln f(t_i) + (1-p)\ln t_i + \ln\Big(\frac{g(t_i)}{\Delta G}\Big) \\ -\ln\Big(\frac{f(t_i)}{\Delta F}\Big)\Big]\int_{t_1}^{t_2} xf(x)dx \;=\; (1-p)\int_{t_1}^{t_2} xf(x)\ln(x-\alpha)dx \\ +\int_{t_1}^{t_2} xf(x)\ln f(x)dx - \lambda\theta\int_{t_1}^{t_2} x^{p+1}f(x)dx.\end{aligned} \tag{3.17}$$

Differentiating (3.17) with respect to $t_i,\ i=1,2$, and after some algebraic calculations, we obtain

$$g(t_i) = ct_i^{(p-1)}e^{-\lambda\theta t_i^p},\ i=1,2 \text{ and } c>0.$$

Hence the required result follows. This completes the proof. □

Remark 3.1. In particular, for $p=1,2$, the Theorem 3.6 provides characterization results of exponential distribution with cdf $F(x)=1-e^{-\lambda x}$, $x>0$, $\lambda>0$ and Rayleigh distribution with cdf $F(x)=1-e^{-\lambda x^2}$, $x>0,\lambda>0$, respectively.

Hereafter, we present results which characterize uniform and power distributions.

Theorem 3.7. *Let X and Y be two absolutely continuous nonnegative random variables as described in Theorem 3.4 and satisfying PRHRM with proportionality constant $\theta>0$. Then the following relationship of the form*

$$\begin{aligned}D_{KL}^w(X||Y;t_1,t_2) \;+\; \mu_X(t_1,t_2)\Big[\ln f(t_i) + (1-\theta)\ln(t_i-\alpha) + \ln\Big(\frac{h_i^Y(t_1,t_2)}{h_i^X(t_1,t_2)}\Big)\Big] \\ = (1-\theta)G_{Z^*}^w(t_1,t_2) + E(X\ln f(X)|t_1<X<t_2),\quad i=1,2,\end{aligned} \tag{3.18}$$

where $G_{Z^*}^w(t_1,t_2)=E(X\ln(X-\alpha)|t_1<X<t_2)$ and $\alpha<t_1<t_2<\beta$ holds if and only if X follows uniform distribution in the interval (α,β).

Proof: The "if part" is straightforward. To prove the "only if part", assume that (3.18) holds for $i=1$ and 2. Using (1.4), the above relation (3.18) further reduces to

$$\begin{aligned}\int_{t_1}^{t_2} xf(x)\ln\Big(\frac{f(x)/\Delta F}{g(x)/\Delta G}\Big)dx + \Big[\ln f(t_i) \;+\; (1-\theta)\ln(t_i-\alpha) + \ln\Big(\frac{g(t_i)}{\Delta G}\Big) \\ -\ln\Big(\frac{f(t_i)}{\Delta F}\Big)\Big]\int_{t_1}^{t_2} xf(x)dx \;=\; (1-\theta)\int_{t_1}^{t_2} xf(x)\ln(x-\alpha)dx \\ +\int_{t_1}^{t_2} xf(x)\ln f(x)dx,\end{aligned} \tag{3.19}$$

for $i = 1, 2$. Differentiating (3.19) with respect to $t_i, i = 1, 2$ and then simplifying further, we obtain

$$g(t_i) = c(t_i - \alpha)^{(\theta-1)},\ i = 1, 2 \text{ and } c > 0,$$

which gives the required result. This completes the proof. □

Theorem 3.8. *Let X and Y be two absolutely continuous nonnegative random variables as mentioned in Theorem 3.4 and satisfying PRHRM with proportionality constant $\theta > 0$. Then for $c > 0$, the following relationship of the form*

$$\begin{aligned} D^w_{KL}(X||Y; t_1, t_2) \ + \ \mu_X(t_1, t_2) & \left[\ln f(t_i) + (1 - c\theta) \ln t_i + \ln\left(\frac{h^Y_i(t_1, t_2)}{h^X_i(t_1, t_2)}\right)\right] \\ &= (1 - c\theta) G^w_X(t_1, t_2) + E(X \ln f(X)|t_1 < X < t_2), \quad i = 1, 2, \end{aligned} \tag{3.20}$$

where $G^w_X(t_1, t_2)$ is given by (2.5) holds if and only if X follows power distribution with cdf $F(x) = (\frac{x}{b})^c,\ 0 < x < b, c > 0$.

Proof: The "if part" is straightforward. To prove the "only if part", let us assume that (3.20) holds. Using (1.4), (3.20) further reduces to

$$\begin{aligned} &\int_{t_1}^{t_2} x f(x) \ln\left(\frac{f(x)/\Delta F}{g(x)/\Delta G}\right) dx \ + \ \left[\ln f(t_i) + (1 - c\theta) \ln t_i + \ln\left(\frac{g(t_i)}{\Delta G}\right) - \ln\left(\frac{f(t_i)}{\Delta F}\right)\right] \\ &\quad \times \int_{t_1}^{t_2} x f(x) dx \ = \ (1 - c\theta) \int_{t_1}^{t_2} x f(x) \ln x dx + \int_{t_1}^{t_2} x f(x) \ln f(x) dx. \end{aligned} \tag{3.21}$$

Differentiating (3.21) with respect to $t_i,\ i = 1, 2$, we get

$$g(t_i) = k t_i^{(c\theta-1)},\ i = 1, 2 \text{ and } k > 0,$$

which follows the required result. This completes the proof. □

4. MONOTONE TRANSFORMATIONS

In this section, we analysis the effect of the doubly truncated weighted KLD given by (1.4) under strictly monotone transformations. The following theorem is a generalization of the Theorem 4.13 of Yasaei Sekeh *et al.* (2013).

