▩ 연구논문

Estimation of Mean Residual Life under Random Censorship Model Using Partial Moment Approximation+

Byung Gu Park
Dept. of Statistics, Kyungpook National University
Jae Man Lee · Young Joon Cha
Dept. of Statistics, Andong National University

Abstract

In this paper we propose a parametric and a nonparametric small sample estimators for the mean residual life(MRL) under the random censorship model using the partial moment approximation. We also compare the proposed nonparametric estimator with the well-known nonparametric MRL estimator based on Kaplan-Meier estimator of the survival function, and present the efficiency of the nonparametric method relative to the Weibull model for small samples.

1. Introduction

Let $X_1, X_2, \cdots, X_n$ be a random sample from a life distribution function $F(x)$ on $[0, \infty)$. Let $S_F(x) = 1 - F(x)$ denote the survival function. Suppose that $F(0) = 0$ and mean $\mu \equiv E(X)$ is finite. Then the MRL function at age x is defined as

$$e(x) \equiv E(X - x \mid X > x)$$

+ This paper was supported (in part) by Non Directed Research Fund, Korea Research Foundation, 1993.

$$= \begin{cases} \dfrac{\int_x^{\infty} S_F(u)du}{S_F(x)}, & S_F(x) \neq 0 \\ 0, & S_F(x) = 0 \end{cases} \qquad (1.1)$$

Let $Y_1, Y_2, \cdots, Y_n$ be a random sample from a censoring distribution function $G(y)$. Let $S_G(y) = 1-G(y)$. Define $Z_i = \min(X_i, Y_i)$ and $\delta_i = I(X_i \le Y_i)$ for $i = 1, \cdots, n$, where $I(A)$ denotes the indicator function for the set A. Under the random censorship model, X_i is assumed to be independent of Y_i for each i, $S(x) = P(Z > x) = S_F(x)S_G(x)$ for any F and G.

In reliability and survival analyses the MRL function plays an important role and has wide range of applications.

Hall and Wellner(1981), and Guess and Proschan(1988) gave a good review of theory and applications for the MRL. Yang(1977, 1978), Ghorai et al.(1982), and Ghorai and Rejtö(1987) studied the asymptotic properties of the several estimators of the MRL. Through development of the approximate equation, Choobineh and Branting(1986) devised a simple alternative expression for semivariance. Choobineh and Park(1990) proposed a nonparametric small sample estimator of the MRL using the partial moment approximation and compared with the empirical MRL estimator.

In this paper, for the MRL function, an approximation of the partial moment applies to the random censorship model. In Section 2, we propose a parametric MRL estimator for the Weibull model and a nonparametric MRL estimator under the random censorship model using the partial moment approximation. In Section 3, to illustrate the proposed MRL estimators, we take an artificial data set. In Section 4, through the Monte Carlo simulation study, we compare a new nonparametric estimator with well-known nonparametric estimator $\hat{e}_{KM}(x)$ based on Kaplan-Meier estimator of the survival function, which is a censored version of the empirical MRL estimator, and present the efficiency of the nonparametric method relative to the Weibull model.

2. Estimation of MRL under Random Censorship Model

The MRL function, defined in (1.1), can be rewritten as

$$e(x) = E(X-x \mid X > x)$$

$$= \frac{\int_x^\infty t\,dF(t)}{S_F(x)} - x$$

$$= \frac{\mu - \int_0^x t\,dF(t)}{S_F(x)} - x \qquad (2.1)$$

where $\int_0^x t\,dF(t)$ is defined as the first partial moment of a random variable X about the origin over $(0, x)$ for fixed x. Choobineh and Park (1990) showed that

$$F(x)\left(\mu - \left(\frac{1-F(x)}{F(x)}\right)^{1/2}\sigma\right) \qquad (2.2)$$

is an approximation to $\int_0^x t\,d\mathrm{F}(t)$, where μ and σ^2 are the mean and variance of X, respectively. Therefore, the partial moment approximation of $e(x)$ is given by

$$e_p(x) = \mu + \left(\frac{F(x)}{1-F(x)}\right)^{1/2}\sigma - x. \qquad (2.3)$$

