

М.А. Сиренко, П.Ф. Тарасенко

ТОЧНЫЕ И АСИМПТОТИЧЕСКИЕ ЗНАКОВЫЕ ТЕСТЫ ДЛЯ ПАРАМЕТРОВ НЕЛИНЕЙНОЙ КВАНТИЛЬНОЙ РЕГРЕССИИ С МАРКОВСКИМИ ШУМАМИ

Одним из привлекательных свойств знаковых статистических процедур является возможность строить точные тесты для проверки простых гипотез о параметрах регрессионных моделей. В данной работе этот подход распространяется на случай нелинейной модели с зависимыми шумами. Рассматривается модель многоквантильной регрессии, что позволяет совместно проверять гипотезы как о параметрах регрессии, так и о параметрах масштаба шума.

Ключевые слова: знаковые статистические процедуры; нелинейная регрессия; квантильная регрессия; Марковский процесс; проверка простых гипотез; точные тесты.

Знаковые статистические процедуры позволяют строить тесты для проверки простых гипотез о параметрах линейных моделей в непараметрической постановке задачи. Это объясняется тем, что при гипотезе распределение знаков невязок известно, поэтому такие тесты позволяют точно контролировать уровень значимости. Знаковые методы сначала были разработаны для меданной регрессии [1], т.е. для частного случая квантильной регрессии, для которой намного раньше были развиты методы, основанные на минимизации взвешенных модулей невязок (см. основополагающую работу [2]). В частности, на основе метода наименьших модулей в [3] были предложены методы анализа симметричной двухквантильной регрессии. Позднее были предложены знаковые процедуры для квантильной регрессии [4, 5], в том числе для зависимых шумов [6].

В данной работе рассматривается ряд обобщений знакового подхода к проверке простых гипотез для параметров квантильной регрессии. Во-первых, мы рассматриваем случай нелинейной модели. Во-вторых, рассматривается многоквантильная регрессия, поэтому от знаков невязок мы переходим к индикаторам их попадания в межквантильные интервалы, как это предлагается в [7]. В-третьих, рассматриваются зависимые наблюдения, когда ошибки образуют стационарный Марковский процесс с дискретным временем.

Рассмотрение методов проверки простых гипотез является первым шагом для построения процедур оценивания параметров на основе подхода, который, как принято считать, впервые был использован в статье [8], а также для построения процедур проверки линейных гипотез, которые более интересны для практических приложений.

1. Постановка задачи

Рассмотрим нелинейную модель регрессии с зависимыми случайными ошибками вида

$$y_t = g_t(\theta) + \varepsilon_t, \quad t = \overline{1, n}, \quad (1)$$

где $g_t(\theta)$ – заданная при каждом t непрерывно дифференцируемая функция от параметров $\theta = (\theta_1, \dots, \theta_T)' \in \mathbb{R}^T$. Случайные отклонения ε_t образуют стационарный Марковский процесс $(r-1)$ -го порядка. Одномерные функции распределения $P\{\varepsilon_t < x\}$ неизвестны и не обязательно совпадают при разных t , но имеют несколько совпадающих квантилей заданных уровней.

Подробнее: пусть конечный набор смежных интервалов $C_1(\mu), \dots, C_K(\mu)$ образует разбиение \mathbb{R}^1 , при этом $P\{\varepsilon_t \in C_k(\mu)\} = p_k$, $k = \overline{1, K}$, где вероятности p_k заданы. Параметры μ определяют масштаб одномерного распределения шумов, так как влияют на ширину межквантильных интервалов. Границы

интервалов $C_k(\mu) = \langle c_{k-1}(\mu), c_k(\mu) \rangle$ (угловая скобка может означать либо открытую, либо закрытую границу) зависят от неизвестных параметров μ линейно $c_k(\mu) = a_k + d'_k \mu$, $k = \overline{1, K-1}$, $c_0(\mu) = -\infty$, $c_K(\mu) = +\infty$. При этом a_k и d_k заданы, а допустимые параметры образуют множество $\left\{ \mu : (d_k - d_{k-1})' \mu + (a_k - a_{k-1}) > 0, k = \overline{2, K-1} \right\}$.

Наиболее интересными являются самые простые случаи симметричной двухквантильной и трехквантильной регрессии, когда единственный параметр μ равен половине интерквантильного размаха. Для двухквантильной регрессии $K = 3$, $c_1(\mu) = -\mu$, $c_2(\mu) = \mu$, $p_1 = p_3 = p$, $p_2 = 1-2p$. Для трехквантильной регрессии $K = 4$, $c_1(\mu) = -\mu$, $c_2(\mu) = 0$, $c_3(\mu) = \mu$, $p_1 = p_4 = p$, $p_2 = p_3 = (1-2p)/2$. В эту же модель укладывается одноквантильная регрессия, при этом параметры μ отсутствуют, $K = 2$, $c_1(\mu) = 0$, $p_1 = p$, $p_2 = 1-p$. Во всех трех моделях вероятность p задана.

