Applications of Mathematics

Chanchal Kundu

Characterizations based on length-biased weighted measure of inaccuracy for truncated random variables

Applications of Mathematics, Vol. 59 (2014), No. 6, 697–714

Persistent URL: http://dml.cz/dmlcz/143995

CHARACTERIZATIONS BASED ON LENGTH-BIASED WEIGHTED MEASURE OF INACCURACY FOR TRUNCATED RANDOM VARIABLES

Chanchal Kundu, Rae Bareli

(Received February 17, 2013)

Abstract. In survival studies and life testing, the data are generally truncated. Recently, authors have studied a weighted version of Kerridge inaccuracy measure for truncated distributions. In the present paper we consider weighted residual and weighted past inaccuracy measure and study various aspects of their bounds. Characterizations of several important continuous distributions are provided based on weighted residual (past) inaccuracy measure.

Keywords: characterization; entropy; weighted residual (past) inaccuracy; proportional (reversed) hazard model

MSC 2010: 60E15, 62N05, 20B10

1. Introduction

The concept of weighted distribution introduced by Rao [25] has many applications in different areas of statistics such as reliability, survival analysis, forestry, ecology, survey sampling and several other fields. Weighted distributions arise when the observations generated from a stochastic process are recorded with some weight function. Many well-known probability models, such as probability density functions of order statistics, record values, or the proportional (reversed) hazard model can be considered weighted distributions. Jain et al. [15], Gupta and Kirmani [14] and Nanda and Jain [23] used the weighted distribution in many practical problems to model unequal sampling probabilities. Let $w(x)$ be a nonnegative function of x such that $E(w(X))$ is finite. Then the corresponding probability density function of the

The work is supported by Department of Science and Technology, Government of India (No. SR/FTP/MS-016/2012).

weighted random variable X^w is given by

$$f^w(x) = \frac{w(x)f(x)}{E(w(X))}.$$

The basic problem in using a weighted distribution as a tool for modeling is the identification of the appropriate weight function that fits the data. When $w(x) = x$, i.e., the weight function depends on the length of the unit of interest, X^w is said to be a length-biased or a size-biased random variable with probability density function

$$f^*(x) = \frac{xf(x)}{E(X)}, \quad x > 0 \text{ and } E(X) < \infty.$$

Then the length-biased distribution function and the survival function are defined as $F^*(t) = E(X)^{-1} \int_0^t xf(x)\,\mathrm{d}x$ and $\overline{F}^*(t) = E(X)^{-1} \int_t^\infty xf(x)\,\mathrm{d}x$, respectively. These functions characterize weighted distributions that arise in clinical trials, reliability, queuing models, survival analysis and population studies where a proper sampling frame is absent. In such situations, items are sampled at rate proportional to their length so that larger values of the quantity being measured are sampled with higher probabilities. See, for details, Cox [4] and Patil and Ord [24].

Recently, the application of Kerridge's [16] inaccuracy measure as a generalization of Shannon's [27] entropy has attracted increasing attention. It has been extensively used as a useful tool for measurement of error in experimental results. Let X and Y be two absolutely continuous nonnegative random variables with distribution functions $F(x)$, $G(x)$ and probability density functions $f(x)$, $g(x)$, respectively. If $F(x)$ is the actual distribution corresponding to the observations and $G(x)$ is the distribution function assigned by the experimenter, then the inaccuracy measure is defined as

$$H(f,g) = -\int_0^\infty f(x)\ln g(x)\,\mathrm{d}x. \tag{1.1}$$

It has applications in statistical inference, estimation and coding theory. See, for more details, Smitha [28]. The dissimilarity between $f(x)$ and $g(x)$, which may represent the income distributions of two groups or regions or two different economic models, is measured by distance or divergence. One important divergence measure due to Kullback-Leibler [17] is given by

$$D(f\|g) = \int_0^\infty f(x)\ln\frac{f(x)}{g(x)}\,\mathrm{d}x,$$

which represents the expected uncertainty contained in $g(x)$ with respect to $f(x)$. With this definition, Kullback-Leibler divergence measure can be written as

$$D(f\|g) = H(f,g) - H(f),$$

where $H(f)$ is the well-known Shannon's entropy given by

$$H(f) = -\int_0^\infty f(x) \ln f(x)\,\mathrm{d}x, \tag{1.2}$$

which can also be obtained from (1.1) for $g(x) = f(x)$. It measures the expected uncertainty contained in $f(\cdot)$ about the predictability of an outcome of X. If the ratio $g(x)/f(x)$ is far from unity, i.e., difference in the distribution is large, then both Kullback-Leibler divergence and Kerridge inaccuracy measure will increase. $D(f||g)$ vanishes for $g(x) = f(x)$, which in turn gives $H(f,g) = H(f)$, i.e., there is no inaccuracy and we are left only with uncertainty measured by Shannon.

However, in some practical situations, such as reliability or mathematical neurobiology, a shift-dependent information measure is desirable. An important feature of the human visual system is that it can recognize objects in a scale and translation invariant manner. Achieving this desirable behavior using biologically realistic networks is a challenge (cf. Wallis [30]). Indeed, knowing that a device fails to operate, or a neuron fails to release spikes in a given time-interval, yields a relevantly different information from the case when such an event occurs in a different equally wide interval. In some cases we are thus led to resort to a shift-dependent information measure that, for instance, assigns different measures to such distributions.

In agreement with Di Crescenzo and Longobardi [7], the weighted measure of inaccuracy is given by

$$H^w(f,g) = -\int_0^\infty x f(x) \ln g(x)\,\mathrm{d}x, \tag{1.3}$$

which yields a 'length-biased' shift-dependent inaccuracy measure assigning greater importance to larger values of X. The following example illustrates the role of weighted inaccuracy measure in the case of random lifetimes.