Theorem 4.1. *Let X and Y be two absolutely continuous nonnegative random variables with pdfs f and g, and cdfs F and G, respectively. Consider* *two bijective functions ϕ_1 and ϕ_2, which are strictly monotone and differentiable. Then for all $0 \le t_1 < t_2 < +\infty$, we have*

$$D^{w}_{KL}(\phi_1(X)||\phi_2(Y); t_1, t_2) = \begin{cases} D^{w,\phi_1}_{KL}(X||\phi_1^{-1}(\phi_2(Y)); \phi_1^{-1}(t_1), \phi_1^{-1}(t_2)), \\ \qquad \text{if } \phi_1 \text{ and } \phi_2 \text{ are strictly increasing,} \\ D^{w,\phi_1}_{KL}(X||\phi_1^{-1}(\phi_2(Y)); \phi_1^{-1}(t_2), \phi_1^{-1}(t_1)), \\ \qquad \text{if } \phi_1 \text{ and } \phi_2 \text{ are strictly decreasing,} \end{cases}$$

where

$$D^{w,\phi}_{KL}(X||Y; t_1, t_2) = \int_{t_1}^{t_2} \phi(x) \frac{f(x)}{\Delta F} \ln\Big(\frac{f(x)/\Delta F}{g(x)/\Delta G}\Big) dx. \tag{4.1}$$

Proof: Assume that $\phi_1(x)$ and $\phi_2(x)$ are strictly increasing functions. Under this condition, the pdfs and cdfs of $\phi_1(X)$ and $\phi_2(Y)$ can be obtained as

$$f_{\phi_1}(x) = \frac{f(\phi_1^{-1}(x))}{\phi_1'(\phi_1^{-1}(x))} \quad \text{and} \quad F_{\phi_1}(x) = F(\phi_1^{-1}(x)) \tag{4.2}$$

and

$$g_{\phi_2}(x) = \frac{f(\phi_2^{-1}(x))}{\phi_2'(\phi_2^{-1}(x))} \quad \text{and} \quad G_{\phi_2}(x) = G(\phi_2^{-1}(x)), \tag{4.3}$$

respectively. Moreover, the pdf and the cdf of $\phi_1^{-1}(\phi_2(X))$ are respectively given by

$$g_{\phi_1^{-1}(\phi_2)}(x) = \frac{g(\phi_2^{-1}(\phi_1(x)))\phi_1'(x)}{\phi_2'(\phi_2^{-1}(\phi_1(x)))} \quad \text{and} \quad G_{\phi_1^{-1}(\phi_2)}(x) = G(\phi_2^{-1}(\phi_1(x))). \tag{4.4}$$

Applying (4.2) and (4.3) in (1.4), we obtain

$$\begin{aligned} D^{w}_{KL}(\phi_1(X)||\phi_2(Y); t_1, t_2) \;=\; & \int_{t_1}^{t_2} x \frac{f(\phi_1^{-1}(x))/\phi_1'(\phi_1^{-1}(x))}{\Delta F^{\phi_1}} \\ & \times \ln\left(\frac{f(\phi_1^{-1}(x))/(\phi_1'(\phi_1^{-1}(x))\Delta F^{\phi_1})}{g(\phi_2^{-1}(x))/(\phi_2'(\phi_2^{-1}(x))\Delta G^{\phi_2})}\right) dx, \end{aligned} \tag{4.5}$$

where $\Delta F^{\phi_1} = F(\phi_1^{-1}(t_2)) - F(\phi_1^{-1}(t_1))$ and $\Delta G^{\phi_2} = G(\phi_2^{-1}(t_2)) - G(\phi_2^{-1}(t_1))$. Further, using the transformation $u = \phi_1^{-1}(x)$ in (4.5), we get

$$\begin{aligned} D^{w}_{KL}(\phi_1(X)||\phi_2(Y); t_1, t_2) \;=\; & \int_{\phi_1^{-1}(t_1)}^{\phi_1^{-1}(t_2)} \phi_1(u) \frac{f(u)}{\Delta F^{\phi_1}} \\ & \times \ln\left(\frac{f(u)/(\phi_1'(u)\Delta F^{\phi_1})}{g(\phi_2^{-1}(\phi_1(u)))/(\phi_2'(\phi_2^{-1}(\phi_1(u)))\Delta G^{\phi_2})}\right) du. \end{aligned} \tag{4.6}$$

Hence from (4.6), the first part of the theorem follows. The second part can be proved similarly and hence omitted. This completes the proof of the theorem. □

Remark 4.1. Note that when $t_1 \to 0$ (for fixed t_2) and $t_2 \to \infty$ (for fixed t_1), Theorem 4.1 reduces to Theorem 4.13 of Yasaei Sekeh *et al.* (2013).

