$L_1$-Statistical Procedures and Related Topics
IMS Lecture Notes – Monograph Series (1997) Volume 31

# $L_1$-test procedures for detection of change

**M. Hušková**

*Charles University, Prague, Czech Republic*

*Abstract*: $L_1$-type test procedures for detection of a change in linear models are proposed, their properties are studied under the null hypothesis (no change).

*Key words*: $L_1$-test procedures, linear models, change point.

AMS subject classification: 62G20, 62E20.

## 1 Introduction

The problem to detect and to identify changes in statistical models has attracted a number of researchers in the last two decades. Using various principles they have proposed a number of statistical procedures that are sensitive w.r.t. detection of changes, have studied their (mostly) limit properties and, also, have applied to real data sets.

The problem of detection and identification of changes in statistical models is known as the *change point problem* (mostly for case of changes in location models), *disorder problem* or *testing the constancy of regression relationship over time*. These problems arise in a number of applications (economic modelling, quality control, biology, medicine, meteorology and ecology among others).

We shall consider here the following regression model with possible change after an unknown time point $m$:

$$Y_i = \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{x}_i \boldsymbol{\delta}_n I\{i > m\} + E_i, \qquad i = 1, ..., n, \tag{1}$$

where $\mathbf{x}_i = (x_{i1}, .., x_{ip})^T$, $x_{i1} = 1$, $i = 1, ..., n$, are known regression vectors, $m(< n), \boldsymbol{\beta} = (\beta_1, ...\beta_p)^T, \boldsymbol{\delta}_n = (\delta_{n1}, ..., \delta_{np})^T$ are unknown parameters, $E_1, ..., E_n$ are i.i.d. random variables with common distribution function $F$. $I\{A\}$ denotes the indicator of the set $A$.

The model corresponds to the situation when up to an unknown $m$ the observations follow the regression model with the regression parameter $\beta$ and then the model changes to the regression model with the regression parameter $\beta + \delta_n$. The parameter $m$ is called the *change point.*

The problem of our interest is to test

$$H_o : m = n \text{ against } H_1 : m < n.$$

The authors usually apply either the likelihood ratio principle or the Bayesian approach. The first principle leads to max-type procedures the other gives sum-type procedures.

First, we shall describe likelihood ratio and related procedures when the distribution of the error terms $F$ is $N(0, \sigma^2)$, $\sigma^2 > 0$ known. It will give motivation how to develop $L_1$- procedures.

Assume that $e_i'$s are i.i.d. with distribution $N(0, \sigma^2)$, $\sigma^2 > 0$ known, the likelihood ratio principle leads to the test statistics

$$T_{n,LSE} = \max_{p<k<n-p} \{-\sum_{i=1}^{k} \rho_{LSE}(Y_i - x_i^T \beta_{k,LSE}) \tag{2}$$

$$- \sum_{i=k+1}^{n} \rho_{LSE}(Y_i - x_i^T \beta_{k,LSE}^*) + \sum_{i=1}^{n} \rho_{LSE}(Y_i - x_i^T \beta_{n,LSE})\}/\sigma^2,$$

where $\rho_{LSE}(x) = x^2$, $x \in R^1$, $\beta_{k,LSE}$ and $\beta_{k,LSE}^*$ are the least squares estimators of the regression parameters based on $X_1, ..., X_k$ and $X_{k+1}, ...., X_n$, respectively, i.e.,

$$\beta_{k,LSE} = \mathbf{C}_k^{-1} \sum_{i=1}^{k} \mathbf{x}_i Y_i, \qquad \beta_{k,LSE}^* = \mathbf{C}_k^{*-1} \sum_{i=k+1}^{n} \mathbf{x}_i Y_i$$

with

$$\mathbf{C}_k = \sum_{i=1}^{k} \mathbf{x}_i \mathbf{x}_i^T, \qquad \mathbf{C}_k^* = \sum_{i=k+1}^{n} \mathbf{x}_i \mathbf{x}_i^T. \tag{3}$$

The test statistics $T_{n,LSE}$ can be expressed equivalently as:

$$T_{n,LSE} = \max_{p<k<n-p} \{\left(\beta_{k,LSE} - \beta_{k,LSE}^*\right)^T (\mathbf{C}_k^{-1} + \mathbf{C}_k^{*-1})^{-1} \tag{4}$$

$$\left(\beta_{k,LSE} - \beta_{k,LSE}^*\right)/\sigma^2\}$$

and

$$T_{n,LSE} = \max_{p<k<n-p} \left\{\mathbf{S}_{k,LSE}^T (\mathbf{C}_k^{-1} \mathbf{C}_n \mathbf{C}_k^{*-1}) \mathbf{S}_{k,LSE}/\sigma^2\right\}, \tag{5}$$

where

$$\mathbf{S}_{k,LSE} = \sum_{i=1}^{k} \mathbf{x}_i (Y_i - \mathbf{x}_i^T \boldsymbol{\beta}_{n,LSE}). \tag{6}$$

