European Journal of Scientific Research
ISSN 1450-216X Vol.32 No.3 (2009), pp.391-397

http://www.eurojournals.com/ejsr.htm

Properties of Kaplan-Meier Estimator: Group Comparison of Survival Curves

Atif Akbar
Corresponding Author, Department of Statistics
Bahauddin Zakariya University, Multan, Paksitan
E-mail: atifakber@yahoo.co.uk

G. R. Pasha
Department of Statistics, Bahauddin Zakariya University, Multan, Paksitan

Syeda Farrukh Huma Naqvi
Department of Statistics, Bahauddin Zakariya University, Multan, Paksitan

Abstract

Kaplan and Meier (1958) first brought solution of a problem to estimate the survival curve in a simple way while considering the right censored data. Results from simulation study, with reference to baseline hazards of three distributions i.e. gamma, weibull and lognormal distributions, gives that for small sample size and high censoring rate the Kaplan-Meier (KM) estimator gives less reliable results. For high censoring rate with large sample size, we found reliable results for KM as established for moderate sample sizes with low censoring rate. For group comparison of survival curves, we used Log-rank test. The significant difference of groups of survival curves were found for large sample size with higher censoring rate and for moderate sample sizes with lower censoring rate and also results were confirmed with Log-rank test and Wilcoxon test.

Keywords: Right censored data, survival curves, Kaplan-Meier estimator, Log-rank test, Wilcoxon test.

1. Introduction

Kaplan and Meier (1958) were the first who carried out the solution of a problem to estimate the survival curve in a simple way while considering the right censoring. It calculated as how many subjects in the sample survived to time t in percentage.

Kaplan-Meier (KM) or product limit (PL) estimate has found very useful in survival analysis. Its applications found in survival analysis by many statisticians.

The Kaplan-Meier or Product Limit estimator $\hat{S}(t)$ of the survival function $S(t) = P(T > t)$ can be defined as (Cox and Oakes, 1984; Rosner, 1995, etc.).

$$\hat{S}(t) = \prod_{t_i \le t} \left(1 - \frac{d_i}{n_i}\right) \tag{1.1}$$

Where t_i are distinct times $t = t_i$, n_i be the size of risk set at time t_i For $i < r$, $n_{i+1} = n_i - d_i - c_i$, d_i is the number who failed at time $t = t_i$ and c_i is the number who are censored at time t with $t_i \le t \le t_{i+1}$.

Chen et al (1982) obtained exact small sample results for KM estimator under a model of proportional hazards and found the bias of KM estimator which was not large under proportional hazard model. Rossa and Zielinski (2002) discussed the disadvantage of KM estimator associated with small and moderate sample size. They proved that mean square error and mean absolute deviation of there proposed estimator was significantly smaller than original KM estimator. Very little work was found on the discussion of small sample properties of KM estimator.

Discussions on the main assumptions of KM estimator and their violation were carried out by many authors like completion methods used to complete the KM estimator beyond the last observation when it is censored was studied by Efron (1967), Gill (1980) and Chen and Phadia (2006).

Log-Rank test is used to test the difference between survival curves of two groups or to check that the curves are statistically equivalent. Log-rank test can only be used for comparing groups defined by categorical covariates. It gives the bad performance when the two survival functions are crossing over. The Log-rank test is a large-sample chi-square test that used as its test criterion compares a statistic that provides an overall comparison of the Kaplan-Meier curves. The Log-rank test is based on a comparison of the observed number of deaths in group 1, and the expected if the survival were equal in the two groups. The Log-Rank teat statistics is (see Kleinbaum, 1995 and Klein and Moeschberger, 1997):

$$\text{Log-rank statistic} = \frac{(O_1 - E_1)^2}{\operatorname{var}(O_1 - E_1)} \tag{1.2}$$

$$\text{Where } \operatorname{var}(O_1 - E_1) = \frac{n_1 \times n_2 \times d \times (n-d)}{n^2 \times (d-1)} \tag{1.3}$$

The Log-Rank statistic approaches to chi-square distribution with 1 degree of freedom.

