STATISTICA, anno LXXX, n. 3, 2020

DYNAMIC INFORMATION VOLATILITY FUNCTION

Rajesh Ganapathi
Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, Kerala, India
Sreenarayanapurath Madhavan Sunoj[1]
Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, Kerala, India
Paduthol Godan Sankaran
Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, Kerala, India

1. INTRODUCTION

Let X be an absolutely continuous random variable with cumulative distribution function $F(t) = P(X \leq t)$ and survival function $\bar{F}(t) = P(X > t) = 1 - F(t)$. Then, a measure of uncertainty defined by Shannon (1948), is given by

$$\xi(X) = -E(\log f(X)) = -\int_{\mathscr{R}} (\log f(x)) f(x) dx, \quad (1)$$

where f is the probability density function of X. Eq. (1) measures the expected uncertainty contained in f about the predictability of an outcome of X, known as Shannon differential entropy measure, and for the estimation of (1), we refer to Beirlant *et al.* (1997), Lake (2009) and the references therein. Since $-\log f(X)$ is a random variable and $\xi(X)$ is the expectation of $-\log f(X)$, Liu (2007) introduced a new statistical measure, the variance of $-\log f(X)$, known as information volatility, given by

$$\begin{aligned} \eta(X) &= \mathrm{Var}(-\log f(X)) = \mathrm{Var}(\log f(X)) \\ &= E(\log f(X))^2 - (E(-\log f(X))^2 \\ &= E(\log f(X))^2 - (\xi(X))^2. \quad (2) \end{aligned}$$

$\eta(X)$ measures the variability of the uncertainty associated in the probability distribution of X. That is, $\eta(X)$ measures how far the uncertainty of X is deviated from the average uncertainty based on $\xi(X)$. We now compute $\eta(X)$ for some important reliability models.

[1] Corresponding Author. E-mail: smsunoj@cusat.ac.in

EXAMPLE 1. *When X follows equilibrium distribution with probability density function $f_E(x)=\frac{\bar{F}(x)}{\mu}$ where $\mu=E(X)<\infty$, then*

$$\eta_E(X)=\frac{1}{\mu}\int_0^{\infty}\left(\log\bar{F}(x)\right)^2\bar{F}(x)dx-\left(\frac{\zeta(X)}{\mu}\right)^2,$$

where $\zeta(X)=-\int_0^{\infty}\left(\log\bar{F}(x)\right)\bar{F}(x)dx$, the cumulative residual entropy (CRE) of X due to Rao et al. (2004).

EXAMPLE 2. *In survival studies, the most widely used semi-parametric model is the proportional hazard rate model (PHM). When X and Y satisfy PHM,*

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

or equivaently

$$\bar{F}_Y(x)=\left(\bar{F}_X(x)\right)^{\theta},\ \theta>0,$$

where $\bar{F}_X(x)$, $\bar{F}_Y(x)$, $h_X(x)=\frac{f_X(x)}{\bar{F}_X(x)}$ and $h_Y(x)=\frac{f_Y(x)}{\bar{F}_Y(x)}$ denote respectively the survival functions and hazard rate functions of X and Y, with probability density function $g_Y(x)=\theta\left(\bar{F}_X(x)\right)^{\theta-1}f_X(x)$. Then the information volatility of Y is obtained as

$$\begin{aligned}\eta(Y) &= Var(-\log g(Y))=Var(\log g(Y))\\ &= E\left(\log g(Y)-E\left(\log g(Y)\right)\right)^2\\ &= \eta(X)+(\theta-1)^2-2(\theta-1)Cov\left(\log f(X),\log\bar{F}(X)\right).\end{aligned} \tag{3}$$

In particular, when X follows standard exponential distribution, then

$$\eta(Y)=\eta(X)+(\theta-1)(\theta-3).$$

$\eta(Y)$ has a more meaningful explanation in a series system. Consider a series system consisting of n components with i.i.d lifetimes X_i, $i=1,2,...,n$. Then the lifetime of the system is $Y=min(X_1,X_2,...,X_n)$, with probability density function $g=n(\bar{F})^{n-1}f$ and survival function $\bar{G}=(\bar{F})^n$. Clearly, X_i and Y satisfy Cox PH model. Then the information volatility of the series system is $\eta(Y)$ in (3), where $\theta>0$ is replaced by $n\geq 1$, an integer.