Under the random censorship model, it is well known that the Kaplan-Meier (1958) estimator, $\hat{F}_{KM}(x)$, of the distribution function $F(x)$ is represented as

$$\hat{F}_{KM}(x) = \begin{cases} 1 - \Pi_{i:Z_{(i)} \le x} \left(\frac{n-i}{n-i+1}\right)^{\delta_{(i)}} & \text{if } x \le Z_{(n)} \\ 1 & \text{if } x > Z_{(n)},\ \delta_{(n)} = 1, \\ \text{undefined} & \text{if } x > Z_{(n)},\ \delta_{(n)} = 0 \end{cases}$$

where $Z_{(i)}$ is the ith order statistic based on $Z_1, \cdots, Z_n$ and (i) is the corresponding censoring indicator.

Let $\hat{\mu} = \int_0^\infty x\,d\hat{F}_{KM}(x)$ and $\hat{\sigma}^2 = \int_0^\infty x^2\,d\hat{F}_{KM}(x) - \hat{\mu}^2$. Substituting $\hat{F}_{KM}(x)$, $\hat{\mu}$ and $\hat{\sigma}$, respectively, for $F(x)$, μ, and σ in the expression of $e_p(x)$, we obtain a new nonparametric MRL estimator $\hat{e}_{1p}(x)$ as follows;

$$\hat{e}_{1p}(x) = \begin{cases} \hat{\mu} + \left(\frac{\hat{F}_{KM}(x)}{1-\hat{F}_{KM}(x)}\right)^{1/2}\hat{\sigma} - x & \text{if } x < x_{(n)} \\ 0 & \text{if } x \ge x_{(n)} \end{cases} \qquad (2.4)$$

where $x_{(n)} = \max_{1\le i\le n} x_i$. The MRL estimator $\hat{e}_{1p}(x)$ is simple in form and uses only familiar statistics, namely, mean, variance, and cumulative probability below the critical value.

We consider the Weibull model with the scale parameter α and shape parameter β under the random censorship model. Then, the likelihood function is

$$L(\alpha, \beta) = \prod_{i=1}^{n} \left[\frac{\beta}{\alpha}\left(\frac{z_i}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{z_i}{\alpha}\right)^{\beta}\right)\right]^{\delta_i} \left[\exp\left(-\left(\frac{z_i}{\alpha}\right)^{\beta}\right)\right]^{1-\delta_i}. \qquad (2.5)$$

The log likelihood function and maximum likelihood estimators are

$$\log L(\alpha, \beta) = r\log\beta - r\beta\log\alpha + (\beta-1)\sum_{i\in D}\log z_i - \sum_{i=1}^{n}\left(\frac{z_i}{\alpha}\right)^{\beta}$$

and

$$\sum_{i=1}^{n} z_i^{\hat\beta}\log z_i \Big/ \sum_{i=1}^{n} z_i^{\hat\beta} - \frac{1}{\hat\beta} - \frac{1}{r}\sum_{i\in D}\log z_i = 0$$

$$\hat\alpha = \left(\frac{1}{r}\sum_{i=1}^{n} z_i^{\hat\beta}\right)^{1/\hat\beta}, \qquad (2.6)$$

where $r = \sum\delta_i$ denote the number of observed lifetimes and D denote the set consisting of those individuals for which $\delta_i = 1$.

For the Weibull model, the expectation and the variance are $\mu = \alpha\Gamma(1+1/\beta)$ and $\sigma^2 = \alpha^2[\Gamma(1+2/\beta) - (\Gamma(1+1/\beta))^2]$, respectively, where Γ is the gamma function defined by the definite integral $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$. Consequently, we have a parametric MRL estimator $\hat{e}_{2p}(x)$ as follows:

$$\hat{e}_{2p}(x) = \hat\alpha\Gamma(1+1/\hat\beta) + \left(\frac{\hat{F}(x)}{1-\hat{F}(x)}\right)^{1/2}\hat\alpha[\Gamma(1+2/\hat\beta) - (\Gamma(1+1/\hat\beta))^2]^{1/2} - x, \qquad (2.7)$$

where $\hat{F}(x) = 1 - \exp\left(-\left(\frac{x}{\hat\alpha}\right)^{\hat\beta}\right)$.

