Предельные теоремы для сумм регрессионных остатков при множественном упорядочении регрессоров

Артем Павлович Ковалевский, Михаил Георгиевич Чебунин

НГТУ, НГУ

artyom.kovalevskii@gmail.com

Вторая конференция Математических центров России

МИАН, 11.11.2022

Let $(\mathbf{X}_i, \mathbf{Y}_i)$, $i = 1, 2 \ldots$, be the independent copies of a random vector $(\mathbf{X}, \mathbf{Y})$ such that
$\mathbf{X}_1 = (X_{1,1}, \ldots, X_{1,d_1})$ takes values in $[0,1]^{d_1}$,
$\mathbf{Y}_1$ takes values in $\mathbb{R}^{d_2}$.

The distribution function (copula) of $\mathbf{X}$ is
$C(\mathbf{u}) = \mathbf{P}(\mathbf{X}_1 \leq \mathbf{u}) = \mathbf{P}\left(X_{1,1} \leq u_1, \ldots, X_{1,d_1} \leq u_{d_1}\right)$,
$\mathbf{u} = (u_1, \ldots, u_{d_1}) \in [0,1]^{d_1}$.
We assume that there is copula density $c(\mathbf{u})$, that is,

$$C(\mathbf{u}) = \int_{\mathbf{v} \leq \mathbf{u}} c(\mathbf{v}) d\mathbf{v}, \quad \mathbf{u} \in [0,1]^{d_1}. \tag{1}$$

We study the asymptotic behavior of the random field

$$\mathbf{Q}_n(\mathbf{u}) = \sum_{j=1}^{n} \mathbf{Y}_j \mathbf{1}\left(\mathbf{X}_j \le \mathbf{u}\right), \quad \mathbf{u} \in [0,1]^{d_1}.$$

Let $\mathbf{m}(\mathbf{u}) = \mathbf{E}(\mathbf{Y} \mid \mathbf{X} = \mathbf{u})$, $\mathbf{u} \in [0,1]^{d_1}$, and
$\mathbf{f}(\mathbf{u}) = \int_{\mathbf{0}}^{\mathbf{u}} \mathbf{m}(\mathbf{v}) c(\mathbf{v}) d\mathbf{v}$.
Let

$$\sigma^2(\mathbf{u}) = \mathbf{E}\left\{(\mathbf{Y} - \mathbf{m}(\mathbf{X}))^T (\mathbf{Y} - \mathbf{m}(\mathbf{X})) \mid \mathbf{X} = \mathbf{u}\right\}$$

be the conditional covariance matrix of $\mathbf{Y}$,
$\sigma(\mathbf{u})$ be the positive definite matrix such that $\sigma(\mathbf{u})\sigma(\mathbf{u})^T = \sigma^2(\mathbf{u})$.

All our limit fields and processes are continuous, so we use the uniform metric. Let $\|\cdot\|$ denote the Euclidean norm in the corresponding space. Our random field $\mathbf{Q}_n$ takes values in the space $B([0,\,1]^{d_1};\,\mathbf{R}^{d_2})$ of bounded measurable functions with the Borel σ-algebra $\mathcal{B}$. This space is not separable, so we note that the random field $\mathbf{Q}_n$ takes values in its subset D with the smaller σ-algebra $\mathcal{D}$. This σ-algebra is generated by the d_1-dimensional analog of Skorohod metrics, see Straf (1972). So, let D be the uniform closure, in the space $B([0,\,1]^{d_1};\,\mathbf{R}^{d_2})$, of the vector subspace of simple functions (that is, linear combinations of step functions).

For $\mathbf{x}, \mathbf{y} \in D$, let the "Skorohod" distance be

$$d(\mathbf{x}, \mathbf{y}) = \inf\{\min(||\mathbf{x} - \mathbf{y}\lambda||_S, ||\lambda||_S) : \ \lambda \in \Lambda\},$$

where

$$||\mathbf{x} - \mathbf{y}\lambda||_S = \sup\{||\mathbf{x}(\mathbf{t}) - \mathbf{y}(\lambda(\mathbf{t}))||, \ \mathbf{t} \in [0,\ 1]^{d_1}\},$$

$||\lambda||_S = \sup\{||\lambda(\mathbf{t}) - \mathbf{t}||, \ \mathbf{t} \in [0,\ 1]^{d_1}\}$, Λ is the group of all transformations λ of the form $\lambda(t_1, \ldots, t_{d_1}) = (\lambda_1(t_1), \ldots, \lambda_{d_1}(t_{d_1}))$, where each $\lambda_i : [0,1] \to [0,1]$ is continuous, strictly increasing, and fixes 0 and 1.

With respect to the corresponding metric topology, D is complete and separable, and it's Borel σ-algebra $\mathcal{D}$ coincides with the σ-algebra generated by coordinate mappings, see Section 3 of Bickel $\&$ Wichura (1971) for details.
The process $\mathbf{Q}_n$ takes values in D. So, it is $\mathcal{D}$-measurable.
We define the weak convergence as follows (cf. Dudley, 1967): $\lim_{n\to\infty} \mathbf{E}f(\mathbf{Q}_n) = \mathbf{E}f(\mathbf{Q})$ for every bounded, continuous, and $\mathcal{D}$-measurable function $f : D \to \mathbf{R}^{d_2}$.

Davydov & Zitikis (2008) in their Proposition 1 have proved that the limitation of the class of functions f to the uniform continuity (instead of the continuity) gives the same definition of the weak convergence (that is, these definitions are equivalent).
We use the symbol $\Rightarrow$ to denote the weak convergence of random fields in the sense that has been mentioned above. We use the same symbol for the weak convergence of random variables and the weak convergence of stochastic processes in the uniform topology.

