Cox example

2020-04-08

Step 1: the consistency of Cox PH model v.s. confounder/non-controled covariates

Example 1 (Independent censoring)

We denote $T_i, i = 1, ..., N$ are the independent, identically, distributed (iid) lifetimes, whose CDF and PDF are denoted as F, f, respectively, and corresponding survival function $S_T(t) = P(T > t)$; the censoring time is defined as $C_i, i = 1, ..., N$. C_is are also iid, with CDF denoted as G and PDF denoted as g and $S_C(t) = P(C > t)$. We set the censors happen on the right and the ovserved time is $Z_i = T_i \wedge C_i$, whose CDF is H and PDF is h. The $\delta_i = I_{[T_i \leq C_i]}$ is the status indicator, which shows whether subject i is censored ($\delta_i = 0$) or not ($\delta_i = 0$). The corresponding hazard function of lifetime is λ_F and cumulative hazard function is Λ_F.

Let's just consider a two dimension covariates scenario, i.e. the $S_T(t)$ and $S_C(t)$ are associated with x_1, x_2. Suppose the death time and censoring time follow Cox PH models, then,

- The hazard function, cumulative hazard function and survival function for death time T:

$$\lambda_T(t|x) = \lambda_0(t)\exp(\beta_1 x_1 + \beta_2 x_2)$$
$$\Lambda_T(t|x) = \Lambda_0(t)\exp(\beta_1 x_1 + \beta_2 x_2)$$
$$S(t|x) = S_0(t)^{\exp(\beta_1 x_1 + \beta_2 x_2)}$$

- The hazard function, cumulative hazard function and survival function for censoring time C:

$$\lambda_C(t|x) = \lambda_0(t)\exp(\gamma_1 x_1 + \gamma_2 x_2)$$
$$\Lambda_C(t|x) = \Lambda_0(t)\exp(\gamma_1 x_1 + \gamma_2 x_2)$$
$$S_C(t|x) = S_0(t)^{\exp(\gamma_1 x_1 + \gamma_2 x_2)}$$

  where the survival time and censoring time functions share the same baseline function and both associate with two covariates x_1, x_2.

Denote the $m()$ function as:

$$m(t,x) = E(\delta|Z = t, X = x) = P(\delta = 1|Z = t, X = x) = \frac{\lambda_T(t|x)}{\lambda_Z(t|x)} = \frac{\lambda_T(t|x)}{\lambda_T(t|x) + \lambda_C(t|x)}$$

Data set generation

- 1. Simulate the covariates for n subjects: $X = (x1, x2) \sim MVN(\mu, \Sigma)$, where X is n by 2 matrix and $x1, x2$ are n by 1 vectors. For each subject, the covariates is x_{i1}, x_{i2}.
  - $X = (x1, x2) \sim MVN(\mu, \Sigma)$, $\mu = (1, 2)$, $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, where $\rho = 0, 0.2, 0.5, 0.9$.
  - (1) $\beta = (\beta_1, \beta_2) = \underline{(2, 0.1)}, \gamma = (\gamma_1, \gamma_2) = (0.1, 0.2)$
  - (2) $\beta = (\beta_1, \beta_2) = \underline{(0.2, 0.1)}, \gamma = (\gamma_1, \gamma_2) = (0.1, 0.2)$
- 2. For subject i, calculate the βX_i. Simulate the survival time T_i from below CDF by using the inverse probablity sampling.

$$S_{T,i}(t) = S_{0,i}(t)^{\exp(\beta X_i)} = \exp(-t \times \exp(\beta X_i))$$

- 3. Simulate the censoring time C_i from below CDF by using the inverse probablity sampling.

$$S_{C,i}(t) = S_{0,i}(t)^{\exp(\gamma X_i)} = \exp(-t \times \exp(\gamma X_i))$$

- 4. Calculate observed time $Z_i = T_i \wedge C_i$ and indicator $\delta_i = I(T_i < C_i)$.
- Then we get the dataset with death time, censoring time, observed time, censoring indicator, and covariates. The times are calculated one by one based on the subject's covariates values.

Data size: $n = 1000$; Censoringship proportion:

Models

- Model 1: Fit the Cox PH model with both x_1 and x_2
- Model 2: Fit the Cox PH model with only x_1 (x_2 is unobserved)
- Model 3: Fit the Cox PH model with only x_2 (x_1 is unobserved)

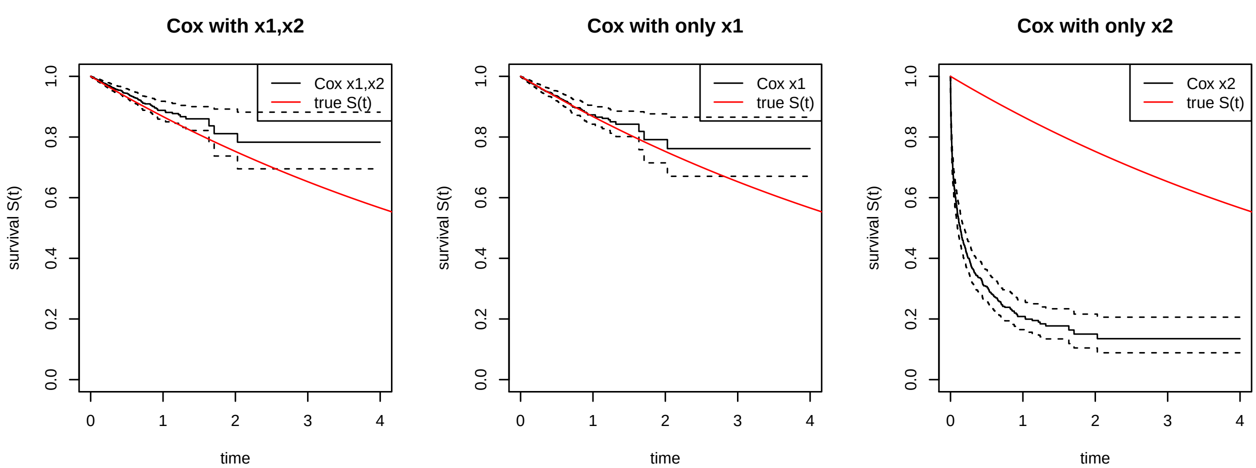
Whether model 2 and model 3 can consistently estimate the survival function? Let's show it from several survival curves.

