ROBUST'98, 225 – 231

L_1 – PROCEDURES FOR CHANGE POINT PROBLEM

Tomáš VÍŠEK[1]

University of Economics, Prague

Abstrakt. Cílem článku je studium L_1-postupů (testů a odhadů bodu změny) pro model změny polohy rozdělení posloupnosti nezávislých náhodných veličin. Uvažujeme model

$$\begin{aligned} Y_i &= \mu + \epsilon_i, & i &= 1, \ldots, m, \\ Y_i &= \mu + \delta_n + \epsilon_i, & i &= m+1, \ldots, n, \end{aligned}$$

kde $1 \le m \le n$ je neznámý bod změny, μ a $\mu + \delta_n$ jsou neznámé parametry před a po změně. Za předpokladu normality rozdělení chyb, můžeme odvodit test založený na metododě maximální věrohodnosti. Na tento přístup můžeme pohlížet jako na L_2 metodu. Záměnou L_2-vzdálenosti za L_1 získáme několik L_1-testů a odhadů bodu změny. Jejich rozdělení je odvozeno.

Резюме. Статья приносит формулиробку L_1-метода (тест и оценку точки изменения) для параметра сдвига последовательности независимих случайных переменных (проблема разладки). Будем рассматривать модель

$$\begin{aligned} Y_i &= \mu + \epsilon_i, & i &= 1, \ldots, m, \\ Y_i &= \mu + \delta_n + \epsilon_i, & i &= m+1, \ldots, n, \end{aligned}$$

где $1 \le m \le n$ неизвестная точка изменения, μ и $\mu + \delta_n$ неизвестные параметри перед и после изменения. Если мы предположим нормальное распределение ошибок, потом можно найти тест для изменения, который основан на максимуме вероятностного отношения. Этот метод можно учитывать как L_2. Изменением L_2-норми на L_1 можно получить несколько L_1-тестов и оценок точки изменения. Их распределение тоже в статье получено.

1. Main results

The paper deals with L_1-procedures for detection of a change in location models. We follow up the paper of Hušková [2], who proposed procedures for testing the constancy of regression relationship over time. The main aim of this paper is to deal with these tests for the special case of location model and to simplify the assumptions in this easier case. We consider the following model:

$$\begin{aligned} Y_i &= \mu + \epsilon_i \ , & i &= 1, \ldots, m, \\ Y_i &= \mu + \delta_n + \epsilon_i \ , & i &= m+1, \ldots, n, \end{aligned} \tag{1}$$

where $1 \le m \le n$ is an unknown change point, μ and $\mu + \delta_n$ are unknown parameters before and after change and $\delta_n \neq 0$. Further, ϵ_i , $\forall i = 1, \ldots, n$ are independent random variables with common distribution function F such

[1] Tato práce vznikla s podporou grantu GA ČR 1163/201/97.

that, $F(0) = \frac{1}{2}$. We want to test whether there is a change in the parameter of the distribution function of Y_i. Namely,

$$H_0 : m = n \qquad \text{against} \qquad H_1 : m < n \ .$$

Procedures for regression case are obtained in [2] by substitution of L_1 norm instead of L_2 one. Different L_1-type estimators of the change point m are introduced and studied in [1]. We use notation:

$$\rho_{L_1}(x) = |x| \quad , \quad x \in R \ ,$$

$$\psi_{L_1}(x) = \begin{cases} -1 \ , & x < 0 \ , \\ 0 \ , & x = 0 \ , \\ 1 \ , & x > 0 \ . \end{cases}$$

The L_1-estimator of parameter μ based on $Y_1, \ldots, Y_k$ is defined as

$$\mu_{k,L_1} = \arg\min_{\mu} \sum_{i=1}^{k} \rho_{L_1}(Y_i - \mu).$$

It can be easily checked, that μ_{k,L_1} is median of observations $Y_1, \ldots, Y_k$. Equivalently, we denote μ^*_{k,L_1} median of $Y_{k+1}, \ldots, Y_n$. Now we can simplify test statistics introduced in [2] and obtain the following ones:

$$T^{(1)}_{n,L_1} = \max_{1<k<n} \left\{ 2\hat{f}\left(F^{-1}(1/2)\right) \left(\sum_{i=1}^{n} |Y_i - \mu_{n,L_1}| - \sum_{i=1}^{k} |Y_i - \mu_{k,L_1}| - \sum_{i=k+1}^{n} |Y_i - \mu^*_{k,L_1}| \right) \right\} \tag{2}$$

