Chapter 3

MEASURE OF INACCURACY

3.1 Introduction

The identification of an appropriate probability distribution for lifetimes is one of the basic problems encountered in reliability theory. Although several methods such as probability plots, goodness of fit procedures etc are available in literature to find an appropriate model followed by the observations, they fail to provide an exact model. A method to attain this goal is to utilize a suitable characteristic property of the model. A property $P$ is said to be characteristic to a distribution if $P$ holds under the distributional assumption and the only distribution for which $P$ holds is the underlying distribution. Thus a characterization theorem enables one to uniquely determine the distribution. Most of the work on characterization of distributions in the reliability context centers around the hazard rate or the mean residual life function. In a variant approach, Ebrahimi (1996) proposed the residual entropy function as a useful tool to analyze the stability of a component/system. Following this several papers appeared employing information measures like time dependent Kullback-Leibler directed distance and their generalizations in characterizing life distributions. As pointed out in the chapter two, the inaccuracy measure can be thought of as a generalization of Shannon’s entropy. Therefore, there is a scope for extending the results based on Shannon’s entropy and its modifications as applicable for the inaccuracy measure to suit the context of reliability. Motivated by this, in the present chapter we look into the problem of characterization of probability distributions using truncated versions of the inaccuracy measure.

Recently, Nair and Gupta (2007) have extended the inaccuracy measure defined by equation (2.52) to the truncated situation as given in equation (2.54). For the random variable, $X - t | X > t$ the truncated inaccuracy measure has the form
Measure of inaccuracy

\[ I(F,G; t) = \int \frac{f(x)}{F(t)} \log \left( \frac{g(x)}{G(t)} \right) dx. \]

It may be observed that \( I(F,G; t) \) considered above can be decomposed as

\[ I(F,G; t) = H(F; t) + D(F,G; t), \quad (3.1) \]

where \( H(F; t) \) is the residual entropy function defined in equation (2.33) and \( D(F,G; t) \) is the modified Kullback-Leibler divergence measure considered in equation (2.48). If \( F(x) \) is the actual distribution corresponding to the observations and \( G(x) \) is the distribution function assigned by the experimenter, equation (3.1) asserts that the residual inaccuracy measure \( I(F,G; t) \) is the sum of the residual entropy function of \( F(x) \), which measures uncertainty, and a measure of discrimination between \( F(x) \) and \( G(x) \). For convenience in the sequel, we denote \( I(F,G; t) \) by \( I(t) \).

3.2 Characterization of probability distributions using the functional form of inaccuracy measure

In this section we consider the problem of characterizing distribution functions \( F(x) \) and \( G(x) \) based on given functional forms for the inaccuracy measure \( I(t) \).

**Theorem: 3.1**

Let \( X \) and \( Y \) be two non-negative continuous random variables with distribution functions \( F(x) \) and \( G(x) \) respectively and \( I(t) \) be as defined in equation (2.54). Further assume that \( I(t) \) is independent of \( t \) for all \( t > 0 \). Then \( F(x) \) is exponential if and only if \( G(x) \) is exponential.

**Proof**

Let \( I(t) = c \), where \( c \) is a positive constant.
Further assume that $F(x)$ is exponential with survival function

$$\overline{F}(x) = e^{-\alpha x}, \alpha > 0, x \geq 0.$$  \hspace{1cm} (3.2)

Using the relation (2.56) namely

$$h_2(t) = \frac{I(t) + h_2(t)}{I(t) + \log h_2(t)}$$  \hspace{1cm} (3.3)

we get

$$\alpha (c + \log h_2(t)) = h_2(t).$$

Differentiating the above equation with respect to $t$, we get

$$h_2'(t) \left( \alpha (h_2(t))^{-1} - 1 \right) = 0.$$

This gives either $h_2'(t) = 0$ or $h_2(t) = \alpha$. In either case $h_2(t) = \alpha$ is a constant. Since the constancy of hazard rate is characteristic to the exponential distribution, one can conclude that $G(x)$ is also exponential.

Conversely assume $\overline{G}(x) = e^{-\beta x}, \beta > 0$.

From equation (3.3)

$$h_2(x) = \frac{\beta}{c + \log \beta}.$$

Using equation (2.6) we get

$$\overline{F}(x) = \exp \left( \frac{-\beta x}{c + \log \beta} \right).$$

This shows that $F(x)$ is exponential.
Remark: 3.1

If $\overline{F}(x)$ and $\overline{G}(x)$ are the survival functions of two random variables following the exponential distribution, it is immediate that $\overline{G}(x) = [\overline{F}(x)]^\theta$ for some $\theta$. In other words $\overline{G}(x)$ is the proportional hazards model of $\overline{F}(x)$ and in this situation $I(t)$ is constant. However, it is not necessary that $\overline{G}(x)$ is the proportional hazards model of $\overline{F}(x)$ or $F(x)$ is exponential for $I(t)$ to be a constant, as the next result shows.

