Quaderni di Statistica
Vol. 12, 2010

Illness-death model for clustered data

Marialuisa Restaino
Dipartimento di Scienze Economiche e Statistiche, Università di Salerno
E-mail: mlrestaino@unisa.it

Summary: The aim of this paper is to develop the three-State Illness-Death model taking into account the clustered nature of data. For this purpose, the Frailty Model has been applied in Multi-State framework, assuming that the underlying process is time-homogeneous Markovian. The unknown parameters of the model have been estimated both in presence of complete observations and in case of censoring. Thus, the impact and the effects of censoring on the estimation have been analyzed. The relevance of the proposed model is illustrated by means of a simulation study, in order to validate the model assumptions and the performance of the proposed estimators.

Keywords: Illness-Death Model, Clustered Data, Censoring, Time-Homogeneous Markov hypothesis.

1. Introduction

Traditionally, statistical methods for survival data have involved estimation of survival probabilities, considering the transition from the initial state to a single ultimate state or endpoint. However, more than one endpoint can be defined in empirical settings, such as in epidemiology and medicine, in social and political sciences, in economics and finance and in actuarial science. In such cases, *Multi-State Models* (Hougaard, 2001), considered as a generalization of survival analysis, are applied to model the subject's evolvement among a finite number of states over time.

The assumptions frequently made in these frameworks are that the life histories of the individuals under study are all statistically independent and all individuals sampled into a study are subject to the same risk.

But there are situations in which the data are observed in clusters, and observations from the same cluster usually share some unobserved characteristics and tend to be correlated.

Although this association should be taken into account in the analysis of survival times, it is often ignored, leading to inefficient and biased estimates. The starting point of a model that takes into account the correlation within cluster is a *proportional hazards model*, representing the effects of observed covariates on survival time, adding in a cluster-specific random effect that represents unobserved influences common to all members of a cluster. The resulting model is the so-called *Frailty model* (Andersen *et al.*, 1993; Clayton, 1978; Clayton and Cuzick, 1985; Hougaard, 1984, 1986; Vaupel *et al.*, 1979).

Moreover, in many studies, the experimental unit is alive when the data are evaluated, but his complete lifetime is not known at that time. Therefore, it is only known that his lifetime exceeds his age at the time of evaluation. This leads to have partial observations, i.e. censored observations. Since the presence of censored data is a technical problem, the censoring time cannot be ignored.

The aim of the paper is to develop a new model, which would apply the Frailty model on Multi-State framework. In particular, a three-state *Frailty* Illness-Death Model has been proposed and the impact and the effects of censoring on the parameters' estimates have been investigated. For deriving the transition probability matrix, the hypothesis underlying the process is assumed to be Time-Homogeneous Markovian, i.e. transition intensities are assumed to be constant over time.

The unknown parameters have been estimated both for the *full model*, i.e. in presence of complete observations, and for the *partial model*, i.e. when there are censored observations. In case of censoring, two different hypotheses have been made. The first one assumes that a censored observation will directly reach the absorbing state at the end of study. Instead, in the second one, a censored observation is assumed to reach the absorbing state, at the end of the study period, going through the different states.

The paper is organized as follows. The three-State Frailty Illness-Death Model has been described in Section 2. Then the maximum like-

lihood estimators for the model with complete observations and for the model with censored observations have been derived in Section 3 and in Section 4, respectively. The results of a simulation study are presented and discusses in Section 5. The final section will give some concluding remarks.

2. Frailty Illness-Death Model

Let $\{\widetilde{X}^{vm}(t), t \in [0,\tau], \tau < \infty, v = 1,\ldots,n_m, m = 1,\ldots,M\}$ be a stochastic process denoting the state of a subject v in cluster m at time t, where M is the number of clusters ($m = 1,\ldots,M$). Each cluster has a number of individuals equal to n_m, so that $\sum_{m=1}^{M} n_m = n$.

Assume that the process $\widetilde{X}(\cdot)$ takes integer values in a finite state space $\mathcal{L} = (1,\ldots,L)$, where the state 1 is the healthy state and the state L is the absorbing state. In our case, $L = 3$. Assume that each subject entries into study at state 1, at time 0, i.e. $P[\widetilde{X}^{vm}(0) = 1] = 1$.
Let

- $\mathbf{Z}^{vm}(t)$ be a vector of p covariates for individual v in cluster m. It is assumed that they are not a random variable and do not depend on time, i.e. $\mathbf{Z}^{vm}$;
- C^{vm} be a right-censoring time for individual v in cluster m. Assume that the censoring is not informative and is not a random variable;
- $\widetilde{X}^{vm}(C^{vm})$ is the state occupying by a subject v in cluster m at censoring time.