Введем обозначения для совместных вероятностей:

$$P^{(l)}(k_1, \dots, k_l) = P\{\varepsilon_{t-l+1} \in C_{k_1}(\mu), \dots, \varepsilon_t \in C_{k_l}(\mu)\}, \quad l = \overline{1, r}, \quad k_1, \dots, k_r = \overline{1, K}, \quad (2)$$

$$\tilde{P}^{(l)} = \left\{ P^{(l)}(k_1, \dots, k_l) : k_1, \dots, k_l = \overline{1, K} \right\}, \quad l = \overline{1, r}.$$

В частности, $\tilde{P}^{(1)} = \{p_1, \dots, p_K\} = \{P^{(1)}(1), \dots, P^{(1)}(K)\}$.

Неизвестными параметрами являются параметры θ , μ и набор $\tilde{P}^{(r)}$, причем в составе $\tilde{P}^{(r)}$ есть линейно зависимые вероятности. Обозначим через Q вектор, образованный некоторым набором линейно независимых вероятностей из состава $\tilde{P}^{(r)}$.

Рассмотрим структурные преобразования $j_l(i_1, \dots, i_l) = 1 + \sum_{j=1}^l (i_j - 1)K^{l-j}$, $l = \overline{1, r}$, которые осуществляют развертку r -мерного набора вероятностей $\tilde{P}^{(l)}$ в одномерный вектор $P^{(l)}$ по правилу $[P^{(l)}]_{j_l(i_1, \dots, i_l)} = P^{(l)}(i_1, \dots, i_l)$. Здесь и далее $[A]_j$ означает j -ю строку матрицы или j -й элемент вектора. С помощью структурной матрицы G и вектора D можно осуществить переход от свободных вероятностей Q к $P^{(r)} = D + GQ$. Это преобразование может учитывать не только условие нормировки, заданные одномерные вероятности, условия стационарности, но и симметрию конечномерных распределений, если это необходимо. Кроме того, далее будем использовать структурные матрицы F_i , $i = \overline{1, r-1}$, которые обеспечивают переходы к вероятностям меньшего порядка $P^{(i)} = F_i P^{(i+1)}$.

В этих условиях, обозначив истинные параметры через $v = (\theta', \mu', Q')$, а гипотетические через $v_0 = (\theta'_0, \mu'_0, Q'_0)$, сформулируем задачу проверки простой гипотезы H_0 о параметрах моделей (1) и (2) против сложной альтернативы H_1 :

$$H_0 : v = v_0, \quad H_1 : v \neq v_0. \quad (3)$$

В качестве признаков для построения статистических процедур будем использовать индикаторы $\bar{s} = (s_1, \dots, s_n)$ – номера интервалов, в которые попадают невязки $y_t - g_t(\theta_0)$, т.е. $s_t = s_t(\theta_0, \mu_0) = s(y_t - g_t(\theta_0), \mu_0)$, где $s(u, \mu_0) = k$ для $u \in C_k(\mu_0)$.

Тест для проверки гипотез (3) построим по принципу максимума отношения правдоподобия, отбирая в критическую область те параметры, которые доставляют наибольшие значения нормы его градиента в гипотетической точке:

$$\nabla_v L(\bar{s} | v, v_0) \Big|_{v=v_0} = \frac{\nabla_v P(\bar{s} | v)}{P(\bar{s} | v_0)} \Big|_{v=v_0}. \quad (4)$$

Здесь $L(\bar{s} | v, v_0) = P(\bar{s} | v) / P(\bar{s} | v_0)$ – функция отношения правдоподобия индикаторных признаков.

2. Градиент правдоподобия индикаторных признаков

При поиске выражений для правой части в равенстве (4) воспользуемся представлением

$$P(s_1, \dots, s_n | v) = P(s_1 | v) P(s_2 | s_1, v) \dots P(s_{r-1} | s_1, s_2, \dots, s_{r-2}, v) \prod_{t=r}^n P(s_t | s_{t-r+1}, \dots, s_{t-1}, v), \quad (5)$$

чтобы получить разложение Тейлора $P(\bar{s} | v) = P(\bar{s} | v_0) + \nabla'_v P(\bar{s} | v_0)(v - v_0) + o(\|v - v_0\|)$. Для данного разложения требуется непрерывно дифференцируемая параметризация семейства конечномерных распределений по μ и Q . Для любой такой параметризации имеем

$$\begin{aligned} P(s_1 | v) &= P(s_1 | v_0) + \nabla'_v P(s_1 | v_0)(v - v_0) + o(\|v - v_0\|), \\ P(s_2 | s_1, v) &= P(s_2 | s_1, v_0) + \nabla'_v P(s_2 | s_1, v_0)(v - v_0) + o(\|v - v_0\|), \dots \\ \dots P(s_t | s_{t-r+1}, \dots, s_{t-1}, v) &= P(s_t | s_{t-r+1}, \dots, s_{t-1}, v_0) + \nabla'_v P(s_t | s_{t-r+1}, \dots, s_{t-1}, v_0) \times \\ &\quad \times (v - v_0) + o(\|v - v_0\|), \quad t = \overline{r, n}. \end{aligned} \quad (6)$$

