Б. Е. Бродский

Структурные сдвиги и единичные корни: различение моделей нестационарности временных рядов

В статье рассмотрена проблема тестирования гипотез о статистической нестационарности (структурный сдвиг, единичный корень) в моделях одномерных временных рядов. Предложен новый метод различения гипотез неизвестного структурного сдвига и единичного корня и исследованы его свойства для модели зависимых наблюдений. Доказана теорема о сходимости к нулю вероятностей ошибочных решений предложенного метода с увеличением объема выборки. Свойства метода изучены в имитационном эксперименте с выборками зависимых наблюдений. Рассмотрены практические приложения метода к задачам тестирования гипотез структурного сдвига и единичного корня в моделях эконометрических временных рядов.

1. Введение

Одной из наиболее сложных проблем современной эконометрической теории нестационарных временных рядов является различение моделей структурного сдвига и единичного корня. Как замечает Перрон [Perron (2005)] в обзоре современного состояния этой теории, присутствие структурного сдвига резко снижает мощность известных тестов на единичные корни (например, ADF (Augmented Dickey-Fuller test) [Dickey, Fuller (1979, 1981)] и KPSS [Kwiatkowski, Phillips, Schmidt, Shin (1992)]), и наоборот, тестирование структурного сдвига в моделях с единичным корнем крайне затруднено. По мнению Перрона, это происходит из-за того, что большинство статистических тестов, предназначенных для решения задач обнаружения единичного корня и структурного сдвига в эконометрических данных, обладают выраженной статистической «сенситивностью» к этим видам нестационарности наблюдений. Проблема в том, чтобы предложить тесты, избирательно реагирующие на тот или иной тип нестационарности данных, или тесты, «проявляющие» различные «отпечатки» этих видов нестационарности. На сегодняшний день такие «сенситивные тесты» не разработаны (см. [Perron (2005)]).

Вместе с тем проблема различения указанных видов нестационарности наблюдений продолжает вызывать значительный исследовательский интерес. По сути дела, она стала актуальной уже после знаменитой статьи [Nelson, Plosser (1982)], в которой на 14 макроэкономических временных рядах было продемонстрировано, что ADF-тест не позволяет отвергнуть гипотезу о наличии стохастического тренда в данных. Эти выводы были подвергнуты критике в работе [Rappoport, Reichlin (1989)] и особенно в статьях [Perron (1989, 1994)], в которых было показано, что присутствие структурного сдвига в данных чаще всего не позволяет эффективно диагностировать наличие стохастического тренда в полученной выборке.

Работы Перрона [Perron (1989, 1994)] базировались на следующей спецификации регрессионной модели временного ряда:

$$y_t = \mu + \beta t + \alpha y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-i} + \varepsilon_t .$$

Если в данных присутствует структурный сдвиг по уровню (μ), то оценка авторегрессионного коэффициента α в асимптотике сдвинута к единице. Это означает, что ADF-тест с модифицированным порогом по работе [MacKinnon (1996)] будет диагностировать наличие стохастического тренда в данных.

В последующих работах [Montanes, Reyes (1999, 2000)] было показано, что аналогичные проблемы возникают с тестом Филлипса–Перрона [Phillips, Perron (1988)] на единичные корни.

Ситуация еще более усложняется, если моменты структурных сдвигов в выборке данных неизвестны. В работах [Christiano (1992)], [Banerjee et al. (1992)], [Zivot, Andrews (1992)], [Perron, Vogelsang (1992)], [Perron (1997)] было показано, что в этом случае мощность известных тестов на единичные корни резко снижается. Некоторые предварительные подходы к решению этой проблемы были намечены в современных работах [Perron, Yabu (2005)], [Kim, Perron (2005)].

В целом следует признать, что проблема эффективного различения нестационарностей типа единичного корня и структурного сдвига не получила окончательного решения в общем контексте неизвестных моментов структурных сдвигов и статистической зависимости наблюдений.

В данной работе исследуется статистическая эффективность **двухшаговой процедуры различения нестационарности стохастического тренда и единичного корня**, основанной на модифицированной статистике Колмогорова–Смирнова. На первом шаге тестируется *гипотеза стационарности выборки* против совместной альтернативы нестационарности наблюдений (структурный сдвиг или единичный корень). На втором шаге в случае принятия гипотезы нестационарности выборки тестируется *гипотеза наличия структурного сдвига* (нескольких структурных сдвигов) против альтернативы единичного корня. Показано, что вероятности ошибочной классификации этих видов нестационарности данных в рамках предложенной процедуры экспоненциально стремятся к нулю с увеличением объема выборки.

