4.1 Introduction

It is well known that the knowledge and use of various methods for information coding and transmission play a vital role in understanding and modeling many aspects of biological system features. As explained in the first Chapter, Shannon's entropy plays an important role in the context of information theory. Since Shannon's entropy is not sufficient enough for the study of the remaining life of system that have survived for some units of time, Ebrahimi and Pellerey (1995) proposed a new measure of uncertainty called residual entropy, a measure that plays a vital role in left truncated data sets. However, Di Crescenzo and Longobardi (2002) showed that in many realistic situations, uncertainty is not necessarily related to the future but can also be refer to the past. For instance, if at time $t$, a system that is observed only at certain pre assigned inspection times is found to be down, then the uncertainty of the system life relies on the past, i.e., on which instant in $(0,t)$ it has failed. Based on this idea, Di Crescenzo and Longobardi (2002) introduced the past entropy over $(0,t)$. They showed the necessity of past entropy using an example and discussed its relationship with residual entropy and studied the monotonic behaviors of it. Let the rv $X$ denote the lifetime of a component/system or of living organism, then past entropy of $X$ at time $t$ is defined as

---
*Some of the results in this Chapter have been communicated to two International Journals.*
\[ R(t) = -\int_0^t \frac{f(x)}{F(t)} \left( \log \frac{f(x)}{F(t)} \right) dx. \]  

(4.1)

Note that (4.1) can be rewritten as

\[ \bar{H}(t) = 1 - \frac{1}{F(t)} \int_0^t (\log \lambda(x)) f(x) dx. \]

(4.2)

Recently, Nanda and Paul (2006) proved some ordering properties based on past entropy and some sufficient conditions for these orders to hold. They also introduced a non-parametric class based on past entropy and studied its properties and examined it under the discrete setup. However, Di Crescenzo and Longobardi (2006) introduced the notion of weighted residual and past entropies and studied its properties and monotone behavior of it. In view of the usefulness of measure of uncertainty (4.1) in past time, in the present chapter, we further explore the same and also define a new conditional measure and study its properties. In Section 4.4, we study Renyi’s entropy for the past lifetime and proved some theorems arising out of it. Further, in Section 4.5, we extend these concepts in the context of weighted models and also study some ordering and aging properties based on these measures. In the final two sections, we discuss some measures of discrimination proposed by Di Crescenzo and Longobardi (2004) and Asadi et al. (2005) and study its applications in the context of weighted models.

4.2 Properties

Differentiating (4.1) with respect to \( t \), we get

\[ \bar{H}'(t) = \lambda(t) \left[ 1 - \bar{H}(t) - \log \lambda(t) \right]. \]

(4.3)

Nanda and Paul (2006) proved that if \( X \) has absolutely continuous distribution function \( F(t) \) and an increasing past entropy \( \bar{H}(t) \), then \( \bar{H}(t) \) uniquely determines \( F(t) \).

Next we prove a characterization theorem for the power distribution using the functional relationships between reversed hazard rate and past entropy.

Theorem 4.1: Let \( X \) be a non-negative rv admitting an absolutely continuous df such that \( E(X) < \infty \) and let \( \bar{H}(t) \) be defined as in (4.1). The relationship
\[ H(t) = k - \log \lambda(t) ; \quad 0 < k < 1 \quad (4.4) \]

where \( k \) is a constant, holds for all \( t > 0 \) if and only if \( X \) follows power distribution with df

\[ F(t) = \left( \frac{t}{b} \right)^c , \quad 0 < t < b , \ b, c > 0 . \quad (4.5) \]

**Proof:** Suppose the relation (4.4) holds. Differentiating (4.4) with respect to \( t \) implies that

\[ H'(t) = - \frac{\lambda'(t)}{\lambda(t)} . \quad (4.6) \]

From (4.6) and (4.3) we get

\[ \lambda'(t) + \lambda^2(t)(1-k) = 0 . \quad (4.7) \]

Divide each term of (4.7) by \( \lambda^2(t) \), we get

\[ \frac{\lambda'(t)}{\lambda^2(t)} = k - 1 . \quad (4.8) \]

Putting \( u(t) = \frac{1}{\lambda(t)} \), (4.8) becomes

\[ - \frac{du}{dt} = (k - 1) . \quad (4.9) \]

Solving the differential equation (4.9), we obtain \( u(t) = At + B \), where \( A = (1-k) \). This implies \( \lambda(t) = \frac{1}{(At + B)} \), now from the uniqueness property of reversed hazard rate, we obtain the required result (4.5).

Conversely when (4.5) holds, substituting (4.5) in (4.2) and on direct calculation, we obtain (4.4) with \( k = \frac{(c-1)}{c} \).

The following Theorems characterize the exponential distribution and exponential family of distributions using the possible relationships between RHR, EIT and the past entropy.
Theorem 4.2: For the rv $X$ considered in Theorem 4.1, with $\lim_{t \to 0} F(t) = 0$, the relation

$$\tilde{H}(t) + \log \tilde{\lambda}(t) = -C \tilde{r}(t),$$

(4.10)

where $C (>0)$ is a constant holds for all $t \geq 0$ if and only if $X$ follows exponential distribution with distribution function (2.18).

Proof: Suppose the relation (4.10) holds, then differentiating (4.10) with respect to $t$

$$\tilde{H}'(t) + \frac{\tilde{\lambda}'(t)}{\tilde{\lambda}(t)} = -C \tilde{r}'(t).$$

(4.11)

Using (4.3), (1.22) and (4.10), we obtain

$$\tilde{\lambda}'(t) + C\tilde{\lambda}(t) + \tilde{\lambda}^2(t) = 0.$$  

(4.12)

Now solving (4.12) following the similar steps as that of the Theorem 4.1, we obtain

$$\tilde{\lambda}(t) = \frac{C \exp(-Ct)}{1 - \exp(-Ct)}.$$  

Using the uniqueness property of $\tilde{\lambda}(t)$, we get the required model (2.18). Substitution of (2.18) in (4.2) and by direct calculation we obtain the converse part of the theorem with $C = \tilde{\lambda}$.

