Checking the sampling method of Tsiatis's example

2019-09-20

In the Tsiatis's paper, the related functions are:

Function	Description	Expression
$P(T < t, C < c)$	Joint CDF	$1 + exp(-\lambda t - \mu c - \theta tc) - exp(-\lambda t) - exp(-\mu c)$
$f(t, c)$	Joint PDF	$(\lambda\mu - \theta + \lambda\theta t + \mu\theta c + \theta^2 tc)exp(-\lambda t - \mu c - \theta ct)$
$f_t(t)$	Marginal PDF of T	$\lambda exp(-\lambda t)$
$S_t(t)$	Survival function of T	$exp(-\lambda t)$
$f_c(c)$	Marginal PDF of C	$\mu exp(-\mu c)$
$S_c(c)$	$P_c(C > c)$	$exp(-\mu c)$
$S_x(t)$	$P(T > t, C > t)$	$exp(-\lambda t - \mu t - \theta t^2)$
$\psi(t)$	$\int_t^\infty f(t, c)dc$	$(\lambda + \theta t)exp(-\lambda t - \mu t - \theta t^2)$

The sampling process:

- We may sample $t \sim Exp(\lambda)$ first (e.g. sample 100 subjects from an exponential distribution with parameter λ)
- For each of the sampled t (one of the 100), we can sample c from the $F(c|t)$ distribution, based on inverse transform sampling idea. Where:

$$f(c|t) = f(t, c)/f(t) = \frac{(\lambda\mu - \theta + \lambda\theta t + \mu\theta c + \theta^2 tc)}{\lambda} exp(-\mu c - \theta ct)$$

$$F(c|t) = \int_0^c f(u|t)du = 1 - \frac{\lambda + \theta c}{\lambda} exp(-\mu c - \theta ct)$$

Then the pairs of (t, c) are sampled from the $f(t, c)$ distribution.

However, since the the close form of the inverse function of $F(c|t)$ cannot be calculated, we can estimate the inverse value approximately:

- Given a value $x \sim U(0, 1)$, we can find the $F(c_1|t)$ that is most close to the value x, where $c_1 \in \mathbb{N}$
- Given c_1, we can find the c_2, where $c_2 \in [c_1 - 1, c_1 + 1]$ by 0.1, that has $F(c_2|t)$ value most close to the x
- Given c_2, we can find the $c_3 \ldots$
- The three digits of the true c_0 value that makes $F(c_0|t) = x$ is estimated by this method.

Examples:

Try more scenarios to illustrate how the estimations work. For seeds were randomly selected. And the first table in each scenario is the table to show the mean absolute difference between true survival time and the estimated survival time by applying Kaplan-Meier method, Slud mehtod and the corrected Slud method. The second table in each scenario is to the mean absolute difference between true survival time and the estimated survival time within first 50 percents subjects, to avoid the unstable estimation in the tail.