2. Постановка задачи

Рассмотрим следующие модели наблюдений на вероятностном пространстве $(\Omega, \mathbb{F}, P)$:

- H_0: *модель стационарного временного ряда*:

$$e_i = \rho e_{i-1} + u_i, \quad e_0 = 0, \quad |\rho| < 1, \quad i = 1, \ldots, N, \tag{1}$$

где u_i — последовательность независимых одинаково распределенных наблюдений с $Eu_i = 0$, $Eu_i^2 = \sigma^2$ (здесь E обозначает математическое ожидание по мере P).

Процесс наблюдений:

$$y_i = e_i .$$

Далее будем предполагать, что для последовательности u_i выполнено *равномерное условие Крамера*: существует константа H такая, что для всех i

$$Ee^{tu_i} < \infty, \ |t| < H.$$

В силу центрированности последовательности u_i это условие эквивалентно следующему: существуют числа $T > 0$, $g > 0$ такие, что для всех i

$$Ee^{tu_i} < \exp\left(\frac{1}{2}\sigma^2 g t^2\right), \ |t| < T; \qquad (2)$$

- H_1: *модель структурного сдвига*: процесс «состояния» e_i, $i = 1, 2, \ldots$, описывается уравнением (1), а процесс наблюдений равен

$$y_i = hI\{i \geq [\theta N]\} + e_i,$$

где $I(A)$ — характеристическая функция множества A; $0 < \alpha \leq \theta \leq 1 - \alpha < 1$, $h \neq 0$, а параметр структурного сдвига θ неизвестен;

- H_2: *модель единичного корня*:

$$e_i = e_{i-1} + u_i, \ \ e_0 = 0, \ \ i = 1, \ldots, N,$$
$$y_i = e_i.$$

Основная статистика, которая будет использоваться для тестирования гипотез H_0, H_1, H_2, имеет вид

$$Z_N(n) = \frac{1}{N^2}\left(n\sum_{j=1}^{N} y_j - N\sum_{j=1}^{n} y_j\right), \ n = 1, \ldots, N.$$

Зафиксируем параметр $0 < \alpha < 1/2$. Решающее правило, основанное на статистике

$$T_N = \max_{[\alpha N] \leq n \leq [(1-\alpha)N]} |Z_N(n)|$$

и пороговой границе $C > 0$, имеет вид

$$d_N = \begin{cases} 0, & \text{если } T_N \leq C \text{ — принимается } H_0, \\ 1, & \text{если } T_N > C \text{ — принимается } H_1 \vee H_2. \end{cases}$$

На этом заканчивается первый этап метода. На втором этапе происходит различение альтернатив H_1 и H_2.

В ситуации единственного момента структурного сдвига $n_0 = [\theta N]$ тестирование гипотез H_1, H_2 осуществляется следующим образом:

1. Пусть на первом шаге метода выполнено неравенство $T_N > C$ и максимум модуля статистики $Z_N(n)$ достигается в точке $\hat{n}_1$ (в случае нескольких точек максимума $Z_N(n)$ выбирается наименьшая из них).
2. На втором шаге фиксируется параметр $0 < \delta < \alpha$ и формируются две подвыборки: $X_1(\hat{n}_1) = \{y_1, \ldots, y_{\hat{n}_1 - [\delta N]}\}$ и $X_2(\hat{n}_1) = \{y_{\hat{n}_1 + [\delta N]}, \ldots, y_N\}$.
3. Для каждой из этих подвыборок рассчитываются соответствующие величины $T_N(X_1)$ и $T_N(X_2)$ — максимумы модуля статистики $|Z_N(\cdot)|$, построенной по этим подвыборкам.

4. Для подвыборок X_1 и X_2 рассчитываются пороговые границы C_1 и C_2 (расчетные формулы приведены в четвертом разделе).

5. Решающее правило на втором шаге метода таково:

$$d_2(N)=\begin{cases}1, & T_N(X_1)<C_1 \text{ или } T_N(X_2)<C_2,\\ 2 & \text{в противном случае.}\end{cases}$$

При $d_2(N)=1$ принимается гипотеза H_1, при $d_2(N)=2$ — гипотеза H_2.

Таким образом, первый этап предлагаемого метода состоит в тестировании гипотезы стационарности наблюдений против совместной альтернативы единичного корня и стохастического тренда. При отклонении гипотезы стационарности выполняется второй этап метода, на котором тестируется гипотеза структурного сдвига против альтернативы единичного корня.

3. Основные результаты

Качество методов тестирования моделей целесообразно оценивать вероятностями ошибок $P_i(H_j)$, $i,\ j=0,1,2,\ i\neq j$. Здесь $P_i(H_j)$ обозначает вероятность принятия гипотезы H_j, когда в действительности верна H_i.

