

On Discrete Residual and Past Measures of Inaccuracy

Vikas Kumar, Richa Thapliyal, and H.C. Taneja

*Department of Applied Mathematics,
Delhi Technological University,
Bawana Road, Delhi-110042, India*

vikas_iitr82@yahoo.co.in, richa31aug@gmail.com, hctaneja@rediffmail.com

Received 14 April 2012

Accepted 27 February 2013

The present communication considers a discrete dynamic measure of inaccuracy between two residual and past lifetime distributions. Under the assumption that the true distribution F and reference distribution G satisfy the proportional hazard model (PHM) and proportional reversed hazard model (PRHM), it has been shown that the proposed measures determines the lifetime distribution uniquely. A relation between Kerridge inaccuracy, residual inaccuracy and past inaccuracy has been explored.

Keywords: Kerridge inaccuracy, dynamic inaccuracy, residual entropy, residual lifetime, 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 as

$$H(f) = - \int_0^\infty f(x) \log f(x) dx, \quad (1.1)$$

and

$$H(f; g) = - \int_0^\infty f(x) \log g(x) dx, \quad (1.2)$$

respectively. In case $g(x) = f(x)$, then (1.2) reduces to (1.1).

The measures (1.1) and (1.2) are not applicable to a system which has survived for some unit of time. Ebrahimi (1996) considered the entropy of the residual lifetime $X_t = [X - t \mid X > t]$ as a

residual measure of uncertainty given by

$$H(f; t) = - \int_t^\infty \frac{f(x)}{\bar{F}(t)} \log \frac{f(x)}{\bar{F}(t)} dx. \quad (1.3)$$

Extending the information measure (1.3), a residual measure of inaccuracy, refer to Taneja *et al.* (2009), is given as

$$H(f, g; t) = - \int_t^\infty \frac{f(x)}{\bar{F}(t)} \log \frac{g(x)}{\bar{G}(t)} dx. \quad (1.4)$$

Here $g(x)$ can be interpreted as some reference p.d.f., whereas $f(x)$ is the baseline p.d.f. in context with a system under consideration. Further, the residual measure of inaccuracy given by (1.4) has been extended to weighted (length biased) residual inaccuracy measure, refer to Kumar *et al.* (2010).

There are many situations where a continuous time is inappropriate for describing the lifetime of devices and other systems, refer to Xekalaki (1983). Some specific situations worth mentioning are: (i) actuaries and bio-statisticians are interested in the lifetimes of persons or organisms, measured in months, weeks, or days, (ii) in case of equipment operating in cycles the random variable of interest is the successful number of cycles before the failure. For instance, the number of flashes in a car-flasher prior to failure of the device, (iii) a discrete life distribution is a natural choice where failure occurs only due to incoming shocks, for example, in weapons reliability, the numbers of rounds fired until failure is more important than age at failure. Further discrete lifetime also occurs through grouping or finite measurement of continuous time phenomena. Since there is a limit on the precision of any measurement, it can be arguably said that sample from a continuous distribution exist only in theory.

Let X be a discrete lifetime random variable taking values on $N = \{1, 2, 3, \dots, n\}$, $1 \leq n < \infty$, with probability $P(X = i) = p_i$, $0 \leq p_i \leq 1$, $\sum_{i=1}^n p_i = 1$. Let $F(i) = P(X \leq i) = \sum_{j=0}^{j=i} p_j$, be the lifetime distribution function and $\bar{F}(i) = 1 - F(i-1)$, $i = 1, 2, \dots$ the corresponding survival function. Shannon (1948) introduced the measure of information (or, uncertainty) associated with X given as

$$H(P) = - \sum_{i=1}^n p_i \log p_i. \quad (1.5)$$

Next, consider a situation where an experimenter needs to predict the probabilities of various outcomes in an experiment. Let Y be the reference (or, predicted) random variable taking values on $N = \{1, 2, 3, \dots, n\}$, $1 \leq n < \infty$ with probability $P\{Y = i\} = q_i$, $0 \leq q_i \leq 1$, $\sum_{i=1}^n q_i = 1$, while the actual distribution is given by $P\{X = i\} = p_i$, that is the experimenter assumes that the probability of the i^{th} outcome is q_i whereas the true probability is p_i . Then the average inaccuracy in the statement of the observer, as proposed by Kerridge (1961), can be measured by