Theorem: 3.2

For the random variables $X$ and $Y$ considered in Theorem 3.1, assume that $I(t)$, defined in equation (2.54), is independent of $t$ for all $t > 0$. Then $F$ has the finite range distribution specified by the survival function

$$\overline{F}(x) = \left(\frac{c + \log \alpha - \log x}{c + \log \alpha - \log k}\right)^\alpha \frac{\alpha}{k} e^x < X < \alpha e^x,$$

if and only if $G$ has the Pareto distribution with survival function

$$\overline{G}(x) = \left(\frac{k}{x}\right)^\alpha, \; x > k > 0, \; \alpha > 0.$$

The proof of the theorem is analogous to that of Theorem 3.1 and hence omitted.

Remark: 3.2

The difference between the models used in Theorem 3.1 and Theorem 3.2, both giving constant inaccuracy, is that whereas in Theorem: 3.1 the decomposition (3.1) yields constant values for $H(F; t)$ and $D(F, G; t)$ while these measures are functions of $x$ in Theorem:3.2.
A random variable $X$ has the Generalized Pareto Distribution (GPD) if its survival function has the form [Lai and Xie (2006)]

$$
\bar{F}(x) = \left(1 + \frac{ax}{b}\right)^{-\left(\frac{1}{a}\right)}, \ x > 0, a > -1, b > 0.
$$

(3.4)

The importance of this distribution in reliability modeling lies in the fact that it has a linear mean residual life in the form $m(x) = b + ax$. Further the family is rich in the sense that it contains the Lomax distribution ($a > 0$), rescaled beta ($-1 < a < 0$), the exponential ($a \to 0$) and the uniform distribution. Hall and Wellner (1981) have established that the Generalized Pareto Distribution is the only family of distributions that has linear mean residual life function.

The next theorem provides a characterization result for the family of distributions specified in equation (3.4) based on a functional form for the inaccuracy measure.

**Theorem: 3.3**

Let $X$ and $Y$ be two non-negative continuous random variables with distribution functions $F(x)$ and $G(x)$ respectively and $I(t)$ denote the truncated inaccuracy measure. Assume that $I(t)$ is a linear function of $t$. Then $X$ follow the generalized Pareto distribution if and only if $Y$ follow the exponential distribution.

**Proof**

Assume that $I(t) = a + bt, b \neq 0$ and that $Y$ follow the exponential distribution with parameter $\alpha$.

Using equation (3.3), we get

$$
h_i(t) = \frac{b + \alpha}{a + bt + \log \alpha}.
$$

Using equation (2.6) namely
\[ F(x) = \exp \left( -\int_0^x h(t) \, dt \right), \]

we get

\[ F(x) = \left( 1 + \frac{bx}{a + \log \alpha} \right)^{-\left(1 + \frac{\alpha}{b}\right)}, \]

where \( A = a + \log \alpha \).

(3.5)

Hence \( X \) follows the generalized Pareto distribution.

Conversely let \( X \) follow the generalized Pareto distribution with survival function (3.4). Then by direct computations we get

\[ h_1(x) = \frac{b + \alpha}{A + bx}. \]

Equation (3.3) now becomes

\[ \frac{b + \alpha}{A + bt} = \frac{b + h_z(t)}{a + bt + \log h_z(t)} \]

or

\[ \frac{a + bt + \log \alpha}{b + \alpha} = \frac{a + bt + \log h_z(t)}{b + h_z(t)}. \]

(3.6)

Since the last equation holds for all \( t > 0 \) and the left side is linear in \( t \), the right side also must be linear. Thich is possible only when \( h_z(t) = \) a constant. Thus \( G \) is exponential.

**Note:**

When \( G \) follows the exponential distribution with parameter \( \alpha \) and \( F \) follows the exponential distribution with parameter \( \frac{\alpha - 1}{\log \alpha} \) then

\[ I(t) = m(t) = \frac{\log \alpha}{\alpha - 1}, \]

where \( m(t) \) is the mean residual life function of \( F \).
Regarding the inaccuracy measures for past life, defined in (2.55), one can obtain similar results. In the next theorem we look into the form of $F(x)$ and $G(x)$ when the truncated inaccuracy measure for past life, $I^*(t)$ is independent of $t$.

