ESTIMATION OF JUMP REGRESSION FUNCTION

Qiu, Peihua
Department of Statistics and Operations Research, Fudan University

Asano, Chooichiro
Research Institute of Fundamental Information Science, Kyushu University→Department of Information System Science, Soka University

Li, Xianping
Department of Statistics and Operations Research, Fudan University

https://doi.org/10.5109/13419

出版情報：Bulletin of informatics and cybernetics. 24 (3/4), pp.197-212, 1991-03. Research Association of Statistical Sciences

Bulletin of Informatics and Cybernetics, Vol. 24, No. 3~4, 1991

ESTIMATION OF JUMP REGRESSION FUNCTION

By

Peihua QIU*, Chooichiro ASANO and Xianping LI***

Abstract

Qui [1] discussed the estimation problem of jump regression functions which were divided into eight types. L^2-consistent estimates of two types of them were obtained. This paper studies further this topics and obtains L^2-consistent estimates of the other four types. For the last two types, the authors also put forwards an estimate. A little numerical results are also given.

1. Introduction

Qiu [1] discussed the estimation problem of jump regression function $f(t)$.

$$f(t) = g(t) + \sum_{i=1}^{p} d_i I_{(s_i, s_{i+1})}(t), \quad t \in [0, 1], \tag{1}$$

where $g(t)$ is continuous, p is the number of jumps, $\{s_i\}_1^p$ are jump points, $\{d_i - d_{i-1}\}_1^p$ are jump magnitudes, $d_0 = 0$, $s_0 = 0$, $s_{p+1} = 1$. According to the information about the number of jumps, their locations and jump magnitudes, we can divide the jump regression functions into eight types. Qiu [1] obtained L^2-consistent estimates of the first two types, namely, p, $\{s_i\}_1^p$ are known, $\{d_i - d_{i-1}\}_{i=1}^p$ are known or unknown. It also discussed the case when the locations of jumps have some indeterminacy. In this paper, we will discuss the other types of jump regression functions. The idea of difference kernel estimate is suggested in Section 2 and with it we obtain a kind of L^2 and a.s. consistent estimates of jump locations and jump magnitudes of the (III) and (IV) jump regression functions. Using hypothesis testing, we gain a kind of L^2 and a.s. consistent estimates of the unknown parts of the (V) and (VI) jump regression functions in Section 3. We also put forwards a kind of estimates of the (VII) and (VIII) jump regression functions in this section. In the end, a little numerical results are given in Section 4.

2. The (III) and (IV) Jump Regression Functions and Difference Kernel Estimate

In this section, we will discuss the (III) and (IV) jump regression functions,

* Department of Statistics and Operations Research, Fudan University, Shanghai 200433, China.
** Research Institute of Fundamental Information Science, Kyushu University, Fukuoka 812, Japan.
At present, Department of Information System Science, Soka University, Hachiouji, Tokyo, Japan.

namely, the number of jumps is known, the jump locations are unknown and the jump magnitudes are known or unknown. If we can resolve the estimation problem of the (IV), we can resolve the (III) naturally. First of all, assume that regression function $f(t)$ has one jump point $0 < \zeta < 1$. Now we want to estimate ζ and the corresponding jump magnitude C with observations $\{y_i\}_1^n$.

Suppose two kernels $K_I(x)$ and $K_2(x)$ satisfy:

(i) $K_1(x) = 0$ for $x \overline{\in}[0, 1]$; $K_2(x) = 0$ for $x \overline{\in} [-1, 0]$,

(ii) $K_i(x) \geq 0, \quad i = 1, 2, \quad x \in R^1$,

(iii) $\int_1^{-1} K_i(x)dx = 1, \quad i = 1, 2.$

Define

$$M_n(t) = \sum_{i=1}^n y_i \int_{t_{i-1}}^{t_i} h_n^{-1} \left[K_1\left(\frac{z-t}{h_n}\right) - K_2\left(\frac{z-t}{h_n}\right)\right] dz, \tag{2}$$

$$0 < h_n \leq t \leq 1 - h_n, \quad \lim_{n\to\infty} nh_n = \infty, \quad \lim_{n\to\infty} h_n = 0,$$

$$|M_n(t_n^*)| \triangleq \max_{h_n \leq t \leq 1-h_n} |M_n(t)|.$$

In the condition of the continuities of $K_i(x)$, $i = 1, 2$, t_n^* exists. Of course, the points which have the property of t_n^* would be more than one. In such a case, t_n^* is the arbitrary one of them.

The t_n^* and $M_n(t_n^*)$ are taken to be the estimates of ζ and C, respectively. We call them Difference Kernel Estimates. In the following, we will illustrate that t_n^* and $M_n(t_n^*)$ are a.s. consistent estimates of ξ and C.

LEMMA 1. *(1) Assume $K(.)$ is boundary, $E|\varepsilon_I|^p < \infty$, $p \geq 2$.*

$$\delta_\eta = \max_{1\leq i\leq n} (t_i - t_{i-1}) = O(n^{-1}),\ h_n = n^{-d},\ 0 < d < 1 - \frac{1}{p}.$$

If regression function $g(t)$ is continuous on $[0, 1]$, then for $\forall t \in (0, 1)$,

$$gn(t) \triangleq \sum_{i=1}^n y_i \int_{t_{i-1}}^{t_i} h_n^{-1} \left[K\left(\frac{z-t}{h_n}\right)\right] dz \text{ converges to } g(t) \text{ almost surely.}$$

(2) Furthermore, if $g(t) \in Lip(1)$, $K(.) \in Lip(1)$, for $0 < a \leq b \leq 1$,

$$\| g_n(t) - g(t) \|_{[a,b]} \xrightarrow{n\to\infty} 0, \text{ a.s.} \quad ,\text{where } \| f(t)|_{[a,b]} \triangleq \max_{a\leq t\leq b} f(t)|.$$

The proof can be found in Cheng and Ling [3]. From Lemma 1 (1), we have

$$\lim_{n\to\infty} M_n(\xi) = C, \text{ a.s.} \tag{3}$$

Denote

$$M_n(t) = M_n^{(1)}(t) - M_n^{(2)}(t),$$

where $$M_n^{(j)}(t) = \sum_{i=1}^{n} y_i \int_{t_{i-1}}^{t_i} h_n^{-1} K_j \left(\frac{z-t}{h_n}\right) dz, \qquad j = 1, 2.$$

$$\begin{cases} f_1(t) = f(t), & t \in [0, \xi], \\ f_2(t) = f(t), & t \in [\xi, 1]. \end{cases}$$