Example 1.1. Let X_1 and X_2 denote random lifetimes of two components with probability density functions $f_1(x) = 2x$, $x \in (0,1)$, and $f_2(x) = 2(1-x)$, $x \in (0,1)$, respectively. By simple calculations, we have $H(f_1,f_2) = H(f_2,f_1) = 3/2 - \ln 2$. But,

$$H^w(f_1,f_2) = \frac{11}{9} - \frac{2}{3}\ln 2 \quad \text{and} \quad H^w(f_2,f_1) = \frac{5}{18} - \frac{1}{3}\ln 2.$$

That is, the inaccuracy measure of the observer for the observations X_1 (or X_2) taking X_2 (or X_1) as corresponding assigned outcomes by the experimenter are identical, while $H^w(f_1,f_2) > H^w(f_2,f_1)$, i.e., weighted inaccuracy of the observer for X_1, X_2 is higher than that for X_2, X_1.

Weighted measures of inaccuracy for residual and past lifetime distributions have also been proposed in the literature. Motivated by the above example we consider weighted inaccuracy for truncated random variables. The rest of the paper is arranged as follows. In Section 2 we provide characterizations of several useful continuous distributions based on weighted residual inaccuracy measure. We also study the bounds of the weighted residual inaccuracy measure and its monotonic transformations. In Section 3 we study the same for weighted past inaccuracy measure. Some characterization results are also provided based on this measure. In conclusion, some discussion concerning empirical inaccuracy measure is made in Section 4.

Throughout this paper, the words *increasing* and *decreasing* are not used in strict sense unless otherwise specified.

2. Characterizations based on weighted residual inaccuracy measure

In the literature, the problem of characterizing probability distributions has been investigated by many researchers, see, for instance, Galambos and Kotz [10] and Azlarov and Volodin [2]. The standard practice in modeling statistical data is either to derive the appropriate model based on the physical properties of the system or to choose a flexible family of distributions and then find a member of the family that is appropriate to the data. In both situations it is helpful if we find characterization theorems that explain the distribution. In fact, characterization approach is very appealing to both theoreticians and applied researchers. In this section we provide characterizations of several useful continuous distributions based on weighted residual inaccuracy measure.

First we review some properties of the weighted residual inaccuracy measure. Kumar et al. [19] introduced the notion of weighted residual inaccuracy at time t of a random variable X as the differential weighted inaccuracy of the left truncated random variable $[X-t \mid X>t]$ given by

$$H^w(f,g;t) = -\int_t^\infty x\frac{f(x)}{\overline{F}(t)}\ln\Big(\frac{g(x)}{\overline{G}(t)}\Big)\,\mathrm{d}x \tag{2.1}$$

and studied various aspects of this measure in analogy with weighted residual entropy. The following theorem, due to Kumar et al. [19], provides a lower bound for the weighted residual inaccuracy measure in terms of $h_G(t)=g(t)/\overline{G}(t)$, the hazard rate of Y, and the conditional mean of X given by

$$m_X(t) = E[X \mid X>t] = \frac{1}{\overline{F}(t)}\int_t^\infty xf(x)\,\mathrm{d}x.$$

For more applications of $m_X(t)$ in insurance and economics, one may refer to Furman and Zitikis [9]. For completeness we give a brief outline of the proof.

Theorem 2.1. *If the hazard rate function $h_G(t)$ is decreasing in t, then*

$$H^w(f,g;t) \geqslant -m_X(t)\ln h_G(t). \tag{2.2}$$

Proof. Note that (2.1) can alternatively be written as

$$H^w(f,g;t) = -\frac{1}{\overline{F}(t)}\int_t^\infty xf(x)\ln h_G(x)\,\mathrm{d}x - \frac{1}{\overline{F}(t)}\int_t^\infty xf(x)\ln\Bigl(\frac{\overline{G}(x)}{\overline{G}(t)}\Bigr)\,\mathrm{d}x. \tag{2.3}$$

Using the fact that $\ln(\overline{G}(x)/\overline{G}(t)) \leqslant 0$ for $x \geqslant t$, and by the assumption $\ln h_G(x) \leqslant \ln h_G(t)$, we have

$$\begin{aligned} H^w(f,g;t) &\geqslant -\frac{1}{\overline{F}(t)}\int_t^\infty xf(x)\ln h_G(x)\,\mathrm{d}x \\ &\geqslant -\frac{\ln h_G(t)}{\overline{F}(t)}\int_t^\infty xf(x)\,\mathrm{d}x. \end{aligned}$$

Hence, the result follows. $\square$

Remark 2.1. In order to characterize the distributions which attain the lower bound of the weighted residual inaccuracy measure as given in the above theorem, let us assume that $H^w(f,g;t) = -m_X(t)\ln h_G(t)$. Then, differentiating with respect to t and simplifying, we get $g'(t)/g(t) = 0$, which in turn gives via $(\mathrm{d}/\mathrm{d}t)h_G(t) = (h_G(t))^2$ that $h_G(t)$ cannot be decreasing, constant or zero. So the inequality of (2.2) is strict.

In the following theorem we provide an upper bound for the weighted residual inaccuracy measure. The proof is immediate from (2.3), and hence omitted.

