Набережночелнинский институт

федерального государственного автономного образовательного

учреждения высшего образования

«КАЗАНСКИЙ (ПРИВОЛЖСКИЙ)

ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Кафедра

«Бизнес-информатика и математические методы в экономике»

Методы эконометрического моделирования и анализа социально-экономических явлений

Электронный образовательный ресурс

Набережные Челны, 2018

УДК 330.43(075.8)
ББК 65в6я73
Р 64

Рецензенты: Исавнин А.Г., зав. каф. математических методов в экономике Набережночелнинского института КФУ, д.ф.-м.н., профессор
Павликов С.В., профессор кафедры ЕНД Набережночелнинского института (факультета) КНИТУ-КАИ, д.ф.-м.н.

Розенцвайг А.К. Методы эконометрического моделирования и анализа социально-экономических явлений: Электронный образовательный ресурс / А.К. Розенцвайг -. – Наб.Челны. - Набережночелнинский институт Казанского федерального университета. – 2018. -122 с.

Электронный образовательный ресурс содержит основные категории, понятия эконометрики как науки, объединяющей качественные методы экономической теории с количественными методами теории вероятностей, математической статистики и экономической статистики. Рассмотрены характерные особенности и проблемы методологии эконометрического моделирования, с которыми приходится сталкиваться экономистам при работе с эконометрическими методами. Особое внимание уделено анализу и построению взаимосвязей экономических переменных с учетом временной структуры статистических данных, что позволяет адекватно отразить их в форме математических и эконометрических моделей реальных экономических явлений и процессов.

Предназначено для самостоятельной работы по изучению эконометрики и эконометрического моделирования статистике, изучаемой студентами дневной и заочной формы обучения специальностей 230700 «Прикладная информатика» и 080500 «Бизнес-информатика» (Интеллектуальные методы бизнес-аналитики).

ВВЕДЕНИЕ

Постоянно усложняющиеся процессы и явления реальной экономической жизни привели к необходимости создания и совершенствования особых методов их изучения и анализа. При этом широкое распространение получило использование методов математического моделирования и количественного анализа. На базе последних выделилось и сформировалось одно из направлений экономических исследований - эконометрика.

Формально «эконометрика» означает «измерения в экономике». Однако область исследований данной дисциплины гораздо шире. *Эконометрика* - это наука, в которой на базе реальных статистических данных строятся, анализируются и совершенствуются математические модели реальных экономических явлений. Обоснование и интерпретация самих моделей является задачей экономической теории, а эконометрика позволяет найти количественное подтверждение либо опровержение того или иного существующего экономического закона либо новой теоретической гипотезы. Одним из важнейших направлений эконометрики является построение прогнозов по различным экономическим показателям.

Эконометрика как научная дисциплина зародилась и получила развитие на основе слияния экономической теории, математической экономики, экономической и математической статистики. Действительно, предмет ее исследования - экономические явления. Но в отличие от экономической теории эконометрика делает упор на количественные, а не на качественные аспекты этих явлений.

Например, экономическая теория утверждает, что спрос на товар с ростом его цены убывает. Но при этом практически неисследованным остается вопрос, как быстро и по какому закону происходит это убывание. Эконометрика отвечает на этот вопрос для каждого конкретного вида товара, места и времени его реализации.

Изучение экономических процессов (взаимосвязей) в эконометрике осуществляется через математические (эконометрические) модели. В этом состоит ее родство с математической экономикой. Но если математическая экономика строит и анализирует обобщенные модели без использования реальных числовых значений, то эконометрика концентрируется на изучении моделей, обоснованных экономической теорией, и сопоставлении их с реальными статистическими данными.

Одной из основных задач экономической статистики является сбор, обработка и представление экономических данных в наглядной форме: в виде таблиц, графиков, диаграмм. Эконометрика также активно пользуется этим инструментарием, но идет дальше, применяя его для анализа теоретически обоснованных экономических взаимосвязей и прогнозирования.

Мощным инструментом эконометрических исследований является аппарат математической статистики. Действительно, большинство экономических показателей носит характер случайных величин, предсказать точные значения которых практически невозможно. Например, весьма сложно предвидеть доход или потребление какого-либо индивидуума, объемы экспорта и импорта страны в течение следующего года и т.д.

Связи между экономическими показателями также не могут носить строгий функциональный характер, а допускают наличие случайных неконтролируемых отклонений (особенно это касается макроэкономических данных). Вследствие этого использование методов математической статистики в эконометрике является естественным и вполне обоснованным. Однако в силу специфики получения статистических данных в экономике (например, в экономике невозможно проведение управляемого эксперимента) эконометрике приходится создавать свои собственные наработки и специальные приемы анализа, которые в математической статистике обычно не рассматривают.

При эконометрическом исследовании имеют место две стороны проблемы обеспечения необходимого качества его результатов - качественная и количественная. Качественная сторона заключается в установлении соответствия между построенной эконометрической моделью и лежащими в ее основе положениями экономической теории. Другая - количественная - состоит в обеспечении наиболее полного соответствия между количественными характеристиками модели и статистической информацией, характерной для поведения изучаемых социально - экономических явлений и процессов в реальных условиях.

Эконометрическое моделирование связано с обоснованием аналитических зависимостей для анализа конкретных экономических показателей. Для этого используют эконометрические методы исследования, связанные с проверкой, обоснованием, оцениванием количественных закономерностей и качественных утверждений (гипотез) в микро- и макроэкономике на основе анализа статистических данных. Результатом такого исследования являются экономические модели и оценки их параметров, проверка гипотез о свойствах экономических показателей и формах их связи. Полученные результаты экономического анализа используют, в частности, для прогнозирования и принятия обоснованных экономических решений.

1. МОДЕЛИРОВАНИЕ И ИСПОЛЬЗОВАНИЕ МОДЕЛЕЙ В РАЗВИТИИ ЭКОНОМИЧЕСКОЙ ТЕОРИИ

Предметом макроэкономической теории является изучение макроэкономических явлений, которые не связаны с какой-то одной отраслью экономики, а имеют отношение ко всем отраслям экономики и должны получить общее (макроэкономическое) объяснение. Макроэкономика рассматривает как изменение объёмов производства и занятости в долгосрочной перспективе (экономический рост), так и их краткосрочные колебания, которые образуют циклы деловой активности.

Макроэкономика на сегодняшний день не является сложившейся законченной дисциплиной, и изучение ключевых вопросов макроэкономики продолжается и сегодня. Изучая макроэкономику, нужно принимать во внимание то, что по некоторым вопросам существует несколько теорий, которые пытаются с разных точек зрения объяснить то или иное явление. Следует также обращать внимание на предпосылки, на которых базируется та или иная теория, и оценивать адекватность этих предпосылок в каждой конкретной ситуации, к которой принято решение применить ту или иную теорию.

Все современные макроэкономические концепции имеют микроэкономическое обоснование: в их основе лежат определенные поведенческие микроэкономические модели, результаты которых агрегируются и затем исследуются на макроуровне. Основным проблемным местом остается теория агрегирования, которая также активно развивается. Агрегирование необходимо не только в теории, но и на практике (при сборе и обработке статистических данных, которые составляют основу для эмпирического анализа).

В макроэкономике рассматривают такие агрегированные экономические переменные, как совокупный выпуск, потребление, инвестиции, экспорт и импорт, уровень цен и так далее. Принято также рассматривать следующие агрегированные рынки: рынок товаров, рынок труда и рынок активов.

Современная экономическая теория, как на микро-, так и на макроуровне, включает математические модели и методы а качестве естественного, необходимого элемента. Использование математики в экономике позволяет, во-первых, выделить и формально описать наиболее важные, существенные связи экономических переменных и объектов: изучение столь сложного объекта предполагает высокую степень абстракции. Во-вторых, из четко сформулированных исходных данных и соотношений методами дедукции можно получать выводы, адекватные изучаемому объекту в той же мере, что и сделанные предпосылки. В-третьих, методы математики и статистики позволяют индуктивным путем получать новые знания об объекте: оценивать форму и параметры зависимостей его переменных, в наибольшей степени соответствующие имеющимся наблюдениям. Наконец, в-четвертых, использование языка математики позволяет точно и компактно излагать положения экономической теории, в общей форме формулировать ее понятия и выводы.

Экономические модели. Для изучения различных экономических явлений и процессов специалисты используют их упрощенные формальные описания, называемые экономическими моделями. Примерами экономических моделей являются модели потребительского спроса, модели фирмы, модели экономического роста, модели равновесия на товарных, факторных и финансовых рынках и многие другие. Строя модели, экономисты-теоретики выявляют существенные факторы, определяющие исследуемое явление и отбрасывают детали, несущественные для решения поставленной проблемы. Достигаемая при этом формализация основных особенностей функционирования экономических объектов позволяет оценить возможные последствия воздействия на них и использовать такие оценки в управлении и принятии решений.

Как обычно строится экономическая модель?

1. Формулируются предмет и цели исследования.

2. В рассматриваемой экономической системе выделяются структурные или функциональные элементы, соответствующие данной цели, выявляются наиболее важные качественные характеристики этих элементов.

3. Словесно, качественно описываются взаимосвязи между элементами модели.

4. Вводятся символические обозначения для учитываемых характеристик экономического объекта и формализуются, насколько возможно (или необходимо), взаимосвязи между ними. Тем самым, формулируется математическая модель.

5. Проводятся расчеты по математической модели и анализ полученного решения.

Математическая структура модели и ее интерпретация/ Следует различать математическую структуру модели и ее содержательную интерпретацию. Рассмотрим следующие два простых примера.

Пример 1. Пусть требуется определить, какую сумму следует положить в банк при заданной ставке процента (20% годовых), чтобы через год получить 12000 рублей?

Вводя формальные обозначения для величин, фигурирующих в задаче:

начальная сумма денег – M_0,

конечная сумма денег – M_1,

ставка процента - R

и, записывая соотношение между ними

$$M_1 = M_0\left[1+\frac{R}{100}\right],$$

найдем требуемую величину из решения основного уравнения модели

$$M_0 = M_1 / \left[1+\frac{R}{100}\right] = \frac{12000}{1{,}2} = 10000\,руб.$$

Пример 2. Пусть требуется определить, каков был объем выпуска продукции завода, если в результате технического перевооружения средняя производительность труда увеличилась на 20%, и завод стал выпускать 12000 единиц продукции.

Вводя формальные обозначения для величин, фигурирующих в задаче:

начальный выпуск - Q_0,

конечный выпуск - Q_1,

процент прироста производительности - R,

и, записывая соотношение между ними (вытекающее из определения средней производительности труда Q/L)

$$Q_1 = Q_0 \frac{L_1}{L_0} = Q_0 \left[1 + \frac{(L_1 - L_0)}{L_0} \right] = Q_0 \left[1 + \frac{R}{100} \right]$$

найдем искомую величину из решения основного уравнения модели

$$Q_0 = Q_1 / \left[1 + \frac{R}{100} \right] = \frac{12000}{1{,}2} = 10000 \, \textit{ед.продукции.}$$

Сравнивая полученные модели и результаты, нетрудно заметить, что математическая форма модели

$$X_1 = X_0 / \left[1 + \frac{R}{100} \right]$$

и даже числовые значения входящих в нее величин в обоих случаях одинаковы. Однако в каждом примере экономическая ситуация, описываемая моделью, экономическая интерпретация самой модели и результатов расчета на основании модельных соотношений являются различными. Таким образом, одни и те же математические модели и методы могут быть использованы для решения принципиально различных экономических задач.

Роль моделей в экономической теории и принятии решений. Экономические модели позволяют выявить особенности функционирования экономического объекта и на основе этого предсказывать будущее поведение объекта при изменении каких-либо параметров. Предсказание будущих изменений, например, повышение обменного курса, ухудшение экономической конъюнктуры, падение прибыли может опираться лишь на интуицию. Однако при этом могут быть упущены, неправильно определены или неверно оценены важные взаимосвязи экономических показателей. К тому же они могут проявлять себя в данной рассматриваемой ситуации совершенно по-другому.

Обычно выбор модели опирается на базовые положения экономической теории, знания об общем характере изучаемых зависимостей на предыдущих этапах исследований, некоторые субъективные предположения (теоретические гипотезы). Поэтому из всех возможных формулировок модели отбирается та, которая в наибольшей степени соответствует реальным статистическим данным, а также сути происходящих экономических процессов.

В модели все взаимосвязи переменных формулируются в явном виде и могут быть оценены количественно, что позволяет получить более качественный и надежный прогноз. Для любого экономического субъекта возможность анализа и прогнозирования изменений в текущей ситуации означает, прежде всего, получение лучших результатов или избежание потерь, как на уровне фирм, так и в государственной политике.

Неполнота экономической модели. По своему определению любая экономическая модель абстрактна и, следовательно, неполна. Выделяя наиболее существенные факторы, определяющие теоретические закономерности функционирования рассматриваемого экономического объекта, она абстрагируется от ряда других, менее известных факторов. Однако, несмотря на свою относительно более редкую применимость, они в некоторых конкретных случаях являются реальными и в совокупности могут определять не только отклонения в поведении объекта, но и само его поведение.

Так в простейшей модели спроса считается, что величина спроса на какой-либо товар определяется его ценой и доходом потребителя. На самом же деле на величину спроса оказывает также влияние ряд других факторов: вкусы и ожидания потребителей, цены на другие товары, воздействие рекламы, моды и так далее. Обычно предполагают, что все факторы, не учтенные явно в экономической модели, оказывают на объект относительно малое результирующее воздействие в интересующем нас аспекте. Состав факторов, учтенных в модели, а также и ее структура могут быть уточнены в ходе совершенствования модели по мере уточнения и расширения области ее применения.

Основные типы математических моделей. Математические модели, используемые в экономике, можно подразделять на классы по ряду признаков, относящихся к особенностям моделируемого объекта, цели моделирования и используемого инструментария. Это модели макро- и микроэкономические, теоретические и прикладные, оптимизационные и равновесные, статические и динамические.

Макроэкономические модели описывают экономику как единое целое, связывая между собой укрупненные материальные и финансовые показатели: ВНП, потребление, инвестиции, занятость, процентную ставку, количество денег и другие. Микроэкономические модели описывают взаимодействие структурных и функциональных составляющих экономики, либо поведение отдельной такой составляющей в рыночной среде. Вследствие разнообразия типов экономических элементов и форм их взаимодействия на рынке, микроэкономическое моделирование занимает основную часть экономико-математической теории. В микроэкономическом моделировании наиболее серьезные теоретические результаты в последние годы получены при исследовании стратегического поведения фирм в условиях олигополии с использованием аппарата теории игр.

Теоретические модели позволяют изучать общие свойства экономики и ее характерных элементов дедукцией выводов из формальных предпосылок. Прикладные модели дают возможность оценить параметры функционирования конкретного экономического объекта и сформулировать рекомендации для принятия практических решений. К прикладным относятся прежде всего эконометрические модели, оперирующие числовыми значениями экономических переменных и позволяющие статистически значимо оценивать их на основе имеющихся наблюдений.

В моделировании рыночной экономики особое место занимают равновесные модели. Они описывают такие состояния экономики, когда результирующая всех сил, стремящихся вывести ее из данного состояния, равна нулю. В нерыночной экономике неравновесие по одним параметрам (например, дефицит) компенсируется другими факторами (черный рынок, очереди и т.п.). Равновесные модели дескриптивны, описательны. В нашей стране долгое время преобладал нормативный подход в моделировании, основанный на оптимизации. Оптимизация в теории рыночной экономики присутствует в основном на микроуровне (максимизация полезности потребителем или прибыли фирмой); на макроуровне результатом рационального выбора поведения экономическими субъектами оказывается некоторое состояние равновесия.

В моделях статических описывается состояние экономического объекта в конкретный момент или период времени; динамические модели включают взаимосвязи переменных во времени. В статических моделях обычно зафиксированы значения ряда величин, являющихся переменными в динамике, - например, капитальных ресурсов, цен и т.п. Динамическая модель не сводится к простой сумме ряда статических, а описывает силы и взаимодействия в экономике, определяющие ход процессов в ней. Динамические модели обычно используют аппарат дифференциальных и разностных уравнений, вариационного исчисления.

Детерминированные модели предполагают жесткие функциональные связи между переменными модели. Стохастические модели допускают наличие случайных воздействий на исследуемые показатели и используют инструментарий теории вероятностей и математической статистики для их описания.

Математическая экономика и эконометрика. Математическая экономика - раздел экономической науки, занимающийся анализом свойств и решений математических моделей явлений и процессов без использования реальных экономических данных. В некоторых случаях эти модели могут рассматриваться как часть математической теории на стыке с экономической наукой. Математическая экономика отделяется обычно от эконометрики, занимающейся статистической оценкой и анализом экономических зависимостей и моделей на основе изучения реальных эмпирических данных.

В математической экономике исследуются теоретические модели, основанные на определенных формальных предпосылках (линейность, выпуклость, монотонность и т.п. зависимости, конкретные формулы взаимосвязи величин). Математическая экономика, вообще говоря, не занимается изучением степени обоснованности того, что данная зависимость имеет тот или иной вид (например, что величина потребления является линейной возрастающей функцией дохода), - это оставляется для эконометрики. Задачей математической экономики является изучение вопроса о существовании решения модели, условиях его стационарности, не отрицательности или наличия других свойств. Это обычно осуществляется, как и в математике, путем дедуктивного получения следствий (теорем) из априорно сделанных предпосылок (аксиом).

Эконометрика предлагает совокупность методов анализа связей между различными экономическими показателями (факторами) на основании реальных статистических данных с использованием аппарата теории вероятностей и математической статистики. При помощи этих методов можно выявлять новые, ранее неизвестные связи, уточнять или отвергать гипотезы о существовании определенных связей между экономическими показателями, предлагаемые существующей экономической теорией.

Полем деятельности эконометрики является разработка различных методов *оценивания, подбора значений параметров* экономических моделей по имеющимся статистическим данным. Там же и методы использования оцененной модели для целей *экономического прогнозирования* и проведения рациональной экономической политики. Кроме того, встречаются ситуации, когда в распоряжении исследователя нет ясной экономической теории, описывающей поведение интересующих его отдельных экономических показателей и связи между различными показателями. Методы эконометрики дают возможность *подбора подходящей модели*, адекватной имеющимся реальным статистическим данным.

Разумеется, предметная область, методология и инструментарий экономической науки не исчерпываются подходами математической экономики и эконометрики - обычно в экономических исследованиях используются также методы качественного анализа, индуктивные, эвристические подходы, перемежающиеся с элементами математической экономики и эконометрики. Таким образом, математическая экономика выступает и как самостоятельный раздел экономической науки, и как один из ее инструментов. При этом разделы математической экономики, исследовавшиеся ранее в чисто теоретическом плане, все больше становятся теоретической базой и элементами прикладных исследований.

Традиционная эконометрическая методология. В последние годы подверглась критике доминировавшая в течение многих лет эконометрическая методология. Согласно этой методологии, полученные результаты считаются тем лучше, чем сильнее оказываются коррелированными величины, которые должны быть связаны между собой согласно существующей экономической теории. Чем точнее предсказания соответствуют наблюдаемым данным (где точность предсказаний, как правило, измеряется с помощью коэффициента детерминации), тем соответственно более значимыми оказываются полученные оценки с точки зрения *t-Стьдента* или *F-Фишера* статистик.

В литературе, написанной в то время, когда этот подход преобладал, отводят значительное место тому, как наиболее эффективным образом организовать перебор потенциальных объясняющих переменных, чтобы наилучшим образом предсказать объясняемую переменную, чтобы коэффициент детерминации был как можно большим, а *F*-статистика как можно более значимой. В некоторые статистические пакеты встроена такая автоматическая процедура перебора. Есть также программы, позволяющие быстро и удобно перебрать набор стандартных функциональных форм зависимостей между парой переменных (корни, логарифмы, обратные величины и т. д.).

Если получены настораживающие диагностики в критериях спецификации, таких как критерий Дарбина-Уотсона, то исследователь, следующий этой традиционной методологии вместо того, чтобы пересмотреть модель, данные и т. п., воспринимает это как сигнал к применению более продвинутых методов оценивания, которые бы позволили справиться с обнаруженными проблемами. В качестве примера можно привести корректирование регрессионной модели на наличие сериальной автокорреляции с помощью метода Кохрана-Оркатта. Сюда же, по-видимому, следует отнести применение гребневой регрессии в случае мультиколлинеарности объясняющих переменных.

Для описываемого подхода характерно стремление, во что бы то ни стало получить "наилучший" результат, вместо стремления получить результат осмысленный и надежный. Эвристическая ценность методов перебора с целью поиска наилучшей модели может быть высока. Однако очень важно понимать, что исследователя на этом пути подстерегает большая опасность. Поскольку исследователь стремится максимально улучшить точность подбора, согласие предсказаний модели и реальных данных, то в проводимой после получения оценок проверке значимости переменных должны применяться модифицированные критические границы. Если этого не делать, то номинальная значимость переменных, оказывается преувеличенной.

Другой важный момент состоит в том, что если используемые при оценивании гипотезы неверны, то полученные оценки должны рассматриваться как бессмысленные, и высокий уровень коэффициента детерминации не может придать им больше достоверности. Кроме того, следует помнить, что выбор модели с максимальным коэффициентом детерминации не имеет под собой достаточно прочного теоретического обоснования. Например, если зависимая переменная в двух регрессионных моделях имеет разную функциональную форму, то их нельзя сравнивать при помощи коэффициента детерминации.

Современный подход к эконометрическому моделированию. Следует отдавать предпочтение тем моделям, которые проходят диагностические критерии, хотя, может быть, и имеют низкий коэффициент детерминации, перед теми моделями, которые имеют высокий коэффициент детерминации, однако диагностические критерии говорят о нарушении основополагающих гипотез, необходимых для того, чтобы обосновать применяемые методы оценивания.

Важное правило современного эконометрического моделирования заключается во всесторонней проверке оцениваемой модели на предмет нарушения тех или иных допущений или предположений. Перечислим основные критерии, которые применяются в настоящее время для того, чтобы проверить, насколько правильно специфицирована регрессионная модель, прежде чем оценивать ее параметры по реальным статистическим данным соответствующего экономического объекта или явления.

- Критерий пропущенных переменных. Этот критерий является критерием добавления той переменной, которая, как подозревается, могла быть пропущена. Многие другие диагностические критерии также используют критерии добавления переменных.

- Критерии функциональной формы. Критерий Рамсея RESET (Regression specification error test).
- Критерии структурных изменений (критерий Чоу, CUSUM и CUSUMSQ) и критерии выбросов.
- Критерии автокорреляции остатков (критерий Дарбина-Уотсона, критерий Бреуши-Годфри (альтернативный критерий Дарбина), точечно-оптимальные критерии Кинга), Кохрана-Оркатта, Хилдрета-Лу.
- Критерии экзогенности регрессоров (критерий Дарбина-Ву-Хаусмана).
- Критерий первых разностей и другие критерии преобразования данных.
- Невложенные критерии (nonnested tests).
- Критерии стационарности переменных (тест Дики-Фуллера – ADF).
- Критерии гетероскедастичности ошибки (тест ранговой корреляции Спирмена, тест Парка, Глейзера, Голдфелда-Квандта).
- Критерии нормальности. Сюда же можно отнести показатели, которые не связаны с проверкой статистических гипотез и не выявляют неправильность спецификации модели, но указывают на недостоверность оценок и их потенциально плохую робастность:
- Показатели влиятельности наблюдений (DFFITS, DFBETAS).
- Показатели мультиколлинеарности (число обусловленности и т. п.).

Общий порядок использования процедур диагностических критериев состоит в следующем. Каждому из критериев соответствует статистика, которая является функцией данных. В предположении, что использованная вероятностная модель верна, можно теоретически вывести распределение данной статистики (зачастую распределение выводится из асимптотической теории и известно только приближенно).

Процедура проверки модели состоит в том, что если полученная на основе имеющегося набора данных статистика выходит за некоторый установленный заранее доверительный интервал, то нулевая гипотеза отклоняется. В случае отклонения нулевой гипотезы делается вывод, что принятые допущения неверны, т.е. модель специфицирована неверно.

Доверительный интервал задается, как правило, указанием критической границы. Вероятность того, что статистика выйдет за пределы доверительного интервала, заданного данной критической границей, и, тем самым, будет отклонена верная нулевая гипотеза, называют уровнем значимости. Понятно, что полученной на основе имеющегося набора данных статистике соответствует некоторый уровень значимости. Этот уровень значимости сам по себе можно рассматривать как статистику и использовать для проверки нулевой гипотезы. Чаще всего на практике используют 5%-ю границу. Если получен уровень значимости менее 5%, то нулевая гипотеза отклоняется и делается вывод о том, что модель специфицирована неверно.

2. ЭКОНОМИЧЕСКОЕ МОДЕЛИРОВАНИЕ И СТАТИСТИЧЕСКИЕ МЕТОДЫ

Изучение реальных явлений и процессов в экономике представляет собой исследование (проверки, обоснование, оценивание) количественных модельных закономерностей и качественных утверждений (гипотез) на основе анализа характеризующих их статистических данных, Применяемые при этом методы являются составной частью эконометрики - науки, изучающей экономические явления с количественной точки зрения. Эконометрика устанавливает и исследует количественные закономерности в социально-экономических явлениях и процессах формализованными методами теории вероятности и математической статистики. Поэтому в зависимости от характера предметной области их необходимо адаптировать к обработке экономических данных, которые отражают их неформальное, предметное содержание.

Закономерности в экономике выражаются в виде связей и зависимостей экономических показателей, которые не могут полностью отражать математические модели их поведения. Такие зависимости и модели должны быть верифицированы с помощью реальных статистических данных, с учетом реальных внутренних механизмов связи и случайных факторов. Модель, которая получена и апробирована на основе анализа статистических данных, может не соответствовать представлениям экономической теории. Такое поведение модели говорит о необходимости ее уточнения и развития.

Особенно важен эконометрический анализ в макро и микроэкономике, где взаимосвязи величин зачастую неочевидны и изменчивы. Нередко встречается ситуация, когда модель перестает "работать" в связи с появлением или активизацией какого-то фактора, и это способствует развитию макроэкономической теории. Поэтому предлагаемый материал "привязан" к макроэкономическим проблемам и моделям. Эконометрический анализ дает возможность обосновать и уточнить форму зависимостей в рассматриваемых макроэкономических моделях, лучше понять механизмы взаимосвязи макроэкономических показателей.

Основной задачей экономического исследования является анализ и построение *взаимосвязей экономических переменных.* Изучение таких взаимосвязей осложнено тем, что они не являются строгими, функциональными зависимостями. Во-первых, всегда очень трудно выявить все основные факторы, влияющие на данную переменную. Во-вторых, многие такие воздействия являются случайными, то есть содержат случайную составляющую. В-третьих, экономисты, как правило, располагают ограниченным набором данных статистических наблюдений, которые к тому же содержат различного рода ошибки.

Математическая статистика (то есть теория обработки и анализа данных) и ее применение в экономике - эконометрика - позволяют строить экономические модели и оценивать их параметры, проверять гипотезы о свойствах экономических показателей и формах их связи. В конечном счете, это служит основой для экономического анализа и прогнозирования, создавая возможность для принятия обоснованных экономических решений.

Любое экономическое исследование всегда предполагает объединение формальной теории (экономической модели) и реальной практики (статистических данных). Теоретические модели используют для описания и объяснения наблюдаемых социально-экономических явлений и процессов. Статистические данные собирают как для обоснования и уточнения существующих моделей, так и для эмпирического построения новых моделей, расширяющих и углубляющих экономическую теорию.

Введение случайного компонента в экономическую модель. Обычно предполагают, что все факторы, не учтенные явно в экономической модели, оказывают на объект некоторое результирующее воздействие, величина которого неизвестна заранее и может быть описана как случайная функция. Для ее описания в модель добавляют (обычно аддитивным образом) случайную величину ε. В качестве примера рассмотрим модель спроса

$$q = f(p,I) + \varepsilon,$$

где q - количество блага, p - цена, I - доход потребителя.

Переменная ε учитывает не только влияние всех прочих факторов (цен на другие товары, изменений конъюнктуры, погоды и т. д.), влияющих на потребление, но не учтенных явно в функции спроса. Существует еще целый ряд причин, в частности, неправильный выбор математической функции $f(p,I)$, включение в нее не всех существенных в данном случае факторов.

Поскольку при установлении таких взаимосвязей приходится иметь дело с выборочными данными, то нужно принимать во внимание ошибки выборки, связанные с характером собранных данных и рядом трудностей измерений микро- и макроэкономических показателей. Поэтому в экономике рассматривают не функциональные, а *статистические зависимости*. Нахождение, оценка и анализ таких зависимостей, представление модельных выражений и оценка их параметров составляют предмет эконометрического исследования.

Статистические данные и стохастическая модель.

Эконометрическая модель. Введение случайного компонента в экономическую модель приводит к тому, что взаимосвязь остальных ее переменных перестает быть строго детерминированной и становится стохастической, что и наблюдается в реальной действительности. Это отчасти делает модель доступной для эмпирической проверки на основе статистических данных о конкретном экономическом объекте. Если проверка показала адекватность модели, то иногда удается оценить параметры функционирования конкретного экономического объекта и сформулировать рекомендации для принятия практических решений. Работа с эконометрическими моделями требует использования инструментария оценивания и статистической проверки модели ("наука" моделирования), а также решения проблем выбора типа модели, набора объясняющих переменных и вида связей между ними ("искусство" моделирования).

