Proportional reversed hazards model for discrete data

Deemat Mathew*
National Remote Sensing Centre, Nagpur, India - deematcm@gmail.com

Krishna Chaitanya
Presidency Uniersity, Kolkata, India

Abstract

Many times a product lifetime can be described through a non-negative integer valued random variable. In this paper, we propose a proportional reversed hazards model for discrete data analogous to the version for continuous data. Some ageing properties of the model are discussed. Stochastic comparison of pair of random variables that follow the model are also done. A new test based on U-statistics is developed for testing that the proportionality parameter in the proposed model is 1 and the asymptotic properties of the proposed test are studied.
Keywords: Discrete reliability, Proportional reversed hazards model, U-Statistics..

1 Introduction

The literature on reliability theory mainly deals with non-negative absolutely continuous random variables. However, quite often we come across situations where the product life can be described through non-negative integer valued random variables. For example, a piece of equipment operates in cycles and the experimenter observes the number of cycles successfully completed prior to failure. A frequently referred example is a Xerox machine whose life length would be the total number of copies it produces before the failure. Another example is the length of the hospital stay of patients who were hospitalized due to an accident (Nawata et al., 2009). The analysis of lifetime data in discrete time is also concerned with the data measured in intervals (Jiang, 2010). For recent discussions about discrete reliability, see Bracquemond and Gaudoin (2003), Nanda and Sengupta (2005), Chen and Manatunga (2007), Yu (2007), Bebbington et al. (2012),Khorashadizadeh et al. (2013), Lai (2013), Noughabi et al. (2013) and Almalki and Nadarajah (2014) and the references therein. So it is necessary to develop tools analogous to the continuous case, for studying the discrete lifetime data.
Nawata et al. (2009) analyzed the length of the hospital stay of patients hospitalized for cataract and related diseases. Since the length of stay may depend on several factors like sex, age, prevalence of other diseases like diabetes, mental strength etc. Nawata et al. (2009) considered a proportional hazards model (PHM) in discrete setup to incorporate such factors. Detailed discussions on the problems related to PHM can be found in Chen and Manatunga (2007).
As parallel to PHM model, the proportional reversed hazards model (PRHM) has become popular in recent times. For survey and some properties of PRHM in continuous domain, see Gupta and Gupta (2007) and references therein. Nanda and Paul (2003) developed a test for reversed hazard rates under the assumption of proportionality of reversed hazard rates. Recent work by Nanda and Das (2011) enlightened the role of PRHM models in modeling lifetime distributions which has potential application in system reliability.
Motivated by the above mentioned studies, we propose a PRHM for discrete data which is analogous to the continuous case using the modified definition of reversed hazard rate. The main advantage of the proposed model is that it can incorporate continuous covariates as well.
The rest of the article is organized as follows. In Section 2, propose a proportional reversed hazards model for discrete data parallel to the version for continuous data. We also discuss the preservation of ageing properties under this model. Stochastic comparison of pairs of random variables following this model are also given. In Section 3, we develop a non-parametric test for testing the proportionality constant of the proposed model. The asymptotic properties of the test statistic are also studied. Finally, in Section 4, we summarize our findings and mention some open problems.

2 Proportional Hazards Model

Let X be a discrete random variable with support $N = \{1, 2, ...\}$ or a subset thereof. Suppose $p(x) = P(X = x)$, $F(x) = P(X \le x)$ and $\bar{F}(x) = P(X > x)$ denote the probability mass function, the distribution function and the reliability function of X, respectively. Note that $\bar{F}(0) = 1$. The modified reversed hazard rate of X denoted by $\lambda^*(.)$ is defined as

$$\lambda^*(x) = \ln \frac{F(x)}{F(x-1)}, \quad x = 1, 2, \ldots. \tag{1}$$

Note that $\lambda^*(.)$ is not a probability but is additive for parallel system as in the continuous case. The cumulative reversed hazard rate function $\Lambda(.)$ is given by