2. Simulation

This simulation study is made up from the simulation schemes of Heller and Simonoff (1992) and Xie and Liu (2000). Survival time is generated from $t_i = (\beta_1 X_{i1} + \beta_2 X_{i2}) \times d_i$, where $i = 1,2,...,n$. Here d has taken as lognormal (0,1), gamma (1,1) and weibull (1,1). X_1 is group indicator variable and taking values 0 and 1 generated from Brenoulli distribution with 0.5 probabiility, X_2 has taken from Uniform (0,1) and $\beta_1 = \beta_2 = 1$. To deal with different level of censoring C_i, censoring times are generated from exponential (λ = 1.5, 2.5, 3.5, 4.5). To illustrate the method of Cox PH model, the time is right censored and consists of $Z_i = \min(t_i, c_i)$ and indicator variable of uncensored and censored survival time as $\delta_i = 0$ when $t_i \le c_i$ and $\delta_i = 1$ when $t_i > c_i$.

There are two groups generated from X_1 variable which has taken values
$X_1 = 0$ for first group and $X_1 = 1$ for second group.
We operate simulation study 1000 times and use MINITAB 13.1.

3. Data Analysis

The following section presents the data manipulation for simulated data as generated by the scheme note above.

3.1. Kaplan- Meier Results

We present the necessary calculations for Kaplan-Meier method with reference to baseline hazards of three distributions. The calculated results are given for various sample sizes i.e. n = 25, 50, 100, 200 and 500 in following tables.

Table 3.1: Kaplan-Meier results for gamma (1, 1) and for censoring exponential (λ = 1.5, 2.5, 3.5, 4.5)

	n=25		n=50		n=100		n=200		n=500	
	X_1=0	X_1=1	X_1=0	X_1=1	X_1=0	X_1=1	X_1=0	X_1=1	X_1=0	X_1=1
Censoring: Exponential (1.5)										
C	24%	44%	22%	40%	29%	35%	25%	37%	25%	40%
UC	20%	12%	25%	12%	24%	13%	25%	13%	23%	12%
Mean	1.093	1.397	1.1270	2.1600	1.4898	2.1930	1.5489	2.5790	1.554	3.293
SE	0.2660	0.256	0.1930	0.2784	0.2140	0.2351	0.2027	0.2347	0.221	0.248
Median	1.019	2.154	1.1390	2.8850	1.3040	2.6970	1.0593	3.0410	1.138	3.191
Censoring: Exponential (2.5)										
C	18%	34%	20%	33%	22%	29%	19%	31%	20%	33%
UC	28%	22%	28%	20%	31%	20%	31%	19%	28%	18%
Mean	2.22	2.073	1.2257	3.1780	1.5020	2.9870	1.7131	3.4160	1.651	3.780
SE	0.2802	0.3673	0.2231	0.5690	0.1972	0.3834	0.2241	0.3632	0.211	0.273
Median	1.431	2.574	0.9831	3.0300	1.2174	2.9940	1.0851	3.1780	1.138	3.186
Censoring: Exponential (3.5)										
C	16%	30%	9%	32%	18%	27%	17%	28%	16%	28%
UC	34%	24%	36%	23%	36"%	22%	33%	22%	32%	24%
Mean	1.322	3.182	1.3330	3.3390	1.5845	3.5910	1.6353	3.6320	1.683	4.057
SE	0.3286	0.591	0.2816	0.6080	0.2221	0.9170	0.1969	0.4029	0.134	0.338
Median	0.999	4.161	1.0137	3.6830	1.2716	3.1520	1.1135	3.0860	1.099	3.024
Censoring: Exponential (4.5)										
C	12%	28%	13%	28%	14%	25%	13%	25%	13%	26%
UC	32%	28%	36%	27%	38%	24%	36%	26%	35%	27%
Mean	1.396	4.520	1.5380	3.1740	1.6860	3.7920	1.6590	4.0830	1.675	4.432
SE	0.3599	1.003	0.3525	0.4312	0.2339	0.5383	0.1941	0.4452	0.129	0.406
Median	1.078	4.302	1.3230	3.3240	1.6850	2.9560	1.0966	3.2130	1.100	2.989