The following Lemma 1 generalizes the result of the first part of Theorem 2.1(1) by Davydov and Egorov (2000) to random fields.

Lemma 1

If $\mathbf{E}\|\mathbf{Y}\|^2 < \infty$ *then* $\widetilde{\mathbf{Q}}_n = \frac{\mathbf{Q}_n - n\mathbf{f}}{\sqrt{n}} \Rightarrow \mathbf{Q}$, *a centered Gaussian field with covariance*

$$K(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{E}\mathbf{Q}^T(\mathbf{u}_1)\mathbf{Q}(\mathbf{u}_2) = \int_{\mathbf{0}}^{\min(\mathbf{u}_1, \mathbf{u}_2)} \sigma^2(\mathbf{v})c(\mathbf{v})d\mathbf{v}$$

$$+ \int_{\mathbf{0}}^{\min(\mathbf{u}_1, \mathbf{u}_2)} \mathbf{m}^T(\mathbf{v})\mathbf{m}(\mathbf{v})c(\mathbf{v})d\mathbf{v}$$

$$- \int_{\mathbf{0}}^{\mathbf{u}_1} \mathbf{m}^T(\mathbf{v})c(\mathbf{v})d\mathbf{v} \int_{\mathbf{0}}^{\mathbf{u}_2} \mathbf{m}(\mathbf{v})c(\mathbf{v})d\mathbf{v},$$

$\mathbf{u}_1$, $\mathbf{u}_2 \in [0, 1]^{d_1}$.

Proof of Lemma 1.

For simplicity, we consider the case $d_1 = 2$ since the construction of the proof given below can be easily extended to the case $d_1 > 2$.
Now let $d_2 = 1$, we will generalize it to $d_2 \geq 1$ using the Cramer-Wold theorem.
Thus, we consider a random field

$$Q_n(\mathbf{u}) = \sum_{j=1}^{n} Y_j \mathbf{1}\left(X_j^{(1)} \leq u^{(1)}, X_j^{(2)} \leq u^{(2)}\right), \ \mathbf{u} \in [0,1]^2.$$

Let us define the partition of the unit square $[0,1]^2$ into N^2 parts as follows. Let $u_0^{(1)} = 0 < u_1^{(1)} < u_2^{(1)} < \ldots < u_N^{(1)} = 1$ be a partition of the interval [0,1], such that

$$\int_{(u_{i-1}^{(1)},0)}^{(u_i^{(1)},1)} \left(\mathbf{E}\left(Y^2 \mid \mathbf{X} = \mathbf{v}\right) + \mathbf{E}Y^2\right) c(\mathbf{v}) d\mathbf{v} = 2\mathbf{E}Y^2/N, \ i = 1, 2, \ldots, N.$$

For any fixed $i = 1, 2, \ldots, N$ let $u_{i,0}^{(2)} = 0 < u_{i,1}^{(2)} < u_{i,2}^{(2)} < \ldots < u_{i,N}^{(2)} = 1$ be another partition of the interval [0,1] (see Pic.1) such that

$$\int_{(u_{i-1}^{(1)}, u_{i,j-1}^{(2)})}^{(u_i^{(1)}, u_{i,j}^{(2)})} \left(\mathbf{E}\left(Y^2 \mid \mathbf{X} = \mathbf{v}\right) + \mathbf{E}Y^2\right) c(\mathbf{v}) d\mathbf{v} = 2\mathbf{E}Y^2/N^2,$$

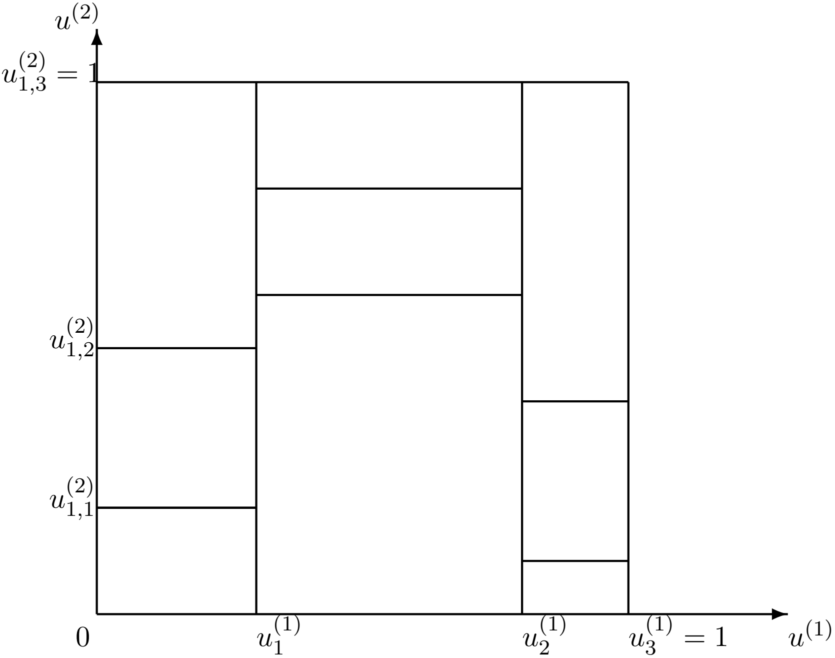
$j = 1, 2, \ldots, N.$

So we have $N+1$ points $u_i^{(1)}$ in the first coordinate and not greater then $(N-1)^2+1$ different points $u_{i,j}^{(2)}$ in the second coordinate. For any $\mathbf{u}=(u^{(1)},u^{(2)})\in[0,1]^2$, there are indexes $i^{(1)},i_1^{(2)},i_2^{(2)},j_1^{(2)},j_2^{(2)}\in\{0,1,\ldots,N\}$ such that $u_{i^{(1)}-1}^{(1)}\le u^{(1)}\le u_{i^{(1)}}^{(1)}$ and