Экономические данные: перекрестные данные (cross-section data) и временные ряды (time series). Статистические данные в эконометрике являются основой для выявления и обоснования эмпирических закономерностей. Без конкретных количественных данных, характеризующих реальное функционирование исследуемого экономического объекта, невозможно определить практическую значимость применяемой экономической модели. Это справедливо даже в том случае, когда целью является выявление преимущественно только качественных закономерностей.

Экономические данные обычно делят на два вида*: перекрестные или пространственные данные (cross-section data*) и *временные ряды (time series).*

Перекрестные или данные - это набор данных показателей экономических переменных, полученные для разных однотипных объектов (фирм, регионов) примерно в однотипных условиях. При этом либо все данные относятся к одному и тому же моменту времени, либо их времени я принадлежность несущественна. *Под пространственной выборкой* понимают серию из n независимых наблюдений системы ($p+1$)-ой независимых величин $(X_1, X_2, \ldots, X_p, Y)$.

Временные ряды - это упорядоченная выборка наблюдений, характеризующих один и тот же объект, но в различные моменты времени. Здесь важны не только сами наблюдаемые значения, но и порядок следования их друг за другом. Модели временных рядов оказываются обычно сложнее моделей пространственной выборки, так как наблюдения в этом случае нельзя считать независимыми друг от друга.

Данные временных рядов характеризуются определенными зависимостями и закономерностями их последовательных значений. Например, между собой могут быть связаны последовательные отклонения от общей тенденции развития; в этих связях экономических показателей могут присутствовать задержки (временные лаги) и т. д. Это обусловливает необходимость специальных методов их обработки и более сложного теоретического анализа по сравнению с данными перекрестных выборок.

Цели и методы сбора статистических данных. Целью сбора экономических данных является получение информационной базы для анализа и принятия решений в условиях появления конкретных проблем. Естественно, что анализ данных и принятие решений проводится на основе какой-либо интуитивной (неявной) или количественной (явной) экономической модели проблемной ситуации. Поэтому собирают именно те данные, которые необходимы для обоснования соответствующей модели.

Существуют различные методы сбора экономических данных: путем опроса, анкетирования и интервьюирования, получения официальной статистической отчетности и т.д. В большинстве стран существуют статистические органы, занимающиеся сбором, обработкой, распространением и публикацией важнейших данных.

Этой деятельностью занимаются также многие специализированные государственные и научные учреждения, а также коммерческие и частные агентства. В Российской Федерации официальными статистическими стандартами являются формы, методы сбора и обработки данных, устанавливаемые Государственным комитетом по статистике.

Подготовка статистических данных и использование их в модели. При подготовке статистических данных для работы с экономической моделью возникают две проблемы. Во-первых, могут отсутствовать необходимые для модели данные. Во-вторых (если все данные есть), нужно правильно отобрать их для конкретной модели так, чтобы они были согласованы и имели общую методическую базу оценки. При отсутствии нужных данных они нередко могут быть рассчитаны по имеющимся данным. Например, если отсутствуют данные о темпе инфляции (INF), но имеются данные о дефляторе валового внутреннего продукта (DEF), то инфляция (по ВВП) может быть рассчитана (в %):

$$INF = (DEF / DEF_{-1} - 1)\,100\%$$

(индекс -1 означает предшествующий год).

Если все нужные данные есть, то для модели необходимо их преобразовать в некоторый взаимно согласованный набор. Если это данные в денежном выражении, то это должны быть всюду одни и те же текущие, либо фиксированные (одного и того же года), денежные единицы. Реальным объемным показателям (то есть в фиксированных ценах) должны соответствовать реальные относительные показатели (например, процентные ставки нужно скорректировать на темп инфляции). В соответствии с решаемой задачей определяют и обобщающие показатели: валовой национальный продукт (ВНП, или GNP), валовой внутренний продукт (ВВП, или GDP), валовые внутренние или национальные сбережения, дефлятор ВНП или ВВП или ряд других. Например, если речь идет о внутреннем производстве и о влиянии на него внутренних инвестиций, то в качестве обобщающего показателя, на который влияют эти инвестиции, должен выступать ВВП, а не ВНП.

3. ОСОБЕННОСТИ ПОСТРОЕНИЯ ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ

3.1. Основные этапы эконометрического исследования

Построение эконометрической модели — центральная проблема любого эконометрического исследования, поскольку ее «качество» определяет достоверность и обоснованность результатов анализа тенденций развития, прогнозов рассматриваемых социально-экономических процессов, а также вытекающих из них выводов, в том числе и по вопросам разработки необходимых управленческих мероприятий.

В эконометрических исследованиях обычно предполагается, что закономерности моделируемого процесса складываются под влиянием ряда других явлений, факторов. Обобщенную форму эконометрической модели, описывающей закономерности развития такого процесса, обозначенного переменной y, в зависимости от уровней, воздействующих на него внешних явлений, факторов X_i, $i = 1, 2,..., p$, можно представить следующим уравнением:

$$y = f(b_0, b_1, \Lambda, b_p, X_1, X_2, \Lambda, X_p) + \varepsilon, \qquad (3.1)$$

выражающий вид и структуру взаимосвязей между уровнями переменных y, и X_i; ε, — случайная ошибка модели, в отношении свойств и характеристик которой, как это будет показано далее, выдвигается ряд дополнительных предположений.

Эконометрика допускает различные предположения относительно «статистического» содержания факторных переменных X_i , в то время как переменная y согласно (3.1) всегда рассматривается как случайная величина. Использование той или иной интерпретации значений переменных эконометрических моделей, как правило, не вносит принципиальные изменения в процедуры их построения, в методы оценки их параметров, но часто сказывается на свойствах полученных результатов.

Выражение (3.1) определяет лишь общий вид эконометрической модели. В конкретных эконометрических исследованиях могут использоваться специальные типы моделей, каждый из которых имеет свои характерные особенности. Эти типы обычно можно классифицировать на основе двух признаков. Во-первых, по виду факторов X_i , во-вторых, по свойствам ошибки модели.

В частности, в моделях регрессии классического типа обычно используются факторы, независимые между собой. Также предполагается, что ошибка модели имеет свойства «белого шума» — случайного процесса с нулевым математическим ожиданием, постоянной конечной дисперсией и нулевой корреляцией между ее разновременными значениями. Это означает, что в ряду ошибки ε отсутствуют автокорреляционные связи.

Модели могут различаться и по характеру связей факторов с переменной у. По этому признаку их разделяют на линейные и нелинейные модели. Эконометрические модели могут различаться и по свойствам своих параметров (модели с постоянной и переменной структурой), и по целому ряду других признаков.

Характерная особенность эконометрического исследования заключается в том, что зачастую априорно наиболее подходящий для рассматриваемого процесса тип модели определить не представляется возможным. Но при этом, как правило, на основе содержательного анализа рассматриваемого явления обычно удается выделить приемлемые альтернативные варианты модели и сформировать их исходные предпосылки. По результатам этапов эконометрического исследования эти варианты уточняются, и среди них выбирается тот, который в большей степени соответствует рассматриваемому процессу, явлению.

В общем случае процедуру построения эконометрической модели можно разделить на несколько взаимосвязанных между собой этапов. На каждом из них последовательно решаются задачи, которые имеют следующее содержание:

1. *Спецификация модели* - анализ специфических свойств рассматривае-

мых явлений и процессов (предметной области) и обоснование типа моделей, наиболее подходящих для их описания. Отметим, что в общем случае целями этого этапа являются:

1.1. Обоснование формы модели, выражаемой общим математическим уравнением (системой уравнений), связывающим включенные в модель переменные и содержащим неизвестные параметры (коэффициенты).

1.2. Выбор рационального состава включаемых в модель переменных и определение количественных характеристик, отражающих их уровни в прошлые периоды времени (на однородных объектах некоторой совокупности — территориях, предприятиях и т.п.).

2. *Параметризация модели* - оценка параметров выбранного варианта модели на основании исходных статистических данных, выражающих уровни показателей (переменных) на пространственной совокупности однородных объектов или в различные моменты времени.

3. *Верификация модели* - проверка качества построенной модели и обоснование вывода о целесообразности ее использования в ходе дальнейшего эконометрического исследования, а также для объяснения поведения изучаемых экономических показателей, прогнозирования и предсказания их поведения.

4. При выводе о нецелесообразности использования построенной эконометрической модели в дальнейших исследованиях следует вернуться к первому (или какому-либо другому этапу) и попытаться выбрать более качественную модификацию модели (другой вариант модели).

Выделенные этапы построения моделей достаточно условны и отражают циклический характер современных экономических исследований: от экономической теории к моделированию; от моделирования к совершенствованию теории и более глубокому пониманию сути происходящих процессов; от понимания сути к осуществлению продуманной и целенаправленной экономической политики. Состав используемых на них процедур, приемов и методов, их очередность зависят от типа разрабатываемой эконометрической модели, особенностей исследуемых экономических процессов, свойств исходных данных и т.п.

3.2. Обоснование формы эконометрической модели.

Основные подходы к решению проблем первого этапа исследования в значительной степени базируются на методах содержательного анализа закономерностей рассматриваемых процессов, подкрепляемых по мере необходимости методами общей и математической статистики. Дело в том, что обычно в практических исследованиях функциональный вид эконометрической модели может быть не известным. Часто рассматривают несколько альтернативных их вариантов, среди которых необходимо выбрать наиболее подходящий как с точки зрения требований экономической теории, так и необходимой точности аппроксимации функциональным выражением (3.1) исходного ряда измеренных значений зависимой переменной y.

В этой связи, прежде чем подойти к решению задач первого этапа, необходимо сформировать хотя бы предварительные исходные предпосылки экономического и математического содержания в отношении вида функциональной зависимости. Здесь можно отметить два возможных подхода. В одних случаях составом переменных X_i, $i = 1, 2,..., p$ и формой зависимости (1.1), отражают общепринятую экономическую концепцию. В других - выявляют эмпирические взаимосвязи между ними в ходе конкретных исследований статистических данных.

Примером первого подхода может служить двухфакторная модель Кобба-Дугласа

$$y = b_0 \cdot X_1^{b1} \cdot X_2^{b2} + \varepsilon, \tag{3.2}$$

которую применяют в макроэкономических исследованиях взаимосвязи между объемом полученного валового национального продукта (y) и используемыми производственными ресурсами (X_1 – основные фонды, X_2 – затраты живого труда). Для параметров этой модели известна содержательная экономическая интерпретация, их можно сопоставлять с имеющимися результатами аналогичных исследований.

Второй подход связан с привлечением исходных статистических данных для построения эконометрической модели. В общем случае ими являются известные наборы (массивы) значений зависимой переменной y и независимых факторов X_i. Различают два принципиально различных типа исходных информационных массивов - статический и динамический.

Статический массив представляет собой значения результирующей (зависимой, объясняемой и т.п.) переменной y и влияющих на нее факторов (независимых, объясняющих переменных) X_i, имевших место у объектов однородной совокупности в определенный период времени. Пример таких объектов - однотипные промышленные предприятия (заводы одной отраслевой направленности). В качестве y в практических исследованиях часто рассматривают показатели производительности труда, объемов выпускаемой продукции и некоторые другие. В качестве X_i - влияющие на уровень этих показателей факторы - объемы используемых фондов, численность и квалификация рабочей силы и т.п.

Приведем другой пример статической информации, характерной для социальных исследований. В качестве y часто рассматривают показатели заболеваемости (смертности) населения в регионах страны. Их уровень в каждом из регионов определяют значения независимых факторов, отражающих достигнутый материальный уровень жизни, климатические условия, состояние окружающей среды и т.п. В этом случае необходимая для построения эконометрической модели информация собирается по совокупности регионов страны за определенный фиксированный промежуток времени, например год.

В общем случае будем считать, что необходимая для построения эконометрической модели базового типа (1.1) статическая информация выражается следующими массивами взаимосвязанных (взаимосоответствующих) наборов данных: Такие наборы называют *пространственными данными (cross-sectional data)* или пространственной выборкой:

$$
\begin{array}{ccccccc}
y_1 & x_{11} & x_{21} & \Lambda & x_{i1} & \Lambda & x_{p1} \\
y_2 & x_{12} & x_{22} & \Lambda & x_{i2} & \Lambda & x_{p2} \\
\Lambda & \Lambda & \Lambda & \Lambda & \Lambda & \Lambda & \Lambda \quad \Lambda \\
y_j & x_{1j} & x_{2j} & \Lambda & x_{ij} & \Lambda & x_{pj} \\
\Lambda & \Lambda & \Lambda & \Lambda & \Lambda & \Lambda & \Lambda \quad \Lambda \\
y_n & x_{1n} & x_{2n} & \Lambda & x_{in} & \Lambda & x_{pn}
\end{array},
$$

где y_j — уровень зависимой переменной на j-ом объекте совокупности; x_{ij} — уровень i-го фактора на j-ом объекте статистической совокупности; $j = 1, 2, \ldots, n$; $i = 1, 2, \ldots, p$.

Динамическую информацию, которая связывает значения некоторой зависимой переменной y в моменты времени t, называют *временным (динамическим) рядом (time-series data).*

Исходная информация для построения эконометрических моделей может быть и смешанного типа. Если выборка выражает уровни интересующих нас переменных x_{it} по группе заводов за ряд лет, ее называют *панельными данными(panel data).*

В случае, когда периодов времени наблюдений T больше, чем число объектов p, панельные данные называют *объединенным временным рядом (pooled time-series).* Они связывают значения некоторой зависимой переменной y в моменты времени t со значениями независимых переменных (факторов) x_{it}, рассматриваемых в те же моменты времени, $t = l, 2, \ldots, T$.

В ходе содержательного анализа явление часто рассматривается на качественном уровне, и оперируют достаточно обобщенными понятиями, например, заболеваемость, уровень медицинского обслуживания, качество и уровень жизни, климат, качество рабочей силы и т.п. Часто эконометрическая модель строится именно для выражения закономерности, существующей между явлениями. При построении модели используется исходная информация, наборы показателей, которые уже выражают эти зависимости, их свойства, тенденции в виде количественных характеристик. Следовательно, модельное выражение должно представлять эти реально существующие взаимосвязи.

Для традиционных направлений исследований проблема обоснования состава показателей обычно считается решенной. Например, в макроэкономических исследованиях производительности труда, обычно рассматриваются регламентированные, уже устоявшиеся наборы показателей, значения которых публикуются в статистических сборниках, научных отчетах и других официальных издания Госкомстата. Такие как выработка на одного работающего как показатель, выражающий явление «производительность труда», объемы ВВП - показатель результативности экономики), объемы основных дов - показатель уровня материальной обеспеченности производственного процесса, экономики) и ряд других.

Вместе с тем в ряде областей эконометрических исследований такие системы показателей не могут быть сформированы столь однозначно. Часто одно и то же явление может быть выражено альтернативными вариантами показателей.

В отсутствие объективных данных в эконометрических исследованиях допускается замена одного показателя другим, косвенно отражающим то же явление. Например, среднедушевой доход как показатель материального уровня жизни может быть заменен среднегодовым товарооборотом на одного жителя региона. Неправильный выбор показателя, представляющего рассматриваемое явление в модели, может существенно повлиять на качество эконометрической модели.

3.3. Выбор совокупности факторов эконометрической модели

Проблема обоснования «оптимального» набора факторов обычно решается на основе как содержательного (теоретического), так и количественного (статистического) анализа тенденций рассматриваемых процессов.

На этапе содержательного анализа решается вопрос о целесообразности включения в модель тех или иных факторов, исходя из их экономического смысла. В макроэкономических исследованиях состав факторов, как правило, определяется на основании допущений экономической теории. Пример — двухфакторные производственные функции типа Кобба-Дугласа, которые строятся в предположении, что объем выпуска (производства) экономической системы в основном зависит от размеров используемых основных фондов и количества затраченного труда. Функция типа Кобба-Дугласа учитывает теоретическое предположение о постоянной эластичности выпуска по каждому из производственных факторов.

На этапе содержательного анализа обычно решается проблема установления самого факта наличия взаимосвязей между явлениями. Однако каждое из явлений может быть представлено разными наборами факторами и даже их комбинациями. Поэтому в ряде исследований на основании содержательного анализа однозначно состав независимых переменных модели определить практически невозможно. Могут существовать их альтернативные наборы.

Например, для исследования закономерностей динамики производительности труда на заводе могут быть отобраны следующие факторы: объем основных фондов, энерговооруженность труда, фондовооруженность труда, численность рабочей силы, ее квалификация. При этом квалификация как явление может выражаться разными показателями, например, средним уровнем образования работников, их усредненным квалификационным разрядом и т.п. Кроме того, можно ожидать, что показатели энерговооруженности, фондовооруженности труда, объема основных фондов характеризуют одно и то же явление — уровень материально-технической оснащенности производственного процесса. Таким образом, некоторые из рассматриваемых в таком исследовании показателей, выражающих количественные характеристики независимых переменных, относятся к сходным явлениям.

Факторы, выражающие одну и ту же причину, могут быть тесно взаимосвязаны между собой. Так, уровень розничного товарооборота в основном зависит от среднедушевого дохода; концентрация загрязняющих веществ — от объемов их выбросов. Вследствие этого одновременное включение таких факторов в модель вряд ли целесообразно, поскольку, таким образом, одна и та же причина будет учтена дважды. В результате в общем случае на этапе обоснования эконометрической модели решается задача выбора наиболее предпочтительного состава независимых факторов среди ряда альтернативных вариантов.

Можно выделить два основных подхода к решению этой проблемы. Первый предполагает априорное (до построения модели) исследование характера и силы взаимосвязей между рассматриваемыми переменными, по результатам которого в модель включаются факторы, наиболее значимые по своему «непосредственному» влиянию на зависимую переменную y. И, наоборот, из модели исключаются факторы, которые, либо малозначимы с точки зрения силы своего влияния на эту переменную, либо их сильное влияние на нее обусловлено индуцированными взаимосвязями с другими переменными.

В основе «априорного» подхода лежат следующие предположения.

1. Сильное влияние фактора на зависимую переменную должно подтверждаться определенными количественными характеристиками, важнейшей является их парный линейный коэффициент корреляции, выборочное значение которого рассчитывается на основании имеющейся информации.

Логика использования коэффициента парной корреляции при отборе значимых факторов на практике состоит в следующем. Если значение коэффициента корреляции достаточно велико, т.е. превосходит некоторый эмпирический рубеж (на практике 0,5-0,6), то можно говорить о наличии существенной линейной связи между переменными y и X_i, или о достаточно сильном влиянии X_i на y. Чем больше абсолютное значение r_{yxi}, тем сильнее это влияние (положительное или отрицательное, в зависимости от знака коэффициента парной корреляции).

2. Если два и более факторов выражают одно и то же явление, то, как правило, между ними также должна существовать достаточно сильная взаимосвязь. На это может указать выборочное значение их парного коэффициента корреляции. На практике взаимосвязь между факторами признается существенной, если их коэффициент корреляции достигает величины 0,8-0,9. В таких ситуациях один из этих факторов целесообразно исключить из модели, чтобы одна и та же причина не учитывалась дважды. Однако такое исключение следует проводить только в тех случаях, когда факторы выражают одно и то же явление.

Приведенные рубежные значения (в первом случае — 0,5-0,6; во втором — 0,8-0,9) достаточно условны. В каждом конкретном случае они устанавливаются индивидуально. Значительно усложняет проблему отбора факторов явление *ложной корреляции*, которое характеризуется достаточно высокими по абсолютной величине значениями коэффициентов парной корреляции с содержательной точки зрения между собой никак не связанных факторов. Иными словами, большие значения парных коэффициентов корреляции могут иметь место и в тех случаях, когда тенденции рассматриваемых процессов совпали случайно, при отсутствии между ними взаимосвязи, обоснованной представлениями соответствующей экономической теории.

Ложная корреляция может помешать при построении «правильной» модели по двум причинам. Во-первых, в модель случайно могут быть введены незначимые с содержательной точки зрения факторы, характеризующиеся значимыми величинами коэффициента парной корреляции. Во-вторых, из модели могут быть исключены значимые с точки зрения влияния на y факторы, в отношении которых ошибочно признана гипотеза о том, что они выражают то же явление, что и другой фактор (факторы), уже включенный в эту модель.

Среди основных причин включения в модель переменных с ложной корреляцией часто называют ненадежность информации, используемой при определении значений факторов в различные моменты времени, трудности формализации факторов, имеющих качественный характер, неустойчивость тенденций изменения рассматриваемых переменных, неправильную форму взаимосвязи между ними и т.п. Основной путь, придерживаясь которого можно избежать ошибок, связанных с понятием «ложной корреляции», связан с проведением качественного анализа проблемы, направленного на обоснование адекватного ей содержания и формы модели.

Второй подход к отбору независимых факторов — можно назвать апостериорным — предполагает первоначально включить в модель все отобранные на этапе содержательного анализа факторы. Уточнение их состава в этом случае производится на основе анализа характеристик качества построенной модели и силы влияния каждого из факторов на зависимую переменную.

Если фактор X_i признается незначимым, его целесообразно удалить из модели. Эта операция приводит к уменьшению общего количества независимых переменных в модели. Таким образом, на практике используют следующую поэтапную процедуру построения окончательного варианта модели на основе апостериорного подхода:

1. В исходный вариант модели включаются все факторы, отобранные в ходе содержательного анализа проблемы. Рассчитывают значения оценок коэффициентов модели, их среднеквадратические ошибки и значения критериев Стьюдента.

2. Из модели удаляют незначимый фактор, характеризующийся наименьшим значением критерия Стьюдента, при условии, что он статистически незначим и формируют новый вариант модели с уменьшенным на один числом факторов.

Заметим, что в модели может быть несколько незначимых факторов. Однако все их одновременно удалять не следует. Возможно, что недостаточная значимость большинства факторов обусловлена влиянием «наихудшего» из незначимых факторов и на следующем шаге расчетов они окажутся значимыми.

3. Процесс отбора факторов считают законченным, когда остающиеся в модели факторы являются значимыми, если полученный вариант модели удовлетворяет и другим критериям ее качества, то процесс построения модели можно считать завершенным в целом.

В противном случае попытаются сформировать другой альтернативный вариант модели, отличающийся от предыдущего либо составом факторов, либо формой их взаимосвязи с зависимой переменной.

Каждый из этих подходов имеет свои преимущества и недостатки. «Априорный» путь отбора факторов не обладает достаточной обоснованностью. Он в большей степени использует «прямые» количественные индикаторы «силы» взаимосвязей между рассматриваемыми величинами и не принимает во внимание в полной мере особенности комплексного влияния независимых факторов на переменную *y* т.е. своеобразные эффекты «эмерджентности» такого влияния.

Этот эффект выражается в том, что совокупное воздействие нескольких факторов на переменную *y*, может значительно отличаться от суммы воздействий каждого из них именно в силу наличия внутренних взаимосвязей между независимыми переменными. Вместе с тем использование априорного подхода часто позволяет уточнить некоторые предварительные альтернативные варианты наборов независимых факторов, проверить исходные предпосылки модели относительно правильности выбора формы взаимосвязей между ними.

«Апостериорный» подход к отбору факторов, на первый взгляд, предпочтительнее из-за того, что целесообразность включения в модель каждого из факторов определяется на основании всего комплекса взаимосвязей между переменными. Однако когда общее количество факторов достаточно велико, нет никаких гарантий того, что множество несущественных, а то и ложных взаимосвязей между ними не будет превалировать над основными связями. В результате может оказаться, что в числе первых кандидатов на исключение будут «названы» наиболее важные, значимые с точки зрения влияния на переменную *y*, факторы. Поэтому в сложных случаях, т.е. при наличии большого числа отобранных для включения в модель на этапе содержательного анализа факторов, полезно сочетать при обосновании их «оптимального» состава оба подхода, как априорный, так и апостериорный.

4. МНОЖЕСТВЕННАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ

Множественный регрессионный анализ является расширением парного регрессионного анализа. Он применяется в тех случаях, когда поведение объясняемой, зависимой переменной необходимо связать с влиянием более чем одной факторной, независимой переменной. Хотя определенная часть многофакторного анализа представляет собой непосредственное обобщение понятий парной регрессионной модели, при выполнении его может возникнуть ряд принципиально новых задач.

Так, при оценке влияния каждой независимой переменной необходимо уметь разграничивать ее воздействие на объясняемую переменную от воздействия других независимых переменных. При этом множественный корреляционный анализ сводится к анализу парных, частных корреляций. На практике обычно ограничиваются определением их обобщенных числовых характеристик, таких как частные коэффициенты эластичности, частные коэффициенты корреляции, стандартизованные коэффициенты множественной регрессии.

Затем решаются задачи спецификации регрессионной модели, одна из которых состоит в определении объема и состава совокупности независимых переменных, которые могут оказывать влияние на объясняемую переменную. Хотя это часто делается из априорных соображений или на основании соответствующей экономической (качественной) теории, некоторые переменные могут в силу индивидуальных особенностей изучаемых объектов не подходить для модели. В качестве наиболее характерных из них можно назвать *мультиколлинеарность* или *автокоррелированность* факторных переменных.

4.1. Оценка параметров модельной регрессии с помощью метода(МНК)

В данном разделе полагается, что рассматривается модель регрессии, которая специфицирована правильно. Обратное утверждение, если исходные предположения оказались неверными, можно установить только на основании качества полученной модели. Следовательно, этот этап является исходным для проведения множественного регрессионного анализа даже в самом сложном случае, поскольку только он, а точнее его результаты могут дать основания для дальнейшего уточнения модельных представлений. В таком случае выполняются необходимые изменения и дополнения в спецификации модели, и анализ повторяется после уточнения модели до тех пор, пока не будут получены удовлетворительные результаты.

На любой экономический показатель в реальных условиях обычно оказывает влияние не один, а несколько и не всегда независимых факторов. Например, спрос на некоторый вид товара определяется не только ценой данного товара, но и ценами на замещающие и дополняющие товары, доходом потребителей и многими другими факторами. В этом случае вместо парной регрессии $M(Y/_{X=x}) = f(x)$ рассматривается множественная регрессия

$$M(Y/_{X1=x1,\, X2=x2,\, \dots,\, Xp=Xp}) = f(x_1, x_2, \dots, x_p) \qquad (4.1)$$

Задача оценки статистической взаимосвязи переменных Y и X_1, X_2, ..., X_P формулируется аналогично случаю парной регрессии. Уравнение множественной регрессии может быть представлено в виде

$$\boldsymbol{Y} = f(\boldsymbol{B}, \boldsymbol{X}) + \boldsymbol{\varepsilon}, \qquad (4.2)$$

где $\boldsymbol{X}$ — вектор независимых (объясняющих) переменных; $\boldsymbol{B}$ — вектор параметров уравнения (подлежащих определению); $\boldsymbol{\varepsilon}$ - случайная ошибка (отклонение); $\boldsymbol{Y}$ — зависимая (объясняемая) переменная.

Предполагается, что для данной генеральной совокупности именно функция f связывает исследуемую переменную Y с вектором независимых переменных $\boldsymbol{X}$.

Рассмотрим наиболее обоснованную теорией вероятностей и математической статистики и самую простую для статистического анализа и экономической интерпретации модель множественной линейной регрессии. Для этого имеются, по крайней мере, две существенные причины.

Во-первых, уравнение регрессии является линейным, если система случайных величин ($X_1, X_2, ..., X_P, Y$) имеет совместный нормальный закон распределения. Предположение о нормальном распределении может быть в ряде случаев обосновано с помощью предельных теорем теории вероятностей. Часто такое предположение принимается в качестве гипотезы, когда при последующем анализе и интерпретации его результатов не возникает явных противоречий.

Вторая причина, по которой линейная регрессионная модель предпочтительней других, состоит в том, что при использовании ее для прогноза риск значительной ошибки оказывается минимальным.

Теоретическое линейное уравнение регрессии имеет вид:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + ... + b_p X_p + \varepsilon, \tag{4.3}$$

или для индивидуальных наблюдений с номером i:

$$y_i = b_0 + b_1 x_{i1} + b_2 X_{i2} + ... + b_p X_{ip} + \varepsilon_i \tag{4.4}$$

где $i = 1, 2, ..., n$.

Здесь $\boldsymbol{B} = (b_0, b_1, \quad, b_P)$ — вектор размерности (p+1) неизвестных параметров b_j, $j = 0, 1, 2, ..., p$, называется j-ым теоретическим коэффициентом регрессии (частичным коэффициентом регрессии). Он характеризует чувствительность величины Y к изменению X_j. Другими словами, он отражает влияние на условное математическое ожидание $M(Y/_{X1 = x1, X2 = x2, ..., Xp = xp})$ зависимой переменной Y объясняющей переменной X_j при условии, что все другие объясняющие переменные модели остаются постоянными. b_0 — свободный член, определяющий значение Y в случае, когда все объясняющие переменные X_j равны нулю.

После выбора линейной функции в качестве модели зависимости необходимо оценить параметры регрессии.