**Theorem: 3.4**

For the random variables $X$ and $Y$ considered in Theorem 3.3, assume that $I^*(t)$, defined in equation (2.55) is independent of $t$ for all $t > 0$. Then $Y$ follows the power distribution specified by

$$G(x) = x^c, \quad 0 < x < 1, \quad c > 0,$$

if and only if $X$ has a finite range distribution given by

$$F(x) = \left(1 - \frac{\log x}{\log c + k}\right)^c, \quad 0 < x < 1. \quad (3.7)$$

**Proof**

Let $I^*(t) = k$, where $k$ is a constant and that $Y$ be distributed as in equation (3.7). From equation (2.57), we get

$$\lambda_t(x) = c \left[x(\log c - \log x + k)\right]^{-1}. \quad (3.8)$$

This gives

$$F(x) = \exp \left(-\int_x^1 \lambda_t(t) dt\right)$$

$$= \exp \left(-\int_x^1 \frac{c}{t \left(\log \left(\frac{c}{t}\right) + k\right)} dt\right). \quad (3.9)$$

Take $\log \left(\frac{c}{t}\right) = u$, then equation (3.9) becomes
Measure of inaccuracy

\[ F(x) = \exp \left\{ - \int_{\log e}^{\log e} \frac{c}{u + k} \, du \right\} \quad (3.10) \]

Solving equation (3.10), we get equation (3.8). The proof of the converse is analogous to that of Theorem 3.3 and hence omitted.

\subsection{3.3 Characterization of probability distributions under the proportional hazards model assumption}

In this section we look into the problem of characterization of probability distributions by the form of the inaccuracy measure under the assumption that \( \overline{F}(x) \) and \( \overline{G}(x) \) satisfy the condition for being a proportional hazards model. Assume that

\[ \overline{G}(x) = \left( \overline{F}(x) \right)^{\theta}, \quad \theta > 0. \]

Then equation (2.56) becomes

\[ (I(t) + \log(\theta h_1(t)))h_1(t) = I'(t) + \theta h_1(t) \]

or

\[ I'(t) = h_1(t)(I(t) + \log(\theta h_1(t)) - \theta). \quad (3.11) \]

Multiplying both sides by \( \overline{F}(t) \), the above equation can be written as

\[ \frac{d}{dt} \left( \overline{F}(t) I(t) \right) = f(t)(\log(\theta h_1(t)) - \theta) \quad (3.12) \]

Integrating equation (3.12) over the range \( (t, \infty) \), we get

\[ I(t) = -\int_{t}^{\infty} \frac{f(x)}{\overline{F}(t)} (\log(\theta h_1(x)) - \theta) \, dx. \]

Hence, if \( (Y, \overline{G}) \) is the proportional hazards model of \( (X, \overline{F}) \), we have the relationship
\[ I(t) = E \left( \left( \theta - \log(\theta h_t(x)) \right) \mid X > t \right). \]  

(3.13)

**Remark: 3.3**

By direct calculation, the modified Kullback-Leibler divergence measure \( D(F,G;t) \), discussed in Chapter 2, is a constant under the proportional hazards model assumption. Hence, from the decomposition of \( I(t) \) in equation (3.1), the inaccuracy in the proportional hazards model at different time points varies with

\[ H(t) = 1 - E \left( \log h(x) \mid X > t \right), \]  

(3.14)

only. Consequent to this and the characterization of proportional hazard models by the constancy of \( D(F,G;t) \) given in Ebrahimi and Kirmani (1996,b), we can write

\[ I(t) = H(t) + K(\theta), \]  

(3.15)

where \( K(\theta) \) is a constant, independent of \( t \). From equation (3.15), we have

\[ I'(t) = H'(t). \]

Hence \( H(t) \) is an increasing function of \( t \) if and only if \( I(t) \) is increasing in \( t \). Belzunce et al. (2004) proved that an increasing \( H(t) \) determines \( F(t) \) uniquely. Thus, we have the following result.

**Theorem: 3.5**

If \( X \) has an absolutely continuous distribution function \( F(x) \) and an increasing inaccuracy measure \( I(t) \), then \( F(x) \) is uniquely determined by \( I(t) \).

**Remark: 3.4**

Functional form of \( I(t) \) characterizing various continuous distributions is given in Table 3.1 given below.
In Table 3.1, $A(\theta) = \theta - 1 - \log \theta$.