We extend $f_1(t)$ to be $f_1^*(t)$, where the domain is $\left[-\frac{1-\xi}{2}, \frac{1+\xi}{2}\right]$, namely $f_1^*(t) = f_1(t)$ for $t \in [0, \xi]$. If $f_1^*(t) \in Lip(1)$, then from Lemma 1 (2),

$$\| M_n^*(t) \|_{[0,\xi]} \le \| M_n^{*(1)}(t) - f_1^*(t) \|_{[0,\xi]} + \| M_n^{*(2)}(t) - f_1^*(t) \|_{[0,\xi]} \xrightarrow{n\to\infty} 0, \text{ a.s.}$$

where $M_n^*(t)$ is defined like (2) with the regression $f_1^*(t)$, $M_n^*(t) = M_n^{*(1)} - M_n^{*(2)}(t)$, the meaning of $M_n^{*(i)}(t)$ is like $M_n^{(i)}(t)$, $i = 1, 2$.

$$\therefore \quad \| M_n(t) \|_{(h_n, \xi-h_n)} = \| M_n^*(t) \|_{(h_n, \xi-h_n)} \le \| M_n^*(t) \|_{[0,\xi]} \xrightarrow{n\to\infty} 0, \text{ a.s.}$$

For the same reason,

$$\| M_n(t) \|_{(\xi+h_n, 1-h_n)} \xrightarrow{n\to\infty} 0, \text{ a.s.}$$

Combining the two parts above with (3), we have $\xi - h_n \le t_n^* \le \xi + h_n$, a.s., when n is large enough. Then $\lim_{n\to\infty} t_n^* = \xi$, a.s.

Without loss of generality, we assume $C > 0$.

$$M_n(t_n^*) = \sum_{i=1}^{n} y_i \int_{t_{i-1}}^{t_i} h_n^{-1} \left[K_1\left(\frac{z-t_n^*}{h_n}\right) - K_2\left(\frac{z-t_n^*}{h_n}\right)\right] dz.$$

Obviously,

$$M_n(t_n^*) \ge \begin{cases} \sum_{i=1}^{n} \tilde{y}_i h_n^{-1} \int_{t_{i-1}}^{t_i} \left[K_1\left(\frac{z-t_n^*}{h_n}\right) - K_2\left(\frac{z-t_n^*}{h_n}\right)\right] dz, & \xi - h_n \le t_n^* \le \xi, \\ \sum_{i=1}^{n} \underset{\sim}{y}_i h_n^{-1} \int_{t_{i-1}}^{t_i} \left[K_1\left(\frac{z-t_n^*}{h_n}\right) - K_2\left(\frac{z-t_n^*}{h_n}\right)\right] dz, & \xi < t_n^* \le \xi + h_n, \end{cases}$$

where $$\tilde{y}_i = y_i - C1_{t_i > \xi}, \quad \underset{\sim}{y}_i = y_i + C1_{t_i \le \xi}, \quad i = \overline{1,n}.$$

If $f(t) - C1_{t \ge \xi} \triangleq \tilde{f}(t) \in Lip(1)$, then

$$\left|\sum_{i=1}^{n} \tilde{y}_i \int_{t_{i-1}}^{t_i} h_n^{-1} \left[K_1\left(\frac{z-t_n^*}{h_n}\right) - K_2\left(\frac{z-t_n^*}{h_n}\right)\right] dz \right| \le \max \left|\sum_{i=1}^{n} \underset{\sim}{y}_i \int_{t_{i-1}}^{t_i} h_n^{-1} \left[K_1\left(\frac{z-t}{h_n}\right) - K_2\left(\frac{z-t}{h_n}\right)\right] dz \right| \xrightarrow{n\to\infty} 0, \text{ a.s. }, \text{ for } \frac{\xi}{2} \le t \le \frac{1+\xi}{2}.$$

For the same reason,

$$\left|\sum_{i=1}^{n} y_i|_{t_{i-1}}^{t_i} h_n^{-1}\left[K_1\left(\frac{z-t_n^*}{h_n}\right)-K_2\left(\frac{z-t_n^*}{h_n}\right)\right]dz\right| \xrightarrow{n\to\infty} 0, \quad \text{a.s.}$$

$$\therefore \lim_{n\to\infty} M_n(t_n^*) \geq 0, \quad \text{a.s.} \tag{4}$$

and also

$$\therefore \lim_{n\to\infty} |M_n(t_n^*)| \geq \lim_{n\to\infty} |M_n(\xi)| = C \tag{5}$$

From (4) and (5), we have

$$\lim_{n\to\infty} M_n(t_n^*) \geq C, \quad \text{a.s.} \tag{6}$$

on the other hand, $M_n(t_n^*)$ is not greater than

$$\sum_{i=1}^{n} \tilde{y}_i h_n^{-1}\int_{t_{i-1}}^{t_i}\left[K_1\left(\frac{z-t_n^*}{h_n}\right)-K_2\left(\frac{z-t_n^*}{h_n}\right)\right]dz + C\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} h_n^{-1}K_1\left(\frac{z-t_n^*}{h_n}\right)dz,$$

$$\text{for,} \quad \xi - h_n \leq t_n^* \leq \xi;$$

$$\sum_{i=1}^{n} \underset{\sim}{y}_i h_n^{-1}\int_{t_{i-1}}^{t_i}\left[K_1\left(\frac{z-t_n^*}{h_n}\right)-K_2\left(\frac{z-t_n^*}{h_n}\right)\right]dz + C\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} h_n^{-1}K_2\left(\frac{z-t_n^*}{h_n}\right)dz,$$

$$\text{for,} \quad \xi < t_n^* \leq \xi + h_n$$

Obviously,

$$\lim_{n\to\infty}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} h_n^{-1}K_j\left(\frac{z-t_n^*}{h_n}\right)dz = 1, \text{ a.s. } j = 1, 2. \tag{7}$$

$$\therefore \lim_{n\to\infty} M_n(t_n^*) \geq C, \quad \text{a.s.}$$

From (6) and (7), we have $\lim_{n\to\infty} M_n(t_n^*) = C$, a.s. For the $C < 0$ case, we can discuss similarly. Therefore we obtain:

THEOREM 1. *$K_i(x)$ is boundary function and satisfy conditions (i)~(iii), $i = 1, 2$, $E|\varepsilon_1|^p > \infty$, $p \geq 2$, $\delta_n = \max_{1\leq i\leq n} (t_i - t_{i-1}) = O(n^{-1})$, $h_n = n^{-d}$, $0 < d < 1 - \frac{1}{p}$.*

$f_i(t) \in Lip(1)$, $K_i(x) \in Lip(1)$, $i = 1, 2$, then

(i) *t_n^* is a.s. consistent estimate of ξ;*

(ii) *Furthermore, if $f(t) - CI_{t>\xi} \in Lip(1)$, then $M_n(t_n^*)$ is a.s. consistent estimate of C.*

REMARK 1. Without loss of generality, assume $t_i - t_{i-1} = \frac{1}{n}$, $i = \overline{1,n}$

Then the number of items in (2) is $2nh_n$.

