Conjoint Degradation Model of Disablement for Survival and Longitudinal Data Measured with Errors in Presence of Covariates

V. Bagdonavičius[1] and M. Nikulin[2]

[1] *Department of Mathematical statistics, University of Vilnius, Vilnius, Lithuania*

[2] *EA 2961 "Mathematical Statistics and Its Applications", Victor Segalen University Bordeaux 2, France & Laboratory of Statistical Methods, Steklov Mathematical Institute, Saint Petersburg, Russia*

Abstract

The paper considers the semiparametric analysis of several new degradation and failure time regression models without and with time depending covariates. These joint models for survival and longitudinal data measured with errors can be applied in studies of longevity, aging and degradation in survival analysis, biostatistics, epidemiology, demography, biology and reliability.

Keywords: aging, censored data, consistency, degradation, degradation process, disability,failure, longevity, longitudinal data, measurement error, noise, nonparametric estimation, optimal estimator, path model, regression parameters, reliability, semiparametric estimation, survival analysis, traumatic event, Wiener process, Wulfsohn-Tsiatis model.

AMS classification: 62F10, 62J05, 62GO5, 62N05, 60M10, 62G30

1. Introduction

The purpose of this paper is at first the construction of a statistical accelerated degradation model of disablement in the elderly to verify that a hierarchical relationship exists between the concepts of Activities Daily Living (ADL), Instrumental Activities of Daily Living (IADL) and mobility and to use this model to study the evolution of disability. The concepts of ADL, IADL and mobility have been developed and widely used for the assessment of the consequences of disease. ADL include basic activities of personal care, IADL functions are concerned with a person's ability to perform activities in a given environment and according to sex-linked roles, and the definitions of mobility vary from basic function to more complex level of activities. The Rosow-Breslow (1966) scale includes items which explore gross mobility. The disability is defined as a reduction or loss of functional capacity or activity resulting from an impairments. The PAQUID (Personnes Agées QUID) study is an epidemiological study of functional and cerebral ageing conducted in 75 randomly chosen parishes in two administartive areas "Gironde" and "Dordogne" in the South-Western France since 1988/89. To be included at baseline subjects has to be aged 65 or more on the 1 January of 1998. 3777 persons gave their written consent to participate and were visited at home by specially trained psychologists for the baseline interview and were followed-up 3 and 5 years after in order to identify cases of *dementia*. The general descriptions of the PAQUID study

one can see, for example, in Dartigues and al. (1991), Barberger-Gateau and al (1992, 2000). Following Katz and al. (1970), Spector (1990) and Rosow and Breslow (1966) functional assessment variables used in the present study included:
1. Five ADL items: bathing, dressing, going to toilet, transfer and feeding.
2. Five IADL items for both sexes: ability to use telephone, shopping, mode of transportation, responsibility for own medication, ability to handle finances.
3. Three items from the Rosow scale; walk up and down to second floor, walk half a mile, and do heavy work around the house.

For each of these three domains a subject was considered as "*dependent*" if he could not perform at least one activity of the domain without a given level of help. Dependency for each domain was represented by a single item coded 0 (fully independent subject for the domain) versus 1. In Guttman's approach an unidimensional scale is in which the subject's responses to the items would place individuals in perfect order. According to this perfect scale the sum of correct responses to the items can be used to recreate all the responses of a subject in a single digit. In this case a single disability scale is obtained by summing up the responses to each of the three dichotomized disability items for all subjects fitting into the hierarchy:
subjects without any dependency were coded 0,
those only in th Rosow scale were coded 1,
those dependent in Rosow and IADL scales not ADL scale were coded 2,
and those dependent were coded 3.

Zdorova-Cheminade (2003) studied by simulation the considered model as statistical degradation model of disablement in the elderly to verify that a hierarchical relationship exists between the concepts of Activities Daily Living, Instrumental Activities of Daily Living and mobility and to use this model to study the evolution of disability. The cumulative disability scale was used to describe the degradation process in time. In longitudinal analysis an additional level was considered to the disability index to take death into account. It is evident that this approach can be used in many other medical studies where a degradation is observed, especially in oncology.

To construct a mathematical model we shall use the terminology used in theory of reliability, (see, Meeker and Escobar (1998), Gavrilov and Gavrilova (2003), Bagdonavicius and Nikulin (2001)). Functioning of a unit is characterized by its degradation process and by the random moments of its potential failures. We call a failure of a unit natural if the degradation attains some critical level. Other failures are called traumatic. These can be related with production defects, caused by mechanical damages or by fatigue of components etc.

The intensities of the traumatic failures depend on degradation. As a rule these intensities are increasing functions of degradation values. Suppose that the lifetime of a unit is determined by the degradation process $Z(t)$ and the moment of its potential traumatic failure T. For example, $Z(t)$ may be the value of tire wear at the moment when a tire has run t km (in this case "time" is the tire run), the size of fatigue crack, the size of failure-causing conducting filament of chlorine-copper compound in a printed-circuit board, luminosity of light emitting diode at the moment t (Meeker and Escobar (1998), Bagdonavičius and Nikulin (2002)), etc.

Denote by T^0 the moment of non-traumatic failure, i.e. the moment when the degradation attains some critical value z_0. The moment of the unit's failure is

$$\tau = T^0 \wedge T.$$

2. Modeling the degradation-failure time process

We suppose that the real degradation process is modeled by the general path model (Meeker and Escobar (1998))

$$Z_r(t) = g(t, A); \tag{1}$$

here $A = (A_1, \ldots, A_r)$ is a random vector with positive components and the distribution function F_A, and g is a specified continuously differentiable increasing in t function. The typical form of the degradation curves is

$$g(t, a) = e^{a_1}(1+t)^{a_2}. \tag{2}$$

In the particular case of linear degradation $a_2 = 0$.

Denote by h the function inverse to g with respect to the first argument. Evidently, it is continuously differentiable and increasing in t. Moreover,

$$T^0 = h(z_0, A). \tag{3}$$

The observed degradation process $Z = Z(t), t \geq 0$ may be slightly different from the real degradation process.

Suppose that the values of the real degradation process are measured at time moments $t_1, \ldots, t_m$.

Model 1 (Degradation with measurement errors). *The observed degradation values are*

$$Z(t_j) = Z_r(t_j)\, U(t_j), \tag{4}$$

where $e_j = \ln U(t_j)$ *are i.i.d. random variables,* $e_j \sim N(0, \sigma^2)$.

Model 2 (Degradation with the noise). *The observed degradation process is*

$$Z(t) = Z_r(t)\, U(t), \tag{5}$$

where

$$V(t) = \ln U(t) = \sigma W(c(t)), \tag{6}$$

W *is the standard Wiener process independent on* A, *and* $c : [0, \infty \to [0, \infty)$, $c(0) = 0$, *is a specified continuous and increasing time function,* $c(0) = 0$.

For any $t > 0$ the median of the random variable $U(t)$is 1.

The cumulative distribution function (c.d.f.) of the real degradation at the moment t is

$$F_{r0}(z \mid t) = P(g(t, A) \leq z) = \int\limits_{g(t,a) \leq z} dF_A(a), \tag{7}$$

the mean real degradation attained at the moment t is

$$m_{r0}(t) = Eg(t, A) = \int\limits_{\mathcal{A}} g(t, a) dF_A(a), \tag{8}$$

where $\mathcal{A}$ is the set of possible values of the random vector A.

Note that the characteristics (7) and (8) have sense only in the ideal situation when the probability of failure until the moment t is zero. So we shall modify them defining the c.d.f. and other characteristics of the degradation at any moment t as conditional given survival up to this moment.

Let T be the time to traumatic failure.

Model 3. *The conditional survival of T given the real degradation process has the form:*

$$S_T(t \mid A) = P\{T > t \mid g(s, A), 0 \le s \le t\} = \exp\Big\{-\int_0^t \lambda_0(s,\alpha)\lambda(g(s,A))ds\Big\}; \qquad (9)$$

here λ is the unknown intensity function, $\lambda_0(s,\alpha)$ being from a parametric family of hazard functions.

Note that the function λ is defined on the set of degradation values, not on the time scale.