Remark 4.2. Consider $\phi_1(x) = F(x)$ and $\phi_2(x) = G(x)$. Here, both $F(x)$ and $G(x)$ are strictly increasing in their supports. Also, consider $\phi_1(x) = \overline{F}(x)$ and $\phi_2(x) = \overline{G}(x)$, which are strictly decreasing in supports. Clearly, $\phi_1(x)$ and $\phi_2(x)$ satisfy the assumptions of Theorem 4.1. Thus as an application of the Theorem 4.1, we get

$$D^w_{KL}(F(X)||G(Y); t_1, t_2) = D^{w,F}_{KL}(X||F^{-1}(G(Y)); F^{-1}(t_1), F^{-1}(t_2)) \tag{4.7}$$

and

$$D^w_{KL}(\bar{F}(X)||\bar{G}(Y); t_1, t_2) = D^{w,\bar{F}}_{KL}(X||\bar{F}^{-1}(\bar{G}(Y)); \bar{F}^{-1}(t_2), \bar{F}^{-1}(t_1)). \tag{4.8}$$

The following proposition is due to Theorem 4.1. It provides the effects of the doubly truncated weighted KLD under affine transformations.

Proposition 4.1. *Let X and Y be two nonnegative absolutely continuous random variables with pdfs f and g, and cdfs F and G, respectively. Define $\phi_1(X) = a_1X + b_1$ and $\phi_2(Y) = a_2Y + b_2$, where $a_1, a_2 > 0$ and $b_1, b_2 \geq 0$ are constants. Then for $t_1 > b_2$ and $b_2 \geq b_1$,*

$$D^w_{KL}(\phi_1(X)||\phi_2(Y); t_1, t_2) = D^{w,\phi_1}_{KL}\Big(X||\frac{a_2}{a_1}Y + \frac{b_2 - b_1}{a_1}; \frac{t_1 - b_1}{a_1}, \frac{t_2 - b_1}{a_1}\Big). \tag{4.9}$$

Remark 4.3. Under the assumptions as described in Proposition 4.1, the right hand side expression given by (4.9) can be written further as

$$\begin{aligned} D^w_{KL}(\phi_1(X)||\phi_2(Y); t_1, t_2) \;=\; & a_1 D^w_{KL}\Big(X||\frac{a_2}{a_1}Y + \frac{b_2 - b_1}{a_1}; \frac{t_1 - b_1}{a_1}, \frac{t_2 - b_1}{a_1}\Big) \\ & + b_1 D_{KL}\Big(X||\frac{a_2}{a_1}Y + \frac{b_2 - b_1}{a_1}; \frac{t_1 - b_1}{a_1}, \frac{t_2 - b_1}{a_1}\Big). \end{aligned} \tag{4.10}$$

In particular, if we consider $\phi_1(x) = \phi_2(x) = \phi(x)$, then Theorem 4.1 reduces to the following result. We omit the proof as it follows from that of the Theorem 4.1.

Theorem 4.2. *Let X and Y be two absolutely continuous nonnegative random variables as described in Theorem 4.1. Assume that $\phi(x)$ is strictly monotone and differentiable function. Then for all $0 \leq t_1 < t_2 < +\infty$, we have*

$$D^w_{KL}(\phi(X)||\phi(Y); t_1, t_2) = \begin{cases} D^{w,\phi}_{KL}(X||Y; \phi^{-1}(t_1), \phi^{-1}(t_2)), \\ \qquad\qquad \text{if } \phi(x) \text{ is strictly increasing,} \\ D^{w,\phi}_{KL}(X||Y; \phi^{-1}(t_2), \phi^{-1}(t_1)), \\ \qquad\qquad \text{if } \phi(x) \text{ is strictly decreasing,} \end{cases}$$

where $D^{w,\phi}_{KL}(X||Y; t_1, t_2)$ is given by (4.1).

Proposition 4.2. *Let X and Y be two nonnegative absolutely continuous random variables with pdfs f and g, and cdfs F and G, respectively. Define $\phi_1(x) = \phi_2(x) = \phi(x) = ax + b$, where $a > 0$ and $b \geq 0$ are constants. Then for $t_1 > b$*

$$D^w_{KL}(\phi(X)||\phi(Y); t_1, t_2) = D^{w,\phi}_{KL}\Big(X||Y; \frac{t_1 - b}{a}, \frac{t_2 - b}{a}\Big). \tag{4.11}$$

5. INEQUALITIES AND BOUNDS

In this section, we obtain various inequalities and bounds for doubly truncated weighted KLD given by (1.4) in terms of other measures, which may be useful in mathematical statistics, ergodic theory and other scientific fields. Most of these depend on the measures (2.1)–(2.3) and (2.5)–(2.7). Several authors studied these measures and obtained various results. For some results on these measures, we refer to Navarro and Ruiz (1996), Misagh and Yari (2011), Sankaran and Sunoj (2004) and Kundu (2017).