Theorem 2.2. *If the hazard rate function $h_G(t)$ is increasing in t, then*

$$H^w(f,g;t) \leqslant -m_X(t)\ln h_G(t) - \frac{1}{\overline{F}(t)}\int_t^\infty xf(x)\ln\Bigl(\frac{\overline{G}(x)}{\overline{G}(t)}\Bigr)\,\mathrm{d}x. \tag{2.4}$$

Example 2.1. Let X be a nonnegative random variable with probability density function

$$f(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \text{otherwise}, \end{cases}$$

and Y is uniformly distributed over $(0,1)$. Then $h_G(t) = 1/(1-t)$, which is increasing in t, $m_X(t) = 2(t^2+t+1)/(3(t+1))$ and $H^w(f,g;t) = 2(t^2+t+1)\ln(1-t)/(3(t+1))$. Note that the right hand side of (2.4) is greater than $2(t^2+t+1)\ln(1-t)/(3(t+1))$. It is easily seen that (2.4) is fulfilled.

R e m a r k 2.2. Proceeding analogously as in Remark 2.1, we can show that the equality in (2.4) holds if and only if Y follows exponential distribution.

Below, we study the weighted residual inaccuracy measure under monotonic transformations in analogy with Di Crescenzo and Longobardi [7].

Theorem 2.3. *Let X and Y be two absolutely continuous nonnegative random variables. Suppose $\varphi(x)$ is strictly monotonic, continuous and differentiable function with derivative $\varphi'(x)$. Then*

$$H^w(\varphi(X),\varphi(Y);t) = \begin{cases} H^{w,\varphi}(X,Y;\varphi^{-1}(t)) \\ \quad +E[\varphi(X)\ln\varphi'(X) \mid X > \varphi^{-1}(t)], & \varphi \text{ strictly increasing}, \\ \overline{H}^{w,\varphi}(X,Y;\varphi^{-1}(t)) \\ \quad +E[\varphi(X)\ln -\varphi'(X) \mid X \leqslant \varphi^{-1}(t)], & \varphi \text{ strictly decreasing}, \end{cases}$$

where

$$H^{w,\varphi}(X,Y;t) = -\int_t^\infty \varphi(x)\frac{f(x)}{\overline{F}(t)}\ln\frac{g(x)}{\overline{G}(t)}\,\mathrm{d}x$$

and

$$\overline{H}^{w,\varphi}(X,Y;t) = -\int_0^t \varphi(x)\frac{f(x)}{F(t)}\ln\frac{g(x)}{G(t)}\,\mathrm{d}x,$$

which is the weighted past inaccuracy measure corresponding to weight function $\varphi(x)$ as discussed in the next section.

R e m a r k 2.3. For two absolutely continuous nonnegative random variables X and Y

$$H^w(aX,aY;t) = aH^w\Big(f,g;\frac{t}{a}\Big) + m_X\Big(\frac{t}{a}\Big)a\ln a$$

for all $a > 0$ and $t > 0$. Furthermore, for all $b > 0$ and $t > b$

$$H^w(X+b,Y+b;t) = H^w(f,g;t-b) + bH(f,g;t-b),$$

where $H(f,g;t) = -\int_t^\infty (f(x)/\overline{F}(t))\ln(g(x)/\overline{G}(t))\,\mathrm{d}x$ is the residual inaccuracy measure given by Taneja et al. [29].

In order to provide characterization results we define the proportional hazard rate model (PHRM, cf. Cox [3]), proportional reversed hazard rate model (PRHRM, cf. Gupta et al. [12]) and the geometric vitality function (cf. Nair and Rajesh [22]). Let X and Y be two random variables with hazard rate functions $h_F(t)$, $h_G(t)$ and reversed hazard rate functions $\varphi_F(t)$ $(= f(t)/F(t))$, $\varphi_G(t)$, respectively.

Definition 2.1. Two random variables X and Y are said to satisfy the PHRM, if there exists $\theta > 0$ such that $h_G(t) = \theta h_F(t)$, or equivalently, $\overline{G}(t) = [\overline{F}(t)]^\theta$, for some θ.

Definition 2.2. Two random variables X and Y are said to satisfy the PRHRM, if there exists $\theta > 0$ such that $\varphi_G(t) = \theta\varphi_F(t)$. Or, equivalently, $G(t) = [F(t)]^\theta$, for some θ.

Definition 2.3. The geometric vitality function of a left truncated random variable is given by

$$\mathcal{G}_X(t) = E[\ln X \mid X > t]$$

and the corresponding weighted version of it is given by $\mathcal{G}_X^w(t) = E[X \ln X \mid X > t]$, provided $E(\ln X)$ is finite.

The PHRM model has been widely used in analyzing survival data; see, for instance, Cox [5], Ebrahimi and Kirmani [8], Gupta and Han [13], and Nair and Gupta [21]. The PRHRM model is flexible enough to accommodate both monotonic and non-monotonic failure rates even though the baseline failure rate is monotonic. See Sengupta et al. [26], Di Crescenzo [6] or Gupta and Gupta [11] for some results on this model. For more properties and applications of the geometric vitality function, one may refer to Nair and Rajesh [22].

Now we provide characterization theorems for some continuous distributions using hazard rate, conditional mean, weighted geometric vitality function and weighted residual inaccuracy measure under PHRM and PRHRM. Below, we characterize the uniform distribution.

Theorem 2.4. *Let X and Y be two absolutely continuous random variables satisfying PRHRM with proportionality constant θ (> 0). A relation of the form*

$$H^w(f,g;t) + m_X(t)\ln h_G(t) = (1-\theta)[\mathcal{G}_Z^w(t) - m_X(t)\ln(t-\alpha)], \tag{2.5}$$

where $\mathcal{G}_Z^w(t) = E[X\ln(X-\alpha) \mid X > t]$ and $\alpha < t < \beta$, holds if and only if X denotes the random lifetime of a component with uniform distribution over (α,β).

