KYBERNETIKA — VOLUME *43* (2007), NUMBER 1, PAGES 61 – 74

M–ESTIMATION IN NONLINEAR REGRESSION FOR LONGITUDINAL DATA

Martina Orsáková

The longitudinal regression model $Z_i^j = m(\theta_0, \mathbb{X}_i(T_i^j)) + \varepsilon_i^j$, where Z_i^j is the jth measurement of the ith subject at random time T_i^j, m is the regression function, $\mathbb{X}_i(T_i^j)$ is a predictable covariate process observed at time T_i^j and ε_i^j is a noise, is studied in marked point process framework. In this paper we introduce the assumptions which guarantee the consistency and asymptotic normality of smooth M-estimator of unknown parameter θ_0.

Keywords: M-estimation, nonlinear regression, longitudinal data

AMS Subject Classification: 62M10, 62F10, 60G55

1. INTRODUCTION

One branch of statistical methodology dealing with repeated measurements are longitudinal studies. The distinguishing feature of a longitudinal data is that the outcome of interest is measured repeatedly over time in the same subjects, with the general objective of characterizing change in the outcome over time, and studying factors which contribute to the mean level and to the change. We deal with longitudinal studies where the number of subjects is generally large relative to the number of time points. We consider parametric transitions model (see Diggle, Heagerty, Liang and Zeger [5]), where the objective of the analysis is to characterize the conditional mean of the current response given past outcomes, as a function of time, as well as covariates, the interaction of time and these covariates.

We will not be speaking about the studies which either treat times of measurements as fixed by the study design or, in an observational setting, assume that the measurement times are unrelated to the interest, and from a statistical modeling perspective can therefore be treated as if they had been fixed in advance. We will consider an event that can occur at irregularly spaced intervals in time, known as waiting times, such that no two events occur simultaneously. The sequence of events is known as a point process. The number of events in given time intervals, the waiting time between successive events, or, more generally, the intensity of occurrence of events may be of interest. We will measure a variable at each time an event occurs, yielding a marked point process. In this study we deal with regression model with repeated measurements in time of event occurrence, where the regression function is known function with unknown parameters.

This paper is an applied work based on theory of marked point processes. We have followed closely the approach of Scheike [23] and extended it for M-estimators. Scheike derived a consistent and asymptotic normal estimator for the unknown parameter in nonlinear regression model for longitudinal data with counting process measurement times. He used weighted least square method. The standard least square method is unstable if outliers are present in the data. Outlying data can distort an parameter estimation. The M-estimators try to reduce the effect of outliers by replacing the square function by another one.

The plan is as follows. Section 2 presents a model more precisely and the assumptions are enumerated. Section 3 contains proofs of consistency and asymptotic normality of rescaled M-estimator. Finally the properties of M-estimator are shown for the case of unknown conditional variance of the noise.

2. NOTATION, MODEL, ASSUMPTIONS

2.1. Model specification

Let $(\Omega, \mathcal{A}, P)$ be a probability space. (In what follows all $o_p(.)$ and $O_p(.)$ are understood with respect to this P.) We shall consider the nonlinear regression model

$$Z_i^j = m(\theta_0, \mathbb{X}_i(T_i^j)) + \varepsilon_i^j, \quad i = 1, \dots, n, \quad j = 1, \dots, N_t^i, \tag{1}$$

where Z_i^j is the jth measurement on the ith subject at random time T_i^j in the interval of observation $[0, t]$, $m(\theta_0, .)$ is a known regression function depending on an unknown d-dimensional parameter θ_0, $\mathbb{X}_i(T_i^j)$ a p-dimensional covariate process of subject i at time T_i^j and ε_i^j is noise. Let $\mathcal{B}$ denote the Borel σ-field on $\mathbb{R}$. For $A \in \mathcal{B}$, define the counting process

$$N_s^i(A) = \sum_j I(Z_i^j \in A) I(T_i^j \le s)$$

and the associated marked point process $P^i(\mathrm{d}s \times \mathrm{d}z)$

$$P^i([0, s] \times A) = N_s^i(A), \ s \ge 0, \ A \in \mathcal{B}.$$

Define further the history of the subjects, that is, the history of the marked point processes, as

$$\mathcal{F}_u = \sigma(N_s^i(A) : \ s \le u, \ A \in \mathcal{B}, \ i = 1, \dots, n) \vee \mathcal{A}.$$

The σ-algebra $\mathcal{A}$ is independent of the former one and represents knowledge prior to time 0. We further need to define the σ-field $\mathcal{F}_{T_i^j-} = \sigma((Z_i^m, T_i^m) : \ T_i^m < T_i^j; \ T_i^j) \vee \mathcal{A}$ that contains the information just prior to observation of a jump size.

Define further

$$N_s^i = N_s^i(\mathbb{R}),$$

the counting process associated with the T_i^j's. It is assumed that no two of the counting processes N_s^i jump at the same time. We assume that $\mathbb{X}_i(s)$ is predictable with respect to the history $\mathcal{F}_s$ and that N_s^i has a random intensity $\lambda_s^i > 0$, that is, $\lambda_s^i\,\mathrm{d}s$ is the probability of a jump in the time interval $(s, s+\mathrm{d}s]$. Aalen's (1975,1978) multiplicative model $\lambda_s^i = \alpha(s)Y^i(s)$ (with $\alpha(s)$ deterministic and $Y^i(s)$ predictable) is often used in applications.

