Youngnam Statistical Letters
June, 1991, Vol. 2, 11-21

Mean Residual Life Times

Sang-Bock Lee* and Byung-Gu Park**

ABSTRACT

A different approach to the evaluation of mean residual life function under the random censorship model is presented. For small sample sizes, the performances between the proposed estimator and other estimators for mean residual life function are compared in terms of bias and mean square error via a Monte Carlo study.

1. Introduction and Notations

Let $x_1, \cdots, x_n$ be independent and identically distributed random survival times with a common distribution function $F(x)$ on $[0, \infty)$ with $F(0) = 0$ and mean μ. Let $S_F(x) = 1 - F(x)$ denote the survival function. Then the mean residual life function (MRLF) at age x is defined as

$$\begin{aligned} e(x) &= E[\, X - x \mid X > x \,] \\ &= \frac{\int_x^\infty S_F(u)\, du}{S_F(x)} \end{aligned} \qquad (1)$$

and $e(x) = 0$ whenever $S_F(x) = 0$. Yang(1978) proposed an estimator $\widehat{e}(x)$ of MRLF $e(x)$ as

$$\widehat{e}(x) = [S_n(x)]^{-1} \int_x^\infty S_n(v)\, dv \qquad (2)$$

for $v \leq X_{(n)}$, and showed that $\widehat{e}(x)$ is asymptotically unbiased and uniformly strong consistent. Also she proved that the empirical process

$$n^{1/2} \{ \widehat{e}(x) - e(x) \} \qquad (3)$$

* Department of Statistics, Hyosung Women's University, Taegu, Korea.
** Department of Statistics, Kyungpook National University, Taegu, Korea.

for $x \leq M \leq T_F = \inf\{ x \mid F(x) = 1 \} < \infty$, converges weakly to a zero mean Gaussian process. Burke, Csorgo and Horvath (1981) obtained strong approximations of the empirical process in (3) without assuming $T_F < \infty$. But they pointed out that their results are unsatisfied under the random censorship model.

Now, let $Y_1, \cdots, Y_n$ be random censoring times with a 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 $i = 1, \cdots, n$. Under the random censorship model X_i is assumed to be independent of Y_i for each i,

$$S(x) = P(Z > x) = S_F(x) S_G(x) \tag{4}$$

for any F and G. In genral, X and Y are not directly observable, but one observes only $(Z_1, \delta_1), \cdots, (Z_n, \delta_n)$. The problem of estimating S_F and some functionals of S_F are based on the data $(Z_1, \delta_1), \cdots, (Z_n, \delta_n)$.

Let the truncated MRLF $e_M(x)$ be defined by

$$e_M(x) \;=\; \frac{\int_x^M S_F(u)\, du}{S_F(x)}\;, \tag{5}$$

where $x \leq M \leq T_H = \inf\{z \mid H(z) = 1\}$. Then, the estimator $\widehat{e}_M(x)$ of a truncated MRLF $e_M(x)$ is defined by

$$\widehat{e}_M(x) = \int_x^M \widehat{S}_F(u)\, du \;/\; \widehat{S}_F(x), \tag{6}$$

where $\widehat{S}_F$ is some estimator of $S_F = 1 - F$.

Yang(1977) proposed estimators $\widehat{e}_M^{KM}(x)$ *and* $\widehat{e}_M^{NA}(x)$ of MRLF by using the Kaplan–Meier estimator and the Nelson–Aalen estimator. She also studied strong consistency and weak convergency of the estimators.

In this paper, we propose an estimator $\widehat{e}_M^{SV}(x)$ of MRLF $e_M(x)$ as using the Susarla–Van Ryzin(1980) estimator $\widehat{S}_F^{SV}(x)$ of survival function S_F, where

$$\widehat{S}_F^{SV}(x) = \frac{n-k}{n} \prod_{\{j; Z_{(j)} \leq x\}} \left[\frac{n-i+2}{n-i+1}\right]^{(1-\delta_j)} \tag{7}$$

