УДК 519.234.3

В. Б. Горяинов

ЛОКАЛЬНО НАИБОЛЕЕ МОЩНЫЕ РАНГОВЫЕ КРИТЕРИИ НЕЗАВИСИМОСТИ НАБЛЮДЕНИЙ В МОДЕЛИ ПРОСТРАНСТВЕННОЙ АВТОРЕГРЕССИИ

Для процесса пространственной авторегрессии порядка (1,1) построены локально наиболее мощные ранговые критерии для проверки независимости наблюдений. Показано, что при нулевой гипотезе статистики предложенных критериев являются свободными от распределения и асимптотически нормальными.

E-mail: b-goryiainov@mail.ru

Ключевые слова: пространственная авторегрессия, ранговые методы, локально наиболее мощные критерии.

Введение. В работе проведены исследования случайного авторегрессионного поля X_{ij}, описываемого уравнением

$$X_{ij} = \theta_{10}X_{i-1,j} + \theta_{01}X_{i,j-1} + \theta_{11}X_{i-1,j-1} + \varepsilon_{ij},\ i,j = 0,1,2,\ldots, \quad (1)$$
$$X_{ij} = 0 \text{ для любых } i < 0 \text{ или } j < 0,$$

где $\theta = (\theta_{10}, \theta_{01}, \theta_{11})$ — вектор авторегрессионных коэффициентов, а ε_{ij} — независимые одинаково распределенные случайные величины с нулевым математическим ожиданием $\mathsf{E}\varepsilon_{ij} = 0$.

Такие поля широко используются в технике, экономике и естественных науках [1]. До недавнего времени существовавшие методы проверки гипотез в модели (1) основывались на принципе максимального правдоподобия в предположении нормальности наблюдений [2, 3]. Как правило, такие методы чувствительны к засорению выборки резко выделяющимися наблюдениями.

В работах [4, 5] были построены локально наиболее мощные (ЛНМ) критерии проверки гипотез о коэффициентах θ, основанные только на знаках наблюдений.

В настоящей работе получены ранговые критерии проверки гипотезы H^0 о независимости наблюдений в модели (1). Предполагается, что распределение ε_{ij} принадлежит известному параметрическому семейству. Показано, что при H^0 распределение статистик построенных ранговых критериев не зависит от распределения ε_{ij} и асимптотически нормально.

Постановка задачи. Рассмотрим поле (1), где ε_{ij} — независимые одинаково распределенные случайные величины с функцией распределения $F(x)$ и плотностью $f(x)$, $\theta = (\theta_{10}, \theta_{01}, \theta_{11})$ — неизвестный вектор параметров.

Пусть $a = (a_{10}, a_{01}, a_{11})$ — некоторый известный вектор. Рассмотрим задачу проверки гипотезы

$$H^0 : \theta = 0$$

против односторонних альтернатив вида

$$H^+{}_a : \quad \theta = \Delta a, \ \Delta > 0,$$

$$H_a^- : \quad \theta = \Delta a, \ \Delta < 0,$$

и двусторонней альтернативы

$$H_a : \quad \theta = \Delta a, \ \Delta \neq 0.$$

Пусть $X = \{X_{ij}\}, i = 1, \ldots, m, j = 1, \ldots, n$ — матрица наблюдений поля (1).

Обозначим R_{ij} ранг X_{ij} в последовательности

$$X_{11}, \ldots, X_{m1}, \ldots, X_{1n}, \ldots, X_{mn}.$$

Отметим, что матрица рангов наблюдений $R = \{R_{ij}\}$ принадлежит множеству $\mathcal{M}$ матриц размера $m \times n$, элементы которых являются перестановками множества $\{1, 2, \ldots, mn\}$. На основе информации только о матрице R требуется построить оптимальные критерии проверки гипотез о параметре θ. Оптимальность критериев будем понимать в следующем смысле.