Proposition 5.1. *Let X and Y be two nonnegative absolutely continuous random variables with pdfs f and g, and cdfs F and G, respectively. Suppose that $D^w_{KL}(X||Y; t_1, t_2)$ is increasing (decreasing) in t_1 (for fixed t_2) and t_2 (for fixed t_1). Then*

$$D^w_{KL}(X||Y; t_1, t_2) \geq (\leq) \Big(\frac{h^Y_1(t_1, t_2)}{h^X_1(t_1, t_2)} - 1\Big)\mu_X(t_1, t_2) - t_1 \ln\Big(\frac{h^Y_1(t_1, t_2)}{h^X_1(t_1, t_2)}\Big)$$

and

$$D^w_{KL}(X||Y; t_1, t_2) \leq (\geq) \Big(\frac{h^Y_2(t_1, t_2)}{h^X_2(t_1, t_2)} - 1\Big)\mu_X(t_1, t_2) - t_2 \ln\Big(\frac{h^Y_2(t_1, t_2)}{h^X_2(t_1, t_2)}\Big).$$

Proof: Under the given condition, the required inequalities follow from (3.1) and (3.2), and hence, omitted. □

The following proposition provides bounds of (1.4) in terms of GCM, GFR functions and doubly truncated weighted inaccuracy measure.

Proposition 5.2. *Let X and Y be two random variables as described in Proposition 5.1. If $f(x)$ is increasing (decreasing) in $x > 0$, then*

$$D^w_{KL}(X||Y; t_1, t_2) \geq (\leq)\ \mu_X(t_1, t_2) \ln h^X_1(t_1, t_2) + I^w(X||Y; t_1, t_2)$$

and

$$D^w_{KL}(X||Y; t_1, t_2) \leq (\geq)\ \mu_X(t_1, t_2) \ln h^X_2(t_1, t_2) + I^w(X||Y; t_1, t_2).$$

Proof: Let $f(x)$ be increasing (decreasing) in $x > 0$. Then for $t_1 < x < t_2$, we have

$$\frac{f(t_1)}{\Delta F} \le (\ge) \frac{f(x)}{\Delta F} \le (\ge) \frac{f(t_2)}{\Delta F}. \tag{5.1}$$

Moreover, $g(x)/\Delta G$ is positive. Thus from (5.1) we get

$$\ln\Big(\frac{f(t_1)}{\Delta F}\Big/\frac{g(x)}{\Delta G}\Big) \le (\ge) \ln\Big(\frac{f(x)}{\Delta F}\Big/\frac{g(x)}{\Delta G}\Big) \le (\ge) \ln\Big(\frac{f(t_2)}{\Delta F}\Big/\frac{g(x)}{\Delta G}\Big). \tag{5.2}$$

Multiplying (5.2) by $x\frac{f(x)}{\Delta F}$ and then integrating from t_1 to t_2 with respect to x, the required inequalities follow. $\square$

Proposition 5.3. *Let X and Y be two random variables as described in Proposition* 5.1. *If $g(x)$ is increasing (decreasing) in $x > 0$, then*

$$D^w_{KL}(X||Y; t_1, t_2) \ \ge (\le) \ -\mu_X(t_1, t_2) \ln h^Y_2(t_1, t_2) - S^w_X(t_1, t_2)$$

and

$$D^w_{KL}(X||Y; t_1, t_2) \ \le (\ge) \ -\mu_X(t_1, t_2) \ln h^Y_1(t_1, t_2) - S^w_X(t_1, t_2).$$

Proof: Proof follows analogous to that of the Proposition 5.2. Hence omitted. $\square$

Below, in Proposition 5.4, we give new inequalities for $D^w_{KL}(X||Y; t_1, t_2)$ involving a pair of likelihood ratio ordered random variables.

Proposition 5.4. *Let X and Y be two random variables as described in Proposition* 5.1. *If $X \ge^{lr} Y$, then*

$$\mu_X(t_1, t_2) \ln\Big(\frac{h^X_1(t_1, t_2)}{h^Y_1(t_1, t_2)}\Big) \le D^w_{KL}(X||Y; t_1, t_2) \le \mu_X(t_1, t_2) \ln\Big(\frac{h^X_2(t_1, t_2)}{h^Y_2(t_1, t_2)}\Big).$$

Proof: Under the given condition, $f(x)/g(x)$ is increasing in $x > 0$. Then for $t_1 < x < t_2$, we have $\frac{f(t_1)}{g(t_1)} \le \frac{f(x)}{g(x)} \le \frac{f(t_2)}{g(t_2)}$. Thus from (1.4) we obtain

$$\begin{aligned} D^w_{KL}(X||Y; t_1, t_2) \ &\le \ \int_{t_1}^{t_2} x\frac{f(x)}{\Delta F} \ln\Big(\frac{f(t_2)/\Delta F}{g(t_2)/\Delta G}\Big) dx \\ &= \ \mu_X(t_1, t_2) \ln\Big(\frac{h^X_2(t_1, t_2)}{h^Y_2(t_1, t_2)}\Big) \end{aligned} \tag{5.3}$$

and

$$\begin{aligned} D^w_{KL}(X||Y; t_1, t_2) \ &\ge \ \int_{t_1}^{t_2} x\frac{f(x)}{\Delta F} \ln\Big(\frac{f(t_1)/\Delta F}{g(t_1)/\Delta G}\Big) dx \\ &= \ \mu_X(t_1, t_2) \ln\Big(\frac{h^X_1(t_1, t_2)}{h^Y_1(t_1, t_2)}\Big). \end{aligned} \tag{5.4}$$

