

I. Некоторые эконометрические методы

Оглавление

ФУНКЦИОНАЛЬНАЯ ФОРМА РЕГРЕССИОННОЙ МОДЕЛИ	2
✓ Тестирование правильности спецификации регрессионной модели.....	2
✓ Линейные и нелинейные модели.....	3
✓ Выбор между альтернативными функциональными формами	7
ФИКТИВНЫЕ ПЕРЕМЕННЫЕ КАК РЕГРЕССОРЫ	10
✓ Общие соображения	10
✓ Использование фиктивных переменных для проверки однородности наблюдений и прогнозирования	12
✓ Использование фиктивных переменных в моделях с временными рядами	14
✓ Спектральный анализ и регрессия	16
МОДЕЛИ С КАЧЕСТВЕННОЙ ЗАВИСИМОЙ ПЕРЕМЕННОЙ	18
✓ Модели с бинарной зависимой переменной	18
✓ Модель выбора. Пробит и логит	19
✓ Оценка качества модели и проверка гипотез	20
✓ Множественные модели с качественными зависимыми переменными	23
ДИНАМИЧЕСКАЯ СПЕЦИФИКАЦИЯ РЕГРЕССИОННОЙ МОДЕЛИ	25
✓ Модель распределенного лага	25
✓ Динамические регрессионные модели. Авторегрессионная модель с распределенным лагом.....	28
ИНТЕГРИРОВАННЫЕ ПРОЦЕССЫ, ЛОЖНАЯ РЕГРЕССИЯ И КОИНТЕГРАЦИЯ	32
✓ Стационарные и нестационарные случайные процессы.....	32
✓ Ложная регрессия.....	34
✓ Тестирование стационарности.....	36
✓ Коинтеграция. Регрессии с интегрированными переменными	41
✓ Оценивание коинтеграционной регрессии: подход Энгла-Грейнджера	43
✓ Коинтеграция в динамических системах: подход Йохансена	45
✓ Литература по единичным корням и коинтеграции	49
ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ	51

Функциональная форма регрессионной модели

Необходимость изменить функциональную форму модели возникает, если неверна одна из следующих гипотез, выполнение которых требуется для того, чтобы обычный метод наименьших квадратов (ОМНК) в применении к регрессионной модели $Y_i = X_i\beta + \varepsilon_i$ ($i = 1, \dots, N$) давал хорошие результаты:¹

1. Ошибки имеют нулевое математическое ожидание, или, что то же самое, мат. ожидание зависимой переменной является линейной комбинацией регрессоров:

$$E(\varepsilon_i) = 0, \quad E(Y_i) = X_i\beta.$$

2. Ошибки гомоскедастичны, т. е. имеют одинаковую дисперсию для всех наблюдений:

$$V(\varepsilon_i^2) = E(\varepsilon_i^2) = \sigma^2.$$

Тестирование правильности спецификации регрессионной модели

Если ошибка имеет ненулевое мат. ожидание, то оценки ОМНК окажутся смещенными. Другими словами, в ошибке осталась детерминированная (неслучайная) составляющая, которая может быть функцией входящих в модель регрессоров, что и означает, что функциональная форма выбрана неверно. Заметить эту ошибку спецификации можно на глаз с помощью графиков остатков по “подозрительным” переменным: регрессорам и их функциям (в т. ч. произведениям разных регрессоров), расчетным значениям и их функциям. Остатки дают представления об ошибках, поэтому они должны в правильно заданной регрессии

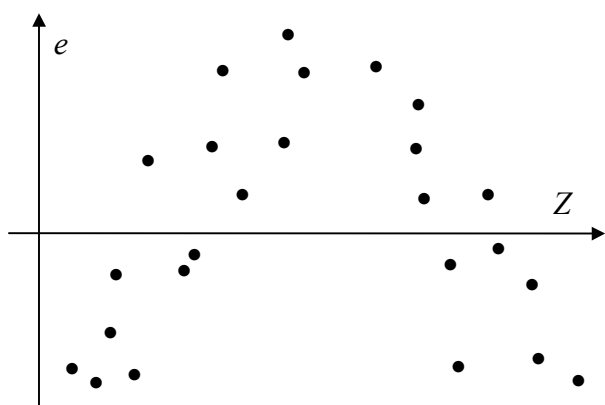


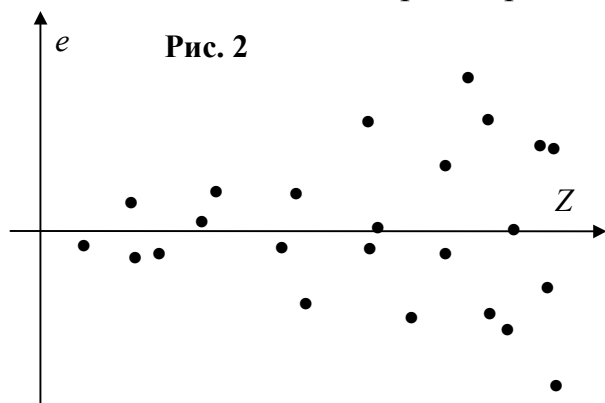
Рис. 1

иметь везде нулевое среднее. Если остатки (e), например, для каких-то значений некоторой переменной Z в среднем больше нуля, а для каких-то — меньше, то это служит признаком неправильно специфицированной модели (см. Рис. 1).

Остатки дают представления об ошибках, поэтому они должны в правильно заданной регрессии

¹ Это, конечно, не все обычно принимаемые гипотезы, а только те, которые интересны с точки зрения рассматриваемой темы.

Похожим образом обнаруживается и гетероскедастичность (отсутствие гомоскедастичности). Она проявляется в том, что разброс остатков меняется в зависимости от некоторой переменной Z (см. Рис. 2)



Дисперсия ошибок может меняться в зависимости от регрессоров и их функций, расчетных значений и их функций. Формальный тест можно провести с помощью вспомогательной регрессии — регрессии квадратов остатков по “подозрительным” переменным и константе. Соответствующая статистика

— обычная F -статистика для гипотезы о равенстве нулю коэффициентов при всех переменных кроме константы, выдаваемая любым статистическим пакетом.

Ошибки в спецификации функциональной формы обнаруживаются также тестами на автокорреляцию остатков, такими как статистика Дарбина-Уотсона, если наблюдения упорядочены по какому-либо признаку, например, по порядку возрастания одного из регрессоров. Понятно, что это тест неформальный.

Линейные и нелинейные модели

Линейная форма модели в целом является более предпочтительной. Линейные модели оцениваются более простым методом наименьших квадратов. При выполнении некоторого набора гипотез оценки ОМНК для линейной модели обладают рядом хороших свойств, не выполняющихся для оценок нелинейной модели, это же относится к распределениям оценок и различным статистикам.

В линейной регрессионной модели мат. ожидание зависимой переменной — это линейная комбинация регрессоров с неизвестными коэффициентами, которые и являются оцениваемыми параметрами модели. Такая модель является **линейной по виду**. В матричной форме ее можно записать как $Y = X\beta + \varepsilon$. Не обязательно, чтобы влияющие на Y факторы входили в модель линейно. Регрессорами могут быть любые точно заданные (не содержащие неизвестных параметров) функции исходных факторов — это не меняет свойств ОМНК. Важно, чтобы модель была **линейной по параметрам**. Бывает, что

модель записана в виде, который нелинеен по параметрам, но преобразованием уравнения регрессии и переобозначением параметров можно привести ее к линейному виду. Такую модель называют **внутренне линейной**.

Поясним введенные понятия на примерах. Модель $Y = \alpha + \beta X_1 X_2 + \varepsilon$ нелинейна по X_1 и X_2 , но линейна по параметрам, и можно сделать замену $X = X_1 X_2$, так что модель примет линейный вид: $Y = \alpha + \beta X + \varepsilon$. Модель $Y = \exp(\alpha + \beta x + \varepsilon)$ нелинейна по виду, но сводится к линейной логарифмированием обеих частей: $\ln Y = \alpha + \beta x + \varepsilon$. В этой новой модели зависимой переменной будет уже $\ln Y$. Модель $Y = (a - 1)(b + X) + \varepsilon$ нелинейна по параметрам a и b , но сводится к линейной заменой параметров $\alpha = (a - 1)b$ и $\beta = a - 1$. Тогда $Y = \alpha + \beta X + \varepsilon$.

Для применения метода наименьших квадратов важно, чтобы ошибка была **аддитивной**, то есть, чтобы зависимая переменная являлась суммой своего математического ожидания и ошибки. Об этом следует помнить, производя преобразования модели. Например, модель $Y = \alpha X^\beta + \varepsilon$ нельзя преобразовать в линейную по параметрам с аддитивной ошибкой. Аналогичную **модель с мультипликативной ошибкой** $Y = \alpha X^\beta \varepsilon$ можно преобразовать к виду $\ln Y = \ln \alpha + \beta \ln X + \ln \varepsilon$ или $\tilde{Y} = \tilde{\alpha} + \beta \tilde{X} + \tilde{\varepsilon}$ где $\tilde{Y} = \ln Y$, $\tilde{\alpha} = \ln \alpha$, $\tilde{X} = \ln X$, $\tilde{\varepsilon} = \ln \varepsilon$. Однако следует отметить, что вследствие преобразования распределение ошибки изменилось. Если $\tilde{\varepsilon}$ оказывается нормально распределенной, это значит, что ε имела логнормальное распределение.

Экономическая теория оперирует моделями разных типов. Некоторые из них дают регрессионные уравнения линейного вида, некоторые – нелинейного. Рассмотрим это на примере однородных производственных функций. Самая популярная производственная функция – **функция Кобба-Дугласа** – легко приводится к линейному виду логарифмированием:

$$Y = \alpha K^\beta L^{1-\beta} \Rightarrow$$

$$\ln Y - \ln L = \ln \alpha + \beta (\ln K - \ln L),$$

где Y – выпуск продукции, K — капитал, L — труд.

Функция с постоянной эластичностью замены (ПЭЗ) дает внутренне нелинейное уравнение регрессии:

$$Y = \alpha (\beta K^\rho + (1-\beta)L^\rho)^{1/\rho}.$$

Достаточно гладкую функцию вблизи некоторой точки можно разложить в ряд Тейлора, получив тем самым линейную форму модели. Так, при ρ

→ 0 функция с постоянной эластичностью замены совпадает с функцией Кобба-Дугласа. Если же приблизить функцию ПЭЗ в точке $\rho = 0$ разложением в ряд Тейлора до членов первого порядка, то получается так называемая **транслоговая производственная функция**:

$$\ln Y - \ln L = \ln \alpha + \beta (\ln K - \ln L) + \gamma (\ln K - \ln L)^2,$$

$$\text{где } \gamma = \frac{1}{2} \rho \beta (1 - \beta).$$

Разложение в ряд Тейлора дает **полиномиальную форму модели**. В полиномиальную регрессионную модель могут входить не только первые степени исходных переменных, но и их одночлены различных степеней: степени этих переменных и **члены взаимодействия** (произведения степеней двух или более различных переменных).

Может случиться: что “истинная” модель бывает настолько нелинейной, что полиномиальное приближение становится неудовлетворительным — количество оцениваемых параметров было бы слишком большим. Тогда приходится пожертвовать удобствами ОМНК и использовать нелинейный МНК или другие методы. Есть также много других причин, по которым предпочтительнее использовать внутренне нелинейную функциональную форму. Например, функция ПЭЗ, рассмотренная выше, включает в себя как частные случаи при разных значениях параметра ρ сразу несколько популярных видов производственных функций: функцию Кобба-Дугласа, линейную функцию (с полной взаимозаменяемостью факторов) и функцию леонтьевского типа (с полной взаимодополняемостью факторов). Оценив ее, можно сделать вывод о том, к какому из этих трех видов ближе “истинная” функция.

Кроме натуральных степеней исходных переменных можно использовать и другие функции от них. Это и уже встречавшиеся выше логарифмы и т. п.: $\ln X$, \sqrt{X} , $1/X$, e^X , $1/(1+e^{-X})$ (логиста) и др. Интересной функцией является **преобразование Бокса-Кокса**: $\frac{X^\alpha - 1}{\alpha}$. При $\alpha \rightarrow 0$ она стремится к $\ln X$. При других значениях это некоторая степень X (с точностью до линейного преобразования). В этом отношении преобразование Бокса-Кокса схоже с функцией ПЭЗ. Оно также похоже на нее в том отношении, что дает внутренне нелинейную модель. Обычно исследователь обладает достаточной свободой при выборе функциональной формы модели. Но важно, чтобы при этом не нарушались те условия, которые необходимы для хорошей работы применяемых методов оценивания. Нужно не забывать проводить проверку пра-

вильности спецификации модели и исправлять модель, когда получена плохая диагностика, например, добавлять одночлены более высоких степеней в полиномиальную модель.

Рассмотрим, как может помочь изменение функциональной формы в борьбе с гетероскедастичностью. Многие экономические переменные таковы, что размер отклонений, с ними связанных, зависит от величины этих переменных (например, пропорционален), а величина эта в выборке колеблется в широких пределах (изменяется в несколько раз). Возникающая при этом гетероскедастичность снижает эффективность оценок параметров. Объяснить потерю эффективности можно следующим образом. В методе наименьших квадратов все наблюдения выступают в одинаковых "весовых категориях", и поэтому в оценках непропорционально мало используется информация от наблюдений с меньшей дисперсией. Тем самым происходит потеря информации. Поэтому, например, нехорошо в регрессию включать временные ряды для номинальных показателей, если в рассматриваемой стране высокая инфляция, или использовать непреобразованную модель в случае выборки стран, в которой есть и большие, и малые страны (США наряду с Исландией). Обычно применяют два вида преобразований. Рассмотрим их на примере функции потребительского спроса кейнсианского типа: $C = \alpha I + \beta X + \varepsilon$, где C — потребление, I — доход, X — символизирует прочие факторы. Разумно предположить, что среднеквадратическое отклонение ошибки прямо пропорционально I .

1) **Нормирование**. Пронормировать рассматриваемую модель можно, разделив ее на I :

$$C/I = \alpha + \beta X/I + \varepsilon / I.$$

Можно использовать для нормировки (взвешивания) и переменную, не входящую в модель. Обозначим ее N :

$$C/N = \alpha + \beta X/N + \varepsilon / N.$$

Нормирование равнозначно использованию взвешенного метода наименьших квадратов. Как веса для номинальных величин можно использовать уровень цен, получив тем самым реальные величины. Как веса для стран можно использовать население, получив тем самым среднедушевые показатели (потребление на душу населения и т. п.).

