УДК 519.234.3 DOI: 10.18698/1812-3368-2016-6-4-15

ЗНАКОВЫЕ КРИТЕРИИ ПРОВЕРКИ ГИПОТЕЗЫ О ПОРЯДКЕ УРАВНЕНИЯ В МОДЕЛИ СКОЛЬЗЯЩЕГО СРЕДНЕГО

В.Б. Горяинов[1] vb-goryainov@bmstu.ru
Е.Р. Горяинова[2] el-goryainova@mail.ru

[1] МГТУ им. Н.Э. Баумана, Москва, Российская Федерация
[2] Национальный исследовательский университет «Высшая школа экономики», Москва, Российская Федерация

Аннотация

Построен знаковый критерий проверки гипотезы о порядке уравнения скользящего среднего. Найдено асимптотическое распределение статистики критерия, которое оказалось центральным χ^2-распределением при основной гипотезе и нецентральным χ^2-распределением при альтернативной. Знание асимптотического распределения при альтернативной гипотизе позволяет рассчитывать асимптотическую относительную эффективность построенного знакового критерия по отношению к известным критериям. Приведен пример вычисления асимптотической относительной эффективности построенного знакового критерия по отношению к классическому критерию, основанному на выборочном коэффициенте ковариации. Определены значения асимптотической относительной эффективности для нормального распределения, двойного экспоненциального распределения (распределения Лапласа) и загрязненного нормального распределения (распределения Тьюки). Показано, что при засорении обновляющего процесса гауссовскими выбросами эффективность этого критерия может быть сколь угодно большой по сравнению с традиционным критерием, основанным на выборочном коэффициенте корреляции

Ключевые слова

Модель скользящего среднего, гипотеза о порядке уравнения, знаковый критерий, распределение Тьюки

Поступила в редакцию 25.04.2016

Введение. Одной из основных задач теории случайных процессов является идентификация описывающих их моделей. Для процессов, описываемых уравнением авторегрессии скользящего среднего (ARMA-процессов), идентификация сводится к оцениванию коэффициентов соответствующего разностного уравнения и к определению порядка этого уравнения. Классические методы идентификации, предполагающие гауссовость наблюдаемых процессов, рассмотрены, например, в работах [1–5]. Однако, как показала практика, предположения о гауссовости обычно нарушаются, что может снизить эффективность классических методов. Это обстоятельство привело к появлению альтернативных робастных методов, которые не теряют эффективность при отклонении распределения вероятности наблюдаемого случайного процесса от гауссовского [6–9]. Наиболее распространены из робастных методов оценивания коэффициентов уравнения ARMA-процесса ранговые методы [10, 11], метод наименьших модулей [12], знаковые методы [13, 14] и методы, обобщающие метод максимального правдоподобия [15].

Другая важная задача теории ARMA-процессов — задача определения порядка соответствующего разностного уравнения. Для гауссовских процессов эта задача была решена в работе [16]. Однако построенные в указанной работе критерии имеют низкую эффективность при нарушении предположения о гауссовости обновляющего процесса. Критерии, более устойчивые к предположению о нарушении гауссовости, были предложены в работе [17] и основываются на рангах остатков наблюдений. В настоящей работе для процесса скользящего среднего построен знаковый критерий проверки гипотезы о порядке разностного уравнения. Показано, что при засорении обновляющего процесса гауссовскими выбросами эффективность этого критерия может быть сколь угодно большой по сравнению с традиционным критерием, основанным на выборочном коэффициенте корреляции.

Постановка задачи. Рассмотрим устойчивую модель скользящего среднего

$$u_i = \varepsilon_i + \alpha_1\varepsilon_{i-1} + \ldots + \alpha_q\varepsilon_{i-q}, \quad i = 0, \pm 1, \pm 2, \ldots, \tag{1}$$

где $u_1, \ldots, u_n$ — наблюдения; ε_i — независимые одинаково распределенные случайные величины с неизвестной плотностью распределения вероятности $f(x)$; $\alpha = (\alpha_1, \ldots, \alpha_q)$ — неизвестный вектор параметров. Устойчивость означает, что корни характеристического уравнения

$$M(z) = \sum_{j=0}^{q} \alpha_j z^{q-j} = 0, \quad \alpha_0 = 1, \tag{2}$$

лежат внутри единичного круга. Предположим, что на функцию распределения $F(x)$ случайных величин ε_i наложены следующие условия:

$$F(0) = 1/2, \quad E\varepsilon_1 = 0, \quad E|\varepsilon_1|^{1+r} < \infty, \quad 0 < r \le 1. \tag{3}$$