Theorem 4.3: Let $\lim_{t \to 0} \log a(t) f(t) = 0$, $\lim_{t \to 0} F(t) = 0$, $\bar{m}_p(t) = E \left\{ P(X) \left| X \leq t \right. \right\}$ and $E \left\{ P(X) \right\} < \infty$, then the past entropy of a non-negative rv satisfies a relation of the form

$$\tilde{H}(t) + \log \tilde{\lambda}(t) = P(t) - \bar{m}_p(t) + \partial \tilde{r}(t),$$  

(4.13)

where $P(t)$ is any function of $t$ holds for all $t \geq 0$ if and only if the pdf of $X$ belongs to exponential family (1.28).

Proof: Assume that (4.13) holds. On differentiating (4.13) with respect to $t$, we obtain

$$\tilde{H}'(t) + \frac{\tilde{\lambda}'(t)}{\tilde{\lambda}(t)} = P'(t) - \bar{m}_p'(t) + \partial \tilde{r}'(t).$$

(4.14)

From the definition of $\bar{m}_p(t)$, we have
Using (4.3), (1.22), (4.15) and (4.13) we get (4.14) as

\[ \dot{\lambda}(t) - \left( \dot{P}(t) + \Theta \right) \lambda'(t) + \lambda^2(t) = 0. \]  

(4.16)

Dividing each term by \( \lambda^2(t) \) and putting \( u(t) = \frac{-1}{\lambda(t)} \), (4.16) becomes

\[ \frac{du}{dt} + \left( \theta + \dot{P}(t) \right) u(t) + 1 = 0. \]  

(4.17)

Solving the differential equation (4.17), we get \( \lambda(t) = \frac{\exp(P(t) + \Theta t)}{\int_0^t \exp(P(x) + \Theta x) \, dx} \). Now from the uniqueness property of \( \lambda(t) \) we obtain (1.28).

Conversely, substitution of (1.28) in (4.1) and on simplification, we get

\[ \bar{H}(t) = -\theta \bar{m}(t) - E(C(X) | X \leq t) - D(\Theta) - \log F(t). \]  

(4.18)

Add and subtract \( \log f(t) \) in (4.18) yields (4.13) with \( P(t) = C(t) \).

Our next results provide characterization theorems for the Pareto I distribution and log exponential family of distributions using a functional relationship between the past entropy and geometric vitality function in past time denoted by \( \log G(t) = E(\log X | X \leq t) \).

**Theorem 4.4:** Let \( X \) be a non-negative rv in the support \([k, \infty)\), \( k > 0 \), admitting an absolutely continuous df such that \( E(\log X) < \infty \) and \( \log G(t) = E(\log X | X \leq t) \). Then the relationship

\[ \bar{H}(t) + \log \lambda(t) = C \log \left( \frac{G(t)}{t} \right), \]  

(4.19)

where \( C > 1 \) is a constant holds for all \( t \geq 0 \) if and only if \( X \) follows Pareto I distribution with cdf
\[ F(t) = 1 - \left( \frac{k}{t} \right)^c ; \quad t > k, \ k, c > 0. \]  

(4.20)

**Proof:** Assuming (4.19) and differentiating with respect to \( t \), we have

\[ \bar{H}(t) + \frac{\dot{\lambda}(t)}{\lambda(t)} = C \frac{d}{dt} \left( \log \bar{G}(t) \right) - \frac{C}{t}. \]  

(4.21)

Substituting (4.3), (4.19) and \( \frac{d}{dt} \left( \log \bar{G}(t) \right) = \dot{\lambda}(t) \left( \log t - \log \bar{G}(t) \right) \), and on simplification, (4.21) implies

\[ \dot{\lambda}(t) + \frac{\dot{\lambda}(t)}{\lambda(t)} + \frac{C}{t} = 0. \]  

(4.22)

Now following the similar steps as that of the Theorem 4.2, the solution of the differential equation (4.22) is \( \dot{\lambda}(t) = \frac{(C-1)}{t^c \left( k^{-c+1} - t^{-c+1} \right)} \). From the uniqueness property of \( \dot{\lambda}(t) \), we obtain (4.20).

To prove the converse part, assume (4.20). From a direct calculation we obtain

\[ \bar{H}(t) = \log F(t) - \log ck^c + (c + 1) \log \bar{G}(t). \]  

(4.23)

Add and subtract \( \log f(t) \) in (4.23) and on simplification, we get (4.19).

**Theorem 4.5:** For the rv considered in Theorem 4.1, let \( \lim_{t \to 0} \log Q(t)f(t) = 0 \), \( \log \bar{m}_Q(t) = E \left( \log Q(X) | X \leq t \right) \) and \( E \left( \log Q(X) \right) < \infty \), a relation of the form

\[ \bar{H}(t) + \log \dot{\lambda}(t) = \log Q(t) - \log \bar{m}_Q(t) - \theta \log \left( \frac{\bar{G}(t)}{t} \right) \]  

(4.24)

where \( Q(t) \) is any function of \( t \), holds for all \( t \geq 0 \) if and only if the pdf of \( X \) belongs to log exponential family with probability density function (1.29).

**Proof:** Assuming (4.24) for all \( t \geq 0 \), using the similar steps as that of Theorem (4.4), we have the proof.
4.3 Conditional measure of uncertainty for past lifetime

In continuation of the measure of uncertainty of residual lifetime (1.35) proposed by Ebrahimi and Pellerey (1995), Sankaran and Nair (1999) introduced a conditional measure of uncertainty (1.36) which is defined in Chapter 1. Analogous to $M(t)$ defined in (1.36), for a non-negative rv $X$, we define a conditional measure of uncertainty for the past life as

$$M(t) = E\left(-\log f(X) | X \leq t\right)$$

$$= -\frac{1}{F(t)} \int_0^t f(x) \log f(x) dx. \quad (4.25)$$

Clearly $M(t)$ gives the measure of uncertainty of the past lifetime of a unit. Using (4.25) and (4.1), $M(t)$ can be directly related to $\bar{H}(t)$ and $\lambda(t)$ through the following relationships

$$M(t) = \bar{H}(t) - \log F(t) \quad (4.26)$$

and

$$\lambda(t) = \bar{H}'(t) - M'(t). \quad (4.27)$$

We now give a characterization theorem for the exponential distribution using the conditional measure of uncertainty for the past life defined by (4.25) and the right truncated conditional moment $\bar{m}(t)$.