Scenario 1: $\lambda = 1, \theta = 1, \mu = 1$

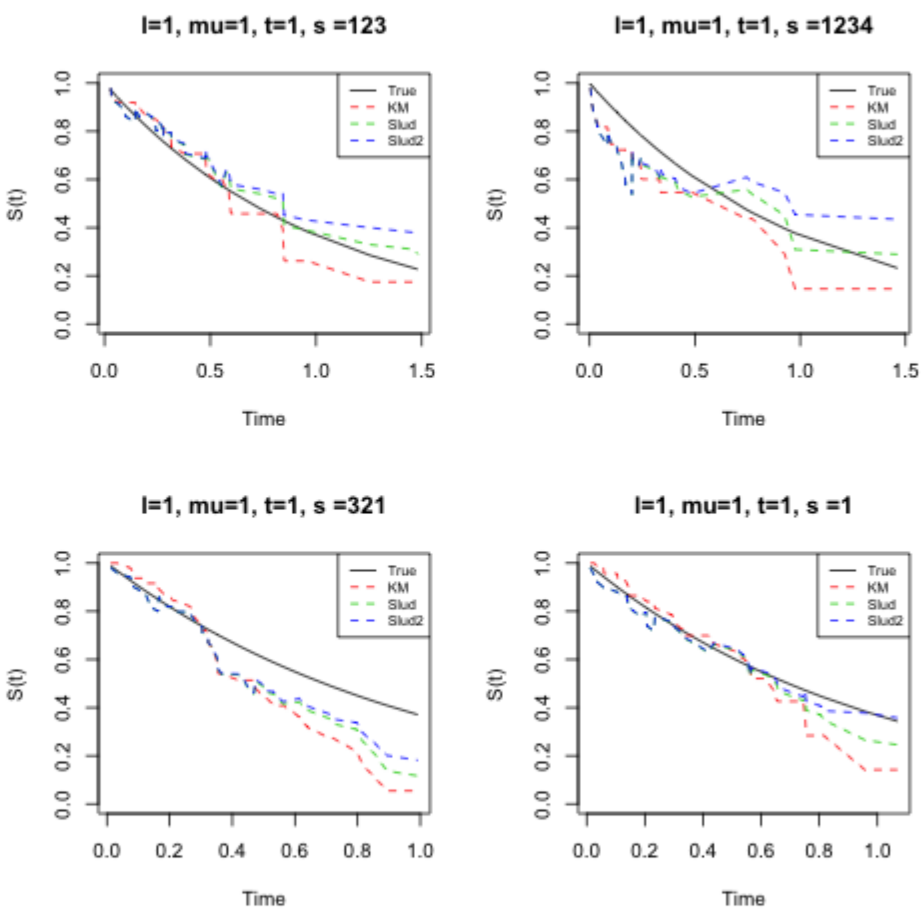
Sample size = 50

Table 2: lambda = 1, mu = 1, theta = 1

	KM	Slud1	Slud2
n = 50, seed = 123	0.042	0.041	0.051
n = 50, seed = 1234	0.116	0.129	0.133
n = 50, seed = 321	0.110	0.086	0.078
n = 50, seed = 1	0.048	0.035	0.027

Table 3: 50% of data: lambda = 1, mu = 1, theta = 1

	KM	Slud1	Slud2
n = 50, seed = 123	0.038	0.034	0.035
n = 50, seed = 1234	0.114	0.130	0.130
n = 50, seed = 321	0.036	0.021	0.021
n = 50, seed = 1	0.027	0.031	0.031