Пусть имеется n наблюдений вектора объясняющих переменных $\boldsymbol{X} = (1, X_1, X_2, ..., X_P)$ и зависимой переменной Y:

$$(1, x_{i1}, x_{i2}, ..., x_{ip}, y_i), \ i = 1, 2, ..., n.$$

Для того чтобы однозначно можно было бы решить задачу отыскания параметров $b_0, b_1, \dots, b_P$ (т.е. найти некоторый наилучший вектор $\boldsymbol{B}$), должно выполняться неравенство $n > p + 1$. Если это неравенство не будет выполняться, то существует бесконечно много различных векторов параметров, при которых линейная формула связи между $\boldsymbol{X}$ и Y будет абсолютно точно соответствовать имеющимся наблюдениям. При этом, если $n = p + 1$, то оценки коэффициентов вектора $\boldsymbol{B}$ рассчитываются единственным образом — путем решения системы $p + 1$ линейного уравнения:

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + ... + b_p x_{ip}, \tag{4.5}$$

где $i = 1, 2, ..., n$.

Например, для однозначного определения оценок параметров уравнения регрессии $Y = b_0 + b_1 X_1 + b_2 X_2$ достаточно иметь выборку из трех наблюдений $(1, x_{i1}, x_{i2}, y_i)$, $i = 1, 2, 3$. В этом случае найденные значения параметров b_0, b_1, b_2 определяют такую плоскость $Y = b_0 + b_1 X_1 + b_2 X_2$ в трехмерном пространстве, которая пройдет именно через имеющиеся три точки.

С другой стороны, добавление в выборку к имеющимся трем наблюдениям еще одного приведет к тому, что четвертая точка $(x_{41}, x_{42}, x_{43}, y_4)$ практически всегда будет лежать вне построенной плоскости (и, возможно, достаточно далеко). Это потребует определенной переоценки параметров.

Таким образом, вполне логичен следующий вывод: если число наблюдений больше минимально необходимой величины, т.е. $n > p + 1$, то уже нельзя подобрать линейную форму, в точности удовлетворяющую всем наблюдениям. Поэтому возникает необходимость оптимизации, т.е. оценивания параметров $b_0, b_1, \ldots, b_P$, при которых формула регрессии дает наилучшее приближение одновременно для всех имеющихся наблюдений.

В данном случае число $\nu = n - p - 1$ называется числом степеней свободы. Нетрудно заметить, что если число степеней свободы невелико, то статистическая надежность оцениваемой формулы невысока. Например, вероятность надежного вывода (получения наиболее реалистичных оценок) по трем наблюдениям существенно ниже, чем по тридцати. Считается, что при оценивании множественной линейной регрессии для обеспечения статистической надежности требуется, чтобы число наблюдений превосходило число оцениваемых параметров, по крайней мере, в 3 раза.

Прежде чем перейти к описанию алгоритма оценивания коэффициентов регрессии, отметим желательность выполнимости ряда теоретических предпосылок МНК, которые позволят обосновать характерные особенности регрессионного анализа в рамках классической линейной многофакторной модели.

4.2. Теоретические предпосылки метода МНК.

1°. *Математическое ожидание случайного отклонения ε_i равно нулю для всех наблюдений:*

$$M(\varepsilon_i) = 0, \quad i = 1, 2, \ldots, n.$$

2°. *Наличие гомоскедастичности* (постоянство дисперсии случайных отклонений). Дисперсия случайных отклонений ε_i должна быть постоянной: $D(\varepsilon_i) = D(\varepsilon_j) = \sigma^2$ для любых наблюдений с номером i и j.

3°. *Отсутствие автокорреляции.* Случайные отклонения ε_I и ε_J не должны зависеть друг от друга для всех $i \neq j$.

$$Cov(\varepsilon_i, \varepsilon_j) = \begin{cases} 0, & если\ i \neq j, \\ \sigma^2, & если\ i = j. \end{cases}$$

4°. Случайное отклонение должно быть независимым от объясняющих переменных:

$$Cov(\varepsilon_i, X_{ij}) = 0.$$

5°. Модель эмпирической регрессии должна являться линейной относительно параметров. Это ограничение не распространяется на факторные переменные.

6°. *Отсутствие мультиколлинеарности.* Между объясняющими переменными должна отсутствовать строгая (сильная) линейная зависимость.

7°. *Случайные величины -ошибки* ε_i, $i = 1, 2, ..., n$, должны иметь нормальный закон распределения ($\varepsilon_i \sim N(0, \sigma_\varepsilon)$).

Выполнимость данной предпосылки важна для проверки статистических гипотез и построения интервальных оценок.

Как и в случае парной регрессии, истинные значения параметров b_j с помощью случайной выборки получить невозможно. В этом случае вместо теоретического уравнения регрессии (4.3) оценивается так называемое *эмпирическое уравнение регрессии.* Эмпирическое уравнение регрессии представим в виде:

$$Y = \hat{b}_0 + \hat{b}_1 X_1 + \hat{b}_2 X_2 + ... + \hat{b}_p X_p + e\,. \tag{4.6}$$

Здесь $\hat{b}_0, \hat{b}_1, ..., \hat{b}_p$ - оценки теоретических значений $b_0, b_1, ..., b_p$ коэффициентов регрессии *(эмпирические коэффициенты регрессий); e* - эмпирическая оценка неизвестного случайного отклонения ε. Для отдельных наблюдений имеем:

$$y_i = \hat{b}_0 + \hat{b}_1 x_{i1} + \hat{b}_2 x_{i2} + ... + \hat{b}_p x_{ip} + e_i \tag{4.7}$$

Оцененное уравнение в первую очередь должно описывать общую закономерную тенденцию изменения зависимой переменной Y. При этом необходимо иметь возможность оценить случайные отклонения измеренных значений y_i от таких неслучайных расчетных значений.

По данным выборки объема n: $(1, x_{i1}, x_{i2}, ..., x_{ip}, y_i)$, $i = 1, 2, ..., n$, требуется оценить значения параметров b_j вектора $\boldsymbol{B}$, т.е. провести параметризацию выбранной модели (здесь x_{ij}, $j = 0, 1, 2, ..., p$ значение переменной Xj в i-ом наблюдении).

При выполнении перечисленных выше предпосылок МНК относительно ошибок ε_i оценки $\hat{b}_0, \hat{b}_1, ..., \hat{b}_p$ коэффициентов $b_0, b_1, ..., b_p$ множественной линейной регрессии с помощью МНК являются несмещенными, эффективными и состоятельными (т.е. BLUE-оценками).

На основании (4.7) отклонение e_i значения зависимой переменной Y от модельного значения $\hat{y}_i$, соответствующего уравнению регрессии в i-ом наблюдении ($i = 1, 2, ..., n$), рассчитывается по формуле:

$$e_i = y_i - \hat{y}_i\,, \quad \hat{y}_i = \hat{b}_0 + \hat{b}_1 x_{i1} + \hat{b}_2 x_{i2} + ... + \hat{b}_p x_{ip}\,. \tag{4.8}$$

Для оценки параметров уравнения множественной линейной регрессии наиболее распространенным методом является метод наименьших квадратов (МНК). Этот метод использует минимизацию суммы квадратов отклонений наблюдаемых значений зависимой переменной Y от ее расчетных величин $\hat{Y}$, получаемых с помощью модельного уравнения регрессии:

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \to min.$$

По МНК для нахождения оценок минимизируется следующая функция, квадратичная относительно коэффициентов регрессии $b_0, b_1, ..., b_p$:

$$S(\hat{b}_0, \hat{b}_1, ..., \hat{b}_p) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}\left(y_i - \left(\hat{b}_0 + \sum_{j=1}^{p}\hat{b}_j x_{ij}\right)\right)^2. \qquad (4.9)$$

Данная функция является квадратичной относительно неизвестных величин b_j, $j = 0, 1, ..., p$. Она ограничена снизу, следовательно, имеет минимум. Необходимым условием минимума функции $S(b_0, b_1, ..., b_p)$ является равенство нулю всех ее частных производных по b_j. Частные производные квадратичной функции (2.9) являются линейными функциям относительно искомых оценок коэффициентов регрессии:

$$\frac{\partial S(\hat{b}_0, \hat{b}_1, ..., \hat{b}_p)}{\partial b_0} = 2\sum_{i=1}^{n}\left(y_i - \left(\hat{b}_0 + \sum_{j=1}^{p}\hat{b}_j x_{ij}\right)\right)(-1),$$

$$\frac{\partial S(\hat{b}_0, \hat{b}_1, ..., \hat{b}_p)}{\partial b_j} = 2\sum_{i=1}^{n}\left(y_i - \left(\hat{b}_0 + \sum_{j=1}^{p}\hat{b}_j x_{ij}\right)\right)(-x_{ij}), \qquad (4.10)$$

где $j = 1, 2, ..., p$.

Приравнивая их к нулю, получаем нормальную систему $p + 1$ линейных уравнений с $p + 1$ неизвестными оценками коэффициентов регрессии, что является одним из достоинств метода МНК. Такая система имеет обычно единственное решение:

$$\sum_{i=1}^{n}\left(y_i - \left(\hat{b}_0 + \sum_{j=1}^{p}\hat{b}_j x_{ij}\right)\right) = 0,$$

$$\sum_{i=1}^{n}\left(y_i - \left(\hat{b}_0 + \sum_{j=1}^{p}\hat{b}_j x_{ij}\right)\right)(x_{ij}) = 0, \qquad (4.11)$$

где $j = 1, 2, ..., p$.

В исключительных случаях, когда столбцы системы линейных уравнений линейно зависимы, она имеет бесконечно много решений или не имеет решения вовсе. Однако данные реальных статистических наблюдений к таким исключительным случаям практически никогда не приводят.

Система линейных уравнений относительно неизвестных оценок параметров линейной модели имеет следующий вид:

$$\begin{cases}
\hat{b}_0 \cdot n + \hat{b}_1 \sum_{i=1}^{n} x_{i1} + \hat{b}_2 \sum_{i=1}^{n} x_{i2} + \ldots + \hat{b}_p \sum_{i=1}^{n} x_{ip} = \sum_{i=1}^{n} y_i \\
\hat{b}_0 \sum_{i=1}^{n} x_{i1} + \hat{b}_1 \sum_{i=1}^{n} x_{i1}^2 + \hat{b}_2 \sum_{i=1}^{n} x_{i2}x_{i1} + \ldots + \hat{b}_p \sum_{i=1}^{n} x_{ip}x_{i1} = \sum_{i=1}^{n} y_i x_{i1} \\
\hat{b}_0 \sum_{i-1}^{n} x_{i2} + \hat{b}_1 \sum x_{i2}x_{i1} + \hat{b}_2 \sum_{i=1}^{n} x_{12}^2 + \ldots + \hat{b}_p \sum_{i=1}^{p} x_{ip}x_{i2} = \sum_{i=1}^{n} y_i x_{i2} \\
\ldots \\
\hat{b}_0 \sum_{i=1}^{n} x_{ip} + \hat{b}_1 \sum_{i=1}^{n} x_{ip}x_{i1} + \hat{b}_2 \sum_{i=1}^{n} x_{ip}x_{i2} + \ldots + \hat{b}_p \sum_{i-1}^{n} x_{ip}^2 = \sum_{i=1}^{n} y_i x_{ip}\ .
\end{cases}$$

После деления всех уравнений системы на объем выборки n все суммарные величины преобразуются в соответствующие средние величины:

$$\begin{cases}
\hat{b}_0 + \hat{b}_1 \overline{X}_1 + \hat{b}_2 \overline{X}_2 + \ldots + \hat{b}_p \overline{X}_p = \overline{Y} \\
\hat{b}_0 \overline{X}_1 + \hat{b}_1 \overline{X_1^2} + \hat{b}_2 \overline{X_2 X_1} + \ldots + \hat{b}_p \overline{X_p X_1} = \overline{YX}_1 \\
\hat{b}_0 \overline{X}_2 + \hat{b}_2 \overline{X_2 X_1} + \hat{b}_2 \overline{X_2^2} + \ldots + \hat{b}_p \overline{X_p X_2} = \overline{YX}_2 \\
\ldots \\
\hat{b}_0 \overline{X}_p + \hat{b}_1 \overline{X_p X_1} + \hat{b}_2 \overline{X_p X_2} + \ldots + \hat{b}_p \overline{X_p^2} = \overline{YX}_p .
\end{cases} \tag{4.12}$$

Из первого уравнения можно определить величину коэффициента регрессии $\hat{b}_0$:

$$\hat{b}_0 = \overline{Y} - \hat{b}_1 \overline{X}_1 - \hat{b}_2 \overline{X}_2 - \Lambda - \hat{b}_p \overline{X}_p\ .$$

Подставляя его в уравнение (2.8), получим следующую форму записи эмпирического линейного уравнения множественной регрессии:

$$\hat{y}_i = \overline{Y} + \hat{b}_1(x_{i1} - \overline{X}_1) + \hat{b}_2(x_{i2} - \overline{X}_2) + \ldots + \hat{b}_p(x_{ip} - \overline{X}_p). \tag{4.13}$$

Нормальную систему линейных уравнений МНК (2.11) наиболее наглядно можно представить с помощью векторно-матричной формы записи.

4.3. Оценивание коэффициентов множественной линейной регрессии.

Представим данные наблюдений и соответствующие коэффициенты в матричной форме.

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1p} \\ 1 & x_{21} & x_{22} & \ldots & x_{2p} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{np} \end{pmatrix}, \quad B = \begin{pmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \ldots \\ \hat{b}_p \end{pmatrix}, \quad e = \begin{pmatrix} e_1 \\ e_2 \\ \ldots \\ e_n \end{pmatrix}.$$

Здесь $\boldsymbol{Y}$ — n - мерный вектор-столбец наблюдений зависимой переменной $\boldsymbol{Y}$; $\boldsymbol{X}$ — матрица размерности n х $(p + 1)$, в которой i - я строка ($i = 1, 2, \ldots, n$)

представляет наблюдение вектора значений независимых переменных X_1, X_2, ..., X_P; единица соответствует переменной при свободном члене b_o, $\boldsymbol{B}$ — вектор-столбец размерности $(p+1)$ параметров уравнения регрессии (2.8); $\boldsymbol{e}$ — вектор-столбец отклонений выборочных (реальных) значений y_i зависимой переменной Y (2.7) от значений $\hat{y}_i$ размерности n, получаемых из модельного уравнения регрессии:

$$\hat{y}_i = X \cdot \hat{B} = \hat{b}_0 + \hat{b}_1 x_{i1} + \hat{b}_2 x_{i2} + ... + \hat{b}_p x_{ip} . \quad (4.14)$$

Сумма квадратов отклонений МНК в матричном виде запишется следующим образом:

$$e^T \cdot e \to min \quad или \quad S(\hat{B}) = (Y - X \cdot \hat{B})^T (Y - X \cdot \hat{B}) \to min.$$

Условие экстремума:

$$S'\left(\hat{B}\right) = 0 . \quad (4.15)$$

Частные производные по параметрам в матричной форме вычисляются следующим образом:

$$S'_B(\hat{B}) = (Y - X \cdot \hat{B})^T (Y - X \cdot \hat{B}) = (Y^T - \hat{B}^T X^T)(Y - X\hat{B}) =$$

$$= Y^T Y - Y^T X \hat{B} - \hat{B}^T X^T Y + \hat{B}^T X^T X \hat{B} =$$

$$= Y^T Y - 2 \cdot \hat{B}^T X^T Y + \hat{B}^T X^T X \hat{B}.$$

$$S'_B(\hat{B}) = -2 \cdot X^T Y + 2 \cdot X^T X \hat{B} = 0 \Rightarrow (X^T X)\hat{B} = (X^T Y)$$

Таким образом, получим в матричной форме оценки параметров линейной регрессии:

$$\hat{B} = (X^T X)^{-1}(X^T Y). \quad (4.16)$$

Для вычисления их не нужно составлять и решать нормальную систему линейных уравнений МНК. Достаточно выполнить указанные в формуле (2.16) алгебраические операции в матричной форме над результатами исходных, выборочных наблюдений $\boldsymbol{X}$ и $\boldsymbol{Y}$. В частности, при выполнении контрольной работы такие расчеты достаточно легко выполняются с помощью Мастера функций табличного процессора Microsoft Exel.

4.4. Интерпретация параметров и уравнения линейной регрессии.

Интерпретация – содержательное объяснение в терминах предметной области – результатов анализа экономического явления или объекта, представленного статистическими (выборочными данными), является одной их самых важных задач регрессионного анализа. Так, рассматривая полученные оценки параметров уравнения регрессии, можно сказать, изменение фактора X_i на одну единицу своего измерения ведет к изменению объясняемой переменной Y на $\hat{b}_i$ единиц измерения этой переменной. Направление ее изменения определяется знаком коэффициента перед фактором X_i.

При этом единицы, в которых измерены выборочные значения переменных Y и X_i, влияют на величину оценок параметров регрессии $\hat{b}_i$. Нужно обязательно фиксировать, в каких единицах измерены значения всех переменных, прежде чем заменять слово «единица» конкретными названиями: тонны, рубли и т.п. Отсюда следует, что коэффициенты регрессии перед различными факторами нельзя сравнивать друг с другом.

Все другие более общие показатели характера влияния факторов на объясняемую переменную, не зависящие от масштаба их измерения, такие как стандартизованные коэффициенты регрессии и коэффициенты эластичности, получают на основе этих оценок параметров $\hat{b}_i$.

Параметр $\hat{b}_0$ представляет оценку значения объясняемой переменной Y при нулевых значениях факторных переменных. Она может иметь или не иметь экономический смысл в зависимости от характера конкретной ситуации.

При интерпретации модельного уравнения регрессии важно отмечать его следующие характерные особенности. Во-первых, $\hat{b}_0$ и $\hat{b}_i$ являются только оценками неизвестных констант b_0 и b_i истинной, теоретической регрессии, которая к тому же не обязательно является линейной. Кроме того, величина и качество статистических оценок зависят от правильности выбора самого метода оценивания. В частности предпосылки МНК обосновывают условия, в которых можно получить «лучшие» оценки параметров модельной регрессии.

Во-вторых, эмпирическое уравнение регрессии отражает общую закономерную тенденцию, представленную выборочными данными. Но при этом каждое отдельное наблюдение подвержено ошибкам измерения и случайным воздействиям со стороны неконтролируемых факторов. Следовательно, расчетные значения объясняемой переменной не могут быть детерминированными и их нужно дополнять характеристиками вариации, такими как стандартные ошибки или доверительные интервалы параметров регрессии.

И, наконец, в-третьих, правильность интерпретации зависит от правильного полноты совокупности факторов и выбора модельного представления статистической связи. Это связано с включением в уравнение всех статистически значимых объясняющих переменных, а также выбором формы эмпирической функции регрессии. Если форма уравнения является линейной, то можно использовать такие характеристики линейной статистической связи, как коэффициенты парной, частной и множественной корреляции. Но в случае, когда реальная функция регрессии нелинейная, они не могут адекватно отражать силы влияния факторов.

5. АНАЛИЗ КАЧЕСТВА УРАВНЕНИЯ РЕГРЕССИИ

Выявление лучшего варианта эконометрической модели обычно осуществляется сравнением соответствующих им качественных характеристик, которые можно рассчитать на основе исходной статистической информации, содержащейся в векторе *Y*, матрице *X*, и новой расчетной информации, появляющейся после построении каждого из вариантов модели. Основным условием высокого «качества» модели является обоснованность математической формы уравнения эмпирической регрессии. Важную роль при этом играет как состав включенных в него независимых переменных, так и характер их взаимосвязей с зависимой переменной y, которые в совокупности определяют причины ее изменчивости.

Сопоставление новой расчетной информации, полученной после оценки параметров модельной регрессии с исходной статистической информацией, позволяет установить, насколько удалось реализовать это условие на практике.

5.1. Характеристики и критерии качества эконометрических моделей.

Ведущая роль при определении характеристик качества эконометрической модели принадлежит ряду ее «выборочной» ошибки e_i, $i = 1, 2, \dots, n$, которая формируется с использованием найденных оценок ее параметров как

$$e_i = y_i - \hat{y}_i,$$

где $\hat{y}_i$ — расчетное значение переменной y_i при известных значениях независимых переменных X_j - x_{ij}, где $i = l, 2, \dots, n$; $j = 0, 1, 2, \dots, p$. Так для линейной модели (2.7) значения $\hat{y}_i$ определяются на основании следующего выражения:

$$\hat{y}_i = \hat{b}_0 + \hat{b}_1 x_{i1} + \hat{b}_2 x_{i2} + \dots + \hat{b}_p x_{ip}$$

Для каждого набора оценок параметров $\hat{b}_0, \hat{b}_1, \dots, \hat{b}_p$ того или иного варианта модели, описывающей рассматриваемый процесс, рассчитывается «свой» ряд ошибки e_i, который можно интерпретировать как ряд оценок ее истинных, но неизвестных значений ε теоретической регрессии (4.4).

В общем случае «качество» эконометрической модели оценивается с помощью различных характеристик. Самой простой из них является *средняя ошибка аппроксимации*, которая вычисляется как среднее отклонение расчетных значений от результатов фактических измерений. Совокупность отклонений $e_i = y_i - \hat{y}_i$ можно рассматривать как абсолютные ошибки аппроксимации, а их абсолютные относительные величины

$$\left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100\%$$

как относительные ошибки аппроксимации.

Чтобы получить общее представление о качестве модели, из относительных отклонений по каждому наблюдению вычисляют среднюю ошибку аппроксимации $\bar{A}$ как простую среднюю арифметическую:

$$\overline{A} = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100\% . \qquad (5.1)$$

Считается, что допустимый предел ошибки не превышает 8 – 10%.

Другой характеристикой качества модельного уравнения регрессии является несмещенная оценка дисперсии случайных отклонений S_e^2:

$$S_e^2 = \frac{\sum_{i=1}^{n} e_i^2}{n-p-1} = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-p-1}, \qquad (5.2)$$

где p – число объясняющих переменных, факторов. Корень квадратный из оценки дисперсии обозначается как S_e и называется *стандартной ошибкой регрессии*.

Ошибка модельной регрессии во многом предопределена тем, что оценки $\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_p$ рассчитывают по данным случайных измерений, и они являются случайными значениями величин b_0, b_1, …, b_P - неизвестных коэффициентов регрессии. Насколько хорошим оказывается соответствие между ними, насколько приемлемым можно считать «качество» полученной модели регрессии.

Надежность случайных оценок устанавливают также с помощью определения оценок их дисперсий (стандартных ошибок). Кроме того, строят доверительные интервалы для теоретических значений и проверяют статистические гипотезы о значимости отличия их эмпирических величин от ожидаемых, теоретических значений.

5.2. Оценка качества коэффициентов модельной регрессии.

Оценки коэффициентов множественной линейной регрессии в матричной форме (5.2) определяются следующим образом:

$$\hat{B} = (X^T X)^{-1} (X^T Y).$$

Чтобы оценить ошибку оценки коэффициентов регрессии, представленных матрицей $\hat{B}$, подставим в правую часть формулы теоретические значения объясняемой переменной $Y = XB + \varepsilon$:

$$\hat{B} = (X^T X)^{-1} (X^T Y) = (X^T X)^{-1} X^T (XB + \varepsilon) = (X^T X)^{-1} (X^T X) B +$$
$$+ (X^T X)^{-1} X^T \varepsilon = B + (X^T X)^{-1} X^T \varepsilon.$$

Таким образом, ошибка полученной оценки $\Delta \hat{B}$ имеет вид:

$$\Delta \hat{B} = \hat{B} - B = (X^t X)^{-1} X^T \varepsilon.$$

Дисперсия случайной многомерной величины $\hat{B}$ определяется с помощью ковариационной матрицы $\Sigma_{\hat{B}}$:

$$\Sigma_{\hat{B}} = M((\hat{B} - B)(\hat{B} - B)^T) =$$

$$= M[(X^TX)^{-1}X^T\varepsilon)((X^TX)^{-1}X^T\varepsilon)^T] = M[(X^TX)^{-1}X^T\varepsilon\varepsilon^TX(X^TX)^{-1}].$$

В силу того, что объясняющие переменные X_J не являются случайными величинами, их можно вынести за знак математического ожидания:

$$\Sigma_{\hat{B}} = (X^TX)^{-1}X^TM(\varepsilon \cdot \varepsilon^T)X(X^TX)^{-1}. \qquad (5.4)$$

Матрица $M(\varepsilon \cdot \varepsilon^T)$ представляет собой ковариационную матрицу неизвестных случайных отклонений $\boldsymbol{\varepsilon}$ теоретической регрессии:

$$M(\varepsilon\varepsilon^T) = \Sigma_\varepsilon = \begin{pmatrix} M(\varepsilon_1^2) & M(\varepsilon_1\varepsilon_2) & \Lambda & M(\varepsilon_1\varepsilon_n) \\ M(\varepsilon_2\varepsilon_1) & M(\varepsilon_2^2) & \Lambda & M(\varepsilon_2\varepsilon_n) \\ . \; . \; . & . \; . \; . & . \; . \; . & . \; . \; . \\ M(\varepsilon_n\varepsilon_1) & M(\varepsilon_n\varepsilon_2) & \Lambda & M(\varepsilon_n^2) \end{pmatrix}.$$

В силу предпосылки МНК 2° все диагональные элементы одинаковы $M(\varepsilon_i^2) = D(\varepsilon_i) = \sigma_\varepsilon^2$, а все остальные равны нулю в силу предпосылки 3°. Таким образом, ковариационная матрица случайных ошибок $\Sigma_\varepsilon = \sigma_\varepsilon^2 \cdot E_n$, а выражение (5.4) принимает следующий вид:

$$\Sigma_{\hat{B}} = \sigma_\varepsilon^2(X^TX)^{-1}X^TE_nX(X^TX)^{-1} = \sigma_\varepsilon^2(X^TX)^{-1}. \qquad (5.5)$$

Неизвестное значение дисперсии случайного отклонения теоретической регрессии σ_ε^2 заменяется соответствующей несмещенной выборочной оценкой S_e^2 (3.2). Следовательно, по выборке мы можем определить только выборочные оценки дисперсий коэффициентов эмпирической регрессии $\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_p$, которые являются диагональными элементами матрицы $\Sigma_{\hat{B}}$:

$$S_{\hat{b}_j}^2 = S_e^2(X^TX)_{jj}^{-1}, \qquad (5.6)$$

где через $(X^TX)_{jj}^{-1}$ обозначены диагональные элементы обратной матрицы $(X^TX)^{-1}$, $j = 0, 1, 2, \ldots, p$.

Стандартные ошибки коэффициентов регрессии вычисляются по формулам:

$$S_{\hat{b}_j} = S_e\sqrt{(X^TX)_{jj}^{-1}}, \qquad (5.7)$$

где $j = 0, 1, \ldots, p$.

5.3. Доверительные интервалы коэффициентов регрессии.

Для построения интервальной оценки неизвестных коэффициентов регрессии b_i вводится случайная величина $\dfrac{\hat{b}_i - b_i}{S_{bi}} = t_{bi}$ - стандартизованный коэффициент регрессии, имеющая распределение Стьюдента с числом степеней свободы $\nu = n - p - 1$.

При заданном уровне значимости α доверительный интервал записывается следующим образом:

$$\left| \frac{\hat{b}_i - b_i}{S_{\hat{b}i}} \right| = \left| t_{\hat{b}i} \right| < t_{\alpha,\nu} \,,$$

где $t_{\alpha,\nu}$ - табличное значение t-критерия Стьюдента

Из данного неравенства следует:

$$\hat{b}_i - S_{\hat{b}i} \cdot t_{\alpha,\nu} < b_i < \hat{b}_i + S_{\hat{b}i} \cdot t_{\alpha,\nu}$$

$$\hat{b}_i - \Delta\hat{b}_i < b_i < \hat{b}_i + \Delta\hat{b}_i \qquad (5.8)$$

где $S_{\hat{b}i}$ и $\Delta\hat{b}_i = S_{\hat{b}i} \cdot t_{\alpha,n-p-1}$ - стандартная и предельная ошибки выборочных оценок $\hat{b}_i$ соответственно.

5.4. Стандартная ошибка и доверительные интервалы регрессии.

Дисперсия многомерной случайной величины $\hat{Y} = X \cdot \hat{B}$ определяется с помощью ковариационной матрицы $\Sigma_{\hat{B}}$:

$$\Sigma_{\hat{Y}} = M((X\hat{B} - XB)(X\hat{B} - XB)^T) =$$

$$= M[X(\hat{B} - B)(\hat{B} - B)^T)X^T] = X \cdot M[(\hat{B} - B)(\hat{B} - B)] \cdot X^T = X \cdot \Sigma_{\hat{B}} X^T,$$

где матрица $\Sigma_{\hat{B}}$ определена формулой (5.5).

Таким образом, получаем окончательное выражение в матричной форме:

$$\Sigma_{\hat{Y}} = X \cdot S_e^2 (X^T X)^{-1} X^T = S_e^2 \cdot X(X^T X)^{-1} X^T. \qquad (5.9)$$

Выборочные оценки дисперсий i-го значения эмпирической регрессии $S^2_{\hat{y}i}$, соответствующего i-му набору значений факторов $(1, x_{i1}, x_{i2}, \ldots, x_{ip})$ в i-ой строке матрицы $\boldsymbol{X}$ исходных данных, располагаются по диагонали матрицы $\Sigma_{\hat{Y}}$:

$$S^2_{\hat{y}i} = S_e^2 \cdot X_i (X^T X)^{-1} X_i^T. \qquad (5.10)$$

Стандартные ошибки оценок значений регрессии вычисляются по формулам:

$$S_{\hat{y}i} = S_e \cdot \sqrt{X_i (X^T X)^{-1} X_i^T}\,, \qquad (5.11)$$

где $i = 1, 2, ..., n$.