Проверим гипотезу H_0 о том, что наблюдения $u_1, \ldots, u_n$ описываются схемой (1), в которой $\alpha_j = \alpha_j^0$, $j = 1, \ldots, q$, где $\alpha_1^0, \ldots, \alpha_q^0$ — неизвестные коэффициенты. Альтернативная гипотеза будет заключаться в том, что наблюдения $u_1, \ldots, u_n$ также описываются схемой

$$u_i = \varepsilon_i + \alpha_1\varepsilon_{i-1} + \ldots + \alpha_m\varepsilon_{i-m}, \quad i = 0, \pm 1, \pm 2, \ldots, \tag{4}$$

где $m > q$, $\alpha_j = \alpha_j^0 + \dfrac{K_j}{\sqrt{n}}$, $j = 1, \ldots, m$, $\alpha_{q+1}^0 = \ldots = \alpha_m^0 = 0$; $K_1, \ldots, K_m$ — некоторые постоянные.

Следовательно, хотим проверить гипотезу о том, что порядок уравнения скользящего среднего равен q, против альтернативной гипотезы, состоящей в том, что истинный порядок m, $m > q$.

Приступим к построению соответствующего критерия. Для удобства изменим нумерацию наблюдений и предположим, что наблюдаются величины $u_{1-q}, \dots, u_n$.

Пусть $\hat{\alpha} = (\hat{\alpha}_1, \dots, \hat{\alpha}_q)$ любые $\sqrt{n}$-состоятельные оценки неизвестных параметров α, т. е. $\sqrt{n}(\hat{\alpha} - \alpha) = O_p(1)$ при $n \to \infty$. Если в условии (3) $r = 1$ (существует конечный второй момент $E\varepsilon_1^2$), то в качестве параметра $\hat{\alpha}$ можно использовать, например, оценку наименьших квадратов, или M-оценку [18]. Если выполнено условие (3) с $r < 1$, то для симметричных распределений можно применять Ra-оценку, полученную в работе [19].

Примем $\hat{\alpha}_{q+1} = \dots = \hat{\alpha}_m = 0$, $\hat{\varepsilon}_{1-m} = 0, \dots, \hat{\varepsilon}_0 = 0$,

$$\hat{\varepsilon}_k = u_k - \sum_{j=1}^{q} \hat{\alpha}_j \hat{\varepsilon}_{k-j}, \quad k = 1, \dots, n,$$

$$\gamma_t(\hat{\alpha}) = \frac{1}{\sqrt{n}} \sum_{k=t+1}^{n} \operatorname{sign}(\hat{\varepsilon}_k \hat{\varepsilon}_{k-t}), \quad t = 1, 2, \dots,$$

$$\gamma_t = \frac{1}{\sqrt{n}} \sum_{k=t+1}^{n} \operatorname{sign}(\varepsilon_k \varepsilon_{k-t}), \quad t = 1, 2, \dots$$

Определим последовательность a_k, $k \geq 1 - m$, рекуррентной формулой

$$a_k = -\sum_{j=1}^{m} \alpha_j a_{k-j}, \quad k > 0, \tag{5}$$

с начальными условиями

$$a_{1-m} = \dots = a_{-1} = 0, \qquad a_0 = 1. \tag{6}$$

Обозначим $E\varepsilon_1^- = \int\limits_{-\infty}^{0} x f(x) dx.$

Теорема 1. Если верна гипотеза H_1 и выполнены условия (3), то при $n \to \infty$ для всех $t = 1, 2, \dots$

$$\gamma_t(\hat{\alpha}) - \gamma_t = 4 f(0) E\varepsilon_1^- \left(\sum_{j=1}^{q} \sqrt{n} (\hat{\alpha}_j - \alpha_j) a_{t-j} + \sum_{j=q+1}^{m} K_j a_{t-j} \right) + o_p(1).$$