Finally, it is assumed that the noise terms have conditional mean and variance given by

$$\mathrm{E}\,(\varepsilon_i^j|\mathcal{F}_{T_i^j-}) = 0,$$
$$\mathrm{E}\,((\varepsilon_i^j)^2|\mathcal{F}_{T_i^j-}) = \sigma^2(\mathbb{X}_i(T_i^j)),$$

where $\sigma^2(.)$ is deterministic, continuous and bounded. The conditional distribution of $\frac{\varepsilon_i^j}{\sigma(\mathbb{X}_i(T_i^j))}$ is denoted

$$F_{T_i^j}(z) = P\left(\frac{\varepsilon_i^j}{\sigma(\mathbb{X}_i(T_i^j))} \le z\middle|\mathcal{F}_{T_i^j-}\right),$$

therefore, the conditional distribution of Z_i^k is

$$P(Z_i^j \le z|\mathcal{F}_{T_i^j-}) = F_{T_i^j}\left(\frac{z - m(\theta_0, \mathbb{X}_i(T_i^j))}{\sigma(\mathbb{X}_i(T_i^j))}\right).$$

The rescaled M-estimator of the parameter θ_0 is defined as a minimizer of

$$\begin{aligned} L_n(\theta,t) &:= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}\rho\left(\frac{Z_i^j - m(\theta, \mathbb{X}_i(T_i^j))}{\sigma(\mathbb{X}_i(T_i^j))}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\int_0^t\int H_i(\theta,s,z)P^i(\mathrm{d}s\times\mathrm{d}z), \end{aligned} \tag{2}$$

where $\rho(.)$ is a real continuous function and

$$H_i(\theta,s,z) = \rho\left(\frac{z - m(\theta, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right).$$

In the case of differentiability of the respective functions, the estimator $\hat\theta$ will be preferred, defined as a solution to the equations

$$\frac{\partial}{\partial\theta_k}L_n(\theta,t) = 0,\ \ k = 1,\dots,d \tag{3}$$

that is

$$\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}\psi\left(\frac{Z_i^j - m(\theta, \mathbb{X}_i(T_i^j))}{\sigma(\mathbb{X}_i(T_i^j)))}\right)\cdot\left(\frac{\partial}{\partial\theta_k}m(\theta, \mathbb{X}_i(T_i^j))\right)\cdot\frac{1}{\sigma(\mathbb{X}_i(T_i^j))} = 0,\ \ k = 1,\dots,d,$$

where $\psi(.) = \rho'(.)$. We make the natural assumption that

$$\begin{aligned} &\mathrm{E}\left(\psi\left(\frac{\varepsilon_i^j}{\sigma(\mathbb{X}_i(T_i^j))}\right)\middle|\mathcal{F}_{T_i^j-}, T_i^j = s\right) \\ \int\psi\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}F_s\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right) &= 0. \end{aligned} \tag{4}$$

Before giving the asymptotic results about the estimator of θ_0, some further conditions are stated.

2.2. Assumptions

To facilitate reading we divided the assumptions into several groups.

(A) Let $m(\theta, x)$ be three times differentiable in a neighborhood $B(\theta_0, K)$ around parameter θ_0. Assume that $\psi(.)$ is two times differentiable and $\psi''(.)$ is absolutely continuous function. Assume further

$$\mathrm{E}\left(\int_0^t \int \left[\frac{\partial}{\partial\theta_k} H_i(\theta_0, s, z)\right]^2 \lambda_s^i \mathrm{d}F_s\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s\right) < \infty,$$

$$\mathrm{E}\left(\int_0^t \int \left[\frac{\partial^2}{\partial\theta_k\partial\theta_l} H_i(\theta_0, s, z)\right]^2 \lambda_s^i \mathrm{d}F_s\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s\right) < \infty$$

for all $k, l = 1, \dots d$.

(B) Let us set

$$\gamma_1(\theta_0, \mathbb{X}_i(s)) := \int \psi'\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}F_s\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right),$$

$$\gamma_2(\theta_0, \mathbb{X}_i(s)) := \int \left[\psi'\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\right]^2\mathrm{d}F_s\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right),$$

$$\gamma_3(\theta_0, \mathbb{X}_i(s)) := \int \left[\psi\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\right]^2\mathrm{d}F_s\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)$$

and assume that

$$\gamma_2(\theta_0, \mathbb{X}_i(s)) < \infty,$$
$$\gamma_3(\theta_0, \mathbb{X}_i(s)) < \infty.$$

Further we assume that there exist non-negative definite symmetric matrices Σ_I, Σ_U such that as $n \to \infty$

$$\left(\frac{1}{n}\sum_{i=1}^n \int_0^t m_k'(\theta_0, \mathbb{X}_i(s)) \cdot m_l'(\theta_0, \mathbb{X}_i(s)) \cdot \frac{\lambda_s^i}{\sigma^2(\mathbb{X}_i(s))} \cdot \gamma_1(\theta_0, \mathbb{X}_i(s))\mathrm{d}s\right)_{k,l=1}^d \xrightarrow[n\to\infty]{p} \Sigma_I, \quad (5)$$

$$\left(\frac{1}{n}\sum_{i=1}^n \int_0^t m_k'(\theta_0, \mathbb{X}_i(s)) \cdot m_l'(\theta_0, \mathbb{X}_i(s)) \cdot \frac{\lambda_s^i}{\sigma^2(\mathbb{X}_i(s))} \cdot \gamma_3(\theta_0, \mathbb{X}_i(s))\mathrm{d}s\right)_{k,l=1}^d \xrightarrow[n\to\infty]{p} \Sigma_U \quad (6)$$

and

$$\frac{1}{n^2}\sum_{i=1}^{n}\int_0^t [m_k'(\theta_0,\mathbb{X}_i(s))\cdot m_l'(\theta_0,\mathbb{X}_i(s))]^2\cdot\lambda_s^i\cdot\frac{1}{\sigma^4(\mathbb{X}_i(s))}\cdot\gamma_2(\theta_0,\mathbb{X}_i(s))\,\mathrm{d}s$$
$$+\frac{1}{n^2}\sum_{i=1}^{n}\int_0^t [m_{k,l}''(\theta_0,\mathbb{X}_i(s))]^2\cdot\lambda_s^i\cdot\frac{1}{\sigma^2(\mathbb{X}_i(s))}\cdot\gamma_3(\theta_0,\mathbb{X}_i(s))\,\mathrm{d}s \xrightarrow[n\to\infty]{p} 0$$

for all $k,l=1,\ldots,d$.