2) **Логарифмирование**. Прологарифмировав уравнение $C = \alpha I + \varepsilon$ при $\sqrt{V(\varepsilon)} \ll C$ можно получить следующее линейное приближение:

$$\ln C = \ln \alpha + \ln I + \varepsilon / \alpha I.$$

Вряд ли можно привести теоретические возражения и против того, чтобы сразу использовать **линейную в логарифмах модель** (эта форма модели сокращенно называется **логлинейной**), например,

$$\ln C = \alpha + \beta \ln I + \varepsilon.$$

"Кандидатами" на логарифмирование в первую очередь служат те переменные, которые заведомо могут принимать только положительные значения. Один из их признаков, это то, что, как правило, интересуются относительными приростами таких переменных, а не абсолютными приростами. В экономике это следующие величины: физические объемы благ, цены, стоимостные показатели, различные индексы.

Как итог, перечислим основные функциональные формы регрессионной модели (без учета ошибки) с примерами.

Функциональная форма	Пример
Линейная	$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2$
Полиномиальная	$Y = \alpha_0 + \alpha_1 X + \alpha_2 X^2 + \alpha_3 X^3$
	$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_{11} X_1^2 + \alpha_{22} X_2^2 + \alpha_{12} X_1 X_2$
Логлинейная (линейная в логарифмах)	$\ln Y = \alpha_0 + \alpha_1 \ln X$
Мультипликативная	$Y = \alpha_0 X_1^{\alpha_1} X_2^{\alpha_2}$
Нормированная	$Y/N = \alpha_0 + \alpha_1 X/N$

Возможны различные комбинации этих форм. Например, часто встречается **полулогарифмическая форма**:

$$\ln Y = \alpha + \beta X, \text{ или } Y = \alpha + \beta \ln X,$$

$$\text{или } \ln Y = \alpha + \beta \ln X + \gamma Z.$$

Выбор между альтернативными функциональными формами

Самый распространенный способ выбора между альтернативными моделями — выбор на основе точности подбора. В качестве показателя точности подбора обычно используется коэффициент детерминации (R^2). Не следует забывать, что этот показатель можно использовать для сравнения только моделей с одной и той же зависимой переменной. Чтобы учитывать при выборе

простоту модели, делают поправку на количество регрессоров. Это дает коэффициент детерминации скорректированный на количество степеней свободы (\tilde{R}^2).

Оценки метода наименьших квадратов являются одновременно и оценками метода максимального правдоподобия. Поэтому предлагается сравнивать модели на основе максимума логарифмической функции правдоподобия ($\hat{\ell}$). Если учесть при этом количество наблюдений (N) и ввести “штраф” за большое количество регрессоров (k), то получится **информационный критерий Акаике** (Akaike information criterion):

$$\text{AIC} = -2/N (\hat{\ell} - k).$$

Чем меньше AIC, тем лучшей считается модель.

Существует и другой подход к выбору между моделями. Одна из моделей предполагается истинной, т.е. принимается за нулевую гипотезу, и тестируется против некоторой альтернативной гипотезы, спецификация которой зависит от альтернативной модели. По сути дела, осуществляется тестирование функциональной формы “нулевой” модели.

Если одна из моделей является частным случаем другой модели (англ. nested), то в качестве “нулевой” берется более узкая модель, а альтернативой служит более широкая. В случае линейной регрессии применяется соответствующая F-статистика, а в случае нелинейной — одна из χ^2 -статистик: статистика Вальда, множителя Лагранжа или отношения правдоподобия. Если же модели не входят одна в другую (nonnested), то любая из них принимается за нулевую и дополняется за счет информации, содержащейся в другой модели, так, чтобы “нулевая” модель была частным случаем этой расширенной. Здесь уже можно применить один из вышеупомянутых тестов. Если нулевая гипотеза отвергается, то это означает, что альтернативная модель содержит какую-то информацию, не содержащуюся в “нулевой” модели.

Тестов такого рода предложено очень много. Опишем только концептуально наиболее простые.

Сначала рассмотрим случай, когда обе сравниваемые модели линейны и зависимая переменная одна и та же. **J-тест** заключается в том, что в “нулевую” модель добавляется в качестве еще одного регрессора расчетные значения из альтернативной модели. Проверяется гипотеза о равенстве ко-

ээффициента при дополнительном регрессоре нулю с помощью соответствующей t-статистики.

Похожий тест состоит в том, что в “нулевую” модель добавляют из альтернативной модели все те регрессоры, которые не содержатся в нулевой и проверяют гипотезу о равенстве коэффициентов при дополнительных регрессорах нулю с помощью соответствующей F-статистики. В этом тесте обе сравниваемые модели содержатся в расширенной модели.

Один из тестов для сравнения моделей с разными зависимыми переменными — **Р_F-тест**. Пусть две сравниваемые модели заданы следующими уравнениями:

$$Y_1 = f_1(Y) = X_1 \beta_1 + \varepsilon_1,$$

$$Y_2 = f_2(Y) = X_2 \beta_2 + \varepsilon_2.$$

Например, $f_1(Y) = Y$, $f_2(Y) = \ln Y$ или $f_2(Y) = Y/W$ (“взвешенная” зависимая переменная). В “нулевую” модель в Р_F-тесте добавляется регрессор, равный разности расчетных значений из альтернативной модели и приведенных к тому же виду расчетных значений из “нулевой” модели. Так, в первую модель нужно добавить

$$X_2 \hat{\beta}_2 - f_2(f_1^{-1}(X_1 \hat{\beta}_1)) = \hat{Y}_2 - f_2(f_1^{-1}(\hat{Y}_1)).$$

Пусть, к примеру, $f_1(Y) = Y$, а $f_2(Y) = \ln(Y)$. Тогда в первую модель добавляют $\hat{Y}_2 - \ln(\hat{Y}_1)$, а во вторую — $\hat{Y}_1 - \exp(\hat{Y}_2)$.

Если отвергаются обе модели, то это должно означать, что каждая из них содержит информацию, не содержащуюся в другой, и следует попытаться как-то соединить две модели в одну. Если обе модели не отвергаются, то это означает, что с точки зрения данного теста они эквивалентны.

Фиктивные переменные как регрессоры

Общие соображения

Термин “фиктивные переменные” используется как противоположность “значащим” переменным, показывающим уровень количественного показателя, принимающего значения из непрерывного интервала. Как правило, фиктивная переменная — это индикаторная переменная, отражающая качественную характеристику. Чаще всего применяются бинарные фиктивные переменные, принимающие два значения, 0 и 1, в зависимости от определенного условия. Например, в результате опроса группы людей 0 может означать, что опрашиваемый — мужчина, а 1 — женщина. К фиктивным переменным иногда относят регрессор, состоящий из одних единиц (т.е. константу, свободный член), а также временной тренд.

Фиктивные переменные, будучи экзогенными, не создают каких-либо трудностей при применении ОМНК. Фиктивные переменные являются эффективным инструментом построения регрессионных моделей и проверки гипотез.

Пример. (Проверка гипотезы о равенстве средних в двух выборках в предположении равенства дисперсий)

Нулевая гипотеза состоит в том, что случайные величины в двух выборках имеют одинаковое математическое ожидание. Альтернативная гипотеза состоит в том, что математические ожидания равны только в пределах выборок, но не между выборками. Предполагается, что величины нормально распределены и дисперсии одинаковы для всех наблюдений. Объединим две выборки в одну. Пусть Y_i — вектор наблюдений для данной величины, D_i — фиктивная переменная принимающая значение 0 для первой выборки и 1 для второй выборки. Тогда для проверки гипотезы оценим регрессионную модель:

$$Y_i = \alpha + \beta D_i + \varepsilon_i.$$

Нулевая гипотеза: $\beta = 0$. Альтернативная гипотеза: $\beta \neq 0$. Такую гипотезу можно проверить с помощью t-статистики для коэффициента β . $\hat{\alpha}$ будет оценкой мат. ожидания для первой выборки, $\hat{\alpha} + \hat{\beta}$ для второй.

— — —

Предположим, что математическое ожидание зависимой переменной в регрессии увеличивается на некоторую фиксированную величину, если вы-

полняется определенное условие. Пусть, например для выборки предприятий одной отрасли оценивается производственная функция Кобба-Дугласа. Есть гипотеза, что для частных предприятий в этой отрасли производство при тех же труде и капитале выше, чем для государственных. Введем переменную D_i , которая принимает значение 0 для государственных предприятий и 1 для частных. Регрессионное уравнение будет иметь вид:

$$\ln Y_i - \ln L_i = \alpha_0 + \alpha_1 D_i + \beta (\ln K_i - \ln L_i).$$

Если коэффициент α_1 значимо положителен, то гипотезу нельзя отвергнуть.

Еще одна область применения фиктивных переменных — когда предполагается, что коэффициенты при “значащих” переменных меняются в зависимости от некоторого условия.

Пусть в приведенной модели $\beta = \beta_0$ для гос. предприятий и $\beta = \beta_1$ для частных. Тогда модель запишется в виде:

$$\begin{aligned} \ln Y_i - \ln L_i = & \alpha_0 + \alpha_1 D_i + \beta_0 (\ln K_i - \ln L_i) + \\ & + (\beta_1 - \beta_0) D_i (\ln K_i - \ln L_i). \end{aligned}$$

Заменив параметры, получаем линейную относительно параметров модель.

В регрессионное уравнение может войти несколько фиктивных переменных. Рассмотрим два условия: А и В. Пусть D_i^A — индикатор условия А ($D_i^A = 1$, если выполнено условие А, и 0 — если нет), D_i^B — индикатор условия В. Тогда $D_i^{AB} = D_i^A D_i^B$ — индикатор одновременного выполнения условий А и В. Эти три переменные будут точно описывать, в каком состоянии находится “мир” для данного наблюдения. Следует отметить, что четвертая фиктивная переменная (индикатор того, что одновременно не выполнены условия А и В) будет излишней в регрессии, уже включающей константу. Если ее добавить в регрессию, то матрица регрессоров будет вырожденной.

Дисперсионный анализ с фиксированными эффектами может быть сведен к регрессионному анализу с фиктивными регрессорами. Проверке гипотез с помощью ковариационного анализа будет соответствовать проверка гипотезы о равенстве нулю коэффициентов при соответствующей группе фиктивных переменных.

Использование фиктивных переменных для проверки однородности наблюдений и прогнозирования

Приведенную выше модель для двух типов предприятий

$$\ln Y_i - \ln L_i = \alpha_0 + \alpha_1 D_i + \beta_0 (\ln K_i - \ln L_i) + (\beta_1 - \beta_0) D_i (\ln K_i - \ln L_i).$$

можно использовать для проверки гипотезы о том, что коэффициенты регрессии разные для гос. предприятий и для частных. Гипотеза проверяется с помощью F-теста на добавление переменных D_i и $D_i (\ln K_i - \ln L_i)$.

В общем случае пусть наблюдения разбиты на две группы — I^1 и I^2 . Матрица регрессоров X распадается на две матрицы регрессоров X^1 и X^2 соответственно, а зависимая переменная Y — на Y^1 и Y^2 соответственно. Нулевая гипотеза состоит в том, что наблюдения порождены моделью $Y = X\beta + \varepsilon$. Альтернативная гипотеза состоит в том, что первая группа наблюдений порождена моделью $Y^1 = X^1\beta^1 + \varepsilon^1$, а вторая группа наблюдений — моделью $Y^2 = X^2\beta^2 + \varepsilon^2$, причем $\beta^1 \neq \beta^2$.

Введем фиктивную переменную D , такую что $D_i=0$ при $i \in I^1$ и $D_i=1$ при $i \in I^2$. Если все ошибки имеют одинаковую дисперсию, то гипотезу можно проверить с помощью регрессии Y по $Z = [X \ X*D]$. Здесь $X*D$ обозначает прямое произведение матрицы X на D , так что i -я строка матрицы Z равна $Z_i = [X_i, D_i X_i]$.

Тест на равенство коэффициентов регрессии в двух выборках, называют **тестом Чоу**. Нулевая гипотеза проверяется с помощью F-статистики для гипотезы о том, что коэффициенты при всех добавленных переменных равны нулю.

Еще одно использование фиктивных переменных — проверка гипотезы о том, что некоторое наблюдение принадлежит к той же выборке, что и все остальные наблюдения. Пусть i^* — номер этого наблюдения. Введем фиктивную переменную D , такую что $D_i=0$ при $i \neq i^*$ и $D_{i^*}=1$. Добавим эту переменную в исходную регрессию. Нужной нам статистикой будет F- или t-статистика для гипотезы о том, что коэффициент при добавленной переменной равен нулю. Если нулевая гипотеза отвергается, то соответствующее наблюдение следует считать **выбросом**. Назовем этот тест тестом для выбросов.

Тот же тест можно провести для небольшой группы наблюдений I^* . Требуется добавить регрессию по одной фиктивной переменной описанного вида для каждого из наблюдений $i \in I^*$. Нужной нам статистикой будет F-

статистика для гипотезы о том, что коэффициенты при всех добавленных переменных одновременно равны нулю.

Фиктивные переменные, которые равны нулю для всех наблюдений кроме одного, обладают тем свойством, что при добавлении их в регрессию соответствующий остаток зануляется.

Если в тесте Чоу одна из двух выборок содержит мало наблюдений (не больше количества регрессоров), то остатки в этой выборке должны зануляться при применении ОМНК. В этом случае тест Чоу совпадает с описанным только что тестом для выбросов.

Рассмотрим теперь использование фиктивных переменных для прогнозирования. Пусть мы оценили некоторую регрессию ($Y = X\beta + \varepsilon$) и у нас имеются дополнительные наблюдения, для которых известна матрица регрессоров (X^*), но неизвестны значения зависимой переменной (Y^*). Предсказания находятся по формуле $X^*\hat{\beta}$, где $\hat{\beta}$ — оценки ОМНК из регрессии Y по X . Эти предсказания можно найти с помощью следующей регрессионной модели:

$$\begin{bmatrix} Y \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} X & \mathbf{O} \\ X^* & \mathbf{I} \end{bmatrix} \begin{bmatrix} \beta \\ \beta^* \end{bmatrix} + \begin{bmatrix} \varepsilon \\ \varepsilon^* \end{bmatrix}.$$

Вместо неизвестной зависимой переменной здесь стоят нули, и добавлены фиктивные переменные, каждая из которых равна нулю для соответственного добавочного наблюдения. Оценки β будут совпадать с $\hat{\beta}$, а оценки β^* будут равны $\hat{\beta}^* = -X^*\hat{\beta}$, то есть будут равны предсказаниям со знаком минус. Стандартные ошибки предсказаний будут равны стандартным ошибкам оценок $\hat{\beta}^*$, полученным из той же регрессии.