Обозначим Q критическую область рангового критерия, т. е. такое подмножество в $\mathcal{M}$, что если матрица R принадлежит Q, то гипотеза H^0 отклоняется. Через $P_{mn}(Q, \Delta)$ обозначим функцию мощности рангового критерия, определяемую как вероятность отклонения гипотезы H^0, когда H^0 не верна:

$$P_{mn}(Q, \Delta) = \mathsf{P}\{R \in Q | \text{ верна альтернатива } \theta = \Delta a\}.$$

Пусть $P_{mn}(Q, \Delta)$ дифференцируема в точке 0 по Δ. Определим ЛНМ ранговый критерий для проверки гипотезы H^0 против односторонней альтернативы $H^+{}_a$ как критерий, имеющий функцию мощности $P_{mn}(Q, \Delta)$, наиболее круто возрастающую по переменной Δ в правосторонней окрестности точки 0. Это означает, что критическая область Q ЛНМ рангового критерия должна быть выбрана так, чтобы величина $\dfrac{dP_{mn}(Q, \Delta)}{d\Delta}$ при $\Delta = 0$ была максимальна. Совершенно аналогично определим ЛНМ ранговый критерий для проверки гипотезы H^0 против односторонней альтернативы H_a^- как критерий, имеющий минимальное значение $\dfrac{dP_{mn}(Q, \Delta)}{d\Delta}$ при $\Delta = 0$.

Локально наиболее мощные критерии. Начнем с построения ЛНМ рангового критерия для проверки гипотезы H^0 против альтернативы H_a^+. Так как

$$P_{mn}(Q, \Delta) = \sum_{r \in Q} P_{mn}(r, \Delta),$$

где $P_{mn}(r, \Delta) = P\{R = r|$ верна альтернатива $\theta = \Delta a\}$, то величина $\left.\dfrac{dP_{mn}(Q, \Delta)}{d\Delta}\right|_{\Delta=0}$ будет наибольшей, если в критическую область Q последовательно, вплоть до достижения заданного уровня значимости, включаются матрицы r, имеющие наибольшие значения $\dfrac{dP_{mn}(r, \Delta)}{d\Delta}$ в точке $\Delta = 0$. Поэтому искомая критическая область Q будет равна

$$Q = \left\{ r : \left.\frac{dP_{mn}(Q, r)}{d\Delta}\right|_{\Delta=0} > C \right\},$$

где постоянная C определяется уровнем значимости α критерия, т.е. находится из условия $\mathsf{P}\{R \in Q\} = \alpha$ при гипотезе H^0.

Для построения ЛНМ рангового критерия нужно знать поведение функции мощности, а значит, и $P_{mn}(r, \Delta)$ в окрестности $\Delta = 0$.

Для плотности $f(x)$, удовлетворяющей введенным ниже условиям (8), (9), определим функцию меток

$$\varphi(x) = -\frac{f'(x)}{f(x)} \tag{2}$$

и сами метки

$$a_{mn}(i, j) = \mathsf{E}[\varphi(\varepsilon^{(i)})\varepsilon^{(j)}] = \mathsf{E}[\varphi(F^{-1}(U^{(i)}))F^{-1}(U^{(j)})], \quad i, j = 1, \ldots, mn, \tag{3}$$

где $\varepsilon^{(1)}, \ldots, \varepsilon^{(mn)}$ и $U^{(1)}, \ldots, U^{(mn)}$ — элементы вариационного ряда из распределения с плотностью $f(x)$ и равномерного распределения на $[0, 1]$ соответственно;

$$F^{-1}(u) = \inf\{x : F(x) \geq u\}. \tag{4}$$

На множестве матриц $\mathcal{M}$ и множестве индексов

$$\mathcal{I} = \{(1, 0), (0, 1), (1, 1)\}$$

определим статистики

$$z_{pq}(r) = \sum_{k=p+1}^{m} \sum_{l=q+1}^{n} a_{mn}(r_{kl}, r_{k-p,l-q}), \qquad r \in \mathcal{M}, \qquad (p, q) \in \mathcal{I}, \tag{5}$$

$$z(r) = a_{10}z_{10}(r) + a_{01}z_{01}(r) + a_{11}z_{11}(r). \tag{6}$$

Теорема 1. *Пусть плотность $f(x)$ независимых одинаково распределенных случайных величин ε_{ij} удовлетворяет следующим условиям:*

$$\mathsf{E}\varepsilon_{11} = 0; \tag{7}$$

$$\mathsf{E}|\varepsilon_{11}| < \infty; \tag{8}$$

$$\int\limits_{-\infty}^{\infty} |f'(x)|\, dx < \infty; \tag{9}$$

$$|f(x) - f(y)| < C|x - y| \text{ для любых } x, y \text{ из } \mathbb{R}, \quad C > 0. \tag{10}$$

Тогда при $\Delta \to 0$

$$P_{mn}(r, \Delta) = \frac{1}{(mn)!}\Big(1 + \Delta z(r)\Big) + o(\Delta).$$

Доказательство теоремы 1 приведено в приложении.

