ЛЕКЦИОННЫЕ И МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ

Анализ временных рядов

Канторович Г.Г.

В этом номере публикуются очередные три лекции курса «Анализ временных рядов». Они посвящены методологии применения расширенного теста Дикки–Фуллера (ADF-test) для проверки наличия единичного корня, или, другими словами, определения типа нестационарности временного ряда. Рассматривается зависимость распределения так называемого t-отношения от наличия в составе регрессоров свободного члена и/или тренда, а также от лаговой структуры исследуемого ряда. В качестве методики предлагается процедура Доладо и его соавторов. Рассматривается случай кратных единичных корней. Десятая лекция посвящена исследованию наличия единичных корней при возможном наличии структурного скачка в параметрах модели. Обсуждаются результаты Перрона для экзогенного времени скачка и Зивота с Эндрюсом – для эндогенного. Приводятся примеры конкретных экономических рядов. В последующих лекциях предполагается рассмотреть модели взаимосвязи стационарных и нестационарных временных рядов.

Лекция 8

В прошлой лекции мы упоминали распределение Дикки–Фуллера двух типов: типа F-отношения и типа t-отношения $\frac{\hat{\gamma}}{s.e.(\hat{\gamma})}$. Второе из отношений рассматривалось для различных моделей. Один раз – для модели без линейного тренда, а во второй раз мы использовали таблицу распределения для случая, когда в модель включен линейный тренд. Тест Дикки–Фуллера предназначен для того, чтобы различить временные ряды типа TS и DS. В соответствии с нулевой гипотезой H_0 исследуемый ряд принадлежит к типу DS. По альтернативной гипотезе он может быть типа TS, но одновременно быть нестационарным – иметь детерминированный тренд, или не иметь тренда – быть стационарным. Как уже упоминалось, спецификация модели в прямой и альтернативной гипотезе, т.е. включение или не включение в модели свободного члена и/или детерминированного тренда, влияет на распределение t-отношения. Рассмотрим это подробнее.

Канторович Г.Г. – профессор, к. физ.-мат. н., проректор ГУ–ВШЭ, зав. кафедрой математической экономики и эконометрики.

Дикки и Фуллер [4] начинали с исследования уравнения: $\Delta X_t = \gamma X_{t-1} + \varepsilon_t$. В этом случае при условии, что выполнена гипотеза H_0, т.е. что $\gamma = 1$, распределение t-отношения (отношения оценки коэффициента γ, полученной МНК, к оценке его среднеквадратичного отклонения) называется DF-распределением. Включив в модель свободный член, получим модель $\Delta X_t = \alpha + \gamma X_{t-1} + \varepsilon_t$. А дополнительное включение линейного тренда дает модель $\Delta X_t = \alpha + \beta t + \gamma X_{t-1} + \varepsilon_t$. Оказалось, что во всех трех случаях распределение t-отношения $\frac{\hat{\gamma}}{s.e.(\hat{\gamma})}$ выражается через интегралы от винеровского процесса, но по-разному. Все три распределения принято связывать с именами Дикки и Фуллера. Однако эти распределения разные и зависят от того, какие добавочные регрессоры входят в уравнение. Обозначим критические величины для первого распределения τ_0, для второго распределения τ_μ, для третьего распределения τ_τ.

Взаимное расположение этих критических значений для одного и того же уровня значимости и одного и того же числа степеней свободы схематически показано на рис. 8.1. Также на рисунке показано критическое значение распределения Стьюдента для того же уровня значимости и того же числа степеней свободы. Все эти значения отрицательные. Впрочем, в литературе их иногда приводят со знаком плюс, что обычно не приводит к недоразумениям.

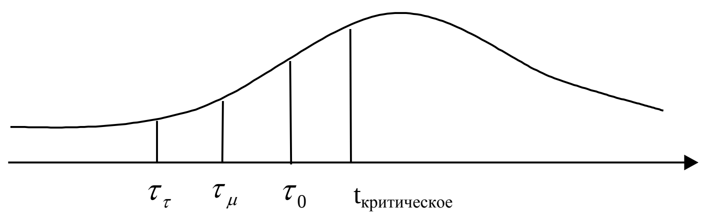

Рис. 8.1.

Прежде чем двигаться дальше зададимся следующим вопросом. Из результатов Нельсона и Канга [11] мы знаем об опасности принятия процесса типа I(1) за процесс типа I(0) и применения к нему обычных процедур метода наименьших квадратов. Если мы будем применять к рядам типа DS обычные формулы регрессионного анализа, то получим кажущиеся зависимости, кажущиеся тренды, т.е. некий статистический феномен вместо экономической взаимосвязи, которую ищем. Может показаться, что достаточно обращаться с процессом типа TS как с процессом типа DS, т.е. брать от него первые разности и дальше оперировать с рядом первых разностей. Ведь при поиске взаимозависимости между процессами y и z можно перейти к рядам Δy и Δz, которые останутся процессами типа I(0), и строить обычную регрессию. При этом может оказаться, что один из рядов (или оба) уже были стационарными и не требовали взятия первой разности. Приводит ли такое *излишнее* взятие разностей к негативным последствиям? Другими словами, существуют ли негативные последствия, если мы примем ряд типа TS за ряд типа DS? Эта ситуация может встретиться при применении метода Бокса–Дженкинса всякий раз, когда мы переоценили величину параметра *d*.

Рассмотрим опасность излишнего взятия разностей на конкретном примере. Пусть рассматривается процесс MA(1), стационарный процесс, имеющий вид: $x_t = \varepsilon_t + \alpha\varepsilon_{t-1} = (1+\alpha L)\varepsilon_t$. Если, не зная, что он уже стационарный, взять от процесса первую разность, то получим:

$$y_t = \Delta x_t = (1-L)x_t = (1-L)(1+\alpha L)\varepsilon_t = (1-(1-\alpha)L - \alpha L^2)\varepsilon_t.$$

Получившаяся модель теперь относится к типу MA(2), оценивать нужно уже два параметра, а не один, но сейчас более важен другой аспект. Оценим обычным способом коэффициенты модели $y_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}$. Посмотрим, какова остаточная дисперсия процессов x_t и y_t. В первом случае дисперсия будет равна: $Var(x_t) = (1+\alpha^2)\sigma_\varepsilon^2$. Во втором случае, поскольку взята излишняя разность, один из корней соответствующего характеристического уравнения равен единице. Это означает, что получившийся процесс необратим, для него не существует AR(∞) представления, и ясно, что это само по себе приведет к сложностям в оценивании. Дисперсия этого процесса будет равна $Var(y_t) = (1+\theta_1^2+\theta_2^2)\sigma_u^2$. Или можно записать по-другому: $Var(y_t) = [1+(1-\alpha)^2+\alpha^2]\sigma_\varepsilon^2$. Невооруженным глазом видно, что $Var(y_t) > Var(x_t)$, т.е. взятие излишней разности приводит к увеличению дисперсии, в том числе и остаточной дисперсии после применения метода наименьших квадратов. Мы позже рассмотрим пример, в котором численно убедимся, что это увеличение действительно имеет место. Иногда этот эффект наряду с поведением выборочной автокорреляционной функции может служить некоторым неформальным критерием для ответа на вопрос, нужно брать последовательную разность или нет. Если при переходе к разности дисперсия возрастает, то, скорее всего, этого делать не надо.

Таким образом, неверное установление типа процесса ведет к негативным последствиям при ошибке «в обе стороны».

Теперь вернемся к *unit root test*. Специфицируем процесс типа TS в виде: $x_t = \alpha + \beta t + \varepsilon_t$. Попробуем выписать такую модель, в которой различие между типами TS и DS будет выражаться в значениях коэффициентов. Для этого заменим в модели белый шум ε_t на процесс u_t, подчиняющийся марковской схеме первого порядка $u_t = \rho u_{t-1} + \varepsilon_t$. Тогда нулевая гипотеза, заключающаяся в том, что ряд принадлежит типу DS, эквивалентна тому, что $\rho = 1$, а альтернативная – тому, что $\rho < 1$.

