International Journal of Mathematics and Statistics Invention (IJMSI)
E-ISSN: 2321 - 4767 P-ISSN: 2321 - 4759
www.ijmsi.org Volume 7 Issue 2 || February , 2019 ||PP-26-32

Inference in a general alternative for the Cox model and its application in clinical trials

Mohamed Ali HAFDI
IMI laboratory Member, Professor at High school of technology Laayoune, Ibn Zohr university
e-mail: m.hafdi@uiz.ac.ma

ABSTRACT: *In this paper I propose a new model as a general alternative of the Cox model which can be applied in case of a monotonic hazard ratio and also when the cross effect of hazard rates (also the survival function) is observed. I give a semiparametric estimation of parameters base on modified partial likelihood (MPL). I derive the limit distribution of the MPL estimator and I investigate a finite samples properties of this estimation by simulation. Real data examples are considered in the end of this paper.*
KEYWORDS: *Cox model, Partial likelihood, Survival function, Proportionality assumption.*

Date of Submission: 05-08-2019 Date of Acceptance: 20-08-2019

I. INTRODUCTION

The Cox model is widely used in in several fields for describing the relationship beteween the hazard rates and associeted covariates. A basic assumption in this model is the proportionality of the effects of the covariates on the hazard rates, which means that the hazard ratio for two units is constant on time. In many data analysis this assumption is not valid because the hazard ratio can vary with time as in case of observed cross-effect of hazard rates as in the well known data concerning effects of chemotherapy and radiotherapy on the survival times of gastric cancer patients (Stablein and Koutrouvelis (1985)).
In this paper a semiparametric model is proposed, which generalizes the Cox model. I describe my model in the next section. In section 3 the modified partial likelihood (see Bagdonavičius and Nikulin (2002)) is used to estimate the model parameters and is performed by a simulation study in section 4. As an application a real example is considered in section 5.

II. A GENERAL ALTERNATIVE

Let $S_x(t)$ and $\lambda_x(t)$ be the survival and hazard rate functions under a p-dimensional covariate $x = (x_1, ..., x_p)$. Denote by
$\Lambda_x(t) = \int_0^t \lambda_x(u)du = -log(S_x(t))$, $t \geq 0$, the cumulative hazard rate under x. The Cox model express the hazard rate according to x as follow

$$\lambda_x(t) = e^{\beta^T x}\lambda_0(t), \tag{1}$$

where β is a vector of unknown parameters and $\lambda_0(t)$ stands for an unknown baseline hazard function. Under model (1) the cumulative hazard rate under x has the forme

$$\Lambda_x(t) = e^{\beta^T x}\Lambda_0(t),$$

where $\Lambda_0(t) = \int_0^t \lambda_0(t)dt$ is the cumulative baseline hazard function which supposed verifiying $\Lambda_0(\infty) = \infty$. One of the main assumptions in the model (1) is of course the proportionality, that the ratio of two hazard under a covariates x and y is constant in time,

$$\frac{\lambda_x(t)}{\lambda_y(t)} = e^{\beta^T (x-y)}.$$

Several variants and generalization of Cox model where proposed, see for example Aalen (1980), Hsieh (2001), Bagdonavičius and Nikulin (1999,2002). In this paper the suggested model is defined as:

$$\lambda_x(t) = e^{\beta^T x} (1 + \Lambda_0(t))^{e^{-\gamma^T x}-1} \lambda_0(t). \tag{2}$$

It implies that for different covariates x and y

$$\frac{\lambda_x(t)}{\lambda_y(t)} = e^{\beta^T (x-y)} (1 + \Lambda_0(t))^{e^{-\gamma^T x}-e^{-\gamma^T y}} .$$

Note that $c_0 = e^{\beta^T (x-y)}$, so the ratio of hazard rates has the following properties:

(i) If $\gamma^T x < \gamma^T y$ then the ratio of hazard increase with time of c_0 to ∞.

(ii) If $\gamma^T x > \gamma^T y$ then the ratio of hazard decrease with time of c_0 to 0.