Пусть теперь Y^* становятся известными. Интересно было бы проверить, насколько фактические значения отличаются от предсказанных. Оказывается, можно воспользоваться аналогичной регрессией, в которой слева вместо нулей стоят Y^* :

$$\begin{bmatrix} Y \\ Y^* \end{bmatrix} = \begin{bmatrix} X & \mathbf{O} \\ X^* & \mathbf{I} \end{bmatrix} \begin{bmatrix} \beta \\ \beta^* \end{bmatrix} + \begin{bmatrix} \varepsilon \\ \varepsilon^* \end{bmatrix}.$$

Оценки коэффициентов при фиктивных переменных в этом случае будут равны ошибкам предсказаний $\hat{\beta}^* = Y^* - X^*\hat{\beta}$. **Тест на адекватность предсказаний** проводится как тест на одновременное равенство коэффициентов

при фиктивных переменных нулю: $\beta^* = 0$. Очевидно, что этот тест совпадает с тестом для выбросов.

Использование фиктивных переменных в моделях с временными рядами

В регрессионных моделях с временными рядами используется три основных вида фиктивных переменных:

1) Переменные-индикаторы принадлежности наблюдения к определенному периоду — для моделирования скачкообразных структурных сдвигов. Границы периода (моменты “скачков”) должны быть установлены из априорных соображений. Например, 1, если наблюдение принадлежит периоду 1941-45 гг. и 0 в противном случае. Это пример использования для моделирования временного структурного сдвига. Постоянный структурный сдвиг моделируется переменной равной 0 до определенного момента времени и 1 для всех наблюдений после этого момента времени.

2) Сезонные переменные — для моделирования сезонности. Сезонные переменные принимают разные значения в зависимости от того, какому месяцу или кварталу года или какому дню недели соответствует наблюдение.

3) Линейный временной тренд — для моделирования постепенных плавных структурных сдвигов. Эта фиктивная переменная показывает, какой промежуток времени прошел от некоторого “нулевого” момента времени до того момента, к которому относится данное наблюдение (координаты данного наблюдения на временной шкале). Если промежутки времени между последовательными наблюдениями одинаковы, то временной тренд можно составить из номеров наблюдений.

Фиктивные переменные помогают отразить тот факт, что коэффициенты линейной регрессии могут меняться во времени. В простейшем случае изменится константа, а тем самым и мат. ожидание зависимой переменной.

Пусть исходная модель имеет вид $Y_t = \alpha + \beta X_t + \varepsilon_t$ и предполагается, что α линейно зависит от фиктивной переменной F_t : $\alpha_t = \alpha_0 + \alpha_1 F_t$. Тогда уравнение изменится следующим образом: $Y_t = \alpha_0 + \alpha_1 F_t + \beta X_t + \varepsilon_t$, оставаясь линейным по параметрам.

Коэффициенты при значащих переменных тоже могут быть подвержены изменениям. Проинтерпретировать это можно так, что сила их влияния на независимую переменную меняется со временем.

Например, в рассмотренном уравнении может быть $\beta_t = \beta_0 + \beta_1 F_t$. Тогда $Y_t = \alpha + \beta_0 X_t + \beta_1 F_t X_t + \varepsilon$. Эта модель также остается линейной по параметрам. Коэффициент β_1 показывает, как исходный коэффициент β зависит от времени. С помощью соответствующей t-статистики можно проверить гипотезу, что $\beta_1 = 0$ (β не меняется со временем).

Можно предложить следующий тест на стабильность коэффициентов модели во времени. Для его проведения нужно добавить в уравнения произведения всех исходных регрессоров и фиктивной переменной. Например, в модель $Y_t = \alpha + \beta_1 X_{t1} + \beta_2 X_{t2} + \varepsilon$ следует добавить регрессоры F_t , $X_{t1}F_t$ и $X_{t2}F_t$. Если коэффициенты при добавочных переменных значимы в совокупности (применяем F-статистику), то нельзя отвергнуть гипотезу о том, что коэффициенты изменяются со временем.

Тест Чоу представляет собой частный случай описанного теста. Для временных рядов тест Чоу — это тест на то, что в определенный момент времени произошло скачкообразное изменение коэффициентов регрессии.

Временной тренд отличается от бинарных фиктивных переменных тем, что имеет смысл использовать его степени: t^2 , t^3 и т. д. Они помогают моделировать гладкий, но нелинейный тренд. (Бинарную переменную нет смысла возводить в степень, потому что в результате получится та же самая переменная.)

Можно также комбинировать три указанных вида фиктивных переменных, создавая переменные “взаимодействия” соответствующих эффектов. Пусть Y — квартальные данные по некоторому показателю. Его поведение можно смоделировать, представляя мат. ожидание как комбинацию линейного тренда и сезонности.

$$Y_t = \alpha_0 + \alpha_1 t + \beta_1 Q_{t1} + \beta_2 Q_{t2} + \beta_3 Q_{t3} + \gamma_1 Q_{t1}t + \gamma_2 Q_{t2}t + \gamma_3 Q_{t3}t + \varepsilon_t,$$

где t — тренд, Q_i — квартальные сезонные переменные

$$Q_{ij} = \begin{cases} 1, & \text{если } j\text{-й квартал} \\ 0, & \text{иначе} \end{cases}$$

Q_{t4} не нужно вводить в эту регрессию, так как есть константа, а $Q_{t4}t$ не нужно вводить в регрессию, так как есть временной тренд t .

Если все $\gamma_j \neq 0$, то это означает, что структура сезонности линейно изменяется со временем.

Комбинация рассмотренных фиктивных переменных позволяет моделировать еще один эффект — изменение наклона тренда с определенного мо-

мента. Помимо тренда в регрессию следует тогда ввести следующую переменную: в начале выборки до некоторого момента времени она равна 0, а вторая ее часть представляет собой временной тренд (1, 2, 3 и т. д. в случае одинаковых интервалов между наблюдениями).

Регрессионные модели с фиктивными переменными являются альтернативой ARIMA-моделям и регрессионным моделям с AR- или MA-процессом в ошибке. В первом случае изменение мат. ожидания во времени можно назвать детерминированным трендом, во втором — стохастическим (строго говоря термин “стохастический тренд” употребляют только по отношению к нестационарным процессам). Решить, какой вид модели применять, сложно. Дело в том, что трудно отличить (в случае малых выборок), когда случайная величина имеет линейный детерминированный тренд со стационарными отклонениями от него, а когда она формируется нестационарным авторегрессионным процессом. То же самое верно для выбора способа моделирования сезонности.

Использование фиктивных переменных имеет следующие преимущества:

1) Интервалы между наблюдениями не обязательно должны быть одинаковыми. В выборке могут быть пропущенные наблюдения.

2) Коэффициенты при фиктивных переменных легко интерпретировать, они наглядно представляют структуру динамического процесса.

3) Для оценивания модели не приходится выходить за рамки классического метода наименьших квадратов.

Спектральный анализ и регрессия

Спектральный анализ можно осуществить с помощью гармонических фиктивных переменных (гармонического тренда). Пусть рассматриваются помесечные данные. Тогда сезонные колебания можно моделировать, используя следующий набор фиктивных переменных:

$$S_{ik} = \sin(2\pi t/k), \quad C_{ik} = \cos(2\pi t/k), \quad k=1, \dots, 6.$$

При $k=1$ период колебаний равен 12 месяцам, при $k=2$ — 6 месяцам, при $k=3$ — 4 месяцам, при $k=5$ — 2,4 месяцам, при $k=6$ — 2 месяцам.

Включение в регрессию полного набора ($k=1, \dots, 6$) рассматриваемых переменных эквивалентно включению набора месячных бинарных фиктивных переменных ($M_{ij} = 1$, если j -й месяц и 0 в противном случае). Гармонические переменные следует применять в том случае, если предполагается, что се-

зонность может быть гладкой. В этом случае высокочастотные гармоники (с коротким периодом) не включают в регрессию, например, берут только $k=1, 2$.

Одна из возможных содержательных интерпретаций такого подхода состоит в том, что гармоники с более длинным периодом моделируют долгосрочные (перманентные) эффекты, а с коротким — краткосрочные.

Модели с качественной зависимой переменной

Модели с качественной зависимой переменной как правило возникают, когда экономика рассматривается на очень дезагрегированном уровне. Обычно это ситуация, когда некоторая экономическая единица (субъект), делает выбор между двумя и более возможными альтернативами. В качестве примера можно привести выбор предприятия: внедрять какую-то новую технологию или нет.

Модели с бинарной зависимой переменной

В бинарную модель входит зависимая переменная Y , принимающая два значения (обычно 0 и 1), а также регрессоры X , которые содержат факторы, определяющие выбор. Обычная линейная регрессионная модель не подходит для описания этой ситуации. Она предполагает, что зависимая переменная имеет непрерывное распределение, а здесь необходимо, чтобы она имела дискретное распределение.

Пусть, например, рассматривается выбор группы людей: быть безработным или работать. Построенная линейная регрессия будет предсказывать абсурдные значения Y – дробные, отрицательные и большие единицы. Математическое ожидание остатков при этом будет даже асимптотически зависеть от X .

Вообще говоря, предсказывать результат выбора можно было бы и по результатам линейной регрессии: если \hat{Y} (расчетное значение Y) больше $1/2$, то берем 1, если \hat{Y} меньше $1/2$, то берем 0. Хотя это соображение и не улучшает модель, оно подсказывает, какой может быть более адекватная модель.

С формальной точки зрения требуется найти модель, которая порождала бы дискретное распределение, зависящее от X , которое бы хорошо описывало данные. Поскольку для бинарной зависимости переменной распределение будет бинарным, то оно полностью определяется вероятностью получения единицы (как функцией X), которая совпадает с математическим ожиданием, если переменная принимает значения 0 и 1:

$$\begin{aligned} E(Y | X) &= \text{Prob}(Y = 1 | X) \cdot 1 + \text{Prob}(Y = 0 | X) \cdot 0 = \\ &= \text{Prob}(Y = 1 | X). \end{aligned}$$

Геометрически задача состоит в том, чтобы найти гиперплоскость, которая бы в определенном смысле наилучшим образом разделяла две группы наблюдений (соответствующие 0 и 1) в пространстве регрессоров. Может

случиться, что такая плоскость будет не единственной; это происходит при идеальном разделении.

Модель выбора. Пробит и логит

Предлагается два вида моделей выбора, которые могли бы порождать интересующее нас распределение зависимой переменной: *пороговая модель* и *модель, основанная на полезности альтернатив*. Идея пороговой модели уже обрисована выше. Предполагается, что в основе выбора лежит ненаблюдаемая переменная \tilde{Y} , математическое ожидание которой является линейной комбинацией набора регрессоров X : $\tilde{Y} = X\beta + \varepsilon$. Наблюдается только дискретная величина Y , которая связана с \tilde{Y} следующим образом: если \tilde{Y} больше некоторой пороговой величины C , то $Y = 1$, если меньше, то $Y = 0$. Как обычно предполагается, что ошибки ε_i имеют нулевое математическое ожидание, одинаково распределены и независимы. Величину C можно принять равной нулю.

Другая модель предполагает, что выбор осуществляется на основе ненаблюдаемой полезности альтернатив $u(Y, X)$. Если $u(1, X) > u(0, X)$, то выбираем 1, если $u(0, X) < u(1, X)$, то выбираем 0. В простейшем случае полезность является линейной функцией регрессоров:

$$u(1, X) = u_1 = X\beta_1, \quad u(0, X) = u_0 = X\beta_0.$$

Чтобы модель была вероятностной, предполагается, что есть отклоняющие факторы, так что $u_1 = X\beta_1 + \varepsilon_1$, $u_0 = X\beta_0 + \varepsilon_0$. Эта модель сводится к пороговой, если взять $\tilde{Y} = u_1 - u_0 = X(\beta_1 - \beta_0) + \varepsilon_1 - \varepsilon_0 = X\beta + \varepsilon$, а в качестве порога — ноль. Выведем теперь из распределения ε распределение \tilde{Y} , а из распределения \tilde{Y} — распределение Y .²

Есть два удобных вида распределения, которые обычно используют для описания отклонения ε .

1. Логистическое распределение.³

² Для симметричного распределения $1 - F(-X) = F(X)$.

³ Если ε_1 и ε_0 имеют распределение Вейбулла и независимы, то ε имеет логистическое распределение. Если ε_1 и ε_0 имеют нормальное распределение и независимы, то ε тоже имеет нормальное распределение.

Плотность логистического распределения равна $\frac{e^z}{(1 + e^z)^2}$ (см. Рис. 3), а функция распределения равна $\frac{1}{1 + e^{-z}}$ (ее называют логистой). Модель с бинарной зависимой переменной с логистически распределенным отклонением называют **ЛОГИТ**. Для логита $E(Y | X) = 1 - \frac{1}{1 + e^{X\beta}} = \frac{1}{1 + e^{-X\beta}}$.

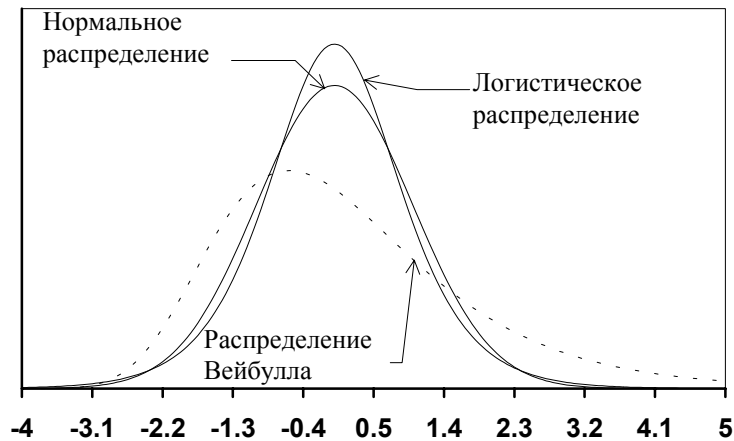


Рис. 3

2. Нормальное распределение.

