O’ZBEKISTON RESPUBLIKASI

OLIY VA O’RTA MAXSUS TA’LIM VAZIRLIGI

TOSHKENT MOLIYA INSTITUTI

Предмет, метод, задачи и основные понятия основы эконометрики

Тузувчи: Н.Рашитова

Toshkent – 2017

1-ТЕМА	
ПРЕДМЕТ, МЕТОД,ЗАДАЧИ И ОСНОВНЫЕ ПОНЯТИЯ ОСНОВЫ ЭКОНОМЕТРИКИ	
ПЛАН:	
1.1.	**Ведение в эконометрику. Предмет дисциплины эконометрика**
1.2.	**Понятие об эконометрической модели**
1.3.	**Эконометрическое моделирования**

1.1.Ведение в эконометрику. Предмет дисциплины эконометрика

Буквально интерпретированное, эконометрика означает "экономическое измерение». Хотя измерение является важной частью эконометрики, сфера эконометрики гораздо шире, как можно видеть из следующих цитат:

Эконометрика, результат определенного взгляда на роль экономики, состоит в применении математической статистики в экономических данных кредитовать эмпирическую поддержку моделей, построенных по математической экономике и получить численные результаты.

> … эконометрика может быть определена как количественный анализ реальных экономических явлений на основе параллельного развития теории и наблюдений, связанных с соответствующими методами логического вывода.

- Первые попытки количественных исследований в экономике относятся к XVII в. Они были связаны с представителями нового направления в экономической теории —политической арифметики. У. Петти,Ч. Давенант, Г. Кингиспользовали конкретные экономические данные в своих исследованиях, в первую очередь, при расчёте национального дохода. Это направление пробудило поиск экономических законов, по аналогии с физическими, астрономическими и другими естественнонаучными законами. При этом существование неопределённости в экономике ещё не осознавалось.

- Важным этапом возникновения эконометрики явилось развитие статистической теории в трудах Ф. Гальтона, К. Пирсона, Ф. Эджворта. Эти учёные предопределили первые применения парной корреляции. Так, Дж. Э. Юл определял связь между уровнем бедности и формами помощи бедным. Г. Хукер же измерял связь между уровнем брачности и благосостоянием, в котором использовалось несколько индикаторов благосостояния, также он исследовал временные ряды экономических переменных.
- С 1830-х годов наиболее развитые страны стали испытывать необъяснимые с точки зрения экономической науки того времени потрясения — упадок деловой активности, возникновение массовой безработицы. Быстрое промышленное развитие и урбанизация выявили огромный пласт нерешённых социальных проблем. Уже в конце XIX в. неоклассическая теория стала восприниматься как слишком удалённая от действительности. Теория могла стать убедительной в том случае, если она бы смогла объяснить изменения, происходящие в экономике. Для её практического применения требовались количественные выражения базовых экономических терминов.

Эконометрика может быть определена как социальная наука, в которой инструменты экономической теории, математики и статистические выводы применяются к анализу экономических явлений.

Искусство эконометристом заключается в нахождении множества предположений, которые являются одновременно достаточно конкретными и достаточно реалистично, чтобы позволить ему принять наилучшее преимущество из доступных ему данных.

Эконометристы … являются положительным помощь в попытке рассеять плохое имиджем экономики (количественных или иным образом) в качестве предмета, в котором открываются пустые коробки, предполагая существование Can-открывалки раскрыть содержание которых любые десять экономисты интерпретируют в 11 направлениях.

Изучение вводного курса статистики является обязательным условием для любого серьезного курса эконометрики. И вот почему. Курс эконометрики преследует две цели. Во-первых, необходимо показать, как различные количественные методы могут быть использованы для моделирования статистических данных. Это относительно просто. Вторая цель — выработать понимание статистических свойств этих методов, а также того, почему они работают в одних случаях, но не работают в других. Это гораздо сложнее. Поэтому именно такой навык является реально востребованным, — здесь нужны хорошие базовые знания теории статистики. Если вы не изучали статистику, то вам на время следует отложить эту книгу в сторону и вернуться к ней, изучив основы статистики.

1.2.Понятие об эконометрической модели

Вводный курс статистики обычно востребован в различных дисциплинах. По этой причине некоторые его темы не имеют большого значения для эконометрики. За исключением глав, посвященных выборкам, оценкам и гипотезам, другие главы, которые имеют отношение к курсам бизнеса или психологии, во многом расходятся с тематикой эконометрики. Далее перечислен список тем из теории статистики, которые нужно знать для изучения эконометрики.

Описательная статистика. Распределения частот и графическое представление, включая гистограммы (но не в виде деревьев); линейные графики; меры положения и вариации (среднее значение, медиана, мода, дисперсия и стандартное отклонение).

Этот обычный материал не должен вызывать затруднения.

Вероятность. Пространство событий, относительная частота, понятие вероятности; предельная и условная вероятности. Для наших целей простого понимания этих явлений будет достаточно.

Теорему Байеса знать не обязательно *учайные переменные, распределения вероятностей и ожидания.* Этот материал требует внимания. Тем не менее, нет смысла изучать все статистические распределения, которые интересны статистикам. Вам нужно понимать свойства нормального распределения, а также биномиального распределения. Остальные распределения знать не обязательно.

Выборки. Достаточно знания простой случайной выборки. Вам не следует тратить время на стратификацию или кластеры, хотя вы наверняка встречались с этими темами, если проводили опросы. Вы не обязаны знать о выборках без замещения.

Оценки. Понимание разницы между способом оценивания и оценкой имеет большое значение. Вы обязаны знать о несмещенности, дисперсии и оценивании дисперсии.

Статистический вывод. Вы обязательно должны глубоко понимать смысл статистических выводов. Вам нужно знать разницу между ошибками первого рода и второго рода, понимать, что такое уровень значимости теста, а также логику использования одностороннего или двустороннего теста. Вы должны уметь проводить *t*- и .Г-тесты, но знать математические формулы *t*- и *F*-распределений необязательно. Вам должны быть известны доверительные интервалы. Обязательно уметь применять все эти понятия для проверки гипотез. Необязательно тратить время на проверку гипотез, относящихся к разностям средних значений выборок, или гипотез, относящихся к выборочным про порциям.

Дисперсионный анализ. Желательно знать эту тему, но это не самое важное.

После изучения этих тем вводного курса статистики, вероятно, можно сосредоточиться на введении в регрессионный анализ; поскольку для многих студентов, изучивших курс статистики, эконометрика не будет отдельным курсом, полезно изучить введение в эту тему. Возможно, это так, но для наших целей это пока еше преждевременно и будет ненужным дублированием.

Обзор этих тем не может заменить изучения курса статистики. Он делается лишь для того, чтобы дать возможность повторить и закрепить статистические понятия, которые особенно нужны для эконометрического анализа. Скорее всего, вам будете не известна только одна тема — асимптотические свойства способов оценивания (свойства, когда выборка становится очень большой). Эта тема имеет для нас большое значение. Наш обзор не касается проверки гипотез. Принципы проверки гипотез рассмотрены в контексте регрессионного анализа, но если вы никогда не были знакомы с этой темой, вам следует изучить ее самостоятельно, до того, как вы займетесь эконометрикой

Метод эконометрических исследований направлена, по существу, в сочетании экономической теории и фактических измерений, с использованием теории и методики статистического вывода в качестве моста пирса.

Почему отдельная дисциплина?

Как свидетельствуют предыдущие определения, эконометрика является сплавом экономической теории, математической экономики, экономической статистики и математической статистики. Тем не менее, тема заслуживает того, чтобы изучать в своем собственном праве по следующим причинам.

Экономическая теория делает заявления или гипотезы, которые в основном качественный характер. Например, микроэкономическая теория утверждает, что, при прочих неизменных, снижение цены товара, как ожидается, увеличить требуемое количество этого товара. Таким образом, экономическая теория постулирует отрицательная или обратная зависимость между ценой и количеством требуемой от товара. Но сама теория не дает числовую меру соотношения между ними; то есть, она не говорит, сколько количество будет идти вверх или вниз в результате определенного изменения цены товара. Это работа эконометрист обеспечить такие численные оценки. Иными словами, эконометрика дает эмпирическое содержание к большей экономической теории.

Основной задачей математической экономики является выражение экономической теории в математической форме (уравнений) без учета измеримость или эмпирической проверки теории. Эконометрика, как уже отмечалось ранее, в основном заинтересованы в эмпирической проверке экономической теории. Как мы увидим, эконометрист часто использует математические уравнения, предложенные математического экономиста, но ставит эти уравнения в такой форме, что они поддаются эмпирической проверке. И это преобразование математических уравнений в эконометрических требует большой изобретательности и практических навыков.

Экономическая статистика в основном касается сбора, обработки и представления экономических данных в виде графиков и таблиц. Это рабочие места экономической статистике. Это он или она, кто в первую очередь отвечает за сбор данных о валовом национальном продукте (ВНП), занятость, безработица, цены, и так далее. Данные, собранные таким образом составляют исходные данные для эконометрического работы. Но экономический статистик не идти дальше, не будучи связаны с использованием собранных данных для проверки экономических теорий. Конечно, тот, кто делает это становится эконометрист.

Хотя математическая статистика предоставляет множество инструментов, используемых в торговле, эконометрист часто необходимы специальные методы с учетом уникальной природы большинства экономических данных, а именно, что данные не генерируются в результате контролируемого эксперимента. Эконометрист, как метеоролога, как правило, зависит от данных, которые не могут управляться непосредственно. Как Спанос правильно замечает:

В эконометрики модельер часто сталкивается с наблюдательными в отличие от экспериментальных данных. Это имеет два важных последствия для эмпирического моделирования в эконометрики. Во-первых, моделист требуется освоить самые разные навыки, чем те, которые необходимы для анализа экспериментальных данных. , , , Во-вторых, разделение коллектора данных и аналитику данных требует модельер для ознакомления / себя полностью с природой и структурой данных о которых идет речь.

1.3. Эконометрическое моделирования

Грубо говоря, традиционные эконометрические методологии протекает по следующим направлениям:

1. Постановка теории или гипотезы.
2. Уточнение математической модели теории.
3. Уточнение статистическую или эконометрическую модель.
4. Получение данных.
5. Оценка параметров эконометрической модели.
6. Тестирование гипотез.
7. Прогнозирование или предсказание.
8. Использование модели для целей управления или политики.

Чтобы проиллюстрировать предыдущие шаги, рассмотрим хорошо известную кейнсианской теории потребления.

1***. Постановка теории или гипотезы***

Кейнс утверждал:

Основной психологический закон ... в том, что мужчины [женщины] расположены, как правило, и в среднем, чтобы увеличить их потребление, как их доход увеличивается, но не так сильно, как увеличение их доходов.

Короче говоря, Кейнс предположил, что предельная склонность к потреблению (MPC), скорость изменения потребления за единицу (скажем, доллар) изменение дохода, больше нуля, но меньше, чем 1.

2. Уточнение математической модели потребления

Хотя Кейнс постулировал положительную взаимосвязь между потреблением и доходом, он не уточнил точную форму функциональной связи между ними. Для простоты, математический экономист может предложить следующий вид функции кейнсианского потребления:

$$Y = \beta_1 + \beta_2 X \quad 0 < \beta_2 < 1 \quad (I.3.1)$$

где Y = расходы на потребление и X = доход, и где β_1 и β_2, известный как параметры модели, являются, соответственно, коэффициент наклона и пересечения.

Коэффициент наклона в 2 измеряет ПДК. Геометрически уравнение I.3.1 является таким, как показано на рисунке I.1. Это уравнение, в котором говорится, что потребление линейно связано с доходом, является примером математической модели взаимосвязи между потреблением и доходом, который называется функцией потребления в экономике. Модель представляет собой просто набор математических уравнений. Если модель имеет только одно уравнение, как и в предыдущем примере, она называется моделью с одним уравнением, в то время как, если она имеет более чем одно уравнение, известно как модель множественного уравнения (последний будет рассмотрен позже в книге).

В уравнении. (I.3.1) переменная появляется слева от знака равенства, называется зависимой переменной и переменной (B1) на правой стороне называется независимым, или пояснительная переменная (s). Таким образом, в функции кейнсианской потребления, уравнение. (I.3.1), потребление (расходы) является зависимой переменной и источником дохода является объясняющей переменной.

3. Спецификация эконометрической модели потребления

Чисто математическая модель функции потребления данной в уравнении. (I.3.1) имеет ограниченный интерес к эконометристом, ибо она предполагает, что существует точная или детерминированным взаимосвязь между потреблением и доходом. Но отношения между экономическими переменными, как правило, неточны. Таким образом, если мы должны были получить данные о потребительских расходах и располагаемых (т.е. после уплаты) дохода образца, скажем, 500 американских семей и построить эти данные на миллиметровой бумаге с потребительским расходам на вертикальной оси и располагаемого дохода по горизонтали оси, мы не будем ожидать, что все 500 наблюдений лежать точно на прямой линии уравнения. (I.3.1), потому что, в дополнение к доходу, другие переменные влияют на потребительские расходы. Например, размер семьи, возраст членов в семье, семейной религии и т.д., могут оказать некоторое влияние на потребление.

Для обеспечения неточных отношений между экономическими переменными, эконометрист изменил бы функцию детерминированный потребления в уравнении. (I.3.1) следующим образом:

$$Y = \beta 1 + \beta 2X + U \ (I.3.2)$$

где У, известный как нарушения или ошибки, термин, является случайной (стохастической) переменной, которая имеет четко определенные вероятностные свойства. Термин возмущение у вполне может представлять все те факторы, которые влияют на потребление, но не учитываются в явном виде.

Уравнение I.3.2 является примером эконометрической модели. Технически, это является примером линейной регрессионной модели, которая является основной проблемой этой книги. Эконометрическая функция потребления предполагает, что зависимой переменной Y (потребление) линейно связано с объясняющей переменной X (дохода), но о том, что отношения между двумя не является точным; она подвержена индивидуальной изменчивости.

Эконометрическая модель функции потребления можно изобразить, как показано на рисунке I.2.

4. Получение данных

Для оценки эконометрической модели, приведенной в формуле. (I.3.2), то есть, чтобы получить численные значения B и B2, нам нужны данные. Хотя мы будем иметь больше говорить о решающем значении данных для экономического анализа в следующей главе, а пока давайте посмотрим на данные, приведенные в таблице I.l, которые относятся к экономике США за период 1960-2005 гг. Y переменная в этой таблице представляет собой совокупность (для экономики в целом) личные потребительские расходы (PCE), а переменная X является валовой внутренний продукт (ВВП), показатель совокупного дохода, как измеряется миллиардами 2000 долларов. Таким образом, данные в "реальных" условиях; то есть, они измеряются в постоянных (2000) ценах. Данные приведены на рисунке I.3 (см Рисунок I.2). В настоящее время безнадзорности линия, проведенная на рисунке.

5. Оценка эконометрической модели

Теперь, когда у нас есть данные, наша следующая задача состоит в том, чтобы оценить параметры функции потребления. Численные оценки параметров дают эмпирическое содержание в функции потребления. Фактическая механика оценки параметров будут обсуждаться в главе 3. В настоящее время, обратите внимание, что статистический метод регрессионного анализа является основным инструментом, используемым для получения оценок. Используя эту технику, и данные, приведенные в таблице I.l, получаем следующие оценки B и B2, а именно, -299.5913 и 0.7218. Таким образом, оценочная функция потребления:

$$Yt = -299.5913 + 0.7218Xt \quad (I.3.3)$$

Шапочка на Y указывает на то, что она является оценкой. Оценочная функция потребления (т.е. линия регрессии) показана на рисунке I.3.

ТАБЛИЦА I.1

Данные по Y (Личные потребительские расходы) и X (валовой внутренний продукт, 1960-2005), как в 2000 году миллиарды долларов

Год	***PCE(У)***	***CDP(X)***
1960	***1597.4***	***2501.8***
1961	***1630.3***	***2560.0***
1962	***1711.1***	***2715.2***

1963	***1781.6***	***2834.0***
1964	***1888.4***	***2998.6***
196S	***200 7.7***	***3191.1***
1966	***2121.8***	***3399.1***
1967	***2185.0***	***3484.6***
1968	***2310.5***	***3652.7***
1969	***2396.4***	***3765.4***
1970	***2451.9***	***3771.9***
1971	***2545.5***	***3898.6***
1972	***2701.3***	***4105.0***
1973	***2833.8***	***4341.5***
1974	***2812.3***	***4319.6***
197S	***2876.9***	***4311.2***
1976	***3035.5***	***4540.9***
1977	***3164.1***	***4750.5***
1978	***3303.1***	***5015.0***
1979	***3383.4***	***5173.4***
1980	***3374.1***	***5161.7***
1981	***3422.2***	***5291.7***
1982	***3470.3***	***5189.3***
1983	***3668.6***	***5423.8***
1984	***3863.3***	***5813.6***
1985	***4064.0***	***6053.7***
1986	***4228.9***	***6263.6***
1987	***4369.8***	***6475.1***
1988	***4546.9***	***6742.7***
1989	***4675.0***	***6981.4***
1990	***4770.3***	***7112.5***
1991	***4778.4***	***7100.5***

1992	***4934.8***	***7336.6***
1993	***5099.8***	***7532.7***
1994	***5290.7***	***7835.5***
199S	***5433.5***	***8031.7***
1996	***5619.4***	***8328.9***
1997	***5831.8***	***8703.5***
1998	***6125.8***	***9066.9***
1999	***6438.6***	***9470.3***
2000	***6739.4***	***9817.0***
2001	***6910.4***	***9890.7***
2002	***7099.3***	***10048.8***
2003	***7295.3***	***10301.0***
2004	***7577.1***	***10703.5***
2005	***7841.2***	***11048.6***

Как показано на рисунке I.3 показывает, линия регрессии соответствует данным достаточно хорошо в том, что точки данных очень близки к линии регрессии. Из этого рисунка мы видим, что за период 1960-2005 коэффициента наклона (т.е. ПДК) составляла около 0,72, предполагая, что в течение периода выборки увеличение реальных доходов одного доллара привело, в среднем, к увеличению около 72 центов в реальных расходов на потребление. Мы говорим, в среднем, потому что отношения между потреблением и доходом является неточным; как это видно из рис I.3, не все точки лежат точно на линии регрессии. Проще говоря, мы можем сказать, что, по нашим данным, в среднем, или среднее, потребительские расходы выросли примерно на 72 центов за увеличения на один доллар в реальном доходе.

6. Проверка гипотез

Предполагая, что встроенная модель является достаточно хорошим приближением к действительности, мы должны разработать соответствующие критерии, чтобы выяснить, является ли получены оценки, скажем, уравнение I.3.3 согласуются с ожиданиями теории, которая проходит испытания. В соответствии с "Позитивный" экономистов, как Милтон Фридман, теории или гипотезы, которая не является проверяемым путем обращения к эмпирические данные не могут быть допустимы в рамках научного исследования.

Как было отмечено ранее, Кейнс ожидал, что MPC будет положительным, но меньше 1. В нашем примере мы обнаружили, что MPC составляет около 0,72. Но прежде, чем мы примем этот вывод, как подтверждение теории кейнсианский потребления, мы должны спросить ли эта оценка значительно ниже единицы, чтобы убедить нас в том, что это не случайность или особенность частых данных, которые мы использовали. Другими словами, это 0,72 статистически меньше, чем 1? Если да, то это может поддержать теорию Кейнса.

Такое подтверждение или опровержение экономических теорий на основе выборки данных основывается на ветви статистической теории, известной как статистического вывода (проверка гипотез). В этой книге мы увидим, как на самом деле этот процесс проводили умозаключение.

7. Прогнозирование или предсказывание

Если выбранная модель не опровергает гипотезу или рассматриваемой теории, мы можем использовать его, чтобы предсказать будущую стоимость зависимого, или прогноз, переменная Y на основе известной или ожидаемой будущей стоимости пояснительная или предсказателем, переменная X.

Чтобы проиллюстрировать это, предположим, мы хотим предсказать расходы среднего уровня потребления на 2006 год значение ВВП за 2006 г. составил 11319.4 млрд долларов.

Полагая этот показатель ВВП на правой части уравнения. (I.3.3), получим:

или около 7870 млрд долларов. Таким образом, принимая во внимание значение ВВП, среднее или среднее, прогноз, расходы на потребление составляет около 7870 млрд долларов. Фактическое значение расхода на потребление сообщили в 2006 году составил 8044 млрд долларов. Расчетная модель уравнения. (I.3.3), таким образом, занижены по сравнению с фактическими расходами на потребление около 174 млрд долларов. Мы могли бы сказать, что ошибка прогноза составляет около 174 000 000 000 долларов, что составляет около

1,5 процента от фактического значения ВВП за 2006 год Когда мы полностью обсудим модель линейной регрессии в последующих главах, мы попытаемся выяснить, если такая ошибка "маленький" или "большой". Но что важно сейчас, следует отметить, что такие ошибки прогноза неизбежны, учитывая статистический характер нашего анализа.