If (i) $|K_i(x)| \le M,\ i = 2,$ (M is an arbitrary positive constant.),

(ii) $E|\varepsilon_1|^{2q} < \infty,\ q > 1$, then

$$E|M_n(t_n^*)|^{2q} \leqslant (2nh_n)^{2q-1}\ 2nh_n E|y_1|^{2q}\left(\frac{2M}{nh_n}\right)^{2q} = 2^{2q}E|y_i|^{2q}(2M)^{2q} < \infty\,.$$

Therefore $\{|M_n(t_n^*)|^2,\ n \ge 1\}$ are uniformly integrable. Then $M_n(t_n^*) \xrightarrow[n\to\infty]{L_2} C$, and $t_n^* \xrightarrow[n\to\infty]{L_2} \xi$.

Thus, the spline correction theorem is true in this case, which was given in Qiu [1].

REMARK 2. The non-zero value domains of definition of $K_1(x)$ and $K_2(x)$ can be both extended to $[-T, T]$, where T is an arbitrary positive constant. Namely, $K_1(x) = 0$ when $x\overline{\in}[0, T]$; $K_2(x) = 0$ when $x\overline{\in}[-T, 0]$.

REMARK 3. The non-zero value domains of definition of $K_1(x)$ and $K_2(x)$ must be unilateral in order to guarantee $\lim_{n\to\infty} M_n(\xi) = C$, a.s.

In fact

$$\begin{aligned}
M_n(\xi) &= \sum_{i=1}^{n} y_i\int_{t_{i-1}}^{t_i} h_n^{-1}\left[K_1\left(\frac{z-\xi}{h_n}\right) - K_2\left(\frac{z-\xi}{h_n}\right)\right]dz \\
&= \sum_{i=1}^{n} \tilde{y}_i\int_{t_{i-1}}^{t_i} h_n^{-1}\left[K_1\left(\frac{z-\xi}{h_n}\right) - K_2\left(\frac{z-\xi}{h_n}\right)\right]dz + C\int_{t_r}^{1} h_n^{-1}\left[K_1\left(\frac{z-\xi}{h_n}\right)\right. \\
&\quad \left.-K_2\left(\frac{z-\xi}{h_n}\right)\right]dz \sim C\int_{t_r}^{1} h_n^{-1}\left[K_1\left(\frac{z-\xi}{h_n}\right) - K_2\left(\frac{z-\xi}{h_n}\right)\right]dz \\
&= C\int_0^1 h_n^{-1} K_1\left(\frac{z-\xi}{h_n}\right) dz - C\int_0^{t_r} h_n^{-1} K_1\left(\frac{z-\xi}{h_n}\right)dz - C\int_{t_r}^{1} h_n^{-1} K_2\left(\frac{z-\xi}{h_n}\right)dz \\
&= C\int_{-\xi/h_n}^{(1-\xi)/h_n} K_1(y)dy - C\int_{-\xi/h_n}^{(1-\xi)/h_n} K_1(y)dy - C\int_{(t_r-\xi)/h_n}^{(1-\xi/h_n} K_2(y)dy \\
&\xrightarrow{n\to\infty} C - C\int_{-\infty}^{0} K_1(y)dy - C\int_0^{\infty} K(y)dy
\end{aligned}$$

where t_r satisfy $t_r < \xi,\ t_{r+1} \ge \xi$. In order to guarantee $\lim_{n\to\infty} M_n(\xi) = C$, a.s., it is necessary that $\int_{-\infty}^{0} K_1(y)dy = \int_0^{\infty} K_2(y)dy = 0$. Because of the continuities of $K_1(y)$ and $K_2(y)$, we have $K_1(y) \equiv 0$ when $y \in (0, \infty)$. Namely, the non-zero value domains of definition of $K_1(x)$ and $K_2(x)$ are unilateral.

REMARK 4.

$$\sum_{i=1}^{n} \tilde{y}_i h_n^{-1}\int_{t_{i-1}}^{t_i}\left[K_1\left(\frac{z-t_n^*}{h_n}\right) - K_2\left(\frac{z-t_n^*}{h_n}\right)\right]dz + \frac{C}{h_n}\int_{t_r}^{t_n^*+h_n} K_1\left(\frac{z-t_n^*}{h_n}\right)dz,$$

$$\xi - h_n \le h_n^{-1} \le \xi;$$

$$M_n(t_n^*) \approx \sum_{i=1}^{n} \underset{\sim}{y}_i h_n^{-1} \int_{t_{i-1}}^{t_i} \left[K_1\left(\frac{z-t_n^*}{h_n}\right) - K_2\left(\frac{z-t_n^*}{h_n}\right) \right] dz + \frac{C}{h_n} \int_{t_n^*-h_n}^{t_{r+1}} K_2\left(\frac{z-t_n^*}{h_n}\right) dz,$$

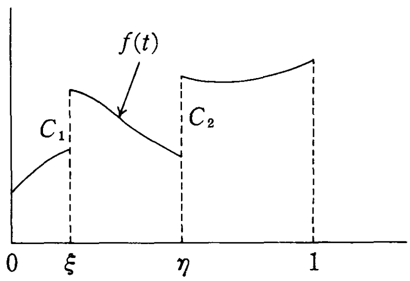
$$\xi < t_n^* \leq \xi + h_n,$$

where t_r satisfy $t_r < \xi$, $t_{r+1} \geq \xi$, because of

$$\frac{C}{H_n} \int_{t_r}^{t_n^*+h_n} K_1\left(\frac{z-t_n^*}{h_n}\right) dz \sim \frac{C}{h_n} \int_{\xi}^{t_n^*+h_n} K_1\left(\frac{z-t_n^*}{h_n}\right) dz = C \int_{(\xi-t_n^*)/h_n}^{1} K_1(y)dy,$$

$$\frac{C}{h_n} \int_{t_n^*-h_n}^{t_{r+1}} K_2\left(\frac{z-t_n^*}{h_n}\right) dz \sim \frac{C}{h_n} \int_{t_n^*-h_n}^{\xi} K_2\left(\frac{z-t_n^*}{h_n}\right) dz = C \int_{-1}^{(\xi-t_n^*)/h_n} K_y(y)dy.$$

Therefore if $\dfrac{\xi - t_n^*}{h_n} \xrightarrow{n\to\infty} 0$, a.s., namely $\xi - t_n^* = o(h_n)$, then $M_n(t_n^*) \underset{n\to\infty}{\overset{a.s.}{\longrightarrow}} C$.