Чтобы убедиться в этом, запишем: $x_{t-1} = \alpha + \beta(t-1) + u_{t-1}$. Теперь умножим левую и правую части на ρ, что дает следующее выражение:

$$\rho x_{t-1} = \rho\alpha + \rho\beta(t-1) + \rho u_{t-1}.$$

Вычитая, получаем: $x_t = [\alpha(1-\rho) + \rho\beta] + \beta(1-\rho)t + \rho x_{t-1} + \varepsilon_t$. Это соотношение включает свободный член, линейный тренд, авторегрессионый член и белый шум. При выполнении нулевой гипотезы $\rho = 1$ процесс принимает вид:

$$x_t = \beta + x_{t-1} + \varepsilon_t .$$

Мы получили случайное блуждание с дрейфом, если $\beta \neq 0$. Если же $\rho < 1$, то в модели присутствует тренд и авторегрессионый член, причем авторегрессионная часть стационарна. Можно сделать два вывода. Во-первых, нам действительно удалось показать, что рассматриваемая модель охватывает оба наших случая. Во-вторых − что можно относить к классу TS такие процессы, где вместо белого шума ε_t стоит стационарный процесс типа ARMA.

В соответствии с полученным соотношением принято рассматривать следующую модель для проверки наличия единичного корня: $x_t = \mu + bt + \rho x_{t-1} + \varepsilon_t$. Если оценить эту модель методом наименьших квадратов, то нужно проверить сформулированную нулевую гипотезу: $H_0 : \rho = 1$. Если бы при нулевой гипотезе ряд остался стационарным, то при предположении о нормальности ε_t надо просто проверить равенство коэффициента b единице, для чего сравнить t-статистику для этого коэффициента с квантилями t-распределения Стьюдента. Чтобы еще ближе свести процедуру к привычной, вычтем из обеих частей x_{t-1}. Тогда получается уравнение $\Delta x_t = \mu + bt + \underbrace{(\rho - 1)}_{\gamma} x_{t-1} + \varepsilon_t$. Получаем, что надо проверить следующую нулевую гипотезу против альтернативной:

$$H_0 : DS \Leftrightarrow \gamma = 0$$
$$H_1 : TS \Leftrightarrow \gamma < 0 \ .$$

Поскольку $\rho > 1$ соответствует нестационарному процессу взрывного типа, который исключен из рассмотрения, используется односторонний тест. При обычной регрессии для проверки этих гипотез применяется t-отношение, т.е. $t = \dfrac{\hat{\gamma}}{\hat{\sigma}_\gamma}$, которое сравнивается с критическим значением распределения Стьюдента. Однако мы уже видели, что при выполнении гипотезы H_0 распределение t-отношения, не подчиняется распределению Стьюдента, а подчиняется распределению, которое сейчас принято называть распределением Дикки−Фуллера. Поэтому схема проверки гипотезы такая же, только мы должны сравнивать статистику с квантилями не t-распределения, а DF-распределения. И решающее правило тоже чрезвычайно простое. Если $|\tau| > \tau_{кр}$, выборочная статистика расположена левее табличного значения (мы «работаем» на левом хвосте распределения), то мы отвергаем нулевую гипотезу и считаем, что ряд относится к типу TS, а если выборочное значение лежит правее табличного значения, то ряд − типа DS. Мы уже знаем, что табличные значения распределения Стьюдента и нормального распределения лежат правее для одного и того же уровня значимости.

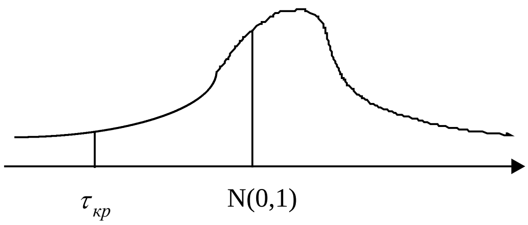

Рис. 8.2.

Оказалось, что вид распределения, а следовательно его квантили, критические значения существенно зависят от того, включены ли в оцениваемую модель свободный член и/или тренд. Мы начали с уравнения $x_t = \alpha + \beta t + u_t$, где u_t подчиняется марковской схеме. То есть мы сразу включили в модель свободный член и тренд. Далеко не всегда необходимо вводить тренд в уравнение. Если начать с уравнения $y_t = \alpha + u_t$, в котором нет тренда, то в оцениваемом уравнении не только тренд пропадет, но и в свободном члене пропадет одно слагаемое. Поэтому в предположении гипотезы H_0 у процесса нет дрейфа. Действительно, если выполнена гипотеза H_0, это означает, что $\rho = 1$, и в уравнении пропадет свободный член. У нас будет случайное блуждание без дрейфа, совсем простое. Но тогда оказывается, что у t-отношения другое распределение. Поэтому нужно рассчитать уже другое критическое значение. Аналогично, если мы исключим еще и α из исходной модели, что тоже может быть, то получим еще одно, несколько иное распределение. Поэтому в зависимости от того, какую модель типа DS мы специфицируем: модель с линейным трендом и дрейфом, модель только с дрейфом или модель без дрейфа, у нас получаются три разных случая, три разных таблицы распределения. Существует несколько систем обозначений, мы будем пользоваться такими: $\tau_0, \tau_\mu, \tau_\tau$.

При практическом применении теста Дикки–Фуллера возникает вопрос, каким образом выбрать вид модели, или, что тоже самое, как правильно специфицировать прямую и альтернативную гипотезы. Во всех случаях мы проводим различие между рядами типа TS и DS соответственно, но внутри каждого из типов мы должны выбрать подходящий к исследуемым данным конкретный вид модели. Если выборочная статистика попадает между критическими значениями, то возникает неясность, может быть мы неправильно специфицировали модель и поэтому пытаемся сделать вывод по неправильному критическому значению. Так как у нас существует разнообразие основной и альтернативных гипотез, т.е. точных спецификаций DS и TS процесса, то как быть в этом случае? Понятно, что можно сделать глобальный перебор всех возможных моделей, но ведь вывод по разным моделям может быть противоречивым. И еще один аспект. Мы рассмотрели ситуацию, когда u_t подчиняется схеме AR(1), но случайная составляющая может описываться автокорреляционной схемой более высокого порядка. Начнем с распространения DF-теста на процессы, когда u_t подчиняется ARMA более высокого порядка.

Для простоты выкладок рассмотрим процесс $x_t = u_t$, опустив без потери общности и тренд, и свободный член для уменьшения громоздкости выкладок. Пусть u_t подчиняется авторегрессионной схеме второго порядка:

$$u_t = \rho_1 \cdot u_{t-1} + \rho_2 \cdot u_{t-2} + \varepsilon_t .$$

Если мы хотим исключить u_t, то должны выписать $\rho_1 \cdot x_{t-1} = \rho_1 \cdot u_{t-1}$ и $\rho_2 \cdot x_{t-2} = \rho_2 u_{t-2}$ и затем вычесть эти выражения из исходного. В результате получим: $x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \varepsilon_t$. Мы видим, что сделанное предположение о виде авторегрессионной зависимости случайного процесса u_t эквивалентно тому, что процесс x_t также подчиняется схеме AR(2).

Прежде всего, нужно выразить в терминах коэффициентов этого уравнения ρ_1 и ρ_2 нулевую гипотезу о том, что ряд относится к типу DS, т.е. имеет единичный корень. Перепишем уравнение в операторном виде: $(1 - \alpha_1 - \alpha_2 L^2) x_t = \varepsilon_t$. Что значит, что уравнение имеет единичный корень? Если подставить L=1, получается $1 - \alpha_1 - \alpha_2 = 0$. Следовательно, нулевая гипотеза принимает вид:

$$\rho_1 + \rho_2 = 1.$$

Далее вычтем из обеих частей x_{t-1}, получаем:

$$\Delta x_t = (\rho_1 - 1) x_{t-1} + \rho_2 x_{t-2} + \varepsilon_t .$$

Добавим и вычтем $\rho_2 x_{t-1}$ в правой части. Получаем:

$$\Delta x_t = (\rho_1 + \rho_2 - 1) x_{t-1} - \rho_2 \Delta x_{t-1} + \varepsilon_t .$$

Теперь нулевая гипотеза о том, что ряд принадлежит типу DS или что у него есть единичный корень, сводится к проверке равенства нулю коэффициента $(\rho_1 + \rho_2 - 1) = 0$. Это соотношение позволяет переформулировать нулевую гипотезу $H_0 : DS \Leftrightarrow \rho_1 + \rho_2 - 1 = 0$. То есть если построить регрессию Δx_t на x_{t-1} и на приращение того же x_{t-1}, то можно стандартным образом проверить нулевую гипотезу. Для этого нужно рассчитать t-отношение и сравнить его величину с табличным значением нужного распределения. Что изменилось от того, что процесс подчиняется схеме AR(2)? В уравнении прибавился регрессор, являющийся конечной разностью – приращением Δx_{t-1}. А если бы была модель AR(3)? Вновь при наличии единичного корня сумма коэффициентов должна быть равна единице. Мы бы сделали аналогичное преобразование, и в преобразованном уравнении добавилось бы новое приращение Δx_{t-2}. Наличие у исследуемого процесса MA-части не вносит ничего принципиально нового.