◀ Определим последовательность $\hat{a}_k$, $k \geq 1 - m$, рекуррентной формулой $\hat{a}_k = -\sum\limits_{j=1}^{m} \hat{\alpha}_j \hat{a}_{k-j}$, $k > 0$, с начальными условиями $\hat{a}_{1-m} = \dots = \hat{a}_{-1} = 0$, $\hat{a}_0 = 1$. Отметим, что $\hat{\varepsilon}_n = \sum_{k=0}^{n-1} \hat{a}_k u_{n-k}$, $n > 0$, $\hat{\varepsilon}_{1-m} = \ldots = \varepsilon_0 = 0$. Поскольку уравнение (1) предполагают устойчивым, то имеет место представление $\varepsilon_n = \sum_{k=0}^{\infty} a_k u_{n-k}$, где коэффициенты a_k, $k \geq 0$, удовлетворяют рекуррентному соотношению (5) с начальными условиями (6), и ряд $\sum_{k=0}^{\infty} a_k$ сходится абсолютно. Обозначим

$$\hat{\eta} = \hat{\alpha} - \alpha = (\hat{\alpha}_1 - \alpha_1, \ldots, \hat{\alpha}_m - \alpha_m)^{\text{т}};$$

$$v_{k-j} = \sum_{i=1}^{k-1} a_{i-j} \varepsilon_{k-i}, \quad \hat{v}_{k-j} = \sum_{i=1}^{k-1} \hat{a}_{i-j} \varepsilon_{k-i}, \;\; j = 1, \ldots, m, \;\; k = 2, \ldots, n;$$

$$V_{k-1} = (v_{k-1}, \ldots, v_{k-m})^{\text{т}}, \quad \hat{V}_{k-1} = (\hat{v}_{k-1}, \ldots, \hat{v}_{k-m})^{\text{т}}, \quad k = 2, \ldots, n;$$

$$\theta_{k-1}(\hat{\eta}) = -\hat{\eta}^{\text{т}} (\hat{V}_{k-1} - V_{k-1}) + \sum_{i=k}^{\infty} a_i u_{k-i}, \quad k = 2, \ldots, n;$$

$$\delta\varepsilon_k = \varepsilon_k - \hat{\varepsilon}_k, \qquad k > -m.$$

Отметим, что

$$\delta\varepsilon_k = \hat{\eta}^{\text{т}} V_{k-1} + \theta_{k-1}(\hat{\eta}), \qquad k = 2, \ldots, n. \tag{7}$$

Обозначим

$$\Delta_k(x) = \begin{cases} 1, & \text{если} \quad \varepsilon_k < x; \\ 0, & \text{если} \quad \varepsilon_k \geq x. \end{cases}$$

Тогда $\operatorname{sign}(\hat{\varepsilon}_k) = 1 - 2I(\hat{\varepsilon}_k < 0) = 1 - 2\Delta_k(\delta\varepsilon_k)$. Найдем разность $\gamma_t(\hat{\alpha}) - \gamma_t$. Имеем

$$\gamma_t(\hat{\alpha}) - \gamma_t = \frac{1}{\sqrt{n}} \sum_{k=t+1}^{n} (1 - 2\Delta_{k-t}(\delta\varepsilon_{k-t}))(1 - 2\Delta_k(\delta\varepsilon_k)) -$$

$$- \frac{1}{\sqrt{n}} \sum_{k=t+1}^{n} (1 - 2\Delta_{k-t}(0))(1 - 2\Delta_k(0) = 4S_1 + 4S_2 + o_p(1), \quad n \to \infty,$$

где

$$S_1 = \frac{1}{\sqrt{n}} \sum_{k=t+1}^{n} (\Delta_{k-t}(\delta\varepsilon_{k-t})\Delta_k(\delta\varepsilon_k) - \Delta_{k-t}(0)\Delta_k(0)) -$$

$$- \frac{f(0)}{\sqrt{n}} \hat{\eta}^{\text{т}} \sum_{k=t+1}^{n} \Delta_{k-t}(0) V_{k-1} - \frac{1}{\sqrt{n}} \sum_{k=t+1}^{n} (\Delta_k(\delta\varepsilon_k) - \Delta_k(0));$$

$$S_2 = f(0) \sum_{j=1}^{m} \sqrt{n} (\hat{\alpha}_j - \alpha_j) \frac{1}{n} \sum_{k=t+1}^{n} \xi_k;$$

$$\xi_k = \Delta_{k-t}(0)\sum_{i=1}^{k-1} a_{i-j}\varepsilon_{k-i}, \quad k = t+1,\dots,n.$$

Из независимости и одинаковой распределенности ε_k, $k = 0, \pm 1, \pm 2,\dots,$ следует, что

$$\mathrm{E}\xi_k = a_{t-j}\mathrm{E}(\Delta_k(0)\varepsilon_{k-i}) = a_{t-j}\mathrm{E}\varepsilon_1^-, \quad k = t+1,\dots,n.$$

Поэтому на основании закона больших чисел при $n\to\infty$ имеем

$$\frac{1}{n}\sum_{k=t+1}^{n}\xi_k = a_{t-j}\mathrm{E}\varepsilon_1^- + o_p(1), \quad j = 1,\dots,m, \quad t>0.$$