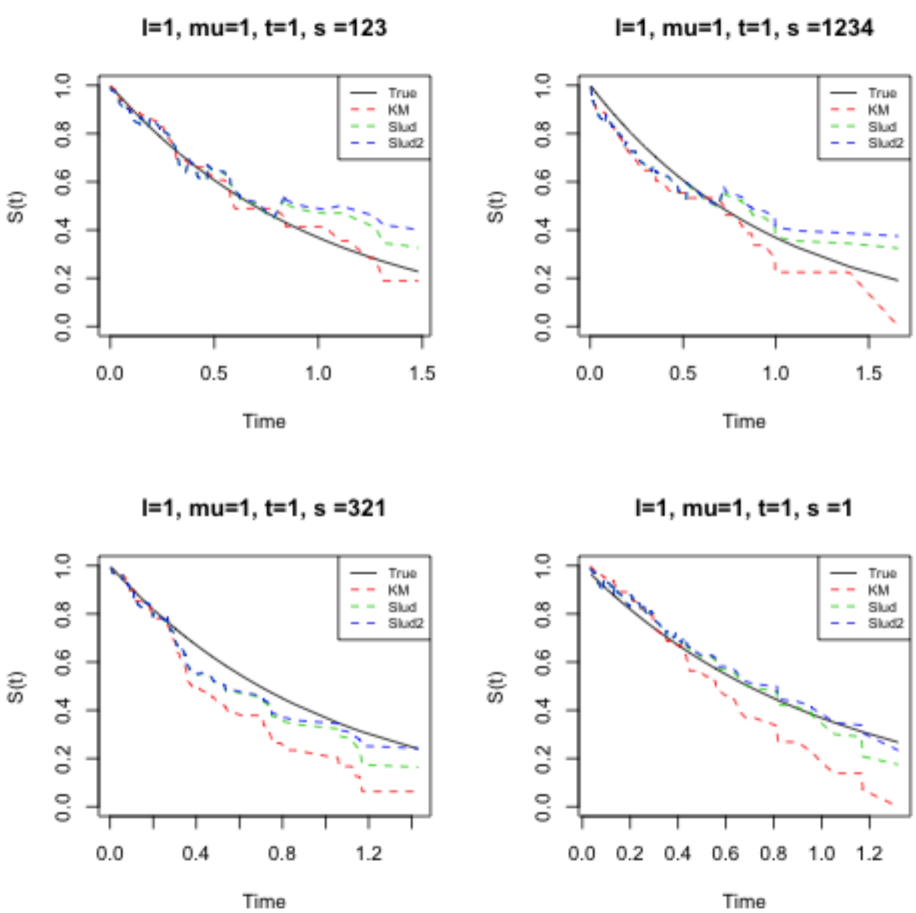
Sample size = 100

Table 4: lambda = 1, mu = 1, theta = 1

	KM	Slud1	Slud2
n = 100, seed = 123	0.022	0.035	0.041
n = 100, seed = 1234	0.064	0.054	0.057
n = 100, seed = 321	0.086	0.047	0.041
n = 100, seed = 1	0.057	0.028	0.029

Table 5: 50% of data: lambda = 1, mu = 1, theta = 1

	KM	Slud1	Slud2
n = 100, seed = 123	0.022	0.025	0.025
n = 100, seed = 1234	0.063	0.061	0.060
n = 100, seed = 321	0.017	0.014	0.014
n = 100, seed = 1	0.044	0.031	0.031