REMARK 5. Now, we discuss the case when the regression function has more then one jumps. Without loss of generality, assume the regression function has two jumps. Let $|M_n(t_n^1)| = \max_{h_n \leq t \leq 1-h_n} |M_n(t)|$, $|M_n(t_n^2)| = \max_{h_n \leq t \leq 1-h_n} |M_n(t)|$, where $\overline{t} \,\overline{\in}\, (t_n^1 - 2h_n, t_n^1 + 2h_n)$ and t_n^1 and t_n^2 are arbitrary points which satisfy the relations above. $t_n^{(1)} < t_n^{(2)}$ are order statistics of t_n^1 and t_n^2. We take $t_n^{(1)}$, $t_n^{(2)}$ as the estimates of ζ and η, the two jump points, respectively. Like the case of unique jump, $t_n^{(1)}$, $t_n^{(2)}$, $M_n(t_n^{(1)})$, $M_n(t_n^{(2)})$ are a.s. and L^2 consistent estimates of ξ and η, C_1 and C_2 respectively under some conditions.

Cheng [2] has the conclusion that

LEMMA 2. *Suppose $U_1, U_2, \ldots, U_n$ are i.i.d. random variables. $U_1 \sim N(a, \sigma^2)$, $\{L \triangleq L_n\}_1^\infty$ are positive integers which satisfy the conditions that $\lim_{n\to\infty} \dfrac{L}{n} = 0$ and*

$$\lim_{n\to\infty} \frac{(\log n)^2}{L} = 0. \text{ Define}$$

$$V_m = \left(\sum_{m-L+1}^{m} U_i - \sum_{m-2L+1}^{m-L} U_i \right) / \sqrt{2L}, \qquad m = \overline{2L,n}.$$

$$W_n = \max \{ [V_m| : m = \overline{2L,n} \},$$

$$A_n(x) = \left\{ 2 \log\left(\frac{3n}{2L} - \right) \right\}^{-\frac{1}{2}} \left\{ x + 2 \log \left(\frac{3n}{2L} - 3 \right) + \frac{1}{2} \log \log \left(\frac{3n}{2L} - 3 \right) - \frac{1}{2} \log \pi \right\}.$$

Then $\lim_{n\to\infty} P\left(\frac{W_n}{\sigma} \leq A_n(x) \right) = \exp(-2e^{-x}), \qquad -\infty < x < \infty.$

In our case, define

$$X_m = \left(\sum_{m-L+1}^{m} Y_i - \sum_{m-2L+1}^{m-1} {}_i \right) / \sqrt{2L}, \qquad m = \overline{2L,n}$$

$$= \left(\sum_{m-L+1}^{m} \varepsilon_i - \sum_{m-2L+1}^{m-L} \varepsilon_i \right) / \sqrt{2L} + \left(\sum_{m-L+1}^{m} f(t_i) - \sum_{m-2L+1}^{m-L} f(t_i) \right) / \sqrt{2L}$$

$$\triangleq X'_m + C_m .$$

If $f(t)$ is differentiable on [0, 1] and $|f'(t)| \leq M < \infty$ $t \in (0, 1)$, then

$$\left| \sum_{m-L+1}^{m} f(t_i) - \sum_{m-2L+1}^{m-L} f(t_i) \right| / \sqrt{2L} = \left| \sum_{i=0}^{L-1} \left[f(t_{m-i}) - f(t_{m-L-i}) \right] \right| / \sqrt{2L}$$

$$\leq M \frac{L^2}{n} / \sqrt{2L} = \frac{\sqrt{2}}{2} M n^{-1} L^{\frac{3}{2}}$$

$$\xi_n \triangleq \sup \{ |X_m|, m = \overline{2L,n} \} \leq \sup \{ |X'_m|, m = \overline{2L,n} \} + \frac{\sqrt{2}}{2} M n^{-1} L^{\frac{3}{2}}$$

$$\triangleq \xi'_n + \frac{\sqrt{2}}{2} M n^{-1} L^{\frac{3}{2}} .$$

If $\varepsilon_1 \sim N(0, \sigma^2)$, then according to Lemma 2,

$$\lim_{n\to\infty} P\left(\frac{\xi'_n}{\sigma} \geq A_n(x) \right) = 1 - \exp(-2e^{-x}), \qquad -\infty < x < \infty.$$

Obviously,

$$P\left(\frac{\xi_n}{\sigma} \geq A_n(x) \right) \leq P\left(\frac{\xi'_n + \frac{\sqrt{2}}{2} M n^{-1} L^{\frac{3}{2}}}{\sigma} \geq A_n(x) \right) \tag{8}$$

$$= P\left\{\frac{\left(\xi_n' + \frac{\sqrt{2}}{2}Mn^{-1}L^{\frac{1}{2}}\right)\sqrt{2\log\left(\frac{3n}{2L} - 3\right)}}{\sigma} \geq x + 2\log\left(\frac{3n}{2L}\right) - 3 + \frac{1}{2}\log\log\left(\frac{3n}{2L} - 3\right) - \frac{1}{2}\log\pi\right\}.$$

If
$$\lim_{n\to\infty} \frac{\sqrt{2}}{2}Mn^{-1}L^{\frac{3}{2}}\sqrt{2\log\left(\frac{3n}{2L} - 3\right)} = 0, \tag{9}$$

then
$$\lim_{n\to\infty} P\left(\frac{\xi_n' + \frac{\sqrt{2}}{2}Mn^{-1}L^{\frac{3}{2}}}{\sigma} \geq A_n(x)\right) = 1 - \exp(-2e^{-x}), \quad x \in R^1. \tag{10}$$

On the other hand, $|X_m| \geq |X_m'| - |C_m| \geq |X_m'| - \frac{\sqrt{2}}{2}Mn^{-1}L^{\frac{3}{2}}$.

$$\therefore\ \xi_n \geq \xi_n' - \frac{\sqrt{2}}{2}Mn^{-1}L^{\frac{3}{2}}.$$

$$\therefore\ P\left(\frac{\xi_n}{\sigma} \geq A_n(x)\right) \geq P\left(\frac{\xi_n' - \frac{\sqrt{2}}{2}Mn^{-1}L^{\frac{3}{2}}}{\sigma} \leq A_n(x)\right). \tag{11}$$

Under the condition of (9), we have

$$\lim_{n\to\infty} P\left(\frac{\xi_n' - \frac{\sqrt{2}}{2}Mn^{-1}L^{\frac{3}{2}}}{\sigma} \geq A_n(x)\right) = 1 - \exp(-2e^{-x}), \quad x \in R^1. \tag{12}$$

Combining (8), (10), (11) and (12), we have

$$\lim_{n\to\infty} P\left(\frac{\xi_n}{\sigma} \geq A_n(x)\right) = 1 - \exp(-2e^{-x}), \quad x \in R^1.$$

From this, we obtain:

THEOREM 2. *Suppose* $\varepsilon_1 \sim N(0, \sigma^2)$, $\{L = L_n\}_1^\infty$ *are positive integers which satisfy* (*i*) $\lim_{n\to\infty} L/n = 0$, (*ii*) $\lim_{n\to\infty} (\log n)^2/L = 0$.