Liu (2007) showed that unlike the Shannon entropy $\xi(X)$, the information volatility function $\eta(X)$ possesses some interesting mathematical properties. For example, $\eta(X)$ is invariant under linear transformation, while $\xi(X)$ is only shift invariant. Also, $\eta(X)$ is invariant under affine transformations, implying that it is independent of the mean and variance of X while $\xi(X)$ is only independent of the mean. Therefore, if a family of distributions can be described solely by the mean and the variance, then its $\eta(X)$ is a constant for the whole family. For uniform distribution $\eta(X)=0$, normal $\eta(X)=\frac{1}{2}$, and for exponential distribution $\eta(X)=1$. Further, Liu (2007) proved that the Shannon entropy, $E(-\log f(X))$ is not a 'good' statistic in the sense that it is not possible to find discrete distributions converging weakly to the distribution in such a way that the corresponding Shannon entropies converge, whereas the information volatility $\eta(X)$ has this convergence property. For more properties of $\eta(X)$, we refer to Liu (2007). Motivated with these, in the present note we further study various properties of $\eta(X)$ in the context of reliability modelling.

The paper is organized as follows. In Section 2 we introduce a dynamic information volatility function and studied its basic properties. In Section 3, we prove some ageing properties and characterizations based on it. Concluding remarks are reported in Section 4.

2. Dynamic residual information volatility function

In reliability and life testing, data are generally truncated, where one has information about the current age of the component under consideration. Clearly, $\eta(X)$ in (2) is not appropriate to measure the information volatility, and hence modify it to take the current age into account so that the modified measure becomes a function of current age t and thus dynamic. If X represents the lifetime of a component or system, then dynamic residual information volatility (DRIV) function is defined as follows.

DEFINITION 3. *For a non-negative random variable X, the dynamic residual information volatility based on the residual random variable $X_t=(X-t|X>t)$, $t>0$ is defined as*

$$\begin{aligned}\eta(t) &= \eta(X;t)=Var(-\log f(X_t))\\ &= E(-\log f(X_t))^2-(E(-\log f(X_t)))^2\\ &= \int_t^{\infty}\left(\log\frac{f(x)}{\bar F(t)}\right)^2\frac{f(x)}{\bar F(t)}dx-(\xi(t))^2,\end{aligned} \tag{4}$$

where $\xi(t)=E(-\log f(X_t))=-\frac{1}{\bar F(t)}\int_t^{\infty}(\log f(x))f(x)dx+\log\bar F(t)$ denote the expected residual entropy of X_t due to Ebrahimi (1996) based on (1).

Eq. (4) further simplifies to

$$\eta(t)=\frac{1}{\bar F(t)}\int_t^{\infty}(\log f(x))^2 f(x)dx-\left(\xi(t)-\log\bar F(t)\right)^2. \tag{5}$$

$\eta(t)$ in (5) measures the information volatility of the residual random variable X_t. Clearly, when $t\to 0$, $\eta(t)$ reduces to $\eta(X)$.

From (5), we also have

$$\eta(t) \leq \frac{1}{\bar{F}(t)} \int_t^{\infty} (\log f(x))^2 f(x) dx.$$

EXAMPLE 4. *When X follows Uniform (a,b), then $\eta(t)=0$.*

REMARK 5. *$\eta(t)=0$ for Uniform (a,b) implies that dynamic information volatility can be taken as a measure of variability in the random variables. In other words, deviation of $\eta(t)$ from 0 produces volatility of random variables, thus $\eta(t)$ moves away from 0 indicates more volatile.*

EXAMPLE 6. *When X follows exponential with probability density function $f(x) = \lambda e^{-\lambda x}; x \geq 0, \lambda > 0$, we have $\eta(t) = 1 = \eta(X)$.*

EXAMPLE 7. *Suppose X follows Pareto II distribution with survival function $\bar{F}(x) = (1+ax)^{-c}; x \geq 0; c, a > 0$. Then $\eta(t) = \left(\frac{c+1}{c}\right)^2 > 1$.*

EXAMPLE 8. *When X is beta (finite range) distribution with survival function $\bar{F}(x) = (1-px)^d, 0 \leq x \leq \frac{1}{p}, p, d > 0$,, we get $\eta(t) = \left(\frac{d-1}{d}\right)^2 < 1$.*

REMARK 9. *It is to be noted that for the exponential distribution with parameter λ, $\xi(t) = 1 - \log\lambda$, independent of t, while $\eta(t) = 1$, independent of both λ and t. In the case of Pareto II given in Example 7 and beta (finite range) given in Example 8, $\xi(t)$ assume log-linear functions of t involving both the parameters, however, $\eta(t)$ takes only constant values in terms of its shape parameters. Thus, even if the residual entropy function of Pareto II and beta densities takes log-linear behaviour, the dynamic residual information volatility function $\eta(t)$ assumes constant values, and greater (less) than unity. Thus $\eta(t)$ provides as a quantitative measure in determining the volatility of a random variable.*

To illustrate the role of dynamic behaviour of information volatility (4), we propose the following example.