Di Cresenzo and Longobardi (2004) has characterized proportional reversed hazards model (See Section 2.4) using Kullback-Leibler measure of discrimination between past lifetime distributions. In this context, the Kullback-Leibler divergence for past life simplifies to

$$
D^*(F,G;t) = \phi - 1 - \log \phi,
$$

which is independent of $t$. Thus the inaccuracy measure of past life is

$$
I^*(t) = \phi - 1 - \log \phi + H(F;t), \quad (3.16)
$$

where $H(F;t)$ is the past entropy as defined in (2.35). Equation (3.16) can also be written as

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$F(x)$</th>
<th>$I(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalized Pareto (exponential, Lomax, re-scaled beta)</td>
<td>(1 + \frac{ax}{b} \left(1 + \frac{x}{a} \right)^{-\frac{1}{\beta}}); $x &gt; 0$</td>
<td>(\log(b + at) + (a + 1)^{-1} - \log(a + 1) + A(\theta))</td>
</tr>
<tr>
<td>Power (Uniform $c=1$)</td>
<td>$1 - x^c$, $0 &lt; x &lt; 1.$</td>
<td>(\frac{c - 1}{c} + \log\left(\frac{1 - t^c}{c}\right) + \frac{(c - 1)t^c}{1 - t^c}\log t + A(\theta))</td>
</tr>
<tr>
<td>Burr XII</td>
<td>((1 + x^c)^{-k}); $x &gt; 0.$</td>
<td>(k^{-1}\left[\log(ke) - c^{-1}\log(1 + r) + \frac{c - 1}{e} \sum_{r=1}^{1+t} \frac{(-1)^{r-1}}{1 + r} + A(\theta)\right])</td>
</tr>
<tr>
<td>Exponential geometric (Adamilis &amp; Loukas (1998))</td>
<td>((1 - p)e^{-\lambda t} \left(1 - pe^{-\lambda t}\right)^{-1})</td>
<td>(2 - \log \lambda + p^{-1}e^{\lambda t}\log\left[(1 - p)e^{-\lambda t}\right]).</td>
</tr>
</tbody>
</table>
\[ I^*(t) = E\left(\phi - \log(\phi \lambda(x)) \middle| X \leq t\right). \] (3.17)

Equation (3.17) can now be written as

\[ I^*(t) = \int_0^t \frac{f(x)}{F(t)} \left(\phi - \log(\phi \lambda(x))\right) dx. \]

Differentiating the above expression and rearranging the terms, we have the relationship

\[ \lambda_1(t) \left( I^*(t) + \log(\phi \lambda_1(t)) - \phi \right) = -I^*(t). \] (3.18)

Arguing as in the Theorem 3.5 and using equation (3.18), we have the following theorem.

**Theorem: 3.6**

If \( X \) has an absolutely continuous distribution function \( F(x) \) and an increasing inaccuracy measure \( I^*(t) \), then \( I^*(t) \) uniquely determines \( F(x) \).

**Remark: 3.5**

For the power distribution specified by

\[ F(x) = x^c, \ 0 < x < 1, \ c > 0, \]

by direct computation using equation (3.16), we get

\[ I^*(t) = \log t - c^{-1} - \log(c\phi) + \phi. \] (3.19)

Observing that the above equation is increasing in \( t \), the functional form of \( I^*(t) \) given in equation (3.19) holds if and only if \( X \) follow the power distribution.

We now characterize some common failure time distributions using a possible relationship between \( I(t) \) and the mean residual life function \( m(t) \), reviewed in Chapter 2.
Theorem: 3.7

If $X$ is a non-negative random variable admitting an absolutely continuous distribution function $F$ and $(Y, G)$ is the proportional hazards model of $(X, F)$, then the relationship

$$I(t) = A(\theta) + \log m(t),$$

(3.20)

where $A(\theta)$ is a real valued function of $\theta$, independent of $t$, and $m(t)$ is the mean residual life function of $X$ holds for every $t > 0$ if and only if $X$ has the generalized Pareto distribution.

Proof

Under the conditions of the theorem, when $X$ has generalized Pareto distribution, using equation (3.11), we get

$$I(t) = (a+1)^{-1} - \log (a+1) + \theta - 1 - \log \theta + \log(b+at),$$

which is of the form (3.20).

Conversely let $I(t)$ be as in equation (3.20).

Differentiating equation (3.20) with respect to $t$, we get

$$I'(t) = \frac{m'(t)}{m(t)},$$

(3.21)

Further using equations (3.11), (3.20) and (3.21), we have

$$\frac{m'(t)}{m(t)} = h_1(t) \left( C(\theta) + \log(h_1(t)m(t)) \right),$$

(3.22)

where $C(\theta) = A(\theta) + \log \theta - \theta$.