$$u_{i_1^{(2)},j_1^{(2)}}^{(2)}=\max_{i,j}\{u_{i,j}^{(2)}\le u^{(2)}\},$$

$$u_{i_2^{(2)},j_2^{(2)}}^{(2)}=\min_{i,j}\{u_{i,j}^{(2)}\ge u^{(2)}\},$$

so $u_{i_1^{(2)},j_1^{(2)}}^{(2)}\le u^{(2)}\le u_{i_2^{(2)},j_2^{(2)}}^{(2)}$.
Denote $\mathbf{u}^l=(u_{i^{(1)}-1}^{(1)},u_{i_1^{(2)},j_1^{(2)}}^{(2)})$ and $\mathbf{u}^u=(u_{i^{(1)}}^{(1)},u_{i_2^{(2)},j_2^{(2)}}^{(2)})$ (Pic. 2).

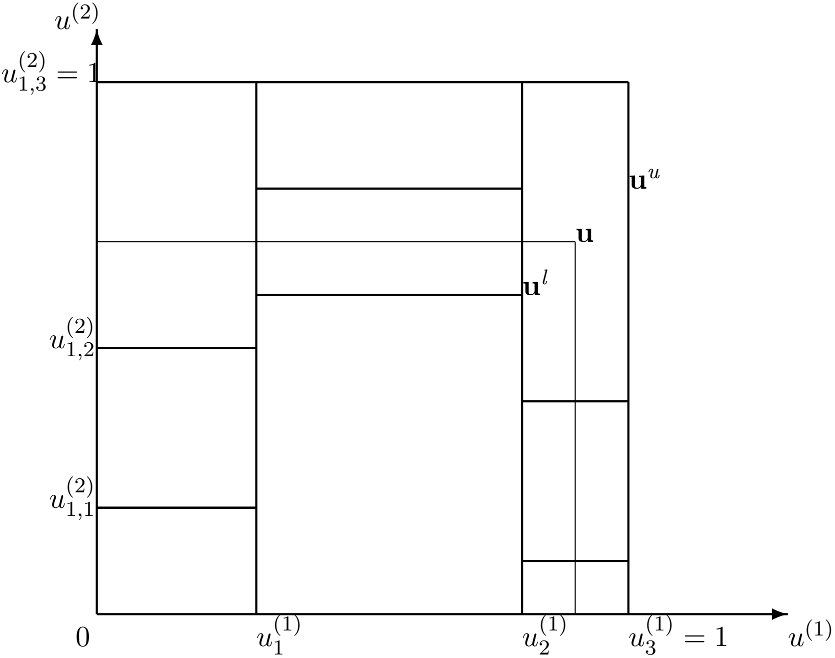

$u^{(2)}$
$u^{(2)}_{1,3} = 1$
$u^{(2)}_{1,2}$
$u^{(2)}_{1,1}$
0
$u^{(1)}_1$
$u^{(1)}_2$
$u^{(1)}_3 = 1$
$u^{(1)}$

Define the metric entropy with bracketing for the special separable pseudometric space (S, ρ), where

$$S = \left\{ h_{\mathbf{u}}(\mathbf{x}, y) = y\mathbf{1}(\mathbf{x} \leq \mathbf{u}), \quad \mathbf{u} \in [0,1]^2 \right\}$$
$$\rho^2 \left(h_{\mathbf{u_1}}, h_{\mathbf{u_2}} \right) = \mathbf{E} \left(h_{\mathbf{u_1}}(\mathbf{X}, Y) - h_{\mathbf{u_2}}(\mathbf{X}, Y) \right)^2$$

Let $S(\delta) = \{s_1, s_2, \ldots, s_M\} \subseteq S$ be such that for some random variables $f^l(s_i)$ and $f^u(s_i), i \leq M$, the following conditions are valid. For any $s \in S$ there exists $s_i \in S(\delta)$ such that

$$\rho(s, s_i) \leq \delta,$$
$$f^l(s_i) \leq f(s, \mathbf{X}, Y) \leq f^u(s_i) \quad a.s.,$$
$$\rho\left(f^u(s_i), f^l(s_i) \right) \leq \delta, \quad i \leq n.$$


$u^{(2)}$
$u^{(2)}_{1,3} = 1$
$\mathbf{u}^u$
$\mathbf{u}$
$\mathbf{u}^l$
$u^{(2)}_{1,2}$
$u^{(2)}_{1,1}$
0
$u^{(1)}_1$
$u^{(1)}_2$
$u^{(1)}_3 = 1$
$u^{(1)}$

Then $H^B(\delta, S, \rho) = \min\{M : S(\delta) \subseteq S\}$ is called the metric entropy with bracketing.
According to Ossiander's (1987) theorem, to prove Lemma for $\widetilde{Q}_n$ we must show that

$$\int_0^1 \left(H^B(t, S, \rho)\right)^{1/2} dt < \infty.$$

To prove it we show that

$$H^B(\delta, S, \rho) \leq C \log \left(\frac{1}{\delta}\right).$$

First notice that
$f\left(h_{\mathbf{u}}, \mathbf{X}, Y\right) = h_{\mathbf{u}}(\mathbf{X}, Y) - Eh_{\mathbf{u}}(\mathbf{X}, Y), \mathbf{u} \in [0,1]^2$ is a separable random process satisfying the conditions $(2.1) - (2.3)$ of Ossiander's work, moreover $\widetilde{Q}_n = n^{-1/2} \sum_{i=1}^{n} f\left(h_{\mathbf{u}}, \mathbf{X}_i, Y_i\right)$.