(iii) If $\gamma = 0$ then we have the Cox model.

We remark that if $c_0 < 1$ in (i) or if $c_0 > 1$ in (ii) then the hazard rates (also the survival functions) intersect in a single point t_0.

III. SEMIPARAMETRIC ESTIMATION

Semiparametric estimation of unknown parameters in Cox model was developed by Cox (1972), Tsiasis (1981), Andersen and Gill (1982). We use here a modified partial likelihood approach (Bagdonavičius and Nikulin (1999)) to estimate the parameters in the proposed model.

Suppose that n patients are observed. The ith of them is observed under the covariate x_i. Denote by T_i and C_i the failure and censoring times for the ith patient and set

$$X_i = \min(T_i, C_i), \;\; \delta_i = \mathbf{1}_{\{T_i \le C_i\}},$$

$$N_i(t) = \mathbf{1}_{\{T_i \le t, \delta_i = 1\}}, \;\; Y_i(t) = \mathbf{1}_{\{X_i \ge t\}},$$

where $\mathbf{1}_A$ denotes the indicator of the event A. Then $N(t) = \sum_{i=1}^n N_i(t)$ and $Y(t) = \sum_{i=1}^n Y_i(t)$ are the numbers of observed failures in the interval $[0, t]$ and patients at risk just before the moment t, respectively. We suppose that failure times T_i are absolutely continuous random variables.

The partial likelihood function (PL) (see Andersen and others (1993)) adapted to model (2)

$$L(\theta) = \prod_{i=1}^n \left(\int_0^\infty \frac{g\left(x_i, \theta, \Lambda_0(v)\right) dN_i(v)}{\sum\limits_{j=1}^n Y_j(v) g\left(x_j, \theta, \Lambda_0(v)\right)} \right)^{\delta_i} \tag{3}$$

where $\theta = \left(\beta^T, \gamma^T\right)^T$ and $g(x, \theta, u) = e^{\beta^T x} \left(1+u\right)^{e^{-\gamma^T x} - 1}$, depends on unknown cumulative baseline hazard rates $\Lambda_0(t)$.

Let us consider the *modified partial likelihood function* (see Bagdonavičius and Nikulin (1999)) (MPL):

$$\tilde{L}(\theta) = \prod_{i=1}^n \left(\int_0^\infty \frac{g\left(x_i, \theta, \tilde{\Lambda}_0(v,\theta)\right) dN_i(v)}{\sum\limits_{j=1}^n Y_j(v) g\left(x_j, \theta, \tilde{\Lambda}_0(v,\theta)\right)} \right)^{\delta_i} \tag{4}$$

where the the random function $\tilde{\Lambda}_0(t, \theta)$ which is formulated using the Doob-Meier decomposition, is obtained recurrently from the equations

$$\tilde{\Lambda}_0(t, \theta) = \int_0^t \frac{dN(v)}{S^{(0)}(v-, \tilde{\Lambda}_0, \theta)}, \tag{5}$$

taking $S^{(0)}(v, \tilde{\Lambda}_0, \theta) = \sum\limits_{i=1}^n g(x_i(v), \tilde{\Lambda}_0(v, \theta), \theta) Y_i(v)$ and $\tilde{\Lambda}_0(0, \theta) = 0$.

For fixed θ the "estimator" $\tilde{\Lambda}_0$ can be found recurrently. Really, let $T_1^* < ... < T_r^*$ be observed and ordered distinct failure times, $r \le n$. Note by d_l the number of failures at the moment T_l. Then

$$\tilde{\Lambda_0}(0; \theta) = 0, \quad \tilde{\Lambda_0}(T_1^*; \theta) = \frac{d_1}{S^{(0)}(0, \tilde{\Lambda}_0, \theta)},$$

$$\tilde{\Lambda_0}(T_{l+1}^*; \theta) = \tilde{\Lambda_0}(T_l^*; \theta) + \frac{d_{l+1}}{S^{(0)}(T_l^*, \tilde{\Lambda}_0, \theta)}) \tag{6}$$