3. An Example

In order to illustrate an idea what the three MRL estimators look like, we take an artificial data set ⟨Table 1⟩ as follows:

(i) The underlying survival distribution is a Weibull with scale parameter 1 and shape parameter 4.

(ii) The censoring distribution is an exponential with failure rate 0.4, where the censoring rate is 10%.

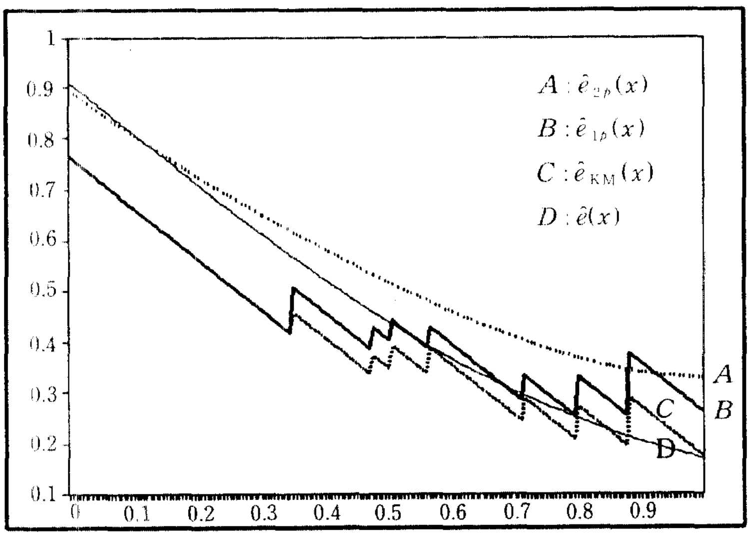
(iii) The sample size is 10.

⟨Figure 1⟩ gives a graphical presentation to illustrate $\hat{e}_{KM}(x)$, $\hat{e}_{1p}(x)$, $\hat{e}_{2p}(x)$, $\hat{e}_{2p}(x)$, and $e(x)$.

⟨ Table 1 ⟩ An Artificial Data Set

(0.3499, 1)	(0.4723, 1)	(0.5046, 1)	(0.5605, 1)	(0.7108, 1)
(0.7764, 0)	(0.7959, 1)	(0.8754, 1)	(1.0774, 1)	(1.2625, 1)

* 1 and 0 are described as uncensored and censored, respectively.



⟨ Figure 1 ⟩ MRL Estimators $\hat{e}_{KM}(x)$, $\hat{e}_{1p}(x)$, and $\hat{e}_{2p}(x)$ with true MRL $e(x)$

4. Numerical Illustration and Conclusion

In this section we intend to compare the performances of three estimators $\hat{e}_{KM}$

(x), $\hat{e}_{1p}(x)$, $\hat{e}_{2p}(x)$ for the MRL in small sample of censored data through Monte Carlo study. Furthermore, Monte Carlo simulations were carried out to investigate the effects of varing the survival distributions, censoring rates, and sample sizes. Trials were done 1000 times. For each combinations of survival distributions (Weibull distribution with scale parameter $\alpha = 1$ and shape parameter β ($\beta = 4, 10, 15, 20$), mission time t (t : $S_F(t) = 0.9(0.1)0.5$)), sample sizes n ($n = 3, 5, 7, 10$), and 30% censoring rates, the mean squared error(MSE) and bias were computed.

〈Tables 2〉 illustrates the following findings:

(i) The proposed estimators $\hat{e}_{1p}(x)$ and $\hat{e}_{2p}(x)$ do very well if the underlying distribution is symmetric and the sample size is small.

(ii) $\hat{e}_{2p}(x)$ is very good for the initial stage of the survival function.