Combining (5.3) and (5.4), we obtain the required inequalities. $\square$

Proposition 5.5. *Let X and Y be two random variables as described in Proposition* 5.1. *Then*

$$D^w_{KL}(X||Y;t_1,t_2) \geq \mu_X(t_1,t_2)\ln\Big(\frac{\mu_X(t_1,t_2)}{\mu_Y(t_1,t_2)}\Big).$$

Proof: The result follows from the log-sum inequality and hence omitted. □

Proposition 5.6. *Let X and Y be two random variables as described in Proposition* 5.1. *Then*

$$D^w_{KL}(X||Y;t_1,t_2) \leq E\left(X\frac{f(X)/\Delta F}{g(X)/\Delta G}\Big|t_1<X<t_2\right) - \mu_X(t_1,t_2). \tag{5.5}$$

Proof: The proof follows from the inequality $\ln x \leq x-1$, for all $x>0$. Hence it is omitted. □

Hereafter, we consider three nonnegative random variables X_1, X_2 and X_3 and obtain bounds of $D^w_{KL}(X_1||X_3;t_1,t_2)$, $D^w_{KL}(X_1||X_2;t_1,t_2)$ and $D^w_{KL}(X_2||X_3;t_1,t_2)$.

Proposition 5.7. *Let X_1, X_2 and X_3 be three nonnegative absolutely continuous random variables with pdf's $f_1(x), f_2(x)$ and $f_3(x)$, respectively. The corresponding cdf's are $F_1(x)$, $F_2(x)$ and $F_3(x)$. If $X_1 \geq^{lr} X_2$, then*

$$D^w_{KL}(X_1||X_3;t_1,t_2) \;\geq\; \mu_{X_1}(t_1,t_2)\ln\Big(\frac{h_1^{X_1}(t_1,t_2)}{h_1^{X_2}(t_1,t_2)}\Big) + ID^w(t_1,t_2)$$

and

$$D^w_{KL}(X_1||X_3;t_1,t_2) \;\leq\; \mu_{X_1}(t_1,t_2)\ln\Big(\frac{h_2^{X_1}(t_1,t_2)}{h_2^{X_2}(t_1,t_2)}\Big) + ID^w(t_1,t_2),$$

where $ID^w(t_1,t_2) = I^w(X_1||X_3;t_1,t_2) - I^w(X_1||X_2;t_1,t_2)$.

Proof: Using $X_1 \geq^{lr} X_2$ and $t_1 < x$, we obtain $f_1(x) \geq f_2(x)f_1(t_1)/f_2(t_1)$. Thus, from (1.4)

$$\begin{aligned} D^w_{KL}(X_1||X_3;t_1,t_2) &= \int_{t_1}^{t_2} x\frac{f_1(x)}{\Delta F_1}\ln\Big(\frac{f_1(x)/\Delta F_1}{f_3(x)/\Delta F_3}\Big)dx \\ &\geq \int_{t_1}^{t_2} x\frac{f_1(x)}{\Delta F_1}\ln\Big(\frac{(f_2(x)f_1(t_1))/(\Delta F_2\Delta F_1)}{(f_3(x)f_2(t_1))/(\Delta F_3\Delta F_2)}\Big)dx \\ &= \mu_{X_1}(t_1,t_2)\ln\Big(\frac{h_1^{X_1}(t_1,t_2)}{h_1^{X_2}(t_1,t_2)}\Big) + ID^w(t_1,t_2). \end{aligned} \tag{5.6}$$

The upper bound can be obtained similarly. This completes the proof. □

Proposition 5.8. *Let X_1, X_2 and X_3 be three random variables as described in Proposition 5.7. If $X_2 \geq^{lr} X_3$, then*

$$D^w_{KL}(X_1||X_2;t_1,t_2) \;\geq\; \mu_{X_1}(t_1,t_2)\ln\Big(\frac{h_2^{X_3}(t_1,t_2)}{h_2^{X_2}(t_1,t_2)}\Big) + IS_1^w(t_1,t_2)$$

and

$$D^w_{KL}(X_1||X_2;t_1,t_2) \;\leq\; \mu_{X_1}(t_1,t_2)\ln\Big(\frac{h_1^{X_3}(t_1,t_2)}{h_1^{X_2}(t_1,t_2)}\Big) + IS_1^w(t_1,t_2),$$

where $IS_1^w(t_1,t_2) = I^w(X_1||X_3;t_1,t_2) - S^w(X_1;t_1,t_2)$.

Proof: Under $X_2 \geq^{lr} X_3$, for $x < t_2$, we have $f_2(x) \leq f_3(x)f_2(t_2)/f_3(t_2)$. Thus applying this inequality in (1.4) and after some simplifications, lower bound can be obtained. The upper bound can be obtained similarly. This completes the proof. □

Proposition 5.9. *Let X_1, X_2 and X_3 be three random variables as described in Proposition 5.7. If $X_1 \geq^{lr} X_3$, then*

$$D^w_{KL}(X_2||X_3;t_1,t_2) \;\geq\; \mu_{X_2}(t_1,t_2)\ln\Big(\frac{h_1^{X_1}(t_1,t_2)}{h_1^{X_3}(t_1,t_2)}\Big) + IS_2^w(t_1,t_2)$$

and

$$D^w_{KL}(X_2||X_3;t_1,t_2) \;\leq\; \mu_{X_2}(t_1,t_2)\ln\Big(\frac{h_2^{X_1}(t_1,t_2)}{h_2^{X_3}(t_1,t_2)}\Big) + IS_2^w(t_1,t_2),$$

where $IS_2^w(t_1,t_2) = I^w(X_2||X_1;t_1,t_2) - S^w(X_2;t_1,t_2)$.