for $l = 1, ..., r-1$.
So the modified score functions associated to (4) are

$$\tilde{U}_j(\theta) = \sum_{i=1}^{n} \int_0^{\infty} \{w_j^{(i)}(u, \theta, \tilde{\Lambda}_0) - E_j(u, \theta, \tilde{\Lambda}_0)\}\, dN_i(u), \tag{7}$$

where $w_j^{(i)}(t, \theta, \Lambda_0) = \frac{\partial}{\partial \beta_j} \log\{\lambda_i(t, \theta)\} = x_j^{(i)}$, $w_{p+j}^{(i)}(t, \theta, \Lambda_0) = \frac{\partial}{\partial \gamma_j} \log\{\lambda_i(t, \theta)\} = -x_j^{(i)} e^{-\gamma^T x^{(i)}} \ln(1 + \Lambda_0(t))$, for any $j \in \{1, ..., p\}$,

$$E_j(v, \theta, \Lambda_0) = \frac{S_j^{(1)}(v, \theta, \Lambda_0)}{S^{(0)}(v, \theta, \Lambda_0)},\ S_j^{(1)}(v, \theta, \Lambda_0) = \sum_{i=1}^{n} w_j^{(i)}(v, \theta, \Lambda_0) Y_i(v) g\left(x_i, \theta, \Lambda_0(v)\right),$$

for any $j \in \{1, ..., 2p\}$.

The estimator of the survival function under any value x of the covariate is

$$\hat{S}_x(t) = \exp\left\{-e^{(\hat{\beta}+\hat{\gamma})^T x} \left(\left(1 + \hat{\Lambda}_0(t)\right)^{e^{-\gamma^T x}} - 1\right)\right\}, \tag{8}$$

with $\hat{\Lambda}_0(t) = \tilde{\Lambda}_0(t, \hat{\theta})$ and $\hat{\theta}$ is the maximum modified partial likelihood estimator of θ.

IV. ASYMPTOTIC PROPERTIES OF THE ESTIMATORS

Suppose that maximal time given for experiment is $\tau \in]0, +\infty[$ and all items which did not fail and were not censored before τ, are censored at this moment.
Denote by θ_0 the true value of θ under the model (2), $\|A\| = sup_{i,j} |a_{ij}|$ the norm of the matrix $A = (a_{ij}$, $A^{\otimes 2}$ the product AA^T,

$$g_1 = \frac{\partial}{\partial \theta} g,\ w^{(i)} = \left(w_j^{(i)}\right)_{j=1,...,2p},\ S^{(1)} = \left(S_j^{(1)}\right)_{j=1,...,2p},\ E = (E_j)_{j=1,...,2p},$$

$$S_\star^{(0)}(v, \theta) = \sum_{i=1}^{n} Y_i(v) g(x_i, \Lambda_0(v), \theta) g_2(x_i, \Lambda_0(v), \theta),$$

$$S_\star^{(1)}(v, \theta) = \sum_{i=1}^{n} g_1(x_i, \Lambda_0(v), \theta) g_2(x_i, \Lambda_0(v), \theta),$$

$$S^{(2)}(v, \theta) = \sum_{i=1}^{n} \frac{\partial w^{(i)}(v, \theta)}{\partial \theta} Y_i(v) g(x_i, \Lambda_0(v), \theta),$$

$$S_\star^{(2)}(v, \theta) = \sum_{i=1}^{n} g(x_i, \Lambda_0(v), \theta) g_3(x_i, \Lambda_0(v), \theta),$$

where g_2 and g_3 denotes respectively the partial derivatives of g and g_1 with respect to the second argument.