Доверительные интервалы для неизвестной функции регрессии $\overline{y}_i$ строятся также с помощью случайной величины $\frac{\hat{y}_i - \overline{y}_i}{S_{\hat{y}i}} = t_{\hat{y}i}$ - стандартизованная переменная, имеющей распределение Стьюдента с числом степеней свободы $\nu = n - p - 1$, i - номер измерения.

$$\left| \frac{\hat{y}_i - \overline{y}_i}{S_{\hat{y}i}} \right| = \left| t_{\hat{y}i} \right| \leq t_{\alpha,\nu}$$

Из данного неравенства следует:

$$\hat{y}_i - \Delta\hat{y}_i < \overline{y}_i < \hat{y}_i + \Delta\hat{y}_i, \qquad (5.12)$$

где $S_{\hat{y}i}$ и $\Delta\hat{y}_i = S_{\hat{y}i} \cdot t_{\alpha,\nu}$ - стандартная и предельная ошибки расчетного значения $\hat{y}_i$ соответственно.

5.5. Статистическая значимость уравнения регрессии.

Проверить статистическую значимость уравнения множественной регрессии означает установить, соответствует ли регрессионная модель, принятая для объяснения взаимосвязи между переменными, исходным статистическим данным. Или, другими словами, достаточно ли включенных в уравнение регрессии факторов для описания поведения объясняемой переменной на основе имеющихся выборочных данных.

Проверка значимости уравнения регрессии производится с помощью метода статистического анализа – дисперсионного анализа. Оценивание качества модельного уравнения регрессии с помощью *F-критерия* Фишера состоит в проверке гипотезы H_0 о статистической значимости уравнения регрессии. Это является обоснованием реальности показателя статистической связи. В случае, когда нулевая гипотеза отвергается, влияние включенных в регрессию факторов на объясняемую переменную преобладает над ее изменениями в силу других, не принятых во внимание причин.

Для этого сравнивают фактическое значение критерия $F_{эмп}$ с критическим, табличным значением $F_{табл}$.

$$F_{эмп} = \frac{Q_{рег} / k_1}{Q_e / k_2} = \frac{R^2_{YX1X2}}{1 - R^2_{YX1X2}} \cdot \frac{n - p - 1}{p},$$

где p - число объясняющих переменных, или факторов, включенных в модель, $k_1 = p$ и $k_2 = n - p - 1$ - число степеней свободы большей (числителя) и меньшей (знаменателя) дисперсий соответственно.

$$F_{табл} = F_{\alpha,k1,k2} = F(\alpha, p, n - p - 1).$$

Здесь α - уровень значимости критерия.

Если $F_{эмп} > F_{табл}(\alpha, k_1, k_2)$, то гипотеза H_0 о случайной природе статистической связи отклоняется. С вероятностью $(1-\alpha)$ делается заключение о статистической значимости уравнения регрессии в целом и показателя тесноты связи $R_{YX1X2...Xp}$, которые сформировались под совместным воздействием факторов X_1, X_2, …, X_p, и которое нет оснований считать случайным. В противном случае оснований для отклонения гипотезы H_0 нет и данная статистическая связь статистически незначима.

6. УРАВНЕНИЕ РЕГРЕССИИ В СТАНДАРТИЗОВАННОЙ ФОРМЕ

6.1. Стандартизованные переменные.

В эконометрике часто используется обобщенный подход к определению параметров множественной регрессии (4.13) с исключенным коэффициентом $\hat{b}_0$:

$$\hat{y} = \overline{Y} + \hat{b}_1(X_1 - \overline{X}_1) + \hat{b}_2(X_2 - \overline{X}_2) + ... + \hat{b}_p(X_p - \overline{X}_p).$$

Разделим обе части уравнения на стандартное отклонение объясняемой переменной S_Y и представим его в виде:

$$\frac{\hat{y} - \overline{Y}}{S_Y} = \frac{\hat{b}_1}{S_Y}(X_1 - \overline{X}_1) + \frac{\hat{b}_2}{S_Y}(X_2 - \overline{X}_2) + ... + \frac{\hat{b}_p}{S_Y}(X_p - \overline{X}_p).$$

Разделим и умножим каждое слагаемое на стандартное отклонение соответствующей факторной переменной, чтобы перейти к стандартизованным (центрированным и нормированным) переменным:

$$\frac{\hat{y} - \overline{Y}}{S_Y} = \frac{\hat{b}_1 S_{X1}}{S_Y}\frac{(X_1 - \overline{X}_1)}{S_{X1}} + \frac{\hat{b}_2 S_{X2}}{S_Y}\frac{(X_2 - \overline{X}_2)}{S_{X2}} + ... + \frac{\hat{b}_p S_{Xp}}{S_Y}\frac{(X_p - \overline{X}_p)}{S_{Xp}}$$

где новые переменные обозначены как

$$\hat{t}_{\hat{y}} = \frac{\hat{y} - \overline{Y}}{S_Y},\ t_{x1} = \frac{X_1 - \overline{X_1}}{S_{X1}}, ..., t_{xp} = \frac{X_p - \overline{X_p}}{S_{Xp}}.$$

Все стандартизованные переменные имеют нулевую среднюю величину и одинаковую дисперсию, равную единице.

Уравнение регрессии в стандартизованной форме имеет вид:

$$\hat{t}_y = \hat{\beta}_1 \cdot t_{x1} + \hat{\beta}_2 \cdot t_{x2} + ... \hat{\beta}_p t_{xp}\ ,, \qquad (6.1)$$

где $\hat{\beta}_1 = \frac{\hat{b}_1 S_{X1}}{S_Y}$, $\hat{\beta}_2 = \frac{\hat{b}_2 S_{X2}}{S_Y}$, ..., $\hat{\beta}_p = \frac{\hat{b}_p S_{Xp}}{S_Y}$ - стандартизованные коэффициенты множественной регрессии.

Стандартизованные коэффициенты регрессии $\hat{\beta}_i$ отличаются от коэффициентов $\hat{b}_i$ обычной, естественной формы тем, что их величина не зависит масштаба измерения объясняемой и объясняющих переменных модели. Кроме того, между ними существует простая взаимосвязь

$$\hat{b}_1 = \frac{\hat{\beta}_1 S_Y}{S_{X1}}, \hat{b}_2 = \frac{\hat{\beta}_2 S_Y}{S_{X2}}, \ldots, \hat{\beta}_p = \frac{\hat{\beta}_p S_Y}{S_{Xp}}. \tag{6.2}$$

Она дает другой способ вычисления коэффициентов $\hat{b}_i$ по известным значениям $\hat{\beta}_i$, более удобный в случае, например, двухфакторной регрессионной модели.

6.2. Система уравнений МНК в стандартизованных переменных.

Оказывается, что для вычисления коэффициентов стандартизованной регрессии нужно знать только парные коэффициенты линейной корреляции. Чтобы показать каким образом это делается, исключим из нормальной системы уравнений МНК неизвестную $\hat{b}_0$ с помощью первого уравнения. Умножая первое уравнение на $(-\overline{X}_1)$ и почленно складывая его со вторым уравнением, получим:

$$\hat{b}_1\left(\overline{X_1^2} - \overline{X}_1^2\right) + \hat{b}_2\left(\overline{X_2 X_1} - \overline{X}_2 \overline{X}_1\right) + \Lambda + \hat{b}_p\left(\overline{X_p X_1} - \overline{X}_p \overline{X}_1\right) =$$

$$= \overline{YX}_1 - \overline{Y} \cdot \overline{X}_1.$$

Заменяя обозначениями дисперсии и ковариаций выражения в скобках

$$\left(\overline{X_1^2} - \overline{X}_1^2\right) = S_{X1}^2, \quad \left(\overline{X_2 X_1} - \overline{X}_2 \overline{X}_1\right) = cov(X_2, X_1) = S_{X1X2},$$

$$\left(\overline{X_p X_1} - \overline{X}_p \overline{X}_1\right) = S_{XpX1}, \quad \overline{YX}_1 - \overline{Y} \cdot \overline{X}_1 = S_{YX1},$$

перепишем второе уравнение в удобном для дальнейшего упрощения виде:

$$\hat{b}_1 S_{X1}^2 + \hat{b}_2 S_{X2X1} + \Lambda + \hat{b}_p S_{XpX1} = S_{YX1}.$$

Разделим обе части этого уравнения на стандартное отклонение переменных S_Y и S_{X1}, а каждое слагаемое разделим и умножим на стандартное отклонение переменной, соответствующей номеру слагаемого:

$$\frac{\hat{b}_1 S_{X1}}{S_Y} \frac{S_{X1}^2}{S_{X1} S_{X1}} + \frac{\hat{b}_2 S_{X2}}{S_Y} \frac{S_{X2X1}}{S_{X2} S_{X1}} + \Lambda + \frac{\hat{b}_p S_{Xp}}{S_Y} \frac{S_{XpX1}}{S_{Xp} S_{X1}} = \frac{S_{YX1}}{S_Y S_{X1}}.$$

Вводя характеристики линейной статистической связи:

$$r_{X1X2} = \frac{S_{X1X2}}{S_{X1} S_{X2}}, \quad \Lambda \quad r_{X1Xp} = \frac{S_{X1Xp}}{S_{X1} S_{Xp}}, \quad r_{X1Y} = \frac{S_{X1Y}}{S_{X1} S_Y},$$

и стандартизованные коэффициенты регрессии

$$\hat{\beta}_1 = \frac{\hat{b}_1 S_{X1}}{S_Y},\ \hat{\beta}_2 = \frac{\hat{b}_2 S_{X2}}{S_Y},\ \ldots,\ \hat{\beta}_p = \frac{\hat{b}_p S_{Xp}}{S_Y},$$

получаем:

$$\hat{\beta}_1 + \hat{\beta}_2 r_{X1X2} + \hat{\beta}_3 r_{X1X3} + \Lambda\ \hat{\beta}_p r_{X1Xp} = r_{X1Y}.$$

После аналогичных преобразований всех остальных уравнений ,нормальная система линейных уравнений МНК (4.12) принимает следующий, более простой вид:

$$\begin{cases} \hat{\beta}_1 \quad + \hat{\beta}_2 \cdot r_{X1X2} + \hat{\beta}_3 \cdot r_{X1X3} + \ldots + \hat{\beta}_p \cdot r_{X1Xp} = r_{X1Y} \\ \hat{\beta}_1 \cdot r_{X2X1} + \hat{\beta}_2 \quad + \hat{\beta}_3 \cdot r_{X2X3} + \ldots + \hat{\beta}_p \cdot r_{X2Xp} = r_{X2Y} \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ \hat{\beta}_1 \cdot r_{XpX1} + \hat{\beta}_2 \cdot r_{XpX2} + \hat{\beta}_3 \cdot r_{XpX3} + \ldots + \hat{\beta}_p \quad = r_{X2Y}. \end{cases} \tag{6.3}$$

6.3. Параметры стандартизованной регрессии

Стандартизованные коэффициенты регрессии в частном случае модели с двумя факторами определяются из следующей системы уравнений:

$$\begin{cases} \hat{\beta}_1 \quad + \hat{\beta}_2 \cdot r_{X1X2} = r_{X1Y} \\ \hat{\beta}_1 r_{X1X2} \quad + \hat{\beta}_2 \cdot = r_{X2Y} \end{cases} \tag{6.4}$$

Решая эту систему уравнений, находим:

$$\hat{\beta}_1 = \frac{r_{X1Y} - r_{X1X2} \cdot r_{x2Y}}{1 - r_{X1X2}^2}, \tag{6.5}$$

$$\hat{\beta}_2 = \frac{r_{X2Y} - r_{X1X2} \cdot r_{X1Y}}{1 - r_{X1X2}^2}. \tag{6.6}$$

Подставив найденные значения коэффициентов парной корреляции в уравнения (6.4) и (6.5), получим $\hat{\beta}_1$ и $\hat{\beta}_2$. Затем с помощью формул (6.2) нетрудно вычислить оценки коэффициентов $\hat{b}_1$ и $\hat{b}_2$, а затем, при необходимости, вычислить оценку $\hat{b}_0$ по формуле $\hat{b}_0 = \overline{Y} - \hat{b}_1 \overline{X}_1 - \hat{b}_2 \overline{X}_2$.

7. ЭКОНОМИЧЕСКИЙ АНАЛИЗ НА ОСНОВЕ МНОГОФАКТОРНОЙ РЕГРЕССИОННОЙ МОДЕЛИ

7.1 Коэффициенты стандартизованной регрессии

Стандартизованные коэффициенты регрессии показывают, на сколько стандартных отклонений S_Y изменится в среднем объясняемая переменная Y, если соответствующая объясняющая переменная X_i изменится на величину S_{Xi} одного ее стандартного отклонения при сохранении неизменной величину среднего уровня всех остальных факторов.

В силу того, что в стандартизованной регрессии все переменные заданы как центрированные и нормированные случайные величины, коэффициенты β_i сравнимы между собой. Сравнивая их друг с другом, можно ранжировать соответствующие им факторы X_i по силе воздействия на объясняемую переменную Y. В этом состоит основное преимущество стандартизованных коэффициентов регрессии от коэффициентов b_i регрессии в естественной форме, которые несравнимы между собой.

Эта особенность стандартизованных коэффициентов регрессии позволяет использовать при отсеве наименее значимых факторов X_i с близкими к нулю значениями их выборочных оценок $\hat{\beta}_i$. Решение об исключении их из модельного уравнения линейной регрессии принимается после проверки статистических гипотез о равенстве нулю его средней величины, равной.

7.2. Средние и частные коэффициенты эластичности

В экономических исследованиях для количественного представления результатов множественного анализа также часто применяют коэффициенты эластичности. Коэффициент эластичности характеризует изменение в процентах объясняемой переменной при изменении факторной переменной на 1% и определяется для зависимости $\hat{y} = f(x)$ как

$$Э_{y/x} = \frac{d\hat{y}/\hat{y}}{dx/x} = f'(x)\frac{x}{\hat{y}}.$$

Он имеет четкую экономическую интерпретацию для степенной функции $\hat{y} = a \cdot x^b$. Производная этой функции $f'(x) = a \cdot b \cdot x^{b-1}$. Оказывается, что коэффициент эластичности при этом равен

$$Э_{y/x} = a \cdot b \cdot x^{b-1}\frac{x}{a \cdot x^b} = \frac{a \cdot b \cdot x^b}{a \cdot x^b} = b$$

и имеет постоянную величину.

Обобщение понятия эластичности на другие формы регрессионной связи приводит к тому, что $Э_{y/x}$ оказывается зависящим от величины факторной переменной. Так для линейной регрессии $\hat{y} = a + b \cdot x$ эластичность принимает вид:

$$Э_{y/x} = b \cdot \frac{x}{\hat{y}} = b \cdot \frac{x}{a + bx}.$$

Поскольку коэффициент эластичности для линейной функции не является постоянной величиной, то вводят средний показатель эластичности:

$$\bar{Э}_{y/x} = b \cdot \frac{\bar{X}}{\bar{Y}}.$$

Для характеристики относительной силы влияния каждого фактора в многофакторной регрессии используют частные коэффициенты эластичности:

$$Э_{Y/X_i} = \frac{\partial \hat{Y}/y}{\partial X_i / X_i} = f'(X_1, X_2, \Lambda, X_p)\frac{X_i}{\hat{Y}}$$

В частном случае двухфакторной линейной зависимости

$$\hat{Y} = \hat{b}_0 + \hat{b}_1 \cdot X_1 + \hat{b}_2 \cdot X_2$$

для характеристики влияния X_1 и X_2 на Y рассчитывают частные коэффициенты эластичности:

$$\overline{Эл}_{Y/X1} = \hat{b}_1 \cdot \frac{\overline{X_1}}{\bar{Y}},$$

$$\overline{Эл}_{Y/X2} = \hat{b}_2 \cdot \frac{\overline{X_2}}{\bar{Y}}.$$

Экономическая интерпретация этой статистической взаимосвязи является следующей. При увеличении фактора X_1 на 1% от его среднего уровня объясняемая переменная Y изменяется на $\hat{b}_1/\bar{Y}$,% своего от среднего уровня, а при увеличении фактора X_2 на 1% от его среднего уровня объясняемая переменная Y изменяется на $\hat{b}_2/\bar{Y}$,% от своего среднего уровня.

Направление изменения объясняемой переменной (ее знак) зависит от знака коэффициента при соответствующем факторе.

7.3. Коэффициенты парной, частной и множественной корреляции. Наиболее обоснованной характеристикой статистической связи является выборочный коэффициент линейной корреляции. При исследовании двухфакторной модели $\hat{Y} = \hat{b}_0 + \hat{b}_1 \cdot X_1 + \hat{b}_2 \cdot X_2$ линейные коэффициенты парной корреляции определяются следующим:

$$r_{X1Y} = \frac{\overline{X_1Y} - \overline{X_1} \cdot \overline{Y}}{S_{X1} \cdot S_Y},$$

$$r_{X2Y} = \frac{\overline{X_2Y} - \overline{X_2} \cdot \overline{Y}}{S_{X2} \cdot S_Y},$$

$$r_{X1X2} = \frac{\overline{X_1X_2} - \overline{X_1} \cdot \overline{X_2}}{S_{X1} \cdot S_{X2}}.$$

Анализируя значения коэффициентов парной корреляции, приходим к выводам относительно силы и направления взаимосвязи между каждой парой переменных. Однако очевидно, что влияние каждой из факторных переменных на объясняемую переменную не может выражаться также просто, как в случае парной регрессии, в силу их взаимосвязи между собой.

Для характеристики влияния каждого фактора множественной регрессии, аналогичной стандартизованным коэффициентам регрессии, на основе парных коэффициентов линейной корреляции рассчитывают частные коэффициенты корреляции. Они также используются при отборе факторов: целесообразность включения того или иного фактора в модель можно обосновывать величиной показателя частной корреляции.

Частные коэффициенты корреляции характеризуют влияние фактора X_i на объясняемую переменную Y при устранении влияния других факторов, включенных в уравнение регрессии. В зависимости от их числа вычисляют частные корреляции различного порядка.

В случае двухфакторной модели частные коэффициенты корреляции первого порядка вычисляются непосредственно с помощью парных коэффициентов корреляции по формулам:

$$r_{X1Y/X2} = \frac{r_{X1Y} - r_{X1X2} \cdot r_{X2Y}}{\sqrt{(1 - r_{X1X2}^2)(1 - r_{X2Y}^2)}},$$

$$r_{X2Y/x1} = \frac{r_{X2Y} - r_{X1x2} \cdot r_{X1Y}}{\sqrt{(1 - r_{X1X2}^2)(1 - r_{X1Y}^2)}},$$

$$r_{X1X2/y} = \frac{r_{X1X2} - r_{X1Y} \cdot r_{X2Y}}{\sqrt{(1 - r_{X1Y}^2)(1 - r_{X2Y}^2)}}.$$

Если сравнить значения коэффициентов парной и частной корреляции, то приходим к выводу, что из-за сильной связи X_1 и Y, умеренной связи X_2 и Y коэффициенты r_{X1Y} и $r_{X1Y/X2}$ отличаются незначительно, r_{X2Y} и $r_{X2Y/X1}$ - значительно. Выводы о направлении связи на основе этих коэффициентов совпадают.

Когда общее число факторов больше двух и равно p, то возможны частные коэффициенты корреляции не только первого, но и второго, третьего, …, *(p-1)* порядка. Так влияние фактора X_1 на переменную Y можно оценить при разных способах исключения влияния других факторов: по одному, по два, …, всех *(p-2)* – ух факторов из числа входящих в регрессию. Хотя частная корреляция различного порядка и может представлять интерес при изучении различных взаимосвязей факторных переменных, на практике ограничиваются вычислением частных корреляций самого высокого порядка. Этими показателями обычно принято дополнять эмпирические модели множественной регрессии

Практическая значимость эмпирического уравнения множественной регрессии оценивается помощью коэффициента множественной корреляции $R_{YX1X2...Xp}$ или его квадрата $R_{YX1X2...Xp}^2$ – коэффициента множественной детерминации.

Коэффициент множественной корреляции характеризует тесноту связи рассматриваемого набора факторов с объясняемой переменной, или оценивает тесноту совместного влияния факторов на результат. Его величина не может быть меньше, чем величина наибольшего коэффициента парной корреляции:

$$R_{YX1X2\ldots Xp} \geq \max_{i=1,2,\ldots,p}(r_{XiY}).$$

При правильном выборе факторов величина коэффициента множественной корреляции должен существенно отличаться от коэффициентов парной корреляции. Целесообразность включения в состав регрессии нового фактора можно обосновать, сравнивая коэффициенты множественной и парной корреляции с его участием.

Коэффициент множественной корреляции можно рассчитать несколькими способами, это полезно делать для контроля правильности вычислений:

1) наиболее просто это сделать для уравнения линейной регрессии в стандартизованной форме:

$$R_{YX1X2\ldots Xp} = \sqrt{\beta_1 \cdot r_{X1Y} + \beta_2 \cdot r_{X2Y} + \Lambda + \beta_p r_{XpY}}\ .$$

2) используя матрицу парных коэффициентов корреляции:

$$R_{YX1X2} = \sqrt{1 - \frac{|R'|}{|R|}}\ ,$$

где $|R'| = \begin{vmatrix} 1 & r_{YX1} & r_{YX2} & \Lambda & r_{YXp} \\ r_{X1Y} & 1 & r_{X1X2} & \Lambda & r_{X1Xp} \\ r_{X2Y} & r_{X2X1} & 1 & \Lambda & r_{X2Xp}\ X \\ \Lambda & \Lambda & \Lambda & \Lambda & \Lambda \\ r_{XpY} & r_{XpX1} & r_{XpX2} & \Lambda & 1 \end{vmatrix}$ - определитель матрицы коэффициентов парной корреляции;

$|R| = \begin{vmatrix} 1 & r_{X1X2} & \Lambda & r_{X1Xp} \\ r_{X2X1} & 1 & \Lambda & r_{X2Xp} \\ \Lambda & \Lambda & \Lambda & \Lambda \\ r_{XpX1} & r_{XpX2} & \Lambda & 1 \end{vmatrix}$ - определитель матрицы коэффициентов парной корреляции между факторами, или межфакторного взаимодействия.

3) как корень квадратный из коэффициента множественной детерминации:

$$R_{YX1X2\ldots Xp} = \sqrt{R^2_{YX1X2\ldots Xp}} = \sqrt{\frac{S^2_{\hat{Y}}}{S^2_Y}} = \sqrt{\frac{S^2_Y - S^2_e}{S^2_Y}} =$$

$$= \sqrt{1 - \frac{S^2_e}{S^2_Y}} = \sqrt{1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_{(X1X2\ldots Xp)i})^2}{\sum_{i=1}^{n}(y_i - \overline{Y})^2}}\ .$$

На основании значения коэффициента множественной регрессии делается вывод о том, что зависимостью переменной y от факторных переменных объясняется R^2 процентов вариации, представленной исходными данными. Прочие факторы, не включенные в модель, составляют соответственно (1- R^2)% от общей вариации объясняемой переменной y.

8. АНАЛИЗ СТАЦИОНАРНЫХ ВРЕМЕННЫХ РЯДОВ

Часто экономические показатели, представленные временными рядами, имеют настолько сложную структуру, что моделирование таких рядов путем построения моделей тренда, сезонности и применения других традиционных подходов не приводит к удовлетворительным результатам. Во временном ряду ошибок остаются статистические зависимости, которые можно моделировать.

В последнее время большое внимание уделяется моделированию стационарных временных рядов. Это объясняется тем, что многие временные ряды могут быть приведены к стационарному виду после операции выделения тренда, фильтрации сезонной компоненты или взятия разности. Как правило, ряд ошибок — это стационарный ряд.

Наиболее распространенные модели стационарных рядов это *модели авторегрессии и модели скользящего среднего.*

8.1. Основные характеристики временных рядов.

Временной ряд может рассматриваться как одна частная реализация стохастического процесса. Специальный случай стохастических процессов представляют стационарные процессы, когда отдельная реализация служит «полномочным представителем» всей совокупности возможных реализаций случайного процесса. Рассмотрим формальное определение стационарности.

Ряд называется *строго стационарным (strictly stationary)* или *стационарным в узком смысле*, если совместное распределение n наблюдений $y_{tl}, y_{t2},\dots, y_{tn}$ такое же, как и для n наблюдений, смещенных на промежуток времени τ . Таким образом, свойства строго стационарного временного ряда не зависят от начала отсчета времени.

Исследователя обычно интересует не все распределение; а средние значения и ковариации. Поэтому на практике чаще пользуется *понятие слабой стационарности (weak stationary)* или *стационарности в широком смысле*.

В этом случае стационарность временного ряда связывается с требованиями того, чтобы он имел числовые характеристики (среднее, дисперсию и ковариацию), которые не зависят от момента времени t:

$$M(y_t) = M(y_{t+\tau}) = \mu;$$
$$D(y_t) = M[y_t - M(y_t)]^2 = M[y_{t+\tau} - M(y_{t+\tau})]^2 = \gamma(0);$$
$$Cov(y_t, y_{t+\tau}) = M[(y_t - \mu)(y_{t+\tau} - \mu)] = \gamma(\tau)$$

При дальнейшем изложении под термином «стационарность» будет рассматриваться слабая стационарность. Это означает, что автоковариация будет зависеть только с сдвига по времени τ и не будет зависеть от t.

Здесь ковариация $Cov(y_t, y_{t+\tau})$ называется автоковариацией (и обозначается $\gamma(\tau)$, так как характеризует статистическую связь между уровнями одного и того временного ряда, отстоящими на τ временных тактов. При анализе изменения $\gamma(\tau)$ в зависимости от временного сдвига τ принято говорить об *автоковариационной функции (autocovariation function)*. Очевидно, что при $\tau = 0$ автоковариационная функция $\gamma(0)$ равна дисперсии временного ряда $Cov(y_t, y_t) = D(y_t)$.

С понятием автоковариационной функции тесно связана *автокорреляционная функция, АКФ (autocorrelation function, ACF):*

$$\rho(\tau) = Cov(y_t, y_{t+\tau})/D(y_t) = \gamma(\tau)/\gamma(0).$$

Значения АКФ также характеризуют тесноту (степень) статистической связи между уровнями временного ряда, разделенными τ временными тактами. Однако в отличие от значений автоковариационной функции они безразмерны, не зависят от масштаба измерения уровней исследуемого временного ряда и подчиняются ограничению

$$|\rho(\tau)| \le 1, \text{ (очевидно, что } \gamma(0) = 1).$$

Из условия стационарности следует, что $\rho(\tau) = \rho(-\tau)$, поэтому при рассмотрении поведения АКФ ограничиваются лишь областью положительных значений τ.

На практике значения АКФ статистически оцениваются по имеющимся уровням временного ряда. Выборочная оценка коэффициента автокорреляции $r(\tau)$ может быть определена следующим образом:

$$r(\tau) = \frac{\frac{1}{n-\tau}\sum_{t=1}^{n-\tau}(y_t - \bar{y})(y_{t+\tau} - \bar{y})}{\frac{1}{n}\sum_{t=1}^{n}(y_t - \bar{y})^2}$$

где n - длина временного ряда $y_1, y_2, ..., y_n$; τ - временной сдвиг ($\tau = 1, 2, ..., n - 1$); $\bar{y}$ - оценка среднего значения, найденная по формуле:

$$\bar{y} = \frac{\sum_{t=1}^{n} y_t}{n}.$$

Отметим, что числитель выражения $r(\tau)$ представляет выборочную оценку коэффициента автоковариации. Иногда график АКФ, отражающий изменение $r(\tau)$ в зависимости от сдвига τ *называют коррелограммой (correlogram).*

Очевидно, что с увеличением значения лага τ число пар наблюдений (n - τ), используемых для расчета, уменьшается. Поэтому в практических руководствах рекомендуется поддерживать соотношение $\tau \leq n/4$.

Для стационарного временного ряда с увеличением τ автокорреляционная функция должна демонстрировать свойство монотонного убывания по абсолютной величине, так как взаимосвязь между уровнями ряда с ростом τ ослабевает. Однако это условие может нарушаться для выборочных значений АКФ.

Идея перенесения частной корреляции, «очищенной» взаимосвязи, на временные ряды находит свое выражение в *частной автокорреляционной функции, ЧАКФ (partial autocorrelation function, PACE)*. С помощью ЧАКФ измеряется корреляция между уровнями ряда y_t, и $y_{t+\tau}$, разделенными τ временными тактами, при исключении влияния на эту взаимосвязь всех промежуточных уровней ряда $y_{t+1}, y_{t+2}, \ldots, y_{t+\tau-1}$.

Например, коэффициент частной автокорреляции $\rho_{|}(\tau)$ при τ = 2 будет определять корреляцию между уровнями временного ряда, разделенными двумя тактами времени, при условии, что значения промежуточных уровней зафиксированы на среднем уровне:

$$\rho_{ч}(2) = \rho(y_t , y_{t+2} \,|\, y_{t+1} = \mu).$$

В эконометрических пакетах присутствует возможность по строения графика ЧАКФ, на котором показаны выборочные оценки коэффициентов частной автокорреляции в зависимости от лагов τ.

Очевидно, что коэффициент частной автокорреляции $\rho_{ч}(1)$ для лага τ = 1 будет равен коэффициенту автокорреляции $\rho(1)$, так как при этом значении τ нет промежуточных лагов. Но при $\tau > 1$ уже появятся отличия в значениях этих коэффициентов.