Proof: It is given that $X_1 \geq^{lr} X_3$. Therefore, for $x > t_1$, we have $f_3(x) \leq f_1(x)f_3(t_1)/f_1(t_1)$. Applying this in (1.4), we get the lower bound of $D^w_{KL}(X_2||X_3;t_1,t_2)$. The upper bound can be obtained similarly. This completes the proof. □

6. NUMERICAL EXAMPLES

In this section, we consider examples for the verification of few of the results obtained in Section 5. To verify the Proposition 5.2, we consider the following example and present numerical values of the lower and upper bounds of the doubly truncated weighted KLD.

Example 6.1. Suppose that X follows power distribution and Y follows U-quadratic distribution in the interval $(0,1)$ with pdfs $f(x) = cx^{c-1}$, $c > 0$, $0 < x < 1$ and $g(x) = 12(x - \frac{1}{2})^2$, $0 < x < 1$, respectively. Here, $f(x)$ is decreasing in x for $c < 1$ and increasing in x for $c > 1$. In Table 1 and Table 2, we present numerical values of the lower bounds (LB) and upper bounds (UB) of the doubly truncated weighted KLD for different values of t_1 and t_2 for $c = 0.5$ and 1.5, respectively.

Table 1: Bounds of $D^w_{KL}(X||Y;t_1,t_2)$ for $c=0.5$.

(t_1,t_2)	LB	$D^w_{KL}(X\|Y;t_1,t_2)$	UB	(t_1,t_2)	LB	$D^w_{KL}(X\|Y;t_1,t_2)$	UB
(0.1,0.4)	0.022421	0.077094	0.184155	(0.4,0.5)	0.405266	0.428923	0.455370
(0.1,0.5)	0.252138	0.321949	0.473062	(0.4,0.7)	0.467240	0.532709	0.619190
(0.1,0.9)	0.233992	0.360127	0.710058	(0.4,0.9)	0.546199	0.649186	0.802994
(0.2,0.5)	0.309069	0.369426	0.464263	(0.5,0.6)	0.445497	0.469392	0.495566
(0.2,0.7)	0.356237	0.449858	0.622275	(0.5,0.7)	0.439779	0.485807	0.540251
(0.2,0.8)	0.296777	0.406123	0.620246	(0.5,0.9)	0.426941	0.513659	0.629808

Table 2: Bounds of $D^w_{KL}(X||Y;t_1,t_2)$ for $c=1.5$.

(t_1,t_2)	LB	$D^w_{KL}(X\|Y;t_1,t_2)$	UB	(t_1,t_2)	LB	$D^w_{KL}(X\|Y;t_1,t_2)$	UB
(0.1,0.4)	0.048524	0.185502	0.232704	(0.4,0.5)	0.455614	0.483096	0.505924
(0.1,0.5)	0.434464	0.634879	0.694851	(0.4,0.7)	0.362879	0.458389	0.518698
(0.1,0.9)	0.014652	0.519028	0.628185	(0.4,0.9)	0.280465	0.458343	0.550633
(0.2,0.5)	0.458314	0.570112	0.623685	(0.5,0.6)	0.395761	0.422765	0.445968
(0.2,0.7)	0.404517	0.620202	0.701468	(0.5,0.7)	0.340430	0.398162	0.441841
(0.2,0.8)	0.213175	0.487130	0.581533	(0.5,0.9)	0.227166	0.356106	0.435718

To illustrate Proposition 5.3, we consider the following example.

Example 6.2. Let X and Y be two random variables with pdfs $f(x)=1,\ 0<x<1$ and $g(x)=b(1-x)^{b-1},\ 0<x<1, b>0$, respectively. Note that $g(x)$ is increasing in x for $b<1$ and decreasing in x for $b>1$. In Table 3 and 4, we present the numerical values of the bounds of the doubly truncated weighted KLD for $b=0.2$ and $b=1.2$, respectively.

Table 3: Bounds of $D^w_{KL}(X||Y;t_1,t_2)$ for $b=0.2$.

(t_1,t_2)	LB	$D^w_{KL}(X\|Y;t_1,t_2)$	UB	(t_1,t_2)	LB	$D^w_{KL}(X\|Y;t_1,t_2)$	UB
(0.1,0.4)	−0.042188	−0.006974	0.038905	(0.4,0.5)	−0.033416	−0.000816	0.032219
(0.1,0.5)	−0.074668	−0.012759	0.066401	(0.4,0.7)	−0.163019	−0.006694	0.141966
(0.1,0.9)	−0.532245	−0.044619	0.346645	(0.4,0.9)	−0.547117	−0.003054	0.384598
(0.2,0.5)	−0.068887	−0.007283	0.062713	(0.5,0.6)	−0.050187	−0.000755	0.047997
(0.2,0.7)	−0.193724	−0.020447	0.159375	(0.5,0.7)	−0.128847	−0.002608	0.116350
(0.2,0.8)	−0.315074	−0.027845	0.239444	(0.5,0.9)	−0.521607	0.005471	0.379679

Table 4: Bounds of $D^w_{KL}(X||Y;t_1,t_2)$ for $b=1.2$.