Assumptions A :

a)- Suppose that exist a neighbourhood Θ of θ_0 and continuous on Θ uniformly in $t \in [0, \tau]$ and bounded on $\Theta \times [0, \tau]$ functions $s^{(k)}(v, \theta)$, $s_\star^{(k)}(v, \theta)$, such that $s^{(0)}(v, \theta_0) > 0$ on $[0, \tau]$ and

$$\sup_{\theta \in \Theta, v \in [0, \tau]} \|\frac{1}{n} S^{(k)}(v, \theta) - s^{(k)}(v, \theta)\| \longrightarrow 0 \ \text{ as } \ n \longrightarrow \infty,$$

$$\sup_{\theta \in \Theta, v \in [0, \tau]} \|\frac{1}{n} S_\star^{(k)}(v, \theta) - s_\star^{(k)}(v, \theta)\| \longrightarrow 0 \ \text{ as } \ n \longrightarrow \infty,\ (k = 0, 1, 2),$$

$$\sup_{\theta\in\Theta, v\in[0,\tau]} \| \frac{\partial E(v,\theta}{\partial\theta} - \frac{\partial e(v,\theta)}{d\theta} \| \xrightarrow{P} 0, \; e(v,\theta) = \frac{s^{(1)}(v,\theta)}{s^{(0)}(v,\theta)}.$$

b)- $\Lambda_0(\tau) < \infty$.
Put $e_\star(v,\theta) = s_\star^{(0)}(v,\theta)/s^{(0)}(v,\theta), h(t,\theta) = \exp(\int_0^\tau e_\star(v,\theta) d\Lambda_0(v))$,

$$h1(v,\theta) = \frac{s^{(1)}(v,\theta)s_\star^{(0)}(v,\theta) - s^{(0)}(v,\theta)s_\star^{(1)}(v,\theta)}{s^{(0)}(v,\theta)} - s_\star^{(2)}(v,\theta),$$

$$w(v,\theta) = e(v,\theta) - \frac{1}{h(v,\theta)s^{(0)}(v,\theta)} \int_v^\tau h_1(s,\theta)h(s,\theta)d\Lambda_0(s).$$

c)- The symmetrical matrix

$$\Sigma_1(\theta_0) = -\int_0^\tau \left(s^{(2)}(u,\theta_0) - \frac{de(u,\theta_0)}{d\theta} s^{(0)}(u,\theta_0) \right) d\Lambda_0(u)$$

is positive definite.
d)- $\int_0^\tau J(v)(w(v,\theta_0) - E(v,\theta_0))^{\otimes 2} s^{(0)}(u,\theta_0) d\Lambda_0(v) < \infty$.

Theorem *Under Assumptions A we have*

$$n^{1/2}(\hat\theta - \theta_0) \xrightarrow{D} N(0, \Sigma_1^{-1}(\theta_0)).$$

Sketch of the proof. Similarly as in Bagdonaviçus and Nikulin(1999)

$$n^{1/2}(\hat\theta - \theta_0) = \left(-\frac{1}{n}\frac{d\tilde U(\theta_0)}{d\theta} \right)^{-1} n^{-1/2}\tilde U(\theta_0) + o_p(1), \tag{9}$$

$$n^{-1/2}\tilde U(\theta_0) = n^{-1/2}\sum_{i=1}^n \int_0^\tau J(u)(w^{(i)}(u,\theta_0) - w(u,\theta_0))dM_i(u) + o_p(1),$$

and

$$\begin{aligned} -\frac{1}{n}\frac{d\tilde U(\theta_0)}{d\theta} &= -\frac{1}{n}\int_0^\tau \left(S^{(2)}(u,\theta_0) - \frac{dE(u,\theta_0)}{d\theta} S^{(0)}(u,\theta_0) \right) d\Lambda_0(u) \\ -\frac{1}{n}\sum_{j=1}^n \int_0^\tau & \left(\frac{dw_j(u,\theta_0)}{d\theta} - \frac{dE(u,\theta_0)}{d\theta} \right) dM_j(u) + o_p(1) \xrightarrow{Pr} \Sigma_1(\theta_0), \end{aligned} \tag{10}$$

where M_j is the martingale process obtained from the DOOB-MEIER decomposition.
The predictable variation for $n^{-1/2}\tilde U(\theta_0)$ is