Модель с нормально распределенным отклонением называют **ПРОБИТ**.

Для пробита

$$E(Y | X) = \int_{-\infty}^{X\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

Логистическое распределение очень похоже на нормальное. Различить, когда следует применять логит, а когда — пробит, в малых выборках невозможно. Оценки коэффициентов β отличаются множителем, который практически постоянен.

Оценка качества модели и проверка гипотез

Пробит и логит обычно оценивают методом максимального правдоподобия. Существуют также упрощенные методы, использующие сгруппированные наблюдения. Предположим, что методом максимального правдоподобия получен вектор оценок $\hat{\beta}$. Как в этом случае можно судить о качестве модели и проверять гипотезы?

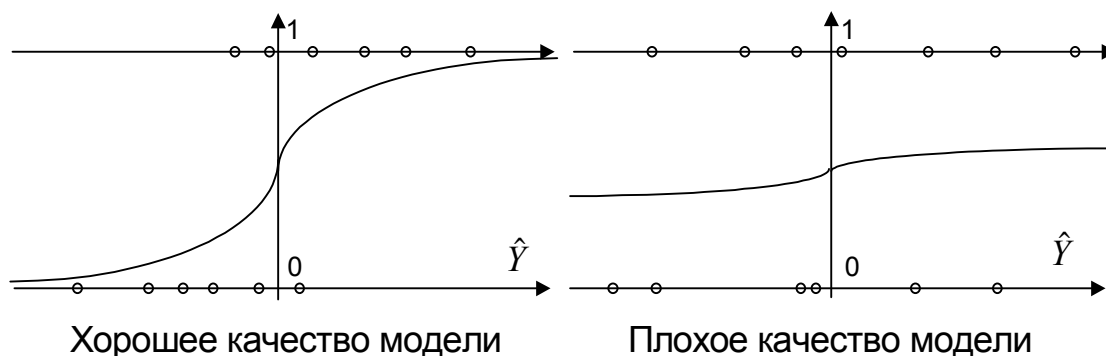


Рис. 4

Величину $\hat{Y} = X\hat{\beta}$ можно назвать по аналогии с линейной регрессией расчетными значениями. Она является оценкой математического ожидания ненаблюдаемой величины \tilde{Y} , сравнивая которую с нулем делают выбор между 0 и 1. Уравнение $\hat{Y}(X) = 0$ задает ту гиперплоскость, которой разделяются две группы точек — те точки, для которых предсказано $Y_i = 0$, и те точки, для которых предсказано $Y_i = 1$ (с помощью критерия $\hat{Y}_i < 0 \Rightarrow 0$, а $\hat{Y}_i > 0 \Rightarrow 1$). Поэтому наглядно о качестве модели можно судить по диаграмме соответствующих точек по Y : чем лучше разделены две группы точек, тем более качественна модель. О качестве модели можно судить также по графику оценки $E(Y)$ по \hat{Y} ($\frac{1}{1+e^{-\hat{Y}}}$ по \hat{Y}). Этот график в случае “хорошей” модели должен

быть "крутой" в нуле. (См. Рис. 4)

На этих двух графиках слева внизу и справа вверху расположены правильно предсказанные точки, а слева вверху и справа внизу — неправильно. То же самое можно представить таблицей:

		Предсказано		
		0	1	Сумма
На самом деле	0	×	×	×
	1	×	×	×
	Сумма	×	×	×

Понятно, что "хорошая" модель должна давать высокий процент правильных предсказаний.

Для проверки набора ограничений на параметры удобно использовать статистику отношения правдоподобия $LR = 2 (\ell(\hat{\beta}) - \ell(\hat{\beta}_R))$, где

$$\ell = \sum_{i=1}^N [Y_i \ln P_i + (1 - Y_i) \ln (1 - P_i)]$$
 — логарифмическая функция правдоподобия,

подобия,

$\hat{\beta}$ — оценка методом максимума правдоподобия без ограничений,

$\hat{\beta}_R$ — оценка при ограничениях.

Эту же статистику можно использовать для построения показателя качества модели, аналогичного F-статистике для линейной регрессии. Это статистика для проверки гипотезы о том, что коэффициенты при всех регрессорах, кроме константы, равны одновременно нулю. Соответствующая статистика отношения правдоподобия равна $LR_0 = 2 (\ell(\hat{\beta}) - \ell_0)$, где ℓ_0 — максимум логарифмической функции правдоподобия для константы. Она распределена асимптотически как χ^2 с $k-1$ степенями свободы, где k — количество параметров в исходной модели, включая константу. Величина 1 получается следующим образом. Пусть N — общее количество наблюдений, n_0 — количество наблюдений, для которых $Y_i = 0$, n_1 — количество наблюдений, для которых $Y_i = 1$. Тогда предсказанная вероятность появления $Y_i = 1$ в модели с одной константой будет равна для всех наблюдений n_1/N . Отсюда $\ell_0 = n_0 \ln n_0 + n_1 \ln n_1 - N \ln N$. Еще один показатель качества модели, основанный на максимуме функции правдоподобия — информационный критерий Акаике:

$$AIC = -\frac{2}{N} (\ell(\hat{\beta}) - k).$$

Для моделей с бинарной зависимой переменной можно сконструировать и некий аналог коэффициента детерминации — псевдо- R^2 :

$$R^2 = \frac{\sum_{i=1}^N (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^N (\hat{Y}_i - \bar{Y})^2 + N\sigma^2},$$

где \bar{Y} — среднее \hat{Y}_i , σ^2 — дисперсия ошибки ε , которая равна 1 для пробита и $\frac{\pi^2}{3}$ для логита.

Множественные модели с качественными зависимыми переменными

В этом подразделе будет говориться о логите, хотя это верно и для пробита. Множественный логит является логическим продолжением бинарного. Он возникает, когда рассматривается выбор между более чем двумя альтернативами. Существует два основных типа множественных моделей: упорядоченный логит и собственно множественный логит. Упорядоченный логит развивает пороговую модель, а собственно множественный логит — модель выбора по полезности.

Упорядоченный логит имеет дело с альтернативами, которые можно расположить в определенном порядке. Например, это могут быть оценки, полученные на экзамене, или качество товара, которое может характеризоваться сортом от "высшего" до "третьего". Будем предполагать, что альтернативы пронумерованы от 0 до S . Переменная Y принимает значение s , если выбрана альтернатива s . Предполагается, что в основе выбора лежит ненаблюдаемая величина $\tilde{Y} = X\beta + \varepsilon$. $Y = 0$ выбирается, если \tilde{Y} меньше нижнего (первого) порогового значения, $Y = 1$, если \tilde{Y} попадает в промежуток от первого до второго порогового значения и т. д.; $Y = S$ выбирается, если \tilde{Y} превышает верхнее пороговое значение:

$$Y_i = \begin{cases} 0, & \tilde{Y} \leq \gamma_1 \\ 1, & \gamma_1 < \tilde{Y} \leq \gamma_2 \\ \vdots & \\ S, & \tilde{Y} > \gamma_S \end{cases}$$

Если ε имеет логистическое распределение, то логарифмическая функция распределения равна

$$\begin{aligned} \ell &= \sum_{i \in I_0} \ln(\text{Prob}(Y_i = 0)) + \sum_{i \in I_1} \ln(\text{Prob}(Y_i = 1)) + \dots + \\ &+ \sum_{i \in I_S} \ln(\text{Prob}(Y_i = S)) = \\ &= \sum_{i \in I_0} \ln\left(\frac{1}{1 + e^{X_i\beta}}\right) + \sum_{i \in I_1} \ln\left(\frac{1}{1 + e^{X_i\beta - \gamma_2}} - \frac{1}{1 + e^{X_i\beta}}\right) + \dots + \end{aligned}$$

$$+ \sum_{i \in I_s} \ln \left(\frac{1}{1 + e^{\gamma_s - X_i \beta}} \right).$$

Эту величину следует максимизировать по β и γ . В результате получается оценка максимума правдоподобия.

Если альтернативы не упорядочены, то предполагается, что выбор делается на основе функции полезности $u(Y, Z)$. Обозначим $u_s(Z) = u(s, Z)$. В линейной модели $u_s = Z_s \beta_s + \varepsilon_s$, где Z_s – матрица регрессоров, β_s – неизвестные параметры. Обычно делают одно из двух упрощающих допущений: либо что регрессоры для всех альтернатив одни и те же: $u_s = Z \beta_s + \varepsilon_s$, либо что функция имеет один и тот же вид, а меняются только факторы, определяющие выбор, т.е. $u_s = Z_s \beta + \varepsilon_s$. Y_i выбирается равным s , если $u_s(Z_i) > u_t(Z_i) \forall s \neq t$. В множественном логите принимается, что ошибки ε_s имеют распределение Вейбулла. Распределение Вейбулла⁴ в стандартной форме имеет функцию распределения $F(X) = e^{-e^{-X}}$ (см. Рис. 3). Распределение Вейбулла обладает следующими важными для рассматриваемой модели свойствами: максимум нескольких величин, распределенных по Вейбуллу, также распределен по Вейбуллу, а разность двух величин, распределенных по Вейбуллу, имеет логистическое распределение. Используя эти свойства, можно вывести, что в многомерном логите

$$\text{Prob}(Y_i = s) = P_i^s = \frac{e^{Z_s \beta_s}}{\sum_{t=0}^S e^{Z_t \beta_t}}.$$

Вероятности не изменятся, если числитель и знаменатель нормировать, разделив на $e^{Z_0 \beta_0}$:

$$P_i^s = \frac{e^{Z_s \beta_s - Z_0 \beta_0}}{1 + \sum_{t=1}^S e^{Z_t \beta_t - Z_0 \beta_0}}$$

Если принимается, что $\beta_s = \beta \forall s$, то удобно обозначить $Z_s - Z_0 = X_s$ ($s = 1, \dots, S$), а если $Z_s = X \forall s$, то $\beta_s - \beta_0$ можно заменить на β_s .

⁴ Распределение Вейбулла также иногда называют распределением экстремального значения первого рода. Кроме того, именем Вейбулла называют и другие распределения, поэтому может возникнуть путаница.

Динамическая спецификация регрессионной модели

В этом разделе рассматривается, как можно построить модель, в которой переменными являются временные ряды. Основным понятием, употребляемым при разговоре о регрессионной модели для временных рядов, является понятие **лага**. В буквальном смысле по-английски lag — запаздывание. Под лагом некоторой переменной понимают ее значение в предыдущие периоды времени. Например, для переменной Y_t лагом в τ периодов будет $Y_{t-\tau}$. В векторном виде лаг переменной Y принято записывать как $Y_{-\tau}$. В терминологии имеется некоторая неоднозначность. Часто лагом называют величину $t - \tau$. Кроме того, лагом называют структуру, т.е. форму, в которой входят в модель лаги некоторой переменной.

Другой способ обозначения лага — с помощью **лагового оператора**. Его обозначают буквой L (иногда B). Лаговый оператор — это линейный оператор. С ним можно обращаться как с переменной, но он должен стоять *перед* той переменной, к которой применяется. LX обозначает X_{-1} , $L^\tau X = X_{-\tau}$. Если применить многочлен от лага $f(L) = a_n L^n + \dots + a_1 L + a_0$ к переменной X , то получится

$$f(L)X = \left(\sum_{i=0}^n a_i L^i \right) X = \sum_{i=0}^n a_i (L^i X) = \sum_{i=0}^n a_i X_{-i}.$$

Другой постоянно используемый оператор — оператор разности или абсолютного прироста Δ , который определяется как $1 - L$, так что $\Delta X = X - X_{-1}$. Вторая разность — дважды взятый оператор Δ : $\Delta^2 = (1 - L)^2 = 1 - 2L + L^2$ и т. д.

Модель распределенного лага

Часто при моделировании экономических процессов на зависимую переменную влияют не только текущие значения объясняющего фактора, но и его лаги. Типичным примером являются капиталовложения: они всегда дают результат с некоторым лагом.

Модель **распределенного лага** можно записать следующим образом:

$$Y = \alpha + \sum_{\tau=0}^q \beta_\tau X_{-\tau} + \varepsilon = \alpha f(L) X + \varepsilon,$$

где q — величина наибольшего лага, $f(z) = \sum \beta_\tau z^\tau$ — многочлен. Коэффициенты β_τ показывают структуру лага и называются весами. Оценивание этой модели может быть затруднено проблемой мультиколлинеарности. Такое случается, если величина X_t мало меняется со временем (если X_t — случайный процесс, то это означает автокорреляцию). При этом невозможно точно оценить структуру лага; хотя возможно точно оценить сумму весов β_τ . Последнюю можно вычленить из модели следующим образом:

$$Y = \alpha + \beta_\Sigma X + \sum_{\tau=1}^q \beta_\tau (X_{-\tau} - X) + \varepsilon, \quad \text{где } \beta_\Sigma = \sum_{\tau=0}^q \beta_\tau.$$

Это пример преобразования формы регрессионной модели с временными рядами.

В случае мультиколлинеарности лаговых переменных обычно на лаговую структуру накладывают какое-нибудь ограничение, чтобы уменьшить количество оцениваемых коэффициентов. Одна из возможных структур лага — это **полиномиальный лаг**, веса которого задаются полиномом от величины лага τ :

$$\beta_\tau = \gamma_0 + \gamma_1 \tau + \gamma_2 \tau^2 + \dots + \gamma_p \tau^p = \sum_{s=0}^p \gamma_s \tau^s, \quad \tau = 0, \dots, q.$$

где p — степень многочлена. Простейший полиномиальный лаг — линейный. Для него $\beta_\tau = \gamma_0 + \gamma_1 \tau$. Его структуру можно представить на следующей диаграмме (Рис. 5).

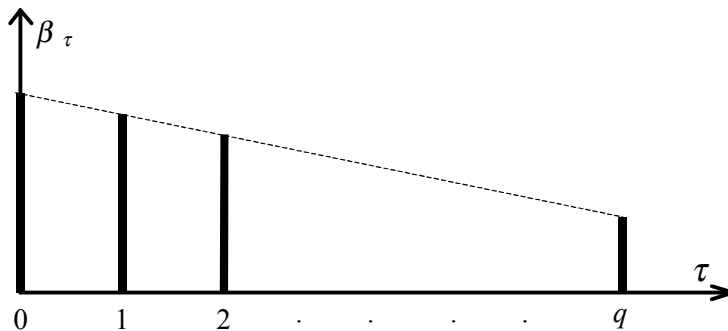


Рис. 5

Полиномиальный лаг накладывают на модель $q - p$ линейных ограничений. Понятно при этом, что если модель была линейной, то она и останется линейной. Рассмотрим, каким образом ее можно оценить.

