Applied Mathematics & Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/amisl/050101

Concavity of the Distribution Function Using Interval Censored Data with Covariates

Bader Aljawadi∗

Department of Mathematics, Hebron University, Palestine

Received: 11 Dec. 2015, Revised: 6 May 2016, Accepted: 8 May 2016
Published online: 1 Jan. 2017

Abstract: In many real applications in survival analysis, estimation of the distribution function and hence the survival function is common in practice, where the problem of estimating a smooth shape-constrained distribution function has recently received some attention. In this article an interesting proposition is built on the assumption that the distribution function of the random variable (failure time) is a concave function. Where the concavity of the distribution function is discussed in the presence of covariates considering interval censoring model. It is shown that concavity of the distribution function is well defined under the proposed situation.

Keywords: Interval censoring, Hazard function, Covariates, Concavity, Taylor approximation

1 Introduction

In many situations in survival analysis, one may interested in the distribution of times to event data which might be used to extract the distribution of event time variable based on the incomplete or censored data sets. Elaborations on models incorporate censoring mechanisms can be found in the literature (See, e.g. [11] for models that incorporating these typical mechanisms which prevent time event times from being observed directly). However, one of the most common censoring models in survival analysis is right censoring, where the observed information is whether the event time occurs before an observed censoring point (exact failure time), or we observe the censoring time point given the information that the exact failure time occurred beyond a censoring time. However, right censoring is very common in several different applications such as cancer clinical trials, where survival function and hence the distribution function can be obtained through different techniques such as Nelson-Aalen and Kaplan-Meier estimators. Another well-known censoring model in survival analysis is interval censoring, where, the event time is not exactly observed, and the only observed information is that it belongs to an available interval ([1,2,10,13]).

However, assume that $T_1, T_2, \ldots, T_n$ be independent random event times with unknown distribution function $F \in (0,\infty)$, here $F(T) = 0$ means that the event of interest will never occur at all. Note that $F(T) = 0$ equivalent to $S(T) = 1$ since $F(T) = 1 - S(T)$. In interval censoring, the set of event times is not observed and instead of that, we only have finitely m inspection points $0 < \tau_{i1} < \tau_{i2} < \ldots < \tau_{im} < \infty, i = 1,\ldots,n$, and each unit i inspected at each of the assigned inspection points and determining if the interested event occurred at that point or not. More formally, the observed information is:

$$X_{ij} = 1\{\tau_{i,j-1} < T_i < \tau_{i,j}\}, \quad \text{for } 1 \leqslant j \leqslant (m_i + 1), \text{ where } \tau_{i,0} = 0 \text{ and } \tau_{i,(m_i+1)} = \infty$$

Under various censoring models, one of the main interests is estimation of the distribution function which might be obtained parametrically or nonparametrically. Where in parametric approach it is assumed that the distribution of survival data is known and some common distributions can be used to represent this function such as Weibull and Gompertz functions. On the other hand, the nonparametric approach is also common in practice, where, the non-parametric estimation of the survival function is mostly concerned with detecting the trend in the data set without strong parametric assumptions on its form ([4,5]). However, some techniques is suited to estimate the survival function and hence the distribution

∗ Corresponding author e-mail: baderj@hebron.edu

function under various censoring models such as Kaplan-Meier estimator which is suitable for right censoring case and Turnbull estimator which is suited for interval censoring case ([8]).

Another way of constructing adaptive tools to obtain the nonparametric maximum likelihood estimator of the distribution function is well-understood with respect to the distribution function computation as well as its asymptotic properties, where this function can be obtained by deriving the convex minorant of a suitable function depending on the data set ([6,7]).

Furthermore, Inference under shape constraints has been considered for recent activities. Where in survival analysis researchers have paid more attention on the estimation of a smooth distribution function under constraints that this function satisfies certain qualitative properties, such as concavity and monotonicity on certain subsets of its domain. [4] considered three nonparametric estimators of the distribution function based on mixed-case interval censored data when the covariates are excluded from the data set, they assumed that the distrbition function of event times is concave or unimodal and they proposed some algorithms for the computation of the derived estimators. Furthermore, [12] proposed two methods to find shape-constrained density estimates, where these methods can be used for univariate or higher-dimensional kernel density estimation with shape constraints. So, this approach is appealing for a main reason which is that extraction of alternative nonparametric estimators of the distribution function is typically require such properties ([3]). However, this article gives an overview for one of the most important shape constraints which is concavity of the distribution function estimator in case of interval censored data when some covariates are available. In section 2 the likelihood function in case of interval censoring with covariates is constructed and in section 3, it will be shown that the concave maximum likelihood estimators are well defined in the interval censoring model.