EXAMPLE 10. *Let X have Pareto I distribution with survival function $\bar{F}(x) = x^{\beta}$, $x > 1$, $\beta > 0$. Then the residual entropy function $\xi(t) = \log t + \frac{(\beta+1)}{\beta} - \log\beta$ and information volatility function using (2), $\eta(X) = 0$, while its dynamic information volatility based on (4) becomes*

$$\begin{aligned} \eta(t) = \quad & (\log\beta + \beta\log t)^2 + \left(\frac{(\beta+1)}{\beta}\right)^2 \left[\beta^2 (\log t)^2 + 2\beta \log t + 1\right] \\ & - \frac{2(\beta+1)}{\beta}(\log\beta + \beta\log t)(\beta\log t + 1) - \left(\log t + \frac{(\beta+1)}{\beta} - \log\beta\right)^2, \end{aligned}$$

a function of t. The graph of $\xi(t)$ and $\eta(t)$ given in Figures 1, 2, and 3 depict the behaviour of $\eta(t)$ and $\xi(t)$ functions with respect to different values of β for Pareto I model.

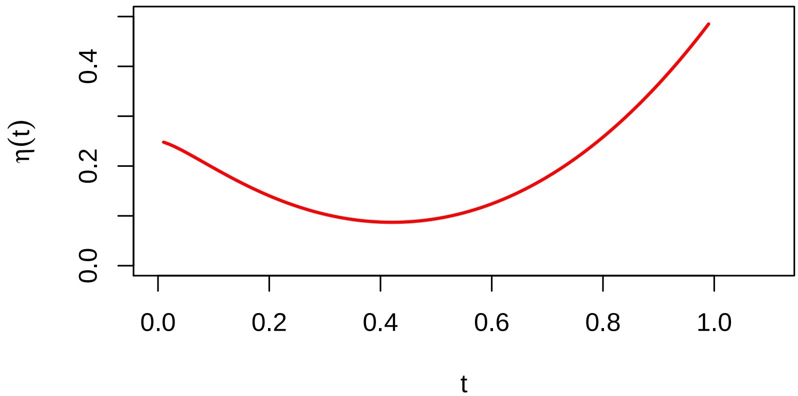

Figure 1 – Plot of $\eta(t)$ for Pareto I model with $\beta = 2$.

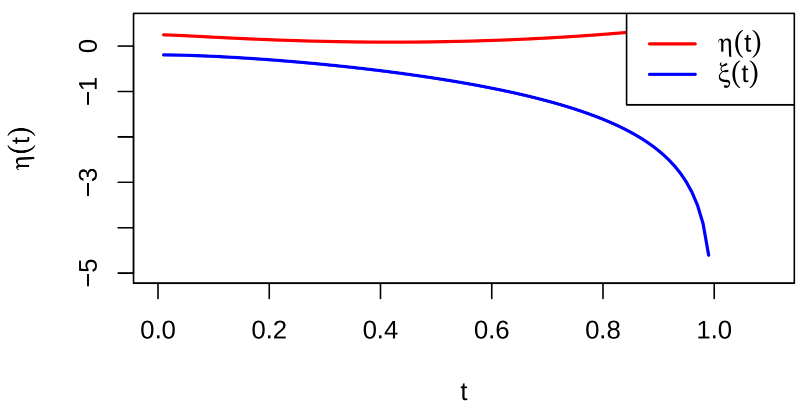

Figure 2 – Plot of $\xi(t)$ and $\eta(t)$ for Pareto I model with $\beta = 2$.