Horváth (1995) among others derived the limit distribution of $T_{n,LSE}$ under $H_o$ and showed that if mild assumptions are satisfied then under $H_o$, as $n \to \infty$,

$$\max \Big\{ \mathbf{S}_{k,LSE}^T (\mathbf{C}_k^{-1} \mathbf{C}_n \mathbf{C}_k^{*-1}) \mathbf{S}_{k,LSE} / \sigma^2, \\ k = [n/\log n], ..., n - [n/\log n] \Big\} / T_{n,LSE} \to 0.$$

which means that asymptotically even under $H_0$ the terms with $k$ "small" or close to $n$ dominate the others. To avoid to this unpleasant property some modifications were proposed. Namely, the class of test statistics depending on a suitable weight function $q$ was introduced:

$$T_{n,LSE}(q) = \max_{1 \le k < n} \Big\{ \frac{\mathbf{S}_{k,LSE}^T \mathbf{C}_n^{-1} \mathbf{S}_{k,LSE}}{q^2(k/n)\sigma^2} \Big\}. \tag{7}$$

Typical choices of the weight function $q$ are the following

$$q(t) = (t(1-t))^{-1/2} \qquad t \in (a_1, a_2) \tag{8}$$

$$q(t) = 0 \qquad \text{otherwise,}$$

where $0 < a_1 < a_2 < 1$, or

$$q(t) = (t(1-t))^{-\gamma} \qquad t \in (0, 1/2), \tag{9}$$

with $\gamma \in [0, 1/2)$.

Some authors (e.g. Jandhyala and MacNeill, 1989; Ploberger and Kr¨amer, 1992) suggested to apply procedures based on properly standardized partial sums of the $LSE$-residuals:

$$S_{k,LSE}^o = \sum_{i=1}^{k} (Y_i - \mathbf{x}_i^T \boldsymbol{\beta}_{n,LSE}), \; k = 1, 2, ..., n \tag{10}$$

They proposed a computationally feasible procedure:

$$T_{n,LSE}^o(q) = \max_{1 \le k < n} \Big\{ \frac{|S_{k,LSE}^o|}{\sqrt{n}\sigma q(k/n)} \Big\}, \tag{11}$$

where $q$ is a weight function. Another type of procedures is based on moving sums (MOSUM) of the LSE-residuals. They are defined by:

$$T_{n,LSE}^*(G) = \max_{G < k < n} \Big\{ \frac{1}{\sqrt{G}} |S_{k,LSE}^o - S_{k-G,LSE}^o| / \sigma \Big\} \tag{12}$$

and

$$T^{**}_{n,LSE} = \max_{G<k<n-G} \Big\{ \frac{1}{\sqrt{G}} |S^o_{k+G,LSE} - 2S^o_{k,LSE} + S^o_{k-G,LSE}|/\sigma \Big\} \qquad (13)$$

Bayesian type of test statistics have the form

$$T^B_{n,LSE}(v) = \sum_{k=1}^{n-1} v(k/n) \Big\{ \mathbf{S}^T_{k,LSE} \mathbf{C}_n^{-1} \mathbf{S}_{k,LSE}/\sigma^2 \Big\}^{1/2}. \qquad (14)$$

where $v(1/n), ..., v((n-1)/n)$ represent priors.

Inspite that the procedures were developed for normally distributed random errors they can be applied also for nonnormally distributed random errors with zero mean and finite absolute moment of the order $2+\Delta(\Delta > 0)$ (in some cases a finite second moment suffices).

If $\sigma^2$ is unknown it is recommended to estimate it by

$$\widehat{\sigma}^2_{n,LSE} = \frac{1}{n-p} \min_{1 \le k \le m} \Big\{ \sum_{i=1}^{k} (Y_i - \mathbf{x}_i^T \beta_{k,LSE})^2 + \sum_{i=k+1}^{n} (Y_i - \mathbf{x}_i^T \beta^*_{k,LSE})^2 \Big\}.$$

and plug into the above statistics.

Typically large values of the introduced test statistics indicate that the null hypothesis $H_o$ fails. The exact distributions even under $H_o$ of the above introduced test statistics are unknown. The limit distributions were derived under mild assumptions on the distribution of the error terms $E_i's$ and on the design points $\mathbf{x}_1, ..., \mathbf{x}_n$, which enable to get the approximations for the critical values.

The test procedures corresponding to $T_{n,LSE}$ and $T_{n,LSE}(q)$ were studied by a number of authors, e.g. Quandt (1958, 1960), Worsley (1983), Kim and Siegmund (1989), Gombay and Horváth (1994), Horváth (1995), Antoch and Hušková (1992). Well known is the paper by Brown, Durbin and Evans (1975) devoted to the procedures based on recursive residuals. Bayesian type procedures were proposed and studied by Broemling and Tsurumi (1987) and Jandhyala and MacNeill (1989, 1991, 1992). Procedures based on partial sums of $LSE$- residuals were investigated, e.g., by Jandhyala and MacNeill (1989) and Ploberger and Krämer (1992). Hackl (1980) deeply studied procedures based on moving sums (MOSUM). A number of applications in econometrics is contained in Hackl (1989) and Hackl and Westlund (1991). Horváth, Hušková and Serbinowska (1995) considered the case when the change can occur in the regression parameters and/or in the scale $\sigma$.