2 The Likelihood Function

In interval censoring model and for a group of data set consists of n observations then, the exact failure time is not fully observed and the only known information is that it belongs to an observed interval such that $T_i \in [l_i, r_i], \forall i = 1, \ldots, n$, where l_i and r_i are the left and right endpoints of the observed interval respectively. Therefore, the general form of the likelihood function for interval censoring model is given as follows:

$$L(F) = \prod_{i=1}^{n} [F(r_i) - F(l_i)]$$

However, since the covariates involved in the analysis then, a link function might be employed to investigate the effect of covariates on the proposed procedure, where in such situation the Cox model can be used as a result of the direct combination between hazard and survival functions. Where for a given covariates vector $Z = (z_1, z_2, \ldots, z_n)$ then the hazard function $\lambda(t)$ can be defined as follows:

$$\begin{aligned}\lambda(t) &= \lim_{dt\to 0} \frac{P(t \leqslant T \leqslant t+dt/T \geqslant t)}{dt} \\ &= \lim_{dt\to 0} \frac{P(t \leqslant T \leqslant t+dt)}{dtP(T \geqslant t)} \\ &= \frac{f(t)}{S(t)} \end{aligned} \tag{1}$$

The Cox hazard model given the covariates vector Z is defined as

$$\lambda(t) = \lambda_{\circ}(t/Z) exp(\beta^T Z) \tag{2}$$

where $\lambda_{\circ}(t)$ is the baseline hazard function and β^T is the parameters vector. Thus, based on the expressions in $(1)$and (2) then the survival function can be defined as

$$\begin{aligned}\lambda(t) &= \frac{f(t)}{S(t)} = -\frac{d}{dt} logS(t) \\ S(t) &= exp\Big(-\int_0^t \lambda(x)dx\Big) \\ &= exp\Big(-\int_0^t \lambda_{\circ}(x/Z)exp(\beta^T Z)dx\Big) \\ &= exp\Big(-\Lambda_{\circ}(t/Z)exp(\beta^T Z)\Big) \end{aligned} \tag{3}$$

where $\Lambda_{\circ}(t/Z)$ is the baseline cumulative hazard function.

The cumulative baseline hazard function is need to be estimated using a reasonable technique such as Breslow or Scheik and Zhang estimators. Recently [2] has proposed the Taylor approximation to estimate the baseline hazard function which is most attractive since it may produce more accurate and smooth estimation of the underlined function and hence, this technique will be employed in this article as it will shown in the sequel. Thus, the log-likelihood function when the survival function is replaced by the expression in equation (3) can be written as follows:

$$\begin{aligned} l(\beta) = \sum_{i=1}^{n} log\Big[&exp(-\Lambda_{\circ}(l_i)exp(\beta^T Z)) - \\ &exp(-\Lambda_{\circ}(r_i)exp(\beta^T Z))\Big] \end{aligned} \tag{4}$$

2.1 Taylor Approximation

The Taylor series, which is more general case of the Maclaurin series is mainly used for approximation functions. However, this technique is proposed to estimate the cumulative baseline hazard function even though several specifications for this function are common, such as in the parametric settings some well known distributions can be used such as Weibull function. On the other hand, alternative semi-parametric approach can be used considering that the baseline hazard function to be piecewise constant which leads to the iterative convex minorant (ICM) technique proposed by [14]. But this approach has a main drawback which is that it may produce the baseline hazard function as a step function and hence not continuous, [2]. Thus, and to overcome the drawbacks of some of the proposed techniques for baseline hazard function, the Taylor approximation will employed and the likelihood ratio test can be used as a model selection tool to obtain the optimal order for Taylor series. However, let's consider the Taylor series of the baseline hazard function with q^{th} order as

$$\lambda_\circ(t/\phi) = \sum_{m=0}^{q} \frac{\lambda_m}{m!} t^m \tag{5}$$

where ϕ is the Taylor series parameters vector such that $\phi = (\lambda_\circ, \lambda_1, ..., \lambda_q)$.