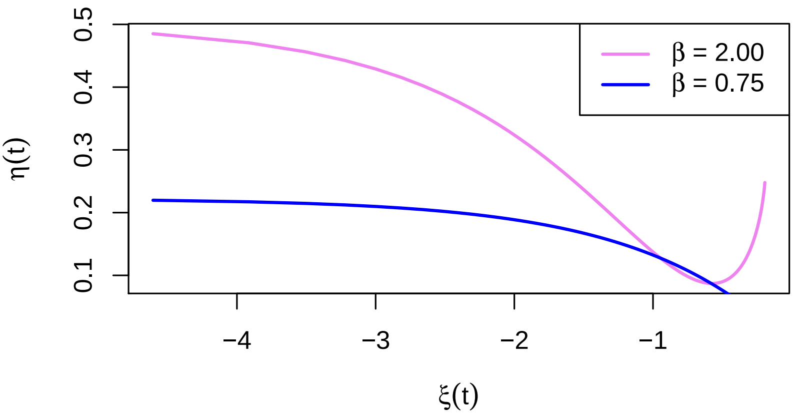

Figure 3 – Plot of $\eta(t)$ against $\xi(t)$ for Pareto I model with $\beta = 0.75, 2$.

It is well known that Shannon entropy has additive property. That is, when the random variables X and Y are independent, for the two-dimensional version of (1), we have

$$\xi(X,Y) = -E(\ln f(X,Y)) = \xi(X) + \xi(Y).$$

The dynamic information volatility function possess the same property. When X and Y are independent, for the two-dimensional version of (4),

$$\eta_{X_s,Y_t}(s,t) = \text{Var}\left(-\ln f_{X_s,Y_t}(s,t)\right) = \eta_{X_s}(s) + \eta_{Y_t}(t).$$

In the following theorem we obtain the expression for $\eta(X)$ for a monotone transformation.

THEOREM 11. *If $\phi(.)$ is an increasing and differentiable function, then*

$$\begin{aligned} \eta(\phi(X);t) = \quad & \eta\left(X;\phi^{-1}(t)\right) + Var\left(\log\phi'(X)|X > \phi^{-1}(t)\right) \\ & -2Cov\left(\log f_{X_{\phi^{-1}(t)}}(X), \log\phi'(X)|X > \phi^{-1}(t)\right), \end{aligned} \tag{6}$$

where $f_{X_{\phi^{-1}(t)}}(x) = \frac{f(x)}{\bar{F}_X(\phi^{-1}(t))}$.

PROOF. Let $Y=\phi(X)$. Then $f_Y(y)=f_X(x)\left|\frac{dx}{dy}\right|=\frac{f_X(\phi^{-1}(y))}{\phi'(\phi^{-1}(y))}$. Using (4), $\eta(t)$ of Y becomes

$$\begin{aligned}\eta(Y;t) &= \frac{1}{\bar{F}_X(\phi^{-1}(t))}\int_t^\infty\left(\log\frac{f(\phi^{-1}(x))}{\phi'(\phi^{-1}(x))\bar{F}_X(\phi^{-1}(t))}\right)^2\frac{f(\phi^{-1}(x))}{\phi'(\phi^{-1}(x))}dx\\ &\quad-(\xi_Y(t))^2\\ &= \frac{1}{\bar{F}_X(\phi^{-1}(t))}\int_{\phi^{-1}(t)}^\infty\left(\log\frac{f(y)}{\phi'(y)\bar{F}_X(\phi^{-1}(t))}\right)^2 f(y)dy-(\xi_Y(t))^2,\\ &= \int_{\phi^{-1}(t)}^\infty\left(\log f_{X_{\phi^{-1}(t)}}(y)-\log\phi'(y)\right)^2 f_{X_{\phi^{-1}(t)}}(y)dy-(\xi_Y(t))^2,\end{aligned} \tag{7}$$

where $\xi_Y(t)=\int_{\phi^{-1}(t)}^\infty\left(\log f_{X_{\phi^{-1}(t)}}(y)-\log\phi'(y)\right)f_{X_{\phi^{-1}(t)}}(y)dy$ is the entropy of Y. Now expanding (7) and simplifying we obtain (6). □

COROLLARY 12. *When $\phi(X)=aX+b$, then (6) reduces to*

$$\eta(aX+b;t)=\eta\left(X;\frac{t-b}{a}\right)-2\log a\,\xi\left(X;\frac{t-b}{a}\right).$$

COROLLARY 13. *For $\phi(.)$ defined in Theorem 11, it is yielded the following bound*

$$\eta\left(X;\phi^{-1}(t)\right)-\eta(\phi(X);t)\le 2Cov\left(\log f_{X_{\phi^{-1}(t)}}(X),\log\phi'(X)|X>\phi^{-1}(t)\right).$$

3. PROPERTIES OF $\eta(t)$

In this section, we prove some results on certain ageing classes and characterizations of distributions based on $\eta(t)$.