Существует еще одно применение расчетной модели уравнения. (I.3.3). Предположим, что президент решит предложить снижение налога на прибыль. Каков будет эффект такой политики в отношении доходов и, таким образом, на потребительские расходы и в конечном итоге на работу?

Предположим, что в результате предлагаемого изменения политики, инвестиционных расходов расширяются. Что будет влияние на экономику? Как показывает макроэкономической теории, изменение следующие виды доходов, скажем, стоит доллара изменения инвестиционных расходов определяется дохода мультипликатора M, который определяется как

M = 1

1 - MPC (I.3.5)

Если мы используем MPC 0,72, полученного в формуле. (I.3.3), этот множитель становится равным примерно M = 3,57. То есть, увеличение (уменьшение) доллар инвестиций в конечном итоге приведет к более чем трехкратное увеличение (уменьшение) дохода; обратите внимание, что требуется время для мультипликатор для работы.

Критическое значение в этом вычислении является MPC, для множитель зависит от него. И эта оценка MPC может быть получена из моделей, таких как выход уравнения. (I.3.3). Таким образом, количественная оценка ПДК дает ценную информацию для целей политики. Зная ПДК, можно прогнозировать дальнейший ход доходов, расходов на потребление и занятость после изменения фискальной политики правительства.

8. Применение модели для управления или политики целей

Предположим, что мы имеем оценочную функцию потребления, приведенные в формуле. (I.3.3). Предположим далее, что правительство считает, что потребительские расходы около 8750 (миллиарды 2000 долларов) будет держать уровень безработицы на нынешнем уровне около 4,2 процента (в начале 2006 года). Какой уровень дохода будет гарантировать целевой суммы расходов на потребление?

Если результаты регрессии приведены в формуле. (I.3.3) представляется разумным, простая арифметика показывает, что

которая дает X = 12537, приблизительно. То есть, уровень дохода около 12537 (миллиардов) долларов, дана MPC около 0,72, будет производить расходование около 8750 млрд долларов.

Поскольку эти расчеты показывают, по оценкам, модель может быть использована для контроля или политики, целей. Путем соответствующего сочетания фискальной и денежно-кредитной политики, правительство может манипулировать управляющей переменной X для получения желаемого уровня целевой переменной Y.

Рисунок I.4 обобщает анатомии классического эконометрического моделирования.

Выбор среди конкурирующих моделей

Когда правительственное учреждение (например, министерство торговли США) собирает экономические данные, такие как, что показанные в таблице I.1, это не обязательно иметь какую-либо экономическую теорию в виду. Как же тогда можно узнать, что данные действительно поддерживают кейнсианскую теорию потребления? Это потому, что функция кейнсианского потребления (т.е. линия регрессии), показанная на рисунке I.3 чрезвычайно близки к реальным точкам данных? Возможно ли, что другая модель потребления (теория), возможно, в равной степени соответствовать данным, а?

Например, Милтон Фридман разработал модель потребления, называется доход.

Как классифицирует схема на рисунке I.5 предполагает, эконометрику, можно разделить на две основные категории: теоретическая эконометрика и прикладная эконометрика. В каждой категории можно подойти к этому вопросу в классическом или байесовской традициях.

Теоретическая эконометрика касается разработки соответствующих методов измерения экономических отношений специализируемая на эконометрических моделях. В этом аспекте, эконометрика сильно опирается на математической статистике. Например, один из методов, широко используемый, является метод наименьших квадратов. Теоретические эконометрики должны изложить предположения этого метода, его свойства, и что происходит с этими свойствами, когда один или несколько предположений метода не FUL заполняли.

В прикладных эконометрики мы используем инструменты теоретических эконометрики для изучения некоторых специальных поле экономики и бизнеса, такие как функции производства, инвестиционной функции, функций спроса и предложения, теории портфеля и т.д.

Эта книга касается в основном с развитием эконометрических методов, их предположениям, их использования, а также их ограничения. Эти методы проиллюстрированы примерами из различных областей экономики и бизнеса. Но это не книга прикладной эконометрики в том смысле, что он вникает глубоко в какой-либо конкретной поля экономического применения.

Список литературы

Основная литература:	1. Christopher Dougherty. Introduction to Econometrics. Oxford University Press, 2011. – 573 p. 2. Gujarati D.N. Basic Econometrics. McGraw-Hill, 5th edition, 2009. – 922 p. 3.Абдуллаев О.М., Ходиев Б.Ю., Ишназаров А.И. Эконометрика. Учебник. –Т.: Fan va texnologiya. 2007. – 612 с. 4. Абдуллаев О.М., Жамалов М.С. Эконометрическое моделирование. Учебник. –Т.: Fan va texnologiya. 2010. – 612 с.
Дополнительная литература	1. Greene W.H.Econometric Analysis. Prentice Hall. 7th edition, 2011. -1232p. 2.Валентинов В,А.Эконометрика: Учебник. –М.:ИТК»Дашков и К»,2009–367с. 3.Кремер Н.Ш.Эконометрика:Учебник–М.:ЮНИТИ-ДАНА,2008. –562с. 4.Айвазян С.А. Прикладная статистика и основы эконометрика:Учебник. –М.ЮНИТИ,2007. –345с. 5.Елисеева И.И., Курышева С.В.и др. Эконометрика:Учебник–М.:Финансы и статистика, 2007. –260с. 6.Habibullayev I Iqtisodiy matematik usullsr va modellar: oʻquv qoʻllamna / Oʻzbekiston Respublikasi Oliy va oʻrta maxsus ta`lim vazirligi/–Toshkent:“Tafakkur-Boʻstoni”, 2012.112 b.
Интернет ресурсы:	**www.ziyonet.uz** **www.lex.uz** **www.stat.uz** **www.mehnat.uz** **www.mf.uz** **www.tfi.uz** **www.mfer.uz** **www.conomical.uz** **www.cisstat.com** **www.gks.ru** **www.unstats.un.org**

2-ТЕМА

ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ И ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ В ЭКОНОМЕТРИКЕ

ПЛАН:

2.1.Экономические информации и их обработка

2.2. **Экономические показатели и их типы.**

2.3.**Требование по построению эконометрических моделей и формирование исходной информации.**

2.1.Экономические информации и их обработка

Эконометрическая информация формируется в результате статистического наблюдения экономических объектов.

Слово «информация» в переводе с латинского языка означает осведомленность, давать сведения о чем-либо.

Статистическая информация (статистические данные) — первичный статистический материал, формирующийся в процессе статистического наблюдения, который затем подвергается систематизации, сводке, обработке, анализу и обобщению. Она является фундаментом эконометрического исследования.

Статистическая информация может быть прежде всего получена из различных публикаций. Главными источниками опубликованной статистической информации являются издания органов государственной

Статистическое наблюдение — это начальная стадия экономико-статистического исследования.

статистики. Если нет соответствующих данных в статистических сборниках, то можно получить их самим, т.е. провести статистическое наблюдение.

Она представляет собой научно организованную работу по собиранию массовых первичных данных о явлениях и процессах общественной жизни.

Важность этого этапа исследования определяется тем, что использование только объективной и достаточно полной информации, полученной в результате статистического наблюдения, на последующих этапах исследования в состоянии обеспечить научно обоснованные выводы о характере и закономерностях развития изучаемого объекта.

Однако не всякое собирание сведений может быть названо статистическим наблюдением. Статистическим можно назвать лишь такое наблюдение, которое обеспечивает регистрацию устанавливаемых фактов в учетных документах для последующего их обобщения.

Собираемые данные должны отвечать двум требованиям : достоверности и сопоставимости. ***Достоверность*** — это соответствие данных тому, что есть на самом деле. Вся методика, организация и техника проведения статистического наблюдения должны быть нацелены на обеспечение достоверных данных.

Чтобы данные об отдельных явлениях можно было обобщать, они должны быть сопоставимы друг с другом : собираться в одно и то же время, по единой методике. Кроме того, должна быть обеспечена сравнимость с прошлыми исследованиями, чтобы можно было понять, как изменяется явление.

2.2. Экономические показатели и их типы.

Получаемая в процессе статистического наблюдения информация об отдельных единицах статистической совокупности характеризует их, как правило, с различных сторон. Однако обобщающую характеристику по совокупности в целом можно получить, систематизируя и обобщая полученную информацию, а также сводку, являющуюся второй стадией статистического исследования, в процессе которого осуществляется научная обработка собранного материала. В результате этого этапа индивидуальные данные превращаются в упорядоченную систему статистических показателей.

Таким образом, *статистическая сводка* — систематизация единичных фактов, позволяющая перейти к обобщающим показателям, относящимся ко всей изучаемой совокупности и ее частям, и осуществлять анализ и прогнозирование изучаемых явлений и процессов.

Сводка статистической информации, как правило, не ограничивается получением общих итогов по изучаемой совокупности. Чаще всего исходная информация на этой стадии статистической работы систематизируется, образуются отдельные статистические совокупности, т.е. осуществляется статистическая группировка. Причем различающиеся между собой единицы статистической совокупности по значениям изучаемого признака можно объединить в группы (по их сходству или различию в существенном отношении).

Другими словами, *группировка* — это процесс образования однородных групп на основе расчленения статистической совокупности на части или объединение изучаемых единиц в частные совокупности по существенным для них признакам.

Результатом осуществления этого двуединого процесса является разделенный на группы объект наблюдения.

Результаты сводки и группировки материалов статистического наблюдения оформляются в виде статистических рядов распределения и таблиц.

Статистические ряды распределения представляют собой упорядоченное расположение единиц изучаемой совокупности на группы по группировочному признаку.

Ряды распределения, образованные по качественным признакам, называют *атрибутивными*. Следующая таблица является примером статистического ряда распределения по атрибутивному признаку.

Т а б л и ц а 2.1

Распределение продавцов магазина по категориям

Группы продавцов по категориям	Число продавцов (чел.)	В % к итогу
Первая	50	25
Вторая	100	50
Третья	50	25
И т о г о	200	100

При группировке ряда по количественному признаку получаются *вариационные ряды*. При этом вариационные ряды по способу построения бывают *дискретными* (прерывными), основанными на прерывной вариации признака (например, число касс в магазине, комнат в квартире), и *интервальными* (непрерывными), базирующимися на непрерывно изменяющемся значении признака, имеющими любые (в том числе и дробные) количественные выражения (величина фонда оплаты труда, объем товарооборота).

Вариационные ряды состоят из двух элементов: варианты и частоты.

Варианта — это отдельное значение варьируемого признака, которое он принимает в ряду распределения. *Частотами* называются численности отдельных вариант или каждой группы вариационного ряда. Частоты, выраженные в долях единицы или в процентах к итогу, называются *частостями*.

Сумма частот составляет объем ряда распределения.

Теперь рассмотрим дискретный ряд распределения (табл. 2.2):

Т а б л и ц а 2.2

Распределение магазинов по числу товарных секций

Число товарных секций	Число магазинов	В % к итогу
1	3	6
2	10	20
3	15	30
4	12	24
5	7	14
6	3	6
И т о г о	50	100

Здесь в первом столбце приведены варианты, во втором – частоты, а в третьем – частости.

Характер распределения изображается графически в виде *полигона распределения*, представленного на рис. 2.1.

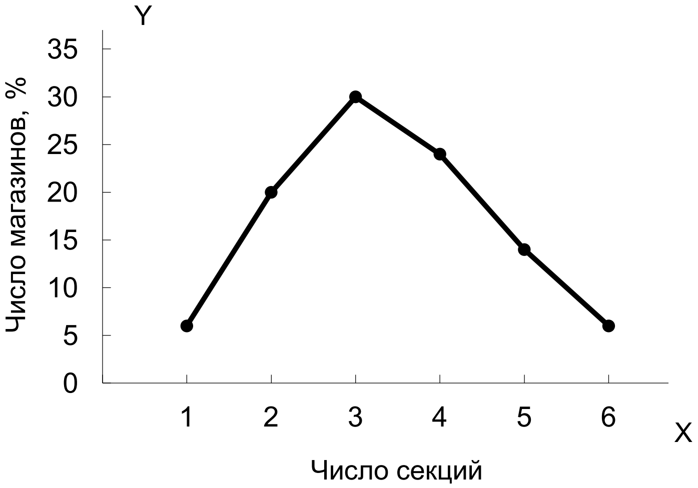


Рис. 2.1. Полигон распределения магазинов по числу товарных секций

Далее рассмотрим интервальный ряд распределения (табл. 2.3):

Т а б л и ц а 2.3

Распределение рабочих завода по выработке

Выработка рабочих (тыс. сум)	Число рабочих (чел.)	В % к итогу
80 – 100	5	10
100 – 120	10	20
120 – 140	20	40
140 – 160	10	20
160 – 180	5	10
И т о г о	50	100

Интервальный ряд распределения, так же как и дискретный, помогает выявить структуру изучаемого явления. Приведенные в табл. 2.3 данные свидетельствуют о составе продавцов по уровню производительности труда.

Интервальный ряд распределения изображается графически в виде *гистограммы*. При ее построении на оси абсцисс откладывают интервалы ряда, высота которых равна частотам, отложенным на оси ординат. Над осью абсцисс строятся прямоугольники, площадь которых соответствует величинам произведений интервалов на их частоты. Данные табл. 2.3 представлены на рис. 2.2.

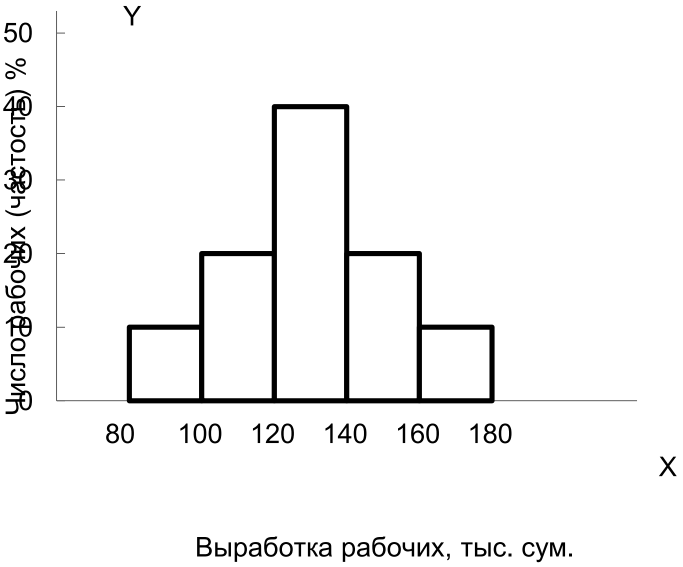


Рис. 2.2. Гистограмма распределения рабочих по выработке

2.3.Требование по построению эконометрических моделей и формирование исходной информации.

Результаты сводки и группировки материалов наблюдения, как правило, представляются в виде статистических таблиц. Это наиболее рациональная форма представления результатов статистической сводки.

Статистическая таблица — система строк и столбцов, в которых в определенной последовательности и связи излагается статистическая информация о социально-экономических явлениях.

Т а б л и ц а 2.4

Продажа некоторых продуктов питания продовольственными магазинами города

Товарные группы	Продано (тыс. сум.)	
	2011	2012
Мясо и птица	12,8	13,9
Колбасные изделия и копчености всякие	14,0	13,9
Рыба всякая и сельди	2,0	2,4
Молоко и молочные продукты	8,83	8,78

Важное значение при изучении различных процессов имеет графическое изображение статистической информации. Графики являются самой эффективной формой представления данных с точки зрения их восприятия. С помощью графиков достигается наглядность характеристики структуры, динамики, взаимосвязи явлений, их сравнения.

Статистический график представляет собой чертеж, на котором при помощи условных геометрических фигур (линий, точек или других символических знаков) изображаются статистические данные.

Диаграмма представляет собой чертеж, на котором статистическая информация изображается посредством геометрических фигур или символических знаков. Диаграммы бывают разных видов: линейные, радиальные, точечные, плоскостные, объемные, фигурные.

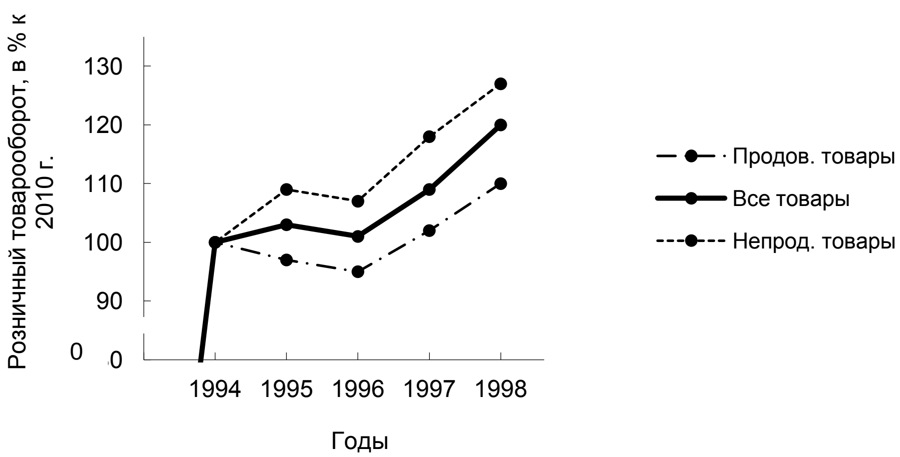


Рис. 2.3. Динамика розничного товарооборота в государственной и кооперативной торговле области (в % к 2010 году)

В статистике наибольшее применение имеют *линейные диаграммы*.

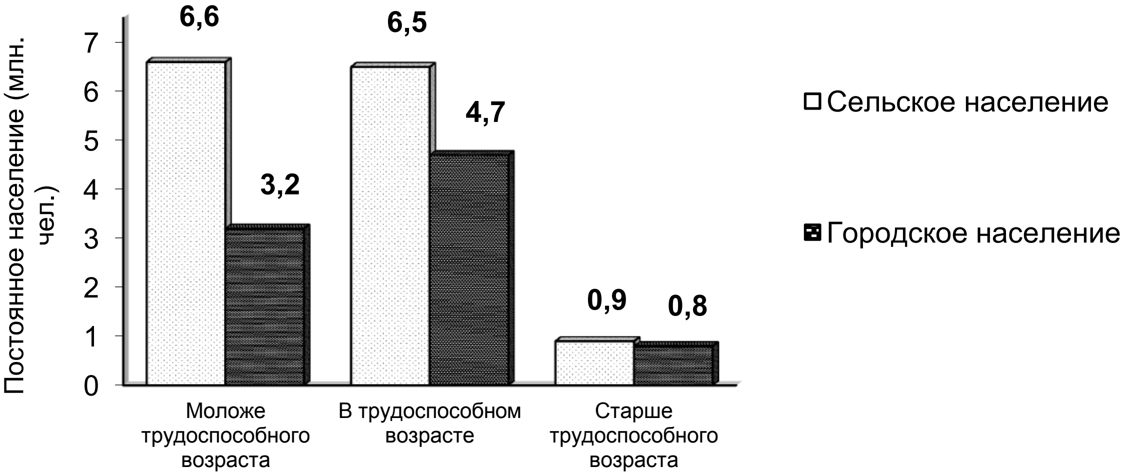


Рис. 2.4. Численность постоянного населения страны

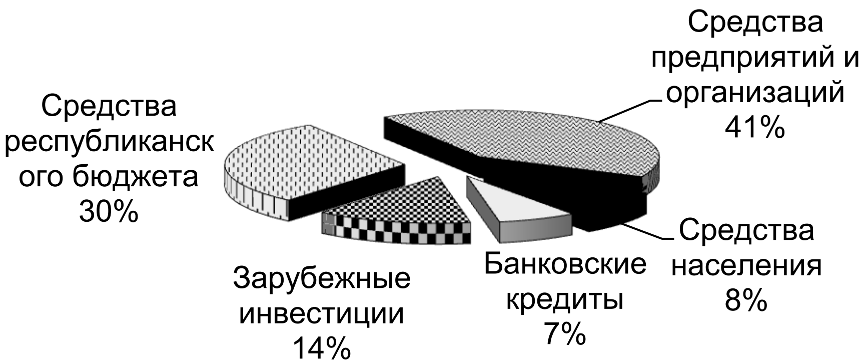


Рис. 2.5. Диаграмма структуры капиталовложений по источникам финансирования

в регионе за период с января по июнь 2016 года

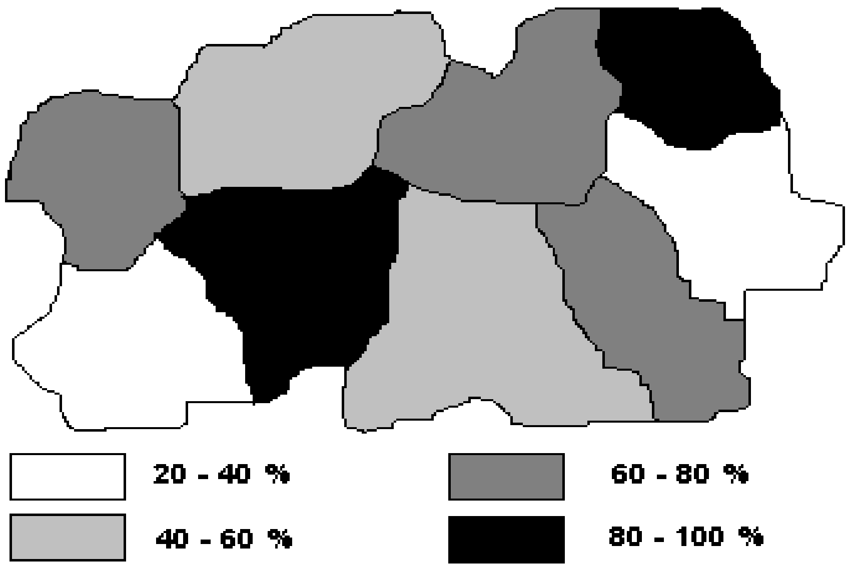


Рис. 2.6. Картограмма, характеризующая процент машинного сбора хлопка по районам

Система совместных, одновременных уравнений (или структурная форма модели) обычно содержит эндогенные и экзогенные переменные.