Примером стационарности служит *"белый шум"* (white noise). *Процессом белого шума ("белым шумом", "чисто случайным временным рядом")* называют стационарный временной ряд, в данном случае случайная величина ε_t, которая удовлетворяет следующим условиям:

$M(\varepsilon_t) = 0$;

$M(\varepsilon_t, \varepsilon_t) = D(\varepsilon_t) = \gamma(0) = \sigma_0^2 > 0$;

$M(\varepsilon_t, \varepsilon_{t\pm\tau}) = Cov(\varepsilon_t, \varepsilon_{t\pm\tau}) = \gamma(\tau) = 0$ при $\tau \neq 0$,

(причем постоянная дисперсия σ_0^2 не зависит от t).

В качестве примера "белого шума" можно привести остатки, рассматриваемые в классической линейной регрессионной модели. В случае нормального распределения остатков они образуют гауссовский "белый шум". Следует подчеркнуть, что "белый шум" не обязательно должен подчиняться нормальному распределению, в то же время нормальный (гауссовский) "белый шум" удовлетворяет всем этим условиям.

В заключение отметим, что вид АКФ и ЧАКФ оказывает существенную помощь в выборе моделей, описывающих поведение анализируемых временных рядов.

8.2. Авторегрессионные модели.

Рассмотрим класс авторегрессионных моделей, сокращенно обозначаемых *АР(р) -модели* (число в скобках указывает порядок авторегрессии) или, как принято в специальной англоязычной литературе и мировой статистической практике, *AR(p) - models (AutoRegressive processes of order p).*

В авторегрессии каждое значение ряда находится в линейной зависимости от предыдущих значений. Если анализируемый динамический процесс зависит от значений, отстоящих до p временных лагов назад, то это авторегрессионный процесс порядка, т.е. $AR(p)$:

$$y_t=\alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + \varepsilon_t$$

или

$$(1 - \alpha_1 B + \alpha_2 B^2 + \ldots + \alpha_p B^p)\, y_t = \Phi(B)\, y_t = \varepsilon_t \qquad (8.1)$$

где B — оператор сдвига, т.е. преобразование ряда, смещающее его на один временной лаг, $\Phi(B)$ — оператор авторегресии.

При этом для выполнения условия стационарности все корни многочлена $\Phi(B)$ должны лежать вне единичного круга, т.е. все корни соответствующего характеристического уравнения должны быть по модулю больше 1 и различны.

Рассмотрим простейший вариант линейного авторегрессионного процесса - *модель авторегрессии 1-го порядка* - AR(1), или *марковский процесс.*

Эта модель может быть представлена в виде:

$$y_t = \alpha y_{t-1} + \varepsilon_t \qquad (8.2)$$

где α — числовой коэффициент, $|\alpha| < 1$, ε_t — последовательность случайных величин, образующих “белый шум”.

Основные свойства марковского процесса:

1. $M(y_t) = 0$;
2. $D(y_t) = \sigma_0^2 / (1 - \alpha^2)$;
3. $Cov(y_t, y_{t\pm\tau}) = \alpha^\tau D(y_t)$;
4. $\rho(y_t, y_{t\pm\tau}) = \rho(\tau) = \alpha^\tau$.

Представим $y_t = \alpha y_{t-1} + \varepsilon_t$ в виде:

$$y_t = \alpha(\alpha y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \alpha^2 y_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t = \alpha^2(\alpha y_{t-3} + \varepsilon_{t-2}) + + \alpha \varepsilon_{t-1} + \varepsilon_t =$$

$$\alpha^3 y_{t-3} + \alpha^2 y_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t = \ldots =$$

$$\varepsilon_t + \alpha \varepsilon_{t-1} + \alpha^2 y_{t-2} + \ldots = \sum_{\tau=0}^{\infty} \alpha^\tau \varepsilon_{t-\tau}\ ; \qquad (8.3)$$

Очевидно, что y_t зависит от всех предыдущих, но не от будущих случайных величин ε_t . Отсюда с учетом характера “белого шума”. непосредственно следует, что $M(y_t) = 0$.

Также получим выражение для дисперсии марковского процесса AR(1), с помощью выражения (8.3):

$$D(y_t) = \sigma_0^2\ (1 + \alpha^2 + \alpha^4 + \ldots) = \sigma_0^2/(1 - \alpha^2). \qquad (8.4)$$

Сумма бесконечной геометрической прогрессии записана при условии $|\alpha| < 1$. Отсюда видно, что при значении α близком к значениям, равным ± 1 дисперсия ряда будет намного больше σ_0^2- дисперсии белого шума. Следовательно, если последовательные значения ряда сильно коррелированны, то даже незначительные возмущения будут порождать размашистые колебания.

Чтобы показать свойства 3 и 4, умножим обе части уравнения (8.2) на y_{t-1} и возьмем математическое ожидание:

$$M(y_t\, y_{t-1}) = \alpha M(y_{t-1}\, y_{t-1}) + M(\varepsilon_t)\, M(y_{t-1})$$

где второе слагаемое $M(\varepsilon_t\, y_{t-1})$ записано с учетом некоррелируемости значений ряда с любыми будущими случайными величинами ε_t. Окончательная запись этого соотношения

$$Cov(y_t,\, y_{t-1}) = \alpha\, D(y_t).$$

Таким образом,

$$\alpha = Cov(y_t,\, y_{t-1})/\ D\ (y_t),$$

т.е. α - коэффициент автокорреляции первого порядка (определяет значение коэффициента парной корреляции между соседними уровнями ряда):

$$\rho(1) = \alpha.$$

Можно показать, что

$$Cov(y_t,\, y_{t\pm\tau}) = \alpha^{\tau}\, D(y_t),$$

а, следовательно,

$$\rho(y_t,\, y_{t\pm\tau}) = \alpha^{\tau} = \rho(\tau).$$

Поэтому степень тесноты корреляционной связи между членами последовательности экспоненциально убывает по мере их взаимного удаления друг от друга во времени.

Все автокорреляции марковского процесса можно выразить через автокорреляцию первого порядка:

$$\rho(\tau) = \rho\,(1)\ \rho\,(\tau - 1) = \rho(1)^2 \rho(\tau - 2) = \ldots = \rho\,(1)^{\tau} = \alpha^{\tau}.$$

Значения *частной автокорреляционной функции* равны нулю для всех лагов $k > 2$, что может быть использовано при подборе модели. Этот результат верен для теоретической частной автокорреляционной функции и может не выполняться для выборочной автокорреляционной функции. Однако, если выборочные частные корреляции статистически незначимо отличаются от нуля при $k > 2$, то использование модели $AR(1)$ не противоречит исходным данным.

Условие стационарности ряда для AR(1) определяется требованием к коэффициенту α: $|\alpha| < 1$. Это же самое, что корень z_0 уравнения $1- \alpha z = 0$ - частного случая характеристического уравнения для общего линейного процесса авторегрессии, должен быть по абсолютной величине больше 1.

Процесс с параметром $|\alpha| > 1$ является нестационарным. Такие ряды маловероятны в реальных финансово-экономических задачах, так как это подразумевает взрывные ряды, а давление экономической среды не позволяет показателям принимать бесконечно большие значения.

Из авторегрессионных процессов выше первого порядка в экономической практике часто встречаются так называемые *процессы Юла.* Они описываются с помощью модели AR(2):

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t \tag{8.5}$$

Определим выражение для вычисления значений *автокорреляционной функции* $\rho(\tau)$, сокращенно АКФ, для любого значения сдвига ряда τ. Для этого опять умножим обе части уравнения (8.5) на $y_{t-\tau}$:

$$y_t y_{t-\tau} = \alpha_1 y_{t-1} y_{t-\square} + \alpha_2 y_{t-2} y_{t-\square} + \varepsilon_t y_{t-\tau} \tag{8.6}$$

и возьмем математическое ожидание:

$$M(y_t y_{t-\tau}) = \gamma(\tau) = \alpha_1 \gamma(\tau-1) + \alpha_2 \gamma(\tau-2). \tag{8.7}$$

Очевидно, что $M(\varepsilon_t\ y_{t-\tau}) = 0$,так как $y_{\ t-k}$ зависит от $\varepsilon_{t-\tau}$ и предшествующих значений, а не от будущих значений ε_t,. Разделим обе части равенства (3.7) на $\gamma(0)$:

$$\rho(\tau) = \alpha_1 \rho\,(\tau-1) + \alpha_2\, \rho\,(\tau-2). \tag{8.8}$$

Это выражение позволяет вычислить значение АКФ для различных значений лагов τ. Подставим последовательно в (8. 8) значение $\tau = 1$ и $\tau = 2$.

С учетом того, что $\rho(0) = 1$, а $\rho(-1) = \rho(1)$, получим

$$\rho(1) = \alpha_1 + \alpha_2\, \rho(1);$$
$$\rho(2) = \alpha_1\, \rho(1) + \alpha_2. \tag{8.9}$$

Эта система называется *системой Юла—Уокера (Yule— Walker)* для $AR(2)$.

Если разрешить эту систему относительно α_1 и α_2 ,то получим выражения:

$$\alpha_1 = \rho(1)[1 - \rho(2)]/\ [1 - \rho^2(1)];$$
$$\alpha_2 = [\rho(2) - \rho^2(1)]/\ [1 - \rho^2(1)]. \tag{8.10}$$

Выразим из системы (8.9) два первых значения АКФ:

$$\rho(1) = \alpha_1/(\ 1- \alpha_2);$$
$$\rho(2) = \alpha_2 + \alpha_1^2\ /(\ 1 - \alpha_2). \tag{8.11}$$

Определив $\rho(1)$ и $\rho(2)$, можно вычислить любые последующие значения АКФ с помощью (8.8).

Получим соотношение, связывающее между собой дисперсию ряда y_t и дисперсию белого шума ε_t, равную σ_0^2. Для этого умножим уравнение $AR(2)$ на y_t:

$$y_t y_t = \alpha_1 y_{t-1} y_t + \alpha_2 y_{t-2} y_t + \varepsilon_t\ (\ \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t\ )$$

Возьмем математическое ожидание:

$$\gamma(0) = \alpha_1 \gamma(\lambda) + \alpha_2 \gamma(2) + \sigma_0^2.$$

Очевидно, что $M(\varepsilon_t\ ,\ y_{t-1}\ ) = 0$, $M(\varepsilon_t\ ,\ y_{t-2}\ ) = 0$, так как y_t не зависит от будущих значений ε_t. Следовательно,

$$\sigma_0^2 = \gamma(0) - \alpha_1 \gamma(\lambda) - \alpha_2 \gamma(2).$$

От коэффициентов автоковариации перейдем к автокорреляционным коэффициентам, умножив и разделив правую часть уравнения на $\gamma(0)$:

$$\sigma_0^2 = \gamma(0)\,[\,1-\alpha_1\gamma(1)-\alpha_2\gamma(2)]. \qquad (8.12)$$

Подставим в него выражения (3.11) для $\rho(1)$ и $\rho(2)$ через величины α_1 и α_2:

$$\gamma(0) = \frac{(1-\alpha_2)\,\sigma_0^2}{(1+\alpha_2)\left((1-\alpha_2)^2-\alpha_1^2\right)}.$$

Отсюда, учитывая, что дисперсия должна быть положительна, получаем условия стационарности процесса $AR(2)$. Условия стационарности ряда y, также могут быть получены с учетом (6.14) из требований

$$|\,\rho(1)| < 1,\ \ |\rho(2)| < 1.$$

Отметим, что те же самые условия получаются из требования, чтобы все корни соответствующего характеристического уравнения $1 - \alpha_1 z - \alpha_2 z^2 = 0$ лежали вне единичного круга.

Условия стационарности процесса *AR(2)* могут быть записаны в виде

$$|\,\alpha_2| < 1;$$
$$\alpha_1 + \alpha_2 < 1;$$
$$\alpha_2 - \alpha_1 < 1. \qquad (8.13)$$

ЧАКФ для процесса *AR(p)* будет иметь ненулевые значения лишь при $k \le p$, а начиная с лага $k = p + 1$ теоретическая ЧАКФ равна нулю. Это свойство становится ключевым при подборе порядка p авторегрессионной модели для конкретных экономических временных рядов.

Рассмотренные свойства авторегрессионных моделей и изучение АКФ (ACF) и ЧАКФ (PACF) позволяют сформулировать следующие *практические рекомендации* по их идентификации:

1. У моделей *AR(p)* значения коэффициентов АКФ экспоненциально затухают (либо монотонно, либо попеременно меняя знак).

2. ЧАКФ для модели *AR(p)* имеет выбросы (пики) на первых) лагах, а значения коэффициентов для лагов, больших порядка авторегрессии, статистически незначимы.

8.3. Модели скользящего среднего.

Модель скользящего среднего предполагает, что в ошибках модели в предшествующие периоды сосредоточена информация обо всей предыстории ряда. В этой модели каждое новое значение - среднее между текущей флуктуацией и несколькими (в частности, одной) предыдущими ошибками.

Модели скользящего среднего порядка q, обозначаемые *CC(q)*, в англоязычной литературе *MA(q) (Moving Average models)*, имеют вид:

$$y_t = \varepsilon_t - \theta_1\,\varepsilon_{t-1} - \theta_2\,\varepsilon_{t-2} - \ldots - \theta_q\,\varepsilon_{t-q}\,, \qquad (8.14)$$

где ε_t — “белый шум”.

Широко распространены в статистической практике модели скользящего среднего 1-го (q = 1) и второго порядка (q = 2):

$$\text{MA(1)}: y_t = \varepsilon_t - \theta\,\varepsilon_{t-1}\,; \tag{8.15}$$

$$\text{MA(2)}: y_t = \varepsilon_t - \theta_1\,\varepsilon_{t-1} - \theta_2\,\varepsilon_{t-2}\,. \tag{8.16}$$

Рассмотрим модель скользящего среднего 1-го порядка — *MA*(1). Преобразуем (8.15), последовательно выражая $\varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3}$ и т.д.:

$$\varepsilon_t = y_t + \theta\,\varepsilon_{t-1} = y_t + \theta\,(y_{t-1} - \theta\,\varepsilon_{\tau-2}) = y_t + \theta\,y_{t-1}$$
$$+\theta^2(y_{t-2} + \theta\,\varepsilon_{t-3}) = y_t+\theta\,y_{t-1} + \theta^2\,y_{t-2} +\theta^3(y_{t-3} + \theta\,\varepsilon_{t-4}) =$$
$$= y_t+\theta\,y_{t-1} + \theta^2\,y_{t-2} +\theta^3\,y_{t-3} + \ldots$$

Это выражение можно переписать в виде:

$$y_t = \varepsilon_t - \sum_{k=1}^{\infty}\theta^k y_{t-k}\,. \tag{8.17}$$

Таким образом, ряд y_t, сгенерированный моделью *MA*(1), может быть представлен также в виде модели авторегрессии бесконечного порядка. В моделях скользящего среднего *MA*(q) не требуется накладывать никаких ограничений на параметры $\theta_1, \theta_2, ..., \theta_q$ для обеспечения стационарности ряда. Однако, если в модели MA(1) параметр θ по абсолютной величине больше или равен 1, то текущее значение y_t в соответствии с (8.17) будет зависеть от своих прошлых значений $y_{t-1}, y_{t-2}, ...$, берущихся с весами, бесконечно растущими по мере удаления в прошлое. Чтобы избежать этого, надо, чтобы веса в (6.21) образовывали сходящийся ряд, т.е. чтобы $|\theta| < 1$.

Отметим, что подобно тому, как ряд, генерированный моделью скользящего среднего первого порядка *MA*(1), может быть представлен в виде модели авторегрессии бесконечного порядка *AR*(∞), также существует представление A*R*(1) в виде *MA*(∞). При этом на параметры процесса *AR*(p) не накладываются никакие условия для того, чтобы этот процесс был обратимым. Но для стационарности процесса корни его характеристического уравнения должны лежать вне единичного круга. В то же время параметры процесса *MA(q)* не должны удовлетворять никаким условиям для стационарности, однако для обратимости корни его характеристического уравнения

$$1 - \theta_1 z - \theta_2 z^2 - ... - \theta_q z^q = 0.= 0$$

должны лежать вне единичного круга.

Найдем выражение для АКФ процесса *MA(q)*. Для этого представим y_{t-k} в виде соотношения (3.14):

$$y_{t-k} = \varepsilon_{t-k} - \theta_1\varepsilon_{t-k-1} - \theta_2\varepsilon_{t-k-2} - \ldots - \theta_q\varepsilon_{t-k-q}. \tag{8.18}$$

Перемножим соответственно левые и правые части уравнений (6.18) и (6.22), а затем возьмем математическое ожидание от получившегося выражения. При этом следует учесть, что элементы белого шума ε_{t1} и ε_{t2} не коррелируют при $t_1 \neq t_2$.

Тогда выражение для ковариации $M(y_t\,y_{t-\tau}) = \gamma(\tau)$ примет вид:

$$\gamma(\tau)=\begin{cases}\sigma_0^2\left(1+\theta_1^2+\theta_2^2+\ldots+\theta_q^2\right) & при\ \ \tau=0;\\ \sigma_0^2\left(-\theta_\tau+\theta_1\theta_{\tau+1}+\theta_2\theta_{\tau+2}+\ldots+\theta_{q-\tau}\theta_q\right), & при\ \ \tau=1,2,\ldots,q\\ 0 & при\ \ \tau>q.\end{cases} \qquad (8.19)$$

АКФ получается путем деления (8.19) на дисперсию процесса $\gamma(0)$:

$$\rho(\tau)=\begin{cases}\dfrac{-\theta_\tau+\theta_1\theta_{\tau+1}+\theta_2\theta_{\tau+2}+\ldots+\theta_{q-\tau}\theta_q}{1+\theta_1^2+\theta_2^2+\ldots+\theta_q^2} & при\ \ \tau=1,2,\ldots,q\\ 0, & при\ \ \tau>q\end{cases} \qquad (8.20)$$

Таким образом, АКФ – функция процесса *MA(q)* равна нулю для всех значений τ, больших порядка q. Это важное характеристическое свойство модели.

На практике наиболее часто используют частный случай модели — модель скользящего среднего 1-го порядка MA(1):

$$y_t = \varepsilon_t - \theta\,\varepsilon_{t-1}$$

где ε_t — “белый шум”.

Как уже было показано ранее, для обратимости процесса необходимо выполнение условия $|\theta| < 1$.

Очевидно, что $M(y_t) = 0$; $D(y_t) = \sigma_0^2$.

АКФ согласно (3.20) определяется выражением

$$\rho(\tau)=\begin{cases}\dfrac{-\theta}{1+\theta^2} & при\,\tau=1\\ 0 & при\,\tau\geq 2\end{cases}. \qquad (8.21)$$

ЧАКФ $\rho_ч(\tau)$ задается выражением

$$\rho_ч(\tau)=-\theta^\tau\,\frac{1-\theta^2}{1-\theta^{2(\tau+1)}} \qquad (8.22)$$

Поведение ЧАКФ определяется затухающей экспонентой. Если значение $\rho(1)$ положительно, то параметр $\theta < 0$, следовательно, $\rho_ч(\tau)$ осциллирует с переменным знаком. Если значение $\rho(1)$ отрицательно, то параметр $\theta > 0$, следовательно, все значения $\rho_ч(\tau)$ отрицательны.

Отмеченные свойства моделей скользящих средних позволяют сформулировать следующие *практические рекомендации* по их идентификации.

Для моделей MA(1):

• автокорреляционная функция имеет выброс (пик) при лаге, равном 1, а остальные значения статистически незначимы;

• частная автокорреляционная функция экспоненциально затухает (либо монотонно, либо осциллируя, т.е. меняя знак).

Для моделей MA(2):

• автокорреляционная функция имеет выбросы (пики) на лагах, равных 1 и 2, а остальные значения статистически незначимы;

• частная автокорреляционная функция имеет форму синусоиды или экспоненциально затухает.

8.4. Модели авторегрессии со скользящими средними в остатках

На практике для наглядности описания анализируемого экономического процесса в модель могут быть включены как члены, описывающие авторегрессионные составляющие, так и члены, моделирующие остаток в виде процесса скользящих средних. Такой процесс называется - *АРСС (p, q)* или, как принято в англоязычной литературе, *AutoRegressive-Moving Average (ARMA (p, q))*. Параметры p и q определяют соответственно порядок авторегрессионной составляющей и порядок скользящих средних.

Модель *ARMA(p, q)* имеет вид:

$$y_t=\alpha_1 y_{t-1} + \alpha_2 y_{t-2} + {}_{\dots}+\alpha_p y_{t-p} + \varepsilon_p - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q}\ . \qquad (8.23)$$

Такая модель может интерпретироваться как линейная множественная регрессия. В качестве объясняющих переменных в ней выступают предыдущие значения самой зависимой переменной, а в качестве регрессионного остатка - скользящие средние из элементов белого шума.

Чтобы процесс (8.23) был стационарным, необходимо и достаточно, чтобы все корни характеристического уравнения *AR(p)*-процесса лежали вне единичного круга:

$$1 - \alpha_1 z - \alpha_2 z^2 - \dots - \alpha_p z^p = 0. \qquad (8.24)$$

Аналогично, для обратимости процесса (8.23) необходимо и достаточно, чтобы все корни характеристического уравнения процесса MA(q) лежали вне единичного круга:

$$1 - \alpha_1 z - \alpha_2 z^2 - \dots - \alpha_q z^q = 0. \qquad (8.25)$$

Простейший смешанный процесс ARMA(1,1):

$$y_t=\alpha_1 y_{t-1} + \varepsilon_p - \theta_1\varepsilon_{t-1} \qquad (8.26)$$

Это уравнение можно преобразовать к виду:

$$y_t+\alpha_1 y_{t-1} = \varepsilon_p - \theta_1\varepsilon_{t-1} \qquad (8.27)$$

Стационарность процесса ARMA(1,1) обеспечивается условием $|\alpha| < 1$, а обратимость, в свою очередь, гарантируется выполнением условия $|\theta| <1$.

Автоковариационные функции процесса ARMA(1,1):

$$\gamma(0) = \frac{\left(1+\theta^2-2\alpha\theta\right)\sigma_0^2}{1-\alpha^2}, \qquad (8.28)$$

$$\gamma(1) = \frac{(1-\alpha\theta)(\alpha-\theta)\sigma_0^2}{1-\alpha^2}. \qquad (8.29)$$

Значение автоковариационной функции для лага τ больше 1 определяется следующим рекуррентным соотношением:

$$\gamma(\tau) = \alpha\gamma(\tau-1) \text{ при } \tau > 1. \qquad (8.30)$$

Следовательно, значения АКФ будут иметь вид

$$\rho(1) = \frac{(1-\alpha\theta)(\alpha-\theta)}{1+\theta^2-2\alpha\theta} \qquad (8.31)$$

$$\rho(\tau) = \alpha\,\rho(\tau-1) = \alpha^{\tau-1}\,\rho(1) \text{ при } \tau > 1. \qquad (8.32)$$

Из (8.31),(8.32) видно, что, хотя выражение для $\rho(1)$ отличается от соответствующего выражения для процесса *AR*(1), соотношение между $\rho(1)$ и последующими значениями *АКФ* такое же. Таким образом, для процесса *ARMA*(1,1) значения *АКФ* будут экспоненциально убывать от значения $\rho(1)$, причем если α положительна, — то монотонно, если отрицательна, — то знакопеременно.

Поведение *ЧАКФ* определяется начальным значением $\rho_ч(1)$, после которого функция экспоненциально убывает. Если θ положительна, то функция убывает монотонно, если отрицательна, — то знакопеременно.

Исследования показывают, что при использовании в экономических задачах модели *ARMA*(*p, q*), потребностям практики, как правило, удовлетворяют следующие пять видов этой модели, представленных в таблице.

Свойства автокорреляционных *(АКФ)*
и частных автокорреляционных *(ЧАКФ)* функций

Функция	*ARMA(1,0)*	*ARMA(2,0)*	*ARMA(0,1)*	*ARMA(0,2)*	*ARMA(1,1)*
АКФ	Экспоненциально затухает (монотонно или знакопеременно)	Экспоненциально затухает или имеет форму синусоидальной волны	Выброс (пик) на лаге 1	Выбросы (пики) на лагах 1 и 2	Экспоненциально затухает от значения $\rho(1)$ (монотонно или знакопеременно)
ЧАКФ	Выброс (пик) на лаге 1	Выбросы (пики) на лагах 1 и 2	Экспоненциально затухает (монотонно или знакопеременно)	Экспоненциально затухает или имеет форму синусоидальной волны	Экспоненциально убывает от значения $\rho_ч(1)$ (монотонно или знакопеременно)

В этой таблице указаны свойства *АКФ* и *ЧАКФ* для этих моделей. Очевидно, что модели *ARMA*(1,0) и *ARMA*(2,0) — это рассмотренные ранее модели *AR*(1), *AR*(2), а модели *ARMA*(0,1) и *ARMA*(0,2) соответствуют моделям скользящего среднего *MA*(1) и *MA*(2).

8.5. Методология Бокса—Дженкинса.

Экономические временные ряды за редким исключением *нестационарны.* Нестационарность чаще всего проявляется в наличии зависящей от времени неслучайной составляющей *f(t).* Если случайный остаток, полученный вычитанием из исходного ряда его неслучайной составляющей *f(t),* представляет собой стационарный временной ряд, то исходный ряд называется *нестационарным однородным.*

Для описания таких рядов используется модель *авторегрессии — проинтегрированного скользящего среднего АРПСС(p, d, q)* или в англоязычном варианте *AutoregRessive Integrated Moving Average mode, (ARIMA-model).* В специальной литературе она также известна как *модель Бокса—Дженкинса,* по имени авторов, разработавших этот подход и методологию моделирования.

Модель *ARIMA* используется для описания временных рядов, обладающих следующими свойствами:

1) ряд включает (аддитивно) составляющую *f(t)*, имеющую вид алгебраического полинома;

2) ряд, получившийся после применения к нему *d раз* процедуры последовательных разностей, может быть описан моделью *ARMA(p, q).*

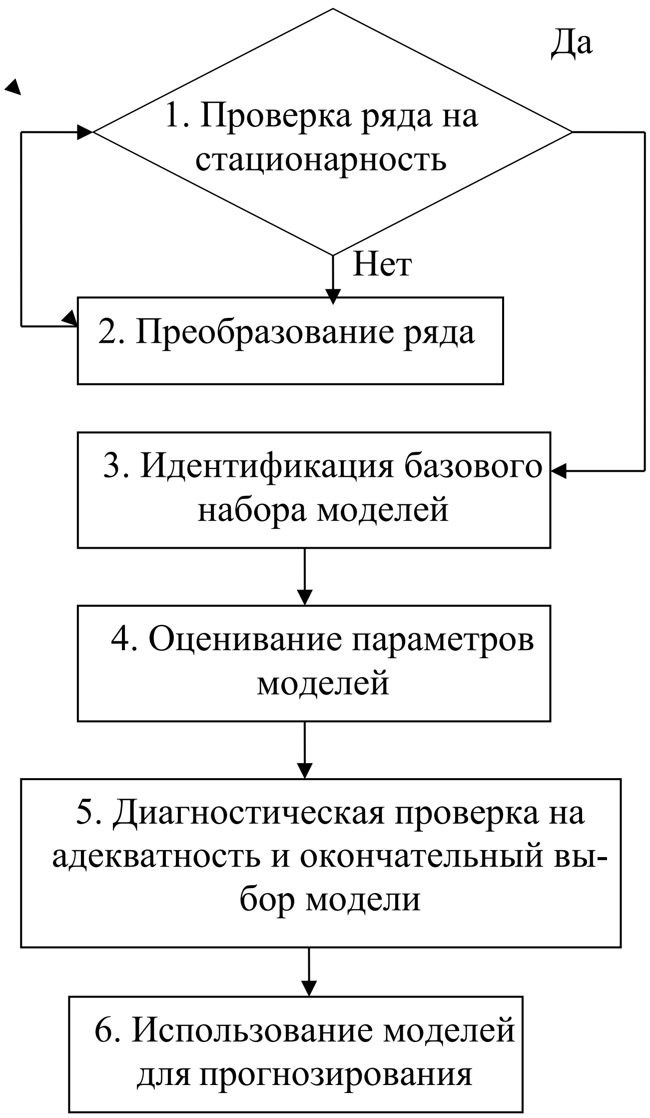

Рис.1. Схема подбора модели ARIMA

В эконометрических исследованиях наиболее распространены модели *ARIMA(p, d, q),* значения параметров которой не превышают числа 2. При этом параметры p и q определяют порядок авторегрессионной составляющей и порядок скользящего среднего (аналогично модели *ARMA(p, q))* соответственно, а параметр d — порядок разности (дискретной производной).

Временной ряд после взятия последовательных разностей (дискретного дифференцирования k-го порядка) может оказаться стационарным, удовлетворяющим модели *ARMA*(p, *q):*

$$\Delta^k y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + \varepsilon_p - \theta_1 \varepsilon_{t-1} - \\ - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}, \qquad (8.33)$$

где $\Delta^k y_t$, - k-*я* последовательная разность уровней ряда y_t.

Очевидно, что из модели для $\Delta^k y_t$ легко получить модель для исходного ряда. Например, для модели Δy_t это можно сделать с помощью выражения

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

Поясним содержание некоторых блоков, показанных на рис. 1.

Сначала *(в блоке* 1, 2) необходимо получить стационарный ряд. При тестировании исходных данных на стационарность прежде всего используется визуальный анализ графика. Например, уже на этом этапе можно обнаружить ярко выраженную трендовую или сезонную составляющую.