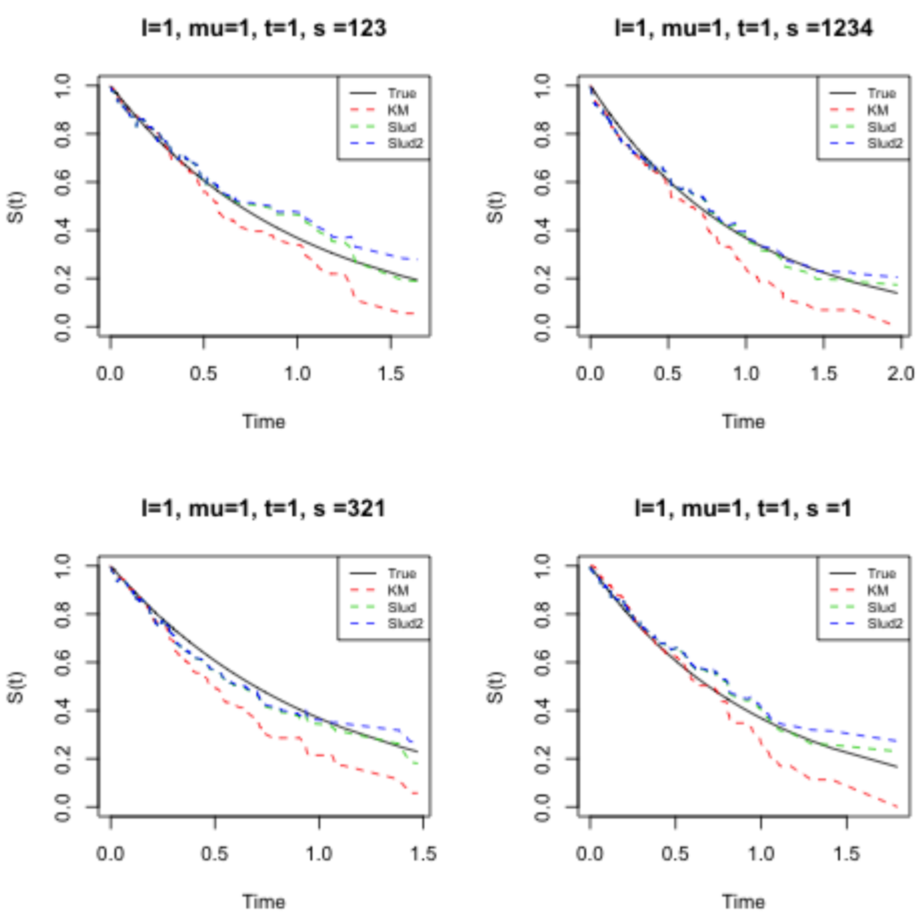
Sample size = 200

Table 6: lambda = 1, mu = 1, theta = 1

	KM	Slud1	Slud2
n = 200, seed = 123	0.030	0.023	0.027
n = 200, seed = 1234	0.043	0.026	0.027
n = 200, seed = 321	0.066	0.028	0.026
n = 200, seed = 1	0.031	0.024	0.026

Table 7: 50% of data: lambda = 1, mu = 1, theta = 1

	KM	Slud1	Slud2
n = 200, seed = 123	0.012	0.018	0.019
n = 200, seed = 1234	0.029	0.035	0.034
n = 200, seed = 321	0.016	0.016	0.016
n = 200, seed = 1	0.024	0.011	0.011