P r o o f. The *if part* is obtained from (2.1). To prove the converse, let us assume that (2.5) holds. Then by definition we can write

$$\begin{aligned}
-\int_t^\infty & x f(x) \ln \frac{g(x)}{\overline{G}(t)}\,\mathrm{d}x + \ln h_G(t)\int_t^\infty x f(x)\,\mathrm{d}x \\
&= (1-\theta)\Big[\int_t^\infty x\ln(x-\alpha) f(x)\,\mathrm{d}x - \ln(t-\alpha)\int_t^\infty x f(x)\,\mathrm{d}x\Big].
\end{aligned}$$

Differentiating with respect to t, we get after some algebraic calculations

$$g(t) = k(t-\alpha)^{\theta-1}, \quad k > 0 \text{ (constant)},$$

which gives the required result. $\square$

Next, we give a theorem which characterizes the power distribution.

Theorem 2.5. *For two absolutely continuous random variables X and Y satisfying PRHRM with proportionality constant θ (> 0), the relation*

$$H^w(f,g;t) + m_X(t)\ln h_G(t) = (1-c\theta)[\mathcal{G}_X^w(t) - m_X(t)\ln t], \tag{2.6}$$

for all $0 < t < b$, characterizes the power distribution

$$F(t) = \begin{cases} \left(\dfrac{t}{b}\right)^c, & 0 < t < b,\ b, c > 0 \\ 0, & \textit{otherwise.} \end{cases} \tag{2.7}$$

Proof. If X follows the power distribution as given in (2.7), then (2.6) is obtained from (2.1). To prove the converse, let us assume that (2.6) holds. Then differentiating with respect to t, we get, after some algebraic calculations,

$$g(t) = kt^{c\theta-1}, \quad k > 0 \text{ (constant)},$$

which gives the required result. $\square$

Below, we characterize exponential distribution under PHRM.

Theorem 2.6. *Let X and Y be two absolutely continuous random variables satisfying PHRM with proportionality constant θ (> 0). A relation of the form*

$$H^w(f,g;t) + m_X(t)\ln h_G(t) = \lambda\theta[m_{X^2}(t) - tm_X(t)], \tag{2.8}$$

where $m_{X^2}(t) = E[X^2 \mid X > t]$, the conditional expectation of X^2, holds for all $t > 0$ if and only if X follows exponential distribution with mean $1/\lambda$.

Proof. The *if part* is straightforward. To prove the converse, let us assume that (2.8) holds. Then differentiating with respect to t, we get, after some algebraic calculations,

$$g(t) = k\mathrm{e}^{-\lambda\theta t}, \quad k > 0 \text{ (constant)},$$

which gives the required result. $\square$

Next, we provide characterization of Weibull and Rayleigh distributions.

Theorem 2.7. *Let X and Y be two absolutely continuous random variables satisfying PHRM with proportionality constant θ (>0). A relation of the form*

$$\begin{aligned} H^w(f,g;t) + m_X(t)\ln h_G(t) \\ = (1-p)[\mathcal{G}_X^w(t) - m_X(t)\ln t] + \theta[m_{X^{p+1}}(t) - t^p m_X(t)], \end{aligned} \tag{2.9}$$

where $m_{X^{p+1}}(t) = E[X^{p+1} \mid X > t]$, the conditional expectation of X^{p+1}, holds for all $t > 0$, $p > 0$ if and only if X follows Weibull distribution

$$F(t) = 1 - \mathrm{e}^{-t^p}, \quad t > 0,\ p > 0.$$

Proof. The *if part* is straightforward. To prove the converse, let us assume that (2.9) holds. Then differentiating with respect to t, we get, after some algebraic calculations,

$$g(t) = kt^{(p-1)}\mathrm{e}^{-\theta t^p}, \quad k > 0 \text{ (constant)},$$

which gives the required result. □

Corollary 2.1. *Let X and Y be two absolutely continuous random variables satisfying PHRM with proportionality constant θ (>0). A relation of the form*

$$H^w(f,g;t) + m_X(t)\ln\Big(\frac{h_G(t)}{t}\Big) + \mathcal{G}_X^w(t) = \theta[m_{X^3}(t) - t^2 m_X(t)],$$

where $m_{X^3}(t) = E[X^3 \mid X > t]$, the conditional expectation of X^3, holds for all $t > 0$ if and only if X follows Rayleigh distribution $\overline{F}(t) = \mathrm{e}^{-t^2}$, $t > 0$.

Now we consider Pareto-type distributions which are flexible parametric models and play an important role in reliability, actuarial science, economics, finance, and telecommunications. Arnold [1] proposed a general version of this family of distributions called Pareto-IV distribution having the cumulative distribution function

$$F(x) = 1 - \Big[1 + \Big(\frac{x-\mu}{\beta}\Big)^{1/\gamma}\Big]^{-\alpha}, \quad x > \mu, \tag{2.10}$$

where $-\infty < \mu < \infty$, $\beta > 0$, $\gamma > 0$, and $\alpha > 0$ are location, scale, inequality, and shape parameters, respectively. This distribution is related to many other families of distributions. For example, setting $\alpha = 1$, $\gamma = 1$ and $(\gamma = 1, \mu = \beta)$ in (2.10), one at a time, we obtain Pareto-III, Pareto-II, and Pareto-I distributions, respectively. Also, taking $\mu = 0$ and $\gamma \to 1/\gamma$ in (2.10), we obtain Burr-XII distribution.

Now we consider Pareto-type distributions for characterization under PHRM. Below, we provide characterization of Pareto-I distribution.