Подставив (6) в (5), после перегруппировки получим

$$\begin{aligned} \frac{P(\bar{s} | v)}{P(\bar{s} | v_0)} &= 1 + \left(\frac{\nabla'_v P(s_1 | v_0)}{P(s_1 | v_0)} + \frac{\nabla'_v P(s_2 | s_1, v_0)}{P(s_2 | s_1, v_0)} + \dots + \frac{\nabla'_v P(s_{r-1} | s_1, \dots, s_{r-2}, v_0)}{P(s_{r-1} | s_1, \dots, s_{r-2}, v_0)} + \right. \\ &\quad \left. + \sum_{t=r}^n \frac{\nabla'_v P(s_t | s_{t-r+1}, \dots, s_{t-1}, v_0)}{P(s_t | s_{t-r+1}, \dots, s_{t-1}, v_0)} \right) \times (v - v_0) + o(\|v - v_0\|). \end{aligned}$$

Отсюда

$$\nabla_v L(\bar{s} | v, v_0) \Big|_{v=v_0} = \frac{\nabla_v P(s_1 | v_0)}{P(s_1 | v_0)} + \frac{\nabla_v P(s_2 | s_1, v_0)}{P(s_2 | s_1, v_0)} + \dots + \frac{\nabla_v P(s_{r-1} | s_1, \dots, s_{r-2}, v_0)}{P(s_{r-1} | s_1, \dots, s_{r-2}, v_0)} + \sum_{t=r}^n \frac{\nabla_v P(s_t | s_{t-r+1}, \dots, s_{t-1}, v_0)}{P(s_t | s_{t-r+1}, \dots, s_{t-1}, v_0)}.$$

Последнее выражение можно преобразовать таким образом, чтобы градиент отношения правдоподобия был выражен через совместные вероятности порядков r и $r-1$:

$$\nabla_v L(\bar{s} | v, v_0) \Big|_{v=v_0} = \sum_{t=r}^n \frac{\nabla_v P(s_{t-r+1}, \dots, s_t | v_0)}{P(s_{t-r+1}, \dots, s_t | v_0)} - \sum_{t=r+1}^n \frac{\nabla_v P(s_{t-r+1}, \dots, s_{t-1} | v_0)}{P(s_{t-r+1}, \dots, s_{t-1} | v_0)}. \quad (7)$$

Перейдем к задаче определения градиентов $\nabla_v P(s_{t-r+1}, \dots, s_t | v) \Big|_{v=v_0}$ и $\nabla_v P(s_{t-r+1}, \dots, s_{t-1} | v) \Big|_{v=v_0}$. Для этого необходимо получить производные по каждому из входящих в v параметров. Несмотря на то что заданным значениям v соответствует целое множество распределений на $P(\bar{s} | v)$, следующая теорема показывает, что для произвольной непрерывно дифференцируемой параметризации распределения $P(\bar{s} | v)$ по параметрам μ и Q вид градиентов не зависит от способа параметризации. Поэтому они могут использоваться для дальнейших суждений о локальных изменениях функции правдоподобия.

Теорема 1. Пусть существует r -мерная непрерывная плотность распределения $f_t(x_1, \dots, x_r)$ случайных величин $\varepsilon_{t-r+1}, \dots, \varepsilon_t$, $t = \overline{r, n}$. Тогда для произвольной непрерывно дифференцируемой параметризации распределения $P(\bar{s} | v)$ по параметрам μ и Q имеет место

$$\begin{aligned} \nabla_{\theta} P(s_{t-r+1}, \dots, s_t | v) \Big|_{v=v_0} &= \sum_{i=1}^r [P_{r,i}(C_{s_{t-r+1}}, \dots, C_{s_t} | c_{s_{t-r+i-1}}) f_{t-r+i}(c_{s_{t-r+i-1}}) - \\ &\quad - P_{r,i}(C_{s_{t-r+1}}, \dots, C_{s_t} | c_{s_{t-r+i}}) f_{t-r+i}(c_{s_{t-r+i}})] \cdot \nabla_{\theta} g_{t-r+i}(\theta_0). \end{aligned} \quad (8)$$

$$\begin{aligned} \nabla_{\theta} P(s_{t-r+1}, \dots, s_{t-1} | v) \Big|_{v=v_0} &= \sum_{i=1}^{r-1} [P_{r-1,i}(C_{s_{t-r+1}}, \dots, C_{s_{t-1}} | c_{s_{t-r+i-1}}) f_{t-r+i}(c_{s_{t-r+i-1}}) - \\ &\quad - P_{r-1,i}(C_{s_{t-r+1}}, \dots, C_{s_{t-1}} | c_{s_{t-r+i}}) f_{t-r+i}(c_{s_{t-r+i}})] \cdot \nabla_{\theta} g_{t-r+i}(\theta_0). \end{aligned} \quad (9)$$

$$\nabla_{\mu} P(s_{t-r+1}(\theta_0, \mu_0), \dots, s_t(\theta_0, \mu_0) | v) \Big|_{v=v_0} = \sum_{i=1}^r \left[P_{r,i}(C_{s_{t-r+1}}, \dots, C_{s_t} | c_{s_{t-r+1}-1}) f_{t-r+i}(c_{s_{t-r+1}-1}) d_{s_{t-r+1}-1} - \right. \\ \left. - P_{r,i}(C_{s_{t-r+1}}, \dots, C_{s_t} | c_{s_{t-r+1}}) f_{t-r+i}(c_{s_{t-r+1}}) d_{s_{t-r+1}} \right]. \quad (10)$$