$$H(P, Q) = - \sum_{i=1}^n p_i \log q_i. \quad (1.6)$$

The modification of Shannon's entropy as a measure of uncertainty in residual lifetime distribution for discrete random variable has drawn attention of many researchers. Rajesh and Nair (1998) have

proposed

$$H(P; j) = - \sum_{k=j}^{\infty} \frac{p(k)}{\bar{F}(j)} \log \frac{p(k)}{\bar{F}(j)}, \quad (1.7)$$

as residual lifetime information measure in discrete domain and showed that $H(P; j)$ uniquely determines the distribution function $F(t)$ and characterized the same. Generalizing the concept of Rajesh and Nair (1998), Belzunce (2004) extended γ -entropy, refer to Ord *et al.* (1981), for the residual lifetime in discrete domain, and showed that $H_\gamma(P; j)$ uniquely determines the lifetime distribution. Nanda and Paul (2006) have extended this to entropy of order β and have studied their properties. In this communication we propose and study the measures of dynamic (residual and past both) inaccuracy, when the distributions P and Q are discrete domain, and satisfy *proportional hazard model* (PHM), refer to Cox (1959), or *proportional reversed hazard model* (PRHM), refer to Gupta and Gupta (2007). In Section 2, we consider the concept of PHM and PRHM in case of discrete domain. In Section 3, we propose and characterize the residual inaccuracy measure under proportional hazard model (PHM). In Section 4, we propose and characterize the past inaccuracy measure under proportional reversed hazard model (PRHM).

2. Discrete Case Hazard Models

The hazard rate and the reverse hazard rate functions for a random variable X with non-negative integral support are defined by

$$h(i) = \frac{p(i)}{\bar{F}(i)}, \quad i = 1, 2, \dots \quad (2.1)$$

and

$$r(i) = \frac{p(i)}{F(i)}, \quad i = 1, 2, \dots \quad (2.2)$$

respectively. In the discrete setup both the hazard rate and the reversed hazard rate can be interpreted as a probability which is not the case in the continuous case. If X represents the lifetime of a component then $h(i)$ is the probability that the component will fail at time $X = i$ given that it has survived up to the time before i . Similarly, $r(i)$ is the probability that the component will fail at time $X = i$, given it is known to have failed before i . But many of the properties of the hazard rate and reversed hazard rate which hold in the continuous case do not hold in the discrete case. Specifically the definitions (2.1) and (2.2) don't lead respectively to PHM and PRHM in discrete domain. Xie, Gaudoin and Bracquemond (2002) have redefined the hazard rate and reversed hazard rate functions in discrete setup as follows:

$$\lambda(i) = \log \frac{\bar{F}(i-1)}{\bar{F}(i)}, \quad i = 1, 2, \dots \quad (2.3)$$

and

$$\mu(i) = \log \frac{F(i)}{F(i-1)}, \quad i = 1, 2, \dots \quad (2.4)$$

It is easy to verify that the functions defined by (2.1) and (2.3) satisfy

$$h(i) = 1 - \exp[-\lambda(i)], \quad (2.5)$$

and defined by (2.2) and (2.4) lead to

$$r(i) = 1 - \exp[-\mu(i)], \quad \forall i = 1, 2, \dots \quad (2.6)$$

Further we note that the both hazard rates (as well as reversed hazard rates) preserve the same monotonicity property.