Также в методике Бокса-Дженкинса рекомендуется проводить анализ *АКФ* (или *ЧАКФ*). Быстрое убывание значений выборочной *АКФ* — простой критерий стационарности (аналогичное поведение должна демонстрировать и *ЧАКФ*).

Для перехода к стационарному ряду традиционно применяют оператор взятия последовательных разностей (процедуру дискретного дифференцирования). Для определения значения параметра d (порядка разности) может быть использован эвристический критерий $S^2(k)$. Для вычисления его по формуле

$$S^2(k) = \frac{\frac{1}{n-k}\sum_{t=1}^{n-k}(\Delta^k y_t)^2}{C_{2k}^k}$$

необходимо вычислить последовательные разности $k = 1, 2, \ldots, n$ для исходного ряда $y_1, y_2, \ldots, y_n$ как

$$\Delta^k y_t = \Delta^{k-1} y_t - \Delta^{k-1} y_{t-1}.$$

Начиная с некоторого значения $k = k_0$, величина $S^2(k)$ начинает стабилизироваться, оставаясь примерно на одном уровне при дальнейшем увеличении k. Тогда оценка порядка дискретной производной нестационарного временного ряда при построении *ARIMA*-модели $d = k_0 – 1$.

Также о том, что необходимая для стационарности ряда степень разности достигнута, будет свидетельствовать быстрое затухание *АКФ*. В практических исследованиях d, как правило, не превышает 2.

В блоке 3 после получения стационарного ряда исследуется характер поведения выборочных *АКФ* и *ЧАКФ* и выдвигаются гипотезы о значениях параметров p (порядок авторегрессии) и q (порядок скользящего среднего). При этом следует иметь в виду, что выборочные корреляционные функции могут не демонстрировать детального сходства с теоретическими значениями. Например, умеренно большие значения выборочной *АКФ* могут наблюдаться после затухания теоретической функции, а также могут наблюдаться всплески, не имеющиеся в теоретической функции. Поэтому для идентификации модели нужно использовать только главные черты *АКФ*, не обращая внимания на более тонкие детали. На выходе блока 3 может формироваться базовый набор, включающий 1,2 или даже большее число моделей.

В блоке 4 после осуществления идентификации моделей необходимо оценить их параметры. В современных эконометрических ППП используются разные подходы (МНК, нелинейный МНК, метод максимального правдоподобия (ММП)). Все эти оценки при больших объемах выборок асимптотически эквивалентны.

В блоке 5 для проверки каждой пробной модели на адекватность анализируется ее ряд остатков. У адекватной модели остатки должны быть похожими на белый шум, т.е. их выборочные автокорреляции не должны существенно отличаться от нуля.

При проверке значимости коэффициентов *АКФ* используются два подхода:

• проверка значимости каждого коэффициента автокорреляции отдельно;

• проверка значимости множества коэффициентов автокорреляции как группы.

Первый подход опирается на работу Бартлетта, показавшего, что если модель адекватна исходным данным и ошибки представляют собой белый шум, то распределение коэффициентов автокорреляции приближается к нормальному закону с нулевым математическим ожиданием и дисперсией $1/n$, т.е. к $N(0; 1/n)$. Поэтому если выборочный коэффициент автокорреляции ρ_k с вероятностью P = 1 - α выходит за границы доверительного интервала $\pm\, t_\alpha/\sqrt{n}$, то нулевая гипотеза о равенстве нулю коэффициента ρ_k отвергается.

Второй подход опирается на Q-статистику Бокса—Пирса, позволяющую проверить равенство нулю сразу τ первых значений АКФ остатков.

Q-статистика определяется как

$$Q = n\sum_{k=1}^{\tau}\rho_k^2\,. \tag{8.34}$$

При нулевой гипотезе об отсутствии автокорреляции статистика Q имеет Хи-квадрат распределение с $\nu = \tau - p - q$ степенями свободы, где $p,\ q$ — параметры *ARMA*-модели.

Если $Q > \chi^2_{кр}$, то как группа первые τ коэффициентов автокорреляции значимы (т.е. не все $\rho_1, \rho_2, ..., \rho_\tau$ равны нулю). В некоторых эконометрических пакетах включена модификация этого подхода — тест Бокса-Льюнга. Соответствующая статистика в этом случае определяется выражением:

$$\tilde{Q} = n(n+2)\sum_{k=1}^{\tau}\rho_k^2\,/(n-k)\,. \tag{8.35}$$

$\tilde{Q}$ имеет такое же асимптотическое распределение, как и Q, однако ее распределение ближе к $\chi^2_{кр}$ для конечных выборок. В практических руководствах рекомендуется рассматривать τ = n/4 (но не более 50). Кроме того, при построении модели *ARIMA* необходимо проверить значимость коэффициентов (по t-критерию). При этом модель не должна содержать лишних параметров, т.е. уменьшение числа параметров будет способствовать появлению значимой автокорреляции остатков.

Если в результате проверки несколько моделей оказываются адекватны исходным данным, то при окончательном выборе следует учесть два требования:

• повышение точности (качество подгонки модели);

• уменьшение числа параметров модели.

Для оценки качества модели эконометрических пакетах прикладных программ используются *АКФ* и *ЧАКФ* остаточного ряда, статистические характеристики оценок параметров модели, представительный набор статистических критериев.

Однако окончательный выбор модели все же должен оставаться за исследователем-экономистом, хорошо представляющим предметную область. Иногда можно построить две модели, которые одинаково хорошо соответствуют данным на ретроспективном участке (например, модель с порядком дифференцирования $d=1$ и AR-членами и модель с большим порядком дифференцирования и MA-членами). Однако предлагаемые прогнозы у этих моделей могут существенно различаться. Поэтому окончательный выбор между такими моделями должен опираться на представления исследователя о виде нестационарности исходного ряда, о характере его трендовой составляющей.

Успех применения мощного, гибкого, но, в то же время, сложного аппарата модели *ARIMA* во многом зависит от опыта и квалификации исследователя, а компьютерные программы автоматического выбора вида модели способны только снизить трудоемкость сложных вычислений.

9. ПРОБЛЕМЫ ЭКОНОМЕТРИЧЕСКОГО МОДЕЛИРОВАНИЯ

Характерные особенности методологии эконометрического моделирования, с которыми приходится сталкиваться экономисту при работе с эконометрическими методами, сформулированы одним из ведущих современных исследователей - эконометриков Дэвидом Хендри. следующим образом:

“Проблемы в эконометрии многочисленны и разнообразны. Экономика - это сложный, динамический, многомерный и эволюционирующий объект, поэтому изучать ее трудно. Как общество, так и общественная система изменяются со временем, законы меняются, происходят технологические инновации, поэтому найти в этой системе инварианты непросто. Временные ряды коротки, сильно агрегированы, разнородны, нестационарны, зависят от времени и друг от друга, поэтому мы имеем мало эмпирической информации для изучения. Экономические величины измеряются неточно, подвержены значительным позднейшим исправлениям, а важные переменные часто не измеряются или вообще не могут быть измерены. Поэтому все наши выводы неточны и ненадежны. Экономические теории со временем меняются, соперничающие объяснения сосуществуют друг с другом, и поэтому надежная теоретическая основа для моделей отсутствует. И среди самих эконометристов, по-видимому, нет согласия по поводу того, как следует заниматься их предметом”.

При использовании эконометрических методов исследователь сталкивается с огромным количеством сложных проблем, что делает полученные им результаты недостаточно надежными и спорными. Поэтому должны быть веские причины, побуждающие использовать эти методы. Отметим некоторые из них.

Во-первых, не существует альтернативы статистическим методам в поиске общих закономерностей, связывающих наблюдаемые эмпирические факты. Поскольку любое измерение связано со случайными ошибками, то применение статистики неизбежно, кроме тех случаев, когда ошибки пренебрежимо малы. Результаты экономических измерений нельзя отнести к разряду точных чисел. Кроме того, многие экономические факторы являются ненаблюдаемыми и не измеряемыми, либо данные о них отсутствуют, и эти ненаблюдаемые величины приходится считать случайными. Таким образом, хочет этого исследователь или нет, он всегда получает оценки, имеющие некоторое случайное распределение. Если не использовать статистические методы, то вероятностные свойства получаемых оценок остаются неизвестными.

Предположим, что исследователь получил оценку определенной величины, которая оказалась положительным числом. Может ли он быть уверенным, что полученный результат значим, а не является случайным совпадением? Эконометрические методы позволяют формально проверить гипотезу о знаке полученной оценки.

Не будет преувеличением утверждать, что любой вывод о наличие наблюдаемой закономерности делается, осознано или неосознанно, исходя из того, насколько вероятно было бы такое сочетание данных при отсутствии связи. Если такое случайное совпадение представляется маловероятным, тогда с уверенностью делается вывод о наличии связи. Но описанная процедура есть ни что иное, как неформальное описание проверки статистической гипотезы определенного вида. И, по-видимому, предпочтительнее применять менее субъективные формальные процедуры, что переводит получаемые результаты на универсальный язык статистики и стандартизирует их.

Чрезвычайную важность применения унифицированных процедур демонстрирует нам история физики. В экспериментальной физике этот процесс выразился в использовании стандартизированного лабораторного оборудования, в использовании стандартных методов проведения экспериментов и измерений. По сути дела речь шла о возникновении широко принятой культуры проведения экспериментов и измерений. Экономика точно также нуждается в унификации процедур, применяемых при обработке данных и моделировании.

Во-вторых, применение эконометрии не исключает применения других прикладных методов. Эконометрические методы следует рассматривать скорее как своеобразный инструмент, чем как самостоятельное прикладное моделирование.

Эконометрические методы позволяют не только получить оценки, но и проверить гипотезы, лежащие в основе используемой модели экономического явления, выявить, какие гипотезы нарушаются, и указать, в каком направлении следует модифицировать модель. Большинство серьезных проблем, с которыми приходится сталкиваться при эконометрическом моделировании, на самом деле носят общий характер для экономической науки, поскольку они связаны с характером самого предмета исследований — экономических процессов.

Невозможность использования эконометрических методов и применение вследствие этого каких-то других методов является с этой точки зрения проблемой, а не достоинством.

Другой вопрос заключается в том, следует ли вообще моделировать экономические процессы. Можно предложить следующий аргумент: люди всегда, рассуждая об экономике, держат в голове некоторую неформализованную “модель” или “теорию”. Отличие ученого-экономиста заключается в том, что он в явном виде выписывает ту модель, с которой работает. Поэтому может выявить, на каких именно предположениях он основывается, может увидеть слабые места своей концепции.

Постоянство механизмов

Одно из условий, на которое опирается эконометрическое моделирование, состоит в том, что функциональное соотношение не меняется в течение рассматриваемого периода. Однако это условие часто нереалистично, особенно в случае, когда приходится иметь дело с переходной экономикой. Это обычная проблема, с которой экономист сталкивается при исследовании экономических явлений и процессов с изменчивой структурой. Как бы то ни было, приходится делать предположение о неизменности формы модели, иначе моделирование не было бы возможно.

Один из возможных способов учета структурных сдвигов состоит в использовании различного рода *сконструированных* переменных, таких как, фиктивные переменные и тренды. Включение в эконометрическую модель трендов позволяет учитывать изменения во всех коэффициентах регрессионного уравнения: свободном члене и коэффициентах при “экономических” переменных. Фиктивные переменные (принимающие только два значения — 0 и 1) позволяют учесть резкие структурные скачки.

Кроме того, использование фиктивных переменных и гармонических трендов (синусов и косинусов) позволяет учесть в модели сезонные колебания. Если предположить, что сезонность имеет детерминированный характер, то ее можно смоделировать, добавив в уравнение регрессии компоненту следующего вида:

$$\mathrm{d}_1 M_1 + \mathrm{d}_2 M_2 + \ldots + \mathrm{d}_{12} M_{12}.$$

Здесь $M_1, \ldots, M_{12}$ — сезонные месячные переменные.

Все же эти методы не позволяют адекватно учесть изменения, если неизвестен их характер или момент изменения (в случае скачка). Особенно большие проблемы создают структурные сдвиги для прогнозирования. Если резкое изменение в параметрах экономического явления или процесса произошло в течение исследуемого периода, то это изменение можно заметить и учесть в модели. Если же неожиданное изменение произойдет после исследуемого периода, то сделанные прогнозы окажутся неверными.

Недостаточный набор данных

Имеющихся реальных данных может быть недостаточно для того, чтобы определить функциональную связь между переменными, либо они недостаточно варьируются, чтобы можно было отделить влияние одного фактора от влияния другого. Последняя проблема получила в эконометрическом моделировании название “мультиколлинеарности”. В отличие от естественных экспериментальных наук, у отдельного исследователя, изучающего социально-экономические явления и процессы, как правило, нет возможности сколько-нибудь заметно на них повлиять.

Чтобы восполнить недостаток данных реальных наблюдений или измерений, исследователю приходится делать некоторые априорные допущения, зачастую недостаточно обоснованные. Как правило, функциональная форма модели заранее неизвестна. В этом случае хорошим выходом из положения было бы использование непараметрических методов оценивания. Однако для применения таких методов необходим значительно больший набор данных. Поэтому на практике, как правило, предполагают, что зависимость между двумя переменными линейна. Часто линейная зависимость дает хорошую аппроксимацию гладкой зависимости в некоторой небольшой окрестности, но, вообще говоря, нет никакой гарантии, что “истинная” зависимость не окажется сильно нелинейной как раз в том интервале, к которому относятся данные.

При применении статистических методов следует помнить, что постулируемые свойства, как правило, носят асимптотический характер, то есть проявляются в пределе, при стремлении количества наблюдений к бесконечности. В частности, если в линейной регрессии в качестве регрессоров используются лаги зависимой переменной, то, даже если выполнены стандартные предположения регрессионного анализа, полученные оценки будут состоятельными, но окажутся смещенными.

Проблема ложной регрессии

Для того, чтобы получить высокий коэффициент детерминации, достаточно, чтобы в зависимой переменной и в регрессоре имелся тренд и динамика трендов до некоторой степени совпала. Коэффициент детерминации, как правило, оказывается высоким в регрессии одного растущего показателя по другому растущему показателю.

С другой стороны, коэффициент детерминации, как правило, бывает низким в регрессии одного процесса типа “белый шум” по другому такому же процессу.

Двумя основными причинами наличия “тренда” во временных рядах являются:

- детерминированная составляющая (тогда говорят о детерминированном тренде),
- нестационарность (тогда говорят о стохастическом тренде).

Наличие детерминированного тренда может приводить к появлению ложной регрессии. Пусть, например Y_t и X_t порождаются процессами $Y_t = a + b_t + \mathrm{e}_t$,

$X_t = c + d_t + x_t$, где e_t, x_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.

Если существует стационарная линейная комбинация нестационарных случайных процессов, то эти процессы называют коинтегрированными. Коинтегрированность гарантирует (по крайней мере, асимптотически, то есть для больших выборок), что не возникнет ложная регрессия. Теория коинтеграции — это новый, быстро развивающийся раздел современной эконометрики.

Для оценивания моделей с нестационарными, но коинтегрированными переменными, вообще говоря, следует использовать специальные методы. К сожалению, методы оценивания коинтеграционных регрессий сложны с точки зрения реализации, и способы проверки их спецификации плохо разработаны. Поэтому, несмотря на указанные недостатки, обычный метод наименьших квадратов остается наиболее мощным инструментом эконометрики.

Причинный анализ в эконометрике

Особенной осторожности требует причинная интерпретация полученных закономерностей. Достаточно часто случается, что один исследователь получает результаты, которые он интерпретирует как причинное воздействие одной переменной на другую. Но находятся оппоненты, которые указывают, что те же результаты можно получить вследствие обратного причинного воздействия, либо воздействия на эти переменные третьего фактора. При этом они приводят достаточно веские теоретические доводы.

Одним из самых известных примеров является количественная теория денег. Ее слабой стороной является предположение об экзогенности эмиссии. По тем же причинам положительную связь между темпом инфляции и политической нестабильностью можно интерпретировать двояко: высокая инфляция может вызывать политическую нестабильность, но с другой стороны, можно утверждать, что политическая нестабильность часто приводит к возникновению инфляции.

Приведем пример альтернативной интерпретации, говорящей о возможном влиянии неучтенного третьего фактора. Рядом исследователей получена сильная отрицательная корреляция между темпом инфляции в стране и темпом экономического роста в ней. Из этого делается вывод о том, что инфляция отрицательно влияет на экономический рост. Альтернативная интерпретация заключается в том, что неблагоприятные шоки предложения могут вызывать как усиление инфляции, так и сокращение роста. Кроме того, правительства, проводящие политику, отрицательно влияющие на темпы экономического роста, — такую как протекционизм, большие бюджетные дефициты, — также, скорее всего, проводят и политику, стимулирующую инфляцию.

Особенно большую сложность в интерпретации причинных связей в экономике создает эффект ожиданий. Здесь наиболее наглядно проявляется недостаточность рассуждений по правилу “после этого, значит, вследствие этого”. Учитывая указанные трудности, можно утверждать, что любое суждение о причинности в экономике очень субъективно и опирается на множество гипотез, в правильности многих из которых нельзя быть полностью уверенным.

Классические направления математической статистики — корреляционный и регрессионный анализ — практически с момента своего появления широко применяются для выявления самых разнообразных причинных связей. Даже если отвлечься от эффектов ложной корреляции и ложной регрессии и принять, что связь действительно существует, то все равно следует понимать, что полученные оценки сами по себе не могут говорить ничего о виде и направлении связи. На уравнения регрессии нельзя смотреть шаблонно как на структурные уравнения экономических моделей, непосредственно представляющие причинные процессы.

Выводы из полученных оценок можно делать, только опираясь на предварительные допущения и гипотезы, вытекающие из существующей экономической теории, то есть на положения, не вытекающие из рассматриваемых наблюдений. Так корреляционный анализ статистических данных устанавливает только наличие связи, а когда исследователь-экономист делает выводы о направлении ее, экономические аргументы становятся решающими.

10. ПОРЯДОК ВЫПОЛНЕНИЯ И ОФОРМЛЕНИЯ ТИПОВОГО ПРИМЕРА.

Необходимо исследовать зависимость объясняемой переменной Y от факторных переменных X_1 и X_2, используя множественный регрессионный анализ. Для этого необходимо выполнить:

1. Вычисление средних величин (выполнение на компьютере, остальные разделы – с помощью калькулятора).
2. Вычисление коэффициентов парной корреляции.
3. Вычисление стандартизованных коэффициентов множественной линейной регрессии и ранжирование с их помощью факторных переменных.
4. Обоснование формы и оценка параметров линейной множественной регрессии.
5. Построение множественной линейной регрессии в естественной форме.
6. Вычисление стандартной ошибки регрессии.
7. Вычисление частных коэффициентов корреляции и частных коэффициентов эластичности.
8. Вычисление коэффициента множественной регрессии и индекса множественной корреляции.
9. Проверка гипотезы о статистической значимости полученного уравнения множественной регрессии
10. Вычисление доверительных интервалов параметров регрессии при уровне значимости $\alpha = 0{,}1$ *или* $0{,}05$.

Каждый из них должен включать:

а) описание предметной области, изложение теоретических предпосылок и обоснование расчетных формул;

б) выполнение расчетов;

в) обсуждение полученных результатов.

Имеются следующие выборочные данные:

№ п/п	X_1	X_2	Y
1	28	-69	298
2	30	-67	310
3	29	-62	295
4	30	-77	316
5	31	-78	311
6	29	-63	308
7	30	-72	309
8	30	-79	306

№ п/п	X_1	X_2	Y
25	28	-69	298
26	30	-67	310
27	29	-62	295
28	30	-77	316
29	31	-78	311
30	29	-63	308
31	30	-72	309
32	30	-79	306

9	29	-74	307
10	29	-68	293
11	30	-67	303
12	29	-72	300
13	30	-73	311
14	29	-66	302
15	29	-71	299
16	29	-66	304
17	29	-76	309
18	29	-73	297
19	30	-67	315
20	29	-69	297
21	29	-59	308
22	30	-76	304
23	31	-78	318
24	32	-82	337
33	29	-74	307
34	29	-68	293
35	30	-67	303
36	29	-72	300
37	30	-73	311
38	29	-66	302
39	29	-71	299
40	29	-66	304
41	29	-76	309
42	29	-73	297
43	30	-67	315
44	29	-69	297
45	29	-59	308
46	30	-76	304
47	31	-78	318
48	32	-82	337

где X_1 и X_2 - объясняющие или факторные переменные; Y- объясняемая переменная; n = 48 - число выполненных измерений, или объем выборки.

Исследование должно включать в себя наряду с получением оценок статистических характеристик многомерной пространственной выборки обоснование применяемых методов их вычисления, а также элементы экономического анализа.

Выполнение численных расчетов необходимо сопровождать обсуждением экономического содержания полученных результатов. В конце работы на отдельном листе прилагается список использованной методической и учебной литературы.

Вычисления всех средних величин, которые необходимы для проведения последующего анализа, можно оформить в форме таблицы Microsoft Exel:

Таблица 1

№	x_{i1}	x_{i2}	y_i	x_{i1}^2	x_{i2}^2	y_i^2	$x_{i1}x_{i2}$	$x_{i1}y_i$	$x_{i2}y_i$
Σ	$\sum x_{i1}$	$\sum x_{i2}$	$\sum y_i$	$\sum x_{i1}^2$	$\sum x_{i2}^2$	$\sum y_i^2$	$\sum x_{i1}x_{i2}$	$\sum x_{i1}y_i$	$\sum x_{i2}y_i$
Σ/n	$\overline{X}_1$	$\overline{X}_2$	$\overline{Y}$	$\overline{X_1^2}$	$\overline{X_2^2}$	$\overline{Y^2}$	$\overline{X_1X_2}$	$\overline{X_1Y}$	$\overline{X_2Y}$

На основе средних величин получают следующие показатели:

1. Статистические дисперсии:

$S_Y^2 = \overline{Y^2} - \overline{Y}^2 = 93953{,}00 - 306{,}21^2 = 189{,}457$;

$S_{X_1}^2 = 1{,}423$; $S_{X_2}^2 = 38{,}177$;

Стандартные (среднеквадратические) отклонения:

$$S_{X_1} = \sqrt{1{,}423} = 1{,}19;\ S_{X_2} = 6{,}18;\ S_Y = 13{,}76.$$

Ковариации:

$$S_{YX_1} = \overline{Y \cdot X_1} - \overline{Y} \cdot \overline{X_1} = 9105{,}23 - 29{,}69 \cdot 306{,}21 = 14{,}67;$$

$$S_{YX_2} = -50{,}83; \quad S_{X_1X_2} = -4{,}51.$$

2. Линейные коэффициенты парной корреляции:

$$r_{YX_1} = 0{,}8934;\ r_{YX_2} = -0{,}5977;\ r_{X_1X_2} = -0{,}6177.$$

Анализируя значения коэффициентов парной корреляции, необходимо сделать следующие выводы:

Связанные переменные	**Значения коэффициентов**	**Теснота связи**	**Направление связи**
X_1 и Y	0,8934	сильная	прямая
X_2 и Y	-0,5977	умеренная	обратная
X_1 и X_2	-0,6177	умеренная	обратная

3. Стандартизованные коэффициенты регрессии:

$$\hat{\beta}_1 = 0{,}8433;\ \hat{\beta}_2 = -0{,}0819.$$

Сравнивая модули значений $\hat{\beta}_1$ и $\hat{\beta}_2$, приходим к выводу, что сила влияния фактора X_1 на объясняемую переменную Y намного больше, чем сила влияния фактора X_2.

4. Выборочные оценки значений коэффициентов множественной линейной регрессии:

$$\hat{b}_1 = 9{,}703;\ \hat{b}_2 = -0{,}182;$$

$$\hat{b}_0 = \overline{Y} - \hat{b}_1 \overline{X_1} - \hat{b}_2 \overline{X_2} = 306{,}21 - 9{,}73 \cdot 29{,}69 + 0{,}18 \cdot 69{,}90 = 30{,}908.$$

5. Уравнение множественной линейной регрессии в естественной форме:

$$\hat{Y} = \hat{b}_0 + \hat{b}_1 \cdot X_1 + \hat{b}_2 \cdot X_2 = 4{,}614 + 9{,}730 \cdot X_1 - 0{,}182 \cdot X_2$$

6. Вычисление стандартной ошибки регрессии S_e:

Таблица 2

№ п/п	x_{i1}	x_{i2}	Y_i	$\hat{Y}_i$	$e_i = Y_i - \hat{Y}_i$	e_i^2
1	28	-69	298	289,63	8,37	70,12
2	30	-67	310	308,72	1,28	1,64
3	29	-62	295	298,08	-3,08	9,48
4	30	-77	316	310,54	5,46	29,76
5	31	-78	311	320,46	-9,46	89,42
6	29	-63	308	298,26	9,74	94,84
7	30	-72	309	309,63	-0,63	0,40
8	30	-79	306	310,91	-4,91	24,10
9	29	-74	307	300,27	6,73	45,32
10	29	-68	293	299,17	-6,17	38,11
11	30	-67	303	308,72	-5,72	32,73
12	29	-72	300	299,90	0,10	0,01
13	30	-73	311	309,81	1,19	1,40
14	29	-66	302	298,81	3,19	10,18
15	29	-71	299	299,72	-0,72	0,52
16	29	-66	304	298,81	5,19	26,95
17	29	-76	309	300,63	8,37	70,02
18	29	-73	297	300,09	-3,09	9,52
19	30	-67	315	308,72	6,28	39,43
20	29	-69	297	299,36	-2,36	5,55
21	29	-59	308	297,53	10,47	109,58
22	30	-76	304	310,36	-6,36	40,48
23	31	-78	318	320,46	-2,46	6,03
24	32	-82	337	330,92	6,08	37,02
25	30	-79	316	310,91	5,09	25,92
26	30	-64	302	308,17	-6,17	38,11
27	31	-77	320	320,27	-0,27	0,08
28	29	-64	291	298,44	-7,44	55,41
29	29	-71	290	299,72	-9,72	94,49
30	32	-74	324	329,46	-5,46	29,77
31	29	-63	304	298,26	5,74	32,93
32	29	-68	292	299,17	-7,17	51,46
33	29	-63	309	298,26	10,74	115,31
34	30	-61	303	307,63	-4,63	21,40
35	30	-67	307	308,72	-1,72	2,96
36	31	-76	329	320,09	8,91	79,36
37	32	-73	336	329,27	6,73	45,24
38	31	-77	310	320,27	-10,27	105,56
39	30	-64	301	308,17	-7,17	51,46

40	27	-63	270	278,80	-8,80	77,48
41	32	-73	338	329,27	8,73	76,14
42	28	-60	283	287,98	-4,98	24,85
43	30	-63	305	307,99	-2,99	8,95
44	32	-82	332	330,92	1,08	1,18
45	30	-70	304	309,27	-5,27	27,75
46	28	-60	293	287,98	5,02	25,15
47	29	-73	299	300,09	-1,09	1,18
48	27	-66	283	279,35	3,65	13,33
Σ	1425	-3355	14698	14698,00	0,00	1798,07
Σ/n	29,69	-69,90	306,21	306,21	0,00	37,46

$$S_e^2 = \overline{e^2} - \bar{e}^2 = 37{,}46 - 0 = 37{,}46; \qquad S_e = \sqrt{37{,}46} = 6{,}12.$$

Частные коэффициенты линейной корреляции:

$$r_{X1Y/X2} = 0{,}832; \; r_{X2Y/X1} = -0{,}144; \; r_{X1X2/Y} = -0{,}216$$

Сравнение с парными коэффициентами корреляции показывает, что существует слабая взаимосвязь между факторными переменными. Но она сильно проявилась на связи Y и X_2, уменьшив характеристику связи с -0,598 до -0,144. На связь Y и X_1 она повлияла незначительно (0,893 и 0,832 соответственно).

7. Средние частные коэффициенты эластичности:

$$\overline{Эл}_{y/x_1} = 90{,}944, \quad \overline{Эл}_{y/x_2} = 0{,}041..$$

При увеличении фактора X_1 на 1% от его среднего уровня объясняемая переменная Y возрастает на 0,944 % от среднего уровня. При увеличении фактора X_2 на 1% от его среднего уровня объясняемая переменная Y возрастает на 0,041 % от своего среднего уровня.

8. Характеристика совокупного влияния всех факторов на объясняемую переменную - коэффициент множественной корреляции R_{YX1X2}:

- для уравнения регрессии в стандартизованной форме:

$$R_{yx1x2} = \sqrt{\beta_1 \cdot r_{x1y} + \beta_2 \cdot r_{x2y}} = 0{,}8957$$

- как корень квадратный из коэффициента множественной детерминации:

$$R_{yx1x2} = \sqrt{R^2_{yx1x2}} = \sqrt{\frac{S^2_{регр}}{S^2_y}} = \sqrt{\frac{S^2_y - S^2_e}{S^2_y}} = \sqrt{1 - \frac{S^2_e}{S^2_y}} = 0{,}8957.$$

На основании значения коэффициента множественной регрессии делается вывод, что зависимость объясняемой переменной y от факторов x_1 и x_2 характеризуется как тесная, в которой 80,2% вариации y определяются вариацией данных факторов. Прочие факторы, не включенные в модель, составляют соответственно только 19,8 % от общей вариации y.

