Proceedings of the World Congress on Engineering 2012 Vol I
WCE 2012, July 4 - 6, 2012, London, U.K.

On Dynamic Cumulative Residual Inaccuracy Measure

HC Taneja and Vikas Kumar *

Abstract—**An alternate measure of entropy based on distribution function rather than the density function of a random variable X, called the cumulative residual entropy (CRE) was proposed by Rao et al. (2004). In this communication the concept of CRE has been extended to cumulative residual inaccuracy (CRI) and then to a dynamic version of it. A characterization problem for the proposed dynamic inaccuracy measure has been studied under proportional hazard model. Three specific lifetime distributions exponential, Pareto and the finite range have been characterized using the proposed dynamic inaccuracy measure.**

Keywords: entropy, cumulative residual entropy, inaccuracy, hazard rate , mean residual life function, proportional hazard model.

1 Introduction

Let X and Y be two non-negative random variables representing time to failure of two systems with p.d.f. respectively $f(x)$ and $g(x)$. Let $F(x) = P(X \leq x)$ and $G(Y) = P(Y \leq y)$ be failure distributions, and $\bar{F}(x) = 1 - F(x)$, $\bar{G}(x) = 1 - G(x)$ be survival functions of X and Y respectively . Shannon's (1948) measure of uncertainty associated with the random variable X and Kerridge measure of inaccuracy (1961) are given by

$$H(f) = -\int_0^\infty f(x) \log f(x) dx \ . \tag{1}$$

and

$$H(f;g) = -\int_0^\infty f(x) \log g(x) dx \ . \tag{2}$$

respectively. In case $g(x) = f(x)$, then (2) reduces to (1).

The measures (1) and (2) are not applicable to a system which has survived for some unit of time. Ebrahimi (1996) considered the entropy of the residual lifetime $X_t = [X - t|X > t]$ as a dynamic measure of uncertainty given by

$$H(f;t) = -\int_t^\infty \frac{f(x)}{\bar{F}(t)} log \frac{f(x)}{\bar{F}(t)} dx \ . \tag{3}$$

Extending the dynamic measure of information, a dynamic measure of inaccuracy, refer to Taneja et al. (2009) is given as

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

Rao et al. (2004) introduced an alternate measure of entropy called cumulative residual entropy(CRE) of a random variable X defined as

$$\xi(F) = -\int_0^\infty \bar{F}(x) \log \bar{F}(x) dx \ , \tag{5}$$

where $\bar{F}(x) = 1 - F(x)$ is the survival function.

This measure is based on cumulative distribution function (CDF) rather than probability density function, and is thus, in general more stable since the distribution function is more regular because it is defined in an integral form unlike the density function which is defined as the derivative of the distribution. Some general results regarding this measure have been studied by Rao (2005), Drissi et al. (2008) and Navarro et al. (2009).

Asadi and Zohrevand (2007) have defined the dynamic cumulative residual entropy (DCRE) as the cumulative residual entropy of the residual lifetime $X_t = [X - t|X > t]$. This is given by

$$\xi(F;t) = -\int_t^\infty \frac{\bar{F}(x)}{\bar{F}(t)} \log \frac{\bar{F}(x)}{\bar{F}(t)} dx \ . \tag{6}$$

In this communication in Section 2, we define a cumulative residual inaccuracy measure. In Section 3, we propose a dynamic cumulative residual inaccuracy measure. In Section 4, we prove that dynamic measure determines the lifetime distribution functions uniquely, and characterize three specific lifetime distributions in this context.

2 Cumulative Residual Inaccuracy

If $\bar{F}(x)$ and $\bar{G}(x)$ are survival functions of lifetime random variables X and Y respectively, then the cumulative residual inaccuracy (CRI) is defined as

$$\xi(F;G) = -\int_0^\infty \bar{F}(x) \log \bar{G}(x) dx \ . \tag{7}$$

*Department of Applied Mathematics, Delhi Technological University, (Formerly Delhi College of Engineering) Bawana Road, Delhi-110042, India. Email: hctaneja@rediffmail.com

When these two distributions coincide, the measure (7) reduces to the cumulative residual entropy (5).

If the two random variables X and Y satisfy the proportional hazard model (PHM), refer to Cox (1959) and Efron (1981), that is, if $\lambda_G(x) = \beta\lambda_F(x)$, or equivalently

$$\bar{G}(x) = [\bar{F}(x)]^{\beta}, \tag{8}$$

for some constant $\beta > 0$, then obviously the cumulative residual inaccuracy (7) reduces to a constant multiple of the cumulative residual entropy (5).