Theorem 2.8. *Let X and Y be two absolutely continuous random variables satisfying PHRM with proportionality constant θ (>0). Then a relation*

$$H^w(f,g;t) + m_X(t)\ln h_G(t) = (\alpha\theta+1)[\mathcal{G}_X^w(t) - m_X(t)\ln t], \tag{2.11}$$

holds for all $t>\beta$ if and only if X follows Pareto-I distribution

$$F(t) = 1 - \Big(\frac{\beta}{t}\Big)^{\alpha}, \quad t>\beta,\ \alpha,\beta>0.$$

Proof. The *if part* is straightforward. To prove the converse, let us assume that (2.11) holds. Then, differentiating with respect to t, we get, after some algebraic calculations,

$$g(t) = kt^{-(\alpha\theta+1)}, \quad k>0 \text{ (constant)},$$

which gives the required result. □

We conclude this section by characterizing Pareto-II distribution. The proof is similar to that of Theorem 2.8 and hence omitted.

Theorem 2.9. *Let X and Y be two absolutely continuous random variables satisfying PHRM with proportionality constant θ (>0). Then a relation*

$$H^w(f,g;t) + m_X(t)\ln h_G(t) = (\alpha\theta+1)[\mathcal{G}_Z^w(t) - m_X(t)\ln(t-\mu+\beta)],$$

where $\mathcal{G}_Z^w(t) = E[X\ln(X-\mu+\beta)\mid X>t]$ holds for all $t>\mu$ if and only if X follows Pareto-II distribution

$$F(t) = 1 - \Big[1 + \Big(\frac{t-\mu}{\beta}\Big)\Big]^{-\alpha}, \quad t>\mu.$$

3. Characterizations based on weighted past inaccuracy measure

In this section we consider the same distributions as in the previous section for characterization based on weighted past inaccuracy measure. First we review some properties of the weighted past inaccuracy measure.

Kumar and Taneja [18] introduced the notion of weighted past inaccuracy measure of a random variable X truncated above some t as

$$\overline{H}^w(f,g;t) = -\int_0^t x\frac{f(x)}{F(t)}\ln\Big(\frac{g(x)}{G(t)}\Big)\,\mathrm{d}x \tag{3.1}$$

and studied various aspects of this measure in analogy with weighted past entropy. In agreement with Theorem 4.2 of Kumar and Taneja [18], a sharper upper bound for the weighted past inaccuracy measure is given in the following theorem.

Theorem 3.1. *If $\varphi_G(t)$ is decreasing in t, then*

$$\overline{H}^w(f,g;t) \leqslant -\overline{m}_X(t)\ln\varphi_G(t) + \frac{1}{F(t)}\int_0^t xf(x)\ln\Big(\frac{G(t)}{G(x)}\Big)\,\mathrm{d}x, \tag{3.2}$$

where $\overline{m}_X(t) = E[X \mid X < t]$, the conditional mean of the right truncated random variable $[X \mid X < t]$.

P r o o f. Note that (3.1) can be written as

$$\overline{H}^w(f,g;t) = -\frac{1}{F(t)}\int_0^t xf(x)\ln\varphi_G(x)\,\mathrm{d}x - \frac{1}{F(t)}\int_0^t xf(x)\ln\Big(\frac{G(x)}{G(t)}\Big)\,\mathrm{d}x.$$

Hence we obtain the result by using the fact that $\ln\varphi_G(x) \geqslant \ln\varphi_G(t)$, for $x \leqslant t$, if $\varphi_G(\cdot)$ is a decreasing function. $\square$

E x a m p l e 3.1. Let X and Y be two nonnegative random variables as given in Example 2.1. Then $\varphi_G(t) = 1/t$, which is decreasing in t, $\overline{m}_X(t) = \frac{2}{3}t$ and $\overline{H}^w(f,g;t) = \frac{2}{3}t\ln t$, $t > 0$. Note that the right-hand side of (3.2) is $\frac{2}{3}t\ln t + \frac{2}{9}t$. Hence, (3.2) is fulfilled.

R e m a r k 3.1. In order to characterize the distributions which attain the upper bound of the weighted past inaccuracy measure as given in the above theorem, let us assume that the equality in (3.2) holds. Then, differentiating with respect to t and simplifying we get $\varphi_G(t) = $ constant, which contradicts the fact that X is nonnegative random variable. So the inequality of (3.2) is strict.

In the following theorem we provide a lower bound for the weighted past inaccuracy measure. The proof is analogous to Theorem 4.2 of Kumar and Taneja [18] but for completeness we give a brief outline of the proof.

Theorem 3.2. *For two absolutely continuous nonnegative random variables X and Y,*

$$\overline{H}^w(f,g;t) \geqslant \overline{m}_X(t) - \frac{1}{F(t)}\int_0^t xf(x)\varphi_G(x)\,\mathrm{d}x. \tag{3.3}$$

P r o o f. From (3.1), we have

$$\begin{aligned}\overline{H}^w(f,g;t) &= -\frac{1}{F(t)}\int_0^t xf(x)\ln\varphi_G(x)\,\mathrm{d}x + \frac{1}{F(t)}\int_0^t xf(x)\ln\Big(\frac{G(t)}{G(x)}\Big)\,\mathrm{d}x\\ &\geqslant -\frac{1}{F(t)}\int_0^t xf(x)\ln\varphi_G(x)\,\mathrm{d}x\\ &\geqslant \frac{1}{F(t)}\int_0^t xf(x)(1-\varphi_G(x))\,\mathrm{d}x\\ &\geqslant \overline{m}_X(t) - \frac{1}{F(t)}\int_0^t xf(x)\varphi_G(x)\,\mathrm{d}x,\end{aligned}$$

where the second-last inequality follows from the fact that $\ln x \leqslant x-1$ for $x>0$. $\square$

E x a m p l e 3.2. Let X and Y be two nonnegative random variables as given in Example 2.1. Then $\overline{H}^w(f,g;t) = \frac{2}{3}t\ln t$ and the right-hand side of (3.3) is $\frac{2}{3}t-1$, for $t\in(0,1)$. Denote $\psi(t)=\frac{2}{3}t(\ln t-1)+1$, which is decreasing in t with $\psi(1)>0$. Hence it is easily seen that (3.3) is fulfilled.