**Theorem 4.6:** For a rv $X$ considered in Theorem 4.1 with $\lim_{t \to 0} t f(t) = 0$ and $M(t)$ as defined in (4.25). A relation of the form

$$M(t) - \frac{1}{\mu} \bar{m}(t) = k \quad (4.28)$$

where $k$ is a constant, is satisfied for all $t \geq 0$ if and only if $X$ have an exponential distribution with distribution function (2.18).

**Proof:** Assume that the relation (4.28) holds, then by substituting (4.26) and the definition of $\bar{m}(t)$, we get
\[-\frac{1}{F(t)} \int_0^t f(x) \log f(x) dx - \frac{1}{\mu F(t)} \int_0^t xf(x) dx = k. \tag{4.29}\]

Multiply both sides of (4.29) by \( F(t) \) and on differentiation with respect to \( t \) using the condition, we obtain \( f(t) = c \exp \left( \frac{-t}{\mu} \right) \). Applying the boundary conditions, we have \( c = \mu \).

Conversely, when \( X \) is specified by exponential distribution (2.18), from direct calculation of \( \tilde{M}(t) \) using (4.25), we obtain (4.28) with \( k = -\log \lambda \).

In the following theorems we prove certain characterizations to some well-known distributions viz power and Pareto I and families of distributions such as exponential and log exponential using the functional form of \( \tilde{M}(t) \) and \( \log G(t) \).

**Theorem 4.7:** Let \( X \) be a non-negative rv having an absolutely continuous df with \( E(\log X) < \infty \) and \( \log G(t) \) is defined as in Theorem 4.4. Then a relationship

\[ \tilde{M}(t) + (c - 1) \log G(t) = k, \tag{4.30} \]

where \( k \) is a constant, holds for all \( t \geq 0 \) if and only if \( X \) follows power distribution (4.5).

**Proof:** suppose that the relation (4.30) holds. Using (4.26) and the definition of \( \log G(t) \) we get

\[ -\frac{1}{F(t)} \int_0^t \log f(x) f(x) dx + \frac{(c - 1)}{F(t)} \int_0^t \log xf(x) dx = k. \tag{4.31} \]

Now proceeding the similar steps as that of the Theorem 4.6, the remaining part of the theorem can be proved. A direct substitution of (4.5) in (4.26) gives the converse part of the theorem.

**Theorem 4.8:** For a rv \( X \) defined in Theorem 4.7 with a support \( [k, \infty) \), \( k > 0 \), a relation of the form
\[ \bar{M}(t) - (c - 1) \log \bar{G}(t) = K, \]  

(4.32)

where \( K \) is a constant and \( c > 1 \), is satisfied for all \( t > 0 \) if and only if \( X \) follows a Pareto I distribution (4.20).

### 4.4 Renyi’s entropy for past lifetime

As pointed out in Chapter I, Renyi’s entropy measure for the residual life also being a measure of uncertainty of component. Based on the past life of a system, Asadi et.al (2005) defined the Renyi entropy for the past lifetime \( X | X \leq t \) as

\[
\bar{I}_\beta(t) = \frac{1}{(1 - \beta)} \log \int_0^t \frac{f_\beta(x)}{F_\beta(t)} dx. 
\]

(4.33)

As a measure of uncertainty, \( \bar{I}_\beta(t) \) can be used to describe the physical characteristics of the failure mechanism and so characterization theorems using this concept helps one to determine the lifetime distribution through the knowledge of the form of the Renyi entropy for the past life \( \bar{I}_\beta(t) \). Now (4.33) can be rewritten as

\[
(1 - \beta)\bar{I}_\beta(t) = \log \left( \int_0^t f_\beta(x) dx \right) - \beta \log F(t). 
\]

(4.34)

The following Theorem characterizes power distribution using the functional relationship between Renyi’s past entropy and the reversed hazard rate.

**Theorem 4.9**: For a rv \( X \) defined in Theorem 4.1 with Renyi entropy for past life \( \bar{I}_\beta(t) \) is defined in (4.34), then a relationship

\[
\bar{I}_\beta(t) = K - \log \lambda(t), \tag{4.35}
\]

where \( K \) is a constant holds if and only if \( X \) follows a power distribution with cdf (4.5).

**Proof**: Assume that (4.35) holds. Using (4.33) and on simplification, (4.35) implies

\[
\log \int_0^t \frac{f_\beta(x)}{F_\beta(t)} dx = K_1 - \log \left( \lambda(t) \right)^{(1-\beta)}, \text{ where } K_1 = K(1 - \beta).
\]
i.e.
\[ \int_0^1 \frac{f^\beta(x)}{F^\beta(t)} dx = K^* - (\lambda(t))^{(\beta-1)} \]

or
\[ \int_0^1 f^\beta(x) dx = K^* f^{\beta-1}(t) F(t), \text{ where } K^* = \exp(K_t). \quad (4.36) \]

Differentiating (4.36) using the assumption \( \lim_{t \to 0} f^\beta(t) = 0 \) with respect to \( t \), we get
\[ f^\beta(t) = K^* \left( f^\beta(t) + (\beta-1)f^{\beta-2}(t)f'(t)F(t) \right). \quad (4.37) \]

Divide each term of (4.37) by \( f^{\beta-2}(t)F^2(t) \) and on simplification using \( \lambda(t) = \frac{f(t)}{F(t)} - \lambda^2(t) \), we obtain
\[ (1-K^*\beta)\lambda^2(t) - K^*(\beta-1) \frac{d}{dt} \lambda(t) = 0. \quad (4.38) \]

Solving the differential equation (4.38), we obtain the required result. The converse part is obtained by direct calculation.