2. Consistency of $\widehat{e}_M^{SV}(x)$

First, we consider the strong consistency of $\widehat{e}_M^{SV}(x)$. From now on, we denote

$$\sup\{|f(u)| \mid x \le u \le M\} \quad by \quad \|f\|_M$$

for any function f on $(x,\ M)$, and denote some constants as $c_1, c_2, \cdots$. From Section 1, one can get that

$$\begin{aligned}
&|\widehat{e}_M^{SV}(x) - e_M(x)| \qquad (8)\\
&= [\widehat{S}_F^{SV}(x) S_F(x)]^{-1}\\
&\times\ |S_F(x) \int_x^M \widehat{S}_F^{SV}\, du - \widehat{S}_F^{SV}(x) \int_x^M S_F\, du|\\
&= [\widehat{S}_F^{SV}(x) S_F(x)]^{-1}\ |S_F(x) \int_x^M [\widehat{S}_F^{SV} - S_F]\, du\\
&+\ [S_F(x) - \widehat{S}_F^{SV}(x)] \int_x^M S_F(x)\, du|\\
&\le\ [\widehat{S}_F^{SV}(x)]^{-1}\ \{S_F(x) \int_x^M |\widehat{S}_F^{SV}(x) - S_F|\, du\\
&+\ |S_F(x) - \widehat{S}_F^{SV}(x)| \int_x^M S_F\, du\}\\
&\le\ [\widehat{S}_F^{SV}(x) S_F(x)]^{-1}\ \{S_F(x) \int_x^M |S_n W_n - S S_G^{-1}|\, du\\
&+\ |S_F(x) - \widehat{S}_F^{SV}(x)| \int_x^M S_F\, du\}\\
&\le\ o[n^{1/3}]\{M S_G^{-1}(M) \|S_n - S\|_M\ +\ M\|S_n W_n - S_G^{-1}\|_M\\
&+\ |S_F(x) - \widehat{S}_F^{SV}(x)| \int_x^M S_F\, du\}\\
&\le\ o[n^{1/3}]\{I + II + III\}.
\end{aligned}$$

The 2^{nd} inequality follows by a triangular inequality after adding and substracting the integral $\int_x^M S_G^{-1} S_n\, du$, and from the condition (A3) of Surala and Van Ryzin(1980) one can get 3^{rd} inequality.

Now one can observe that

$$I\ =\ O\Big[\frac{M(\log\log n)}{\sqrt{n} S_G(M)}\Big] \qquad (9)$$

by the law of iterated logarithm (Susarla and Van Ryzin(1980)). to deal with II, it can be observed that,

$$II \le c_1 M||S_G^{-1} S_n(\ln W_n - \ln S_G^{-1})||_M + c_2 M||S_G^{-1} S_n(\ln W_n - \ln S_G^{-1})^2||_M \tag{10}$$

since $S_n \le 1$. A rate for strong convergence of strong II to zero can be obtained by three conditions of Susarla and Van Ryzin(1980);

(C1) $\sum_{i=1}^{\infty} a_n^2 \ / \ n^2 S^4(M) < \infty$, for positive constants a_n and for $0 < 2a < 1$ and $0 < 2\beta < 1 < p$,

(C2) $\sum_{n=1}^{\infty} a_n^p \ / \ S^{2p}(M) n^{\beta p} < \infty$,

(C3) $\underline{\lim}\ n^{\alpha} S(M) > 0$.

They obtained a rate for

$$M||S_G^- 1 S_n(\ln W_n - \ln S_G^- 1)||_M \to 0 \quad a.s., \tag{11}$$

as $o(\ S_G^{-1}(M) \max(a_n^{-1},\ \log n \ / \ n^{(1-2\alpha)/2}\ )$ under the above conditions (C1), (C2) and (C3). Finally, they also found that

$$III = o(\log n \ / \ n^{-1/2}). \tag{12}$$

Combining (9), (10), (11) and (12), one can get the following theorem without proof.

Theorem 2.1 (Strong convergency of $e_M(x)$)
Suppose that the conditions (C1), (C2) and (C3) are satisfied. Then for any $x \in [\ 0,\ M\ )$,

$$\hat{e}_M^{SV}(x) - e_m(x) = o(\ n^{1/3} S_G^- 1(M) M \cdot \max(a_n^- 1, \log n \ / \ n^{(1-2\alpha)/2})\ ) \quad a.s.,$$

where $M \le T_H = \inf\{x \mid H(x) = 1\}$.

Remark 2.1. Examples satisfying conditions (C1), (C2) and (C3) are given in Susarla and Van Ryzin(1980).

3. Comparisons of Estimators for $e_M(x)$

In this section we compare the performances of three estimators $\hat{e}_M^{KM}(x)$, $\hat{e}_M^{SV}(x)$ and $\hat{e}_M^{NA}(x)$ for $e_M(x)$ in terms of bias and mean square error(MSE) via a Monte Carlo study. The random censorship model was adopted: $X_1, \cdots, X_n$ are the true survival times and $Y_1, \cdots, Y_n$ are independent and identically distributed with absolutely continuous distributions F, G, respectively, as those of the first section.