Таким образом, при $n\to\infty$

$$S_2 = f(0)\mathrm{E}\varepsilon_1^-\left(\sum_{j=1}^{q}\sqrt{n}(\hat\alpha_j-\alpha_j)a_{t-j} + \sum_{j=q+1}^{m}K_j a_{t-j}\right) + o_p(1), \quad t>0.$$

Применяя метод перехода от верхней грани по континууму значений η на отрезке $[-n^{-\mu}, n^{-\mu}]$ к верхней грани по конечным множествам [20]

$$\eta_s = -n^{-\mu} + s2n^{-\mu}3^{-m_n}, \quad s = 0,1,\dots,3^{m_n}, \quad m_n \sim \frac{\mu\ln n}{4m\ln 3},$$

получаем, что при всех μ таких, что

$$\max\left(\frac{2m}{4m+1}, \ \frac{1}{2}\frac{r+1}{2r+1}\right) < \mu < \frac{1}{2},$$

справедливо $\sup\limits_{|\eta|\le n^{-\mu}}|z(\eta)| \xrightarrow{\mathrm{P}} 0, \ n\to\infty.$

В силу ограниченности по вероятности $\sqrt{n}\hat\eta$ для любого $\varepsilon>0$ и определенного выше μ

$$\mathrm{P}\{S_1>\varepsilon\} \le \mathrm{P}\left\{S_1>\varepsilon, |\eta|\le n^{-\mu}\right\} + \mathrm{P}\left\{|\eta|>n^{-\mu}\right\} \le$$
$$\le \mathrm{P}\left\{\sup_{|\eta|\le n^{-\mu}}|z(\eta)|>\varepsilon\right\} + \mathrm{P}\left\{\sqrt{n}\,|\hat\eta|>n^{1/2-\mu}\right\} \to 0, \quad n\to\infty,$$

что и завершает доказательство теоремы. ►

Из теоремы 1 вытекает следствие.

Следствие 1. Если верна гипотеза H_0, то при $n\to\infty$

$$\gamma_t(\hat\alpha) - \gamma_t = 4f(0)\mathrm{E}\varepsilon_1^-\sum_{j=1}^{q}\sqrt{n}\left(\hat\alpha_j - \alpha_j^0\right)a_{t-j} + o_p(1), \quad n\to\infty, \quad t>0. \qquad (8)$$

Величины γ_t, $t=1,2,\ldots$, в любом конечном наборе совместно асимптотически при $n\to\infty$ нормальны со средним 0 и дисперсией 1, а также независимы. Однако в силу (8) для величин $\gamma_t(\hat\alpha)$ это неверно, их распределение и при гипотезе H_0 зависит от параметров $\alpha_1,\ldots,\alpha_q$ и оценок $\hat\alpha_1,\ldots,\hat\alpha_q$. Чтобы избежать этого затруднения, введем величины

$$\hat c_t^0=\gamma_t(\hat\alpha)+\alpha_1^0\gamma_{t-1}(\hat\alpha)+\ldots+\alpha_q^0\gamma_{t-q}(\hat\alpha),\quad t=q+1,\ldots,$$

и

$$c_t^0=\gamma_t+\alpha_1^0\gamma_{t-1}+\ldots+\alpha_q^0\gamma_{t-q},\quad t=q+1,\ldots$$

Согласно (8), при гипотезе H_0 и $n\to\infty$ для всех $t=q+1,\ldots$ имеем

$$\hat c_t^0-c_t^0=4f(0)\mathrm{E}\varepsilon_1^-\sum_{j=1}^{q}\sqrt{n}\left(\hat\alpha_j-\alpha_j^0\right)\sum_{i=0}^{q}\alpha_i^0a_{t-i-j}+o_p(1)=o_p(1),\quad n\to\infty,\tag{9}$$

так как в силу определения величин a_k, $k>0$, $\sum\limits_{i=0}^{q}\alpha_i^0a_{t-i-j}=0$ при любых $j=1,\ldots,q$ и $t=q+1,\ldots$

Очевидно, что величины c_t^0, $t=q+1,\ldots,m$, совместно асимптотически при $n\to\infty$ нормальны с нулевым средним и ковариационной матрицей

$$R_{ts}=\begin{cases}0, & |t-s|>q;\\ \sum\limits_{j=0}^{q-|t-s|}\alpha_j^0\alpha_{j+|t-s|}^0, & |t-s|\le q.\end{cases}\tag{10}$$

Из (9) следует, что при гипотезе H_0 величины $\hat c_t^0$, $t=1,\ldots,m$, имеют такое же асимптотическое распределение.