$$\Lambda(x) = \sum_{k=x+1}^{\infty} \lambda^*(k) = -\ln F(x). \tag{2}$$

Hence the function $\lambda^*(.)$ determines the distribution of X uniquely.
Consider a parallel system with n independent and identically distributed components with lifetimes X_i, $i = 1, 2, ..., n$ and distribution function $F(.)$ and reversed hazard rate λ_F^*. Then the lifetime of the system $Z = \max(X_1, X_2, ..., X_n)$ has the distribution function $F_{(n)}(.)$ given by

$$F_{(n)}(x) = [F(x)]^n. \tag{3}$$

Then it can be seen that the reversed hazard rate corresponding to Z is proportional to that of X as

$$\lambda_{F_n}^*(x) = n\lambda_F^*(x).$$

Hence we can define the PRHM in discrete setup as follows. Let X and Y be two non negative random variables denoting the lifetime of two systems with distribution function $F(.)$ and $G(.)$ respectively. Also let $\bar{F}(x) = P(X > x)$ and $\bar{G}(x) = P(Y > x)$ be the respective reliability functions, and $\lambda_F^*(.)$ and $\lambda_G^*(.)$ the corresponding reversed hazard rates. Suppose X and Y are related through

$$G(x) = [F(x)]^\theta, \quad \theta > 0. \tag{4}$$

In view of the definition of reversed hazard rate given in (1), the model specified in (4) turns out to be the proportional reversed hazards model. That is, the reversed hazard rate $\lambda_G(.)$ corresponding to Y is proportional to that of X. The probability mass function of Y can be expressed as

$$\begin{aligned} g(y) &= G(y) - G(y-1) \\ &= [F(y)]^\theta - [F(y-1)]^\theta \\ &= [F(y)]^\theta \left[1 - \frac{F(y-1)^\theta}{F(y)^\theta}\right] \\ &= [F(y)]^\theta \left[1 - e^{-\theta\lambda^*(y)}\right]. \end{aligned} \tag{5}$$

Next we study the preservation of ageing properties under the proposed model (4). The following lemma will be useful in proving some of the results.

Lemma 1 *Let ϕ be a real function on R. Let x_1, x_2, y_1 and y_2 be such that $x_1 \le y_1 \le y_2$ and $x_1 \le x_2 \le y_2$.*
(i) If ϕ is convex , then

$$\frac{\phi(y_1) - \phi(x_1)}{y_1 - x_1} \le \frac{\phi(y_2) - \phi(x_2)}{y_2 - x_2}.$$

(ii) If ϕ is concave , then

$$\frac{\phi(y_1) - \phi(x_1)}{y_1 - x_1} \ge \frac{\phi(y_2) - \phi(x_2)}{y_2 - x_2}.$$

Theorem 1 *The random variable X is IRHR(DRHR) and the reversed hazard rate of Y is proportional to that of X, then Y is IRHR(DRHR).*

Theorem 2 *The random variable X is NBU(NWU) and the reversed hazard rate of Y is proportional to that of X, $\theta > 1$ ($\theta \lneq 1$) then Y is NBU (NWU).*

Proof: Suppose that X has NBU then for all $x, t \geq 0$ we have

$$R(x+t) \leq R(x)R(t).$$

That is

$$F(x+t) - F(x) \geq F(t) - F(x)F(t). \tag{6}$$

Let $x_1 = F(x)F(t)$, $x_2 = F(x)$, $y_1 = F(t)$ and $y_2 = F(x+t)$. Then, we have $x_1 \leq y_1 \leq y_2$ and $x_1 \leq x_2 \leq y_2$. Hence by Lemma 1, for $\phi(x) = x^\theta$ and $\theta > 1$ we have the inequality

$$\frac{[F(t)]^\theta - [F(x)F(t)]^\theta}{F(t) - F(x)F(t)} \leq \frac{[F(x+t)]^\theta - [F(x)]^\theta}{F(x+t) - F(x)},$$

which gives

$$[F(t)]^\theta - [F(x)]^\theta [F(t)]^\theta \leq [F(x+t)]^\theta - [F(x)]^\theta.$$

Rewriting this in terms of $R_Y(x)$, we get

$$R_Y(x+t) \leq R_Y(x)R_Y(t). \tag{7}$$

Hence, Y is NBU for $\theta > 1$. By using the second part of the Lemma 1, proof for NWU case follows.