$$\nabla_{\mu} P(s_{t-r+1}(\theta_0, \mu_0), \dots, s_{t-1}(\theta_0, \mu_0) | v) \Big|_{v=v_0} = \sum_{i=1}^{r-1} \left[P_{r-1,i}(C_{s_{t-r+1}}, \dots, C_{s_{t-1}} | c_{s_{t-r+1}-1}) f_{t-r+i}(c_{s_{t-r+1}-1}) d_{s_{t-r+1}-1} - \right. \\ \left. - P_{r-1,i}(C_{s_{t-r+1}}, \dots, C_{s_{t-1}} | c_{s_{t-r+1}}) f_{t-r+i}(c_{s_{t-r+1}}) d_{s_{t-r+1}} \right]. \quad (11)$$

$$\nabla_Q P(s_{t-r+1}(\theta_0, \mu_0), \dots, s_t(\theta_0, \mu_0) | v) \Big|_{v=v_0} = [G]_{j_r(s_{t-r+1}, \dots, s_t)}', \quad (12)$$

$$\nabla_Q P(s_{t-r+1}(\theta_0, \mu_0), \dots, s_{t-1}(\theta_0, \mu_0) | v) \Big|_{v=v_0} = [F_{r-1}G]_{j_{r-1}(s_{t-r+1}, \dots, s_{t-1})}'. \quad (13)$$

Здесь $f_t(\cdot)$ – плотность распределения ε_t , $P_{r,i}(C_{s_{t-r+1}}, \dots, C_{s_t} | c_k)$ – условная вероятность попадания вектора $\|\varepsilon_s : s = \overline{t-r+1, t}; s \neq t-r+i\|$ в параллелепипед, образованный интервалами $\{C_{s_j} : j = \overline{t-r+1, t}; j \neq t-r+i\}$ при условии $\varepsilon_{t-r+i} = c_k$.

Доказательство. Для доказательства (8) обозначим $c_t(k | \mu_0, \theta_0, \theta) = c_k + d'_k \mu_0 - g_t(\theta) + g_t(\theta_0)$. Тогда

$$P(s_{t-r+1}, \dots, s_t | v) = \frac{c_t(s_t | \mu_0, \theta_0, \theta)}{c_t(s_t - 1 | \mu_0, \theta_0, \theta)} \int \dots \int \frac{c_{t-r+1}(s_{t-r+1} | \mu_0, \theta_0, \theta)}{c_{t-r+1}(s_{t-r+1} - 1 | \mu_0, \theta_0, \theta)} f_t(x_1, \dots, x_r) dx_1 \dots dx_r.$$

Из этого следует (8). Аналогично получаем (9).

Для получения градиентов по параметрам μ воспользуемся тем, что совместные вероятности полностью определяются параметрами Q . Поэтому если рассмотреть произвольную непрерывно дифференцируемую параметризацию конечномерных распределений по μ , то при любых v выполняется

$$\nabla_{\mu} P(s_{t-r+1}(\theta, \mu), \dots, s_t(\theta, \mu) | v) = \nabla_{\mu} P^{(r)}_{s_{t-r+1}(\theta, \mu), \dots, s_t(\theta, \mu)} = 0. \quad (14)$$

С другой стороны,

$$\nabla_{\mu} P(s_{t-r+1}(\theta, \mu), \dots, s_t(\theta, \mu) | v) \Big|_{v=v_0} = P'_{\mu}(s_{t-r+1}(\theta, \mu), \dots, s_t(\theta, \mu) | v) \Big|_{v=v_0} - \\ - \sum_{i=1}^r \left[P_{r,i}(C_{s_{t-r+1}}, \dots, C_{s_t} | c_{s_{t-r+1}-1}) f_{t-r+i}(c_{s_{t-r+1}-1}) d_{s_{t-r+1}-1} - \right. \\ \left. - P_{r,i}(C_{s_{t-r+1}}, \dots, C_{s_t} | c_{s_{t-r+1}}) f_{t-r+i}(c_{s_{t-r+1}}) d_{s_{t-r+1}} \right], \quad (15)$$

где P'_{μ} – градиент по параметризации самого распределения без учета зависимости аргументов от μ . Нас интересует градиент при гипотетических параметрах $\nabla_{\mu} P(s_{t-r+1}(\theta_0, \mu_0), \dots, s_t(\theta_0, \mu_0) | v) \Big|_{v=v_0} = P'_{\mu}(s_{t-r+1}(\theta_0, \mu_0), \dots, s_t(\theta_0, \mu_0) | v) \Big|_{v=v_0}$, который благодаря непрерывной дифференцируемости параметризации семейства распределений равен $P'_{\mu}(s_{t-r+1}(\theta, \mu), \dots, s_t(\theta, \mu) | v) \Big|_{v=v_0}$. Поэтому из (14) и (15) получаем (10). Аналогично можно получить (11) для градиента правдоподобия меньшего порядка.

Воспользуемся теперь произвольной параметризацией семейства распределений по Q :

$$\nabla_Q P(s_{t-r+1}(\theta, \mu), \dots, s_t(\theta, \mu) | v) \Big|_{v=v_0} = \nabla_Q P^{(r)}_{s_{t-r+1}(\theta, \mu), \dots, s_t(\theta, \mu)} \Big|_{v=v_0} = [G]_{j_r(s_{t-r+1}, \dots, s_t)}'. \quad (16)$$

Нас интересует градиент $\nabla_Q P(s_{t-r+1}(\theta_0, \mu_0), \dots, s_t(\theta_0, \mu_0) | v) \Big|_{v=v_0}$, который совпадает с (16), поскольку параметризация непрерывно дифференцируема. Отсюда имеем (12) и аналогично получаем (13).