Then write $f(h_{\mathbf{u}}, \mathbf{x}, y)$ as follows

$$f(h_{\mathbf{u}}, \mathbf{x}, y) = y_+ \mathbf{1}(\mathbf{x} \leq \mathbf{u}) - y_- \mathbf{1}(\mathbf{x} \leq \mathbf{u}) - \int_{\mathbf{0}}^{\mathbf{u}} m_+(\mathbf{v}) c(\mathbf{v}) d\mathbf{v}$$

$$+ \int_{\mathbf{0}}^{\mathbf{u}} m_-(\mathbf{v}) c(\mathbf{v}) d\mathbf{v}.$$

Let

$$f^u = Y_+ \mathbf{1}(\mathbf{X} \leq \mathbf{u}^u) - Y_- \mathbf{1}(\mathbf{X} \leq \mathbf{u}^l)$$

$$- \int_{\mathbf{0}}^{\mathbf{u}^l} m_+(\mathbf{v}) c(\mathbf{v}) d\mathbf{v} + \int_{\mathbf{0}}^{\mathbf{u}^u} m_-(\mathbf{v}) c(\mathbf{v}) d\mathbf{v},$$

$$f^l = Y_+ \mathbf{1}(\mathbf{X} \leq \mathbf{u}^l) - Y_- \mathbf{1}(\mathbf{X} \leq \mathbf{u}^u)$$

$$- \int_{\mathbf{0}}^{\mathbf{u}^u} m_+(\mathbf{v}) c(\mathbf{v}) d\mathbf{v} + \int_{\mathbf{0}}^{\mathbf{u}^l} m_-(\mathbf{v}) c(\mathbf{v}) d\mathbf{v}.$$

Then

$$f^u - f^l = |Y|\mathbf{1}(\mathbf{u}^l \le \mathbf{X} \le \mathbf{u}^u) + \int_{\mathbf{u}^l}^{\mathbf{u}^u} |m(\mathbf{v})|c(\mathbf{v})d\mathbf{v}.$$

We use the Cauchy-Bunyakovsky and Jensen inequalities, as well as simple algebraic inequalities, and obtain that (see Davydov and Egorov (2000) for details)

$$\left(E\left(f^u - f^l\right)^2\right)^{1/2} \le \left(\int_{\mathbf{u}^l}^{\mathbf{u}^u} E\left(Y^2 \mid \mathbf{X} = \mathbf{v}\right) c(\mathbf{v})d\mathbf{v}\right)^{1/2}$$

$$+ \left(EY^2 \int_{\mathbf{u}^l}^{\mathbf{u}^u} c(\mathbf{v})d\mathbf{v}\right)^{1/2} \le 2\left(\int_{\mathbf{u}^l}^{\mathbf{u}^u} \left(E\left(Y^2 \mid \mathbf{X} = \mathbf{v}\right) + EY^2\right) c(\mathbf{v})d\mathbf{v}\right)$$

$$\le 2\sqrt{2N \cdot 2EY^2/N^2} = 4\sqrt{EY^2/N} = \delta.$$

In the last inequality, we used the fact that the region of integration is included in $2N$ rectangles from the constructed partition of the unit square.

Summing up these equalities we get

$$M < (N+1)^3 \leq \left(\left[\frac{16EY^2}{\delta^2} \right] + 2 \right)^3.$$

Hence by Ossiander's theorem

$$\widetilde{Q}_n \Rightarrow Q$$

where Q is the Gaussian field,
$EQ(\mathbf{u}) = 0, \mathbf{E}(Q(\mathbf{u}_1), Q(\mathbf{u}_2)) = K(\mathbf{u}_1, \mathbf{u}_2)$. Elementary calculations show that

$$K(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{E}(Y^2 \mathbf{1}(\mathbf{X} \leq \mathbf{u}_1, \mathbf{X} \leq \mathbf{u}_2))$$

$$- \int_{\mathbf{0}}^{\mathbf{u}_1} m(\mathbf{v}) c(\mathbf{v}) d\mathbf{v} \int_{\mathbf{0}}^{\mathbf{u}_2} m(\mathbf{v}) c(\mathbf{v}) d\mathbf{v}.$$

This construction of the proof can be easily extended to the case $d_1 > 2$ by splitting the corresponding integral into pieces of size $2\mathbf{E}Y^2/N^{d_1}$. So we use Theorem 3.1 of Ossiander (1987) and get the weak convergence $Q_n \Rightarrow Q$ for $d_2 = 1$.
From the tightness of Q_n for $d_2 = 1$ we get the tightness of $\mathbf{Q}_n$ because the tightness is coordinatewise. According to the Cramer-Wold theorem since $\mathbf{Q}_n(\mathbf{u})$ is linear, we obtain the convergence of finite-dimensional distributions for any $d_2 \geq 1$.
The proof is complete.