Survival plot

$\beta = (\beta_1, \beta_2) = (2, 0.1)$, $\gamma = (\gamma_1, \gamma_2) = (0.1, 0.2)$

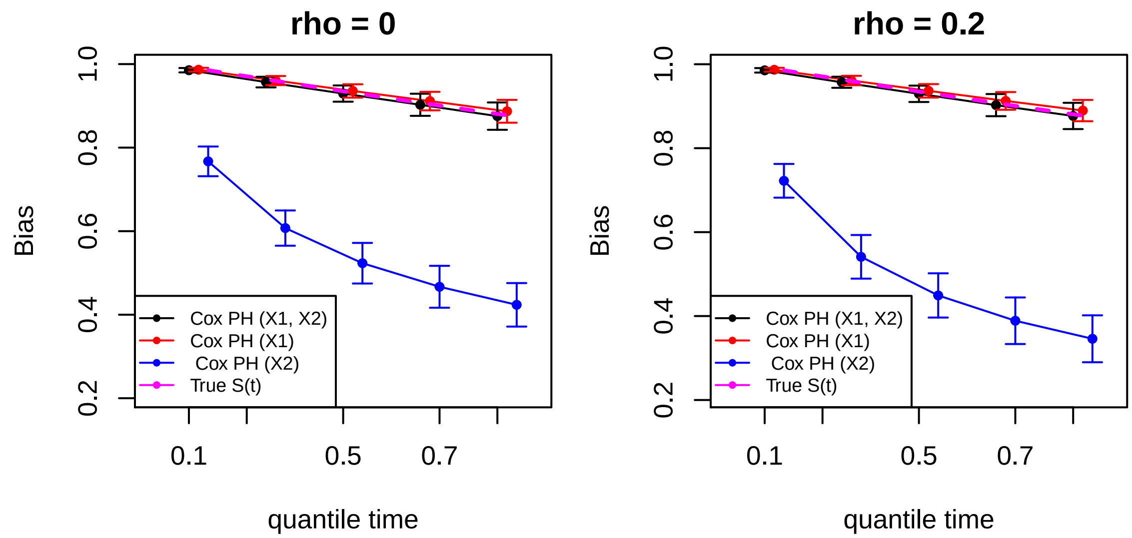
Survival function $S(T > t|X) = \exp(-t)^{\exp(\beta X)}$ Observed time function $S(Z > t|X) = S(T > t|X)S(C > t|X) = \exp(-t)^{\exp(\beta X)} \exp(-t)^{\exp(\gamma X)}$

Draw the survival plots at the point (-1, 0.5):

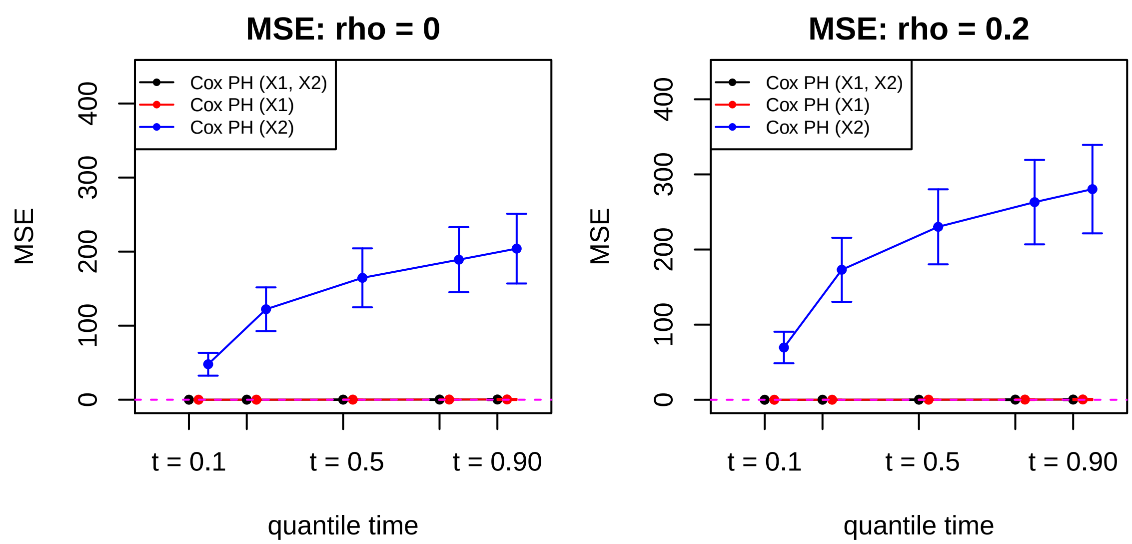


Let's repete the above procedure for 100 times. The estimation value and standard deviations are shown below:

$\beta = (\beta_1, \beta_2) = (2, 0.1)$



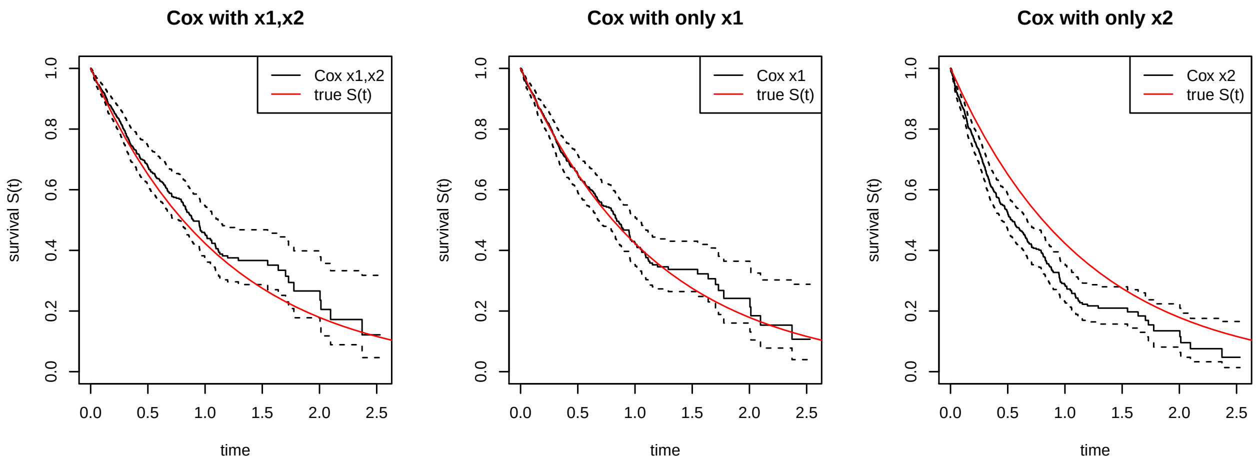

```
## Warning in arrows(c(0.1, 0.25, 0.5, 0.75, 0.9), qtablem[, 1] - 1.96 *
## qtablesd[, : zero-length arrow is of indeterminate angle and so skipped

## Warning in arrows(c(0.1, 0.25, 0.5, 0.75, 0.9), qtablem[, 1] - 1.96 *
## qtablesd[, : zero-length arrow is of indeterminate angle and so skipped

## Warning in arrows(c(0.1, 0.25, 0.5, 0.75, 0.9) + 0.025, qtablem[, 2] - 1.96
## * : zero-length arrow is of indeterminate angle and so skipped

## Warning in arrows(c(0.1, 0.25, 0.5, 0.75, 0.9) + 0.025, qtablem[, 2] - 1.96
## * : zero-length arrow is of indeterminate angle and so skipped

## Warning in arrows(c(0.1, 0.25, 0.5, 0.75, 0.9), qtablem[, 1] - 1.96 *
## qtablesd[, : zero-length arrow is of indeterminate angle and so skipped

## Warning in arrows(c(0.1, 0.25, 0.5, 0.75, 0.9), qtablem[, 1] - 1.96 *
## qtablesd[, : zero-length arrow is of indeterminate angle and so skipped

## Warning in arrows(c(0.1, 0.25, 0.5, 0.75, 0.9) + 0.025, qtablem[, 2] - 1.96
## * : zero-length arrow is of indeterminate angle and so skipped
```



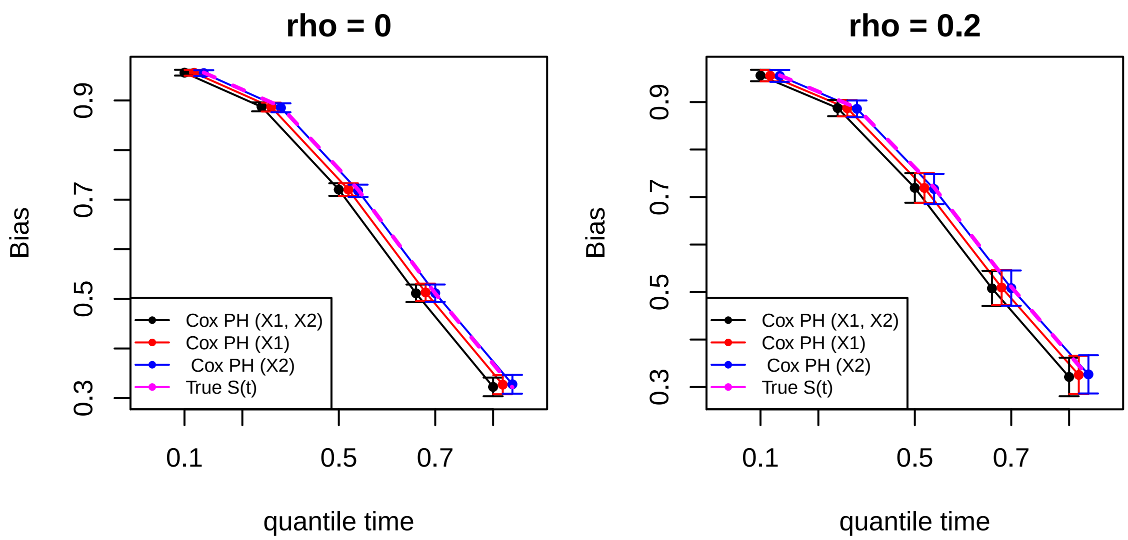
$\beta = (\beta_1, \beta_2) = (0.2, 0.1), \gamma = (\gamma_1, \gamma_2) = (0.1, 0.2)$

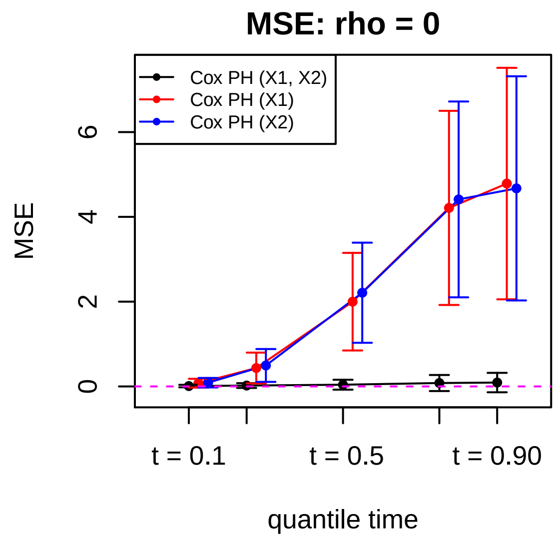
Survival function $S(T > t|X) = \exp(-t)^{\exp(\beta X)}$ Observed time function $S(Z > t|X) = S(T > t|X)S(C > t|X) = \exp(-t)^{\exp(\beta X)} \exp(-t)^{\exp(\gamma X)}$

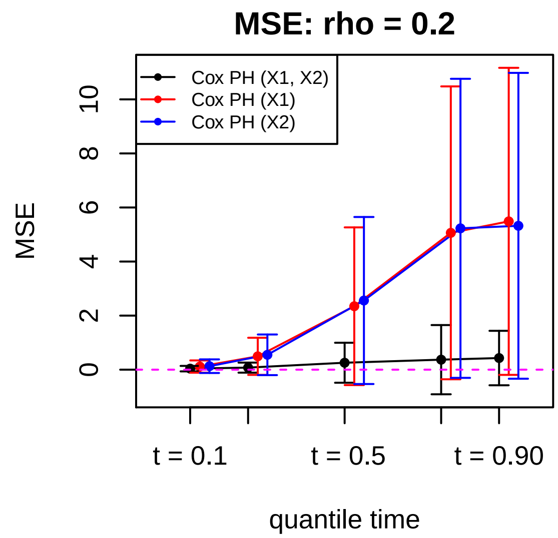
Draw the survival plots at the point (-1, 0.5):



