International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 6 (Special Issue), 2007, pp 223-230

Characterizations of Proportional Hazard Models by Properties of Information Measures

N. Unnikrishnan Nair and R. P. Gupta
Dept. of Mathematics and Statistics
Dalhousie University, Canada
E-mail: nunair2004@yahoo.com

[Received November 14, 2005; Revised June 20, 2007; Accepted September 5, 2007]

Abstract

In the present paper, we establish some characterizations of proportional hazard and reversed hazard models based on generalized relative entropy and measure of inaccuracy, by the property that these measures are independent of the point of truncation of the models involved.

Keywords and Phrases: Proportional hazard models, relative entropy, measure of inaccuracy, generalized Pareto distribution.

AMS Classification: 62E, 62N05.

1 Introduction

The concept of proportional hazard (PH) models is well known in literature by its successful application in a variety of areas like reliability, survival analysis, medicine, biology, economics *etc.* Consider a non-negative random variable X with continuous survival function $\overline{F}(x) = P(X > x)$ and hazard rate

$$h_X(x) = \frac{d \log \overline{F}(x)}{dx}.$$

If Y is another non-negative random variable with hazard rate proportional to that of X *viz.*

$$h_Y(x) = \theta h_X(x),\ \theta > 0$$

or equivalently

$$\overline{G}(x) = \overline{F}^{\theta}(x) \tag{1}$$

where $\overline{G}(.)$ is the survival function of Y, then $(Y,\overline{G})$ is called the PH model corresponding to $(X,\overline{F})$. One can also consider the reversed hazard rate

$$\lambda_X(x) = \frac{d\log F(x)}{dx}$$

and define proportional reversed hazard (PRH) model (Y,G) related to (X,F) as

$$\lambda_Y(x) = \theta\lambda_X(X) \quad \text{or} \quad G(x) = F^{\theta}(x) \tag{2}$$

where $F(.)$ and $G(.)$ are the distribution functions of X and Y respectively. In spite of the popularity enjoyed by the proportional models, only a few attempts are made to develop characteristic properties that enable identification of these models in a real situation. One such result established by (4) makes use of the (7) directed information distance (also called relative entropy, directed divergence and cross entropy function in different contexts)

$$H(F,G) = \int f(x)\log\frac{f(x)}{g(x)}dx \tag{3}$$

between two distributions with distribution functions F and G and density functions f and g. They have shown that the distributions of X and Y truncated below at some point $t > 0$ with relative entropy

$$\overline{H}(F,G;t) = \int_t^{\infty} \frac{f(x)}{\overline{F}(t)}\log\Big(\frac{f(x)}{\overline{F}(t)}\Big/\frac{g(x)}{\overline{G}(t)}\Big)dx \tag{4}$$

is independent of t if and only if $(Y,\overline{G})$ is the PH model of $(X,\overline{F})$. Following this (8) studied the relative entropy when the range of integration is $(0,t)$ and characterized the PRH model through the property that

$$H(F,G;t) = \int_0^{t} \frac{f(x)}{F(t)}\log\Big(\frac{f(x)}{F(t)}\Big/\frac{g(x)}{G(t)}\Big) \tag{5}$$

does not contain t. In Section 2 we extend these results by considering a more general relative entropy measure and deduce the above results as special cases. We also compare the two measures (4) and (5) with reference to PR and PRH formations. The importance of PH models (Lehman alternative) and several results are given in (5). For interpretation of the truncated versions in (4) and (5) in the context of reliability as residual life and past life distributions see (3) and (2). Another concept of interest in statistical inference concerning PH models is measure of inaccuracy. If the experimenter assigns distribution function G whereas the true underlying distribution is F, then the inaccuracy of the assignment G is measured by

$$I(F,G) = -\int f(x)\log g(x)dx. \tag{6}$$

See (6) for properties of $I(F,G)$ and its applications to inference. We prove characterization of proportional models using (6) in Section 3.

2 Characterization by Extended Relative Entropy

In this section, we consider the extended relative entropy measure defined by (see (9))

$$H_r(F,G) = \frac{1}{r}\log\int\left[\frac{f(x)}{g(x)}\right]^r f(x)dx, r > -1. \tag{7}$$

For the finiteness of H_r it is assumed that $E_X\left(\frac{f(x)}{g(x)}\right)^r < \infty$. The motivation for suggesting (7) arises from
(a) it reduces to Kullback-Leibler measure as $r \to 0$.
(b) H_r is more sensitive to events with higher probabilities than H and therefore the former reveals uncertainty better than the latter.