### 4.5 Weighted models

In this section, we study the usefulness of these uncertainty measures viz past entropy (4.1), conditional measure of uncertainty for the past life (4.25) and the Renyi’s entropy for the past life (4.33) in the context of weighted distributions. The mathematical relationships between the weighted and original variables for (4.1), (4.25) and (4.33) are given by

\[ \tilde{H}^v(t) = 1 - \left[ F(t)\tilde{m}_w(t) \right]^{-1} \int_0^1 w(x)f(x) \log \left( \frac{w(x)\lambda(x)}{\tilde{m}_w(x)} \right) dx, \quad (4.39) \]

\[ \tilde{M}^v(t) = -\left[ F(t)\tilde{m}_w(t) \right]^{-1} \int_0^1 w(x)f(x) \log \left( \frac{w(x)f(x)}{E(w(X))} \right) dx \quad (4.40) \]
and
\[
\tilde{H}^\ast(t) = \frac{1}{(1-\beta)[\tilde{m}_\ast(t)]^\beta} \int \log \left[ E\left( w\beta(X)j'(\beta-1)(X)|X \leq t\right) \right] \, d\tilde{M}(t),
\]
(4.41)

where \( \tilde{H}^\ast(t), \tilde{M}^\ast(t) \) and \( \tilde{I}^\ast(t) \) respectively denote the past entropy, conditional measure of uncertainty for the past life and Renyi’s entropy for the past life for the weighted rv \( X_w \).

**Remark 4.1:** When the weight function \( w(t) = t^\alpha \), the model reduces to size-biased model.

The following theorem characterizes the exponential and log exponential families of distributions using the weighted conditional measure for past life and the weighted form of \( \log G(t) \).

**Theorem 4.10:** Let \( X_w \) be a weighted rv with weight function \( w(t) \) and \( \log G''(t) = E(\log X_w | X_w \leq t) \). Assume \( \lim_{t \to 0} \frac{\log V(t) f(t)}{t} = 0 \), then a relationship
\[
\tilde{M}^\ast(t) = \log U(\theta) - \theta \log \tilde{G}^\ast(t) - E\left( \log V(X_w) | X_w \leq t \right),
\]
(4.42)

where \( U \) and \( V \) are any functions of \( \theta \) and \( x_w \) respectively, satisfies if and only if the pdf of \( X \) belongs to one parameter log exponential family (1.29).

**Proof:** For the one-parameter log exponential family (1.29), we have
\[
f''(t) = \frac{w(t)t^\theta C(t)}{A(\theta)E\left(w(X)\right)} = \frac{t^\theta C^\ast(t)}{A'(\theta)}, \quad \text{where } C^\ast(t) = w(t)C(t) \text{ and } A'(\theta) = A(\theta)E\left(w(X)\right).
\]

Now proceeding the similar steps as that of the Theorem 4.3, we obtain the result.

**Theorem 4.11:** For a rv considered in Theorem 4.10 and assume that \( \lim_{t \to 0} B(t) f(t) = 0 \). A relation of the form
\[
\tilde{M}^\ast(t) + \theta \tilde{m}^\ast(t) = -A(\theta) - E\left(B(X_w) | X_w \leq t \right),
\]
(4.43)
where \( A \) and \( B \) are any functions of \( \theta \) and \( x_w \), is satisfied for all \( t \geq 0 \) if and only if the pdf of \( X \) belongs to one parameter exponential family (1.28).

Proof: For the weighted rv \( X_w, f''(t) = \exp \left( \theta t + C'(t) + D'(\theta) \right) \), belongs to the exponential family with \( C'(t) = \left( \log w(t) \right) + C(t) \) and \( D'(\theta) = D(\theta) - \log E \left( w(X) \right) \). Rest of the proof is similar to that of Theorem 4.3.

### 4.6 Some new classes of distributions

Recently Di Crescenzo and Longobardi (2002) observed that even if the Shannon's entropy of two components with lifetimes \( X \) and \( Y \) are same, the expected uncertainty contained in the conditional density of \( X \) given \( X \leq t \) (i.e., past entropy of \( X \)) is different from that contained in the conditional density of \( Y \) given \( Y \leq t \) (i.e., past entropy of \( Y \)). Motivated from this, Nanda and Paul (2006) defined the following ordering based on past entropy.

**Definition 4.1** (Nanda and Paul (2006)) Let \( X \) and \( Y \) be two random variables denoting the lifetimes of two components. Then \( X \) is said to be greater than \( Y \) in past entropy order (written as \( X \preceq PE Y \)) if \( \bar{H}_X(t) \leq \bar{H}_Y(t) \) for all \( t > 0 \).

**Definition 4.2** (Nanda and Paul (2006)): A rv \( X \) is said to have increasing (decreasing) uncertainty of life (or increasing (decreasing) past entropy) if \( \bar{H}(t) \) is increasing (decreasing) in \( t \geq 0 \).

**Theorem 4.12**: Let \( X_w \) be a weighted rv with the weight function \( w(t) \) if (a) \( \frac{w(t)}{\bar{m}_w(t)} \) is decreasing and (b) \( X \) is DRHR, then \( X_w \) has increasing past entropy (IPE).

**Proof**: Using (2.9), and from the conditions (a) and (b) implies \( \lambda''(t) \) is DRHR. Now using Theorem 3.1 in Nanda and Paul (2006) (i.e., If \( X \) is DRHR then \( X \in IPE \)), we get \( X_w \) is IPE. Nanda and Paul (2006) have shown with an example that IPE property does
not imply DRHR property. Using this argument the converse of the Theorem 3.1 does not hold.

Theorem 4.12 can be illustrated by the following example.

Example 4.1: Let $X$ be a non-negative rv following a power distribution with df (4.5) and weight function $w(t) = t$. Then this rv satisfies the conditions given in Theorem 4.14 and hence the past entropy of its weighted version is increasing. Figure 4.1 shows that the increasing nature of past entropy of weighted version of power distribution for $c = 2$ and $t \in (0,100)$.

![Figure 4.1: Plot of $H^*(t)$ against $t \in (0,100)$ when $w(t) = t$ and $c = 2$](image)

Theorem 4.13: (i) If $w(t)$ is increasing and $w(t)\lambda(t)$ is decreasing then $X_w$ is IPE.

(ii) If $w(t)h(t)$ is decreasing then $X_w$ is IPE.

Proof: Using the theorems given in Bartoszewicz and Skolimowska (2006) (see Chapter 1) and Nanda and Paul (2006, Theorem 3.1), we can prove (i) and (ii).

Example 4.2: Consider a rv having exponential distribution with weight function $w(t) = t$. $X$ satisfies the conditions of the Theorem 4.15 (i) and hence the past entropy of its length-biased rv is increasing.