Из теоремы 1 вытекают следующие теоремы, определяющие вид ЛНМ критериев.

Теорема 2. *Пусть выполнены условия (7)–(10) и R — матрица рангов наблюдений поля (1). Тогда ЛНМ ранговый критерий отклоняет H^0 в пользу H_a^+, если*

$$z(R) > C^+, \tag{11}$$

и принимает в противном случае. Постоянная C^+ определяется уровнем значимости критерия.

Теорема 3. *Пусть выполнены условия (7)–(10) и R — матрица рангов наблюдений поля (1). Тогда ЛНМ ранговый критерий отклоняет H^0 в пользу H_a^-, если*

$$z(R) < C^-, \tag{12}$$

и принимает в противном случае. Постоянная C^- определяется уровнем значимости критерия.

Теоремы 2 и 3 позволяют естественным образом определить ранговый критерий для проверки H^0 против двусторонней альтернативы H_a на уровне значимости α как объединение двух односторонних критериев, проверяющих на уровне значимости $\alpha/2$ альтернативы H_a^+ и H_a^-. В этом случае при выполнении условий (7)–(10) гипотеза H^0 отклоняется в пользу H_a, если

$$|z(R)| > C, \tag{13}$$

и принимается в противном случае. Постоянная C определяется уровнем значимости критерия.

Для вычисления меток $a_{mn}(i, j)$ нужна совместная плотность $f_{ij}(x, y)$ порядковых статистик $\varepsilon^{(i)}$ и $\varepsilon^{(j)}$, $1 \le i < j \le mn$, из распределения с плотностью $f(x)$ и функцией распределения $F(x)$. Известно

[6, § 2.2], что

$$f_{ij}(x,y) = \\ = \begin{cases} K_{ij} F^{i-1}(x) f(x) [F(y) - F(x)]^{j-i-1} f(y) [1 - F(y)]^{mn-j}, & \text{если } x \le y, \\ 0, & \text{если } x > y, \end{cases}$$

где

$$K_{ij} = \frac{(mn)!}{(i-1)!(j-i-1)!(mn-j)!}, \qquad 1 \le i < j \le mn.$$

Отсюда, в частности, следует, что условия (8), (9) достаточны для существования меток (3).

Отметим, что при H^0 все $(mn)!$ различных значений R равновероятны. Поэтому распределение $z(R)$ при H^0 не изменится, если матрица рангов R будет вычисляться по наблюдениям X поля (1) в предположении, что плотность инновационного поля ε_{ij} будет отличаться от плотности $f(x)$, порождающей метки (3) статистик (5). При этом мощность критериев (11) – (13), вообще говоря, уменьшится.

Асимптотическая нормальность статистик ЛНМ критериев. Для практического применения критериев (11) – (13) нужно знать распределение статистики $z(R)$ при гипотезе H^0.

Для небольших m и n квантили статистики $z(R)$ можно оценить методом Монте-Карло. Если же m и n велики, то следующая теорема позволяет для распределения $z(R)$ применить нормальную аппроксимацию.

Теорема 4. *Пусть выполнены условия* (7)–(10), $\sigma^2 = \mathsf{E}\varepsilon_{11}^2 < \infty$ *и* f *имеет конечное количество информации Фишера*

$$I(f) = \int\limits_{-\infty}^{\infty} \left(\frac{f'(x)}{f(x)} \right)^2 f(x)\, dx < \infty. \tag{14}$$

Тогда статистики $\dfrac{1}{\sqrt{mn}} z_{pq}(R)$ *асимптотически нормальны с нулевым математическим ожиданием и дисперсией* $\sigma^2 I(f)$:

$$\frac{1}{\sqrt{mn}} z_{pq}(R) \overset{ac}{\sim} \mathcal{N}\big(0, \sigma^2 I(f)\big).$$

Доказательство теоремы 4 приведено в приложении.

Пример 1 (*нормальное распределение*). Пусть

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad F(x) = \Phi(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt.$$

Тогда $\varphi(x) = x$, и для всех $1 \leq i < j \leq mn$

$$a_{mn}(i,j) = \mathsf{E}[\varepsilon^{(i)}\varepsilon^{(j)}] = \mathsf{E}[\Phi^{-1}(U^{(i)})\Phi^{-1}(U^{(j)})] =$$

$$= K_{ij} \int\limits_0^1 du \int\limits_u^1 \Phi^{-1}(u)\Phi^{-1}(v) u^{i-1} [v-u]^{j-i-1} [1-v]^{mn-j}\, dv.$$

Так как $I(f) = 1$, то $\dfrac{1}{\sqrt{mn}} z_{pq}(R) \overset{ac}{\sim} \mathcal{N}(0,1)$.