$$\begin{aligned} < n^{-1/2}\tilde U(\theta_0) > &= \frac{1}{n}\sum_{i=1}^n \int_0^\tau J(u)(w_i(u,\theta_0) - w(u,\theta_0))^{\otimes 2} u(x_i, \Lambda_i(u,\theta_0),\theta_0) Y_i(u) d\Lambda_0(u) \\ &= -\frac{1}{n}\int_0^\tau J(u)(S^{(2)}(u,\theta_0) - \frac{dE(u,\theta_0)}{d\theta} S^{(0)}(u,\theta_0)) d\Lambda_0(u) \\ &\quad + \frac{1}{n}\int_0^\tau J(u)(w(u,\theta_0) - E(u,\theta_0))^{\otimes 2} S^{(0)}(u,\theta_0) d\Lambda_0(u). \end{aligned}$$

Under assumption d)-,

$$\frac{1}{n}\int_0^\tau J(u)(w(u,\theta_0) - E(u,\theta_0))^{\otimes 2} S^{(0)}(u,\theta_0) d\Lambda_0(u) \xrightarrow{Pr} 0, \; as\ n \to \infty,$$

so

$$< n^{-1/2}\tilde{U}(\theta_0) > \xrightarrow{Pr} \Sigma_1(\theta_0).$$

Under assumptions A and in the same way as it is shown in Bagdonaviçus and Nikulin(1999), we have

$$n^{-1/2}\tilde{U}(\theta_0) \xrightarrow{D} N(0, \Sigma_1(\theta_0)). \tag{11}$$

From (9)-(11) the proof is complete.

V. SIMULATION STUDY

In this section we consider the model given in (2) with time-independent univariate covariates $x = x1$ and bivariate covariate $x = (x_1, x_2)$ where x_1 and x_2 are independent and generated from Uniform[-2,2] distribution. The failure time T_x is generated from a following model:

$$\lambda_x(t) = e^{\beta^T x}\,(1+t)^{e^{-\gamma^T x}-1},$$

which is a particular case of model (2) ($\Lambda_0(t) = t$). We take for $\theta = (\beta_1, \gamma_1)$ three choices (2,0) (Cox model hold), (2,-1),(2,1). In bivariate covariate cases, we take $\theta = (\beta_1, \beta_2, \gamma_1, \gamma_2)$ two vectors (2,1,0,0) (Cox model hold) and (2,1,1,1). For n number of units we consider two values 100 and 200. The censoring time C_x is considered constant ($C_x = d$) and generated independently of T_x. The constant d is calculated from the formulae $p = S_x(d)$ to garantie the chosen censoring probability p. Two values considered of p are $p = 0$ (a complete data) and $p = 0.2$ (censoring percentage is 20%). We note that The simulation consists of 2000 replication for each of considered cases. The results are resumed in the two following tables. Note that the values in parentheses are variances of parameter estimates.

From table 1 and table 2 the MPL estimates $\hat{\theta}$ work well in all considered cases. They are nearly unbiased and have a small variance in all cases. Their bias and variance decreases and increases respectively in function of n and p. It means that $\hat{\theta}$ converge on average and quadratic average and the speed of convergence depends on p.

		$\beta = 2$	$\gamma = 0$	$\beta = 2$	$\gamma = -1$	$\beta = 2$	$\gamma = 1$
n	p	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}$
100	0	1.9321 (0.1005)	-0.2866 (0.3748)	2.0652 (0.0900)	-0.9479 (0.3342)	2.0274 (0.0371)	0.9995 (0.0086)
	0.2	1.9953 (0.1011)	-0.2196 (0.3946)	2.1098 (0.1040)	-0.7173 (0.4234)	2.0333 (0.0391)	0.9985 (0.0089)
200	0	1.9500 (0.0548)	-0.1871 (0.2441)	2.0350 (0.0470)	-0.9721 (0.2190)	2.0127 (0.0173)	0.9990 (0.0038)
	0.2	1.9838 (0.0508)	-0.1539 (0.2662)	2.0805 (0.0511)	-0.7838 (0.3205)	2.0087 (0.0267)	0.9934 (0.0054)