Assume that $\widetilde{X}^{vm}(t)$ and C^{vm} are independent ($m = 1,\ldots,M$, $v = 1,\ldots,n_m$), conditional on covariate $\mathbf{Z}^{vm}$. Thus, the observed process is given by $X^{vm}(t) = \min[\widetilde{X}^{vm}(t), \widetilde{X}^{vm}(C^{vm})]$ and the observed data for individual v in cluster m are $\{X^{vm}(t), \delta^{vm}(t), t \in [0,\tau]\}$, where $\delta^{vm}(t) = I[C^{vm} \geq t]$ indicates whether the process is observed at time t, for $m = 1,\ldots,M$, $v = 1,\ldots,n_m$.
Suppose that for each individual v, u discrete times are observed, $t_0,\ldots,t_u$,

with $0 \leq t_0 < t_1 < \cdots < t_{u-1} < \infty$. Let $N_{ij}(t_g)$ be the number of subjects being in state i at t_{g-1} and in j at time t_g, for $g = 1, \ldots, u$.
Consider a model in which the transition intensity hazard function partly depends on an unobservable random variable, which acts multiplicatively on the hazard, so that large values of the variable increase the hazard. This random variable is the *frailty term* γ_m, assumed to be i.i.d. with density function $f(\cdot; \phi)$.

Conditionally on the frailty term γ_m, the process $X^{vm}(\cdot)$ is assumed to be independent. The censoring time is not informative on the frailties γ_m.

Moreover, assume that the process $X(\cdot)$ is Time-Homogeneous Markovian, i.e. the transition intensities do not depend on time, but they are constant over time. Therefore, the transition intensity from state i to state j for individual v in cluster m is given by:

$$\lambda_{ij}^{vm} = \lambda_{ij,0} \exp\{\boldsymbol{\beta}_{\boldsymbol{ij}}^{\boldsymbol{s}\,T} \mathbf{Z}^{vm} + \gamma_m\} \tag{1}$$

for $i, j = 1, \ldots, L$, $s = 1, \ldots, p$ and for $i \neq j$, where $\lambda_{ij,0}$ is the *baseline hazard intensity* of transition from state i to state j, $\mathbf{Z}^{vm}$ is a p-dimensional vector of covariates of individual v in cluster m, γ_m is the *frailty term* for cluster m, which takes into account the within-cluster correlation and

$$\boldsymbol{\beta}_{\boldsymbol{ij}}^{\boldsymbol{s}} = (\beta_{ij}^1, \beta_{ij}^2, \ldots, \beta_{ij}^p)^T.$$

Thus, let $\boldsymbol{\lambda}^{vm}$ be the $(L \times L)$ transition intensity matrix for individual v in cluster m, with entries equal to (1), for $i, j = 1, \ldots, L$, and for $i \neq j$, and entries equal to $\lambda_{ii}^{vm} = -\sum_{i \neq j} \lambda_{ij}^{vm}$ for $i = 1, \ldots, L$, that is $\boldsymbol{\lambda}^{vm} = \{\lambda_{ij}\}^{vm}$.

For $0 \leq s \leq t$, let $\mathbf{P}^{vm}(s,t)$ be the $(L \times L)$ transition probability matrix for individual v in cluster m, with entries:

$$p_{ij}^{vm}(s,t) = P[X^{vm}(t) = j | X^{vm}(s) = i] \tag{2}$$

$$p_{ii}^{vm}(s,t) = 1 - \sum_{i \neq j} p_{ij}^{vm}(s,t), \tag{3}$$

for $i, j = 1, \ldots, L$, with $L = 3$.

According to the Time-Homogeneous Markov assumption, each tran-

sition probability $p_{ij}^{vm}(s,t)$ depends only on $t-s$, that is:

$$p_{ij}^{vm}(s,t) = p_{ij}^{vm}(0,t-s) = p_{ij}^{vm}(t-s), \tag{4}$$

for $i,j = 1,\ldots,L$, with $L = 3$.

Then, the transition probabilities can be simply expressed in terms of the transition intensities through the Kolmogorov equation (Cox and Miller, 1965; Kalbfleisch and Lawless, 1985).

By the light of data generating process, presented previously, the maximum likelihood parameters estimators are derived, both in presence of complete observations (Section 3) and in presence of censored observations (Section 4). In the second case, two different hypotheses on censored observations are made. The first one (*H1*) assumes that an observation censored at a given time t in a given state i, will directly reach the absorbing state, at the end of the study. The second hypothesis (*H2*) is that a censored observation would reach the absorbing state, at the end of the study period, going through the different states.

3. Parameters estimation in presence of complete observations

Firstly, consider the case in which there are not censored observations. Let $p_{ij}^{'v}$ be the observed ([1]) transition probabilities of individual v, for $v = 1,\ldots,n$ and $i,j = 1,2,3$.

The likelihood is given by:

$$L(\boldsymbol{\xi}) = \prod_{v=1}^{n}\prod_{g=1}^{u}\prod_{i,j=1}^{L}\{p_{ij}^{'v}[(t_g - t_{g-1})|\boldsymbol{\xi}]\}^{N_{ij}(t_g)}. \tag{5}$$

where $\boldsymbol{\xi} = (\boldsymbol{\lambda}_0, \boldsymbol{\beta})$ is the vector of the unknown parameters, with $\boldsymbol{\lambda}_0 = (\lambda_{12,0}, \lambda_{13,0}, \lambda_{23,0})$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}_{12}, \boldsymbol{\beta}_{13}, \boldsymbol{\beta}_{23})$.