$$T^{(2)}_{n,L_1} = \max_{1<k<n} \left\{ 4\hat{f}^2\left(F^{-1}(1/2)\right) \frac{k(n-k)}{n} \left(\mu_{k,L_1} - \mu^*_{k,L_1}\right)^2 \right\} \tag{3}$$

$$T^{(3)}_{n,L_1} = \max_{1<k<n} \left\{ \frac{n}{k(n-k)} \left(\sum_{i=1}^{k} \psi_{L_1}(Y_i - \mu_{n,L_1}) \right)^2 \right\} \ , \tag{4}$$

where $\hat{f}\left(F^{-1}\left(\frac{1}{2}\right)\right)$ is an estimator of $f\left(F^{-1}\left(\frac{1}{2}\right)\right)$, F^{-1} and f denote the quantile function and the density.

Next, we introduce some modifications. Namely, we define the weighted type test statistic:

$$T_{n,L_1}(q) = \max_{1 \le k < n} \left\{ \frac{\left(\sum_{i=1}^{k} \psi_{L_1}(Y_i - \mu_{n,L_1})\right)^2}{nq^2(k/n)} \right\} , \tag{5}$$

where

$$q(t) = \begin{cases} \sqrt{t(1-t)}\ , & t \in (a_1, a_2)\ , \\ 0\ , & \text{otherwise} \end{cases} \tag{6}$$

and $0 < a_1 < a_2 < 1$ are given, or

$$q(t) = q_\gamma(t) = (t(1-t))^\gamma \quad \text{and} \quad \gamma \in (0, \frac{1}{2}) \text{ is given.} \tag{7}$$

Next, we have statistics based on moving sum of residuals (MOSUM):

$$T^*_{n,L_1}(G) = \max_{G<k<n} \left\{ \frac{1}{\sqrt{G}} \left| \sum_{i=k-G+1}^{k} \psi_{L_1}(Y_i - \mu_{n,L_1}) \right| \right\} \tag{8}$$

and

$$T^{**}_{n,L_1}(G) = \max_{G<k<n-G} \left\{ \frac{1}{\sqrt{2G}} \left| \sum_{i=k+1}^{k+G} \psi_{L_1}(Y_i - \mu_{n,L_1}) - \sum_{i=k-G+1}^{k} \psi_{L_1}(Y_i - \mu_{n,L_1}) \right| \right\}. \tag{9}$$

And finally, we consider Bayesian type of statistic

$$T^B_{n,L_1} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n-1} \nu(\frac{k}{n}) \left| \sum_{i=1}^{k} \psi_{L_1}(Y_i - \mu_{n,L_1}) \right| \ , \tag{10}$$

where $\nu(k/n)$ represent priors. All these statistics are easy to compute and interpret. Particulary statistics $T^{(3)}_{n,L_1}, T_{n,L_1}(q), T^*_{n,L_1}(G)$ and T^B_{n,L_1} are based on difference of number of observations over and under the whole sample median , concerning observations $Y_1, \dots, Y_k$ (respectively $Y_{k-G+1}, \dots, Y_k$ for $T^*_{n,L_1}(G)$). The exact distributions of the above statistics under the null hypothesis are hard to derive. Therefore we usually use some approximations.

We simplify the assumptions of theorems in [2] for the case of location model.

Assumptions:

T(i) Random variables $Y_1, \dots, Y_n$ follow the model (1) with $n = m$ and F has median 0 and Lipsichtz of order $\gamma_1 \in (0, 1]$ and strictly positive density f at the median.

T(ii) $\hat{f(0)}$ be an estimator of $f(0)$ such that, as $n \to \infty$,

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

Now we formulate following theorem:

Theorem 1. *Let assumption T(i) be satisfied then*

$$\lim_{n\to\infty} P\left(a(\log n)(T^{(3)}_{n,L_1})^{1/2} \le t + b(\log n)\right) = \\ = \exp\{-2\exp\{-t\}\}\ ,\ t \in R\ , \tag{11}$$

where

$$a(y) = \sqrt{(2\log y)}\ ,\ b(y) = 2\log y + \frac{\log\log y - \log\pi}{2}.$$

If, moreover, T(ii) is satisfied then (11) remains true if $T^{(3)}_{n,L_1}$ *is replaced with* $T^{(1)}_{n,L_1}$ *or* $T^{(2)}_{n,L_1}$. *Under T(i) also*