* C = Censored observations, UC = Uncensored observations, SE = Standard Error

When we have taken the censoring distribution as the exponential (λ = 1.5) then for group one (i.e. X_1= 0), the censoring, for Gamma distribution, is 25%, the mean values of Kaplan-Meier estimate increase as sample size increases. For second group (i.e. X_1 = 1), the censoring now become 39% also the mean values of Kaplan-Meier estimate for second group increase as sample sizes increases. Comparatively, as we see, that the censoring rate for second group, is higher than the first group.

When censoring distribution is taken as the exponential (λ = 2.5) then for group one, the censoring, for Gamma distribution, is 20% and for second group the censoring become 32%. Same pattern of results of mean and median values for all sample sizes is followed as for exponential (λ = 1.5).

When censoring distribution is taken as the exponential (λ = 3.5) then for group one, the censoring, for Gamma distribution, is 15% and for second group the censoring become 30%. Same pattern of results of mean and median values for all sample sizes is followed as for exponential (λ = 1.5) and exponential (λ = 2.5).

For censoring distribution is taken as exponential (λ = 4.5), the censoring for group one is 12% and for group two it becomes 25%.

Same results are followed by lognormal and Weibull distributions for different sample sizes and censoring rates.

For baseline hazard distributions, i.e. lognormal, Weibull and Gamma and for the censoring distribution exponential with parameter value equals to 1.5, lognormal distribution produce maximum censoring, moderate in Weibull and minimum in gamma. And same situation arises when exponential distribution bears parameter values, 2.5, 3.5 and 4.5.

KM results are same for all censoring proportions and also for lognormal, Weibull and gamma distributions and the results are provided herein.

Table 3.2: Kaplan-Meier results for Lognormal (0, 1) and for censoring exponential (λ = 1.5, 2.5, 3.5, 4.5)

	n=25		**n=50**		**n=100**		**n=200**		**n=500**	
	X_1=0	**X_1=1**	**X_1=0**	**X_1=1**	**X_1=0**	**X_1=1**	**X_1=0**	**X_1=1**	**X_1=0**	**X_1=1**
Censoring: Exponential (1.5)										
C	32%	40%	32%	42%	31%	43%	33%	42%	32%	42%
UC	20%	8%	17%	5%	19%	8%	18%	8%	19%	7%
Mean	1.4670	2.0230	1.7270	2.4620	1.9940	2.8430	1.8010	3.3260	2.079	4.048
SE	0.2434	0.3683	0.2529	0.2472	0.2655	0.2260	0.1633	0.2620	0.177	0.240
Median	1.6100	2.7700	1.9710	3.4710	1.7160	3.3910	1.5927	4.3010	1.716	4.723
Censoring: Exponential (2.5)										
C	28%	36%	26%	36%	23%	39%	26%	36%	25%	38%
UC	24%	12%	26%	12%	27%	11%	24%	14%	25%	11%
Mean	1.6810	2.9560	2.2250	3.4980	1.8860	4.3410	2.2742	4.1070	2.321	5.062
SE	0.3303	0.4675	0.3967	0.4840	0.2271	0.4266	0.2693	0.3690	0.183	0.379
Median	1.8130	3.7760	1.9320	3.9180	1.6360	5.2220	1.7170	4.5500	1.670	4.820
Censoring: Exponential (3.5)										
C	24%	32%	24%	31%	21%	37%	21%	33%	21%	33%
UC	28%	12%	26%	19%	29%	15%	28%	18%	30%	17%
Mean	2.0450	3.5690	2.5510	4.3060	2.2160	4.6020	2.2900	85.9000	2.454	5.312
SE	0.4867	0.8420	0.5251	0.7670	0.3094	0.4896	0.2531	0.3609	0.199	0.363
Median	1.7790	4.1920	2.0040	4.1420	1.6104	5.2500	1.6578	4.1410	664.000	4.550
Censoring: Exponential (4.5)										
C	20%	28%	18%	29%	19%	30%	19%	32%	19%	30%
UC	32%	20%	33%	20%	33%	19%	30%	20%	33%	19%
Mean	2.2190	4.0980	2.3020	4.5420	2.1970	5.4280	2.5130	5.6750	2.573	5.474
SE	0.5002	0.8380	0.4145	0.6950	0.3415	0.7081	0.3113	1.0370	0.217	0.379
Median	1.8790	4.3370	1.8720	4.7700	1.5740	5.3550	1.6397	4.5160	1.713	4.313

