Advances and Applications in Statistics

Published Online: January 2015
Available online at http://www.pphmj.com/journals/adas.htm
Volume 43, Number 2, 2014, Pages 107-118

ESTIMATION OF A CHANGE-POINT IN THE WEIBULL REGRESSION MODEL

Oscar Palmeros-Rojas[1], **José A. Villaseñor-Alva**[1] **and Francisco S. Tajonar-Sanabria**[2]

[1]Colegio de Postgraduados, Campus Montecillo
Programa de Estadística
Carretera México-Texcoco Km. 36.5, Montecillo
Texcoco, Estado de México, C.P. 56230, México
e-mail: jvillasr@colpos.mx

[2]Facultad de Ciencias Físico-Matemáticas-BUAP
Avenida San Claudio y 18 Sur
Col. San Manuel, Puebla, Puebla, C.P. 72570
México

Abstract

In some lifetime studies abrupt changes in the hazard function are observed due to overhauls, major operations or specific maintenance activities. In such situations it is of interest to detect the location in time where such changes occur. The Weibull regression hazard model with a single change-point is proposed, which is a generalization of the well-known exponential model. Statistical inference is focused on the problem concerning the change-point, taking into account covariates and random censorship. A likelihood-based approach is used for obtaining estimators for a change-point as well as for the hazard and regression parameters for the proposed model. Then the main goal is

Received: August 29, 2014; Accepted: October 14, 2014

2010 Mathematics Subject Classification: 62N02.

Keywords and phrases: empirical processes, Glivenko-Cantelli and Donsker classes, hazard function, maximum likelihood estimators, consistency.

to prove the consistency of these estimators by relying on modern empirical process theory.

1. Introduction

The literature about change-point problems is quite rich due to applications in diverse fields such as industry, reliability studies, analyzing biomedical data, medical follow-up studies, etc. See Chen and Gupta [8]. In some lifetime studies, abrupt changes in the hazard function are observed due to overhauls, major operations or specific maintenance activities. In such situations it is important to detect the location in time where such changes occur.

Some forms of the hazard function have been studied, the oldest and most popular is the exponential model. This model was first considered by Miller [12]; however, he did not consider inference on the change-point.

Dupuy [9] developed one of the first change-point exponential models that incorporates covariables in the hazard function. He considered the piecewise constant hazard function proposed by Wu et al. [3]. Dupuy studied the hazard model

$$\lambda_0(t; Z) = (\alpha + \theta 1_{\{t>\tau\}})\exp\{(\beta + \gamma 1_{\{t>\tau\}})' Z\}.$$

Using a maximum likelihood approach he proposed estimators of a change-point, hazard and regression parameters.

Here a more general model is introduced, which considers the Weibull hazard function as a baseline in the Cox hazard model with a change-point, random right hand censorship and covariates **Z**, are included. It is assumed that **Z** is a vector of baseline covariates whose effects do not change over time; however, this assumption can be relaxed to using time-dependent covariates of the form $\mathbf{Z}(t)$. Covariates depending on time are observed in some epidemiological applications.

In this paper, we consider the following combined parametric model for the hazard function with a change-point given by

$$h(t \mid \mathbf{Z}) = \begin{cases} \lambda\zeta(\lambda t)^{\zeta-1} \exp\{\boldsymbol{\beta}'\mathbf{Z}(t)\}; & t \leq \tau; \\ \\ \eta\zeta(\eta t)^{\zeta-1} \exp\{(\boldsymbol{\beta} + \boldsymbol{\gamma})'\mathbf{Z}(t)\}; & t > \tau; \end{cases} \tag{1}$$

where τ is the unknown change-point in time, $\lambda > 0$, $\eta > 0$, $\zeta > 0$ and $\boldsymbol{\beta} \in \mathbb{R}^q$, $\boldsymbol{\gamma} \in \mathbb{R}^q \setminus \{\mathbf{0}\}$ which are unknown regression coefficients.

