Empirical likelihood test in a posteriori change-point nonlinear model

Gabriela Ciuperca · Zahraa Salloum

Abstract In this paper, in order to test whether changes have occurred in a nonlinear parametric regression, we propose a nonparametric method based on the empirical likelihood. Firstly, we test the null hypothesis of no-change against the alternative of one change in the regression parameters. Under null hypothesis, the consistency and the convergence rate of the regression parameter estimators are proved. The asymptotic distribution of the test statistic under the null hypothesis is obtained, which allows to find the asymptotic critical value. On the other hand, we prove that the proposed test statistic has the asymptotic power equal to 1. These theoretical results allows find a simple test statistic, very useful for applications. The epidemic model, a particular model with two change-points under the alternative hypothesis, is also studied. Numerical studies by Monte-Carlo simulations show the performance of the proposed test statistic, compared to an existing method in literature.

Keywords Change-point · Nonlinear parametric model · Empirical likelihood test · Asymptotic behaviour.

1 Introduction

We consider a classical model of parametric nonlinear regression:

$$Y_i = f(X_i; \beta) + \varepsilon_i, \quad i = 1, \ldots, n,$$

where a possible change in the regression parameters could occurs. This is called, change-point problem.
Change-point detection problems fall in two categories. The first type is a posteriori: after that the \( n \) all observations are realized, we study if, a certain moment \( k \in \{2, \cdots, n-1\} \), the model (parameter \( \beta \), to be more precise) is changed:

\[
Y_i = \begin{cases} 
  f(X_i; \beta_1) + \varepsilon_i & i = 1, \cdots, k \\
  f(X_i; \beta_2) + \varepsilon_i & i = k + 1, \cdots, n.
\end{cases}
\]  

(2)

The second type of change-points model is sequential (a priori), where the change detection is performed in real time. If in the first \( k \) observations no change in the parameter regression has occurred, at observation \( k + 1 \) we test that there is no change in the model: \( Y_i = f(X_i; \beta) + \varepsilon_i \), for all \( i = 1, \cdots, k+1 \), against the hypothesis that the model has the form:

\[
Y_i = f(X_i; \beta) + \varepsilon_i \\
Y_{k+1} = f(X_{k+1}; \beta^*) + \varepsilon_{k+1},
\]

(3)

with \( \beta \neq \beta^* \).

In this paper, we consider a posteriori change-point problem.

For the two types of problems, the number of publications in the last years is every extensive. Let us mention some references concerning the sequential change-point problem. If the function \( f \) is linear, \( f(x, \beta) = x^t \beta \), in the papers [12], [13], the CUSUM method is used to find a test statistic for detecting the presence or absence of a change. The results have been generalized by [7] for a nonlinear model. We can also mention the papers [14], [17], [18] for the sequential detection of a change-point.

For a posteriori change-point problem, in order to detect a change-point presence, model (1) is tested against model (2). The non-identifiability of model under the null hypothesis makes classical test techniques unusable. In most articles in the literature, the authors propose criteria: see for example [20], [5], [28]. Various hypothesis tests have been proposed only for the linear models. The likelihood-ratio test method is used in [2] and [15]. A non-parametric approach based on Empirical Likelihood (EL) for testing a change in a linear model is considered by [16]. Always using the EL method, the papers [31], [30] construct the confidence region for the coefficient difference of a two-sample linear regression model. For a linear quantile model, [23] proposes two types of statistics: one based on the subgradient and an another based on Wald statistic. For a generalized linear models, a method based on maximum of score statistics is used in [1] to test the change in the regression parameters.

In this paper, we consider the change-point problem in a general nonlinear model, by the EL method. Then, the framework of [16] is generalized. One of the major difficulties for nonlinear model (beside the linear model approach) is that, for finding the test statistic, the corresponding score functions depend on the regression parameters, and above all, the analytical form of these derivatives is unknown. On the other hand, for linear models, many proofs are based on the convexity of the regression function with respect to the parameter regression, then, the extreme value of a convex function is attained on the boundary. These two factors lead to a more difficult theoretical study of the
test statistics for nonlinear model. Another difficulty to study the properties of the test statistic, for detecting a change in model, is due to the dependence on the change-points of the regression parameter estimator. To the authors’ knowledge, the only paper which studies a hypothesis test in a change-point nonlinear model is [4] for very smooth nonlinear functions, using the least square method. But the least square method, in respect to the EL method, has the disadvantage that is less efficient for outliers data. This occurs in the case of fatter tailed distributions of the error term. Moreover, we will see in Section 2 that the considered assumptions in [4] are stronger than in the present paper.

Note that the paper [11] tests the structural stability in a nonlinear model by a generalized method of moments, and not a change in the regression parameters.

I would emphasize that in the present paper, we have obtained an interesting result concerning the numerical simulations. The EL test outperforms the change detection by least square (LS) test proposed by [4]. The LS test does not work when the change-point is off-centred in the measurement interval. The proposed EL test does not this defect.

The paper is organized as follows. We first construct in Section 2 a statistic, in order to test the change in the regression parameters of the nonlinear model. The asymptotic behaviour of the test statistic under the null hypothesis as well as under the alternative hypothesis is studied. A particular case of two change-point model, the epidemic model, is considered in Section 3. In Section 4 simulations results illustrate the performance of the proposed test, concerning the empirical size, the empirical power and the estimation of the time of change, in particular when the error distribution is not Gaussian, when it has outliers or a large standard deviation. Some lemmas and their proofs are given in the last section (Appendix, Section 5).

2 Test of a change-point

In this section, for a nonlinear model we are going to test the hypothesis that there is no change in the parameters of model (1) against the hypothesis that the parameters change from $\beta_1$ to $\beta_2$ at an unknown observation $k$, i.e. the model (2).

2.1 Hypothesis, notations, assumptions

All throughout the paper, $C$ denotes a positive generic constant which may take different values in different formula or even in different parts of the same formula. All vectors are column and $v^t$ denotes the transposed of $v$. All vectors and matrices are in bold. Concerning the used norms, for a $m$-vector $v = (v_1, \cdots, v_m)$, let us denote by $\|v\|_1 = \sum_{j=1}^{m} |v_j|$ its $L_1$-norm
and \( \|v\|_2 = (\sum_{j=1}^{n} v_j^2)^{1/2} \) its \( L_2 \)-norm. For a matrix \( M = (a_{ij})_{1 \leq i \leq m_1, 1 \leq j \leq m_2} \), we denote by \( \|M\|_1 = \max_{j=1,\ldots,m_2} (\sum_{i=1}^{m_1} |a_{ij}|) \), the subordinate norm to the vector norm \( \|\cdot\|_1 \). Let \( \frac{L}{n \to \infty} \), \( \frac{P}{n \to \infty} \), \( \frac{a.s.}{n \to \infty} \) represent convergence in distribution, in probability and almost sure, respectively, as \( n \to \infty \).

For coherence, we try to use the some notations as in the paper [10], where the linear model was considered. This will allow to highlight the difficulties and results due to the nonlinearity.

For each observation \( i \), \( Y_i \) is the response variable, \( X_i \) is a \( p \times 1 \) random vector of regressors with distribution function \( H(x) \), with \( x \in \mathcal{T} \), \( \mathcal{T} \subseteq \mathbb{R}^p \), and \( \varepsilon_i \) is the error.

The continuous random vector sequence \( (X_i, \varepsilon_i)_{1 \leq i \leq n} \) is independent identically distributed (i.i.d), with the same joint distribution as \( (X, \varepsilon) \). For all \( i, \varepsilon_i \) is independent of \( X_i \).

The regression function \( f : \mathcal{T} \times \Gamma \to \mathbb{R} \), with \( \mathcal{T} \subseteq \mathbb{R}^p \) and \( \Gamma \subseteq \mathbb{R}^d \), is known up to a parameter \( \beta = (\beta_1, \ldots, \beta_d) \). The parameter set \( \Gamma \) is supposed compact.

In following, for \( x \in \mathcal{T} \) and \( \beta \in \Gamma \), we use notation \( f(x, \beta) \equiv \partial f(x, \beta)/\partial \beta \), \( f(x, \beta) \equiv \partial^2 f(x, \beta)/\partial \beta^2 \).

With regard to the random variable \( \varepsilon \) we make following assumption :

(A1) \( E[\varepsilon_i] = 0 \) and \( E[\varepsilon_i^2] < \infty \), for all \( i = 1, \cdots, n \).

The regression function \( f : \mathcal{T} \times \Gamma \to \mathbb{R} \) and the random vector \( X \) satisfy the conditions :

(A2) for all \( x \in \mathcal{T} \) and for \( \beta \in \Gamma \), the function \( f(x, \beta) \) is thrice differentiable in \( \beta \) and continuous on \( \mathcal{T} \).

(A3) \( (|\partial f(x, \beta)|)_{1 \leq j, k \leq d} \) and \( (|\partial^2 f(x, \beta)|)_{1 \leq j, k, l \leq d} \) are bounded for any \( x \in \mathcal{T} \) and \( \beta \) in a neighbourhood of \( \beta^0 \).

(A4) \( E[||f(X, \beta)||_1] < \infty \), \( E[||f(X, \beta) f^t(X, \beta)||_1] < \infty \) and \( E[||\partial^2 f(X, \beta)||_1] < \infty \), for all \( 1 \leq j, k \leq d \) and \( \beta \) in a neighbourhood of \( \beta^0 \).

Assumptions (A3), (A4) are standard conditions, which are used in nonlinear models, for example see book [23]. We remark that assumption (A4) is weaker than the corresponding assumption employed in paper [3], where the least square method is used to test \( H_0 \) against \( H_1 \). The assumptions of the paper [3] are: \( E[\varepsilon_i^2] < \infty \), \( \sup_{\beta} E[|f(X, \beta)|^{2s}] < \infty \), \( \sup_{\beta} E[||f(X, \beta)||_2]^{2s} < \infty \) and for all \( 1 \leq j, k \leq d \), \( \sup_{\beta} E[||\partial^2 f(X, \beta)||_1]^{s} < \infty \), for some \( s > 2 \).

We are interested in testing of the null hypothesis of no change in the model (4). Then the model has the form (1), that is

\[
H_0 : \beta_1 = \beta_2 = \beta.
\]
The alternative hypothesis assumes that one change occurs in the regression parameters, that is

\[ H_1 : \beta_1 \neq \beta_2. \]  

(5)

Let \( \beta^0 \) denote the true (unknown) of the parameter \( \beta \) under hypothesis \( H_0 \) and \( \beta'_1, \beta^0_2 \) (also unknown) the true parameters under hypothesis \( H_1 \).

In addition to the notations introduced above, let us consider the following \( d \)-random vectors

\[ g(X_i, \beta) = g_i(\beta) = f(X_i, \beta)[y_i - f(X_i, \beta)]. \]

We remark that, under the hypothesis \( H_0 \), we have \( g_i(\beta^0) = f(X_i, \beta^0) \epsilon_i \), for all \( i = 1, \ldots, n \) and \( E[g_i(\beta^0)] = 0 \). Consider also the \( d \times d \) matrix

\[ V = E[f(X_i, \beta^0) \epsilon_i f(X_i, \beta^0)^t]. \]

Then \( \text{Var}(\epsilon_i f(X_i, \beta^0)) = \sigma^2 V \).

In order to introduce the empirical likelihood, let \( y_1, \ldots, y_k, y_{k+1}, \ldots, y_n \) be observations for the random variables \( Y_1, \ldots, Y_k, Y_{k+1}, \ldots, Y_n \). For more details concerning empirical likelihood method, the reader can refer to [21].

Consider the following sets \( I \equiv \{1, \ldots, k\} \) and \( J \equiv \{k+1, \ldots, n\} \), which contain the observation subscripts of the two segments for the model (2). Corresponding to these sets, let be the probability vectors \( (p_1, \ldots, p_k) \) and \( (q_{k+1}, \ldots, q_n) \). These vectors contained the probability to observe the value \( y_i \) (respectively \( y_j \)) for the dependent variable \( Y_i \) (respectively \( Y_j \)) : \( p_i \equiv \mathbb{P}[Y_i = y_i] \), for \( i = 1, \ldots, k \) and \( q_j \equiv \mathbb{P}[Y_j = y_j] \), for \( j = k+1, \ldots, n \). Obviously, these probabilities satisfy the relations \( \sum_{i \in I} p_i = 1 \) and \( \sum_{j \in J} q_j = 1 \).

2.2 Test statistics

Under hypothesis \( H_0 \) given by (4), the profile empirical likelihood (EL) for \( \beta \) is

\[
\mathcal{R}_{0,nk}(\beta) = \sup_{(p_1, \ldots, p_k)} \sup_{(q_{k+1}, \ldots, q_n)} \left\{ \prod_{i \in I} p_i \prod_{j \in J} q_j; \sum_{i \in I} p_i = 1, \sum_{j \in J} q_j = 1, \sum_{i \in I} p_i g_i(\beta) = \sum_{j \in J} q_j g_j(\beta) = 0_d \right\},
\]

with \( 0_d \) the \( d \)-vector with all components zero. Without constraints \( \sum_{i \in I} p_i g_i(\beta) = 0_d \), the maximum of \( \prod_{i \in I} p_i \prod_{j \in J} q_j \) are attained for \( p_i = k^{-1}, q_j = (n-k)^{-1} \), respectively. Then, the profile EL ratio for \( \beta \) has the form

\[
\mathcal{R}'_{0,nk}(\beta) = \sup_{(p_1, \ldots, p_k)} \sup_{(q_{k+1}, \ldots, q_n)} \left\{ \prod_{i \in I} k p_i \prod_{j \in J} (n-k) q_j; \sum_{i \in I} p_i = 1, \sum_{j \in J} q_j = 1, \sum_{i \in I} p_i g_i(\beta) = \sum_{j \in J} q_j g_j(\beta) = 0_d \right\}.  
\]

(6)
Similarly, under hypothesis $H_1$ given by (5), the profile EL is

$$R_{1,nk}(\beta_1, \beta_2) = \sup_{(p_1, \ldots, p_k, l_2)} \sup_{(q_{k+1}, \ldots, q_n)} \left\{ \prod_{i \in I} p_i \prod_{j \in J} q_j; \sum_{i \in I} p_i = 1, \sum_{j \in J} q_j = 1, \sum_{i \in I} p_i g_i(\beta_1) = 0_d, \sum_{j \in J} q_j g_j(\beta_2) = 0_d \right\}.$$

Then, the profile EL ratio for $\beta_1, \beta_2$ has the form

$$R'_{1,nk}(\beta_1, \beta_2) = \sup_{(p_1, \ldots, p_k, l_2)} \sup_{(q_{k+1}, \ldots, q_n)} \left\{ \prod_{i \in I} p_i \prod_{j \in J} (n-k)q_j; \sum_{i \in I} p_i = 1, \sum_{j \in J} q_j = 1, \sum_{i \in I} p_i g_i(\beta_1) = 0_d, \sum_{j \in J} q_j g_j(\beta_2) = 0_d \right\}.$$

Thus, using an idea similar to the maximum likelihood test for testing $H_0$ against $H_1$, we consider the profile EL ratio

$$\frac{R'_{0,nk}(\beta)}{R_{1,nk}(\beta_1, \beta_2)} = \frac{R'_{0,nk}(\beta)}{R'_{1,nk}(\beta_1, \beta_2)},$$

but, under this form, it has a complicated expression. In order to find a simpler form for the test statistic, we will study the denominator behaviour of the process given by (7).

The following result is a generalization of the nonparametric version of the Wilks theorem. More specifically, under $H_1$ due to the observation independence, on each segment we have a Wilks theorem. Then, we prove that, under $H_1$, the profile EL ratio for $\beta_1, \beta_2$ has a $\chi^2$ asymptotic distribution.