(C) Let G be a such function that for $\theta\in B(\theta_0,r)$

$$\left|\frac{\partial^3}{\partial\theta_j\partial\theta_k\partial\theta_l}H_i(\theta,t,z)\right|\le G(t,z),$$
$$\left(\frac{1}{n}\sum_{i=1}^{n}\int_0^t\int G(s,z)\cdot\lambda_s^i\cdot\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s\right)=O_p(1),$$
$$\left(\frac{1}{n^2}\sum_{i=1}^{n}\int_0^t\int G^2(s,z)\cdot\lambda_s^i\cdot\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s\right)\xrightarrow[n\to\infty]{p} 0.$$

3. PROPERTIES OF M–ESTIMATION

3.1. Properties of M-estimation for known $\sigma^2(.)$

Theorem 3.1. Assume that the assumptions (A), (B), (C) are satisfied. Then there exists a consistent solution of (3) (a sequence of the solutions of the equations (3), $\{\widehat{\theta}_n\}$, such that $||\widehat{\theta}_n-\theta_0||=o_p(1)$ as $n\to\infty$) and provides a local minimum of (2) with probability tending to one.

Proof. (The proof follows as in Theorem 1 of Scheike [23].) Write the system of equations (3) as $U_n(\theta,t)=0$ and make a Taylor expansion for $\theta\in B(\theta_0,r)$ around true value θ_0

$$\begin{aligned}U_n(\theta,t) &= U_n(\theta_0,t)+d_{\theta-\theta_0}U_n(\theta_0,t)+\frac{1}{2}\cdot d^2_{\theta-\theta_0}U_n(\theta^*,t)\\ &= U_n(\theta_0,t)+(\theta-\theta_0)^T\cdot I_n(\theta_0,t)+\frac{1}{2}\cdot d^2_{\theta-\theta_0}U_n(\theta^*,t)\end{aligned}$$

where $\theta^*\in$ line (θ,θ_0) (line (θ,θ_0) denotes the line segment between θ and θ_0), $I_n(\theta,t):=\left(\frac{\partial^2}{\partial\theta_k\partial\theta_l}L_n(\theta_0,t)\right)_{k,l=1}^d$ and $d^2_{\theta-\theta_0}U_n(\theta^*,t)=\sum_{k_1\ge0,\ \ldots k_d\ge0}^{k_1+\ \ldots k_d=2}\frac{1}{k_1!\ldots k_d!}\frac{\partial^2U_n(\theta^*,t)}{\partial\theta_1^{k_1}\ldots\partial\theta_d^{k_d}}$ $(\theta_1-\theta_{01})^{k_1}\ldots(\theta_d-\theta_{0d})^{k_d}$.

To prove the proposition, the plan is to show that

$$U_n(\theta_0, t) \xrightarrow[n\to\infty]{p} 0, \tag{7}$$

$$I_n(\theta_0, t) \xrightarrow[n\to\infty]{p} \Sigma_I, \tag{8}$$

$$\frac{\partial^2 U_n(\theta,t)}{\partial\theta_1^{k_1}\ldots\partial\theta_d^{k_d}} = O_p(1), \quad k_1 \geq 0, \ \ldots k_d \geq 0, \quad k_1 + \ \ldots k_d = 2 \tag{9}$$

for all $\theta \in B(\theta_0, r)$. Then it follows that $\widehat{\theta}_n$ provides a local minimum with probability tending to one, and is a consistent solution of (3).

To show (7) we need to establish that the process $U_n(\theta_0, t)$ is a martingale. Note that due to the assumption (4) you can write

$$\begin{aligned} &U_n(\theta_0, t)\\ = & -\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}\psi\left(\frac{Z_i^j - m(\theta_0, \mathbb{X}_i(T_i^j))}{\sigma(\mathbb{X}_i(T_i^j))}\right)\cdot\frac{1}{\sigma(\mathbb{X}_i(T_i^j))}\cdot m_k'(\theta_0, \mathbb{X}_i(T_i^j))\right)_{k=1}^{d}\\ = & -\frac{1}{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}\psi\left(\frac{Z_i^j - m(\theta_0, \mathbb{X}_i(T_i^j))}{\sigma(\mathbb{X}_i(T_i^j))}\right)\cdot\frac{1}{\sigma(\mathbb{X}_i(T_i^j))}\cdot m_k'(\theta_0, \mathbb{X}_i(T_i^j))\right.\\ &\left. -\sum_{i=1}^{n}\int_0^t\int\psi\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\cdot\frac{m_k'(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\cdot\lambda_s^i \mathrm{d}F_s\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s\right)_{k=1}^{d}. \end{aligned}$$