Denote $X_{n,1}^{(k)} \leq X_{n,2}^{(k)} \leq \cdots \leq X_{n,n}^{(k)}$, $1 \leq k \leq d_1$, the order statistics of the k-th column of matrix X, and $\mathbf{Y}_{n,1}^{(k)}, \mathbf{Y}_{n,2}^{(k)}, \ldots, \mathbf{Y}_{n,n}^{(k)}$ the corresponding values of the vectors $\mathbf{Y}_i$. The random vectors $\left(\mathbf{Y}_{n,i}^{(k)}, i \leq n\right)$ are called induced order statistics (concomitants). Using the asymptotics of $\mathbf{Q}_n(\mathbf{u})$, we study the asymptotics of $d_1 \times d_2$-dimensional process of sums of induced order statistics under different orderings

$$\mathbf{Z}_n(t) = \left(\sum_{j=1}^{[nt]} \mathbf{Y}_{n,j}^{(1)}, \sum_{j=1}^{[nt]} \mathbf{Y}_{n,j}^{(2)}, \ldots, \sum_{j=1}^{[nt]} \mathbf{Y}_{n,j}^{(d_1)} \right), \ t \in [0,1].$$

Let $\mathbf{e}_{k,t} = (1, \ldots, 1, t, 1, \ldots, 1)$ the vector in $[0,1]^{d_1}$ with k-th coordinate being t and other coordinates being 1.

Lemma 2 generalizes the result of Theorem 2.1(2) by Davydov and Egorov (2000) to multiple ordering but under additional assumption $\mathbf{m} \equiv \mathbf{0}$.

Lemma 2

If $\mathbf{E}\|\mathbf{Y}\|^2 < \infty$, $\mathbf{m} \equiv \mathbf{0}$ *then* $\widetilde{\mathbf{Z}}_n = \frac{\mathbf{Z}_n}{\sqrt{n}} \Rightarrow \mathbf{Z}$, *a centered Gaussian* $(d_1 \times d_2)$*-dimensional process with covariance matrix function* $\mathbf{E}\mathbf{Z}^T(t_1)\mathbf{Z}(t_2) = (K(\mathbf{e}_{k_1,t_1}, \mathbf{e}_{k_2,t_2}))_{k_1,k_2=1}^{d_1}$,

$$K(\mathbf{e}_{k_1,t_1}, \mathbf{e}_{k_2,t_2}) = \mathbf{E}\mathbf{Q}^T(\mathbf{e}_{k_1,t_1})\mathbf{Q}(\mathbf{e}_{k_2,t_2})$$

$$= \int_{\mathbf{0}}^{\min(\mathbf{e}_{k_1,t_1}, \mathbf{e}_{k_2,t_2})} \sigma^2(\mathbf{v}) c(\mathbf{v}) d\mathbf{v}.$$

Results for linear regression

Let $(\mathbf{X}_i, \xi_i, \eta_i) = (X_{i1}, \ldots, X_{id_1}, \xi_{i1}, \ldots, \xi_{i,d_2-1}, \eta_i)$ be independent and identically distributed random vector rows, $i = 1, \ldots, n$. All components of a raw can be dependent and $X_{i1}, \ldots, X_{id_1}$ have copula (so their marginal distributions are uniform on [0, 1]) and (1) is true.

Rows $(\mathbf{X}_i, \xi_i, \eta_i)$ form matrix (X, ξ, η).

We assume a linear regression hypothesis H_0:

$$\eta_i = \xi_i \theta + \varepsilon_i = \sum_{j=1}^{d_2-1} \xi_{ij}\theta_j + \varepsilon_i, \qquad (2)$$

$\{\varepsilon_i\}_{i=1}^n$ and $\{(\mathbf{X}_i, \xi_i)\}_{i=1}^n$ are independent, $\mathbf{E}\,\varepsilon_1 = 0$, $\mathbf{Var}\,\varepsilon_1 > 0$.

Vector $\theta = (\theta_1, \ldots, \theta_{d_2-1})^T$ and constant $\mathbf{Var}\,\varepsilon_1$ are unknown.

We consider d_1 orderings of rows of the matrix (X, ξ, η) in acsending order of columns of X.
The result of d_1 orderings is a sequence of d_1 matrices $(X^{(j)}, \xi^{(j)}, \eta^{(j)})$ with rows $(\mathbf{X}_i^{(j)}, \xi_i^{(j)}, \eta_i^{(j)}) = (X_{i1}^{(j)}, \ldots, X_{id_1}^{(j)}, \xi_{i1}^{(j)}, \ldots, \xi_{i,d_2-1}^{(j)}, \eta_i^{(j)})$, $j = 1, \ldots, d_1$.

Let $\widehat{\theta}$ be LSE:

$$\widehat{\theta} = (\xi^T \xi)^{-1} \xi^T \eta.$$

It does not depend on the order of rows.

Let $h^{(j)}(x) = \mathbf{E}\{\xi_1 | X_{1j} = x\}$ be conditional expectations, $L^{(j)}(x) = \int\limits_0^x h^{(j)}(s)\,ds$ be *induced theoretical generalised Lorentz curves* (see Davydov and Egorov (2000)),

$$b_j^2(x) = \mathbf{E}\left((\xi_1 - h^{(j)}(x))^T(\xi_1 - h^{(j)}(x)) \mid X_{1j} = x\right)$$

be matrices of conditional covariances.
Let $G = \mathbf{E}\xi_1^T\xi_1$. Then

$$\int_0^1 \left(b_j^2(x) + (h^{(j)}(x))^T h^{(j)}(x)\right)\, dx = G$$

for any $j = 1, \dots, d_1$.
Let $\widehat{\varepsilon}_i^{(j)} = \eta_i^{(j)} - \xi_i^{(j)}\widehat{\theta}$ be regression residuals, $\widehat{\Delta}_k^{(j)} = \sum\limits_{i=1}^k \widehat{\varepsilon}_i^{(j)}$ be its partial sums, $\widehat{\Delta}_0^{(j)} = 0$.