$\beta = (\beta_1, \beta_2) = (0.2, 0.1)$







Example 2 (Dependent censoring (Tsiatis example))

The Tsiatis proposed an example of dependent censorship.

$$P(T > t, C > s) = \exp(-\lambda t - \mu s - \theta ts)$$

$$S_T(t) = P(T > t) = \exp(-\lambda t)$$

$$S_C(s) = P(C > s) = \exp(-\mu s)$$

If we add covariates in the joint model, we can replace λ with $\lambda' = k\exp(\beta X)$, replace μ with $\mu' = k\exp(\gamma X)$. Therefore, the joint distribution with covariates is:

$$P(T > t, C > s|X) = \exp(-k\exp(\beta X)t - k\exp(\gamma X)s - \theta ts)$$

And the marignal distributions are

$$S_T(t|X) = P(T > t|X) = \exp(-k\exp(\beta X)t) = \exp(-kt)^{\exp(\beta X)} = S_0(t)^{\exp(\beta X)}, \text{ where } S_0(t) = \exp(-kt)$$

$$S_C(s|X) = P(C > s|X) = \exp(-k\exp(\gamma X)s) = \exp(-ks)^{\exp(\gamma X)} = S_0(s)^{\exp(\gamma X)}$$

Therefore the marignal distributions $S_T(t|X)$ and $S_C(t|X)$ have the forms of Cox PH model.

Data set generation

- 1. Simulate the covariates for n subjects: $X = (x1, x2) \sim MVN(\mu, \Sigma)$, where X is n by 2 matrix and $x1, x2$ are n by 1 vectors. For each subject, the covariates is x_{i1}, x_{i2}.
  - $\beta = (\beta_1, \beta_2) = (0.2, 0.1), \gamma = (\gamma_1, \gamma_2) = (0.1, 0.2)$
  - $X = (x1, x2) \sim MVN(\mu, \Sigma)$, $\mu = (1, 2)$, $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, where $\rho = 0, 0.2, 0.5, 0.9$.
- 2. For subject i, calculate the βX_i. Simulate the survival time T_i from below CDF by using the inverse probablity sampling.

$$S_{T,i}(t) = S_{0,i}(t)^{\exp(\beta X_i)} = \exp(-t \times \exp(\beta X_i))$$

- 3. Simulate the censoring time C_i from below the conditional CDF by using the inverse probablity sampling.

$$P(C > t|T = t, X = x)$$

- 4. Calculate observed time $Z_i = T_i \wedge C_i$ and indicator $\delta_i = I(T_i < C_i)$.
- Then we get the dataset with death time, censoring time, observed time, censoring indicator, and covariates. The times are calculated one by one based on the subject's covariates values.

Data size: $n = 1000$; Censoringship proportion:

Models

- Model 1: Fit the Cox PH model with both x_1 and x_2
- Model 2: Fit the Cox PH model with only x_1, (x_2 is unobserved)
- Model 3: Fit the Cox PH model with only x_2, (x_1 is unobserved)

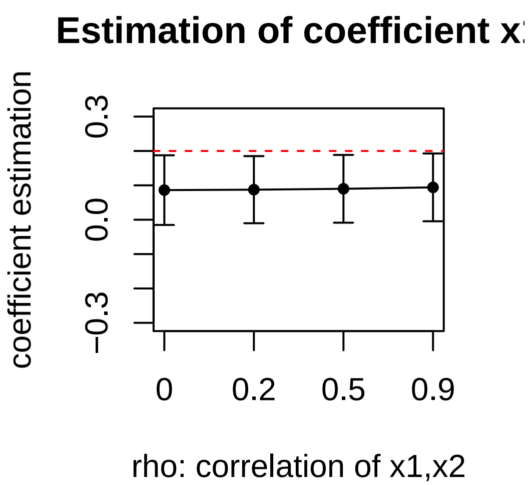
The estimated covariates table:

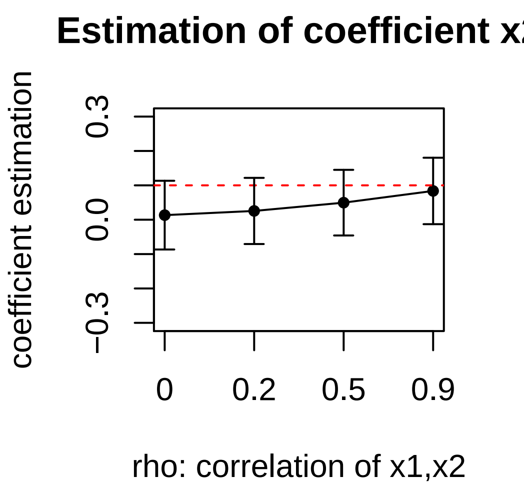
- Cox model fitted with only X1, the β_1 is;
- Cox model fitted with only X2, the β_2 is;

Table 1: Covariates estimation

	X1 (true = 0.2)			X2 (true = 0.1)		
rho	mean	sd	signifcant	mean	sd	signifcant
0.0	0.0861	0.0517	0.48	0.0133	0.0510	0.12
0.2	0.0874	0.0498	0.51	0.0255	0.0491	0.11
0.5	0.0899	0.0503	0.51	0.0496	0.0487	0.17
0.9	0.0941	0.0503	0.53	0.0837	0.0493	0.50