Подставим выражения для β_τ в исходную модель.

$$\sum_{\tau=0}^q \beta_{\tau} X_{-\tau} = \sum_{\tau=0}^q \left(\sum_{s=0}^p \gamma_s \tau^s \right) X_{-\tau} = \sum_{s=0}^p \gamma_s \sum_{\tau=0}^q \tau^s X_{-\tau} = \sum_{s=0}^p \gamma_s Z_s.$$

Получим новую модель

$$Y = \alpha + \sum_{s=0}^p \gamma_s Z_s + \varepsilon$$

с преобразованными регрессорами $Z_s = \sum_{\tau=0}^q \tau^s X_{-\tau}$. Оценив γ_s надо подставить

их в формулу для весов β_{τ} .

При оценивании модели с ограничениями на структуру лага, нужно проверить, правильно ли наложены ограничения. С помощью соответствующей F-статистики можно сравнить ее с исходной, неограниченной, моделью, поскольку она является ее частным случаем. Модель

$$Y = \alpha + \sum_{s=0}^q \gamma_s Z_s + \varepsilon$$

эквивалентна исходной модели с точностью до линейных преобразований, поэтому достаточно проверить гипотезу о том, что последние $q - p$ коэффициентов в ней ($\gamma_{p+1}, \dots, \gamma_q$) равны нулю.

Часто принимают, что веса на концах полиномиальной лаговой структуры равны нулю. Это требование накладывает на модель дополнительные ограничения.

Еще один популярный вид структуры лага — **экспоненциальный (геометрический) лаг**. Его веса задаются следующими соотношениями:

$$\beta_{\tau} = \beta_0 \delta^{\tau}, \quad \tau = 0, \dots, \infty, \quad \text{где } 0 < \delta < 1.$$

Веса геометрического лага убывают экспоненциально с увеличением лага (Рис. 6).

Сумма весов в этой модели равна

$$\beta_{\Sigma} = \sum_{\tau=0}^{\infty} \beta_{\tau} = \sum_{\tau=0}^{\infty} \beta_0 \delta^{\tau} = \frac{\beta_0}{1-\delta}.$$

К модели с геометрическим лагом можно применить **преобразование Койка** (Koyck transformation). Проведем его с использованием лаговых операторов.

$$Y = \sum_{\tau=0}^{\infty} \beta_0 \delta^{\tau} L^{\tau} X + \varepsilon = \beta_0 \sum_{\tau=0}^{\infty} (\delta L)^{\tau} X + \varepsilon = \beta_0 \frac{1}{1-\delta L} X + \varepsilon.$$

Отсюда $(1-\delta L)Y = \beta X + (1-\delta L)\varepsilon$ или, по определению лагового оператора,

$$Y - \delta Y_{-1} = \beta X + \varepsilon - \delta \varepsilon_{-1}.$$

Еще одна проблема, возникающая при оценивании модели распределенного лага, — найти величину наибольшего лага. Самый простой способ — взять неограниченную модель с достаточно большим лагом и проверять гипотезы по “отсечению хвоста” с помощью t и F -статистик.

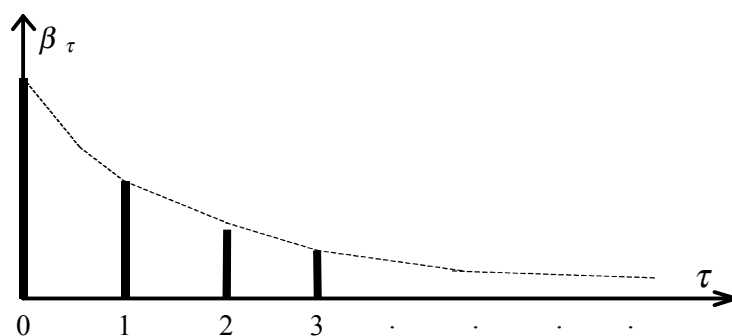


Рис. 6

Динамические регрессионные модели. Авторегрессионная модель с распределенным лагом

Динамическая регрессия — это такая регрессия, в которой в качестве регрессоров используются лаги зависимой переменной. Рассмотрим достаточно общую модель с одной независимой переменной — **авторегрессионную модель с распределенным лагом**. Ее можно записать в следующем виде:

$$Y = \alpha + \sum_{k=1}^p \beta_k Y_{-k} + \sum_{L=0}^q \gamma_L X_{-L} + \varepsilon,$$

где первая сумма представляет собой авторегрессионный член — распределенный лаг зависимой переменной, вторая сумма — распределенный лаг независимой переменной. Сокращенно эту модель обозначают ADL(p, q) (от английского autoregressive distributed lag).

В операторной форме:

$$Y = \alpha + Lf(L)Y + g(L)X + \varepsilon, \text{ где } f(.) \text{ и } g(.) \text{ — многочлены,}$$

или

$$h(L)Y = \alpha + g(L)X + \varepsilon, \text{ где } h(L) = 1 - Lf(L).$$

В частности, ADL(1,1) имеет вид

$$Y = \alpha + \beta_1 Y_{-1} + \gamma_0 X + \gamma_1 X_{-1} + \varepsilon.$$

Рассмотрим некоторые часто встречающиеся динамические модели, являющиеся частными случаями ADL-модели.

Модель ADL(0, q) — это модель распределенного лага, рассмотренная в предыдущем параграфе, так что в правой части нет лагов зависимой переменной.

Модель геометрического распределенного лага после преобразования Койка — это ADL(1, 0) с MA(1)-ошибкой и ограничением, что коэффициент при Y_{-1} равен параметру MA-процесса (δ) с обратным знаком:

$$Y = (1 - \delta)\alpha + \delta Y_{-1} + \beta_0 X + (\varepsilon - \delta \varepsilon_{-1}).$$

Авторегрессионную модель AR(p) можно считать ADL($p, 0$) с ограничением $\beta_0 = 0$. В этой модели переменная в левой части зависит только от своих собственных лагов:

$$Y = \alpha + \sum_{k=1}^p \beta_k Y_{-k} + \varepsilon.$$

В экономике субъекты не сразу могут приспособиться к меняющимся условиям — это происходит постепенно. Нужно время на обучение, переход на новые технологии, изменение условий долгосрочных контрактов и т.д. Эти процессы можно моделировать с помощью **модели частичного приспособления**

$$Y^D = b_0 + b_1 X_{-1} + \varepsilon,$$

$$\Delta Y = Y - Y_{-1} = \mu(Y^D - Y_{-1}),$$

где Y^D — желаемый уровень величины Y , μ — скорость приспособления ($0 < \mu \leq 1$). Если $\mu = 1$, то приспособление происходит мгновенно и всегда $Y^D = Y$.

Исключив ненаблюдаемую переменную Y^D , модель приводят к виду, удобному для оценивания:

$$Y = \mu b_0 + (1 - \mu)Y_{-1} + \mu b_1 X_{-1} + \mu \varepsilon.$$

Это ADL(1, 1) с коэффициентом при текущем значении X равным нулю.

Чтобы ввести в экономические модели ожидания экономических субъектов в простейшем случае используют модель **адаптивных ожиданий**. Адаптивные ожидания некоторой величины формируются только на основе прошлых значений этой величины. Например, пусть Y зависит от ожиданий величины X (X^E):

$$Y = \alpha_0 + \alpha_1 X^E + \varepsilon.$$

Ошибка в ожиданиях в предыдущий период приводит к корректировке ожиданий:

$$\Delta X^E = X^E - X^E_{-1} = \theta (X - X^E_{-1}).$$

Здесь θ — скорость приспособления ожиданий ($0 < \theta \leq 1$). Если $\theta = 1$, то ожидания всегда равны действительной величине X : $X^E = X$.

Решить разностное уравнение для ожиданий проще всего с использованием лагового оператора. Схему корректировки ожиданий можно записать как

$$(1 - (1 - \theta)L) X^E = \theta X, \text{ откуда}$$

$$X^E = \frac{\theta}{1 - (1 - \theta)L} X = \theta \sum_{\tau=0}^{\infty} (1 - \theta)^\tau X_{-\tau}.$$

Исключив ненаблюдаемые ожидания X^E , получим модель с геометрическим распределенным лагом.

Преобразование Койка дает другую форму модели адаптивных ожиданий — ADL(1, 0) с MA(1)-ошибкой и ограничением на коэффициенты:

$$(1 - (1 - \theta)L) Y = \theta \alpha_0 + \alpha_1 \theta X + (1 - (1 - \theta)L) \varepsilon.$$

В динамических регрессионных моделях важно различие между долгосрочной и краткосрочной динамикой (англ. long-run и short-run). Рассмотрим в долгосрочном аспекте модель ADL(1,1):

$$Y = \alpha + \beta_1 Y_{-1} + \gamma_0 X + \gamma_1 X_{-1} + \varepsilon.$$

Пусть установились стационарные уровни X и Y . Обозначим их X^* и Y^* .

Тогда

$$Y^* = \alpha + \beta_1 Y^* + \gamma_0 X^* + \gamma_1 X^*.$$

Уравнение

$$Y^* = \frac{\alpha}{1-\beta} + \frac{\gamma_0 + \gamma_1}{1-\beta} X^* = \alpha' + \lambda X^*$$

описывает **долгосрочное стационарное состояние** экономического процесса. Здесь $\lambda = \frac{\gamma_0 + \gamma_1}{1-\beta}$ — коэффициент долгосрочного влияния X на Y . Если Y и X — логарифмы исходных переменных, то λ — долгосрочная эластичность.

Модель ADL(1,1) можно привести к виду, который отражает краткосрочную динамику экономической системы. В этом виде модель называется **моделью исправления ошибок**, сокращенно ЕСМ (англ. error-correction model):

$$\begin{aligned} \Delta Y &= \alpha - (1 - \beta_1) Y_{-1} + \gamma_0 \Delta X + (\gamma_0 + \gamma_1) X_{-1} + \varepsilon \\ &= \alpha + \gamma_0 \Delta X - (1 - \beta_1) (Y_{-1} - \lambda X_{-1}) + \varepsilon. \end{aligned}$$

Предполагается, что если в предыдущий период переменная Y отклонилась от своего долгосрочного значения $\alpha' + \lambda X$, то член $Y_{-1} - \lambda X_{-1}$ корректирует динамику в нужном направлении. Для того, чтобы это происходило, необходимо выполнение условия $\beta_1 < 1$.

Бывает, что из теории явления известно, что $\lambda = 1$, тогда $\beta_1 + \gamma_0 + \gamma_1 = 1$. Часто именно такую модель называют ЕСМ.

Модели частичного приспособления и адаптивных ожиданий являются частными случаями модели исправления ошибок — не только формально математически, но и по экономическому содержанию. Например, модель частичного приспособления в форме ЕСМ выглядит как

$$\Delta Y = \mu b_0 - \mu (Y_{-1} - b_1 X_{-1}).$$

Интегрированные процессы, ложная регрессия и коинтеграция

Стационарные и нестационарные случайные процессы.

Чтобы проиллюстрировать различие между стационарными и нестационарными случайными процессами, рассмотрим авторегрессию первого порядка (AR(1)), т.е. авторегрессию, содержащую один лаг зависимой переменной:

$$Y_t = \mu + \rho Y_{t-1} + \varepsilon_t, \quad t = (-\infty, \dots, 0, 1, \dots, +\infty)$$

(предполагаем, что $\varepsilon_t \sim \text{IID}(0, \sigma_\varepsilon^2)$ — независимые одинаково распределенные случайные величины с нулевым мат. ожиданием и дисперсией σ_ε^2).

Слабое определение **стационарности** требует, чтобы математическое ожидание Y_t было постоянным (или нулевым), а ковариации не зависели от времени, только от лага:

$$Y_t = \text{const} (= 0), \quad \text{var}(Y_t) = \sigma_Y^2 = \text{const}, \quad \text{cov}(Y_t, Y_{t-\tau}) = c_\tau.$$

Покажем, что если $|\rho| < 1$, то процесс AR(1) будет стационарным. Решая уравнение авторегрессионной модели, получим

$$Y = \frac{\mu}{1-\rho} + \sum_{i=0}^{\infty} \rho^i \varepsilon_{-i}.$$

Мат. ожидание Y переменной постоянно: $E(Y) = \frac{\mu}{1-\rho}$. Второй член — это взвешенная сумма ошибок (геометрический распределенный лаг). Условие $|\rho| < 1$ гарантирует, что дисперсия этой суммы, а следовательно, и дисперсия Y конечна:

$$\sigma_Y^2 = \sum_{i=0}^{\infty} \rho^{2i} \text{var}(\varepsilon) = \sigma_\varepsilon^2 \frac{1}{1-\rho^2}.$$

Найдем также автоковариации процесса:

$$\text{cov}(Y, Y_{-\tau}) = \sum_{i=0}^{\infty} \rho^{-\tau} \rho^{2i} \sigma_\varepsilon^2 = \sigma_\varepsilon^2 \frac{\rho^{-\tau}}{1-\rho^2}.$$

Таким образом, рассматриваемый процесс слабо стационарен. На самом деле, поскольку ошибки ε_t одинаково распределены, то он стационарен и в сильном смысле.

Вывод изменится, если рассмотреть процесс с определенного момента времени, например, с $t = 1$. Предположим, что Y_0 — детерминированная ве-

личина. В этом случае процесс AR(1) не будет стационарным по данному выше определению. Дисперсия Y и автоковариации будут зависеть от t :

$$\text{var}(Y_t) = \sigma_{Y_t}^2 \quad \text{cov}(Y_t, Y_{t-\tau}) = c_{\tau t}.$$

Однако со временем такой процесс (если только $|\rho| < 1$) все больше приближается к стационарному. Его можно назвать асимптотически стационарным.

При $|\rho| > 1$ это будет **“взрывной” процесс**. Влияние прошлых ошибок в нем не угасает, и все более усиливается со временем. Мы не будем рассматривать такие процессы.

Авторегрессионный процесс первого порядка при $\rho = 1$ называют **случайным блужданием**. Если $\mu = 0$, то это случайное блуждание в собственном смысле слова, а при $\mu \neq 0$ это **случайное блуждание с дрейфом**.