Theorem 3 *Let X_1 and X_2 be discrete random variables with distribution functions $F_1(.)$ and $F_2(.)$ respectively. Suppose Y_1 and Y_2 are two random variables such that the reversed hazard rate of Y_i is proportional to that of X_i, $i = 1, 2$, with same proportionality constant θ. Then $X_1 \geq_{lr} X_2$, implies $Y_1 \geq_{lr} Y_2$, provided $\langle F_1(x) \rangle$ and $\langle F_2(x) \rangle$ are log convex and log concave sequences respectively.*

Proof: Let $g_1(.)$ and $g_2(.)$ be the density functions of Y_1 and Y_2 respectively. Also suppose that $\lambda_1^*(.)$ and $\lambda_2^*(.)$ are reverse hazard rate functions of X_1 and X_2 respectively. Let $X_1 \geq_{lr} X_2$, then by Dewan and Kattumannil (2011)

$$X_1 \geq_{rhr} X_2,$$

which implies

$$\left(\frac{F_1(x)}{F_2(x)}\right)^\theta \uparrow \text{ in } x \text{ for } \theta > 0. \tag{8}$$

Since $\langle F_1(x) \rangle$ is log convex we have $\lambda_1^*(x+1) \geq \lambda_1^*(x)$ and since $\langle F_2(x) \rangle$ is log concave, $\lambda_2^*(x+1) \leq \lambda_2^*(x)$. Hence we have

$$(\lambda_1^*(x) - \lambda_2^*(x)) \uparrow \text{ in } x \tag{9}$$

$$\Rightarrow e^{\theta(\lambda_1(x) - \lambda_2(x))} \uparrow \text{ in } x. \tag{10}$$

Hence

$$\frac{g_1(x)}{g_2(x)} \uparrow \text{ in } x.$$

That is, $Y_1 \geq_{lr} Y_2$.

Theorem 4 *Let X_1 and X_2 be discrete random variables with distribution functions $F_1(.)$ and $F_2(.)$ respectively. Suppose Y_1 and Y_2 are two random variables such that the reversed hazard rate of Y_i is proportional to that of X_i, $i = 1, 2$, with same proportionality constant θ. Then $X_1 \geq_{rhr} X_2$, implies $Y_1 \geq_{rhr} Y_2$.*

Proof: Suppose $X_1 \geq_{rhr} X_2$. Then

$$\ln \frac{F_{X_1}(x)}{F_{X_1}(x-1)} \geq \ln \frac{F_{X_2}(x)}{F_{X_2}(x-1)}.$$

Theorem 5 *Let X_1 and X_2 be discrete random variables with distribution functions $F_1(.)$ and $F_2(.)$ respectively. Suppose Y_1 and Y_2 are two random variables such that the reversed hazard rate of Y_i is proportional to that of X_i, $i = 1, 2$, with same proportionality constant θ. Then $X_1 \geq_{st} X_2$, implies $Y_1 \geq_{st} Y_2$.*

Proof: Suppose $X_1 \geq_{st} X_2$. Then

$$F_{X_1}(x) \quad \leq \quad F_{X_2}(x).$$

3 Test for Proportionality Parameter

In this section, we develop a distribution free test for comparing lifetimes of two systems. Let $X_1, X_2, ..., X_m$ and $Y_1, Y_2, ..., Y_n$ be random samples from F and G respectively. Consider the model

$$\lambda^*_G(x) = \theta\lambda^*_F(x), \;\; \forall x.$$

We consider the problem of testing the null hypothesis

$$H_0 : \theta = 1,$$

against the alternative hypothesis

$$H_1 : \theta > 1.$$

We are interested in one way alternatives as the case with $\theta < 1$ can be solved with the old definition of reversed hazard rate. In the context of proportional reversed hazards model, the above testing problem can be restated as