Пример 2 (*двойное экспоненциальное распределение*).

Пусть

$$f(x) = \frac{1}{2} e^{-|x|}, \qquad F(x) = \begin{cases} \dfrac{1}{2} e^{x}, & \text{если } x < 0, \\ 1 - \dfrac{1}{2} e^{-x}, & \text{если } x \geq 0. \end{cases}$$

Тогда $\varphi(x) = \operatorname{sign} x$,

$$F^{-1}(u) = \operatorname{sign}(1-2u) \ln(1 - |2u-1|), \quad \varphi(F^{-1}(u)) = \operatorname{sign}(2u-1).$$

Поэтому для всех $i, j = 1, \ldots, mn$,

$$a_{mn}(i,j) = \mathsf{E}[\operatorname{sign}(\varepsilon^{(i)})\varepsilon^{(j)}] =$$

$$= \mathsf{E}[\operatorname{sign}(2U^{(i)} - 1) \operatorname{sign}(1 - 2U^{(j)}) \ln(1 - |2U^{(j)} - 1|)],$$

и, например, при $1 \leq i < j \leq mn$

$$a_{mn}(i,j) = K_{ij} \int\limits_0^1 du \int\limits_u^1 \operatorname{sign}(2u-1) \operatorname{sign}(1-2v) \ln(1 - |2v-1|) \times$$

$$\times u^{i-1} [v-u]^{j-i-1} [1-v]^{mn-j}\, dv.$$

Так как $I(f) = 1$, то $\dfrac{1}{\sqrt{mn}} z_{pq}(R) \overset{ac}{\sim} \mathcal{N}(0,1)$.

Пример 3 (*логистическое распределение*). Пусть

$$f(x) = \frac{e^{-x}}{(1+e^{-x})^2}, \qquad F(x) = \frac{1}{1+e^{-x}}.$$

Тогда

$$\varphi(x) = \frac{1-e^{-x}}{1+e^{-x}}, \qquad F^{-1}(u) = \ln\left(\frac{u}{1-u}\right), \qquad \varphi(F^{-1}(u)) = 2u - 1,$$

$$a_{mn}(i,j) = \mathsf{E}\left[\frac{1 - e^{-\varepsilon^{(i)}}}{1 + e^{-\varepsilon^{(i)}}} \varepsilon^{(j)}\right] = \mathsf{E}\left[(2U^{(i)} - 1) \ln\left(\frac{U^{(j)}}{1 - U^{(j)}}\right)\right],$$

$$i, j = 1, \ldots, mn,$$

и, например, для всех $1 \le i < j \le mn$

$$a_{mn}(i,j) = K_{ij} \int_0^1 du \int_u^1 (2u-1) \ln\left(\frac{v}{1-v}\right) u^{i-1}[v-u]^{j-i-1}[1-v]^{mn-j}\, dv.$$

Так как $I(f) = \dfrac{1}{3}$, то $\dfrac{1}{\sqrt{mn}} z_{pq}(R) \overset{ac}{\sim} \mathcal{N}\left(0, \dfrac{1}{3}\right)$.

Выводы. Построены локально наиболее мощные ранговые критерии для проверки гипотезы H_0 о независимости наблюдений в модели пространственной авторегрессии. Статистики построенных критериев при H_0 не зависят от распределения инновационного поля и асимптотически нормальны. Указан явный вид статистик критериев для нормального, двойного экспоненциального и логистического инновационных полей.

Приложение. ***Доказательство теоремы 1***. Для удобства изложения далее всюду для произвольной матрицы C размера $m \times n$ тем же символом C будем обозначать вектор $C = (c_1, \ldots, c_N)$ размерности $N = mn$, элементы которого совпадают с элементами матрицы C, упорядоченными по столбцам:

$$C = (c_{11}, \ldots, c_{m1}, \ldots, c_{1n}, \ldots, c_{mn}),$$

так что равенство $c_{st} = c_k$ будет означать, что $k = m(t-1) + s$, т.е.