Earlier work on change-point regression models has been mainly focused on the constant hazard model, Nguyen et al. [7] and Loader [4]; linear models, Horvath [10]; and change-point problems in hazard rate models without random censoring, Yao [13]. Properties of the estimators of the change-points have been investigated via simulations but no theoretical results are available, Matthews and Farewell [5], Loader [4].

In this paper, we propose to use a profile maximum likelihood approach for obtaining estimators for τ and the parameters in model (1). Then the main goal is to prove that these estimators are consistent.

This work is structured as follows: In Section 2, notation and terminology are established; in Section 3, the construction of profile maximum likelihood estimators for the parameters in model (1) is given; in Section 4, the consistency of the estimators is proved; in Section 5, some conclusions are established.

2. Preliminaries

Consider a random sample of n subjects. Let $\tilde{T}_i$ denote the underlying failure time for the ith subject, which can be potentially censored by a random censoring time C_i. For each individual, the random variable $S_i = \min(\tilde{T}_i, C_i)$ is the observation on the failure time. Let $\Delta_i = I(\tilde{T}_i \leq C_i)$ be the censorship indicator, which takes on value 1 if the failure event is observed, and 0 in any other case. Let $\mathbf{Z}_i(t)$ be a q-dimensional left-continuous covariate process with right hand limits (caglad), associated with the occurrence event of interest. Assume $\tilde{T}_i$ and C_i are independent, but both depend on the vector of covariables $\mathbf{Z}_i(t)$. Then the observations, consist of independent identically distributed triplets $T_i = (S_i, \Delta_i, \mathbf{Z}_i(t))$, for $i = 1, ..., n$.

According to model (1), the density of the random variable T_i is given by

$$f_\xi(t_i)$$

$$= \begin{cases} \exp\left\{-\int_0^{t_i} \lambda\zeta(\lambda t_i)^{\zeta-1} e^{\boldsymbol{\beta}'\mathbf{Z}_i(u)} du\right\}\{\lambda\zeta(\lambda t_i)^{\zeta-1} e^{\boldsymbol{\beta}'\mathbf{Z}_i(t_i)}\}; & t_i \le \tau; \\ \exp\left[-\int_0^{\tau} \lambda\zeta(\lambda\tau)^{\zeta-1} e^{\boldsymbol{\beta}'\mathbf{Z}_i(u)} du - \int_\tau^{t_i} \eta\zeta(\eta t_i)^{\zeta-1} e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})'\mathbf{Z}_i(u)} du\right] & \\ \quad \times \{\lambda\zeta(\eta u)^{\zeta-1} e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})'\mathbf{Z}_i(t_i)}\}; & t_i > \tau; \end{cases} \tag{2}$$

where $\boldsymbol{\xi} = (\tau, \varphi')$, $\varphi = (\lambda, \eta, \zeta, \boldsymbol{\beta}, \boldsymbol{\gamma})'$.

3. Estimation Procedure

Assume that the parameters of model (1) belong respectively to bounded subsets of $\mathbb{R}$, $\lambda \in \mathcal{A} \subset \mathbb{R}^+$, $\eta \in \mathcal{B} \subset \mathbb{R}^+$, $\zeta \in \mathcal{C} \subset \mathbb{R}^+$. For the regression parameters we have $\boldsymbol{\beta} \in \mathcal{D} \subset \mathbb{R}^q$ and $\boldsymbol{\gamma} \in \mathcal{E} \subset \mathbb{R}^q \backslash \{\mathbf{0}\}$. Suppose that $\tau \in \mathcal{I}$, where $\mathcal{I}$ can be either a known interval $[\tau_0, \tau_1]$, the union of known intervals $\bigcup_{i=1}^k [a_i, b_i]$ or a set of time points $\{\tau_0, ..., \tau_k\}$. Suppose that a real valued vector of parameters $\xi_0 = (\tau_0, \varphi'_{\mathbf{0}})'$ is contained in $\Phi = \mathcal{I} \times \mathcal{A} \times \mathcal{B} \times \mathcal{C} \times \mathcal{D} \times \mathcal{E}$, such that $\lambda_0 \ne \eta_0$ and $\boldsymbol{\gamma}_0 \ne \mathbf{0}$; so that a change actually occurs. Finally suppose that the covariate process $\mathbf{Z}(t)$ is uniformly bounded in variation (this assumption ensures that the covariate function has not too many fluctuations), with positive definite covariance matrix $\Sigma(\mathbf{Z}(t))$. Let $F_0(t) \equiv F_{\xi_0}(t)$, which denotes the distribution function and $\mu_0 \equiv E_{\xi_0}(t)$, the expectation of T_i in model (1), both under the true parameter values.