Along the same line $M-$ type and $R-$ (rank based) tests were developed and studied. Some results on the $M-$ procedures for changes in regression models can be found, e.g., in Sen (1984) and Hušková (1990a,b, 1994 a,b). $R-$ type test procedures were studied by Sen (1980, 1982) and Hušková (1994b).

## 2 $L_1$ procedures

It is easily seen that the test statistics $T_{n,LSE}$'s are functions of least squares estimators and of the $LSE$-residuals $Y_i - \mathbf{x}_i^T \boldsymbol{\beta}_{LSE}$, where $\boldsymbol{\beta}_{LSE}$ is a least squares estimator, therefore $T_{n,LSE}$'s can be viewed as the $L_2$- type test statistic.

Now, along this line the $L_1$-procedures will be developed. Namely, we replace the $LSE$ estimators by $L_1$ estimators, $\rho_{LSE}$ by $\rho_{L_1}(x) = |x|$, $x \in R^1$, and the $LSE$-residuals by $L_1$-residuals $\psi_{L_1}(Y_i - \mathbf{x}_i^T \boldsymbol{\beta}_{L_1})$, where $\boldsymbol{\beta}_{L_1}$ is an $L_1$ estimator $\boldsymbol{\beta}$ and $\psi_{L_1}(x) = -1$, $x < 0$, $\psi_{L_1}(x) = 0$, $x = 0$, $\psi_{L_1}(x) = 1$, $x > 0$.

From three equivalent expressions for $T_{n,LSE}$ ((2), (4), (5)) we get three different test statistics. Namely,

$$T_{n,L_1}^{(1)} = \max_{p<k<n-p} \{2\widehat{f}(F^{-1}(1/2))\Big( - \sum_{i=1}^{k} |Y_i - x_i^T \boldsymbol{\beta}_{k,L_1}| \qquad (15)$$

$$- \sum_{i=k+1}^{n} |Y_i - x_i^T \boldsymbol{\beta}_{k,L_1}^*| + \sum_{i=1}^{n} |Y_i - x_i^T \boldsymbol{\beta}_{n,L_1}|\Big)\},$$

$$T_{n,L_1}^{(2)} = \max_{p<k<n-p} \Big\{4\widehat{f}^2(F^{-1}(1/2))\Big(\boldsymbol{\beta}_{k,L_1} - \boldsymbol{\beta}_{k,L_1}^*\Big)^T \qquad (16)$$

$$(\mathbf{C}_k^{-1} + \mathbf{C}_k^{*-1})^{-1}\Big(\boldsymbol{\beta}_{k,L_2} - \boldsymbol{\beta}_{k,L_1}^*\Big)\Big\}$$

and

$$T_{n,L_1}^{(3)} = \max_{p<k<n-p} \Big\{\mathbf{S}_{k,L_1}^T (\mathbf{C}_k^{-1} \mathbf{C}_n \mathbf{C}_k^{*-1}) \mathbf{S}_{k,L_1}\Big\}, \qquad (17)$$

where $\widehat{f}(F^{-1}(1/2))$ is an estimator of $f(F^{-1}(1/2))$, $F^{-1}$ and $f$ denote the quantile function and the density, respectively,

$$\mathbf{S}_{k,L_1} = \sum_{i=1}^{k} \mathbf{x}_i \psi_{L_1}(Y_i - \mathbf{x}_i^T \boldsymbol{\beta}_{n,L_1}), \qquad (18)$$

and $\boldsymbol{\beta}_{k,L_1}$ and $\boldsymbol{\beta}_{k,L_1}^*$ are the $L_1$- estimators of the regression parameters based on $X_1, ..., X_k$ and $X_{k+1}, ...., X_n$, respectively, i.e., they are defined as solutions of the minimization problems

$$\min \sum_{i=1}^{k} |Y_i - x_i^T \mathbf{t}|, \ \mathbf{t} \in R^p \qquad (19)$$

and

$$\min \sum_{i=k+1}^{n} |Y_i - x_i^T \nu|, \ \nu \in R^p, \tag{20}$$

respectively.

The statistic $T^{(1)}_{n,L_1}$ is a likelihood ratio type statistic, $T^{(2)}_{n,L_1}$ is a Wald type test statistic and $T^{(3)}_{n,L_1}$ is a score type test statistic.

Computational feasibility of the statistic $T^{(3)}_{n,L_1}$ is evident. The statistics $T^{(1)}_{n,L_1}, T^{(2)}_{n,L_1}$ depend on the estimator of $f(F^{-1}(1/2))$ and also on the estimators $\beta_{k,L_1}$ and $\beta^*_{k,L_1}$, $k = 1, ..., n$. Quality of the estimator of $f(F^{-1}(1/2))$ strongly influence the quality of the test itself.