Let $\widehat{Z}_n^{(j)} = \{\widehat{Z}_n^{(j)}(t),\ 0 \le t \le 1\}$ be a piecewise linear random function with nodes

$$\left(\frac{k}{n},\ \frac{\widehat{\Delta}_k^{(j)}}{\sqrt{n\mathbf{Var}\varepsilon_1}}\right),\quad k = 0,\ 1, \ldots,\ n.$$

From Theorem 1 (Kovalevskii, 2020) we have

Theorem 3

If matrix G exists and is non-degenerate and H_0 is true then $\widehat{Z}_n^{(j)} \Longrightarrow \widehat{Z}^{(j)}$ for any $j = 1, \ldots, d_1$. Here $\widehat{Z}^{(j)}$ is a centered Gaussian process with continuous a.s. sample paths and covariance function

$$\widehat{K}_{jj}(s,t) = \min(s,t) - L^{(j)}(s)G^{-1}(L^{(j)}(t))^T,\quad s,t \in [0,1].$$

We prove that the d_1-dimensional process $\widehat{Z}_n = (\widehat{Z}_n^{(j)}, j = 1, \ldots, d_1)$ has a Gaussian limit.

Theorem 4

If matrix G exists and is non-degenerate and H_0 is true then $\widehat{Z}_n \Longrightarrow \widehat{Z}$. Here $\widehat{Z}$ is a centered d_1-dimensional Gaussian process with continuous a.s. sample paths and covariance matrix function $\widehat{K}(s,t) = \left(\widehat{K}_{ij}(s,t)\right)_{i,j=1}^{d_1}$,

$$\widehat{K}_{ij}(s,t) = \mathbf{P}(X_{1i} \le s, X_{1j} \le t) - L^{(i)}(s) G^{-1} (L^{(j)}(t))^T, \;\; s,t \in [0,1].$$

We now describe the application of this result to testing the hypothesis of linear dependence. Let $d_2 - 1 = d_1$. Let $(\xi_i, \eta_i) = (\xi_{i1}, \ldots, \xi_{i,d_1}, \eta_i)$ be independent and identically distributed random vector rows, $i = 1, \ldots, n$. In addition, we assume that the column ξ_{i,d_1} consists of ones:

$$\xi_{i,d_1} \equiv 1. \tag{3}$$

We want to test the linear dependence (2). To do this we estimate the parameters θ and $\mathbf{Var}\,\varepsilon_1$, sort the data in ascending order of each of the first $d_1 - 1$ columns of the regressor and calculate processes of the sums of regression residuals. We use the quantile functions $F^{-1}_{\xi_{1j}}$ to apply Theorem 2. So we assume that $\xi_{ij} = F^{-1}_{\xi_{1j}}(X_{ij})$, $i = 1, \ldots, n$, $j = 1, \ldots, d_1 - 1$. If the matrix G exists and is non-degenerate then under the true hypothesis H_0 we are in the conditions of Theorem 2.

From (3) we have $\widehat{Z}_n(1) = \mathbf{0}$. So we can use a statistics of omega squared type and calculate its limiting distribution by lines of Chakrabarty et al. (2020):

$$\omega_n^2 = \sum_{j=1}^{d_1-1} \int_0^1 \left(\widehat{Z}_n^{(j)}(t)\right)^2 dt \Rightarrow \omega^2 = \sum_{j=1}^{d_1-1} \int_0^1 \left(\widehat{Z}^{(j)}(t)\right)^2 dt.$$

We estimate the covariance function in Theorem 2 from empirical data. An estimate for $\mathbf{P}(X_{1i} \le s,\ X_{1j} \le t)$ is

$$\frac{1}{n} \sum_{k=1}^{n} \mathbf{1}\{\xi_{ki} \le \xi^{(i)}_{[ns],i},\ \xi_{kj} \le \xi^{(j)}_{[nt],j}\}.$$

It converges to the probability uniformely on (s,t) in $[0,1]^2$. We estimate functions $L^{(j)}$ and matrix G by their empirical counterparts.

Results for compound Poisson fields

Let $A \subset \mathbb{R}^{d_1}$ be a compact Borel set with Lebesque measure $0 < \mu(A) < \infty$, and there is a compound Poisson field in A with intensity $\nu > 0$. That is, the number $\eta = \eta_\nu$ of points $\mathbf{X}_i$ in A has Poisson distribution with parameter $\nu\mu(A)$, points are distributed unformly in A under the condition of fixed N.
Any point $\mathbf{X}_i$ is associated with mark $\mathbf{Y}_i$ that takes values in $\mathbf{R}^{d_2}$, $0 < d_2 < \infty$. Vectors $(\mathbf{X}_i, \mathbf{Y}_i)$ are mutually independent and identically distributed copies of a random vector $(\mathbf{X}, \mathbf{Y})$ such that $\mathbf{X} = (X^{(1)}, \dots, X^{(d_1)})$ takes values in A, $\mathbf{Y}$ takes values in $\mathbb{R}^{d_2}$.
Matrix (X, Y) has random number η of rows $(\mathbf{X}_i, \mathbf{Y}_i)$.