The log-likelihood function for the unknown parameters $\boldsymbol{\xi} = (\tau, \lambda, \eta, \zeta, \boldsymbol{\beta}', \boldsymbol{\gamma}')'$, based on the observed data $T_i = (s_i, \Delta_i, \mathbf{Z}_i(t))$, $i = 1, .., n$ is

$$\mathcal{L}_n(\boldsymbol{\xi})$$

$$= \begin{cases} \sum_{t_i \leq \tau} \Big\{ N_i(\tau)[\ln(\lambda\zeta) + (\zeta - 1)\ln(\lambda t_i) + \beta' \mathbf{Z}_i(t)] \\ -\int_0^t \lambda\zeta(\lambda u_i)^{\zeta-1} e^{\boldsymbol{\beta}'\mathbf{Z}_i(u)} du \Big\}; & t_i \leq \tau; \\ \sum_{t_i > \tau} \Big\{ [N_i(\infty) - N_i(\tau)][\ln(\eta\zeta) + (\zeta - 1)\ln(\eta t_i) + (\boldsymbol{\beta} + \boldsymbol{\gamma})' \mathbf{Z}_i(t)] \\ -\int_0^\tau \lambda\zeta(\lambda u)^{\zeta-1} e^{\boldsymbol{\beta}'\mathbf{Z}_i(u)} du - \int_\tau^t \eta\zeta(\eta u_i)^{\zeta-1} e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})'\mathbf{Z}_i(u)} du \Big\}; & t_i > \tau; \end{cases} \tag{3}$$

where $N_i(t) = 1_{\{T_i \leq t, \Delta_i = 1\}}$ is a counting process for the failure of individual i.

To obtain the maximum likelihood estimator of the vector of parameters $\boldsymbol{\xi_0}$, the following procedure is used. First, let $\tau \in \mathcal{I}$ be a fixed value. Then the estimator $\hat{\varphi}$ of $\varphi_{\mathbf{0}}$ is taken as the value that maximizes the log-likelihood function $\mathcal{L}_n(\boldsymbol{\xi})$ on $\mathcal{A} \times \mathcal{B} \times \mathcal{C} \times \mathcal{D} \times \mathcal{E}$. Hence, an estimator for τ_0 is

$$\hat{\tau} = \sup\{\tau \in \mathcal{I} \,|\, \max(\mathcal{L}_n(\tau, \hat{\varphi}(\tau)), \mathcal{L}_n(\tau\pm, \hat{\varphi}(\tau \pm))) = \sup_{\tau \in \mathcal{I}} \mathcal{L}_n(\tau, \hat{\varphi}(\tau))\},$$

where $\mathcal{L}_n(\tau\pm, \hat{\varphi}(\tau \pm))$ is the left or right limit of $\mathcal{L}_n$ in τ. Finally, the maximum likelihood estimator for $\boldsymbol{\varphi}$ is obtained by taking $\hat{\varphi} = \hat{\varphi}(\hat{\tau})$.