(t_1,t_2)	LB	$D^w_{KL}(X\|Y;t_1,t_2)$	UB	(t_1,t_2)	LB	$D^w_{KL}(X\|Y;t_1,t_2)$	UB
(0.1,0.4)	−0.009386	0.002084	0.010888	(0.4,0.5)	−0.007930	0.000329	0.008478
(0.1,0.5)	−0.015747	0.004043	0.019521	(0.4,0.7)	−0.033325	0.003839	0.042921
(0.1,0.9)	−0.069415	0.028401	0.150307	(0.4,0.9)	−0.080490	0.016423	0.152438
(0.2,0.5)	−0.015039	0.002460	0.017861	(0.5,0.6)	−0.011771	0.000417	0.012774
(0.2,0.7)	−0.036351	0.008604	0.051923	(0.5,0.7)	−0.027794	0.001945	0.033504
(0.2,0.8)	−0.052349	0.014472	0.086279	(0.5,0.9)	−0.081046	0.012505	0.144275

The following example shows a case in which Proposition 5.4 is fulfilled.

Example 6.3. Consider two nonnegative random variables X and Y with pdfs $f(x) = \frac{2}{3}(1+x),\ 0 < x < 1$ and $g(x) = \frac{2}{3}(2-x),\ 0 < x < 1$, respectively. By straightforward calculations, it is not hard to verify that $X \geq^{lr} Y$. In Table 5, we present the numerical values of the lower and upper bounds of the doubly truncated weighted KLD between X and Y.

Table 5: Bounds of $D^w_{KL}(X||Y; t_1, t_2)$.

(t_1, t_2)	LB	$D^w_{KL}(X\|Y; t_1, t_2)$	UB	(t_1, t_2)	LB	$D^w_{KL}(X\|Y; t_1, t_2)$	UB
(0.1,0.3)	−0.028607	0.005274	0.278213	(0.5,0.7)	−0.080397	0.006264	0.081120
(0.1,0.6)	−0.126393	0.034815	0.122128	(0.5,0.8)	−0.131348	0.014560	0.134047
(0.1,0.8)	−0.229447	0.071593	0.225767	(0.5,0.9)	−0.189889	0.026835	0.196978
(0.3,0.4)	−0.023699	0.001387	0.023540	(0.6,0.7)	−0.043674	0.001614	0.043969
(0.3,0.7)	−0.136517	0.023788	0.136517	(0.6,0.8)	−0.094577	0.006683	0.096309
(0.3,0.9)	−0.248611	0.056871	0.255551	(0.6,0.9)	−0.153075	0.015624	0.158454

In this part of the paper, we provide an example in support of the Proposition 5.7.

Example 6.4. Let X and Y be two nonnegative random variables as described in Example 6.3. Consider another random variable Z with pdf $f_3(x) = \frac{1}{2\sqrt{1-x}},\ 0 < x < 1$. Here, $X \geq^{lr} Y$. The lower and upper bounds of $D^w_{KL}(X||Y; t_1, t_2)$ obtained in the Proposition 5.7 are presented in Table 6.

Table 6: Bounds of $D^w_{KL}(X||Y; t_1, t_2)$.

(t_1, t_2)	LB	$D^w_{KL}(X\|Y; t_1, t_2)$	UB	(t_1, t_2)	LB	$D^w_{KL}(X\|Y; t_1, t_2)$	UB
(0.1,0.2)	−0.011780	0.000238	0.009483	(0.4,0.6)	−0.074033	−0.001039	0.000091
(0.1,0.5)	−0.107631	0.000543	0.061938	(0.4,0.7)	−0.127975	−0.003167	0.094956
(0.1,0.7)	−0.230205	−0.004876	0.113178	(0.4,0.9)	−0.274740	−0.012652	0.175896
(0.2,0.4)	−0.046788	0.000172	0.035488	(0.6,0.7)	−0.045791	−0.000503	0.041852
(0.2,0.5)	−0.085598	−0.000282	0.058566	(0.6,0.8)	−0.103484	−0.002223	0.087403
(0.2,0.8)	−0.275617	−0.011008	0.146067	(0.6,0.9)	−0.172764	−0.004064	0.138765

Real Data: We consider two real data sets, which represent the failure times of the air conditioning system of two different air planes (see Bain and Engelhardt, 1991, p. 101). The data sets are given below:

Data Set I (*Plane 7912*): 1, 3, 5, 7, 11, 11, 11, 12, 14, 14, 14, 16, 16, 20, 21, 23, 42, 47, 52, 62, 71, 71, 87, 90, 95, 120, 120, 225, 246, 261.

Data Set II (*Plane 7911*): 33, 47, 55, 56, 104, 176, 182, 220, 239, 246, 320.

The above data sets, Data Set I and Data Set II can be fitted as exponential distributions with parameters (hazard rates) λ_1 and λ_2, respectively. We assume that due to some reasons, the data in the interval $[50, 200]$ are observed. Based on this assumption, the unknown parameters can be estimated. For this purpose, we use the method of maximum likelihood. Here, the estimated values of the parameters are $\hat{\lambda}_1 = 0.026029$ and $\hat{\lambda}_2 = 0.005611$. From (1.4), we get $\hat{D}^w_{KL}(X||Y; 50, 200) = 2.5069$.