R e m a r k 3.2. Proceeding analogously as in Remark 3.1, we can show that there is no nonnegative random variable which attains the lower bound of the weighted past inaccuracy measure and the inequality of (3.3) is strict.

In analogy with Theorem 2.3, we obtain results on weighted past inaccuracy measure under monotonic transformations.

Theorem 3.3. *Let X and Y be two absolutely continuous nonnegative random variables. Suppose $\varphi(x)$ is strictly monotonic, continuous and differentiable function with derivative $\varphi'(x)$. Then*

$$\overline{H}^w(\varphi(X),\varphi(Y);t) = \begin{cases} \overline{H}^{w,\varphi}(X,Y;\varphi^{-1}(t)) \\ \quad +E[\varphi(X)\ln\varphi'(X) \mid X\leqslant\varphi^{-1}(t)], & \varphi \text{ strictly increasing},\\ H^{w,\varphi}(X,Y;\varphi^{-1}(t)) \\ \quad +E[\varphi(X)\ln-\varphi'(X) \mid X>\varphi^{-1}(t)], & \varphi \text{ strictly decreasing},\end{cases}$$

where $\overline{H}^{w,\varphi}(X,Y;t)$ and $H^{w,\varphi}(X,Y;t)$ are as defined in Theorem 2.3.

R e m a r k 3.3. For two absolutely continuous nonnegative random variables X and Y

$$\overline{H}^w(aX, aY; t) = a\overline{H}^w\Big(f, g; \frac{t}{a}\Big) + \overline{m}_X\Big(\frac{t}{a}\Big) a \ln a$$

for all $a > 0$ and $t > 0$. Furthermore, for all $b > 0$ and $t > b$

$$\overline{H}^w(X + b, Y + b; t) = \overline{H}^w(f, g; t - b) + b\overline{H}(f, g; t - b),$$

where $\overline{H}(f, g; t) = -\int_0^t (f(x)/F(t)) \ln(g(x)/G(t))\,\mathrm{d}x$ is the past inaccuracy measure given by Kumar et al. [20].

Now we provide characterization theorems for the same distributions as considered in the previous section using reversed hazard rate, conditional mean, weighted geometric vitality function and weighted past inaccuracy measure under PHRM and PRHRM. Recall that weighted geometric vitality function of a right truncated random variable is given by $\overline{\mathcal{G}}^w_X(t) = E[X \ln X \mid X < t]$. Below, we characterize the uniform distribution.

Theorem 3.4. *Let X and Y be two absolutely continuous random variables satisfying PRHRM with proportionality constant θ (> 0). A relation of the form*

$$\overline{H}^w(f, g; t) + \overline{m}_X(t) \ln \varphi_G(t) = (1 - \theta)[\overline{\mathcal{G}}^w_Z(t) - \overline{m}_X(t) \ln(t - \alpha)], \tag{3.4}$$

where $\overline{\mathcal{G}}^w_Z(t) = E[X \ln(X - \alpha) \mid X < t]$ and $\alpha < t < \beta$, holds if and only if X denotes the random lifetime of a component with uniform distribution over (α, β).

P r o o f. The *if part* is obtained from (3.1). To prove the converse, let us assume that (3.4) holds. Then by definition we can write

$$-\int_0^t xf(x) \ln \frac{g(x)}{G(t)}\,\mathrm{d}x + \ln \varphi_G(t) \int_0^t xf(x)\,\mathrm{d}x$$
$$= (1 - \theta)\Big[\int_0^t x \ln(x - \alpha) f(x)\,\mathrm{d}x - \ln(t - \alpha) \int_0^t xf(x)\,\mathrm{d}x\Big].$$

Differentiating with respect to t, we get after some algebraic calculations

$$g(t) = k(t - \alpha)^{\theta - 1}, \quad k > 0 \text{ (constant)},$$

which gives the required result. □

Next, we give a theorem which characterizes the power distribution. The proof follows the same lines as that of Theorem 2.5.

Theorem 3.5. *For two absolutely continuous random variables X and Y satisfying PRHRM with proportionality constant θ (> 0), the relation*

$$\overline{H}^w(f,g;t) + \overline{m}_X(t)\ln\varphi_G(t) = (1-c\theta)[\overline{\mathcal{G}}^w_X(t) - \overline{m}_X(t)\ln t],$$

for all $0 < t < b$, characterizes the power distribution as given in (2.7).

Below, we characterize exponential distribution under PHRM. The proof is similar to that of Theorem 2.6 and hence omitted.

Theorem 3.6. *Let X and Y be two absolutely continuous random variables satisfying PHRM with proportionality constant θ (> 0). A relation of the form*

$$\overline{H}^w(f,g;t) + \overline{m}_X(t)\ln\varphi_G(t) = \lambda\theta[\overline{m}_{X^2}(t) - t\overline{m}_X(t)],$$

where $\overline{m}_{X^2}(t) = E[X^2 \mid X < t]$, the conditional expectation of X^2, holds for all $t > 0$ if and only if X follows exponential distribution with mean $1/\lambda$.