Нет смысла рассматривать случайное блуждание, начавшееся бесконечно давно, поскольку за бесконечное время процесс “уходит в бесконечность”, его дисперсия становится бесконечной.

Для процесса, начавшегося в момент $t = 1$ имеем:

$$Y_t = \mu t + \sum_{i=1}^t \varepsilon_i + Y_0, \quad E(Y_t) = \mu t + Y_0.$$

Таким образом, константа (“дрейф”) в авторегрессионной записи процесса приводит к появлению *линейного тренда* в Y_t . Дисперсия равна

$$\text{var}(Y_t) = t\sigma_\varepsilon^2.$$

Она возрастает бесконечно со временем.

Случайное блуждание является примером авторегрессионного процесса с **единичным корнем**. Он называется так по следующей причине. Запишем AR(1) с помощью лагового оператора:

$$(1 - \rho L)Y_t = \mu + \varepsilon_t.$$

В левой части этого уравнения первый множитель — многочлен первой степени от лага. Корень этого многочлена равен $1/\rho$. При $\rho=1$ корень многочлена равен 1.

В случае авторегрессионного процесса произвольного порядка имеем

$$f(L)Y_t = \mu + \varepsilon_t.$$

Если все корни многочлена $f(\cdot)$ по модулю больше 1, то есть лежат за пределами единичного круга на комплексной плоскости, то процесс стационарен. Если один из корней лежит в пределах единичного круга, то процесс “взрывной”. Если же $k > 0$ корней лежат на единичной окружности, а осталь-

ные — за ее пределами, то процесс нестационарный, но не “взрывной” и о нем говорят, что он имеет k единичных корней.

Первые разности ΔY_t авторегрессионного процесса первого порядка с $\rho=1$ есть просто ошибки ε_t , т.е. первые разности стационарны. Нестационарный процесс, первые разности которого стационарны называют **интегрированным** первого порядка и обозначают $I(1)$. Стационарный процесс обозначают $I(0)$. Если k -е разности случайного процесса стационарны, то его называют интегрированным k -го порядка и обозначают $I(k)$.

Рассмотрим, например, процесс

$$z_t = \sum_{i=1}^t Y_i, \text{ где } Y_t = Y_{t-1} + \varepsilon_t.$$

Он будет $I(2)$, то есть вторые разности ($\Delta^2 z_t$) стационарны.

Ложная регрессия

Очень часто экономические процессы бывают нестационарными. В качестве примера можно привести объем производства, уровень цен. Уровень безработицы как процент трудоспособного населения это, с другой стороны, пример стационарной переменной. В данном случае термин “стационарность” употреблен не в строгом смысле. Скорее подразумевается, что дисперсия процесса ограничена.

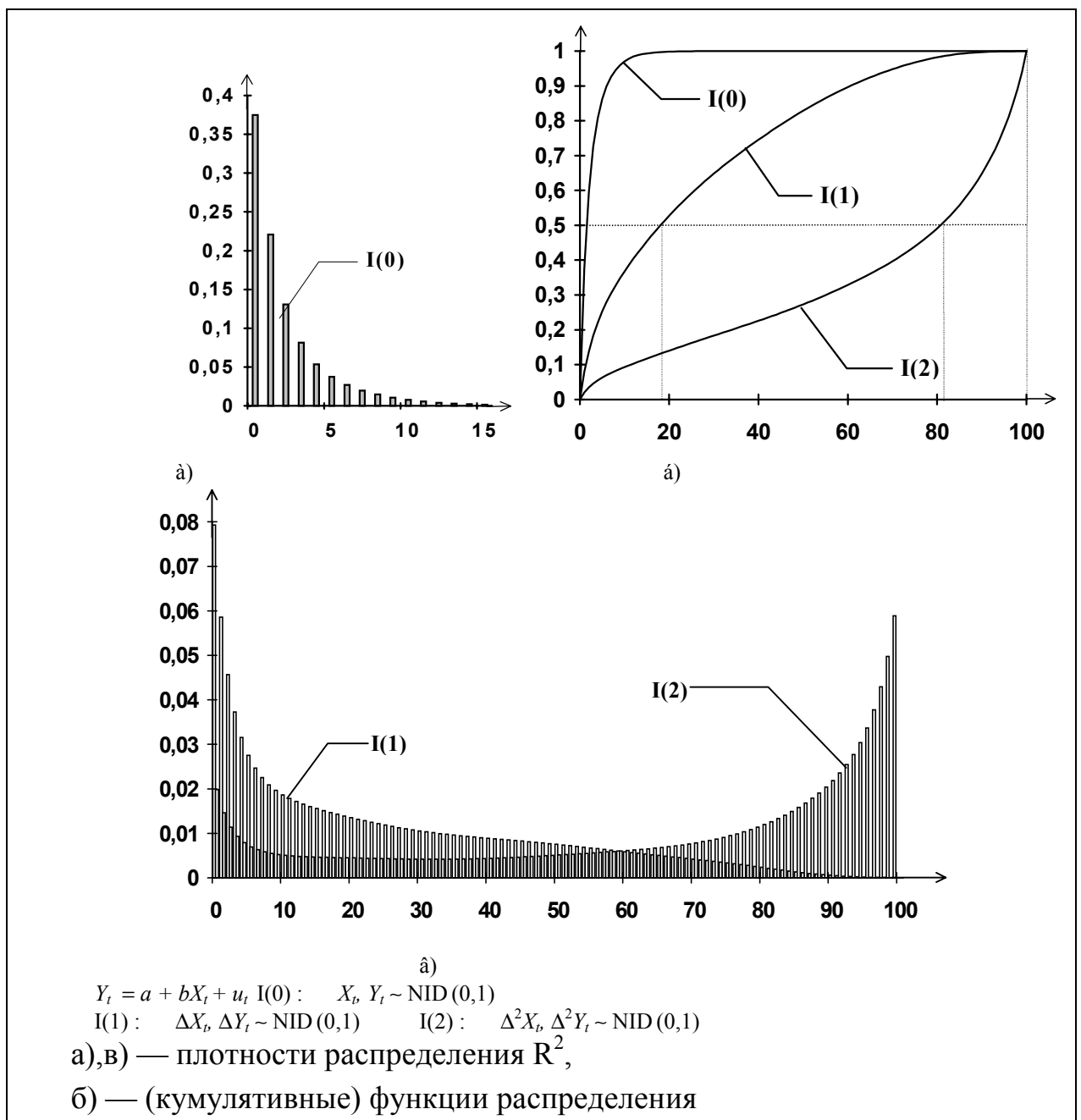
Стационарность регрессоров является очень важным условием при оценивании регрессионных моделей. Если модель неверно специфицирована, и некоторые из переменных, которые в нее неправильно включены, являются $I(1)$, то полученные оценки будут очень плохими. Они не будут обладать свойством состоятельности, то есть не будут сходиться по вероятности к истинным значениям параметров по мере увеличения размеров выборки. Привычные показатели, такие как коэффициент детерминации R^2 , t -статистики, F -статистики, будут указывать на наличие связи там, где на самом деле ее нет. Такой эффект называют **ложной регрессией**.

Показать эффект ложной регрессии можно с помощью метода Монте-Карло. Сгенерируем достаточно много раз два случайных блуждания с независимыми нормально распределенными ошибками ($\varepsilon_t, \xi_t \sim \text{NID}(0, 1)$):

$$Y_t = Y_{t-1} + \varepsilon_t, \quad X_t = X_{t-1} + \xi_t.$$

Оценив достаточно много раз регрессию Y_t по константе и X_t вида $Y_t = a + bX_t + u_t$ мы получим экспериментальное распределение различных статистик. Например, эксперименты Монте-Карло показывают, что t-статистика для b при 50 наблюдениях и номинальном уровне значимости 5% в действительности отвергает верную гипотезу об отсутствии связи примерно в 75% случаев. Вместо того, чтобы использовать 5%-ю критическую границу $t_{5\%} \approx 2$ нужно использовать $t_{5\%} \approx 1,2$.

На рисунке показаны распределения коэффициента детерминации R^2 (в процентах) при длине выборки в 50 наблюдений. Хотя процессы независи-



мы, но регрессия с большой вероятностью дает высокий коэффициент детерминации из-за нестационарности. Два независимых $I(1)$ -процесса примерно в половине случаев дают коэффициент детерминации превышающий 20%. Для $I(2)$ -процессов примерно в половине случаев коэффициент детерминации превышает 80% !

То же самое, хотя и в меньшей степени, можно наблюдать и в случае двух стационарных $AR(1)$ -процессов с коэффициентом автокорреляции ρ близким к 1. Отличие заключается в том, что здесь ложная связь асимптотически (при стремлении размеров выборки к бесконечности) исчезает, а в случае $I(1)$ -процессов — нет. Все же проблема остается серьезной, поскольку на практике экономист имеет дело с конечными и часто довольно малыми выборками.

О процессе типа случайного блуждания без дрейфа говорят как о **стохастическом тренде**, поскольку влияние каждой ошибки не затухает со временем.

Наличие обычного детерминированного тренда также может приводить к появлению ложной регрессии. Пусть, например Y_t и X_t порождаются процессами $Y_t = a + bt + \varepsilon_t$, $X_t = c + dt + \xi_t$, где ε_t , ξ_t — независимые, одинаково распределенные ошибки. Регрессия Y_t по константе и X_t может иметь высокий коэффициент детерминации и этот эффект только усиливается с ростом размера выборки. К счастью, с “детерминированным” вариантом ложной регрессии достаточно легко бороться. В рассматриваемом случае достаточно добавить в уравнение тренд в качестве регрессора, и эффект ложной регрессии исчезает.

Тестирование стационарности

С осознанием опасности применения ОМНК к нестационарным рядам, появилась необходимость в тестах, которые позволили бы отличить стационарный процесс от нестационарного.

К неформальным методам тестирования стационарности можно отнести визуальный анализ графиков спектральной плотности и автокорреляционной функции.

В настоящее время самым популярным из формальных тестов является тест, разработанный Дики и Фуллером (DF). Базовый порождающий данные процесс (ПДП), который они использовали, — авторегрессионный процесс первого порядка:

$$y_t = \rho y_{t-1} + \varepsilon_t. \quad (A1)$$

При $\rho=1$ это случайное блуждание. Конечно, вряд ли экономическая переменная может быть описана процессом (A1). Более реалистично было бы предположить наличие в этом процессе константы и тренда:

$$y_t = \mu_0 + \rho y_{t-1} + \varepsilon_t. \quad (A2)$$

$$y_t = \mu_0 + \mu_1 t + \rho y_{t-1} + \varepsilon_t. \quad (A3)$$

$$y_t = \mu_0 + \mu_1 t + \mu_2 t^2 + \rho y_{t-1} + \varepsilon_t. \quad (A4)$$

Нулевая гипотеза в тесте Дики-Фуллера состоит в том, что ряд нестационарен и имеет один единичный корень ($\rho=1$) (и при этом $\mu_i=0$), альтернативная — что ряд стационарен ($\rho<1$):

$$H_0 : \rho=1, \mu_i=0 \quad H_A : \rho<1.$$

Здесь $i=0$, если оценивается (A2), $i=1$, если оценивается (A3), и $i=2$, если оценивается (A4).

Предполагается, что ошибки ε_t некоррелированы. Это предположение очень важно, без него тест не будет работать!

Для получения статистики, с помощью которой можно было бы проверить нулевую гипотезу, Дики и Фуллер предложили оценить авторегрессию и взять из нее обычную t-статистику для гипотезы о том, что $\rho=1$. При этом тест является *односторонним*, поскольку альтернатива $\rho>1$, соответствующая “взрывному” процессу, не рассматривается.

Необычность DF заключается в том, что с помощью одной t-статистики проверяется гипотеза сразу о двух коэффициентах.⁵ Если мы в регрессии (A3) отвергли нулевую гипотезу, то принимаем альтернативную гипотезу, что процесс описывается уравнением (A3) с $\rho<1$, то есть это стационарный вокруг линейного тренда процесс. В противном случае имеем нестационарный процесс ($\rho=1$), описываемый уравнением (A2), то есть случайное блуждание с дрейфом, но без временного тренда в уравнении авторегрессии.

Часто встречается несколько иная интерпретация этой особенности данного теста: проверяется гипотеза $H_0 : \rho=1$ против гипотезы $H_A : \rho<1$, и *оцениваемая регрессия не совпадает с порождающим данные процессом*, каким он предполагается согласно альтернативной гипотезе. Так, чтобы проверить нулевую гипотезу для ПДП типа (A2) нужно построить регрессию (A3) или

⁵ Если применить другую параметризацию, то об одном.

(A4). Аналогично для тестирования ПДП типа (A3) нужно оценить регрессию (A4). Однако приведенная ранее интерпретация более точная.

Поскольку полученная статистика имеет нестандартное распределение, для ее использования требуются специальные таблицы. Эти таблицы были получены численно методом Монте-Карло. Все эти статистики получены на основе одного и того же ПДП (A1) с $\rho = 1$, но с асимптотической точки зрения годятся и для других ПДП, несмотря на наличие мешающих параметров, которые приходится оценивать.

Чтобы удобно было использовать стандартные регрессионные пакеты, уравнения регрессии преобразуются так, чтобы зависимой переменной была первая разность. В случае (A1) имеем уравнение ($\phi = \rho - 1$):

$$\Delta y_t = \phi y_{t-1} + \varepsilon_t.$$

Будем обозначать статистику, получаемую в результате оценивания регрессии (A1) τ_{nc} , в результате оценивания регрессии (A2) — τ_c , в результате оценивания регрессии (A3) — τ_{ct} и в результате оценивания регрессии (A4) — τ_{ctt} . Это означает, соответственно, что в регрессии нет константы (nc), есть только константа (c), есть константа и линейный временной тренд (ct), есть константа, линейный тренд и квадратичный тренд (ctt). (Дики и Фуллер использовали другие обозначения, здесь используются обозначения Мак-Киннона).

Следующая таблица показывает, какую статистику можно применять в какой ситуации.