**Theorem 1** Suppose that assumptions (A1)-(A3) hold. Under the hypothesis $H_1$, we have

$$-2 \log R_{1,nk}(\beta_1, \beta_2) \xrightarrow{n \to \infty} \chi^2(2d).$$

**Proof.** Under hypothesis $H_1$, on the first segment generated by the observations for $i \in I$, the profile EL function for $\beta_1$, for fixed $k$, is

$$\sup_{(p_1, \ldots, p_k); 0 \leq p_i \leq 1, \sum_{i \in I} p_i = 1, \sum_{i \in I} p_i g_i(\beta_1) = 0_d} \left\{ \prod_{i \in I} k p_i; 0 \leq p_i \leq 1, \sum_{i \in I} p_i = 1, \sum_{j \in J} q_j g_j(\beta_2) = 0_d \right\}.$$

Applying the Lagrange multiplier method, using the paper [22], we have

$$p_i = \frac{1}{k(1 + \lambda_i^T g_i(\beta_1))},$$

where $\lambda_i \in \mathbb{R}^d$ the Lagrange multiplier.

Similarly, the profile EL function on the second segment generated by the observations for $j \in J$, is

$$\sup_{(q_{k+1}, \ldots, q_n); 0 \leq q_j \leq 1, \sum_{j \in J} q_j = 1, \sum_{j \in J} q_j g_j(\beta_2) = 0_d} \left\{ \prod_{j \in J} (n-k)q_j; 0 \leq q_j \leq 1, \sum_{j \in J} q_j = 1, \sum_{j \in J} q_j g_j(\beta_2) = 0_d \right\}.$$
This function is maximized for \( q_j = (n - k - \lambda^1_i g_i(\beta_2))^{-1} \), with \( \lambda_2 \in \mathbb{R}^p \) the Lagrange multiplier. Then the empirical log-likelihood ratio statistic can be written

\[
-2 \log R_{1,nk}(\beta_1, \beta_2) = 2 \sum_{i \in I} \log [1 + \lambda^1_i g_i(\beta_1)] + 2 \sum_{j \in J} \log [1 - \lambda^2_j g_j(\beta_2)].
\]  

(8)

In view of Theorem 2 of [22], using assumptions (A1), (A2) and (A3), each sum of the right-hand side of (8) converges in law to \( \chi^2(d) \). Taking into account that the two terms of relation (8) involved two independent sets of random vectors we obtain the theorem.

Consequently of this theorem, under hypothesis \( H_1 \) the denominator of the EL ratio given by (7), is not asymptotically depend on the parameters \( \beta_1 \) and \( \beta_2 \). Then, from now on, we are going to consider that test statistic \(-2 \log R_{1,nk}(\beta)\).

Taking into account the expression of \( R_{1,nk}(\beta) \) given by (6) and using the Lagrange multiplier method, we have that maximizing \(-2 \log R_{1,nk}(\beta)\) is equivalent to maximizing the following statistic with respect to \( \beta, \eta_1, \eta_2, \lambda_1, \lambda_2 \)

\[
\sum_{i \in I} [\log p_i - n \lambda^1_i p_i g_i(\beta)] + \sum_{j \in J} [\log q_j + n \lambda^2_j q_j g_j(\beta)] + \eta_1 \left(\sum_{i \in I} p_i - 1\right) + \eta_2 \left(\sum_{j \in J} q_j - 1\right),
\]  

(9)

where \( \beta \in \Gamma, \eta_1, \eta_2 \in \mathbb{R} \) and \( \lambda_1, \lambda_2 \in \mathbb{R}^d \).

Since the derivatives of (9) with respect to \( p_i, q_j \) are null, using a similar argument as in the proof of Theorem 1 we obtain that

\[
p_i = \frac{1}{k + n \lambda^1_i g_i(\beta)}, \quad q_j = \frac{1}{n - k - n \lambda^2_j g_j(\beta)}.
\]  

(10)

Then, the statistic \(-2 \log R_{1,n,k,0}(\beta)\) becomes

\[
2 \sum_{i \in I} \log [1 + \frac{n}{k} \lambda^1_i g_i(\beta)] + 2 \sum_{j \in J} \log [1 - \frac{n}{n - k} \lambda^2_j g_j(\beta)].
\]  

(11)

Taking into account relation (10) for the probabilities \( p_i \) and \( q_j \), the derivative with respect to \( \beta \) of (11) is \( 2n \left(\sum_{i \in I} p_i \lambda^1_i g_i(\beta) - \sum_{j \in J} q_j \lambda^2_j g_j(\beta)\right) = 0 \), with \( g_i(\beta) \) the \( d \times d \) matrix of the derivatives of vector \( g_i(\beta) \) with respect to \( \beta \), for \( i = 1, \ldots, k \). In order to have single parameters \( \lambda_i \), we restrict the study to a particular case, when \( \lambda_1 \) and \( \lambda_2 \) satisfy the constraint

\[
V_{1n}(\beta) \lambda_1 = V_{2n}(\beta) \lambda_2,
\]

with

\[
V_{1n}(\beta) \equiv k^{-1} \sum_{i \in I} g_i(\beta), \quad V_{2n}(\beta) \equiv (n - k)^{-1} \sum_{j \in J} g_j(\beta).
\]
In the case of the true parameter $\beta^0$, this two last matrices are denoted $V^0_{1n} \equiv V_{1n}(\beta^0)$ and $V^0_{2n} \equiv V_{2n}(\beta^0)$. Considering this constraint, statistic $V_{1n}$ becomes

$$2 \sum_{i \in I} \log \left[ 1 + \frac{n}{k} \lambda_i^t g_i(\beta) \right] + 2 \sum_{j \in J} \log \left[ 1 - \frac{n}{n-k} \lambda_j^t V_{1n}(\beta) V_{2n}^{-1}(\beta) g_j(\beta) \right].$$ \hspace{1cm} (12)

On the other hand, in order that the parameters belong a bounded set, in the place of $k$, we consider $\hat{\theta}_{nk} \equiv k/n$, and we denote statistic $V_{1n}$ by $Z_{nk}(\theta_{nk}, \lambda, \beta)$. Under hypothesis $H_1$, if $k_0$ is the point where the model change, we denote $\theta_{nk} = k_0/n$.

Similar to the classical maximum likelihood test, but for models without change-points, we will study the maximum of empirical log-likelihood test statistic. For this, we calculate the score functions of test statistic $V_{1n}$

$$\phi_{1n}(\theta_{nk}, \lambda, \beta) \equiv \frac{\partial Z_{nk}(\theta_{nk}, \lambda, \beta)}{\partial \lambda} = \sum_{i \in I} \frac{g_i(\beta)}{\theta_{nk}} + \lambda_i^t g_i(\beta) - \sum_{j \in J} \frac{V_{1n}(\beta) V_{2n}^{-1}(\beta) g_j(\beta)}{1 - \theta_{nk} - \lambda_j^t V_{1n}(\beta) V_{2n}^{-1}(\beta) g_j(\beta)}. \hspace{1cm} (13)$$

$$\phi_{2n}(\theta_{nk}, \lambda, \beta) \equiv \frac{\partial Z_{nk}(\theta_{nk}, \lambda, \beta)}{\partial \beta} = \sum_{i \in I} \frac{\dot{g}_i(\beta) \lambda_i^t}{\theta_{nk} + \lambda_i^t g_i(\beta)} - \sum_{j \in J} \frac{\partial (V_{1n}(\beta) V_{2n}^{-1}(\beta) g_j(\beta))/\partial \beta}{1 - \theta_{nk} - \lambda_i^t V_{1n}(\beta) V_{2n}^{-1}(\beta) g_j(\beta)} \lambda_i^t. \hspace{1cm} (14)$$

Then, solving the system $\phi_{1n}(\theta_{nk}, \lambda, \beta) = 0_d$ and $\phi_{2n}(\theta_{nk}, \lambda, \beta) = 0_d$, the obtained solutions $\hat{\lambda}(\theta_{nk})$ and $\hat{\beta}(\theta_{nk})$ are the maximizers of the statistic $V_{1n}$. We so obtain the profile maximum empirical likelihood function $Z_{nk}(\theta_{nk}, \hat{\lambda}(\theta_{nk}), \hat{\beta}(\theta_{nk}))$, which depends only on the change-point parameter $\theta_{nk}$.

We emphasise that, compared with a linear model, in our case, matrix $V_{1n}(\beta)$, $V_{2n}(\beta)$ and derivative $g(\beta)$ depend on $\beta$. These, besides the nonlinearity of $g(\beta)$ involve difficulties in the study of the statistic $Z_{nk}(\theta_{nk}, \lambda, \beta)$ and of the solutions $\hat{\lambda}(\theta_{nk}), \hat{\beta}(\theta_{nk})$.

2.3 Asymptotic behaviour of the test statistic

In this section, for the probabilities given by (10), under the constraint $V_{1n}(\beta) \lambda_1 = V_{2n}(\beta) \lambda_2$, we will first prove that $k_{p_1}, (n-k)q_j$, can be framed by two strictly positive constants. This implies that the test statistic $Z_{nk}(\theta_{nk}, \hat{\lambda}(\theta_{nk}), \hat{\beta}(\theta_{nk}))$ is well defined.
Empirical likelihood test in a posteriori change-point nonlinear model

Properties established for $\hat{\lambda}(\theta_{nk})$ and $\hat{\beta}(\theta_{nk})$, solutions of (12), will allow to consider instead of (12), a more simple test statistic, given by relation (26). Next, we will study the asymptotic behaviour of this statistic, firstly under the hypothesis $H_0$ and next under $H_1$.

2.3.1 Asymptotic behaviour under $H_0$

In order to study asymptotic behaviour of $\hat{\lambda}(\theta_{nk})$ and $\hat{\beta}(\theta_{nk})$, we will first study $k p_i$, for $i \in I$, and $(n-k)q_j$, for $j \in J$, with $p_i$, $q_j$ given by (10). More exactly, we show that, if $\beta$ in the neighbourhood of $\beta^0$, $k p_i$ and $(n-k)q_j$ can be framed by two strictly positive and bounded constants, with probability close to one.

**Proposition 1** Let the $\eta$-neighbourhood of $\beta^0$, $V_\eta(\beta^0) = \{ \beta \in \Gamma; \| \beta - \beta^0 \|_2 \leq \eta \}$, with $\eta \to 0$. Under hypothesis $H_0$, suppose that assumptions (A1)-(A4) hold. Then we have

(i) For all $i \in I$, for all $\epsilon > 0$, there exist two constants $M_1, M_2 > 0$, such that, for all $\beta \in V_\eta(\beta^0)$,

$$\mathbb{P}\left[ \frac{1}{M_2} \leq \frac{1}{1 + \frac{\lambda}{\theta_{nk}} g_i(\beta)} \leq \frac{1}{M_1} \right] \geq 1 - \epsilon. \quad (15)$$

(ii) For all $j \in J$, for all $\epsilon > 0$, there exist two constants $M_3, M_4 > 0$, such that, for all $\beta \in V_\eta(\beta^0)$,

$$\mathbb{P}\left[ \frac{1}{M_4} \leq \frac{1}{1 - \frac{\lambda}{\theta_{nk}} V_{1n}(\beta)(V_{2n}(\beta))^{-1} g_j(\beta)} \leq \frac{1}{M_3} \right] \geq 1 - \epsilon. \quad (16)$$

**Proof.** (i) We consider the following decomposition for the Lagrange multiplier: $\lambda = \rho \phi$, such that $\rho \geq 0$ and $\| \phi \|_1 = 1$. Lemma 2 implies that, there exists $M_1 > 0$, such that

$$\frac{1}{1 + \frac{\lambda}{\theta_{nk}} g_i(\beta)} \geq \frac{1}{1 + \frac{\rho}{\theta_{nk}} \| \phi \|_1 g_i(\beta)} \geq \frac{1}{1 + \frac{\rho}{\theta_{nk}} \| g_i(\beta) \|_1} \geq \frac{1}{M_2},$$

with probability close to 1, that is, for all $\epsilon > 0$,

$$\mathbb{P}\left[ \frac{1}{1 + \frac{\rho}{\theta_{nk}} \| g_i(\beta) \|_1} \geq \frac{1}{M_2} \right] \geq 1 - \epsilon. \quad (17)$$

For the right-hand side of relation (15), we assume the contrary, that is, there exists $M_2 > 0$ such that

$$\sup_{i \in I, \beta \in \Gamma} \frac{1}{1 + \frac{\lambda}{\theta_{nk}} g_i(\beta)} \geq \frac{1}{M_2}.$$ 

This is equivalent to the fact that there exists $M_5 > 0$, such that

$$\inf_{i \in I, \beta \in \Gamma} \frac{\lambda^t}{\theta_{nk}} g_i(\beta) \leq - M_5.$$
Since $\lambda = \rho \phi$, $\rho > 0$, and $0 < \theta_{nk} < 1$, therefore exists $M_6 > 0$ such that
\[
\inf_{i \in I, \beta \in \Gamma} \phi_i^t g_i(\beta) \leq -M_6. \tag{18}
\]

On the other hand, we have that $\inf_{i \in I, \beta \in \Gamma} \phi_i^t g_i(\beta) \geq -\inf_{i \in I, \beta \in \Gamma} \|g_i(\beta)\|_1$, with probability 1. Taking into account relation (18), there exists $M_6 > 0$ such as $-\inf_{i \in I, \beta \in \Gamma} \|g_i(\beta)\|_1 \leq -M_6$ again too $\sup_{i \in I, \beta \in \Gamma} \|g_i(\beta)\|_1 \geq M_6$, which is in contradiction with relation (2). Then, the relation (18) holds.

By the following result, we show that $\hat{\lambda}(\theta_{nk})$ and $\hat{\beta}(\theta_{nk})$, the solutions of the score equations $\phi_{1n}(\theta_{nk}, \lambda, \beta) = 0_d$ and $\phi_{2n}(\theta_{nk}, \lambda, \beta) = 0_d$, have suitable properties. More precisely, we show that $\|\hat{\lambda}(\theta_{nk})\|_2 \to 0$, as $n \to \infty$ and that $\hat{\beta}(\theta_{nk})$ is a consistent estimator of $\beta^0$, under hypothesis $H_0$. We also obtain their convergence rate. This will allow us to propose a simpler test statistics instead of $Z_{nk}(\theta_{nk}, \lambda, \beta)$.

**Theorem 2** Suppose that the assumptions (A1)-(A4) hold. Under the hypothesis $H_0$, we have $\hat{\lambda}(\theta_{nk}) = \min \{\theta_{nk}, 1 - \theta_{nk}\}O_P((n \min (\theta_{nk}, 1 - \theta_{nk}))^{-1/2})$ and $\hat{\beta}(\theta_{nk}) - \beta^0 = O_P((n \min (\theta_{nk}, 1 - \theta_{nk}))^{-1/2})$.

**Proof.** The structure of the proof is similar to that of linear model (Lemma A1 of [16]) but important modifications and supplementary results are necessary, due to the model nonlinearity. Without loss of generality, we assume that $\min\{\theta_{nk}, 1 - \theta_{nk}\} = \theta_{nk}$. The other case is similar.