Scenario 2: $\lambda = 0.1, \mu = 0.2, \theta = 0.02$

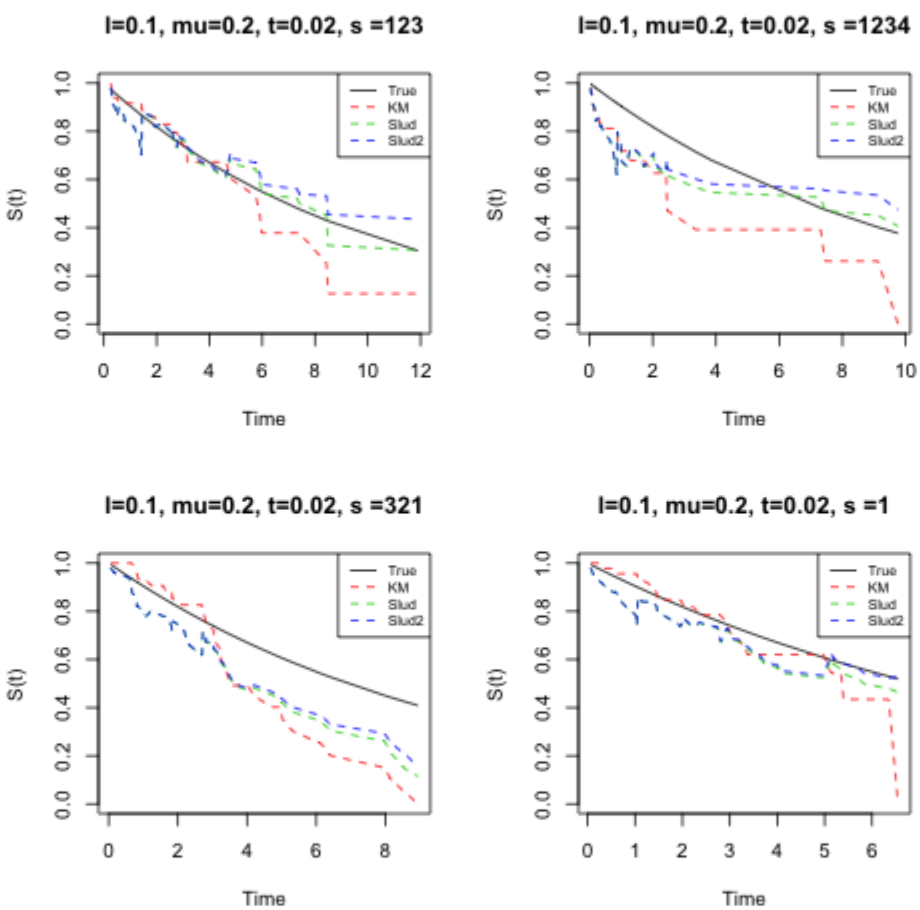
Sample size = 50

Table 8: lambda = 0.1, mu = 0.2, theta = 0.02

	KM	Slud1	Slud2
n = 50, seed = 123	0.046	0.045	0.051
n = 50, seed = 1234	0.152	0.145	0.146
n = 50, seed = 321	0.109	0.121	0.113
n = 50, seed = 1	0.050	0.065	0.058

Table 9: 50% of data: lambda = 0.1, mu = 0.2, theta = 0.02

	KM	Slud1	Slud2
n = 50, seed = 123	0.028	0.061	0.062
n = 50, seed = 1234	0.102	0.146	0.146
n = 50, seed = 321	0.036	0.079	0.079
n = 50, seed = 1	0.024	0.071	0.071



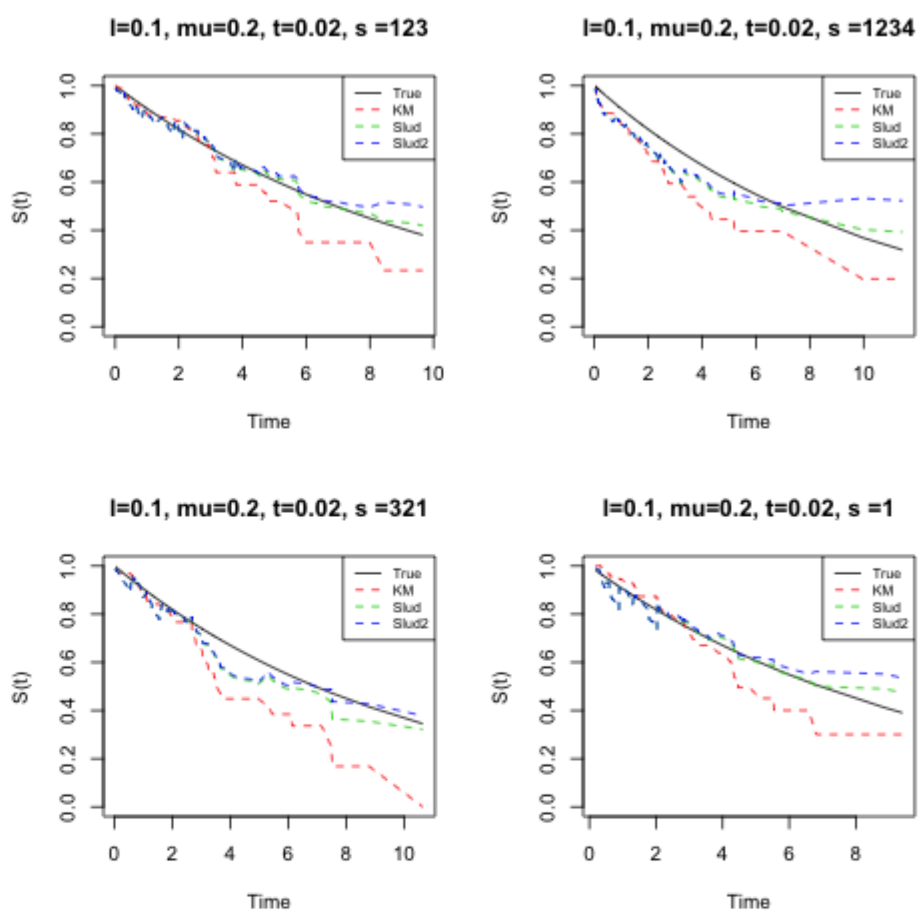
Sample size = 100

Table 10: lambda = 0.1, mu = 0.2, theta = 0.02

	KM	Slud1	Slud2
n = 100, seed = 123	0.046	0.022	0.026
n = 100, seed = 1234	0.102	0.069	0.068
n = 100, seed = 321	0.084	0.040	0.036
n = 100, seed = 1	0.047	0.024	0.029