Эндогенные неременные обозначены в приведенной ранее системе одновременных уравнений как **y.** Это зависимые переменные, число которых равно числу уравнений в системе.

Экзогенные переменные обозначаются обычно как **х.** Это предопределенные переменные, влияющие на эндогенные переменные, но не зависящие от них.

При построении эконометрических моделей необходимо обратит на следующие моменты.

При построении моделей выявляются существенные факторы, определяющие исследуемое явление и отбрасываются детали, несущественные для решения поставленной проблемы.

С одной стороны, модели должны быть доступны для изучения, в силу чего они не должны быть слишком сложными — значит, они неминуемо будут лишь упрощенными копиями. Но с другой стороны, выводы, полученные при их изучении, следует распространить на объекты-оригиналы, следовательно, модель должна отражать существенные черты изучаемого реального объекта.

Список литературы

Основная литература:	1. Christopher Dougherty. Introduction to Econometrics. Oxford University Press, 2011. – 573 p. 2. Gujarati D.N. Basic Econometrics. McGraw-Hill, 5th edition, 2009. – 922 p. 3.Абдуллаев О.М., Ходиев Б.Ю., Ишназаров А.И. Эконометрика. Учебник. –Т.: Fan va texnologiya. 2007. – 612 с. 4. Абдуллаев О.М., Жамалов М.С. Эконометрическое моделирование. Учебник. –Т.: Fan va texnologiya. 2010. – 612 с.
Дополнительная литература	1. Greene W.H.Econometric Analysis. Prentice Hall. 7th edition, 2011. -1232p. 2.Валентинов В,А.Эконометрика: Учебник. –М.:ИТК»Дашков и К»,2009–367с. 3.Кремер Н.Ш.Эконометрика:Учебник–М.:ЮНИТИ-ДАНА,2008. –562с. 4.Айвазян С.А. Прикладная статистика и основы эконометрика:Учебник. –М.ЮНИТИ,2007. –345с. 5.Елисеева И.И., Курышева С.В.и др. Эконометрика:Учебник–М.:Финансы и статистика, 2007. –260с. 6.Habibullayev I Iqtisodiy matematik usullsr va modellar: o‘quv qo‘llamna / O‘zbekiston Respublikasi Oliy va o‘rta maxsus ta`lim vazirligi/–Toshkent:“Tafakkur-Bo‘stoni”, 2012.112 b.
Интернет ресурсы:	www.ziyonet.uz **www.lex.uz** www.stat.uz **www.mehnat.uz** **www.mf.uz** www.tfi.uz www.mfer.uz www.conomical.uz **www.cisstat.com** **www.gks.ru** **www.unstats.un.org**

3-ТЕМА	
ПАРНЫЙ КОРРЕЛЯЦИОННО-РЕГРЕССИОННЫЙ АНАЛИЗ	
ПЛАН:	
3.1.	**Спецификация модели.**
3.2.	**Линейная регрессия и корреляция: смысл и оценка параметров.**
3.3.	**Оценка значимости параметров линейной регрессии и корреляции.**
3.4.	**Интервальный прогноз на основе линейного уравнения регрессии.**
3.5.	**Нелинейная регрессия. Корреляция для нелинейной регрессии.**
3.6.	**Средняя ошибка аппроксимации.**

3.1.Спецификация модели.

Любое эконометрическое исследование начинается со *спецификации модели,* т. е. с формулировки вида модели исходя из соответствующей теории связи между переменными.

В первую очередь из всего круга факторов, влияющих на результативный признак, необходимо выделить наиболее существенно влияющие факторы. Парная регрессия достаточна, если имеется доминирующий фактор, который и используется в качестве объясняющей переменной. В этом случае нужно знать, какие остальные факторы предполагаются неизменными, возможно, в дальнейшем их придется учесть в модели и от простой регрессии перейти к множественной.

Уравнение простой регрессии характеризует связь между двумя переменными, которая проявляется как некоторая закономерность лишь в среднем по совокупности наблюдений. В уравнении регрессии корреляционная по сути связь признаков представляется в виде функциональной связи, выраженной соответствующей математической функцией. Практически в каждом отдельном случае величина y складывается из двух слагаемых:

$$y_j = \tilde{y}_{xj} + \varepsilon_j$$

где y, — фактическое значение результативного признака;

y_{xj} — теоретическое значение результативного признака, найденное исходя из соответствующей математической функции связи y и x, т. е. из уравнения регрессии;

Ej - случайная величина, характеризующая отклонения реального значения результативного признака от теоретического, найденного по уравнению регрессии.

Случайная величина ε, или *возмущение,* включает влияние не учтенных в модели факторов, случайных ошибок и особенностей измерения. Ее присутствие в модели обусловлено тремя источниками: спецификацией модели, выборочным характером исходных данных, особенностями измерения переменных.

К ошибкам спецификации будет относиться не только неправильный выбор той или иной математической функции для $\tilde{y}_x$, но и недоучет в уравнении регрессии какого-либо существенного фактора, т. е. использование парной регрессии вместо множественной. Наряду с ошибками спецификации могут иметь место ошибки выборки, поскольку исследователь чаще всего работает с выборочными данными при установлении закономерной связи между признаками. Ошибки выборки имеют место и в силу неоднородности данных в исходной статистической совокупности, что, как правило, бывает при изучении экономических процессов. Если совокупность неоднородна, то уравнение регрессии не имеет практического смысла. Для получения хорошего результата обычно исключают из совокупности единицы с аномальными значениями исследуемых признаков.

И в этом случае результаты регрессии представляют собой выборочные характеристики. Использование *временной информации* также представляет собой выборку из всего множества хронологических дат. Изменив временной интервал, можно получить другие результаты регрессии.

Наибольшую опасность в практическом использовании методов регрессии представляют *ошибки измерения.* Если ошибки спецификации можно уменьшить, изменяя форму модели (вид математической формулы), а ошибки выборки — увеличивая объем исходных данных, то ошибки измерения практически сводят на нет все усилия по количественной оценке связи между признаками. Особенно велика роль ошибок измерения при исследовании на макроуровне. Так, в исследованиях спроса и потребления в качестве объясняющей переменной широко используется «доход на душу населения». Вместе с тем статистическое измерение величины дохода сопряжено с рядом трудностей и не лишено возможных ошибок, например в результате наличия сокрытых доходов.

Приведем еще один пример: в настоящее время органы государственной статистики получают балансы предприятий, достоверность которых никто не подтверждает. Последующее обобщение такой информации может содержать ошибки измерения. Исследуя, например, в качестве результативного признака прибыль предприятий, мы должны быть уверены, что предприятия показывают в отчетности адекватные реальной действительности величины.

Предполагая, что ошибки измерения сведены к минимуму, основное внимание в эконометрических исследованиях уделяется *ошибкам* *спецификации модели.*В парной регрессии спецификация модели связана с выбором вида математической функции, а в множественной — также с отбором факторов, включаемых в модель.

3.2. Линейная регрессия и корреляция: смысл и оценка параметров.

Линейная регрессия находит широкое применение в эконометрике ввиду четкой экономической интерпретации ее параметров. Линейная регрессия сводится к нахождению уравнения вида:

$$\tilde{y}_x = a + b * x + \varepsilon$$

Уравнение такого вида позволяет по заданным значениям фактора *x* иметь теоретические значения результативного признака подстановкой в него фактических значений фактора x.

Построение линейной регрессии сводится к оценке ее параметров — *a* и *b*. Оценки параметров линейной регрессии могут быть найдены разными методами. Можно обратиться к полю корреляции и, выбрав на графике две точки, провести через них прямую линию, затем по графику найти значения параметров. Параметр *a* определим как точку пересечения линии регрессии с осью *oy,* а параметр *b* оценим исходя из угла наклона линии регрессии как *dy/dx,* где *dy* — приращение результата *y,* а *dx* — приращение фактора *x.*

Классический подход к оцениванию параметров линейной регрессии основан на *методе наименьших квадратов* (МНК).

Можно воспользоваться следующими формулами для определения параметров значений *a* и *b:*

$$a = \overline{y} - b\overline{x}$$

$$b = \frac{\sum(x-\overline{x})(y-\overline{y})}{\sum(x-\overline{x})^2}$$

$$b = \frac{\overline{yx} - \overline{y} * \overline{x}}{x^2 - \overline{x}^2}$$

Параметр ***b*** называется коэффициентом регрессии. Его величина показывает среднее изменение результата с изменением фактора на одну единицу. Знак при коэффициенте регрессии ***b*** показывает направление связи: при $b > 0$ — связь прямая, а при $b < 0$ — связь обратная.

Возможность четкой экономической интерпретации коэффициента регрессии сделала линейное уравнение регрессии достаточно распространенным в эконометрических исследованиях.

Формально *a* — значение *y* при $x = 0$. Если признак-фактор *x* не имеет и не может иметь нулевого значения, то трактовка свободного члена *a* не имеет смысла. Параметр *a* может не иметь экономического содержания. Попытки экономически интерпретировать параметр *a* могут привести к абсурду, особенно при $a<0$.

Интерпретировать можно лишь знак при параметре ***a***. Если $a > 0$, то относительное изменение результата происходит медленнее, чем изменение фактора. Иными словами, вариация результата меньше вариации фактора — коэффициент вариации по фактору *x* выше коэффициента вариации для результата *y: Vx> Vy.*

Уравнение регрессии всегда дополняется показателем тесноты связи. При использовании линейной регрессии в качестве такого показателя выступает

линейный коэффициент корреляции r_{xy} . Имеются разные модификации формулы линейного коэффициента корреляции, например:

$$r_{xy} = \frac{\sum((x-\bar{x})(y-\bar{y}))}{\sqrt{\sum(y-\bar{y})^2 * \sum(x-\bar{x})^2}} \qquad \text{или} \qquad r_{xy} = b\frac{\sigma_x}{\sigma_y}$$

$$\sigma_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} \qquad \sigma_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n}}$$

Как известно, линейный коэффициент корреляции находится в границах **$-1 < r_{xy} < 1$.**

Если коэффициент регрессии $b > 0$, то $0 < r_{xy} < 1$, и, наоборот,

при $b<0$ $-1<r_{xy}<0$.

Следует иметь в виду, что величина линейного коэффициента корреляции оценивает тесноту связи рассматриваемых признаков в ее линейной форме. Поэтому близость абсолютной величины линейного коэффициента корреляции к нулю еще не означает отсутствия связи между признаками. При иной спецификации модели связь между признаками может оказаться достаточно тесной.

Для оценки качества подбора линейной функции рассчитывается квадрат линейного коэффициента корреляции r^2_{xy}, называемый ***коэффициентом детерминации***. Коэффициент детерминации характеризует долю дисперсии результативного признака *y,* объясняемую регрессией, в общей дисперсии результативного признака:

$$r^2_{yx} = \frac{\sigma^2{}_y объясн.}{\sigma^2{}_y общ}$$

Соответственно величина 1 - r^2 характеризует долю дисперсии *y,* вызванную влиянием остальных не учтенных в модели факторов.

Величина коэффициента детерминации является одним из критериев оценки качества линейной модели. Чем больше доля объясненной вариации, тем соответственно меньше роль прочих факторов и, следовательно, линейная модель хорошо аппроксимирует исходные данные.

Линейный коэффициент корреляции как измеритель тесноты линейной связи признаков логически связан не только с коэффициентом регрессии *b,* но и с коэффициентом эластичности, который является показателем силы связи, выраженным в процентах. При линейной связи признаков *x* и *y* средний коэффициент эластичности в целом по совокупности определяется как

$$Э = b\frac{\bar{x}}{\bar{y}}$$

т. е. его формула по построению близка к формуле линейного коэффициента корреляции $r_{yx} = b\frac{\sigma_x}{\sigma_y}$

Несмотря на схожесть этих показателей, измерителем *тесноты связи* выступает линейный коэффициент корреляции *(*r_{yx}*),* а коэффициент регрессии *(b)* и коэффициент эластичности (Э) — показатели *силы связи:* коэффициент регрессии является *абсолютной* мерой, так как имеет единицы измерения, присущие изучаемым признакам *у* и *х,* а коэффициент эластичности — *относительным* показателем силы связи, потому что выражен в процентах.

3.3. Оценка значимости параметров линейной регрессии и корреляции.

После того как уравнение линейной регрессии найдено, проводится оценка значимости как уравнения в целом, так и отдельных его параметров.

Оценка значимости уравнения регрессии в целом дается с помощью F-критерия Фишера. При этом выдвигается нулевая гипотеза, что коэффициент регрессии равен нулю, т. е. $b = 0$, и, следовательно, фактор x не оказывает влияния на результат y.

Непосредственному расчету F-критерия предшествует анализ дисперсии. Центральное место в нем занимает разложение общей суммы квадратов отклонений переменной y от среднего значения y на две части — «объясненную» и «остаточную»:

$$\sum(y-\bar{y})^2 = \sum(\tilde{y}_x - \bar{y})^2 + \sum(y - \tilde{y}_x)^2$$

Любая сумма квадратов отклонений связана с числом степеней свободы df {degrees of *freedom),* т.е. с числом свободы независимого варьирования признака. Число степеней свободы связано с числом единиц совокупности **n** и с числом определяемых по ней констант.

Существует равенство между числом степеней свободы общей, факторной и остаточной суммами квадратов. Число степеней свободы остаточной суммы квадратов при линейной регрессии составляет **n-2**. Число степеней свободы для общей суммы квадратов определяется числом единиц, и поскольку мы используем среднюю вычисленную по данным выборки, то теряем одну степень свободы, т. е. *dfобщ* = ***n*** **- 1**. Итак, имеем два равенства:

1) $\sum(y-\bar{y})^2 = \sum(\tilde{y}_x - \bar{y})^2 + \sum(y - \tilde{y}_x)^2$

2) n - 1 = 1 + (n - 2).

Разделив каждую сумму квадратов на соответствующее ей число степеней свободы, получим *средний квадрат отклонений* или *дисперсию на одну степень свободы* D.

$$D_{общ} = \frac{\sum(y-\bar{y})^2}{n-1};\; D_{факт} = \frac{\sum(\tilde{y}_x-\bar{y})^2}{1};\, D_{ост} = \frac{\sum(y-\tilde{y}_x)^2}{n-2}$$

Определение дисперсии на одну степень свободы приводит дисперсии к сравнимому виду. Сопоставляя факторную и остаточную дисперсии в расчете на одну степень свободы, получим величину F-отношения, т. е. критерий *F:*

$$F = \frac{D_{факт}}{D_{ост}}$$

Если нулевая гипотеза H_o справедлива, то факторная и остаточная дисперсии не отличаются друг от друга. Если H_o несправедлива, то факторная дисперсия превышает остаточную в несколько раз. Табличное значение F-критерия — это максимальная величина отношения дисперсий, которая может иметь место при случайном расхождении их для данного уровня вероятности наличия нулевой гипотезы. Вычисленное значение F-отношения признается достоверным (отличным от единицы), если оно больше табличного. В этом случае нулевая гипотеза об отсутствии связи признаков отклоняется и делается вывод о существенности этой связи. Если же величина *F* окажется меньше табличной, то вероятность нулевой гипотезы выше заданного уровня (например, 0,05) и она не может быть отклонена без риска сделать неправильный вывод о наличии связи. В этом случае уравнение регрессии считается статистически незначимым:

$F_{факт} < F_{табл}$, H_0 не отклоняется.

Величина F-критерия связана с коэффициентом детерминации r^2. Тогда значение F-критерия можно выразить следующим образом:

$$F = \frac{r^2}{1-r^2}(n-2)$$

В линейной регрессии обычно оценивается значимость не только уравнения в целом, но и отдельных его параметров. С этой целью по каждому из параметров определяется его стандартная ошибка: m_b и m_a.

Стандартная ошибка коэффициента регрессии параметра $\boldsymbol{m_b}$ рассчитывается по формуле:

$$m_b = \sqrt{\frac{\sum(y-\tilde{y}_x)^2/(n-2)}{\sum(x-\bar{x})^2}}$$

Отношение коэффициента регрессии к его стандартной ошибке дает t-статистику, которая подчиняется статистике Стьюдента при (n-2) степенях свободы. Эта статистика применяется для проверки статистической значимости коэффициента регрессии и для расчета его доверительных интервалов.

Для оценки значимости коэффициента регрессии его величину сравнивают с его стандартной ошибкой, т. е. определяют фактическое значение t-критерия Стьюдента, которое затем сравнивают с табличным значением при определенном уровне значимости *а* и числе степеней свободы (n - 2).

$$t_b = \frac{b}{m_b}$$

Стандартная ошибка параметра ***a*** определяется по формуле:

$$m_a = \sqrt{\frac{\sum(y-\tilde{y}_x)^2}{n-2}} * \sqrt{\frac{\sum x^2}{n*\sum(x-\bar{x})^2}}$$

Значимость линейного коэффициента корреляции проверяется на основе величины ошибки коэффициента корреляции m_r:

$$m_r = \sqrt{\frac{1-r^2}{n-2}}$$

Фактическое значение t-критерия Стьюдента определяется:

$$t_r = \frac{r_{yx}}{\sqrt{1-r^2}} * \sqrt{n-2}$$

Таким образом, проверка гипотез о значимости коэффициента регрессии и корреляции равносильна проверке гипотезы о значимости линейного уравнения регрессии.

3.4.Интервальный прогноз на основе линейного уравнения регрессии.

В прогнозных расчетах по уравнению регрессии определяется предсказываемое y_p значение как точечный прогноз y_x при $x_p = x_к$, т. е. путем подстановки в линейное уравнение регрессии соответствующего значения x. Однако точечный прогноз явно нереален, поэтому он дополняется расчетом стандартной ошибки y_x, т. е. $m_{\tilde{y}_x}$ и соответственно мы получаем интервальную оценку прогнозного значения y^*:

$$\tilde{y}_x - m_{\tilde{y}_x} \le y^* \le \tilde{y}_x + m_{\tilde{y}_x}$$

$$m_{\tilde{y}} = D_{ост} * \sqrt{\frac{1}{n}} + \sqrt{\frac{(x_k - \bar{x})^2}{\sum (x-\bar{x})^2}}$$

Рассмотренная формула стандартной ошибки предсказываемого среднего значения y при заданном значении x_k характеризует ошибку положения линии регрессии. Величина стандартной ошибки m_{yx} достигает минимума при x_k = x среднему и возрастает по мере того, как «удаляется» от x среднего в любом направлении. Иными словами, чем больше разность между x_k и x среднего, тем больше ошибка m_{yx} , с которой предсказывается среднее значение y для заданного значения x_k.Можно ожидать наилучшего результата прогноза, если признак-фактор ч находится в центре области наблюдения x, и нельзя ожидать хороших результатов прогноза при удалении x_k от x среднего. Если же значения xk оказывается за пределами наблюдаемых значений x, используемых при построении линейной регрессии, то результаты прогноза ухудшаются в зависимости от того, насколько x_k отклоняется от области наблюдаемых значений фактора x.

Для прогнозируемого значения (y теор) 95-%-ные доверительные интервалы при заданном x_k определяются выражением:

$$\Delta\tilde{y}_{xk} \pm t_{табл} * m_{\tilde{y}_x}$$

Интервал достаточно широк прежде всего за счет малого объема наблюдений.

При прогнозировании на основе уравнения регрессии следует помнить, что величина прогноза зависит не только от стандартной ошибки индивидуального значения *y,* но и от точности прогноза значения фактора X. Его величина может задаваться на основе анализа других моделей исходя из конкретной ситуации, а также анализа динамики данного фактора.

3.5. Нелинейная регрессия. Корреляция для нелинейной регрессии.

Если между экономическими явлениями существуют нелинейные соотношения, то они выражаются с помощью соответствующих нелинейных функций.

