Independent setting

2019-12-15

Example setting:

The survival time follows an exponential distribution $T \sim EXP(\theta)$, the censoring time follows an extreme distribution with parameter θ. And T and C are independent, random censoring.

$$P(T > x) = S_t(x) = e^{-\theta x}, \quad P(C > y) = S_c(y) = e^{-(e^{\theta y}-1)}$$

Then

$$f_T(t) = \frac{\partial}{\partial t}(1 - S_T(t)) = \frac{\partial}{\partial t}(1 - e^{-\theta t}) = \theta e^{-\theta t}$$
$$f_{t,c}(x, y) = \theta^2 e^{-e^{\theta c}+\theta c-\theta t+1}$$
$$S_Z(t) = P(T > t, C > t) = e^{-\theta t} e^{-(e^{\theta t}-1)} = e^{-e^{\theta t}-\theta t+1}$$
$$f_Z(t) = \frac{\partial}{\partial t}(1 - S_Z(t)) = 1 - e^{-e^{\theta t}-\theta t+1} = \theta(1 + e^{\theta t})e^{-e^{\theta t}-\theta t+1}$$
$$\psi(t) = \int_t^\infty f(t, c)dc = \int_t^\infty \theta^2 e^{-e^{\theta c}+\theta c-\theta t+1} dc = \theta e^{-e^{\theta t}-\theta t+1}$$

Therefore, the $m()$ function is:

$$m(t) = \frac{\lambda_F(t)}{\lambda_H(t)} = \frac{f_T(t)}{S_T(t)} / \frac{f_Z(t)}{S_Z(t)} = \frac{\theta e^{-\theta t}}{e^{-\theta t}} / \frac{\theta(1 + e^{\theta t})e^{-e^{\theta t}-\theta t+1}}{e^{-e^{\theta t}-\theta t+1}} = \frac{1}{1 + e^{\theta t}}$$

And for the $\rho()$ function,

$$\begin{aligned}\rho =& \frac{f(t)/\psi(t) - 1}{S(t)/S_x(t) - 1} \\ =& \frac{\theta e^{-\theta t}/(\theta e^{-e^{\theta t}-\theta t+1}) - 1}{e^{-\theta t}/e^{-e^{\theta t}-\theta t+1} - 1} \\ =& 1\end{aligned}$$

For the censoring percentage

$$\begin{aligned}P(T < C) =& \int_0^\infty \int_0^y f_{t,c}(x, y)dxdy \\ =& \int_0^\infty \int_0^y \theta^2 e^{-e^{\theta y}+\theta y-\theta x+1} dxdy \\ =& \int_0^\infty \theta(e^{\theta y} - 1)e^{1-e^{\theta y}} \\ =& 1 - e\Gamma(0, 1) \\ \approx& 0.4\end{aligned}$$

For the simulation, 500 subjects were generated. Their survival time T and censoring time C were randomly generated from the above exponential distribution and extreme distribution. The observed time Z and status δ are then calculated

$$Z = T \wedge C, \delta = I(T < C)$$

500 iterations were taken.

The estimation of $90\%, 75\%, 50\%, 25\%, 10\%$ quantiles were reported.

Results

The mean value of the estimations of the quantiles of the KM, exponential m(), Dikta method 1 and Dikta method 2 were calculated.

The cells in the table are the mean value of the estimations over 500 iterations minus the true quantiles value i.e (0.9, 0.75, 0.5, 0.25, 0.1)

Table 1: Mean absolute difference between estimated and true S()