By the definition of the profile empirical likelihood ratio $\mathcal{R}^0_{nk}(\beta)$, we have the following constraints
\[
0_d = \sum_{i \in I} p_i g_i(\beta) = \sum_{j \in J} q_j g_j(\beta). \tag{19}
\]

We recall that, under hypothesis $H_0$, the expression of $p_i$ is given by (10), and it is equal to $(\theta_{nk} + n\lambda g_i(\beta))^{-1}$, for $i = 1, \ldots, n\theta_{nk}$. Then, by elementary calculations, we obtain
\[
0_d = \frac{1}{n\theta_{nk}} \sum_{i \in I} g_i(\beta) - \frac{1}{n\theta_{nk}} \sum_{i \in I} \frac{g_i(\beta)g_i^t(\beta)}{\lambda(\beta)} \lambda(\beta). \tag{20}
\]

Let us make the remark that we denote $\lambda$ by $\lambda(\beta)$ in order to indicate that for each value of $\beta$, solution of (20), we will have a different value for $\lambda$. We take $\beta = \beta^0 + (n\theta_{nk})^{-r} 1_d$, with $1_d$ the $d$-vector with all components 1 and $r > 0$ will be specified later. Therefore, $\|\beta - \beta^0\|_2 = (n\theta_{nk})^{-r} \to 0$, as $n\theta_{nk} \to \infty$.

For the first sum of the right-hand side of (20), by Lemma 3, we have
\[
\frac{1}{n\theta_{nk}} \sum_{i \in I} g_i(\beta) = O_P((n\theta_{nk})^{-1/2}) + V^0_{1n}(\beta - \beta^0) + o_P(\beta - \beta^0).
\]
Now, we consider the second term of the right-hand side of relation (20). From Proposition 1, we have that for all $\epsilon > 0$, there exists $M_1, M_2 > 0$, such that

$$P\left[\frac{1}{M_1} \sum_{i \in I} g_i(\beta)g_i^t(\beta) \leq \sum_{i \in I} \frac{g_i(\beta)g_i^t(\beta)}{\lambda_i(\beta)} \leq \frac{1}{M_2} \sum_{i \in I} g_i(\beta)g_i^t(\beta)\right] < \epsilon.$$ 

This implies that, in order to study the second term of the right-hand side of relation (20), we must study only $(n\theta_{nk})^{-1} \sum_{i \in I} g_i(\beta)g_i^t(\beta)$. By a Taylor’s expansion of $g_i(\beta)$ in a neighbourhood of $\beta^0$, using an argument similar to the one used for the first term of (20), together with the assumption (A3), we obtain

$$\frac{1}{n\theta_{nk}} \sum_{i \in I} g_i(\beta)g_i^t(\beta) = \frac{1}{n\theta_{nk}} \sum_{i \in I} g_i(\beta^0)g_i^t(\beta^0)(1 + o_P(1)). \tag{21}$$

Taking into account Lemma 3 and relation (21), the relation (20) becomes

$$0_d = \left[O_P((n\theta_{nk})^{-1/2}) + V_{1n}(\beta-\beta^0) - \frac{1}{n\theta_{nk}} \sum_{i=1}^{n\theta_{nk}} g_i(\beta^0)g_i^t(\beta^0)\lambda(\beta) \right](1 + o_P(1)) \tag{22}$$

We consider a constant $r$ such that $1/3 < r < 1/2$. If $\beta = \beta^0 + (n\theta_{nk})^{-r}1_d$, then $(\beta - \beta^0)^t1_d > 0$, and if $\beta = \beta^0 - (n\theta_{nk})^{-r}1_d$ then $(\beta - \beta^0)^t1_d < 0$. Then, the relation (22) implies

$$\lambda(\beta^0 \pm (n\theta_{nk})^{-r}1_d) = \pm \left[\theta_{nk} \left(1 - \sum_{i=1}^{n\theta_{nk}} g_i(\beta^0)g_i^t(\beta^0)\right)^{-1} V_{1n}(\beta^0)^{-r}1_d + O_P((n\theta_{nk})^{-1/2}) \right][1 + o_P(1)]. \tag{23}$$

For the observations $j \in J$, let us consider the function $v: \Gamma \rightarrow \mathbb{R}^d$ defined by

$$v(\beta) = \sum_{j \in J} q_j g_j(\beta) = \frac{1}{n - n\theta_{nk}} \sum_{j \in J} \frac{g_j(\beta)}{1 - \lambda_j(\beta)} V_{1n}(\beta) V_{2n}^{-1}(\beta) g_j(\beta).$$

Note that $v(\hat{\beta}(\theta_{nk})) = 0_d$. For $v(\beta)$, we have the following decomposition

$$\frac{V_{1n}(\beta) V_{2n}^{-1}(\beta)}{n(1 - \theta_{nk})^2} \sum_{j \in J} \frac{g_j(\beta)g_j^t(\beta)}{1 - \lambda_j(\beta)} V_{1n}(\beta) V_{2n}^{-1}(\beta) g_j(\beta) \lambda(\beta) + \frac{1}{n(1 - \theta_{nk})} \sum_{j \in J} g_j(\beta).$$

To facilitate writing, we consider the following $d \times d$ squares matrices, defined by

$$D_{1n}^0 = \frac{1}{n\theta_{nk}} \sum_{i \in I} g_i(\beta^0)g_i^t(\beta^0), \quad D_{2n}^0 = \frac{1}{n - n\theta_{nk}} \sum_{j \in J} g_j(\beta^0)g_j^t(\beta^0). \tag{24}$$
As for the observations $i \in I$, we obtain, similarly as for relation \((22)\),

\[
v(\beta) = [V_{2n}^0(\beta - \beta^0) + (1 - \theta_{nk})^{-1}V_{1n}^0]D_{2n}^0\lambda(\beta) + O_P((n(1 - \theta_{nk}))^{-1/2})]
\cdot (1 + o_P(1)).
\]

Replacing $\lambda(\beta)$ by the value obtained in \((23)\), we obtain

\[
v(\beta) = [V_{2n}^0(\beta - \beta^0) + (\theta_{nk})(1 - \theta_{nk})^{-1}V_{1n}^0]D_{2n}^0(D_{1n}^0)^{-1}V_{1n}^0(\beta - 
\beta^0) + O_P((n(1 - \theta_{nk}))^{-1/2}) + O_P((n\theta_{nk})^{-1/2})](1 + o_P(1)).
\]

Because $\beta = \beta^0 \pm (n\theta_{nk})^{-r}1_d$, $1/3 \leq r < 1/2$ and min\{\(\theta_{nk}, 1 - \theta_{nk}\)\} = $\theta_{nk}$, then $v(\beta)$ becomes

\[
[(V_{2n}^0 + \frac{\theta_{nk}}{1 - \theta_{nk}}V_{1n}^0(V_{2n}^0)^{-1}D_{2n}^0(D_{1n}^0)^{-1}V_{1n}^0)(\beta - \beta^0) + O_P((n\theta_{nk})^{-1/2})](1 + o_P(1)).
\]

\[(25)\]

This implies that $v(\beta^0 + (n\theta_{nk})^{-r}1_d)$ and $v(\beta^0 - (n\theta_{nk})^{-r}1_d)$ have a different signs, component by component. Moreover, because $v$ contains continuous functions in the neighbourhood of $\beta^0$, there exists a $\beta$ such that $v(\beta) = 0_d$.

But since $v(\beta(\theta_{nk})) = 0_d$, we have that $\beta(\theta_{nk}) \in [\beta^0 - (n\theta_{nk})^{-r}1_d, \beta^0 + (n\theta_{nk})^{-r}1_d]$, which implies, because $r < 1/2$, that $\beta(\theta_{nk}) - \beta^0 = O_P((n\theta_{nk})^{-r}) \geq O_P((n\theta_{nk})^{-1/2})$. This last relation, together with the relation \((23)\), since $\beta(\theta_{nk}) - \beta^0$ is the coefficient of a matrix strictly positive, implies that in order to have $v(\beta(\theta_{nk})) = 0_d$, we must have $\beta(\theta_{nk}) - \beta^0 = O_P((n\theta_{nk})^{-1/2})$. Considering this result, for the relation \((23)\), we obtain $\lambda(\beta(\theta_{nk})) = \theta_{nk}O_P((n\theta_{nk})^{-1/2})$.

The theorem is completely proved.

\[\blacksquare\]

Remark 1 In view of the proof of Theorem 2 under hypothesis $H_0$, we can consider instead of $Z_{nk}(\theta_{nk}, \lambda, \beta)$, given by \((22)\), the following modified statistic

\[
T_{nk}(\theta_{nk}, \lambda, \beta) = 2\sum_{i \in I} \log(1 + \frac{1}{\theta_{nk}}\lambda'g_i(\beta)) + 2\sum_{j \in J} \log(1 - \frac{1}{1 - \theta_{nk}}\lambda'g_j(\beta)).
\]

\[(26)\]

Because the regression function is nonlinear, in order to the maximum empirical likelihood always exists, we consider that the parameter $\theta_{nk} \in [\Theta_{1n}, \Theta_{2n}] \subset (0, 1)$, such that $n\Theta_{1n} \to \infty$, $n(1 - \Theta_{2n}) \to \infty$, as $n \to \infty$ for example. The reader can find a discussion concerning the possible values of $\Theta_{1n}$, $\Theta_{2n}$ in the papers \([32], [19]\).

Finally, the test statistic for testing the hypothesis $H_0$ against $H_1$ is

\[
\tilde{T}_n \equiv \max_{\theta_{nk} \in [\Theta_{1n}, \Theta_{2n}]} T_{nk}(\theta_{nk}, \hat{\lambda}(\theta_{nk}), \hat{\beta}(\theta_{nk})).
\]

\[(27)\]

Then, we can consider as estimator for the time of change $k^0$, the maximum empirical likelihood estimator: $\hat{k}_n \equiv n\hat{\theta}_n \equiv n\min\{\theta_{nk}; \theta_{nk} = \arg\max_{\theta_{nk} \in [\Theta_{1n}, \Theta_{2n}]} T_{nk}(\theta_{nk}, \hat{\lambda}(\theta_{nk}), \hat{\beta}(\theta_{nk}))\}$. Recall that $\hat{\lambda}(\theta_{nk})$ and $\hat{\beta}(\theta_{nk})$ are the solutions of the score equations \((13)\) and \((14)\).
The following result gives the asymptotic distribution of the test statistic \( \tilde{T}_n \) given by (27), under the null hypothesis of no-change. For this purpose, we consider functions:

\[
A(x) \equiv (2 \log x)^{1/2}, \quad D(x) = 2 \log x + \log \log x, \quad u(n) = \frac{1 - \theta_{1n} \theta_{2n}}{\sigma_{1n}(1 - \theta_{2n})} \to \infty \text{ as } n \to \infty.
\]

**Theorem 3** Under the assumptions (A1)-(A4), if the hypothesis \( H_0 \) is true, then we have, for all \( t \in \mathbb{R} \)

\[
\lim_{n \to \infty} P\{ A(\log u(n))(\tilde{T}_n)^{1/2} \leq t + D(\log u(n)) \} = \exp(-e^{-t}). \tag{28}
\]

**Proof.** The structure of the proof is similar to that of Theorem 1.3.1 of [9], with the modification that the Lemma 5 is used instead of Theorem 1.1.1 of [9], and the Theorem A.3.4 of [9] instead of Corollary A.3.1 of [9]. The details are omitted. \(\blacksquare\)

**Corollary 1** Consequence of this theorem, for a fixed size \( \alpha \in (0,1) \), we can deduce the critical test region:

\[
(\tilde{T}_n)^{1/2} \geq \frac{-\log(-\log \alpha) + D(\log u(n))}{A(\log u(n))}.
\]

Using Theorem 3 in applications is quite complicated. First, because we must first solve equation system (13) and (14) where the nonlinearity in parameter \( \beta \) up to and including in matrices \( V_{1n}(\beta), V_{2n}(\beta), V_{-1n}(\beta) \) causes numerical difficulties and long computation time. Moreover, then it must then find \( \theta_{nk} \) that maximizes statistic (27). We can propose an approached form for the test statistic much simpler to use in practice, but which preserves the theoretical properties of (27).

**Remark 2** Taking into account the last relation of Lemma 5, Theorem 3 implies that, in practice, for testing the hypothesis \( H_0 \) against \( H_1 \), we will use an approximate form

\[
T(\theta_{nk}) = n\sigma^{-2}\theta_{nk}(1-\theta_{nk})(W_{1n}^0 - W_{2n}^0)^T V^{-1}(W_{1n}^0 - W_{2n}^0)(1+oP(1)), \tag{29}
\]

where

\[
W_{1n}^0 = \frac{1}{n\theta_{nk}} \sum_{i \in I} g_i(\beta^0), \quad W_{2n}^0 = \frac{1}{n(1-\theta_{nk})} \sum_{j \in J} g_j(\beta^0). \tag{30}
\]

Since \( \beta^0 \) is unknown, in applications, we replace it with a consistent estimator, for example, the ordinary least square estimator, denoted by \( \hat{\beta}_{LS} \). Under \( H_0 \), error variance \( \sigma^2 \) is estimated by \( n^{-1} \sum_{i=1}^{n} [Y_i - f(X_i, \hat{\beta}_{LS})]^2 \) and matrix \( V \) by \( n^{-1} \sum_{i=1}^{n} f(X_i, \hat{\beta}_{LS}) f^T(X_i, \hat{\beta}_{LS}) \).

The approximated maximum empirical likelihood estimator for the time of change \( k^0 \) is

\[
\hat{k}_n = n\hat{\theta}_n = n \min\{ \hat{\theta}_{nk}; \hat{\theta}_{nk} = \arg \max_{\theta_{1n} \leq \theta_{nk} \leq \theta_{2n}} T(\theta_{nk}) \}.
\]
2.3.2 Asymptotic behaviour of $T_{nk}$ and $\hat{T}_{nk}$ under $H_1$

We consider now that the hypothesis $H_1$ is true. We will first prove that the maximum of statistic $T_{nk}$ converges to the maximum of its limit distribution. Then, we will show that statistic test $\hat{T}_{nk}$ is consistent (it has asymptotic power equal to 1). If $k^0$ is the true time of change, we denote by $\theta_{n0} = k^0/n$ and we suppose that $\theta_0 \equiv \lim_{n \to \infty} \theta_{n0}$, where $\theta_0 \in (0,1)$.

For $x \in \mathcal{T}$ and $e \in \mathcal{R}$, let $F_x(e)$ and $G_x(e)$ the conditional distributions of $g(X_i, \beta)$ when $X_i = x$ for $i \in I$ and $j \in J$, respectively. Let $\mathbb{I}_{\{\cdot\}}$ the indicator function. Recall that, the distribution function of $X$ is $H(x)$. For $x$ and $\theta$ fixed, we define

$$dP_x(e) \equiv (\theta \mathbb{I}_{\{\theta \leq \theta_0\}} + \theta_0 \mathbb{I}_{\{\theta > \theta_0\}})dF_x(e) + (\theta - \theta_0)\mathbb{I}_{\{\theta > \theta_0\}}dG_x(e),$$

$$dQ_x(e) \equiv ((1 - \theta) \mathbb{I}_{\{\theta \geq \theta_0\}} + (1 - \theta_0)\mathbb{I}_{\{\theta < \theta_0\}})dG_x(e) + (\theta - \theta)\mathbb{I}_{\{\theta < \theta_0\}}dF_x(e),$$

$$dR_x(e) \equiv \mathbb{I}_{\{\theta < \theta_0\}}dF_x(e) + \mathbb{I}_{\{\theta > \theta_0\}}dG_x(e).$$

Since under $H_0$, we proved that instead of EL statistic $[12]$, we can consider statistic $[26]$, let us define the following statistic

$$A_{nk}(\theta_{nk}) = T_{nk}(\theta_{nk}, \tilde{\lambda}(\theta_{nk}), \tilde{\beta}(\theta_{nk}))/\left(2n\right), A_n(0) = A_n(1) = 0,$$

with $T_{nk}$ given by relation $[26]$, and $\tilde{\lambda}(\theta_{nk}), \tilde{\beta}(\theta_{nk})$ solutions of the system

$$\begin{align*}
\frac{\partial T_{nk}(\theta_{nk}, \lambda, \beta)}{\partial \lambda} & = \sum_{i \in I} \frac{g_{i}(\beta)}{\lambda + g_{i}(\beta)} - \sum_{j \in J} \frac{1}{1 - \theta_{nk} - \lambda g_{j}(\beta)} = 0, \\
\frac{\partial T_{nk}(\theta_{nk}, \lambda, \beta)}{\partial \beta} & = \sum_{i \in I} \frac{g_{i}(\beta)}{\lambda + g_{i}(\beta)} - \sum_{j \in J} \frac{1}{1 - \theta_{nk} - \lambda g_{j}(\beta)} = 0.
\end{align*}$$

(32)

By a similar proof to that of Theorem $[2]$ under $H_0$, we have that

$$\tilde{\lambda}(\theta_{nk}) = \min\{\theta_{nk}, 1 - \theta_{nk}\}O_{F}(n\min\{\theta_{nk}, 1 - \theta_{nk}\}^{-1/2}),$$

$$\tilde{\beta}(\theta_{nk}) - \beta^0 = O_{F}(n\min\{\theta_{nk}, 1 - \theta_{nk}\}^{-1/2}).$$

(33)

For any $\lambda$ and $\beta$, let the function $K : \mathcal{T} \times \mathcal{R} \times (0, 1)$ defined by

$$K(x, e, \theta) = \theta + \lambda f(x, \beta)e - f(x, \beta) + f(x, \beta^0).$$

Let also

$$\psi(\theta, \lambda, \beta) = \int_{\mathcal{T}} \left( \int_{\mathcal{R}} \log K(x, e, \theta)dP_x(e) + \int_{\mathcal{R}} \log(1 - K(x, e, \theta))dQ_x(e) \right) dH(x) - \theta \log \theta - (1 - \theta) \log(1 - \theta).$$

(34)

We will prove by Theorem $[3]$ that $\psi$ is the limit process of $A_{nk}$, under $H_1$.