На первом шаге процедуры тестируется *гипотеза стационарности* H_0 против совместной альтернативы $H_1\vee H_2$, поэтому здесь используются вероятности ошибок $P_0(H_1\vee H_2)$, $P_1(H_0)$, $P_2(H_0)$.

Справедлива следующая теорема.

Теорема 1. *Пусть* $0<C<\sigma|h|\alpha^2$. *Тогда вероятности ошибок* $P_0(H_1\vee H_2)$, $P_1(H_0)$, $P_2(H_0)$ *стремятся к нулю с увеличением объема выборки* N.

Более того, при гипотезе H_1 *нормированная оценка момента структурного сдвига* $\hat{n}/N$ *сходится почти наверное к истинному параметру структурного сдвига* θ *при увеличении объема выборки* N.

Доказательство. Рассмотрим свойства статистики $Z_N(\cdot)$ при различных альтернативах. Вначале запишем для гипотезы H_0:

$$e_i=\rho(\rho e_{i-2}+u_{i-1})+u_i=\ldots=\sum_{k=0}^{i-1}u_{i-k}\rho^k.$$

Поэтому

$$\sum_{j=1}^{n}e_j=\sum_{j=1}^{n}\sum_{k=0}^{j-1}u_{j-k}\rho^k=\sum_{m=1}^{n}u_m\sum_{k=0}^{n-m}\rho^k.$$

При гипотезе H_0 статистика $Z_N(i)$ имеет вид

$$Z_N(i)=\left(i\sum_{m=1}^{N}u_m\frac{1-\rho^{N-m+1}}{1-\rho}-N\sum_{m=1}^{i}u_m\frac{1-\rho^{i-m+1}}{1-\rho}\right)\Bigg/N^2=$$

$$=\frac{1}{(1-\rho)N^2}\left[\sum_{m=1}^{i}u_m\left[i(1-\rho^{N-m+1})-N(1-\rho^{i-m+1})\right]+\sum_{m=i+1}^{N}u_m i(1-\rho^{N-m+1})\right].$$

В силу независимости u_i

$$E_0 Z_N^2(i) \le \frac{\sigma^2}{(1-\rho)N}\left(\sum_{m=1}^{i}(N-i)^2 + \sum_{m=i+1}^{N} i^2\right) = \sigma^2 \frac{i(N-i)}{N^3(1-\rho)^2},$$

поэтому

$$\max_i E_0 Z_N^2(i) \le O\left(\frac{1}{N}\right).$$

Отсюда, используя условие Крамера для последовательности u_i, получим

$$P_0\{H_1 \vee H_2\} = P_0\left\{\max_n |Z_N(n)| > C\right\} \le \begin{cases} \exp\left(-\dfrac{(1-\rho)^2 C^2 N}{2\sigma^2 g}\right), & 0 < C < NTg\sigma^2, \\ \exp\left(-\dfrac{(1-\rho)CNT}{2}\right), & C \ge NTg\sigma^2. \end{cases}$$

При гипотезе H_2 статистика $Z_N(i)$ принимает вид

$$Z_N(i) = \left[i\sum_{m=1}^{N} u_m(N-m+1) - N\sum_{m=1}^{i} u_m(i-m+1)\right]\Big/ N^2 = \left[\sum_{m=1}^{i} u_m(N-i)(m-1) + \sum_{m=i+1}^{N} u_m i(N-m+1)\right]\Big/ N^2.$$

Поэтому при гипотезе H_2 в силу независимости u_i имеем

$$E_2 Z_N^2(i) = \frac{\sigma^2}{N^4}\left[\sum_{m=1}^{i}(N-i)^2(m-1)^2 + \sum_{m=i+1}^{N} i^2(N-m+1)^2\right] =$$

$$= \frac{\sigma^2}{N^4}\left[(N-i)^2\sum_{k=0}^{i-1} k^2 + i^2\sum_{k=1}^{N-i} k^2\right] = \frac{i(N-i)}{6N^3}(2i(N-i)+1)\sigma^2.$$

Поэтому для любого $t \in (0, 1)$ существует предел

$$\lim_{N\to\infty} \frac{1}{[Nt]} E_2 Z_N^2([Nt]) = \frac{t(1-t)^2}{3}\sigma^2 = \sigma_t^2.$$

С использованием результата работы [Peligrad (1992)] имеем

$$U = \frac{1}{\sigma_t \sqrt{N}} Z_N([Nt]) \to W(t),$$

где $W(t)$ — винеровский процесс на (0, 1).