Table 3.3: Kaplan-Meier results for Weibull (1, 1) and for censoring exponential (λ = 1.5, 2.5, 3.5, 4.5)

	n=25		n=50		n=100		N=200		n=500	
	X_1=0	X_1=1	X_1=0	X_1=1	X_1=0	X_1=1	X_1=0	X_1=1	X_1=0	X_1=1
Censoring: Exponential (1.5)										
C	30%	30%	25%	38%	29%	38%	22%	38%	26%	38%
UC	22%	14%	25%	15%	23%	12%	24%	14%	25%	13%
Mean	0.9980	1.7250	1.2835	1.9920	1.5226	2.2950	1.5222	3.1200	1.560	3.311
SE	0.1895	0.3266	0.2030	0.2841	0.1837	0.2103	0.1709	0.3486	0.139	0.260
Median	1.8470	2.1880	1.2910	2.6020	1.5690	3.2180	1.1832	3.3530	1.113	3.392
Censoring: Exponential (2.5)										
C	20%	32%	23%	34"%	22%	31%	19%	35%	19%	32%
UC	30%	18%	27%	19%	29%	19%	29%	19%	31%	19%
Mean	1.1850	1.6870	1.3640	2.4950	1.6849	3.6180	1.8600	3.3590	1.541	3.730
SE	0.2525	0.3642	0.2308	0.3770	0.2248	0.4949	0.2821	0.3005	0.118	0.295
Median	1.2580	2.2660	1.2590	3.3750	1.4316	3.5370	1.0988	3.6410	1.111	3.036
Censoring: Exponential (3.5)										
C	14%	30%	14%	29%	19%	28%	17%	28%	16%	27%
UC	40%	20%	36%	22%	32%	23%	31%	24%	34%	22%
Mean	1.4820	2.5490	1.4640	3.7110	1.6433	3.4110	1.5876	3.9760	1.662	4.003
SE	0.4099	0.4780	0.2773	0.7390	0.2203	0.4237	0.1729	0.4465	0.126	0.384
Median	1.3360	3.3380	1.1507	3.6330	1.3297	3.6530	1.1990	3.1410	1.138	3.028
Censoring: Exponential (4.5)										
C	16%	26%	14%	26%	15%	25%	12%	27%	14%	24%
UC	36%	20%	38%	24%	34%	26%	36%	25%	37%	26%
Mean	1.6170	3.0170	1.7360	3.5030	1.7943	4.3390	1.6533	4.2060	1.692	4.181
SE	0.5100	0.5500	0.3602	0.6311	0.2636	0.6498	0.1961	0.5055	0.130	0.347
Median	1.4810	3.8500	1.1490	3.5140	1.3688	3.9460	1.1694	3.2030	1.130	2.964

3.2. Group Comparison of Survival Curves

The method of Log-rank test for comparison of Kaplan-Meier survival curves of two groups with reference to baseline hazards of three distributions is illustrated by given table. The calculated results are given for various sample sizes i.e. *n* =25, 50, 100, 200 and 500 in following table,

Table 3.4: Group Comparison of Kaplan-Meier Curves for Gamma (1, 1) and for Censoring Exponential (λ = 1.5, 2.5, 3.5, 4.5)