	With true m()				With estimated m()			
Quantile	KM	Exp m()	Dikta 1	Dikta 2	KM	Exp m()	Dikta 1	Dikta 2
theta = 0.2								
t0.1	0.01185	0.00814	0.00815	0.00815	0.01185	0.01077	0.01080	0.01081
t0.25	0.01708	0.01134	0.01135	0.01136	0.01708	0.01698	0.01690	0.01685
t0.5	0.02335	0.01438	0.01438	0.01442	0.02335	0.02335	0.02315	0.02303
t0.75	0.04036	0.01970	0.02079	0.02221	0.04036	0.04381	0.04843	0.05172
t0.9	0.09632	0.06909	0.10000	0.10000	0.09632	0.01688	0.10000	0.10000
theta = 0.8								
t0.1	0.01181	0.00814	0.00814	0.00815	0.01181	0.01073	0.01075	0.01076
t0.25	0.01712	0.01139	0.01141	0.01142	0.01712	0.01702	0.01694	0.01689
t0.5	0.02339	0.01443	0.01444	0.01447	0.02339	0.02346	0.02326	0.02314
t0.75	0.04045	0.01974	0.02076	0.02215	0.04045	0.04410	0.04862	0.05190
t0.9	0.09611	0.06904	0.10000	0.10000	0.09611	0.01701	0.10000	0.10000
theta = 1								
t0.1	0.01181	0.00814	0.00814	0.00815	0.01181	0.01073	0.01075	0.01076
t0.25	0.01712	0.01139	0.01141	0.01142	0.01712	0.01702	0.01694	0.01689
t0.5	0.02339	0.01443	0.01444	0.01447	0.02339	0.02346	0.02326	0.02314
t0.75	0.04045	0.01975	0.02076	0.02215	0.04045	0.04415	0.04862	0.05190
t0.9	0.09611	0.06902	0.10000	0.10000	0.09611	0.01701	0.10000	0.10000
theta = 2								
t0.1	0.01181	0.00814	0.00814	0.00815	0.01181	0.01073	0.01075	0.01076
t0.25	0.01712	0.01139	0.01141	0.01142	0.01712	0.01702	0.01694	0.01689
t0.5	0.02339	0.01443	0.01444	0.01447	0.02339	0.02346	0.02326	0.02314
t0.75	0.04045	0.01978	0.02076	0.02215	0.04045	0.04419	0.04862	0.05190
t0.9	0.09611	0.06900	0.10000	0.10000	0.09611	0.01702	0.10000	0.10000
theta = 5								
t0.1	0.01175	0.00814	0.00814	0.00815	0.01175	0.01069	0.01072	0.01073
t0.25	0.01709	0.01136	0.01138	0.01139	0.01709	0.01698	0.01690	0.01685
t0.5	0.02339	0.01446	0.01446	0.01449	0.02339	0.02345	0.02326	0.02314
t0.75	0.04042	0.01993	0.02080	0.02219	0.04042	0.04434	0.04861	0.05188
t0.9	0.09641	0.06897	0.10000	0.10000	0.09641	0.01700	0.10000	0.10000

To make the table easy to look at, I used the column 2,3,4,5 to divide the column 1, column 7,8,9,10 to divide column 6.

The values that are less than 1 are showing that the methods have less bias than the KM.

Table 2: Mean absolute difference between estimated and true S()

	With true m()				With estimated m()			
Quantile	KM	Exp m()	Dikta 1	Dikta 2	KM	Exp m()	Dikta 1	Dikta 2

theta = 0.2								
t0.1	1	0.68687	0.68756	0.68791	1	0.90901	0.91119	0.91210
t0.25	1	0.66387	0.66472	0.66523	1	0.99416	0.98926	0.98639
t0.5	1	0.61576	0.61604	0.61744	1	1.00011	0.99168	0.98655
t0.75	1	0.48809	0.51510	0.55012	1	1.08536	1.19975	1.28138
t0.9	1	0.71734	1.03824	1.03824	1	0.17527	1.03824	1.03824
theta = 0.8								
t0.1	1	0.68933	0.68988	0.69022	1	0.90876	0.91077	0.91166
t0.25	1	0.66568	0.66659	0.66708	1	0.99433	0.98957	0.98672
t0.5	1	0.61695	0.61741	0.61881	1	1.00316	0.99462	0.98953
t0.75	1	0.48799	0.51332	0.54766	1	1.09026	1.20186	1.28290
t0.9	1	0.71833	1.04048	1.04048	1	0.17703	1.04048	1.04048
theta = 1								
t0.1	1	0.68957	0.68988	0.69022	1	0.90877	0.91077	0.91166
t0.25	1	0.66568	0.66659	0.66708	1	0.99434	0.98957	0.98672
t0.5	1	0.61707	0.61741	0.61881	1	1.00304	0.99462	0.98953
t0.75	1	0.48815	0.51332	0.54766	1	1.09132	1.20186	1.28290
t0.9	1	0.71814	1.04048	1.04048	1	0.17702	1.04048	1.04048
theta = 2								
t0.1	1	0.68969	0.68988	0.69022	1	0.90863	0.91077	0.91166
t0.25	1	0.66549	0.66659	0.66708	1	0.99433	0.98957	0.98672
t0.5	1	0.61707	0.61741	0.61881	1	1.00304	0.99462	0.98953
t0.75	1	0.48900	0.51332	0.54766	1	1.09246	1.20186	1.28291
t0.9	1	0.71791	1.04048	1.04048	1	0.17704	1.04048	1.04048
theta = 5								
t0.1	1	0.69280	0.69304	0.69337	1	0.91004	0.91204	0.91294
t0.25	1	0.66499	0.66597	0.66646	1	0.99405	0.98933	0.98647
t0.5	1	0.61802	0.61799	0.61934	1	1.00260	0.99427	0.98926
t0.75	1	0.49317	0.51470	0.54893	1	1.09699	1.20270	1.28374
t0.9	1	0.71539	1.03727	1.03727	1	0.17634	1.03727	1.03727