Следовательно, в наиболее общем случае, когда случайное возмущение относится к типу ARMA(p,q), нужно исследовать следующее уравнение:

$$\Delta x_t = \gamma x_{t-1} + \sum_{i=1}^{p} w_i \Delta x_{t-i} + \sum_{i=1}^{q} \delta_i \varepsilon_{t-i} .$$

И если бы у процесса в исходном уравнении был свободный член и линейный тренд, вид уравнения сохранился бы. То есть общая модель, которую надо оценивать методом наименьших квадратов, принимает вид:

$$\Delta x_t = \alpha + \beta t + \gamma x_{t-1} + \sum_{i=1}^{p} w_i \Delta x_{t-i} + \sum_{i=1}^{q} \delta_i \varepsilon_{t-i} .$$

Нулевая гипотеза о том, что процесс относится к типу DS, по смыслу эквивалентна тому, что $\gamma = 0$. А альтернативная гипотеза, что ряд относится к типу TS, означает, что $\gamma < 0$.

Ключевой вопрос: изменило ли добавление регрессоров распределение t-отношения для коэффициента γ? Мы уже видели, что добавление свободного члена и/или линейного тренда приводило к изменению распределения и его критических значений. Оказалось, что наличие приращений и запаздывающих значений случайного возмущения не меняет распределения, и мы можем пользоваться теми же таблицами Мак-Киннона.

Если в модели $\Delta x_t = \alpha + \beta t + \gamma x_{t-1} + \sum_{i=1}^{p} w_i \Delta x_{t-i} + \sum_{i=1}^{q} \delta_i \varepsilon_{t-i}$ присутствуют и свободный член, и тренд, то нулевую гипотезу надо проверять, используя статистику τ_τ, если – только свободный член α, то статистику τ_μ, если нет ни того, ни другого, то надо использовать статистику τ_0. Этот тест носит название *Augmented Dickey-Fuller test* из-за того, что в уравнении появились приращения, и, как мы только что указали, те же самые статистики $\tau_0, \tau_\mu, \tau_\tau$ «работают» для этой расширенной модели. Стандартная аббревиатура его названия: ADF-тест. По-русски название этого теста обычно переводят: расширенный тест Дикки–Фуллера. Напомним, что его использование основано на предположении, что процесс u_t подчиняется ARMA(p,q), а это значит, имеет постоянную дисперсию. Второе существенное предположение состоит в том, что мы знаем точные значения p и q.

В ADF-тесте мы проверяем значимость только одного единственного коэффициента. Все остальные коэффициенты нас сейчас не интересуют, они являются вспомогательными. Для применимости ADF-теста важно проверить, что дисперсия случайного возмущения ε_t постоянна, т.е. наличие гомоскедастичности возмущений. Для макроэкономических рядов этот вопрос не столь важен, а для финансовых важен чрезвычайно, потому что в них волатильность часто меняется и дисперсия не всегда постоянна. В случае гетероскедастичности возмущений ADF-тест уже не применим.

И вторая важная проблема: как правило, мы не знаем значений p и q. Выяснилось, что, к сожалению, ADF-тест весьма чувствителен к правильному выбору параметров p и q. Моделирование методом Монте-Карло показало, что если число приращений в ADF-тесте согласовано с длиной реализации ряда, распределение t-отношения для коэффициента γ является распределением Дикки–Фуллера. Следовательно, количество лагов, которое нужно включать в модель при применении ADF-теста, должно быть увязано с длиной реализации.

Для выбора числа лагов, включаемых при применении ADF-теста, было предложено несколько эвристических критериев, проверенных практикой и моделированием Монте-Карло. Например, было предложено выбирать количество лагов следующим образом [16]: число лагов выбирается равным $\left[T^{1/3}\right]$, где [x] означает целую часть числа x. Для квартальных данных было предложено [17] выбирать число лагов по формуле: $\left[4\left(T/100\right)^{1/4}\right]$, и для месячных данных – $\left[12\left(T/100\right)^{1/4}\right]$. Наконец, Диболд и Нерлов показали [6], что на практике хорошо работает приближение $\left[T^{1/4}\right]$.

Получаем простое правило. В макроэкономических рядах, если у вас от 81 до 256 точек, то нужно включать 3 лага. Если у вас меньше 81 точки, то два лага. В российской макроэкономике других случаев у вас не будет. Если вы имеете дело с данными по финансовым рядам, то там может быть другая ситуация. Интересно, что мощность критерия зависит не от числа наблюдений, к чему мы привыкли, но еще от «*спэна*», т.е. от общей продолжительности ряда. У вас есть возможность выбора. Если есть выбор: использовать 500 наблюдений на коротком интервале или 500 – на длинном, лучше брать на длинном. Популярный эконометрический пакет Econometric Views включает в себя некоторый алгоритм выбора числа лагов. К сожалению, мне не удалось выяснить, как именно он это выбирает.

В литературе [8] встречается еще один прием по выбору надлежащего числа лагов при применении ADF-теста. Предлагается оставлять такое количество лагов, при котором все оценки МНК-коэффициентов при приращениях будут статистически значимы по обычному t-распределению Стьюдента.

На практике использование разного числа лагов может привести к разным выводам о типе процесса. Обычно мы стараемся учитывать трудность однозначного определения числа лагов и доверять результату, когда он устойчив к изменению лага, но в некотором разумном диапазоне. То, что иногда трудно различить тип ряда, TS или DS, не должно удивлять, ведь если для процесса AR(1) $\rho = 0{,}998$, т.е. близко к единице, то на конечных отрезках очень трудно отличить этот стационарный процесс от случайного блуждания. Пакет Econometric Views ищет ADF-уравнение с тем количеством лагов, которое вы ему укажете, или по умолчанию применяет встроенный алгоритм.

ADF-тест допускает возмущение вокруг тренда или в отсутствии тренда в виде общего процесса ARMA(p,q). Вторая возможность, которую стоит рассмотреть, это возможность наличия гетероскедастичности случайной составляющей ε_t. Наличие гетероскедастичности или волатильность более характерны для финансовых рядов. Филлипс и Перрон предложили [14] так называемый *непара-*

метрический тест Филлипса–Перрона (PP-test). Правда, практика его применения показала его чрезвычайно малую мощность, так что в настоящее время он почти не применяется, хотя и включен в состав ряда специализированных компьютерных пакетов, в частности в Econometric Views. Тест Филлипса–Перрона использует уравнение без приращений, т.е. в виде: $x_t = \mu + bt + \rho x_{t-1} + \varepsilon_t$. Но в качестве случайного возмущения допускается любой стационарный, но может быть гетероскедастичный и автокоррелированный процесс.

Для учета более общей структуры случайного возмущения в тесте Филлипса–Перрона рассчитывается следующая статистика:

$$Z(\tau_\mu) = \tau_\mu\left(\hat{\sigma}/\hat{\sigma}_{\tau l}\right) - \frac{1}{2}\left(\hat{\sigma}_{\tau l}^2 - \hat{\sigma}^2\right) \cdot T \cdot \left\{\hat{\sigma}_{\tau l}^2 \sum_{l=2}^{T}(x_{l-1} - \bar{x}_{-1})^2\right\}^{-1/2}.$$

Здесь, $\hat{\sigma}^2$ – оценка дисперсии остатков от регрессии, $\bar{x}_{-1} = \frac{1}{T-1}\sum_{t=1}^{T-1} x_t$; $\hat{\sigma}_{\tau l}^2 = \frac{1}{T}\sum_{l=1}^{T} e_l^2 + \frac{2}{T}\sum_{j=1}^{l} w_{jl} \cdot \sum_{l=j+1}^{T} e_l e_{l-j}$. Веса $w_{jl} = 1 - \frac{j}{l+1}$ здесь подобраны так, чтобы обеспечивать положительность оценки дисперсии $\hat{\sigma}_{\tau l}^2$. Параметр l можно выбирать, используя вышеприведенные рекомендации Шверта. Интересно, что статистика Z подчиняется тому же распределению Дикки–Фуллера, и мы можем воспользоваться теми же самыми таблицами Мак–Киннона.

Лекция № 9

Продолжим исследование наличия единичных корней или порядка интеграции временного ряда. Следует сказать, что, несмотря на наличие ряда тестов, в частности теста ADF, процедура исследования не так проста. Еще раз аккуратно выпишем, какие гипотезы проходят проверку.