Various combinations of two survival distributions S_F's, exponential(*Exp*) and Weibull(*Weib*), and two censoring distributions S_G's, exponential and uniform(*Unif*), have been simulated with different censoring patterns (10%, 30%) and different sample sizes ($n = 30,\ 50,\ 100$).

The given x's as conditionals considered in simulation were obtained by inverse of true survival function S_F, *i.e.*, $x = S_F^{-1}(1),\ S_F^{-1}(.9), \cdots, S_F^{-1}(.1)$.

Replication was done 500 times. For each values of x, the mean, bias, and MSE of $\hat{e}_M(x)$'s were computed. The standard error(s.e.) was also obtained for each MSE. We can summarize the design of simulation as the following Table 1.

Table 1. Design of Simulations

Distribution Survival/Censoring	Sample size n	Censoring Rate	Inverse Quentile $x = S^{-1}(x)$
			1.0
$Exp(1)/Exp(\lambda)$	30		.9
		10%	.8
$Exp(1)/Unif(\lambda)$			.7
	50		.6
			.5
		30%	.4
$Weib(1.15, 2)/Exp(\lambda)$			.3
	100		.2
			.1

Tables 2 (a)–(c) summarize the results of this simulation for $x = S_F(.7)$, $S_F(.5)$ and $S_F(.4)$, and sample sizes n=30, 50, and 100 with different censoring proportions (about 10%, 30%).

From Table 3.2, one can observe the following facts:

(1) As censoring proportion increases, or equivalently, as inverse quantile of survival function increases, MSE and bias become increased.

(2) Three estimators for MLRF tend to underestimate as either of censoring propotion is increased or inverse quantile of survival function S_F increased.

(3) The estimator $\widehat{e}_M^{SV}(x)$ may slightly underestimate and the estimator $\widehat{e}_M^{NA}(x)$ may be slightly overestimated.

(4) As sample increases, the estimator $\widehat{e}_M^{KM}(x)$ has a tendency to have positive bias near at right.

A change of censoring distribution from exponential to uniform, or a change of true distribution from exponential to Weibull gives no essential change in results.

Table 2. Comparisons of MSE's for $\widehat{e}_M^{KM}(x)$, $\widehat{e}_M^{SV}(x)$ and $\widehat{e}_M^{NA}(x)$

(a)-1 When $S_F(x) = Exp(1),\ S_G(x) = Exp(x)$ $\lambda =\ .111$

(10% censoring)

n	Estimator	x(S(x))								
			.357(.7)			.693(.5)			.916(.4)	
		MEAN	BIAS	MSE	MEAN	BIAS	MSE	MEAN	BIAS	MSE
	KME[1]	.977	-.023	.056	.972	-.028	.090	.965	-.035	.125
30	SVE[2]	.963	-.037	.052	.952	-.048	.082	.951	-.049	.112
	NAE[3]	1.038	.038	.068	1.046	.046	.110	1.07	.047	.152
	KME	.998	-.002	.034	.997	-.003	.048	.89	-.011	.063
50	SVE	.987	-.013	.029	.981	-.019	.045	.972	-.028	.057
	NAE	1.046	.046	.041	1.054	.054	.058	1.055	-.055	.074
	KME	1.001	.001	.017	1.006	.006	.025	1.034	.034	.032
100	SVE	.997	-.003	.017	.996	.004	.024	.992	-.008	.030
	NAE	1.032	.032	.020	1.043	.043	.029	1.046	.046	.038

(a)-2 $\lambda = .429$ (30 % censoring).

n	Estimator	x(S(x))								
			.357(.7)			.693(.5)			.916(.4)	
		MEAN	BIAS	MSE	MEAN	BIAS	MSE	MEAN	BIAS	MSE
	KME	.918	-.082	.088	.871	-.129	.134	.843	-.157	.171
30	SVE	.882	-.118	.065	.844	-.156	.095	.836	-.164	.120
	NAE	.976	-.032	.093	.926	-.074	.138	.901	-.099	.177
	KME	.935	-.064	.050	.911	-.089	.081	.885	-.115	.113
50	SVE	.903	-.097	.044	.878	-.122	.069	.856	-.143	.091
	NAE	.976	-.024	.053	.960	-.040	.087	.939	-.061	.120
	KME	.961	-.038	.027	.948	-.052	.043	.932	-.068	.060
100	SVE	.935	-.065	.023	.918	-.082	.036	.901	-.099	.048
	NAE	.990	-.010	.029	.983	-.014	.046	.973	-.027	.063