Theorem 4.14: If $w(t) \leq (\geq) E\left(w(X)\right) \leq (\geq) E\left(w(X) \mid X \leq t\right)$, then $X \leq (\geq) X_w$.

Proof: Assume that

$$w(t) \leq E\left(w(X)\right) \leq E\left(w(X) \mid X \leq t\right),$$

(4.44)
Some measures of uncertainty in past lifetime

holds. Then we have

\[ \frac{w(t)}{E\left( w(X) \mid X \leq t \right)} \leq 1 \]
(4.45)

\[ \frac{E\left( w(X) \mid X \leq t \right)}{E\left( w(X) \right)} \geq 1 \]
(4.46)

and

\[ \frac{w(t)}{E\left( w(X) \right)} \leq 1. \]
(4.47)

Multiplying (4.45), (4.46) and (4.47) by \( \lambda(t), F(t) \) and \( f(t) \) respectively and using (2.9), (2.8) and (1.1), we get

\[ \lambda^*(t) \leq \lambda(t) \]
(4.48)

\[ F^*(t) \geq F(t) \]
(4.49)

and

\[ f^*(t) \leq f(t). \]
(4.50)

Substituting (4.48), (4.49) and (4.50) in (4.1) we get \( \tilde{H}^*(t) \geq \tilde{H}(t) \).

The following result is direct from the definitions 4.1 and 4.2.

**Result 4.1:** If \( X \geq X^*_w \) and \( X \geq X^*_w \) then

(i) \( X \) is IPE implies \( X^*_w \) is IPE.

(ii) \( X^*_w \) is DPE implies \( X \) is DPE.

In connection with the ordering based on past entropy, we define the following order based on the conditional measure of uncertainty.

**Definition 4.3:** Let \( X \) and \( Y \) be two random variables denoting the lifetimes of two components, then \( X \) is said to be greater than \( Y \) in conditional measure of uncertainty life order (written as \( X \geq CMUL Y \)) if \( \bar{M}_X(t) \leq \bar{M}_Y(t) \) for all \( t > 0 \).
**Definition 4.4:** A df $F(t)$ is said to have increasing (decreasing) conditional measure of uncertainty life if $\bar{M}(t)$ is increasing (decreasing) in $t \geq 0$.

**Theorem 4.15:** (1) If $F(t)$ is ICMUL, then $X$ has IPE. The converse is not true always.

**Proof:** The first part is direct from (4.26) and to prove the converse consider the following example

**Example 4.3:** Let $X$ be a non-negative rv having df

$$F(t) = \begin{cases} 
\frac{t^2}{2}; & 0 \leq t < 1 \\
\frac{t^2 + 2}{6}; & 1 \leq t < 2 \\
1; & t \geq 2
\end{cases}$$

For this distribution, the past entropy $\bar{H}(t)$ and $\bar{M}(t)$ are given by

$$\bar{H}(t) = \begin{cases} 
\log \left( \frac{t^2}{2} \right) + \frac{1}{2}; & 0 \leq t \leq 1 \\
\log \left( \frac{t^2 + 2}{6} \right) + \left( \frac{t^2 - 1}{t^2 + 2} \right) \log 3 - \left( \frac{t^2}{t^2 + 2} \right) \log t + \frac{1}{2}; & 1 \leq t \leq 2 \\
\frac{1}{2} \log 3 - \frac{2}{3} \log 2 + \frac{1}{2}; & t \geq 2
\end{cases}$$

and

$$\bar{M}(t) = \begin{cases} 
\frac{1}{2} \log t; & 0 \leq t \leq 1 \\
\left( \frac{t^2 - 1}{t^2 + 2} \right) \log 3 - \left( \frac{t^2}{t^2 + 2} \right) \log t + \frac{1}{2}; & 1 \leq t \leq 2 \\
\frac{1}{2} \log 3 - \frac{2}{3} \log 2 + \frac{1}{2}; & t \geq 2
\end{cases}$$

For this distribution, $\bar{H}(t)$ is increasing in $t$ (see Nanda and Paul (2006)), but $\bar{M}(t)$ is not increasing in $t \in [0,1]$ as shown in figure 4.2.
Theorem 4.16: If \( w(t) \leq (\geq) E\left( w(X)|X \leq t \right) \), then \( X \geq (\leq) X_w \).

Proof: Proof is similar to that of Theorem 4.16.

Similarly we can define the following order based on Renyi's entropy.

Definition 4.5: A df \( F(.) \) is said to be increasing (decreasing) Renyi's past entropy if \( \overline{I}_\beta(t) \) is increasing (decreasing) for all \( t \geq 0 \).

Theorem 4.17: If \( f(t) \) is increasing (decreasing) in \( t \), then

\[
\overline{I}_\beta(t) \leq (\geq) -\log \lambda(t) \text{ for all } \beta .
\] (4.51)

Proof: When \( f(t) \) is increasing, then \( f(x) \leq f(t) \) for all \( x \leq t \). Using (4.34), we have

\[
\overline{I}_\beta(t) \leq (1-\beta)^{-1} \log \int_{\beta}^{1} \frac{f^{(\beta-1)}(t)}{F^{(\beta-1)}(t)} \frac{f(x)}{F(t)} \, dx , \text{ for all } x \leq t .
\] (4.52)

On simplification, (4.52) implies (4.51). In a similar manner when \( f(t) \) is decreasing, then the inequality is reversed.

Theorem 4.18: When \( f(t) \) and \( w(t) \) are increasing (decreasing), then

\[
\overline{I}_\beta(t) \leq (\geq) \overline{I}_{\beta-}(t)
\]
**Proof:** Assume that \((1-\beta)(\bar{I}_a(t) - \bar{I}_{k_a}(t)) = \log \left( \frac{A}{B} \right)\), where \(A = \int_0^t (f(x))^\beta \, dx\) and \(B = \int_0^t \left( \frac{f''(x)}{F''(t)} \right)^\beta \, dx\).