The model states that the conditional hazard rate $\lambda_T(t \mid A)$ at the moment t given the degradation $g(s, A), 0 \le s \le t$, has the form

$$\lambda_T(t \mid A) = \lambda_0(t,\alpha)\lambda(g(t,A)). \qquad (10)$$

The term $\lambda(g(t, A))$ shows the influence of degradation on the hazard rate, the term $\lambda_0(t,\alpha)$ shows the-influence of time on the hazard rate not explained by degradation. If, for example, $\lambda_0(t,\alpha) = (1 + t)^\alpha, e^{\alpha t}$, then $\alpha = 0$ corresponds to the case when the hazard rate at any moment t is a function of the degradation level at this moment.

Wulfsohn and Tsiatis (1997) considered the so called joint model for survival and longitudinal data measured with error, given by

$$\lambda_T(t \mid A) = \lambda_0(t)e^{\beta(A_1+A_2t)}$$

with bivariate normal distribution of of (A_1, A_2). The difference: in our model the function λ, characterizing the influence of degradation on the hazard rate, is non-parametric, in the Wulfsohn-Tsiatis model this function is parametric. On the other hand, the baseline hazard rate λ_0 (it is proportional to the hazard rate which should be observed if the degradation would be absent) is parametric in our model and non-parametric in Wulfsohn-Tsiatis model. We consider the case when the distribution of A is not specified.

Let us consider the model (9) and set

$$\Lambda(z) = \int_0^z \lambda(y)dy. \qquad (11)$$

The conditional survival function (5) can be rewritten in the following form

$$S_T(t \mid A) = \exp\Big\{-\int_{g(0,A)}^{g(t,A)} \lambda_0(h(z,A),\alpha)h'(z,A)\, d\Lambda(z)\Big\}, \qquad (12)$$

where h is the function inverse to g with respect to the first argument.

3. Survival and degradation characteristics

The survival function and the mean of the time-to-falure τ are

$$S_\tau(t) = P(\tau > t) = \int\limits_{g(t,a)\le z_0} S_T(t \mid a)dF_A(a), \tag{13}$$

and

$$e_\tau = E(\tau) = \int\limits_{\mathcal{A}} \{h(z_0,a)S_T(h(z_0,a) \mid a)+$$

$$\int\limits_{g(0,a)\wedge z_0}^{z_0} \lambda_0(h(z,a))h(z,a)h'(z,a)S_T(h(z,a) \mid a)d\Lambda(z)\}dF_A(a), \tag{14}$$

respectively.

Set

$$\delta = \begin{cases} 0, & \text{if} \quad \tau = T^{(0)}, \\ 1, & \text{if} \quad \tau = T. \end{cases} \tag{15}$$

Important reliability characteristics are the probability of the traumatic failure in the interval $[0;t]$:

$$P^{(tr)}(t) = P\{\tau \le t,\ \delta = 1\} = 1 - \int\limits_{\mathcal{A}} S_T(t \wedge h(z_0,a)) \mid a)dF_A(a) \tag{16}$$

and the probability of the non-traumatic failure in the interval $[0;t]$:

$$P^{(0)}(t) = \{\tau \le t,\ \delta = 0\} = \int\limits_{g(t,a)\ge z_0} S_T(h(z_0,a)) \mid a)dF_A(a). \tag{17}$$

Now we are able to modify the characteristics (7) and (8) to the case of the real situation.

The c.d.f. of the real degradation at the moment t given the survival until this moment is:

$$F_r(z \mid t) = P(g(t,A) \le z \mid \tau > t) = \frac{1}{S_\tau(t)} \int\limits_{g(t,a)\le z} S_T(t \mid a)dF_A(a), \tag{18}$$

if $z \le z_0$, and $F_r(z \mid t) = 1$, if $z > z_0$. The mean real degradation attained at the moment t given the survival until this moment is

$$m_r(t) = E(g(t,A) \mid \tau > t) = \frac{1}{S_\tau(t)} \int\limits_{\mathcal{A}} \{g(t,a) \wedge z_0 - z_*\}S_T(t \mid a)dF_A(a), \tag{19}$$

where z_* is the minimal possible value of degradation.

For any $B \subset \mathbf{R}, B_0 \subset \mathbf{R}$, $t_1 \leq s < t$ the probability that at the moment t the value of the real degradation will be in the set B given that at the moment s it is in the set B_0 and given the survival until t, is

$$P(g(t,A) \in B \mid g(s,A) \in B_0, T > t)$$

$$= \int\limits_{g(t,a)\in B, g(s,a)\in B_0} S_T(t \mid a)dF_A(a \mid t) / \int\limits_{g(s,a)\in B_0} S_T(t \mid a)dF_A(a \mid t). \tag{20}$$

Denote by

$$T^0(z) = h(z, A)$$

the time needed to attain the level z $(z \leq z_0)$ of the real degradation. The conditional survival function and the conditional mean of $T^0(z)$ given the survival until the degradation attains the level z and given survival until the moment $T^0(z)$ are

$$P(T^0(z) > t \mid T > T^0(z)) = \frac{\int\limits_{g(t,a)<z} S_T(h(z,a) \mid a)dF_A(a)}{\int\limits_{\mathcal{A}} S_T(h(z,a) \mid a)dF_A(a)}, \tag{21}$$

and

$$E(T^0(z) \mid T > T^0(z)) = \frac{\int\limits_{\mathcal{A}} h(z_0,a)S_T(h(z,a) \mid a)dF_A(a)}{\int\limits_{\mathcal{A}} S_T(h(z,a) \mid a)dF_A(a)}, \tag{22}$$

respectively.

4. Modeling covariate effects on degradation and survival

Suppose units are observed under the covariate $x = x(t) = (x_1(t), \ldots, x_m(t)), t \geq 0$, and the values of the real degradation process are measured at time moments $t_1, \ldots, t_m$.

In this case, following Bagdonavičius and Nikulin (2001), the real degradation process is modeled by

$$Z_r(t \mid x) = g(\varphi(t, \beta, x), A), \tag{23}$$

where

$$\varphi(t, \beta, x) = \int_0^t e^{\beta x^T(u)} du, \tag{24}$$

is the link function, $\beta = (\beta_1, \ldots, \beta_m)$ is the vector of the regression parameters.

If the covariate is constant over time then

$$\varphi(t, \beta, x) = e^{\beta x^T} t. \tag{25}$$

Model 4 (Degradation with measurement errors). *The observed degradation values are*

$$Z(t_j \mid x) = Z_r(t_j \mid x)\, U(t_j), \tag{26}$$

where $e_j = \ln U(t_j)$ *are i.i.d. random variables,* $e_j \sim N(0, \sigma^2)$.

Model 5 (Degradation with the noise). *The observed degradation process is*

$$Z(t \mid x) = Z_r(t \mid x)\, U(t),$$

where

$$V(t) = \ln U(t) = \sigma W(c(t)), \tag{27}$$

W is the standard Wiener process independent on A, and $c : [0, \infty \to [0, \infty)$, is a specified increasing and continuous time function, $c(0) = 0$.

The c.d.f. of the real degradation at the moment t given the covariate $x(s), 0 \le s \le t$, is

$$F_{r0}(z \mid t, x) = \int\limits_{g(\varphi(t,\beta,x),a) \le z} dF_A(a), \tag{28}$$

the mean real degradation attained at the moment t is

$$m_{r0}(t, x) = \int\limits_{\mathcal{A}} g(\varphi(t, \beta, x), a) dF_A(a). \tag{29}$$

As in the case without covariates these characteristics have sense only in the ideal situation when the probability of failure until the moment t is zero.

Similarly as in (9), we model the survival distrubution:

Model 6 *The probability of survival until the moment t given the real degradation process $Z_r(s \mid x), 0 \le s \le t$ is*

$$S_T(t \mid a, x) = P\{T > t \mid A = a, x(s), 0 \le s \le t\} =$$

$$\exp\Big\{ - \int_0^t \lambda_0(\varphi(s, \beta, x), \alpha) \lambda(g(\varphi(t, \beta, x), a)) ds \Big\}, \tag{30}$$

where the λ is an unknown intensity function, $\lambda_0(s, \alpha)$ being from a parametric family of hazard functions.