Различают два класса нелинейных регрессий:

• регрессии, нелинейные относительно включенных в анализ объясняющих переменных, но линейные по оцениваемым параметрам;

• регрессии, нелинейные по оцениваемым параметрам.

Примером нелинейной регрессии по включенным в нее объясняющим переменным могут служить следующие функции:

• полиномы разных степеней: $y = a + b*x + c*x^2+e$

равносторонняя гипербола $y = a$ + b/x+e.

К нелинейным регрессиям по оцениваемым параметрам относятся функции:

• степенная $y = a \cdot x^b \cdot e$;

• показательная $y = a \cdot b^x \cdot e$;

• экспоненциальная $y = e^{a + bx} \cdot E$.

Нелинейная регрессия по включенным переменным не имеет никаких сложностей для оценки ее параметров. Они определяются, как и в линейной регрессии, методом наименьших квадратом (МНК), ибо эти функции линейны по параметрам.

Среди нелинейных функций, которые могут быть приведены к линейному виду, в эконометрических исследованиях очень широко используется степенная функция $y = a \cdot x^b \cdot e$. Это связано с тем, что параметр b в ней имеет четкое экономическое истолкование, т. е. является *коэффициентом эластичности.* Это значит, что величина коэффициента b показывает, на сколько процентов изменится в среднем результат, если фактор изменится на 1 %.

$$Э = \frac{(b + 2cx)x}{a + bx + cx^2}$$

Уравнение нелинейной регрессии, так же как и в линейной зависимости, дополняется показателем корреляции, а именно индексом корреляции *(R)*

$$R = \sqrt{1 - \frac{\sum (y - \tilde{y}_x)^2}{\sum (y - \bar{y})^2}}$$

Величина данного показателя находится в границах: $О \le R \le 1$; чем ближе к единице, тем теснее связь рассматриваемых при признаков, тем более надежно найденное уравнение регрессии.

Поскольку в расчете индекса корреляции используется соотношение факторной и общей суммы квадратов отклонений, R^2 имеет тот же смысл, что и коэффициент детерминации. В специальных исследованиях величину R^2 для нелинейных связей называют индексом детерминации.

Оценка статистической значимости индекса корреляции проводится так же, как и оценка значимости коэффициента корреляции.

Индекс детерминации R^2 используется для проверки статистической значимости в целом уравнения нелинейной регрессии по F-критерию Фишера.

$$F = \frac{R^2}{1-R^2} * \frac{n-m-1}{m}$$

где n — число наблюдений;

m - число параметров при переменных x.

Величина m характеризует число степеней свободы для факторной суммы квадратов, а (n-m-1) — число степеней свободы для остаточной суммы квадратов.

Для степенной функции $y_x = a \cdot x^b$ значение $m = 1$ и формула F-критерия примет тот же вид, что и при линейной зависимости:

$$F = \frac{R^2}{1-R^2} * (n-2)$$

Для параболы:

$$F = \frac{R^2}{1-R^2} * \frac{n-3}{2}$$

Индекс детерминации R^2 можно сравнивать с коэффициентом детерминации для обоснования возможности применения линейной функции. Чем больше кривизна линии регрессии, тем величина коэффициента детерминации r^2 меньше индекса детерминации R^2. Близость этих показателей означает, что нет необходимости усложнять форму уравнения регрессии и можно использовать линейную функцию. Практически если величина $(R^2—r^2)$ превышает 0,1, то предположение о линейной форме связи считается оправданным. В противном случае проводится оценка существенности различия между R^2 и r^2, вычисленных по одним и тем же исходным данным, через t-критерий Стьюдента:

$$t = \frac{R^2 - r^2}{m_{|R-r|}}$$

$$m_{|R-r|} = 2 * \sqrt{\frac{(R^2 - r^2) - (R^2 - r^2)^2 * (2 - (R^2 + r^2))}{n}}$$

Если $t_{факт} > t_{табл}$, то различия между рассматриваемыми показателями корреляции существенны и замена нелинейной регрессии уравнением линейной функции невозможна. Практически если величина $t < 2$, то различия между R и r несущественны, и, следовательно, возможно применение линейной регрессии, даже если есть предположения о некоторой нелинейности рассматриваемых соотношений признаков фактора и результата.

3.6. Средняя ошибка аппроксимации.

Фактические значения результативного признака отличаются от теоретических, рассчитанных по уравнению регрессии, т. е у и y_x. Чем меньше эти отличия, тем ближе теоретические значения к эмпирическим данным, тем лучше качество модели. Величина отклонений фактических и расчетных значений результативного признака *(у - y_x)* по каждому наблюдению представляет собой ошибку аппроксимации. В отдельных случаях ошибка аппроксимации может оказаться равной нулю. Отклонения *(у -* $\tilde{y}_x$) несравнимы между собой, исключая величину, равную нулю. Для сравнения используются величины отклонений, выраженные в процентах к фактическим значениям. Поскольку *(у - y_x)* может быть величиной как положительной, так и отрицательной, ошибки аппроксимации для каждою наблюдения принято определять в процентах по модулю.

Отклонения *(у — y_x)* можно рассматривать как абсолютную ошибку аппроксимации, а

$$\left|\frac{(y-\tilde{y}_x)}{y}\right|*100$$

как относительную ошибку аппроксимации. Для того чтобы иметь общее суждение о качестве модели из относительных отклонений по каждому наблюдению, находят среднюю ошибку аппроксимации как среднюю арифметическую простую.

$$A=\frac{1}{n}*\sum\left|\frac{(y-\tilde{y}_x)}{y}\right|*100$$

Возможно и другое определение средней ошибки аппроксимации:

$$A=\frac{100}{\bar{y}}*\sqrt{\frac{\sum(y-\tilde{y}_x)^2}{n}}$$

Список литературы

Основная литература:	1. Christopher Dougherty. Introduction to Econometrics. Oxford University Press, 2011. – 573 p. 2. Gujarati D.N. Basic Econometrics. McGraw-Hill, 5th edition, 2009. – 922 p. 3.Абдуллаев О.М., Ходиев Б.Ю., Ишназаров А.И. Эконометрика. Учебник. –Т.: Fan va texnologiya. 2007. – 612 с. 4. Абдуллаев О.М., Жамалов М.С. Эконометрическое моделирование. Учебник. –Т.: Fan va texnologiya. 2010. – 612 с.
Дополнительная литература	1. Greene W.H.Econometric Analysis. Prentice Hall. 7th edition, 2011. -1232p. 2.Валентинов В,А.Эконометрика: Учебник. –М.:ИТК»Дашков и К»,2009–367с. 3.Кремер Н.Ш.Эконометрика:Учебник–М.:ЮНИТИ-ДАНА,2008. –562с. 4.Айвазян С.А. Прикладная статистика и основы эконометрика:Учебник. –М.ЮНИТИ,2007. –345с. 5.Елисеева И.И., Курышева С.В.и др. Эконометрика:Учебник–М.:Финансы и статистика, 2007. –260с. 6.Habibullayev I Iqtisodiy matematik usullsr va modellar: o‘quv qo‘llamna / O‘zbekiston Respublikasi Oliy va o‘rta maxsus ta`lim vazirligi/–Toshkent:“Tafakkur-Bo‘stoni”, 2012.112 b.
Интернет ресурсы:	www.ziyonet.uz **www.lex.uz** www.stat.uz **www.mehnat.uz** **www.mf.uz** www.tfi.uz www.mfer.uz www.conomical.uz **www.cisstat.com**

	www.gks.ru **www.unstats.un.org**

4-ТЕМА	
МНОГОФАКТОРНЫЙ ЭКОНОМЕТРИЧЕСКИЙ АНАЛИЗ	
ПЛАН:	
4.1.	**Понятие множественной регрессии и классическая линейная модель множественной регрессии. Спецификация модели**
4.2.	**Оценка параметров уравнения множественной регрессии методом наименьших квадратов**
4.3.	**Множественная и частная корреляция**
4.4.	**Оценка надежности результатов множественной регрессии и корреляции**
4.5.	**Регрессионные модели с переменой структурой (фиктивные переменные)**
4.6.	**Проверка модели на наличие гетероскедастичности (метод Гольдфельда-*Квандта)***

4.1. Понятие множественной регрессии и классическая линейная модель множественной регрессии. Спецификация модели

Парная регрессия может дать хороший результат при моделировании, если влиянием других факторов, воздействующих на объект исследования, можно пренебречь. Например, при построении модели потребления того или иного товара от дохода исследователь предполагает, что в каждой группе дохода одинаково влияние на потребление таких факторов, как цена товара, размер семьи и ее состав. Вместе с тем исследователь никогда не может быть уверен в справедливости данного предположения. Для того чтобы иметь правильное представление о влиянии дохода на потребление, необходимо изучить их корреляцию при неизменном уровне других факторов. Решение такой задачи предполагает отбор единиц совокупности с одинаковыми значениями всех других факторов, кроме дохода. Этот путь приводит к планированию эксперимента — методу, который используется в химических, физических, биологических исследованиях. Экономист в отличие от экспериментатора-естественника лишен возможности регулировать другие факторы. Поведение отдельных экономических переменных контролировать нельзя, т.е. не удается обеспечить равенство всех прочих условий для оценки влияния одного исследуемого фактора. В этом случае следует попытаться выявить влияние других факторов, введя их в модель, т.е. построить уравнение множественной регрессии

$$y = a + b_x * x_x + b_2 * x_2 + ... + b_p * x_p + e.$$

Такого рода уравнение может применяться при изучении потребления. Тогда коэффициенты *bj*— частные производные потребления *y*по соответствующим факторам xj.

$$b_1 = \frac{dy}{dx_1}, b_2 = \frac{dy}{dx_2}, b_p = \frac{dy}{dx_p}$$

Множественная регрессия широко используется в решении проблем спроса, доходности акций, при изучении функции издержек производства, в макроэкономических расчетах и целого ряда других вопросов эконометрики.

Основная цель множественной регрессии — построить модель с большим числом факторов, определив при этом влияние каждого из них в отдельности, а также совокупное воздействие их на моделируемый показатель.

4.2.Оценка параметров уравнения множественной регрессии методом наименьших квадратов

Построение уравнения множественной регрессии начинается с выбора спецификации модели. Она включает в себя два вопроса: отбор факторов и выбор вида уравнения регрессии.

Включение в уравнение множественной регрессии того или иного набора факторов связано прежде всего с представлением исследователя о природе взаимосвязи моделируемого показателясдругими экономическими явлениями. Факторы, включаемые во множественную регрессию, должны отвечать следующим требованиям:

1) быть количественно измеримы. Если необходимо включить в модель качественный фактор, не имеющий количественного измерения, то нужно придать ему количественную определенность (например, в модели урожайности качество почвы задается в виде баллов; в модели стоимости объектов недвижимости учитывается место нахождения недвижимости: районы могут быть проранжированы;

2) не должны быть коррелированы между собой и тем более находиться в точной функциональной связи.

Включение в модель факторов с высокой интеркорреляцией, когда $r_{уч1} < r_{X1X2}$, для зависимости

$$y = a + b_x * xj + b_2 * x_2 + e,$$

может привести к нежелательным последствиям — система нормальных уравнений может оказаться плохо обусловленной и повлечь за собой неустойчивость и ненадежность оценок коэффициентов регрессии.

Если между факторами существует высокая корреляция, то нельзя определить их изолированное влияние на результативный показатель, и параметры уравнения регрессии оказываются *неинтерпретируемыми.* Так, в уравнении $y = a + Ь_x * x_x + Ь_г *x_2 + е$ предполагается, что факторы x_xи x_2независимы друг от друга, т.е.. r_{x1x2}. = 0. Тогда можно говорить, что параметр b_xизмеряет силу влияния фактора x_1на результат y при неизменном значении фактора x_2. Если же r_{x1x2} = 1, то с изменением фактора x1, фактор x2 не может оставаться неизменным. Отсюда b_xи b_2нельзя интерпретировать как показатели раздельного влияния x_1и x_2 *на* y.

Включаемые во множественную регрессию факторы должны объяснить вариацию зависимой переменной. Если строится модель с набором pфакторов, то для нее рассчитывается показатель детерминации R^2, который фиксирует долю объясненной вариации результативного признака за счет рассматриваемых в регрессии p факторов. Влияние других не учтенных в модели факторов оценивается как $(1\text{-}R^2)$ с соответствующей остаточной дисперсией S^2.

При дополнительном включении в регрессию $(p+1)$ фактор коэффициент детерминации должен возрастать, а остаточная дисперсия уменьшаться.

Если этого не происходит и данные показатели практически не отличаются друг от друга, то включаемый в анализ фактор x_p,, не улучшает модель и является лишним. Насыщение модели лишними факторами не только не снижает величину остаточной дисперсии и не увеличивает показатель детерминации, но и приводит к статистической незначимости параметров регрессии по t-критерию Стьюдента.

Несмотря на то, что теоретически регрессионная модель позволяет учесть любое число факторов, практически в этом нет необходимости. Отбор факторов проводится на основе качественного теоретико-экономического анализа. Однако теоретический анализ часто не позволяет однозначно ответить на вопрос о количественной взаимосвязи рассматриваемых признаков и целесообразности включения фактора в модель. Поэтому отбор факторов обычно проводится в две стадии: на первой отбираются факторы исходя из сути проблемы; на второй — на основе матрицы показателей корреляции и определения t-статистики для параметров регрессии.

4.3.Множественная и частная корреляция

Коэффициенты интеркорреляции (т.е. корреляции междуобъясняющими переменными) позволяют исключать из модели дублирующие факторы. Считается, что две переменных явно *коллинеарны,* т. е. находятся между собой в линейной зависимости, если $r_{xixj} \geq 0.7$

Поскольку одним из условий построения уравнения множественной регрессии является независимость действия факторов, *т.е..* R_{xixj}= 0, коллинеарность факторов нарушает это условие. Если факторы явно коллинеарны, то они дублируют друг друга и один из них рекомендуется исключить из регрессии. Предпочтение при этом отдается не фактору, более тесно связанному с результатом, а тому фактору, который при достаточно тесной связи с результатом имеет наименьшую тесноту связи с другими факторами. В этом требовании проявляется специфика множественной регрессии как метода исследования комплексного воздействия факторов в условиях их независимости друг от друга.

Включение в модель мультиколлинеарных факторов нежелательно по следующим причинам:

• затрудняется интерпретация параметров множественной регрессии как характеристик действия факторов в «чистом» виде, ибо факторы коррелированы; параметры линейной регрессии теряют экономический смысл;

• оценки параметров ненадежны, обнаруживают большие стандартные ошибки и меняются с изменением объема наблюдений (не только по величине, но и по знаку), что делает модель непригодной для анализа и прогнозирования.

Для оценки мультиколлинеарности факторов может использоваться определитель матрицы парных коэффициентов корреляции между факторами.

Если бы факторы не коррелировали между собой, то матрица парных коэффициентов корреляции между ними была бы единичной, поскольку все недиагональные элементы *были бы* равны нулю.

Если же между факторами существует полная линейная зависимость и все коэффициенты корреляции равны единице, то определитель такой матрицы равен нулю.

Чем ближе к нулю определитель матрицы межфакторной корреляции, тем сильнее мультиколлинеарность факторов и надежнее результаты множественной регрессии. И, наоборот, чем ближе к единице определитель матрицы межфакторной корреляции, тем меньше мультиколлинеарность факторов.

Отбор факторов, включаемых в регрессию, является одним и из важнейших этапов практического использования методов регрессии.

4.4.Оценка надежности результатов множественной регрессии и корреляции

Наиболее широкое применение получили следующие методы построения уравнения множественной регрессии:

- метод исключения;
- метод включения
- шаговый регрессионный анализ.

При отборе факторов рекомендуется пользоваться следующим правилом: число включаемых факторов обычно в 6—7 раз меньше объема совокупности, по которой строится регрессия. Если это соотношение нарушено, то число степеней свободы остаточной вариации очень мало. Это приводит к тому, что параметры уравнения регрессии оказываются статистически незначимыми, а F-критерий меньше табличного значения.

Как и в парной зависимости, используются разные виды уравнений множественной регрессии: линейные и нелинейные

Ввиду четкой интерпретации параметров наиболее широко используются линейная и степенная функции. В линейной множественной регрессии $y_x = a + b_x * x_x + b_2 * x_2 + ... + b_p * x_p$параметры при x называются *коэффициентами «чистой» регрессии.* Они характеризуют среднее изменение результата с изменением соответствующего фактора на единицу при неизмененном значении других факторов, закрепленных на среднем уровне. (функция потребления)

Свободный член уравнения множественной линейной регрессии ее (параметр *а)* вбирает в себя информацию о прочих не учитываемых в модели факторах. Его величина экономическойитерпретации не имеет. Формально его значение предполагает то значение y, когда все $x = 0$, что практически не бывает.

В степенной функции

$$\tilde{y}_x = a * x_1^{b1} * x_2^{b2} ... * x_p^{bp}$$

коэффициенты *bj*являются коэффициентами эластичности. Они показывают, на сколько процентов в среднем изменяется результат с изменениемсоответствующего фактора на 1 % при неизменности действиядругих факторов. Этот вид уравнения регрессии получил наибольшее распространение в производственных функциях, в исследованиях спроса и потребления.

Стандартные компьютерные программы обработки регрессионного анализа позволяют перебирать различные функции и выбрать ту из них, для которой остаточная дисперсия и ошибка аппроксимации минимальны, а коэффициент детерминации максимален. Однако чем сложнее функция, тем менее интерпретируемы ее параметры.

Параметры уравнения множественной регрессии оцениваются как и в парной регрессии, методом наименьших квадратов. При его применении строится система нормальных уравнений, решение которой и позволяет получить оценки параметров регрессии. Ее решение может быть осуществлено методом определителей.Возможен и иной подход к определению параметров множественной регрессии, когда на основе матрицы парных коэффициентов корреляции строится уравнение регрессии в стандартизованном масштабе:

$$t_y = \beta_1 * t_{x1} + \beta_2 * t_{x2} + \ldots + \beta_p * t_{x\,p} + \varepsilon$$

где t_y, t_{x1},…t_{xp}- стандартизированные переменные

$$t_y = \frac{y - \bar{y}}{\sigma_y}, t_{xi} = \frac{x_i - \bar{x}_i}{\sigma_{x\,i}}$$

для которых среднее значение равно нулю: ty=txi=0, среднее квадратическое отклонение равно единице (сигмы), *B*-стандартизированные коэффициенты регрессии

Применив МНК к уравнению множественной регрессии стандартизованном масштабе, после соответствующих преобразований получим систему нормальных уравнений вида

$$\begin{aligned}
r_{yx1} &= \beta_1 + \beta_2 * r_{x2x1} + \beta_3 * r_{x3x1} + \ldots + \beta_p * r_{xpx1} \\
r_{yx2} &= \beta_1 * r_{x2x1} + \beta_2 + \beta_3 * r_{x3x2} + \ldots + \beta_p * r_{xpx2} \\
r_{yx\,p} &= \beta_1 * r_{xpx1} + \beta_2 * r_{xpx2} + \beta_3 * r_{xpx3} + \ldots + \beta_p
\end{aligned}$$

Решая ее методом определителей, найдем параметры — стандартизованные коэффициенты регрессии. Стандартизованные коэффициенты регрессии показывают на сколько сигм изменится в среднем результат, если соответствующий фактор *Xj*изменится на одну сигму при неизменном среднем уровне других факторов. В силу того, что все переменныезаданы как центрированные и нормированные, стандартизованные коэффициенты регрессии сравнимы между собой. Сравнивая их друг с другом, можно ранжировать факторы по силе их воздействия на результат. В этом основное достоинство стандартизованных коэффициентов регрессии в отличие от коэффициентов «чистой» регрессии, которые несравнимы между собой.

В парной зависимости стандартизованный коэффициент регрессии есть не что иное, как линейный коэффициент корреляции. Подобно тому, как в парной зависимости коэффициенты регрессии и корреляции связаны между собой, так и во множественной регрессии коэффициенты «чистой» регрессии b_tсвязаны со стандартизованными коэффициентами регрессии B_t, а именно

$$b_i = \beta_i \frac{\sigma_y}{\sigma_{xi}}$$

Это позволяет от уравнения регрессии в стандартизированном масштабепереходить к уравнению регрессии в натуральном масштабе переменных. Параметрa определяется как

$$a = \bar{y} - b_1 * \bar{x}_1 - b_2 * \bar{x}_2 - \ldots - b_p * \bar{x}_p$$

Содержание стандартизованных коэффициентов регрессии позволяет использовать их при отсеве факторов — из модели исключаются факторы с наименьшим значением B.

3. Ранжирование факторов, участвующих в множественной линейной регрессии, может быть проведено через стандартизованные коэффициенты регрессии. Эту же цель можно достичь с помощью частных коэффициентов корреляции для линейных связей. При нелинейной взаимосвязи исследуемых признаков эту функцию выполняют частные индексы детерминации. Кроме того, частные показатели корреляции широко используются при отборе факторов: целесообразность включения того или иного фактора в модель доказывается величиной показателя частной корреляции.