Теорема доказана.

3. Знаковые тесты

Формально искомый тест для проверки гипотез (3) может иметь вид

$$\left\| \nabla_v L(\bar{s} | v, v_0) \Big|_{v=v_0} \right\|^2 > \text{const}, \quad (17)$$

где $\|\cdot\|$ – некоторая подходящая норма вектора, а формулы (7)–(13) дают выражение для тестовой статистики. Нетрудно убедиться, что если выбрать $p_1 = \dots = p_K = 1/K$ и гипотетические параметры Q_0 определяют равномерное r -мерное распределение (т.е. $[Q_0]_i = K^{-r}$ и гипотеза состоит в том, что зависимость отсутствует), то тест (17) будет локально наиболее мощным против любой линейной одномерной односторонней альтернативы, поскольку знаменатель отношения правдоподобия $P(\bar{s} | v_0)$ превращается в константу. В остальных случаях приходится опираться на логическую обоснованность принципа максимума отношения правдоподобия.

В качестве альтернативы для (17) можно рассматривать тест вида

$$\left\| \nabla_v P(\bar{s} | v) \Big|_{v=v_0} \right\|^2 = \left\| P(\bar{s} | v_0) \cdot \nabla_v L(\bar{s} | v, v_0) \Big|_{v=v_0} \right\|^2 > \text{const}, \quad (18)$$

который является локально наиболее мощным против любой линейной одномерной односторонней альтернативы.

Относительно тестов (17) и (18) остаются вопросы, на которые необходимо дать ответ. Во-первых, в (8)–(11) присутствуют неизвестные величины, которые требуют обоснованной замены. Во-вторых, необходимо указать, каким образом можно определять критические значения и какие нормы вектора при этом использовать.

Начнем с ответа на первый вопрос. Чтобы избавиться от неизвестных величин, входящих в состав (17), используем метод, аналогичный [7], – заменим их доступными нам величинами. Так, игнорируя некоторые эффекты зависимости, можно заменить

$$P_{r,i}(C_{s_{t-r+1}}, \dots, C_{s_t} | c_k) \text{ на } P^{(r,i)}(k_1, \dots, k_r) = \sum_{s=1}^K P^{(r)}(k_1, \dots, k_{i-1}, s, k_{i+1}, \dots, k_r).$$

После этого неизвестными останутся величины

$$(f_t(c_{k-1}) - f_t(c_k)) / p_k \text{ и } (f_t(c_{k-1})d_{k-1} - f_t(c_k)d_k) / p_k,$$

замена которых специальными весами $B_1(k)$ и $B_2(k)$ подробно обсуждается в [7]. Например, для $K = 2$ (квантильная регрессия) $B_1 = \{-1/p, 1/(1-p)\}$. При $K = 3$ (симметричная двухквантильная регрессия) $B_1 = \{-1, 0, 1\}$, $B_2 = \{1, -2p/(1-2p), 1\}$. Для $K = 3$ (симметричная трехквантильная регрессия) $B_1 = \{-A, -\alpha, \alpha, A\}$, $B_2 = \{A, -1, -1, A\}$, где $A = (1-2p)/2p$, α – априорная догадка о величине $(f_t(0) - f_t(\mu)) / f_t(\mu)$.

Заметим, что похожие «разумные» замены часто встречаются при конструировании непараметрических процедур. В ранговом анализе полученные похожим образом веса рангов называют метками (scores). Важными свойствами меток являются их нулевые средние $\sum_{k=1}^K B_1(k) p_k = 0$, $\sum_{k=1}^K B_2(k) p_k = 0$ и нулевая ковариация $\sum_{k=1}^K B_1(k) B_2(k) p_k = 0$.

В результате замен из (17) и (18) получаем тесты вида

$$\left\| \xi_n(\bar{s} | v_0) \right\|^2 > \text{const}, \quad (19)$$

$$\left\| P(\bar{s} | v_0) \cdot \xi_n(\bar{s} | v_0) \right\|^2 > \text{const}, \quad (20)$$

где $\xi_n(\bar{s} | v_0)$ – векторная статистика, которая является модифицированным и нормированным градиентом отношения правдоподобия и задается выражениями

$$\xi_n(\bar{s} | v) = n^{-1/2} \sum_{t=r}^n \sum_{i=1}^r W_{t,i}(\bar{s} | v),$$

$$W_{t,i}(\bar{s}|\mathbf{v}) = \begin{bmatrix} \nabla_{\theta} g_{t-r+i}(\theta) R_{t,1}^{(r,i)}(\mathbf{v}) \\ R_{t,2}^{(r,i)}(\mathbf{v}) \\ \frac{[G]_{j_r(s_{t-r+1}(\mathbf{v}), \dots, s_t(\mathbf{v}))}' - (1-\delta_{t,r}) \frac{[F_{r-1}G]_{j_{r-1}(s_{t-r+1}(\mathbf{v}), \dots, s_{t-1}(\mathbf{v}))}' }{P^{(r-1)}(s_{t-r+1}(\mathbf{v}), \dots, s_{t-1}(\mathbf{v}))}}}{P^{(r)}(s_{t-r+1}(\mathbf{v}), \dots, s_t(\mathbf{v}))} \end{bmatrix}, \quad (21)$$