The probability (30) can be rewritten in the following form

$$S_T(t \mid a, x) = \exp\Big\{ - \int_{g(\varphi(0,\beta,x),a)}^{g(\varphi(t,\beta,x),a)} \lambda_0(h(z, a), \alpha) \psi'(h(z, a), \beta, x) h'(z, a)\, d\Lambda(z) \Big\}, \tag{31}$$

where $\psi(s, \beta, x)$ is the inverse function of $\varphi(t, \beta, x)$ with respect to the first argument, ψ_1' is the derivative of ψ with respect to the first argument, and Λ is defined by the formula (11).

In the case of constant covariates

$$\psi(t, \beta, x) = e^{-\beta x^T} t, \quad \psi'(t, \beta, x) = e^{-\beta x^T}.$$

The given model implies that the survival function and the mean of the random variable τ given the covariate x are

$$S_\tau(t \mid x) = P(\tau > t \mid x(s), 0 \le s \le t) = \int\limits_{g(\varphi(t,\beta,x),a) < z_0} S_T(t \mid a, x) dF_A(a). \tag{32}$$

The probability of the traumatic failure in the interval $[0;t]$under the covariate x is:

$$P^{(tr)}(t \mid x) = P\{\tau \le t, \delta = 1 \mid x(s), 0 \le s \le t\} =$$
$$1 - \int_{\mathcal{A}} S_T(t \wedge \psi(h(z_0, a), \beta, x)) \mid a, x) dF_A(a) \tag{33}$$

and the probability of non-traumatic failure in the interval $[0;t]$ under the covariate x is

$$P^{(0)}(t \mid x) = \{\tau \le t, \delta = 0 \mid x(s), 0 \le s \le t\} =$$
$$\int_{g(\varphi(t,\beta,x)),a) \ge z_0} S_T(\psi(h(z_0, a), \beta, x)) \mid a, x) dF_A(a), \tag{34}$$

where δ is given by (15).

The c.d.f. of the real degradation at the moment t given the survival until this moment is

$$F_r(z \mid t, x) = P(Z_r(t \mid x) \le z \mid x(s), 0 \le s \le t, \tau > t)$$
$$= \frac{1}{S_\tau(t \mid x)} \int_{g(\varphi(t,\beta,x),a) \le z} S_T(t \mid a, x) dF_A(a), \tag{35}$$

and the mean real degradation attained at the moment t given the survival until this moment is

$$m_r(t, x) = EZ_r(t \mid x) = \frac{1}{S_T(t \mid x)} \int_{\mathcal{A}} (g(\varphi(t, \beta, x), a) \wedge z_0 - z_*) S_T(t \mid a, x) dF_A(a). \tag{36}$$

For any $B \subset \mathbf{R}, B_0 \subset \mathbf{R}$, the probability that at the moment t the value of the real degradation will be in the set B given that at the moment s it is in the set B_0 and given $x(u), 0 \le u \le t$, is

$$P(Z_r(t \mid x) \in B \mid Z_r(s \mid x) \in B_0, \tau > t, x(u), 0 \le u \le t)$$
$$= \frac{\int_{g(\varphi(t,\beta,x),a) \in B, g(\varphi(s,\beta,x),a) \in B_0} S_T(t \mid a, x) dF_A(a)}{\int_{g(\varphi(s,\beta,x),a) \in B_0} S_T(t \mid a, x) dF_A(a)}. \tag{37}$$

Denote by

$$T^0(z, x) = \psi(h(z, A), \beta, x) \tag{38}$$

the time needed to attain the level z of the real degradation under the covariate x. The conditional survival function and the conditional mean of $T^0(z, x)$ given the survival until $T^0(z, x)$, and given $x(u), 0 \le u \le t$, are

$$P(T^0(z, x) > t \mid \tau > T^0(z, x), x(u), 0 \le u \le t) =$$
$$\frac{\int_{g(\varphi(t,\beta,x),a) < z} S_T(\psi(h(z, a), \beta, x) \mid a, x) dF_A(a)}{\int_{\mathcal{A}} S_T(\psi(h(z, a), \beta, x) \mid a, x) dF_A(a)}, \tag{39}$$

and

$$E(T^0(z) \mid \tau > T^0(z), x(u), 0 \le u \le t) =$$

$$\frac{\int\limits_{\mathcal{A}} \psi(h(z,a),\beta,x) S_T(\psi(h(z,a),\beta,x) \mid a,x) dF_A(a)}{\int\limits_{\mathcal{A}} S_T(\psi(h(z,a),\beta,x) \mid a,x) dF_A(a)}, \tag{40}$$

respectively.

5. Semiparametric estimation of degradation and survival characteristics: models without covariates

Assume that the c.d.f. F_A and the intensities λ are completely unknown.

Fix the degradation measurement moments $t_{i,1}, \dots, t_{i,m_i}$ of the ith unit ($i = 1 \dots, n$). If the failure time τ_i of this unit occurs in the interval $[t_{i,j_i}, t_{i,j_i+1})$ ($j_i = 1, \dots, m_i; t_{m_i+1} = \infty$) then the values $Z_{i1}, \dots, Z_{i,j_i}$ of the degradation process Z_i of the ith unit are observed at the time moments $t_{i1}, \dots, t_{i,j_i}$. Set

$$Y_{ij} = \ln Z_{ij}, \quad Y_i = (Y_{i1}, \dots, Y_{ij_i})^T.$$

Then given $A_i = a_i$, j_i

$$Y_i \sim N(\mu_i, \sigma^2 \Sigma_i),$$

where

$$\mu_i = (\mu_{i1}, \dots, \mu_{i,j_i})^T, \quad \mu_{ij} = \mu_{ij}(a_i) = \ln g(t_{ij}, a_i), \quad \Sigma_i = \| s_{ikl} \|_{j_i \times j_i},$$

$$s_{ikl} = \mathbf{1}_{\{k=l\}} \quad \text{(first model)},$$

$$s_{ikl} = c_{ik} \wedge c_{il}, \quad \text{(second model)}, \quad c_{ij} = c(t_{ij}). \tag{41}$$

Denote by b_{ikl} the elements of the inverse matrix Σ_i^{-1}, and by N the number of units such that $j_i \ge r$ (usually $r = 2$).

The predictors $\hat{A}_i$ of the random vectors A_i are found minimizing with respect to $a_1, \dots, a_N$ the quadratic form

$$\sum_{i=1}^{N} (Y_i - \mu_i(a_i))^T \, \Sigma_i^{-1} \, (Y_i - \mu_i(a_i)). \tag{42}$$

Suppose that the function $g(t, a)$ is differentiable with respect to a and set $m = \sum_{i=1}^{N} j_i$. The estimator $\hat{\sigma}^2$ of the parameter σ^2 is found maximizing with respect to σ^2 the conditional likelihood function

$$L(\sigma^2 \mid \hat{A}_1, \dots, \hat{A}_N) =$$

$$\frac{1}{(2\pi)^{m/2} \sigma^m} \prod_{i=1}^{N} | \Sigma_i |^{-1/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{N} (Y_i - \mu_i(\hat{A}_i))^T \, \Sigma_i^{-1} \, (Y_i - \mu_i(\hat{A}_i)) \right\}. \tag{43}$$

The minimization of the quadratic form (42) and the maximization of the conditional likelihood function gives the following equations for $\hat{\sigma}^2$ and $\hat{A}_i$ computing:

$$\sum_{k=1}^{j_i} \frac{\partial}{\partial \hat{A}_i} \ln g(t_{ik}, \hat{A}_i) \sum_{l=1}^{j_i} \{Y_{il} - \ln g(t_{il}, \hat{A}_i)\} b_{ikl} = 0; \tag{44}$$

$$\hat{\sigma}^2 = \frac{\hat{c}}{m} \tag{45}$$

where

$$\hat{c} = \sum_{i=1}^{N} (Y_i - \mu_i(\hat{A}_i))^T \ \Sigma_i^{-1} \ (Y_i - \mu_i(\hat{A}_i)). \tag{46}$$