Without loss of generality we will assume that the covariates $\mathbf{Z}(t)$ do not change over time. The proof when $\mathbf{Z}(t)$ is a vector of time-varying covariates is quite similar. For a given τ, we estimate λ, η, ζ, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ by considering the corresponding score functions:

$$\frac{\partial \mathcal{L}_n(\xi)}{\partial \lambda} = \begin{cases} \sum_{t_i \le \tau} \Delta_i \left[\frac{\zeta}{\lambda}\right] - (\lambda t_i)^{\zeta - 1} t_i e^{\boldsymbol{\beta}'\mathbf{Z}_i}; & t_i \le \tau; \\ \sum_{t_i > \tau} -\zeta(\lambda\tau)^{\zeta} \tau e^{\boldsymbol{\beta}'\mathbf{Z}_i}; & t_i > \tau; \end{cases}$$

$$\frac{\partial \mathcal{L}_n(\xi)}{\partial \zeta} = \begin{cases} \sum_{t_i \le \tau} \Delta_i \left[\frac{1 + \zeta \ln(\lambda t_i)}{\zeta}\right] - (\lambda t_i)^{\zeta} \ln(\lambda t_i) e^{\boldsymbol{\beta}'\mathbf{Z}_i}; & t_i \le \tau; \\ \sum_{t_i > \tau} \Delta_i \left[\frac{1 + \zeta \ln(\eta t_i)}{\zeta}\right] - (\lambda\tau)^{\zeta} \ln(\lambda\tau) e^{\boldsymbol{\beta}'\mathbf{Z}_i} & \\ -e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})'\mathbf{Z}_i}[(\eta t_i)^{\zeta} \ln(\eta t_i) - (\eta\tau)^{\zeta} \ln(\eta\tau)]; & t_i > \tau; \end{cases}$$

$$\frac{\partial \mathcal{L}_n(\xi)}{\partial \eta} = \left\{ \sum_{t_i > \tau} \Delta_i [\zeta/\eta] - e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})'\mathbf{Z}_i}[\zeta(\eta t_i)^{\zeta-1} t_i - \zeta(\eta\tau)^{\zeta-1}\tau]; \; t_i > \tau; \right.$$

$$\frac{\partial \mathcal{L}_n(\xi)}{\partial \gamma} = \left\{ \sum_{t_i > \tau} \Delta_i Z_i - Z_i e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})'\mathbf{Z}_i}[(\eta t_i)^{\zeta} - (\eta\tau)^{\zeta}]; \; t_i > \tau; \right.$$

$$\frac{\partial \mathcal{L}_n(\xi)}{\partial \beta} = \begin{cases} \sum_{t_i \le \tau} \Delta_i Z_i - Z_i (\lambda t_i)^{\zeta} e^{\boldsymbol{\beta}'\mathbf{Z}_i}; & t_i \le \tau; \\ \sum_{t_i > \tau} \Delta_i Z_i - Z_i (\lambda\tau)^{\zeta} e^{\boldsymbol{\beta}'\mathbf{Z}_i} & \\ \qquad -Z_i e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})'\mathbf{Z}_i}[(\eta t_i)^{\zeta} - (\eta\tau)^{\zeta}]; & t_i > \tau. \end{cases}$$

Notice that for each $\tau \in \mathcal{I}$ given, the estimates of parameters λ, η, ζ, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are obtained by solving the score equations; however, due to the non-linearity of the score functions, in order to obtain their respective solutions it is necessary to use numerical algorithms along with an iterative method. This numerical solution can be implemented with the available software at present time.

4. Consistency of the Estimators

Empirical processes techniques will be used in order to prove that the profile ML-estimators converge in probability to the parameter values.

Let us define the succession $Y_n(\xi) = n^{-1}(L_n(\xi) - L_n(\xi_{\mathbf{0}}))$, and the function $\mu_n(\xi) = \mathbb{E}_0[Y_n(\xi)]$, where ξ and $\xi_{\mathbf{0}}$ are defined as before. The following result proves that the class of functions defined by (2) is Glivenko-Cantelli (definition and properties of this class of functions can be seen in van der Vaart and Wellner [2] and van der Vaart [1]), which is equivalent to show that Y_n converges uniformly to $\mu_n(\xi)$.