Then consider, for a fixed $\theta \in (0, 1)$, $\tilde{\lambda}(\theta), \tilde{\beta}(\theta)$ the solutions to the following score equations

$$\begin{align*}
z_1(\theta, \lambda, \beta) & = \int_{\mathcal{T}} \left( \int_{\mathcal{R}} \frac{g(x, \beta)}{K(x, e, \theta)}dP_x(e) - \int_{\mathcal{R}} \frac{g(x, \beta)}{1 - K(x, e, \theta)}dQ_x(e) \right) dH(x) = 0, \\
z_2(\theta, \lambda, \beta) & = \int_{\mathcal{T}} \left( \int_{\mathcal{R}} \frac{\lambda g(x, \beta)}{K(x, e, \theta)}dP_x(e) - \int_{\mathcal{R}} \frac{\lambda g(x, \beta)}{1 - K(x, e, \theta)}dQ_x(e) \right) dH(x) = 0.
\end{align*}$$

(35)
where, $z_1(\theta, \lambda, \beta) = \partial \psi(\theta, \lambda, \beta) / \partial \lambda$ and $z_2(\theta, \lambda, \beta) = \partial \psi(\theta, \lambda, \beta) / \partial \beta$.

We require the following assumptions for the next theorems:

(A5) The matrix
\[
\begin{pmatrix}
\frac{\partial^2 z_2(\theta, \lambda, \beta)}{\partial \lambda} & \frac{\partial^2 z_2(\theta, \lambda, \beta)}{\partial \beta} \\
\frac{\partial^2 z_1(\theta, \lambda, \beta)}{\partial \lambda} & \frac{\partial^2 z_1(\theta, \lambda, \beta)}{\partial \beta}
\end{pmatrix}
\]
is nonsingular for all $\theta \in (0, 1)$.

(A6) The two following integrals are applied component by component of the corresponding matrix.
\[
\int_{\mathbb{R}} \int_{\mathbb{R}} \left( \frac{g(x, \beta)}{K(x, \beta)} \right) \left( \frac{g'(x, \beta)}{|1-K(x, \beta)|} \right) d(F_x(e) + G_x(e)) < \infty,
\]
\[
\int_{\mathbb{R}} \int_{\mathbb{R}} \left( \frac{g(x, \beta)}{K(x, \beta)} \right) \left( \frac{g'(x, \beta)}{|1-K(x, \beta)|} \right) d(F_x(e) + G_x(e)) < \infty.
\]
Let $g_i(x, \beta)$ the $l$-th component of the vector $g(x, \beta)$. For all $1 \leq j, k \leq d$, we suppose that
\[
\int_{\mathbb{R}} \int_{\mathbb{R}} \left( \frac{1}{K(x, \beta)} \right) \left( \frac{\partial^2 g_i(x, \beta)}{\partial \beta_j \partial \beta_k} \right) d(F_x(e) + G_x(e)) < \infty.
\]

(A7) The functions $f(x, \beta)$ and $\hat{f}(x, \beta)$ are equicontinuous in $\beta$ on $\Gamma$.

Remark 3 A sufficient condition for the equicontinuity of the functions $f(x, \beta)$ and $\hat{f}(x, \beta)$ is that they are Lipschitzian with respect to $\beta$ on $\Gamma$.

Following theorem shows that if $\theta_{nk}$ converges to the true value $\theta_0$, then the maximum of the modified EL test statistic converges to the maximum of its limit distribution.

Theorem 4 Under the alternative hypothesis $H_1$, if the assumptions (A5)-(A7) are satisfied, $\theta_0 \in (0, 1)$ and $\lim_{n \to \infty} \theta_{nk} = \theta \in [0, 1]$, then $\Lambda_{nk}(\theta_{nk}) \xrightarrow{n \to \infty} \psi(\theta, \tilde{\lambda}(\theta), \tilde{\beta}(\theta))$, where $\psi(\theta, \tilde{\lambda}(\theta), \tilde{\beta}(\theta))$ is a strictly increasing function on $(0, \theta_0)$ decreasing on $(\theta_0, 1)$ and $\max_{0 < \theta < 1} \psi(\theta, \tilde{\lambda}(\theta), \tilde{\beta}(\theta)) = \psi(\theta_0, \tilde{\lambda}(\theta_0), \tilde{\beta}(\theta_0)) = 0$.

Proof. We will prove this theorem in three steps.

Step 1. We first prove that, for all fixed $\theta \in (0, 1)$, we have
\[
\arg \max_{(\lambda, \beta)} T_{nk}(\theta, \lambda, \beta) \xrightarrow{n \to \infty} \arg \max_{(\lambda, \beta)} \psi(\theta, \lambda, \beta).
\]

Obviously, by the law of large numbers, for all $(\theta, \lambda, \beta) \in (0, 1) \times \mathbb{R} \times \Gamma$, we have $(2n)^{-1} T_{nk}(\theta, \lambda, \beta) \xrightarrow{n \to \infty} \psi(\theta, \lambda, \beta)$. On the other hand, by the assumption (A5), $\arg \max_{(\lambda, \beta)} \psi(\theta, \lambda, \beta)$ is the unique solution of the system (35). Seen the assumptions (A6) and (A7), the function $(2n)^{-1} T_{nk}(\theta, \lambda, \beta)$ is equicontinuous and bounded in $\lambda$ and $\beta$. Then, using Theorem 1.12.1 of [27], we have that the convergence of $(2n)^{-1} T_{nk}(\theta, \lambda, \beta)$ to $\psi(\theta, \lambda, \beta)$ is uniform in $(\lambda, \beta)$. Taking
into account that the solution of system (32) is unique, we obtain relation (36).

Step 2. We show that

$$\max_{\theta_{nk}} \Lambda_{nk} \xrightarrow{n \to \infty} \max_{\theta} \psi(\theta, \tilde{\lambda}(\theta), \tilde{\beta}(\theta)), \tag{37}$$

with $\tilde{\lambda}(\theta)$ and $\tilde{\beta}(\theta)$ the solutions of score equations (35). By similar calculations as in the proof of Theorem 2, taking into account the Step 1, we can show that, for $\theta = \lim_{n \to \infty} \theta_{nk}$, we have

$$\Lambda_{nk} = \frac{1}{n} \log \left[ \frac{1}{n} \sum_{\theta \in \Theta} \left[ (1 - \theta) - \hat{\lambda}'(\theta) g_{j}(\hat{\beta}(\theta)) \right] \right]$$

with $\hat{\lambda} = \frac{1}{n} \sum_{\theta \in \Theta} \left[ (1 - \theta) - \hat{\lambda}'(\theta) g_{j}(\hat{\beta}(\theta)) \right]$ and $\hat{\beta}(\theta)$ the solutions of score equations (35). By similar calculations as in the proof of Theorem 2, taking into account the Step 1, we can show that, for $\theta = \lim_{n \to \infty} \theta_{nk}$, we have

$$\Lambda_{nk} = \frac{1}{n} \log \left[ \frac{1}{n} \sum_{\theta \in \Theta} \left[ (1 - \theta) - \hat{\lambda}'(\theta) g_{j}(\hat{\beta}(\theta)) \right] \right]$$

with $\Lambda_{nk} \xrightarrow{n \to \infty} \psi(\theta, \hat{\lambda}(\theta), \hat{\beta}(\theta))$, where $\theta = \lim_{n \to \infty} \theta_{nk}$.

For $\theta \notin \{0, 1, \theta_0\}$, partial derivative $\partial \psi(\theta, \lambda, \beta)/\partial \theta$ becomes

$$\partial \psi(\theta, \lambda, \beta) = \int_{\mathbb{R}} \int_{\mathbb{R}} \left[ \log K(x, e, \theta) \mathbb{I}_{\{\theta < \theta_0\}} \int_{\mathbb{R}} K(x, e, \theta) dF(e) \right]$$

$$- \int_{\mathbb{R}} \left[ \log(1 - K(x, e, \theta)) \mathbb{I}_{\{\theta < \theta_0\}} \int_{\mathbb{R}} (1 - K(x, e, \theta)) dG(e) \right]$$

$$+ \log(1 - \theta) - \log(1 - \theta). \tag{38}$$

On the other hand, we have that $dR_\lambda(e) = \mathbb{I}_{\{\theta < \theta_0\}} dF(e) + \mathbb{I}_{\{\theta > \theta_0\}} dG(e)$, hence

$$\partial \psi(\theta, \lambda, \beta) = \int_{\mathbb{R}} \int_{\mathbb{R}} \left[ \log K(x, e, \theta) - \log(1 - K(x, e, \theta)) \right] dR_\lambda(e) dH(x).$$

Because $\tilde{\lambda}'(\theta) g(x, \tilde{\beta}(\theta)) = K(x, e, \theta) - \theta$ and $z_2(\theta, \tilde{\lambda}(\theta), \tilde{\beta}(\theta)) = 0_d$, we obtain

$$\int \left[ (1 - K(x, e, \theta)) dP_x(e) + \int (1 - \frac{1}{K(x, e, \theta)}) dQ_x(e) \right] dH(x) = 0. \tag{39}$$

Since $\int_{\mathbb{R}} dP_x(e) + \int_{\mathbb{R}} dQ_x(e) \right] dH(x) = 1$, relation (38) becomes

$$1 - \theta \int \left[ \frac{dP_x(e)}{K(x, e, \theta)} - \frac{dQ_x(e)}{1 - K(x, e, \theta)} \right] dH(x) - \int \left[ \frac{dP_x(e)}{K(x, e, \theta)} - \frac{dQ_x(e)}{1 - K(x, e, \theta)} \right] dH(x) = 0. \tag{40}$$

This relation is true for all $\theta \in (0, 1)$. If we take $\theta = 0$ and afterward $\theta = 1$, relation (39) implies

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{dP_x(e)}{K(x, e, \theta)} dH(x) = \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{dQ_x(e)}{1 - K(x, e, \theta)} dH(x) = 1. \tag{40}$$
The relation (37) is proved in a similar way as the proof of Theorem 3.2 of [16], using relations (39) and (41).

Step 3. Similar as in the proof of Theorem 3.2 of [16], we prove that 
\[ \psi(\theta_0, \hat{\lambda}(\theta_0), \hat{\beta}(\theta_0)) = 0, \]
and that for all \( \gamma \in (0, \min(\theta_0, 1 - \theta_0)) \), we have
\[ \max_{|k-n\theta_0| \geq n\gamma} A_{nk}(\theta_{nk}) \xrightarrow{a.s.} \max_{n \to \infty} \psi(\theta, \hat{\lambda}(\theta), \hat{\beta}(\theta)). \quad (41) \]
Which implies \( \lim_{n \to \infty} \mathbb{P}[\arg \max_k A_{nk}(\theta_{nk}) - \theta_0 | \geq \gamma] = 0. \)

Corollary 2 The proof of Theorem 4 implies that the estimator of \( \theta_0 \) defined by
\[ \hat{\theta}_n \equiv \min \{ \hat{\theta}_{nk}; \hat{\theta}_{nk} = \arg \max_{\theta_{nk} \in \Theta_{1n}, \Theta_{2n}} T_{nk}(\theta_{nk}, \hat{\lambda}(\theta_{nk}), \hat{\beta}(\theta_{nk})) \} \]
satisfies the property that \( \hat{\theta}_n - \theta_n \xrightarrow{P} 0 \). Taking into account Remark 3, we have also \( \hat{\theta}_n - \theta_n \to 0 \) in probability, with \( \hat{\theta}_n \) the estimator of \( \theta_0 \) defined in Remark 3.

We prove by the following theorem that test statistic \( \tilde{T}_n \) given by (27) has the asymptotic power equal to 1.

Theorem 5 Under assumptions (A1)-(A7), the power of the empirical likelihood ratio test \( \tilde{T}_n \) converges to 1.

Proof. By the proof of Theorem 4, we have under hypothesis \( H_1 \)
\[ \max_{\theta_{nk} \in \Theta_{1n}, \Theta_{2n}} T_{nk}(\theta_{nk}, \hat{\lambda}(\theta_{nk}), \hat{\beta}(\theta_{nk})) \xrightarrow{a.s.} 0. \]
Taking into account relation (33), we have
\[ \tilde{T}_n \xrightarrow{a.s.} 0. \quad (42) \]
In the other hand, if we suppose that the hypothesis \( H_0 \) is true, by Theorem 3, we have for all \( t \in \mathbb{R} \),
\[ \mathbb{P}[A(\log u(n))\tilde{T}_n^2 \leq t + D(\log u(n))] = \exp(-e^{-t}). \]
Taking in the last relation \( t = -\log \log u(n) \), we obtain
\[ \lim_{n \to \infty} \mathbb{P}[A(\log u(n))(\tilde{T}_n/n)^2 \leq - \frac{\log \log u(n)}{\sqrt{n}} + \frac{D(\log u(n))}{\sqrt{n}}] = 0. \quad (43) \]
The theorem follows.

We emphasise that, similar results to Theorems 3, 4 and 5 were obtained by other authors for simpler models. The reader can find the corresponding results for the test to detect a change in distribution sequence in [22] and for detecting a change in the parameters of a linear model in [16].
3 Extension to a particular two change-points model

In this section, we consider the epidemic model. The epidemic linear model by empirical likelihood test was considered in paper [19]. In a previous paper, [29] detect an epidemic alternative in the mean value of a sequence of independent normal random variables by various test statistics: likelihood ratio, recursive residual, score-like, semi-likelihood ratio. The works [24] and [25] studied by likelihood ratio, the epidemic changes in the mean of a sequence of exponential random variables and respectively, in the shape parameter of a sequence of gamma random variables.