4.5.Регрессионные модели с переменой структурой (фиктивные переменные)

Частные коэффициенты (или индексы) корреляции характеризуют тесноту связи между результатом и соответствующим фактором при устранении влияния других факторов, включенных в уравнение регрессии.

Показатели частной корреляции представляют собой отношение сокращения остаточной дисперсии за счет дополнительного включения в анализ нового фактора к остаточной дисперсии, имевшей место до введения его в модель.

Если выразить остаточную дисперсию через показатель детерминации $S^2_{ocm} = \sigma^2_y (1 - r^2)$, то формула коэффициента частной корреляции примет вид:

$$r_{yx2 \bullet x1} = \sqrt{1 - \frac{1 - R^2_{yx1x2}}{1 - r^2_{yx1}}}$$

Данные показатели частной корреляции принято называть коэффициентами (индексами) частной корреляции первого порядка, ибо они фиксируют тесноту связи двух переменных и закреплении (элиминировании влияния) одного фактора. Если рассматривается регрессия с числом факторов *p,* то возможны частные коэффициенты корреляции не только первого, но и второго, третьего, ..., (р — 1) порядка, т. е. влияние фактора *x1* можно оценить при разных условиях независимости действия других факторов:

$r_{yx1 \bullet x2}$ -при постоянном действии фактора x_2;

$r_{yx1 \bullet x2x3}$- при постоянном действии факторов x_2и x_3;

$r_{yx1 \bullet x2 \ldots xp}$ - при неизменном действии всех факторов, включенных в уравнение регрессии.

Сопоставление коэффициентов частной корреляции разных порядков по мере увеличения числа включаемых факторов показывает процесс «очищения» связи результативного признаке с исследуемым фактором.

Порядок частного коэффициента корреляции определяется количествомфакторов, влияние которых исключается. Например, $r_{yx1\bullet x2}$ — коэффициент частной корреляции *первого порядка.* Соответственно коэффициенты парной корреляции называются коэффициентами *нулевого порядка.* Коэффициенты частной корреляции *более высоких порядков* можно найти через коэффициентной корреляции более низких порядков по рекуррентной формуле:

$$r_{yxi\bullet x1x2...xp} = \frac{r_{yxi\bullet x1x2...xp-1} - r_{yxp\bullet x1x2...xp-1} * r_{xixp\bullet x1x2...xp-1}}{\sqrt{(1-r^2_{yxp\bullet x1x2...xp-1})*(1-r^2_{xixp\bullet x1x2...xp-1})}}$$

Придвух факторах и i = 1 данная формула примет вид:

$$r_{yx1\bullet x2} = \frac{r_{yx1} - r_{yx2} * r_{x1x2}}{\sqrt{(1-r^2_{yx2})*(1-r^2_{x1x2})}}$$

*С*оответственно при i = 2 и двух факторах частный коэффициент корреляции y с фактором x_2можно определить по формуле

$$r_{yx2\bullet x1} = \frac{r_{yx2} - r_{yx1} * r_{x1x2}}{\sqrt{(1-r^2_{yx1})*(1-r^2_{x1x2})}}$$

Для уравнения регрессии с тремя факторами частные коэффициенты корреляции *второго порядка* определяются на основе частных коэффициентов корреляции первого порядка. Так, можно исчислить три частных коэффициента корреляции второго порядка:

$$r_{yx1\bullet x2x3}; r_{yx2\bullet x1x3}; r_{yx3\bullet x1x2}$$

каждый из которых определяется по рекуррентной формуле. Например, при i=1 имеем формулу для расчета:

$$r_{yx1\bullet x2x3} = \frac{r_{yx1\bullet x2} - r_{yx3\bullet x2} * r_{x1x3\bullet x2}}{\sqrt{(1 - r^2_{yx3\bullet x2})(1 - r^2_{x1x3\bullet x2})}}$$

Рассчитанные по рекуррентной формуле частные коэффициенты корреляции изменяются в пределах от —1 до +1, а по формулам через множественные коэффициенты детерминации от 0 до 1. Сравнение их друг с другом позволяет ранжировать факторы по тесноте их связи с результатом. Частные коэффициент корреляции, подтверждают ранжировку факторов по их воздействию на результат, на основе стандартизованных коэффициентов регрессии (B-коэффициентов) в отличие от последних дают конкретную меру тесноты связи каждого фактора с результатом в чистом виде.

В эконометрике частные коэффициенты корреляции обычно не имеют самостоятельного значения. В основном их используют на стадии формирования модели, в частности в процедуре отсева факторов. Так, при построении многофакторной модели, например методом исключения переменных на первом шаге определяется уравнение регрессии с полным набором факторов и рассчитывается матрица частных коэффициентов корреляции. На втором шаге отбирается фактор с наименьшей и несущественной по *t*-критерию Стьюдента величиной показателя частной корреляции. Исключив его из модели, строят новое уравнение регрессии. Процедура продолжается до тех пор, пока не окажется, что все частные коэффициенты корреляции существенно отличаются от нуля. Если исключен несущественный фактор, то множественные коэффициенты детерминации на двух смежных шагах построения регрессионной модели почти не отличаются друг от друга.

4. Значимость уравнения множественной регрессии в целом, так же как и в парной регрессии, оценивается с помощью *F*-критерия Фишера:

$$F = \frac{D_{факт}}{D_{ост}} = \frac{R^2}{1-R^2} * \frac{n-m-1}{m}$$

Dфакт - факторная сумма квадратов на одну степень свободы;

R^2— коэффициент (индекс) множественной детерминации;

n — число наблюдений;

m — число параметров при переменных *x* (в линейной регрессиисовпадает с числом включенных в модель факторов);

Dост, - остаточная сумма квадратов на одну степень свободы.

Оценивается значимость не только уравнения в целом, но и фактора, дополнительно включенного в регрессионную модель. Необходимость такой оценки связана с тем, что не каждый фактор, вошедший в модель, может существенно увеличивать долю объясненной вариации результативного признака. Кроме того при наличии в модели нескольких факторов они могут вводиться в модель в разной последовательности. Ввиду корреляции междуфакторами значимость одного и того же фактора может быть разной в зависимости от последовательности введения его в модель. Мерой для оценки включения фактора в модель служит частный F-критерий, т. е. F_{xi}

Частный F-критерий построен на сравнении прироста факторной дисперсии, обусловленного влиянием дополнительно включенного фактора, с остаточной дисперсией на одну степень свободы по регрессионной модели в целом.

С помощью частного F-критерия можно проверить значимость всех коэффициентов регрессии в предположении, что каждый соответствующий фактор x_i- был введен в уравнение множественной регрессии последним.

Частный F-критерий оценивает значимость коэффициентов чистой регрессии. Зная величину *Fxi*, можно определить и t-критерий для коэффициента регрессии при i-м факторе, t_{bi}, а именно $t_{bi} = \sqrt{F_{xi}}$.

Оценка значимости коэффициентов чистой регрессии поt-критерию Стьюдента может быть проведена и без расчета частных F-критериев. В этом случае, как и в парной регрессии, длякаждого фактора используется формула

$$t_{bi} = \frac{b_i}{m_{bi}}$$

где b_i, - коэффициент чистой регрессии при факторе *xi*

m_{bi}- средняя квадратическая ошибка коэффициента регрессии b_i,.

Для уравнения множественной регрессии $\tilde{y} = a + b_1 x_1 + b_2 x_2 + ...b_p x_p$ средняя квадратическая ошибка коэффициента регрессии может быть определена по следующей формуле:

$$m_{bi} = \frac{\sigma_y \sqrt{1 - R^2_{yx1...xp}}}{\sigma_{x1} \sqrt{1 - R^2_{xix1...xp}}} * \frac{1}{\sqrt{n - m - 1}}$$

Где σ_y — среднее квадратическое отклонение для признака у;

$R^2_{yx1...xp}$ — коэффициент детерминации для уравнения множественнойрегрессии;

σ_{x1}-среднее квадратическое отклонение для признака x_i,;

$R^2_{xix1...xp}$ - коэффициент детерминации для зависимости фактора *xi со* всеми другими факторами уравнения множественной регрессии;

n-m-1 — число степеней свободы для остаточной суммы квадратовотклонений.

Если величина частного F-критерия выше табличного значения, то это означает одновременно не только значимость рассматриваемого коэффициента регрессии, но и значимость частного коэффициента корреляции.

Проверка надежности парных линейных коэффициентов корреляции при помощи t-критерия Стьюдента производиться по формуле: $t_{yx1} = \frac{r_{yx1}}{m_r}$

$$mr = \sqrt{\frac{1 - r_{yx1}^2}{n - 2 - 1}}$$

Проверка надежности частных коэффициентов корреляции проводиться по формулам:

$$t_{yx1\bullet x2} = \frac{r_{yx1\bullet x2}}{m_r}$$

где $m_r = \sqrt{\frac{1 - r_{yx1\bullet x2}^2}{n - m - 1}}$

5. В большинстве случаев в модель регрессии включаются количественные факторные переменные. Однако при проведении некоторых исследований может возникнуть необходимость во включении в модель регрессии качественных факторных переменных. Это могут быть разного рода атрибутивные признаки, такие, например, как профессия, пол, образование, возраст, климатические условия, принадлежность к определенному региону. Для того, чтобы ввести такие переменные в регрессионную модель, им должны быть присвоены те или иные *цифровые метки,* т. е. качественные переменные необходимо преобразовать в количественные. Такого вида сконструированные переменные в эконометрикепринято называть *фиктивными переменными.*

Качественные признаки могут приводить к неоднородности исследуемой совокупности, что может быть учтено при моделировании двумя путями:

- регрессия строится для каждой качественно отличной группы единиц совокупности, т.е. для каждой группы в отдельности, чтобыпреодолеть неоднородность единиц общей совокупности;

- общая регрессионная модель строится для совокупности в целом, учитывающей неоднородность данных. В этом случае в регрессионную модель вводятся фиктивные переменные, т.е. строится регрессионная модель с переменной структурой, отражающей неоднородность данных.

Фиктивная переменная – это атрибутивная, или качественная, факторная переменная, которая представлена с помощью определенного цифрового кода.

Модель регрессии, включающая в себя в качестве факторной переменной фиктивную переменную, называется моделью регрессии с переменной структурой.

В качестве примера модели регрессии с переменной структурой можно привести модель регрессии размера заработной платы от стажа работников с различным образованием (среднее, среднее специальное и высшее).

Для включения факторной переменной в модель регрессии используются только две фиктивные переменные, потому что количество фиктивных переменных в модели регрессии должно быть на единицу меньше чем значений качественной переменной.

Общий вид модели с фиктивной переменной:

$$Y=a+bx+cz^2+e$$

Система уравнений для нахождения параметров линейной модели с фиктивной переменной составляется при помощи метода наименьших квадратов.

$$\begin{cases} \sum y = na + b\sum x + c\sum z, \\ \sum yx = a\sum x + b\sum x^2 + c\sum zx, \\ \sum yz = a\sum z + b\sum xz + c\sum z^2. \end{cases}$$

Решив систему, получаем уравнение регрессии. Далее проверяется значимость найденного уравнения и делается вывод о том, улучшился ли результат модели с введением в нее фиктивной переменной (R^2).

6.*Гетероскедастичность*- это предположение о неоднородности дисперсий случайных ошибок модели регрессии.

Случайная ошибкамодели регрессии - этовеличина отклонения в модели линейной множественной регрессии:

$$\varepsilon = y - \tilde{y}$$

где ε - остатки модели регрессии.

Одно из условий нормальной линейной модели множественной регрессии заключается в том, что дисперсия случайной ошибки модели регрессии является постоянной для всех наблюдений. Данное условие называется ***гомоскедастичностью*** дисперсий случайных ошибок модели регрессии.

4.6.Проверка модели на наличие гетероскедастичности (метод Гольдфельда-*Квандта)*

Гомоскедастичность*-** это предположение о постоянстве дисперсии случайной ошибки $\boldsymbol{e}_i$ для всех ***iнаблюдений модели регрессии.

Но на практике условие гомоскедастичности случайной ошибки, или остатков модели регрессии, не всегда выполняется.

Последствия гетероскедастичности остатков модели регрессии:

1) оценки нормальной линейной модели регрессии остаются несмещенными и состоятельными, но теряется эффективность;

2)появляется вероятность неверного вычисления оценок стандартных ошибок коэффициентов модели регрессии, что может привести к утверждению неверной гипотезы о значимости коэффициентов регрессии и значимости модели регрессии в целом. Обнаружить гетероскедастичность остатков модели регрессии можно путем проверки гипотез.

При малом объеме выборки, что наиболее характерно для эконометрических исследований, для оценки гетероскедастичности может использоваться метод Гольдфельда-Квандта. Для того, чтобы оценить нарушение гомоскедастичности, необходимо провести *параметрический тест,* который включает в себя несколько этапов:

1 этап.

Упорядочение***n*** наблюдений по мере возрастания переменной ***x***.

2 этап.

Исключение из рассмотрения***C*** центральных наблюдений; при этом ***(n-C):2>p***, где ***p*** – число оцениваемых параметров. Из экспериментальных расчетов, для случая одного фактора рекомендовано при ***n=30*** принимать***C=8***.

3 этап.

Разделение совокупности из ***(n-C)*** наблюдений на две группы (соответственно с малыми и большими значениями фактора ***x***) и определение по каждой из групп уравнений регрессии.

4 этап.

Определение остаточной суммы квадратов для первой (S_1) и второй (S_2) групп и нахождение их отношения: ***R=S1:S2***, где ***S1>S2***.

При выполнении нулевой гипотезы о гомоскедастичности отношение R будет удовлетворять F-критерию с ***(n-C-2p):2*** степенями свободы для каждой остаточной суммы квадратов.

Если $\mathbf{F_{факт}} > \mathbf{F_{теор}}$, то основная гипотеза отклоняется, и в основной модели регрессии присутствует гетероскедастичность, зависящая от факторной переменной ***x***.

Если $\mathbf{F_{факт}} < \mathbf{F_{теор}}$, то основная гипотеза принимается, и гетероскедастичность в основной модели регрессии не зависит от факторной переменной ***x***.

Чем больше величина **R** превышает табличное значение **F**-критерия, тем более нарушена предпосылка о равенстве дисперсий.

Список литературы

Основная литература:	1. Christopher Dougherty. Introduction to Econometrics. Oxford University Press, 2011. – 573 p. 2. Gujarati D.N. Basic Econometrics. McGraw-Hill, 5^{th} edition, 2009. – 922 p. 3.Абдуллаев О.М., Ходиев Б.Ю., Ишназаров А.И. Эконометрика. Учебник. –Т.: Fan va texnologiya. 2007. – 612 с. 4. Абдуллаев О.М., Жамалов М.С. Эконометрическое моделирование. Учебник. –Т.: Fan va texnologiya. 2010. – 612 с
Дополнительная литература	1. Greene W.H.Econometric Analysis. Prentice Hall. 7^{th} edition, 2011. -1232p. 2.Валентинов В,А.Эконометрика: Учебник. –М.:ИТК»Дашков и К»,2009–367с. 3.Кремер Н.Ш.Эконометрика:Учебник–М.:ЮНИТИ-ДАНА,2008. –562с. 4.Айвазян С.А. Прикладная статистика и основы эконометрика:Учебник. –М.ЮНИТИ,2007. –345с. 5.Елисеева И.И., Курышева С.В.и др. Эконометрика:Учебник–М.:Финансы и статистика, 2007. –260с.

	6.Habibullayev I Iqtisodiy matematik usullsr va modellar: o‘quv qo‘llamna / O‘zbekiston Respublikasi Oliy va o‘rta maxsus ta`lim vazirligi/–Toshkent:“Tafakkur-Bo‘stoni”, 2012.112 b.
Интернет ресурсы:	www.ziyonet.uz **www.lex.uz** www.stat.uz **www.mehnat.uz** **www.mf.uz** www.tfi.uz www.mfer.uz www.conomical.uz **www.cisstat.com** **www.gks.ru** **www.unstats.un.org**

5-ТЕМА
ОЦЕНКА КАЧЕСТВА ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ
ПЛАН:
5.1. Оценка существенности параметров линейной регрессии и корреляции
5.2. Интервалы прогноза по линейному уравнению регрессии

5.1.Оценка существенности параметров линейной регрессии и корреляции

Оценка значимости уравнения регрессии в целом дается с помощью F-критерия Фишера. При этом выдвигается нулевая гипотеза, что коэффициент регрессии равен нулю, т. е. $b= 0$, и, следовательно, фактор x не оказывает влияния на результат y.

Непосредственному расчету F-критерия предшествует анализ дисперсии. Центральное место в нем занимает разложение общей суммы квадратов отклонений переменной y от среднего значения y на две части — «объясненную» и «необъясненную»:

$$\sum(y-\bar{y})^2 = \sum(y_x-\bar{y})^2 + \sum(y-y_x)^2$$

Общая сумма квадратов отклонений	Сумма квадратов отклонений, объяснённая регерссией	Остаточная сумма квадратов отклонений

Общая сумма квадратов отклонений индивидуальных значений результативного признака y от среднего значения $У$ вызвана влиянием множества причин. Условно разделим всю совокупность причин на две группы: *изучаемый фактор x* и *прочие факторы.* Если фактор не оказывает влияния на результат, то линия регрессии на графике параллельна оси ox и y — y. Тогда вся дисперсия результативного признака обусловлена воздействием прочих факторов и общая сумма квадратов отклонений совпадет с остаточной. Если же прочие факторы не влияют на результат, то y связан с x функционально и остаточная сумма квадратов равна нулю. В этом случае сумма квадратов отклонений, объясненная регрессией, совпадает с общей суммой квадратов.

Поскольку не все точки поля корреляции лежат на линии регрессии, то всегда имеет место их разброс как обусловленный влиянием фактора x, т. е. регрессией y по x, так и вызванный действием прочих причин (необъясненная вариация). Пригодность линии регрессии для прогноза зависит от того, какая часть общей вариации признака y приходится на объясненную вариацию. Очевидно, что если сумма квадратов отклонений, обусловленная регрессией, будет больше остаточной суммы квадратов, то уравнение регрессии статистически значимо и фактор x оказывает существенное воздействие на результат y. Это равносильно тому, что коэффициент детерминации r^2 будет приближаться к единице.

Любая сумма квадратов отклонений связана с числом степеней свободы (*df*– *degrees of freedom*), т. е. с числом свободы независимого варьировани признака. Число степеней свободы связано с числом единиц совокупности : с числом определяемых по ней констант.

Применительно к исследуемой проблеме число степеней свободы должно показать, сколько независимых отклонений из n возможных [(y_1 — $\bar{y}$), (y_2 - $\bar{y}$), (y_n — $\bar{y}$)] требуется для образования данной суммы квадратов. Так, для общей суммы квадратов $\sum(y-\bar{y})^2=0$ требуется (n-1) независимых отклонений, ибо по совокупности из n и единиц после расчета среднего уровня свободно варьируют лишь (и — 1) число отклонений. Например, имеем ряд значений у. 1, 2, 3, 4, 5. Среднее из них равно 3, и тогда п отклонений от среднего составят: —2; —1; 0; 1; 2. Так как $\sum(y-\bar{y})=0$, то свободно варьируют лишь четыре отклонения, а пятое отклонение может быть определено, если предыдущие четыре известны.

При расчете объясненной или факторной суммы квадратов

$\sum(\ \hat{y}_x - \bar{y})^{\ 2}$ используются теоретические (расчетные) значения результативного признака $\hat{y}_x$, найденные по линии регрессии:

$$\hat{y}_x = a + b * x$$

В линейной регрессии $\sum.\ \widehat{(}y_x - \bar{y})^{\ 2}\ = b^2 * \sum(x - \bar{x}\ \ )^2$ В этом нетрудно убедиться, обратившись к формуле линейного коэффициента корреляции:

$$r_{xy} = b\frac{\sigma_x}{\sigma_y} \qquad (3.1.)$$

Из формулы (3.1.) видно, что

$$r_{xy}{}^2 = b^2\frac{\sigma^2{}_x}{\sigma^2{}_y} \qquad (3.2.)$$

Соответственно сумма квадратов отклонений, обусловленных линейной регрессией, составит:

$$\sum(y_x - \bar{y})^2 = b^2 \cdot \sum(x - \bar{x})^2$$

Поскольку при заданном объеме наблюдений по x и y факторная сумма квадратов при линейной регрессии зависит только от одной константы коэффициента регрессии b, то данная сумма квадратов имеет одну степень свободы. К этому же выводу придем, если рассмотрим содержательную сторону расчетного значения признака *y,* т. е.$\hat{y}_x$. Величина $\hat{y}_x$ определяется по уравнению линейной регрессии:. $\hat{y}_{x=a+b*x}$ Параметр *a* можно *определить как.* $\bar{y}\ \ - \ b\ *\bar{x}$.