Differentiating (5) with respect to t gives,

$$\eta'(t)+2\xi'(t)\left(\xi(t)-\log\bar{F}(t)\right)$$

$$=h(t)\left[\eta(t)-(\log f(t))^2+\left(\xi(t)-\log\bar{F}(t)-1\right)^2-1\right]. \tag{8}$$

But due to Ebrahimi (1996), $\xi'(t)=h(t)(\xi(t)+\log h(t)-1)$, then (8) further simplifies to

$$\eta'(t)=h(t)\left(\eta(t)-(\xi(t)+\log h(t))^2\right). \tag{9}$$

Definition 14. *$\bar{F}$ has increasing (decreasing) uncertainty of residual life IURL (DURL) if $\xi(t)$ is increasing (decreasing) in t (Ebrahimi, 1996).*

Definition 15. *$\bar{F}$ has increasing (decreasing) residual information volatility IRIV (DRIV) if $\eta(t)$ is increasing (decreasing) in t.*

In some situations it will be difficult to compute the $\eta(t)$, however, bounds of the same can be obtained. The following theorem proves to this effect.

Theorem 16. *If $\bar{F}$ is IRIV (DRIV), then $\eta(t) \geq (\leq)(\xi(t) + \log h(t))^2$.*

Corollary 17. *When $\bar{F}$ is both IRIV (DRIV) and DFR (IFR), then we obtain a more sharper bound than in Theorem 16, given by $\eta(t) \geq (\leq) 1$*

Proof. It has been shown that IFR (DFR) implies DURL (IURL) (Ebrahimi, 1996). Therefore, when $\bar{F}$ is DFR (IFR), $\xi(t) + \log h(t) \geq (\leq) 1$, hence the result. □

Remark 18. *When $\bar{F}$ is both IRIV and DRIV, $\eta(t) = 1$, for the exponential distribution. Thus exponential distribution is the boundary of IRIV and DRIV classes.*

Theorem 19. *IFR $\Rightarrow$ IRIV. But the dual class need not imply.*

Proof. When X is IFR, it is DURL and hence $\xi(t) + \log h(t) \leq 1$, then (9) becomes,

$$\eta(t)h(t) \leq \eta'(t) \tag{10}$$

implies that $\eta'(t) \geq 0$ as both $\eta(t)$ and $h(t)$ non-negative, thus IRIV. However, when X is IURL, reversing the inequality in (10), which in turn gives $\eta'(t) \leq \eta(t)h(t)$. Since $\eta(t)h(t) \geq 0$, one cannot conclude to DRIV. □

The following is a characterization to $\eta(t)$ connecting some well-known lifetime models.

Theorem 20. *The relationship X, $\eta(t) = C$ holds if and only if X follows Pareto II, exponential and finite range distributions according as $C \geq = \leq 1$.*

Proof. Assume that $\eta(t) = C$ holds. Then using (7) we have

$$(\xi(t) + \log h(t))^2 = C,$$

or equivalently

$$\xi(t) + \log h(t) = K,$$

where $K = C^{\frac{1}{2}}$, is a characterization to the required models (see Sankaran and Gupta, 1999). □

The following theorem establishes a relationship for $\eta(t)$ in terms of variance of $h(t)$ and covariance function of $h(t)$ and entropy function $H(X)$.

THEOREM 21. *For the random variable X, $\eta(t)$ admits the relationship*

$$\begin{aligned}\eta(t) &= Var(\log h(X)|X>t)+2Cov(\log h(X),H(X)|X>t)\\ &\quad +Var(H(X)|X>t)\\ &= Var(\log h(X)|X>t)+2Cov(\log h(X),H(X)|X>t)+1\end{aligned}$$

where $H(t)=\log\bar{F}(t)$ denote the cumulative hazard rate of X.