When \(w(t)\) and \(f(t)\) are increasing then \(f''(t) \geq f(t)\), which implies

\[
\log \left( \frac{A}{B} \right) \leq \log \left[ \frac{\int_0^t (f(x))^\beta \, dx}{\int_0^t (f(x))^\beta \, dx} \right] \leq \log \left( \frac{F''(t)}{F(t)} \right)^\beta \quad \text{(4.53)}
\]

From (4.53) we get

\[
\log \left( \frac{A}{B} \right) \leq \log \left( \frac{F''(t)}{F(t)} \right)^\beta = \beta \log \left( \frac{F''(t)}{F(t)} \right) \quad \text{(4.54)}
\]

But when \(w(t)\) is increasing (decreasing), then \(F''(t) \leq (\geq) F(t)\) (Sunoj and Maya (2006)). Therefore from (4.54), we get \(\beta \log \left( \frac{F''(t)}{F(t)} \right) \leq \beta \log 1 = 0\). This implies that

\((1-\alpha)(\bar{I}_a(t) - \bar{I}_{k_a}(t)) \leq 0, \) therefore \(\bar{I}_a(t) \leq \bar{I}_{k_a}(t)\).

The following tables give the measures of uncertainty for various distributions.

*Table: 4.1 \(\bar{H}(t)\) for various distributions*

<table>
<thead>
<tr>
<th>Distribution</th>
<th>pdf</th>
<th>(\bar{H}(t))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exponential</strong></td>
<td>(\lambda \exp(-\lambda t)); (t, \lambda &gt; 0)</td>
<td>(1 - \log \left( \frac{\lambda}{1 - \exp(-\lambda t)} \right) - \left( \frac{\lambda t \exp(-\lambda t)}{1 - \exp(-\lambda t)} \right))</td>
</tr>
<tr>
<td><strong>Pareto I</strong></td>
<td>(ck^{\alpha}t^{-(\alpha+1)}); (t \geq k, k, c &gt; 0)</td>
<td>(\frac{c+1}{c} - \log \left( \frac{ck^\alpha}{1 - t_k^\alpha} \right) + \left( \frac{(c+1)t_k^\alpha}{1 - t_k^\alpha} \right))</td>
</tr>
<tr>
<td>Distribution</td>
<td>PDF</td>
<td>( M(t) )</td>
</tr>
<tr>
<td>--------------</td>
<td>-----</td>
<td>-------------</td>
</tr>
<tr>
<td>Pareto II</td>
<td>( pq^p (t+q)^{-(p+1)} ); ( t &gt; 0, q, p &gt; 0 )</td>
<td>( \frac{p+1}{p} \log \left( \frac{pq^p}{1-(\frac{q}{t+q})^p} \right) - \frac{(a+1)}{1-(\frac{c}{t+q})^a} \log \left( \frac{q}{1-(\frac{q}{t+q})^p} \right) \log(t+q) )</td>
</tr>
<tr>
<td>Beta</td>
<td>( dR^{-d}(R-t)^{(d-1)} ); ( 0 &lt; t &lt; R, d &gt; 0 )</td>
<td>( \frac{1-d}{d} \log\left( \frac{dR^{-d}}{1-(\frac{R}{R-t})^{-d}} \right) - \frac{c-1}{1-(\frac{R}{R-t})^c} \log \left( \frac{R}{1-(\frac{R}{R-t})^{-d}} \right) \log(R-t) )</td>
</tr>
<tr>
<td>Power</td>
<td>( \frac{ct^{(c-1)}}{b^c} ); ( 0 \leq t \leq b, b, c &gt; 0 )</td>
<td>( \frac{c-1}{c} \log\left( \frac{c}{t} \right) )</td>
</tr>
</tbody>
</table>

**Table: 4.2** \( \tilde{M}(t) \) and \((1-\beta)I_\kappa(\beta,t)\) for various distributions

<table>
<thead>
<tr>
<th>Pdf</th>
<th>( \tilde{M}(t) )</th>
<th>((1-\beta)I_\kappa(\beta,t))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda \exp(-\lambda t) ; ) ( t, \lambda &gt; 0 )</td>
<td>( 1 - \log \lambda - \left( \frac{\lambda t \exp(-\lambda t)}{1 - \exp(-\lambda t)} \right) )</td>
<td>( (\beta-1) \log \lambda - \log \beta + \frac{1 - \exp(-\lambda \beta t)}{(1 - \exp(-\lambda t))^\beta} )</td>
</tr>
<tr>
<td>( ck^t t^{(c+1)} ; ) ( t \geq k, k, c &gt; 0 )</td>
<td>( \left( \frac{c+1}{c} \right) - \log \left( \frac{ck^t}{1 - (\frac{k}{t})^c} \right) + \left( \frac{(c+1)(\frac{k}{t})^c}{1 - (\frac{k}{t})^c} \right) \log \left( \frac{k}{t} \right) )</td>
<td>( \log \left( c^{\beta} k^{(1-\beta)} \right) - \log \left( (c+1) \beta-1 \right) + \log \left( \frac{1 - (\frac{k}{t})^{(c+1)-1}}{(1 - (\frac{k}{t})^c)^\beta} \right) )</td>
</tr>
<tr>
<td>( pq^p (t+q)^{-(p+1)} ; ) ( t &gt; 0, q, p &gt; 0 )</td>
<td>( \frac{p+1}{p} \log \left( \frac{pq^p}{1-(\frac{q}{t+q})^p} \right) - \frac{(a+1)}{1-(\frac{c}{t+q})^a} \log \left( \frac{q}{1-(\frac{q}{t+q})^p} \right) \log(t+q) )</td>
<td>( \log \left( p^{\beta} q^{(1-\beta)} \right) - \log \left( (p+1) \beta \right) + \log \left( \frac{1 - (\frac{q}{t+q})^{(p+1)-1}}{(1 - (\frac{q}{t+q})^p)^\beta} \right) )</td>
</tr>
<tr>
<td>( dR^{-d}(R-t)^{(d-1)} ; ) ( 0 &lt; t &lt; R, d &gt; 0 )</td>
<td>( \frac{1-d}{d} \log\left( \frac{dR^{-d}}{1-(\frac{R}{R-t})^{-d}} \right) - \frac{d-1}{1-(\frac{R}{R-t})^{-d}} \log \left( \frac{R}{1-(\frac{R}{R-t})^{-d}} \right) \log(R-t) )</td>
<td>( \log \left( d^{\beta} R^{(1-\beta)} \right) + \log \left( (d-1) \beta+1 \right) + \log \left( \frac{1 - (\frac{R}{R-t})^{(d-1)+1}}{(1 - (\frac{R}{R-t})^{-d})^\beta} \right) )</td>
</tr>
<tr>
<td>( \frac{ct^{(c-1)}}{b^c} ; ) ( 0 \leq t \leq b, b, c &gt; 0 )</td>
<td>( \frac{c-1}{c} \log c + c \log b - (c-1) \log t )</td>
<td>( \beta \log c - \log((c-1) \beta+1) + (1-\beta) \log t )</td>
</tr>
</tbody>
</table>
4.7 Measures of discrimination

In this section, we discuss some measures of discrimination proposed by Di Crescenzo and Longobardi (2004) and Asadi et al. (2005). Further we derive the bounds and inequalities for the comparison of weighted distributions and their unweighted counterparts using these measures.