According to martingale transform theorem (see Boel, Varaiya, Wong [4], p. 1010) $U_n(\theta_0, t)$ is a martingale with $\mathrm{E}(U_n(\theta_0, t)) = 0$ (define further 0-martingale to be a martingale with mean 0) with predictable quadratic variation process

$$\begin{aligned} &\frac{1}{n^2}\sum_{i=1}^{n}\int_0^t\int\left(\frac{\partial}{\partial\theta_k}H_i(\theta_0, s, z)\right)\cdot\left(\frac{\partial}{\partial\theta_l}H_i(\theta_0, s, z)\right)\cdot\lambda_s^i\,\mathrm{d}F_s\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s\\ = &\frac{1}{n^2}\sum_{i=1}^{n}\int_0^t\int\psi^2\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\cdot\frac{m_k'(\theta_0, \mathbb{X}_i(s))\cdot m_l'(\theta_0, \mathbb{X}_i(s))}{\sigma^2(\mathbb{X}_i(s))}\\ &\cdot\lambda_s^i\,\mathrm{d}F_s\left(\frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s\\ = &\frac{1}{n^2}\sum_{i=1}^{n}\int_0^t\gamma_3(\theta_0, \mathbb{X}_i(s))\cdot\frac{m_k'(\theta_0, \mathbb{X}_i(s))\cdot m_l'(\theta_0, \mathbb{X}_i(s))}{\sigma^2(\mathbb{X}_i(s))}\cdot\lambda_s^i\,\mathrm{d}s \xrightarrow[n\to\infty]{p} 0, \end{aligned}$$

by assumption (B). Then Lenglart's inequality, see Jacod and Shiryaev [9], p. 35, gives that

$$\sup_{s\leq t}|U_n(\theta_0, s)| \xrightarrow[n\to\infty]{p} 0 \tag{10}$$

and (7) follows.
To establish (8), write

$$I_n(\theta_0,t) = \frac{1}{n}\sum_{i=1}^{n}\Bigg(M^i\left(\frac{\partial^2}{\partial\theta_k\partial\theta_l}L_n(\theta_0,s)\right)_t + \int_0^t\int \frac{\partial^2}{\partial\theta_k\partial\theta_l}L_n(\theta_0,s)\cdot\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s\Bigg)_{k,l=1}^{d} \qquad (11)$$

where

$$M^i\left(\frac{\partial^2}{\partial\theta_k\partial\theta_l}L_n(\theta_0,s)\right)_t = \frac{\partial^2}{\partial\theta_k\partial\theta_l}L_n(\theta_0,t) - \int_0^t\int \frac{\partial^2}{\partial\theta_k\partial\theta_l}L_n(\theta_0,s)\cdot\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s.$$

Note that the first term in (11) is a 0-martingale with a compensator that is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\int_0^t\int \psi'\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\cdot\frac{m_k'(\theta_0,\mathbb{X}_i(s))\cdot m_l'(\theta_0,\mathbb{X}_i(s))}{\sigma^2(\mathbb{X}_i(s))} \cdot\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s$$
$$-\frac{1}{n}\sum_{i=1}^{n}\int_0^t\int \psi\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\cdot\frac{m_{kl}''(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\cdot\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s$$

and quadratic predictable variation process that is given as

$$\frac{1}{n^2}\sum_{i=1}^{n}\int_0^t\int \left[\psi'\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\right]^2\cdot\frac{[m_k'(\theta_0,\mathbb{X}_i(s))\cdot m_l'(\theta_0,\mathbb{X}_i(s))]^2}{\sigma^4(\mathbb{X}_i(s))} \cdot\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s$$
$$+\frac{1}{n^2}\sum_{i=1}^{n}\int_0^t\int \left[\psi\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\right]^2\cdot\frac{[m_{kl}''(\theta_0,\mathbb{X}_i(s))]^2}{\sigma^2(\mathbb{X}_i(s))} \cdot\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s$$
$$= \frac{1}{n^2}\sum_{i=1}^{n}\int_0^t \gamma_2(\theta_0,\mathbb{X}_i(s))\cdot\frac{1}{\sigma^4(\mathbb{X}_i(s))}\cdot[m_k'(\theta_0,\mathbb{X}_i(s))\cdot m_l'(\theta_0,\mathbb{X}_i(s))]^2\cdot\lambda_s^i\,\mathrm{d}s$$
$$+\frac{1}{n^2}\sum_{i=1}^{n}\int_0^t \gamma_3(\theta_0,\mathbb{X}_i(s))\cdot\frac{1}{\sigma^2(\mathbb{X}_i(s))}\cdot[m_{kl}''(\theta_0,\mathbb{X}_i(s))]^2\cdot\lambda_s^i\,\mathrm{d}s \xrightarrow[n\to\infty]{p} 0$$

by assumption (B). So by Lenglarts's inequality the first term of (11) converges in probability to zero. The second term is equal to

$$\left(\frac{1}{n}\sum_{i=1}^{n}\int\limits_0^t \gamma_1(\theta_0,\mathbb{X}_i(s))\cdot\frac{1}{\sigma^2(\mathbb{X}_i(s))}\cdot m_k'(\theta_0,\mathbb{X}_i(s))\cdot m_l'(\theta_0,\mathbb{X}_i(s))\cdot\lambda_s^i\,\mathrm{d}s\right)_{k,l=1}^{d} \xrightarrow[n\to\infty]{p} \Sigma_I.$$

Finally we will show (9), i. e. for all $\theta\in B(\theta_0,K)$, $k_1\geq 0,\ \ldots,k_d\geq 0,\ k_1+\ \ldots+k_d=2$

$$\frac{\partial^2 U_n(\theta,t)}{\partial\theta_1^{k_1}\ldots\partial\theta_d^{k_d}}=O_p(1).$$

The assumption (C) gives that

$$\left|\frac{\partial^2 U_{nl}(\theta,t)}{\partial\theta_j\theta_k}\right| := \left|\frac{\partial^3 L_n(\theta,t)}{\partial\theta_j\theta_k\theta_l}\right| \leq \frac{1}{n}\sum_{i=1}^{n}\int\limits_0^t\int G(s,z)P^i(\mathrm{d}s\times\mathrm{d}z).$$