Ряд относится к типу $DS \Rightarrow H_0 : x_t = x_{t-1} + \varepsilon_t$ – случайное блуждание без дрейфа, нестационарный ряд типа I(1);

Ряд относится к типу $TS \Rightarrow H_1 : x_t = \rho x_{t-1} + \varepsilon_t$; $|\rho| < 1$ – стационарный ряд, у него даже нет никакого тренда, это самый простейший вид.

Если $x_t = x_{t-1} + \varepsilon_t$, то общее решение разностного уравнения будет иметь следующий вид: $x_t = x_0 + \sum_{i=1}^{t} \varepsilon_i$.

У нас может быть и более сложная ситуация, когда к ряду типа DS относится ряд вида: $H_0 : x_t = \mu + x_{t-1} + \varepsilon_t$, случайное блуждание с дрейфом. Для него общее решение будет иметь вид: $x_t = x_0 + \mu t + \sum_{i=1}^{t} \varepsilon_i$. Стоит отметить, что процессы $x_t = \rho x_{t-1} + \varepsilon_t$ и $x_t = \mu + x_{t-1} + \varepsilon_t$ ведут себя очень по-разному и не являются реальными альтернативами друг другу. В самом деле, нулевая гипотеза соответствует нестационарному процессу с трендом, а альтернативная – стационарному процессу с нулевым математическим ожиданием и без тренда. Более подходящая альтернативная гипотеза имеет вид: $H_1 : x_t = \alpha + \beta t + \varepsilon_t$, также содержащий линейный тренд. В результате мы получали ADF-тест в следующем виде: $\Delta x_t = \alpha + \beta t + (\rho - 1)x_{t-1} + \sum_{i=1}^{p} \delta_i \Delta x_{t-i} + \varepsilon_t$. Построение такой регрессии – это и есть проверка ADF-теста. При этом статистика, которую мы используем для проверки гипотезы, – это стандартное t-отношение для коэффициента $(\rho - 1)$, т.е. $t = \dfrac{\hat{\rho} - 1}{\hat{\sigma}_\rho} \sim \tau_\tau$. Эта статистика имеет распределение, таблицы которого рассчитаны Мак-Кинноном. Дополнительный регрессор βt меняет предельное распределение, которое первоначально было в тесте DF. Это распределение выражается через стохастические интегралы немного по-другому, и его критические значения получили обозначение τ_τ.

Это означает, что статистика будет разной в зависимости от того, равно β нулю или $\beta \neq 0$. Более того, мы знаем, что если и $\alpha = 0$, то распределение вновь меняется. Если мы строили уравнение при $\beta = 0$, т.е. не включали добавочный трендовый регрессор, критические значения обозначались τ_μ. И, наконец, если одновременно и $\alpha = 0$, и $\beta = 0$, т.е. мы не включаем их в качестве регрессоров в оцениваемую модель, то тогда соответствующие критические значения получили обозначение τ_0. Это просто одна из возможных систем обозначений.

Сложившаяся ситуация отличается от обычной линейной множественной регрессии. Там добавление дополнительного регрессора не меняет распределение используемой t-статистики (лишь меняется число степеней свободы). Теперь это не так. Мы заранее не знаем, какие регрессоры на самом деле включать в модель, а какие нет, и оказываемся в дополнительно неопределенной ситуации. Нужна некоторая методика применения ADF-теста.

Рассмотрим процедуру, предложенную Доладо, Дженкинсоном и Сосвилла-Риверо [7]. Назначение ее все то же, мы хотим разделить ряды типа TS и DS, но не знаем, надо включать тренды в модель или не надо. Напомним, что вопрос заключается в том, как приводить ряд к стационарному виду: с помощью построения регрессии с линейным трендом или с помощью перехода к первым разностям. Процедура состоит из следующих шагов.

1. Оцениваем ADF уравнение вида:

$$\Delta x_t = \alpha + \beta t + (\rho - 1)x_{t-1} + \sum_{i=1}^{p} \delta_i \Delta x_{t-i} + \varepsilon_t,$$

т.е. включаем в него и свободный член, и линейный тренд. При этом нулевая гипотеза имеет вид $H_0 : DS \Rightarrow \rho = 1$. Если $\rho = 1$, то ряд фактически относится к новому типу, который мы почти не исследовали, в нем наличествуют два типа тренда. Пусть для простоты p=0, тогда получаем процесс вида $x_t = \alpha + \beta \cdot t + x_{t-1} + \varepsilon_t$. Он на самом деле содержит два тренда. Будем называть линейный тренд, соответствующий ряду типа TS, *детерминированным трендом*, а случайное блуждание с дрейфом − *стохастическим трендом*, т.е. случайно меняющимся трендом. Ряд нового типа содержит оба тренда: и стохастический, и детерминированный. Общее решение такого уравнения имеет вид:

$$x_t = x_0 + \sum(\alpha + \beta \cdot t + \varepsilon_t) = x_0 + \alpha \cdot t + \beta \cdot \frac{t(t+1)}{2} + \sum \varepsilon_t .$$

Таким образом, процесс содержит уже квадратичный тренд. Напомним, что $\sum t = \dfrac{t(t+1)}{2}$, это просто арифметическая прогрессия. Сравнение величины t-отношения для коэффициента при x_{t-1} с критическим значением распределения Дикки−Фуллера может привести к двум возможностям. И выводы зависят от того, попадаем ли мы в критическую область отвержения нулевой гипотезы или не попадаем. Если мы попадаем в критическую область, т.е. t-отношение (с учетом знака) удовлетворяет неравенству $t < \tau_\tau$, тогда наш ряд не имеет единичного корня, т.е. является типа TS. Если же $t > \tau_\tau$, то прежде чем сделать вывод о типе ряда, нужно ответить на вопрос, правильно ли мы включили в наше уравнение линейный тренд βt. Для этого предположим, что нулевая гипотеза выполнена и инкорпорируем эту гипотезу в исходное уравнение, получая модель без регрессора x_{t-1}: $\Delta x_t = \alpha + \beta t + \sum_{i=1}^{p} \delta_i \Delta x_{t-i} + \varepsilon_t$.

2. Оценим уравнение $\Delta x_t = \alpha + \beta t + \sum_{i=1}^{p} \delta_i \Delta x_{t-i} + \varepsilon_t$. Это уравнение заведомо не имеет единичных корней. Поэтому вопрос о том, значим или не значим статистически коэффициент β, наклон линейного тренда, может быть решен стандартной t-статистикой для коэффициента при тренде. Если коэффициент β не значим, т.е. $\beta = 0$, то делаем вывод, что мы зря включили трендовую переменную в уравнение ADF-теста. Если же коэффициент β значим, то наш вывод на первом шаге о том, что ряд принадлежит типу DS, получил подтверждение, и процедура окончена. Если же коэффициент β не значим, то надо перейти к следующему шагу.

3. Поскольку правомерность включения линейного тренда в спецификацию ADF-теста не подтвердилась, переходим к ADF-тесту без тренда, т.е. оцениваем регрессию $\Delta x_t = \alpha + (\rho - 1)x_{t-1} + \sum_{i=1}^{p} \delta_i \Delta x_{t-i} + \varepsilon_t$. Рассчитываем опять t-статистику для коэффициента при регрессоре x_{t-1}, но теперь ее нужно сравнивать с τ_μ, по-

тому что анализируется уже другое уравнение. Если $t < \tau_\mu$, то наш вывод состоит в том, что нет единичного корня и нет линейного тренда. Если $t > \tau_\tau$, то мы начинаем проверять правомерность включения свободного члена α в спецификацию ADF-теста. Строим регрессию $\Delta x_t = \alpha + \sum_{i=1}^{p} \delta_i \Delta x_{t-i} + \varepsilon_t$ и проверяем значимость свободного члена по обычной t-статистике и распределению Стьюдента. Часто длина реализации составляет 70–80, тогда можно проверять гипотезу по нормальному распределению. Если мы считаем, что α не значимо, то не должны его включать в уравнение для ADF теста. В этом случае оцениваем регрессию $\Delta x_t = (\rho - 1)x_{t-1} + \sum_{i=1}^{p} \delta_i \Delta x_{t-i} + \varepsilon_t$. Для проверки нулевой гипотезы о стационарности используем критические значения τ_0.

Проиллюстрируем применение этой процедуры примером ранее рассматриваемого нами ряда индекса деловой активности для Великобритании (UK FTA All Share) [9]. Количество лагов было выбрано равным трем.