Заменяя неизвестные $\alpha_1^0,\ldots,\alpha_q^0$ $\sqrt{n}$-состоятельными оценками $\hat\alpha_1,\ldots,\hat\alpha_q$, получаем статистики

$$\hat c_t=\gamma_t(\hat\alpha)+\hat\alpha_1\gamma_{t-1}(\hat\alpha)+\ldots+\hat\alpha_q\gamma_{t-q}(\hat\alpha),\quad t=q+1,\ldots,m,$$

которые при гипотезе H_0 совместно асимптотически при $n\to\infty$ нормальны с нулевым средним и ковариационной матрицей (10).

При гипотезе H_1 в силу теоремы 1 запишем

$$\hat c_t^0-c_t^0=4f(0)\mathrm{E}\varepsilon_1^-\sum_{j=q+1}^{m}K_j\sum_{i=0}^{q}\alpha_i^0a_{t-i-j}+o_p(1)=\begin{cases}4f(0)\mathrm{E}\varepsilon_1^-K_t+o_p(1), & t=q+1,\ldots,m;\\ o_p(1), & t=q+1,\ldots\end{cases}$$

Аналогично предыдущему случаю получаем, что при гипотезе H_1 величины $\hat c_t$, $t=q+1,\ldots,m$, совместно асимптотически при $n\to\infty$ нормальны со средними

$$\delta_t=\begin{cases}4f(0)\mathrm{E}\varepsilon_1^- K_t, & t=q+1,\ldots,m;\\ 0, & t=q+1,\ldots\end{cases}$$

и ковариациями (10).

Введем вектор $\hat{C}=(\hat{c}_{q+1},\ldots,\hat{c}_m)^{\mathrm{т}}$. Пусть $\hat{R}=(\hat{R}_{ts})$, где $\hat{R}$ получается заменой в R неизвестных $\alpha_1^0,\ldots,\alpha_q^0$ оценками $\hat{\alpha}_1,\ldots,\hat{\alpha}_q$. Статистикой критерия для проверки гипотезы H_0 возьмем статистику $\hat{C}^{\mathrm{т}}\hat{R}^{-1}\hat{C}$. Таким образом, можно сформулировать следующую теорему.

Теорема 2. Если выполнены условия (3), то при гипотезе H_0 статистика $\hat{C}^{\mathrm{т}}\hat{R}^{-1}\hat{C}$ асимптотически при $n\to\infty$ распределена как χ^2 с $m-q$ степенями свободы, а при гипотезе H_1 как нецентральный χ^2 с $m-q$ степенями свободы и параметром нецентральности $(\delta_{q+1},\ldots,\delta_m)^{\mathrm{т}}R^{-1}(\delta_{q+1},\ldots,\delta_m)$.

Согласно теореме 2, проверить гипотезу H_0 против альтернативной гипотезы H_1 можно с помощью статистики $\hat{C}^{\mathrm{т}}\hat{R}^{-1}\hat{C}$, отклоняя гипотезу H_0 в пользу гипотезы H_1 на уровне значимости γ, если $\hat{C}^{\mathrm{т}}\hat{R}^{-1}\hat{C}>\chi^2_{1-\gamma}(m-q)$, где $\chi^2_{1-\gamma}(m-q)$ — квантиль χ^2-распределения с $m-q$ степенями свободы. Кроме того, знание распределения статистики $\hat{C}^{\mathrm{т}}\hat{R}^{-1}\hat{C}$ при альтернативной гипотезе H_1 позволяет сравнивать эффективность этого критерия с другими критериями, как это видно из приведенного ниже примера.

Пример. Сравним эффективность построенного знакового критерия с классическим критерием, основанным на выборочном коэффициенте корреляции. Пусть для простоты $q=0$, $m=1$, т. е. гипотеза H_0 состоит в том, что u_k — последовательность типа белого шума, а при альтернативной гипотезе H_1 процесс u_k подчиняется уравнению скользящего среднего первого порядка

$$u_i=\varepsilon_i+\frac{K}{\sqrt{n}}\varepsilon_{i-1},\ \ i=0,\ \pm1,\ \pm2,\ldots$$