Now, again as above, one gets:

$$\frac{1}{n}\sum_{i=1}^{n}\int\limits_0^t\int G(s,z)P^i(\mathrm{d}s\times\mathrm{d}z)-\frac{1}{n}\sum_{i=1}^{n}\int\limits_0^t\int G(s,z)\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s$$

is a 0-martingale with predictable quadratic variation process that is equal to

$$\frac{1}{n^2}\sum_{i=1}^{n}\int\limits_0^t\int G^2(s,z)\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s$$

and according to assumption (C) converges in probability to zero. Lenglart's inequality gives that

$$\sup_{u\leq t}\frac{1}{n}\sum_{i=1}^{n}\left|\int\limits_0^u\int G(s,z)P^i(\mathrm{d}s\times\mathrm{d}z)-\int\limits_0^u\int G(s,z)\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s\right|\xrightarrow[n\to\infty]{p}0$$

and together with assumption (C) $\frac{1}{n}\sum\limits_{i=1}^{n}\int\limits_0^t\int G(s,z)\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s=O_p(1)$ we get (9). □

The next theorem gives asymptotic normality for a consistent solution of the equations (3).

Theorem 3.2. Under the assumptions (A), (B), (C) and with $\widehat{\theta}_n$ consistent solution of (3), it holds that

$$\sqrt{n}\,(\widehat{\theta}_n-\theta_0)\xrightarrow[n\to\infty]{\mathcal{D}}N(0,\Sigma),$$

where $\Sigma = \Sigma_I^{-1}\Sigma_U\Sigma_I^{-1}$.

Further $\hat{\Sigma}_U := \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{N_t^i} \psi^2 \left(\frac{Z_i^j - m(\widehat{\theta}_n, \mathbb{X}_i(T_i^j))}{\sigma(\mathbb{X}_i(T_i^j))} \right) \cdot \frac{m_k'(\widehat{\theta}_n, \mathbb{X}_i(T_i^j)) \cdot m_l'(\widehat{\theta}_n, \mathbb{X}_i(T_i^j))}{\sigma^2(\mathbb{X}_i(T_i^j))} \right)_{k,l=1}^{d}$ and $-I(\widehat{\theta}_n, t)$ provide consistent estimates of Σ_U and Σ_I, respectively.

P r o o f. Making a Taylor expansion around the true parameter θ_0 one gets

$$U_n(\widehat{\theta}_n, t) = U_n(\theta_0, t) + (\widehat{\theta}_n - \theta_0)^T \cdot I_n(\theta^*, t)$$

where $\theta^* \in$ line $(\widehat{\theta}_n, \theta_0)$, so since $\widehat{\theta}_n$ is a solution, this equation states that

$$-I_n(\theta^*, t)^{-1} \cdot U_n(\theta_0, t) = \widehat{\theta}_n - \theta_0.$$

The theorem follows if one can show

$$\sqrt{n}\, U_n(\theta_0, t) \xrightarrow[n\to\infty]{\mathcal{D}} N(0, \Sigma_U), \tag{12}$$

$$I_n(\theta^*, t) \xrightarrow[n\to\infty]{p} \Sigma_I \text{ for } \widehat{\theta}_n \xrightarrow[n\to\infty]{p} \theta_0. \tag{13}$$

The first condition (12) follows by the martingale convergence theorem, see Jacod and Shiryaev [9], p. 476 or Andersen and others [2], p. 83. The quadratic predictable variation process of $\sqrt{n}U_n(\theta_0, t)$ is equal to

$$\frac{1}{n} \sum_{i=1}^{n} \int_0^t \gamma_3(\theta_0, \mathbb{X}_i(s)) \cdot \frac{m_k'(\theta_0, \mathbb{X}_i(s)) \cdot m_l'(\theta_0, \mathbb{X}_i(s))}{\sigma^2(\mathbb{X}_i(s))} \cdot \lambda_s^i \,\mathrm{d}s$$

which converges to Σ_U according to the assumption (6). Then together with the result (10) one gets that

$$\left\langle \sqrt{n}U_n(\theta_0, t) I\{|U_n(\theta_0, t)| > \varepsilon\} \right\rangle \xrightarrow[n\to\infty]{p} 0.$$

Finally, condition (13) follows from the Taylor expansion (similarly as in the last part of previous proof). □

Further result can be used for testing a simple hypothesis H: $\theta = \theta_0$ against the composite hypothesis G: $\theta \in \Theta$. Let $W_p(\Sigma)$ denote the Wishart distribution corresponding to a p-dimensional normal distribution $N_p(0, \Sigma)$. Define $R_n = n(L_n(\widehat{\theta}_n, t) - L_n(\theta_0, t))$.

Theorem 3.3. Under the hypothesis H and under the assumptions of the Theorem 3.2. it holds that

$$R_n \xrightarrow[n\to\infty]{\mathcal{D}} W_p(1/2\Sigma_I^{-1/2}\Sigma_U\Sigma_I^{-1/2}).$$

P r o o f. The proof can be again carried out by repeating the steps of the proof of Theorem 3 of [23]. □

3.2. Properties of M-estimation for unknown $\sigma^2(.)$

The conditional variance $\sigma^2(.)$ is usually unknown. The natural choice is replacing $\sigma^2(.)$ by its estimator $\widehat{\sigma_n^2}(.)$. To be able to apply asymptotic properties of the M-statistics on the estimator $\widehat{\theta}_n$, we need to know something about the behaviour of the sum

$$\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}\psi\left(\frac{Z_i^j - m(\theta,\mathbb{X}_i(T_i^j))}{\widehat{\sigma_n}(\mathbb{X}_i(T_i^j))}\right)\cdot\left(m_k'(\theta,\mathbb{X}_i(T_i^j))\right)_{k=1}^{d}\cdot\frac{1}{\widehat{\sigma_n}(\mathbb{X}_i(T_i^j))}.$$

It seems to be obvious if the derivative of the function ρ exists and is sufficiently smooth that for consistent estimator of $\sigma^2(.)$ we can rely on properties proved in the previous section. Indeed under some additional assumptions the theorems will be still valid.