It is easy to verify that if X and Y are two discrete random variables with non-negative integral support with probability distribution functions F and G respectively, then the hazard rate function defined by (2.3) satisfy the PHM

$$\overline{G}(i) = \{\overline{F}(i)\}^\beta, \quad \beta > 0, \quad (2.7)$$

equivalently

$$\lambda^*(i) = \beta \lambda(i); \quad (2.8)$$

and the reversed hazard rate function defined by (2.4) satisfy the PRHM

$$G(i) = \{F(i)\}^\beta, \quad \beta > 0 \quad (2.9)$$

equivalently

$$\mu^*(i) = \beta \mu(i). \quad (2.10)$$

3. Discrete Residual Inaccuracy Measure and Characterization Problem

In sequel to measure of inaccuracy (1.6), we propose

$$H(P, Q; j) = - \sum_{k=j}^{\infty} \frac{p(k)}{\overline{F}(j)} \log \frac{q(k)}{\overline{G}(j)}, \quad (3.1)$$

as a discrete residual inaccuracy associated with two residual lifetime distributions P and Q . When $j = 0$, then (3.1) becomes (1.6), and when $p(k) = q(k), \forall k$ then (3.1) becomes (1.7), the discrete measure of uncertainty.

Example 3.1. Consider a random variable X with probability mass function (pmf)

$$P(X = i) = p_i = \frac{2i}{n(n+1)}, \quad i = 1, 2, \dots, n,$$

and another random variable Y with pmf

$$P(Y = i) = q_i = \frac{1}{n}, \quad i = 1, 2, \dots, n.$$

Then we have survival function corresponding to X and Y given as

$$\begin{aligned} \overline{F}(i) &= 1 - F(i-1) = 1 - \frac{i(i-1)}{n(n+1)}, \\ \overline{G}(i) &= 1 - G(i-1) = \frac{n-i+1}{n}, \end{aligned}$$

respectively. Substituting these in (3.1) and simplifying, we obtain the discrete residual inaccuracy measure

$$H(P, Q; j) = \log(n - j + 1) \left\{ \frac{(n + j)(n - j + 1)}{n(n + 1) - j(j - 1)} \right\}. \quad (3.2)$$

Theorem 3.1. *Let X and Y be two discrete random variables with non-negative integral support satisfying the proportional hazard model (2.7). If $H(P, Q; j)$ is increasing and finite, then $H(P, Q; j)$ uniquely determines the survival function $\bar{F}(j)$ of the random variable X .*

Proof. We have

$$H(P, Q; j) = - \sum_{k=J}^{\infty} \frac{p(k)}{\bar{F}(j)} \log \frac{q(k)}{\bar{G}(j)},$$

which is equivalent to

$$\sum_{k=J}^{\infty} p(k) \log q(k) = \bar{F}(j) \log \bar{G}(j) - \bar{F}(j) H(P, Q; j). \quad (3.3)$$

For $j + 1$ we obtain

$$\sum_{k=J+1}^{\infty} p(k) \log q(k) = \bar{F}(j+1) \log \bar{G}(j+1) - \bar{F}(j+1) H(P, Q; j+1). \quad (3.4)$$

Denoting $H(P, Q; j)$ by $H(j)$, and subtracting (3.3)-(3.4), we get

$$p(j) \log q(j) = \bar{F}(j) \{ \log \bar{G}(j) - H(j) \} - \bar{F}(j+1) \{ \log \bar{G}(j+1) - H(j+1) \}.$$

Substituting $p(j) = \bar{F}(j) - \bar{F}(j+1)$, and $q(j) = \bar{G}(j) - \bar{G}(j+1)$, we obtain

$$\begin{aligned} & \{ \bar{F}(j) - \bar{F}(j+1) \} \log \{ \bar{G}(j) - \bar{G}(j+1) \} \\ &= \bar{F}(j) [\log \bar{G}(j) - H(j)] - \bar{F}(j+1) \{ \log \bar{G}(j+1) - H(j+1) \}. \end{aligned} \quad (3.5)$$

Dividing by $\bar{F}(j)$ both side and simplifying we get

$$\begin{aligned} & \left\{ \frac{\bar{F}(j+1)}{\bar{F}(j)} \right\} \log \frac{\bar{G}(j+1)}{\bar{G}(j)} + \left\{ 1 - \frac{\bar{F}(j+1)}{\bar{F}(j)} \right\} \log \left\{ 1 - \frac{\bar{G}(j+1)}{\bar{G}(j)} \right\} \\ &= -H(j) - \frac{\bar{F}(j+1)}{\bar{F}(j)} H(j+1). \end{aligned} \quad (3.6)$$