Отсюда получим для вероятности ошибки

$$P_2\{H_0\} = P_2\left\{\max_{[\alpha N]\le n\le[(1-\alpha)N]} |Z_N(n)| < C\right\} \le P\left\{\max_{[\alpha N]\le n\le[(1-\alpha)N]} \frac{1}{\sigma_t\sqrt{N}} |Z_N(n)| < \frac{C}{\alpha\sqrt{N}}\left(\frac{3}{1-\alpha}\right)^{1/2}\right\} \to 0 \ \text{ при } \ N\to\infty.$$

При гипотезе H_1 статистика $Z_N(i)$ равна

$$Z_N(i) = a_\theta\left(\frac{i}{N}\right) + \eta_N(i),$$

где

$$a_\theta(t)=\begin{cases} t(1-\theta)h, & 0\le t\le\theta,\\ \theta(1-t)h, & \theta<t\le 1;\end{cases}$$

$$\eta_N(i)=\frac{1}{(1-\rho)N^2}\left[\sum_{m=1}^{i}u_m\left[i(1-\rho^{N-m+1})-N(1-\rho^{i-m+1})\right]+\sum_{m=i+1}^{N}u_m i(1-\rho^{N-m+1})\right].$$

Поэтому при $C<\sigma|h|\theta(1-\theta)$ получим для вероятности ошибки

$$P_1\{H_0\}=P_1\left\{\max_i|Z_N(i)|\le C\right\}\le P_1\left\{\max_i|\eta_N(i)|\ge\sigma|h|\theta(1-\theta)-C\right\}.$$

Обозначим $\kappa=\sigma|h|\theta(1-\theta)-C$. Заметим, что для последней вероятности справедлива экспоненциальная оценка сверху, аналогичная доказанной ранее для гипотезы H_0:

$$P_1\left\{\max_i|\eta_N(i)|\ge\kappa\right\}\le\begin{cases}\exp\left(-\dfrac{(1-\rho)^2\kappa^2N}{2\sigma^2g}\right), & 0<\kappa<NTg\sigma^2,\\ \exp\left(-\dfrac{(1-\rho)\kappa NT}{2}\right), & \kappa\ge NTg\sigma^2.\end{cases}$$

Таким образом, все вероятности ошибок для предложенного теста стремятся к нулю с увеличением объема выборки N.

Докажем теперь второе утверждение теоремы. Пусть $\hat{n}$ — оценка момента структурного сдвига, полученная на первом этапе метода. Тогда для любого $0<\varepsilon<\alpha$ рассмотрим событие $\left\{\left|\dfrac{\hat{n}}{N}-\theta\right|>\varepsilon\right\}$. Имеем

$$P_1\left\{\left|\frac{\hat{n}}{N}-\theta\right|>\varepsilon\right\}\le P_1\left\{\max_i|\eta_N(i)|>\frac{1}{2}\varepsilon|h|\min(\theta,1-\theta)\right\}.$$

Пусть $\kappa_0=\dfrac{1}{2}\varepsilon|h|\min(\theta,1-\theta)$. Для вероятности в правой части неравенства справедлива экспоненциальная оценка сверху, доказанная ранее:

$$P_1\left\{\max_i|\eta_N(i)|\ge\kappa_0\right\}\le\begin{cases}\exp\left(-\dfrac{(1-\rho)^2\kappa_0^2N}{2\sigma^2g}\right), & 0<\kappa_0<NTg\sigma^2,\\ \exp\left(-\dfrac{(1-\rho)\kappa_0 NT}{2}\right), & \kappa_0\ge NTg\sigma^2.\end{cases}$$

Из этой экспоненциальной оценки следует, что нормированная оценка момента структурного сдвига $\hat{n}/N$ сходится почти наверное к истинному параметру структурного сдвига θ при увеличении объема выборки N.

Теорема 1 доказана.

На втором шаге метода происходит *различение гипотез* H_1 и H_2. На содержательном уровне возможность различения этих гипотез основана на различных свойствах решающей статистики $Z_N(\cdot)$ при гипотезах структурного сдвига и единичного корня.

Справедлива следующая теорема.

Теорема 2. *Вероятности ошибочных решений на втором шаге метода стремятся к нулю при увеличении объема выборки N.*

Идея доказательства теоремы 2 состоит в следующем. Из теоремы 1 следует, что нормированная оценка момента структурного сдвига $\hat{n}_1/N$ стремится почти наверное к истинному параметру структурного сдвига θ с увеличением объема выборки N. Поэтому для любого $0<\delta<\alpha$ эта нормированная оценка попадает в δ-окрестность параметра θ при достаточно большом N с вероятностью, близкой к единице.