$$(T_{n,L_1}(q))^{1/2} \xrightarrow{D} \sup_{0<t<1}\left\{\frac{|B(t)|}{q(t)}\right\},$$

$$(T^B_{n,L_1}(\nu)) \xrightarrow{D} \int_0^1 \nu(t)|B(t)|,$$

where $\{B(t); t \in (0,1)\}$ *is Brownian bridge and* q *is a weight function defined by (6) or (7) and* $\nu(t) = (q(t)\sqrt{t(1-t)})^{-1}$. *If moreover, as* $n \to \infty$,

$$\frac{G}{n} \to 0\ ,\ \frac{n^{2/3}\log n}{G} \to 0$$

then

$$\lim_{n\to\infty} P\left(a(\log\frac{n}{G})(T^{*}_{n,L_1}(G))^{1/2} \le t + b(\log\frac{n}{G}) + \log 2\right) = \\ = \exp\{-2\exp\{-t\}\}\ ,\ t \in R$$

and

$$\lim_{n\to\infty} P\left(a(\log\frac{n}{G})(T^{**}_{n,L_1}(G))^{1/2} \le t + b(\log\frac{n}{G}) + \log 3\right) = \\ = \exp\{-2\exp\{-t\}\}\ ,\ t \in R\ .$$

Proof: All the assertions are special case of Theorems 2.1., 2.2. and 2.3. in [2]. We only need to verify the assumptions (ii), (iii) and (iv) in cited paper. It is easy, because matrix $C_l = l$, $\forall l$, we have $\lim_{n\to\infty} \frac{1}{n}C_{[nt]} = \lim_{n\to\infty} \frac{[nt]}{t} = t$, that is $C = 1$. Further $\frac{1}{k}C_k - C = 0$ and $\frac{1}{n-k}C^*_k - C = 0$ and the only regressor $x_i = 1$, $\forall i$, that is $\frac{1}{k}\sum_{i=1}^{k}||x_i||^3 + \frac{1}{n-k}\sum_{i=k+1}^{n}||x_i||^3 = 2.$ □

So now we have derived some tests for the presence of change in model (1). But if we reject null hypothesis we would like to know the time of change too. That is some estimators of m. Here are some suggestions that arise directly from the test statistics.

$$(12)\qquad m^{(1)}_{n,L_1} = \arg\max_{1<k<n} \left\{ \left(\sum_{i=1}^{n} |Y_i - \mu_{n,L_1}| - \sum_{i=1}^{k} |Y_i - \mu_{k,L_1}| - \sum_{i=k+1}^{n} |Y_i - \mu^*_{k,L_1}| \right) \right\} ,$$

$$(13)\qquad m^{(2)}_{n,L_1} = \arg\max_{1<k<n} \left\{ \frac{k(n-k)}{n} \left(\mu_{k,L_1} - \mu^*_{k,L_1} \right)^2 \right\} ,$$

$$(14)\qquad m^{(3)}_{n,L_1} = \arg\max_{1<k<n} \left\{ \frac{n}{k(n-k)} \left(\sum_{i=1}^{k} \psi_{L_1}(Y_i - \mu_{n,L_1}) \right)^2 \right\} ,$$

$$(15)\qquad m_{n,L_1}(q) = \arg\max_{1\le k<n} \left\{ \frac{\left(\sum_{i=1}^{k} \psi_{L_1}(Y_i - \mu_{n,L_1}) \right)^2}{nq^2(k/n)} \right\} ,$$

and

$$(16)\qquad m^{**}_{n,L_1}(G) = \arg\max_{G<k<n-G} \left\{ \frac{1}{\sqrt{2G}} \left| \sum_{i=k+1}^{k+G} \psi_{L_1}(Y_i - \mu_{n,L_1}) - \sum_{i=k-G+1}^{k} \psi_{L_1}(Y_i - \mu_{n,L_1}) \right| \right\} .$$

Now the distributions of these estimators under the alternative are of interest. Similarly, as for the statistics itself the exact distribution is hard to obtain. So we are interested in asymptotic behaviour.

First we need some assumptions on the alternatives:

M(i) Random variables Y_i follow the model 1 and F has median 0 and continuos and strictly positive density f at the median.

M(ii) There exists $\eta \in (0,1)$ such that $m = [n\eta]$.

M(iii) $\delta_n \to 0$, $|\delta_n|\sqrt{\frac{n}{\log\log n}} \to \infty$, as $n \to \infty$.