Let us consider the following model for the paths $g(t, A_i)$:

$$g(t, A_i) = e^{A_{i1}} (1+t)^{A_{i2}}. \tag{47}$$

Set $\mathbf{1} = (1, \dots, 1)^T_{j_i}$, $C_i = (c_{i1}, \dots, c_{i,j_i})^T$. In such a case we have

$$\hat{A}_{i1} = \frac{c_i d_i - e_i f_i}{c_i^2 - b_i e_i}, \quad \hat{A}_{i2} = \frac{c_i f_i - b_i d_i}{c_i^2 - b_i e_i}, \tag{48}$$

and

$$\hat{\sigma}^2 = \frac{\hat{c}}{m}, \tag{49}$$

where

$$\begin{aligned} \hat{c} &= \sum_{i=1}^{N} (g_i + b_i \hat{A}_{1i}^2 + e_i \hat{A}_{2i}^2 + 2 c_i \hat{A}_{1i} \hat{A}_{2i} - 2 f_i \hat{A}_{1i} - 2 d_i \hat{A}_{2i}) \\ b_i &= \mathbf{1}^T \Sigma_i^{-1} \mathbf{1}, \quad c_i = C_i^T \Sigma_i^{-1} \mathbf{1}, \quad d_i = Y_i^T \Sigma_i^{-1} C_i, \\ e_i &= C_i^T \Sigma_i^{-1} C_i, \quad f_i = Y_i^T \Sigma_i^{-1} \mathbf{1}, \quad g_i = Y_i^T \Sigma_i^{-1} Y_i. \end{aligned} \tag{50}$$

The conditional mean of the estimator $\hat{\sigma}^2$ given N and j_i is:

$$E(\hat{\sigma}^2 \mid j_i, i = 1, \dots, N) = \frac{m+2}{m} \sigma^2,$$

so $\hat{\sigma}^2$ is a consistent estimator of σ^2.

Set

$$V_{ij} = e_{ij} \quad \text{(first model)}, \quad V_{ij} = \sigma W_i(c_{ij}) \quad \text{(second model)},$$
$$V_i = (V_{i1}, \dots, V_{i,j_i})^T.$$

Replacing Y_{ij} by their expressions $Y_{ij} = A_{i1} + A_{i2} c_{ij} + V_{ij}$ in the formulas (48), we obtain that

$$\hat{A}_i = A_i + \varepsilon_i,$$

where

$$\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i1}), \quad \varepsilon_{i1} = \frac{c_i v_i - e_i u_i}{c_i^2 - b_i e_i}, \quad \varepsilon_{i2} = \frac{c_i u_i - b_i v_i}{c_i^2 - b_i e_i},$$
$$u_i = V_i^T \Sigma_i^{-1} \mathbf{1}, \quad v_i = V_i^T \Sigma_i^{-1} C_i.$$

Given j_i the vector $V_i \sim N(0, \sigma^2 \Sigma_i)$, so the bi-dimensional vector ε_i is normally distributed with zero mean and the density

$$\varphi_i(x;\sigma^2) = \frac{1}{2\pi\sigma_{i1}\sigma_{i2}\sqrt{1-\rho_i^2}} \exp\left\{-\frac{1}{2(1-\rho_i^2)}\left[\frac{x_1^2}{\sigma_{i1}^2} - 2\rho_i\frac{x_1x_2}{\sigma_{i1}\sigma_{i2}} + \frac{x_2^2}{\sigma_{i2}^2}\right]\right\},$$

which depends only on σ^2, because

$$\sigma_{i1}^2 = Var(\varepsilon_{i1}) = \frac{e_i}{b_ie_i - c_i^2}\sigma^2, \quad \sigma_{i2}^2 = Var(\varepsilon_{i2}) = \frac{b_i}{b_ie_i - c_i^2}\sigma^2,$$

$$cov(\varepsilon_{i1}, \varepsilon_{i2}) = -\frac{c_i}{b_ie_i - c_i^2}\sigma^2,$$

and the correlation coefficient

$$\rho_i = -\frac{c_i}{\sqrt{b_ie_i}}.$$

Note that the means of the random variables A_i and $\hat{A}_i$ are equal:

$$E(\hat{A}_i) = E(A_i),$$

and

$$Eg(t, \hat{A}_i) = m_{r0}(t)\exp\{\frac{1}{2}[\sigma_{i1}^2 + 2\sigma_{i1}\sigma_{i2}\rho_i \ln(1+t) + \sigma_{i2}^2 \ln^2(1+t)]\}.$$

So a consistent estimator of the mean real degradation $m_{r0}(t)$ is

$$\hat{m}_{r0}(t) = \frac{\sum\limits_{i=1}^{N} g(t, \hat{A}_i)}{\sum\limits_{i=1}^{N} \exp\{\frac{1}{2}[\hat{\sigma}_{i1}^2 + 2\hat{\sigma}_{i1}\hat{\sigma}_{i2}\hat{\rho}_i \ln(1+t) + \hat{\sigma}_{i2}^2 \ln^2(1+t)]\}},$$

and a consistent estimator of the mean observed degradation m_{o0} (for the model 2) is

$$\hat{m}_{o0}(t) = e^{\frac{1}{2}\hat{\sigma}^2 c^2(t)}\hat{m}_{r0}(t),$$

where

$$\hat{\sigma}_{i1}^2 = \frac{e_i}{b_ie_i - c_i^2}\hat{\sigma}^2, \quad \hat{\sigma}_{i2}^2 = \frac{b_i}{b_ie_i - c_i^2}\sigma^2, \quad \hat{\rho}_i = -\frac{c_i}{\sqrt{b_ie_i}}.$$

Lemma 1. *If $t_{ij} = q_i(j/m_i)$, $j = 1, \ldots, m_i$, where q_i is a continuous increasing on $[0,1]$ function, $q_i(0) = 0, q_i(1) \le c < \infty$. Under the first model almost surely*

$$\lim_{m_i \to \infty} \sigma_{i1}^2 = \lim_{m_i \to \infty} \sigma_{i2}^2 = 0.$$

Proof. It is sufficient to prove that almost surely

$$b_i - \frac{c_i^2}{e_i} = \sum_{j=1}^{j_i} c_{ij}^2 - \frac{1}{n}(\sum_{j=1}^{j_i} c_{ij})^2 \xrightarrow{a.s.} \infty, \quad e_i - \frac{c_i^2}{b_i} = j_i - \frac{(\sum\limits_{j=1}^{j_i} c_{ij})^2}{\sum\limits_{j=1}^{j_i} c_{ij}^2} \xrightarrow{a.s.} \infty.$$

The function $r_i(x) = c(q_i(x))$ is continuous and increasing on $(0,1)$, $r_i(0) = 0$. We have

$$\frac{1}{j_i}(b_i - \frac{c_i^2}{e_i}) = \frac{1}{j_i}\sum_{j=1}^{j_i} r_i^2(j/j_i) - (\frac{1}{j_i^2}\sum_{j=1}^{j_i} r_i(j/j_i))^2 \stackrel{a.s.}{\to}$$

$$\int_0^{a_i} r_i^2(x)dx - (\int_0^{a_i} r_i(x)dx)^2 > 0,$$

where

$$a_i = q_i^{-1}(E(\tau_i \wedge q_i(1)).$$

So $b_i - c_i^2/e_i \stackrel{a.s.}{\to} \infty$.

Analogously

$$\frac{1}{j_i}(e_i - \frac{c_i^2}{b_i}) \stackrel{a.s.}{\to} 1 - \frac{(\int_0^{a_i} r_i(x)dx)^2}{\int_0^{a_i} r_i^2(x)dx} \in (0,1),$$

so $e_i - c_i^2/b_i \stackrel{a.s.}{\to} \infty$.

The proof is complete.

Remark 1. *The lemma is true in the particular case $q_i(x) = x$, i.e. when $t_{ij} = j/m_i$ are equidistant.*

Set

$$\tilde{F}_A(a) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{A_i \le a\}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{A_{i1} \le a_1, \ldots, A_{ir} \le a_r\}}. \tag{51}$$

and let us consider the statistic

$$\hat{F}_A(a) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{\hat{A}_i \le a\}} \tag{52}$$

as the estimator of F_A. If the conditions of the Lemma 1 are verified then under the model 1 the difference $\hat{F}_A - \tilde{F}_A$ converges in probability to zero uniformly on $\mathcal{A}$ as $\min m_i \to \infty$.

Simulation results show that if the noise is not very large (i.e. the variance σ^2 is not large) then the estimators $\hat{F}_A$ and $\tilde{F}_A$ are close and for the model 2.

The estimator $\tilde{F}_A$ is the best estimator of the c.d.f. F_A. If m_i are not very small (1 model) or σ^2 is not large (models 1 and 2), then $\hat{F}_A$ is a good estimator of F_A and the c.d.f. of the real degradation at the moment t may be estimated by

$$\hat{F}_{r0}(z \mid t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{g(t,\hat{A}_i) \le z\}},$$

Let us consider estimation of Λ. Suppose at first that $\sigma = \alpha = 0$.