The first derivatives of the log-likelihood with respect to vector of

[1] The term observed means that the frailty term γ_m is assumed to be observed and is not to be estimated.

unknown parameters $\boldsymbol{\xi}$ are given by:

$$S(\boldsymbol{\xi}) = \frac{\partial \log L(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = \sum_{v=1}^{n} \sum_{g=1}^{u} \sum_{i,j=1}^{L} N_{ij}(t_g) \frac{\partial p_{ij}^{'v}[(t_g - t_{g-1})|\boldsymbol{\xi}]}{\partial \boldsymbol{\xi}} \frac{1}{p_{ij}^{'v}[(t_g - t_{g-1})|\boldsymbol{\xi}]}. \quad (6)$$

Following the paper of Kalbfleisch and Lawless (1985), the first derivatives of $\mathbf{P}^{'v}(s,t)$ according to the vector of parameters $\boldsymbol{\xi}$ are given by:

$$\frac{\partial}{\partial \boldsymbol{\xi}} \mathbf{P}^{'v}(s,t) = \mathbf{A}^{'v} \mathbf{V}^{'v} (\mathbf{A}^{'v})^{-1}, \quad (7)$$

with $\mathbf{A}^{'v}$ the $(L \times L)$ observed matrix (whose j-*th* column is the eigenvector associated with $d_j^{'}$ eigenvalue $(j = 1, \ldots, L)$), where:

$$\boldsymbol{V'^{v}} = \begin{pmatrix} g_{11} t \exp\{d_1^{'} t\} & g_{12} \frac{(\exp\{d_1^{'} t\} - \exp\{d_2^{'} t\})}{d_1^{'} - d_2^{'}} & g_{13} \frac{(\exp\{d_1^{'} t\} - \exp\{d_3^{'} t\})}{d_1^{'} - d_3^{'}} \\ 0 & g_{22} t \exp\{d_2^{'} t\} & g_{23} \frac{(\exp\{d_2^{'} t\} - \exp\{d_3^{'} t\})}{d_2^{'} - d_3^{'}} \\ 0 & 0 & g_{33} t \exp\{d_3^{'} t\} \end{pmatrix}$$

and with g_{ij} the (i,j) entry of matrix for each individual, defined as:

$$\mathbf{G}^{'v} = (\mathbf{A}^{'v})^{-1} \frac{\partial \boldsymbol{\lambda}^{'v}}{\partial \boldsymbol{\xi}} \mathbf{A}^{'v}, \quad (8)$$

and $\boldsymbol{\lambda}^{'v}$ is the $(L \times L)$ observed transition intensity matrix for individual v.

4. Parameters estimation in presence of censored observations

In this section, the observed likelihood when observations are partially known, i.e. censored, is given. If one observation is censored, this means that an individual gets out of the study before the end of study period, and is treated as missing, because the exact time of leaving is unknown.

Since the interest is on the transition probability $p_{ij}^{'v}$, it is analyzed how these probabilities change in presence of censoring, by using the results available in Survival Analysis literature (Collett, 2003).
Recalling that C^v is the censoring time for individual v ($v = 1, \ldots, n$) and $\delta^v(t) = I[C^v \geq t]$ indicates whether the process is observed at time t, for individual v, the two different hypotheses (*H1* and *H2*) introduced in (2) are considered.

Firstly assume that an observation is censored at time t_g ($g = 1, \ldots, u$) in state i ($i = 1, 2, 3$). In other words, the subject v is observed to be in state i at time t_{g-1} and gets out of the study at time t_g.

Assume that each censored observation in state i will directly reach the absorbing state, i.e. state 3, at the end of the study T (*H1*).
Let $N_i(t_g)$ be the number of subjects who are still alive in state i at time t_{g-1} and censored at time t_g.

The likelihood is given by:

$$L(\boldsymbol{\xi}) = \prod_{v=1}^{n} \prod_{i,j} \prod_{g=1}^{r} \{p_{ij}^{'v}[(t_g - t_{g-1})|\boldsymbol{\xi}]\}^{N_{ij}(t_g - t_{g-1})} \{p_{i3}^{'v}[(T - t_{g-1})|\boldsymbol{\xi}]\}^{N_i(t_g)}. \tag{9}$$

where $\boldsymbol{\xi} = (\boldsymbol{\lambda}_0, \boldsymbol{\beta})$, $\boldsymbol{\lambda}_0 = (\lambda_{12,0}, \lambda_{13,0}, \lambda_{23,0})$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}_{12}, \boldsymbol{\beta}_{13}, \boldsymbol{\beta}_{23})$.

The first derivatives of log-likelihood according to the vector of parameters are given by:

$$\begin{aligned} S(\boldsymbol{\xi}) &= \frac{\partial \log L(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \\ &= \sum_{v=1}^{n} \sum_{i,j} \sum_{g=1}^{r} \left\{ N_{ij}(t_g - t_{g-1}) \frac{\partial \{p_{ij}^{'v}[(t_g - t_{g-1})|\boldsymbol{\xi}]\}}{\partial \boldsymbol{\xi}} \right. \\ &\frac{1}{\{p_{ij}^{'v}[(t_g - t_{g-1})|\boldsymbol{\xi}]\}} + \\ &\left. + N_i(t_g) \frac{\partial \{p_{i3}^{'v}[(T - t_{g-1})|\boldsymbol{\xi}]\}}{\partial \boldsymbol{\xi}} \frac{1}{\{p_{i3}^{'v}[(T - t_{g-1}|\boldsymbol{\xi}]\}} \right\}. \end{aligned} \tag{10}$$

and they are computed by means of (7).

