Journal of Korean
Data & Information Science Society
2004, Vol. 15, No. 4, pp. 965～971

Nonparametric Estimation of Mean Residual Life by Partial Moment Approximation under Proportional Hazard Model[1]

Young Joon Cha[2]

Abstract

In this paper we consider several nonparametric estimators for the mean residual life by using the partial moment approximation under the proportional hazard model. Also we compare the magnitude of mean square error of the proposed nonparametric estimators for mean residual life under the proportional hazard model.

Keywords : Mean Residual Life, Partial Moment Approximation, Proportional Hazard Model

1. Introduction

In human and survival studies as well as in life-testing experiments in the engineering sciences, the proportional hazard model(PHM) has received some special attention since the paper of Koziol and Green(1976). Chen, Hollander, and Langberg(1982) computed the small sample bias and variance of the Kaplan-Meier estimator for the PHM. Under this PHM, Ebrahimi(1985) and Cheng and Lin(1987) introduced the nonparametric estimators for the survival function and examined the asymptotic properties for the survival function estimators. Further summaries in PHM are provided by Lawless(2002).

Estimation of the survival function is very important in survival analysis. Equally important also are estimation of the mean residual life(MRL). Hall and Wellner(1981), and Guess and Proschan(1988) gave a good review of theory and applications for the mean residual life function. Choobineh and Branting(1986)

---

1) This Research was supported by Andong National University Research Fund in 2003
2) Professor, Department of Information Statistics, Andong National University
e-mail : yjcha@andong.ac.kr

computed a simple expression for semivariance using the approximate equation. Choobineh and Park(1990) proposed the small sample estimator of the MRL using the partial moment approximation and compared with the empirical MRL estimator. Park, Lee, and Cha(1994), the partial moment approximation applied to the mean residual life function for the random censorship model.

This paper formulates, in Section 2, the nonparametric estimators for the mean residual life function using the partial moment approximation based on the proportional hazard model. We utilize the technique of Choobineh and Park(1990) to provide the nonparametric estimators for the mean residual life function based on the proportional hazard model. The accuracy of the nonparametric estimators by means of simulation is presented in Section 3.

2. Nonparametric MRL Estimation by Partial Mement Approximation based on PHM

Let $X_1, X_2, \cdots, X_n$ be i.i.d random variables with common distribution function $F(x)$ on $[0, \infty)$ with $F(0)=0$ and finite mean $\mu = E(X)$. Let $S_F = 1-F(x)$ denote the survival function. The mean residual life function at age x is defined as $e(x) = \int_x^\infty S_F(u)du/S_F(x) - x$ and $e(x)=0$ whenever $S_F(x)=0$. Let $Y_1, Y_2, \cdots, Y_n$ be i.i.d censoring random variables with common distribution function $G(y)$. Let $S_G(y) = 1-G(y)$. Define $Z_i = min(X_i, Y_i)$ and $\delta_i = I(X_i \leq Y_i)$ for $i=1, \cdots, n$, where $I(A)$ denotes the indicator for the set A. In this paper we consider the proportional hazard model $S_G(t) = (S_F(t))^{\beta}$ with for each t and an unknown positive constant β. In actuarial science and reliability analysis, the proportional hazard model has received some special attention since the paper of Koziol and Green(1976). The MRL function can be rewritten as $e(x) = \left(\mu - \int_x^\infty tdF(t)\right)/S_F(x) - x$, where $\int_0^x tdF(t)$ is defined as the first partial moment of a random variable X about the origin over $(0, x)$ for fixed x. To get partial moment approximation of MRL at age t as $e(x) = \mu + (F(x)/1-F(x))^{1/2}\sigma - x$, we utilized the result of Choobineh and Park(1990).

In this section we consider Kaplan-Meier type and Nelson-Aalen type mean residual life estimators using the partial moment approximation based on PHM. Hence, we can easily construct two nonparametric estimators $\hat{e}_{km}(x)$, $\hat{e}_{na}(x)$ for the mean residual life by the partial moment approximation under PHM, respectively as follows :

$$\hat{e}_{km}(x) = \int_0^\infty x d\hat{F}_{KM}(x) + \left(\frac{\hat{F}_{KM}(x)}{1-\hat{F}_{KM}(x)}\right)^{1/2} * \left(\int_0^\infty x^2 d\hat{F}_{KM}(x) - \left(\int_0^\infty x d\hat{F}_{KM}(x)\right)^2\right)^{1/2} - x$$