These assumptions correspond to local change alternative, when the change δ_n is small and decreases to zero.

First theorem concerns distribution of $m_{n,L_1}(q_\gamma)$ and special case $m^{(3)}_{n,L_1}$.

Theorem 2. *Let assumptions M(i), M(ii) and M(iii) be satisfied. Then*

$$4f^2(0)\delta_n^2(m_{n,L_1}(q_\gamma) - m) \xrightarrow{D} \min\left\{ z \in R; \max_{t\in R}\{W(t) - |t|g(t)\} = W(z) - |z|g(z) \right\} ,$$

where $0 \le \gamma \le \frac{1}{2}$ *and*

$$g(t) = \begin{cases} (1-\gamma)(1-\eta) - \gamma\eta \ , & t < 0 \ , \\ (1-\gamma)\eta + \gamma(1-\eta) \ , & t > 0 \end{cases}$$

and

$$W(t) = \begin{cases} W_1(-t) \ , & t < 0 \ , \\ W_2(t) \ , & t > 0 \ , \end{cases}$$

with $\{W_1(t), t > 0\}$ *and* $\{W_2(t), t > 0\}$ *being independent Wiener processes.*

Particularly, we have

$$4f^2(0)\delta_n^2(m^{(3)}_{n,L_1} - m) \xrightarrow{D} \min\left\{ z \in R; \max_{t \in R}\left\{ W(t) - \frac{|t|}{2} \right\} = W(z) - \frac{|z|}{2} \right\}.$$

Proof: Because $\frac{1}{n}\sum_{i=1}^n |\psi_{L_1}(Y_i - \mu_{n,L_1})|^{2+\Delta} \le 1 = O_p(1)$ and B.1 - B.3 are satisfied by M(i) - M(iii), we only need to verify assumption B.4 and B.5 in [3] and we can use Theorem 32.3.1 of that paper.

Now because $\lambda_1(t,\upsilon) = \int -\psi_{L_1}(x-t)dF(x-\upsilon) = 2F(t-\upsilon) - 1$ and $\lambda_{2+\Delta}(t,\upsilon) = 1$, $\forall \Delta \ge 0$. So $\lambda_1^{(1)}(t,\upsilon) = \frac{\partial \lambda_1(t,\upsilon)}{\partial t} = 2f(t-\upsilon)$ and under M(i) it is positive for t in neighborhood of 0 and $\upsilon = 0$, $\lambda_1(t,\upsilon)$ continuous in neighborhood of $(0,0)$, $\lambda_1(0,0) = 0$, and $\lambda_1^{(1)}(0,0) = -\frac{\partial \lambda_1(t,\upsilon)}{\partial \upsilon} = -2f(t-\upsilon)$. We verified assumption C.1, C.2 and C.3 of Theorem 32.3.3 and so the assumptions B.4 and B.5 with $b = 2f(0)$ and $c = 1$. □

Theorem 3. *Let assumptions M(i), M(ii) and M(iii) be satisfied and, moreover, as* $n \to \infty$,

$$G \to \infty \ , \qquad \frac{n^{2/(2+\Delta)}\log n}{G} \to 0$$

and

$$\frac{|\delta_n|\sqrt{G}}{\sqrt{\log\log n}} \to \infty.$$

Then

$$4f^2(0)\delta_n^2(m^{**}_{n,L_1}(G) - m) \xrightarrow{D}$$

$$\min\left\{ z \in R; \max_{t \in R}\left\{ W(t) - \frac{|t|}{\sqrt{6}} \right\} = W(z) - \frac{|z|}{\sqrt{6}} \right\},$$

where

$$W(t) = \begin{cases} W_1(-t) \ , & t < 0 \ , \\ W_2(t) \ , & t > 0 \ , \end{cases}$$

with $\{W_1(t), t > 0\}$ *and* $\{W_2(t), t > 0\}$ *being independent Wiener processes.*

Proof: The same as in the preceding case. □

References

[1] Bai J.: Least Absolute Deviation Estimation of a Shift, *Econometric Theory, 11, (1995), 403 - 436*.

[2] Hušková M.: L_1-Test Procedures for Detection of Change, *In Y.Dodge (ed.): L_1-Statistical Procedures and Related Topics, (1997), 57 - 70*.

[3] Hušková M.: Limit Theorems for M-Processes Via Rank Statistics Processes, *In N.Balakrishnan (ed.): Advances in Combinatorial Methods and Applications to Probability and Statistics, (1997), 521 - 534*.