We assume under alternative hypothesis, denoted $H_2$, that the model have two change-points $k_1$ and $k_2$ ($1 < k_1 < k_2 < n$), such that the model of the first and the third segment is the same. More specifically, the regression model can be written

$$H_2: \mathbf{Y}_i = \begin{cases} f(\mathbf{X}_i, \beta_1) + \varepsilon_i & i = 1, \ldots, k_1 \\ f(\mathbf{X}_i, \beta_2) + \varepsilon_i & i = k_1 + 1, \ldots, k_2 \\ f(\mathbf{X}_i, \beta_1) + \varepsilon_i & i = k_2 + 1, \ldots, n. \end{cases} \tag{44}$$

Therefore, we want to test the null hypothesis $H_0$ of no-change given by (1), against the alternative hypothesis $H_2$ given by (44).

Under the hypothesis $H_2$, we consider the following two sets, $I' = \{1, \ldots, k_1, k_2, \ldots, n\}$ and $J' = \{k_1 + 1, \ldots, k_2\}$. We define the corresponding probability vectors $(u_1, \ldots, u_{k_1}, u_{k_2+1}, \ldots, u_n)$ and $(v_{k_1+1}, \ldots, v_{k_2})$, where $u_i \equiv P[Y_i = y_i]$ and $v_j \equiv P[Y_j = y_j]$ denotes the probability to observe the value $y_i$ (respectively $y_j$), for the dependent variable $Y_i$ (respectively $Y_j$), for $i \in I'$ and $j \in J'$. Obviously, these probabilities satisfy the relations $\sum_{i \in I'} u_i = 1$ and $\sum_{j \in J'} v_j = 1$.

Under hypothesis $H_0$, the profile EL ratio for $\beta$ is

$$U'_{0,n,k_1,k_2}(\beta) = \sup_{(u_1, \ldots, u_{k_1}, u_{k_2+1}, \ldots, u_n) \ (v_{k_1+1}, \ldots, v_{k_2})} \ \left\{ \prod_{i \in I'} (n - k_2 + k_1) u_i \prod_{j \in J'} (k_2 - k_1) v_j; \right.$$  

$$\sum_{i \in I'} u_i = \sum_{j \in J'} v_j = 1, \sum_{i \in I'} u_i g_i(\beta) = \sum_{j \in J'} v_j g_j(\beta) = 0 \right\}.$$

Under hypothesis $H_2$, the profile EL ratio for $\beta_1, \beta_2$ has the form

$$U'_{1,n,k_1,k_2}(\beta_1, \beta_2) = \sup_{(u_1, \ldots, u_{k_1}, u_{k_2+1}, \ldots, u_n) \ (v_{k_1+1}, \ldots, v_{k_2})} \ \left\{ \prod_{i \in I'} (n - k_2 + k_1) u_i \prod_{j \in J'} (k_2 - k_1) v_j; \right.$$  

$$\sum_{i \in I'} u_i g_i(\beta_1) = \sum_{j \in J'} v_j g_j(\beta_2) = 0 \right\}.$$

Then, in order to test $H_0$ against $H_2$, we consider the profile EL ratio $U'_{0,n,k_1,k_2}(\beta)/U'_{1,n,k_1,k_2}(\beta_1, \beta_2)$. 

\[ \text{null hypothesis, denoted } H_0, \text{ that the model have two change-points } k_1 	ext{ and } k_2 (1 < k_1 < k_2 < n), \text{ such that the model of the first and the third segment is the same. More specifically, the regression model can be written } \]

\[ H_2: \mathbf{Y}_i = \begin{cases} f(\mathbf{X}_i, \beta_1) + \varepsilon_i & i = 1, \ldots, k_1 \\ f(\mathbf{X}_i, \beta_2) + \varepsilon_i & i = k_1 + 1, \ldots, k_2 \\ f(\mathbf{X}_i, \beta_1) + \varepsilon_i & i = k_2 + 1, \ldots, n. \end{cases} \tag{44} \]

Therefore, we want to test the null hypothesis $H_0$ of no-change given by (1), against the alternative hypothesis $H_2$ given by (44).

Under the hypothesis $H_2$, we consider the following two sets, $I' = \{1, \ldots, k_1, k_2, \ldots, n\}$ and $J' = \{k_1 + 1, \ldots, k_2\}$. We define the corresponding probability vectors $(u_1, \ldots, u_{k_1}, u_{k_2+1}, \ldots, u_n)$ and $(v_{k_1+1}, \ldots, v_{k_2})$, where $u_i \equiv P[Y_i = y_i]$ and $v_j \equiv P[Y_j = y_j]$ denotes the probability to observe the value $y_i$ (respectively $y_j$), for the dependent variable $Y_i$ (respectively $Y_j$), for $i \in I'$ and $j \in J'$. Obviously, these probabilities satisfy the relations $\sum_{i \in I'} u_i = 1$ and $\sum_{j \in J'} v_j = 1$.

Under hypothesis $H_0$, the profile EL ratio for $\beta$ is

$$U'_{0,n,k_1,k_2}(\beta) = \sup_{(u_1, \ldots, u_{k_1}, u_{k_2+1}, \ldots, u_n) \ (v_{k_1+1}, \ldots, v_{k_2})} \ \left\{ \prod_{i \in I'} (n - k_2 + k_1) u_i \prod_{j \in J'} (k_2 - k_1) v_j; \right.$$  

$$\sum_{i \in I'} u_i = \sum_{j \in J'} v_j = 1, \sum_{i \in I'} u_i g_i(\beta) = \sum_{j \in J'} v_j g_j(\beta) = 0 \right\}.$$

Under hypothesis $H_2$, the profile EL ratio for $\beta_1, \beta_2$ has the form

$$U'_{1,n,k_1,k_2}(\beta_1, \beta_2) = \sup_{(u_1, \ldots, u_{k_1}, u_{k_2+1}, \ldots, u_n) \ (v_{k_1+1}, \ldots, v_{k_2})} \ \left\{ \prod_{i \in I'} (n - k_2 + k_1) u_i \prod_{j \in J'} (k_2 - k_1) v_j; \right.$$  

$$\sum_{i \in I'} u_i g_i(\beta_1) = \sum_{j \in J'} v_j g_j(\beta_2) = 0 \right\}.$$

Then, in order to test $H_0$ against $H_2$, we consider the profile EL ratio $U'_{0,n,k_1,k_2}(\beta)/U'_{1,n,k_1,k_2}(\beta_1, \beta_2)$.
Similarly as in Section 2 when we tested a single change-point, using Lagrange multipliers, we obtain that under hypothesis $H_0$, the probabilities $u_i$, $v_j$ are

$$u_i = \frac{1}{(n-k_2+k_1) + n\lambda_1^* g_i(\beta)} \quad v_j = \frac{1}{(k_2-k_1) - n\lambda_2^* g_j(\beta)}.$$ \hfill (45)

Using the similar arguments as in the proof of Theorem 1 we deduce that the asymptotic distribution of $-2 \log U_{1,n,k_1,k_2}^2(\beta_1, \beta_2)$ is $\chi^2(3d)$ and then we can consider the test statistic $-2 \log U_{0,n,k_1,k_2}^2(\beta)$. We restricted to the case where $\lambda_1$ and $\lambda_2$ satisfy the constraint $\tilde{V}_{1n}(\beta) \lambda_1 = \tilde{V}_{2n}(\beta) \lambda_2$, with $\tilde{V}_{1n}(\beta) = (n + k_1 - k_2)^{-1} \sum_{i \in I'} \tilde{g}(\beta)$ and $\tilde{V}_{2n}(\beta) = (k_2 - k_1)^{-1} \sum_{j \in J'} \tilde{g}(\beta)$. In this case, considering the parameter $\hat{\theta}_{n,k_1,k_2} = n^{-1}(n-k_2+k_1)$, that depends on two change-points $k_1$, $k_2$, we will consider the test statistic

$$2 \sum_{i \in I'} \log \left[ 1 + \frac{1}{\hat{\theta}_{n,k_1,k_2}} \lambda_1^* g_i(\beta) \right] + 2 \sum_{j \in J'} \log \left[ 1 - \frac{1}{1 - \hat{\theta}_{n,k_1,k_2}} \lambda_1^* \tilde{V}_{1n}(\beta) \tilde{V}_{2n}^{-1}(\beta) g_j(\beta) \right].$$ \hfill (46)

Let us denote by $\hat{\lambda}(\hat{\theta}_{n,k_1,k_2})$, $\hat{\beta}(\hat{\theta}_{n,k_1,k_2})$ the solutions of the score equations of this random process equal to zero. We can show, as in Section 2, that statistic (46) is, under hypothesis $H_0$, asymptotically equivalent to the statistic

$$U_{n,k_1,k_2}(\hat{\theta}_{n,k_1,k_2}, \lambda, \beta) = 2 \sum_{i \in I'} \log \left[ 1 + \frac{1}{\hat{\theta}_{n,k_1,k_2}} \lambda_1^* g_i(\beta) \right] + 2 \sum_{j \in J'} \log \left[ 1 - (1 - \hat{\theta}_{n,k_1,k_2})^{-1} \lambda_1^* g_j(\beta) \right].$$

Then, we will consider for testing null hypothesis $H_0$ against $H_2$ the test statistic

$$\max_{1 < k_1 < k_2 < n} \left\{ U_{n,k_1,k_2}(\hat{\theta}_{n,k_1,k_2}, \hat{\lambda}(\hat{\theta}_{n,k_1,k_2}), \hat{\beta}(\hat{\theta}_{n,k_1,k_2})) \right\}.$$

In the case when $k_1$ or $k_2 - k_1$ have a small value, the maximum empirical likelihood may not exist. In this case, the proposed test may not detect the presence of change in the model. For the empirical likelihood maximum always exists, we consider two natural numbers $\Theta_{n1}$ and $\Theta_{n2}$, such as $\Theta_{n1} < k_1 < k_2 < n - \Theta_{n2}$. Finally, the test statistic for testing $H_0$ against $H_2$ becomes

$$\max_{\Theta_{n1} < k_1 < k_2 < n - \Theta_{n2}} \left\{ U_{n,k_1,k_2}(\hat{\theta}_{n,k_1,k_2}, \hat{\lambda}(\hat{\theta}_{n,k_1,k_2}), \hat{\beta}(\hat{\theta}_{n,k_1,k_2})) \right\}.$$

In order to facilitate the practical utilization of the test statistic, we can easily obtain the corresponding statistic given in Remark 2 by relation (29).

4 Simulation study

In this section, we report a simulation study by Monte Carlo method, in order to evaluate the performance of the proposed test statistics. Firstly, for a fixed theoretical size, we calculate the critical value of the test statistics,
for different values of \( n \). Afterward, by Monte Carlo technique, we calculate empirical test size, empirical test power and the estimation of the change-point localisation. This study was conducted firstly for a nonlinear model with a single change-point and secondly for an epidemic model. The obtained results by the proposed test statistic are compared with those obtained from LS method, proposed by [4].

All simulations were performed using the R language. The program codes are available from the authors.

We consider the nonlinear function

\[
 f(x, \beta) = a \frac{1 - x^b}{b},
\]

with \( \beta = (a, b) \in [-100, 100] \times [0.1, 20] \). The same model was considered in [6], where the model was estimated by the penalized least absolute deviation method.

### 4.1 Model with a single change-point

For the nonlinear function of (47), the following two-phase (one change-point) nonlinear model is considered under \( H_1 \)

\[
 Y_i = a_1 \frac{1 - X_i^{b_1}}{b_1} \mathbb{I}_{i \leq k_0} + a_2 \frac{1 - X_i^{b_2}}{b_2} \mathbb{I}_{i > k_0} + \varepsilon_i, \quad i = 1, \ldots, n \quad (48)
\]

with \( X_i = i/1000, n = 1000 \) and true value of parameters \( a_1^0 = 10, b_1 = 2, a_2^0 = 7, b_2 = 1.75 \). Under hypothesis \( H_0 \), the true parameters are \( a^0 = 10, b^0 = 2 \).

The change absence against one-change in model is tested using the (approximated) maximum empirical likelihood statistic \( T(\theta_{nk}) \) given by (29).

In order to calculate the empirical test size, an without change-point model is considered and we count, the number of times, on the Monte Carlo replications when we obtain \( \max_{\theta} T^{1/2}(\theta_{nk}) \geq c_\alpha \). For a fixed size \( \alpha \in (0, 1) \), critical value \( c_\alpha \) is calculated in accordance with Corollary [11]:

\[
 c_\alpha = \frac{-\log(-\log \alpha) + D(\log u(n))}{A(\log u(n))}.
\]

For theoretical size \( \alpha = 0.05 \), we first calculate critical values \( c_\alpha \), varying the sample size \( n \) from 200 to 1000 (see Table 1).

For model (48) with Gaussian standardized errors, 500 Monte Carlo replications were performed. We also present in Table 1 the empirical power, using statistic test (29). For different position of change-point, the asymptotic test power is 1. We fix sample size \( n = 1000 \), theoretical test size \( \alpha = 0.05 \) and we vary the error distribution. In order to calculate the empirical size of test (type I error probabilities), 500 Monte Carlo replications are realized for different error distributions: \( \varepsilon_i \sim \mathcal{N}(0, 1) \),
Empirical likelihood test in a posteriori change-point non linear model

Table 1 Critical values \( c_{\alpha} \), for \( \alpha=0.05 \). Empirical power on 500 Monte Carlo replications, when \( \varepsilon \sim N(0,1) \).

| \( n \) |  \( k_0 \) | \( c_{\alpha} \) | power |
|---|---|---|---|
| 1000 | 600 | 1.544 | 1 |
| 800  | 500 | 1.492 | 1 |
| 600  | 400 | 1.434 | 1 |
| 400  | 250 | 1.340 | 1 |
| 200  | 75  | 1.133 | 1 |

\[ \varepsilon_i = 2 \exp(2) - 1, \quad \varepsilon_i = 1/\sqrt{6} (\chi^2(3) - 3) \quad \text{and} \quad \varepsilon_i = 2/\sqrt{6} t(6), \] where \( N(0,1) \), \( \exp(2) \), \( \chi^2(3) \) and \( t(6) \) are standard normal distribution, exponential distribution with mean 1/2, chi-square distribution with degree of freedom 3 and Student distribution with degree of freedom 6, respectively. In all cases, except for Student distribution (when the empirical size is slightly larger than 0.05), the empirical size is 0 (see Table 2). For the same four error distributions, but for model with a change-point in \( k_0 \), by 500 Monte Carlo model replications, for different change-point location: \( k_0 \in \{200,400,600,800\} \), we obtain that the empirical power is 1, in any case.