Подставив выражение параметра a в линейную модель, получим:

$$\widehat{y_x}\ =\ \ \bar{y} - b * \bar{x} + b * x = \bar{y} - b * (x - \bar{x})$$

Отсюда видно, что при заданном наборе переменных у и х расчетное значение $\hat{y}_x$ является в линейной регрессии функцией только одного параметра — коэффициента регрессии. Соответственно и факторная сумма квадратов отклонений имеет число степеней свободы, равное 1.

Существует равенство между числом степеней свободы общей, факторной и остаточной суммами квадратов. Число степеней свободы остаточной суммы квадратов при линейной регрессии составляет n — 2. Число степеней свободы для общей суммы квадратов определяется числом единиц, и поскольку мы используем среднюю вычисленную по данным выборки, то теряем одну степень свободы, т. е. $df_{общ} = n\text{-}1$ •

Итак, имеем два равенства:

$$\sum (y - \bar{y})^2 = \sum (\hat{y}_x - \bar{y})^2 + \sum (y - \hat{y}_x)2$$

$$n - 1 = 1 + (n - 2) \tag{3.3.}$$

Разделив каждую сумму квадратов на соответствующее ей число степеней свободы, получим средний квадрат отклонений, или, что то же самое, дисперсию на одну степень свободы D.

$$D = \frac{\sum (y - \bar{y})^2}{n - 1};$$

$$D = \frac{\sum (\hat{y}_x - \bar{y})^2}{1}$$

$$D = \frac{\sum (y - \hat{y}_x)^2}{n - 2}$$

Определение дисперсии на одну степень свободы приводит дисперсии к сравнимому виду. Сопоставляя факторную и остаточную дисперсии в расчете на одну степень свободы, получим величину F-отношения (F-критерий) $F = \frac{D\ факт}{D\ ост}$ (3.4.)

где F—критерий для проверки нулевой гипотезы $H_0 = D_{факт} = D_{ост}$

Если нулевая гипотеза справедлива, то факторная и остаточная дисперсии не отличаются друг от друга. Для H_0 необходимо опровержение, чтобы факторная дисперсия превышала остаточную в несколько раз. Английским статистиком Снедекором разработаны таблицы критических значений F-отношений при разных уровнях существенности нулевой гипотезы и различном числе степеней свободы.

Табличное значение F-критерия — это максимальная величина отношения дисперсий, которая может иметь место при случайном их расхождении для данного уровня вероятности наличия нулевой гипотезы.

Вычисленное значение F-от ношения признается достоверным (отличным от единицы), если оно больше табличного. В этом случае нулевая гипотеза об отсутствии связи признаков отклоняется и делается вывод о существенности этой связи: $F_{факт} \succ F_{табл}$ H_0 отклоняется.

Если же величина окажется меньше табличной $F_{факт} \prec F_{табл}$, то вероятность нулевой гипотезы выше заданного уровня (например, 0,05) и она не может быть отклонена без серьезного риска сделать неправильный вывод о наличии связи. В этом случае уравнение регрессии считается статистически незначимым. H_0 не отклоняется.

В рассматриваемом примере:

$\sum(y-\bar{y})^2 = \sum y^2 - n\cdot\bar{y}^2 = 99700 - 7\cdot 110^2 = 15000$ общая сумма квадратов

$\sum(y_x-\bar{y})^2 = b^2\sum(x-\bar{x})^2 = 36,84^2\cdot(80-7\cdot(22\div 7)^2) =$ фактическая сумма квадратов

$\sum(y-y_x)^2 = 15000 - 14735 = 265$ Остаточная сумма квадратов

$D_{факт}$=14735

$D_{ост}$=265/5=53

F=14735/53=278

$F_{\alpha=0,05} = 6,61;\ F_{\alpha=0,01} = 16,26$

Поскольку $F_{факт} \succ F_{табл}$ как при 1%-ном, так и при 5%-ном уровне значимости, то можно сделать вывод о значимости уравнения регрессии (связь доказана).

Величина F-критерия связана с коэффициентом детерминации r^2. Факторную сумму квадратов отклонений можно представить как

$$\sum (y_x - \overline{y})^2 = r^2 \cdot \sigma_y^2 \cdot n$$

а остаточную сумму квадратов — как

$$\sum (y - y_x)^2 = (1 - r^2) \cdot \sigma_y^2 \cdot n.$$

Тогда значение F-критерия можно выразить как

$$F = \frac{r^2}{1 - r^2} \cdot (n - 2) \quad (3.5.)$$

В нашем примере r^2 = 0,982. Тогда $F = \frac{0,982}{1 - 0,982} \cdot (7 - 2) = 273$ (некоторое несовпадение с предыдущим результатом объясняется ошибками округления).

Оценка значимости уравнения регрессии обычно дается в виде таблицы дисперсионного анализа (табл. 1.1).

Таблица 1.

Дисперсионный анализ результатам регрессии

Источники вариации	Число степеней свободы	Сумма квадратов отклонений	Дисперсия на одну степень свободы	F- отношение	
				фактическое	табличное
Общая	6	15000	-	-	-
Объясненная	1	14735	14735	278	6,61
Остаточная	5	265	53	1	-

В линейной регрессии обычно оценивается значимость не только уравнения в целом, но и отдельных его параметров. С этой целью по каждому из параметров определяется его стандартная ошибка: m_b и m_a.

Стандартная ошибка коэффициента регрессии определяется по формуле

$$m_b = \sqrt{\frac{\sum (y - \hat{y}_x)^2 / (n - 2)}{\sum (x - \bar{x})^2}}.$$

где $S^2 -$ *остаточная дисперсия на одну степень свободы.*

Величина стандартной ошибки совместно с t-распределением Стьюдента при $n-2$ степенях свободы применяется для проверки существенности коэффициента регрессии и для расчета его доверительных интервалов.

Для оценки существенности коэффициента регрессии его величина сравнивается с его стандартной ошибкой, т. е. определяется фактическое значение критерия Стьюдента: $t_b = \frac{b}{m_b}$, которое затем сравнивается с табличным значением при определенном уровне значимости a и числе степеней свободы $(n — 2)$.

В рассматриваемом примере фактическое значение t-критерия для коэффициента регрессии составило:

$$t_b = \frac{b}{m_b}$$

Этот же результат получим, извлекая квадратный корень из найденного ранее F-критерия, т.е.

$$t_b = \sqrt{F} = \sqrt{278} = 16{,}67$$

Стандартная ошибка параметра а определяется по формуле:

$$m_a = \sqrt{\frac{\sum (y-\hat{y}_x)^2}{n-2} \cdot \frac{\sum x^2}{n \cdot \sum (x-\bar{x})^2}}. \quad (3.6.)$$

Процедура оценивания существенности данного параметра не отличается от рассмотренной выше для коэффициента регрессии; вычисляется t-критерий: $t_a = \frac{a}{m_a}$, его величина сравнивается с табличным значением при $df=n-2$степенях свободы.

Значимость линейного коэффициента корреляции проверяется на основе величины ошибки коэффициента корреляции m_r:

$$m_r = \sqrt{\frac{1-r^2}{n-2}} \quad (3.7.)$$

Фактическое значение t-критерия Стьюдента определяется как

$$t_r = \frac{r}{\sqrt{1-r^2}} \cdot \sqrt{n-2} \quad \mathbf{(}3.8.\mathbf{)}$$

Данная формула свидетельствует, что в парной линейной регрессии $t_r^{\,2} = F$,

ибо, как уже указывалось $F = \frac{r^2}{1-r^2} \cdot (n-2)$. Кроме того, $t_b^{\,2} = F$ Следовательно,

$t_r^{\,2} = t_b^2$.

Таким образом, проверка гипотез о значимости коэффициентов регрессии и корреляции равносильна проверке гипотезы о существенности линейного уравнения регрессии.

В рассматриваемом примере t_r не совпало с t_b в результате ошибок округлений. Величина t_r = 16,73 значительно превышает табличное значение 2,57 при a = 0,05. Следовательно, коэффициент корреляции существенно отличен от нуля и зависимость является достоверной.

Рассмотренная формула оценки коэффициента корреляции рекомендуется к применению при большом числе наблюдений и если r не близко к + 1 или —1. Если же величина коэффициента корреляции близка

к + 1, то распределение его оценок отличается от нормального или распределения Стьюдента, так как величина коэффициента корреляции ограничена значениями от —1 до +1. Чтобы обойти это затруднение, Р. Фишером было предложено для оценки существенности r ввести вспомогательную величину Z, связанную с коэффициентом корреляции следующим отношением:

$$z = \frac{1}{2} \cdot \ln \frac{1+r}{1-r} \quad (3.9.)$$

При изменении r от — 1 до +1 величина zизменяется от —∞ до +∞, что соответствует нормальному распределению. Математический анализ доказывает, что распределение величины z мало отличается от нормального даже при близких к единице значениях коэффициента корреляции. Стандартная ошибка величины z определяется по формуле

$$m_z = \frac{1}{\sqrt{n-3}} \quad (3.10.)$$

Далее выдвигаем нулевую гипотезу H_0, которая состоит в том, что корреляция отсутствует, т. е. теоретическое значение коэффициента корреляции равно нулю. Коэффициент корреляции значимо отличен от нуля, если - $\frac{z}{m_z} = t_z \succ t_{\alpha=0,05}$ т. е. если фактическое значение t_z превышает его табличное значение на уровне значимости a = 0,05 или a = 0,01.

Ввиду того, что *r и z* связаны между собой приведенным выше соотношением, можно вычислить критические значения *z*, соответствующие каждому из значений *Z*. Таблицы критических значений r разработаны для уровней значимости 0,05 и 0,01 и соответствующего числа степеней свободы. Критические значения *z* предполагают справедливость нулевой гипотезы, т. е. r мало отлично от нуля. Если фактическое значение коэффициента корреляции по абсолютной величине превышает табличное, то данное значение r считается существенным. Если же z оказывается меньше табличного, то фактическое значение r несущественно.

5.2.Интервалы прогноза по линейному уравнению регрессии

В прогнозных расчетах по уравнению регрессии определяется предсказываемое (y_p) значение как точечный прогноз $\hat{y}_x$ при x_p=x_k, т. е. путем подстановки в уравнение регрессии $\hat{y}_x = a + b * x$ соответствующего начения x. Однако точечный прогноз явно не реален. Поэтому он дополняется расчетом стандартной ошибки $\hat{y}_x$ т. е. $m_{\hat{y}x}$ и соответственно интервальной оценкой прогнозного значения (y^*)

$$y_x - m_{y_x} \leq y^* \leq y_x + m_{y_x}$$

Чтобы понять, как строится формула для определения величин стандартной ошибки$\hat{y}_x$, обратимся к уравнению линейной регрессии: $y_x = a + bx$. Подставим в это уравнение выражение параметра *a:*

$$a = \overline{y} - b \cdot \overline{x}$$

тогда уравнение регрессии примет вид:

$$a = \overline{y} - b \cdot \overline{x} + b \cdot \overline{x} = \overline{y} + b \cdot (x - \overline{x})$$

Отсюда вытекает, что стандартная ошибка m_{y_x} зависит от ошибки $\bar{y}$ и ошибки коэффициента регрессии b, т.е.:

$$m_{y_x}{}^2 = m_y^2 + m_b^2 (x - \bar{x})^2 \quad (3.11.)$$

Из теории выборки известно, что $m_y{}^2 = \frac{\sigma^2}{n}$ Используя в качестве оценки $\boldsymbol{\sigma^2}$ остаточную дисперсию на одну степень свободы $\boldsymbol{S^2}$**,** получим формулу расчета ошибки среднего значения переменной ***y:***

$$m_y{}^2 = \frac{S^2}{n} \quad (3.12)$$

Ошибка коэффициента регрессии, как уже было показано, определяется формулой:

$$m_y{}^2 = \frac{S^2}{\sum (x - \bar{x})^2}$$

Считая, что прогнозное значение фактора $\boldsymbol{x_p = x_k}$**,** получим следующую формулу расчета стандартной ошибки предсказываемого по линии регрессии значения, т. е.

$$m_{y_x}{}^2 = \frac{S^2}{n} + \frac{S^2}{\sum (x - \bar{x})^2} \cdot (x_k - \bar{x}) = S^2 \cdot \left(\frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x - \bar{x})^2} \right) \quad (3.13.)$$

Соответственно $m_{\hat{y}x}$ имеет выражение:

$$m_y{}^2 = S \cdot \sqrt{\frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x - \bar{x})^2}} \quad (3.14.)$$

Рассмотренная формула стандартной ошибки предсказываемого среднего значения у при заданном значении $\boldsymbol{x_к}$ характеризует ошибку положения линии регрессии. Величина стандартной ошибки $m_{\hat{y}x}$ **,** как видно из формулы, достигает минимума при $x_k = \bar{x}$**,** и возрастает по мере того, как «удаляется» от $\boldsymbol{\bar{x}}$ в любом направлении. Иными словами, чем больше разность между $\boldsymbol{x_к}$ и $\boldsymbol{x}$**,** тем больше ошибка $m_{\hat{y}x}$ с которой предсказывается среднее значение ***y*** для заданного значения $\boldsymbol{x_к}$. Можно ожидать наилучшие результаты прогноза, если признак-фактор $\boldsymbol{x}$ находится в центре области наблюдений $\boldsymbol{x}$ и нельзя ожидать хороших результатов прогноза при удалении $\boldsymbol{x_к}$ от $\bar{x}$. Если же значение $\boldsymbol{x_к}$ оказывается за пределами наблюдаемых значений $\boldsymbol{x}$, используемых при построении линейной регрессии, то результаты прогноза ухудшаются в зависимости от того, насколько $\boldsymbol{x_к}$ отклоняется от области наблюдаемых значений фактора х.

Для нашего примера $m_{\hat{y}x}$ составит :

$$m_{\hat{y}x} = \sqrt{53(\frac{1}{7} + \frac{(x_{k} - 3{,}143)^2}{10{,}857})}$$

При $x_k = \bar{x}$

$$m_{\hat{y}x} = \sqrt{57:7} = 2{,}75$$

При $x_k = 4$

$$m_{\hat{y}x} = \sqrt{53(\frac{1}{7} + \frac{(4-3{,}143)2}{10{,}857})} = 3{,}34$$

Соответственно $m_{\hat{y}x}$ составит эту же величину и при $\boldsymbol{x_к}$ = 2,286. Для прогнозируемого значения $\hat{\boldsymbol{y}}_{\boldsymbol{x}}$ 95%-ные доверительные интервалы при заданном $\boldsymbol{x_к}$ определяются выражением

$$\hat{y}_{x_k \pm \ t_a * m_{\hat{y}_x}}$$

Т.е. $\hat{y}_{xk} \pm 2{,}57 * 3{,}34$ *или* $\hat{y}_{xk} \pm 8{,}58$

При $x_k = 4$, прогнозное значение у составит:

$$y_{p=} 5{,}79 + 36{,}84 * 4 = 141{,}57$$

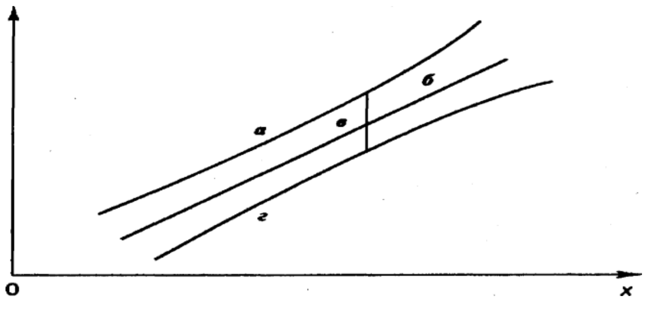
Которое представляет собой точечный прогноз.

Прогноз линии регрессии в интервале составит :

$$132{,}99 \leq \hat{y}_{xk} \leq 150{,}15$$

На графике доверительные границы для $\boldsymbol{y_x}$ представляют собой гиперболы, расположенные по обе стороны от линии регрессии (рис.2.1.).

Рис. 2.1. показывает, как изменяются пределы в зависимости от изменения $\boldsymbol{x}_к$: две гиперболы по обе стороны от линии регрессии определяют 95 % -ные



доверительные интервалы для среднего значения ***y*** при заданном значении ***x.***

Однако фактические значения ***y*** варьируют около среднего значения $\hat{y}_x$**.** Индивидуальные значения ***y*** могут отклоняться от $\hat{y}_x$на величину случайной ошибки е, дисперсия которой оценивается как остаточная дисперсия на одну степень свободы S^2 . Поэтому ошибка предсказываемого индивидуального значения ***y*** должна включать не только стандартную ошибку $m_{\hat{y}x}$но и случайную ошибку ***S.***

Рис. 2.4. Доверительный интервал линии регрессии: a — верхняя доверительная граница; $б$ - линия регрессии; $в$ - доверительный интервал для $\hat{y}_X$ при x_K\ $г$ — нижняя доверительная граница

Средняя ошибка прогнозируемого индивидуального значения у $m_{y\,i\,(x_k)}$ составит :

$$m_{yi(xk)} = S * \sqrt{1 + \frac{1}{n} + \frac{(x_{k\,-}\ \bar{x})^2)}{\sum(x - \overline{x})^2}}$$

(3.15.)

По данным рассматриваемого примера получим:

$$m_{yi(xk)} = 53 * \sqrt{1 + \frac{1}{7} + \frac{(4-3{,}143)^2}{10{,}857}} = 8{,}01$$

Доверительные интервалы прогноза индивидуальных значений;; при $\boldsymbol{x_к}$**— *4*** с вероятностью 0,95 составят: 141,57 ± 2,57 • 8,01, или 141,57 ± 20,59, это означает, что 120,98 ≤$\boldsymbol{y_p}$ ≤ 162,16.

Интервал достаточно широк прежде всего за счет малого объема наблюдений.

При прогнозировании на основе уравнения регрессии следует помнить, что величина прогноза зависит не только от стандартной ошибки индивидуального значения ***y***, но и от точности прогноза значения фактора ***x***. Его величина может задаваться на основе анализа других моделей исходя из конкретной ситуации, а также из анализа динамики данного фактора.

Рассмотренная формула средней ошибки индивидуального значения признака ***y***($m_{yi(xk)}$) может быть использована также для оценки существенности различия предсказываемого значения исходя из регрессионной модели и выдвинутой гипотезы развития событий.

Предположим, что в нашем примере с функцией издержек выдвигается предположение, что в предстоящем году в связи со стабилизацией экономики при выпуске продукции в 8 тыс. ед. затраты на производство не превысят 250 млн сум. Означает ли это действительно изменение найденной закономерности или же данная величина затрат соответствует регрессионной модели?

Чтобы ответеть на этот вопрос, найдем точечный прогноз при x=8, т.е.

$$y_x = 8 = -5{,}79 + 36{,}84 * 8 = 288{,}93$$

Предполагаемое же значение затрат, исходя из экономической ситуации, - 250,0. Для оценки существенности различия этих величин определим среднюю ошибку прогнозируемого индивидуального значения :

$$m_{yi(xk)} = S * \sqrt{1 + \frac{1}{n} + \frac{(x_{k-}\ \bar{x})^2)}{\sum(x-\bar{x})^2}} = m_{yi(xk)} = 53 * \sqrt{1 + \frac{1}{7} + \frac{(8-3{,}143)^2}{10{,}857}} = 13{,}26$$

Сравним ее с величиной предполагаемого снижения издержек производства, т. е. 38,93:

$$t = \frac{-38{,}93}{13{,}26} = -2{,}93$$

Поскольку оценивается значимость только уменьшения затрат, то используется односторонний t-критерий Стьюдента. При ошибке в 5 % с пятью степенями свободы $t_{таблц}$= 2,015. Следовательно, предполагаемое уменьшение затрат значимо отличается от прогнозируемого по модели при 95 %-ном уровне доверия. Однако если увеличить вероятность до 99 %, при ошибке в 1 % фактическое значение t-критерия оказывается ниже табличного 3,365, и рассматриваемое различие в величине затрат статистически не значимо.