(*iii*) $\lim_{n\to\infty} n^{-1}L^{\frac{3}{2}}\sqrt{\log\left(\frac{3n}{2L} - 3\right)} = 0$. *$f(t)$ is differentiable over* $[0, 1]$ *and* $|f'(t)| \leq M < \infty$, $t \in (0, 1)$. *Then*

$$\lim_{n\to\infty} P\left(\frac{\xi_n}{\sigma} \leq A_n(x)\right) = \exp(-2e^{-x}), \quad x \in R^1.$$

Therefore, we can test the hypothesis H_0: $C = 0 \leftrightarrow H_1$: $C \neq 0$ under the condition that $f(t) - Cl_{t>\xi}$ is differentiable over [0, 1] and its derivative is boundary. The concrete test procedure is that for significance level $\alpha \in (0,1)$, $x(\alpha) = -\log\left(-\frac{1}{2}\log(1-\alpha)\right)$ is the solution of the equation $\exp(-2e^{-x}) = 1 - \alpha$ and we reject HO if and only if $\xi_n > \sigma A_n(x(\alpha))$. This test has the asymptotic significance level α.

The discussion about $\beta(\theta)$, interval estimate of jump point, Th2, Th4, Th5 and many others in Cheng [2] are right in our case after some simple remedies.

3. Estimation of Other Four Types of Jump Regression Functions

First of all, we discuss the (v) and (VI) jump regression functions. In this case, the number of jumps is unknown, but jumps can only take place at m positions $0 < \xi_1 < \xi_2 < \ldots < \xi_m < 1$ on [0, 1].

Suppose there are η_i observations on Gauss' $[0, \xi_i]$, $i = \overline{1,m}$
Let's investigate the r.v. array

$$\left\{(\varepsilon_{nI+i} - \varepsilon_{nI-i})/[n^{\varepsilon}]^{\frac{1}{2}}\sqrt{2}\,\sigma,\ i = \overline{1,[n^{\varepsilon}]}\right\}.$$

Firstly $\sum_{i=1}^{[n^{\varepsilon}]} \operatorname{var}\left[\frac{\varepsilon_{nI+i} - \varepsilon_{nI-i}}{[n^{\varepsilon}]^{\frac{1}{2}}\sqrt{2}\sigma}\right] = 1$, and secondly, for $\forall\eta > 0$,

$$\lim_{n\to\infty} \sum_{i=1}^{[n^{\varepsilon}]} \int_{|x|\geq\eta} x^2 dF_{ni}(x), \quad \left(F_{ni}(x) \text{ is d.f. of } \frac{\varepsilon_{nI+i} - \varepsilon_{nI-i}}{[n^{\varepsilon}]^{\frac{1}{2}}\sqrt{2}\sigma}\right)$$

$$= \lim_{n\to\infty} \int_{|y|\geq[n^{\varepsilon}]^{\frac{1}{2}}\eta} y^2\, dG(y), \quad \left(G(y) \text{ is d.f. of } \frac{\varepsilon_2 - \varepsilon_1}{\sqrt{2}\sigma}\right) = 0.$$

The condition for last equation is that ε_1 has the second order moment.

Therefore the r.v. array above satisfy the Linderberg conditions, and

$$\eta_n \triangleq \sum_{i=1}^{[n\varepsilon]} \frac{\varepsilon_{nI+i} - \varepsilon_{nI-i}}{[n^{\varepsilon}]^{\frac{1}{2}}\sqrt{2}\sigma} \xrightarrow{D} N(0, 1).$$

Now, make the hypothesis at ξ_1 that H_0: there is no jump at $\zeta_1 \leftrightarrow H_1$: $C \neq 0$ is the jump magnitude at ξ_1. We take the following procedure to test this hypothesis. For level α_n, $C_{\alpha n/2}$ is generated by $N(0, 1)$, namely $P(|X| > C_{\alpha n/2}) = \alpha_n$, where $X \sim N(0, 1)$. We reject H_0, if

$$|\xi_n| \triangleq \frac{1}{[n^{\varepsilon}]^{\frac{1}{2}}\sqrt{2}\sigma}\left|\sum_{i=1}^{[n^{\varepsilon}]}(Y_{nI+i} - Y_{nI-i})\right| \geq C_{\alpha n/2}.$$

If H_0 is true

$$\frac{1}{[n^{\varepsilon}]^{\frac{1}{2}}\sqrt{2}\sigma}\left|\sum_{i=1}^{[n^{\varepsilon}]}(f(t_{nI+i})-f(t_{nI-i}))\right| \geq \frac{1}{[n^{\varepsilon}]^{\frac{1}{2}}\sqrt{2}\sigma}\sum_{i=1}^{[n^{\varepsilon}]}|f'(p_i)|(t_{nI+i}-t_{nI-i})$$

$$\leq \frac{M[n^{\varepsilon}]([n^{\varepsilon}]+\ )}{[n^{\varepsilon}]^{\frac{1}{2}}\sqrt{2}\,\sigma} = O(n^{(3\varepsilon-2)/2}),$$

where $t_{nI-i} \leq p_i \leq t_{nI+i}$, $i = \overline{1,[n^{\varepsilon}]}$, $|f'(t)| \leq M$, $t \in (0, 1)$.