Lemma 1. *The class of functions* $\mathcal{F} = \{f \mid f : \Phi \to R\}$, *defined by* (2), *is Glivenko-Cantelli*.

Proof. Since every Donsker class is Glivenko-Cantelli (see [2] and [1]), it will be proved that the family of functions $\mathcal{F}$ is Donsker. Since the Donsker property is preserved under addition and product, it is enough to prove that the family of functions $\mathcal{G} = \{g \mid g : \Phi \to R\}$ is Donsker, where $g = 1_{\{T>\tau\}} e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})' \mathbf{Z}} T^{\zeta}$.

We know that the family of functions $\mathcal{H} = \{1_{\{T>\tau\}} e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})' \mathbf{Z}} T\}$ is Donsker, where $\boldsymbol{\beta}, \boldsymbol{\gamma}$ and $\mathbf{Z} \in \mathbb{R}^q$, $\mathbf{Z}$ and T are bounded (see Dupuy [9]).

Let $g : \mathbf{\Phi} \to \mathbb{R}^+$ be the function defined by $g(T) = T^{\zeta - 1}$. Then from example 2.10.10 pp. 192 in van der Vaart and Wellner [2] we have that class $\mathcal{H} \cdot g$ is Donsker, that is, $\mathcal{J} = \{1_{\{T>\tau\}} e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})' \mathbf{Z}} T^{\zeta} \mid \zeta > 0\}$ is Donsker. Given that the addition of Donsker classes is Donsker, then the family of functions $\mathcal{F}$ defined in (2) is Donsker and so $\mathcal{F}$ is also Glivenko-Cantelli. □

In order to prove this result when covariates depend on time, which are denoted by $\mathbf{Z}(t)$, it is assumed that $\mathbf{Z}(t)$ is a caglad and uniformly bounded in variation process, the result of Parner [6] follows since, if W is a caglad uniformly bounded in variation process on $[0, 1]$, then the class $\{\chi_u \mid u \in [0, 1]\}$ of projections $\chi_u : W \to \chi_u(W) = W(u)$ is Donsker and by the continuous mapping theorem, we can show that the class (2) converges weakly in the set of bounded real functions on Φ.

The following lemma asserts that the real valued vector of parameters for model (2) is identifiable.

Lemma 2. *Let* $f_{\xi}(t)$ *be the probability density function defined in equation* (2). *If* $f_{\xi}(t) = f_{\xi_{\mathbf{0}}}(t)$ *almost anywhere under* F_0, *then* $\xi = \xi_{\mathbf{0}}$.

Proof. The proof of this result follows by contradiction. In fact, assume that $\xi \neq \xi_{\mathbf{0}}$ and $f_{\xi}(t) = f_{\xi_{\mathbf{0}}}(t)$. Without loss of generality, assume that $\tau \neq \tau_0$; in other words, either (i) $\tau < \tau_0$ or (ii) $\tau > \tau_0$. Assume that (i) holds (the opposite is treated analogously). By hypothesis, $f_{\xi}(t) = f_{\xi_{\mathbf{0}}}(t)$ almost anywhere. In particular, if $\Delta = 0$,

$$1_{\{t \leq \tau\}} \exp[-(\lambda t)^{\zeta} e^{\boldsymbol{\beta}'\mathbf{Z}}] + 1_{\{t > \tau\}} \exp[-(\lambda t)^{\zeta} e^{\boldsymbol{\beta}'\mathbf{Z}} - e^{(\boldsymbol{\beta}+\gamma)'\mathbf{Z}}[(\eta t)^{\zeta} - (\eta\tau)^{\zeta}]]$$

$$= 1_{\{t \leq \tau_0\}} \exp[-(\lambda_0 t)^{\zeta_0} e^{\boldsymbol{\beta}_0'\mathbf{Z}}]$$

$$+ 1_{\{t > \tau_0\}} \exp[-(\lambda_0 \tau_0)^{\zeta_0} e^{\boldsymbol{\beta}_0'\mathbf{Z}} - e^{(\boldsymbol{\beta}_0+\gamma_0)'\mathbf{Z}}[(\eta_0 t)^{\zeta_0} - (\eta_0\tau_0)^{\zeta_0}]]. \qquad (4)$$