$$H_0 : \lambda^*_G(x) = \lambda^*_F(x), \quad \forall x$$

against the alternative

$$H_1 : \lambda^*_G(x) > \lambda^*_F(x), \quad \text{for some } x.$$

Note that $\theta = 1$ if and only if $\lambda_G(x) = \lambda_F(x)$. But $\theta > 1$ implies the wider class of alternatives

$$H_{11} : h_G(x) > h_F(x).$$

We know that $\lambda^*_G(x) = \theta\lambda^*_F(x)$ if and only if $G(x) = [F(x)]^\theta$ so that under H_1, Under H_1, $\frac{G(x)}{F(x)}$ is non-decreasing in x for all $x > 0$. That is , H_1 holds if and only if, for every $x > y \geq 1$,

$$\delta(x, y) = G(x)F(y) - G(y)F(x) \geq 0.$$

with strict inequality holds for some x and y.
Consider

$$\begin{aligned}
\Delta(F, G) &= \sum_{y=1}^{\infty} \sum_{x=y+1}^{\infty} (G(x)F(y) - G(y)F(x))f(x)g(y) \\
&= \sum_{y=1}^{\infty}\sum_{x=y}^{\infty} G(x)F(y)f(x)g(y) - \sum_{y=1}^{\infty}\sum_{x=y}^{\infty} G(y)F(x)f(x)g(y) \\
&= P[X_1 \leq Y_1 < Y_2 \leq X_2] - P[Y_1 \leq Y_2 < X_1 \leq X_2].
\end{aligned}$$

Clearly, under H_0, $\Delta(F, G) = 0$ and under H_1, $\Delta(F, G)$ is positive. Next we find an estimator of $\Delta(F, G)$.

Let X_1, X_2 and Y_1, Y_2 be random samples from two independent populations having distribution functions F and G respectively. Consider the indicator function

$$h(x_1, x_2, y_1, y_2) = \begin{cases} 1 & \text{if } xyyx \\ -1 & \text{if } yyxx \\ 0 & \text{otherwise,} \end{cases}$$

where the arrangement $xyyx$ represents $\{X_1 \leq Y_1 < Y_2 \leq X_2\} \cup \{X_1 \leq Y_2 < Y_1 \leq X_2\} \cup \{X_2 \leq Y_1 < Y_2 \leq X_1\} \cup \{X_2 \leq Y_2 < Y_1 \leq X_1\}$. Then a U-statistic defined by

$$W = \frac{1}{\binom{n_1}{2}\binom{n_2}{2}} \sum \sum h(x_{i_1}, x_{i_2}, y_{j_1}, y_{j_2}), \tag{11}$$

where summation is over $1 \leq i_1 < i_2 \leq n_1$ and $1 \leq j_1 < j_2 \leq n_2$. is an unbiased estimator of $4\Delta(F, G)$ and is zero under H_0 and positive under H_1.
The test procedure is to reject the null hypothesis H_0 in favour of H_1 for large values of the test statistic W.

3.1 Asymptotic Properties

In this sub-section, we discuss the asymptotic distribution of the test statistic W. The following theorem follows from a result due to Lehmann (1951).

Theorem 6 *Under the alternative hypothesis H_1, the statistic W is a consistent estimator of $\Delta(F, G)$.*

The next result gives the asymptotic distribution of the test statistic.