$$\hat{e}_{na}(x) = \int_0^\infty x d\hat{F}_{NA}(x) + \left(\frac{\hat{F}_{NA}(x)}{1-\hat{F}_{NA}(x)}\right)^{1/2} * \left(\int_0^\infty x^2 d\hat{F}_{NA}(x) - \left(\int_0^\infty x d\hat{F}_{NA}(x)\right)^2\right)^{1/2} - x$$

where

$Z_{(1)} < Z_{(2)} < \cdots < Z_{(n)}$ denote the ordered observed lifetimes, and $\delta_{(1)}, \delta_{(2)}, \cdots \delta_{(n)}$ are their unordered indicator values,

$$1 - \hat{F}_{KM}(t) = \prod_{i:\, Z_{(i)} \le t} \left(\frac{n-i}{n-i+1}\right)^{\delta_{(i)}}, \quad 1 - \hat{F}_{NA}(t) = e^{-\sum_{i:\, Z_{(i)} \le t} \frac{\delta_{(i)}}{n-i+1}}$$

It is well-known (Cheng and Chang(1985)) that the observations and the observed types of failure are independent if and only if $S_G = (S_F(t))^{\beta}$ holds for unknown positive constant β. Also under PHM, they compared a maximum likelihood estimator with Kaplan-Meier estimator for the survival function.

Using Choobineh and Branting(1986) and Cheng and Chang(1985), we can gain the mean residual life estimator as follows :

$$\hat{e}_c(x) = \begin{cases} x d\hat{F}_C(x) + \left(\dfrac{\hat{F}_C(x)}{1-\hat{F}_C(x)}\right)^{1/2} * \\ \left(\int_0^\infty x^2 d\hat{F}_C(x) - \left(\int_0^\infty x d\hat{F}_C(x)\right)^2\right)^{1/2} - x \ , \quad \text{if} \ \ x < x_{(n)} \end{cases}$$

where

$$\hat{F}_C(x) = 1 - (\overline{Z}_n(x))^T, \overline{Z}_n(x) = \frac{1}{n}\sum_{i=1}^n I(Z_i > x), T = \frac{1}{n}\sum_{i=1}^n \delta_i, x_{(n)} = max_{1 \le i \le n} x_i$$

By estimating $I(Z > x, \delta = 1)$, $I(Z > x, \delta = 0)$, and β under PHM, Ebrahimi(1985) derived the nonparametric estimator for the survival function. From the above result and the partial moment approximation, we construct the mean residual life estimator as follows :

$$\hat{e}_e(x) = \begin{cases} \hat{\mu} + \left(\dfrac{\hat{F}_E(x)}{1-\hat{F}_E(x)} \right)^{1/2} \hat{\sigma} - x & \text{if } x < x_{(n)} \\ 0 & \text{if } x \geq x_{(n)} \end{cases}$$

where

$$1-\hat{F}_E(x) = T\exp\left\{T\ln 1/T + T\ln \hat{\overline{U}}(x)\right\} + (1-T)\exp\{T\ln 1/(1-T) + T\ln \hat{\overline{U}}(x)\}$$

$$\hat{\overline{U}}(t) = \frac{1}{n}\sum_{i=1}^{n} I(Z_i > t, \delta_i = 1) \text{ and } \hat{\overline{C}}(t) = \frac{1}{n}\sum_{i=1}^{n} I(Z_i > t, \delta_i = 0)$$

$$\hat{\mu} = \int_0^\infty x d\hat{F}_E(x),\ \hat{\sigma}^2 = \int_0^\infty x^2 d\hat{F}_E(x) - \hat{\mu}^2,\quad T = \frac{1}{n}\sum_{i=1}^{n}\delta_i.$$