Если u_k удовлетворяет уравнению $u_i=\varepsilon_i+\alpha\varepsilon_{i-1}$, $i=0,\pm1,\pm2,\ldots,$ то выборочный коэффициент корреляции

$$r_n=\frac{\sum\limits_{k=2}^{n}u_k u_{k-1}}{\sum\limits_{k=2}^{n}u_k^2}$$

асимптотически нормален с математическим ожиданием $\rho=\dfrac{\alpha}{1+\alpha^2}$ и дисперсией $w^2=1-3\rho^2+4\rho^4$. Обозначим $T_n(\alpha)=\sqrt{n}w^{-1}(r_n-\rho)$. Статистика $T_n(\alpha)$ имеет стандартное нормальное распределение. Классический критерий отклоняет гипотезу $\alpha=\alpha_0$ в пользу альтернативной гипотезы $\alpha\neq\alpha_0$ на уровне значимости γ, если $T_n^2(\alpha_0) > \chi^2_{1-\gamma/2}$, где $\chi^2_{1-\gamma/2}$ — квантиль χ^2-распределения уровня $1-\gamma/2$ с одной степенью свободы. В связи с этим при альтернативной гипотезе H_1 статистика $T_n(\alpha)$ имеет нецентральное χ^2-распределение с параметром нецентральности K^2. Следовательно, асимптотическая относительная эффективность (АОЭ) знакового критерия относительно классического критерия, определяемая как отношение параметров нецентральности, равна $e = \left(4f(0)\mathrm{E}\varepsilon_1^-\right)^2$.

Вычислим АОЭ e для различных распределений обновляющего процесса ε_t. Если процесс ε_t имеет стандартное нормальное распределение плотностью $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$, то эффективность знакового критерия низка, поскольку $e = 4/\pi^2$. Другими словами, знаковому критерию необходимо в $\pi^2/4 \approx 2,4674$ раз больше наблюдений для достижения выводов той же надежности, которые дает классический критерий. Если процесс ε_t имеет двойное экспоненциальное распределение плотностью $f(x) = (1/2)e^{-|x|}$, то $e=1$, что свидетельствует об одинаковой эффективности двух критериев.

Рассмотрим типичную на практике ситуацию, когда процесс ε_t в большинстве своем представляют стандартные нормальные величины, однако с небольшой вероятностью δ среди них попадаются нормальные случайные величины с нулевым средним и большей дисперсией $\tau^2 > 1$. Другими словами, пусть процесс ε_t имеет загрязненное нормальное распределение (распределение Тьюки), плотность распределения вероятности которого имеет вид

$$f(x) = (1-\delta)\frac{1}{\sqrt{2\pi}} e^{-x^2/2} + \delta \frac{1}{\sqrt{2\pi}\tau} e^{-x^2/(2\tau^2)}.$$

Тогда

$$e = \frac{4(\tau+\delta-\tau\delta)^2(1-\delta+\tau\delta)^2}{\pi^2\tau^2}.$$

Значение АОЭ может быть сколь угодно большим, если при $\delta > 0$ и $\tau > 0$ доля загрязнений δ стремится к единице или величина загрязнений τ стремится к бесконечности. Например, $e = 2,1365$ при $\delta = 0,2$ и $\tau = 10$.

Выводы. Определен знаковый критерий проверки гипотезы о порядке уравнения скользящего среднего. Установлено, что асимптотическое распределение статистики критерия является центральным χ^2-распределением при основной гипотезе и нецентральным χ^2-распределением при альтернативной гипотезе. С помощью асимптотического распределения при альтернативной гипотезе можно находить асимптотическую относительную эффективность построенного знакового критерия по отношению к уже известным критериям. Рассмотрен пример вычисления асимптотической относительной эффективности построенного знакового критерия по отношению к классическому критерию, основанному на выборочном коэффициенте ковариации. Определены значения асимптотической относительной эффективности для нормального распределения, двойного экспоненциального распределения (распределения Лапласа) и загрязненного нормального распределения (распределения Тьюки).

ЛИТЕРАТУРА

1. *Schelter B., Winterhalder M., Timmer J.* Handbook of time series analysis: recent theoretical developments and applications. Weinheim: Wiley, 2006. 508 p.

2. *Montgomery D.C., Jennings C.L., Kulahci M.* Introduction to time series analysis and forecasting. Hoboken: Wiley, 2015. 655 p.

3. *Tsay R.S.* Analysis of time series. Hoboken: Wiley, 2010. 667 p.

4. *Rao T.S., Rao S.S., Rao C.R.* Handbook of statistics. Vol. 30. Time series analysis: methods and applications. Amsterdam: Elsevier, 2012. 755 p.

5. *Wilson G.T., Reale M., Haywood J.* Models for dependent time series. Boca Raton: CRC Press, 2015. 334 p.