In order to ensure $\lambda_\circ(t/\phi) \geqslant 0$, the log-hazard function is considered such that

$$log\Big(\lambda_\circ(t/\phi)\Big) = \sum_{m=0}^{q} \frac{\lambda_m}{m!} t^m \tag{6}$$

Then, the cumulative baseline hazard function can be defined as

$$\begin{aligned} \Lambda_\circ(t/\phi) &= \int_0^t \lambda_\circ(y/\phi)dy \\ &= \int_0^t exp\Big(\sum_{m=0}^{q} \frac{\lambda_m}{m!} y^m\Big)dy \end{aligned} \tag{7}$$

Therefore, the baseline cumulative hazard function at the left and right endpoints of the observed intervals can be defined respectively as follows:

$$\begin{aligned} \Lambda_\circ(l_i/\phi) &= \int_0^{l_i} exp\Big(\sum_{m=0}^{q} \frac{\lambda_m}{m!} y^m\Big)dy \\ \Lambda_\circ(r_i/\phi) &= \int_0^{r_i} exp\Big(\sum_{m=0}^{q} \frac{\lambda_m}{m!} y^m\Big)dy,\ \forall i = 1,...,n \end{aligned} \tag{8}$$

Likelihood ratio test:

The likelihood ratio principle can be used in order to obtain best order of this Taylor approximation. Where this technique can be employed by fitting the log likelihood function considering only one parameter $(i.e.\ \theta = \lambda_\circ)$ as well as the covariates vector β, and denote the fitted value as $l_0 = max[l(\hat{\beta}, \hat{\theta})]$. Then, the log likelihood function can be fitted again based on two Taylor parameters $(i.e.\ \theta = (\lambda_\circ, \lambda_1))$ and the covariates vector β and denote this new fitted value by $l_1 = max[l(\hat{\beta}, \hat{\theta})]$. Then if $-2 \times (l_0 - l_1) < \chi^2_{df,1-\alpha}$ for $df = 1$ and a typical value of the significance level α such as 0.05 then the Taylor approximation can be hold based on one parameter only. This procedure can be repeated consequently based on one more variable in each iteration and so on until the optimal order obtained.

3 Concavity of the Distribution Function

To investigate the concavity of the distribution function estimator based on the log likelihood function given in (4), then the Hessian matrix should be constructed and investigate the negative semi-definite property of this matrix. Thus, let H denotes the Hessian matrix of the log likelihood function based on an $(n+m)$ parameters which is defined as follows:

$$H = \begin{vmatrix} \frac{\partial^2 l}{\partial \beta_1^2} & \frac{\partial^2 l}{\partial \beta_1 \partial \beta_2} & \cdots & \frac{\partial^2 l}{\partial \beta_1 \partial \beta_n} & \frac{\partial^2 l}{\partial \beta_1 \partial \lambda_\circ} & \frac{\partial^2 l}{\partial \beta_1 \partial \lambda_1} & \cdots & \frac{\partial^2 l}{\partial \beta_1 \partial \lambda_m} \\ \frac{\partial^2 l}{\partial \beta_2 \partial \beta_1} & \frac{\partial^2 l}{\partial \beta_2^2} & \cdots & \frac{\partial^2 l}{\partial \beta_2 \partial \beta_n} & \frac{\partial^2 l}{\partial \beta_2 \partial \lambda_\circ} & \frac{\partial^2 l}{\partial \beta_2 \partial \lambda_1} & \cdots & \frac{\partial^2 l}{\partial \beta_2 \partial \lambda_m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 l}{\partial \beta_n \partial \beta_1} & \frac{\partial^2 l}{\partial \beta_n \partial \beta_2} & \cdots & \frac{\partial^2 l}{\partial \beta_n^2} & \frac{\partial^2 l}{\partial \beta_n \partial \lambda_\circ} & \frac{\partial^2 l}{\partial \beta_n \partial \lambda_1} & \cdots & \frac{\partial^2 l}{\partial \beta_n \partial \lambda_m} \\ \frac{\partial^2 l}{\partial \lambda_\circ^2} & \frac{\partial^2 l}{\partial \lambda_\circ \partial \lambda_1} & \cdots & \frac{\partial^2 l}{\partial \lambda_\circ \partial \lambda_m} & \frac{\partial^2 l}{\partial \lambda_\circ \partial \beta_1} & \frac{\partial^2 l}{\partial \lambda_\circ \partial \beta_2} & \cdots & \frac{\partial^2 l}{\partial \lambda_\circ \partial \beta_n} \\ \frac{\partial^2 l}{\partial \lambda_1 \partial \lambda_\circ} & \frac{\partial^2 l}{\partial \lambda_1^2} & \cdots & \frac{\partial^2 l}{\partial \lambda_1 \partial \lambda_m} & \frac{\partial^2 l}{\partial \lambda_1 \partial \beta_1} & \frac{\partial^2 l}{\partial \lambda_2 \partial \beta_2} & \cdots & \frac{\partial^2 l}{\partial \lambda_1 \partial \beta_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 l}{\partial \lambda_m \partial \lambda_\circ} & \frac{\partial^2 l}{\partial \lambda_m \partial \lambda_1} & \cdots & \frac{\partial^2 l}{\partial \lambda_m^2} & \frac{\partial^2 l}{\partial \lambda_m \partial \beta_1} & \frac{\partial^2 l}{\partial \lambda_m \partial \beta_2} & \cdots & \frac{\partial^2 l}{\partial \lambda_m \partial \beta_n} \end{vmatrix}$$