Example 2.1 Let X be a non-negative random variable with p.d.f.

$$f_X(x) = \begin{cases} 2x & ; \quad \text{if } 0 \le x < 1 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

and survival function $\bar{F}(x) = 1 - F(x) = (1 - x^2)$, and let the random variable Y be uniformly distributed over $(0, 1)$, then its density and survival functions are respectively given by

$$g_Y(x) = 1 \ \ and \ \ \bar{G}_Y(x) = 1 - x \ , \quad 0 < x < 1.$$

Substituting these in (7) , we obtain the cumulative residual inaccuracy as

$$\xi(F, G) = \frac{7}{18}.$$

Example 2.2 Let a non-negative random variable X be uniformly distributed over (a, b), $a < b$, with density and distribution functions respectively given by

$$f(x) = \frac{1}{b-a} \ \ and \ \ F(x) = \frac{x-a}{b-a}, \quad a < x < b.$$

If the random variables X and Y satisfy the proportional hazard model (PHM), then the distribution function of the random variable Y is

$$\bar{G}(x) = [\bar{F}(x)]^{\beta} = \left[\frac{b-x}{b-a}\right]^{\beta} \quad a < x < b, \ \beta > 0.$$

Substituting these in (7) and simplifying we obtain the cumulative inaccuracy measure as

$$\xi(F; G) = \frac{\beta(b-a)}{4} \ . \tag{9}$$

3 Dynamic Cumulative Residual Inaccuracy

In life-testing experiments normally the experimenter has information about the current age of the system under consideration. Obviously the CRI measure (7) defined above is not suitable in such a situation and should be modified to take into account the current age also. Further, if X is the lifetime of a component, which has already survived upto time t, then the random variable $X_t = [X - t | X > t]$ called the residual lifetime random variable has the survival function

$$\bar{F}_t(x) = \begin{cases} \frac{\bar{F}(x)}{\bar{F}(t)} & ; \quad \text{if } x > t \\ 1 & ; \quad \text{otherwise} \end{cases}$$

and similarly for $\bar{G}_t(x)$. Thus the cumulative inaccuracy measure for the residual lifetime distribution is given by

$$\xi(F, G; t) = -\int_t^{\infty} \bar{F}_t(x) \log \bar{G}_t(x) dx \tag{10}$$

$$= -\int_t^{\infty} \frac{\bar{F}(x)}{\bar{F}(t)} \log \frac{\bar{G}(x)}{\bar{G}(t)} dx \ . \tag{11}$$

Obviously when $t = 0$, then (11) becomes (7).

Example 3.1 Let X be a non-negative random variable with p.d.f.

$$f_X(x) = \begin{cases} 2x & ; \quad \text{if } 0 \le x < 1 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

and the survival function $\bar{F}(x) = 1 - F(x) = (1 - x^2)$, and let the random variable Y be uniformly distributed over $(0, 1)$ with density and survival functions given respectively by

$$g_Y(x) = 1 \ \ and \ \ \bar{G}_Y(x) = 1 - x \ , \quad 0 < x < 1.$$

Substituting these values in (11) , we obtain the dynamic cumulative residual inaccuracy measure as

$$\xi(F, G; t) = \left(\frac{9(1-t) - 2(1-t)^2}{18(1+t)}\right).$$

Example 3.2 Let X and Y be two non-negative random variables with survival functions $\bar{F}(x) = (x+1)e^{-x}$, $x > 0$ and $\bar{G}(x) = e^{-2x}$, $x > 0$. Substituting these values in (11), we obtain

$$\xi(F, G; t) = \left(\frac{6 + 2t}{t + 1}\right).$$

Taking limit as $t \to 0$, we obtain

$$\lim_{t \to 0} \xi(F, G; t) = \xi(F, G) = 6.$$

4 Characterization Problem

The general characterization problem is to determine when the proposed dynamic inaccuracy measure (11) characterizes the distribution function uniquely. It is known that the hazard rate $\lambda_F(t) = \frac{f(t)}{\bar{F}(t)}$ and the mean residual life function $m_F(t) = \frac{\int_t^{\infty} \bar{F}(x)dx}{\bar{F}(t)}$ characterize the distribution and the relation between the two is given by

$$\lambda_F(t) = \frac{1 + m_F'(t)}{m_F(t)} \ . \tag{12}$$

ISBN: 978-988-19251-3-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

We study the characterization problem under the proportional hazard model (8).

Theorem 4.1 Let X and Y be two non negative random variables with survival functions $\bar{F}(x)$ and $\bar{G}(x)$ satisfying the proportional hazard model (8). Let $\xi(F,G;t) < \infty, \forall\, t \geq 0$ be an increasing function of t, then $\xi(F,G;t)$ uniquely determines the survival function $\bar{F}(x)$ of the variable X.