Denote by $\tau_i = T_i \wedge T_i^0$ the failure time of the ith individual, T_i, T_i^0 being the traumatic and non-traumatic failure moments. Set

$$X_i = \min(\tau_i, t_{i,m_i}), \quad \delta_i = \begin{cases} 1, & \text{if } X_i = T_i, \\ 0, & \text{otherwise.} \end{cases}$$

If $\sigma = 0$ and $j_i \geq r$ then the values of A_i can be found and the degradation value $g(T_i, A_i)$ just prior to the failure also can be found. Set

$$Z_i = g(X_i, A_i),$$

and on $[z_*, z_0]$ define the following counting processes:

$$N_i(z) = \mathbf{1}_{\{Z_i \leq z, \delta_i = 1\}}, \quad N(z) = \sum_{i=1}^{n} N_i(z).$$

$N(z)$ is the number of units having a failure before or at the moment when the degradation attains the level z.

Let $\mathcal{F}_z$ be the σ-algebra generated by the random variables

$$A_1, \ldots, A_n \quad \text{and} \quad N_1(y), \ldots, N_n(y) \quad \text{with} \quad y \leq z.$$

Then (see Bagdonavičius,V., Bikelis, A., Kazakevičius, A. and Nikulin, M. (2002, 2003))) the process $N_i(z)$ can be written as the sum

$$N_i(z) = \int_{z_*}^{z} \lambda(y) Y_i(y) dy + M_i(z), \tag{53}$$

where

$$Y_i(z) = \mathbf{1}_{\{Z_i \geq z\}} h'(z, A_i), \tag{54}$$

and M_i is a martingale with respect to the filtration $(\mathcal{F}_z \mid z \geq 0)$.

This implies that the optimal estimator of the cumulative intensity $\Lambda(z)$ is of Nelson-Aalen type (see Andersen, Borgan, Gill and Keiding (1993), pp.177-178):

$$\tilde{\Lambda}(z) = \int_{z_*}^{z} Y^{-1}(y)\, dN(y) = \sum_{Z_i \leq z, \delta_i = 1} Y^{-1}(Z_i) = \sum_{Z_i \leq z, \delta_i = 1} \left(\sum_{Z_j \geq Z_i} h'(z, A_i) \right)^{-1}, \tag{55}$$

where $Y(z) = \sum\limits_{j=1}^{n} Y_j(z)$.

The estimator is correctly defined if $Z_i \leq z$ and $\delta_i = 1$ for some i. If such i do not exist then the estimator is defined as 0.

To get asymptotic properties of $\hat{\Lambda}$ assumptions about the choice of t_{i,m_i} are needed. We assume that t_{i,m_i} are the realizations of (censoring) random variables C_i which are mutually independent, identically distributed and conditionally (given $A_i = a$) independent of (τ_i, δ_i). Let $S_c(t \mid a)$ denote the conditional survival function of C_i:

$$S_c(t \mid a) = P(C_i > t \mid A_i = a).$$

Two additional assumptions are needed. Set

$$c_0(a) = \sup_{z \leq z_0} h'(z, a), \quad c_1(a) = \sup_{z_1 < z_2 \leq z_0} \frac{\mid h'(z_2, a) - h'(z_1, a) \mid}{z_2 - z_1}$$

and

$$b(z) = E[h'(z, A_i) \mathbf{1}_{\{Z_i \geq z\}}] = \int h'(z, a) S_c(h(z, a) \mid a)\, S_T(h(z, a) \mid a)\, dF_A(a).$$

Assumption 1 $E[c_0(A) + c_1(A)] < \infty$.

Assumption 2 $\inf_{z \le z_0} b(z) > 0$.

Similarly as in Bagdonavicius et al (2002,2003) we obtain:

Theoreme 1 *Under Assumptions 1-2, the estimator $\hat{\Lambda}$ is uniformly consistent, i.e.*

$$\sup_{z \le z_0} \mid \hat{\Lambda}(z) - \Lambda(z) \mid \xrightarrow{P} 0,$$

as $n \to \infty$.

Theoreme 2 *If Assumptions 1-2 hold, then the random vector function $\sqrt{n}(\hat{\Lambda} - \Lambda)$ tends in distribution in the space $D[z_*; z_0]$ to a zero mean Gaussian process W with the covariance function*

$$E[W(z)W(z')] = \sigma^2(z \wedge z'),$$

where

$$\sigma^2(z) = \int_{z_*}^{z} \frac{\lambda(y)}{b(y)} dy.$$

The c.d.f F_A is estimated by $\tilde{F}_A$ given by (51). It is well-known that almost surely

$$\sup_a \mid \tilde{F}_A(a) - F_A(a) \mid \to 0,$$

as $n \to \infty$. Moreover, if the function F_A is absolutely continuous then the random function $\sqrt{n}(\hat{F}_A - F_A)$ tends in distribution in the Skorokhod space $D^r[0; \infty]$ to a zero mean Gaussian field W_0 with the covariance function

$$E[W_0(a)W_0(a')] = F_A(a \wedge a') - F_A(a)\, F_A(a');$$

where $a \wedge a' = (a_1 \wedge a'_1, \ldots, a_r \wedge a'_r)$.

Theoreme 3 *If Assumptions 1-2 hold and the function F_A is absolutely continuous then the random vector function $\sqrt{n}(\hat{F}_A - F_A,\, \hat{\Lambda} - \Lambda)$ tends in distribution in the space $D^r[0; \infty] \times D[z_*; z_0]$ to a random vector function with independent components.*

Estimating the survival and degradation characteristics we replace the functions F_A and Λ in the expressions of these characteristics by $\tilde{F}_A$ and $\hat{\Lambda}$. Asymptotic properties of such estimators can be easily found because they are smooth functionals of the pair $(\tilde{F}_A, \hat{\Lambda})$.

If $\alpha \neq 0$ then two step estimation of Λ is possible. If α would be known then similarly as in the case $\alpha \neq 0$ the following estimator of Λ could be used:

$$\Lambda^*(z, \alpha) = \int_{z_*}^{z} Y^{-1}(y, \alpha)\, dN(y) = \sum_{Z_i \le z, \delta_i = 1} Y^{-1}(Z_i, \alpha),$$

where $Y(z) = \sum_{j=1}^{n} Y_j(z)$, and

$$Y_j(z) = \mathbf{1}_{\{Z_j \ge z\}} \lambda_0(h(z, A_j), \alpha) h'(z, A_j).$$

If α is unknown, the estimator $\hat{\alpha}$ is obtained by maximising the loglikelihood function

$$\ln L(\alpha) \asymp \sum_{i=1}^{n} \delta_i \ln \lambda_0(X_i, \alpha) + \sum_{i=1}^{n} \delta_i \ln \lambda^*(Z_i, \alpha) -$$

$$\sum_{i=1}^{n} \int\limits_{g(0,A_i)}^{g(X_i,A_i)} \lambda_0(h(z, A_i))h'(z, A_i)\frac{dN(z)}{Y(z,\alpha)},$$

where

$$\lambda^*(z, \alpha) = \sum_{j=1}^{n} \frac{\delta_j \mathbf{1}_{\{z=Z_j\}}}{Y(z, \alpha)}.$$

The final estimator of Λ is

$$\tilde{\Lambda}(z) = \Lambda^*(z, \hat{\alpha}).$$

If $\sigma \neq 0$ then the values of stochastic processes $Z_i(t)$ instead of $g(t, A_i)$ are observed.