9. Оценивание качества уравнения регрессии с помощью F-критерия Фишера состоит в проверке гипотезы H_0 о статистической значимости уравнения регрессии или показателя тесноты связи. Для этого сравнивают фактическое значение критерия $F_{эмп}$ с критическим, табличным значением $F_{табл}$.

$$F_{эмп} = \frac{R^2_{yx1x2}}{1-R^2_{yx1x2}} \cdot \frac{k_1}{k_2} == \frac{0,8957}{1-0,8957} \cdot \frac{48-2-1}{2} = 193,22,$$

где p - число объясняющих переменных, или факторов, включенных в модель, k_1 и k_2 - число степеней свободы большей (числителя) и меньшей (знаменателя) дисперсий соответственно.

$$F_{табл} = F_{\alpha,k1,k2} = F(\alpha, k_1, k_2) = F(0,05, 2, 45) = 3,25,$$

где $\alpha = 0,05$ - уровень значимости, $k_1 = p = 2$ и $k_2 = n - p - 1 = 45$.

Так как в нашем случае $F_{эмп} > F_{табл} (193,22 > 3,25)$, то гипотеза H_0 о случайной природе статистической связи отклоняется. Имеющиеся статистические данные свидетельствуют о том, что в 95% случаев связь обусловлена влиянием факторов регрессии, а остальные не включенные в нее факторы являются статистически не значимыми. С вероятностью 0,95 делаем заключение о статистической значимости уравнения регрессии в целом и показателя тесноты связи R_{yx1x2}, которые сформировались под неслучайным воздействием факторов X_1 и X_2.

10. Найдем доверительные интервалы для коэффициентов регрессии $\hat{b}_0$, $\hat{b}_1$ и $\hat{b}_2$. Матрица $(X^TX)^{-1}$, необходимая оценки дисперсии параметров регрессии, вычисляется по данным таб.1:

$$(X^T \cdot X)^{-1} = \begin{pmatrix} 13,4317 & -0,5013 & -0,0210 \\ 0,5014 & 0,0234 & 0,0028 \\ 0,0211 & 0,0028 & 0,0009 \end{pmatrix}.$$

Величина критерия Стьюдента определяется с помощью статистической таблицы по входным параметрам: α - уровень значимости критерия и ν - число степеней свободы.

$$t_{\alpha,\nu} = t(0,05, 45) = 2,027.$$

где $\nu = n - p - 1$.

Оценки дисперсий, стандартных и предельных ошибок, а также доверительных интервалов коэффициентов регрессии выполняют следующим образом:

1) $S^2_{\hat{b}0} = S^2_e \cdot ((X^T \cdot X)^{-1})_{11} = 37{,}46 \cdot 13{,}4317 = 503{,}151$, $S_{\hat{b}0} = 22{,}431$

Предельная ошибка $\Delta\hat{b}_0$ выборочной оценки коэффициента регрессии b_0:

$$\Delta\hat{b}_0 = S_{\hat{b}0} \cdot t_{\alpha,\nu} = 22{,}431 \cdot 2{,}027 = 45{,}468$$

Таким образом, интервальная оценка коэффициента b_0:

$$\hat{b}_0 - \Delta\hat{b}_0 < b_0 < \hat{b}_0 + \Delta\hat{b}_0$$

$$30{,}908 - 45{,}468 < b_0 < 30{,}908 + 45{,}468.$$

$$-14{,}560 < b_0 < 76{,}376$$

2) Для интервальной оценки коэффициента регрессии b_1

$S^2_{\hat{b}1} = S^2_e \cdot ((X^T \cdot X)^{-1})_{22} = 37{,}46 \cdot 0{,}0234 = 0{,}877$, $S_{\hat{b}1} = 0{,}936$,

предельная ошибка выборки $\Delta\hat{b}_1 = S_{\hat{b}1} \cdot t_{\alpha,n-p-1} = 0{,}936 \cdot 2{,}027 = 1{,}898$.

Аналогичным образом имеем: $\hat{b}_1 - \Delta\hat{b}_1 < b_1 < \hat{b}_1 + \Delta\hat{b}_1$,

$$7{,}832 < b_1 < 11{,}628.$$

3) Для интервальной оценки коэффициента регрессии b_2

$S^2_{\hat{b}2} = S^2_e \cdot ((X^T \cdot X)^{-1})_{33} = 37{,}46 \cdot 0{,}00087 = 0{,}0326$, $S_{\hat{b}2} = 0{,}181$

предельная ошибка выборки $\Delta\hat{b}_2 = S_{\hat{b}2} \cdot t_{\alpha,n-p-1} = 0{,}181 \cdot 2{,}027 = 0{,}366$.

Аналогичным образом имеем: $\hat{b}_2 - \Delta\hat{b}_2 < b_2 < \hat{b}_2 + \Delta\hat{b}_2$,

$$-0{,}548 < b_2 < 0{,}184.$$

11. ВАРИАНТЫ КОНТРОЛЬНЫХ ЗАДАНИЙ

Вариант 1

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
67	-201	602	82	-237	730	78	-226	697
68	-199	602	70	-209	626	75	-218	672
70	-206	625	83	-243	737	82	-245	729
76	-221	674	80	-239	716	85	-247	763
80	-238	718	76	-221	681	68	-201	608
87	-256	781	81	-238	721	72	-215	643
75	-222	668	80	-238	710	71	-209	637
79	-230	702	76	-223	680	72	-212	638
79	-234	701	70	-207	626	68	-203	611
73	-217	648	79	-237	708	86	-252	773
86	-253	773	74	-216	658	85	-251	764
78	-228	692	77	-228	690	71	-206	632
79	-230	708	65	-193	579	72	-214	641
67	-201	596	80	-234	714	76	-227	675
79	-237	702	79	-229	704	82	-239	735
74	-221	664	72	-210	640	67	-196	599
74	-214	659	82	-243	736	73	-214	647
87	-257	777	86	-251	773	95	-279	847
77	-227	683	71	-207	635	88	-259	782
90	-269	801						

Вариант 2

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
48	99	520	61	128	664	52	112	564
40	83	435	42	88	456	32	67	345
52	106	564	55	117	601	48	97	518
50	107	541	52	110	567	40	82	430
39	79	424	44	94	475	40	81	436
47	100	516	41	82	450	46	92	504
38	80	413	47	97	516	41	90	450
46	96	505	43	91	468	54	110	584
47	98	514	55	118	602	48	96	521
44	97	479	43	87	472	53	110	580
45	92	490	49	104	538	47	97	509
44	90	483	42	89	459	50	102	542
53	108	577	31	71	333	46	101	498
52	107	569	40	86	432	56	112	608
45	96	493	47	97	516	42	93	461
42	86	461	43	89	471	41	84	445
45	98	486	48	101	527	55	112	595
45	97	492	44	93	483	40	80	431
51	108	554	47	99	512	46	101	503
48	100	519	44	93	474	38	83	416
46	92	496	57	120	624	44	89	483
43	89	467	53	107	579	35	79	376

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
49	99	538	36	75	387	54	113	588

<u>Вариант 3</u>

X	**Y**	**Z**	**X**	**Y**	**Z**	**X**	**Y**	**Z**
46	50	411	39	59	524	56	62	500
55	57	491	57	60	511	52	58	464
57	61	508	48	52	422	42	46	373
55	58	491	41	47	367	48	50	423
51	51	454	57	57	507	60	63	539
62	70	552	50	57	449	62	64	554
43	43	377	64	67	567	46	47	413
64	71	567	43	45	381	51	59	450
56	64	496	57	62	507	69	76	620
65	67	580	54	56	478	60	60	531
56	63	502	50	51	444	57	61	511
51	58	458	59	66	529	62	65	551
58	60	516	48	51	427	58	66	518
42	47	374	45	54	404	54	60	483
46	54	405	51	53	450	57	63	506
54	60	485	40	41	358	44	45	392
62	67	554	59	65	530	55	62	494
57	58	512	46	54	404	50	54	445
68	68	610	47	52	417	63	67	566
47	56	421	55	59	486	44	48	391

X	**Y**	**Z**	**X**	**Y**	**Z**	**X**	**Y**	**Z**
69	74	613	49	55	437	51	60	456
65	72	578	64	70	572	47	55	416
56	57	500	43	45	384	67	72	602
57	65	506	55	57	488	60	64	538
73	79	652	40	43	358	53	53	467
54	56	485	67	74	601	57	65	507
57	66	508	57	57	511	41	45	362
63	67	557	48	54	425	58	61	513
58	62	513	46	49	406	54	63	476
57	57	510						

Вариант 4

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
67	207	529	75	225	597	62	195	486
57	171	449	68	207	542	71	220	566
82	246	648	65	201	518	67	205	531
62	191	487	80	241	631	66	207	520
57	172	446	71	219	565	67	201	534
83	255	656	61	183	483	70	215	558
69	213	547	77	239	612	64	197	510
86	261	681	69	209	546	66	206	523
66	203	519	64	201	505	78	241	622
79	242	629	68	208	539	65	197	511

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
82	248	646	52	159	413	63	196	501
73	221	574	59	183	470	69	216	549
65	195	516	63	197	496	59	180	471
55	172	437	56	176	447	71	214	558
61	186	482	62	192	486	59	185	466
64	197	510	67	204	535	64	196	505
59	182	465	73	225	580	68	208	538
77	236	612	65	198	516	71	220	562
49	154	388	70	217	557	76	231	601
63	189	499	74	225	587	77	234	610
71	222	565	73	224	580	52	160	409
69	214	549	77	237	612			

Вариант 5

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
31	318	-37	23	231	-32	28	284	-34
28	280	-36	28	288	-30	26	265	-29
30	305	-32	26	269	-35	29	298	-38
23	234	-30	26	267	-34	26	264	-30
25	255	-32	28	284	-29	23	232	-33
25	258	-33	27	277	-35	27	272	-30
25	258	-27	25	253	-34	29	292	-39
27	272	-28	26	268	-30	24	245	-25

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
31	317	-40	27	270	-35	27	274	-32
25	252	-33	28	289	-30	25	256	-35
28	280	-31	27	278	-35	29	291	-30
25	258	-31	29	295	-36	29	290	-34
30	304	-39	24	246	-33	25	257	-33
28	289	-30	23	238	-24	25	258	-31
28	288	-34	27	279	-29	30	309	-34
31	317	-37	27	273	-32	25	257	-29
22	229	-29	27	273	-31	31	316	-38
27	272	-35	28	280	-38	25	251	-30
26	263	-32	24	244	-32	27	278	-31
24	249	-30	24	243	-32	26	268	-32
25	252	-29	26	269	-28	25	255	-34
22	229	-31						

Вариант 6

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
89	-262	533	88	-256	520	92	-273	550
90	-261	537	96	-288	573	90	-270	537
91	-268	537	92	-269	549	89	-266	530
87	-255	520	92	-276	549	90	-264	531
89	-261	533	91	-266	536	91	-270	536
93	-275	551	85	-248	508	95	-285	562

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
90	-263	537	94	-281	556	92	-268	544
94	-279	558	89	-266	531	92	-268	547
92	-269	550	94	-282	563	91	-272	537
93	-276	548	94	-281	554	91	-269	542
94	-274	560	90	-265	539	92	-273	542
92	-272	551	96	-288	573	97	-283	580
97	-284	575	97	-290	577	92	-269	543
93	-276	555	94	-278	554	96	-284	572
94	-275	556	95	-282	562	92	-270	548
92	-276	547	93	-276	550	93	-271	549
85	-250	501	91	-271	540	92	-271	547
95	-278	566	93	-279	554	97	-289	573
92	-267	543	91	-273	538	99	-294	587
99	-288	586	91	-271	536	92	-271	544
89	-266	526	91	-264	536	91	-271	545
94	-275	562	92	-270	556	93	-279	551
98	-285	587	87	-261	518	93	-277	551
89	-259	530	93	-273	550	94	-280	562
88	-264	520	95	-283	568	90	-263	539
92	-270	544	87	-252	516	97	-289	572
95	-281	564	96	-283	566	95	-278	567
91	-268	541	98	-285	581	94	-279	561
95	-277	568	93	-279	557	91	-265	540

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
93	-273	552	92	-269	551	93	-270	555
90	-262	539	90	-268	534	91	-264	536

Вариант 7

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
45	-177	260	38	-145	223	53	-204	309
38	-146	221	34	-135	203	29	-110	169
58	-226	342	44	-172	254	40	-152	230
46	-184	266	47	-181	275	41	-161	238
40	-152	237	65	-259	387	44	-167	259
41	-155	237	34	-133	196	28	-109	158
58	-230	347	42	-161	244	48	-186	283
41	-155	244	48	-189	285	53	-203	316
51	-199	302	54	-208	318	26	-99	149
59	-227	346	55	-220	323	43	-168	256
62	-244	364	50	-191	291	42	-159	251
44	-172	256	37	-141	218	68	-266	401
47	-184	273	43	-163	252	42	-159	242
43	-170	256	47	-188	279	60	-237	355
47	-183	272	44	-170	254	56	-222	334
55	-218	321	51	-203	300	39	-147	225
61	-240	357	51	-204	299	47	-188	278
47	-180	279	48	-186	286	55	-211	329

X	Y	Z	X	Y	Z	X	Y	Z
58	-232	347	43	-172	256	35	-138	208
40	-158	238	64	-255	381	41	-157	244
47	-183	280	53	-207	308	47	-183	274
46	-184	271	43	-168	252	56	-217	330
42	-162	242	45	-174	268	42	-167	251
50	-198	299	43	-172	252	53	-208	309
45	-171	262	43	-166	249	45	-180	269
43	-170	248	49	-188	287	43	-167	248
42	-161	248	34	-130	201	44	-176	254
59	-232	351	54	-215	314	46	-176	269
42	-163	247	49	-191	293	59	-234	344
39	-150	224	44	-176	256	42	-163	242
33	-123	191	38	-143	222	40	-151	239
64	-256	378	51	-201	304	30	-111	171
41	-162	238						

Вариант 8

X	Y	Z	X	Y	Z	X	Y	Z
46	279	44	57	342	55	56	338	46
40	245	39	45	274	39	48	297	47
58	354	49	55	336	50	44	268	36
35	212	25	46	277	41	49	294	44
53	323	47	53	320	47	50	309	40

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
47	285	42	45	278	44	46	284	36
40	240	32	43	266	38	46	276	41
60	361	53	48	294	43	47	289	41
39	235	29	50	308	47	39	234	31
41	246	39	44	271	35	49	298	43
58	357	54	38	228	33	44	264	37
59	361	55	68	414	66	49	302	45
57	343	50	54	325	44	53	323	51
65	390	55	57	348	54	44	272	41
34	208	28	34	211	28	47	284	43
50	301	42	59	356	52	51	315	41
45	277	36	57	350	50	56	340	52
50	306	47	57	350	50	39	236	33
55	334	52	49	301	41	46	285	38
50	304	44	41	248	37	68	415	59
45	278	39	46	284	36	47	282	39
50	303	44	46	278	38	40	242	38
25	159	23	53	320	45	35	210	29
44	266	43	43	260	34	60	360	53
54	331	47	53	321	48	37	227	35
54	328	48						

Вариант 9

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
77	619	72	74	605	66	67	549	57
76	626	58	77	632	68	69	568	56
62	505	45	69	554	66	66	530	62
83	680	80	72	594	60	71	568	53
83	673	71	76	612	58	71	574	69
79	637	70	69	556	60	77	628	60
77	629	61	84	679	76	78	631	72
81	649	66	77	629	61	77	631	73
71	568	51	72	582	59	67	542	63
83	669	70	76	611	56	73	587	65
73	591	56	78	626	61	67	542	53
81	648	75	87	703	69	83	681	69
82	663	71	79	650	70	67	544	53
77	616	60	80	655	68	76	616	66
82	672	63	70	577	54	71	579	59
83	677	77	80	657	70	74	604	60
80	655	61	75	613	71	73	594	58
77	619	60	69	557	54	75	601	60
69	568	57	77	619	61	80	659	61
78	624	74	76	619	57	72	583	58
65	528	62	81	667	79	66	529	49
71	579	55	69	567	55	80	647	63
69	564	65	83	667	64	72	585	62

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
85	688	68	77	634	70	68	561	61
81	656	62	69	565	62	84	675	77
72	578	56	81	664	80			

Вариант 10

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
72	152	574	69	140	546	71	149	561
72	152	575	70	145	554	73	148	577
73	155	577	73	155	580	69	139	545
68	144	538	68	140	541	73	147	574
71	145	567	70	144	555	68	139	540
77	155	612	69	145	551	67	135	5.35
74	154	584	71	150	563	69	139	544
67	136	532	71	142	567	80	168	634
68	142	536	71	142	566	77	157	615
71	147	561	65	139	517	74	149	584
74	153	590	67	137	532	78	161	615
69	145	548	68	144	542	68	137	535
69	143	548	66	139	520	73	153	580
81	168	645	71	144	558	64	131	506
68	143	536	72	148	572	67	139	530
67	134	534	66	134	522	67	143	531
76	152	607	72	151	571	70	144	551

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
67	139	533	73	150	577	67	142	533
68	141	540	67	134	535	67	137	527
69	146	548	75	155	596	74	153	582
70	146	556	66	139	523	66	139	52!
69	142	549	73	150	582	65	136	510
70	140	559	70	141	555	71	151	566
73	146	583	66	141	524	75	158	596
70	147	557	68	143	534	79	164	626
75	157	597	74	156	583	76	153	606
63	130	498	74	151	586	73	152	576
68	136	543	71	149	562	65	137	514
75	152	596	75	159	599	73	152	581
80	164	634						

Вариант 11

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
7	23	38	6	19	28	6	20	35
11	38	63	5	18	27	10	37	55
10	34	55	10	31	56	14	47	79
4	16	16	10	31	55	12	42	65
11	42	64	13	43	74	6	25	29
12	44	70	15	49	84	6	22	35
9	33	50	6	26	31	8	28	47

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
5	18	22	11	42	58	6	26	28
8	29	44	10	31	54	6	25	31
13	42	76	14	45	78	13	46	77
12	39	71	8	28	47	11	42	59
10	36	58	14	49	79	9	32	45
9	27	45	6	23	35	10	36	58
9	30	49	16	57	95	8	31	45
11	36	65	8	29	39	12	43	66
9	31	44	6	23	33	14	44	75
10	38	55	8	25	41	14	49	76
10	35	54	8	32	40	7	28	38
13	45	75	7	25	39	16	50	90
15	48	89	8	29	38	11	34	63
4	18	22	10	36	59	11	40	59
9	30	48	12	39	69	12	42	70
14	50	76	13	42	70	12	45	64
8	27	40	9	27	52	10	32	58
16	53	92	9	29	49	7	30	33
12	42	65	10	38	57	8	32	46
15	50	81	13	41	77	9	30	46
13	39	75	7	21	32	9	34	50
11	37	62	9	32	50	11	36	63
9	35	45	11	42	58			

Вариант 12

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
40	209	272	40	212	266	41	224	279
40	201	260	39	204	265	40	213	275
42	214	281	41	224	271	41	216	276
41	222	271	42	223	282	40	213	278
42	210	289	41	215	280	41	209	273
41	219	277	42	214	279	42	225	282
40	208	264	40	210	264	42	219	287
41	208	280	41	207	268	40	215	267
41	206	280	42	215	293	40	214	275
39	200	267	40	203	276	41	223	277
38	209	256	41	207	278	40	209	270
39	202	263	40	212	271	41	220	271
42	212	293	42	226	290	41	205	267
40	206	277	40	204	265	40	212	273
40	216	261	42	221	287	41	218	284
41	220	280	41	220	268	39	210	255
41	208	267	40	217	278	40	219	270
40	201	264	41	217	278	40	204	266
41	216	284	40	202	263	41	205	274
39	212	257	41	219	283	41	219	279
41	217	278	39	213	257	40	216	276

Вариант 13

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
48	200	327	52	216	258	65	269	324
46	188	226	65	261	319	70	281	346
62	249	305	65	265	319	58	232	282
63	261	314	49	196	235	70	287	344
51	213	253	69	276	336	63	258	305
61	253	295	55	229	269	60	249	291
62	252	302	47	192	232	77	311	384
56	230	278	61	246	302	63	261	311
50	204	243	53	212	261	49	199	241
54	225	262	58	237	289	59	242	289
53	220	256	68	273	335	65	262	316
47	192	232	48	192	239	65	267	324
66	273	324	52	214	256	61	244	299
58	236	284	52	212	255	64	260	311
58	234	282	60	246	297	66	265	324
70	282	343	54	224	264	58	236	286
59	239	290	67	277	332	63	261	308
62	255	309	57	237	281	71	286	351
51	204	250	61	245	304			

Вариант 14

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
59	420	57	61	435	46	63	460	53
71	509	67	63	457	55	65	462	50
61	435	43	59	422	57	64	456	61
67	469	60	64	454	49	68	485	50
62	449	55	63	458	55	70	490	63
62	450	59	62	449	47	57	409	41
61	437	59	68	486	67	65	472	53
59	422	54	65	468	57	69	502	66
65	463	49	68	478	56	62	436	59
63	455	47	65	463	45	65	457	64
65	472	62	62	441	46	66	475	62
62	448	56	68	491	50	66	474	65
62	443	49	64	450	55	62	452	45
65	462	55	60	432	47	60	435	44
67	484	50	64	453	59	64	465	44
63	442	50	67	478	59	60	431	40
58	419	54	68	481	58	60	432	55
64	456	53	62	438	54	63	446	43
64	451	45	67	487	50	62	444	58
65	465	61	65	473	49	62	437	54
62	439	45	66	464	49	64	458	60
66	466	62	64	458	44	64	453	63
63	453	48	61	431	54	62	443	43

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
57	408	39	62	434	42	56	410	38
55	396	51	69	486	54			

Вариант 15

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
76	304	218	89	364	259	81	326	239
85	340	249	78	318	226	84	338	248
77	315	225	87	353	253	80	326	236
74	303	216	83	332	243	83	338	245
83	341	240	82	331	240	88	361	254
73	297	209	83	332	244	79	317	228
77	316	225	84	341	248	81	325	238
78	318	224	81	324	241	81	332	233
85	348	246	91	364	272	83	339	245
82	329	239	85	343	253	90	368	262
81	325	234	80	320	233	73	300	211
85	349	251	86	345	254	80	323	232
79	320	228	84	345	244	75	303	222
83	337	248	88	355	261	77	313	227
81	330	240	87	356	252	80	320	232
87	357	256	82	333	245	85	341	250
84	337	242	85	347	248	83	338	248
80	321	235	77	310	227	90	362	269

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
81	328	235	79	320	234	84	343	242
81	324	236	81	329	239	78	315	224
74	303	216	90	366	264	75	303	222
85	343	246	86	351	257	85	340	252
80	328	235	75	307	217	77	313	221
75	305	220	79	317	227	93	373	272
83	341	243	93	372	274	78	313	230
75	309	224	84	339	250	74	298	220
71	284	211	82	337	243	78	312	225
77	312	224	85	344	254	80	325	232

Вариант 16

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
25	233	-7	49	443	-9	43	392	-10
47	432	-6	22	201	-6	27	245	-7
42	385	-10	48	440	-2	49	445	-1
38	347	-1	30	274	-2	35	323	-2
30	274	-2	29	261	-8	36	331	-10
30	276	-6	29	270	-9	15	140	-2
32	294	-6	38	345	-2	33	303	-6
38	348	-5	37	334	-8	27	247	-7
34	315	-10	55	498	-9	43	387	-3
34	307	-4	38	342	-7	47	426	-7

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
41	377	-9	27	250	-2	33	304	-6
37	338	-5	44	403	-5	28	257	-9
39	353	-7	48	434	-3	41	377	-5
36	332	-3	35	320	-6	59	535	-1
34	307	-10	35	318	-10	23	208	-5
29	264	-1	28	256	-4	39	357	-4
31	285	-3	39	358	-7	37	337	-3
41	377	-5	38	347	-2	49	447	-2
15	137	-5	35	318	-4	26	238	-7
27	245	-1	40	369	-10	20	182	-3
46	421	-1	44	400	-10	48	437	-9
34	313	-10	30	277	-9	32	292	-10
39	359	-3	43	395	-10	42	386	-3
46	415	-4	24	221	-8	37	338	-5
37	341	-8	31	287	-10	30	277	-4
35	321	-6	21	244	-5	27	247	-4
28	255	-7	34	313	-6	18	170	-5
22	206	-1	24	219	-5	42	380	-7
35	324	-10	43	392	-5	52	473	-8
39	356	-7	49	442	-2	43	392	-1

Вариант 17

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
13	71	88	17	87	117	13	60	76
11	62	72	16	70	110	16	78	102
18	91	123	17	80	117	15	65	97
13	60	90	16	67	97	22	95	149
15	79	93	17	68	117	14	69	85
21	85	135	13	67	82	15	77	95
17	79	112	18	86	112	12	49	64
13	54	75	14	57	88	12	62	74
16	71	110	13	67	71	11	49	63
13	67	74	11	48	60	16	76	109
19	90	115	15	60	102	17	70	105
17	81	116	10	51	69	15	77	92
15	61	87	15	63	93	14	72	83
12	54	76	14	67	93	13	57	82
17	77	118	9	54	59	17	70	101
20	82	120	19	85	127	15	64	91
17	72	107	18	87	107	14	61	91
18	91	115	15	61	101	12	67	78
17	81	116	15	71	92	12	51	72
16	73	102	23	102	153	50	96	134
11	52	64	15	63	93	15	64	94
13	55	75	17	70	105	16	74	110
17	87	101	17	83	102	16	69	101

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
20	88	133	17	76	109	10	40	67
19	85	132	19	86	119	15	61	102

Вариант 18

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
47	482	-103	52	529	-105	56	576	-130
49	504	-111	47	479	-110	48	487	-113
50	513	-109	53	532	-121	52	522	-107
50	500	-111	57	580	-129	50	507	-107
52	526	-124	48	491	-108	51	521	-116
49	493	-101	51	516	-115	51	527	-104
49	505	-114	46	465	-106	53	540	-124
54	552	-125	57	578	-128	49	506	-108
50	503	-114	50	507	-103	50	515	-120
52	529	-109	53	539	-119	51	511	-110
54	546	-122	51	510	-109	59	598	-120
58	590	-124	55	566	-121	52	522	-107
55	550	-122	53	533	-122	51	515	-112
49	491	-109	51	528	-108	52	521	-119
55	567	-116	54	549	-119	51	527	-107
50	512	-117	53	537	-107	52	524	-119
51	512	-105	48	492	-108	51	526	-110
56	564	-127	54	546	-115	52	531	-122

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
50	515	-101	51	527	-120	47	486	-100
53	538	-115	50	500	-118	54	558	-123
49	492	-112	55	559	-127	50	506	-115
54	548	-112	54	556	-128	49	494	-103
59	606	-124	50	503	-116	53	541	-122
52	534	-405	56	563	-132	52	521	-119
57	586	-131	54	547	-128	55	561	-123
54	554	-124	50	516	-107	51	519	-104
52	537	-121	48	496	-110	55	556	-122
50	510	-108						

Вариант 19

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
42	299	205	41	295	199	34	247	161
52	373	250	41	289	196	53	374	255
46	325	227	36	254	176	41	296	200
39	277	187	39	273	191	50	354	243
54	383	267	32	224	152	21	148	99
41	291	198	45	320	218	51	359	245
47	331	229	46	331	229	40	284	194
34	240	168	36	260	170	46	323	222
62	435	300	40	281	197	50	355	242
46	327	229	52	373	257	48	337	233

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
35	251	173	36	252	175	28	198	136
43	310	205	56	401	274	55	386	271
21	152	96	41	288	195	60	420	294
44	310	212	41	296	198	48	342	231
28	198	136	37	265	176	45	320	218
52	365	250	52	371	250	47	329	228
59	416	288	43	306	213	40	283	196
59	413	286	41	296	199	58	414	288
50	353	240	42	294	209	57	400	279
57	404	276	53	377	262	35	252	168
58	409	281	41	293	197	34	238	162
41	288	196	24	174	113	43	301	209
41	288	204	48	337	237	33	239	157
43	310	205	48	340	230	33	233	160
45	322	218	39	279	193	43	302	208
52	369	254	52	372	250	45	322	221
38	273	182	42	298	208	57	399	276
43	302	214	40	287	198	35	254	174
45	323	219	49	343	243	48	345	237
38	272	188	42	302	203	58	408	287
41	287	202	33	240	160	34	239	164
30	216	145	41	295	197	45	316	219

Вариант 20

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
28	298	-69	29	308	-59	27	270	-63
29	291	-73	30	310	-67	30	304	-76
30	301	-68	30	308	-71	29	295	-62
28	283	-60	32	326	-65	28	293	-69
32	337	-82	30	305	-63	30	307	-74
31	311	-78	30	316	-79	32	332	-82
29	306	-68	29	308	-63	30	302	-64
30	307	-71	31	328	-73	30	309	-72
28	293	-60	30	300	-62	29	303	-64
29	291	-64	29	299	-73	30	303	-66
29	307	-74	29	290	-71	27	283	-66
30	306	-73	29	293	-68	32	324	-74
29	298	-66	30	303	-68	30	303	-67
30	310	-72	29	303	-78	29	297	-77
29	292	-68	28	294	-71	31	310	-80
30	311	-73	29	309	-63	29	309	-73
28	298	-58	29	302	-66	30	303	-61
30	315	-75	27	286	-66	29	299	-71
29	303	-63	27	274	-58	30	303	-70
31	329	-76	28	289	-67	29	293	-76
29	309	-76	32	336	-73	31	328	-75
30	306	-77	29	297	-73	31	310	-77