6. *Daraio C., Simar L.* Advanced robust and nonparametric methods in efficiency analysis. New York: Springer, 2007. 260 p.

7. *Huber P., Ronchetti E.M.* Robust statistics. Hoboken: Wiley, 2009. 360 p.

8. *Hettmansperger T.P., McKean J.W.* Robust nonparametric statistical methods. Boca Raton: CRC Press, 2011. 535 p.

9. *Wilcox R.R.* Introduction to robust estimation and hypothesis testing. Amsterdam: Elsevier, 2012. 689 p.

10. *Andrews B.* Rank-based estimation for autoregressive moving average time series models // J. Time Ser. Anal. 2008. Vol. 29. No. 1. P. 51–73.

11. *Goryainov V.B.* Identification of a spatial autoregression by rank methods // Automation and Remote Control. 2011. Vol. 72. No. 5. P. 975–988.

12. *Goryainov V.B.* Least-modules estimates for spatial autoregression coefficients // Journal of Computer and Systems Sciences International. 2011. Vol. 50. No. 4. P. 565–572.

13. *Горяинова Е.Р., Горяинов В.Б.* Знаковые критерии в модели скользящего среднего // Вестник МГТУ им. Н.Э. Баумана. Сер. Естественные науки. 2008. № 1. С. 76–86.

14. *Goryainov V.B., Goryainova E.R.* Nonparametric identification of the spatial autoregression model under a priori stochastic uncertainty // Automation and Remote Control. 2010. No. 2. P. 198–208.

15. *Truquet L., Yao J.* On the quasi-likelihood estimation for random coefficient autoregressions // Statistics. 2012. Vol. 46. No. 4. P. 505–521.

16. *McLeod A.I.* On the distribution and applications of residual autocorrelations in Box-Jenkins models // J. R. Statist. Soc. B. 1978. Vol. 40. P. 296–302.

17. *Hallin M., Puri M.L.* Optimal rank-based procedures for time series analysis: Testing an ARMA model against other ARMA models // The Annals of Statistics. 1988. Vol. 16. P. 402–432.

18. *Maronna R.A., Martin D., Yohai V.* Robust statistics: Theory and methods. Chichester: Wiley, 2006. 403 p.

19. *Bustos O., Fraiman R., Yohai V.J.* Asymptotic behaviour of the estimates based on residual autocovariances for ARMA models // Lecture Notes in Statist. 1984. Vol. 26. P. 26–49.

20. *Mukantseva L.A.* Testing normality in one-dimensional and multi-dimensional linear regression // Theory of Probability and its Applications. 1978. Vol. 22. P. 591–602.

Горяинов Владимир Борисович — д-р физ.-мат. наук, доцент, профессор кафедры «Математическое моделирование» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5).

Горяинова Елена Рудольфовна — канд. физ.-мат. наук, доцент, доцент департамента математики на факультете экономических наук Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ) (Российская Федерация, 101000, Москва, ул. Мясницкая, д. 20).

Просьба ссылаться на эту статью следующим образом:

Горяинов В.Б., Горяинова Е.Р. Знаковые критерии проверки гипотезы о порядке уравнения в модели скользящего среднего // Вестник МГТУ им. Н.Э. Баумана. Сер. Естественные науки. 2016. № 6. С. 4–15. DOI: 10.18698/1812-3368-2016-6-4-15

SIGN TEST FOR HYPOTHESIS ABOUT THE ORDER OF EQUATION IN MOVING AVERAGE MODEL

V.B. Goryainov[1] vb-goryainov@bmstu.ru
E.R. Goryainova[2] el-goryainova@mail.ru

[1] Bauman Moscow State Technical University, Moscow, Russian Federation
[2] National Research University Higher School of Economics, Moscow, Russian Federation

Abstract

The article deals with constructing the sign test for the hypothesis about the order of equation in moving average. We found the asymptotic distribution of the test statistics which appeared to be the central χ^2-distribution under the null hypothesis and the noncentral χ^2-distribution under the alternative one. Knowing the asymptotic distribution makes it possible to calculate the asymptotic relative efficiency of the constructed sign test criterion with respect to the known criteria. In our research we give an example of calculating the asymptotic relative efficiency of the constructed sign test criterion in relation to the classical criterion, based on a sample covariance ratio. Moreover, we determine the values of the asymptotic relative efficiency for a normal distribution, the double exponential distribution (Laplace distribution) and contaminated normal distribution (Tukey distribution). It is shown that if the innovation process in the moving average model is contaminated with Gaussian outliers, the asymptotic relative efficiency of this test can be arbitrarily large compared to the traditional criterion, based on a sample correlation coefficient

Keywords

Moving average model, hypothesis about the order of the equation, sign test, Tukey distribution

REFERENCES

[1] Schelter B., Winterhalder M., Timmer J. Handbook of time series analysis: recent theoretical developments and applications. Weinheim, Wiley, 2006. 508 p.