1. Методом наименьших квадратов была оценена следующая модель:

$$\underset{t}{\Delta x_t} = \underset{(2,31)}{0{,}124} + \underset{(2,29)}{0{,}00025} \cdot t - \underset{(-2,27)}{0{,}0284 x_{t-1}} + \sum_{i=1}^{3} \hat{\delta}_i \cdot \Delta x_{t-i} + e_t$$ (в скобках – t-отношения),

где t-отношение для коэффициента при x_{t-1} равно −2,27. Мы должны это число сравнить со значением τ_τ на уровне значимости 5%. Оно равно −3,49. Мы видим, что −2,27>−3,49. Поэтому мы не можем отвергнуть нулевую гипотезу, которая заключается в том, что ряд принадлежит к типу DS, т.е. что он имеет единичный корень.

2. На втором шаге оцениваем регрессию вида:

$$\underset{t}{\Delta x_t} = -\underset{(-0,27)}{0{,}0023} + \underset{(1,18)}{0{,}00007t} + \sum_{i=1}^{3} \hat{a}_i \cdot x_{t-i} + e_t,$$

т.е. из числа регрессоров убираем регрессор x_{t-1}. Мы сравниваем t-статистику коэффициента при линейном тренде с таблицами нормального распределения. Видно, что коэффициент не значим, следовательно, тренд был включен напрасно, т.е. его включение не согласуется с нашими реальными данными. Поэтому переходим к третьему шагу.

3. Оцениваем регрессию вида: $\underset{t}{\Delta x_t} = \underset{(0,63)}{0{,}0166} - \underset{(-0,38)}{0{,}00176} x_{t-1} + \sum_{i=1}^{3} \hat{b}_i \Delta x_{t-i} + e_t$.

Сравниваем t-отношение при регрессоре x_{t-1} с τ_μ. Поскольку t-отношение очень маленькое, мы не можем отвергнуть гипотезу о наличии единичного корня. Но, может быть, мы зря включили свободный член в нашу модель.

4. Оцениваем модель $\Delta x_t = \underset{(1,78)}{0{,}0067} + \sum_{i=1}^{3} c_i \Delta x_{t-i} + e_t$. Здесь t-статистика, равная 1,78, при сравнении с критической величиной стандартного нормального распределения оказывается значимой на 5-процентном уровне по одностороннему критерию. Раз коэффициент значим, то количество лагов выбрано правильно, эта модель правильно специфицирована.

Общий вывод применения процедуры Доладо, Дженкинсона и Сосвилла-Риверо заключается в том, что ряд содержит единичный корень, он относится к типу DS. Также по ходу ее применения выяснилось, что ряд не содержит линейного тренда.

Если бы на четвертом шаге мы проверяли нулевую гипотезу по двустороннему критерию, то тогда бы должны были сказать, что свободный член не значим, но, скорее всего, пришли бы к тому же самому выводу о наличии единичного корня. Но спецификация модели не содержала бы свободный член.

Мы видим, что установление типа нестационарности ряда не сводится к однократному применению соответствующего теста. Требуется тщательное исследование правильности спецификации тестовой модели. Описанная методика исследования наличия единичного корня не является единственной. Другие подходы можно найти в монографии Паттерсона [12].

У нас осталось еще два момента, связанных с различением между рядами типа DS и TS. Первый вопрос поставил Перрон. Правильно ли то, что мы сравниваем только такие простые модели, у которых детерминированный тренд на всем временном промежутке один и тот же, в то время как стохастический тренд меняется от точки к точке. В качестве более общей альтернативы Перрон предложил рассмотреть следующую модель. У процесса есть линейный тренд, но он может структурно меняться под влиянием каких-то шоков, т.е. детерминированный тренд претерпевает структурные изменения. Ряд с таким трендом конечно нестационарный, но он разбивается на два временных участка, на каждом из которых ряд стационарен или, по крайней мере, относится к типу TS. Но происходит изменение параметра, изменение структуры модели. Имеет смысл рассмотреть такую сегментированную модель в качестве альтернативы стохастическому тренду. При этом ясно, что в некоторых случаях момент такого структурного скачка известен заранее, в других случаях − нет, и что таких «переломов» может быть, вообще говоря, несколько.

И еще один момент. Мы говорили, что ищем единичный корень, но на самом деле пытались различить процессы типа I(0) и I(1). Если процесс содержит несколько единичных корней, то такая возможность пока не рассматривалась. А ведь такой случай не исключен. Мы уже рассматривали процессы ARIMA(p,d,q) и говорили, что параметр d может быть любым числом. Хотя в экономических задачах параметр d больше чем 2 практически не встречается. Но, по крайне мере, процесс может содержать два единичных корня. В этом случае появляется третья возможность: ряд относится к типу I(2). Классическая процедура проверки статистических гипотез проводится таким образом, что третья возможность фактически «объединяется» с альтернативной гипотезой. Это ведет к увеличению ошибки второго рода и потере эффективности статистического теста. Рассмотрим эту проблему подробнее.

В наиболее общем виде ADF-тест имеет вид:

$$\Delta x_t = \alpha + \beta t + (\rho - 1)x_{t-1} + \sum_{i=1}^{p} \delta_i \Delta x_{t-i} + \varepsilon_t .$$

В качестве гипотез мы фактически рассматривали следующие:

$$H_0 : d = 1;$$
$$H_1 : d = 0 .$$

В классической процедуре проверки гипотез ищется критическое множество при фиксированной величине ошибки 1-ого рода $P\{x \in S | H_0\} = \alpha$. Критическое множество строится, чтобы разделить две гипотезы: основную и альтернативную. Есть два случая. Первый, когда мы точно знаем, что либо d=0, либо d=1. То есть у нас может быть максимально один единичный корень. И вторая ситуация, когда у нас может быть ни одного, один или два единичных корня. Что из общих соображений должно произойти в этом случае? В первом случае, если не верна нулевая гипотеза, мы точно знаем, что d=0. А во втором случае, если не верна нулевая гипотеза, то либо d=0, либо d=2. Поэтому критические значения должны измениться, у нас будет другая структура критического множества. Фактически, когда строился ADF-тест, мы проверяли несколько иную гипотезу, что параметр d не больше единицы. А альтернативная гипотеза, что нет ни одного единичного корня.

Предположим, мы хотим проверить, имеет ли процесс x_t два единичных корня. Естественным путем представляется двукратное применение ADF-теста. Если нулевая гипотеза ADF-теста не отвергается, надо рассмотреть ряд Δx_t и исследовать его на наличие единичного корня. Но возникает вопрос: может быть, процесс содержит не два, а три или четыре единичных корня? У нас есть два конкурирующих разумных подхода. Первый заключается в продолжении выше намеченной процедуры. Если обнаруживается наличие единичного корня, то тогда берем первую разность Δx_t и исследуем, нет ли у преобразованного процесса единичного корня. Если есть, то переходим ко второй разности и вновь проверяем наличие единичного корня у преобразованного ряда. И так далее, пока гипотеза о наличии единичного корня не будет отвергнута. При этом подходе мы последовательно проверяем гипотезы о наличии все большего числа единичных корней.

Второй разумный подход состоит в проверке гипотез в обратной последовательности. Сначала выдвигается предположение о максимально возможном числе единичных корней, например равном трем. В этом случае берем вторую разность процесса и проверяем преобразованный ряд на наличие единичного корня. Если у преобразованного ряда нет единичного корня, это означает, что возможность наличия трех единичных корней у исходного процесса исключена. Если нет трех единичных корней, переходим к проверке на наличие двух единичных корней, для чего проверяем наличие единичного корня у первой разности исходного ряда. Этот подход характеризуется последовательной проверкой гипотез о наличии все меньшего числа единичных корней. На каждом шаге основная и альтернативная гипотезы формулируются следующим образом:

$$H_0 : \max \quad k \quad \textit{единичных корней};$$

$$H_1 : \textit{меньше чем} \quad k \quad \textit{единичных корней}.$$

Диккки и Пантула [5] показали, используя моделирование по методу Монте-Карло, что второй подход эффективнее.

Схема его применения для случая, например, трех возможных единичных корней такова.

Применяем обычный ADF-тест для второй разности исходного процесса, т.е. оцениваем регрессию вида $\Delta^2 x_t = \alpha + \beta t + (\rho - 1)\Delta x_{t-1} + \sum_{i=1}^{p} \delta_i \Delta^2 x_{t-i} + \varepsilon_t$. При этом нулевая и альтернативная гипотезы в терминах исходного процесса выглядят следующим образом.

$$H_0 : 2 \textit{ корня};$$

$$H_1 : 1 \textit{ корень}.$$

Стандартным образом проверяем t-статистику, сравнивая ее с τ_τ, и так далее. Если процедура Доладо привела к выводу, что процесс содержит три единичных корня, то процедура окончена. Если же гипотеза о наличии трех единичных корней отвергается, применяем процедуру Доладо к первой разности исходного процесса, что эквивалентно сравнению гипотез о наличии двух и одного единичного корня. Если гипотеза о наличии двух единичных корней отвергается, то переходим к ADF-тесту для исходного ряда.