$$c_{st} = c_{m(t-1)+s}, \qquad s = 1, \ldots, m, \quad t = 1, \ldots, n.$$

Обозначив $f_\Delta(v)$, $v = (v_{11}, \ldots, v_{m1}, \ldots, v_{1n}, \ldots, v_{mn}) = (v_1, \ldots, v_N)$, плотность X при альтернативе $\theta = \Delta a$, получим

$$\mathsf{P}\{R = r\} = \int_{R=r} f_\Delta(v)\, dv.$$

Выразим плотность $f_\Delta(v)$ через плотность

$$f_0(u) = \prod_{s=1}^{m} \prod_{t=1}^{n} f(u_{st})$$

X при гипотезе H^0, т.е. через плотность случайного вектора

$$\varepsilon = (\varepsilon_{11}, \ldots, \varepsilon_{m1}, \ldots, \varepsilon_{1n}, \ldots, \varepsilon_{mn}) = (\varepsilon_1, \ldots, \varepsilon_N).$$

Так как определитель отображения

$$\begin{aligned} u_{st} &= v_{st} - \Delta(a_{10} v_{s-1,t} + a_{01} v_{s,t-1} + a_{11} v_{s-1,t-1}), \\ &\qquad\qquad s = 1, \ldots, m, \quad t = 1, \ldots, n; \\ v_{st} &= 0 \text{ при } s \le 0 \text{ или } t \le 0 \end{aligned}$$

равен единице, то

$$f_\Delta(v) = \prod_{s=1}^{m}\prod_{t=1}^{n} f(v_{st} - \Delta(a_{10}v_{s-1,t} + a_{01}v_{s,t-1} + a_{11}v_{s-1,t-1})) =$$

$$= \prod_{s=1}^{m}\prod_{t=1}^{n} f(h_{st}(v,\Delta)),$$

где для краткости обозначено

$$h_{st}(v,\Delta)) = h_k(v,\Delta)) = v_{st} - \Delta(a_{10}v_{s-1,t} + a_{01}v_{s,t-1} + a_{11}v_{s-1,t-1}),$$

$$k = m(t-1) + s.$$

Поэтому с учетом того, что при H^0 все $N!$ событий $\{R = r\}$, $r \in \mathcal{M}$, равновероятны,

$$\mathsf{P}\{R = r\} = \int\limits_{R=r} \prod_{s=1}^{m}\prod_{t=1}^{n} f(h_{st}(v,\Delta))\,dv = \int\limits_{R=r} \prod_{s=1}^{m}\prod_{t=1}^{n} f(h_{st}(v,0))\,dv +$$

$$+ \int\limits_{R=r} \left(\prod_{s=1}^{m}\prod_{t=1}^{n} f(h_{st}(v,\Delta)) - \prod_{s=1}^{m}\prod_{t=1}^{n} f(h_{st}(v,0)) \right) dv =$$

$$= \frac{1}{N!} + \Delta \sum_{k=1}^{N} \int\limits_{R=r} \frac{f(h_k(v,\Delta)) - f(h_k(v,0))}{\Delta} \times$$

$$\times \left(\prod_{j=k+1}^{N} f(h_j(v,\Delta)) \right) \left(\prod_{i=1}^{k-1} f(h_i(v,0)) \right) dv.$$

Отметим, что почти всюду по v для $k = m(t-1) + s$

$$\lim_{\Delta\to 0} \frac{f(h_k(v,\Delta)) - f(h_k(v,0))}{\Delta} \times$$

$$\times \left(\prod_{j=k+1}^{N} f(h_j(v,\Delta)) \right) \left(\prod_{i=1}^{k-1} f(h_i(v,0)) \right) =$$

$$= \varphi(v_{st})(a_{10}v_{s-1,t} + a_{01}v_{s,t-1} + a_{11}v_{s-1,t-1}) \prod_{s=1}^{m}\prod_{t=1}^{n} f(v_{st}).$$

Используя (10) и интегрируя сначала по $v_N, \ldots, v_{k+1}$ в указанной последовательности, а затем по остальным v_i за исключением $v_{s-1,t}$, $v_{s,t-1}$, $v_{s-1,t-1}$, получаем для $k = m(t-1) + s$

$$\int\limits_{\mathbb{R}^N} \left| \frac{f(h_k(v,\Delta)) - f(h_k(v,0))}{\Delta} \right| \times$$

$$\times \left(\prod_{j=k+1}^{N} f(h_j(v,\Delta)) \right) \left(\prod_{i=1}^{k-1} f(h_i(v,0)) \right) dv \le$$

$$\le C \sum_{(i,j)\in\mathcal{I}} |a_{ij}| \int\limits_{\mathbb{R}^N} |v_{s-i,t-j}| f(v_{s-i,t-j})\, dv_{s-i,t-j} \le C_1 \mathsf{E}|\varepsilon_{11}| < \infty.$$