Using the relation, $h(t)m(t) = 1 + m'(t)$, equation (3.22) becomes

$$m'(t) = \left(1 + m'(t)\right)\left(C(\theta) + \log(1 + m'(t))\right).$$
Differentiating with respect to \( t \) and rearranging the terms we get
\[
m'(t) \left( C(\theta) + \log(1 + m'(t)) \right) = 0,
\]
where \( m'(t) = \frac{dm(t)}{dt} \). Equation (3.23) implies if either \( m'(t) = 0 \) or \( m'(t) = e^{-C(\theta)} - 1 \). In both the cases \( m'(t) \) is a constant. This shows that \( m(t) \) is linear in \( t \). The rest of the proof follows from Hall and Wellner (1981).

**Theorem: 3.8**

Under the conditions of the Theorem 3.7, the relationship
\[
I(t) = K(\theta) - \log h_1(t),
\]
where \( K(\theta) \) is a real valued function of \( \theta \) and \( h_1(t) \) is the hazard rate of \( F \) holds if and only if \( F \) is generalized Pareto distribution with survival function (3.4).

**Proof**

The ‘if’ part of the theorem follows from the expression for \( I(t) \) given in equation (3.11). To prove the only if part, assume that equation (3.24) holds. Differentiating equation (3.24) with respect to \( t \), we get
\[
I'(t) = -\frac{h_1'(t)}{h_1(t)}.
\]

Using equations (3.24) and (3.25), in equation (3.11), we get
\[
-\frac{h_1'(t)}{(h_1(t))^2} = A(\theta),
\]
where \( A(\theta) \) is a function of \( \theta \), independent of \( t \). Equation (3.26) can be written as
\[
\frac{d}{dt} \left( \frac{1}{h_1(t)} \right) = A(\theta).
\]
The solution of the above differential equation is

$$h_1(t) = \frac{1}{Bt + C},$$

where $B = A'(\theta)$ and $C^{-1} = h_1(0)$. This is the hazard rate of the generalized Pareto distribution. Since the distribution function is uniquely determined by the hazard rate, the theorem follows.

A parallel result exists for $I'(t)$ which we state as follows.

**Theorem: 3.9**

Let $X$ be a non-negative random variable admitting an absolutely continuous distribution function $F$ and let $(Y, G)$ is the proportional reversed hazards model of $(X, F)$. Then the relationship

$$I'(t) = K(\phi) - \log \lambda_1(t), \quad (3.27)$$

where $K(\phi)$ is a real function of $\phi$ and $\lambda_1(t)$ is the reversed hazard rate of $X$ holds for all real $t \geq 0$ if and only if $X$ has power distribution with distribution function specified by equation (3.7).

**Proof**

For the power distribution

$$I'(t) = \log t - c^{-1} - \log(c\phi) + \phi,$$

which is of the form equation (3.27) with $K(\phi) = \phi - c^{-1} - \log \phi$ and $\lambda_1(t) = ct^{-1}$. This proves the ‘if’ part.

Conversely assume that the relation (3.27) holds. Differentiating the equation (3.27) with respect to $t$, we get

$$I'(t) = -\frac{\lambda'_1(t)}{\lambda_1(t)}.$$
From equations (3.28) and (3.18), we get
\[
\frac{\dot{\lambda}_i(t)}{\lambda_i(t)} = K(\phi) - \phi + \log \phi. \tag{3.29}
\]

Equation (3.29) can be written as
\[
\frac{-d}{dt} \left( \frac{1}{\lambda_i(t)} \right) = K(\phi) - \phi + \log \phi. \tag{3.30}
\]

Solving the differential equation (3.30), we obtain
\[
\frac{1}{\lambda_i(t)} = A t + B,
\]
where \( A = \phi - \log \phi - K(\phi) \). This gives \( \lambda_i(t) = \frac{1}{A t + B} \), using equation (2.12), we conclude that \( X \) follows power distribution.

The next theorem focus attention on a characterization result for the Gompertz distribution by the form of \( I(t) \) in terms of the vitality function, reviewed in Section 2.1.

**Theorem: 3.10**

Let \( X \) be a non-negative random variable admitting an absolutely continuous distribution function \( F \) and with hazard rate \( h_i(t) \) and let \( G \) be the proportional hazard model of \( F \). Then \( X \) has the Gompertz distribution with survival function
\[
\overline{F}(x) = \exp \left( \frac{-B}{\log C} (C^x - 1) \right); x > 0, B > 0, C > 0, \tag{3.31}
\]
if and only if
\[
I(t) = K(\theta) + \beta v(t), \tag{3.32}
\]
for some real function \( K(\theta) \) and a real constant \( \beta < 0 \).
Proof

By direct calculation using equation (3.31) we get \( h_i(t) = BC' \) and

\[
H(t) = 1 - \left( \frac{1}{F(t)} \right)^{-1} \int_t^\infty \log h_i(x)f(x)\,dx
\]

\[
= 1 - \left( \frac{1}{F(t)} \right)^{-1} \int_t^\infty (\log B + x \log C)f(x)\,dx.
\]

Using the equation (3.15) and the last equation, we get the form of \( J(t) \) as stated in the theorem.