Proof. Rewriting (11) as

$$\xi(F,G;t) = -\frac{1}{\bar{F}(t)}\int_t^{\infty} \bar{F}(x)\log \bar{G}(x)dx + m_F(t)\log \bar{G}(t)\ , \tag{13}$$

where $m_F(t)$ is the mean residual life function. Substituting (8) into (13) gives

$$\xi(F,G;t) = -\frac{\beta}{\bar{F}(t)}\int_t^{\infty} \bar{F}(x)\log \bar{F}(x)dx + \beta m_F(t)\log \bar{F}(t)\ .$$

Differentiating this w.r.t. t both sides, we obtain

$$\begin{aligned}\xi'(F,G;t) &= \beta\log \bar{F}(t)[1+m_F'(t)] - \beta\lambda_F(t)m_F(t) \\ &\quad -\beta\lambda_F(t)\int_t^{\infty}\frac{\bar{F}(x)}{\bar{F}(t)}\log \bar{F}(x)dx\ ,\end{aligned} \tag{14}$$

where $\lambda_F(t)$ is hazard rate function. Substituting (12) and (13) in (14) we obtain

$$\xi'(F,G;t) = \lambda_F(t)\{\xi(F,G;t) - \beta m_F(t)\}\ . \tag{15}$$

Let F_1, G_1 and F_2, G_2 be two sets of the probability distribution functions satisfying the proportional hazard model, that is, $\lambda_{G_1}(x) = \beta\lambda_{F_1}(x)$, and $\lambda_{G_2}(x) = \beta\lambda_{F_2}(x)$, and let

$$\xi(F_1,G_1;t) = \xi(F_2,G_2;t)\ \forall\ \ t \geq 0\ . \tag{16}$$

Differentiating it both sides w.r.t. t, and using (15), we obtain

$$\lambda_{F_1}(t)\{\xi(F_1,G_1;t) - \beta m_{F_1}(t)\} = \lambda_{F_2}(t)\{\xi(F_2,G_2;t) - \beta m_{F_2}(t)\}. \tag{17}$$

If for all $t \geq 0$, $\lambda_{F_1}(t) = \lambda_{F_2}(t)$, then $\bar{F}_1(t) = \bar{F}_2(t)$ and the proof will be over, otherwise, let

$$A = \{t : t \geq 0, and\ \ \lambda_{F_1}(t) \neq \lambda_{F_2}(t)\} \tag{18}$$

and assume the set A to be non empty . Thus for some $t_0 \in A$, $\lambda_{F_1}(t_0) \neq \lambda_{F_2}(t_0)$. Without loss of generality suppose that $\lambda_{F_2}(t_0) > \lambda_{F_1}(t_0)$. Using this, (17) for $t = t_0$ gives

$$\xi(F_1,G_1;t_0) - \beta m_{F_1}(t_0) > \xi(F_2,G_2;t_0) - \beta m_{F_2}(t_0),$$

which implies that

$$m_{F_1}(t_0) < m_{F_2}(t_0),$$

a contradiction. Thus the set A is empty set and this concludes the proof.

Next, we give the characterization theorem.

Theorem 4.2 Let X and Y be two non-negative continuous random variables satisfying the PHM (8). If X is with mean residual life $m_F(t)$, then the dynamic cumulative residual inaccuracy measure

$$\xi(F,G;t) = c\, m_F(t), \quad c > 0 \tag{19}$$

if, and only if
(i) X follows the exponential distribution for $c = \beta$,
(ii) X follows the Pareto distribution for $c > \beta$,
(iii) X follows the finite range distribution for $0 < c < \beta$.

Proof. First we prove the 'if' part.
(i) If X has exponential distribution with survival function $\bar{F}(x) = \exp(-\theta x), \theta > 0$, then the mean residual life function $m_F(t) = \frac{1}{\theta}$. The dynamic cumulative residual inaccuracy measure (11) under PHM is given as

$$\xi(F,G;t) = \frac{\beta}{\theta} = c m_F(t),$$

for $c = \beta$.

(ii) If X follows Pareto distribution with p.d.f.

$$f(x) = \frac{ab^a}{(x+b)^{a+1}}, \quad a\ > 1,\ \ b >\ 0,$$

then the survival function is

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

and the mean residual life is

$$m_F(t) = \frac{\int_t^{\infty}\bar{F}(x)dx}{\bar{F}(t)} = \frac{t+b}{a-1}\ . \tag{20}$$

The dynamic cumulative inaccuracy measure (11), under PHM is given by

$$\xi(F,G;t) = \frac{\beta a(t+b)}{(a-1)^2} = c m_F(t)\ ,$$

for $c = \frac{\beta a}{a-1} > \beta$.