1) KME = $\widehat{e}_M^{KM}(x)$ 2) SVE = $\widehat{e}_M^{SV}(x)$ 3) NAE = $\widehat{e}_M^{NA}(x)$

Table 2 (continued)

(b)-1 When $S_F(x) = Exp(1),\ S_G(x) = Unif(\lambda)$ $\lambda = 9.9$
(10% censoring)

n	Estimator	x(S(x))								
		.357(.7)			.693(.5)			.916(.4)		
		MEAN	BIAS	MSE	MEAN	BIAS	MSE	MEAN	BIAS	MSE
30	KME[1]	.968	-.032	.049	.961	-.039	.082	.950	-.050	.010
	SVE[2]	.954	-.046	.046	.947	-.053	.074	.937	-.063	.091
	NAE[3]	1.025	.025	.056	1.030	.030	.095	1.026	.026	.121
50	KME	.990	-.010	.034	.985	-.015	.039	.980	-.020	.042
	SVE	.977	-.023	.031	.970	-.030	.045	.963	-.037	.062
	NAE	1.034	.034	.039	1.040	.040	.060	1.042	.042	.084
100	KME	1.005	.005	.017	1.005	.005	.017	1.005	.005	.031
	SVE	.993	.007	.016	.994	-.006	.016	.991	.001	.033
	NAE	1.028	.028	.019	1.038	-.038	.019	1.041	.041	.036

(b)-2 $\lambda = 3.3$ (30% censoring).

n	Estimator	x(S(x))								
		.357(.7)			.693(.5)			.916(.4)		
		MEAN	BIAS	MSE	MEAN	BIAS	MSE	MEAN	BIAS	MSE
30	KME	.838	-.162	.071	.761	-.239	.117	.694	-.306	.160
	SVE	.846	-.154	.060	.788	-.212	.093	.744	-.256	.115
	NAE	.869	-.130	.064	.794	-.206	.106	.725	-.275	.146
50	KME	.857	-.143	.049	.800	-.200	.078	.751	-.249	.106
	SVE	.852	-.148	.046	.809	-.191	.068	.773	-.227	.088
	NAE	.881	-.119	.044	.827	-.173	.069	.780	-.220	.095
100	KME	.886	-.113	.026	.843	-.157	.044	.805	-.195	.061
	SVE	.890	-.110	.024	.852	-.148	.038	.822	-.179	.051
	NAE	.901	-.099	.044	.861	-.139	.040	.824	-.176	.055

1) KME = $\widehat{e}_M^{KM}(x)$ 2) SVE = $\widehat{e}_M^{SV}(x)$ 3) NAE = $\widehat{e}_M^{NA}(x)$

Table 2 (continued)

(c)-1 When $S_F(x) = Weib(1.15, 2)$, $S_G(x) = Exp(\lambda)$ $\lambda = .14$
(10% censoring)

n	Estimator	x(S(x))								
		.517(.7)			.721(.5)			.829(.4)		
		MEAN	BIAS	MSE	MEAN	BIAS	MSE	MEAN	BIAS	MSE
30	KME[1]	.433	-.004	.005	.363	-.002	.007	.333	-.002	.009
	SVE[2]	.430	-.007	.007	.360	-.006	.007	.330	-.005	.008
	NAE[3]	.452	.015	.006	.385	.019	.008	.357	.022	.011
50	KME	.438	.001	.003	.368	-.002	.004	.337	.002	.004
	SVE	.436	-.001	.003	.366	.000	.004	.335	.000	.005
	NAE	.452	.015	.004	.384	.018	.005	.385	.018	.006
100	KME	.438	.001	.002	.365	-.001	.003	.340	.005	.002
	SVE	.437	.000	.002	.366	.000	.002	.339	.004	.002
	NAE	.446	.008	.002	.380	.014	.002	.351	.015	.003

(c)-2 $\lambda = .49$ (30% censoring).