$$R_{t,l}^{(r,i)}(\mathbf{v}) = B_l(s_{t-r+i}) \left(L_t^{(r,i)}(\mathbf{v}) - (1-\delta_{t,r})(1-\delta_{i,r}) L_{t-1}^{(r-1,i)}(\mathbf{v}) \right),$$

$$L_t^{(r,i)}(\mathbf{v}) = \frac{P^{(r,i)}(s_{t-r+1}(\mathbf{v}), \dots, s_t(\mathbf{v})) P^{(1)}(s_{t-r+i}(\mathbf{v}))}{P^{(r)}(s_{t-r+1}(\mathbf{v}), \dots, s_t(\mathbf{v}))}.$$

Заметим, что тест (20), даже после замены неизвестных величин весами, остается локально наиболее мощным против линейных односторонних одномерных альтернатив, если $[Q_0]_i = K^{-r}$, но при дополнительном условии, что эти альтернативы отличаются от гипотезы только по параметрам θ , только по μ или только по Q . Однако имитационные эксперименты показывают, что поведение этого теста при альтернативах не всегда удовлетворительно. Кроме того, возникают трудности с поиском его асимптотических критических значений. Поэтому для использования мы рекомендуем тест (19)

Перейдем к ответу на второй из поставленных выше вопросов – о выборе критических значений и способов определения нормы векторных статистик. Как и для случая независимых наблюдений [1, 7], для проверки гипотез (3) можно построить тесты вида (19) и (20) с точным уровнем значимости, поскольку при гипотезе распределение статистик $\|\xi_n(\bar{s}|\mathbf{v}_0)\|^2$ и $\|P(\bar{s}|\mathbf{v}_0) \cdot \xi_n(\bar{s}|\mathbf{v}_0)\|^2$ совпадает с распределением случайных величин $\|\xi_n(\bar{\eta}|\mathbf{v}_0)\|^2$ и $\|P(\bar{\eta}|\mathbf{v}_0) \cdot \xi_n(\bar{\eta}|\mathbf{v}_0)\|^2$ соответственно, где случайный вектор $\bar{\eta} = (\eta_1, \dots, \eta_n)'$ составлен из последовательности случайных величин, образующих стационарный Марковский процесс $(r-1)$ -го прядка с конечным числом состояний $(\eta_t \in \{1, 2, \dots, K\})$ и известными конечномерными вероятностями $P(k_1, \dots, k_r|\mathbf{v}_0)$. В результате процентные точки распределения статистики тестов (19) и (20) можно определить с любой точностью, используя метод Монте-Карло. Заметим, что в вычислении $\xi_n(\bar{\eta}|\mathbf{v}_0)$ не участвуют параметры μ_0 , параметры θ_0 используются только при вычислении градиентов $\nabla_{\theta} g_t(\theta_0)$ в (21), а параметры Q_0 влияют только на величины вероятностей $P^{(r)}$, $P^{(r-1)}$, $P^{(r,i)}$, $P^{(r-1,i)}$.

При больших объемах наблюдений n лучше воспользоваться асимптотическими критическими значениями. Для теста (19) их можно вычислять на основе асимптотической нормальности распределения статистики $\xi_n(\bar{s}|\mathbf{v}_0)$ при гипотезе. Используем для этого обобщение теоремы 7.7.9 из [9] на случай последовательности случайных векторов, которое легко получить, применяя теорему 7.7.7 из [9].

Лемма 1. Пусть числовая последовательность $\{a_t : t \geq 1\}$ и последовательность случайных векторов $\{z_t : t \geq 1\}$ удовлетворяют следующим условиям:

1. Существует целое число $m > 0$ такое, что для любых n и t_1, \dots, t_n ($0 < t_1 < \dots < t_n$) совокупности случайных величин $\{z_{t_1}, \dots, z_{t_n}\}$ и $\{z_1, \dots, z_{t_1-m-1}, z_{t_n+m+1}, \dots\}$ взаимно независимы.

2. $Mz_t = 0$, $t = 1, 2, \dots$

3. $M\|z_t\|^{2+\delta} < M$ для некоторого M и $\delta > 0$, $t = 1, 2, \dots$

4. $|a_t| < L$ для всех t и некоторого $L > 0$.

5. Существует предельная матрица $\Sigma = \lim_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T \sum_{s=1}^T a_t a_s' Mz_t z_s'$.

Тогда случайный вектор $T^{-1/2} \sum_{t=1}^T a_t z_t$ сходится по распределению к $N(0, \Sigma)$.

Теорема 2. Пусть выполняется гипотеза (3) и следующие условия:

1. $\|\nabla_{\theta} g_t(\theta)\| < L$ для всех t и некоторого $L > 0$.

2. Существует предельная матрица $V = \lim_{n \rightarrow \infty} V_n$, где $V_n = M \xi_n(\bar{s} | v) \xi_n'(\bar{s} | v)$.

3. $P^{(r)}(k_1, \dots, k_r) > 0$ при всех k_1, \dots, k_r .