Conversely, we assume that equation (3.32) holds. Differentiating equation (3.32) with respect to \( t \), we get

\[
I'(t) = h_i(t)(I(t) - K(\theta) - \beta t).
\]  
(3.33)

Using equation (3.33) in equation (3.11), we have

\[
\log h_i(t) = \alpha - \beta t, \quad (3.34)
\]

where, \( \alpha = \theta - \log \theta - K(\theta) \). Equation (3.34) can be written as \( h_i(t) = e^{\alpha - \beta t} = BC' \), where \( B = e^\alpha \) and \( C = e^{-\beta} \). This is the hazard rate of the Gompertz distribution. Since the hazard rate uniquely determines the distribution, \( X \) follows Gompertz distribution.

Roy and Mukharjee (1989) examined the concept of averaging of hazard rate and looked into the problem of the characterization of life distributions using this concept. When one is interested in the pattern of failure of a device in a finite interval, instead of examining the nature of failure at each point in the interval this concept become a handy tool to describe the failure pattern. Rajesh (2001) has obtained characterization results for some lifetime distributions using the residual entropy function and the averages of hazard rate. The arithmetic, geometric and harmonic mean of hazard rates for a non-negative random variable \( Y \) with hazard rate \( h_2(t) \) are defined as
Measure of inaccuracy

\[ A^*(x) = \frac{1}{x} \int_0^x h_z(t) \, dt , \]

\[ G^*(x) = \exp \left( \frac{1}{x} \int_0^x \log h_z(t) \, dt \right) , \]  \hspace{1cm} (3.35)

and

\[ H^*(x) = \frac{1}{x} \int_0^x \frac{1}{h_z(t)} \, dt . \]

We now look into the problem of characterization of distributions based on the functional form of \( A^*(x) , G^*(x) \) and \( H^*(x) \) in terms of the residual inaccuracy measure.

**Theorem: 3.11**

Let \( X \) and \( Y \) be two non-negative random variables admitting absolutely continuous distribution functions such that \((Y, G)\) is the proportional hazards model of \((X, F)\). Denote by \( A^*(x) , G^*(x) \) and \( H^*(x) \) are arithmetic, geometric and harmonic mean of hazard rates of \( Y \) and let \( I(t) \) the residual inaccuracy function. The relationship

\[ A^*(t) = G^*(t) = H^*(t) = \exp(\theta - I(t)) , \]  \hspace{1cm} (3.36)

holds for all real \( t > 0 \) if and only if \( X \) follows the exponential distribution.

**Proof**

Assume equation (3.36) holds. This gives

\[ I(t) + \log G^*(t) = \theta . \]  \hspace{1cm} (3.37)

Using equation (3.35), equation (3.37) can be written as

\[ t I(t) + \int_0^t \log h_z(x) \, dx = \theta t . \]  \hspace{1cm} (3.38)
Differentiating with respect to $t$ and using equations (3.11), (3.38) simplifies to

$$I(t) + \log(\theta h_t(t)) = \theta.$$  \hspace{1cm} (3.39)

The equation (3.39) can be written as

$$I'(t) + \frac{h'_t(t)}{h_t(t)} = 0.$$  \hspace{1cm} (3.40)

Using equation (3.11), equation (3.40) becomes

$$h'_t(t) = 0.$$ 

This gives

$$h_t(t) = \lambda, \text{ a constant.}$$

Since the constancy of hazard rate is characteristic to the exponential model, the distribution of $X$ is exponential. From Roy and Mukharjee (1989), the properties

$$A^*(t) = G^*(t) = H^*(t)$$

is characteristic to the exponential model. Hence the sufficiency part holds. Conversely when $X$ follows exponential distribution with parameter $\lambda$, by direct calculations we get

$$A^*(t) = G^*(t) = H^*(t) = \lambda \theta$$

and

$$I(t) = \theta - \log \lambda \theta.$$  

The validity of equation (3.36) can be verified from the above expressions.