If $e^{-\lambda_j} = \frac{\bar{F}(j+1)}{\bar{F}(j)}$, $e^{-\lambda_j^*} = \frac{\bar{G}(j+1)}{\bar{G}(j)}$, then (3.6) becomes

$$-\lambda_j^* e^{-\lambda_j} + (1 - e^{-\lambda_j}) \log(1 - e^{-\lambda_j^*}) = -H(j) + e^{-\lambda_j} H(j+1). \quad (3.7)$$

Using the proportional hazard model (2.8), (3.7) can be written as

$$(1 - e^{-\lambda_j}) \log(1 - e^{-\beta\lambda_j}) + H(j) - \beta\lambda_j e^{-\lambda_j} - e^{-\lambda_j} H(j+1) = 0. \quad (3.8)$$

It is noted that

$$g(x) = (1 - x) \log(1 - x^\beta) + H(j) + \beta x \log x - x H(j+1) = 0 \quad (3.9)$$

has one root $x = e^{-\lambda_j}$ for fixed j . In this case $g(0) = H(j) \geq 0$, $g(1) = H(j) - H(j+1) \leq 0$, and

$$g'(x) = -H(j+1) + \log \frac{x^\beta}{1-x^\beta} - \frac{\beta x^{\beta-1}(1-x)}{(1-x^\beta)} + \beta.$$

Therefore, $g(x)$ first decreases and then increases in $(0, 1)$, with a minimum at

$$x_j = \left\{ \frac{1}{1 + \exp[-H(j+1) - \frac{\beta x^{\beta-1}(1-x)}{(1-x^\beta)} + \beta]} \right\}^{\frac{1}{\beta}},$$

which implies that equation (3.8) has a unique positive solution $e^{-\lambda_j}$ in $(0, 1)$ for all j . Finally, since $1 - e^{-\lambda_j} = h(j)$, where $h(j) = \frac{p(j)}{F(j)}$ is the discrete failure rate, thus $H(P, Q; j)$ uniquely determines the distribution function F . \square

4. Discrete Past Inaccuracy Measure and Characterization Problem

In many realistic situations uncertainty is not necessarily related to the future but can also refer to the past. In this situation the random variable of interest is $iX = [i - X \mid X \leq i]$, known as the *inactivity time*. This is because, once at time X the system fails, and at time i it is observed to be in a failure state, the random time for which the system was down is iX , where X is a discrete random variable. Based on this idea Di Crescenzo and Longobardi (2002) have studied the past entropy given by

$$H^*(P; j) = - \sum_{k=0}^j \frac{p(k)}{F(j)} \log \frac{p(k)}{F(j)}. \quad (4.1)$$

Here we propose a discrete dynamic measure of inaccuracy based on the past entropy over $(0, j)$, defined as

$$H^*(P, Q; j) = - \sum_{k=0}^j \frac{p(k)}{F(j)} \log \frac{q(k)}{G(j)}, \quad (4.2)$$

When $p(k) = q(k)$, $\forall k$ then (4.2) becomes (4.1), the discrete dynamic measure of past entropy given by Di Crescenzo and Longobardi (2002).

Example 4.1. Reconsider the random variable X and Y as defined in Example 3.1, the distribution functions corresponding to X and Y are given by

$$F(j) = \sum_{i=1}^j p_i = \frac{j(j+1)}{n(n+1)},$$

and

$$G(j) = \sum_{i=1}^j q_i = \frac{j}{n},$$

respectively. Substituting these in (4.2) and simplifying, we obtain the discrete past inaccuracy measure

$$H^*(P, Q; j) = \log j. \quad (4.3)$$

Next, we study characterization problem for the discrete past inaccuracy measure under the proportional reversed hazard model. Sengupta *et al.* (1999) illustrated that PRHM leads to a better fit for some data sets than PHM.