Table 11: 50% of data: lambda = 0.1, mu = 0.2, theta = 0.02

	KM	Slud1	Slud2
n = 100, seed = 123	0.019	0.024	0.024
n = 100, seed = 1234	0.059	0.059	0.058
n = 100, seed = 321	0.019	0.024	0.024
n = 100, seed = 1	0.032	0.028	0.028



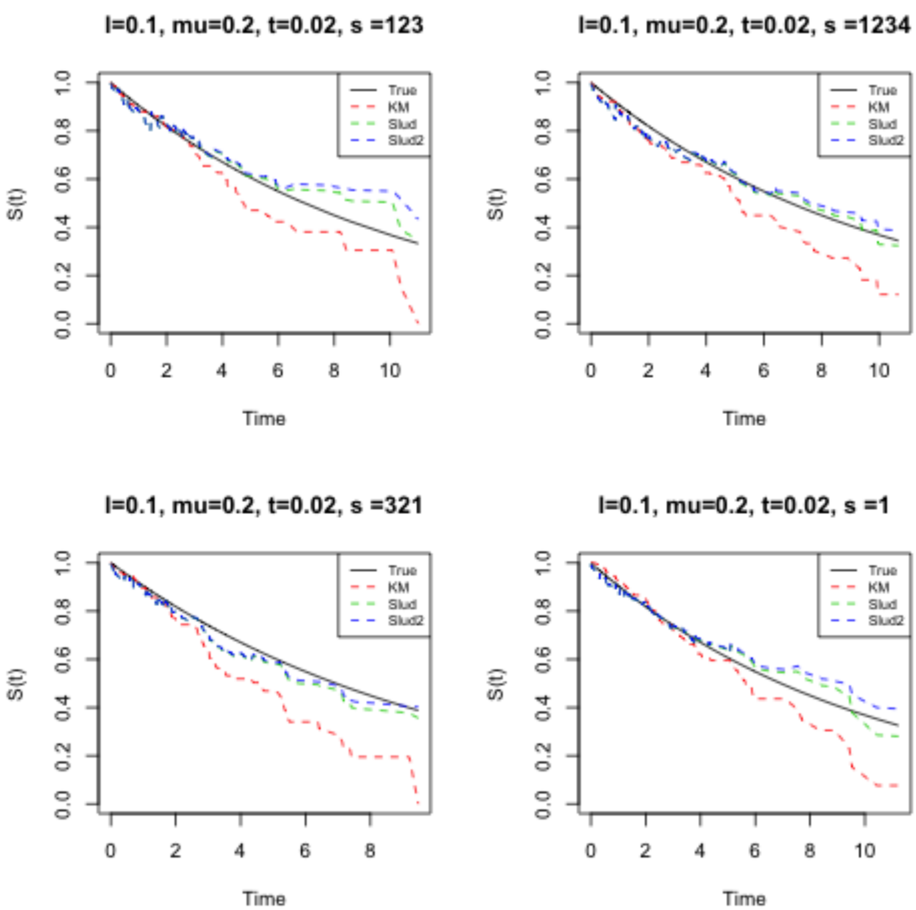
Sample size = 200

Table 12: lambda = 0.1, mu = 0.2, theta = 0.02

	KM	Slud1	Slud2
n = 200, seed = 123	0.035	0.023	0.027
n = 200, seed = 1234	0.055	0.031	0.032
n = 200, seed = 321	0.075	0.029	0.026
n = 200, seed = 1	0.034	0.017	0.020

Table 13: 50% of data: lambda = 0.1, mu = 0.2, theta = 0.02

	KM	Slud1	Slud2
n = 200, seed = 123	0.010	0.023	0.023
n = 200, seed = 1234	0.029	0.038	0.038
n = 200, seed = 321	0.013	0.017	0.017
n = 200, seed = 1	0.020	0.015	0.015