Scheike and Zhang [24] propose the non-parametric estimation of conditional variance which is uniformly consistent. Let $K(.)$ be a kernel function with support on $[-1;1]$, $\int K(u)\,\mathrm{d}u = 1$, and let $b = (b_1,\dots b_p)$ be a p-dimensional bandwidth, $|b| = b_1\dots b_p$, $b\in(0,\infty)^p$. The Nadaraya–Watson type estimator, $\hat{m}(z)$ of $m(z)$ is defined by

$$\hat{m}(z) = \frac{\hat{r}(z)}{\hat{\alpha}(z)},$$

where

$$\hat{r}(z) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i} Z_i^j\frac{1}{|b|}K(z-\mathbb{X}_i(T_i^j),b)$$

and

$$\hat{\alpha}(z) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}\frac{1}{|b|}K(z-\mathbb{X}_i(T_i^j),b).$$

Similarly, the estimator of the variance function, $\sigma^2(.)$, can be estimated by the squared-residual kernel estimator

$$\hat{V}(z) = \frac{V(z)}{\hat{\alpha}(z)} - (\hat{m}(z))^2,$$

where

$$V(z) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}(Z_i^j)^2\frac{1}{|b|}K(z-\mathbb{X}_i(T_i^j),b).$$

Before stating the theorem, let us redefine the notation to make the dependence on σ more explicit. Redefine $L_n(\theta,t)$ to $L_n(\sigma,\theta,t)$, $U_n(\theta,t)$ to $U_n(\sigma,\theta,t)$, $I_n(\theta,t)$ to $I_n(\sigma,\theta,t)$ and $R(\theta,t)$ to $R(\sigma,\theta,t)$.

Theorem 3.4. Under the assumptions

(i) (4) holds,

(ii) $\mathbb{X}_i(s)$ belongs to some compact set C almost surely for all s,

(iii) $\widehat{\sigma_n^2}(.)$ is a uniformly consistent estimator of $\sigma^2(.)$, i.e. $\sup_{x\in C} |\widehat{\sigma_n^2}(x) - \sigma^2(x)| \xrightarrow[n\to\infty]{p} 0$, and further there exists $\varepsilon > 0$ such that $\widehat{\sigma_n}(x) > \varepsilon$ for all $x \in C$,

(iv) $m(\theta, x)$ is three times differentiable in a neighborhood $B(\theta_0, K)$ around the parameter θ_0 and $\psi(.)$ is two times differentiable and $\psi''(.)$ is absolutely continuous,

(v) $\gamma_4(\theta_0, \mathbb{X}_i(s)) := \int \left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)^4 \mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right) < \infty$,

(vi) $\mathrm{E}\frac{1}{n}\sum\int_0^t \lambda_s^i\, \mathrm{d}s = O_p(1)$

there exists a consistent solution of $U(\widehat{\sigma_n}, \theta, t) = 0$, that provides a local minimum of $L(\widehat{\sigma_n}, \theta, t)$.

P r o o f. The proof proceeds similarly as in Theorem 3.1. To prove the proposition, we need to show that

$$(U_n(\widehat{\sigma_n}, \theta_0, t) - U_n(\sigma, \theta_0, t)) \xrightarrow[n\to\infty]{p} 0, \tag{14}$$

$$(I_n(\widehat{\sigma_n}, \theta, t) - I_n(\sigma, \theta_0, t)) \xrightarrow[n\to\infty]{p} 0, \tag{15}$$

$$\left(d^2_{\theta-\theta_0} U_n(\widehat{\sigma_n}, \theta^*, t) - d^2_{\theta-\theta_0} U_n(\sigma, \theta^*, t)\right) \xrightarrow[n\to\infty]{p} 0, \tag{16}$$

where $\theta^* \in$ line (θ, θ_0).

To show (14), start with the Taylor expansion of the $\psi(\widehat{\sigma_n}, \theta_0, s)$ for all $\widehat{\sigma_n}(\mathbb{X}_i(s))$ satisfying the assumption (iii) around true value $\sigma(\mathbb{X}_i(s))$:

$$\begin{aligned} &\psi(\widehat{\sigma_n}, \theta_0, s) \\ &= \psi(\sigma, \theta_0, s) - \psi'(\sigma_*, \theta_0, s) \cdot \frac{z - m(\theta_0, \mathbb{X}_i(s))}{\sigma_*^2(\mathbb{X}_i(s))} \cdot (\widehat{\sigma_n}(\mathbb{X}_i(s)) - \sigma(\mathbb{X}_i(s))) \end{aligned} \tag{17}$$

where $\sigma_*(x) \in$ line $(\widehat{\sigma_n}(x), \sigma(x))$ for all $x \in C$. Let us apply (17) in the formula

$$\begin{aligned} &U_n(\widehat{\sigma_n}, \theta_0, t) \\ &= -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i} \psi\left(\frac{Z_i^j - m(\theta_0, \mathbb{X}_i(T_i^j))}{\widehat{\sigma_n}(\mathbb{X}_i(T_i^j))}\right) \cdot \left(m_k'(\theta_0, \mathbb{X}_i(T_i^j))\right)_{k=1}^{d} \cdot \frac{1}{\widehat{\sigma_n}(\mathbb{X}_i(T_i^j))} \end{aligned}$$