Censoring: Exponential (1.5)					
	n=25	n=50	n=100	n=200	n=500
LR	2.835	5.7280	6.2490	18.1200	43.760
WT	2.418	4.7300	4.3200	16.5500	33.540
P(LR)	0.1881	0.0442	0.0437	0.0010	0.000
P(WT)	0.2435	0.0751	0.1204	0.0080	0.000
Censoring: Exponential (2.5)					
LR	3.029	7.4200	8.5700	19.4300	52.540
WT	3.24	6.3600	6.3000	18.8200	40.500
P(LR)	0.1785	0.0138	0.0593	0.0001	0.000
P(WT)	0.2293	0.0385	0.0858	0.0019	0.000
Censoring: Exponential (3.5)					
LR	3.974	10.7800	9.2800	23.6900	61.750
WT	2.903	7.6200	6.4400	19.9300	49.740
P(LR)	0.1243	0.0102	0.0133	0.0001	0.000
P(WT)	0.2111	0.0137	0.0414	0.0019	0.000
Censoring: Exponential (4.5)					
LR	5.48	9.5500	11.5000	28.6400	65.870
WT	4.39	9.0100	8.4000	24.5100	53.120
P(LR)	0.0676	0.0251	0.0041	0.0003	0.000
P(WT)	0.1286	0.0131	0.0197	0.0014	0.000

When we have taken the censoring distribution as the exponential (λ = 1.5), for group comparison i.e. to test the hypothesis for homogeneity of survival curves of group one and group two; the p-values of log-rank test and Wilcoxon test decrease gradually with the increment of sample size and become 0.000 as sample size become 500. Moreover the p-values for Wilcoxon test decrease rapidly than log-rank test while increasing sample size. For sample sizes 25, 50 and 100 the hypotheses for equality of groups' survival curves are non-significant and significant for sample sizes 200 and 500.

When censoring distribution is taken as the exponential (λ = 2.5) then for sample sizes 25 and 50 the hypotheses for homogeneity of groups' survival curves are non-significant and significant for sample sizes 100, 200 and 500.

When censoring distribution is taken as the exponential (λ = 3.5) then for sample sizes 25 and 50 the hypotheses for homogeneity of groups' survival curves are non-significant and significant for sample sizes 100, 200 and 500.

For censoring distribution exponential (λ = 4.5), for sample sizes 25 the hypotheses for homogeneity of groups' survival curves are non-significant and significant for sample sizes 50,100, 200 and 500.

Same results are followed by lognormal and weibull distributions for all sample sizes and censoring rates.

The results for Weibull and lognormal distributions are summarized as below.

Table 3.5: Group Comparison of Kaplan-Meier Curves for Lognormal (0, 1) and for Censoring Exponential (λ = 1.5, 2.5, 3.5, 4.5)

Censoring: Exponential (1.5)					
	n=25	**n=50**	**n=100**	**n=200**	**n=500**
LR	2.9420	4.9000	10.0600	19.4600	47.250
WT	2.6910	4.2700	10.3300	17.9100	41.510
P(LR)	0.2216	0.2350	0.0277	0.0024	0.000
P(WT)	0.2177	0.1774	0.0137	0.0009	0.000
Censoring: Exponential (2.5)					
LR	3.8610	6.3100	15.2900	18.8600	63.250
WT	3.4310	5.8000	14.7600	21.3900	60.900
P(LR)	0.2486	0.1933	0.0018	0.0004	0.000
P(WT)	0.2055	0.0992	0.0015	0.0001	0.000
Censoring: Exponential (3.5)					
LR	4.3850	4.7700	16.3900	24.0200	70.870
WT	3.9120	4.3900	15.5500	23.5800	70.580
P(LR)	0.2051	0.2154	0.0068	0.0000	0.000
P(WT)	0.1659	0.2110	0.0034	0.0001	0.000
Censoring: Exponential (4.5)					
LR	4.0010	7.8800	19.0000	27.5400	70.370
WT	3.9750	7.1100	18.1300	26.9300	71.780
P(LR)	0.2191	0.0881	0.0004	0.0004	0.000
P(WT)	0.1808	0.0667	0.0008	0.0001	0.000