7. CONCLUDING REMARKS

In this paper, we consider a generalized discrimination measure, which is known as the doubly truncated weighted KLD. We obtain few characterization results based on the proposed measure. These results may be useful in studying various characteristics of a system when its lifetimes fall in an interval. Further, the effect of the affine transformations on the proposed discrimination measure is studied. Several inequalities and bounds are obtained. Finally, few applications with bounds in support of the results are presented.

ACKNOWLEDGMENTS

The authors wish to express their sincere thanks to an anonymous reviewer for some useful comments which have improved the initial version of this manuscript.

REFERENCES

[1] Bain, L.J. and Engelhardt, M. (1991). *Statistical Analysis of Reliability and Life-Testing Models*, 2nd. Edition, Marcel and Dekker, New York.

[2] Belis, M. and Guiasu, S. (1968). A quantitative-qualitative measure of information in cybernetic systems, *IEEE Transactions on Information Theory*, **4**, 593–594.

[3] Cox, D.R. (1972). Regression models and life tables (with Discussion), *Journal of the Royal Statistical Society, Series B*, **34**, 187–220.

[4] Di Crescenzo, A. and Longobardi, M. (2006). On weighted residual and past entropies, *Scientiae Mathematicae Japonicae*, **64**, 255–266.

[5] Ebrahimi, N. and Kirmani, S.N.U.A. (1996). A characterization of the proportional hazards model through a measure of discrimination between two residual life distributions, *Biometrika*, **83**, 233–235.

[6] Gupta, R.C. and Gupta, R.D. (2007). Proportional reversed hazard rate model and its applications, *Journal of Statistical Planning and Inference*, **137**, 3525–3536.

[7] Kasza, J. and Solomon, P. (2015). Comparing score-based methods for estimating Bayesian networks using Kullback–Leibler divergence, *Communications in Statistics – Theory and Methods*, **44**, 135–152.

[8] Kullback, S. and Leibler, R.A. (1951). On information and sufficiency, *Annals of Mathematical Statistics*, **22**, 79–86.

[9] Kundu, C. (2017). On weighted measure of inaccuracy for doubly truncated random variables, *Communications in Statistics – Theory and Methods*, **46**, 3135–3147.

[10] Mirali, M.; Baratpour, S. and Fakoor, V. (2017). On weighted cumulative residual entropy, *Communications in Statistics – Theory and Methods*, **46**, 2857–2869.

[11] Misagh, F. and Yari, G. (2011). On weighted interval entropy, *Statistics and Probability Letters*, **81**, 188–194.

[12] Misagh, F. and Yari, G. (2012). Interval entropy and informative distance, *Entropy*, **14**, 480–490.

[13] Nair, N.U. and Gupta, R.P. (2007). Characterization of proportional hazard models by properties of information measures, *International Journal of Statistical Sciences*, **6**, 223–231.

[14] Nair, K.R.M. and Rajesh, G. (2000). Geometric vitality function and its applications to reliability, IAPQR Transactions, **25**, 1–8.

[15] Navarro, J. and Ruiz, J.M. (1996). Failure rate functions for doubly truncated random variables, *IEEE Transactions on Reliability*, **45**, 685–690.

[16] Nourbakhsh, M.; Yari, G. and Mehrali, Y. (2016). Weighted entropies and their estimations, *Communications in Statistics – Simulation and Computation*, http://www.tandfonline.com/doi/abs/10.1080/03610918.2016.1140776.

[17] Rajesh, G.; Abdul-Sathar, E.I. and Rohini, N.S. (2017). On dynamic weighted survival entropy of order α, *Communications in Statistics – Theory and Methods*, **46**, 2139–2150.

[18] Ruiz, J.M. and Navarro, J. (1996). Characterizations based on conditional expectations of the doubles truncated distribution, *Annals of the Institute of Statistical Mathematics*, **48**, 563–572.

[19] Sankaran, P.G. and Gleeja, V.L. (2008). Proportional reversed hazard and frailty models, *Metrika*, **68**, 333–342.

[20] Sankaran, P.G. and Sunoj, S.M. (2004). Identification of models using failure rate and mean residual life of doubly truncated random variables, *Statistical Papers*, **45**, 97–109.

[21] Sankaran, P.G.; Sunoj, S.M. and Nair, N.U. (2016). Kullback–Leibler divergence: A quantile approach, *Statistics and Probability Letters*, **111**, 72–79.

[22] Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*, Springer Verlag, New York.

[23] Suhov, Y. and Yasaei Sekeh, S. (2015). Weighted cumulative entropies: an extension of CRE and CE, http://arxiv.org/abs/1507.07051v1.

[24] Sunoj, S.M.; Sankaran, P.G. and Maya, S.S. (2009), Characterizations of life distributions using conditional expectations of doubly (interval) truncated random variables, *Communications in Statistics – Theory and Methods*, **38**, 1441–1452.

[25] Yasaei Sekeh, S.; Borzadaran, G.R.M. and Roknabadi, A.H.R. (2013). On Kullback–Leibler dynamic information, http://dx.doi.org/10.2139/ssrn.2344078.