Evaluation the Hessian matrix in its general form is not applicable since the number of covariates unspecified, so to avoid this problem and to investigate the negative semi-definite property of the Hessian matrix, a nice trick can be used which is basically depends on the Taylor approximation theorem that can be used to re-formalize the functions based on the gradients.

Theorem 1(*Taylor Theorem*). *Let $f : [x,y] \to \mathbb{R}, f, f', f'', ..., f^{n-1}$ be continuous on $[x,y]$ and suppose $f^{(n)}$ exists on (x,y). Then there exists $c \in (x,y)$ such that*

$$\begin{aligned} f(c) = f(x) + \nabla f(x)(c-x) + \nabla^2 f(x)\frac{(c-x)^2}{2} + ... + \\ \nabla^{(n-1)} f(x)\frac{(c-x)^{(n-1)}}{(n-1)!} \end{aligned} \tag{9}$$

which is equivalent to

$$f(x+\delta) = f(x) + \nabla f(x)\delta + \frac{1}{2}\nabla^2 f(x)\delta^2 + O(\| \delta \|)^3 \tag{10}$$

where δ is real number and $O(\| \delta \|)^3$ is the error term.

When $f(x)$ is scalar function with parameters vector $x = (\beta_1, ..., \beta_n, \lambda_\circ, \lambda_1, ..., \lambda_m)$ with length $\omega = n+m$, then the first derivative $\nabla f(x)$ is a $1 \times \omega$ matrix, which can be viewed as an ω-dimensional vector-valued function of the ω-dimensional vector x. For the second derivative $\nabla^2 f(x)$, we can take the matrix of partial derivatives of the functions $\nabla f(x)$. We could write it as $\nabla^2 f(x)$ for the moment. Note that $\nabla^2 f(x)$ is an $\omega \times \omega$ matrix which represent H such that

$$H = \nabla^2 f(x)$$

Now the main task is to verify the negative semi definite property for this matrix based on the defined parameters vector x.

Definition 1.*The square matrix N is negative semi-definite if $\forall z \in \Re^n$, then $ZNZ^T \leqslant 0$. If the inequality is strict for all $z \neq 0$, then N is negative definite.*

Based on Taylor theorem and for the parameters vector x, then the log likelihood function given in (4) can be expressed as follows:

$$l(x+\delta)=l(x)+\nabla l(x)\delta+\frac{1}{2}\delta^T\nabla^2 l(x)\delta+O(\|\delta\|)^3 \tag{11}$$

where $\nabla^2 l(x)$ is the Hessian matrix and by definition (1) this matrix is negative semi definite if $\delta^T H\delta \leqslant 0,\ \forall\delta\in\Re^d$. Therefore, it is sufficient to consider $\nabla^2 l(x)$ and follow to show that: $\delta^T\nabla^2 l(x)\delta\leqslant 0$.