The weighted type test statistics are defined by

$$T_{n,L_1}(q) = \max_{1 \le k < n} \Big\{ \frac{\mathbf{S}^T_{k,L_1} \mathbf{C}_n^{-1} \mathbf{S}_{k,L_1}}{q(k/n)} \Big\}, \tag{21}$$

where the weight function $q$ is the same as in $LSE$-case.

Next, we introduce the test statistics based on partial sums of $L_1$-residuals

$$S^o_{k,L_1} = \sum_{i=1}^{k} \psi_{L_1}(Y_i - \mathbf{x}_i^T \beta_{n,L_1}), \ k = 1, 2, ..., n \tag{22}$$

We get the weighted sum type and MOSUM type test statistics

$$T^o_{n,L_1}(q) = \max_{1 \le k < n} \Big\{ \frac{|S^o_{k,L_1}|}{\sqrt{n} q(k/n)} \Big\}, \tag{23}$$

$$T^*_{n,L_1}(G) = \max_{G < k < n} \Big\{ \frac{1}{\sqrt{G}} |S^o_{k,L_1} - S^o_{k-G,L_1}| \Big\}, \tag{24}$$

$$T^{**}_{n,L_1} = \max_{G < k < n-G} \Big\{ \frac{1}{\sqrt{2G}} |S^o_{k+G,L_1} - 2S^o_{k,L_1} + S^o_{k-G,L_1}| \Big\}, \tag{25}$$

where $q$ is a weight function.

Finally, Bayesian type of test statistics have the form

$$T^B_{n,L_1}(v) = \sum_{k=1}^{n-1} v(k/n) \Big\{ 2\mathbf{S}^T_{k,L_1} \mathbf{C}_n^{-1} \mathbf{S}_{k,L_1} \Big\}^{1/2}.$$

where $v(1/n), ..., v((n-1)/n)$ represent priors.

Analogously as in the $L_2$ situation large values indicate that $H_o$ does not hold. The exact distribution of the test statistics even under the null hypothesis can be hardly obtained. The limit distributions under $H_o$ can be derived (see Theorem 1 - Theorem 3 below) which are then useful in getting approximations to the critical values.

Now, we pay attention to the limit behavior of the introduced tests statistics under the null hypothesis.

We consider the assumptions:

(i) Random variables $Y_1, ..., Y_n$ follow the model (1) with $n = m$ and the distribution $F$ has median 0 and Lipschitz of order $\gamma_1 \in (0,1]$ and strictly positive density at the median, i.e., $F^{-1}(1/2) = 0$, $|f(0) - f(x)| \le D|x|^{\gamma_1}$ for some $D > 0$ and all $x$ in a neighbourhood of 0 and $f(0) > 0$.

(ii) $\lim_{n\to\infty} \frac{1}{n}\mathbf{C}_{[nt]} = t\mathbf{C}$, $t \in [0,1]$, for some $\mathbf{C} > 0$.

(iii) There exist $\epsilon \in (0,1)$ and $\gamma_2 > 0$ such that, as $n \to \infty$,

$$||\frac{1}{k}\mathbf{C}_k - \mathbf{C}|| = O(k^{-\gamma_2})$$

$$||\frac{1}{n-k}\mathbf{C}_k^* - \mathbf{C}|| = O((n-k)^{-\gamma_2})$$

uniformly for $1 \le k \le n\epsilon$, where $\mathbf{C}$ is the same as in (ii) and $||.||$ denotes the Eucledian norm.

(iv) As, $n \to \infty$,

$$\max_{1\le k<n}\Big\{\frac{1}{k}\sum_{i=1}^{k}||\mathbf{x}_i||^3 + \frac{1}{n-k}\sum_{i=k+1}^{n}||\mathbf{x}_i||^3\Big\} = O(1).$$

(v) $\widehat{f}(0)$ be an estimator of $f(0)$ such that, as $n \to \infty$,

$$\widehat{f}(0) - f(0) = o_p((\log\log n)^{-1/2}).$$

**Theorem 1** *Let assumptions (i), (ii), (iii) and (iv) be satisfied then*

$$\lim_{n\to\infty} P(a(\log n)(T_{n,L_1}^{(3)})^{1/2} \le t + b_p(\log n)) = \exp\{-2\exp\{-t\}\},\ t \in R^1, \tag{26}$$

*where*

$$a(y) = (2\log y)^{1/2}, \quad b_p(y) = 2\log y + \frac{p}{2}\log\log y - \log(2\Gamma(p/2)),\ y > 1, \tag{27}$$

*and*

$$\Gamma(z) = \int_0^\infty t^{z-1}\exp\{-t\}dt.$$

*If, moreover, (v) is satisfied then the assertion (26) remains true if* $(T_{n,L_1}^{(3)})^{1/2}$ *is replaced by* $(T_{n,L_1}^{(1)})^{1/2}$ *or* $(T_{n,L_1}^{(2)})^{1/2}$.