Table 2 Empirical size for four error distributions on 500 Monte Carlo replications, \( \alpha=0.05 \).

| \( n \) | error distribution | Normal | Exponential | \( \chi^2 \) | Student |
|---|---|---|---|---|
| 200 | 0 | 0 | 0 | 0 |
| 1000| 0 | 0 | 0 | 0 |

As mentioned in Remark 2, one can also estimate the change-point location by EL method. In table 3 we have the summarized results (minimum, maximum, mean, standard-deviation, median) for the estimator \( \hat{k}_n \), given by Remark 2 by 500 Monte Carlo replication. In view of the results presented in Table 3, for different error distribution and for different positions of the change in the interval, we deduce that the proposed estimation method approaches very well the true value \( k_0 \), regardless of the error distribution and of the change-point position on the interval \([1:n]\). Note that, in all situations the median and the mean of the change-point estimations coincide or is very close to the true value.
Table 3  Descriptive statistics for the estimators of the change-point. Model with two phases by EL method, \( n = 1000, \) 500 Monte Carlo replications.

| error distribution | \( k^0 \) | \( \min(k_n) \) | \( \max(k_n) \) | \( \text{mean}(k_n) \) | \( \text{sd}(k_n) \) | \( \text{median}(k_n) \)
|--------------------|----------|----------------|----------------|----------------|----------------|----------------|
| \( \varepsilon_i \sim N(0, 1) \) | 200      | 186            | 205            | 197            | 4              | 199            |
|                    | 400      | 388            | 416            | 400            | 4              | 400            |
|                    | 600      | 585            | 612            | 600            | 6              | 600            |
|                    | 800      | 780            | 819            | 795            | 9              | 797            |
| \( \varepsilon_i \sim \frac{2}{\sqrt{6t}}(6) \) | 200      | 180            | 209            | 199            | 5              | 200            |
|                    | 400      | 390            | 407            | 400            | 4              | 400            |
|                    | 600      | 580            | 611            | 599            | 6              | 600            |
|                    | 800      | 780            | 815            | 797            | 10             | 798            |
| \( \varepsilon_i \sim \frac{2}{\sqrt{(2x)}p(2) - 1} \) | 200      | 183            | 204            | 199            | 4              | 200            |
|                    | 400      | 391            | 412            | 401            | 5              | 400            |
|                    | 600      | 591            | 609            | 600            | 4              | 600            |
|                    | 800      | 780            | 817            | 795            | 5              | 799            |
| \( \varepsilon_i \sim \frac{1}{\sqrt{6(\chi^2(3) - 3)}} \) | 200      | 188            | 210            | 200            | 4              | 200            |
|                    | 400      | 393            | 417            | 401            | 5              | 400            |
|                    | 600      | 585            | 614            | 600            | 6              | 600            |
|                    | 800      | 780            | 807            | 795            | 8              | 797            |

4.2 Epidemic model

For nonlinear function of (47), under hypothesis \( H_2 \), we consider the following three-phase (two change-points) model

\[
Y_i = a_1 \frac{1 - X_i^{b_1}}{b_1} \mathbb{I}_{1 \leq i \leq k_1} + a_2 \frac{1 - X_i^{b_2}}{b_2} \mathbb{I}_{k_1 + 1 \leq i \leq k_2} + a_3 \frac{1 - X_i^{b_1}}{b_1} \mathbb{I}_{k_2 + 1 \leq i \leq n} + \varepsilon_i, \tag{49}
\]

with \( X_i = i/1000, \) \( n = 1500 \) and the true value of parameters \( a_1^0 = 10, b_1^0 = 2, a_2^0 = 7, b_2^0 = 1.75. \) Under null hypothesis \( H_0 \) the true parameters are \( a^0 = 10, b^0 = 2. \)

In Table 4 we give results after 150 Monte Carlo replications in order to calculate the empirical power of test, for \( n = 1500. \) We deduce that empirical size is zero and empirical test power is 1.

Table 4  Empirical powers and empirical size for epidemic model, \( \alpha = 0.05, n = 1500. \)

| \( k_1 \) | \( k_2 \) | power |
|----------|----------|-------|
| no-change |         |       |
| 100      | 900      | 1     |
| 200      | 500      | 1     |
| 400      | 600      | 1     |
| 600      | 900      | 1     |
Table 5: Descriptive statistics for the estimators of the change-point. Model with two phases by LS method, \( n = 1000 \), 500 Monte Carlo replications.

| error distribution | \( k^0 \) | \( k_n \) | \( \min(k_n) \) | \( \max(k_n) \) | \( \text{mean}(k_n) \) | \( \text{sd}(k_n) \) | \( \text{median}(k_n) \) |
|--------------------|---------|---------|-------------|-------------|----------------|-------------|-------------|
| \( \varepsilon_i \sim \mathcal{N}(0,1) \) | 200     | ???    | ???         | ???         | ???           | ???         | ???         |
|                    | 400     | 396    | 400         | 399         | 1             | 400         | 400         |
|                    | 600     | 595    | 605         | 600         | 3             | 600         | 3           |
|                    | 800     | ???    | ???         | ???         | ???           | ???         | ???         |

4.3 Comparison with LS test

On data considered in subsection 4.1 for \( \varepsilon \sim \mathcal{N}(0,1) \) and \( n = 1000 \) we apply the method proposed by [4], where the estimation method and the associated test is by least squares. This study is realized by computing the test statistic \( \sup F(0 : 1) \) given in [4]. Under hypothesis \( H_1 \), when the model has a change-point in \( k^0 = 600 \), we realize 500 Monte Carlo simulations. We obtain that the test statistic value always exceeds the critical value of 12.85 (see [3]). Then, by LS test of [4], the null hypothesis \( H_0 \) is always rejected and hence the power of test is 1. Whereas if we generate the values \( Y_i \) without change-point for gaussian errors, then, the test statistic value of [4] always exceeds critical. Hence the empirical size of the test proposed by [4] is 1, a result significantly worse than that obtained by our test. We note that (see Table 5) if under \( H_1 \) the true change-point is off-centered in the measurement interval, because of the function nonlinearity, then numerical problem arise for the LS estimation method. This is symbolized by “???” in Table 5. The same problem appears when the errors are not gaussian, regardless of the position of the change-point in the measurement interval. In contrast, we have seen that the EL test works for any error distribution and any change-point position.

5 Appendix

The following lemma will be used in the proof of propositions, theorems and of other lemmas.

**Lemma 1** Let \( X = (X_1, \ldots, X_p) \) a random vector (column), with the random variables \( X_1, \ldots, X_p \) not necessarily independent, and \( M = (m_{ij})_{1 \leq i,j \leq p} \), such that \( M = XX^T \). If for \( j=1, \ldots, p \), we have

\[
\text{for all } \eta_j > 0, \text{there exists } \delta_j > 0 \text{ such that } P[|X_j| \geq \delta_j] \leq \eta_j, \tag{50}
\]

then

(i) \( P[\|X\|_1 \geq p \max_{1 \leq j \leq p} \delta_j] \leq \max_{1 \leq j \leq p} \eta_j \),

(ii) \( P\left(\|X\|_2 \geq \sqrt{p} \max_{1 \leq j \leq p} \delta_j\right) \leq \max_{1 \leq j \leq p} \eta_j \),

(iii) \( P\left(\|M\|_1 \geq p \max_{1 \leq i,j \leq p} \left\{\delta_i^2, \delta_j^2\right\}\right) \leq \max_{1 \leq i,j \leq p} \left\{\eta_i^2, \eta_j^2\right\} \),

where \( \|M\|_1 = \max_{1 \leq i \leq p} \left\{\sum_{j=1}^p |m_{ij}| \right\} \) is the subordinate norm to the vector norm \( \|. \|_1 \).
Proof of Lemma 2 (i) Using the relation (51), we can write
\[
P_n(\|x\|_1 \geq p \max_{1 \leq j \leq p} \delta_j) \leq P(p \max_{1 \leq j \leq p} |X_j| \geq \max_{1 \leq j \leq p} \eta_j).
\]

(ii) The relation (51) is equivalent to \(P[X_j^2 \geq \delta_j^2] \leq \eta_j\), which implies that
\[
P_n(\|X\|_2^2 \geq p \max_{1 \leq j \leq p} \delta_j^2) = P(p \max_{1 \leq j \leq p} X_j^2 \geq \max_{1 \leq j \leq p} \delta_j^2) \leq \max_{1 \leq j \leq p} \eta_j.
\]

(iii) For \(1 \leq i, j \leq p\), we have
\[
P[X_i X_j] \geq \max\{\delta_i^2, \delta_j^2\} \leq P[p \max\{X_i^2, X_j^2\} \geq \max\{\delta_i^2, \delta_j^2\}] \leq \max\{\eta_i^2, \eta_j^2\}.
\]
Then, \(P[p \min_{1 \leq i \leq p} m_{ij} \geq \max\{\delta_i^2, \delta_j^2\}] \leq \max\{\eta_i^2, \eta_j^2\}\). Hence, for each \(1 \leq j \leq p\),
\[
P[P[p \min_{1 \leq i \leq p} m_{ij} \geq \max\{\delta_i^2, \delta_j^2\}] \leq \max\{\eta_i^2, \eta_j^2\}.
\]

Lemma 2 Let the \(\eta\)-neighbourhood of \(\beta^0\), \(\mathcal{V}_\eta(\beta^0) = \{\beta \in \Gamma; \|\beta - \beta^0\|_2 \leq \eta\}\), with \(\eta \to 0\). Then, under assumptions (A1)-(A4), for all \(\epsilon > 0\), there exists a positive constant \(M > 0\), such that, for all \(\beta \in \mathcal{V}_\eta(\beta^0),\)
\[
P[\|g(x)\|_1 \geq M] \leq \epsilon.
\]

Proof of Lemma 2 In the following, for simplicity, we denote the functions \(f(x_i, \beta)\) by \(f_i(\beta)\), and \(f(X_i, \beta)\) by \(f_i(\beta)\). The Taylor’s expansion up the order 2 of \(g_i(\beta)\) at \(\beta = \beta^0\) is
\[
gi(\beta) = f_i(\beta^0)\varepsilon_i + \frac{1}{2} M_{1i}(\beta - \beta^0)\varepsilon_i - \frac{1}{2} f_i(\beta^0) f_i'(\beta^0) \varepsilon_i - \frac{1}{6} f_i(\beta^0) (\beta - \beta^0)^{\frac{5}{2}} M_{2i}(\beta - \beta^0) - \frac{1}{4} f_i(\beta^0) f_i'(\beta^0) \varepsilon_i - \frac{1}{12} f_i(\beta^0) (\beta - \beta^0)^{\frac{5}{2}} M_{2i}(\beta - \beta^0),
\]

where
\[
M_{1i} = \left(\frac{\partial^2 f_i(\beta^0)}{\partial \beta_j \partial \beta_k}\right)_{1 \leq j, k \leq d}, \quad M_{2i} = \left(\frac{\partial^2 f_i(\beta^0)}{\partial \beta_j \partial \beta_k \partial \beta_k}\right)_{1 \leq j, k \leq d}
\]
and
\[
\beta_{i,j}^{(1)} = \beta^0 + u_{i,j}k(\beta - \beta^0), \quad \beta_{i,j}^{(2)} = \beta^0 + v_{i,j}k(\beta - \beta^0), \text{ with } u_{i,j}k, v_{i,j}k \in [0, 1].
\]

We note that \(\beta_{i,j}^{(1)}\) and \(\beta_{i,j}^{(2)}\) are random vectors which depend on \(X_i\).

For \(f_i(\beta^0)\varepsilon_i\), because \(X_i\) and \(\varepsilon_i\) are independent, and \(E(\varepsilon_i) = 0\), we have that \(E[f_i(\beta^0)\varepsilon_i] = 0\) and \(Var[f_i(\beta^0)\varepsilon_i] = \sigma^2V\). For the \(j\)-th component of \(f_i(\beta^0)\), by the Bienaymé-Tchebychev’s inequality, for \(1 \leq j \leq d\), for all \(\epsilon_1 > 0\), we have
\[
P[|f_i(\beta^0)\varepsilon_i| \geq \epsilon_1] \leq \frac{\sigma^2}{\epsilon_1^2} V_{jj},
\]

For \(f_i(\beta^0)\varepsilon_i\), because \(X_i\) and \(\varepsilon_i\) are independent, and \(E(\varepsilon_i) = 0\), we have that \(E[f_i(\beta^0)\varepsilon_i] = 0\) and \(Var[f_i(\beta^0)\varepsilon_i] = \sigma^2V\). For the \(j\)-th component of \(f_i(\beta^0)\), by the Bienaymé-Tchebychev’s inequality, for \(1 \leq j \leq d\), for all \(\epsilon_1 > 0\), we have
\[
P[|f_i(\beta^0)\varepsilon_i| \geq \epsilon_1] \leq \frac{\sigma^2}{\epsilon_1^2} V_{jj},
\]
where $V_{jj}$ is the j-th term diagonal of the matrix $V$.

For all $\epsilon > 0$, taking $\epsilon_1 = \sigma \sqrt{6 V_{jj}}/\epsilon$ in (52), we obtain $P\big[|\frac{\partial f_i}{\partial \beta_j}(\beta^0)| \geq \sigma \sqrt{6 V_{jj}}/\epsilon\big] \leq \epsilon/6$. Applying Lemma (i), we obtain, for all $\epsilon > 0$

$$
P\big[\|f_i(\beta^0)\epsilon_i\|_1 \geq \sigma d \max_{1 \leq j \leq d} \sqrt{6 V_{jj}}\big] \leq \epsilon/6. \quad (53)$$

For the second term of the right-hand side of (51), using assumption (A3), we obtain that for $1 \leq j, k \leq d$, for all $\epsilon > 0$ there exists $\epsilon_2 > 0$, such that, $P\big[|\frac{\partial^2 f_i(k^1)}{\partial \beta_j \partial \beta_k}(\beta^0)| \geq \epsilon_2\big] \leq \epsilon/6$. By Lemma (iii), we have that for all $\epsilon > 0$,

$$
P\big[\|M_{1i}\|_1 \geq \epsilon_2\big] \leq \frac{\epsilon}{6}. \quad (54)$$

Using Bienaymé-Tchebychev’s inequality, and assumption (A1), we obtain that for all $C_1 > 0$

$$
P[|\epsilon_i| > C_1] \leq \sigma^2 / C_1. \quad (55)$$

Recall that $\|\beta - \beta^0\|_2 < \eta$, with $\eta \to 0$. Then, using (54) and (55), we can write that, for all $\epsilon > 0$, there exists $\epsilon_2 > 0$ such that, $P\big[\|M_{1i}(\beta - \beta^0)\epsilon_i\|_1 \geq \epsilon_2\big] \leq P\big[\|M_{1i}\|_1 \|\beta - \beta^0\|_1 \geq \epsilon_2\big] \leq P\big[\|M_{1i}\|_1 \geq \epsilon_2/C_1\eta\big] \leq P\big[\|M_{1i}\|_1 \geq \epsilon_2\big] \leq \epsilon/6$. Therefore, for all $\epsilon > 0$, there exists $\epsilon_2 > 0$ such that

$$
P\big[\|M_{1i}(\beta - \beta^0)\epsilon_i\|_1 \geq \epsilon_2\big] \leq \frac{\epsilon}{6}. \quad (56)$$

We consider now the term $\hat{f}_i(\beta^0)^T \hat{f}_i(\beta^0)(\beta - \beta^0)$ of relation (51). By Markov’s inequality, taking also into account assumption (A4), we obtain that $1 \leq j, l \leq d$, for all $\epsilon_3 > 0$, that $P\big[|\frac{\partial^2 f_i(k^1)}{\partial \beta_j \partial \beta_l}(\beta^0)\| \geq \epsilon_3\big] \leq \mathbb{E}\big[|\frac{\partial^2 f_i(k^1)}{\partial \beta_j \partial \beta_l}(\beta^0)\|/\epsilon_3\big]$. We choose, for all $\epsilon > 0, \epsilon_3 = 6 \mathbb{E}\big[|\frac{\partial^2 f_i(k^1)}{\partial \beta_j \partial \beta_l}(\beta^0)\|\big]/\epsilon$. Then, the last relation becomes $P\big[|\frac{\partial^2 f_i(k^1)}{\partial \beta_j \partial \beta_l}(\beta^0)\| \geq 6 \mathbb{E}\big[|\frac{\partial^2 f_i(k^1)}{\partial \beta_j \partial \beta_l}(\beta^0)\|\big]/\epsilon\big] \leq \epsilon/6$. Using Lemma (iii), we obtain

$$
P\big[\|f_i(\beta^0)^T f_i(\beta^0)\|_1 \geq \frac{6d}{\epsilon} \max_{1 \leq j, l \leq d} \mathbb{E}\big[|\frac{\partial^2 f_i(k^1)}{\partial \beta_j \partial \beta_l}(\beta^0)\|\big] \big] \leq \frac{\epsilon}{6}. \quad (57)$$

relation that involves, since for all $C_2 > 0$ we have $\|\beta - \beta^0\|_1 \leq C_2 \eta$ for $\eta \to 0$, that

$$
P\big[\|f_i(\beta^0)^T f_i(\beta^0)(\beta - \beta^0)\|_1 \geq 6d/\epsilon \max_{1 \leq j, l \leq d} \mathbb{E}\big[|\frac{\partial^2 f_i(k^1)}{\partial \beta_j \partial \beta_l}(\beta^0)\|\big]\big] \leq \epsilon/6. \quad (57)$$

Then, for all $\epsilon > 0$

$$
P\big[\|f_i(\beta^0)^T f_i(\beta^0)(\beta - \beta^0)\|_1 \geq \frac{6d}{\epsilon} \max_{1 \leq j, l \leq d} \mathbb{E}\big[|\frac{\partial^2 f_i(k^1)}{\partial \beta_j \partial \beta_l}(\beta^0)\|\big] \big] \leq \frac{\epsilon}{6} \quad (57)$$
For $M_{1i}(\beta - \beta^0)f_i^t(\beta^0)(\beta - \beta^0)$ of relation (61), using assumption (A3) and the Markov’s inequality, we obtain for each $j$-th component $\frac{\partial f_i(\beta^0)}{\partial \beta_j}$ of the vector $f_i(\beta^0)$, for all $\epsilon_4 > 0$, that $P\left[\left|\frac{\partial f_i(\beta^0)}{\partial \beta_j}\right| \geq \epsilon_4\right] \leq \frac{E\left[\frac{\partial f_i(\beta^0)}{\partial \beta_j}\right]}{\epsilon_4}$.