Следовательно, по теореме Лебега о мажорируемой сходимости при $\Delta \to 0$

$$\mathsf{P}\{R = r\} = \frac{1}{N!} +$$
$$+ \Delta \sum_{s=1}^{m} \sum_{t=1}^{n} \int\limits_{R=r} \varphi(v_{st})(a_{10}v_{s-1,t} + a_{01}v_{s,t-1} + a_{11}v_{s-1,t-1}) \times$$
$$\times \prod_{i=1}^{m} \prod_{j=1}^{n} f(v_{ij})\, dv + o(\Delta).$$

Делая замену переменных

$$v_{st} = w_{r_{st}}, \qquad s = 1,\ldots,m, \quad t = 1,\ldots,n,$$

с якобианом, равным единице, и учитывая, что плотность вариационного ряда $\varepsilon^{(1)},\ldots,\varepsilon^{(N)}$ равна $\dfrac{1}{N!}\prod\limits_{k=1}^{N} f(w_k)$ при $w_1 < \cdots < w_N$ и нулю в противном случае, получаем

$$\mathsf{P}\{R = r\} = \frac{1}{N!} +$$
$$+ \frac{1}{N!}\Delta \sum_{s=1}^{m} \sum_{t=1}^{n} \int\limits_{w_1 < \cdots < w_N} \varphi(w_{r_{st}})(a_{10}w_{s-1,t} + a_{01}w_{s,t-1} + a_{11}w_{s-1,t-1}) \times$$
$$\times \prod_{i=1}^{m} \prod_{j=1}^{n} f(w_{ij})\, dw + o(\Delta) =$$
$$= \frac{1}{N!}\Big(1 + \Delta\big(a_{10}z_{10}(r) + a_{01}z_{01}(r) + a_{11}z_{11}(r)\big)\Big) + o(\Delta).$$

Теорема 1 доказана.

Доказательство теоремы 4. Так как $X_{ij} = \varepsilon_{ij}$, $i = 1,\ldots,m$, $j = 1,\ldots,n$, при H^0, то R_k — ранг ε_k в последовательности $\varepsilon_1,\ldots,\varepsilon_N$ или, что то же самое, ранг $U_k = F^{-1}(\varepsilon_k)$ в последовательности $U_1,\ldots,U_N$. Обозначим $\mathfrak{F}_N$ — σ-алгебру, порожденную случайными величинами $\{R_1,\ldots,R_N\}$, $\mathfrak{F}_\infty = \bigcup_{N=1}^{\infty} \mathfrak{F}_N$. В работе [7, п.V.4.1] показано, что $\mathsf{E}\left(U_k - \dfrac{R_k}{N+1}\right)^2 < \dfrac{1}{N}$, откуда следует, что $U_1, \ldots, U_N$, а значит, и

$$\varphi(F^{-1}(U_k))F^{-1}(U_l), \qquad k, l = 1, \ldots, N,$$

являются $\mathfrak{F}_\infty$-измеримыми. Используя теорему Б из [7, п. II.1.2], получаем

$$a_{mn}(r_k, r_l) = \mathsf{E}[\varphi(F^{-1}(U^{(r_k)}))F^{-1}(U^{(r_l)})] = \\ = \mathsf{E}[\varphi(F^{-1}(U_k))F^{-1}(U_l)|R_k = r_k, R_l = r_l].$$

Поэтому для всех $k, l = 1, \ldots, N$ случайная величина

$$a_{mn}(R_k, R_l) = \mathsf{E}[\varphi(F^{-1}(U_k))F^{-1}(U_l)|\mathfrak{F}_N]$$

является $\mathfrak{F}_N$-измеримой, и по лемме [7, п.V.1.4, с. 201] с учетом (14)

$$\lim_{N\to\infty} \mathsf{E}\Big(\mathsf{E}[\varphi(F^{-1}(U_k))F^{-1}(U_l)|\mathfrak{F}_N] - \varphi(F^{-1}(U_k))F^{-1}(U_l)\Big)^2 = 0. \quad (15)$$

Лемма

$$\mathsf{D}[z_{pq}(R)] \le mnC\mathsf{E}[a^2_{mn}(R_2, R_1)], \qquad (p, q) \in \mathcal{I},$$

где постоянная C не зависит от m и n.