(iii) Varing censoring distributions give no changes in the simulation results.

(iv) $\hat{e}_{KM}(x)$ and $\hat{e}_{1p}(x)$ are similar to in all cases, thus we recommend the estimator $\hat{e}_{1p}(x)$ because of simplicity in form.

〈 Table 2 〉 Bias and MSE of $\hat{e}_{KM}(x)$ $\hat{e}_{1p}(x)$, and $\hat{e}_{2p}(x)$ under 30% Censoring Rates

sample size		3						5					
		$\hat{e}_{KM}(x)$		$\hat{e}_{1p}(x)$		$\hat{e}_{2p}(x)$		$\hat{e}_{KM}(x)$		$\hat{e}_{1p}(x)$		$\hat{e}_{2p}(x)$	
β	r	bias	mse	bias	mse	bias	mse	bias	mse	bias	mse	bias	mse
	.5697	−.0215	.0205	−.0179	.0204	.0328	.0047	−.0089	.0150	.0006	.0155	.0339	.0038
	.6873	−.0140	.0199	−.0081	.0202	.0236	.0063	−.0106	.0142	.0029	.0146	.0274	.0052
4.0	.7728	−.0159	.0209	−.0089	.0211	.0160	.0079	−.0110	.0145	.0063	.0149	.0220	.0065
	.8454	−.0241	.0206	−.0165	.0207	.0117	.0096	−.0153	.0159	.0048	.0167	.0196	.0079
	.9124	−.0356	.0199	−.0265	.0199	.0128	.0112	−.0268	.0160	−.0019	.0167	.0224	.0093
	.7985	−.0009	.0038	.0006	.0038	.0155	.0008	−.0016	.0023	.0021	.0023	.0124	.0003
	.8607	.0041	.0036	.0067	.0036	.0132	.0011	.0017	.0022	.0077	.0023	.0102	.0005
10.0	.9020	.0043	.0034	.0076	.0034	.0122	.0014	.0031	.0022	.0112	.0023	.0093	.0007
	.9350	.0028	.0033	.0069	.0034	.0129	.0017	.0015	.0021	.0120	.0024	.0102	.0009
	.9640	−.0024	.0029	.0024	.0031	.0165	.0021	−.0026	.0020	.0108	.0024	.0140	.0012
	.8607	−.0022	.0017	−.0010	.0017	.0109	.0003	.0015	.0010	.0041	.0011	.0095	.0002
	.9048	.0041	.0015	.0061	.0015	.0106	.0004	.0041	.0010	.0084	.0010	.0090	.0003
15.0	.9336	.0043	.0015	.0071	.0015	.0110	.0006	.0051	.0010	.0109	.0011	.0092	.0004
	.9562	.0022	.0014	.0052	.0015	.0125	.0007	.0040	.0009	.0120	.0011	.0108	.0005
	.9759	−.0012	.0013	.0028	.0014	.0160	.0009	.0015	.0009	.0120	.0012	.0142	.0007
	.8936	.0010	.0009	.0020	.0009	.0092	.0002	.0013	.0006	.0035	.0006	.0081	.0001
	.9277	.0036	.0008	.0051	.0008	.0097	.0003	.0034	.0006	.0067	.0006	.0085	.0002
20.0	.9498	.0042	.0008	.0060	.0009	.0106	.0004	.0035	.0005	.0084	.0006	.0093	.0003
	.9670	.0026	.0008	.0051	.0009	.0123	.0005	.0028	.0005	.0096	.0007	.0110	.0004
	.9818	−.0005	.0007	.0025	.0008	.0155	.0007	.0010	.0005	.0103	.0007	.0142	.0005
sample size		7						10					
	.5697	−.0044	.0110	.0080	.0116	.0325	.0031	−.0011	.0075	.0172	.0083	.0303	.0019
	.6873	−.0040	.0107	.0139	.0115	.0270	.0042	−.0010	.0072	.0214	.0081	.0265	.0026
4.0	.7728	−.0083	.0105	.0136	.0111	.0226	.0053	−.0018	.0072	.0233	.0081	.0234	.0032
	.8454	−.0132	.0109	.0145	.0114	.0211	.0064	−.0060	.0071	.0237	.0080	.0232	.0040
	.9124	−.0210	.0117	.0130	.0127	.0248	.0076	−.0150	.0075	.0242	.0083	.0280	.0049
	.7985	.0008	.0017	.0055	.0017	.0121	.0003	.0033	.0012	.0101	.0013	.0125	.0003
	.8607	.0027	.0016	.0097	.0017	.0100	.0004	.0041	.0011	.0131	.0013	.0113	.0004
10.0	.9020	.0022	.0015	.0115	.0016	.0091	.0006	.0026	.0011	.0136	.0013	.0111	.0006
	.9350	−.0022	.0014	.0128	.0017	.0101	.0008	.0011	.0010	.0158	.0013	.0128	.0007
	.9640	−.0036	.0014	.0128	.0019	.0139	.0011	−.0012	.0010	.0206	.0016	.0172	.0010
	.8607	.0008	.0008	.0045	.0008	.0089	.0001	−.0005	.0005	.0042	.0006	.0088	.0001
	.9048	.0030	.0007	.0085	.0007	.0084	.0002	.0007	.0005	.0071	.0005	.0085	.0002
15.0	.9336	.0030	.0006	.0104	.0007	.0087	.0003	.0008	.0005	.0095	.0006	.0090	.0003
	.9562	.0015	.0006	.0117	.0008	.0103	.0004	.0001	.0004	.0122	.0007	.0107	.0004
	.9759	−.0005	.0006	.0129	.0010	.0138	.0005	−.0011	.0004	.0160	.0009	.0143	.0005
	.8936	−.0008	.0004	.0018	.0005	.0066	.0001	.0008	.0003	.0046	.0003	.0060	.0001
	.9277	.0022	.0004	.0064	.0005	.0064	.0001	.0019	.0002	.0071	.0003	.0070	.0001
20.0	.9498	.0019	.0004	.0075	.0004	.0068	.0002	.0022	.0002	.0090	.0003	.0070	.0002
	.9670	.0010	.0004	.0087	.0005	.0082	.0002	.0015	.0002	.0109	.0004	.0090	.0002
	.9818	−.0004	.0004	.0103	.0006	.0111	.0003	.0008	.0002	.0140	.0006	.0120	.0003