Values of t and z will be exhibited which satisfy expression (4) with $\tau \neq \tau_0$. This leads to a contradiction in one of the set conditions in Section 3.

Let z be a realization of $\mathbf{Z}$ and let Λ_z be the set of all $t \in (\tau, \tau_0]$ such that (t, z) does not satisfy expression (4).

Then for almost all observations $(t, z) \in \Lambda_{\boldsymbol{x}}$, except for a set $\mathcal{C}$ with zero measure, we get that $F_0((t, z)) = 0$, with $t \in (\tau, \tau_0]$.

Taking z_1 and z_2, two different realizations of $\mathbf{Z}$, such that neither z_1 nor z_2 are in $\mathcal{C}$ (these values exist, given that $\mathbf{Z}$ is non-degenerate), let $t_1, t_2 \in (\tau, \tau_0]$, with $t_1 \neq t_2$, such that pairs $(t_1, z_1), (t_1, z_2), (t_2, z_1)$ and (t_2, z_2) satisfy $F_0((t_i, z_i)) = 0$ for $i = 1, 2$.

By substituting (t_1, z_1) in (4), we get

$$(\lambda\tau)^{\zeta}e^{\boldsymbol{\beta}'z_1}+e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})'z_1}[(\eta t_1)^{\zeta}-(\eta\tau)^{\zeta}]=(\lambda_0 t_1)^{\zeta_0}e^{\boldsymbol{\beta}_0'z_1}. \tag{5}$$

Similarly, for a pair (t_2, z_1), we get

$$(\lambda\tau)^{\zeta}e^{\boldsymbol{\beta}'z_1}+e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})'z_1}[(\eta t_2)^{\zeta}-(\eta\tau)^{\zeta}]=(\lambda_0 t_2)^{\zeta_0}e^{\boldsymbol{\beta}_0'z_1}. \tag{6}$$

Subtracting (5) from (6), we get

$$\eta^{\zeta}e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})'z_1}(t_2^{\zeta}-t_1^{\zeta})=\lambda_0^{\zeta_0}e^{\boldsymbol{\beta}_0'z_1}(t_2^{\zeta_0}-t_1^{\zeta_0}). \tag{7}$$

Proceeding analogously for (t_1, z_2) and (t_2, z_2), we get,

$$\eta^{\zeta}e^{(\boldsymbol{\beta}+\boldsymbol{\gamma})'z_2}(t_2^{\zeta}-t_1^{\zeta})=\lambda_0^{\zeta_0}e^{\boldsymbol{\beta}_0'z_2}(t_2^{\zeta_0}-t_1^{\zeta_0}). \tag{8}$$

By hypothesis $t_1 \neq t_2$, $\alpha+\theta>0$, $\lambda_0 \neq 0$ and $\alpha_0 \neq 0$, then the quotient of (8) and (7) is $e^{(\boldsymbol{\beta}+\boldsymbol{\gamma}-\boldsymbol{\beta}_0)(z_2-z_1)}=1$, that is,

$$(\boldsymbol{\beta}+\boldsymbol{\gamma}-\boldsymbol{\beta}_0)(z_2-z_1)=0. \tag{9}$$

Considering all pairs of orthogonal vectors to (z_2, z_1) we choose a pair (z_4, z_3) that does not belong to $\mathcal{C}$; for this pair it is true that $(\boldsymbol{\beta}+\boldsymbol{\gamma}-\boldsymbol{\beta}_0)(z_4-z_3)=0$. We follow with this process until obtaining q pairs of vectors, such that none of these pairs belongs to $\mathcal{C}$, and whose difference is linearly independent (that is, we are building a base for $\mathbb{R}^q$), resulting in $(\boldsymbol{\beta}+\boldsymbol{\gamma}-\boldsymbol{\beta}_0)=0$, that is, $\boldsymbol{\beta}+\boldsymbol{\gamma}=\boldsymbol{\beta}_0$.