Доказательство. Так как слагаемые в $z_{pq}(R)$ одинаково распределены, то

$$\mathsf{D}[z_{pq}(R)] = (m-p)(n-q)\mathsf{D}[a_{mn}(R_2, R_1)] + \\ + (m-2p)(n-2q)\mathsf{cov}[a_{mn}(R_2, R_1), a_{mn}(R_3, R_2)] + \\ + 2\big[(m-p)^2(n-q)^2 - (m-p)(n-q) - (m-2p)(n-2q)\big] \times \\ \times \mathsf{cov}[a_{mn}(R_2, R_1), a_{mn}(R_4, R_3)] \le \\ \le N\Big(2\mathsf{D}[a_{mn}(R_2, R_1)] + N\mathsf{cov}[a_{mn}(R_2, R_1), a_{mn}(R_4, R_3)]\Big).$$

Так как при H^0 все значения вектора (R_1, R_2, R_3, R_4) равновероятны, то

$$\mathsf{E}[a_{mn}(R_2, R_1), a_{mn}(R_4, R_3)] = \\ = \frac{1}{N(N-1)(N-2)(N-3)} \sum_{1\le i\ne j\ne k\ne l\le N} a_{mn}(i, j)a_{mn}(k, l) = \\ = \frac{1}{N(N-1)(N-2)(N-3)} \times \\ \times \sum_{1\le i\ne j\le N} a_{mn}(i, j) \left(\sum_{1\le k\ne l\le N} a_{mn}(k, l) - \sum_{\substack{k=1\\k\ne i\\k\ne j}}^{n} a_{mn}(k, i) - \right.$$

$$-\sum_{\substack{l=1\\l\neq i\\k\neq j}}^{n} a_{mn}(k,j)-\sum_{\substack{l=1\\l\neq i\\k\neq j}}^{n} a_{mn}(i,l)-\sum_{\substack{l=1\\l\neq i\\k\neq j}}^{n} a_{mn}(j,l)\Biggr) =$$

$$= \frac{N(N-1)}{(N-2)(N-3)}\mathsf{E}^2[a_{mn}(R_2,R_1)]-$$

$$-\frac{1}{(N-3)}\Bigl(2\mathsf{E}[a_{mn}(R_2,R_1),a_{mn}(R_3,R_2)]+$$

$$+\mathsf{E}[a_{mn}(R_2,R_1),a_{mn}(R_3,R_1)]+\mathsf{E}[a_{mn}(R_3,R_1),a_{mn}(R_3,R_2)]\Bigr).$$

Поэтому

$$\Bigl|\mathsf{cov}[a_{mn}(R_2,R_1),a_{mn}(R_4,R_3)]\Bigr| \le$$

$$\le \frac{2(N-3)}{(N-2)(N-3)}\mathsf{E}^2[a_{mn}(R_2,R_1)]+\frac{4}{(N-3)}\mathsf{E}[a^2_{mn}(R_2,R_1)],$$

откуда следует утверждение леммы.

Определим

$$t_{pq}=\sum_{i=p+1}^{m}\sum_{j=q+1}^{n}\varphi(F^{-1}(U_{ij}))F^{-1}(U_{i-p,j-q}) =$$

$$= \sum_{i=p+1}^{m}\sum_{j=q+1}^{n}\varphi(\varepsilon_{ij})\varepsilon_{i-p,j-q}, \qquad (p,q)\in\mathcal{I}. \quad (16)$$

Так как $R_1,\ldots,R_N$ и $U_1,\ldots,U_N$ при H^0 независимы, то

$$\mathsf{E}\Bigl[(z_{pq}(R)-t_{pq})^2\Bigl|U^{(1)}=u^{(1)},\ldots,U^{(N)}=u^{(N)}\Bigr] =$$

$$= \sum_{i=p+1}^{m}\sum_{j=q+1}^{n} a^*_{mn}(R_{ij},R_{i-p,j-q}),$$

где

$$a^*_{mn}(R_{ij},R_{i-p,j-q}) = a_{mn}(R_{ij},R_{i-p,j-q})-\varphi(F^{-1}(u^{(R_{ij})}))F^{-1}(u^{(R_{i-p,j-q})}).$$