Scenario 3: $\lambda = 0.1, \mu = 0.2, \theta = 0.018$

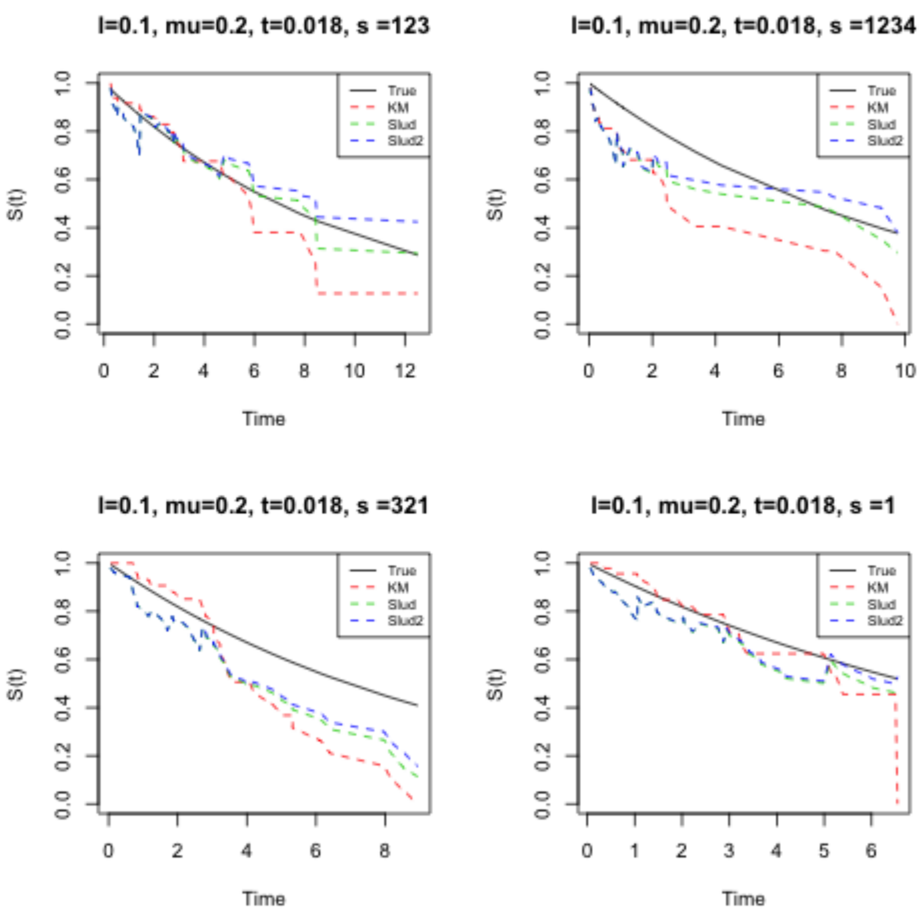
Sample size = 50

Table 14: lambda = 0.1, mu = 0.2, theta = 0.018

	KM	Slud1	Slud2
n = 50, seed = 123	0.045	0.045	0.051
n = 50, seed = 1234	0.155	0.141	0.138
n = 50, seed = 321	0.112	0.111	0.103
n = 50, seed = 1	0.048	0.067	0.059

Table 15: 50% of data: lambda = 0.1, mu = 0.2, theta = 0.018

	KM	Slud1	Slud2
n = 50, seed = 123	0.029	0.061	0.061
n = 50, seed = 1234	0.105	0.134	0.133
n = 50, seed = 321	0.048	0.072	0.072
n = 50, seed = 1	0.023	0.068	0.067