Using the values of $Z_i(t)$ the predictors $\hat{A}_i$ of the random variables A_i are found using the equations (48). Replacing A_i by their predictors $\hat{A}_i$ we obtain the following estimator of the cumulative intensity Λ:

$$\hat{\Lambda}(z) = \int_{z_*}^{z} \frac{d\hat{N}(y)}{\hat{Y}(y)} = \sum_{\hat{Z}_i \leq z, \delta_i = 1} \frac{1}{\hat{Y}(\hat{Z}_i)}, \tag{56}$$

where

$$\hat{Z}_i = g(X_i, \hat{A}_i), \quad \hat{N}(z) = \sum_{i=1}^{n} \hat{N}_i(z), \quad \hat{N}_i(z) = \mathbf{1}_{\{\hat{Z}_i \leq z, \delta_i = 1\}},$$

$$\hat{Y}(z) = \tilde{Y}(z, \hat{\alpha}), \quad \tilde{Y}(z, \alpha) = \sum_{i=1}^{n} \tilde{Y}_i(z, \alpha),$$

$$\tilde{Y}_i(z, \alpha) = \mathbf{1}_{\{\hat{Z}_i \geq z\}} \lambda_0(h(z, \hat{A}_i), \alpha)h'(z, \hat{A}_i),$$

and $\hat{\alpha}$ is obtained by maximizing the loglikelihood function

$$\ln L(\alpha) \asymp \sum_{i=1}^{n} \delta_i \ln \lambda_0(X_i, \alpha) + \sum_{i=1}^{n} \delta_i \ln \tilde{\lambda}(\hat{Z}_i, \alpha) -$$

$$\sum_{i=1}^{n} \int\limits_{g(0,\hat{A}_i)}^{g(X_i,\hat{A}_i)} \lambda_0(h(z, \hat{A}_i), \alpha)h'(z, \hat{A}_i)\frac{d\hat{N}(z)}{\tilde{Y}(z,\alpha)} =$$

$$\sum_{i=1}^{n} \delta_i \ln \lambda_0(X_i, \alpha) + \sum_{i=1}^{n} \delta_i \ln \tilde{\lambda}(\hat{Z}_i, \alpha) - \sum_{i=1}^{n} \sum_{\hat{Z}_j \leq g(X_i, \hat{A}_i), \delta_j = 1} \frac{\lambda_0(h(\hat{Z}_j, \hat{A}_i), \alpha)h'(\hat{Z}_j, \hat{A}_i)}{\tilde{Y}(\hat{Z}_j, \alpha)}, \tag{57}$$

where
$$\tilde\lambda(z,\alpha)=\sum_{j=1}^n\frac{\delta_j\mathbf{1}_{\{z=\hat Z_j\}}}{\tilde Y(z,\alpha)}.$$
Note that
$$h(z,A_j)=e^{-A_{1j}/A_{2j}}z^{1/A_{2j}}-1,\quad h'(z,A_j)=\frac{1}{A_{2j}}e^{-A_{1j}/A_{2j}}z^{1/A_{2j}-1}. \tag{58}$$

The formula (13) implies the following estimator of the probability $S_\tau(t)$ to survive time t:
$$\hat S_\tau(t)=\int\limits_{g(t,a)\le z_0}\hat S_T(t\mid a)d\hat F_A(a)=\frac1n\sum_{i:g(t,\hat A_i)\le z_0}\hat S_T(t\mid\hat A_i),$$
where
$$\hat S_T(t\mid\hat A_i)=\exp\Big\{-\int_{g(0,\hat A_i)}^{g(t,\hat A_i)}h'(z,\hat A_i)\,d\hat\Lambda(z)\Big\}=$$
$$\exp\{-\sum_{j=1}^n\mathbf{1}_{\{\delta_j=1,g(0,\hat A_i)\le\hat Z_j<g(t,\hat A_i)\}}\frac{h'(\hat Z_j,\hat A_i)}{\hat Y(\hat Z_j)}\}.$$
The probabilities $P^{(tr)}(t)$ and $P^{(0)}(t)$ are estimated by
$$\hat P^{(tr)}(t)=1-\frac1n\sum_{i=1}^n\hat S_T(t\wedge h(z_0,\hat A_i)\mid\hat A_i)) \tag{59}$$
and
$$\hat P^{(0)}(t)=\frac1n\sum_{g(t,\hat A_i)\ge z_0}\hat S_T(h(z_0,\hat A_i)\mid\hat A_i), \tag{60}$$
The c.d.f. of the real degradation at the moment t given the survival until this moment is estimated by
$$\hat F_r(z\mid t)=\frac{1}{\hat S_\tau(t)}\int\limits_{g(t,a)\le z}\hat S_T(t\mid a)d\hat F_A(a)=\frac1n\sum_{i:g(t,\hat A_i)\le z}\hat S_T(t\mid\hat A_i)/\hat S_\tau(t),$$
and the mean real degradation attained at the moment t given the survival until this moment is estimated by
$$\hat m_r(t)=\frac{1}{\hat S_\tau(t)}\int\limits_{\mathcal A}(g(t,a)\wedge z_0-z_*)\hat S_T(t\mid a)d\hat F_A(a)=\frac1n\sum_{i=1}^n(g(t,\hat A_i)\wedge z_0-z_*)\hat S_T(t\mid\hat A_i)/\hat S_T(t).$$
The probability that at the moment t the value of the real degradation will be in the set B given that at the moment s it is in the set B_0, is estimated by
$$\hat P(g(t,A)\in B\mid g(s,A)\in B_0,\tau>t)$$
$$=\int\limits_{g(t,a)\in B,g(s,a)\in B_0}\hat S_T(t\mid a)d\hat F_A(a)/\int\limits_{g(s,a)\in B_0}\hat S_T(t\mid a)d\hat F_A(a)=$$

$$\frac{\sum\limits_{i:g(t,\hat{A}_i)\in B, g(s,\hat{A}_i)\in B_0} \hat{S}_T(t \mid \hat{A}_i)}{\sum\limits_{i:g(s,\hat{A}_i)\in B_0} \hat{S}_T(t \mid \hat{A}_i)}.$$

The conditional survival function of $T(z_0)$ and the conditional mean given the survival until the degradation attains the level z_0 are estimated by are estimated by

$$\hat{P}(T^0(z) > t \mid \tau > T^0(z)) = \frac{\int\limits_{g(t,a)<z} \hat{S}_T(h(z,a) \mid a) d\hat{F}_A(a)}{\int\limits_{\mathcal{A}} \hat{S}_T(h(z,a) \mid a) d\hat{F}_A(a)} = \frac{\sum\limits_{i:g(t,\hat{A}_i)<z} \hat{S}_T(h(z,\hat{A}_i) \mid \hat{A}_i)}{\sum\limits_{i=1}^{n} \hat{S}_T(h(z,\hat{A}_i) \mid \hat{A}_i)},$$

and

$$\hat{E}(T^0(z) \mid \tau > T^0(z)) = \frac{\int\limits_{\mathcal{A}} h(z,a)\hat{S}_T(h(z,a) \mid a) d\hat{F}_A(a)}{\int\limits_{\mathcal{A}} \hat{S}_T(h(z,a) \mid a) d\hat{F}_A(a)} = \frac{\sum\limits_{i=1}^{n} h(z,\hat{A}_i)\hat{S}_T(h(z,\hat{A}_i) \mid \hat{A}_i)}{\sum\limits_{i=1}^{n} \hat{S}_T(h(z,\hat{A}_i) \mid \hat{A}_i)},$$

respectively.

6. Estimation of degradation and survival characteristics: models with covariates

Assume that the c.d.f. F_A and the intensities λ are completely unknown.

Suppose that n units are tested , and the ith of them is observed under the covariate $x^{(i)}$. Fix the moments of degradation measurements $t_{i1}, \ldots, t_{i,m_i}$ of this unit. If the failure time τ_i of the ith unit occurs in the interval $[t_{i,j_i}, t_{i,j_i+1})$ ($j_i = 1, \ldots, m_i; t_{m_i+1} = \infty$) then the values $Z_{i1}, \ldots, Z_{i,j_i}$ of the degradation process Z_i of the ith unit are observed at the time moments $t_{i1}, \ldots, t_{i,j_i}$.

Set

$$Y_{ij} = \ln Z_{ij}, \quad Y_i = (Y_{i1}, \ldots, Y_{i,j_i})^T.$$

Then given $A_i = a_i$

$$Y_i \sim N(\mu_i, \sigma^2\Sigma_i),$$

where

$$\mu_i = (\mu_{i1}, \ldots, \mu_{i,j_i})^T, \quad \mu_{ij} = \mu_{ij}(a_i, \beta) = \ln g(\varphi(t_{ij}, \beta, x^{(i)}), a_i),$$

Σ_i is given in (41). Denote by N the number of units such that $j_i \geq r$.