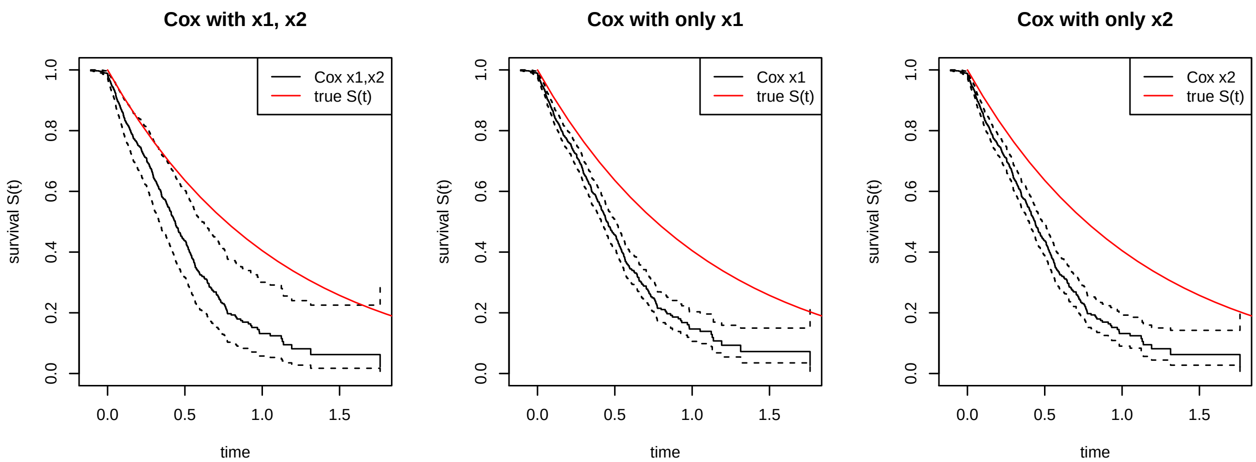
The coefficent plots within 100 repetitions.