PROOF. The proof directly follows from the relationship, $\log f(t) = \log h(t) + \log\bar{F}(t)$. □

Next we introduce a new dynamic measure of information volatility based on the conditional variance. It is defined as

DEFINITION 22. *For a non-negative random variable X, the conditional information volatility is defined as*

$$\begin{aligned}\psi(t) &= Var(-\log f(X)|X>t)\\ &= E\left((-\log f(X)|X>t)^2\right)-(\chi(t))^2\\ &= \frac{1}{\bar{F}(t)}\int_t^\infty (\log f(x))^2 f(x)dx-(\chi(t))^2,\end{aligned}$$

where $\chi(t) = E(-\log f(X)|X>t) = -\frac{1}{\bar{F}t(t)}\int_t^\infty(\log f(x))f(x)dx$ is the conditional measure of uncertainty due to Sankaran and Gupta (1999).

REMARK 23. *It is to be noted that the conditional information volatility function $\psi(t)$ coincides the dynamic information volatility function $\eta(t)$, since $\chi(t)=\xi(t)-\log\bar{F}(t)$.*

Now based on Definition 22, we prove the following characterization result to a general family of distributions. Let $\mathscr{A}$ be a class of absolutely continuous non-constant functions $-\log f(X)$ with derivatives $-\frac{f'(X)}{f(X)}$ defined on the range of X.

THEOREM 24. *For $-\log f(X)\in\mathscr{A}$ satisfying $E\left(g(X)|\frac{f'(X)}{f(X)}|\right)<\infty$,*

$$\psi(t)=(\psi(X))^{\frac{1}{2}}E\left[-\frac{f'(X)}{f(X)}g(X)|X>t\right]+(\chi(X)-\chi(t))(\chi(t)-\log f(t)) \tag{11}$$

if

$$\chi(t)=\xi(X)+(\psi(X))^{\frac{1}{2}}h(t)g(t) \tag{12}$$

or equivalently $f(t)$ satisfies the differential equation

$$\frac{f'(t)}{f(t)} = -\frac{g'(t)}{g(t)} + \frac{\chi(X) - \chi(t)}{(\psi(X))^{\frac{1}{2}}\, g(t)}, \tag{13}$$

holds, where $g(.)$ is a positive Borel-measurable function satisfying the condition (12). Conversely, if there exists a function $-\log f(X)$ for which $\frac{f'(t)}{f(t)} \neq 0$ for all $t > 0$ satisfying (11), then (12) and (13) hold.

PROOF. The proof directly follows from Corollary 2.1 of Nair and Sudheesh (2010), replacing $h(X) = -\log f(X)$. □

4. Concluding Remarks

In this paper, we have introduced a dynamic version of the information volatility function proposed by Liu (2007) and studied its various properties. In particular, we have investigated the usefulness of the measure in the context of the reliability theory. Bounds and relationships with some important reliability ageing classes have been obtained. Some characterization relationships between dynamic information volatility function and certain lifetime distributions were also proved.

References

J. Beirlant, E. J. Dudewicz, L. Györfi, E. C. van der Meulen (1997). *Nonparametric entropy estimation: An overview*. International Journal of Mathematical and Statistical Sciences, 6, no. 1, pp. 17–39.

N. Ebrahimi (1996). *How to measure uncertainty in the residual life time distribution*. Sankhyā A, pp. 48–56.

D. E. Lake (2009). *Nonparametric entropy estimation using kernel densities*. Methods in Enzymology, 467, pp. 531–546.

J. Liu (2007). *Information Theoretic Content and Probability*. Ph.D. thesis, University of Florida.

N. U. Nair, K. K. Sudheesh (2010). *Characterization of continuous distributions by properties of conditional variance*. Statistical Methodology, 7, no. 1, pp. 30–40.

M. Rao, Y. Chen, B. C. Vemuri, F. Wang (2004). *Cumulative residual entropy: A new measure of information*. IEEE Transactions on Information Theory, 50, no. 6, pp. 1220–1228.

P. G. SANKARAN, R. P. GUPTA (1999). *Characterization of lifetime distributions using measure of uncertainty*. Calcutta Statistical Association Bulletin, 49, no. 3-4, pp. 159–166.

C. E. SHANNON (1948). *A mathematical theory of communication*. Bell System Technical Journal, 27, no. 3, pp. 379–423.

SUMMARY

Liu (2007) discussed a new measure, known as the information volatility function to study the variability of the uncertainty contained in a probability distribution. In the present paper, we extend this concept to the residual random variable, a dynamic information volatility function and study its usefulness in reliability modelling. Different ageing and characterization properties of dynamic information volatility function are also derived.

Keywords: Shannon entropy; Information volatility; Reliability measures; Stochastic orders; Characterization.