Тогда случайный вектор $\xi_n(\bar{s} | v)$ сходится по распределению к $N(0, V)$, а случайная величина $\zeta_n^2 = \xi_n'(\bar{s} | v) V_n^{-1} \xi_n(\bar{s} | v)$ сходится по распределению к χ_q^2 , где $q = \dim v$.

Доказательство. Если в лемме 1 положить $a_t z_t = \sum_{i=1}^r W_{t+r-1,i}(\bar{s} | v_0)$, $m = 2r$, то доказательство сводится к проверке равномерной (по t) ограниченности элементов вектора (21), а также проверке свойства $M \xi_n(\bar{s} | v) = 0$. Последнее сразу следует из того, что $\sum_{k=1}^K B_1(k) p_k = 0$, $\sum_{k=1}^K B_l(k) p_k = 0$, если применить эти равенства к вычислению $MR_{l,i}^{(r,i)}(v)$, $l \in \{1, 2\}$, в первом и втором блоке вектора (21). Для третьего блока вопрос сводится к проверке равенства $\sum_{k_1, \dots, k_r} [G]_{j_r(k_1, \dots, k_r)}' = 0$ и $\sum_{k_1, \dots, k_{r-1}} [F_{r-1} G]_{j_{r-1}(k_1, \dots, k_{r-1})}' = 0$. Мы проверяли эти условия численно в большом количестве частных случаев при $r \in \{2, \dots, 5\}$, $K \in \{2, 3, 4\}$, при выполнении условий центральной симметрии k -мерных вероятностей и $k \in \{2, \dots, r\}$ без них.

Теорема доказана.

Таким образом, для ограниченных сходящихся планов в асимптотическом тесте можно использовать статистику $\zeta_n^2 = \xi_n'(\bar{s} | v_0) V_n^{-1} \xi_n(\bar{s} | v_0)$ с критическими значениями распределения хи-квадрат. Для элементов матрицы V_n можно получить аналитические выражения, однако мы не приводим их в силу громоздкости. Отметим только, что матрица V_n зависит от параметров θ и Q , но не от μ , а для линейной модели регрессии остается зависимость только от Q .

Заключение

В данной работе получены точные и асимптотические знаковые тесты для проверки простой гипотезы (3) $H_0: v = v_0$ о параметрах модели многоквантовой регрессии (1) с Марковскими ошибками (1) порядка $r-1$. Рассмотрены наиболее интересные случаи одно-, двух- и трехквантовой регрессии.

В теореме 1 показано, что несмотря на непараметрическую постановку задачи, можно получить выражения для градиента правдоподобия знаковых признаков $\nabla P(\bar{s} | v)$, вид которого не зависит от способа параметризации распределения $P(\bar{s} | v)$. Этот факт позволяет получить тесты (19) и (20), основанные на векторной статистике $\xi_n(\bar{s} | v_0)$, см. (21). Эти тесты являются локально наиболее мощными против линейных односторонних одномерных альтернатив при условии равномерного r -мерного распределения процесса знаковых признаков.

В работе рассмотрены вопрос о вычислении критических значений, обеспечивающих заданный уровень значимости с любой точностью при конечных выборках, а также критические значения, основанные на асимптотическом распределении тестовой статистики (теорема 2) для ограниченных сходящихся планов. В результате рекомендуется тест, интерпретируемый как тест, построенный по принципу максимального отношения правдоподобия, с критической областью вида

$$\zeta_n^2 = \xi_n'(\bar{s} | v_0) V_n^{-1} \xi_n(\bar{s} | v_0) > \text{const},$$

где $V_n = V_n(\theta, Q) = M \xi_n(\bar{s} | v) \xi_n'(\bar{s} | v)$, для которого асимптотическое распределение тестовой статистики является распределением хи-квадрат и не зависит от гипотезы.

Полученный тест представляет интерес как основа для оценивания параметров v по принципу максимального асимптотического p -значения (см. основополагающую работу [8]), т.е.

$$v_n = \arg \min_{v_0} \zeta_n^2(v_0).$$

Заметим, что на практике, особенно для линейных моделей, представляет интерес проверка гипотез вида $H_0: [\theta]_j = [\theta_0]_j$, когда остальные параметры являются мешающими. Такие гипотезы можно рассматривать как линейные и применять для их проверки двухэтапные процедуры. Этот подход подробно описан в [1, 7] применительно к знаковым процедурам.

ЛИТЕРАТУРА

1. Болдин М.В., Симонова Г.И., Тюрин Ю.Н. Знаковый статистический анализ линейных моделей. М.: Наука. Физматлит, 1997. 285 с.
2. Koenker R., Bassett G. Regression quantiles // *Econometrica*. 1978. V. 46. P. 349–359.
3. Chen L.A., Tran L.T., Lin L.C. Symmetric regression quantile and its application to robust estimation for the nonlinear regression model // *J. Statist. Plann. And Infer.* 2004. V. 126. P. 423–440.
4. Coudin E., Dufour J. Finite-sample distribution-free inference in linear median regressions under heteroscedasticity and non-linear dependence of unknown form // *Econometrics Journal*. 2009. V. 12. P. 19–49.
5. Tarassenko P.F., Tarima S.S., Zhuravlev A.V., Singh S. On sign-based regression quantiles // *J. of Statist. Comput. and Simul.* 2015. V. 85, No. 7. P. 1420–1441.
6. Dufour J., Kiviet J. Exact inference methods for first-order autoregressive distributed lag models // *Econometrica*. 1998. V. 82. P. 79–104.
7. Тарасенко П.Ф. Индикаторный статистический анализ. Томск: Изд-во Том. ун-та, 2005. 350 с.
8. Hodges J.L.Jr., Lehmann E.L. Estimates of location based on rank tests // *Ann. Math. Statist.* 1963. V. 34, No. 2. P. 528–611.
9. Андерсон Т. Статистический анализ временных рядов. М.: МИР, 1976. 756 с.