Из леммы следует, что

$$\mathsf{E}\Bigl[(z_{pq}(R)-t_{pq})^2\Bigl|U^{(1)}=u^{(1)},\ldots,U^{(N)}=u^{(N)}\Bigr] \le$$

$$\le mn\mathsf{E}\Bigl[\Bigl(a_{mn}(R_2,R_1)-\varphi(F^{-1}(U_2))F^{-1}(U_1)\Bigr)^2\Bigl|U^{(1)}=$$

$$= u^{(1)},\ldots,U^{(N)}=u^{(N)}\Bigr].$$

Поэтому

$$\mathsf{E}(z_{pq}(R)-t_{pq})^2 \le mn\mathsf{E}\Big(a_{mn}(R_2,R_1)-\varphi(F^{-1}(U_2))F^{-1}(U_1)\Big)^2.$$

Отсюда и из (15) следует

$$\lim_{N\to\infty}\mathsf{E}\frac{1}{mn}\Big(z_{pq}(R)-t_{pq}\Big)^2=0. \tag{17}$$

Так как в сумме (16) каждое слагаемое зависит только от двух других слагаемых этой же суммы, то по центральной предельной теореме для конечно зависимых случайных величин [8, теорема 7.7.5] статистика t_{pq} является асимптотически нормальной. При этом в силу независимости $\varepsilon_{11},\dots,\varepsilon_{mn}$ и условия (7)

$$\mathsf{E}[t_{pq}]=\sum_{i=p+1}^{m}\sum_{j=q+1}^{n}\mathsf{E}[\varphi(\varepsilon_{ij})\varepsilon_{i-p,j-q}]=0,$$

$$\mathsf{D}[t_{pq}]=\sum_{i=p+1}^{m}\sum_{j=q+1}^{n}\mathsf{D}[\varphi(\varepsilon_{ij})\varepsilon_{i-p,j-q}]+$$

$$+2\sum_{i=p+1}^{m}\sum_{j=q+1}^{n}\sum_{\alpha=i+1}^{m}\sum_{\beta=j+1}^{n}\mathsf{cov}[\varphi(\varepsilon_{ij})\varepsilon_{i-p,j-q},\varphi(\varepsilon_{\alpha\beta})\varepsilon_{\alpha-p,\beta-q}]=mnI(f)\sigma^2.$$

Отсюда и из (17) следует утверждение теоремы 4.

СПИСОК ЛИТЕРАТУРЫ

1. R i p l e y B. D. Spatial Statistics (Wiley Series in Probability and Statistics). – Wiley, 1981.
2. T j o s t h e i m D. Statistical Spatial Series Modelling // Advances in Applied Probability. – 1978. Vol. 10. No 1. – P. 130–154.
3. Y a o Q., B r o c k w e l l P. J. Gaussian Maximum Likelihood Estimation for ARMA Models II Spatial Processes // Bernoulli. – 2006. – V. 12. No. 3. – P. 403–429.
4. Г о р я и н о в В. Б., Г о р я и н о в а Е. Р. Знаковые критерии независимости наблюдений в модели пространственной авторегрессии порядка (1,1) // Вестник МГТУ им. Н.Э. Баумана. Сер. Естественные науки. – 2009. – № 2. – С. 115–123.
5. Г о р я и н о в В. Б., Г о р я и н о в а Е. Р. Непараметрическая идентификация пространственной модели авторегрессии в условиях априорной стохастической неопределенности // Автоматика и телемеханика. – 2010. – № 2. – С. 31–41.
6. Д э й в и д Г. Порядковые статистики. – М.: Наука, 1979.
7. Г а е к Я., Ш и д а к З. Теория ранговых критериев. – М.: Наука, 1971.
8. А н д е р с о н Т. Статистический анализ временных рядов. – М.: Мир, 1976.

Статья поступила в редакцию 31.03.2010

Владимир Борисович Горяинов родился в 1961 г., окончил в 1983 г. МГУ им. М.В. Ломоносова. Канд. физ.-мат. наук, доцент кафедры "Математическое моделирование" МГТУ им. Н.Э. Баумана. Автор 23 научных работ в области стохастических дифференциальных уравнений, статистических методов в биологии и медицине.

V.B. Goryainov (b. 1961) graduated from the Lomonosov Moscow State University in 1983. Ph. D. (Phys.-Math.), assoc. professor of "Mathematical Simulation" department of the Bauman Moscow State Technical University. Author of 23 publications in the field of stochastic differential equations, stochastic methods in biology and medicine.