Theorem 2.1. *Let A be the class of absolutely continuous distribution functions supported by the set of non-negative reals. For F, $G \in A$ with corresponding random variables X and Y, the following statements are equivalent*
(i) $(Y, \overline{G})$ is the PH model of $(X, \overline{F})$
(ii) For every real $r > 0$ and $0 < \theta < \frac{1+r}{r}$,

$$\overline{H}_r(F,G;t) = \frac{1}{r}\log\int_t^\infty \left[\frac{f(x)}{F(t)}\Big/\frac{g(x)}{G(t)}\right]^r \frac{f(x)}{F(t)}dx \tag{8}$$

is independent of t for all $t > 0$.

Proof. First we prove that (i) $\Rightarrow$ (ii). From $\overline{G}(x) = \overline{F}^\theta(x)$ and $g(x) = -\theta\overline{F}^{(\theta-1)}(x)$ $f(x)$, we can write (8) as

$$\begin{aligned}\overline{H}_r(F,G;t) &= \frac{1}{r}\log\frac{-\overline{F}^{\theta r}(t)}{\overline{F}^{(r+1)}(t)\theta^r}\int_t^\infty \frac{f(x)dx}{\overline{F}^{(\theta-1)}(x)}\\ &= r^{-1}\log\theta^{-r}(1-r(\theta-1))^{-1}, \quad \theta < \frac{1+r}{r}\end{aligned}$$

which is independent of t. Conversely, assuming constancy of $\overline{H}_r$ say, $\overline{H}_r = K$, (8) reduces to

$$\frac{e^{Kr}\overline{F}^{r+1}(t)}{\overline{G}^r(t)} = \int_t^\infty \frac{f^{r+1}(x)dx}{g^r(x)}.$$

Differentiating,

$$e^{Kr}\frac{(r+1)\overline{F}^r f(t)}{\overline{G}^r(t)} - \frac{r\overline{F}^{(r+1)}(t)g(t)}{\overline{G}^{r+1}(t)} = \frac{f^{r+1}(t)}{g^r(t)},$$

which simplifies to

$$\frac{(r+1)h_Y^r(t)}{h_X^r(t)} - r\frac{h_Y^{r+1}(t)}{h_X^{r+1}(t)} = e^{-Kr}. \tag{9}$$

Writing

$$p(t) = \frac{h_Y(t)}{h_X(t)},$$

we have

$$(r+1)p^r(t) - rp^{r+1}(t) = e^{-Kr}.$$

Differentiating again with respect to t

$$(r+1)rp^{r-1}(t)p'(t) - r(r+1)p^r(t)p'(t) = 0$$

or

$$r(r+1)p'(t)p^{r-1}(t)[1-p(t)] = 0. \tag{10}$$

The solutions for $p(t)$ are, $p(t) = 0$, 1, θ where θ is constant, of which first one is inadmissible and the second gives $\overline{F} = \overline{G}$. Hence $h_Y(t) = \theta h_X(t)$, $\theta > 0$ for all t and this implies (i). □

Remarks:

1. As the form of F is nowhere used in the theorem, it is true of PH models with any base line distribution satisfying the conditions.

2. Reversing the roles of F and G we see that for PH models, $\overline{H}_r(G, F; t) = r^{-1} \log \theta^{r+1}(r\theta - r + \theta)^{-1}$ is also independent of t irrespective of the base line distribution and for all $\theta > \frac{r}{r+1}$. The converse also is true as can be seen by the method of proof of Theorem 2.1.

3. Both $\overline{H}_r(F, G; t)$ and $\overline{H}_r(G, F; t)$ are functions of θ and r. As a function of θ, one can write the relationship $\overline{H}_r(F, G; \theta) = \overline{H}_r(G, F; \frac{1}{\theta})$.

4. $\overline{H}_r(F, G; \theta)$ is an increasing function of r, decreasing function of θ, for $0 < \theta < 1$, attains the minimum value zero at $\theta = 1$, and increasing in the interval $1 < \theta < \frac{1+r}{r}$.