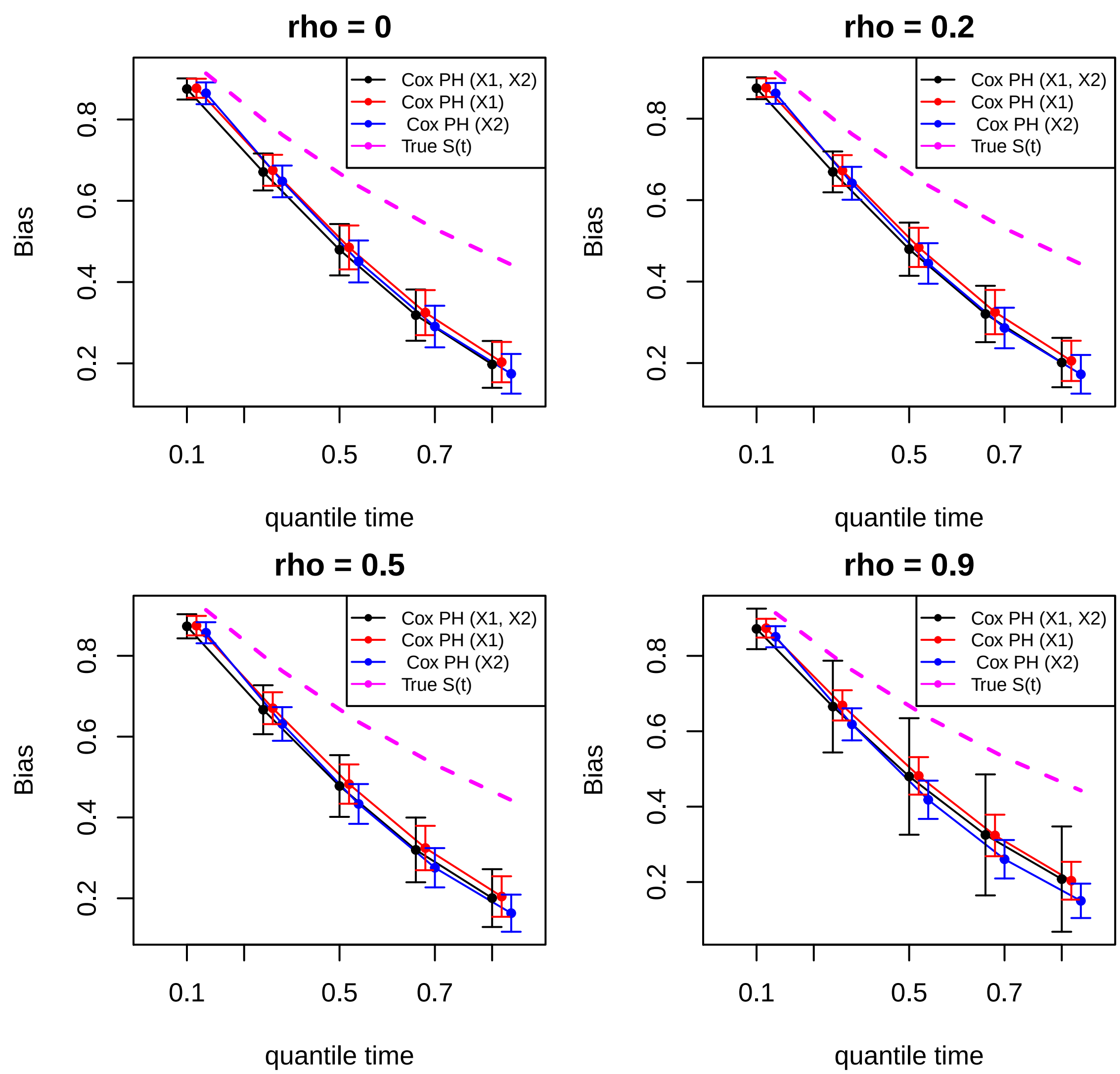


Survival plots

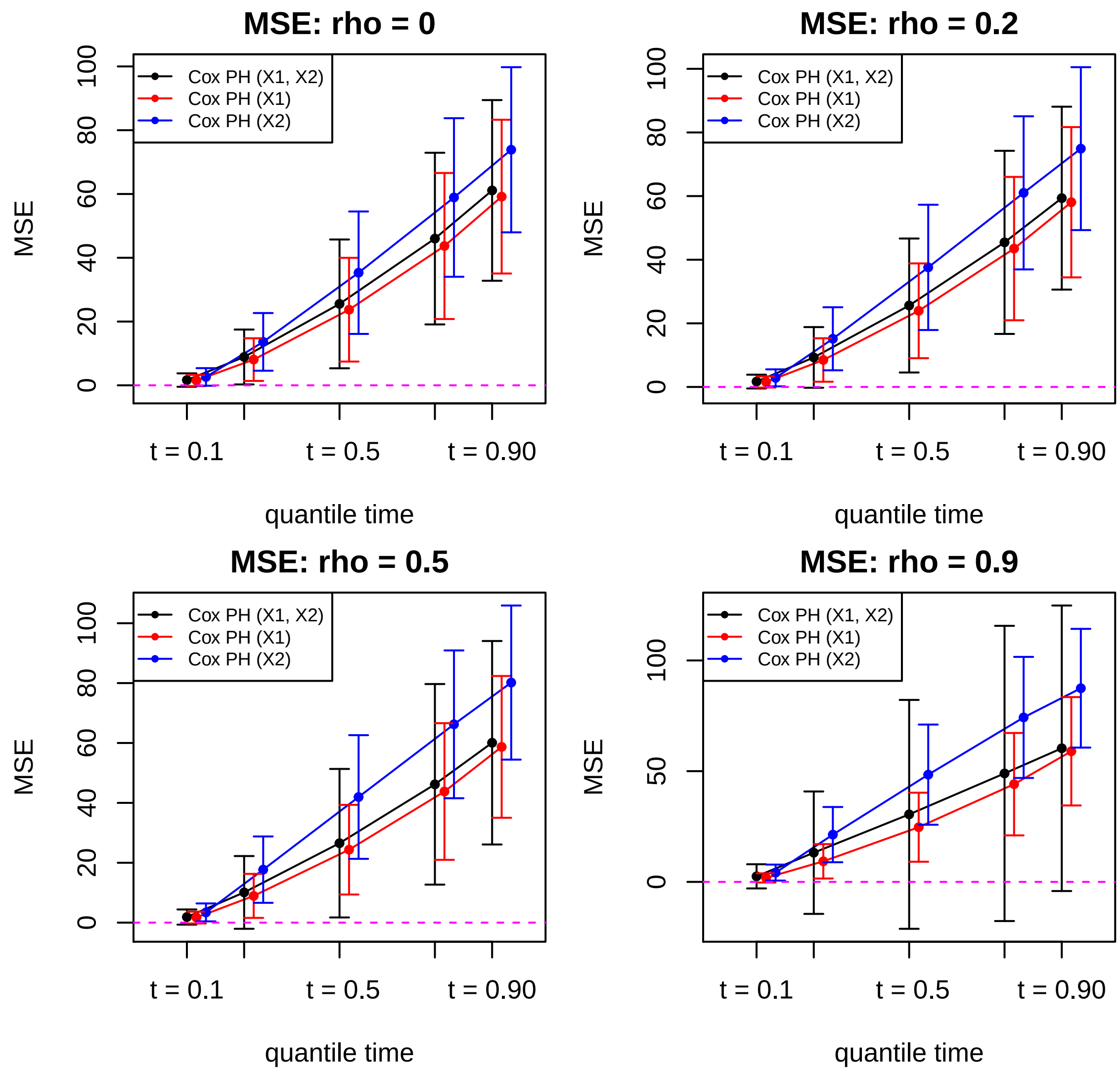
at point (-1,1)



We repete the procedure for 100 times. The estimated value of $S_T(t|x_1 = -1, x_2 = 2)$, $t = 0.1, 0.3, 0.5, 0.9$.



The MSE of the point estimation:



Calculate the bias over the dataset

- For $n = 100$, we just look at the quantile time at $t = 10\%, 25\%, 50\%, 75\%, 90\%$, across the dataset.
  - for each subject in the dataset, calculate the true survival function $S_i(t, x_i)$
  - for each subject in the dataset, calculate the estimated survival functions by above methods, $\hat{S}_i(t, x_i)$.
  - calculate the bias: $\hat{S}_i(t, x_i) - S_i(t, x_i)$
  - then calculate the mean bias for the dataset $\frac{1}{n}\sum_{i=1}^{n}(\hat{S}_i(t, x_i) - S_i(t, x_i))$, which is marked as the bias for the dataset, bias_{data}.
  - for the 100 repetitions, get the mean value of the bias for the dataset $\frac{1}{100}\sum_{i=1}^{100}\text{bias}_{data,i}$
  - For MSE, for each repetition, calculate $\frac{1}{n}\sum_{i=1}^{n}(\hat{S}_i(t, x_i) - S_i(t, x_i))^2$, and then calculate the mean value of the MSE across the 100 repetition.

Table 2: Table of Bias

	cox12	cox1	cox2	m()	m()hat	dikta	dikta hat
rho = 0							
qt=0.1	-25.464	-24.864	-25.45	10.93	10.962	10.852	10.884

qt=0.25	-37.92	-44.315	-45.784	22.561	22.598	22.311	22.349
qt=0.5	-45.367	-80.043	-82.311	54.865	54.743	54.412	54.294
qt=0.75	-45.826	-142.404	-144.538	122.538	121.896	121.591	120.955
qt=0.9	-39.409	-202.738	-203.716	189.76	188.427	187.968	186.642
rho = 0.2							
qt=0.1	-24.458	-24.397	-24.959	10.894	10.936	10.816	10.858
qt=0.25	-37.715	-43.21	-44.624	22.321	22.378	22.129	22.187
qt=0.5	-46.739	-81.964	-84.294	56.57	56.485	56.117	56.036
qt=0.75	-48.149	-141.875	-144.311	121.91	121.312	120.962	120.37
qt=0.9	-41.155	-203.487	-205.219	189.212	187.914	187.418	186.128
rho = 0.5							
qt=0.1	-25.258	-25.324	-25.856	11.479	11.517	11.401	11.439
qt=0.25	-38.581	-44.934	-46.3	22.735	22.773	22.543	22.582
qt=0.5	-49.795	-85.343	-87.65	56.955	56.781	56.501	56.332
qt=0.75	-49.683	-143.538	-146.065	121.572	120.729	120.623	119.788
qt=0.9	-44.448	-203.565	-205.493	188.4	186.678	186.606	184.892
rho = 0.9							
qt=0.1	-25.966	-26.482	-26.989	12.039	12.088	11.96	12.01
qt=0.25	-39.959	-47.551	-48.872	23.504	23.558	23.311	23.367
qt=0.5	-52.832	-88.816	-91.081	57.272	57.082	56.818	56.633
qt=0.75	-52.862	-145.952	-148.497	121.447	120.461	120.497	119.52
qt=0.9	-47.636	-206.298	-208.309	188.101	186.064	186.305	184.279