If $\varepsilon < 2/3$, *then* $\lim_{n\to\infty}\frac{1}{[n^{\varepsilon}]^{\frac{3}{2}}\sqrt{2}}\sum_{i=1}^{[n^{\varepsilon}]}[f(t_{nI+i})-f(t_{nI-i})]=0.$

Therefore $$\frac{1}{[n\varepsilon]^{\frac{3}{2}}\sqrt{2}\sigma}\sum_{i=1}^{[n^{\varepsilon}]}[Y_{nIIi}-Y_{nI-i}] \xrightarrow{D} N(0, 1).$$

The probability of the first error: $r_n \triangleq P_{H_0}(|\xi_n| \geq C_{\alpha n/2})$. If $\lim_{n\to\infty}\alpha_n = 0$, then

$$\lim_{n\to\infty} C_{\alpha n/2} = \infty.$$

$\lim_{n\to\infty} P_{H_0}(|\xi_n| \geq C_{\alpha n/2}) \leq \lim_{n\to\infty} P_{H_0}(|\xi_n| \geq N) = P(|X| \geq N)$, where N is an arbitrarily set positive number.

$$\therefore \lim_{n\to\infty} r_n = 0$$

The probability of the second error

$S_n \triangleq 1 - P_{H_1}(|\xi_n| \geq C_{\alpha n/2}) = 1 - \{P_{H_1}(\eta_n \geq C_{\alpha n/2} - d_n) + P_{H_1}(\eta_n \leq -C_{\alpha n/2} - d_n)\}$,

for $d_n = \frac{1}{[n^{\varepsilon}]^{\frac{1}{2}}\sqrt{2}\sigma}\sum_{i=1}^{[n^{\varepsilon}]}[f(t_{nI+i}) - f(t_{nI-i})] = \frac{[n^{\varepsilon}]}{[n^{\varepsilon}]^{\frac{1}{2}}\sqrt{2}\sigma}C + O(n^{(3\varepsilon-2)/2}) = O(n^{\varepsilon/2})$.

If $\lim_{n\to\infty}\frac{C_{\alpha n}/2}{d_n} = 0$, then $\lim_{n\to\infty} S_n = 0$.

We can make tests at other points similarly and take the number a_n of rejecting H_0 in all m tests as the estimate of the number of jumps p and the corresponding positions as the estimates of the jump positions. If H_1 is received at ε_j, then we use

$$\frac{1}{[n^{\varepsilon}]}\sum_{i=1}^{[n^{\varepsilon}]}[Y(t_{nj+i}) - Y(t_{nj-i})],$$

as the estimate of jump magnitude at ξ_j.

THEOREM 3. *If* (i) ε^1 *has the second order moment* σ^2 *known*,

(ii) $|f'(t)| \leq M < \infty$, $t \in (0, 1)$, (iii) $\varepsilon < 2/3$, and

(iv) $\lim_{n\to\infty} \alpha_n = 0$ *and* $\lim_{n\to\infty} \frac{C_{\alpha n/2}}{n^{\varepsilon/2}} = 0$, *then* $a_n \xrightarrow{L^2} p$.

Furthermore, if

(v) $\varepsilon_1 \sim N(0, \sigma^2)$, (*vi*) $\sum_{n=1}^{\infty} \exp\left\{-\frac{(C\alpha_{n/2}-1)^2}{2}\right\} < \infty$, then $a_n \xrightarrow{a.s.} p$.

PROOF. Denote $a_n \triangleq p_n + q_n$ where P_n is the number of rejecting H_0 at p points which really have jumps and q_n is the number of rejecting H_0 at $m - p$ points which actually have no jumps.
r_{nj} denote the probability of the (I) error at ξj. From the analysis about two types of errors above, we have $\lim_{n\to\infty} r_{ni} = 0$, $i = \overline{1,m}$, under the conditions of (i)−(iv). Without loss of generality, assume the first p points of $\{\xi_j\}_1^m$ really have jumps and the last m − p points have no.
Then $P(q_n = 0) = \prod_{i=p+1}^{m} (1 - r_{ni}) \xrightarrow{n\to\infty} 1$,

$$Eq_n^2 = 1^2P(q_n = 1) + 2^2P(q_n = 2) + \cdots + (m-p)^2P(q_n = m-p) \xrightarrow{n\to\infty} 0.$$

$$\therefore q_n \xrightarrow{L^2} 0. \text{ Similarly, } p_n \xrightarrow{L^2} 0.$$

$$\therefore a_n = p_n + q_n \xrightarrow{L^2} p.$$

If $\varepsilon_1 \sim N(0, \sigma^2)$, then $r_{ni} = P(|\eta_n + d_{ni}| \geq C_{\alpha n/2})$

$$= P(\eta_n \geq C_{\alpha n/2} - d_{ni}) + P(\eta_n \leq -C_{\alpha n2} \leq d_{ni})$$

$$\leq \frac{1}{\sqrt{2\pi}}\frac{1}{C_{\alpha n/2} + d_{ni}} \exp\left\{-\frac{C_{\alpha n/2} - d_{ni})^2}{2}\right\} + \frac{1}{\sqrt{2\pi}}\frac{1}{C_{\alpha n/2} + d_{ni}} \exp\left\{-\frac{(C_{\alpha n/2} + d_{ni})^2}{2}\right\}$$

$$\leq K \exp\left\{-\frac{(C_{\alpha n/2} - 1)^2}{2}\right\},$$

where $d_{ni} = \frac{1}{[n^\varepsilon]^{\frac{1}{2}}\sqrt{2}\sigma}\sum_{i=1}^{[n^\varepsilon]} [f(t_{n_i+j}) - f(t_{n_i-j})]$, and K is a positive constant.
For $0 < \delta < 1$,

$$\sum_{n=1}^{\infty} P(q_n > \delta) = \sum_{n=1}^{\infty}\left(1 - \prod_{i=p+1}^{m}(1 - r_{ni})\right)$$

$$= \sum_{n=1}^{\infty}\sum_{i=p+1}^{m} r_{ni} - \sum_{n=1}^{\infty}\sum_{i,j=p+1}^{\infty} r_{ni}r_{nj} + \cdots + (-1)^{m-p+1}\sum_{n=1}^{\infty}\prod_{i=p+1}^{m} r_{ni}$$

If we can illustrate that

$$\sum_{n=1}^{\infty}\sum_{i=p+1}^{m} r_{ni} < \infty \text{ , then } \sum_{n=1}^{\infty} P(q_n > \delta) < \infty,$$

but
$$\sum_{n=1}^{\infty}\sum_{i=p+1}^{m} r_{ni} \leq (m - p)\, K \sum_{n=1}^{\infty} \exp\left\{ - \frac{(C_{\alpha n/2} - 1)^2}{2} \right\}.$$

Therefore, according to Borel–Cantelli Lemma, $q_n \xrightarrow{\text{a.s.}} 0$ under the condition of (iv). Similarly, $p_n \xrightarrow{\text{a.s.}}$ p.