Лекция № 10

Вернемся к подходу Перрона, упомянутому в предыдущей лекции. Все альтернативы, которые мы до сих пор рассматривали при различении TS и DS моделей, исходили из того, что параметры, которыми модели характеризуются, в частности свободный член и коэффициент при линейном тренде, сохраняются неизменными все время наших наблюдений. Однако исторические макроэкономические ряды подсказывают, что может наблюдаться и другая ситуация. Тренд или свободный член могут скачком изменить свое численное значение, реагируя на так называемые внешние шоки. Такими внешними шоками являются, например, великая депрессия 1929 г. в США, нефтяной шок 1973 г., а теперь и российский дефолт августа 1998 г. Модели с такими скачкообразными изменениями параметров называют моделями со структурными скачками (*structural breaks*). Поскольку стандартный ADF-тест не допускает возможности таких структурных скачков, существует опасность, что скачки будут трактоваться как нестационарность, и нулевая гипотеза не будет отвергнута.

Нам хотелось бы расширить наше рассмотрение на естественные классы моделей, которые содержат скачкообразные изменения параметров. При этом время структурного скачка может быть известным (экзогенный скачок) или неизвестным (эндогенный скачок). Начнем с простой ситуации, когда изменение происходит в некоторый известный момент. В 1989 г. Перрон [13] сумел распространить схему исследования типа ряда (TS или DS) на случай структурных изменений. Если модель типа TS (альтернативная в ADF-тесте) включает свободный член и линейный тренд, то внешний шок может вызвать три типа структурных изменений. В некоторый момент времени может произойти скачок уровня, и линия тренда пойдет дальше с тем же наклоном, т.е. с тем же темпом роста. Может произойти скачок наклона линии тренда без изменения уровня ряда. А может произойти ситуация, когда изменился наклон и одновременно произошел сдвиг. При этом следует исследовать уже кусочно-линейные модели. Вообще говоря, эта модель не является линейной, но на каждом из двух интервалов она линейна.

Перрон специфицировал модели, которые соответствуют такому интуитивному пониманию трех указанных случаев для случая ряда типа TS. Для каждого такого случая он выписал также альтернативную модель типа DS с качественно похожим поведением. Напомню, что в тесте Дикки–Фуллера тип DS соответствует нулевой гипотезе, а тип TS соответствует альтернативной гипотезе:

$$H_0 : DS$$

$$H_1 : TS \ .$$

Перрон рассмотрел 3 типа моделей: A, B и C, которые соответствуют трем различным типам структурных изменений. Сведем эти модели в таблицу.

	$H_1 : TS$	$H_0 : DS$
A	$x_t = \mu_1 + \beta \cdot t + (\mu_2 - \mu_1) \cdot (DU)_t + \varepsilon_t$	$x_t = \mu + x_{t-1} + d \cdot (t_B)_t + \varepsilon_t$
B	$x_t = \mu + \beta_1 \cdot t + (\beta_2 - \beta_1)(DT^*)_t + \varepsilon_t$	$x_t = \mu_1 + x_{t-1} + (\mu_2 - \mu_1)(DU)_t + \varepsilon_t$
C	$x_t = \mu_1 + \beta_1 \cdot t + (\mu_2 - \mu_1)(DU)_t +$	$x_t = \mu_1 + x_{t-1} + d \cdot D(t_B)_t +$
	$+ (\beta_2 - \beta_1)(DT^*)_t + \varepsilon_t$	$+ (\mu_2 - \mu_1) \cdot (DU)_t + \varepsilon_t$

Первая модель (тип A) – это случай, когда происходит изменение свободного члена скачком в некоторый момент времени t_B+1. Следуя Перрону, введем следующие обозначения. $(DU)_t$ – это единый символ, который является обычной dummy-переменной, определенной как: $(DU)_t = \begin{cases} 1, & t > t_B, \\ 0, & t \le t_B \end{cases}$. Эта dummy-переменная несколько непривычно определена, знак «строго больше» означает скачок в момент t_B+1. Причем, видно, что до скачка в модели TS «работает» свободный член уровня μ_1, а после скачка «работает» уровень μ_2. Это совершенно стандартная dummy-переменная.

Для этого же случая (тип A) нулевую модель DS можно написать в следующем виде: $x_t = \mu + x_{t-1} + d \cdot (t_B)_t + \varepsilon_t$, где d – это некий коэффициент, и введена еще одна dummy-переменная $D(t_B)_t = \begin{cases} 1, & t_B + 1 \\ 0, & \textit{в остальных случаях} \end{cases}$. То есть альтернативная гипотеза – это случайное блуждание с дрейфом, как и положено. Дрейф в этой модели равен μ во все моменты времени, кроме момента структурного скачка. Обратите внимание, что $\mu \neq \mu_1$, это разные параметры. Эти модели, обе типа А, описывают одно и то же явление, но по-разному. Одна при детерминированном скачке в случайном члене, а другая при стохастическом. Dummy-переменная обращается в единицу для одного единственного наблюдения. Если описывать «черный вторник» в России, когда курс доллара к рублю скачком возрос, а на следующий день вернулся к практически прежнему значению, то пригодится как раз такая переменная.

Вторая модель (тип В) – это излом наклона линейного тренда. Ее детерминированный вариант нам знаком: $x_t = \mu + \beta_1 \cdot t + (\beta_2 - \beta_1)(DT^*)_t + \varepsilon_t$.

Причем новая dummy-переменная определяется следующим образом:

$$(DT^*)_t = \begin{cases} t - t_B, & t > t_B \\ 0, & t \le t_B \end{cases}.$$

Это обычная смешанная dummy-переменная, тренд с началом отсчета в момент t_B, умноженный на ранее введенную dummy-переменную:

$$(t - t_B)(DU)_t = (DT^*)_t.$$

Просто Перрон ввел для нее специальное обозначение.

Модель типа В для ряда DS означает, что параметр дрейфа изменился и остался измененным, а в модели типа А он изменился на один период и вернулся к старому значению. Для ряда DS модель типа В описывает случайное блуждание с разными дрейфами $x_t = \mu_1 + x_{t-1} + (\mu_2 - \mu_1)(DU)_t + \varepsilon_t$.

Третья модель (тип С) – это самый общий случай, когда происходят оба изменения, и уровня ряда как в модели типа А, и наклона тренда как в модели типа В.

Используя моделирование по методу Монте-Карло, Перрон показал, что если данные сгенерированы одной из моделей из левого столбца, т.е. являются процессом типа TS, то МНК-оценка коэффициента $\hat{\gamma}$ в регрессии $x_t = \mu_1 + \gamma x_{t-1} + \beta t + \varepsilon_t$ все сильнее смещена к единице по мере увеличения величины скачка. Это означает, что ADF-тест будет ошибочно диагностировать наличие единичного корня, подтверждая ранее высказанное опасение. Поскольку смещение не исчезает при увеличении длины ряда, говорят, что ADF-тест несостоятелен при структурном скачке тренда. Для модели типа А, когда присутствует только скачок свободного члена, ADF-тест, вообще говоря, состоятелен (по результатам Перрона), но имеет низкую мощность. Этот результат не является неожиданным, потому что в случае присутствия единичного корня добавка одного регрессора, как мы уже видели, часто меняет выборочную статистику, наверно и здесь так.

Перрон предложил обобщение стратегии тестирования на наличие единичного корня применительно к возможному присутствию структурных скачков. На первом шаге предложенной процедуры исходный ряд «очищается» от тренда со структурным скачком, для чего находятся остатки регрессии исходного ряда на:

а) константу, линейный тренд и переменную $(DU)_t$;

в) константу, линейный тренд и переменную $(DT^*)_t$;

с) константу, линейный тренд, переменную $(DU)_t$ и переменную $(DT^*)_t$, для вышеупомянутых моделей типа А, В и С соответственно. К остаткам от регрессии применяется тест Дикки–Фуллера (не расширенный), но критические точки распределения t-отношения не совпадают со значениями, рассчитанными Фуллером. Поэтому Перрон рассчитал их сам. Оказалось, что эти критические значения больше по абсолютной величине, чем соответствующие значения Дикки–Фуллера, и зависят не от длины реализации T, а от месторасположения момента скачка относительно длины ряда. Попробуйте включить интуитивное понимание. Представьте себе, что имеется длинная серия наблюдений и скачок произошел в самом ее начале, и второй случай, когда скачок произошел в середине. Как вы думаете, когда легче отличить, был скачок или нет? Ответ ясен: во втором случае. Фактически мы сравниваем две модели: на одном временном интервале и на другом. Если интервал маленький, различие как бы не успело набрать силу, нам будет сложно сравнивать. Поэтому критические значения теста Перрона зависят от соотношения $\left(\frac{t_B}{T}\right)$. При любом месте расположения скачка, как уже упоминалось, критические значения статистики Перрона по абсолютной величине больше, чем соответствующие для DF-распределения. И самые большие по модулю критические значения соответствуют $\left(\frac{t_B}{T}\right)$= 0,5. Например, для уровня значимости 5% критическое значение, полученное Перроном для модели типа А при $\left(\frac{t_B}{T}\right)$= 0,5, составило −3,76, а соответствующее значение статистики Дикки–Фуллера составляет −3,41.