Список литературы

Основная литература:	1. Christopher Dougherty. Introduction to Econometrics. Oxford University Press, 2011. – 573 p. 2. Gujarati D.N. Basic Econometrics. McGraw-Hill, 5th edition, 2009. – 922 p. 3.Абдуллаев О.М., Ходиев Б.Ю., Ишназаров А.И. Эконометрика. Учебник. –Т.: Fan va texnologiya. 2007. – 612 с. 4. Абдуллаев О.М., Жамалов М.С. Эконометрическое моделирование. Учебник. –Т.: Fan va texnologiya. 2010. – 612 с
Дополнительная литература	1. Greene W.H.Econometric Analysis. Prentice Hall. 7th edition, 2011. - 1232p. 2.Валентинов В,А.Эконометрика: Учебник. –М.:ИТК»Дашков и К»,2009–367с. 3.Кремер Н.Ш.Эконометрика:Учебник–М.:ЮНИТИ-ДАНА,2008. – 562с. 4.Айвазян С.А. Прикладная статистика и основы эконометрика:Учебник. –М.ЮНИТИ,2007. –345с. 5.Елисеева И.И., Курышева С.В.и др. Эконометрика:Учебник–М.:Финансы и статистика, 2007. –260с. 6.Habibullayev I Iqtisodiy matematik usullsr va modellar: o‘quv qo‘llamna / O‘zbekiston Respublikasi Oliy va o‘rta maxsus ta`lim vazirligi/–Toshkent:“Tafakkur-Bo‘stoni”, 2012.112 b.
Интернет ресурсы:	www.ziyonet.uz **www.lex.uz** www.stat.uz **www.mehnat.uz** **www.mf.uz** www.tfi.uz www.mfer.uz www.conomical.uz **www.cisstat.com**

6-ТЕМА

ЭКОНОМЕТРИЧЕСКАЯ МОДЕЛЬ В ВИДЕ СИСТЕМЫ УРАВНЕНИЙ

ПЛАН:

6.1. Общие понятия о системе уравнений применяемых в эконометрике

6.2. Структурная форма моделей

6.3. Оценка параметров структуры моделей

6.1.Общие понятия о системе уравнений применяемых в эконометрике.

Объектом статистического изучения в социальных науках являются сложные системы. Измерение тесноты связей между переменными, построение изолированных уравнений регрессии недостаточно для описания таких систем и объяснения механизма их функционирования. При использовании отдельных уравнений регрессии, например для экономических расчетов, в большинстве случаев предполагается, что аргументы (факторы) можно изменять независимо друг от друга. Однако это предположение является очень грубым: практически изменение одной переменной, как правило, не может происходить при абсолютной неизменности других. Ее изменение повлечет за собой изменения во всей системе взаимосвязанных признаков. Следовательно, отдельно взятое уравнение множественной регрессии не может характеризовать истинные влияния отдельных признаков на вариацию результирующей переменной. Именно поэтому в последние десятилетия в экономических, биометрических и социологических исследованиях важное место заняла проблема описания структуры связей между переменными системой так называемых одновременных уравнений, называемых также структурными уравнениями. Так, если изучается модель спроса как соотношение цен и количества потребляемых товаров, то одновременно для прогнозирования спроса необходима модель предложения товаров, в которой рассматривается также взаимосвязь между количеством и ценой предлагаемых благ. Это позволяет достичь равновесия между спросом и предложением.

Приведем другой пример.

При оценке эффективности производства нельзя руководствоваться только моделью рентабельности. Она должна бьггь дополнена моделью производительности труда, а также моделью себестоимости единицы продукции.

В еще большей степени возрастает потребность в использовании системы взаимосвязанных уравнений, если мы переходим от исследований на микроуровне к макроэкономическим расчетам. Модель национальной экономики включает в себя систему уравнений: функции потребления, инвестиций заработной платы, а также тождество доходов и т.д. Это связано с тем, что макроэкономические показатели, являясь обобщающими показателями состояния экономики, чаще всего взаимозависимы. Так, расходы на конечное потребление в экономике зависят от валового национального дохода. Вместе с тем величина валового национального дохода рассматривается как функция инвестиций. Система уравнений в эконометрических исследованиях может быть построена по-разному.

Возможна **система независимых уравнений,** когда каждая зависимая переменная (y) рассматривается как функция одного и того же набора факторов (x):

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1m}x_m + \varepsilon_1 \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2m}x_m + \varepsilon_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \ldots\ldots + a_{nm}x_m + \varepsilon_n \end{cases}$$

Набор факторов ***x(*** в каждом уравнении может варьировать. Так, модель вида

$$\begin{cases} y_1 = f(x_1, x_2, x_3, x_4, x_5) \\ y_2 = f(x_1, x_2, x_4, x_5) \\ y_3 = f(x_2, x_3, x_5) \\ y_4 = f(x_3, x_4, x_5) \end{cases}$$

также является системой независимых уравнений с тем лишь отличием, что в ней набор факторов видоизменяется в уравнениях, входящих в систему. Отсутствие того или иного фактора в уравнении системы может быть следствием как экономической нецелесообразности его включения в модель, так и несущественности его воздействия на результативный признак (незначимо значение /-критерия или частного /’-критерия для данного фактора).

Примером такой модели может служить *модель экономической эффективности сельскохозяйственного производства,* где в качестве зависимых переменных выступают показатели, характеризующие эффективность сельскохозяйственного производства, — продуктивность коров, себестоимость 1 ц молока, а в качестве факторов - специализация хозяйства, количество голов на 100 га пашни, затраты труда и т. п.

Каждое уравнение системы независимых уравнений может рассматриваться самостоятельно. Для нахождения его параметров используется метод наименьших квадратов. По существу, каждое уравнение этой системы является уравнением регрессии Поскольку никогда нет уверенности, что факторы полностью объясняют зависимые переменные, то в уравнениях присутствует свободный член л0. Так как фактические значения зависимой переменной отличаются от теоретических на величину случайной ошибки, то в каждом уравнении присутствует величина случайной ошибки.

В итоге система независимых уравнений при трех зависимых переменных и четырех факторах примет вид:

$$\begin{cases} y_1 = a_{01} + a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + \varepsilon_1 \\ y_2 = a_{02} + a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 + \varepsilon_2 \\ y_3 = a_{03} + a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 + \varepsilon_3 \end{cases}$$

Однако если зависимая переменная *у* одного уравнения выступает в виде фактора *х* в другом уравнении, то исследователь может строить модель в виде *системы рекурсивных уравнений:*

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1m}x_m + \varepsilon_1 \\ y_2 = b_{21}y_1 + a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2m}x_m + \varepsilon_2 \\ y_3 = b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots\ldots + a_{3m}x_m + \varepsilon_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = b_{n1}y_1 + b_{n2}y_2 + \ldots\ldots b_{nn-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 + \ldots\ldots + a_{nm}x_m + \varepsilon_n \end{cases}$$

В данной системе зависимая переменная *y* включает в каждое последующее уравнение в качестве факторов все зависимые переменные предшествующих уравнений наряду с набором собственно факторов *x*. Примером такой системы может служить *модель производительности труда и фондоотдачи вида*:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \varepsilon_1 \\ y_2 = b_{21}y_1 + a_{21}x_1 + a_{22}x_2 + .a_{23}x_3 + \varepsilon_2 \end{cases}$$

где *ух* -производительность труда; *у2* — фондоотдача; х1 — фондовооруженность труда; *х2* — энерговооруженность труда; *х3*— квалификация рабочих

Как и в предыдущей системе, каждое уравнение может рассматриваться самостоятельно, и его параметры определяются **методом наименьших квадратов (МНК).**

Наибольшее распространение в эконометрических исследованиях получила система взаимозависимых уравнений. В ней одни и те же зависимые переменные в одних уравнениях входят в левую часть, а в других уравнениях — в правую часть системы:

$$\begin{cases} y_1 = b_{12}y_2 + b_{13}y_3 + \ldots\ldots + b_{1n}y_n + a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1m}x_m + \varepsilon_1 \\ y_2 = b_{21}y_1 + b_{23}y_3 + \ldots\ldots + b_{2n}y_n + a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2m}x_m + \varepsilon_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = b_{n1}y_1 + b_{n2}y_2 + \ldots\ldots b_{nn-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 + \ldots\ldots + a_{nm}x_m + \varepsilon_n \end{cases}$$

Система взаимозависимых уравнений получила название **системы совместных, одновременных уравнений.** Тем самым подчеркивается, что в системе одни и те же переменные (y) одновременно рассматриваются как зависимые в одних уравнениях и как независимые в других.

6.2.Структурная форма моделей

В эконометрике эта система уравнений называется также **структурной формой модели.** В отличие от предыдущих систем каждое уравнение системы одновременных уравнений не может рассматриваться самостоятельно, и для нахождения его параметров традиционный МНК неприменим. С этой целью используются специальные приемы оценивания

Примером системы одновременных уравнений может служить ***модель динамики цены и заработной платы вида***

$$\begin{cases} y_1 = b_{12}y_2 + a_{11}x_1 + \varepsilon_1 \\ y_2 = b_{21}y_1 + a_{22}x_2 + .a_{23}x_3 + \varepsilon_2 \end{cases}$$

где y, — темп изменения месячной заработной платы;

y2 - темп изменения цен;

X1 — процент безработных

X2 — темп изменения постоянного капитала

X3 — темп изменения цен на импорт сырья

Система совместных, одновременных уравнений (или структурная форма модели) обычно содержит эндогенные и экзогенные переменные. Эндогенные переменные обозначены в приведенной ранее системе одновременных уравнений как ***y.*** Это зависимые переменные, число которых равно числу уравнений в системе. Экзогенные переменные обозначаются обычно как ***x.*** Это предопределенные переменные, влияющие на эндогенные переменные, но не зависящие от них.

Простейшая ***структурная форма модели*** *имеет вид:*

$$\begin{cases} y_1 = b_{12}y_2 + a_{11}x_1 + \varepsilon_1 \\ y_2 = b_{21}y_1 + a_{22}x_2 + \varepsilon_2 \end{cases}$$

где *y* — эндогенные переменные;

x - экзогенные переменные.

Классификация переменных на эндогенные и экзогенные зависит от теоретической концепции принятой модели. Экономические переменные могут выступать в одних моделях как эндогенные, а в других как экзогенные переменные. Внеэкономические переменные (например, климатические условия) входят в систему как экзогенные переменные. В качестве экзогенных переменных могут рассматриваться значения эндогенных переменных за предшествующий период времени (лаговые переменные). Так, потребление текущего года (у,) может зависеть не только от ряда экономических факторов, но и от уровня потребления в предыдущем году (у,).

Структурная форма модели позволяет увидеть влияние изменений любой экзогенной переменной на значения эндогенной переменной. Целесообразно в качестве экзогенных переменных выбирать такие переменные, которые могут бьггь объектом регулирования. Меняя их и управляя ими, можно заранее иметь целевые значения эндогенных переменных.

Структурная форма модели в правой части содержит при эндогенных и экзогенных переменных коэффициенты А, и ***aj***(***b***(— коэффициент при эндогенной переменной, ***aj***— коэффициент при экзогенной переменной), которые называются структурными коэффициентами модели. Все переменные в модели, выражены в отклонениях от среднего уровня, т. е. под ***x*** подразумевается ***x*— *x*** а под ***y*** — соответственно ***y* — *y*.** Поэтому свободный член в каждом уравнении системы отсутствует.

Использование МНК для оценивания структурных коэффициентов модели дает, как принято считать в теории, смещенные и несостоятельные оценки. Поэтому обычно для определения структурных коэффициентов модели структурная форма модели преобразуется в ***приведенную форму модели.***Приведенная форма модели представляет собой систему линейных функций эндогенных переменных от экзогенных:

$$\begin{cases} \hat{y}_1 = \delta_{11}x_1 + \delta_{12}x_2 + \ldots\ldots + \delta_{1m}x_m \\ \hat{y}_2 = \delta_{21}x_1 + \delta_{22}x_2 + \ldots\ldots + \delta_{2m}x_m \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \hat{y}_n = \delta_{n1}x_1 + \delta_{n2}x_2 + \ldots\ldots + \delta_{nm}x_m + \varepsilon_n \end{cases}$$

где δ_{ij} - коэффициенты приведенной формы модели.

По своему виду приведенная форма модели ничем не отличается от системы независимых уравнений, параметры которой оцениваются традиционным МНК. Применяя МНК, можно оценить ***8,*** а затем оценить значения эндогенных переменных через экзогенные.

6.3.Оценка параметров структуры моделей

Коэффициенты приведенной формы модели представляют собой нелинейные функции коэффициентов структурной формы модели. Рассмотрим это положение на примере простейшей структурной модели, выразив коэффициенты приведенной формы модели ***(80)*** через коэффициенты структурной модели Ц и ***bt).***Для упрощения в модель не введены случайные переменные.

Для структурной модели вида

$$\begin{cases} y_1 = b_{12}y_2 + a_{11}x_1 \\ y_2 = b_{21}y_1 + a_{22}x_2 \end{cases}$$

приведенная форма модели имеет вид

$$\begin{cases} y_1 = \delta_{11}x_1 + \delta_{12}x_2 \\ y_2 = \delta_{21}x_1 + \delta_{22}x_2 \end{cases}$$

в которой ***y2*** из первого уравнения структурной модели можно выразить следующим образом

$$y_2 = \frac{y_1 - a_{11}x_1}{b_{21}}$$

Отсюда имеем равенство

$$\begin{cases} y_2 = \dfrac{y_1 - a_{11}x_1}{b_{12}} \\ y_2 = b_{21}y_1 + a_{22}x_2 \end{cases}$$

Тогда

$$\frac{y_1 - a_{11} \cdot x_1}{b_{12}} = b_{21}y_1 + a_{22} \cdot x_2$$

Или

$$y_1 - a_{11} \cdot x_1 = b_{12} \cdot b_{21} \cdot y_1 + b_{12} \cdot a_{22} \cdot x_2.$$

В этом случае

$$y_1 - b_{12} \cdot b_{21} \cdot y_1 = a_{11} \cdot x_1 + b_{12} \cdot a_{22} \cdot x_2$$

Или

$$y_1 = \frac{a_{11}}{1 - b_{12} \cdot b_{21}} \cdot x_1 + \frac{a_{22} \cdot b_{12}}{1 - b_{12} \cdot b_{21}} \cdot x_2.$$

Таким образом, мы представили первое уравнение структурной формы модели в виде *уравнения приведенной формы модели*:

$$y_1 = \delta_{11} \cdot x_1 + \delta_{12} \cdot x_2$$

Из уравнения следует, что коэффициенты приведенной формы модели представляют собой нелинейные соотношения коэффициентов структурной формы модели, т. е.

$$\delta_{11} = \frac{a_{11}}{1 - b_{12} \cdot b_{21}}, \quad \delta_{11} = \frac{a_{22} \cdot b_{12}}{1 - b_{12} \cdot b_{21}}$$

Аналогично можно показать, что коэффициенты приведенной формы модели второго уравнения системы и ^2) также нелинейно связаны с коэффициентами структурной модели. Для этого выразим переменную *ух* из второго структурного уравнения модели.

Эконометрические модели обычно включают в систему не только уравнения, отражающие взаимосвязи между отдельными переменными, но и выражения тенденции развития явления, а также разного рода тождества

Так, в 1947 г., исследуя линейную зависимость потребления (с) от дохода ***(у),*** Т.Хавельмо предложил одновременно учитывать тождество дохода. В этом случае модель имеет вид:

$$\begin{cases} C = a + by \\ y = C + x \end{cases}$$

где ***x* —** инвестиции в основной капитал и в запасы экспорта и импорта

*а*и *Ь*— параметры линейной зависимости ***с*** *от у.*

Их оценки должны учитывать тождество дохода в отличие от параметров обычной линейной регрессии.

В этой модели две эндогенные переменные — ***с*** и ***у*** и одна экзогенная переменная х. Система приведенных уравнений составит:

$$\begin{cases} C = A_0 + A_1 x \\ y = B_0 + B_1 x \end{cases}$$

Она позволяет получить значения эндогенной переменной с через переменную *х.*Рассчитав коэффициенты приведенной формы модели

(Ао, А,, В0, В,), можно перейти к коэффициентам структурной модели ***а*** и ***Ь,*** подставляя в первое уравнение приведенной формы выражение переменной *х*из второго уравнения приведенной формы модели. Приведенная форма модели хотя и позволяет получить значения эндогенной переменной через значения экзогенных переменных, аналитически уступает структурной форме модели, так как в ней отсутствуют оценки взаимосвязи между эндогенными переменными

При переходе от приведенной формы модели к структурной исследователь сталкивается с проблемой идентификации. Идентификация — это единственность соответствия между приведенной и структурной формами модели. Рассмотрим проблему идентификации для случая с двумя эндогенными переменными. Как уже отмечалось, КМНК применяется в случае точно идентифицируемой структурной модели. Процедура применения КМНК предполагает выполнение следующих этапов работы. Структурная модель преобразовывается в приведенную форму модели. Для каждого уравнения приведенной формы модели обычным МНК оцениваются приведенные коэффициенты ***(S/J).*** Коэффициенты приведенной формы модели трансформи-руются в параметры структурной модели.

Рассмотрим применение КМНК для простейшей идентифицируемой эконометрической модели с двумя эндогенными и двумя экзогенными переменными

$$\begin{cases} y_1 = b_{12}y_2 + a_{11}x_1 + \varepsilon_1, \\ y_2 = b_{21}y_1 + a_{22}x_2 + \varepsilon_2. \end{cases}$$

Пусть для построения данной моде торой информацией по пяти регионам

Регион	u1	u1	x1	x2
1	2	5	1	3
2	3	6	2	1
3	4	7	3	2
4	5	8	2	5
5	6	5	4	6
Среднее	4	6,2	2,4	3,4

$$\begin{cases} y_1 = \delta_{11}x_1 + \delta_{12}x_2 + u_1 \\ y_2 = \delta_{21}x_1 + \delta_{22}x_2 + u_2 \end{cases}$$

где *u*|, «2 — случайные ошибки приведенной формы модели.

Для каждого уравнения приведенной формы модели применяем традиционный МНК и определяем (5-коэффициенты) Чтобы упростить процедуру расчетов, можно работать с отклонениями от средних уровней, ***т. с. у*=*у* — *уи х* = *х* — *Т.*** Тогда для первого уравнения приведенной формы модели система нормальных уравнений составит:

$$\begin{cases} \sum y_1 x_1 = \delta_{11} \sum x_1^2 + \delta_{12} \sum x_1 x_2 \\ \sum y_1 x_2 = \delta_{11} \sum x_1 x_2 + \delta_{12} \sum x_2^2. \end{cases}$$

Применительно к рассматриваемому примеру, используя отклонения от средних уровней, имеем:

$$\begin{cases} 6 = 5{,}2 \cdot \delta_{11} + 4{,}2 \cdot \delta_{12} \\ 10 = 4{,}2\delta_{11} + 17{,}2\delta_{12}. \end{cases}$$

Решая данную систему, получим следующее первое уравнение приведенной формы модели:

$$y_1 = 0{,}82x_1 + 0{,}373x_2 + u_1$$

Аналогично применяем МНК для второго уравнения приведенной формы модели, получим:

$$\begin{cases} \sum y_2 \cdot x_1 = \delta_{21} \sum x_1^2 + \delta_{22} \sum x_1 \cdot x_2, \\ \sum y_2 \cdot x_2 = \delta_{21} \sum x_1 x_2 + \delta_{22} \sum x_2^2. \end{cases}$$

Система нормальных уравнений составит:

$$\begin{cases} -0{,}4 = 5{,}2 \cdot \delta_{21} + 4{,}2 \cdot \delta_{22} \\ -0{,}4 = 4{,}2 \cdot \delta_{21} + 17{,}2 \cdot \delta_{22}. \end{cases}$$

$$y_2 = -0{,}072 \cdot x_1 - 0{,}00557 \cdot x_2 + u_2.$$

Применительно к нашему примеру имеем

$$\begin{cases} y_1 = 0{,}852 \cdot x_1 + 0{,}373 \cdot x_2 + u_1 \\ y_2 = -0{,}072 \cdot x_1 - 0{,}00557 \cdot x_2 + u_2 \end{cases}$$

Список литературы

Основная литература:	1. Christopher Dougherty. Introduction to Econometrics. Oxford University Press, 2011. – 573 p. 2. Gujarati D.N. Basic Econometrics. McGraw-Hill, 5th edition, 2009. – 922 p. 3. Абдуллаев О.М., Ходиев Б.Ю., Ишназаров А.И. Эконометрика. Учебник. –Т.: Fan va texnologiya. 2007. – 612 с. 4. Абдуллаев О.М., Жамалов М.С. Эконометрическое моделирование. Учебник. –Т.: Fan va texnologiya. 2010. – 612 с
Дополнительная	1. Greene W.H.Econometric Analysis. Prentice Hall. 7th edition, 2011.