**Theorem 2** *Let assumptions (i), (ii) and (iv) be satisfied then, as* $n \to \infty$,

$$(T_{n,L_1}(q))^{1/2} \to^D \sup_{0<t<1}\Big\{\frac{(\sum_{i=1}^{p} B_i^2(t))^{1/2}}{q(k/n)}\Big\}, \tag{28}$$

*and*

$$T^o_{n,L_1}(q)) \to^D \sup_{0<t<1} \Big\{\frac{|B_1(t)|}{q(k/n)}\Big\}, \tag{29}$$

$$T^B_{n,L_1}(v) \to^D \int_0^1 v(t)(\sum_{i=1}^p B_i^2(t))^{1/2} dt, \tag{30}$$

*where* $\{B_j(t); t \in (0,1)\}$ *are independent Brownian bridges and* $q$ *is the weight function defined by either (8) or (9) and* $v(t) = (q(t)\sqrt{t(1-t)})^{-1}$, $t \in (0,1)$.

**Theorem 3** *Let assumptions (i), (ii) and (iv) be satisfied and let, as* $n \to \infty$,

$$G/n \to 0, \qquad G^{-1} n^{2/3} \log n \to 0, \tag{31}$$

*then*

$$\lim_{n\to\infty} P\Big(a(\log(n/G))T^*_{n,L_1}(G) \le t + b_1(\log(n/G) + \log 2\Big) \tag{32}$$

$$= \exp\{-2\exp\{-t\}\}, \; t \in R^1,$$

$$\lim_{n\to\infty} P\Big(a(\log(n/G))T^{**}_{n,L_1}(G) \le t + b_1(\log(n/G) + \log 3\Big) \tag{33}$$

$$= \exp\{-2\exp\{-t\}\}, \; t \in R^1.$$

The assertions of Theorem 1-3 remain true if the $L_1$ -test statistics are replaced by $LSE$ -test statistics and if in assumption (i) the request of zero median is replaced by the request of zero mean and finite absolute moment of order $2 + \Delta$, $\Delta > 0$.

Assertions (26), (32) and (33) are extreme value type theorems. It is known that the convergence in (26), (32) and (33) is rather slow.

The explicit form of the limit distribution in (28), (29) and (30) is known only for some weight function $q$, e.g., for $q$ in (9) with $\gamma = 0$ in Sen (1981) and for $q$ in (8) Siegmund (1987) derived a proper approximation.

Approximation to the critical values corresponding to $T^{(j)}_{n,L_1}$, $j = 1, 2, 3,$, $T^*_{n,L_1}$ and $T^{**}_{n,L_1}$ can be easily calculated using a pocket calculator.

The tests based on either of $T^{(j)}_{n,L_1}$, $j = 1, 2, 3$, $T^B_{n,L_1}$ are consistent for fixed and as well as some local alternatives. Concerning $T_{n,L_1}(q)$, $T^o_{n,L_1}(q)$, $T^*_{n,L_1}$ and $T^{**}_{n,L_1}$ their limit distribution depend on $\delta_n$ and the design matrix. This will be studied in a different paper.

The assumptions (ii) - (iv) imposed on the design matrix are slightly stronger than one usually assumes when studying for example $L_1$ estimators in the model (1) with $\delta_n = 0$.

Concerning the estimator of $f(0)$ we need an estimator that behaves reasonably well not only under the null hypothesis but also under alternatives. Such estimators can be described as follows:

$$\widehat{f}(0) = \frac{\widehat{m}}{n}\widehat{f}_{\widehat{m}}(0) + \frac{n-\widehat{m}}{n}\widehat{f}^*_{\widehat{m}}(0),$$

where $\widehat{f}_{\widehat{m}}(0)$ and $\widehat{f}^*_{\widehat{m}}(0)$ are estimators of $f(0)$ based on $Y_1, ..., Y_{\widehat{m}}$ and $Y_{\widehat{m}+1}, ..., Y_n$ and $\widehat{m}$ is an estimator of possible change point $m$. There is a number of possibilities how to estimate $\widehat{f}_{\widehat{m}}(0)$, $\widehat{f}^*_{\widehat{m}}(0)$ and $\widehat{m}$. Here is one suggestion

$$\widehat{m} = argmax\{||\beta_{k,L_1} - \beta^*_{k,L_1}||;\ k = 1, ..., n\},$$

$$\widehat{f}_{\widehat{m}}(0) = \frac{1}{2\eta\widehat{m}^{1/2}}\sum_{i=1}^{\widehat{m}} I\{\mathbf{x}_i^T\beta_{\widehat{m},L_1} - \widehat{m}^{-1/2}\eta \le Y_i \le \mathbf{x}_i^T\beta_{\widehat{m},L_1} + \widehat{m}^{-1/2}\eta\},$$

where $\eta > 0$ fixed, and $\widehat{f}^*_{\widehat{m}}(0)$ is defined accordingly. Under the assumptions (i),(ii) and (iv) the resulting estimator $\widehat{f}(0)$ has the property requested in the assumption (v).