(iii) In case X follows finite range distribution with p.d.f.

$$f(x) = a(1-x)^{a-1}, \quad a\ > 1,\ \ 0 \leq x \leq\ 1,$$

then the survival function is

$$\bar{F}(x) = 1 - F(x) = (1-x)^a,$$

and the mean residual life is

$$m_F(t) = \frac{1-t}{a+1}.$$

ISBN: 978-988-19251-3-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

The inaccuracy measure (11) under PHM is given by

$$\xi(F,G;t)=\frac{\beta a(1-t)}{(a+1)^2}=cm_F(t)\ ,$$

for $c=\frac{\beta a}{a+1}<\beta$.
This proves the 'if' part.

To prove the only if part, consider (19) to be valid. Using (13) under PHM, it gives

$$-\frac{\beta}{\bar{F}(t)}\int_t^\infty \bar{F}(x)\log\bar{F}(x)dx+\beta m_F(t)\log\bar{F}(t)=cm_F(t)\ .$$

Differentiating it both sides w.r.t. t, we obtain

$$\begin{aligned}\frac{c}{\beta}m'_F(t) &= m'_F(t)\log\bar{F}(t)-\lambda_F(t)m_F(t)+\log\bar{F}(t)\\ &\quad -\lambda_F(t)\frac{1}{\bar{F}(t)}\int_t^\infty \bar{F}(x)\log\bar{F}(x)dx,\end{aligned}$$

$$\begin{aligned}&= m'_F(t)\log\bar{F}(t)-\lambda_F(t)m_F(t)+\log\bar{F}(t)\\ &\quad +\lambda_F(t)[\frac{c}{\beta}m_F(t)-m_F(t)\log\bar{F}(t)]\ .\end{aligned}$$

From (12) put $m'_F(t)=\lambda_F(t)m_F(t)-1$ and simplify, we obtain

$$\lambda_F(t)m_F(t)=\frac{c}{\beta}\ ,$$

which implies

$$m'_F(t)=\frac{c}{\beta}-1\ .$$

Integrating both sides of this w.r.t. t over $(0,\ x)$ yields

$$m_F(x)=(\frac{c}{\beta}-1)x+m_F(0). \tag{21}$$

The mean residual life function $m_F(x)$ of a continuous non-negative random variable X is linear of the form (21) if, and only if the underlying distribution is exponential $for\ c=\beta$, Pareto $for\ c>\beta$, or finite range $for\ 0<c<\beta$, refer to Hall and Wellner (1981). This completes the theorem.

Conclusion By considering the concept of cumulative residual inaccuracy and extending it to its dynamic version, we have characterized certain specific life-time distributions functions like exponential, Pareto and the finite range distributions which play a vital role in reliability modeling. Also the results reported generalize the existing results in context with cumulative residual entropy.

References

[1] Asadi, M., Zohrevand, Y. (2007). On the dynamic cumulative residual entropy. Journal of Stat. Planning and Inference 137,1931-1941.

[2] Cox, DR. (1959). The analysis of exponentially distributed lifetimes with two type of failure, J. Roy. Statist. Soc. Ser. B, 21: 411-421.

[3] Drissi, N., Chonavel, T., Boucher, J.M. (2008). Generalized cumulative residual entropy for distributions with unrestricted supports. Re. Lett. Sign. Proc.

[4] Ebrahimi, N. (1996). How to measure uncertainty about residual lifetime? Sankhya A 58: 48-57, .

[5] Efron, B. (1981). Censored data and the bootstrap , J. Am. Statist. Assoc, 76: 312-319.

[6] Hall, W.J., Wellner, J.A. (1981). Mean residual life. In Statistics and Related Topics (Csrg, et al. eds), North-Holland, Amsterdam, 169-184.

[7] Kerridge, D.F. (1961). Inaccuracy and inference. J. Roy. Statist. Soc. Ser. B 23: 184-194.

[8] Navarro, J., Aguila, Y.D., Asadi, M. (2009). Some new results on the cumulative residual entropy. Journal of Stat. Planning and Inference.

[9] Rao, M., Chen, Y., Vemuri, B.C., Wang, F. (2004). Cumulative residual entropy: a new measure of information . IEEE Transactions on Information Theory 50, 1220-1228.

[10] Rao, M., (2005). More on a new concept of entropy and information. Journal of Theoretical Probability 18,967-981.

[11] Shannon, C.E. (1948). A mathematical theory of communication. Bell System Technical Journal, pp. 279-423 and 623-656.

[12] Taneja, H.C, Kumar, V. Srivastava, R. (2009). A dynamic measure of inaccuracy between two residual lifetime distributions. International Mathematical Forum 4(25):1213-1220.