$$\therefore\ a_n = p_n + q_n \xrightarrow{\text{a.s.}} \text{p}.$$

The a.s. and L^2 consistent estimate a_n of p discussed above requires σ^2 is known. This condition is very severe. In the following, we will construct L^2-consistent estimate of p under the condition that $\varepsilon_1 \sim N(0, \sigma^2)$ and σ^2 is unknown.
Let

$$X_i = \varepsilon_{nI+i} - \varepsilon_{nI-i,}, \quad X_i' = Y_{nI+i} - Y_{nI-i},$$

$$\text{d}_\text{i} = f(t_{nI+i}) - f(t_{nI-i}), \quad i = \overline{1,[n^\varepsilon]},$$

$$\overline{X} = \frac{1}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} X_{i,} \quad \overline{X'} = \frac{1}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} X_i', \quad \overline{d} = \frac{1}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} d_i\,.$$

$$S^2_{[n^\varepsilon]} = \frac{1}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} (X_i - \overline{X})^2 \quad S'^2_{[n^\varepsilon]} = \frac{1}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} (X_i - \overline{X'})^2\,.$$

Then $\overline{X} \sim N\left(0, \frac{2\,\sigma^2}{[n^\varepsilon]}\right)$, and it is independent of $S^2_{[n^\varepsilon]}$. Making hypothesis at ξ_1 that H_0: there is no jump at $\xi_1 \leftrightarrow H_1$: there is jump at ξ_1 and the jump magnitude is $C \neq 0$.

We can see

$$S'^2_{[n^\varepsilon]} = \frac{1}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} X_i'^2 - \overline{X'}^2 = \frac{1}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} (X_i + d_i)^2 - \overline{X} + \overline{d})^2$$

$$= \frac{1}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} X_i^2 + \frac{2}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} X_i d_i + \frac{1}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} d_i^2 - (\overline{X} + \overline{d})^2,$$

When H_0 is true, it is easy to know that

$$\lim_{n\to\infty} (\overline{X} + \overline{d})^2 = 0, \text{ a.s.}, \ \lim_{n\to\infty} \frac{2}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} d_i^2 = 0 \text{ and } \lim_{n\to\infty} \frac{2}{[n^\varepsilon]}\sum_{i=1}^{[n^\varepsilon]} X_i d_i = 0, \text{ a.s.}$$

Thus $\lim_{n\to\infty} S'^2_{[n^\varepsilon]} = 2\,\sigma^2$, a.s.

When H_1 is true, let $D_i - c = d_i'$, $i = \overline{1,[n^\varepsilon]}$, $\overline{d} - c = \overline{d}'$,

Then $$\lim_{n\to\infty} \sup_{1\le i\le[n^\varepsilon]} |d_i'| = 0 \text{ and } \lim_{n\to\infty} \overline{d}' = 0.$$

Similarly $$S_{[n^\varepsilon]}'^2 = \frac{1}{[n^\varepsilon]} \sum_{i=1}^{[n^\varepsilon]} (X_i' - \overline{X}')^2 = \frac{1}{[n^\varepsilon]} \sum_{i=1}^{[n^\varepsilon]} (X_i + d_i - \overline{X} - \overline{d})^2$$

$$= \frac{1}{[n^\varepsilon]} \sum_{i=1}^{[n^\varepsilon]} (X_i + d_i' - \overline{X} - \overline{d}')^2$$

$$= \frac{1}{[n^\varepsilon]} \sum_{i=1}^{[n^\varepsilon]} X_i^2 + \frac{2}{[n^\varepsilon]} \sum_{i=1}^{[n^\varepsilon]} X_i d_i' + \frac{1}{[n^\varepsilon]} \sum_{i=1}^{[n^\varepsilon]} d_i'^2 - (\overline{X} + \overline{d}')^2.$$

As the case when H_0 is true, we can obtain $\lim_{n\to\infty} S_{[n^\varepsilon]}'^2 = 2\sigma^2$, a.s.

Let $\zeta_n = \frac{1}{[n^\varepsilon]^{\frac{1}{2}}} \sum_{i=1}^{[n^\varepsilon]} (Y_{nI+i} - Y_{nI-i})$. As for the above hypothesis test problem, we can take the following procedure: rejecting H_0 when $|\xi_n| \ge \beta_n$, where $\{\beta_n\}$ is a series of positive real numbers which satisfy $\lim_{n\to\infty} \beta_n = \infty$.

When H_0 is true, we can see that

$$\frac{1}{[n^\varepsilon]^{\frac{1}{2}} S_{[n^\varepsilon]}} \left| \sum_{i=1}^{[n^\varepsilon]} [f(t_{nI+i}) - f(t_{nI-i})] \right| = 0\left(n^{\frac{3\varepsilon-2}{2}}\right), \text{ a.s.}$$

If $\varepsilon < 2/3$, *then*

$$\lim_{n\to\infty} \frac{1}{[n^\varepsilon]^{\frac{1}{2}} S_{[n^\varepsilon]}'} \sum_{i=1}^{[n^\varepsilon]} f(t_{nI+i}) - f(t_{nI-i})] = 0, \quad \text{a.s.}$$

The (first) error probability $r_n \triangleq P_{H_0}(|\xi_n| \ge \beta_n)$

$$\lim_{n\to\infty} r_n \le \lim_{n\to\infty} P_{H_0}(|\xi_n| > N) = 2\,\phi(-N),$$

where N is an arbitrary positive number. Therefore $\lim_{n\to\infty} r_n = 0.$

The second error probability

$$\omega_n \triangleq P_{H_I}(|\xi_n| \ge \beta_n) = E\{P_{H_I}(|\xi_n| < \beta|_n S_{[n^\varepsilon]}')\}$$

$$= E\left\{ P_{H_I}\left(\frac{S_{[n^\varepsilon]}'}{\sqrt{2}\sigma}(-\beta_n - h_n) \le \frac{[n^\varepsilon]^{\frac{1}{2}}\overline{X}}{\sqrt{2}\sigma} \le (\beta_n - h_n)\frac{S_{[n^\varepsilon]}'}{\sqrt{2}\sigma} \middle| S_{[n^\varepsilon]}'\right]\right\},$$

where

$$h_n = \frac{1}{[n^\varepsilon]^{\frac{1}{2}} S'_{[n^\varepsilon]}} \sum_{i=1}^{[n^\varepsilon]} f(t_{nI+i}) - f(t_{nI-i})] = \frac{1}{[n^\varepsilon]^{\frac{1}{2}} S'_{[n^\varepsilon]}} \mathrm{C} + \mathrm{O}\left(n^{\frac{3\varepsilon-2}{2}}\right) = O(n^{\varepsilon/2}), \text{ a.s.}$$

If $\lim_{n\to\infty} n^{-\varepsilon/2}\beta_n = 0$, then

$$P_{H_1}\left[\frac{S'_{[n^\varepsilon]}}{\sqrt{2}\sigma}(-\beta_n - h_n) \leq \frac{[n^\varepsilon]^{\frac{1}{2}}\bar{X}}{\sqrt{2}\sigma} \leq (\beta_n - h_n)\frac{S'_{[n^\varepsilon]}}{\sqrt{2}\sigma}\middle| S'_{[n^\varepsilon]}\right], \xrightarrow{n\to\infty} 0, \text{ a.s.}$$

According to Lebesque Control and Convergence Theorem, $\lim_{n\to\infty} W_n = 0$.