Now, the likelihood is given under the second hypothesis (*H2*):

$$L(\boldsymbol{\xi}) = \prod_{v=1}^{n}\prod_{i,j}\prod_{g=1}^{r}\left\{\{p_{ij}^{'v}[(t_g - t_{g-1})|\boldsymbol{\xi}]\}^{N_{ij}(t_g - t_{g-1})}\times \right.$$
$$\left.\times\left[\sum_{k=1}^{3}\{p_{ik}^{'v}[(T - t_{g-1})|\boldsymbol{\xi}]\}w_{ik}(t_{g-1}, t_g)\right]^{N_i(t_g)}\right\}, \qquad (11)$$

where $w_{ik}(t_g)$, for $(k = 1, 2, 3)$, is a weight given by:

$$w_{ik} = \frac{N_{ik}(t_{g-1}, t_g)}{\sum_k N_{ik}(t_{g-1}, t_g)}. \qquad (12)$$

The first derivatives of log-likelihood according to the vector of parameters are given by:

$$\begin{aligned} S(\boldsymbol{\xi}) &= \frac{\partial \log L(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \\ &= \sum_{v=1}^{n}\sum_{i,j}\sum_{g=1}^{r}\left\{\left[N_{ij}(t_g - t_{g-1})\frac{\partial[p_{ij}^{'v}(t_g - t_{g-1})|\boldsymbol{\xi}]}{\partial \boldsymbol{\xi}}\right.\right. \\ &\left.\frac{1}{[p_{ij}^{'v}(t_g - t_{g-1})|\boldsymbol{\xi}]}\right] + \\ &+ N_i(t_g)\frac{\partial \sum_{k=1}^{3}[p_{ik}^{'v}(T - t_{g-1})|\boldsymbol{\xi}]w_{ik}(t_{g-1}, t_g)}{\partial \boldsymbol{\xi}} \\ &\left.\left.\frac{1}{\sum_{k=1}^{3}[p_{ik}^{'v}(T - t_{g-1})|\boldsymbol{\xi}]w_{ik}(t_{g-1}, t_g)}\right\}\right. \end{aligned} \qquad (13)$$

and again, as mentioned above, they are computed by means of (7).

5. Simulation Study

To illustrate the relevance of the proposed approach, in particular to illustrate the incidence and effect of censoring on parameters estimates, a simulation study is performed. Hence, the assumptions underlying the

Illness-Death Frailty model have been evaluated. The finite sample properties of the model proposed have been analyzed.

The Illness-Death Frailty model, as given by (1), with three states and without recovery to healthy state, has been considered. All individuals are in state 1 (*state health*) at time 0. Individuals in state 1 may either develop the illness (*state 2*), or die (*state 3*). Individuals with illness remain in state 2 until death (entering state 3). The hypothesis underlying the model is Time-Homogeneous Markovian.

Simulation has been performed for the model in absence of censoring and for the model in presence of censoring. ([2])
The simulation study has been carried out considering $M = 100$ clusters, three different total number of individuals, ($N_1 = 1000$, $N_2 = 2000$ and $N_3 = 5000$ for the *full model*, and $N_1 = 2000$, $N_2 = 5000$ and $N_3 = 10000$ for the *partial model*), and 100 Monte Carlo runs. The number of individuals within clusters n_m is assumed to be not fixed a priori. The covariates $\mathbf{Z}^{vm}$ are assumed to be time-independent and three dichotomic covariates are considered.

Assume that all subjects are in state 1. Let T_{ij}^{vm} be the time spent in state i before going to state j, distributed as an exponential distribution of parameter λ_{ij}^{vm}, given by (1), for $v = 1, \ldots, n_m$ and $m = 1, \ldots, 100$ (Kalbfleisch and Lawless, 1985). Moreover, T_{11}^{vm} is the time spent in state 1 before going to state 2 or to state 3; T_{22}^{vm} is the time spent in state 2 before going to state 3.

The frailty term is assumed to be distributed as:

$$\gamma_m \sim N(0, \sigma^2) \tag{14}$$

for $m = 1, \ldots, 100$, with $\sigma = 0.05$.

The vector of the parameters to be estimated is given by:

$$\boldsymbol{\xi} = (\lambda_{12,0}, \lambda_{13,0}, \lambda_{23,0}, \beta_{12}^1, \beta_{12}^2, \beta_{12}^3, \beta_{13}^1, \beta_{13}^2, \beta_{13}^3, \beta_{23}^1, \beta_{23}^2, \beta_{23}^3); \tag{15}$$

where $\lambda_{12,0}, \lambda_{13,0}, \lambda_{23,0}$ are the baseline hazard intensities, respectively, for the transition from state 1 to state 2, from state 1 to state 3, and from

[2] Part of the results from the simulation study are illustrated and discussed in the following sections; detailed results are available upon requests from the author.

state 2 to state 3; β_{ij}^1, β_{ij}^2 and β_{ij}^3 are the parameters of transition from state i to state j ($i, j = 1, 2, 3$), for each covariate.

The vector of true values of $\boldsymbol{\xi}$ are set to:

$$\boldsymbol{\xi} = (0.7, 0.6, 0.5, -0.9, 1.2, -0.8, -0.4, 0.7, -0.3, -0.7, 1.0, -0.6). \tag{16}$$

Assume the censoring time C^{vm} is distributed as:

$$C^{vm} \sim \exp(\alpha), \tag{17}$$

for $v = 1, \ldots, n_m$ and $m = 1, \ldots, 100$.