By substituting the last expression in (8), we get $\eta^{\zeta}(t_2^{\zeta}-t_1^{\zeta})=\lambda_0^{\zeta_0}(t_2^{\zeta_0}-t_1^{\zeta_0})$, that is,

$$\eta^{\zeta}t_2^{\zeta}-\lambda_0^{\zeta_0}t_2^{\zeta_0}=\eta^{\zeta}t_1^{\zeta}-\lambda_0^{\zeta_0}t_1^{\zeta_0}. \tag{10}$$

Assume that $\zeta \neq \zeta_0$, this is, (i) $\zeta_0 < \zeta$ or (ii) $\zeta_0 > \zeta$. Without loss of generality, assume that (i) is satisfied, from equation (10), we get

$$\left[\frac{t_2}{t_1}\right]^{\zeta_0}[\eta^{\zeta}t_2^{\zeta-\zeta_0} - \lambda_0^{\zeta_0}] = \eta^{\zeta}t_1^{\zeta-\zeta_0} - \lambda_0^{\zeta_0}. \tag{11}$$

Since $t_1 \neq t_2$, it could be that either (a) $t_1 < t_2$ or (b) $t_1 > t_2$. Assume that (a) is true, then we get that $1^{\zeta_0} < \left[\frac{t_2}{t_1}\right]^{\zeta_0}$.

Given that $\zeta - \zeta_0 > 0$, then substituting the last expression on the right hand side of (11) we obtain

$$\eta^{\zeta}t_2^{\zeta-\zeta_0} - \lambda_0^{\zeta_0} \leq \eta^{\zeta}t_1^{\zeta-\zeta_0} - \lambda_0^{\zeta_0} \Leftrightarrow t_2 \leq t_1.$$

However, this last expression is a contradiction given that (a) was true.

Analogously for (b), when substituting $1^{\zeta_0} > \left[\frac{t_2}{t_1}\right]^{\zeta_0}$ in (11), we have a contradiction. Therefore, it is not possible to have $\zeta_0 < \zeta$. To prove that $\zeta < \zeta_0$ is not possible, an analogous procedure is followed. Hence, we conclude that $\zeta = \zeta_0$. Substituting the last expression in (10), we see that

$$\eta^{\zeta} = \lambda_0^{\zeta}. \tag{12}$$

From $\boldsymbol{\beta} + \boldsymbol{\gamma} = \boldsymbol{\beta_0}$ and substituting (12) in (4), for $(t_1, z_{\mathbf{1}})$ we get

$$(\lambda\tau)^{\zeta}e^{\boldsymbol{\beta}\mathbf{Z_1}} = (\eta\tau)^{\zeta}e^{\boldsymbol{\beta}_0' z_1}. \tag{13}$$

Similarly, for $(t_1, z_{\mathbf{2}})$ we get

$$(\lambda\tau)^{\zeta}e^{\boldsymbol{\beta}\mathbf{Z_1}} = (\eta\tau)^{\zeta}e^{\boldsymbol{\beta}_0' z_1}.$$

Dividing by (13) we get

$$(\boldsymbol{\beta} - \boldsymbol{\beta_0})'(z_{\mathbf{2}} - z_1) = 0. \tag{14}$$