The standard deviation of the estimations of each quantiles are reported in the following table.

Table 3: Standard deviations of the estimated S()

	With true m()				With estimated m()			
Quantile	KM	Exp m()	Dikta 1	Dikta 2	KM	Exp m()	Dikta 1	Dikta 2
theta = 0.2								
t0.1	0.01458	0.01000	0.01001	0.01001	0.01458	0.01316	0.01317	0.01317
t0.25	0.02120	0.01421	0.01423	0.01424	0.02120	0.01886	0.01889	0.01889
t0.5	0.02954	0.01812	0.01818	0.01822	0.02954	0.02699	0.02706	0.02708
t0.75	0.04984	0.02488	0.02677	0.02846	0.04984	0.03878	0.04027	0.04125
t0.9	0.09187	0.00869	0.00000	0.00000	0.09187	0.02130	0.00000	0.00000
theta = 0.8								
t0.1	0.01454	0.01000	0.01001	0.01002	0.01454	0.01313	0.01314	0.01315
t0.25	0.02125	0.01425	0.01427	0.01428	0.02125	0.01890	0.01893	0.01894
t0.5	0.02974	0.01820	0.01827	0.01831	0.02974	0.02723	0.02729	0.02731
t0.75	0.05002	0.02504	0.02675	0.02844	0.05002	0.03918	0.04053	0.04150
t0.9	0.09175	0.00870	0.00000	0.00000	0.09175	0.02156	0.00000	0.00000
theta = 1								
t0.1	0.01454	0.01001	0.01001	0.01002	0.01454	0.01313	0.01314	0.01315

t0.25	0.02125	0.01425	0.01427	0.01428	0.02125	0.01890	0.01893	0.01894
t0.5	0.02974	0.01820	0.01827	0.01831	0.02974	0.02723	0.02729	0.02731
t0.75	0.05002	0.02505	0.02675	0.02844	0.05002	0.03919	0.04053	0.04150
t0.9	0.09175	0.00870	0.00000	0.00000	0.09175	0.02156	0.00000	0.00000
theta = 2								
t0.1	0.01454	0.01001	0.01001	0.01002	0.01454	0.01313	0.01314	0.01315
t0.25	0.02125	0.01425	0.01427	0.01428	0.02125	0.01890	0.01893	0.01894
t0.5	0.02974	0.01821	0.01827	0.01831	0.02974	0.02723	0.02729	0.02731
t0.75	0.05002	0.02507	0.02675	0.02844	0.05002	0.03921	0.04053	0.04150
t0.9	0.09175	0.00871	0.00000	0.00000	0.09175	0.02156	0.00000	0.00000
theta = 5								
t0.1	0.01449	0.01000	0.01000	0.01001	0.01449	0.01310	0.01311	0.01312
t0.25	0.02120	0.01422	0.01424	0.01425	0.02120	0.01889	0.01891	0.01892
t0.5	0.02970	0.01823	0.01828	0.01832	0.02970	0.02727	0.02732	0.02734
t0.75	0.04997	0.02525	0.02682	0.02850	0.04997	0.03935	0.04056	0.04154
t0.9	0.09179	0.00878	0.00000	0.00000	0.09179	0.02152	0.00000	0.00000

To make the table easy to look at, I used the column 2,3,4,5 to divide the column 1, column 7,8,9,10 to divide column 6.

The values that are less than 1 are showing that the methods have less standard deviation than the KM.