Next, we provide characterization of Weibull and Rayleigh distributions. The proof is similar to that of Theorem 2.7 and hence omitted.

Theorem 3.7. *Let X and Y be two absolutely continuous random variables satisfying PHRM with proportionality constant θ (> 0). A relation of the form*

$$\overline{H}^w(f,g;t) + \overline{m}_X(t)\ln\varphi_G(t) = (1-p)[\overline{\mathcal{G}}^w_X(t) - \overline{m}_X(t)\ln t] + \theta[\overline{m}_{X^{p+1}}(t) - t^p\overline{m}_X(t)],$$

where $\overline{m}_{X^{p+1}}(t) = E[X^{p+1} \mid X < t]$, the conditional expectation of X^{p+1}, holds for all $t > 0$, $p > 0$ if and only if X follows Weibull distribution as given in Theorem 2.7.

Corollary 3.1. *Let X and Y be two absolutely continuous random variables satisfying PHRM with proportionality constant θ (> 0). A relation of the form*

$$\overline{H}^w(f,g;t) + \overline{m}_X(t)\ln\left(\frac{\varphi_G(t)}{t}\right) + \overline{\mathcal{G}}^w_X(t) = \theta[\overline{m}_{X^3}(t) - t^2\overline{m}_X(t)],$$

where $\overline{m}_{X^3}(t) = E[X^3 \mid X < t]$, the conditional expectation of X^3, holds for all $t > 0$ if and only if X follows Rayleigh distribution $\overline{F}(t) = \mathrm{e}^{-t^2}$, $t > 0$.

Below, we provide characterization of Pareto-I distribution. The proof is similar to that of Theorem 2.8 and hence omitted.

Theorem 3.8. *Let X and Y be two absolutely continuous random variables satisfying PHRM with proportionality constant θ (> 0). Then a relation*

$$\overline{H}^w(f, g; t) + \overline{m}_X(t) \ln \varphi_G(t) = (\alpha\theta + 1)[\overline{\mathcal{G}}_X^w(t) - \overline{m}_X(t) \ln t],$$

holds for all $t > \beta$ if and only if X follows Pareto-I distribution as given in Theorem 2.8.

Last, we characterize Pareto-II distribution. The proof is similar to that of Theorem 2.9 and hence omitted.

Theorem 3.9. *Let X and Y be two absolutely continuous random variables satisfying PHRM with proportionality constant θ (> 0). Then a relation*

$$\overline{H}^w(f, g; t) + \overline{m}_X(t) \ln \varphi_G(t) = (\alpha\theta + 1)[\overline{\mathcal{G}}_Z^w(t) - \overline{m}_X(t) \ln(t - \mu + \beta)],$$

where $\overline{\mathcal{G}}_Z^w(t) = E[X \ln(X - \mu + \beta) \mid X < t]$ holds for all $t > \mu$ if and only if X follows Pareto-II distribution as given in Theorem 2.9.

4. Conclusion

The investigation of measures of information is an issue of fundamental importance in different areas of science and engineering. In recent years, various authors have shown interest in studying the weighted version of Kerridge inaccuracy measure for truncated distributions. Here we consider weighted residual (past) inaccuracy measure and study their bounds. Characterizations of some commonly used continuous distributions have also been provided. In conclusion, some discussion is made on the empirical version of the inaccuracy measure.

Let $X_1, X_2, \ldots, X_n$ be nonnegative, absolutely continuous independent and identically distributed (iid) random variables, that constitute a random sample from a population having the distribution function $F(x)$. Also, let us consider another random sample $Y_1, Y_2, \ldots, Y_n$ of nonnegative, absolutely continuous iid random variables from $G(x)$. Then the empirical inaccuracy measure is defined as

$$H(\hat{f}_n, \hat{g}_n) = -\int_0^\infty \hat{f}_n(u)\hat{g}_n(u)\,\mathrm{d}u,$$

where $\hat{f}_n$, $\hat{g}_n$ are the empirical densities of the samples. Now we study some statistical interpretations of the inaccuracy measure in connection with maximum likelihood estimate (MLE).

Suppose each sample value X_i, $i = 1, 2, \ldots, n$, is assigned probability $1/n$, then

$$\hat{f}_n(\widetilde{X}) = \frac{1}{n} \sum_{i=1}^{n} \delta(\widetilde{X} - X_i).$$

Also let $G_\theta(\widetilde{X})$ be a statistical model $g(\widetilde{X} \mid \tilde{\theta})$ with unknown parameter $\tilde{\theta}$. Then the empirical version of the inaccuracy measure is

$$\begin{aligned} H(\hat{f}_n, g_\theta) &= -\int_0^\infty \ln g(\widetilde{X} \mid \tilde{\theta}) \left[\frac{1}{n} \sum_{i=1}^{n} \delta(\widetilde{X} - X_i) \right] \mathrm{d}\widetilde{X} \\ &= -\frac{1}{n} \sum_{i=1}^{n} \ln g(X_i \mid \tilde{\theta}) \\ &= -\frac{1}{n} \ln \prod_{i=1}^{n} g(X_i \mid \tilde{\theta}), \end{aligned}$$

which is just the log-likelihood function apart from the factor $(-1/n)$. It can be shown that $E_X[H(\hat{f}_n, g_\theta)] = H(f, g_\theta)$, i.e., the empirical version reduces to the population version for any n by taking its expectation. Hence, from $H(f, g) = H(f) + D(f||g)$, maximizing the likelihood to find the MLE is the same as finding $\tilde{\theta}$ which minimizes $H(f, g_\theta)$ or $D(f||g_\theta)$. It is obvious that the best possible model is the one that fits the data exactly, i.e., when $f(x) = g_\theta(x)$. Therefore, for any general model $g_\theta(x)$

$$H(f, g_\theta) \geqslant H(f),$$

which can also be verified from the nonnegativity property of Kullback-Leibler divergence measure.