3. Simulation Study and Conclusions

In this section, comparisons among $\hat{e}_{km}(x)$, $\hat{e}_{na}(x)$, $\hat{e}_e(x)$, and $\hat{e}_c(x)$ under PHM will be made based on the results of Section 2. To compare the performances of nonparametric estimators $\hat{e}_{km}(x)$, $\hat{e}_{na}(x)$, $\hat{e}_e(x)$, and $\hat{e}_c(x)$ for the MRL function under the PHM, we carry out Monte Carlo simulation. The constant β can be interpreted as the censoring parameter. $\beta = 0$ corresponds to no censoring, and the expected proportion of censored observations increase with β, that is, β represents the odds for censoring. Therefore, simulations study were carried out to investigate the effects of varying censoring rates(50%, 60%, 70%) and sample sizes(30, 50, 100, 200). All simulation trials were done 10,000 times. For each combinations of exponential survival distribution, various censoring odds, and sample sizes, we calculated the mean squared error(MSE) and bias. In order to display the simulation result what the MSE for four MRL estimators look like, we made Figure 1-2. Also we only report in the Figure 1 and 2 for sample size 50 and censoring rate 60%, and sample size 200 and censoring rate 70%, respectively, and in the Table 1 and 2 for sample size 50 and sample size 100, respectively under various censoring rates. From Figure 1-2 and Table 1-2, it is found that, in most cases, proposed mean residual life estimators $\hat{e}_c(x)$ and $\hat{e}_e(x)$ using the partial moment approximation based on PHM performs better than Kaplan-Meier type $\hat{e}_{km}(x)$ and Nelson-Aalen type MRL estimator $\hat{e}_{na}(x)$ based on the criterion of having smaller MSE if censoring rate(CR) is somewhat large.

Also the MSE of $\hat{e}_c(x)$ is less than that of $\hat{e}_e(x)$ if the censoring rates is not small. Or say, the finding indicates that $\hat{e}_c(x)$ is preferred to $\hat{e}_e(x)$ if PHM is valid and the censoring rates is not small.

Figure 1. MSE for sample size 50
Censoring rate 60%

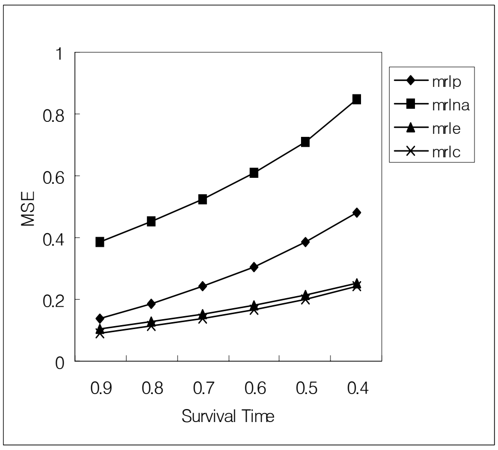
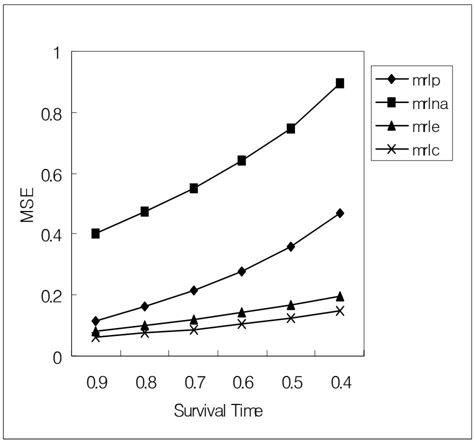


Figure 2. MSE for sample size 200
Censoring rate 70%



* mrlp = $\hat{e}_{km}(x)$, mrlna = $\hat{e}_{na}(x)$, mrle = $\hat{e}_e(x)$, and mrlc = $\hat{e}_c(x)$

Table 1. MSE of $\hat{e}_{km}(x)$, $\hat{e}_{na}(x)$, $\hat{e}_e(x)$, and $\hat{e}_c(x)$ for sample size 50

CR	t:	0.1054	0.2231	0.3567	0.5108	0.6932	0.9163
0.50	$\hat{e}_{km}(x)$	0.0948	0.1268	0.1616	0.2011	0.2507	0.3132
	$\hat{e}_{na}(x)$	0.2179	0.2548	0.2961	0.3425	0.4022	0.4819
	$\hat{e}_e(x)$	0.0726	0.0919	0.1126	0.1369	0.1688	0.2150
	$\hat{e}_c(x)$	0.0766	0.1018	0.1290	0.1620	0.2036	0.2619
0.60	$\hat{e}_{km}(x)$	0.1371	0.1866	0.2414	0.3050	0.3839	0.4791
	$\hat{e}_{na}(x)$	0.3869	0.4541	0.5258	0.6083	0.7117	0.8500
	$\hat{e}_e(x)$	0.1055	0.1288	0.1537	0.1817	0.2143	0.2521
	$\hat{e}_c(x)$	0.0926	0.1150	0.1396	0.1681	0.2019	0.2408
0.70	$\hat{e}_{km}(x)$	0.2244	0.3100	0.4030	0.5128	0.6400	0.7675
	$\hat{e}_{na}(x)$	0.6481	0.7740	0.9045	1.0560	1.2449	1.5210
	$\hat{e}_e(x)$	0.1851	0.2269	0.2678	0.3077	0.3260	0.2903
	$\hat{e}_c(x)$	0.1443	0.1744	0.2044	0.2360	0.2635	0.2619