Отсюда следует, что две подвыборки: $X_1(\hat{n}_1)=\{y_1,\ldots,y_{\hat{n}_1-[\delta N]}\}$ и $X_2(\hat{n}_1)=\{y_{\hat{n}_1+[\delta N]},\ldots,y_N\}$, формируемые на втором этапе метода, в ситуации единственного структурного сдвига с вероятностью, стремящейся к единице, будут классифицироваться как *статистически однородные* (гипотеза H_0).

Вместе с тем при гипотезе единичного корня обе эти подвыборки $X_1(\hat{n}_1)=\{y_1,\ldots,y_{\hat{n}_1-[\delta N]}\}$ и $X_2(\hat{n}_1)=\{y_{\hat{n}_1+[\delta N]},\ldots,y_N\}$ сохранят свойство *статистической неоднородности* (гипотеза H_2). В этом и состоит содержательная идея второго этапа метода: для каждой из этих подвыборок рассчитываются соответствующие величины $T_N(X_1)$ и $T_N(X_2)$ — максимумы модуля статистики $|Z_N(\cdot)|$, построенной по этим подвыборкам. Далее рассчитываются пороговые границы C_1 и C_2 для подвыборок $X_1(\hat{n}_1)=\{y_1,\ldots,y_{\hat{n}_1-[\delta N]}\}$ и $X_2(\hat{n}_1)=\{y_{\hat{n}_1+[\delta N]},\ldots,y_N\}$.

Решающее правило на втором шаге метода таково:

$$d_2(N)=\begin{cases}1, & T_N(X_1)<C_1 \text{ или } T_N(X_2)<C_2,\\ 2 & \text{в противном случае.}\end{cases}$$

Из теоремы 1 следует, что для этого решающего правила вероятности ошибочных решений будут стремиться к нулю при увеличении объема выборки N.

Предложенный метод легко обобщается на случай нескольких моментов структурных сдвигов.

4. Имитационное моделирование

Целью имитационных экспериментов была проверка свойств предложенного теста и качества различения основных рассматриваемых гипотез (стационарности, структурного сдвига и единичного корня) для конечных объемов выборки. При имитационном моделировании рассматривались следующие гипотезы и модели:

- H_0 — *гипотеза стационарности наблюдений*:

$$x_i=\rho x_{i-1}+\sigma v_i,\qquad v_i\sim N(0,1),$$
$$y_i=x_i,\quad i=1,\ldots,N,$$

где N — объем выборки; $-1<\rho<1$, $\sigma>0$; $N(0, 1)$ — стандартное нормальное распределение;

- H_1 — *гипотеза структурного сдвига по среднему*:

$$y_i=x_i+hI(i\le[\theta N]),\quad i=1,\ldots,N,$$

где $0<\theta<1$, $|h|>0$;

- H_2 — *гипотеза единичного корня*:

$$y_i = y_{i-1} + \sigma v_i, \quad i = 1, \dots, N.$$

При тестировании этих гипотез все параметры моделей, а именно σ, ρ, h, θ, предполагались неизвестными. Предполагалось лишь, что $\theta \geq \theta_0 = 0{,}1$, $-0{,}99 \leq \rho \leq 0{,}9$.

Оценивание коэффициента авторегрессии ρ осуществлялось следующим образом. Вначале рассчитывались эмпирические оценки дисперсии и автоковариации 1-го порядка для $n_0 = [\theta_0 N]$:

$$r(0) = \frac{\sum_{i=1}^{n_0} (y_i - \bar{y})}{n_0 - 1},$$

$$r(1) = \frac{\sum_{i=2}^{n_0} (y_i - \bar{y})(y_{i-1} - \bar{y})}{n_0 - 2},$$

где $\bar{y} = \sum_{i=1}^{n_0} y_i / n_0$.

Тогда оценка коэффициента авторегрессии $\hat{\rho} = r(1)/r(0)$, а оценка дисперсии $\hat{\sigma}^2 = \frac{1}{n_0 - 2} \sum_{i=2}^{n_0} (y_i - \hat{\rho} y_{i-1})^2$.

При этих оценках рассчитывалась пороговая граница:

$$C(N) = 1{,}3581 \frac{\hat{\sigma}}{\sqrt{N}} \frac{1 + \frac{|\hat{\rho}|}{4}}{1 - \hat{\rho}},$$

где значение 1,3581 соответствует доверительной вероятности 0,95. Во избежание недоразумений подчеркнем, что выбор пороговой границы, зависящей от объема выборки N, обеспечивает постоянство вероятности ошибки 1-го рода на уровне 0,05.