Acknowledgement. The author thanks the editor and the anonymous reviewer for his/her valuable comments and suggestions on the earlier version of the paper which have led to considerable improvement in the contents.

References

[1] *B. C. Arnold*: Pareto Distributions. Statistical Distributions in Scientific Work 5, International Co-operative Publishing House, Burtonsville, 1983.

[2] *T. A. Azlarov, N. A. Volodin*: Characterization Problems Associated with the Exponential Distribution. Transl. from the Russian. Springer, New York, 1986.

[3] *D. R. Cox*: The analysis of exponentially distributed life-times with two types of failure. J. R. Stat. Soc., Ser. B *21* (1959), 411–421.

[4] *D. R. Cox*: Renewal Theory. Methuens Monographs on Applied Probability and Statistics, Methuen, London; John Wiley, New York, 1962.

[5] *D. R. Cox*: Regression models and life-tables. J. R. Stat. Soc., Ser. B *34* (1972), 187–220.

[6] *A. Di Crescenzo*: Some results on the proportional reversed hazards model. Stat. Probab. Lett. *50* (2000), 313–321.
[7] *A. Di Crescenzo, M. Longobardi*: On weighted residual and past entropies. Sci. Math. Jpn. *64* (2006), 255–266.
[8] *N. Ebrahimi, S. N. U. A. Kirmani*: A characterisation of the proportional hazards model through a measure of discrimination between two residual life distributions. Biometrika *83* (1996), 233–235.
[9] *E. Furman, R. Zitikis*: Weighted premium calculation principles. Insur. Math. Econ. *42* (2008), 459–465.
[10] *J. Galambos, S. Kotz*: Characterizations of Probability Distributions. A Unified Approach with an Emphasis on Exponential and Related Models. Lecture Notes in Mathematics 675, Springer, Berlin, 1978.
[11] *R. C. Gupta, R. D. Gupta*: Proportional reversed hazard rate model and its applications. J. Stat. Plann. Inference *137* (2007), 3525–3536.
[12] *R. C. Gupta, P. L. Gupta, R. D. Gupta*: Modeling failure time data by Lehman alternatives. Commun. Stat., Theory Methods *27* (1998), 887–904.
[13] *R. C. Gupta, W. Han*: Analyzing survival data by PRH models. International Journal of Reliability and Applications *2* (2001), 203–216.
[14] *R. C. Gupta, S. N. U. A. Kirmani*: The role of weighted distributions in stochastic modeling. Commun. Stat., Theory Methods *19* (1990), 3147–3162.
[15] *K. Jain, H. Singh, I. Bagai*: Relations for reliability measures of weighted distributions. Commun. Stat., Theory Methods *18* (1989), 4393–4412.
[16] *D. F. Kerridge*: Inaccuracy and inference. J. R. Stat. Soc., Ser. B *23* (1961), 184–194.
[17] *S. Kullback, R. A. Leibler*: On information and sufficiency. Ann. Math. Stat. *22* (1951), 79–86.
[18] *V. Kumar, H. C. Taneja*: On length biased dynamic measure of past inaccuracy. Metrika *75* (2012), 73–84.
[19] *V. Kumar, H. C. Taneja, R. Srivastava*: Length biased weighted residual inaccuracy measure. Metron *68* (2010), 153–160.
[20] *V. Kumar, H. C. Taneja, R. Srivastava*: A dynamic measure of inaccuracy between two past lifetime distributions. Metrika *74* (2011), 1–10.
[21] *N. U. Nair, R. P. Gupta*: Characterization of proportional hazard models by properties of information measures. International Journal of Statistical Sciences *6* (2007), Special Issue, 223–231.
[22] *K. R. M. Nair, G. Rajesh*: Geometric vitality function and its applications to reliability. IAPQR Trans. *25* (2000), 1–8.
[23] *A. K. Nanda, K. Jain*: Some weighted distribution results on univariate and bivariate cases. J. Stat. Plann. Inference *77* (1999), 169–180.
[24] *G. P. Patil, J. K. Ord*: On size-biased sampling and related form-invariant weighted distributions. Sankhyā, Ser. B *38* (1976), 48–61.
[25] *C. R. Rao*: Linear Statistical Inference and Its Applications. John Wiley & Sons, New York, 1965.
[26] *D. Sengupta, H. Singh, A. K. Nanda*: The proportional reversed hazard model. Technical Report 1999, Indian Statistical Institute, Calcutta.
[27] *C. E. Shannon*: A mathematical theory of communication. Bell Syst. Tech. J. *27* (1948), 379–423, 623–656.
[28] *S. Smitha*: A Study on the Kerridge's Inaccuracy Measure and Related Concepts. Doctoral Dissertation 2010, CUSAT.
[29] *H. C. Taneja, V. Kumar, R. Srivastava*: A dynamic measure of inaccuracy between two residual lifetime distributions. Int. Math. Forum *4* (2009), 1213–1220.

[30] *G. Wallis*: Using spatio-temporal correlations to learn invariant object recognition. Neural Netw. *9* (1996), 1513–1519.

Author's address: *Chanchal Kundu*, Department of Mathematics, Rajiv Gandhi Institute of Petroleum Technology, Rae Bareli 229 316, U.P., India, e-mail: ckundu@rgipt.ac.in.