Three different values of α have been considered: 0.10, 0.20 and 0.30. Firstly, the scenario of the three-state Frailty Illness-Death model has been drawn for the complete model and for the partial model and then the estimates of the parameters have been computed and evaluated.

For each combination of covariates, for each time-point and for each N, the transition counts $N_{ij}(t_{g-1} - t_g)$, i.e. the number of subjects going from one state to another one and the number of censored observations at each state $N_i(t_g)$ and for each value of α have been computed (for $i = 1, 2, 3$ and for $g = 1, \ldots, 5$). ([3]) Thus, the effect of censoring has been observed. At increasing values of α, the mean value of censoring time goes to zero, and the probability of having censored observations become greater, as a consequence of losing information.

Now the distributions of the parameters' estimates are shown in Figure 1([4]) for the *full model*, and in Figures 2 - 5, for the *partial model*, with respect to the two hypothesis on the censored observations. These distributions are given according to the different number of individuals N, the different time-points, and the different values of α.
As demonstrated by Kalbfleisch and Lawless (1985), the maximum likelihood estimates of the vector of parameters are consistent (Figure 1).

Under the assumption *H1*, looking at the estimates of $\lambda_{12,0}$, $\lambda_{13,0}$, β_{12}^1, β_{12}^2, β_{12}^3, fixing the α to be equal to 0.1, it can be observed that the estimates are unbiased and the variance decreases, as number of observations

[3] The tables of both transition counts and the number of censored observations are available upon requests from the author.

[4] Since each box characterized a parameter, the y-scale is different for each parameter.

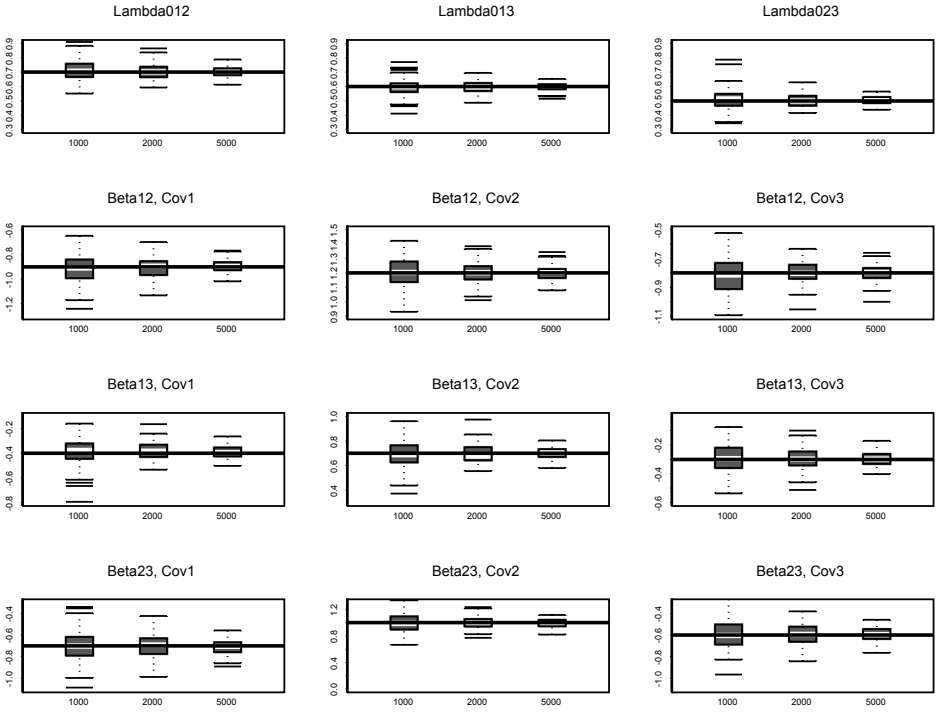


Figure 1. Boxplot of Maximum Likelihood Estimates for the model without censoring.

increases. Then, increasing value of α produces an increment of censored observations and a worsening of estimates (Figure 2). In other words, the estimates tend to be biased and their variability tend to increase.

Besides, the estimates $\lambda_{23,0}$, β_{13}^1, β_{13}^2, β_{13}^3, β_{23}^1, β_{23}^2, β_{23}^3 do not tend to the parameter values, i.e. they are biased. In fact, the parameters estimates of transitions among intermediate states are far from the parameters value, highlighting the restriction of the assumption *H1*, which ignores the intermediate transitions and does not capture this model characteristic (Figure 3).

Therefore, to overcome this limitation, the estimates of the vector of parameters have been computed by considering the second assumption (*H2*), which takes into account the possible intermediate transitions of censored observations, before reaching the absorbing states. It can be

noted how the estimates obtained under the assumption *H2* are better than ones obtained under the assumption *H1*. In fact, also the estimates of the intermediate states ($\lambda_{23,0}$, β^1_{13}, β^2_{13}, β^3_{13}, β^1_{23}, β^2_{23}, β^3_{23}) tend to be unbiased, at increasing the number of number of observations (Figure 4 - 5). Moreover, the increasing of the α values made the estimates worse, i.e. they tend to be biased, confirming the impact of censored observations.