## 3 Proofs

Since the proofs are quite technical we give only a sketch of them. First, we formulate several technical lemmas that are modifications of results proved elsewhere.

**Lemma 1** *Let assumptions (i) - (iv) be satisfied then for any $\eta > 0$ there exist $A_\eta > 0$ and $n_\eta$ such that for all $n \ge n_\eta$*

$$P\Big(\sup\{|\sum_{i=1}^n(\rho_{L_1}(E_i - n^{-1/2}\mathbf{x}_i^T\mathbf{t}) - \rho_{L_1}(E_i) + n^{-1/2}\mathbf{x}_i^T\mathbf{t}\psi_{L_1}(E_i - n^{-1/2}\mathbf{x}_i^T\mathbf{t})) + \frac{f(0)}{n}\mathbf{t}^T\mathbf{C}_n\mathbf{t}|;\ ||\mathbf{t}|| \le D\} \ge A_\eta n^{-v}\Big) < n^{-\eta} \quad (34)$$

*and*

$$P\Big(\sup\Big\{|\sum_{i=1}^n x_{ij}(\psi_{L_1}(E_i - n^{-1/2}\mathbf{x}_i^T\mathbf{t}) - \psi_{L_1}(E_i) + 2f(0)n^{-1/2}\mathbf{x}_i^T\mathbf{t})|;\ ||\mathbf{t}|| \le D\Big\} \ge A_\eta n^{-v}\Big) < n^{-\eta},\ j = 1, ..., p, \quad (35)$$

*for some $v > 0$ and arbitrary $D > 0$.*

**Proof:** The proof of the first assertion is a simple modification of Lemma 1 in Gutenbrunner et al (1993) and Theorem 1 in Hušková (1994c), while the second assertion follows from Theorem 2 in Hušková (1994c). □

**Lemma 2** *Let assumptions (i) - (iv) be satisfied then for any $\eta > 0$ there exists $A_\eta > 0$ and $n_\eta$ such that for all $n \geq n_\eta$*

$$P\Big(||\mathbf{C}_k^{1/2}(\boldsymbol{\beta}_{k,L_1} - \boldsymbol{\beta}) - \frac{1}{2f(0)}\mathbf{C}_k^{-1/2}\sum_{i=1}^{k}\mathbf{x}_i\psi_{L_1}(E_i)|| \tag{36}$$

$$\geq A_\eta k^{-v}\Big) < k^{-\eta}, k \leq n,$$

$$P\Big(||\mathbf{C}_k^{*1/2}(\boldsymbol{\beta}^*_{k,L_1} - \boldsymbol{\beta}) - \frac{1}{2f(0)}\mathbf{C}_k^{*-1/2}\sum_{i=k+1}^{n}\mathbf{x}_i\psi_{L_1}(E_i)|| \tag{37}$$

$$\geq A_\eta (n-k)^{-v}\Big) < (n-k)^{-\eta},\ k < n,$$

*for some $v > 0$ and arbitrary $D > 0$.*

**Proof***:* The proof follows the line of the proofs of Lemma 1 in Gutenbrunner et al (1993), Theorem 4 in Hušková (1994c) and we apply Lemma 1 of the present paper. □

**Lemma 3** *Let assumptions (i) - (iv) be satisfied then*

$$\lim_{n\to\infty} P\Big(a(\log n)\big(\max_{1\leq k<n}\Big\{\Big(\sum_{i=1}^{k}\mathbf{x}_i\psi_{L_1}(E_i) - \mathbf{C}_k\mathbf{C}_n^{-1}\sum_{i=1}^{n}\mathbf{x}_i\psi_{L_1}(E_i)\Big)^T\mathbf{C}_k^{-1}\mathbf{C}_m\mathbf{C}_k^{*-1} \tag{38}$$

$$\Big(\sum_{i=1}^{k}\mathbf{x}_i\psi_{L_1}(E_i) - \mathbf{C}_k\mathbf{C}_n^{-1}\sum_{i=1}^{n}\mathbf{x}_i\psi_{L_1}(E_i)\Big)\Big\}^{1/2} \leq t + b_p(\log n)\Big)$$

$$= \exp\{-2\exp\{-t\}\},\ t \in R^1.$$

**Proof:** The proof follows the line of Theorem 1.1 in Horváth (1995). □

**Lemma 4** *Let assumptions (i), (ii) and (iv) be satisfied then, as $n \to \infty$,*

$$\{\mathbf{C}_n^{-1/2}\sum_{i=1}^{[nt]}\mathbf{x}_i\psi_{L_1}(E_i); t \in (0,1)\} \to^D \{(W_1(t), ..., W_p(t))^T; t \in (0,1)\}, \tag{39}$$