The estimators $\hat{\sigma}^2$, $\hat{\beta}$ of the parameters σ^2, β and the predictors $\hat{A}_i$ of the random vectors A_i are found by maximizing with respect to $\sigma^2, \beta, a_1, \ldots, a_n$ the conditional likelihood function

$$L(\sigma^2, \beta, a_1, \ldots, a_n) =$$

$$\frac{1}{(2\pi)^{m/2}\sigma^m} \prod_{i=1}^{N} \mid \Sigma_i \mid^{-1/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{N} (Y_i - \mu_i(a_i,\beta))^T \; \Sigma_i^{-1} \; (Y_i - \mu_i(a_i,\beta))\right\}. \quad (61)$$

Suppose that the function $g(\varphi(t,\beta,x),a)$ is differentiable with respect to β and a. The maximization of the likelihood function gives the following equations for $\hat\sigma^2$, $\hat\beta$, and $\hat A_i$ computing:

$$\sum_{k=1}^{j_i}\frac{\partial}{\partial A_i}\ln g(\varphi(t_{ik},\hat\beta,x^{(i)}),\hat A_i)\sum_{l=1}^{j_i}\{Y_{il}-\ln g(\varphi(t_{ij},\hat\beta,x^{(i)}),\hat A_i)\}b_{ikl}=0,$$

$$\sum_{i=1}^{N}\sum_{k=1}^{j_i}\frac{\partial}{\partial \beta}\ln g(\varphi(t_{ik},\hat\beta,x^{(i)}),\hat A_i)\sum_{l=1}^{j_i}\{Y_{il}-\ln g(\varphi(t_{ij},\hat\beta,x^{(i)}),\hat A_i)\}b_{ikl},$$

$$\hat\sigma^2=\frac{c(\hat\beta)}{m} \tag{62}$$

where

$$c(\hat\beta)=\sum_{i=1}^{N}(Y_i-\mu_i(\hat A_i,\hat\beta))\;\Sigma_i^{-1}\;(Y_i-\mu_i(\hat A_i,\hat\beta))^T.$$

Set

$$c_{ij}(\beta)=c(\varphi(t_{ij},\beta,x^{(i)})).$$

For the model (47) we have

$$\hat A_{i1}=\frac{c_i(\hat\beta)d_i(\hat\beta)-e_i(\hat\beta)f_i}{c_i^2(\hat\beta)-b_ie_i(\hat\beta)},\quad \hat A_{i2}=\frac{c_i(\hat\beta)f_i-b_id_i(\hat\beta)}{c_i^2(\hat\beta)-b_ie_i(\hat\beta)}, \tag{63}$$

$$\hat\sigma^2=\frac{c(\hat\beta)}{m}, \tag{64}$$

and $\hat\beta$ verifies the system of equations

$$\sum_{i=1}^{N}\hat A_{2i}\sum_{k=1}^{j_i}\psi_{iks}(\hat\beta)\sum_{l=1}^{j_i}\{Y_{il}-\hat A_{1i}-\hat A_{2i}c_{il}(\hat\beta)\}b_{ikl}=0,$$

where

$$c(\hat\beta)=\sum_{i=1}^{n}(g_i+b_i\hat A_{1i}^2+e_i(\hat\beta)\hat A_{2i}^2+2c_i(\hat\beta)\hat A_{1i}\hat A_{2i}-2f_i\hat A_{1i}-2d_i(\hat\beta)\hat A_{2i})$$

$$c_i(\hat\beta)=C_i^T(\hat\beta)\Sigma_i^{-1}\mathbf{1},\quad d_i(\hat\beta)=Y_i^T\Sigma_i^{-1}C_i(\hat\beta),$$

$$e_i(\hat\beta)=C_i^T(\hat\beta)\Sigma_i^{-1}C_i(\hat\beta),\quad \psi_{iks}(\hat\beta)=\frac{\int_0^{t_{ij}}x_s^{(i)}(u)e^{x^{(i)}(u)\hat\beta^T}du}{1+\int_0^{t_{ij}}e^{x^{(i)}(u)\hat\beta^T}du};$$

$$\mathbf{1}=(1,\ldots,1)_{j_i}^T,\quad C_i(\hat\beta)=(c_{i1}(\hat\beta),\ldots,c_{i,j_i}(\hat\beta))^T, \tag{65}$$

and b_i,f_i,g_i are given by (50).

If the vector of covariates $x^{(i)}$ is a step-function , and has the form

$$x^{(i)}(t)=x^{(ik)},\quad \text{if}\quad t\in[t^*_{i,k-1},t^*_{ik})\quad (k=1,\ldots l_i,\ t^*_{i0}=0),$$

then

$$\psi_{iks}(\hat{\beta}) = \frac{\sum\limits_{l=1}^{\nu_{ik}-1} x_s^{(il)} e^{x^{(ik)}\hat{\beta}^T}(t^*_{il} - t^*_{i,l-1}) + x_s^{(i,\nu_{ik})} e^{x^{(i\nu_{ik})}\hat{\beta}^T}(t^*_{i,\nu_{ik}} - t^*_{i,\nu_{ik}-1})}{1 + \sum\limits_{k=1}^{\nu_{ij}-1} e^{x^{(ik)}\hat{\beta}^T}(t^*_{ik} - t^*_{i,k-1}) + e^{x^{(i\nu_{ik})}\hat{\beta}^T}(t^*_{i,\nu_{ik}} - t^*_{i,\nu_{ik}-1})},$$

where $\nu_{ik} = \{l : t_{ik} \in [t^*_{i,l-1}, t^*_{il})\}$.

In particular, if $x^{(i)}$ has only one jump ($l_i = 2$) then

$$\psi_{iks}(\hat{\beta}) = \frac{x_s^{(i1)} e^{x^{(i1)}\beta^T} \min(\tau_i, t^*_{ik}) + x_s^{(i2)} e^{x^{(i2)}\beta^T} \max(t^*_{ik} - \tau_i, 0)}{1 + e^{x^{(i1)}\beta^T} \min(\tau_i, t^*_{ik}) + e^{x^{(i2)}\beta^T} \max(t^*_{ik} - \tau_i, 0)}.$$

The estimator of the c.d.f. $F_A(a)$ has the form (52).

The estimators of the c.d.f. and the mean of the real degradation at the moment t given the covarate $x = x(s), 0 \le s \le t$ is

$$\hat{F}_{r0}(z \mid t, x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{g(\varphi(t,\hat{\beta},x),\hat{A}_i) \le z\}}$$

the mean real degradation attained at the moment t is estimated by

$$\hat{m}_{r0}(t, x) = \frac{1}{n} \sum_{i=1}^{n} g(\varphi(t, \hat{\beta}, x), \hat{A}_i).$$

Similarly as in the case without covariates we define the estimators

$$\hat{\Lambda}(z) = \tilde{\Lambda}(z, \hat{\alpha}),$$

where

$$\tilde{\Lambda}(z, \alpha) = \int_0^z \hat{Y}^{-1}(y, \alpha)\, d\hat{N}(y) = \sum_{\hat{Z}_i \le z, \delta_i = 1} \tilde{Y}^{-1}(\hat{Z}_i, \alpha),$$

$$\hat{Z}_i = g(\varphi(X_i, \hat{\beta}, x^{(i)}), \hat{A}_i), \quad \hat{N}(z) = \sum_{i=1}^{N} \hat{N}_i(z), \quad \hat{N}_i(z) = \mathbf{1}_{\{\hat{Z}_i \le z, \delta_i = 1\}},$$

$$\tilde{Y}(z, \alpha) = \sum_{i=1}^{N} \tilde{Y}_i(z, \alpha), \quad \tilde{Y}_i(z, \alpha) = \mathbf{1}_{\{\hat{Z}_i \ge z\}} \lambda_0(h(z, \hat{A}_i), \alpha) \psi'(h(z, \hat{A}_i), \hat{\beta}, x^{(i)}) h'(z, \hat{A}_i),$$

and $\hat{\alpha}$ is obtained by maximizing the loglikelihood function

$$\ln L(\alpha) \asymp \sum_{i=1}^{N} \delta_i \ln \lambda_0(\varphi(X_i, \hat{\beta}, x^{(i)}), \alpha) + \sum_{i=1}^{n} \delta_i \ln \tilde{\lambda}(\hat{Z}_i, \alpha) -$$