Перрон применил свой подход к историческим макроэкономическим рядам США и выяснил, что допущение возможности структурных скачков изменяет результаты Нельсона и Плоссера, уже не все ряды относятся к типу DS.

Очевидным недостатком изложенного подхода является предположение, что в ответ на внешний шок происходит мгновенное изменение параметра (параметров) модели. Более реалистичным представляется постепенное изменение параметров как реакция на внешний шок. Другими словами, изменение имеет лаговую структуру, может быть, бесконечную. Реакцию первого типа в литературе принято называть аддитивным выбросом (additive outlier, AO), а реакцию второго типа – инновационным выбросом (innovation outlier, IO). В своей работе Перрон предложил для случая инновационного выброса оценить регрессию (для наиболее общего случая – типа С):

$$x_t = \alpha + \theta(DU)_t + \beta t + \delta(DT^*)_t + \gamma x_{t-1} + d \cdot D(t_B)_t + \sum_{i=1}^{k} c_i \Delta x_{t-i} + \varepsilon_t .$$

Нулевой гипотезе наличия единичного корня соответствуют значения параметров $\gamma = 1; \theta = \beta = \delta = 0$, альтернативной гипотезе – $\gamma < 1; \theta, \beta, \delta \neq 0$. Кроме того, при нулевой гипотезе параметр d не должен быть равен 0, в то время как при альтернативной гипотезе он должен быть близок к 0. Для типов В и С число регрессоров соответственно уменьшается. Чрезвычайно удобным для пользователя процедуры Перрона является тот факт, что критические значения, рассчитанные Перроном, являются одними и теми же как для аддитивных, так и для инновационных выбросов.

Одним из обобщений подхода Перрона было распространение его на случай нескольких структурных скачков в разные моменты времени. Раппопорт и Рейхлин [15] рассмотрели случай двух скачков в тренде, а в работах Бая [1] и Бая с Перроном [2] предложенная методика распространена на случай нескольких структурных скачков. Ввиду громоздкости и необходимости набора таблиц новых критических значений большого применения это обобщение пока не нашло.

Слабым местом подхода Перрона является предположение о том, что время структурного скачка известно заранее. Зивот и Эндрюс [18] показали, что момент скачка не может быть предопределен заранее. Они предложили обобщение подхода Перрона на случай эндогенного скачка, т.е. структурного скачка, момент которого оценивается одновременно с оценкой коэффициентов модели. Зивот и Эндрюс полагают, что нет необходимости вводить структурный скачок в нулевую модель, и рассматривают большое по величине изменение уровня просто как осуществление редкого события, соответствующего «хвосту» распределения. Поэтому нулевая гипотеза в их подходе представляет собой случайное блуждание с дрейфом. Альтернативная модель имеет вид:

$$x_t = \alpha + \theta(DU)_t + \beta t + \delta(DT^*)_t(\lambda) + \gamma x_{t-1} + \sum_{i=1}^{k} c_i \Delta x_{t-i} + \varepsilon_t ,$$

сходный с подходом Перрона, но без dummy-переменной $D(t_B)_t$ и с дополнительным параметром λ, равным отношению неизвестного момента скачка к длине ряда. Дальнейшее напоминает поиск на решетке. Строятся альтернативные модели для всевозможных моментов времени скачка от t=2 до t=T−1 и выбирается та из них, для которой t-отношение коэффициента γ принимает наименьшее значение. Не забудьте, эта величина отрицательна.

Поскольку момент скачка оценивается процедурой последовательного перебора, а не известен заранее, критические значения, рассчитанные Перроном, больше не подходят. Зивот и Эндрюс, а также Банерджи с соавторами [3] рассчитали, используя метод Монте-Карло, критические значения для этого подхода. Они лежат еще левее критических значений Перрона. Так, например, 5-процентное критическое значение Перрона для заданного значения параметра λ=0,5 составляет −4,24, а соответствующее значение для оцененного значения параметра равно −5,08. Применив свой подход к все тем же данным Нельсона и Плоссера, Зивот и Эндрюс в четырех из десяти случаях установили в отличие от Перрона наличие единичного корня.

В качестве иллюстрации рассмотрим интересный пример применения этой методики (см. рис. 10.1). Рассмотрим сначала его участок с 1971 по 1988 гг. [9]. На этом участке ряд демонстрирует некоторое подобие наличия структурного скачка. Если применить к этому ряду обычный ADF-тест на наличие единичного корня, то $\tau_\tau = -1{,}84$, а соответствующее критическое значение на 10-процентном уровне значимости равно −3,15, так что мы не можем отвергнуть нулевую гипотезу о наличии единичного корня.

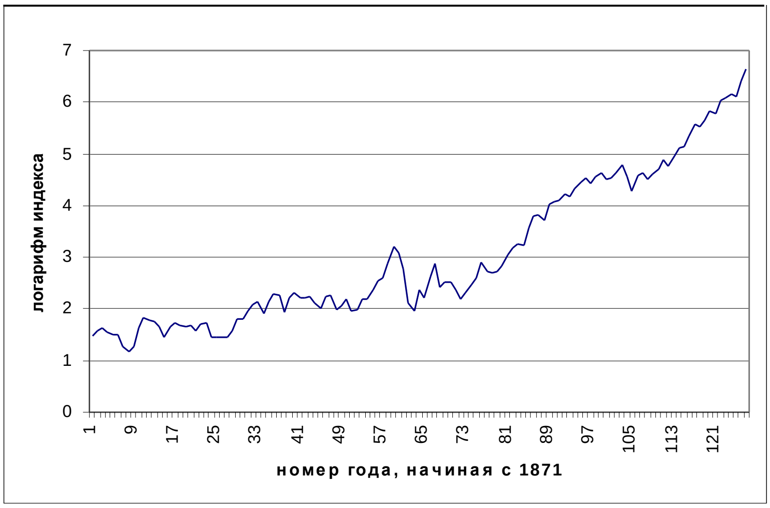

Рис. 10.1. Логарифм среднегодовых значений индекса S&P для США в 1871–1997 гг.[1)]

Следуя Перрону, попробуем включить в рассматриваемые гипотезы наличие структурного скачка. Естественно предположить его в 1929 г. (58-е значение ряда), то, что в Америке называют Great Crash, или великая депрессия. Отношение времени момента структурного скачка к общей длине ряда (118 точек) равно примерно 0,5. Оценка модели типа C дает (в скобках – значения стандартных ошибок):

$$\underset{s.e.}{x_t} = \underset{(0,117)}{0{,}570} + \underset{(0,0017)}{0{,}0070} \cdot t - \underset{(0,068)}{0{,}216}(DU)_t + \underset{(0,003)}{0{,}011}(DT^*)_t - \underset{(0,185)}{0{,}177}\, D(t_B)_t + \underset{(0,062)}{0{,}707}\, x_{t-1} + \underset{(0,092)}{0{,}156}\, \Delta x_{t-1} \cdot$$

Нулевая гипотеза по-прежнему имеет вид: $H_0 : DS \Rightarrow \gamma = 0$. Уравнение ADF-теста записано не в стандартном виде. Чтобы привести его к стандартному виду, надо вычесть из оценки коэффициента единицу, поэтому получаем: $t = \dfrac{0{,}707 - 1}{0{,}062} = -4{,}74$.

Пятипроцентное критическое значение из таблиц Перрона для $\lambda = \dfrac{t_B}{T} = 1/2$ равно −4,24. Поэтому нулевую гипотезу о наличии единичного корня нужно отвергнуть на этом уровне значимости.