Summing up briefly, the presence of censored units leads to results lacking in consistency, when the hypotheses underlying the censoring mechanism are not well specified. Indeed, the first hypothesis *H1* has been showed to be too general, since it is assumed that an observation censored at a generic time t goes directly to the absorbing state, at the end of the study T, without considering the intermediate transitions between that time t and the end of study T. In fact, the estimates of some parameters are unbiased, as number of observations increases, while the estimates of intermediate states' parameters are not.

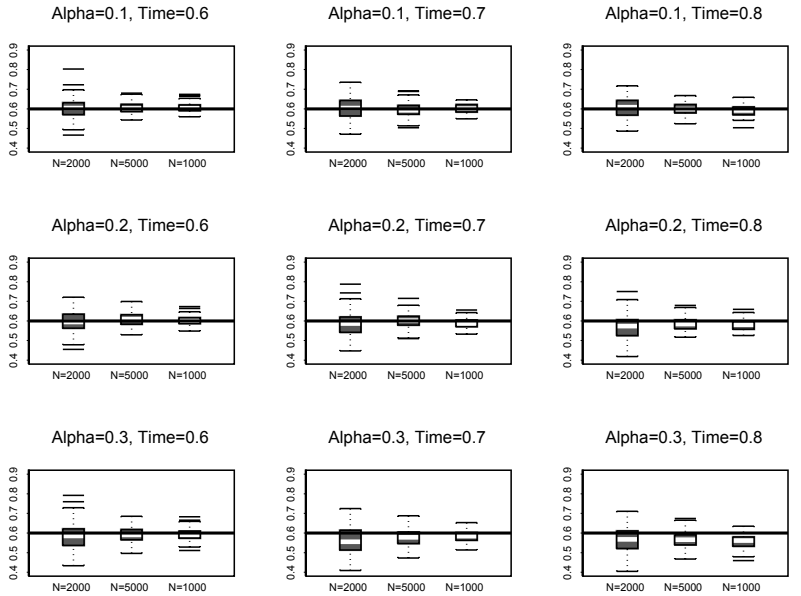
To overcome this limitation, the assumption *H2* has been considered. In this case, the estimates of intermediate transitions also tend to be unbiased. Therefore, the second hypothesis is more suitable than the first one.

Moreover, in both cases, the censoring influences the estimates. Indeed, at increasing values of α, since the probability of having censored observations become greater, the effect of censored observations increases, as a consequence of losing information and of worsening of estimates.

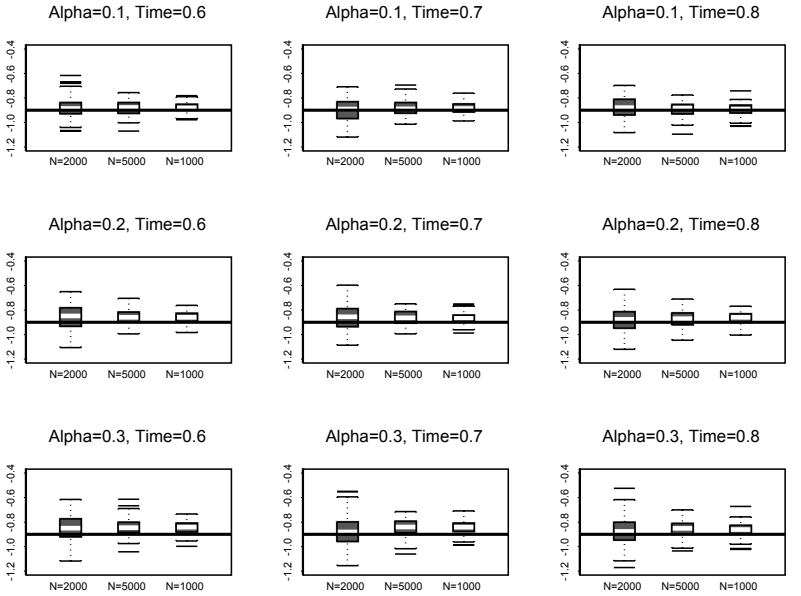
6. Conclusions

The aim of the paper has been to develop a new model, extending the Frailty model to Multi-State framework. The reason is related to the presence of multiple states and nature of data (clustered and censored), which occur in some empirical settings. If these characteristics do not take into account in the model, the estimates could be unbiased and inefficient.

In fact, by results of simulation study, it has been noted how the parameters estimates are consistent in case of complete observations, while the estimates are biased in presence of censored observations, when the hypothesis underlying the likelihood is not well specified.

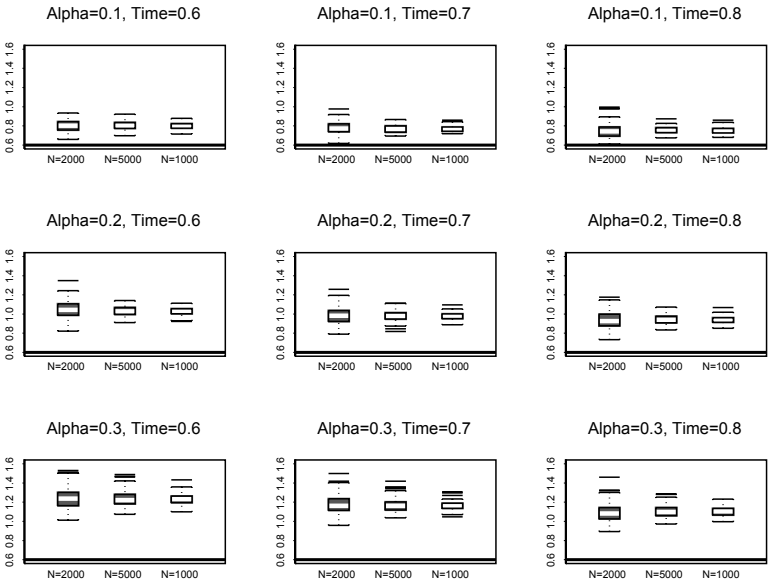


(a)