	ПДП с $\rho = 1$, соответствующий нулевой гипотезе		
Регрессия	A1	A2	A2
A1	τ_{nc}		
A2	τ_c	t	t
A3	τ_{ct}	τ_{ct}	t
A4	τ_{ctt}	τ_{ctt}	τ_{ctt}

В таблице t обозначает обычную t-статистику. Дело в том, что когда регрессия совпадает с ПДП и в регрессии есть детерминированные переменные (константа, тренд), то обычная t-статистика асимптотически имеет стандарт-

ное нормальное распределение и поэтому для проверки гипотезы годятся обычные критические границы. Правда это свойство существенно асимптотическое, и в малых выборках действительный уровень значимости, как показывают имитации Монте-Карло, может сильно отличаться от номинального. Поэтому предпочтительно добавить в регрессию дополнительную переменную и воспользоваться тестом Дики-Фуллера с нестандартными критическими границами, которые хотя и являются тоже асимптотическими, но связаны с меньшими искажениями размера теста.

Из этой таблицы видно, что если можно предположить, что рассматриваемая переменная нестационарна и имеет тренд, то начать тестирование следует с регрессии (A4) и соответствующего теста τ_{ctt} .

Поскольку неизвестно, присутствуют ли в ПДП константа и тренд, то полезно иметь тесты, которые бы позволили проверить соответствующие гипотезы. Такие тесты были предложены Дики и Фуллером. В случае всех этих тестов (в отличие от DF) действительно проверяемая гипотеза совпадает с номинально проверяемой гипотезой (или, согласно альтернативной интерпретации, оцениваемая регрессия *совпадает* с ПДП, каким он предполагается в соответствии с альтернативной гипотезой). По сути дела используются обычные F- и t-статистики для соответствующих гипотез, только критические границы берут другие. Опять же, при получении этих таблиц методом Монте-Карло используется исключительно ПДП (A1) с $\rho = 1$, поэтому тесты являются асимптотическими.

При оценивании регрессии вида (A2) получаем две статистики: t-статистику для гипотезы $\mu_0 = 0$ и F-статистику для гипотезы $\mu_0 = 0$ и $\rho = 1$. При оценивании регрессии вида (A3) получаем четыре статистики: t-статистику для гипотезы $\mu_0 = 0$, t-статистику для гипотезы $\mu_1 = 0$, F-статистику для гипотезы $\mu_1 = 0$ и $\rho = 1$ и F-статистику для гипотезы $\mu_0 = 0$, $\mu_1 = 0$ и $\rho = 1$.

Было бы естественно предположить, что только что описанные F-статистики было бы предпочтительнее использовать, чем ADF-тесты, поскольку действительная гипотеза для них совпадает с номинальной и является как раз той гипотезой, которая и проверяется в ADF-тестах. Однако эти статистики являются двусторонними и, тем самым, не отбрасывают возможность “взрывного” процесса, что должно приводить к потере мощности теста.

Если гипотеза о наличии единичного корня не была отвергнута, то t -статистики для $\mu = 0$ и $\gamma = 0$ могут быть полезны для определения точного вида нестационарного процесса — имеется ли в нем “дрейф” и тренд.

Предположение о том, что переменная следует авторегрессионному процессу первого порядка и ошибки некоррелированы, является, конечно, слишком ограничительным. Тест Дики-Фуллера был модифицирован для авторегрессионных процессов более высоких порядков и получил название дополненного теста Дики-Фуллера (augmented Dickie-Fuller test, ADF).

Базовые уравнения приобретают следующий вид:

$$\Delta y_t = (\rho - 1) y_{t-1} + \sum_{l=1}^L \Delta y_{t-l} + \varepsilon_t. \quad (B1)$$

$$\Delta y_t = \mu_0 + (\rho - 1) y_{t-1} + \sum_{l=1}^L \Delta y_{t-l} + \varepsilon_t. \quad (B2)$$

$$\Delta y_t = \mu_0 + \mu_1 t + (\rho - 1) y_{t-1} + \sum_{l=1}^L \Delta y_{t-l} + \varepsilon_t. \quad (B3)$$

$$\Delta y_t = \mu_0 + \mu_1 t + \mu_2 t^2 + (\rho - 1) y_{t-1} + \sum_{l=1}^L \Delta y_{t-l} + \varepsilon_t. \quad (B4)$$

Распределения этих тестов асимптотически совпадают с соответствующими обычными тестами Дики-Фуллера, и используют те же таблицы. Грубо говоря, роль дополнительной авторегрессионной компоненты сводится к тому, чтобы убрать автокорреляцию из остатков. Процедура тестирования не отличается от описанной выше.

Как показали эксперименты Монте-Карло, тест Дики-Фуллера чувствителен к наличию процесса типа скользящего среднего в ошибке. Эту проблему частично можно снять, добавляя в регрессию достаточно много лагов первой разности (Said and Dickey, 1984). Чтобы тест был состоятельным, требуется увеличивать L с ростом количества наблюдений по определенному закону.

На практике решающим при использовании ADF является вопрос о том, как выбирать L — порядок AR-процесса в оцениваемой регрессии. Можно предложить следующие подходы.

1) Поскольку важно, чтобы остатки были как можно более похожи на “белый шум”, то следует выбирать такое число L , чтобы тест на автокорреляцию остатков показал отсутствие значимой автокорреляции. Поскольку дополнительные лаги не меняют асимптотические результаты, то лучше

взять больше лагов, чем меньше. Однако этот последний аргумент верен только с асимптотической точки зрения.

2) Другой подход состоит в том, чтобы выбирать L на основе обычных t - и F -статистик для соответствующих дополнительных регрессоров.

ADF может давать разные результаты в зависимости от того, каким выбрано количество лагов. Даже добавление лага, который “не нужен” согласно только что приведенным критериям, может резко изменить результат тестирования.

Особую проблему создает наличие сезонной компоненты в переменной. Если сезонность имеет детерминированный характер, то достаточно добавить в регрессию фиктивные сезонные переменные — это не изменяет асимптотического распределения ADF-статистики. Для случая стохастической сезонности также есть специальные модификации теста.

Пока мы рассмотрели тесты $I(1)$ против $I(0)$. Временной ряд может быть интегрированным и более высокого порядка. Как несложно понять, тесты $I(2)$ против $I(1)$ сводятся к рассмотренным, если взять не уровень тестируемого ряда, а первую разность. Аналогично для более высоких порядков интегрирования.

Имитации показали, что следует проверять гипотезы последовательно, начиная с наиболее высокого порядка интегрирования, который можно ожидать априорно. Т. е., сначала следует проверить гипотезу о том, что ряд является $I(2)$, и лишь после этого, если гипотеза была отвергнута, что он является $I(1)$. (См. Dickey and Pantula, 1987.)

Коинтеграция. Регрессии с интегрированными переменными

Как уже говорилось выше, привычные методы регрессионного анализа не подходят, если переменные нестационарны. Однако не всегда при применении МНК имеет место эффект ложной регрессии.

Говорят, что $I(1)$ -процессы Y_{1t} и Y_{2t} являются **коинтегрированными** первого порядка ($CI(1,0)$), если существует их линейная комбинация, которая является $I(0)$, то есть стационарна. То есть $Y_{1t}, Y_{2t} \sim I(1)$, коинтегрированы, если существует коэффициент λ , такой что $Y_{1t} - \lambda Y_{2t} \sim I(0)$. Понятие коинтеграции введено Грейнджером (Granger(1981)).

Понятие коинтеграции тесно связано с моделью исправления ошибки. Коинтегрированные процессы Y_{1t} и Y_{2t} связаны между собой долгосрочным

стационарным соотношением, и следует предположить, что существует некий корректирующий механизм, который при отклонениях возвращает Y_{1t} и Y_{2t} к их долгосрочному отношению.

Если $\lambda=1$, то разность Y_{1t} и Y_{2t} будет стационарной и, грубо говоря, Y_{1t} и Y_{2t} будут двигаться “параллельно” во времени. Следующий рисунок (Рис. 7) изображает две таких коинтегрированных переменных, динамика которых задана моделью исправления ошибки:

$$X_{1t} = X_{1t-1} - 0.2(Y_{1t-1} - Y_{2t-1} + 2) + \varepsilon_{1t}$$

$$Y_{2t} = Y_{2t-1} + 0.5(Y_{1t-1} - Y_{2t-1} + 2) + \varepsilon_{2t}$$

$$\varepsilon_{1t}, \varepsilon_{2t} \sim \text{NID}(0,1).$$

Определение коинтеграции естественным образом распространяется на случай нескольких коинтегрированных переменных произвольного порядка интегрирования. Компоненты n -мерного векторного процесса $Y_t = (Y_{1t}, \dots, Y_{nt})$ называют **коинтегрированными** порядка d, b , что обозначается $Y_t \sim \text{CI}(d, b)$, если (1) Y_{it} является $I(d)$ $i=1, \dots, n$ и (2) существует отличный от нуля вектор β , такой что $Y_t \beta \sim I(d-b)$, $d \geq b > 0$. Вектор β называют **коинтегрирующим вектором**.

В рассмотренном ранее примере коинтеграционный вектор имеет вид $\beta = (-1, \lambda)$. Его можно пронормировать также как $(-1/\lambda, 1)$.

Если переменные в регрессии не стационарны, но действительно связаны друг с другом стационарной линейной комбинацией (модель специфицирована верно), то полученные оценки коэффициентов этой линейной комби-

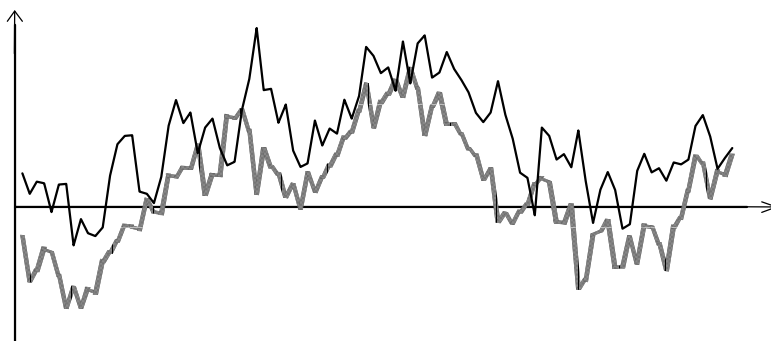


Рис. 7. Два коинтегрированных процесса при $\lambda=1$.

нации будут на самом деле сверхсостоятельными, то есть сходятся по веро-

ятности к истинным коэффициентам со скоростью, пропорциональной не квадратному корню количества наблюдений, как в регрессии со стационарными переменными, а со скоростью, пропорциональной просто количеству наблюдений. Другими словами в обычной регрессии $\sqrt{N}(\hat{\lambda} - \lambda)$ имеет невырожденное асимптотическое распределение, а в регрессии с I(1)-переменными $N(\hat{\lambda} - \lambda)$ имеет невырожденное асимптотическое распределение.

Обычные асимптотические аргументы сохраняют свою силу, если речь идет об оценках параметров краткосрочной динамики в модели исправления ошибок. Таким образом, можно использовать t-статистики, получаемые обычным методом наименьших квадратов, для проверки гипотез о значимости отдельных переменных. Важно помнить, что это относится к оценкам краткосрочных параметров. Этот подход не годится для проверки гипотез о коэффициентах коинтеграционной комбинации.

Оценивание коинтеграционной регрессии: подход Энгла-Грейнджера

Если бы коэффициент λ был известен, то проверка на коинтегрированность была бы эквивалентна проверке $Y_{1t} - \lambda Y_{2t}$ на стационарность. Но в практических проблемах обычно стационарная линейная комбинация неизвестна. Значит, необходимо оценить коинтегрирующий вектор. Следует также проверить, действительно ли этот вектор дает стационарную линейную комбинацию.

Простейшим методом отыскания стационарной линейной комбинации является **метод Энгла-Грейнджера**. Энгл и Грейнджер предложили использовать оценки, полученные из обычной регрессии с помощью метода наименьших квадратов. Одна из переменных должна стоять в левой части регрессии, другая — в правой:

$$Y_{1t} = \lambda Y_{2t} + u_t.$$

Для тестирования стационарности полученной линейной комбинации предлагается применить метод Дики-Фуллера к остаткам из коинтеграционной регрессии. Пусть — \hat{u}_t остатки из этой регрессии. Тест Энгла-Грейнджера проводится с помощью регрессии

$$\hat{u}_t = \rho \hat{u}_{t-1} + \text{остатки}.$$

Распределение t-статистики для гипотезы $\rho=1$ в этой регрессии будет отличаться (даже асимптотически), от распределения DF-статистики, но имеются соответствующие таблицы. Нулевой гипотезой, таким образом, является отсутствие коинтеграции. Если мы отвергаем гипотезу об отсутствии коинтеграции, то это дает уверенность в том, что полученные результаты не являются ложной регрессией.

Игнорирование детерминированных компонент ведет к неверным выводам о коинтеграции. Чтобы этого избежать, в коинтеграционную регрессию следует добавить соответствующие переменные — константу, тренд, квадрат тренда, сезонные фиктивные переменные. Добавление константы, тренда, и квадрата тренда, как и в случае DF, меняет асимптотическое распределение теста Энгла-Грейнджера. Следует помнить, что, в отличие от DF, регрессия, из которой берется t-статистика, остается неизменной, то есть в нее не нужно добавлять детерминированные регрессоры.

В МНК регрессии с коинтегрированными переменными оценки должны быть смещенными из-за того, что в правой части стоит эндогенная переменная, коррелированная с ошибкой. Кроме того, ошибка содержит пропущенные переменные. Коинтеграционная регрессия Энгла-Грейнджера является статической по форме, то есть не содержит лагов переменных. С асимптотической точки зрения не приводит к смещенности оценок, поскольку ошибка является величиной меньшего порядка, чем регрессор, дисперсия которого стремится к бесконечности. Как уже говорилось, оценки на самом деле сверхсостоятельны. Однако в малых выборках смещение может быть существенным.

После того, как найдена стационарная линейная комбинация, можно оценить модель исправления ошибок, которая делает переменные коинтегрированными. В этой регрессии нужно использовать первые разности исходных переменных и остатки из коинтеграционной регрессии, которые будут представлять корректирующий член модели исправления ошибок.

Подчеркнем роль корректирующего члена. До появления метода Энгла-Грейнджера исследователи часто оценивали регрессии в первых разностях, что, хотя и приводило к стационарности переменных, но не учитывался стационарный корректирующий член, то есть регрессионная модель была неверно специфицирована (проблема пропущенной переменной).