We choose, for all $\epsilon > 0$, $\epsilon_4 = 6E\left[\frac{\partial f_i(\beta^0)}{\partial \beta_j}\right]/\epsilon$ and this last relation becomes $P\left[\left|\frac{\partial f_i(\beta^0)}{\partial \beta_j}\right| \geq \frac{6E\left[\frac{\partial f_i(\beta^0)}{\partial \beta_j}\right]}{\epsilon}\right] \leq \epsilon/6$. Applying Lemma 1 (i), for all $\epsilon > 0$ we obtain

$$P\left[\|f_i(\beta^0)\| \geq \frac{6d}{\epsilon} \max_{1 \leq j \leq d} E\left[\frac{\partial f_i(\beta^0)}{\partial \beta_j}\right]\right] \leq \frac{\epsilon}{6}. \quad (58)$$

Using assumption (A3), and relations (54), (58), we can write that

$$P\left[\|M_{1i}(\beta - \beta^0)f_i^t(\beta^0)(\beta - \beta^0)\| \geq \frac{6d}{\epsilon} \max_{1 \leq j \leq d} E\left[\frac{\partial f_i(\beta^0)}{\partial \beta_j}\right]\right] \leq P\left[\|f_i(\beta^0)\| \geq \frac{6d}{\epsilon} \max_{1 \leq j \leq d} E\left[\frac{\partial f_i(\beta^0)}{\partial \beta_j}\right]\right] \leq \epsilon/6.$$

Therefore, for all $\epsilon > 0$,

$$P\left[\|M_{1i}(\beta - \beta^0)f_i^t(\beta^0)(\beta - \beta^0)\| \geq \frac{6d}{\epsilon} \max_{1 \leq j \leq d} E\left[\frac{\partial f_i(\beta^0)}{\partial \beta_j}\right]\right] \leq \frac{\epsilon}{6}. \quad (59)$$

Taking into account assumptions (A3), (A4), by relations (55), (58), we can prove in a similar way as for relation (59) that, for all $\epsilon > 0$,

$$P\left[\|f_i(\beta^0)(\beta - \beta^0)M_{2i}(\beta - \beta^0)\| \geq \frac{6d}{\epsilon} \max_{1 \leq j \leq d} E\left[\frac{\partial f_i(\beta^0)}{\partial \beta_j}\right]\right] \leq \frac{\epsilon}{6}. \quad (60)$$

For the last term on the right-hand side of (61), using assumption (A3), we have that, for all $\beta \in V_0(\beta^0)$, for all $\epsilon > 0$, there exists $\epsilon_5 > 0$, such that $P\left[\|M_{1i}_1\| \|M_{2i}_1\| \geq \epsilon_5\right] \leq \epsilon/6$. Using this relation, we show similarly, then, for all $\epsilon > 0$, there exists $\epsilon_5 > 0$, such that,

$$P\left[\|M_{1i}(\beta - \beta^0)(\beta - \beta^0)M_{2i}(\beta - \beta^0)\| \geq \epsilon_5\right] \leq \frac{\epsilon}{6}. \quad (61)$$

Choosing

$$M = \sup \left\{ \frac{\sigma d}{\epsilon} \max_{1 \leq j \leq d} \sqrt{6V_{jj}}, \epsilon_2, \frac{6d}{\epsilon} \max_{1 \leq j \leq d} \left\{ E\left[\frac{\partial f_i(\beta^0)}{\partial \beta_j}\right], E\left[\frac{\partial f_i(\beta^0)}{\partial \beta_l}\right] \right\}, \epsilon_5 \right\},$$

and combining (53), (55), (57), (59), (60), (61) together, lemma yields. \(\blacksquare\)

**Lemma 3** Under the same assumptions of Theorem 3, we have

$$\frac{1}{n\theta nk} \sum_{i \in I} g_i(\beta) = O_P\left((n\theta nk)^{-1/2}\right) + V_{1i}^0(\beta - \beta^0) + \alpha_P(\beta - \beta^0).$$
Proof of Lemma 3. By the Taylor’s expansion up to the order 3 of \( g_i(\beta) \) at \( \beta = \beta^0 \), we obtain

\[
\frac{1}{n \theta_{nk}} \sum_{i \in I} g_i(\beta) = \frac{1}{n \theta_{nk}} \sum_{i \in I} \hat{f}_i(\beta^0) \epsilon_i + \frac{1}{2 n \theta_{nk}} \sum_{i \in I} \hat{f}_i(\beta^0)(\beta - \beta^0) \epsilon_i
\]

\[
- \frac{1}{2 n \theta_{nk}} \sum_{i \in I} \hat{f}_i(\beta^0) \hat{f}_i^t(\beta^0)(\beta - \beta^0)
\]

\[
- \frac{1}{6 n \theta_{nk}} \sum_{i \in I} \hat{f}_i(\beta^0)(\beta - \beta^0)^{i} M_{2i}(\beta - \beta^0)
\]

\[
- \frac{1}{4 n \theta_{nk}} \sum_{i \in I} \hat{f}_i(\beta^0)(\beta - \beta^0)^{i} \hat{f}_i^t(\beta^0)(\beta - \beta^0)
\]

\[
- \frac{1}{12 n \theta_{nk}} \sum_{i \in I} \hat{f}_i(\beta^0)(\beta - \beta^0)^{i} M_{2i}(\beta - \beta^0)
\]

\[
+ \frac{1}{6 n \theta_{nk}} \sum_{i \in I} M_{i}(\beta - \beta^0)^{i} M_{2i}(\beta - \beta^0),
\]

with \( M_{2i} \) given by Lemma 2 and \( M_i = \left( \sum_{i=1}^{d} \sum_{k=1}^{d} \frac{\partial^2 \hat{f}_i(\beta^0)}{\partial \beta_k \partial \beta_l} (\beta_k - \beta^0_l)(\beta_l - \beta^0_l) \right)_{1 \leq k, l \leq d} \) is a vector of dimension \( (d \times 1) \), where \( \beta_{i,k,l} = \beta^0 + w_{i,k,l}(\beta - \beta^0) \), with \( w_{i,k,l} \in [0, 1] \).

For the first term of the right-hand side of (62), by the central limit theorem, and the fact that \( \mathbb{E}[g_i(\beta^0)] = 0 \), we have

\[
(n \theta_{nk})^{-1} \sum_{i \in I} g_i(\beta^0) = O_P((n \theta_{nk})^{-1/2}). \quad (63)
\]

For the second term of the right-hand side of (62), by the law of large numbers, the term \( (n \theta_{nk})^{-1} \sum_{i \in I} \hat{f}_i(\beta^0)(\beta - \beta^0) \epsilon_i \) converges almost surely to the expected of \( \hat{f}_i(\beta^0)(\beta - \beta^0) \epsilon_i \) as \( n \to \infty \). Furthermore, since \( \epsilon_i \) is independent of \( X_i \) and \( \mathbb{E}[\epsilon_i] = 0 \), we have

\[
\frac{1}{n \theta_{nk}} \sum_{i \in I} \hat{f}_i(\beta^0)(\beta - \beta^0) \epsilon_i = o_P(\beta - \beta^0). \quad (64)
\]

For the third term of the right-hand side of (62), by the law of large numbers and assumption (A4), the term \( (n \theta_{nk})^{-1} \sum_{i \in I} \hat{f}_i(\beta^0)\hat{f}_i^t(\beta^0)(\beta - \beta^0) \) converges almost surely to the expected value of \( \hat{f}_i(\beta^0)\hat{f}_i^t(\beta^0)(\beta - \beta^0) \) as \( n \to \infty \). On the other hand, since \( (n \theta_{nk})^{-1} \sum_{i \in I} \hat{f}_i(\beta^0)\epsilon_i \overset{a.s.}{\to} 0 \), we have

\[
\frac{1}{n \theta_{nk}} \sum_{i \in I} \hat{f}_i(\beta^0)\hat{f}_i^t(\beta^0)(\beta - \beta^0) = -V^0_{1n}(\beta - \beta^0)(1 + o_P(1)). \quad (65)
\]
For the fourth term of the right-hand side of (62), by the law of large numbers, using assumption (A3) and the relation (69), we can write $(6n\theta_{nk})^{-1} \| \sum_{i \in I} f_i(\beta^0) (\beta - \beta^0)^{t} M_{2i}(\beta - \beta^0) \|_1 = O_p(\|\beta - \beta^0\|_2^2)$, which implies

$$\frac{1}{6n\theta_{nk}} \sum_{i \in I} f_i(\beta^0)(\beta - \beta^0)^{t} M_{2i}(\beta - \beta^0) = o_p(\beta - \beta^0).$$

(66)

In the same way, using assumption (A3) and relation (58), we obtain, for the fifth term on the right-hand side of (62), that

$$\frac{1}{4n\theta_{nk}} \sum_{i \in I} f_i(\beta^0)(\beta - \beta^0)^{t} f_i(\beta^0)(\beta - \beta^0) = o_p(\beta - \beta^0).$$

(67)

For the sixth term of the right-hand side of (62), using the assumption (A3), we have

$$\frac{1}{12n\theta_{nk}} \sum_{i \in I} f_i(\beta^0)(\beta - \beta^0)^{t} M_{2i}(\beta - \beta^0) = o_p(\beta - \beta^0).$$

(68)

For $1 \leq j \leq d$, and for any fixed $i$, such that $1 \leq i \leq n \theta_{nk}$, denote by $M_{ij}$ the following random variable designates the $j$-th component of the vector $M_{i}$, such that

$$M_{ij} = \sum_{l=1}^{d} \sum_{k=1}^{d} \frac{\partial^2 f_i(\beta_{i,k}^{(3)})}{\partial \beta_k \partial \beta_l} (\beta_k - \beta_k^0)(\beta_l - \beta_l^0).$$

using assumption (A3), we have with a probability one, $|M_{ij}| \leq C_3\|\beta - \beta^0\|_2^2$. Applying Lemma 1 (i), we obtain

$$\|M_{i}\|_1 \leq C_4\|\beta - \beta^0\|_2^2.$$  

(69)

For the term $(6n\theta_{nk})^{-1} \sum_{i \in I} M_{i} \varepsilon_i$, using relations (65) and (69), we have

$$(6n\theta_{nk})^{-1} \| \sum_{i \in I} M_{i} \varepsilon_i \|_1 \leq (6n\theta_{nk})^{-1} \sum_{i \in I} \|M_{i}\|_1 \|\varepsilon_i\| \leq C_4(6n\theta_{nk})^{-1} n \theta_{nk} \|\beta - \beta^0\|_2^2.$$  

Then,

$$\frac{1}{6n\theta_{nk}} \sum_{i \in I} M_{i} \varepsilon_i = o_p(\beta - \beta^0).$$  

(70)

Finally, for the last term of the right-hand side of (62), using assumption (A3) and relation (69), we obtain with probability 1, $(12n\theta_{nk})^{-1} \| \sum_{i \in I} M_{i}(\beta - \beta^0)^{t} M_{2i}(\beta - \beta^0) \|_1 \leq C_5\|\beta - \beta^0\|_2^2$, which gives,

$$\frac{1}{12n\theta_{nk}} \sum_{i \in I} M_{i}(\beta - \beta^0)^{t} M_{2i}(\beta - \beta^0) = o_p(\beta - \beta^0).$$

(71)

Then, combining relations (63), (64), (65), (66), (67), (68), (70) and (71), we obtain lemma.
Under the same assumptions as in Theorem 3, for all \( q > 0 \), there exist two positive constants \( B = B(q) \), \( T = T(q) \) such that

\[
\begin{align*}
\mathbb{P} & \left[ \max_{\theta < \theta_n \leq 1 - \frac{\theta_n}{n}} \left( \bar{\theta}_n / \log \log n \theta_n \right)^{1/2} \left\| \frac{\hat{\lambda}(\theta_n)}{\min(\theta_n, 1 - \theta_n)} \right\| \geq B \right] \leq q, \\
\mathbb{P} & \left[ \max_{\theta < \theta_n \leq 1 - \frac{\theta_n}{n}} \left( \bar{\theta}_n / \log \log n \theta_n \right)^{1/2} \left\| \tilde{\beta}(\theta_n) - \tilde{\beta}_0 \right\| \geq B \right] \leq q, \\
\mathbb{P} & \left[ n^{-1/2} \max_{\theta < \theta_n \leq 1 - \frac{\theta_n}{n}} \| \bar{\theta}_n \| \left\| \frac{\hat{\lambda}(\theta_n)}{\min(\theta_n, 1 - \theta_n)} \right\| \geq B \right] \leq q, \\
\mathbb{P} & \left[ n^{-1/2} \max_{\theta < \theta_n \leq 1 - \frac{\theta_n}{n}} \| \bar{\theta}_n \| \left\| \tilde{\beta}(\theta_n) - \tilde{\beta}_0 \right\| \geq B \right] \leq q.
\end{align*}
\]

**Proof of Lemma 4** The proof of this lemma is similar to that of Lemma 1.2.2 of [9].

In order, to prove Lemma 5 we consider

\[
R_k = n\alpha^{-2} \theta_n (1 - \theta_n)(W_{1n}^0 - W_{2n}^0)^t V^{-1}(W_{1n}^0 - W_{2n}^0).
\]

Recall that \( V \equiv \mathbb{E}[f(X_i, \beta_0)^t f(X_i, \beta_0)] \), for all \( i = 1, ..., n \).

The results of Lemma 5 are similar to that of Theorem 1.1.1 of [9].

**Lemma 5** Suppose that the assumptions (A1)-(A4) hold. Under the null hypothesis \( H_0 \), for all \( 0 \leq \alpha < 1/2 \) we have

(i) \( n\alpha \max_{\theta_n \in \hat{\theta}_n \subset [0, 1]} [\theta_n (1 - \theta_n)]^{\alpha/2} |Z_n(\theta_n, \hat{\lambda}(\theta_n), \tilde{\beta}(\theta_n)) - R_k| = O_F(1) \).

(ii) \( n\alpha \max_{\theta_n \in \hat{\theta}_n \subset [0, 1]} |Z_n(\theta_n, \hat{\lambda}(\theta_n), \tilde{\beta}(\theta_n)) - R_k| = O_F(n^{-1/2}(\log \log n)^{3/2}) \).