The log likelihood function is:

$$l(x)=\sum_{i=1}^{n} log\Big[exp(-\Lambda_\circ(l_i)exp(\beta^T Z))-exp(-\Lambda_\circ(r_i)exp(\beta^T Z))\Big]$$

where $-\Lambda_\circ(l_i)$ and $-\Lambda_\circ(r_i)$ are the cumulative baseline hazard function at the left and right endpoints of the observed intervals which can be estimated using Taylor approximation given in equation (8) and for simplicity these values will replaced by L_i and R_i respectively. Therefore, the log-likelihood function can be written as:

$$l(x)=\sum_{i=1}^{n} log\Big[exp(-L_i exp(\beta^T Z))-exp(-R_i exp(\beta^T Z))\Big] \tag{12}$$

The two components of the log-likelihood function can be simplified based on the following power series for exponential function such that:

$$\begin{aligned}exp(ax)&=1+ax+\frac{1}{2}(ax)^2+\frac{1}{3!}(ax)^3+\frac{1}{4!}(ax)^4+\dots\\ &=1+ax+\frac{1}{2}(ax)^2+O(\|a\|)^3\end{aligned}$$

Therefore, the expression $exp\Big(-L_i exp(\beta^T Z)\Big),\forall i=1,2,\dots,n$ can be rewritten as follows: $=exp(-L)\Big[1-L(\beta^T Z)-\frac{L(\beta^T Z)^2}{2}$

$$\begin{aligned}exp\Big(-Lexp(\beta^T Z)\Big)&=exp\Big[-L(1+\beta^T Z+\frac{1}{2}(\beta^T Z)^2+\\ &O(\|\beta\|)^3)\Big]\\ &=exp(-L)exp\Big[-L(\beta^T Z+\frac{(\beta^T Z)^2}{2})+\\ &O(\|\beta\|)^3\Big]\\ &=exp(-L)\Big[1-L(\beta^T Z)-\frac{L(\beta^T Z)^2}{2}\\ &+\frac{1}{2}\Big(-L\beta^T Z-\frac{L(\beta^T Z)^2}{2}\Big)^2+O(\|\beta\|)^3)\Big]\end{aligned} \tag{13}$$

Note that $\frac{1}{2}\Big(-L\beta^T Z-\frac{L(\beta^T Z)^2}{2}\Big)^2$ converges to $\frac{1}{2}\Big(L^2(\beta^T Z)^2\Big)$ and hence equation (11) can be written as:

$$\begin{aligned}exp\Big(-Lexp(\beta^T Z)\Big)&=exp(-L)\Big[1-L(\beta^T Z)-\frac{L(\beta^T Z)^2}{2}+\\ &\frac{1}{2}(L^2-L)\beta^T ZZ^T\beta+O(\|\delta\|)^3)\Big]\end{aligned} \tag{14}$$

and in the same manner $exp\Big(-R_i exp(\beta^T Z)\Big),\forall i$ can be written as follows:

$$\begin{aligned}exp(-Rexp(\beta^T Z))&=exp(-R)\Big[1-R(\beta^T Z)-\frac{R(\beta^T Z)^2}{2}+\\ &\frac{1}{2}(R^2-R)\beta^T ZZ^T\beta+O(\|\delta\|)^3)\Big]\end{aligned} \tag{15}$$

However the log likelihood function can be written as follows:

$$\begin{aligned}l(x)&=\sum_{i=1}^{n} log\Bigg[\Bigg(exp(-L_i)-L_i exp(-L_i)\beta^T Z+\\ &\frac{exp(-L_i)}{2}(L_i^2-L_i)\beta^T ZZ^T\beta+O(\|\delta\|)^3\Bigg)-\\ &\Bigg(exp(-R_i)-R_i exp(-R_i)\beta^T Z+\\ &\frac{exp(-R_i)}{2}(R_i^2-R_i)\beta^T ZZ^T\beta+O(\|\delta\|)^3\Bigg)\Bigg]\end{aligned}$$

$$\begin{aligned}&=\sum_{i=1}^{n} log\Bigg[exp(-L_i)-exp(-R_i)+\Big(R_i exp(-R_i)-\\ &L_i exp(-L_i)\Big)\beta^T Z+\frac{1}{2}\Big((L_i^2-L_i)exp(-L_i)-\\ &(R_i^2-R_i)exp(-R_i)\Big)\beta^T ZZ^T\beta+O(\|\delta\|)^3)\Bigg]\end{aligned}$$

An attractive property of rewriting the logarithm expressions is that for any positive random variable a and for any small value $\delta\in\Re$ then

$$\begin{aligned}log(a+\delta)=log(a)+log\Big(1+\frac{\delta}{a}\Big)&=log(a)+\frac{\delta}{a}-\frac{\delta^2}{2a^2}\\ &+O(\|\delta\|)^3\end{aligned} \tag{16}$$