References

[1] Choobineh, F. and Branting, D. (1986), "A Simple Approximation for Semivariance," *European Journal of Operations Research*, Vol. 27, pp. 364−370.

[2] 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, pp. 80−87.

[3] Ghorai, J. K. and Rejtö, L. (1987), "Estimation of Mean Residual Life with Censored Data under the Proportional Hazard Model," *Communications in Statistics-Theory and Methods*, Vol. 16, No. 7, pp. 2097−2114.

[4] Ghorai, J. K. et al. (1982), "Non-Parametric Estimation of Mean Residual Life Time with Censorde Data," *Colloquia Mathematica Societatis* Janos Bolyai, 32. Nonparametric Statistical Inference, pp. 269−291. Edited by B. V. Gnedenok, I. Vincze, M. L. Puri ; North-Holland, Amsterdam.

[5] 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., pp. 215−224.

[6] Hall, W. J. and Wellner, J. A. (1981), "Mean Residual Life," In M. Csörgö, D. A. Dawson, J. N. K. Rao and A. K. Md. E. Saleh, eds., Statistics and Related Topics, pp. 169−184.

[7] Kaplan, E. L. and Meir, P. (1958), "Nonparametric Estimation from Incomplete Observation," *Journal of the American Statistical Association*, Vol. 53, pp. 457−481.

[8] Yang, G. (1977), "Life Expectancy under Random Censorship," *Stochastic Processes and Their Applications*, Vol. 6, pp. 33−39.

[9] Yang, G. (1978), "Estimation of a Biometric Function," *The Annals of Statistics*, Vol. 6, pp. 112−116.