5. For PRH model $G(x) = F^\theta(x)$ a similar result stated in Theorem 2.2 holds. The proof is identical to that of

   Theorem 2.1 with F and G replacing $\overline{F}$ and $\overline{G}$ and is therefore not given.

Theorem 2.2. *The distribution functions F and G in A, (Y, G) is the PRH model corresponding to (X, F) if and only if*

$$H(F, G; t) = \frac{1}{r} \log \int_0^t \left[\frac{f(x)}{F(t)} \Big/ \frac{g(x)}{G(t)}\right]^r \frac{f(x)}{F(t)} dx$$

is independent of t for every real $r > 0$ and $0 < \theta < \frac{1+r}{r}$. The remarks 1 to 4 hold for $H_r(F, G; t)$ also with appropriate modification. Further

$$\overline{H}_r(F, G; t) = H_r(F, G; t) \quad \textit{and} \quad \overline{H}_r(G, F; t) = H_r(G, F; t).$$

3 Characterization by Inaccuracy Measure

Unlike the relative entropy, the property of the measure of inaccuracy (6) depends on the base line distribution. Characterizations of some important life distributions that have simple forms for (6) are discussed in the following theorems.

Theorem 3.1. *Let F, G be distribution functions in A such that $(Y, \overline{G})$ is the PH model of $(X, \overline{F})$. Then the inaccuracy measure truncated below some $t > 0$,*

$$\overline{I}(F, G; t) = -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \log \frac{g(x)}{\overline{G}(t)} dx \tag{11}$$

has the log linear form

$$\overline{I}(F, G; t) = \log \frac{at + b}{\theta(a+1)} - \frac{a + \theta(a+1)}{a+1} \tag{12}$$

for all $t > 0$, $a > -1$ and $b > 0$ if and only if F has generalized Pareto distribution

$$\overline{F}(x) = \left(1 + \frac{ax}{b}\right)^{-(1+\frac{1}{a})}. \tag{13}$$

In particular, the only absolutely continuous distribution on $(0, \infty)$ for which $\overline{I}(F, G; t)$ =constant is the exponential.

Proof. When $(Y, \overline{G})$ is a PH model corresponding to $(X, \overline{F})$, (11) becomes

$$[\log \overline{G}(t) - \overline{I}(t)]\overline{F}(t) = \int_t^\infty f(x) \log g(x) dx \tag{14}$$

with $\overline{I}(t) = \overline{I}(F, G; t)$.
Differentiating (14)

$$\left[\frac{g(t)}{\overline{G}(t)} + \overline{I}'(t)\right]\overline{F}(t) + (\log \overline{G}(t) - \overline{I}(t))f(t) = f(t) \log g(t).$$

Using the definition of the failure rate, the last expression takes the form

$$h_X(t) = \frac{h_Y(t) + \overline{I}'(t)}{\overline{I}(t) + \log h_Y(t)} \tag{15}$$

$$= \frac{\theta h_X(t) + \overline{I}'(t)}{\overline{I}(t) + \log \theta h_X(t)} \tag{16}$$

for the PH model, since in that case $h_Y(t) = \theta h_X(t)$. We note that for the distribution (13) direct calculation using (11) gives (12). To establish the converse we substitute $\overline{I}(t)$ from (12) in (16) to verify

$$h_X(t) \log \frac{at + b}{a+1} + \frac{a + \theta(a+1)}{a+1} h_X(t) + h_X(t) \log h_X(t) = \theta h_X(t) + \frac{a}{at+b}. \tag{17}$$

Differentiating with respect to t and then substituting for $\log h_X(t)$, from (17)

$$\Big[1 + \frac{a}{(at+b)h_X(t)}\Big]\Big[h'_X(t) + \frac{ah_X(t)}{at+b}\Big] = 0.$$

The solutions are

$$h_X(t) = c(at+b)^{-1} \quad \text{and} \quad h_X(t) = -\frac{a}{at+b}$$

where $c > 0$ is a constant of integration. The first solution leads to the generalized Pareto form and the second to the uniform distribution which is also a member of the same family. This completes the proof of the first part.
Finally as a tends to zero in (13)

$$\overline{F}(x) = \exp\big[-\frac{x}{b}\big], \quad x > 0, b > 0$$

and

$$\overline{I}(t) = \log\big(\frac{b}{\theta}\big) - \theta$$

which is independent of t. □

Note:

1. The family of distribution considered in Theorem 3.1 includes the Lomax distribution with survival function

$$\overline{F}(x) = \alpha^{\beta}(x+\alpha)^{-\beta}, \quad x > 0, \alpha > 0, \beta > 0$$

with $a = (\alpha - 1)^{-1}$, $b = a\beta$ and the beta distribution specified by

$$\overline{F}(x) = \big(1 - \frac{x}{R}\big)^{d}, \quad 0 < x < R, d, R > 0$$

for a choice of $a = -(1+d)^{-1}$ and $b = R(1+d)^{-1}$. Uniform distribution is special case when $d = 1$.