**Proof of Lemma 5** For the score function \( \phi_{1n} \) of relation (13), the two terms of the right-hand side are replaced by their decomposition obtained by the relations (22) and (25). On the other hand, we have \( \phi_{1n}(\theta_n, \hat{\lambda}(\theta_n), \tilde{\beta}(\theta_n)) = 0_d \).

Then, we can write

\[
\hat{\lambda}(\theta_n)(1 + o_F(1)) - V_{1n}^0(\beta(\tilde{\beta}(\theta_n) - \beta_0) - \frac{1}{n\sigma_n} \sum_{i \in I} g_i(\beta_0)g_i^t(\beta_0) \hat{\lambda}(\theta_n) + \frac{1}{n\sigma_n} \sum_{i \in I} g_i(\beta_0) + V_{2n}^0(\tilde{\beta}(\theta_n) - \beta_0) - \frac{1}{n\sigma_n} \sum_{i \in I} g_i(\beta_0) + V_{2n}^0(\tilde{\beta}(\theta_n) - \beta_0) - \frac{1}{n\sigma_n} \sum_{i \in I} g_i(\beta_0) + V_{2n}^0(\tilde{\beta}(\theta_n) - \beta_0))](1 + o_F(1)) = 0_d.
\]

Hence,

\[
\hat{\lambda}(\theta_n) = \left( \frac{1}{\sigma_n} D_{1n}^0 + \frac{n\alpha}{1 - \theta_n} (V_{1n}^0(V_{2n}^0)^{-1})^{-1} (V_{1n}^0(V_{2n}^0)^{-1})^{-1} \right) \left( \frac{1}{\sigma_n} D_{2n}^0 - \frac{n\alpha}{1 - \theta_n} \sum_{i \in I} g_i(\beta_0) + V_{1n}^0(V_{2n}^0)^{-1} \sum_{j \in I} g_j(\beta_0) \right) (1 + o_F(1))
\]

with the matrices \( D_{1n}^0 \) and \( D_{2n}^0 \) given by relation (24). On the other hand, by the law of large numbers, we have \( -V_{1n}^0 \xrightarrow{a.s} V \) and \( -V_{2n}^0 \xrightarrow{a.s} V \). Then, \( V_{1n}^0(V_{2n}^0)^{-1} \xrightarrow{a.s} I_d \). Always, by the law of large numbers, \( D_{1n}^0 \) and \( D_{2n}^0 \) converge almost surely to \( \sigma_2 V \) as \( n \to \infty \).

By Theorem 2, we proved that \( \lambda(\theta_n) = \theta_n O_P((n\theta_n)^{-1/2}) \). Then, we obtain

\[
\hat{\lambda}(\theta_n) = \sigma^{-2} \theta_n (1 - \theta_n) V^{-1}(W_{1n}^0 - W_{2n}^0)(1 + o_F(1)) + o_F(\tilde{\beta}(\theta_n) - \beta_0).
\]
The limited development of the statistic $Z_{nk}(\theta_{nk}, \hat{\lambda}(\theta_{nk}), \hat{\beta}(\theta_{nk}))$, specified by the relation \([12]\), in the neighbourhood of $(\lambda, \beta) = (0, \beta^0)$ up to order 2, can be written

\[
\left[\frac{2\hat{\lambda}^t(\theta_{nk})}{\theta_{nk}} \sum_{i \in I} g_i(\beta^0) - \frac{2\hat{\lambda}^t(\theta_{nk})}{1 - \theta_{nk}} V^0_{1n}(V^0_{2n})^{-1} \sum_{j \in J} g_j(\beta^0)\right] - \left[\frac{\hat{\lambda}^t(\theta_{nk})}{(1 - \theta_{nk})^2} V^0_{1n}(V^0_{2n})^{-1}\right]
\]

\[
\cdot \sum_{j \in J} g_j(\beta^0)g_j(\beta^0)V^0_{1n}(V^0_{2n})^{-1} \hat{\lambda}(\theta_{nk}) + \frac{\hat{\lambda}^t(\theta_{nk})}{\theta_{nk}} \sum_{i \in I} g_i(\beta^0)g_i(\beta^0) \hat{\lambda}(\theta_{nk})
\]

\[+ \left[2\hat{\lambda}^t(\theta_{nk})\left(\frac{1}{\theta_{nk}} \sum_{i \in I} g_i(\beta^0) - \frac{1}{1 - \theta_{nk}} V^0_{1n}(V^0_{2n})^{-1} \sum_{j \in J} g_j(\beta^0)\right)(\hat{\beta}(\theta_{nk}) - \beta^0)\right]
\]

\[+ \frac{1}{31}\left[S_1 + 3S_2 + 3S_3 + S_4\right],
\]

where

\[S_1 = \sum_{j=1}^{d} \sum_{l=1}^{d} \sum_{k=1}^{d} \frac{\partial^2 Z_{nk}(\theta_{nk}, \lambda^{(1)}, \beta^{(1)})}{\partial \lambda_{j} \partial \lambda_{k}} (\hat{\beta}_{j} - \beta^0)(\hat{\beta}_{k} - \beta^0),
\]

\[S_2 = \sum_{j=1}^{d} \sum_{l=1}^{d} \sum_{k=1}^{d} \frac{\partial^2 Z_{nk}(\theta_{nk}, \lambda^{(2)}, \beta^{(2)})}{\partial \lambda_{j} \partial \lambda_{k}} (\hat{\lambda}_{j})(\hat{\lambda}_{k} - \beta^0),
\]

\[S_3 = \sum_{j=1}^{d} \sum_{l=1}^{d} \sum_{k=1}^{d} \sum_{a=1}^{d} \frac{\partial^2 Z_{nk}(\theta_{nk}, \lambda^{(a)}, \beta^{(a)})}{\partial \lambda_{j} \partial \lambda_{k} \partial \beta_{a}} (\hat{\lambda}_{j})(\hat{\beta}_{k} - \beta^0)(\hat{\beta}_{a} - \beta^0),
\]

\[S_4 = \sum_{j=1}^{d} \sum_{l=1}^{d} \sum_{k=1}^{d} \sum_{a=1}^{d} \frac{\partial^2 Z_{nk}(\theta_{nk}, \lambda^{(a)}, \beta^{(a)})}{\partial \lambda_{j} \partial \lambda_{k} \partial \beta_{a}} (\hat{\beta}_{a} - \beta^0)(\hat{\lambda}_{j})(\hat{\lambda}_{k} - \beta^0),
\]

where, for $1 \leq j \leq d$, $\hat{\beta}_{j}$ is the j-th component of $\hat{\beta}(\theta_{nk})$, and $\hat{\lambda}_{j}$ is the j-th component of $\hat{\lambda}(\theta_{nk})$. In the expression of $S_1, S_2, S_3, S_4$ we have also, for all $1 \leq j, k, l \leq d$, $\lambda_{jkl} = u^{(a)}_{jkl}(\hat{\beta}(\theta_{nk}) - \beta^0)$, and $\beta_{jkl} = \beta^0 + v^{(a)}_{jkl}(\hat{\beta}(\theta_{nk}) - \beta^0)$, with $u^{(a)}_{jkl}, v^{(a)}_{jkl} \in [0, 1]$ and $a \in \{1, 2, 3, 4\}$.

We note that, the derivative $\partial (V^0_{1n}(\beta)(V^0_{2n}(\beta))^{-1})/\partial \beta$ is considered term by term.

Now, we replace $\hat{\lambda}(\theta_{nk})$ in the relation \([12]\) by the value obtained in \([21]\).

For the first term of \([21]\), using notations given by relation \([20]\), and the fact that $V^0_{1n}(V^0_{2n})^{-1} a \rightarrow I_d$, as $n \rightarrow \infty$, we find that this term is equal to $2n \sigma^{-2} \theta_{nk}(1 - \theta_{nk})(W^0_{1n} - W^0_{2n})V^{-1}(W^0_{1n} - W^0_{2n}) + o_P(\|\hat{\beta}(\theta_{nk}) - \beta^0\|_2)$.

Similarly, for the second term of \([21]\), using notations given by \([21]\), and the fact that $D^0_{1n}$ and $D^0_{2n}$ converge to $\sigma^2 V$, as $n \rightarrow \infty$, we obtain that this term is equal to $n \sigma^{-2} \theta_{nk}(1 - \theta_{nk})(W^0_{1n} - W^0_{2n})V^{-1}(W^0_{1n} - W^0_{2n}) + o_P(\|\hat{\beta}(\theta_{nk}) - \beta^0\|_2)$.

For the third term of \([21]\), we know that, $V^0_{1n} = (n \theta_{nk})^{-1} \sum_{i \in I} g_i(\beta^0)$, and $V^0_{2n} = (n(1 - \theta_{nk}))^{-1} \sum_{j \in J} g_j(\beta^0)$. On the other hand, by the law of large numbers, we have $V^0_{1n}$ and $V^0_{2n}$ converge almost surely to $-V$ as $n \rightarrow \infty$, and $V^0_{1n}(V^0_{2n})^{-1} a \rightarrow I_d$, which implies that the third term of \([21]\) converge almost surely to zero, as $n \rightarrow \infty$.  

\[\text{(73)}\]
By the central limit theorem, we have that 
\[(n(1 – \theta_{nk}))^{-1} \sum_{j \in J} g_j(\beta^0) = O_{\mathbb{P}}((n(1 – \theta_{nk}))^{-1/2}).\] Then, the fourth term of (73) is
\[o_{\mathbb{P}}(n\sigma^{-2}\theta_{nk}(1 – \theta_{nk})(W_{1n}^0 – W_{2n}^0)V^{-1}(W_{1n}^0 – W_{2n}^0)).\]
For the last term of (73), using assumptions (A2)-(A4) and by an elementary calculations, we prove that this term is 
\[o_{\mathbb{P}}(||\hat{\beta}(\theta_{nk}) – \beta^0||_2) + o_{\mathbb{P}}(||\hat{\lambda}(\theta_{nk})||_2) + o_{\mathbb{P}}(||\hat{\lambda}(\theta_{nk})||_2||\hat{\beta}(\theta_{nk}) – \beta^0||_2).\]
Combining the obtained results, we obtain
\[Z(\theta_{nk}, \hat{\lambda}(\theta_{nk}), \hat{\beta}(\theta_{nk})) = n\sigma^{-2}\theta_{nk}(1 – \theta_{nk})(W_{1n}^0 – W_{2n}^0)V^{-1}(W_{1n}^0 – W_{2n}^0)(1 + o_{\mathbb{P}}(1)) + o_{\mathbb{P}}(||\hat{\beta}(\theta_{nk}) – \beta^0||_2) + o_{\mathbb{P}}(||\hat{\lambda}(\theta_{nk})||_2) + o_{\mathbb{P}}(||\hat{\lambda}(\theta_{nk})||_2||\hat{\beta}(\theta_{nk}) – \beta^0||_2).\]
This last relation, together with Lemma 4 imply Lemma 5.

References

1. Antoch, J., Gregoire, G., Jaruskova, D. Detection of structural changes in generalized linear models. *Statistics and Probability Letters*, 69, 315-322, 2004.
2. Bai, J. Likelihood ratio tests for multiple structural changes. *Journal of Econometrics*, 91, 299-323, 1999.
3. Bai, J., Perron, P. Estimating and testing linear models with multiple structural changes. *Econometrica*, 1, 47-48, 1998.
4. Boldea, O., Hall, A.R. Estimation and inference in unstable nonlinear least squares models. *Journal of Econometrics*, 172(1), 158-167, 2013.
5. Ciuperca, G. A general criterion to determine the number of change-points. *Statistics and Probability Letters*, 81(8), 1267-1275, 2011.
6. Ciuperca, G. Penalized least absolute deviations estimation for nonlinear model with change-points. *Statistical Papers*, 52(2), 371-390, 2011.
7. Ciuperca G. Two tests for sequential detection of a change-point in a nonlinear model. *Journal of Statistical Planning and Inference*, 143(10), 1621-1834, 2013.
8. Ciuperca G. Empirical likelihood for nonlinear model with missing responses, *Journal of Statistical Computation and Simulation*, 83(4), 737-756, 2013.
9. Cugnot M., Horváth L. Limit Theorems in Change-point Analysis, John Wiley and Sons (1997).
10. Guan, Z. A semiparametric changepoint model. *Biometrika*, 91(3), 849-862, 2004.
11. Hall A., Sen A. Structural stability testing in models estimated by generalized method of moments, *Journal of Business Economic Statistics*, 17(3), 335-348, 1999.
12. Horváth, L., Hušková, M., Kokoszka, P., Steinebach, J. Monitoring changes in linear models. *Journal of Statistical Planning and Inference*, 126(1), 225-251, 2004.
13. Hušková, M., Kirch, C. Bootstrapping sequential change-point tests for linear regression. *Metrika*, 75, no 5, 673-708, 2012.
14. Lai, T. L., Xing, H. Sequential change-point detection when the pre- and post-change parameters are unknown. *Sequential Analysis*, 29(2), 162–175, 2010.
15. Lee, S., Seo, M.H., Shin, Y. Testing for threshold effects in regression models. *Journal of American Statistical Association*, 106, 220–231, 2011.
16. Liu, Y., Zou, C., Zhang, R. Empirical likelihood ratio test for a change-point in linear regression model. *Communications in Statistics-Theory and Methods*, 37, 2551-2563, 2008.
17. Mei, Y. Sequential change-point detection when unknown parameters are present in the pre-change distribution. *Annals of Statistics* 34(1), 92-122, 2006.
18. Neumeyer, N., Van Keilegom, I. Change-point tests for the error distribution in non-parametric regression. *Scandinavian Journal of Statistics*, 36(3), 518-541, 2009.
19. Ning, W., Paulden, J., Gupta, A. Empirical likelihood ratio test for the epidemic change model. *Journal of Data science*, 10, 107-127, 2012.
20. Nosek, K. Schwarz information criterion based tests for a change-point in regression models. *Statistical Papers*, **51**(4), 915-929, 2010.
21. Owen, A. B. Empirical likelihood. *Chapman & Hall*, 2001.
22. Qin, J., Lawless, J. Empirical likelihood and general estimating equations. *Annals of Statistics*, **22**(4), 900-925, 2004.
23. Qu, Z. Testing for structural change in regression quantiles *Journal of Econometrics*, **146**, 170-184, 2008.
24. Ramanayake, A., Gupta A.K. Tests for an epidemic change in a sequence of exponentially distributed random variables *Biometrical Journal*, **45**, 496-958, 2003.
25. Ramanayake, A. Tests for a change point in the shape parameter of gamma random variables. *Comm. Statist. Theory Methods*, **33**, 821-833, 2004.
26. Seber, G., Wild, C. Nonlinear regression. *Wiley series in probability and mathematical Statistics, Wiley, Hoboken, NJ*, 2003.
27. Van der Vaart A., Wellner J. Weak Convergence and Empirical processes, 1996.
28. Wu, Y. Simultaneous change point analysis and variable selection in a regression problem. *Journal of Multivariate Analysis*, **99**, 2154-2171, 2008.
29. Yao, Q.W. Tests for change-points with epidemic alternatives. *Biometrika*, **80**, 179-191, 1993.
30. Yu, W., Niu, C., Xu, W. An empirical likelihood inference for the coefficient difference of a two-sample linear model with missing response data. *Metrika*, DOI 10.1007/s00184-013-0459-3.
31. Zi, X., Zou, C., Liu, Y. Two-sample empirical likelihood method for difference between coefficients in linear regression model. *Statistical Papers*, **53**(1), 83-93, 2012.
32. Zou C., Liu Y., Qin P., Wang Z. Empirical likelihood ratio test for a change point. *Statistics and Probability Letters*, **77**, 374-382, 2007.