Table 2. MSE of $\hat{e}_{km}(x)$, $\hat{e}_{na}(x)$, $\hat{e}_e(x)$, and $\hat{e}_c(x)$ for sample size 100

CR	t:	0.1054	0.2231	0.3567	0.5108	0.6932	0.9163
0.50	$\hat{e}_{km}(x)$	0.0563	0.0772	0.1000	0.1273	0.1611	0.2052
	$\hat{e}_{na}(x)$	0.1381	0.1625	0.1896	0.2223	0.2642	0.3218
	$\hat{e}_e(x)$	0.0423	0.0555	0.0704	0.0896	0.1141	0.1493
	$\hat{e}_c(x)$	0.0456	0.0638	0.0847	0.1118	0.1471	0.1985
0.60	$\hat{e}_{km}(x)$	0.0905	0.1260	0.1646	0.2100	0.2670	0.3408
	$\hat{e}_{na}(x)$	0.2701	0.3162	0.3658	0.4230	0.4941	0.5885
	$\hat{e}_e(x)$	0.0655	0.0817	0.0990	0.1192	0.1439	0.1766
	$\hat{e}_c(x)$	0.0587	0.0758	0.0947	0.1178	0.1477	0.1860
0.70	$\hat{e}_{km}(x)$	0.1620	0.2276	0.2986	0.3823	0.4896	0.6202
	$\hat{e}_{na}(x)$	0.5165	0.6140	0.7148	0.8290	0.9747	1.1716
	$\hat{e}_e(x)$	0.1256	0.1544	0.1833	0.2145	0.2464	0.2628
	$\hat{e}_c(x)$	0.0961	0.1160	0.1365	0.1590	0.1846	0.2068

References

1. Chen, Y. Y., Hollander, M. and Langberg, N. A.(1982). Small-sample Results for the Kaplan-Meier Estimator, *Journal of American Statistical Association,* Vol. 77, 141-144.
2. Cheng, P. E. and Chang, Y. C.(1985). PLE versus MLE of Survival Functions under a simple Proportional Hazards Model. *Journal of the Chinese Statistical Association,* Vol. 23, Special Issue, 57-73.
3. Cheng, P. E. and Lin, G. D.(1987). Maximum Likelihood Estimation of Survival Function under the Koziol-Green Proportional Hazards Model, *Statistics and Probability Letters,* Vol. 5, 75-80.
4. Choobineh, F. and Branting, D.(1986). A Simple Approximation for Semivariance, *European Journal of Operations Research*, Vol. 27, 364-370.
5. Choobineh, F. and Park, D. H.(1990). A Nonparametric Small Sample Estimator of Mean Residual Life, *Journal of the Korean Statistical Society*, Vol. 19, No.1 80-87.
6. Ebrahimi, N (1985). Nonparametric Estimation of Survival Functions for Incomplete Observations when the Life Time Distribution is proportionally related to the Censoring Time Distribution, *Communications in Statistics-Theory and Methods* Vol. 14, 2887-2898.

7. Guess, F. and Proschan, F.(1988). Mean Residual Life. Theory and Applications, *Handbook of Statistics*, Vol.7, Reliability and Quality Control, P. R. Krishnaiah and C. R. Rao, eds., 215-224.
8. Hall, W. J. and Wellner, J. A.(1981). Mean Residual Life, *In Statistics and Related Topics,* M. Csörgö et al.,, 169-184, North-Hollander, Amsterdam.
9. Koziol, J. and Green, S.(1976). A Cramer-von Mises statistic for randomly censored data. *Biometrika Vol.* 63, 465-481.
10. Lawless, J. F. (2002). *Statistical Models and Methods for Lifetime Data,* 2nd Edition, Wiley.
11. Park, B. G., Lee, J. M. and Cha, Y. J.(1994). Estimation of Mean Residual Life under Random Censorship Model using Partial Moment Approximation. *Journal of the Korean Society for Quality Management* 22(3), 111-118.

[received date : Aug. 2004, accepted date : Nov. 2004]