We can make hypothesis tests similarly at $\xi_2, \ldots, \xi_m$ and use the number of rejecting H_0 of all m tests as the estimate of p. It is easy to prove like theorem 3 that it is a L^2-consistent estimate of p.

In the following, we put forwards a type of estimates of the (VII) and (VIII) jump regression functions, where the number of jumps and jump locations are unknown.

Let $$X_i = \left|\frac{1}{[n^\varepsilon]}\sum_{j=1}^{[n^\varepsilon]} (y_{[n^\varepsilon]+i+j} - y_{[n^\varepsilon]+i-j})\right|, \quad i = \overline{1, (n - 2[n^\varepsilon])},$$

where $0<\varepsilon<1$ is a constant. $X_{(1)} \geqslant X_{(2)} \geqslant \cdots \geqslant X_{(n-2[n^\varepsilon])}$ are order statistics of $\{X_i\}_1^{(n-2[n^\varepsilon])}$. We devide them into two parts at the position where the difference between two neighbouring order statistics is maximal. Without loss of generality, the two parts are supposed as $X_{(1)} \geqslant X_{(2)} \geqslant \cdots \geqslant X_{(L)}$ and $X_{(L+1)} \geqslant X_{(L+2)} \geqslant \cdots \geqslant X_{(n-[n^\varepsilon]}$, where $X_{(L)} - X_{(L+1)} = \max_{1\leq i\leq n-n[n^\varepsilon]-1}(X_{(i)} - X_{(i+1)})$. The corresponding statistics of $\{X_{(i)}\}_1^L$ are $X_{i_1+1}, \cdots, X_{i_1+j_1}; X_{i_2+1}, \cdots, X_{i_2+j_2}; \cdots; X_{i_k+1}, \cdots, X_{i_k+j_k}$ where $i_s + j_s < i_{s+1}$, $s = \overline{1, k-1}$, $\sum_{s=1}^{k} = L$. Select a representative in every group of statistics above which corresponding order statistics is ahead of other's in its group. We denote these representatives as $X_{p_1}, X_{p_2}, \cdots, X_{p_k}$, where $i_L + 1 \leqslant p_L \leqslant i_L + j_L$, $L = \overline{1,k}$. Then we take k as the estimate of the number of jumps, $t_{[n^\varepsilon]} + p_1$, $t_{[n^\varepsilon]} + p_2$, $\cdots$, $t_{[n^\varepsilon]} + p_k$ as the estimates of jump positions and $\frac{1}{[n^\varepsilon]}\sum_{j=1}^{[n^\varepsilon]}(y_{[n^\varepsilon]+p_1+j} - y_{[n^\varepsilon]+p_1-j})$, $\cdots$, $\frac{1}{[n^\varepsilon]}\sum_{j=1}^{[n^\varepsilon]}(y_{[n^\varepsilon]+p_k+j} - y_{[n^\varepsilon]+p_k-j})$ as the estimates of corresponding jump magnitudes.

4. Some Numerical Results

For the Difference Kernel Estimate proposed in section 2, we made some simulations after replacing $M_n(t)$ in (3) by

$$\tilde{M}_n(t) = \frac{1}{nh_n}\sum_{j=1}^{n} y_i\left[K_1\left(\frac{t_i-t}{h_n}\right) - K_2\left(\left(\frac{t_i-t}{h_n}\right)\right)\right].$$

The regression function used is $f(x) = x^2 + 3x + 20\, 1_{[0.5,1]}$ $\varepsilon_i \sim N(0, 1)$, $h_n = 1/1nn$. Use scanning method to obtain t_n^*. The number of scannings is B.
Firstly, we choose

$$K(x) = \begin{cases} 1, & x \in [0, 1], \\ 0, & \text{when } x \text{ at other points; n = 50,} \\ -1, & x \in [-1, 0], \end{cases}$$

where $K(x) = K_1(x) - K_2(x)$. The simulation results with varied B are:

B	10	20	50	100	200
t_{50}	0.5	0.45	0.52	0.49	0.495
$\widetilde{M}_{50}(t^*_{50})$	18.153	18.285	20.069	21.448	21.448

Then I set $B = 10$. The results with the same kernel and varied n are:

n	50	100	200	500	1000
t^*_n	0.5	0.5	0.5	0.5	0.5
$\widetilde{M}_n(t^*_n)$	18.153	19.065	20.211	20.087	20.260

Afterwards, we choose

$$K(x) = \begin{cases} \frac{\pi}{2}\sin(nx), & \text{for} \quad |x|\leq 1, \\ 0 \quad , & \text{for } x \text{ at the other points; } n = 50. \end{cases}$$

The simulation results with varied B are:

B	10	20	50	100
t^*_{50}	0.5	0.5	0.48	0.49
$\widetilde{M}_{50}(t^*_{50})$	20.885	20.885	20.891	20.972

With the same kernel and $B = 10$, the simulation results with varied n are:

n	50	100	500	1000
t^*_n	0.5	0.5	0.5	0.5
$\widetilde{M}_n(t^*_n)$	20.885	20.783	20.326	20.529

From the simulation above, we can draw the following conclusions:

(1) The estimates of jump positions and jump magnitudes derived from $\widetilde{M}_n(t)$ are also consistent. $\widetilde{M}_n(t)$ and $M_n(t)$ are derived from two different forms of kernel estimate of regression function.

(2) The first kernel used in simulation does not satisfy Lip(1) condition. But the results are also satisfactory. So the Lip(1) condition of kernels can probably be cancelled.

(3) It is unnecessary for B to be very large. From theoretical analysis, $B \geqslant 1/h_n$ is enough. The simulation results are coincided with it.

References

[1] Qiu, P.: *Estimation of a kind of jump regression functions*, Chinese Journal of System Science and Mathematics, (To appear), (1990).

[2] Cheng, X. R.: *Hypothesis testing and interval estimation of the model which has only one changepoint*, Scientia Sinica **8A**, (1988), 1–13.

[3] Cheng, K. F. and Lin, P. E.: *Nonparametric estimation of a regression*, Z. Wahrsch, Verw. Gebtete **57**, (1981), 223–233.

Received October 4, 1990