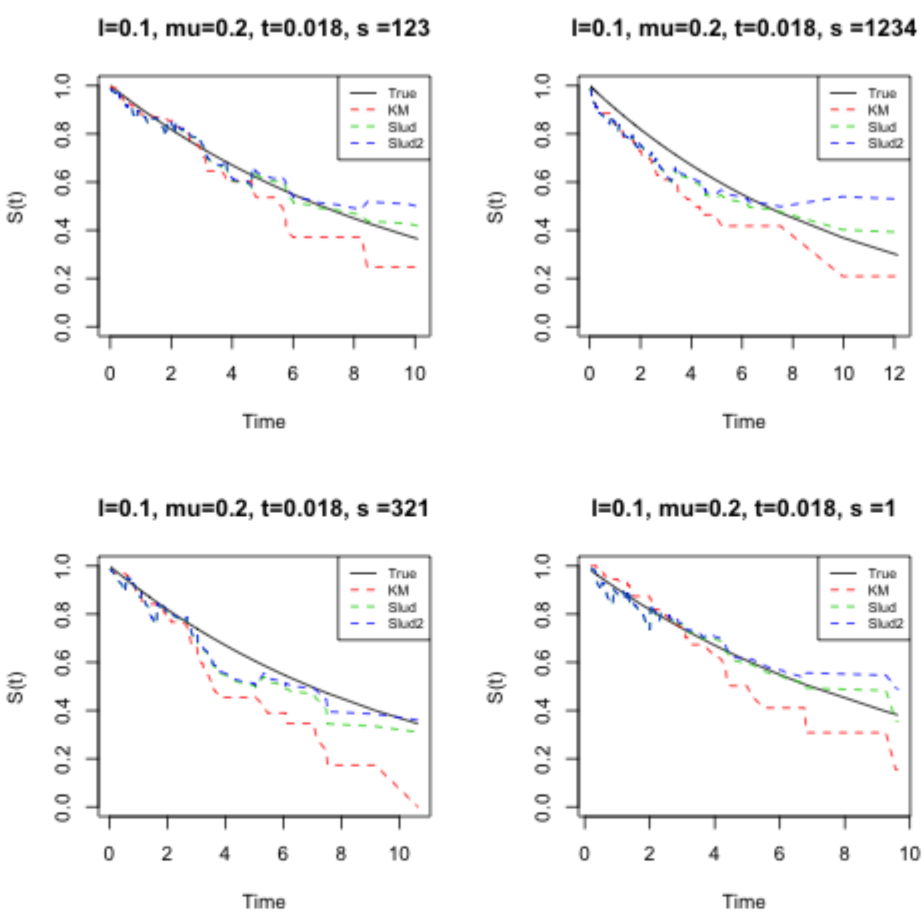
Sample size = 100

Table 16: lambda = 0.1, mu = 0.2, theta = 0.018

	KM	Slud1	Slud2
n = 100, seed = 123	0.044	0.025	0.028
n = 100, seed = 1234	0.093	0.067	0.066
n = 100, seed = 321	0.084	0.045	0.040
n = 100, seed = 1	0.048	0.022	0.027

Table 17: 50% of data: lambda = 0.1, mu = 0.2, theta = 0.018

	KM	Slud1	Slud2
n = 100, seed = 123	0.019	0.024	0.024
n = 100, seed = 1234	0.058	0.061	0.061
n = 100, seed = 321	0.018	0.027	0.027
n = 100, seed = 1	0.033	0.027	0.026



Sample size = 200

Table 18: lambda = 0.1, mu = 0.2, theta = 0.018

	KM	Slud1	Slud2
n = 200, seed = 123	0.032	0.030	0.034
n = 200, seed = 1234	0.047	0.031	0.033
n = 200, seed = 321	0.066	0.026	0.023
n = 200, seed = 1	0.031	0.018	0.021

Table 19: 50% of data: lambda = 0.1, mu = 0.2, theta = 0.018

	KM	Slud1	Slud2
n = 200, seed = 123	0.012	0.028	0.028
n = 200, seed = 1234	0.029	0.039	0.039
n = 200, seed = 321	0.011	0.016	0.016
n = 200, seed = 1	0.019	0.015	0.015