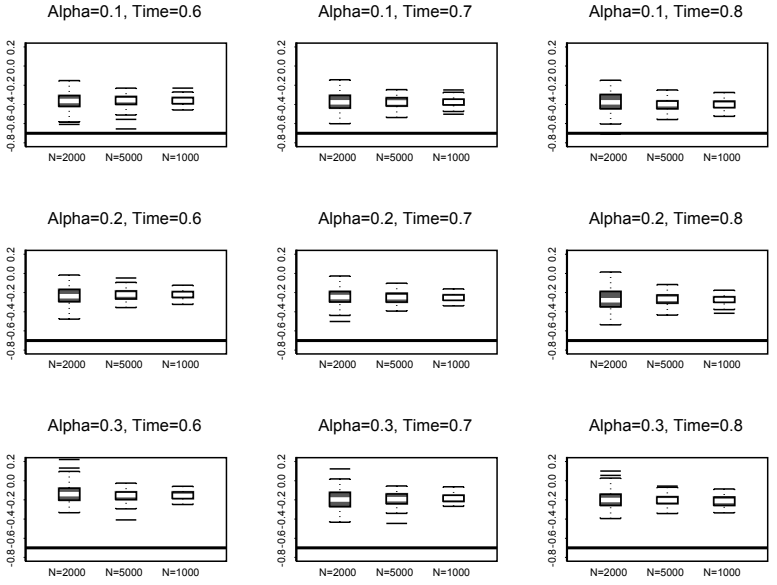


(b)

Figure 2. Model with censoring (assmption H1): $\lambda_{13,0}$ (panel a) and β_{12}^1 (panel b)

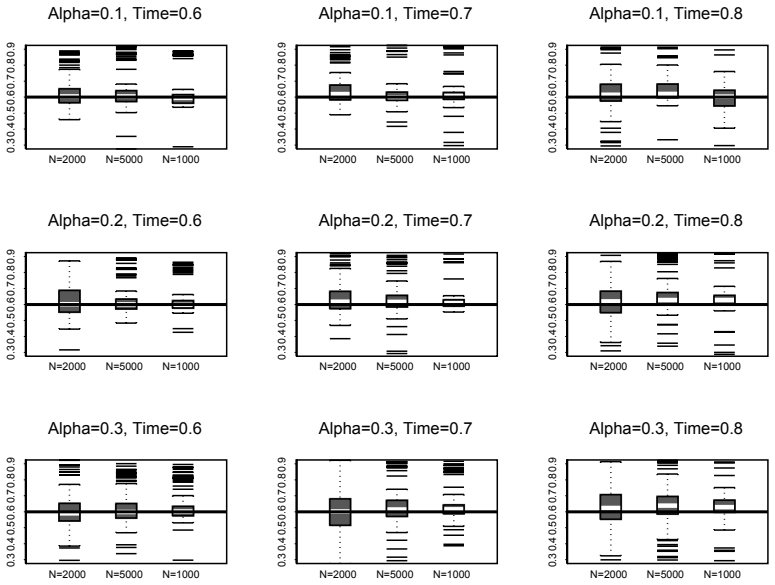


(a)

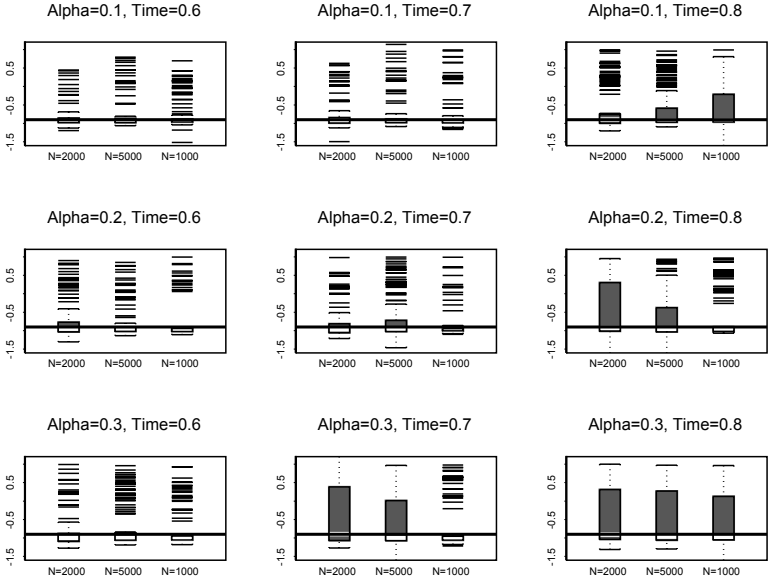


(b)

Figure 3. Model with censoring (assmption H1): $\lambda_{23,0}$ (panel a) and β_{23}^1 (panel b)