Based on this property, and considering that $a=exp(-L)-exp(-R)$ which is greater than *zero* since the hazard function is an increasing function, then the log-likelihood function can be rewritten as follows

$$\begin{aligned}l(x)=\sum_{i=1}^{n}\Bigg[&log\Big(exp(-L_i)-exp(-R_i)\Big)\\ &+log\Big[1+\Big(\frac{R_i exp(-R_i)-L_i exp(-L_i)}{exp(-L_i)-exp(-R_i)}\Big)\beta^T Z\\ &+\frac{1}{2}\Big(\frac{(L_i^2-L_i)exp(-L_i)-(R_i^2-R_i)exp(-R_i)}{exp(-L_i)-exp(-R_i)}\Big)\beta^T ZZ^T\beta\Big]\\ &+O(\|\delta\|)^3\Bigg]\end{aligned}$$

Let

$$a = exp(-L) - exp(-R)$$
$$b = \frac{Rexp(-R) - Lexp(-L)}{exp(-L) - exp(-R)}$$
$$c = \frac{(L^2 - L)exp(-L) - (R^2 - R)exp(-R)}{exp(-L) - exp(-R)}$$

then the log likelihood function is:

$$l(x) = \sum_{i=1}^{n} \left[log(a_i) + log\left(1 + b_i \beta^T Z + \frac{1}{2} c_i \beta^T Z Z^T \beta + O(\| \delta \|)^3\right) \right]$$

Note that $\left(b_i \beta^T Z + \frac{1}{2} c_i \beta^T Z Z^T \beta\right)^2$ converges to $(b_i \beta^T Z)^2$ and hence based on the property in (15) and after simplification this function can be written as:

$$l(x) = \sum_{i=1}^{n} \left[log(a_i) + b_i \beta^T Z + \frac{1}{2} c_i (\beta^T Z Z \beta^T) - \frac{1}{2} b_i^2 \beta^T Z Z \beta^T + O(\| \delta \|)^3 \right]$$
$$= \sum_{i=1}^{n} \left[log(a_i) + b_i \beta^T Z + \frac{1}{2} (c_i - b_i^2) \beta^T Z Z^T \beta + O(\| \delta \|)^3 \right] \quad (17)$$

Comparing the log likelihood function given in (19) with the general expression given in Theorem 1, then it is easy to notice that

$$\nabla l(x) = \sum_{i=1}^{n} b_i Z$$
$$D^2 l(x) = \sum_{i=1}^{n} (c_i - b_i^2) Z Z^T$$

where $D^2 l(x)$ is the Hessian matrix and to verify the negative semi-definite property of this matrix it is sufficient to show that $\sum_{i=1}^{n} (c_i - b_i^2) \leqslant 0$ or the equivalent inequality $\sum_{i=1}^{n} (b_i^2 - c_i) \geqslant 0, \forall i = 1, \dots, n$.

Lemma 1. *For any two real numbers such that* $0 \leqslant L < R$, *and for*

$$b = \frac{Rexp(-R) - Lexp(-L)}{exp(-L) - exp(-R)}$$
$$c = \frac{(L^2 - L)exp(-L) - (R^2 - R)exp(-R)}{exp(-L) - exp(-R)}$$

Then $b^2 > c, \forall L, R \in \Re$

proof:
Let $\Delta = R - L$ which is greater than zero. Then

$$b = \frac{Rexp(-\Delta) - L}{1 - exp(-\Delta)}$$
$$c = \frac{(L^2 - L) - (R^2 - R)exp(-\Delta)}{1 - exp(-\Delta)}$$

and it follows for $1 - exp(-\Delta) > 0$ that $b^2 > c$ is equivalent to:

$$\left(Rexp(-\Delta) - L\right)^2 > \left(1 - exp(-\Delta)\right)\left((L^2 - L) - (R^2 - R)exp(-\Delta)\right)$$

and hence it can be easily conclude that

$$Rexp(-2\Delta) + L + \left(R^2 - 2LR + L^2 - L - R\right)exp(-\Delta) > 0$$
$$Rexp(-2\Delta) + L + \Delta^2 exp(-\Delta) - (L + R)exp(-\Delta) > 0$$
$$Lexp(-2\Delta) + L - 2Lexp(-\Delta) + \Delta exp(-2\Delta) + \Delta^2 exp(-\Delta) - \Delta exp(-\Delta) > 0$$
$$L\left(1 - exp(-\Delta)\right)^2 + \Delta exp(-\Delta)\left(\Delta + exp(-\Delta) - 1\right) > 0$$