Let $X$ and $Y$ be two non-negative random variables admitting absolutely continuous distribution functions $F(t)$ and $G(t)$ respectively, then Kullback and Leibler (1951) extensively studied the concept of directed divergence which gives the discriminations between two populations and it is defined as

$$I(X,Y) = I(F,G) = \int_{0}^{\infty} f(x) \log \frac{f(x)}{g(x)} dx$$

where $f(t)$ and $g(t)$ are the corresponding distribution functions of $X$ and $Y$ respectively. Motivated by this, Ebrahimi and Kirmani (1996) modified (4.55) useful to measure the discrimination between two residual lifetime distributions and is given by

$$I_{x,y}(t) = \int_{0}^{\infty} f(x) \log \left( \frac{f(x) R(t)}{g(x) S(t)} \right) dx ; \quad t > 0 ,$$

where $R(t) = 1 - F(t)$ and $S(t) = 1 - G(t)$. $I_{x,y}(t)$ measures the relative entropy of $(X \mid X > t)$ and $(Y \mid Y > t)$ and it is useful for comparing the residual lifetimes of two items, which have both survived up to time $t$. Along the similar lines of the measure (4.56), Di Crescenzo and Longobardi (2004) defined the information distance between the past lives $(X \mid X \leq t)$ and $(Y \mid Y \leq t)$ as

$$\overline{I}_{x,y}(t) = \int_{0}^{t} f(x) \log \left( \frac{f(x) F(t)}{g(x) G(t)} \right) dx ; \quad t > 0 .$$

Given that at time $t$, two items have been found to be failing, $\overline{I}_{x,y}(t)$ measures the discrepancy between their past lives. Similarly, Renyi divergence between the residual distributions proposed by Asadi et al. (2005) is given by
\[ I_{X,Y}(\beta,t) = \frac{1}{(\beta-1)} \log \int_{x}^{\infty} R^\beta(t) S^{(1-\beta)}(t) \, dx. \] (4.58)

In a similar way, Asadi et al. proposed the Renyi discrimination for the past lives implied by \( F \) and \( G \) as

\[ \overline{I}_{X,Y}(\beta,t) = \frac{1}{(\beta-1)} \log \int_{0}^{f(x)} \left( \frac{f(x)}{g(x)} \right)^{\beta} \left( \frac{F(t)}{G(t)} \right)^{1-\beta} \, dx. \] (4.59)

In view of the wide applicability of the discrimination measures (4.57) and (4.59) in past lifetime, in the present section, we investigate its relationships between original and weighted random variables and prove certain results.

Now for the random variables \( X \) and \( X' \), the measure (4.57) is defined as

\[ \overline{I}_{X,X'}(t) = \int_{0}^{f(x)} \log \left( \frac{f(x)}{F(t)} \right) \, dx. \] (4.60)

The measure (4.60) gives the measure of discrepancy between original and weighted TV and it directly related to past entropy \( H(t) \) through a relation

\[ \overline{I}_{X,X'}(t) = \int_{0}^{f(x)} \log \left( \frac{f(x)}{F(t)} \right) \, dx - H(t). \] (4.61)

Using (1.1) and (2.8) in (4.60), we obtain

\[ \overline{I}_{X,X'}(t) = \log \left[ E \left( w(X) \mid X \leq t \right) \right] - E \left( \log w(X) \mid X \leq t \right). \] (4.62)

For a size-biased model, (4.62) reduces to the form

\[ \overline{I}_{X,X'}(t) = \log \overline{m}^a(t) - \alpha \left( \log \overline{G}(t) \right) \] (4.63)

where \( \overline{m}^a(t) = E \left( X^a \mid X \leq t \right) \) and \( \log \overline{G}(t) = E \left( \log X \mid X \leq t \right) \). When \( \alpha = 1 \), (4.63) reduces to the discrimination measure between the original and length-biased model.
**Theorem 4.19:** If $I_{X,X_w}(t)$ is independent of $t$ if and only if the weight function takes the form $w(t) = (F(t))^{\theta-1}; \theta > 0$.

**Proof:** Suppose that $I_{X,X_w}(t)$ is independent of $t$.

i.e. $I_{X,X_w}(t) = K$, \hspace{1cm} (4.64)

where $K$ is independent of $t$. Using (4.60) and differentiating (4.64) with respect to $t$, we get

$$\lambda(t) \left[ \log \left( \frac{\lambda(t)}{\lambda_w(t)} \right) - k \right] + \lambda_w(t) - \lambda(t) = 0.$$ \hspace{1cm} (4.65)

Divide each term of (4.65) by $\lambda(t)$ yields

$$\left[ \log \frac{\lambda(t)}{\lambda_w(t)} - k \right] \frac{\lambda_w(t)}{\lambda(t)} - 1 = 0.$$ \hspace{1cm} (4.66)

Substitute $u(t) = \frac{\lambda(t)}{\lambda_w(t)}$ and on differentiating (4.66), we get

$$\frac{u'(t)}{u(t)} \left( 1 - \frac{1}{u(t)} \right) = 0$$ \hspace{1cm} (4.67)

which implies that either $u'(t) = 0$ or $u(t) = 1$. But as $X$ and $X_w$ are not equal $u(t) \neq 1$.

So $u'(t) = 0$. Hence we have $u'(t) = 0$, which implies that there exists a non-negative constant $\theta$ such that $\lambda_w(t) = \theta \lambda(t)$. Now using (2.9) we get $w(t) = (F(t))^{\theta-1}; \theta > 0$.