*where $\{W_1(t); t \in (0,1)\}, ..., \{W_p(t); t \in (0,1)\}$ are independent standardized Wiener processes. If, moreover, (31) is fulfilled then*

$$\lim_{n\to\infty} P\Big(a(\log(n/G))\Big\{\max_{1<k\leq n-G}|\sum_{i=k+1}^{k+G}\psi_{L_1}(E_i)|\Big\} \leq t + b_1(\log(n/G)) + \log 2\Big) \tag{40}$$

$$= \exp\{-2\exp\{-t\}\},\ t \in R^1$$

*and*

$$\lim_{n\to\infty} P\Big(a(\log(n/G))\Big\{\max_{G<k\le n-G} \Big|\sum_{i=k+1}^{k+G} \psi_{L_1}(E_i) - \sum_{i=k-G+1}^{k} \psi_{L_1}(E_i)\Big|\Big\} \tag{41}$$

$$\le t + b_1(\log(n/G) + \log 3)\Big) = \exp\{-2\exp\{-t\}\},\ t \in R^1.$$

**Proof:** Since $\sum_{i=1}^{[nt]} \mathbf{x}_i \psi_{L_1}(E_i)$ is the vector of sums of independent random variables with zero mean and finite third absolute moment and since (i), (ii) and (iv) are fulfilled the assertion (39) can be derived using standard arguments. The assertions (40) and (41) are proved, e.g., in Chen (1988). □

**Proof of Theorem 1:** We sketch the proof for $T_{n,L_1}^{(1)}$ only, the proof for $T_{n,L_1}^{(j)}, j = 2, 3$, is omitted because it follows the same line.

Since Lemma 3 it suffices to show that $T_{n,L_1}^{(1)}$ has the same limit distribution as

$$\max_{1\le k<n} L_{n,k}$$

where

$$L_{n,k} = \Big\{\Big(\sum_{i=1}^{k} \mathbf{x}_i \psi_{L_1}(E_i) - \mathbf{C}_k \mathbf{C}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i \psi_{L_1}(E_i)\Big)^T \mathbf{C}_k^{-1} \mathbf{C}_m \mathbf{C}_k^{*-1}$$

$$\Big(\sum_{i=1}^{k} \mathbf{x}_i \psi_{L_1}(E_i) - \mathbf{C}_k \mathbf{C}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i \psi_{L_1}(E_i)\Big)\Big\}^{1/2}.$$

Put

$$V_{n,k} = \Big\{2\widehat{f}(0)\Big(-\sum_{i=1}^{k} |Y_i - x_i^T \beta_{k,L_1}|$$

$$- \sum_{i=k+1}^{n} |Y_i - x_i^T \beta_{k,L_1}^*| + \sum_{i=1}^{n} |Y_i - x_i^T \beta_{n,L_1}|\Big)\Big\}^{1/2},\ k = 1, ..., n-1.$$

Applying Lemma 1 and Lemma 2 we get after tedious but straighforwad calculations that, as $n \to \infty$,

$$\max_{(\log n)^a \le k \le n-(\log n)^a} |L_{n,k} - V_{n,k}| = o_p((\log\log n)^{-1/2}) \tag{42}$$

for all $a > 0$. Moreover, using standard tools we receive also that

$$\max_{1\le k(\log n)^a} (L_{n,k} + V_{n,k}) = o_p(\sqrt{\log\log n}) \tag{43}$$

and

$$\max_{n-(\log n)^a \le k < n} (L_{n,k} + V_{n,k}) = o_p(\sqrt{\log\log n}) \tag{44}$$

for some $a > 0$.

Combining ((42) - (44) we get that the limit distribution of $T^{(1)}_{n,k}$ is the same as $\max_{1\le k<n} L_{n,k}$. The assertion (26) follows. □

**Proof of Theorem 2:** Using Lemma 1 and Lemma 2 we receive that, as $n \to \infty$,

$$\max_{1\le k<n} \Big|\Big|\mathbf{C}_n^{-1/2}(\mathbf{S}_{k,L_1} - \sum_{i=1}^{k} \mathbf{x}_i \psi_{L_1}(E_i)\Big)\Big|\Big| = o_p(1)$$

and

$$\max_{1\le k<n} \Big|n^{-1/2}(S^o_{k,L_1} - \sum_{i=1}^{k} \psi_{L_1}(E_i)\Big)\Big| = o_p(1).$$

The proof can be then finished using classical theorems on the weak convergence of functionals of partial sums of independent random variables. □

**Proof of Theorem 3** By Lemma 1 and Lemma 2 we get after some standard steps that, as $n \to \infty$,

$$\max_{1\le k<n-G} \Big|n^{-1/2}(S^o_{k+G,L_1} - S^o_{k,L_1} - \sum_{i=k+1}^{k+G} \psi_{L_1}(E_i))\Big| = o_p((\log n)^{-1/2}).$$

The assertions (32) and (33) then follows from (40) and (41) in Lemma 4. □

## References


[1] Antoch, J. and Hušková, M. (1992). Change point problem. In *Computational Aspects of Model Choice*, Ed. J. Antoch, pp. 11-38. Heidelberg: Physica Verlag.