Theorem 7 *If $\frac{n_1}{N} \to p \in (0,1)$ as $N = (n_1 + n_2) \to \infty$ and if $\sigma_{22}^2 < \infty$ then $\sqrt{N}W$ converges in law to N(0,σ^2), where $\sigma^2 = \frac{4}{p}\psi_{10} + \frac{4}{1-p}\psi_{01}$ Under H_0, σ^2 is given by (13).*

Proof: Under H_0

$$\begin{aligned} h_{10}(x_1) &= E(h(x_1, X_2, Y_1, Y_2)) \\ &= C + 2F(x-1)\sum_{y=x}^{\infty}[F(y-1)]f(y) - F(x-1) - F^3(x-1), \end{aligned}$$

since X and Y have same distribution under the null hypothesis, where C is a constant independent of x. Hence

$$\psi_{10} = 4V(2F(X-1)\sum_{y=X}^{\infty}[F(y-1)]f(y) - F(X-1) - F^3(X-1)). \tag{12}$$

It is easy to see that under H_0 , ψ_{01} equal to ψ_{10}. Hence, under H_0,

$$\sigma^2 = \frac{4}{p}\psi_{10} + \frac{4}{1-p}\psi_{01} = \frac{4}{p(1-p)}\psi_{10}. \tag{13}$$

We reject the null hypothesis in favour of the alternative hypothesis when

$$\frac{\sqrt{N}W}{\hat{\sigma}} > Z_\alpha,$$

where $\hat{\sigma}$ is the consistent estimator of σ and Z_α is the upper α percentile of $N(0,1)$.

Remark 1 *One can also look at the problem of testing H_0 against the alternatives $H_2 : \theta < 1$ and we reject H_0 in favour of H_2, if*

$$\frac{\sqrt{N}W}{\hat{\sigma}} < -Z_\alpha,$$

4 Conclusions and Future Works

In the present work using the modified reversed hazard rate which can be considered as the discrete analogue of reversed hazard rate in continuous setup, we proposed a proportional reversed hazards model for discrete data. The proposed model can incorporate continuous covariates as well. We studied the ageing properties of the model. We proved certain stochastic orderings under the assumption of proposed model. A test based on U-statistics was developed to test that the proportionality parameter is one. The test statistic follows asymptotically normal distribution and is consistent under relevant alternatives.

References

Almalki, S.J., & Nadarajah, S. (2014). A new discrete modified Weibull distribution. *IEEE Transaction on Reliability.* 63, 68-80.

Bebbington, M., Lai, C. D., Wellington, M., & Zitikis, R. (2012). The discrete additive Weibull distribution: A bathtub-shaped hazard for discontinuous failure data. *Reliability Engineering and System Safety.* 106, 37-44.

Bracquemond, C., & Gaudoin, O. (2003). A survey on discrete lifetime distributions. *International Journal of Reliability, Quality, Safety Engineering.* 10, 69 - 98.

Chen, S., & Manatunga, A. K. (2007). A note on proportional hazards and proportional odds models. *Statistics and Probability Letters.* 77, 981-988.

Gupta, R. C., & Gupta, R. D. (2007). Proportional reversed hazard rate model and its applications, *Journal of Statistical Planning and Inference.* 137, 3525-3536.

Khorashadizadeh, M., Roknabadi, A. H. R., & Borzadaran, G. R. M. (2013). Characterization of life distributions using log-odds rate in discrete ageing, *Communications in Statistics -Theory and Method.* 42, 76-87.

Lai, C. D. (2013). Issues concerning constructions of discrete lifetime models. *Quality Technology and Qualitative Management.* 10, 251-262.

Nanda, A. K., & Sengupta, D. (2005). Discrete life distribution with decreasing hazard. *Sankhya- A.* 55, 164-168.

Nanda, A. K., & Das, S. (2011). Dynamic proportional hazard rate and reversed hazard rate models.*Journal of Statistical Planning and Inference.* 141, 2108-2119.

Nawata, K., Li, M., Ishiguro, A., & Kawabuchi, K. (2009). An analysis of the length of hospital stay for cataract patients in Japan using the discrete-type proportional hazard model. *Mathematics and Computers in Simulation.* 79, 2889-2896.

Noughabi, M. S., Roknabadi, A. H. R., & Borzadaran, G. R. M. (2013). Some discrete lifetime distributions with bathtub-shaped hazard rate Functions, *Quality Engineering*, 25, 225-236.

Yu, Q. (2007). A note on the proportional hazards model with discontinuous data.*Statistics and Probability Letters.* 77, 735-739.