**Theorem: 3.12**

Assume that the conditions of the Theorem 3.11 holds. The relationship

$$I(t) = \left(\frac{a}{(a+1)\theta}\right)tA^*(t) = k,$$ \hspace{1cm} (3.41)

where $k$ is a constant, holds for all $t > 0$, if and only if $F$ has generalized Pareto distribution with survival function (3.4).
Proof

By direct calculation using equation (3.4), we get,

\[ I(t) = \log \left( \frac{b + at}{b} \right) + \left( \frac{a}{(a+1)} + \theta - \log \left( \frac{(a+1)\theta}{b} \right) \right). \]

That is,

\[ I(t) = \log \left( \frac{b + at}{b} \right) + k, \quad (3.42) \]

where \( k = \frac{a}{a+1} + \theta - \log \left( \frac{(a+1)\theta}{b} \right) \), is independent of \( t \). From the definitions of arithmetic mean of hazard rates and the proportional hazards model assumption, we have

\[ t A'(t) = \int_0^t \theta h_i(x) \, dx. \quad (3.43) \]

Since \( X \) follows generalized Pareto distribution, we get \( h_i(t) \) as

\[ h_i(t) = \frac{a+1}{b+at}. \quad (3.44) \]

Using equation (3.44) in equation (3.43), we have

\[ t A'(t) = \left( \frac{a+1}{a} \right) \theta \log \left( \frac{b+at}{b} \right). \quad (3.45) \]

Using equations (3.42) in (3.45) we get (3.41).

Conversely assume equation (3.41) holds. Differentiating this equation with respect to \( t \), we get

\[ I'(t) = \left( \frac{a}{(a+1)\theta} \right) \left( t A''(t) + A'(t) \right). \quad (3.46) \]
But from the definition of $A' (x)$ given in equation (3.35) and using the assumption of the theorem, we have the relationship

$$t A' (t) + A' (t) = \theta h(t). \tag{3.47}$$

Using the equation (3.47), equation (3.46) becomes

$$l(t) = \left(\frac{a}{a+1}\right) h(t). \tag{3.48}$$

From equations (3.11) and (3.48), we have

$$l(t) + \log(\theta h(t)) - \theta = \frac{a}{a+1}. \tag{3.49}$$

Using equation (3.41) in equation (3.49), we get

$$\left(\frac{a t}{(a+1)\theta}\right) A' (t) + \log h(t) = k_1,$$

which is equivalent to

$$\frac{a}{a+1} \int_0^t h(x)dx + \log h(t) = k_1. \tag{3.50}$$

where $k_1 = \frac{a}{a+1} + \theta - \log \theta - k$.

Differentiating equation (3.50) with respect to $t$ we get

$$\frac{h(t)}{h(t)} = \frac{a}{a+1}.$$ \[

The solution of the above equation is $h(t) = \left(c + \frac{a t}{a+1}\right)^{-1}$, where $c > 0$ is the constant of integration. This gives $
\tilde{F}(t) = \left(1 + \frac{a t}{c(a+1)}\right)^{-\left[\frac{1}{a}+1\right]}$, which is the
survival function of generalized Pareto distribution. This completes the proof of the sufficiency part.

### 3.4 Inaccuracy measure for weighted distributions

As mentioned in Section 2.4, the weighted distributions, defined by Rao (1965), find a lot of applications in theoretical statistics. The simplest form of the weighted distribution is the length-biased distribution defined in equation (2.24) namely

\[
f'_L(x) = \frac{xf(x)}{\mu}, \quad x > 0, \quad \mu = E(w(X)) < \infty.
\]  

(3.51)

The inaccuracy for the length-biased random variable \( X_L \) associated to a non-negative random variable \( X \) is,

\[
I_L = -\int_{0}^{\infty} f(x) \log f'_L(x) dx.
\]

Using equation (3.51), the above equation can be written as

\[
I_L = -\int_{0}^{\infty} f(x) \log \left( \frac{xf(x)}{\mu} \right) dx; \quad x > 0, \mu = E(X) < \infty
\]

\[
= -\int_{0}^{\infty} f(x) \log(f(x)) dx - \int_{0}^{\infty} f(x) \log x dx + \log \mu
\]

\[
I_L = H(F) - E(\log X) + \log (E(X)).
\]  

(3.52)

The equation (3.52) shows that the inaccuracy of the length-biased variable can be expressed in the form

\[
I_L = \text{Entropy} - \text{Geometric mean of } X - \text{logarithm of arithmetic mean of } X.
\]