X	Y	Z	X	Y	Z	X	Y	Z
29	292	-69	31	327	-80	30	315	-67
29	297	-63	33	343	-83	30	306	-76
30	307	-67	30	305	-80	30	216	-77
29	304	-66	30	312	-62	30	306	-69
30	316	-67	29	304	-63	30	318	-62
29	307	-71	29	300	-72	30	304	-70
28	289	-62	28	284	-58	31	320	-77
30	301	-64	30	304	-80	30	306	-79
29	297	-69	30	303	-64	29	294	-66
31	318	-78	32	338	-73	30	301	-69

Вариант 21

X	Y	Z	X	Y	Z	X	Y	Z
17	117	97	15	97	80	16	107	77
16	103	94	18	112	104	16	113	84
17	108	97	16	98	86	16	111	94
16	102	90	18	114	94	17	120	91
15	91	75	18	113	91	18	127	103
17	110	99	15	99	70	16	100	94
17	111	88	18	121	99	17	110	97
16	111	82	17	103	85	14	86	79
17	109	83	18	120	101	17	110	90
16	96	81	15	95	85	15	101	76

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
16	96	85	17	116	100	18	110	94
18	108	95	17	116	84	17	112	84
15	97	73	16	112	80	16	99	77
16	108	82	18	117	105	18	108	94
16	102	93	16	104	88	15	90	76
17	109	89	18	121	100	16	110	92
16	98	77	18	112	103	17	121	94
15	95	79	15	98	76	16	96	91
18	127	98	16	111	89	15	99	80
15	96	85	17	120	100	15	108	73
16	106	93	19	128	107			

Вариант 22

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
37	159	58	41	182	72	42	176	72
39	174	63	31	127	51	39	173	60
30	133	59	44	191	71	34	138	56
32	137	59	38	159	63	36	154	71
40	173	67	28	126	55	31	133	52
34	145	53	42	181	81	29	125	44
33	150	53	31	132	46	41	182	77
40	162	69	26	116	42	27	109	47
38	157	68	32	140	58	27	124	51

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
35	140	55	31	143	59	37	154	66
37	157	67	40	163	61	38	157	72
35	149	53	31	126	43	47	198	90
26	114	37	34	139	53	40	176	67
29	118	45	45	185	74	33	149	65
37	166	73	36	161	65	36	154	62
31	134	50	38	153	63	36	150	69
38	154	66	35	150	62	41	179	81
38	159	59	32	129	51	41	172	75
33	151	54	37	157	63	35	150	52
36	144	68	40	161	67	39	169	70
37	158	58	31	134	59	33	133	49
48	202	93	32	146	58	35	147	56
28	129	50	38	154	60	24	115	31
44	178	73	37	167	71	37	160	69
42	180	75	41	182	81	42	168	78
35	142	63	33	139	60	38	153	63
33	136	60	32	128	58	35	159	69
35	141	61	42	178	72	38	162	74
40	169	73						

Вариант 23

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
58	122	339	63	133	373	62	138	365
62	138	353	68	148	396	62	137	352
65	137	377	64	144	375	66	141	378
61	133	349	70	154	419	73	160	429
66	132	394	63	140	372	70	151	412
67	143	385	67	153	396	64	131	373
67	137	395	65	148	385	69	141	407
65	131	389	71	150	423	70	157	405
62	142	353	64	145	376	59	121	334
68	138	405	62	124	368	65	145	385
64	143	365	64	135	383	65	149	377
55	114	310	72	160	421	64	137	376
68	136	395	66	135	394	64	130	379
64	130	368	63	139	362	63	144	358
63	143	361	67	151	386	68	148	388
64	137	383	59	123	338	67	145	396
67	138	401	69	139	404	63	131	373
69	140	398	67	134	388	63	143	372
70	153	401	69	148	405	68	145	404
68	152	388	64	129	364	77	158	450
69	153	408	64	130	374	62	138	353
68	149	388	65	144	376	59	123	343
70	158	417	65	147	372	67	152	387

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
63	126	375	61	128	360	63	139	374
67	149	388	64	137	365	58	131	340
75	158	446	64	137	369	63	128	363
63	130	375	64	140	364	65	148	373
60	137	347	64	147	374	70	151	417
73	159	432	67	148	386	58	130	331
73	150	420	64	140	374	62	138	368
66	138	379	65	145	371	66	140	384
61	137	359	72	151	420			

Вариант 24

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
70	702	471	70	715	477	72	721	487
72	739	492	73	730	509	77	788	528
75	760	506	71	712	488	75	755	524
68	699	464	71	716	484	70	719	482
68	688	457	69	700	474	72	434	486
71	713	478	71	722	493	72	725	493
69	700	475	66	662	454	70	717	470
71	719	490	74	752	514	76	779	529
69	706	480	76	765	522	80	800	553
68	696	457	78	798	543	65	662	443
68	696	470	74	746	507	70	710	475

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
69	703	468	79	797	534	70	718	476
75	755	515	78	787	544	70	702	479
83	838	578	72	737	484	80	804	541
73	748	508	70	708	476	65	655	439
71	715	493	75	753	516	72	739	499
82	831	556	77	775	537	70	707	483
69	709	463	68	695	463	75	752	512
73	734	497	68	697	475	69	702	469
73	736	502	71	713	489	73	747	503
72	736	498	72	730	499	73	744	498
67	683	460	74	755	505	72	729	484
74	741	517	67	684	457	78	785	541
69	696	479	70	706	488	76	777	518
70	705	478	75	756	511	71	711	481
71	719	482	67	678	467	78	780	544
74	751	515	79	808	533	77	784	527
68	682	472	75	756	511			

Вариант 25

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
32	177	56	40	212	72	34	180	64
40	201	69	39	208	63	34	180	66
36	189	56	28	142	55	33	177	61

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
40	213	60	39	196	69	31	159	53
36	180	63	28	143	44	33	182	50
34	185	51	32	162	63	42	223	72
33	177	53	40	200	67	26	130	40
35	175	58	37	191	63	43	224	76
36	189	63	30	163	45	38	205	59
43	229	82	32	177	44	40	205	72
29	153	53	34	187	52	33	181	47
41	208	67	40	208	61	29	147	57
44	221	70	40	213	75	38	202	62
36	186	58	36	196	59	39	210	61
37	185	69	36	199	61	35	181	52
32	173	45	41	206	74	34	187	61
40	211	73	40	204	79	42	218	73
33	177	49	32	172	45	27	145	49
36	187	61	40	219	69	36	197	53
32	179	49						

<u>Вариант 26</u>

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
89	542	-13	89	544	-15	88	536	-8
89	549	-13	91	564	-14	90	549	-13
89	550	-18	88	535	-18	90	542	-4

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
90	556	-12	90	559	-16	91	546	-5
89	550	-20	91	564	-12	88	545	-1
90	540	-18	90	555	-12	90	556	-7
91	553	-17	88	545	-6	88	546	-19
88	544	-2	89	535	-14	87	527	-1
89	541	-14	89	543	-19	92	559	-7
90	556	-19	90	554	-4	89	545	-20
89	541	-10	89	548	-7	88	530	-5
90	552	-4	91	553	-20	90	542	-2
89	548	-15	89	536	-1	89	547	-11
89	534	-13	90	553	-11	90	546	-4
88	528	-16	90	544	-11	92	568	-11
90	553	-15	86	519	-3	90	555	-6
90	551	-1	90	550	-3	89	537	-19
90	553	-4	89	550	-16	88	535	-13

Вариант 27

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
93	298	-194	83	251	-177	90	279	-186
88	281	-178	87	273	-180	89	270	-190
91	283	-199	91	284	-196	81	261	-168
83	254	-178	87	266	-187	87	276	-186
89	271	-188	87	265	-177	88	266	-187

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
86	260	-183	91	280	-191	89	278	-186
85	262	-175	86	274	-186	89	268	-198
88	275	-182	93	286	-206	82	265	-170
85	265	-188	89	278	-192	86	276	-189
87	277	-188	85	269	-172	84	262	-180
82	256	-183	87	268	-186	86	258	-179
91	280	-187	87	269	-192	92	295	-189
86	263	-179	88	264	-181	92	280	-186
86	264	-191	84	263	-185	80	243	-170
87	279	-176	88	282	-192	87	261	-189
91	288	-199	89	273	-196	88	271	-191
84	264	-184	92	284	-201	85	270	-173
82	261	-171	88	267	-195	83	268	-181
83	263	-176	90	274	-197	88	266	-190

Вариант 28

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
39	242	-82	34	215	-71	27	165	-70
36	235	-89	31	198	-70	31	196	-78
35	228	-72	34	209	-72	28	183	-60
20	136	-41	36	225	-82	30	196	-64
31	204	-78	19	132	-40	38	229	-81
41	249	-83	34	218	-74	34	205	-69

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
24	160	-62	28	174	-76	35	221	-80
22	140	-49	30	197	-74	22	143	-63
38	229	-88	24	162	-49	38	242	-92
35	213	-74	17	104	-41	28	181	-69
26	172	-65	27	175	-65	30	199	-77
34	207	-88	29	176	-61	36	225	-84
16	98	-52	33	203	-67	32	201	-68
35	225	-74	32	210	-78	26	159	-58
22	141	-57	30	188	-79	26	163	-72
30	184	-68	41	252	-93	29	181	-73
30	186	-66	44	274	-97	32	206	-72
31	205	-81	22	142	-45	31	205	-80
28	178	-72	33	200	-79	27	174	-63
27	166	-57	38	236	-81			

Вариант 29

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
16	96	-34	13	75	-14	12	66	-24
16	86	-20	16	80	-36	15	84	-18
16	96	-35	15	92	-22	14	71	-21
13	68	-28	17	86	-29	16	86	-25
16	92	-35	15	77	-31	14	85	-23
10	54	-13	16	99	-22	13	78	-33

X	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
16	81	-19	17	94	-34	13	75	-16
15	89	-31	17	104	-32	17	97	-18
14	78	-23	16	86	-29	12	67	-31
13	71	-26	12	74	-15	15	83	-19
12	77	-24	13	75	-17	13	74	-17
13	79	-27	13	84	-33	15	93	-22
15	81	-27	14	88	-22	19	109	-35
14	81	-17	12	67	-30	17	99	-20
13	66	-23	14	78	-27	15	76	-17
14	85	-30	16	87	-19	15	77	-30
17	92	-37	13	74	-27	14	70	-19
13	68	-21	19	107	-27	16	92	-31
18	104	-35	14	76	-34	15	79	-29
15	83	-31	17	102	-18	17	101	-22
13	81	-27	13	79	-28	15	93	-21
17	102	-18	16	86	-36	13	74	-28
15	92	-33	11	70	-20	15	80	-23
13	77	-27	14	77	-16	15	88	-17
16	95	-17	14	81	-24	13	74	-20
15	78	-21	16	87	-31	17	86	-26
16	99	-26	11	71	-24	14	80	-21
15	78	-35	13	81	-18	13	79	-26
15	94	-31	15	76	-19	15	92	-26

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
16	80	-33	16	86	-28	14	86	-26

<u>Вариант 30</u>

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
90	192	172	92	186	164	95	195	173
83	171	147	79	158	143	97	207	187
81	162	142	82	176	146	83	182	146
83	175	149	103	211	193	97	208	192
86	178	166	93	188	181	79	161	146
87	185	157	89	194	176	91	193	171
90	185	168	93	188	182	85	186	152
85	189	167	91	186	165	90	187	160
99	214	191	83	174	150	89	185	177
91	192	168	104	211	194	91	193	174
92	189	166	95	199	176	80	169	149
98	201	193	97	204	178	86	190	158
87	178	156	100	204	199	90	199	167
89	179	167	92	187	173	91	192	179
108	225	212	92	184	173	90	193	178
82	180	148	84	175	154	92	191	171
79	159	142	93	205	180	84	174	148
93	200	182	98	197	183	77	156	148
91	195	168	97	199	193	88	190	158

X	***Y***	***Z***	***X***	***Y***	***Z***	***X***	***Y***	***Z***
98	203	185	104	218	207	89	183	159
90	182	168	84	177	167	81	169	156
87	177	154	90	191	172	92	201	183
85	186	159	93	193	177	93	196	175
88	185	159	92	195	178	108	229	200
103	211	205	76	169	140	78	169	137
91	190	179	90	193	173	98	201	193
94	200	176	90	190	170	99	210	186

Вопросы для самопроверки

1. Базовые понятия теории вероятностей и статистики

1.1. Вероятностный эксперимент, событие, вероятность.
1.2. Случайная величина.
1.3. Числовые характеристики случайных величин.
1.4. Законы распределений случайных величин.
1.5. Таблицы распределений и их применение.
1.6. Взаимосвязь случайных величин.
1.7. Генеральная совокупность и выборка.
1.8. Способы представления и обработки статистических данных.

2. Статистические выводы: оценки и проверка гипотез

2.1. Точечные оценки и их свойства.
2.2. Свойства выборочных оценок.
2.3. Интервальные оценки.
2.4. Статистическая проверка гипотез. Примеры проверки гипотез.

3. Парная линейная регрессия

3.1. Взаимосвязи экономических переменных.
3.2. Суть регрессионного анализа.
3.3. Парная линейная регрессия.
3.4. Метод наименьших квадратов.

4. Проверка качества уравнения регрессии

4.1. Классическая линейная регрессионная модель. Предпосылки метода наименьших квадратов.
4.2. Анализ качества оценок коэффициентов регрессии.
4.3. Проверка гипотез относительно коэффициентов линейного уравнения регрессии.
4.4. Интервальные оценки коэффициентов линейного уравнения регрессии.
4.5. Доверительные интервалы для зависимой переменной.
4.6. Проверка общего качества уравнения регрессии. Коэффициент детерминации R^2.

5. Множественная линейная регрессия

5.1. Оценивание параметров уравнения регрессии с помощью МНК.
5.2. Основные предпосылки применимости МНК.
5.3. Коэффициенты парной, частной и множественной коррелиции.
5.4. Расчет коэффициентов множественной линейной регрессии.
5.5. Дисперсии и стандартные ошибки коэффициентов.
5.6. Интервальные оценки коэффициентов теоретического уравнения регрессии.

5.7. Анализ качества эмпирического уравнения множественной линейной регрессии.
5.8. Проверка статистической значимости коэффициентов уравнения регрессии.
5.9. Проверка общего качества уравнения регрессии.
5.10. Проверка выполнимости предпосылок МНК.
5.11. Статистика Дарбина-Уотсона.

6. Нелинейная регрессия

6.1. Логарифмические (лог-линейные) модели.
6.2. Полулогарифмические модели.
6.3. Обратная модель.
6.4. Степенная модель.
6.5. Показательная модель.

7. Гетероскедастичность

7.1. Суть гетероскедастичности.
7.2. Последствия гетероскедастичности.
7.3. Обнаружение гетероскедастичности.
7.4. Методы смягчения проблемы гетероскедастичности.

8. Автокорреляция

8.1. Суть и причины автокорреляции.
8.2. Последствия автокорреляции.
8.3. Обнаружение автокорреляции.
8.4. Методы устранения автокорреляции.

9. Мультиколлинеарность

9.1. Суть мультиколлинеарности.
9.2. Последствия мультиколлинеарности.
9.3. Определение мультиколлинеарности.
9.4. Методы устранения мультиколлинеарности.

10. Системы одновременных уравнений

11.1. Необходимость использования систем уравнений.
11.2. Составляющие систем уравнений.
11.3. Качество оценок МНК для систем одновременных уравнений.
11.4. Косвенный метод наименьших квадратов (КМНК).
11.5. Инструментальные переменные.
11.6. Проблема идентификации.
11.7. Необходимые и достаточные условия идентифицируемости.
11.8. Оценка параметров систем одновременных уравнений.

Тесты

ЗАДАНИЕ N 1 (*выберите один вариант ответа*)

Эконометрика – это …

ВАРИАНТЫ ОТВЕТОВ:

1) наука, которая дает количественное выражение взаимосвязей экономических явлений и процессов

2) раздел экономической теории, связанный с анализом статистической информации

3) специальный раздел математики, посвященный анализу экономической информации

4) наука, которая осуществляет качественный анализ взаимосвязей экономических явлений и процессов

ЗАДАНИЕ N 2 (*выберите один вариант ответа*)

Коэффициент парной корреляции характеризует …

ВАРИАНТЫ ОТВЕТОВ:

1) тесноту линейной связи между двумя переменными

2) тесноту нелинейной связи между двумя переменными

3) тесноту линейной связи между несколькими переменными

4) тесноту нелинейной связи между несколькими переменными

ЗАДАНИЕ N 3 (*выберите один вариант ответа*)

Фиктивными переменными в уравнении множественной регрессии являются …

ВАРИАНТЫ ОТВЕТОВ:

1) качественные переменные, преобразованные в количественные

2) дополнительные количественные переменные, улучшающие решение

3) комбинации из включенных в уравнение регрессии факторов, повышающие адекватность модели

4) переменные, представляющие простейшие функции от уже включенных в модель переменных

ЗАДАНИЕ N 4 (*выберите один вариант ответа*)

Величина коэффициента регрессии показывает …

ВАРИАНТЫ ОТВЕТОВ:

1) среднее изменение результата при изменении фактора на одну единицу

2) характер связи между фактором и результатом
3) тесноту связи между фактором и результатом
4) тесноту связи между исследуемыми факторами

ЗАДАНИЕ N 5 *(выберите один вариант ответа*)

Метод наименьших квадратов используется для оценивания …
ВАРИАНТЫ ОТВЕТОВ:

1) параметров линейной регрессии
2) величины коэффициента корреляции
3) величины коэффициента детерминации
4) средней ошибки аппроксимации

ЗАДАНИЕ N 6 *(выберите один вариант ответа*)

Несмещенность оценки характеризует …
ВАРИАНТЫ ОТВЕТОВ:

1) равенство нулю математического ожидания остатков
2) наименьшую дисперсию остатков
3) увеличение точности ее вычисления с увеличением объема выборки
4) ее зависимость от объема выборки

ЗАДАНИЕ N 7 *(выберите один вариант ответа*)

Гомоскедастичность подразумевает …
ВАРИАНТЫ ОТВЕТОВ:

1) одинаковую дисперсию остатков при каждом значении фактора
2) рост дисперсии остатков с увеличением значения фактора
3) уменьшение дисперсии остаток с уменьшением значения фактора
4) максимальную дисперсию остатков при средних значениях фактора

ЗАДАНИЕ N 8 *(выберите один вариант ответа*)

Обобщенный метод наименьших квадратов применяется в случае…
ВАРИАНТЫ ОТВЕТОВ:

1) автокорреляции ошибок
2) автокорреляции переменных
3) мультиколлинеарности факторов
4) фиктивных переменных

ЗАДАНИЕ N 9 (*выберите один вариант ответа*)

Коэффициент детерминации рассчитывается для оценки качества…

ВАРИАНТЫ ОТВЕТОВ:

1) подбора уравнения регрессии
2) параметров уравнения регрессии
3) мультиколлинеарных факторов
4) факторов, не включенных в уравнение регрессии

ЗАДАНИЕ N 10 (*выберите один вариант ответа*)

Корреляция подразумевает наличие связи между …

ВАРИАНТЫ ОТВЕТОВ:

1) переменными
2) параметрами
3) случайными факторами
4) результатом и случайными факторами

ЗАДАНИЕ N 11 (*выберите один вариант ответа*)

Число степеней свободы связано с …

ВАРИАНТЫ ОТВЕТОВ:

1) числом единиц совокупности
2) видом уравнения регрессии
3) числом определяемых по совокупности констант
4) характером исследуемых переменных

ЗАДАНИЕ N 12 (*выберите один вариант ответа*)

Критические значения критерия Стьюдента определяются по…

ВАРИАНТЫ ОТВЕТОВ:

1) уровню значимости и одной степени свободы
2) уровню незначимости
3) двум степеням свободы
4) трем и более степеням свободы

ЗАДАНИЕ N 13 (*выберите один вариант ответа*)

Нелинейным является уравнение регрессии нелинейное относительно входящих в него …

ВАРИАНТЫ ОТВЕТОВ:

1) факторов
2) результатов
3) параметров
4) случайных величин

ЗАДАНИЕ N 14 (*выберите один вариант ответа*)

Примером нелинейной зависимости экономических показателей является …
ВАРИАНТЫ ОТВЕТОВ:

1) классическая гиперболическая зависимость спроса от цены
2) линейная зависимость выручки от величины оборотных средств
3) зависимость объема продаж от недели реализации
4) линейная зависимость затрат на производство от объема выпуска продукции

ЗАДАНИЕ N 15 (*выберите один вариант ответа*)

К линейному уравнению нельзя привести …
ВАРИАНТЫ ОТВЕТОВ:

1) $y = a + b \cdot x^c + \varepsilon$
2) $y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + \varepsilon$
3) $y = a + \dfrac{b}{x} + \varepsilon$
4) $y = a + b_1 \cdot x + b_2 \cdot x^2 + \varepsilon$

ЗАДАНИЕ N 16 (*выберите один вариант ответа*)

Величина коэффициента эластичности показывает …
ВАРИАНТЫ ОТВЕТОВ:

1) на сколько процентов изменится в среднем результат при изменении фактора на 1%
2) во сколько раз измениться в среднем результат при изменении фактора в два раза
3) предельно возможное значение результата
4) предельно допустимое изменение варьируемого признака

ЗАДАНИЕ N 17 (*выберите один вариант ответа*)

Тенденция временного ряда характеризует совокупность факторов, …

ВАРИАНТЫ ОТВЕТОВ:

1) оказывающих долговременное влияние и формирующих общую динамику изучаемого показателя

2) оказывающих сезонное воздействие

3) оказывающих единовременное влияние

4) не оказывающих влияние на уровень ряда

ЗАДАНИЕ N 18 (*выберите один вариант ответа*)

Под автокорреляцией уровней временного ряда подразумевается …

ВАРИАНТЫ ОТВЕТОВ:

1) корреляционная зависимость между последовательными уровнями ряда

2) функциональная зависимость между последовательными уровнями ряда

3) корреляционно–функциональная зависимость между последовательными уровнями ряда

4) функциональная зависимость между двумя временными рядами

ЗАДАНИЕ N 19 (*выберите один вариант ответа*)

Аддитивная модель содержит компоненты …

ВАРИАНТЫ ОТВЕТОВ:

1) в виде слагаемых

2) в виде сомножителей

3) в виде их отношений

4) в виде комбинации слагаемых и сомножителей

ЗАДАНИЕ N 20 (*выберите один вариант ответа*)

В стационарном временном ряде трендовая компонента …

ВАРИАНТЫ ОТВЕТОВ:

1) отсутствует

2) присутствует

3) имеет линейную зависимость от времени

4) имеет нелинейную зависимость от времени

ЗАДАНИЕ N 21 (*выберите один вариант ответа*)

Принципиальные сложности применения систем эконометрических уравнений связаны с ошибками…

ВАРИАНТЫ ОТВЕТОВ:

1) спецификации модели
2) оценивания параметров
3) определения случайных воздействий
4) однородности выборочной совокупности

ЗАДАНИЕ N 22 *(выберите один вариант ответа*)

Структурной формой модели называется система …

ВАРИАНТЫ ОТВЕТОВ:

1) взаимосвязанных уравнений
2) независимых уравнений
3) рекурсивных уравнений
4) уравнений с фиксированным набором факторов

ЗАДАНИЕ N 23 *(выберите один вариант ответа*)

В правой части структурной формы взаимозависимой системы могут стоять _______ переменные

ВАРИАНТЫ ОТВЕТОВ:

1) экзогенные
2) лаговые
3) эндогенные
4) нелаговые

ЗАДАНИЕ N 24 *(выберите один вариант ответа*)

Косвенный метод наименьших квадратов применим для …

ВАРИАНТЫ ОТВЕТОВ:

1) идентифицируемой системы одновременных уравнений
2) любой системы одновременных уравнений
3) неидентифицируемой системы уравнений
4) неидентифицируемой системы рекурсивных уравнений

ЛИТЕРАТУРА

1. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. – М.: Дело, 2000. – 400 с.
2. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. Учебник для вузов. – М.: ЮНИТИ, 1998. – 1022 с.
3. Айвазян С.А., Мхитарян В.С. Практикум по прикладной статистике и эконометрике (учебное пособие). – М.: МЭСИ, 1998. – 158 с.
4. Эконометрика: учебник/ И.И. Елисеева, С.В. Курышева, Т.В. Костеева и др.; под ред. И.И. Елисеевой. – 2-е изд., переработ. и доп.- М.: Финансы и статистика, 2007. -576 с.
5. Кремер Н.Ш., Путко Б.А. Эконометрика: Учебник для вузов/ Под ред. проф. Н.Ш.Кремера. – М.: ЮНИТИ – ДАНА, 2002. – 311 с.
6. Доугерти К. Введение в эконометрику. – М.: ИНФРА-М, 1997. – 402 с.
7. Джонстон Дж. Эконометрические методы.- М.: Статистика, 1980.- 446 с.
8. Дуброва Т.А. Статистические методы прогнозирования: Учебное пособие для вузов. – М.: ЮНИТИ-ДАНА, 2003. – 206 с.
9. Замков О.О. Эконометрические методы в макроэкономическом анализе: Курс лекций. – М.: ГУ ВШЭ, 2001. – 122 с.
10. Макаров В.Л., Айвазян С.А., Борисова С.В., Лакалин Э.А. Эконометрическое моделирование. Выпуск 2: Модель экономики России для целей краткосрочного прогноза и сценарного анализа (учебное пособие). – М.: МЭСИ, 2002. – 34 с.
11. Мхитарян В.С., Зелин В.А., Скорик М.А. Моделирование мирового рынка нефти: Пособие для компьютерных исследований по курсу “Эконометрическое моделирование”. – М.: МЭСИ, 2001. – 32 с.
12. Суслов В.И., Ибрагимов Н.М., Талышева Л.П. Цыплаков А. А. Эконометрия. - Новосибирск: ЭФ НГУ, 2003. - 601 с.
13. Эконометрика: Учеб. пособие/ С.А. Бородич. – Мн.: Новое знание, 2001. -408 с.
14. Тихомиров Н.П., Дорохина Е.Ю. Эконометрика: Учебник/ Н.П. Тихомиров, Е.Ю Дорохина. – М.: изд-во “Экзамен”, 2003. – 512 с.
15. Практикум по эконометрике: Учеб. пособие/ И.И. Елисеева, С.В. Курышева, Н.М. Гордеенко и др.; Под ред. И.И.Елисеевой. – М.: Финансы и статистика, 2001. – 1926 с.
16. Балдин К.В., Быстров О.Ф., Соколов М.М. Эконометрика: Учеб. пособие для вузов. – 2-е изд., перераб. и доп. – М.: ЮНИТИ – ДАНА, 2004. – 254 с.

Таблица значений **F**-критерия Фишера при уровне значимости $\alpha = 0{,}05$

k_2 / k_1	1	2	3	4	5	6	8	12	24	∞
1	161,45	199,50	215,72	224,57	230,17	233,97	238,89	243,91	249,04	254,32
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,13	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,06	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,7l	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,80	1,57	1,14
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06
1000	3,85	3,00	2,61	2,38	2,22	2,10	1,95	1,76	1,53	1,03
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

Критические значения ***t***-критерия Стьюдента при уровне значимости α (двух-сторонний)

Число степеней свободы, ν	α		
	0,10	**0,05**	**0,01**
1	6,3138	12,706	63,657
2	2,9200	4,3027	9,9248
3	2,3534	3,1825	5,8409
4	2,1318	2,7764	4,6041
5	2,0150	2,5706	4,0321
6	1,9432	2,4469	3,7074
7	1,8946	2,3646	3,4995
8	1,8595	2,3060	3,3554
9	1,8331	2,2622	3,2498
10	1,8125	2,2281	3,1693
11	1,7959	2,2010	3,1058
12	1,7823	2,1788	3,0545
13	1,7709	2,1604	3,0123
14	1,7613	2,1448	2,9768
15	1,7530	2,1315	2,9467
16	1,7459	2,1199	2,9208
17	1,7396	2,1098	2,8982
18	1,7341	2,1009	2,8784
19	1,7291	2,0930	2,8609
20	1,7247	2,0860	2,8453
21	1,7207	2,0796	2,8314
22	1,7171	2,0739	2,8188
23	1,7139	2,0687	2,8073
24	1,7109	2,0693	2,7969
25	1,7081	2,0595	2,7874
26	1,7056	2,0555	2,7787
27	1,7033	2,0518	2,7707
28	1,7011	2,0484	2,7633
29	1,6991	2,0452	2,7564
30	1,6973	2,0423	2,7500
40	1,6839	2,0211	2,7045
60	1,6707	2,0003	2,6603
120	1,6577	1,9799	2,6174
∞	1,6449	1,9600	2,5758

Отпечатано в Издательско-полиграфическом центре
Набережночелнинского института
Казанского (Приволжского) федерального унверситета

Подписано в печать 05.03.2015
Формат 60х84/16. Печать ризографическая
Бумага офсетная. Гарнитура «Nimes New Roman»
Усл. п. л. 4,6. Уч-изд. л. 4,6
Тираж 100 экз. Заказ №408

423810, г. Набережные Челны, Новый город, проспект Мира, 68/19
Тел./факс (8552) 39-69-99 e-mail: ic-nchi-rpfu@mail.ru