Table 4: Standard deviations of the estimated S()

	With true m()				With estimated m()			
Quantile	KM	Exp m()	Dikta 1	Dikta 2	KM	Exp m()	Dikta 1	Dikta 2
theta = 0.2								
t0.1	1	0.68562	0.68639	0.68679	1	0.90237	0.90332	0.90359
t0.25	1	0.67032	0.67129	0.67182	1	0.89000	0.89108	0.89142
t0.5	1	0.61345	0.61557	0.61694	1	0.91394	0.91603	0.91687
t0.75	1	0.49924	0.53720	0.57108	1	0.77812	0.80805	0.82771
t0.9	1	0.09458	0.00000	0.00000	1	0.23181	0.00000	0.00000
theta = 0.8								
t0.1	1	0.68777	0.68845	0.68884	1	0.90291	0.90380	0.90407
t0.25	1	0.67055	0.67156	0.67209	1	0.88969	0.89078	0.89112
t0.5	1	0.61215	0.61430	0.61565	1	0.91572	0.91774	0.91858
t0.75	1	0.50065	0.53493	0.56859	1	0.78344	0.81032	0.82979
t0.9	1	0.09477	0.00000	0.00000	1	0.23499	0.00000	0.00000
theta = 1								
t0.1	1	0.68806	0.68845	0.68884	1	0.90305	0.90380	0.90407
t0.25	1	0.67059	0.67156	0.67209	1	0.88964	0.89078	0.89112
t0.5	1	0.61222	0.61430	0.61565	1	0.91574	0.91774	0.91859
t0.75	1	0.50074	0.53493	0.56859	1	0.78354	0.81032	0.82979
t0.9	1	0.09478	0.00000	0.00000	1	0.23499	0.00000	0.00000
theta = 2								
t0.1	1	0.68808	0.68845	0.68884	1	0.90300	0.90380	0.90407
t0.25	1	0.67050	0.67156	0.67209	1	0.88958	0.89078	0.89112
t0.5	1	0.61225	0.61430	0.61565	1	0.91576	0.91774	0.91859
t0.75	1	0.50121	0.53493	0.56859	1	0.78401	0.81032	0.82979
t0.9	1	0.09495	0.00000	0.00000	1	0.23500	0.00000	0.00000
theta = 5								
t0.1	1	0.68997	0.69026	0.69066	1	0.90448	0.90513	0.90540

t0.25	1	0.67105	0.67196	0.67248	1	0.89112	0.89212	0.89246
t0.5	1	0.61389	0.61551	0.61687	1	0.91823	0.91977	0.92062
t0.75	1	0.50537	0.53671	0.57040	1	0.78746	0.81167	0.83121
t0.9	1	0.09566	0.00000	0.00000	1	0.23449	0.00000	0.00000

The MSE of each estimation

Table 5: MSE

	With true m()				With estimated m()			
Quantile	KM	Exp m()	Dikta 1	Dikta 2	KM	Exp m()	Dikta 1	Dikta 2
theta = 0.2								
t0.1	0.00021	0.00010	0.00010	0.00010	0.00021	0.00018	0.00018	0.00018
t0.25	0.00045	0.00020	0.00020	0.00020	0.00045	0.00043	0.00042	0.00042
t0.5	0.00087	0.00033	0.00033	0.00033	0.00087	0.00084	0.00083	0.00082
t0.75	0.00248	0.00062	0.00073	0.00087	0.00248	0.00286	0.00345	0.00389
t0.9	0.01234	0.00485	0.01000	0.01000	0.01234	0.00045	0.01000	0.01000
theta = 0.8								
t0.1	0.00021	0.00010	0.00010	0.00010	0.00021	0.00018	0.00018	0.00018
t0.25	0.00045	0.00020	0.00020	0.00020	0.00045	0.00043	0.00042	0.00042
t0.5	0.00088	0.00033	0.00033	0.00034	0.00088	0.00085	0.00084	0.00083
t0.75	0.00250	0.00063	0.00073	0.00086	0.00250	0.00289	0.00346	0.00390
t0.9	0.01230	0.00484	0.01000	0.01000	0.01230	0.00047	0.01000	0.01000
theta = 1								
t0.1	0.00021	0.00010	0.00010	0.00010	0.00021	0.00018	0.00018	0.00018
t0.25	0.00045	0.00020	0.00020	0.00020	0.00045	0.00043	0.00042	0.00042
t0.5	0.00088	0.00033	0.00033	0.00034	0.00088	0.00085	0.00084	0.00083
t0.75	0.00250	0.00063	0.00073	0.00086	0.00250	0.00289	0.00346	0.00390
t0.9	0.01230	0.00484	0.01000	0.01000	0.01230	0.00047	0.01000	0.01000
theta = 2								
t0.1	0.00021	0.00010	0.00010	0.00010	0.00021	0.00018	0.00018	0.00018
t0.25	0.00045	0.00020	0.00020	0.00020	0.00045	0.00043	0.00042	0.00042
t0.5	0.00088	0.00033	0.00033	0.00034	0.00088	0.00085	0.00084	0.00083
t0.75	0.00250	0.00063	0.00073	0.00086	0.00250	0.00290	0.00346	0.00390
t0.9	0.01230	0.00484	0.01000	0.01000	0.01230	0.00047	0.01000	0.01000
theta = 5								
t0.1	0.00021	0.00010	0.00010	0.00010	0.00021	0.00018	0.00018	0.00018
t0.25	0.00045	0.00020	0.00020	0.00020	0.00045	0.00043	0.00042	0.00042
t0.5	0.00088	0.00033	0.00033	0.00034	0.00088	0.00086	0.00084	0.00083
t0.75	0.00250	0.00064	0.00073	0.00087	0.00250	0.00292	0.00346	0.00390
t0.9	0.01237	0.00483	0.01000	0.01000	0.01237	0.00046	0.01000	0.01000