Table 1: Univariate covariate cases

		$\beta_1 = 2$	$\beta_2 = 1$	$\gamma_1 = 0$	$\gamma_2 = 0$	$\beta_1 = 2$	$\beta_2 = 1$	$\gamma_1 = 1$	$\gamma_2 = 1$
n	p	$\hat{\beta}_1$	β_2	$\hat{\gamma}_1$	γ_2	$\hat{\beta}_1$	β_2	$\hat{\gamma}_1$	γ_2
100	0	1.9435 (0.1027)	0.9777 (0.0429)	-0.3865 (0.3865)	-0.1608 (0.1264)	2.0450 (0.0362)	1.0294 (0.0209)	0.9955 (0.0067)	1.0043 0.0086
	0.2	1.9849 (0.0.0992)	0.9932 (0.0463)	-0.2863 (0.3935)	-0.1408 (0.1504)	2.0615 (0.0434)	1.0216 (0.0268)	1.0003 (0.0084)	1.0024 0.0097
200	0	1.980 (0.0502)	0.9854 (0.0191)	-0.1376 (0.1706)	-0.0691 (0.0465)	2.0214 (0.0164)	1.0157 (0.0088)	0.9975 0.0030	1.0029 0.0036
	0.2	1.9846 (0.0493)	0.9918 (0.0195)	-0.1439 (0.1962)	-0.0711 (0.0608)	2.0245 (0.0175)	1.0101 (0.0119)	1.0016 (0.0035)	1.0039 (0.0041))

Table 2: Bivariate covariate cases

VI. ANALYSIS OF RADIO-CHEMOTHERAPY DATA

In this section we give analysis of the two-sample data of Stablein and Koutrouvelis [1] concerning effects of chemotherapy and chemotherapy plus radiotherapy on the survival times of gastric cancer patients. This example is also analysed in Hsieh(2001). The number of patients is 90. By plotting the Kaplan-Meier (KM) estimates of survival functions pertaining to the both treatment groups (Fig. 1), a crossing hazards phenomenon is clearly manifest at a time 781 days. These both estimated curves are visually very different from each other, especially before their crossing. This crossing clearly indicates a violation of the proportional hazards assumption and renders the classical analysis unable to reflect properly the effects of differences between the two treatments. This can also be seen in Fig. 1 in which the two proportional hazards estimated survival curves produced by the partial likelihood approach via the PH model are rather close to each other along the time interval considered. The corresponding estimated parameter β is -0.1059. To accommodate the crossing hazard phenomenon, we performed inference using MPL method of estimation coding 1 for chemo-therapy and 0 for chemo+radio-therapy.

The estimated values obtained by the MPL method of $\theta = (\beta, \gamma)^T$ in model (2) are (-1.9032,-1.3758).

As discussed in Hsieh(2001), one interpretation is that, although the addition of radiotherapy is beneficial to some gastric cancer patients, there is an increased risk of dying associated with the treatment in the two first years or so for other patients. The resulting inference indicates that the radiotherapy would first be detrimental to a patient's survival but becomes beneficial later on.

In Figure 2 we see that the MPL estimated survival functions fit the two nonparametric survival estimates well.