The first term of the inequality is always positive and it is sufficient to show that the other term is also positive. Therefore, assume that:

$$f(\Delta) = \Delta + exp(-\Delta) - 1$$

where Δ is a random variable $\in \Re^+$ then

$$\frac{df}{d\Delta} = 1 - exp(-\Delta) = 0 \Longrightarrow \Delta = 0$$

Thus, the function $f(\Delta)$ is always positive for all Δ since $\Delta \in \Re^+$, and therefore $L\left(1 - exp(-\Delta)\right)^2 + \Delta exp(-\Delta)\left(\Delta + exp(-\Delta) - 1\right)$ is positive for all L and R.

4 Conclusion

In this article, the concavity of the distribution function when covariates involved in the analysis using interval censoring model has been investigated using some well known mathematical techniques especially Taylor approximation which is considered to represent the baseline hazard function. Based on the results shown in this article we can consider estimators of survival function under the shape constraint that the distribution function of event times is concave or unimodal.

Acknowledgement

The author is grateful to Prof. Lutz Dumbgen for his guidance, encouragement and for helpful comments that improved this paper.

References

[1] Aljawadi, B. A., Abu Bakar, M. R. and Ibrahim, N. A., 2012. ”Nonparametric versus Parametric Estimation of the Cure Fraction Using Interval Censored Data” *Communications in statistics-Theory and Methods*, 41(23).
[2] Chen, D. G., Sun, J., Peace, K. E., 2013. ”Interval-Censored Time-to-Event Data: Methods and Applications” *Boca Raton, FL: Chapman and Hall,*.
[3] Cule, M., SAMWORTH, R. and STEWART, M.,2008. ”Maximum likelihood estimation of a multidimensional log-concave density” *Preprint*.

[4] Dmbgen, L., Freitag, S., and Jongbloed, G., 2004. ”Consistency of Concave Regression With an Application to Current Status Data” *Mathematical Methods of Statistics*, 13: 6981.
[5] Dmbgen, L., Rufibach, K., and Dominic, S., 2014. ”Maximum likelihood estimation of a log-concave density based on censored data” *Electronic Journal of Statistics*, 8:1405-1437.
[6] Groeneboom, P., and Wellner, J. A., 1992. ”Information Bounds and Nonparametric Maximum Likelihood Estimation” *Basel: Birkhuser*.
[7] Jongbloed, G., 1998. ”The Iterative Convex Minorant Algorithm for Nonparametric Estimation” *Journal of Computational and Graphical Statistics*, 7: 310321.
[8] Klein, J. P., and Moeschberger, M. L., 2003. ”Survival Analysis: Techniques for Censored and Truncated Data” *(2nd ed.) New York: Springer-Verlag*.
[9] Neuhaus, G., 1999. ”On Testing Tumor Onset Times” *Journal of Statistical Planning and Inference*, 76: 5778.
[10] Sun, J., Zhao, Q., Zhao, X., 2005. ”Generalized log-rank tests for interval-censored failure time data” *Scandinavian Journal of Statistics*, 32: 49-57.
[11] Turnbull, B. T., 1976. ”The Empirical Distribution Function With Arbitrar-ily Grouped, Censored and Truncated Data” *Journal of the Royal Statistical Society*, B, 38: 290-295.
[12] Walter, M. A., 2012. ”Methods for Shape-Constrained Kernel Density Estimation” *Ph.D thesis. The School of Graduate and Postdoctoral Studies, University of Western, Canada*.
[13] Wu, S. F., 2008. ”Interval estimation for a pareto distribution based on a doubly type-II censored sample” *Computational Statistics and Data Analysis*, 52: 3779-3788.
[14] Wei Pan, 1999. ”Extending the Iterative Convex Minorant Algorithm to the Cox Model for Interval-Censored Data” *Journal of Computational and Graphical Statistics*, 8:1, 109-120.

Bader Aljawadi received the PhD degree in applied statistics at University Putra Malaysia (UPM). His research interests are in the areas of survival data analysis including parametric and non-parametric approaches, censoring and truncation in survival data and cure proportion in cancer trials. He has published research articles in reputed international journals of applied mathematics and statistics. He is referee of statistical journals.