2. The inaccuracy measure when F is used in the place of G truncated below $t > 0$,

$$\overline{I}(G, F; t) = -\int_t^{\infty} \frac{g(x)}{\overline{G}(t)} \log \frac{f(x)}{\overline{F}(t)} dx.$$

For the generalized Pareto family

$$\overline{I}(t) = \log \frac{at+b}{a+1} + \frac{(1+2a)b}{\theta(1+a)}.$$

This is also independent of t in the exponential case. A characterization theorem similar to Theorem 3.1 holds in this case also.

3. On the other hand if we truncate the distributions of X and Y above $t > 0$, the resulting inaccuracy measure is

$$I(F, G; t) = -\int_0^t \frac{f(x)}{F(t)} \log \frac{g(x)}{G(t)} dx.$$

Proceeding as in Theorem 3.1, one can find the identity

$$\lambda_X(t) = \frac{\lambda_Y(t) - I'(t)}{\log \lambda_Y(t) + I(t)}.$$

When $I(t)$ is a constant with regard to t, for the PRH model $G(t) = F^\theta(t)$, it follows that $\lambda_X(t)$ should be a constant for all $t > 0$. There is no absolutely continuous distribution on the positive real line with a constant reversed hazard rate (see (1)) and accordingly no characteristic property of the PRH model exist as in the exponential case in Theorem 3.1. However, there is a negative exponential distribution for X in $(-\infty, b)$, $b > 0$ characterizing PRH model with $I(t) = c$, a constant.

4 Conclusions

In this paper we have presented three characterization theorems. Of these, the first two results relate to characterization of the proportional hazard model by the independence of the point of truncation t of an extended relative entropy measure of two absolutely continuous distributions defined over (t, ∞) and an analogous result for the proportional reversed hazard model over the interval $(0, t)$. Some properties of the relative entropy measures under the two models are also given. The generalized Pareto distribution is characterized by the truncated inaccuracy measure of log linear form. This distribution contains the exponential, beta and Lomax distributions as special cases, of which the exponential has an inaccuracy measure that does not depend on the point of truncation.

Acknowledgement

This work was done when first author was visiting Dalhousie University. The research was supported by NSERC grant 4100. The authors thank the referee for his comments on an earlier draft of the paper.

References

[1] Block, H. W., Savits, T. H. and Singh, H. (1998), The reversed hazard rate function, *Probability in the Engineering and Informational Sciences* **12**, 69–90.

[2] Di Cresenzo, A. and Longobardi, M. (2002), Entropy based measures of uncertainty in past life time distibutions, *Journal of Applied Probability* **39**, 439–440.

[3] Ebrahimi, N. (1996), How to measure uncertainty in residual life distributions, *Sankhya, Series A* **58**, 48–56.

[4] Ebrahimi, N. and Kirmani, S. N. U. A. (1996), A characterization of proportional hazard models through a measure of discrimination between two residual life distributions, *Biometrika* **83(1)**, 233–235.

[5] Gupta, R. C., Gupta, P. L. and Gupta, R. D. (1998), Modelling failure time data with Lehmann alternative, *Communications in Statistics, Theory and Methods* **27(4)**, 887–904.

[6] Kerridge, D. F. (1961), Inaccuracy and inference, *Journal of the Royal Statistical Society* **Series B, 23**, 184–194.

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

[8] Nanda, A. K. and Gupta, R. D. (2000), Some properties of the reversed hazard rate function, *Statistical Methods* **3(2)**, 108–124.

[9] Renyi, A. (1961), On measures of entropy and information, *in Proceedings of Fourth Berkely Symposium on Mathematics, Statistics and Probability*, 1960, Vol. 1, University of California Press, pp. 547–561.