The above equation expresses the inaccuracy for the length-biased random variable, \( X_L \) as the difference between entropy and the sum of the geometric mean and the logarithm of the arithmetic mean of \( X \).
The measure of residual inaccuracy, when we assume length-biased model instead of actual density can be expressed in terms of residual inaccuracy, geometric vitality function and mean residual life function of original distribution as follows. By definition

\[ I_L(t) = - \int_0^\infty \frac{f(x)}{F(t)} \log \left( \frac{ug(x)}{F(t)} \right) \, dx \]

\[ = - \int_0^\infty \frac{f(x)}{F(t)} \log \left( \frac{x f(x)}{\mu F(t)} \right) \, dx \]

\[ = M(t) - \log(G(t)) + \log(\mu F(t)) \], \hspace{1cm} (3.53) \]

where \( \log(G(t)) = \frac{1}{F(t)} \int_0^\infty f(x) \log x \, dx \) is the geometric vitality function (See equation (2.19)) and \( M(t) = -\frac{1}{F(t)} \int_0^\infty f(x) \log f(x) \, dx \), is the conditional measure of uncertainty proposed by Sankaran and Gupta (1999), which measures the uncertainty contained in \( f(t) \) about the predictability of the total lifetime of a unit which has survived to age \( t \). \( M(t) \) can be represented as the sum of residual entropy and total hazard rate as

\[ M(t) = H(t) - \log F(t). \]

Using the above representation for \( M(t) \), equation (3.53) becomes,

\[ I_L(t) = H(t) - \log[G(t)] + \log \left( \frac{F(t)}{\mu F(t)} \right). \] \hspace{1cm} (3.54) \]

Using equation (2.25), the above equation can be written as,

\[ I_L(t) = H(t) - \log(G(t)) + \log(t + m(t)). \] \hspace{1cm} (3.55) \]

But using the relationship between the mean residual life function and the vitality function, given in equation (2.18), the above expression can be rewritten as
Measure of inaccuracy

\[ I_{E}(t) = H(t) - \log G(t) + \log \nu(t), \]

where \( \nu(t) \) is the vitality function given in equation (2.17).

Another model of interest in lifetime analysis is the equilibrium distribution discussed in Section 2.4. For the equilibrium distribution, the inaccuracy measure takes the form

\[
I_{E}(t) = -\frac{1}{F(t)} \int_{0}^{\infty} f(x) \left( \log F(x) - \log \left( \frac{F(t) m(t)}{F(t)} \right) \right) dx.
\]

\[
= 1 + \log \left( m(t) F(t) \right) - \log F(t)
\]

\[
= 1 + \log m(t),
\]

where \( m(t) \) is the mean residual life function. Since \( m(t) \) determines the distribution uniquely, \( I_{E}(t) \) determines the distribution \( F \).

In fact

\[
F(x) = -\exp \left( \int_{0}^{x} \exp(1 - I_{E}(t)) dt \right) \exp(1 - I_{E}(x)) I_{E}'(x). \tag{3.56}
\]

Equation (3.56) enables one to characterize distributions by the functional form of \( I_{E}(t) \). The form of \( I_{E}(t) \) which characterizes some distributions is given in the following table.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>( F(x) )</th>
<th>( I_{E}(t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalized Pareto</td>
<td>( \left( 1 + \frac{ax}{b} \right)^{-\left( 1 + \frac{1}{a} \right)}, x &gt; 0 )</td>
<td>( 1 + \log \left( b + at \right) )</td>
</tr>
<tr>
<td>Power</td>
<td>( 1 - x^\alpha, 0 &lt; x &lt; 1 )</td>
<td>( 1 + \log \left( 1 - t \right)^{-\left( 1 - (c + 1) \left( 1 - t^{-\alpha} \right) \right)} )</td>
</tr>
<tr>
<td>Gamma</td>
<td>( \frac{\lambda^\alpha e^{-\lambda t} t^{\alpha-1}}{\Gamma(\alpha)}, x &gt; 0 )</td>
<td>( 1 + \log \left[ \frac{\alpha}{\lambda} - t - \frac{\lambda^{\alpha-1} e^{-\lambda t} t^\alpha}{F(t) \Gamma(\alpha)} \right] )</td>
</tr>
<tr>
<td>Exponential geometric</td>
<td>( (1 - p) e^{-\lambda x} \left( 1 - pe^{-\lambda x} \right)^{-1}, x &gt; 0 )</td>
<td>( 1 + \log \left[ (-\lambda p)^{-\frac{1}{\lambda}} \left( 1 - pe^{-\lambda x} \right) \log \left( 1 - pe^{-\lambda x} \right) \right] )</td>
</tr>
</tbody>
</table>
The characterizations considered in this chapter provide tools for identifying life distributions in terms of different forms of measure of inaccuracy, besides forging relationships between inaccuracy function and basic reliability characteristics. A major difference between the characterizations in this chapter and those in terms of reliability functions currently in use is that, in former, we get a feel of the extent to which the assumed model is inaccurate both in terms of lack of information and mis-specification.