Theorem 4.1. *Let X and Y be two non negative random variables with integral support satisfying the proportional reversed hazard model (PRHM) (2.10). Let $H^*(P, Q; j) < \infty$, be an increasing function of j , then $H^*(P, Q; j)$ uniquely determines the survival function $\bar{F}(j)$ of the variable X .*

Proof. From (4.2) we have

$$\sum_{k=0}^j p(k) \log q(k) = F(j) \log G(j) - F(j) H^*(P, Q; j). \quad (4.4)$$

By replacing j by $j + 1$ in (4.4), we have

$$\sum_{k=0}^{j+1} p(k) \log q(k) = F(j+1) \log G(j+1) - F(j+1) H^*(P, Q; j+1). \quad (4.5)$$

Denoting $H^*(P, Q; j)$ by $H^*(j)$, and subtracting (4.5) from (4.4), we get

$$p(j+1) \log q(j+1) = F(j+1) \{\log G(j+1) - H^*(j+1)\} - F(j) \{\log G(j) - H^*(j)\},$$

Writing $p(j+1) = F(j+1) - F(j)$ and $q(j+1) = G(j+1) - G(j)$, we obtain

$$\begin{aligned} & \{F(j+1) - F(j)\} \log \{G(j+1) - G(j)\} \\ &= F(j+1) [\log G(j+1) - H^*(j+1)] - F(j) \{\log G(j) - H^*(j)\}. \end{aligned} \quad (4.6)$$

which can be expressed as

$$(1 - e^{-\mu_j}) \log(1 - e^{-\mu_j^*}) - \mu_j^* e^{-\mu_j} = e^{-\mu_j} H^*(j) - H^*(j+1), \quad (4.7)$$

where $e^{-\mu_j} = \frac{F(j)}{F(j+1)}$ and $e^{-\mu_j^*} = \frac{G(j)}{G(j+1)}$.

Using the proportional reversed hazard model (2.10), then (4.7) becomes

$$(1 - e^{-\mu_j}) \log(1 - e^{-\beta\mu_j}) + H^*(j+1) - \beta\mu_j e^{-\mu_j} - e^{-\mu_j} H^*(j) = 0. \quad (4.8)$$

It is noted that

$$y_1(x) = (1 - x) \log(1 - x^\beta) + \beta x \log x - x H^*(j) + H^*(j+1) = 0, \quad (4.9)$$

has one root $x = e^{-\mu_j}$, for fixed j .

In this case $y_1(0) = H^*(j+1) \geq 0$, $y_1(1) = H^*(j+1) - H^*(j) \leq 0$ and

$$y'_1(x) = -H^*(j) + \log \frac{x^\beta}{1 - x^\beta} - \frac{\beta x^{\beta-1}(1-x)}{(1-x^\beta)} + \beta.$$

Therefore, $y_1(x)$ first decreases and then increases in $(0, 1)$, with a minimum at

$$x_j = \left\{ \frac{1}{1 + \exp[-H^*(j) - \frac{\beta x^{\beta-1}(1-x)}{(1-x^\beta)} + \beta]} \right\}^{\frac{1}{\beta}},$$

which implies that equation (4.9) has a unique positive solution μ_j in $(0, 1)$ for all j . Finally, since $1 - e^{-\mu_j} = r(j+1)$, where $r(j) = \frac{p(j)}{F(j)}$ is the discrete reversed hazard rate, thus $H^*(P, Q; j)$ uniquely determines the distribution function F . This complete the proof. \square

Remark 4.1. 1. We observe the following relation between the three inaccuracy measures (1.6), (3.1) and (4.2) as considered above.

$$H(P; Q) = \bar{F}(j+1)H(j+1) + F(j)H^*(j) + H[F(j), G(j)], \quad (4.10)$$

where $H[F(j), G(j)] = -F(j)\log G(j) - [1 - F(j)]\log[1 - G(j)]$, corresponds to the Kerridge inaccuracy (1961), and $H(p, 1-p) = -p\log p - (1-p)\log(1-p)$ is the entropy of a Bernoulli distribution.