Сиренко Михаил Андреевич. E-mail: sma@mail.tsu.ru

Тарасенко Петр Феликсович, канд. физ.-мат. наук, доцент, E-mail: ptara@mail.tsu.ru
Томский государственный университет

Поступила в редакцию 28 апреля 2015 г.

Sirenko Mikhail A., Tarassenko Petr F. (Tomsk State University, Russian Federation).

Finite-sample and asymptotic sign-based tests for parameters of non-linear quantile regression with Markov noise.

Keywords: sign-based statistical procedures; non-linear regression; quantile regression; Markov process; finite-sample test.

DOI 10.17223/19988605/32/7

Sign-based statistical procedures are known to be more robust to outliers than least squares and ability to control precise significance level for finite samples when testing simple hypothesis. In this paper, the sign-based approach is extended to the case of non-linear model with dependent noise. The model of multi-quantile regression is considering, which allows to test hypotheses both on the parameters of the regression $y_t = g_t(\theta) + \varepsilon_t$, $t = \overline{1, n}$, and the parameters of the noise ε_t , which forms a stationary Markov process of order $(r-1)$. According to that, the signs of observations are calculated with respect to a set of quantiles (at specified levels) of one-dimensional distribution of ε_t .

The quantiles depend on the unknown parameter μ . For example, in a symmetric two-quantile regression, the quantiles $(-\mu)$ and μ correspond to levels of p and $1-p$, where p is given, and μ is unknown. For the three-quantile regression, the quantiles $(-\mu)$, 0 , μ correspond to levels of p , $1/2$ and $1-p$. In both cases, the parameter μ is a scale parameter of noise.

According to the sign-based approach, the residuals are substituting with indicators of their belonging to the interquantile intervals $\bar{s} = (s_1, \dots, s_n)$, where s_t takes a finite number of values. The unknown parameters in this scheme are $v = (\theta', \mu', Q')'$, where the vector Q contains linearly independent r -dimensional joint probabilities of the states of generated by process of indicators $\{s_t\}$. Since the problem is considering in a nonparametric setting, then each fixed value of parameters μ and Q corresponds to a class of finite-dimensional distributions of the initial process ε_t . However, we can show that for any continuous parameterization of finite-dimensional distributions, all the derivatives of the likelihood of indicators $P(\bar{s} | v)$ are expressed in the same way. In the problem of testing a simple hypothesis $H_0: v = v_0$ it gives the opportunity to build a test based on the principle of maximal likelihood ratio.

In this paper, we consider the problem of calculating the critical values to provide the desired significance level with any accuracy for finite samples, as well as the critical values based on the asymptotic distribution of the test statistic.

The obtained tests can be used as a basis for estimating the parameters v by the principle of maximal p -values, as well as for the development of tests for linear hypothesis.

REFERENCES

1. Boldin, M.V., Simonova, G.I. & Tyurin, Yu.N. (1997) *Znakovyy statisticheskiy analiz lineynykh modeley* [Sign-based Methods in Linear Statistical Models]. Moscow: Nauka. Fizmatlit.
2. Koenker, R. & Bassett, G. (1978) Regression quantiles. *Econometrica*. 46. pp. 349-359.
3. Chen, L.A., Tran, L.T. & Lin, L.C. (2004) Symmetric regression quantile and its application to robust estimation for the nonlinear regression model. *Journal of Statistical Planning and Inference*. 126. pp. 423-440. DOI: 10.1016/j.jspi.2003.09.014
4. Coudin, E. & Dufour, J. (2009) Finite-sample distribution-free inference in linear median regressions under heteroscedasticity and non-linear dependence of unknown form. *Econometrics Journal*. 12. pp. 19-49. DOI: 10.1111/j.1368-423X.2009.00285.x
5. Tarassenko, P.F., Tarima, S.S., Zhuravlev, A.V. & Singh, S. (2015) On sign-based regression quantiles. *Journal of Statistical Computation and Simulation*. 85 (7). pp. 1420-1441. DOI: 10.1080/00949655.2013.875176
6. Dufour, J. & Kiviet, J. (1998) Exact inference methods for first-order autoregressive distributed lag models. *Econometrica*. 82. pp. 79-104. DOI: 10.2307/2998541
7. Tarasenko, P.F. (2005) *Indikatornyy statisticheskiy analiz* [Indicator-based statistical analysis]. Tomsk: Tomsk State University.
8. Hodges, J.L.Jr. & Lehmann, E.L. (1963) Estimates of location based on rank tests. *Ann. Math. Statist.* 34 (2). pp. 528-611.
9. Anderson, T.W. (1971) *The statistical analysis of time series*. John Wiley & Sons, Inc.