Поскольку все коэффициенты, кроме коэффициента при dummy-переменной $D(t_B)_t$, статистически значимы, пересчитаем модель, исключив эту переменную. Построенная модель содержит авторегрессию второго порядка. Как мы уже видели раньше, можно представить модель в эквивалентном виде без авторегрессии, но с автокорреляцией второго порядка у случайной составляющей. Такое представление в этом случае является более удобным для интерпретации. В результате, оценивая, получаем:

[1)] Источник данных: файл S&P500 на сайте http://www.lboro.ac.uk/departments/ec/cup/

$$\underset{s.e.}{x} = \underset{(0,085)}{1,859} + \underset{(0,0047)}{0,0188} \cdot t + \underset{(0,007)}{0,0366}(DT^*)_t - \underset{(0,152)}{0,312}(DU)_t + e_t \cdot$$

Наши остатки стали коррелированными, т.е. $e_t = \underset{(0,093)}{0,932}\, e_{t-1} - \underset{(0,094)}{0,200}\, e_{t-2} + u_t$. Причем $\hat{\sigma}_u = 0,1678$. Что можно заключить по этой модели? До структурного скачка наблюдается значимый линейный тренд, он соответствует приросту индекса на 1,88% в год. Что произошло в момент скачка? С 1930 г. темп роста изменился. Он вырос на 3,66%, и стал 5,55%. Изменение тренда одновременно сопровождалось скачком вниз на 0,312.

К этому же участку ряда была применена методика Зивота и Эндрюса, т.е. дополнительно оценивался момент структурного скачка. В этом случае скачок обнаружился совсем в другой точке, он соответствует 1950 г. И модель после определения времени скачка приняла вид: $\underset{s.e.}{x_t} = \underset{(0,064)}{1,871} + \underset{(0,0024)}{0,0193} t + \underset{(0,0069)}{0,0391}(DT^*)_t + \underset{(0,144)}{0,377}(DU)_t + u_t \cdot$

Здесь u_t тоже подчиняется авторегрессионной схеме второго порядка, мы опускаем это уравнение. Согласно этой модели скачок произошел в 1950 г., и затем темп роста стал больше, чем 1,93%, а точнее возрос на 0,0391. Обратите внимание, что скачок в свободном члене тоже положительный.

Даже на этом примере видна важность учета структурных изменений: если не учитывать структурного скачка, исследуемый ряд относится к типу DS с единичным корнем. Учет структурного скачка позволил описать ряд процессом типа TS. Если искать по формальной процедуре период скачка, то мы нашли другое значение, причем его трудно интерпретировать. Если наложить на график обе модели, то расхождение между ними качественно заметно только между 1929 и 1950 гг. На промежутках до и после этого периода модели хорошо согласуются с рядом и между собой. Если мы занимаемся прогнозированием на ближайшее время, обе эти модели приведут примерно к одному и тому же, если же будем темп роста экстраполировать на длительный период, то, конечно, придем к разным результатам.

Интересно сравнить результаты аналогичного подхода применительно ко всему диапазону с 1871 по 1997 гг. [10]. Стандартный ADF-тест на наличие единичного корня теперь дает величину $\tau_\tau = -1,15$, критическое значение не изменилось, так что по-прежнему мы не можем отвергнуть нулевую гипотезу о наличии единичного корня.

Применив подход Перрона со структурным скачком в 1929 г., получаем следующую модель (в скобках − значения стандартных ошибок):

$$\underset{s.e.}{x_t} = \underset{(0,089)}{0,334} + \underset{(0,0017)}{0,0066} \cdot t - \underset{(0,065)}{0,235}(DU)_t + \underset{(0,003)}{0,012}(DT^*)_t + \underset{(0,181)}{0,184} D(t_B)_t + \underset{(0,058)}{0,731} x_{t-1} + \underset{(0,088)}{0,128} \Delta x_{t-1} \cdot$$

Видно, что коэффициенты модели (за исключением незначимого коэффициента при dummy-переменной $D(t_B)_t$) практически не изменились. Значение t-отношения теперь составило −4,65, и нулевая гипотеза вновь отвергается на 5-процентном уровне значимости.

Пересчитаем модель, исключив переменную $D(t_B)_t$ и перенеся лаговую структуру в остатки. В результате, оценивая, получаем:

$$\underset{s.e.}{x} = \underset{(0,206)}{1,335} + \underset{(0,0053)}{0,0175} t + \underset{(0,0074)}{0,0430}(DT^*)_t - \underset{(0,156)}{0,346}(DU)_t + e_t.$$

Остатки описываются моделью $e_t = \underset{(0,091)}{0,945} e_{t-1} - \underset{(0,092)}{0,177} e_{t-2} + u_t$. По-прежнему $\hat{\sigma}_u = 0,1678$.

По сравнению с аналогичной моделью для 1871–1988 гг. увеличение тренда несколько больше: не на 3,66 %, а на 4,3%. Изменение тренда по-прежнему сопровождается скачком вниз, но теперь на 0,346. По графику можно понять, что причиной этих изменений в параметрах модели является более быстрый рост индекса в 1989–1997 гг. В целом подход Перрона дал сопоставимые результаты для обоих участков.

Совсем по-другому обстоит дело с методикой Зивота и Эндрюса. В этом случае скачок обнаружился в точке, соответствующей 1931 г., что весьма близко к великой депрессии. При этом нулевая модель о наличии единичного корня по-прежнему отвергается. Такое расхождение в оценках времени скачка при добавлении относительно небольшого числа точек ряда ставит интересные вопросы о доверительном интервале этой оценки и о робастности всей процедуры Зивота и Эндрюса, но они выходят за рамки нашего курса.

* *
*

СПИСОК ЛИТЕРАТУРЫ

1. *Bai J.* Estimating Multiple Breaks One at a Time // Econometric Theory. 1997. 13. P. 315–52.
2. *Bai J.* and *Perron P.* Estimating and Testing Linear Models with Multiple Structural Changes // Econometrica. 1998. 66. P. 47–78.
3. *Banerjee A., Lumsdaine R.L.* and *Stock J.H.* Recursive and Sequential Test of the Unit Root and Trend Break Hypothesis: Theory and International Evidence // Journal of Business and Economic Statistics. 1992. 10. P. 271–87.
4. *Dickey D.A.* and *Fuller W.A.* Distribution of the Estimators for Autoregressive Time-Series with a Unit Root // Journal of the American Statistical Assiciation. Vol. 74. 1979. P. 427–431.
5. *Dickey D.A.* and *Pantula S.G.* Determining the Order of Differencing in Autoregressive Processes // Journal of Business and Economic Statistics. 1987. Vol. 5. P. 455–461.
6. *Diebold F.X.* and *Nerlove M.* Unit Roots in Economic Time Series: A Selective Survey // Rhodes G.F. and Fomby T.B.(eds.). Advances in Econometrics. Vol. 8. Greenwich, CT: JAI Press. 1990. P. 3–69.
7. *Dolado J.J., Jenkinson T.* and *Sosvilla-Rivero S.* Cointegration and Unit Roots // Journal of Economic Survey. 1990. Vol. 4. P. 249–73.
8. *Johnston* and *DiNardo J.* Econometric Methods. Fourth Edition. The McGraw-Hill Companies, Inc., 1977.
9. *Mills T.C.* The Econometric Modelling of Financial Time Series. Cambridge University Press, 1993.
10. *Mills T.C.* The Econometric Modelling of Financial Time Series. Second Edition. Cambridge University Press, 1999.

11. *Nelson C.R.* and *Kang H.* Pitfalls in the Use of Time as an Explanatory Variable in Regression // Journal of Business and Economic Statistics. January 1984. Vol. 2. P. 73–82.

12. *Patterson K.* An Introduction to Applied Econometrics: A Time Series Approach. Palgrave, 2000.

13. *Perron P.* The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis // Econometrica. 1989. 57. P. 1361–401.

14. *Phillips P.C.B.* and *Perron P.* Testing for Unit Roots in Time Series Regression // Biometrica. 1988. Vol. 75. P. 335–46.

15. *Rappoport P.* and *Reichlin L.* Segmented Trends and Non-Stationary Time Series // Economic Journal. 1989. 99 (Supplement). P. 168–77.

16. *Said S.T.* and *Dickey D.A.* Testing for Unit Roots in Autoregressive Moving-Average Models with Unknown Order // Biometrica. 1984. Vol. 71. P. 599–607.

17. *Schwert G.W.* Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data // Journal of Monetary Economics. 1987. Vol. 20. P. 73–105.

18. *Zivot E.* and *Andrews D.W.K.* Further Evidence on the Great Crash, the Oil Price Shock? And the Unit Root Hypothesis // Journal of Business and Economic Statistics. 1992. 10. P. 251–70.