В экспериментах вычислялись оценки следующих характеристик качества тестирования:

- на первом шаге метода — вероятности ошибок 1-го и 2-го рода

$$\tilde{\alpha} = P_0\{T_N > C(N)\}, \ \tilde{\beta} = P_{12}\{T_N \leq C(N)\};$$

- на втором шаге метода — вероятности перепутывания гипотез H_1 и H_2

$$\tilde{\gamma} = P_1\{H_2\}, \ \tilde{\delta} = P_2\{H_1\}.$$

Полученные в экспериментах результаты приведены в табл. 1. Оценки вероятностей $\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, \tilde{\delta}$ рассчитывались как средние в $k = 5000$ повторениях имитационных экспериментов.

Прокомментируем полученные результаты. Во-первых, вероятность ошибки 1-го рода $\tilde{\alpha}$ была выбрана на уровне 5%, чем и объясняются стабильно низкие значения этого показателя. Вероятность ошибки 2-го рода $\tilde{\beta}$, как видно из таблицы, быстро убывает при возрастании объема выборки. В табл. 1 приведены выборки данных с отношением сигнал/шум 1:2,

Таблица 1

Оценки вероятностей ошибок в имитационных экспериментах при различном объеме выборки

N		100	200	300	500	700	1000
$h=1$	$\tilde{\alpha}$	0,114	0,056	0,040	0,072	0,056	0,038
$\rho=0{,}3$	$\tilde{\beta}$	0,412	0,282	0,106	0,030	0	0
$\sigma=0{,}5$	$\tilde{\gamma}$	0,05	0,054	0,06	0,04	0,04	0,028
$\theta=0{,}5$	$\tilde{\delta}$	0,0496	0,56	0,52	0,34	0,172	0,090
$h=0{,}5$	$\tilde{\alpha}$	0,082	0,062	0,052	0,040	0,038	0,048
$\rho=0{,}3$	$\tilde{\beta}$	0,64	0,23	0,082	0,026	0,006	0
$\sigma=0{,}5$	$\tilde{\gamma}$	0,05	0,052	0,058	0,044	0,03	0,024
$\theta=0{,}5$	$\tilde{\delta}$	0,50	0,566	0,568	0,354	0,168	0,072
$h=0{,}5$	$\tilde{\alpha}$	0,26	0,124	0,064	0,072	0,038	0,044
$\rho=0{,}7$	$\tilde{\beta}$	0,95	0,86	0,52	0,35	0,136	0,040
$\sigma=0{,}5$	$\tilde{\gamma}$	0,044	0,026	0,028	0,012	0,018	0,020
$\theta=0{,}5$	$\tilde{\delta}$	0,468	0,524	0,518	0,308	0,19	0,090
$h=0{,}5$	$\tilde{\alpha}$	0,046	0,038	0,058	0,032	0,026	0,030
$\rho=-0{,}7$	$\tilde{\beta}$	0,37	0,29	0,088	0,026	0,004	0
$\sigma=0{,}5$	$\tilde{\gamma}$	0,024	0,032	0,058	0,024	0,024	0,032
$\theta=0{,}5$	$\tilde{\delta}$	0,494	0,528	0,502	0,346	0,176	0,084

что является довольно низким показателем. Однако даже в этих сложных условиях тестирования удовлетворительное качество отмечается при объемах выборки порядка 500 наблюдений. В практических задачах отношение сигнал/шум, как правило, гораздо выше (10:100), что позволяет достигать удовлетворительного качества тестирования при объемах выборки 100–150 наблюдений.

Во-вторых, качество тестирования сильно зависит от степени коррелированности наблюдений. При высоких значениях коэффициента автокорреляции (0,7–0,8) удовлетворительное качество тестирования достигается при больших объемах выборки данных (800–1000).

На втором этапе метода происходит различение гипотез структурного сдвига и единичного корня. Из данных, приведенных в табл. 1, видно, что вероятность ошибки при тестировании гипотезы структурного сдвига ($\tilde{\gamma}$), как правило, довольно мала (3–5%), тогда как вероятность ошибки при тестировании гипотезы единичного корня ($\tilde{\delta}$) несколько выше, но также быстро убывает при возрастании объема выборки.

5. Практическое применение

Предложенный метод был использован в задачах проверки стационарности временных рядов эконометрических наблюдений. В качестве этих временных рядов были выбраны важнейшие макроэкономические индикаторы по России за 1994–2007 годы, измеренные с месячным интервалом.