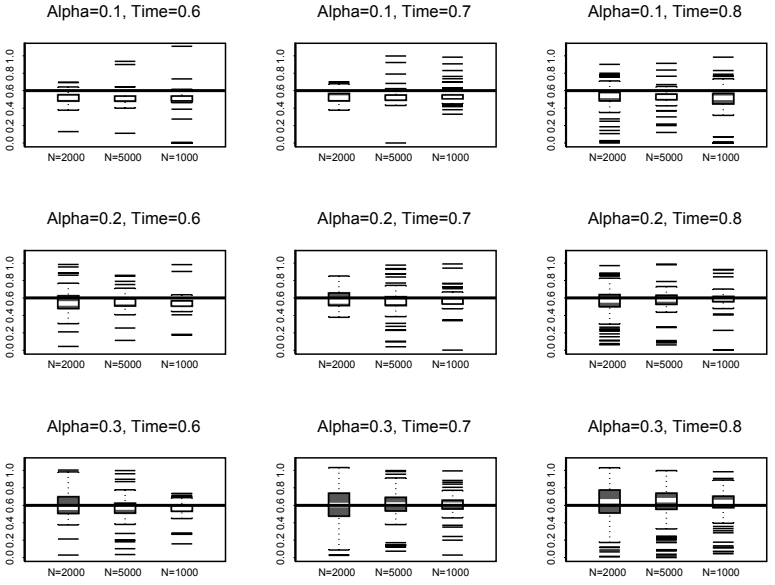


(a)

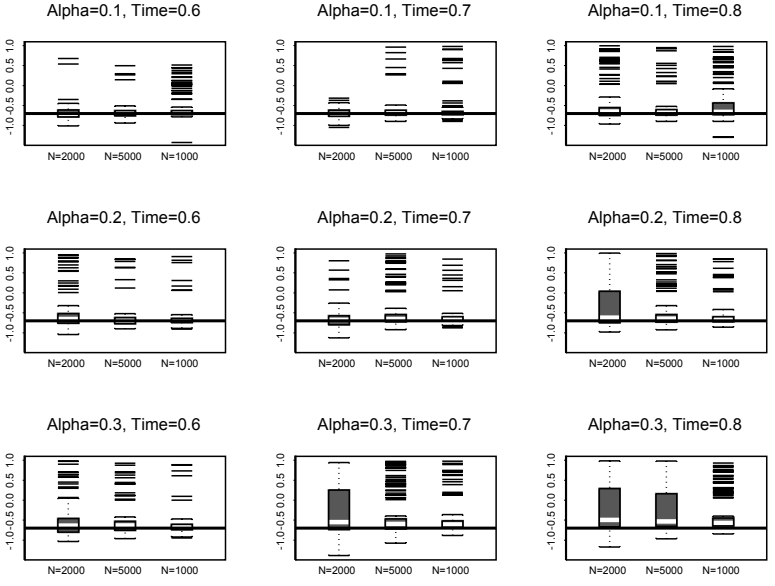


(b)

Figure 4. Model with censoring (assmption H2): $\lambda_{13,0}$ (panel a) and β_{12}^1 (panel b)



(a)



(b)

Figure 5. Model with censoring (assmption H2): $\lambda_{23,0}$ (panel a) and β^1_{23} (panel b)

As possible future developments, the consistency of the estimates under *H2* will be investigated and proved. Then, in order to take into account the effect and the influence of censored observations in the likelihood under the hypothesis *H2*, the construction of the weights $w_{ik}(t_g)$, for $k = 1, 2, 3$ and for $g = 1, \ldots, 5$ need to be changed. In fact, since the weights depend on the vector of parameters ξ, i.e. $w_{ik}((t_g)|\xi)$, they must be estimated as such.

Moreover, further investigation will analyze how the transition intensities and transition probabilities change, not only when the censoring is informative and/or when it is a random variable, but also when time-homogeneous hypothesis does not hold and more appropriate models, such as non-homogeneous approaches, are needed.

References

Andersen P.K., Borgan Ø., Gill R.D., Keiding N. (1993), *Statistical Models based on Counting Processes*, Springer, Berlin.

Clayton D.G. (1978), A model for association in bivariate life tables and its application in epidemiological studies of familiar tendency in chronic disease incidence, *Biometrika*, 65, 141–151.

Clayton D.G., Cuzick J. (1985), Multivariate generalizations of the proportional hazards model (with discussion), *Journal of Royal Statistical Society. Series A*, 148, 82–117.

Collett D. (2003), *Modelling Survival Data in Medical Research*, Chapman & Hall/CRC, London.

Cox D.R., Miller H.D. (1965), *The Theory of Stochastic Processes*, Mertheun, London.

Hougaard P. (1984), Life tables methods for heterogeneous populations: distributions describing the heterogeneity, *Biometrika*, 71, 75–83.

Hougaard P. (1986), Survival models for heterogeneous populations derived from stable derivations, *Biometrika*, 73, 387–396.

Hougaard P. (2001), *Analysis of Multivariate Survival Data*, Springer, Berlin.

Kalbfleisch J.D., Lawless J.F. (1985), The analysis of Panel Data under a Markov Assumption, *Journal of the American Statistical Association*, 80, 863–871.

Vaupel J.W., Manton K.G., Stallard E. (1979), The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography*, 162, 439–454.