(values were timed 1000)

Table 3: Table of Standard deviation

	cox12	cox1	cox2	m()	m()hat	dikta	dikta hat
rho = 0							
qt=0.1	13.699	14.148	14.136	8.792	8.822	8.792	8.822
qt=0.25	19.51	19.33	19.338	14.392	14.451	14.44	14.499
qt=0.5	29.89	30.135	30.479	17.58	17.5	17.58	17.501
qt=0.75	36.494	36.386	37.356	21.559	21.605	21.559	21.604
qt=0.9	36.197	40.917	41.993	21.494	21.656	21.495	21.657
rho = 0.2							
qt=0.1	13.566	13.836	13.776	9.339	9.348	9.339	9.348
qt=0.25	21.783	21.324	21.183	14.473	14.495	14.473	14.495
qt=0.5	31.608	31.297	30.977	18.32	18.247	18.32	18.247
qt=0.75	37.085	39.675	39.515	20.924	21.147	20.925	21.148
qt=0.9	36.656	41.194	42.236	23.133	22.978	23.134	22.978
rho = 0.5							
qt=0.1	13.45	13.547	13.471	9.42	9.434	9.42	9.434
qt=0.25	21.935	21.496	21.386	14.37	14.444	14.37	14.444
qt=0.5	32.128	32.727	32.404	18.663	18.677	18.663	18.678
qt=0.75	38.725	40.75	40.472	21.179	21.533	21.18	21.533
qt=0.9	37.793	41.399	42.368	22.938	22.829	22.939	22.83
rho = 0.9							
qt=0.1	13.887	13.52	13.443	9.514	9.536	9.514	9.536
qt=0.25	21.703	21.772	21.631	14.889	15.007	14.889	15.007
qt=0.5	32.333	31.843	31.619	19.175	19.221	19.175	19.222
qt=0.75	38.112	40.191	39.846	21.708	22.143	21.709	22.144

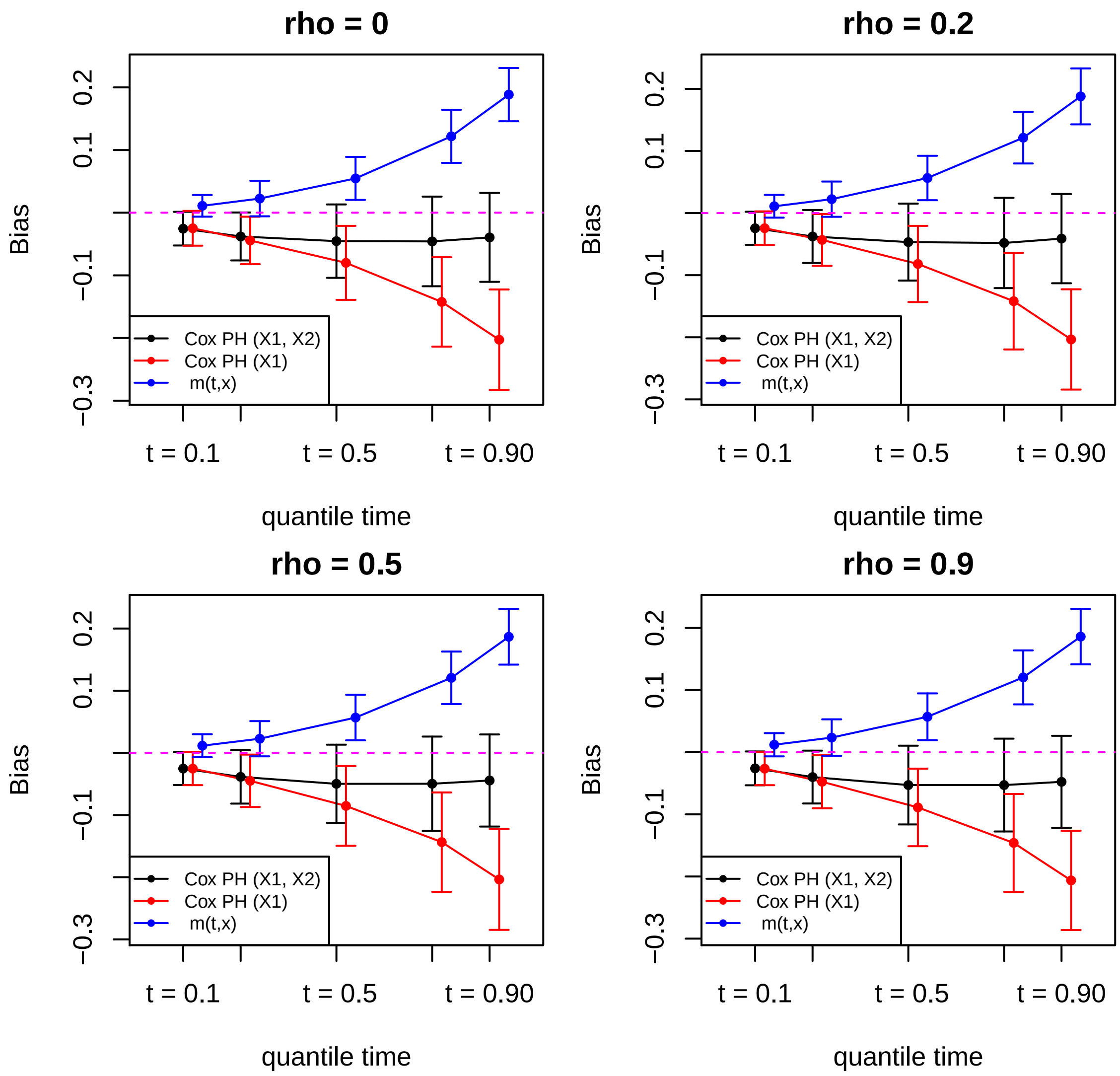
qt=0.9	37.8	40.734	41.564	22.657	22.751	22.658	22.752