We study the asymptotic behavior of the random field

$$\mathbf{Q}_\nu(\mathbf{u}) = \sum_{j=1}^{\eta_\nu} \mathbf{Y}_j \mathbf{1}\left(\mathbf{X}_j \le \mathbf{u}\right)$$

$$= \sum_{j=1}^{\eta_\nu} \mathbf{Y}_j \mathbf{1}\left(X_j^{(1)} \le u^{(1)}, \ldots, X_j^{(d_1)} \le u^{(d_1)}\right),\ \mathbf{u} \in \mathbf{R}^{d_1},$$

as $\nu \to \infty$.
Using the asymptotics of $\mathbf{Q}_\nu(\mathbf{u})$, we study the asymptotics of $d_1 \times d_2$-dimensional process of sums of induced order statistics under different orderings

$$\mathbf{Z}_\nu(t) = \left(\sum_{j=1}^{[\eta_\nu t]} \mathbf{Y}^{(1)}_{\eta_\nu, j}, \sum_{j=1}^{[\eta_\nu t]} \mathbf{Y}^{(2)}_{\eta_\nu, j}, \ldots, \sum_{j=1}^{[\eta_\nu t]} \mathbf{Y}^{(d_1)}_{\eta_\nu, j}\right),\ t \in [0, 1].$$

Let

$$\mathbf{m}(\mathbf{u}) = \mathbf{E}(\mathbf{Y} \mid \mathbf{X} = \mathbf{u})\mathbf{1}\{\mathbf{u} \in A\},$$

$$\mathbf{f}(\mathbf{u}) = \int_{-\infty}^{\mathbf{u}} \mathbf{m}(\mathbf{v})\mathbf{1}(\mathbf{v} \in A)d\mathbf{v},$$

$$\sigma^2(\mathbf{u}) = \mathbf{E}\left\{(\mathbf{Y} - \mathbf{m}(\mathbf{X}))^T(\mathbf{Y} - \mathbf{m}(\mathbf{X})) \mid \mathbf{X} = \mathbf{u}\right\}.$$

Theorem 5

If $\mathbf{E}\|\mathbf{Y}\|^2 < \infty$ *then* $\widetilde{\mathbf{Q}}_\nu = \frac{\mathbf{Q}_\nu - \eta_\nu \mathbf{f}}{\sqrt{\eta_\nu}} \Rightarrow \mathbf{Q}$*, a centered Gaussian field with covariance*

$$K(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{E}\mathbf{Q}^T(\mathbf{u}_1)\mathbf{Q}(\mathbf{u}_2) = \int\limits_{A_{\mathbf{u}_1,\mathbf{u}_2}} \sigma^2(\mathbf{v})d\mathbf{v}$$

$$+ \int\limits_{A_{\mathbf{u}_1,\mathbf{u}_2}} \mathbf{m}^T(\mathbf{v})\mathbf{m}(\mathbf{v})d\mathbf{v} - \int\limits_{A_{\mathbf{u}_1,\mathbf{u}_2}} \mathbf{m}^T(\mathbf{v})d\mathbf{v} \int\limits_{A_{\mathbf{u}_1,\mathbf{u}_2}} \mathbf{m}(\mathbf{v})d\mathbf{v},$$

$A_{\mathbf{u}_1,\mathbf{u}_2} = \{\mathbf{v} \in \mathbf{R}^{d_1} : \ \mathbf{v} \in A, \mathbf{v} \le \mathbf{u}_1, \mathbf{v} \le \mathbf{u}_2\}$.

Theorem 6

If $\mathbf{E}\|\mathbf{Y}\|^2 < \infty$, $\mathbf{m} \equiv \mathbf{0}$ *then* $\widetilde{\mathbf{Z}}_\nu = \frac{\mathbf{Z}_\nu}{\sqrt{\eta_\nu}} \Rightarrow \mathbf{Z}$, *a centered Gaussian* $(d_1 \times d_2)$*-dimensional process with covariance matrix function* $\mathbf{E}\mathbf{Z}^T(t_1)\mathbf{Z}(t_2) = \widetilde{K}(t_1, t_2) = (\widetilde{K}_{ij}(t_1, t_2)^{d_1}_{i,j=1}$,

$$\widetilde{K}_{ij}(t_1, t_2) = \int\limits_{B_{i,t_1,j,t_2}} \sigma^2(\mathbf{v}) d\mathbf{v},$$

$B_{i,t_1,j,t_2} = \{\mathbf{v} \in \mathbf{R}^{d_1} : \ \mathbf{v} \in A, v_i \le t_1, v_j \le t_2\}$.

Multiple ordering of non-homogeneous Poisson process
We interpret the result of the previous section for non-homogeneous Poisson processes. So we take set A of special form,

$$A = \{\mathbf{x} = (x_1, x_2) : \ 0 \le x_1 \le T, 0 \le x_2 \le \lambda(x_1)\},$$

$T > 0$, $\lambda = \{\lambda(t), 0 \le t \le T\} > 0$ is the positive Borel function on $[0, T]$.
We inperpret x_1 as time, $\lambda(x_1)$ as the Poisson parameter at time x_1.
So $(\mathbf{X}_i, \mathbf{Y}_i) = ((X_{i1}, X_{i2}), \mathbf{Y}_i)$.
We suppose that $\mathbf{m}(x_1, x_2)$ does not depend on x_2: Let

$$\mathbf{m}(x_1, x_2) = \mathbf{m}(x_1) = \mathbf{E}(\mathbf{Y}_1 | X_{11} = x_1).$$

We order vectors $(\mathbf{X}_1, \mathbf{Y}_1), \ldots, (\mathbf{X}_\eta, \mathbf{Y}_\eta)$, $\eta = \eta_\nu$, using two different algorythms.
The first one: $(\mathbf{X}_i^*, \mathbf{Y}_i^*) = (X_{i1}^*, X_{i2}^*, \mathbf{Y}_i^*)$ are ordered by time, that is,

$$X_{11}^* < X_{21}^* < \ldots < X_{\eta 1}^*.$$

The second one: $(\mathbf{X}_i^@, \mathbf{Y}_i^@) = (X_{i1}^@, X_{i2}^@, \mathbf{Y}_i^@)$ are ordered by the corresponding intensity, that is,

$$\lambda(X_{11}^@) \le \lambda(X_{21}^@) \le \ldots \le \lambda(X_{\eta 1}^@).$$

If $\lambda(X_{i1}^@) = \lambda(X_{j1}^@)$ then the order of corresponding pairs is random.
So if $\lambda \equiv const$ then all the pairs $(\mathbf{X}_i^@, \mathbf{Y}_i^@)$ are in a random order, this case corresponds to the homogeneous Poisson process.