$$\sum_{i=1}^{N} \int\limits_{g(\varphi(0,\hat{\beta},x^{(i)},\hat{A}_i)}^{g(\varphi(X_i,\hat{\beta},x^{(i)},\hat{A}_i)} \lambda_0(h(z, \hat{A}_i), \alpha) \psi'(h(z, \hat{A}_i), \hat{\beta}, x^{(i)}) h'(z, \hat{A}_i) \frac{d\hat{N}(z)}{\tilde{Y}(z, \alpha)} =$$

$$\sum_{i=1}^{N} \delta_i \ln \lambda_0(\varphi(X_i, \hat\beta, x^{(i)}), \alpha) + \sum_{i=1}^{n} \delta_i \ln \tilde\lambda(\hat Z_i, \alpha)$$

$$-\sum_{i=1}^{N} \sum_{g(\varphi(0,\hat\beta,x^{(i)},\hat A_i) \geq \hat Z_j \leq g(\varphi(X_i,\hat\beta,x^{(i)},\hat A_i),\delta_j=1} \frac{\lambda_0(h(\hat Z_j, \hat A_i), \alpha)\psi'(h(\hat Z_j, \hat A_i), \hat\beta, x^{(i)})h'(\hat Z_j, \hat A_i)}{\tilde Y(\hat Z_j, \alpha)}, \tag{66}$$

where

$$\tilde\lambda(z, \alpha) = \sum_{j=1}^{N} \frac{\delta_j \, \mathbf{1}_{\{z=\hat Z_j\}}}{\tilde Y(z, \alpha)}.$$

The estimators of the survival function

$$\hat S_\tau(t \mid x) = \frac{1}{n} \sum_{ig(\varphi(t,\hat\beta,x)\hat A_i) < z_0} \hat S_T(t \mid \hat A_i, x), \tag{67}$$

where

$$\hat S_T(t \mid \hat A_i, x) =$$

$$\exp\{- \sum_{\delta_j=1, g(\varphi(0,\hat\beta,x),\hat A_i) \leq \hat Z_j < g(\varphi(t,\hat\beta,x),\hat A_i)} \lambda_0(h(\hat Z_j, \hat A_i), \hat\alpha) \, \psi'(h(\hat Z_j, \hat A_i), \hat\beta, x) \, \frac{h'(\hat Z_j, \hat A_i)}{\hat Y(\hat Z_j)}\}.$$

The c.d.f. and the mean of the real degradation at the moment t given the survival until this moment are estimated by

$$\hat F_r(z \mid t, x) = \frac{1}{N} \sum_{i:g(\varphi(t,\hat\beta,x),\hat A_i) \leq z} \hat S_T(t \mid \hat A_i, x)/\hat S_T(t \mid x), \tag{68}$$

$$\hat m_r(t, x) = \frac{1}{N} \sum_{i=1}^{n} g(\varphi(t, \hat\beta, x), \hat A_i) \hat S_T(t \mid \hat A_i, x)/\hat S_T(t \mid x). \tag{69}$$

The probabilities $P^{(tr)}(t \mid x)$ and $P^{(0)}(t \mid x)$ are estimated by

$$\hat P^{(tr)}(t \mid x) = 1 - \frac{1}{N} \sum_{i=1}^{N} \hat S_T(t \wedge \psi(h(z_0, \hat A_i), \hat\beta, x)) \mid \hat A_i, x), \tag{70}$$

$$\hat P^{(0)}(t \mid x) = \frac{1}{N} \sum_{g(\varphi(t,\hat\beta,x)),\hat A_i) \geq z_0} S_T(\psi(h(z_0, \hat A_i), \hat\beta, x)) \mid \hat A_i, x). \tag{71}$$

The probability that at the moment t the value of the real degradation will be in the set B given that at the moment s it is in the set B_0 and given $x(u), 0 \leq u \leq t$, is estimated by

$$\hat P(g(\varphi(t, \beta, x), a) \in B \mid g(\varphi(s, \beta, x), a) \in B_0, \tau > t, x(u), 0 \leq u \leq t)$$

$$= \frac{\sum_{i:g(\varphi(t,\hat\beta,x),\hat A_i) \in B, g(\varphi(s,\hat\beta,x),\hat A_i) \in B_0} \hat S_T(t \mid \hat A_i, x)}{\sum_{i:g(\varphi(s,\hat\beta,x),\hat A_i) \in B_0} \hat S_T(t \mid \hat A_i, x)}. \tag{72}$$

The conditional survival function of $T^0(z,x)$ given the survival until the degradation attains the level z and the surviving until the moment t_1, and given $x(u), 0 \leq u \leq t$, is estimated by

$$\hat{P}(T^0(z) > t \mid \tau > T^0(z), x(u), 0 \leq u \leq t) = \frac{\sum\limits_{i: g(\varphi(t,\hat{\beta},x),\hat{A}_i) < z} \hat{S}_T(\psi(h(z,\hat{A}_i),\hat{\beta},x) \mid \hat{A}_i, x)}{\sum\limits_{i=1}^{n} \hat{S}_T(\psi(h(z,\hat{A}_i),\hat{\beta},x) \mid \hat{A}_i, x)}, \tag{73}$$

and

$$\hat{E}(T^0(z) \mid \tau > T(z_0), x(u), 0 \leq u \leq t) = \frac{\sum\limits_{i=1}^{n} \psi(h(z,\hat{A}_i),\hat{\beta},x)\hat{S}_T(\psi(h(z,\hat{A}_i),\hat{\beta},x) \mid \hat{A}_i, x)}{\sum\limits_{i=1}^{n} \hat{S}_T(\psi(h(z,\hat{A}_i),\hat{\beta},x) \mid \hat{A}_i, x)},$$

respectively.

References

Andersen,P.K., Borgan,O., Gill,R. and Keiding,N. (1993). *Statistical models based on the counting processes*. Springer-Verlag: New York.

Bagdonavičius, V. and Nikulin, M. (2001). Estimation in Degradation Models with Explanatory variables, *Lifetime Data Analysis*, **7**, #1, 85-103.

Bagdonavičius, V. and Nikulin, M. (2002). *Accelerated Life Models*. Chapman & Hall/CRC: Boca Raton.

Bagdonavičius, V., Bikelis, A., Kazakevičius, A. and Nikulin, M. (2002). Non-parametric estimation from simultaneous degradation and failure data, *Comptes Rendus, Academie des Sciences de Paris*, **v. 335**, 183-188.

Bagdonavičius,V., Bikelis, A., Kazakevičius, A. and Nikulin, M. (2003). Estimation from simultaneous degradation and failure time data. In : *Mathematical and Statistical Methods in Reliability.*(B. Lindqvist and Kjell A Doksum, eds.), *World Scientific Publishing*, Series on Quality,Reliability and Engineering Statistics, 7.

Barberger-Gateau P, Rainville C, Letenneur L, Dartigues J.F. (2000). A hierarchical model of domains of disablement in the elderly : a longitudinal approach. *Disability and Rehabilitation*, 22, 308-317.

CTNERHI-INSERM : *Classification Internationale des Handicaps : déficiences, incapacités et désavantages.* (1988). Paris, PUF.

Gavrilov, L.A., Gavrilova, N.S. (2003). The Quest for a General Theory of Aging and Longevity. *SAGE KE: Sciences of Aging Knowledge Environment*, p.1-10 (a review) (sageke.sciencemag./cgi/content/full/sageke;2003/28/re5)

Katz S, Downs T, Cash H, Grotz R. (1970). Progress in development of the index of ADL. *The Gerontologist*, 10, 20-30.

Lawton M, Brody E. Assesment of older people : self maintaining and instrumental activities of daily living. (1969). *The Gerontologist*, 9 : 179-86.

Meeker, W.Q. and Escobar, L. (1998). *Statistical Analysis for Reliability Data.* John Wiley and Sons: New York.

Rosow I, Breslau N. A Guttman health scale for the aged. (1966). *Journal of Gerontology*, 21 : 556-9.

Wulfsohn, M. and Tsiatis, A. (1997). A Joint Model for Survival and Longitudinal Data Measured with Error. *Biometrics*, **53**, 330-339.

Zdorova-Cheminade, O. (2003) *Modélisation du processus d'évolution de l'incapacité chez les personnes agées*, Mémoire de DEA "Epidémiologie et Intervention en Santé Publique", Université Bordeaux 2, Juin 2003, 77 pages.