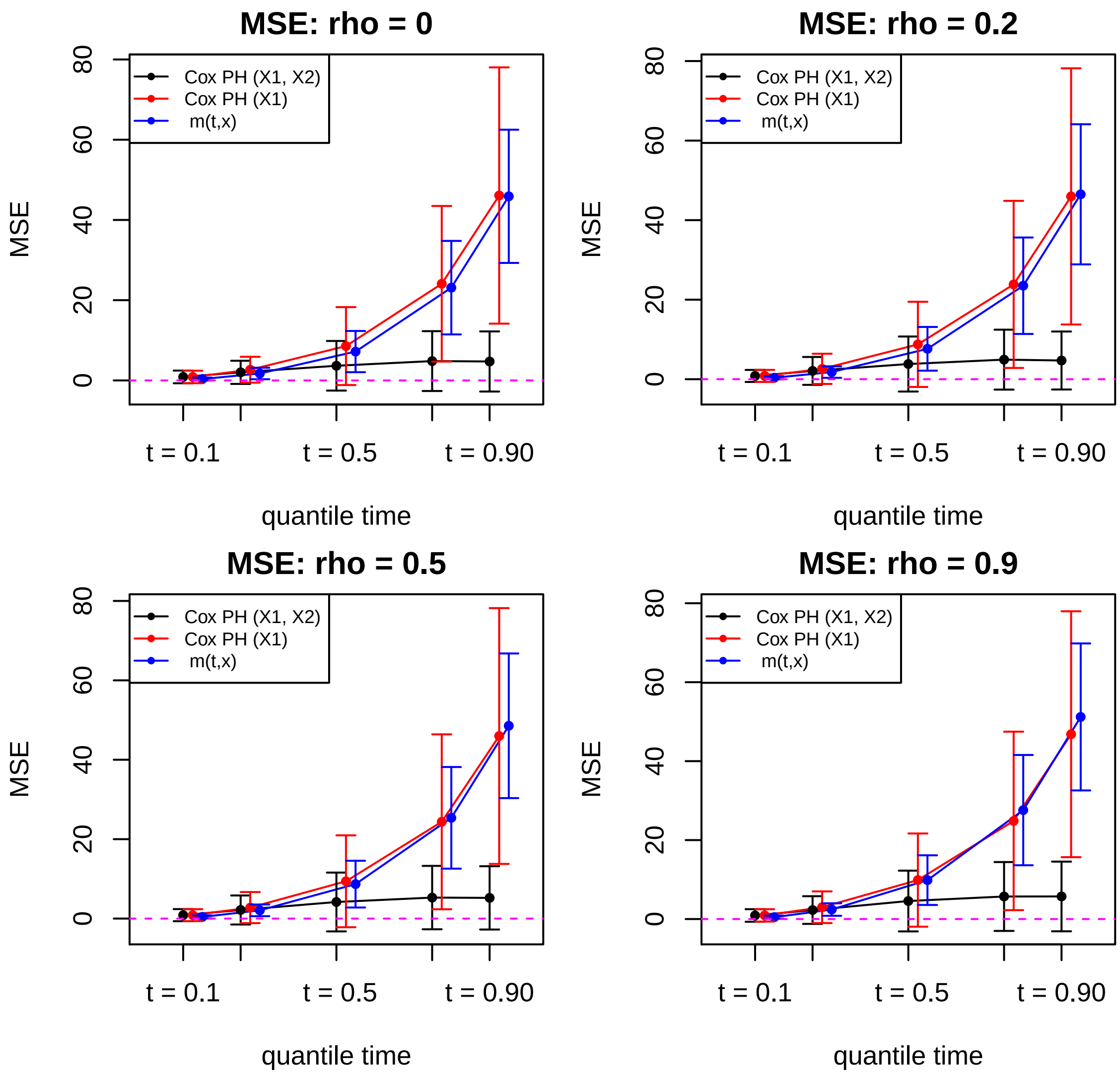
(values were timed 1000)

Table 4: Table of MSE

	cox12	cox1	cox2	m()	m()hat	dikta	dikta hat
rho = 0							
qt=0.1	0.877	0.866	0.893	0.27	0.349	0.269	0.347
qt=0.25	2.016	2.666	2.726	1.215	1.727	1.205	1.717
qt=0.5	3.651	8.56	8.603	5.222	7.189	5.173	7.141
qt=0.75	4.82	24.087	23.984	19.265	23.121	19.034	22.901
qt=0.9	4.72	46.095	45.299	41.545	45.89	40.868	45.251
rho = 0.2							
qt=0.1	0.821	0.83	0.857	0.295	0.37	0.293	0.368
qt=0.25	2.087	2.606	2.679	1.31	1.795	1.301	1.786
qt=0.5	3.818	8.759	8.881	5.793	7.646	5.742	7.597
qt=0.75	4.922	23.837	23.911	19.853	23.506	19.623	23.287
qt=0.9	4.707	45.969	45.649	42.424	46.495	41.749	45.857
rho = 0.5							
qt=0.1	0.857	0.867	0.894	0.329	0.42	0.327	0.418
qt=0.25	2.164	2.764	2.828	1.471	2.07	1.462	2.061
qt=0.5	4.174	9.391	9.494	6.399	8.675	6.347	8.626
qt=0.75	5.285	24.37	24.352	20.894	25.374	20.664	25.158
qt=0.9	5.21	45.978	45.544	43.586	48.555	42.914	47.928
rho = 0.9							
qt=0.1	0.908	0.926	0.955	0.369	0.475	0.367	0.473
qt=0.25	2.278	2.996	3.082	1.71	2.409	1.701	2.4
qt=0.5	4.563	9.87	10.05	7.186	9.856	7.134	9.807
qt=0.75	5.713	24.859	24.937	22.349	27.594	22.119	27.38
qt=0.9	5.724	46.817	46.45	45.393	51.202	44.721	50.584

(values were timed 1000)


rho = 0
rho = 0.2
rho = 0.5
rho = 0.9
Bias
quantile time
t = 0.1
t = 0.5
t = 0.90
0.2
0.1
−0.1
−0.3
Cox PH (X1, X2)
Cox PH (X1)
m(t,x)


MSE: rho = 0
MSE: rho = 0.2
MSE: rho = 0.5
MSE: rho = 0.9
Cox PH (X1, X2)
Cox PH (X1)
m(t,x)
MSE
0
20
40
60
80
t = 0.1
t = 0.5
t = 0.90
quantile time