By the same reasoning as above, $\boldsymbol{\beta} - \boldsymbol{\beta_0} = 0$, that is, $\boldsymbol{\beta} = \boldsymbol{\beta_0}$, hence $\boldsymbol{\gamma} = \mathbf{0}$. This is a contradiction, since by hypothesis $\boldsymbol{\gamma} \neq \mathbf{0}$. This means that $\tau < \tau_0$

does not hold. Similarly, it is proved that $\tau > \tau_0$ cannot be true. Hence, $\tau = \tau_0$. With an analogous procedure, it is proved that $\varphi = \varphi_{\mathbf{0}}$. □

Theorem 3. *Under conditions given in Section* 3, *estimators* $\hat{\tau}$ *and* $\hat{\varphi}$ *for model* (1) *converge in probability to* τ_0 *and* $\varphi_{\mathbf{0}}$, *respectively*.

Proof. Remember that the estimators were obtained through profile maximum likelihood, and thus certain properties of these estimators are used.

Note that $\mu_n(\xi) = n^{-1}\int_\Phi \ln\left[\frac{L(\xi)}{L(\xi_{\mathbf{0}})}\right] f_\xi(Y)dy,$ which is minus the Kullback-Leibler divergence of f_ξ and $f_{\xi_{\mathbf{0}}}$, where $L(\xi)$ is the likelihood function. Since $\mu_n(\xi_{\mathbf{0}}) = 0$, $\xi_{\mathbf{0}}$ is a maximum of μ_n. From Lemma 2, we get that $\xi_{\mathbf{0}}$ is the unique maximum of μ_n. From Lemma 1, Y_n converges uniformly to μ_n, that is, $\hat{\xi}$ converges in probability to $\xi_{\mathbf{0}}$. From this we conclude that $\hat{\tau}$ and $\hat{\varphi}$ converge in probability to τ_0 and $\varphi_{\mathbf{0}}$, respectively. □

5. Conclusions

A more general model than those given by Wu et al. [3], Dupuy [9] and Liu et al. [11] is introduced. Following Dupuy, in this paper it is shown that the profile maximum likelihood estimation method can be extended to estimating the change-point and the hazard and regression parameters in the Weibull regression model. It is also proved that these estimators are consistent.

References

[1] A. W. van der Vaart, Asymptotic Statistics, Cambridge University Press, New York, 1998.

[2] A. W. van der Vaart and J. A. Wellner, Weak Convergence and Empirical Processes, Cambridge University Press Springer, New York, 1996.

[3] C. Q. Wu, L. C. Zhao and Y. H. Wu, Estimation in change-point hazard function models, Statist. Probab. Lett. 63 (2003), 41-48.

[4] C. R. Loader, Inference for a hazard rate change point, Biometrika 78 (1991), 749-757.

[5] D. E. Matthews and V. T. Farewell, On testing for a constant hazard against a change point alternative, Biometrics 38 (1982), 463-468.

[6] E. Parner, Asymptotic theory for the correlated gamma-frailty model, Annals of the Institute of Statistical Mathematics 26 (1998), 183-214.

[7] H. T. Nguyen, G. S. Rogers and E. A. Walker, Estimation in change-point hazard rate models, Biometrika 71 (1984), 299-304.

[8] J. Chen and A. K. Gupta, Parametric Statistical Change Point Analysis with Applications to Genetics, Medicine, and Finance, Birkhauser, Springer, New York, 2012.

[9] J. F. Dupuy, Estimation in a change-point hazard regression model, Statist. Probab. Lett. 76 (2006), 182-190.

[10] L. Horvath, Detecting changes in linear regressions, Statistics 26 (1995), 189-208.

[11] M. Liu, W. Lu and Y. Shao, A Monte Carlo approach for change-point detection in the Cox proportional hazards model, Stat. Med. 27 (2008), 3894-3909.

[12] R. G. Miller, Early failures in life testing, J. Amer. Statist. Assoc. 55 (1960), 491-502.

[13] Y. C. Yao, Maximum likelihood estimation in hazard rate models with a change point, Comm. Statist. A – Theory Methods 15 (1986), 2455-2466.