Theorem 7

If $0 < \sigma^2(\mathbf{x}) < \infty$ *for any* $\mathbf{x} \in A$ *then*

$$\left\{ \eta_\nu^{-1/2} \sum_{i=1}^{[\eta_\nu t]} \Big(\mathbf{Y}_i^* - \mathbf{m}(X_{i1}^*), \mathbf{Y}_i^@ - \mathbf{m}(X_{i1}^@) \Big),\ 0 \le t \le 1, \right\}$$

converges weakly in uniform metrics as $\nu \to \infty$ *to a centered Gaussian* $2d_2$*-dimencional process* $\{(\mathbf{V}_1(t), \mathbf{V}_2(t)), 0 \le t \le 1\}$ *with covariances*

$$\mathbf{E}\mathbf{V}_i^T(t_1)\mathbf{V}_j(t_2) = \int_{G_{ij}(t_1,t_2)} \sigma^2(\mathbf{x})\, d\mathbf{x}, \quad i, j = 1, 2,$$

$$G_{11}(t_1,t_2) = \{(x_1,x_2) :\ 0 \le x_1 \le \min(t_1,t_2), 0 \le x_2 \le \lambda(x_1)\},$$
$$G_{12}(t_1,t_2) = \{0 \le x_1 \le t_1, \lambda(x_1) \le \lambda(t_2), 0 \le x_2 \le \lambda(x_1)\},$$
$$G_{22}(t_1,t_2) = \{\lambda(x_1) \le \min(\lambda(t_1), \lambda(t_2)), 0 \le x_2 \le \lambda(x_1)\}.$$

Bickel, P. J., and Wichura, M. J. , 1971. *Convergence Criteria for Multiparameter Stochastic Processes and Some Applications*, Ann. Math. Stat.**42** (5), 1656–70. MR0383482

Bischoff, W., 1998. *A functional central limit theorem for regression models*, Ann. Stat. **26**, 1398–1410. MR1647677

Brodsky, B., Darkhovsky, B., 2005. *Asymptotically optimal methods of change-point detection for composite hypotheses*, Journal of Statistical Planning and Inference **133** (1), 123–138. MR2162571

Chakrabarty, A., Chebunin, M., Kovalevskii, A. et al., 2020. *A statistical test for correspondence of texts to the Zipf - Mandelbrot law*, Siberian Electronic Mathematical Reports **17**, 1959–1974. MR4239147

Chebunin, M. G., Kovalevskii, A. P., 2021. *Asymptotics of sums of regression residuals under multiple ordering of regressors*, Siberian Electronic Mathematical Reports **18**, No. 2, 1482–1492. DOI 10.33048/semi.2021.18.111.

Csorgo, M., Horváth, L., 1997. *Limit Theorems in Change-Point Analysis*, NY: Wiley. MR2743035

Davydov, Y., Egorov, V., 2000. *Functional limit theorems for induced order statistics*, Mathematical Methods of Statistics **9** (3), 297–313. MR1807096

Davydov, Y., Zitikis, R., 2008. *On weak convergence of random fields*, Annals of the Institute of Statistical Mathematics **60**, 345–365. MR2403523

Dudley, R. M., 1967. *Measures on non-separable metric spaces*, Illinois Journal of Mathematics **11**, 449–453. MR0235087

Kovalevskii, A. P., Shatalin, E. V., 2015. *Asymptotics of Sums of Residuals of One-Parameter Linear Regression on Order Statistics*, Theory of Probability and its Applications **59** (3), 375–387. MR3415974

Kovalevskii, A., Shatalin, E., 2016. *A limit process for a sequence of partial sums of residuals of a simple regression on order statistics*, Probability and Mathematical Statistics **36** (1), 113–120. MR3529343

Kovalevskii, A. P., 2020. *Asymptotics of an empirical bridge of regression on induced order statistics*, Siberian Electronic Mathematical Reports **17**, 954–963. MR4217195

Lorden, G., 1971. *Procedures for Reacting to a Change in Distribution*, Ann. Math. Statist. **42** (6), 1897–1908. MR0309251

MacNeill, I. B., 1978. *Limit processes for sequences of partial sums of regression residuals*, Ann. Prob. **6**, 695–698. MR0494708

MacNeill, I.B., Jandhyala, V.K., Kaul, A., Fotopoulos, S.B., 2020. *Multiple change-point models for time series*, Environmetrics **31** (1), e2593. MR4061127

Ossiander, M., 1987. *A Central limit theorem under metric entropy with L_2 bracketing*, Ann. Prob. **15**, 897–919. MR0893905

Page, E. S., 1954. *Continuous inspection schemes*, Biometrika **41** (1–2), 100–115. MR0088850

Rao, C., 1950. *Sequential Tests of Null Hypotheses*, Sankhya **10** (4), 361–370. MR0039196

Shiryaev, A. N., 1996. *Minimax Optimality of the Method of Cumulative Sum (Cusum) in the Case of Continuous Time*, Russian Mathematical Surveys **51**, 750–751. MR1422244

Shorack, G., Wellner, J., 1986. *Empirical processes with applications to statistics*, Wiley N.-Y. MR3396731

Straf, M.L., 1972. *Weak convergence of stochastic processes with several parameters*, Proc. Sixth Berkely Symp. Math. Statist. Prob. **2**, 187–222. MR0402847