[2] Broemling, L.D. and Tsurumi, H. (1987). Econometrics and Structural change over time (with discussion). Marcel Dekker.

[3] Brown, R.L.,Durbin, J. and Evans, J.M. (1975). Techniques for testing the constancy of regression relationships over time (with discussion). *J. R. Statist. Soc.* B **37** 149–192.

[4] Chen, X. (1988). Inference in a simple change-point problem. *Scientia Sinica* A **31**, 654–667.

[5] Csörgö, M. and Horváth, L. (1997). *Limit Theorems in Change Point Analysis.* New York: Wiley.

[6] Gombay, E. and Horváth, L. (1994). Limit theorems for change in linear regression. *J. Mult. Anal.* **48**, 43–69.

[7] Gutenbrunner, C., Jurečková, J., Koenker, R. and Portnoy, S. (1993). Tests of linear hypotheses based on regression rank scores. *J. Nonpar. Statist.* **2** 307–331.

[8] Hackl, P. (1980). *Testing the Constancy of Regression Relationships over Time.* Göttingen: Vandenhoeck and Ruprecht.

[9] Hackl, P. Ed. (1989). *Statistical Analysis and Forecasting of Economic Structural Change.* IIASA, Springer Verlag.

[10] Hackl, P. and .Westlund, A.H. Eds. (1991). *Economic Structural Change; Analysis and Forecasting.* IIASA, Springer Verlag.

[11] Horváth, L. (1993). The maximum likelihood method for testing changes in parameters of normal observations. *Ann. Statist.* **21** 671–680.

[12] Horváth, L. (1995). Detecting changes in linear regression. *Statistics* **26** 189–208.

[13] Horváth, L. (1993). Limit theorems for the union intersection test. *J. Statist. Plan. Infer.* **44** 133—148.

[14] Hušková, M. (1990a). Some asymptotic results for robust procedures for testing the constancy of regression models over time. *Kybernetika* **26** 392–403.

[15] Hušková, M. (1990b). Asymptotics for robust MOSUM. *Commentationes Mathematicæ Universitatis Carolinæ* **31** 345–356.

[16] Hušková, M. (1994a). Nonrecursive procedures for detecting change in simple linear regression models. In *New Progress in Probability and Statistics,* Eds. M.L. Puri and J.P.Vilaplana, pp. 105–126.

[17] Hušková, M. (1994b). Nonparametric procedures for a change in simple linear regression models. *J. Applied Statistical Sciences* **1** 337–358.

[18] Hušková, M. (1994c). Sequential procedures based on rank regression scores. *J. Nonpar. Statist.* **3** 337–358.

[19] James, B., James, K.L. and Siegmund, D. (1987). Tests for changepoint. *Biometrika* **74** 71—83.

[20] Jandhyala, V.K. and MacNeill, I.B. (1989). Residual partial sum limit process for regression models with applications to detecting parameter changes at unknown times. *Stochastic Processes and their Applications* **33** 309–323.

[21] Jandhyala, V.K. and MacNeill, I.B. (1991). Tests for parameter changes at unknown times in linear regression models. *J. Statist. Plann. Infer.* **27** 291–316.

[22] Jandhyala, V.K. and MacNeill, I.B. (1992). On testing for the constancy of regression coefficients under random walk and change-point

alternatives. *Econometric Theory* **8** 501–517.

[23] Kim, Hyune–Ju and Siegmund, D. (1989). The likelihood ratio test for a change point in simple linear regression. *Biometrika* **76** 409–423.

[24] Maddala, G.S. (1977). *Econometrics.* New York: McGraw-Hill.

[25] Ploberger, W. and Krämer, W. (1992). The CUSUM test with OLS residuals. *Econometrica* **60** 271-285.

[26] Quandt, R.E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *J. Am. Statist. Assoc.* **53** 873–880.

[27] Quandt, R.E. (1960). Tests of the hypothesis that a linear regression system obey two separate regimes. *J. Am. Statist. Assoc.* **55** 342–330.

[28] Sen, P.K. (1980). Asymptotic theory of some tests for a possible change in the regression slope occuring at an unknown time point. *Z. Wahr. Verw. Geb.* **52** 203–218.

[29] Sen, P.K. (1981). *Sequential Nonparametrics.* New York: Wiley.

[30] Sen, P.K. (1982). Asymptotic theory of some tests for constancy of regression relationships over time. *Statistics* **13** 21–31.

[31] Sen, P.K. (1984). Recursive M–tests for the constancy of multivariate regression relationships over time. *Sequential Analysis* **3** 191–211.

[32] Worsley, K.J. (1983). Testing for a two–phase multiple regression. *Technometrics* **25** 35–42.