Следует отметить, что преимущества предложенного метода проявляются начиная с объемов выборки 100–150 наблюдений. Поэтому попытки использования данного метода для тестирования стационарности наблюдений при объемах выборки 20–50 точек будут в большинстве случаев давать некорректные результаты.

В экспериментах были использованы следующие динамические ряды в диапазоне 1994.1–2007.9:

CPI — помесячный индекс инфляции на потребительском рынке (в процентах);
*M*2 — денежный агрегат M2 (трлн руб.);
E — официальный обменный курс доллара (руб./долл.);
PEL — помесячный индекс тарифов на электроэнергию для конечных потребителей (в процентах);
pi — темп инфляции: $pi = CPI/100 - 1$;
eps — темп изменения курса доллара: $eps = E/E(-1) - 1$;
mu — темп изменения денежного агрегата M2: $mu = M2/M2(-1) - 1$;
piel — темп изменения тарифов на электроэнергию для конечных потребителей: $piel = PEL/100 - 1$;
pcum — уровень инфляции (базовый индекс): $pcum = pcum(-1) * CPI/100$;
pcel — уровень тарифов на электроэнергию для конечных потребителей: $pcel = pcel(-1) * PEL/100$;
lpcum — логарифм уровня инфляции: $lpcum = \log(pcum)$;
lpcel — логарифм уровня тарифов на электроэнергию: $lpcel = \log(pcel)$;
*lm*2 — логарифм денежного агрегата M2: $lm2 = \log(M2)$;
lE — логарифм обменного курса доллара: $lE = \log(E)$;
Dlpcum — приращение логарифма уровня цен: $Dlpcum = lpcum - lpcum(-1)$;
Dlpcel — приращение логарифма уровня тарифов на электроэнергию: $Dlpcel = lpcel - lpcel(-1)$;
DlE — приращение логарифма обменного курса доллара: $DlE = lE - lE(-1)$.

Проверка этих динамических рядов на стационарность *с использованием расширенного теста Дики–Фуллера* (ADF) с тремя лагами и без включения константы и тренда дала следующие результаты:

- ряды *pi*, *eps*, *mu*, *piel*, *Dlpcum*, *Dlpcel*, *DlE* стационарны с доверительным уровнем 95%;
- ряды *CPI*, *M*2, *E*, *PEL*, *pcum*, *pcel*, *lpcum*, *lpcel*, *lm*2, *lE* имеют порядок нестационарности $I(1)$ с доверительным уровнем 95%.

Следует отметить, что проверка на наличие структурного сдвига в стандартных эконометрических пакетах для отдельных рядов, как правило, не проводится.

Исследование данных динамических рядов на стационарность *с использованием предложенного в статье метода* дало следующие результаты:

- ряды *pi*, *eps*, *mu*, *piel*, *Dlpcum*, *Dlpcel*, *DlE* стационарны при уровне ошибки 5%. Таким образом, получено полное совпадение результатов с тестом ADF;

- ряды *CPI*, *M2*, *E*, *PEL*, *pcum*, *pcel*, *lpcum*, *lpcel*, *lm2*, *lE* нестационарны при уровне ошибки 5%. Из этих рядов только *pcum*, *lpcum* классифицируются по типу структурного сдвига (в точке 1998.9); для остальных рядов принимается гипотеза наличия единичного корня.

6. Выводы

В работе предложен метод обнаружения и тестирования статистических нестационарностей типа структурного сдвига и единичного корня для одномерных временных рядов. Метод основан на модифицированной статистике Колмогорова–Смирнова и представляет собой двухшаговую процедуру: на первом шаге тестируется *гипотеза стационарности временного ряда* против совместной альтернативы неизвестного структурного сдвига и единичного корня; на втором шаге при отклонении гипотезы стационарности тестируется *гипотеза структурного сдвига* против альтернативы единичного корня.

Для предложенного метода в работе доказаны результаты об экспоненциальной сходимости к нулю вероятностей ошибочных решений в зависимости от возрастающего объема выборки. Для конечных выборок данных свойства предложенной двухшаговой процедуры исследованы методом имитационного моделирования.

В заключение попытаемся дать ответ на некоторые «наивные» вопросы, которые неизбежно возникают при чтении этой статьи:

1. Почему предложенный метод эффективнее широко известных в «узких» эконометрических кругах процедур DF, ADF (Dickey–Fuller), KPSS тестирования единичных корней и стохастических трендов в динамических рядах эконометрических наблюдений, а также методов Чоу [Chow (1960)], CUSUM [Ploberger, Kramer (1992)] и некоторых других для тестирования структурных сдвигов?
2. В работе рассматривается довольно частная модель стационарного временного ряда: $e_i = \rho e_{i-1} + u_i$, $e_0 = 0$, $|\rho| < 1$, $i = 1, \ldots, N$, где u_i — последовательность независимых одинаково распределенных наблюдений с $Eu_i = 0$, $Eu_i^2 = \sigma^2$. Зависят ли результаты, приведенные в статье, от выбора именно этой модели?