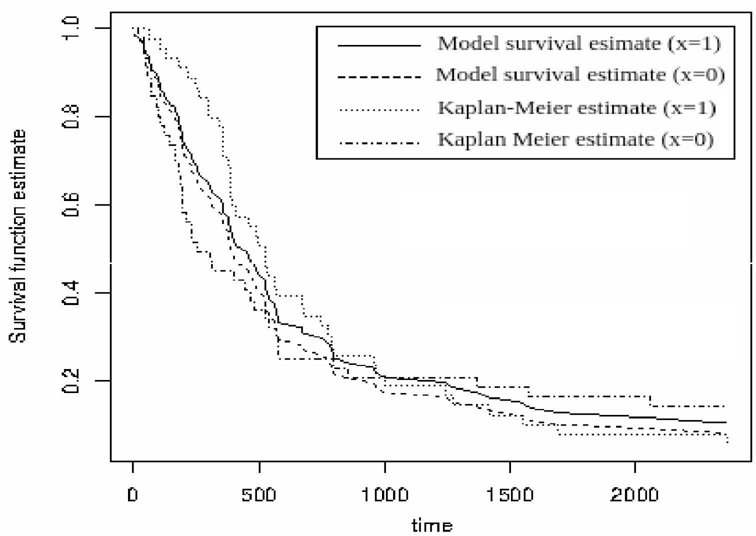

Figure 1: Comparison of Kaplan-Meier and estimates of survival functions under Cox model.

VII. CONCLUSION

The general alternative of the Cox model considered in this paper with the MPL procedure, has proved firstly through the simulation study and secondly through the real data application, its importance to be used in survival analysis fields. It can be also used to test the hypothesis of proportionality of the hazard rates. This will be the subject of next work.

ACKNOWLEDGMENTS

We would like to thank the associate editor and the referees for useful comments and suggestions that improved the paper.

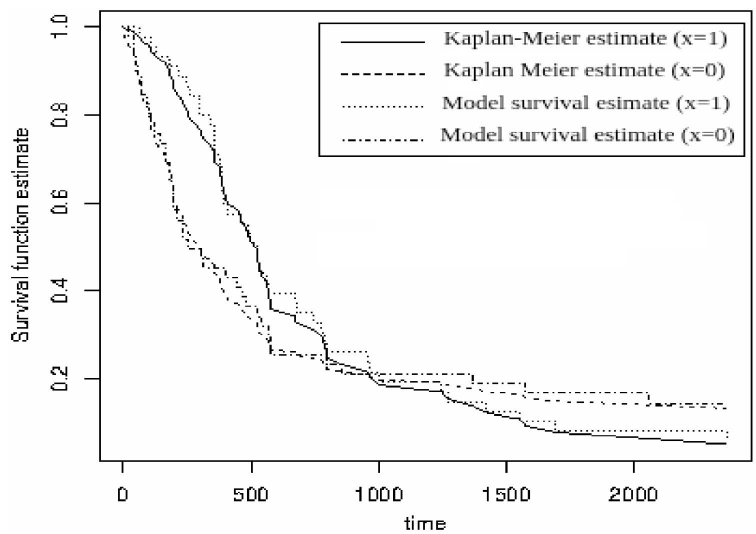

Figure 2: Comparison of Kaplan-Meier and MPL estimates of survival functions.

REFERENCES

[1] Andersen, P. K., Gill, R. D. (1982). Coxs regression model for counting processes: a large sample study. *Ann. Statist.* 10:11001120.

[2] Andersen, P., Borgan, O., Gill, R., Keiding, N. (1993): *Statistical Models Based on Counting Processes*. Springer, New York.

[3] Bagdonavičius V. and Nikulin, M. (1999) Generalized Proportional Hasards Model Based on Modified Partial Likelihood. *Lifetime Data Analysis*, 5, 329- 350.

[4] Bagdonavičius V. and Nikulin M. (2002). *Accelerated Life Models: Modeling and Statistical Analysis.* Boca Raton: Chapman and Hall/CRC.

[5] Cox, D. R. (1972) Regression models and life tables. *J. R. Statist. Soc.*;B, **34**:187-220.

[6] Hsieh, F. (2001) On heteroscedastic hazards regression models: theory and application. *J. R. Statist. Soc.*; B, **63**:63-79.

[7] Stablein, D. M., Koutrouvelis, I. A. (1985) A two sample test sensitive to crossing hazards in uncensored and singly censored data. *Biometrics*; **41**: 643-652.

[8] Tsiatis, A. A. (1981) A large sample study of Coxs regression model. *Ann. Statist.*, 9, 93-108.

Hafdi Mohamed Ali" Inference in a general alternative for the Cox Model and its application in clinical trials" International Journal of Mathematics and Statistics Invention (IJMSI), vol. 07, no. 02, 2019, pp.26-32