[2] Montgomery D.C., Jennings C.L., Kulahci M. Introduction to time series analysis and forecasting. Hoboken, Wiley, 2015. 655 p.

[3] Tsay R.S. Analysis of time series. Hoboken, Wiley, 2010. 667 p.

[4] Rao T.S., Rao S.S., Rao C.R. Handbook of statistics. Vol. 30. Time series analysis: methods and applications. Amsterdam, Elsevier, 2012. 755 p.

[5] Wilson G.T., Reale M., Haywood J. Models for dependent time series. Boca Raton, CRC Press, 2015. 334 p.

[6] Daraio C., Simar L. Advanced robust and nonparametric methods in efficiency analysis. N.Y., Springer, 2007. 260 p.

[7] Huber P., Ronchetti E.M. Robust statistics. Hoboken, Wiley, 2009. 360 p.

[8] Hettmansperger T.P., McKean J.W. Robust nonparametric statistical methods. Boca Raton, CRC Press, 2011. 535 p.

[9] Wilcox R.R. Introduction to robust estimation and hypothesis testing. Amsterdam, Elsevier, 2012. 689 p.

[10] Andrews B. Rank-based estimation for autoregressive moving average time series models. *J. Time Ser. Anal.*, 2008, vol. 29, no. 1, pp. 51–73.

[11] Goryainov V.B. Identification of a spatial autoregression by rank methods. *Automation and Remote Control*, 2011, vol. 72, no. 5, pp. 975–988.

[12] Goryainov V.B. Least-modules estimates for spatial autoregression coefficients. *Journal of Computer and Systems Sciences International*, 2011, vol. 50, no. 4, pp. 565–572.

[13] Goryainova E.R., Goryainov V.B. Sign tests in moving-average model. *Vestn. Mosk. Gos. Tekh. Univ. im. N.E. Baumana, Estestv. Nauki* [Herald of the Bauman Moscow State Tech. Univ., Nat. Sci.], 2008, no. 1, pp. 76–86 (in Russ.).

[14] Goryainov V.B., Goryainova E.R. Nonparametric identification of the spatial autoregression model under a priori stochastic uncertainty. *Automation and Remote Control*, 2010, no. 2, pp. 198–208.

[15] Truquet L., Yao J. On the quasi-likelihood estimation for random coefficient autoregressions. *Statistics*, 2012, vol. 46, no. 4, pp. 505–521.

[16] McLeod A.I. On the distribution and applications of residual autocorrelations in Box-Jenkins models. *J. R. Statist. Soc. B*, 1978, vol. 40, pp. 296–302.

[17] Hallin M., Puri M.L. Optimal rank-based procedures for time series analysis: Testing an ARMA model against other ARMA models. *The Annals of Statistics*, 1988, vol. 16, pp. 402–432.

[18] Maronna R.A., Martin D., Yohai V. Robust statistics: Theory and methods. Chichester, Wiley, 2006. 403 p.

[19] Bustos O., Fraiman R., Yohai V.J. Asymptotic behaviour of the estimates based on residual autocovariances for ARMA models. *Lecture Notes in Statist*, 1984, vol. 26, pp. 26–49.

[20] Mukantseva L.A. Testing normality in one-dimensional and multi-dimensional linear regression. *Theory of Probability and its Applications*, 1978, vol. 22, pp. 591–602.

Goryainov V.B. — Dr. Sci. (Phys.-Math.), Professor of Mathematical Modelling Department, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, Moscow, 105005 Russian Federation).

Goryainova E.R. — Cand. Sci. (Phys.-Math.), Assoc. Professor of Faculty of Economic Sciences, Department of Mathematics, National Research University Higher School of Economics (Myasnitskaya ul. 20, Moscow, 101000 Russian Federation).

Please cite this article in English as:

Goryainov V.B., Goryainova E.R. Sign Test for Hypothesis about the Order of Equation in Moving Average Model. *Vestn. Mosk. Gos. Tekh. Univ. im. N.E. Baumana, Estestv. Nauki* [Herald of the Bauman Moscow State Tech. Univ., Nat. Sci.], 2016, no. 6, pp. 4–15.
DOI: 10.18698/1812-3368-2016-6-4-15