2. For $\beta = 1$, value of x_j for residual and past inaccuracy reduces to

$$x_j = \left\{ \frac{1}{1 + e^{-H(j+1)}} \right\},$$

and

$$x_j = \left\{ \frac{1}{1 + e^{-H^*(j)}} \right\},$$

respectively, these results given by Belzunce *et al.* (2004) and Nanda and Paul (2006) in context with residual and past entropy in discrete domain.

5. Conclusion

So far, discrete life time distributions play only a marginal role in reliability analysis. However, real world life time is either observed in discrete time points or it is measured by discrete quantities. Therefore, the focus of reliability analysis should turn to the more realistic discrete life time distributions with finite support that so far are hardly investigated in academic reliability theory, despite the fact that in real world everything is finite.

The proportional hazards and reversed hazards model are very common in reliability analysis. Here we using these models relevant to situations when lifetimes are discrete and study the characterization problem of discrete dynamic inaccuracy measures.

References

- [1] A. Di Crescenzo, M. Longobardi, Entropy-based measure of uncertainty in past lifetime distributions, *Journal of Applied Probability*, **39** (2002) 434–440.
- [2] A.K. Nanda, P. Paul, Some properties of past entropy and their applications, *Metrika*, **64** (2006) 47–61.
- [3] A.K. Nanda, P. Paul, Some results on generalized residual entropy, *Information Sciences*, **176** (2006) 27–47.
- [4] B. Abraham, P.G. Sankaran, Renyi's entropy for residual life distribution, *Statistical Papers*, **46** (2005) 17–30.
- [5] C.E. Shannon, A mathematical theory of communications, *Bell System Technical Journal*, **27** (1948) 379–423.
- [6] D.F. Kerridge, Inaccuracy and inference, *Journal of Royal Statistical Society, Series B*, **23** (1961) 184–194.

- [7] D.R. Cox, The analysis of exponentially distributed lifetimes with two type of failure, *Journal of Royal Statistical Society, Series B*, **21** (1959) 411–421.
- [8] D. Sengupta, H. Singh, A.K. Nanda, *The proportional reversed hazards model*, Technical Report, Applied Statistics Division, Indian statistical Institute, Kolkata (1999).
- [9] E. Xekalaki, Hazard function and life distributions in discrete time, *Communications in Statistics: Theory and Methods*, **12** (1983) 2503–2509.
- [10] F. Belzunce, J. Navarro, J.M. Ruiz, Y. Aguila, Some results on residual entropy function, *Metrika*, **59** (2004) 147–161.
- [11] G. Rajesh, K.R.M. Nair, Residual entropy function in discrete time, *Far East Journal of Theoretical Statistics*, **2** (1998) 43–57.
- [12] H.C. Taneja, V. Kumar, R. Srivastava, A dynamic measure of inaccuracy between two residual lifetime distributions, *International Mathematical Forum*, **4** (25) (2009) 1213–1220.
- [13] J.K. Ord, G.P. Patil, C. Taillie, The choice of a distribution to describe personal incomes. In: Taillie C., Patil G.P., Baldessari, B.A. (eds.) *Statistical Distributions in Scientific Work*, Vol. 6, D. Reidel Publishing Co., Dordrecht-Boston, Massachusetts, (1981) pp. 193–201.
- [14] M. Xie, O. Gaudoin, C. Bracquemond, Redefining failure rate function for discrete distributions, *International Journal of Reliability Quality Safety Engineering*, **9** (2002) 275–285.
- [15] N. Ebrahimi, How to measure uncertainty about residual lifetime, *Sankhya A*, **58** (2002) 48–57.
- [16] R.C. Gupta, R.D. Gupta, Proportional reversed hazard rate model and its applications, *Journal of Statistical Planning and Inference*, **137** (2007) 3525–3536.
- [17] V. Kumar, H.C. Taneja, R. Srivastava Length biased weighted residual inaccuracy measure, *Metron*, **LXIII 2** (2010) 153–160.