and study both summands in the difference $\Delta U_n(\theta_0, t) := U_n(\widehat{\sigma_n}, \theta_0, t) - U_n(\sigma, \theta_0, t)$ separately. Define the kth component in the $\Delta U_n(\theta_0, t)$ as $\Delta U_n(k, \theta_0, t)$ and define the first summand in the $\Delta U_n(k, \theta_0, t)$ as $\Delta U1_n(k, \theta_0, t)$. Then

$$\begin{aligned} &\Delta U1_n(k, \theta_0, t) \\ &= -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i} \psi(\sigma, \theta_0, T_i^j) \cdot m_k'(\theta, \mathbb{X}_i(T_i^j)) \cdot \frac{1}{\sigma(\mathbb{X}_i(T_i^j))} \cdot \left(\frac{\sigma(\mathbb{X}_i(T_i^j))}{\widehat{\sigma_n}(\mathbb{X}_i(T_i^j))} - 1\right). \end{aligned}$$

Since the $\widehat{\sigma_n^2}(x)$ is a consistent estimator of $\sigma^2(x)$, and $\widehat{\sigma_n}(x) > \varepsilon > 0$ it holds that $\sup_{x\in C}\left|\frac{\sigma(x)}{\widehat{\sigma_n}(x)}\right| \xrightarrow[n\to\infty]{p} 1$ and there exists $K > 0$ such that $\sup_{x\in C}\left|\frac{\sigma(x)}{\widehat{\sigma_n}(x)} - 1\right| < K$.

It follows that $|\Delta U1_n(k,\theta_0,t)| \le K|U_n(k,\theta_0,t)|$ and similarly as in (10) we get $\Delta U1_n(k,\theta_0,t) \xrightarrow[n\to\infty]{p} 0.$

Let us consider the second summand of $\Delta U_n(k,\theta_0,t)$:

$$\Delta U2_n(k,\theta_0,t) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}\psi'(\sigma_*,\theta_0,T_i^j)\cdot\frac{Z_i^j - m(\theta_0,\mathbb{X}_i(T_i^j))}{\sigma_*^2(\mathbb{X}_i(T_i^j))} \cdot m_k'(\theta_0,\mathbb{X}_i(T_i^j))\cdot\frac{\widehat{\sigma_n}(\mathbb{X}_i(T_i^j)) - \sigma(\mathbb{X}_i(T_i^j))}{\widehat{\sigma_n}(\mathbb{X}_i(T_i^j))}.$$

Since $\sigma(x)$, $\sigma_*(x)$, $\widehat{\sigma_n}(x)$ and $m_k'(\theta_0,x)$ are bounded functions and $\psi'(.)$ is absolutely continuous function we can write

$$|\Delta U2_n(k,\theta_0,t)| \le \frac{K}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}\left|\frac{Z_i^j - m(\theta_0,\mathbb{X}_i(T_i^j))}{\sigma(\mathbb{X}_i(T_i^j))}\right|\cdot\left|\widehat{\sigma_n}(\mathbb{X}_i(T_i^j)) - \sigma(\mathbb{X}_i(T_i^j))\right|.$$

We need to show that $|\Delta U2_n(k,\theta_0,t)| \xrightarrow[n\to\infty]{p} 0$. Since the $\widehat{\sigma_n^2}(x)$ is a consistent estimator of $\sigma^2(x)$ it is sufficient to prove that $\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}\left|\frac{Z_i^j - m(\theta_0,\mathbb{X}_i(T_i^j))}{\sigma(\mathbb{X}_i(T_i^j))}\right| < \infty.$ Write

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}\left|\frac{Z_i^j - m(\theta_0,\mathbb{X}_i(T_i^j))}{\sigma(\mathbb{X}_i(T_i^j))}\right| = \frac{1}{n}\sum_{i=1}^{n}\Bigg(M^i\left(\left|\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right|\right)_t + \int_0^t\int\left|\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right|\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s\Bigg) \tag{18}$$

and note that due to the assumptions (v) and (vi) the first term is a 0-martingale with a compensator that is equal to the second term. The quadratic predictable variation process of the martingale is given as

$$\frac{1}{n^2}\sum_{i=1}^{n}\int_0^t\int\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)^2\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s = \frac{1}{n^2}\sum_{i=1}^{n}\int_0^t\lambda_s^i\,\mathrm{d}s \xrightarrow[n\to\infty]{p} 0.$$

Hence, by Lenglarts's inequality the first term of (18) converges in probability to zero. The second term is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\int_0^t\int\left|\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right|\lambda_s^i\,\mathrm{d}F_s\left(\frac{z-m(\theta_0,\mathbb{X}_i(s))}{\sigma(\mathbb{X}_i(s))}\right)\mathrm{d}s \le \frac{1}{n}\sum_{i=1}^{n}\int_0^t\lambda_s^i\,\mathrm{d}s = O_p(1).$$

Now it is easy to show (15) and (16). Since we would just repeat the corresponding technique as above we omit the rest of the proof. □

Theorem 3.5. Under the assumption of Theorem 3.4, Theorem 3.2 and Theorem 3.3 remain true if $\sigma^2(.)$ is replaced by its a uniformly consistent estimator $\widehat{\sigma^2_n}(.)$.