Table 3.6: Group Comparison of Kaplan-Meier Curves for Weibull (1, 1) and for Censoring Exponential (λ = 1.5, 2.5, 3.5, 4.5)

	n=25	n=50	n=100	n=200	n=500
Censoring: Exponential (1.5)					
LR	1.8720	3.9700	7.3300	15.1000	40.650
WT	1.8030	3.0130	4.5900	10.9500	30.840
P(LR)	0.3850	0.1384	0.0996	0.0339	0.000
P(WT)	0.4330	0.2870	0.1356	0.0388	0.000
Censoring: Exponential (2.5)					
LR	2.7110	5.2800	10.2000	20.9400	61.350
WT	2.1150	4.6800	6.2800	16.1000	46.030
P(LR)	0.2425	0.1477	0.0083	0.0021	0.000
P(WT)	0.3067	0.1713	0.0583	0.0051	0.000
Censoring: Exponential (3.5)					
LR	3.0730	7.2000	10.4400	25.0100	57.130
WT	2.5860	5.0300	7.1200	18.4000	43.820
P(LR)	0.2880	0.0326	0.0088	0.0006	0.000
P(WT)	0.3560	0.1053	0.0407	0.0023	0.000
Censoring: Exponential (4.5)					
LR	2.9900	7.2900	10.1500	25.0300	67.960
WT	2.2760	5.3200	7.2100	20.9100	52.710
P(LR)	0.1941	0.0769	0.0116	0.0001	0.000
P(WT)	0.3270	0.1162	0.0424	0.0009	0.000

For baseline hazard distributions, i.e. lognormal, Weibull and Gamma and for the censoring distribution exponential with parameter value equals to 1.5, lognormal distribution produce maximum censoring, moderate in Weibull and minimum in gamma. And same situation arises when exponential distribution bears parameter values, 2.5, 3.5 and 4.5.

References

[1] Chen, Y. Y., Hollander, M. and Langberg, N. A. (1982). Small Sample Results for Kaplan-Meier Estimator. Journal of American Statistical Association, 77, No. 377, pp. 141-144.

[2] Chen, Z. and Phadia, E. (2006). An Optimal Completion of the Product Limit Estimator. Statistics and Probability Letters, 76, pp. 913-919.

[3] Cox, D. R. and Oakes, D. (1984). Analysis of Survival Data. London: Chapman and Hall.

[4] Efron, B. (1967). The Two Sample Problem with Censored Data. Proccedding of the 5th Brekeley Symposium, 4, pp. 831-852.

[5] Gill, R. D. (1980). Censoring and Stochastic Integrals, Math. Centre Tract. 124. Amsterdam: Mathematisch Centrum.

[6] Heller, G. and Simonoff, J. S. (1992). Prediction in Censored Survival Data. A Comparison of the Proportional Hazards and Linear Regression Models, Biometrics, 48, pp. 101-115.

[7] Kaplan, E. L. and Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. JASA, 53, pp. 457-481.

[8] Klein, J. and Moeschberger, M. (1997). Survival Analysis. Techniques for Censored and Truncated data. New York. Springer-Verlag Inc.

[9] Kleinbaum, G. D. (1995). Survival analysis. A Self-Learning Text. Springer-Verlag, New York, Inc.

[10] Rosner, B. (1995). Fundamentals of Biostatistics. 4th ed. Belmont, California: Duxbury Press.

[11] Rossa, A. and Zielinski, R. (2002). A Simple Improvement of the Kaplan-Meier Estimator. Communications is Statistics-Theory and Methods, 31 (1), pp. 147-158.

[12] Xie, J. and Liu, C. (2000). Adjusted Kaplan-Meier Estimator and Log-Rank Test with Inverse Probability of Treatment Weighting for Survival Data. Statist. Med, 00, 1-6.