Несмотря на то, что в модели исправления ошибок используется оценка коинтегрирующего вектора, оценки коэффициентов, полученные из такой

модели будут иметь такие же асимптотические свойства, как если бы коинтегрирующий вектор был точно известен. В частности, можно использовать t -статистики из этой регрессии, поскольку оценки стандартных ошибок являются состоятельными. Это является следствием сверхсостоятельности оценок коинтегрирующего вектора.

Коинтеграция в динамических системах: подход Йохансена

Другой популярный метод нахождения стационарных комбинаций — **метод Йохансена**. Этот метод служит также для тестирования стационарности найденных линейных комбинаций, и по сути дела распространяет методику Дики-Фуллера на случай **векторной авторегрессии** (то есть такой модели, в которой несколько зависимых переменных и зависят они от собственных лагов и от лагов других переменных). Если в обычной авторегрессии мы рассматривали один коэффициент ρ , то здесь следует рассматривать уже матрицу коэффициентов. Предполагается (как и в ADF), что если добавить достаточное число лагов в авторегрессионную модель, то ошибка не будет сериально коррелированной.

Если векторный процесс состоит более чем из двух процессов ($S > 2$), то может существовать несколько коинтегрирующих векторов. Если существует ровно r линейно независимых коинтегрирующих векторов, то говорят, что **ранг коинтеграции** равен r .

Обозначим β матрицу, составленную из таких векторов. Набор коинтегрирующих векторов не является однозначным, на самом деле речь должна идти о коинтеграционном пространстве. Нормировку следует выбирать исходя из экономической теории рассматриваемых процессов.

Метод Йохансена позволяет не только найти матрицу коинтеграционных векторов при данном ранге коинтеграции, но и проверять гипотезы о ранге коинтеграции (количестве коинтегрирующих векторов). Метод непосредственно работает с **векторной моделью исправления ошибок**. Пусть $Y_t = (Y_{1t}, \dots, Y_{nt})$ — векторный процесс (вектор-строка), каждая из компонент которого является $I(1)$ (или $I(0)$). Порождающий данные процесс задается формулой

$$\Delta Y_t = \mu_0 + \mu_1 t + Y_{t-1} \Pi + \Delta Y_{t-1} \Gamma_1 + \dots + \Delta Y_{t-L+1} \Gamma_{L-1} + \varepsilon_t.$$

Предполагается, что ошибки, относящиеся к разным моментам времени, независимы, и $\varepsilon_t \sim N(0, \Omega)$. В модели оцениваются вектор-строка констант μ_0 и коэффициентов при трендах μ_1 , матрицы коэффициентов $\Gamma_1, \dots, \Gamma_{L-1}$ и Π

$(n \times n)$, а также ковариационная матрица Ω . Поскольку по предположению $\Delta Y_t \sim I(0)$, то должно быть выполнено $Y_{t-1} \Pi \sim I(0)$. Ограничения на ранг коинтеграции задаются как ограничения на матрицу Π . При нулевой гипотезе, что ранг коинтеграции равен r , ее можно представить в виде

$$H_0(r): \Pi = \beta \alpha^T,$$

где матрицы α и β имеют размерность $(n \times r)$; β — матрица коинтегрирующих векторов, α — матрица корректирующих коэффициентов. Если $r = 0$, то $\Pi = 0$ и не существует стационарных линейных комбинаций переменных Y_{1t}, \dots, Y_{nt} . В другом крайнем случае, когда $n = r$ любая линейная комбинация этих переменных стационарна, то есть все они $I(0)$.

Для оценивания модели используется метод максимального правдоподобия. При данной матрице β можно получить оценки максимального правдоподобия для остальных неизвестных параметров обычным методом наименьших квадратов. Йохансен показал также, что максимизация функции правдоподобия по β эквивалентна задаче отыскания собственных чисел для некоторой симметричной положительно определенной матрицы. При ранге коинтеграции r выбираются r минимальных собственных чисел. Если расположить собственные числа в порядке возрастания $(\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n)$, то следует выбрать $\lambda_1, \lambda_2, \dots, \lambda_r$. (Йохансен записал ПДП в несколько ином виде, и поэтому у него собственные числа идут в порядке убывания и выбираются r максимальных собственных чисел.) Столбцами матрицы β (коинтегрирующими векторами) будут соответствующие собственные вектора. Конечно, β определяется только с точностью до некоторой нормировки. После того, как найдена оценка максимального правдоподобия $\hat{\beta}$, вычисляются оценки других параметров.

Для проверки гипотез об r используется статистика отношения правдоподобия. **Статистика следа** используется для проверки гипотезы (H_0) о том, что ранг равен r , против гипотезы (H_A) о том, что ранг равен n . Статистика имеет вид

$$LR^{\text{trace}} = -T \sum_{i=r+1}^n \ln(1 - \lambda_i).$$

Тестирование проводится последовательно для $r = n-1, \dots, 0$ и заканчивается, когда нулевая гипотеза не будет отвергнута в первый раз. Можно проводить тестирование в обратном порядке $r = 0, \dots, n-1$. В этом случае тести-

рование заканчивается, когда нулевая гипотеза будет отвергнута в первый раз.

Можно также использовать **статистику максимального собственного числа**, которая используется для проверки гипотезы (H_0) о том, что ранг равен r , против гипотезы (H_A) о том, что ранг равен $r+1$. Эта статистика равна

$$LR^{\lambda\text{-max}} = -\ln(1 - \lambda_{r+1}).$$

Обе статистики имеют нестандартные асимптотические распределения. К счастью, их распределения не зависят от мешающих параметров. Распределение этих статистик зависит только от $n-r$ и от того, как входят в модель константа и тренд.

Можно выделить пять основных случаев, касающихся статуса векторов μ_0 и μ_1 в модели. В порядке перехода от частного к более общему:

Случай 0. $\mu_0 = 0, \quad \mu_1 = 0.$

Случай 1*. $\mu_0 = \gamma\alpha^T, \quad \mu_1 = 0.$

Случай 1. μ_0 произвольный, $\mu_1 = 0.$

Случай 2*. μ_0 произвольный, $\mu_1 = \gamma\alpha^T.$

Случай 2. μ_0 произвольный, μ_1 произвольный.

Здесь γ_0 и γ_1 — вектора-строки длины r . Случай 0 легко понять — константы и тренды в модели полностью отсутствуют. В Случае 1 константа входит в коинтеграционное пространство и, тем самым, в корректирующие механизмы, но не входит в сам процесс Y_t в виде дрейфа. Это легко увидеть, если переписать модель следующим образом.

$$\Delta Y_t = (\gamma_0 + Y_{t-L} \beta) \alpha^T + \Delta Y_{t-1} \Gamma_1 + \dots + \Delta Y_{t-L+1} \Gamma_{L-1} + \varepsilon_t.$$

В Случае 1 μ_0 можно записать как $\mu_0 = \gamma_0 \alpha^T + \mu_0^*$, где γ_0 входит в коинтеграционное пространство, а μ_0^* соответствует дрейфу в векторной модели исправления ошибок. Дрейф в модели исправления ошибок означает, что в Y_t входит линейный тренд. (См. выше рассмотрение простого авторегрессионного процесса с дрейфом.)

Аналогичные рассуждения верны по отношению ко временному тренду в Случаях 2* и 2. В Случае 2* тренд входит в коинтеграционное пространство, но не входит в Y_t в виде квадратичного тренда. В Случае 2 тренд входит и в коинтеграционное пространство, и в Y_t в виде квадратичного тренда.

Методом Монте-Карло получены таблицы LR^{trace} и $LR^{\lambda\text{-max}}$ для всех пяти случаев и нескольких значений $n-r$ (на данный момент имеются таблицы для $n-r=1, \dots, 12$).

Как и в случае ADF очень важным вопросом является выбор длины лага L . Способы по сути дела являются теми же самыми. Для проверки гипотез о длине лага можно использовать тест отношения правдоподобия, который в данном случае имеет обычное распределение χ^2 . Если процесс состоит из n компонент, и проверяется гипотеза о том, что следует увеличить L на единицу то количество степеней свободы соответствующей статистики равно n . Важно также, чтобы отсутствовала автокорреляция остатков.

Метод Йохансена можно использовать также для оценивания моделей с линейными ограничениями на матрицу коинтегрирующих векторов β и на матрицу корректирующих коэффициентов α . Для проверки таких ограничений предлагается использовать все тот же тест отношения правдоподобия, который здесь имеет обычное асимптотическое распределение χ^2 .

Литература по единичным корням и коинтеграции

- Banerjee, A., J.J. Dolado, D.F. Hendry, and G.W. Smith, "Exploring Equilibrium Relationships in Econometrics Through Static Models: Some Monte Carlo Evidence," *Oxford Bulletin of Economics and Statistics*, **48** (1986), 253-277.
- Banerjee, A. J.J. Dolado, J.W. Galbraith and D.F. Hendry, *Co-integration, Error Correction, and the Econometric Analysis of Nonstationary Data*. Oxford: Oxford University Press, 1993.
- Dickey, D.A., W.R. Bell and R.B Miller, "Unit Roots in Time Series Models: Tests and Implications," *American Statistician*, **40** (1986), 12-26.
- Dickey, D.A. and W.A. Fuller, "Distributions of the Estimators for Autoregressive Time Series With a Unit Root," *Journal of American Statistical Association*, **75** (1979), 427-431.
- Dickey, D.A. and S.G. Pantula, "Determining the Order of Differencing in Autoregressive Processes," *Journal of Business and Economic Statistics*, **5** (1987), 455-461.
- Engle, R.F. and C.W.J. Granger, "Co-integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, **55** (1987), 251-276.
- Engle, R.F. and B.S. Yoo, "Forecasting and Testing in Cointegrated Systems," *Journal of Econometrics*, **35** (1987), 143-159.
- Fuller, W.A. *Introduction to Statistical Time Series*. NY: Wiley, 1976.
- Granger C.W.J., "Some Properties of Time Series Data and their Use in Econometric Model Specification," *Journal of Econometrics*, **16** (1981) 121-130.
- Hendry, D.F. "Econometric Modelling with Cointegrated Variables: An Overview," *Oxford Bulletin of Economics and Statistics*, **48** (1986), 201-212.
- Johansen, S., "Statistical Analysis of Cointegration vectors," *Journal of Economic Dynamics and Control*, **12** (1988), 231-254.
- Johansen, S., "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, **59** (1991), 1551-1580.
- Johansen, S., "The Role of the Constant and Linear Terms in Cointegration Analysis of Nonstationary Data," *Econometric Reviews*, **13** (1994), 205-229.
- Johansen, S. and K. Juselius, "Maximum Likelihood Estimation and Inference on Cointegration with Application to the Demand for Money," *Oxford Bulletin of Economics and Statistics*, **52** (1990), 169-210.

- Ouliaris, S., J.Y. Park and P.C.B. Phillips, “Testing for a Unit Root in the Presence of a Maintained Trend,” Ch. 1 in *Advances in Econometrics*, ed. B. Raj, Boston: Kluwer Academic Publishers, 1989.
- Perron, P. “Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach,” *Journal of Economic Dynamics and Control*, **12** (1988), 297-332.
- Phillips, P.C.B., “Time Series Regression with a Unit Root,” *Econometrica*, **55** (1987), 277-301.
- Phillips, P.C.B. and P. Perron, “Testing for a Unit Root in Time Series Regression,” *Biometrika*, **75** (1988) 335-346.
- Said, E.S. and D.A. Dickey, “Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order,” *Biometrika*, **71** (1984), 599-607.
- Sims, C.A., J.H. Stock and M. Watson, “Inference in Linear Time Series Models with some Unit Roots,” *Econometrica*, **58** (1990),113-144.
- Stock, J.H., “Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors,” *Econometrica*, **55** (1987), 1035-1056.
- Stock, J.H. and M. Watson, “Variable Trends in Economic Time Series,” *Journal of Economic Perspectives*, **2** (1988), 147-174.
- Stock, J.H. and M. Watson, “Testing for Common Trends,” *Journal of the American Statistical Association*, **83** (1988), 1097-1107.
- West, K.D., “Asymptotic Normality When Regressors Have a Unit Root,” *Econometrica*, **56** (1988), 1397-1417.

Предметный указатель

A

ADF • 40

ADL • 29

AIC • 8

D

DF • 36

E

ECM • 31

J

J-тест • 8

P

P_E-тест • 9

A

авторегрессионная модель с

распределенным лагом • 28

авторегрессия • 29, 32

векторная • 45

адаптивных ожиданий модель • 30

аддитивная ошибка • 4

Б

бинарная зависимая переменная • 18

Бокса-Кокса преобразование • 5

В

Вейбулла распределение • 24

векторная авторегрессия • 45

взаимодействия члены • 5

взрывной процесс • 33

внутренне линейная модель • 4

выброс • 12

Г

гетероскедастичность • 3

Д

Дики-Фуллера тест • 36

дополненный • 40

динамическая регрессия • 28

долгосрочное стационарное состояние • 31

дрейф • 33

Е

единичный корень • 33

И

интегрированный процесс • 34

информационный критерий Акаике • 8, 22

исправления ошибок модель • 31

векторная • 45

Й

Йохансена метод • 45

К

качественная зависимая переменная • 18

коинтеграции ранг • 45

коинтегрированные процессы • 41

коинтегрирующий вектор • 42

Койка преобразование • 28

Л

лаг • 25

лаговый оператор • 25

линейная по виду модель • 3

линейная по параметрам модель • 3

линейный тренд • 14

логистическое распределение • 19

логит • 20

логлинейная форма модели • 7

ложная регрессия • 34

М

максимального собственного числа

статистика • 47

множественный логит • 23

модель выбора • 19

мультипликативная ошибка • 4

П

полиномиальная форма модели • 5

полиномиальный лаг • 26

полулогарифмическая форма модели • 7

пробит • 20

производственная функция

Кобба-Дугласа • 4

с постоянной эластичностью замены • 4

транслоговая • 5
псевдо- R^2 для моделей с бинарной
зависимой переменной • 22

Р
распределенный лаг • 25

С
сезонные переменные • 14

следа статистика • 46

случайное блуждание • 33

спектральный анализ • 16

стационарность • 32

тестирование • 36

стохастический тренд • 16, 36

Т
тест

Дики-Фуллера • 36

на адекватность предсказаний • 13

на гетероскедастичность • 3

на стационарность • 36

Чоу • 12

У
упорядоченный логит • 23

Ф
фиктивные переменные • 10

функциональная форма

неправильная спецификация • 2

Ч
частичного приспособления модель • 29

Э
Энгла-Грейнджера метод • 43