Proof. The proof follows similarly as above. □

4. DISCUSSION

We have presented in this paper the assumptions under which there exists a consistent and asymptotic normal M-estimator of unknown regression parameter in model with longitudinal data. Let us look at the assumptions more closely. The most severe limitation was given on the form of error penalty function $\rho(.)$ in the M-estimation in (2). Since we used Taylor's expansion for proving characteristics of estimator we required so that $\rho(.)$, $\psi(.)$, $\psi'(.)$,$\psi''(.)$ were absolutely continuous functions. For example ordinary least square estimator meets these requirements but we search for function which is less increasing than square. Since the influence function of an M-estimate is proportional to $\psi(x)$, the function $\psi(x)$ (roughly speaking) measures the influence of a datum on the value of the parameter estimation. For the least square estimator with $\rho(x) = \frac{x^2}{2}$, the influence function is $\psi(x) = x$, that is, the influence of a datum on the estimation increases linearly with the size of its error, which confirms the non-robustness of the least square estimator. Although the set of sufficiently smooth functions of $\psi(.)$ are limited, we still can use e. g. $L_2 - L_1$ function, Cauchy function, Geman and McClure function, Welsch function or Hebert and Leahy function. For the definition of a few commonly used M-estimators see the following table:

Type	$\rho(\mathbf{x})$	$\psi(\mathbf{x})$
$L_2 - L_1$	$2\left(\sqrt{1+\frac{x^2}{2}}-1\right)$	$\frac{x}{\sqrt{1+\frac{x^2}{2}}}$
Cauchy	$\frac{c^2}{2}\log\left(1+\left(\frac{x}{c}\right)^2\right)$	$\frac{x}{1+(\frac{x}{c})^2}$
Geman–Mc Clure	$\frac{x^2/2}{1+x^2}$	$\frac{x}{(1+x^2)^2}$
Welsch	$\frac{c^2}{2}\left(1-e^{-(\frac{x}{c})^2}\right)$	$xe^{-(\frac{x}{c})^2}$
Hebert and Leahy	$\frac{1}{2}\log(1+x^2)$	$\frac{x}{1+x^2}$

These results should be extended also for non-smooth penalty functions. This may be a relevant area for further research that can be inspired with the results and methods of Rubio and Víšek [22].

ACKNOWLEDGEMENT

I would like to express my sincere thanks to my advisor, Dr. Petr Lachout, for his suggestions and partnership. I would also like to thank two referees for their valuable comments.

(Received April 20, 2006.)

REFERENCES

[1] P. K. Andersen and Ø. Borgan: Counting process for life history data: A review. Scand. J. Statistist. *12* (1985), 97–158.
[2] P. K. Andersen, Ø. Borgan, R. D. Gill, and N. Keiding: Statistical Models Based on Counting Processes. Springer, New York 1991.
[3] P. Billingsley: Covergence of Probability Measures. Wiley, New York 1999.
[4] R. Boel, P. Varaiya, and E. Wong: Martingales on jump processes. SIAM J. Control *13* (1975), 999–1061.
[5] P. J. Diggle, P. Heagerty, K. Y. Liang, and S. L. Zeger: Analysis of Longitudinal Data. Oxford Univ. Press, Oxford 2002.
[6] T. R. Flemming and D. P. Harrington: Counting Processes and Survival Analysis. Wiley, New York 1991.
[7] P. J. Huber: Robust Statistics. Wiley, New York 1981.
[8] M. Jacobsen: Statistical Analysis of Counting Processes. Springer, New York 1982.
[9] J. Jacod and A. N. Shiryaev: Limit Theorems for Stochastic Processes. Springer, New York 2002.
[10] V. Jarník: Differential Calculus II (in Czech). Academia, Prague 1984.
[11] J. Jurečková and P. K. Sen: Uniform second order asymptotic linearity of M-statistics in linear models. Statist. Decisions *7* (1989), 263–276.
[12] J. P. Klein and M. L. Moeschberger: Survival Analysis. Springer, New York 1997.
[13] G. Last and A. Brandt: Marked Point Processes on the Real Line, The Dynamic Approach. Springer, New York 1995.
[14] E. Lehmann: Theory of Point Estimation. Wiley, New York 1983.
[15] K. Y. Liang and S. L. Zeger: Longitudinal data analysis using generalized linear models. Biometrika *73* (1986), 13–22.
[16] J. K. Lindsey: Models for Repeated Measurements. Oxford Univ. Press, Oxford 1999.
[17] T. Martinussen and T. H. Scheike: A Non-Parametric Dynamic Additive Regression Model for Longitudinal Data. Research Report, Copenhagen 1998.
[18] T. Martinussen and T. H. Scheike: A Semi-Parametric Additive Regression Model for Longitudinal Data. Research Report, Copenhagen 1998.
[19] M. Orsáková: Models for censored data. In: Proc. 8^{th} Annual Conference of Doctoral Students WDS'99, Faculty of Mathematics and Physics, Charles Univ. Prague 1999.
[20] M. Orsáková: Regression models for longitudinal data. In: Proc. Robust'2000, Union of the Czech Mathematicians and Physicists, Prague 2001, pp. 210–216.
[21] R. Rebolledo: Central limit theorem for local martingales. Z. Wahrsch. verw. Geb. *51* (1980), 269–286.
[22] A. M. Rubio and J. Á. Víšek: A note on asymptotic linearity of M-statistics in nonlinear models. Kybernetika *32* (1996), 353–374.
[23] T. H. Scheike: Parametric regression for longitudinal data with counting process measurement times. Scand. J. Statistist. *21* (1994), 245–263.
[24] T. H. Scheike and M. Zhang: Cumulative regression tests for longitudinal data. Annals Statist. (1998), 1328–1354.
[25] J. Štěpán: The Theory of Probability (in Czech). Academia, Prague 1987.

Martina Orsáková, Department of Probability and Statistics, Faculty of Mathematics and Physics, Charles University in Prague, Sokolovská 83, 186 75, Praha 8. Czech Republic.
e-mail: morsakova@csas.cz