To make the table easy to look at, I used the column 2,3,4,5 to divide the column 1, column 7,8,9,10 to divide column 6.

The values that are less than 1 are showing that the methods have less MSE than the KM.

Table 6: MSE

	With true m()				With estimated m()			
Quantile	KM	Exp m()	Dikta 1	Dikta 2	KM	Exp m()	Dikta 1	Dikta 2

theta = 0.2								
t0.1	1	0.47032	0.47166	0.47239	1	0.85616	0.86098	0.86303
t0.25	1	0.44965	0.45058	0.45124	1	0.94872	0.94020	0.93514
t0.5	1	0.37631	0.37968	0.38267	1	0.96374	0.94752	0.93752
t0.75	1	0.25046	0.29309	0.34885	1	1.15229	1.38967	1.56865
t0.9	1	0.39305	0.81054	0.81054	1	0.03685	0.81054	0.81054
theta = 0.8								
t0.1	1	0.47311	0.47423	0.47492	1	0.85452	0.85908	0.86109
t0.25	1	0.44996	0.45088	0.45152	1	0.94990	0.94142	0.93634
t0.5	1	0.37471	0.37803	0.38094	1	0.96771	0.95161	0.94165
t0.75	1	0.25175	0.29053	0.34567	1	1.15610	1.38416	1.56139
t0.9	1	0.39361	0.81296	0.81296	1	0.03790	0.81296	0.81296
theta = 1								
t0.1	1	0.47350	0.47423	0.47492	1	0.85483	0.85908	0.86109
t0.25	1	0.45000	0.45088	0.45152	1	0.94969	0.94142	0.93634
t0.5	1	0.37480	0.37803	0.38094	1	0.96754	0.95161	0.94165
t0.75	1	0.25179	0.29053	0.34567	1	1.15752	1.38416	1.56139
t0.9	1	0.39341	0.81296	0.81296	1	0.03790	0.81296	0.81296
theta = 2								
t0.1	1	0.47354	0.47423	0.47492	1	0.85479	0.85908	0.86109
t0.25	1	0.44988	0.45088	0.45152	1	0.94947	0.94142	0.93634
t0.5	1	0.37484	0.37803	0.38094	1	0.96743	0.95161	0.94165
t0.75	1	0.25220	0.29053	0.34567	1	1.15960	1.38416	1.56140
t0.9	1	0.39319	0.81295	0.81295	1	0.03791	0.81295	0.81295
theta = 5								
t0.1	1	0.47611	0.47664	0.47731	1	0.85882	0.86261	0.86465
t0.25	1	0.45070	0.45145	0.45204	1	0.95069	0.94275	0.93768
t0.5	1	0.37685	0.37930	0.38206	1	0.97264	0.95648	0.94646
t0.75	1	0.25629	0.29229	0.34739	1	1.16871	1.38558	1.56302
t0.9	1	0.39087	0.80866	0.80866	1	0.03758	0.80866	0.80866

The bias of estimation of our $m()$ function from logistic regression, i.e. the absolute mean value of $\hat{m}(t)-$ true $m(t)$

Table 7: mean absolute difference between hat m() and true m()

0.2	0.8	1	2	5
0.0391791	0.0391817	0.0391817	0.0391817	0.0391232

The estimation of $\hat{\theta}$. The row name shows the true θ value

Table 8: estimated theta from logitic regression

0.2	0.8	1	2	5
0.2121856	0.8487424	1.060928	2.121856	5.30464