Conversely assuming $w(t) = (F(t))^{\theta-1}$ and using (2.8), we obtain

$$F^w(t) = (F(t))^{\theta}.$$ \hspace{1cm} (4.68)

From (4.68) and (4.60) we get that for $t > 0$ $I_{X,X_w}(t) = \theta - 1 - \log \theta$, which is independent of $t$.
**Corollary 4.1:** When $F(t) = t$, then Theorem 4.19 characterizes power distribution with df (4.5).

The discrimination measure (4.59) proposed by Asadi et al. (2005) for the random variables $X$ and $X_{\alpha}$ is defined as

$$
\bar{I}_{x,x_{\alpha}}(\beta, t) = \frac{1}{(\beta - 1)} \log \int_{0}^{t} \left( \frac{f(x)}{F(t)} \right)^{\beta} \left( \frac{f^*(x)}{F^*(t)} \right)^{1-\beta} dx
$$  \hspace{1cm} (4.69)

Using (1.1) and (2.8), (4.69) becomes

$$
\bar{I}_{x,x_{\alpha}}(\beta, t) = \log \left[ E \left( w(X) \mid X \leq t \right) \right] + \frac{1}{(\beta - 1)} \log \left[ E \left( w^{-\beta}(X) \mid X \leq t \right) \right]. \hspace{1cm} (4.70)
$$

For the size-biased model, (4.70) becomes

$$
\bar{I}_{x,x_{\alpha}}(\beta, t) = \log \bar{m}^{\alpha}(t) + \frac{1}{(\beta - 1)} \log \bar{m}^{\alpha(1-\beta)}(t). \hspace{1cm} (4.71)
$$

**Remark 4.2:** When $\beta = 0$, then (4.71) reduces to the measure (4.60).

**Theorem 4.20:** The R{\`e}nyi divergence measure for the past life $\bar{I}_{x,x_{\alpha}}(\beta, t)$ is independent of $t$ if and only if the weight function is $w(t) = (F(t))^{\theta-1}$; $\theta > 0$.

**Proof:** The proof of this theorem is similar to that of Theorem 4.19.

**Corollary 4.2:** When $F(t) = t$, then Theorem 4.20 characterizes power distribution with df (4.5).

### 4.8 Inequalities for measures of discrimination

In this section, we present some results including inequalities and comparisons of discrimination measures for weighted and unweighted or parent distributions. Under some mild constraints, bounds for these measures are also presented here.
Theorem 4.21: If the weight function \( w(t) \) is increasing (decreasing) in \( t > 0 \), then

(a) \[ \overline{I}_{X,X_{w}}(t) \geq (\leq) \log \left( \frac{\lambda(t)}{\lambda^{n}(t)} \right) \]

(b) \[ \overline{I}_{X,X_{w}}(\beta,t) \geq (\leq) \frac{\beta}{(\beta - 1)} \log \left( \frac{\lambda(t)}{\lambda^{n}(t)} \right), \quad \beta \neq 1. \]

Proof: Suppose \( w(t) \) is increasing, then from (1.1) we get \( \frac{f(t)}{f^{n}(t)} \) is decreasing which implies

\[ \frac{f(t)}{f^{n}(t)} \leq \frac{f(x)}{f^{n}(x)} \text{ for all } x \leq t. \]  \hspace{1cm} (4.72)

Now from (4.60) we have

\[ \int_{0}^{t} f(x) \log \left( \frac{f(x)}{f^{n}(x)} \right) dx \geq \int_{0}^{t} f(x) \log \left( \frac{f(t)}{f^{n}(t)} \right) dx \]

which implies that

\[ \overline{I}_{X,X_{w}}(t) \geq \log \left( \frac{\lambda(t)}{\lambda^{n}(t)} \right) \text{ for all } x \leq t. \]

When \( w(t) \) is decreasing then the inequality is reversed.

Proof of (b) is similar to that of (a).

Theorem 4.22: When (i) \( w(t) \) is decreasing (increasing) and (ii) \( X \leq_R \left( \geq_R \right) X_{w} \), then \( \overline{I}_{X,X_{w}}(t) \) is increasing (decreasing) for all \( t > 0 \).

Proof: From the definition (4.60)

\[ \overline{I}_{X,X_{w}}(t) = \log \left( \frac{F^{n}(t)}{F(t)} \right) + \int_{0}^{t} f(x) \log \left( \frac{f(x)}{f^{n}(x)} \right) dx. \]  \hspace{1cm} (4.73)
The first term of (4.73) is increasing using Theorem 2 (see Sunoj and Maya (2006)) and

$$\int_0^T \frac{dx}{F(t)} \log \left( \frac{f(x)}{f''(x)} \right) = \log \mu - \frac{1}{F(t)} \int_0^T f(x) \log w(x) \, dx. \quad (4.74)$$

Now the second term of (4.74) is given by

$$\frac{1}{F(t)} \int_0^T f(x) \log w(x) \, dx = -\log(w(t)) + \frac{1}{w(x)} \int_0^T \frac{w'(x)}{F(t)} F(x) \, dx. \quad (4.75)$$

Differentiating (4.75) with respect to \( t \) and on simplification we get

$$- \frac{d}{dt} \left( \int_0^T \frac{dx}{F(t)} f(x) \log w(x) \right) = -\frac{\lambda(t)}{F(t)} \int_0^T \frac{w'(x)}{w(x)} F(x) \, dx \geq 0. \quad (4.76)$$

Thus (4.73) is the sum of two increasing functions. It implies that \( I_{x,x_s}(t) \) also increasing. Using the similar steps as above, the inequality in the reverse direction can be proved.

**Theorem 4.23:** When \( w(t) \) is increasing (decreasing) and \( \frac{E(X|X \leq t)}{w(t)} \) is increasing (decreasing), then \( I_{x,x_s}(t) \) is increasing (decreasing) for all \( t > 0 \).

**Proof:** When \( w(t) \) is increasing, from Theorem 4.21,

$$I_{x,x_s}(t) \geq \log \left( \frac{\lambda(t)}{\lambda''(t)} \right).$$

Now using (2.9) and the condition given in theorem, we get \( \log \left( \frac{\lambda(t)}{\lambda''(t)} \right) \) increases, which imply the required result. Similarly one can prove the inequality in the reverse direction.