Отвечая на первый вопрос, вновь подчеркнем, что предложенный метод предназначен не только для тестирования гипотезы стационарности наблюдений против альтернатив единичного корня и структурного сдвига, но также для различения гипотез единичного корня и структурного сдвига в случае отклонения нулевой гипотезы о стационарности наблюдений.

Ответ на второй вопрос короче: выбранная модель может обобщаться до произвольной модели ARMA (авторегрессии — скользящего среднего), в том числе с добавлением константы. Общие результаты статьи инвариантны к выбору этой модели.

Список литературы

Banerjee A., Lumsdaine R. L., Stock J. H. Recursive and sequential tests of the unit root and trend-break hypothesis: theory and international evidence // *Journal of Business and Economic Statistics*. 1992. № 10. P. 271–287.

Brodsky B., Darkhovsky B. Non-parametric Statistical Diagnosis: Problems and Methods. KLUWER, Dordrecht, 2000.

Chow G. C. Tests of equality between sets of coefficients in two linear regressions // *Econometrica*. 1960. № 28. P. 591–605.

Christiano L. J. Searching for breaks in GDP // *Journal of Business and Economic Statistics*. 1992. № 10. P. 237–250.

Dickey D. A., *Fuller W. A.* Distribution of the estimators for autoregressive time series with a unit root // *Journal of the American Statistical Association*. 1979. № 74. P. 427–431.

Dickey D. A., *Fuller W. A.* Likelihood ratio statistics for autoregressive time series with a unit root // *Econometrica*. 1981. № 49. P. 1057–1072.

Kim D., *Perron P.* Unit root tests with a consistent break fraction estimator. Department of Economics, Boston University, 2005.

Kwiatkowski D., *Phillips P. C. B.*, *Schmidt P.*, *Shin Y.* Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root // *Journal of Econometrics*. 1992. № 54. P. 159–178.

MacKinnon J. B. Numerical distribution functions for unit roots and cointegration tests // *Journal of Applied Econometrics*. 1996. № 11. P. 601–618.

Montanes A., *Reyes M.* The asymptotic behavior of the Dickey-Fuller tests under the crash hypothesis // *Statistics and Probability Letters*. 1999. № 48. P. 401–409.

Montanes A., *Reyes M.* Structural breaks, unit roots and methods for removing the autocorrelation pattern // *Statistics and Probability Letters*. 2000. № 48. P. 401–409.

Nelson C. R., *Plosser C. I.* Trends and random walks in macroeconomic time series: some evidence and implications // *Journal of Monetary Economics*. 1982. № 10. P. 139–162.

Peligrad M. On the central limit theorem for weakly dependent sequences with a decomposed strong mixing coefficient // *Stochastic Processes and their Applications*. 1992. № 42. P. 181–193.

Perron P. The great crash, the oil price shock, and the unit root hypothesis // *Econometrica*. 1989. № 57. P. 1361–1401.

Perron P. Trend, unit root, and structural change in macroeconomic time series. In: Cointegration for the Applied Economist, Rao B. B. (ed.). Basingstoke: Macmillan Press, 1994. P. 113–146.

Perron P. Dealing with structural breaks. Palgrave Handbook of Econometrics. Vol. 1. Econometric Theory, 2005.

Perron P., *Vogelsang T.* Nonstationarity and level shifts with an application to purchasing power parity // *Journal of Business and Economic Statistics*. 1992. № 10. P. 301–320.

Perron P. Further evidence from breaking trend functions in macroeconomic variables // *Journal of Econometrics*. 1997. № 80. P. 355–385.

Perron P., *Yabu T.* Testing for shifts in trend with an integrated or stationary noise component. Department of Economics, Boston University, 2005.

Phillips P. C. B., *Perron P.* Testing for a unit root in a time series regression // *Biometrika*. 1988. № 75. P. 335–346.

Ploberger W., *Kramer W.* The CUSUM test with OLS residuals // *Econometrica*. 1992. № 60. P. 271–285.

Rappoport P., *Reichlin L.* Segmented trends and non-stationary time series // *Economic Journal*. 1989. № 99. P. 168–177.

Zivot E., Andrews *D.* Further evidence on the Great crash, the oil price shock and the Unit root hypothesis // *Journal of Business and Economic Statistics*. 1992. № 10. P. 251–287.