n	Estimator	x(S(x))								
		.517(.7)			.721(.5)			.829(.4)		
		MEAN	BIAS	MSE	MEAN	BIAS	MSE	MEAN	BIAS	MSE
30	KME	.426	-.011	.008	.359	-.006	.010	.334	-.003	.013
	SVE	.415	-.016	.007	.346	-.019	.009	.322	-.014	.010
	NAE	.456	.019	.009	.383	.017	.012	.359	.022	.015
50	KME	.432	-.005	.005	.362	-.004	.006	.330	-.005	.007
	SVE	.413	-.013	.004	.353	-.013	.006	.322	-.014	.007
	NAE	.448	.011	.005	.381	.015	.007	.351	.014	.009
100	KME	.436	-.001	.002	.367	-.001	.003	.336	.001	.003
	SVE	.431	-.006	.002	.361	-.005	.003	.330	-.005	.003
	NAE	.447	-.009	.002	.379	.013	.003	.342	.007	.004

1) KME= $\widehat{e}_M^{KM}(x)$ 2) SVE = $\widehat{e}_M^{SV}(x)$ 3) NAE = $\widehat{e}_M^{NA}(x)$

4. An Example

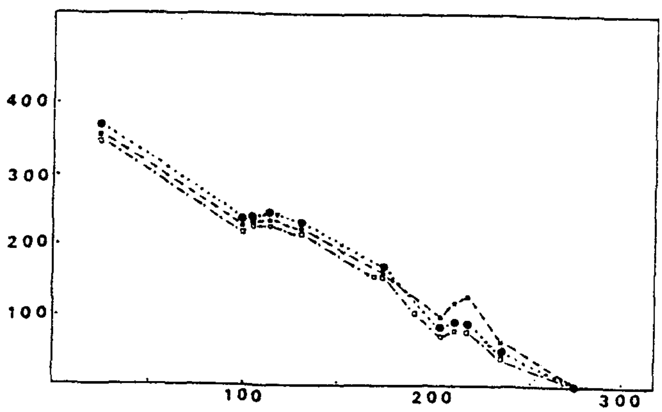
The data used for an illustration are cited from Appendix 1 of Kalbfleisch and Prentice(1980). Approximately 30% of the survival times are censored owing primarily to patients surviving to the time of analysis. Some patients were lost to follow-up because the patient moved or transferred to an institution not participating in the study, though these cases were relatively rare.

Figure 1 shows the data from a clinical trail in the treatment of carcimona of the orthpharynx. From Figure 2, one can see the curves of estimated MRLF's, *i.e.*, $\widehat{e}_M^{KM}(x)$, $\widehat{e}_M^{SV}(x)$ and $\widehat{e}_M^{NA}(x)$.

1	666	1	477	1	308	1	726	1	310
0	1089	0	932	0	1095	0	731	1	.238
0	593	1	446	1	553	1	532	0	154
1	369	1	107	0	854	1	513	0	914
1	105	0	600	1	317	1	407	1	346
1	518	1	395	1	608	1	324	1	275
0	546	1	112	0	182	1	209	1	208
1	174	1	291	0	723	1	498	1	213
1	38	1	128						

* 1 and 0 are descrived uncensored and censored, respectively.

Figure 1. A Clinical Trial in the Treatment of Carcimona of the Orthpharynx (Female 42).



* ● ··; □ · —; ■ — — are stand for $\widehat{e}^{KM}$, $\widehat{e}^{SV}$ and $\widehat{e}^{NA}$, respectively.

Figure 2. Mean Residual Life Times

References

1. Aalen, O. (1976). Nonparametric Inference Inconnect with Multiple Decrement Models. *Scandinavian Journal of Statistics,* 3, 15-17.
2. Burke, M.D., Corgo,S. and Horvath,L. (1981). Strong Approximations of Some Biometric Estimatoıs under Random Censorship. Z.Wahrschein-lichkeitstheoric verw. *Gebiete* 56, 87-112, Springer Verlag.
3. Cox, D.R. (1961). *Renewal Theory.* Methuen, London.
4. Kabfleisch, T.D. and Prentice, R.L. (1980). *The Statistical Analysis of Failure Time Data.* Wiley, New York.
5. Kaplan, E.L. and Meier, P (1958). Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association,* 53, 457-481.
6. Nelson, W. (1972). Theory and Application of Hazard Plotting for Censored Failure Data. *Technometrics,* 14, 945-966.
7. Susarla, V. and Van Ryzin, J. (1980). Large Sample Theory for an Estimator of the Mean Survival Time from Censored Samples, *The Annals of Statistics,* 8, 1002-1016.
8. Yang,G. (1976). Estimation of a Biometric Function. *The Annals of Statistics,* 8, 1002-1016.
9. Yang,G. (1977). Life Expectancy under Random Censorship. *Stochastic Processes and their Applications,* 6, 33-3.