литература	-1232р. 2.Валентинов В,А.Эконометрика: Учебник. –М.:ИТК»Дашков и К»,2009–367с. 3.Кремер Н.Ш.Эконометрика:Учебник–М.:ЮНИТИ-ДАНА,2008. –562с. 4.Айвазян С.А. Прикладная статистика и основы эконометрика:Учебник. –М.ЮНИТИ,2007. –345с. 5.Елисеева И.И., Курышева С.В.и др. Эконометрика:Учебник–М.:Финансы и статистика, 2007. –260с. 6.Habibullayev I Iqtisodiy matematik usullsr va modellar: o‘quv qo‘llamna / O‘zbekiston Respublikasi Oliy va o‘rta maxsus ta`lim vazirligi/–Toshkent:“Tafakkur-Bo‘stoni”, 2012.112 b.
Интернет ресурсы:	www.ziyonet.uz **www.lex.uz** www.stat.uz **www.mehnat.uz** **www.mf.uz** www.tfi.uz www.mfer.uz www.conomical.uz **www.cisstat.com** **www.gks.ru** **www.unstats.un.org**

7-ТЕМА	
ЭКОНОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ ВО ВРЕМЕННЫХ РЯДАХ	
ПЛАН:	
7.1.	**Основные элементы временного ряда**
7.2.	**Автокорреляция уровней временного ряда и выявление его структуры**
7.3.	**Моделирование сезонных и циклический колебаний**

7.1.Основные элементы временного ряда

Можно построить эконометрическую модель, используя два типа исходных данных:

• данные, характеризующие совокупность различных объектов в определенный момент (период) времени;

• данные, характеризующие один объект за ряд последовательных моментов (периодов) времени.

Модели, построенные по данным первого типа, называются пространственными моделями. Модели, построенные на основе второго типа данных, называются моделями временных рядов.

Временной ряд - это совокупность значений какого-либо показателя за несколько последовательных моментов или периодов времени.

Каждый уровень временного ряда формируется под воздействием большого числа факторов, которые условно можно подразделить на три группы:

• факторы, формирующие тенденцию ряда;

• факторы, формирующие циклические колебания ряда;

• случайные факторы.

В отечественной литературе используются два синонима этого термина: «динамический ряд» и «ряд динамики».

При различных сочетаниях в изучаемом явлении или процессе этих факторов зависимость уровней ряда от времени может принимать различные формы. Во-первых, большинство временных рядов экономических показателей имеют тенденцию, характеризующую совокупное долговременное воздействие множества факторов на динамику изучаемого показателя. Очевидно, что эти факторы, взятые в отдельности, могут оказывать разнонаправленное воздействие на исследуемый показатель. Однако в совокупности они формируют его возрастающую или убывающую тенденцию. На рис. S.1 а) показан гипотетический временной ряд, содержащий возрастающую тенденцию.

Во-вторых, изучаемый показатель может быть подвержен циклическим колебаниям. Эти колебания могут носить сезонный характер, поскольку экономическая деятельность ряда отраслей экономики зависит от времени года (например, цены на сельскохозяйственную продукцию в летний период выше, чем в зимний; уровень безработицы в курортных городах в зимний период выше по сравнению с летним). При наличии больших массивов данных за длительные промежутки времени можно выявить циклические колебания, связанные с общей динамикой конъюнктуры рынка, а также с фазой-бизнес цикла, в которой находится экономика страны. На рис. S. 1 б) представлен гипотетический временной ряд, содержащий только сезонную компоненту.

Некоторые временные ряды не содержат тенденции и циклической компоненты, а каждый следующий их уровень образуется как сумма среднего уровня ряда и некоторой (положительной или отрицательной) случайной компоненты. Пример ряда, содержащего только случайную компоненту, приведен на рис. 5.1 в).

Очевидно, что реальные данные не следуют целиком и полностью из каких-либо описанных выше моделей. Чаще всего они содержат все три компоненты. Каждый их уровень формируется под воздействием тенденции, сезонных колебаний и случайной компоненты.

В большинстве случаев фактический уровень временного ряда можно представить как сумму или произведение трендовой, циклической и случайной компонент. Модель, в которой временной ряд представлен как сумма перечисленных компонент, называется аддитивной моделью временного ряда. Модель, в которой временной ряд представлен как произведение перечисленных компонент, называется мультипликативной моделью временного ряда.

Основная задача эконометрического исследования от дельного временного ряда — выявление и придание количественного выражения каждой из перечисленных выше компонент с тем, чтобы использовать полученную информацию для прогнозирования будущих значений ряда или при построении моделей взаимосвязи двух или более временных рядов.

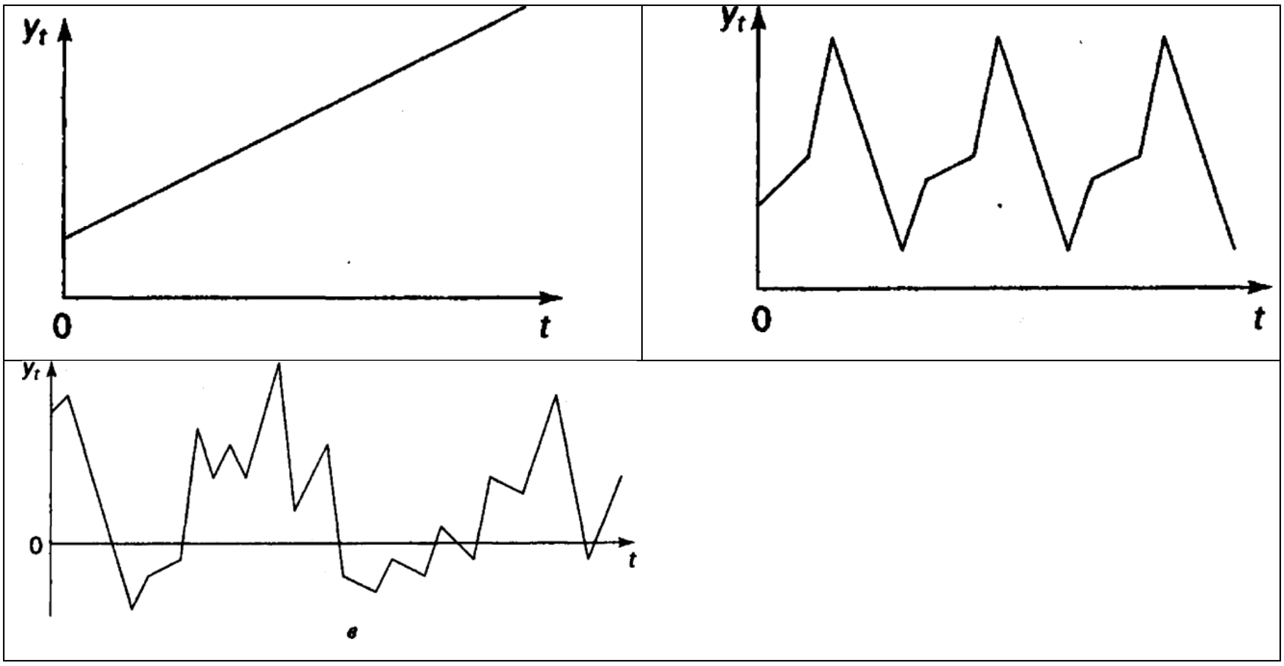


Рис.1. Основные компоненты временного ряда: а - возрастающая тенденция; б - сезонная компонента; в - случайная компонента

7.2.Автокорреляция уровней временного ряда и выявление его структуры

Одним из наиболее распространенных способов моделирования тенденции временного ряда является построение аналитической функции, характеризующей зависимость уровней ряда от времени, или тренда. Этот способ называют аналитическим выравниванием временного ряда.

Поскольку зависимость от времени может принимать разные формы, для ее формализации можно использовать различные виды функций. Для построения трендов чаще всего применяются следующие функции:

- линейный тренд: $\hat{y}_t = a + b \cdot t$;
- гипербола: $\hat{y}_t = a + b / t$;
- экспоненциальный тренд: $\hat{y}_t = e^{a+b \cdot t}$;
- тренд в форме степенной функции $\hat{y}_t = a \cdot t^b$;
- парабола второго и более высоких порядков $\hat{y}_t = a + b_1 \cdot t + b_2 \cdot t^2 + ... + b_k \cdot t^k$.

Параметры каждого из перечисленных выше трендов можно определить обычным МНК, используя в качестве независимой переменной время t=1,2,…,n, а в качестве зависимой переменной – фактические уровни временного ряда y_t. Для нелинейных трендов предварительно проводят стандартную процедуру их линеаризации.

Существует несколько способов определения типа тенденции. К числу наиболее распространенных способов относятся качественный анализ изучаемого процесса, построение и визуальный анализ графика зависимости уровней ряда от времени, расчет некоторых основных показателей динамики. В этих же целях можно использовать и коэффициенты автокорреляции уровней ряда. Тип тенденции можно определить путем сравнения коэффициентов автокорреляции первого порядка, рассчитанных по исходным и преобразованным уровням ряда. Если временной ряд имеет линейную тенденцию, то его соседние уровни y, и y, x тесно коррелируют. В этом случае коэффициент автокорреляции первого порядка уровней исходного ряда должен быть высоким. Если временной ряд содержит нелинейную тенденцию, например, в форме экспоненты, то коэффициент автокорреляции первого порядка по логарифмам уровней исходного ряда будет выше, чем соответствующий коэффициент, рассчитанный по уровням ряда. Чем сильнее выражена нелинейная тенденция в изучаемом временном ряде, тем в большей степени будут различаться значения указанных коэффициентов.

1. Другая формулировка этой формы тренда имеет вид: $\hat{y}_t = a + b \cdot t$;

Выбор наилучшего уравнения в случае, если ряд содержит нелинейную тенденцию, можно осуществить путем перебора основных форм тренда, расчета по каждому уравнению скорректированного коэффициента детерминации Л и выбора уравнения тренда с максимальным значением скорректированного коэффициента детерминации. Реализация этого метода относительно проста при компьютерной обработке данных.

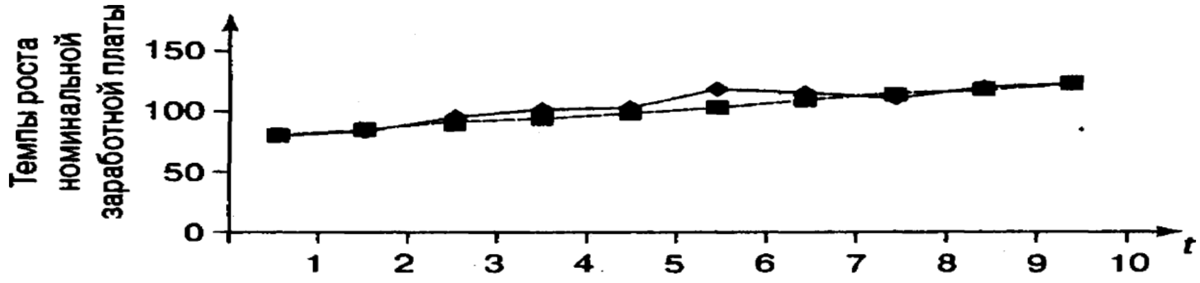
Пример 2. Расчет параметров тренда.

Имеются помесячные данные о темпах роста номинальной заработной платы за 10 месяцев 2015 г. в процентах к уровню декабря 2014 г. (табл. 2). Требуется выбрать наилучший тип тренда и определить его параметры.

Таблица 2

Месяцы	**Темпы роста номинальной месячной заработной платы**	**Месяцы**	**Темпы роста номинальной месячной заработной платы**
Январ	82,9	Июнь	121,6
Февраль	87,3	Июль	118,6
Март	99,4	Августь	114,1
Апрель	104,8	Сентябрь	123,0
Май	107,2	Октябрь	127,3

Построим график данного временного ряда (рис. 3).



Время, мес.

—♦ фактические уровни ряда

—■— уровни ряда, рассчитанные по линейному тренду

Рис. 3. Динамика темпов роста номинальной заработной платы за 10 месяцев 2015 г.

На графике рис. 3 наглядно видно наличие возрастающей тенденции. Возможно существование линейного тренда.

Для дальнейшего анализа определим коэффициенты автокорреляции по уровням этого ряда и их логарифмам (табл.6).

Таблица 6

Автокорреляционная функция временного ряда темпов роста номинальной месячной заработной платы за 10 месяцев 2015 г., % к уровню декабря 2014 г.

Лаг	Автокорреляционная функция	
	по уровням ряда	по логарифмам уровней ряда
1	0,901	0,914
2	0,805	0,832
3	0,805	0,896

Высокие значения коэффициентов автокорреляции первого, второго и третьего порядков свидетельствуют о том, что ряд содержит тенденцию. Приблизительно равные значения коэффициентов автокорреляции по уровням этого ряда и по логарифмам уровней позволяют сделать следующий вывод: если ряд содержит нелинейную тенденцию, то она выражена в неявной форме. Поэтому для моделирования его тенденции в равной мере целесообразно использовать и линейную, и нелинейную функции, например степенной или экспоненциальный тренд.

Для выявления наилучшего уравнения тренда определим параметры основных видов трендов. Результаты этих расчетов представлены в табл. 5.7, согласно данным которой наилучшей является степенная форма тренда, для которой значение скорректированного коэффициента детерминации наиболее высокое. Уравнение степенного тренда можно использовать как в линеаризованном виде, так и в форме исходной степенной функции после проведения операции потенцирования. В исходном виде это уравнение выглядит следующим образом:

$$\hat{y}_t = e^{4,39} \cdot t^{0,193}$$

или

$$\hat{y}_t = 80,32 \cdot t^{0,193}$$

Таблица 3

Уравнения трендов временного ряда темпов роста номинальной месячной заработной платы за 10 месяцев 2015 г., % к уровню декабря 2014 г.

Виды тренда	**Уравнение**	R^2	$\bar{R}^2$
Линейная	$\hat{y}_t = 82,66 + 4,72 \cdot t$ $(0,595)^*$	0,887	0,873
Парабола второго порядка	$\hat{y}_t = 72,9 + 9,599 \cdot t - 0,444 \cdot t^2$ $(2,11)^* (0,187)^*$	0,937	0,920
Показательная	$\widehat{\ln y_t} = 4,39 + 0,193 \ln t$ $(0,017)^*$	$0,939^{**}$	$0,931^{**}$
Экспоненциальная	$\widehat{\ln y_t} = 4,43 + 0,045t$ $(0,006)^*$	$0,872^{**}$	$0,856^{**}$
Гиперболическая	$\hat{y}_t = 1,22,57 - 47,63 / t$ $(8,291)^{\bullet}$	0,758	0,728

*В скобках указаны стандартные ошибки коэффициентов регрессии. Коэффициенты детерминации рассчитаны по линеаризованным уравнениям репрессии.

Во-вторых, в соответствии с интерпретацией параметров ли-нейного тренда каждый последующий уровень ряда есть сумма предыдущего уровня и среднего цепного абсолютного прироста, т. е.

$$\hat{y}_2^{линейн} = \hat{y}_1^{линейн} + b = 87,38 + 4,72 = 92,10;$$
$$\hat{y}_3^{линейн} = \hat{y}_2^{линейн} + b = 92,10 + 4,72 = 96,82 \text{ и т.д.}$$

График линейного тренда приведен на рис. 3.

Параметры экспоненциального тренда имеют следующую интерпретацию. Параметр о - это начальный уровень временного ряда в момент времени t= 0. Величина е* - это средний за единицу времени коэффициент роста уровней ряда.

Для нашего примера уравнение экспоненциального тренда в исходной форме имеет вид:

$$\hat{y}_1^{линейн} = 82,66 + 4,72 \cdot 1 = 87,38;$$
$$\hat{y}_2^{линейн} = 82,66 + 4,72 \cdot 2 = 92,10.$$

Таким образом, начальный уровень ряда в соответствии с уравнением экспоненциального тренда составляет 83,96 (сравните с начальным уровнем 82,66 в линейном тренде), а средний цепной коэффициент роста - 1,046. Следовательно, можно сказать, что темпы роста номинальной месячной заработной платы за 10 месяцев 1999 г. изменялись от уровня 83,96% со средним за месяц цепным темпом роста, равным 104,6%. Иными словами, средний за месяц цепной темп прироста временного ряда составил 4,6%.

По аналогии с линейной моделью расчетные значения уровней ряда по экспоненциальному тренду можно получить как путем подстановки в уравнение тренда значений t=1,2,..., я, так и в соответствии с интерпретацией параметров экспоненциального тренда: каждый его последующий уровень есть произведение предыдущего уровня на соответствующий коэффициент роста:

$y^{эксп}{}_1 = y^{эксп}{}_0 \bullet 1,046 = 83,96 * 1,046 = 87,82;$

$y^{эксп}{}_2 = y^{эксп}{}_1 \bullet 1,046 = 87,82 \bullet 1,046 = 91,87$ и т. д.

При наличии неявной нелинейной тенденции следует дополнять описанные выше методы выбора наилучшего уравнения тренда качественным анализом динамики изучаемого показателя, с тем чтобы избежать ошибок спецификации при выборе вида тренда. Качественный анализ предполагает изучение проблем возможного наличия в исследуемом временном ряде поворотных точек и изменения темпов прироста, или ускорения темпов прироста, начиная с определенного момента (периода) времени под влиянием ряда факторов, и т. д. В случае если уравнение тренда выбрано неверно при больших значениях *t*,результаты анализа и прогнозирования динамики временного ряда с использованием выбранного уравнения будут недостоверными вследствие ошибки спецификации.

Иллюстрация возможного появления ошибки спецификации приводится на рис. 4

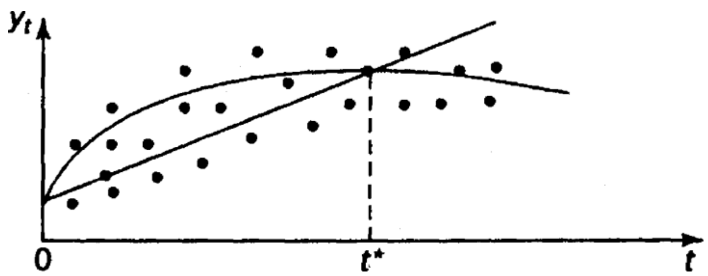


Рис. 4. Ошибка спецификации при выборе уравнения тренда

Если наилучшей формой тренда является парабола второго порядка, в то время как на самом деле имеет место линейная тенденция, то при больших *t* парабола й линейная функция будут по-разному описывать тенденцию в уровнях ряда. При t>t*парабола второго порядка характеризует убывающую тенденцию в уровнях ряда *у,,* а линейная функция - возрастающую.

7.3.Моделирование сезонных и циклический колебаний

Существует несколько подходов к анализу структуры временных рядов, содержащих сезонные или циклические колебания.

Простейший подход - расчет значений сезонной компоненты методом скользящей средней и построение аддитивной или мультипликативной модели временного ряда. Общий вид аддитивной модели следующий:

$$Y = T + S + E \quad (5)$$

Моделирование циклических колебаний в целом осуществляется аналогично моделированию сезонных колебаний, поэтому мы рассмотрим только методы моделирования последних.

Эта модель предполагает, что каждый уровень временного ряда может быть представлен как сумма трендовой (7), сезонной (5) и случайной (E) компонент. Общий вид мультипликативной модели выглядит так:

$$Y = T \cdot S \cdot E \qquad (6)$$

Эта модель предполагает, что каждый уровень временного ряда может быть представлен как произведение трендовой (T), сезонной (S) и случайной (E) компонент. Выбор одной из двух моделей осуществляется на основе анализа структуры сезонных колебаний. Если амплитуда колебаний приблизительно постоянна, строят аддитивную модель временного ряда, в которой значения сезонной компоненты предполагаются постоянными для различных циклов. Если амплитуда сезонных колебаний возрастает или уменьшается, строят мультипликативную модель временного ряда, которая ставит уровни ряда в зависимость от значений сезонной компоненты.

Построение аддитивной и мультипликативной моделей сводится к расчету значений T, S и E для каждого уровня ряда.

Процесс построения модели включает в себя следующие шаги.

1. Выравнивание исходного ряда методом скользящей средней.
2. Расчет значений сезонной компоненты S.
3. Устранение сезонной компоненты из исходных уровней ряда и получение выровненных данных (T+ E) в аддитивной или (T E) в мультипликативной модели.
4. Аналитическое выравнивание уровней (T+ E) или (T- E) и расчет значений T с использованием полученного уравнения тренда.
5. Расчет полученных по модели значений (T+ S) или (T S).

6. Расчет абсолютных и/или относительных ошибок.

Если полученные значения ошибок не содержат автокорреляции, ими можно заменить исходные уровни ряда и в дальнейшем использовать временной ряд ошибок E для анализа взаимосвязи исходного ряда и других временных рядов.

Подробнее методику построения каждой из моделей рассмотрим на примерах.

Призер 5.4. Построение аддитивной модели временного ряда.

Обратимся к данным об объеме потребления электроэнергии жителями района за последние четыре года, представленным в табл. 3.

В примере 5.2 было показано, что данный временной ряд содержит сезонные колебания периодичностью 4. Объемы потребления электроэнергии в осенне-зимний период времени (I и IV кварталы) выше, чем весной и летом (II и III кварталы). По графику, этого ряда (рис. 5.2) можно установить наличие приблизительно равной амплитуды колебаний. Это свидетельствует о возможном существовании в ряде аддитивной модели. Рассчитаем ее компоненты.

Шаг 1. Проведем выравнивание исходных уровней ряда методом скользящей средней. Для этого:

a) просуммируем уровни ряда последовательно за каждые четыре квартала со сдвигом на один момент времени и определим условные годовые объемы потребления электроэнергии (гр. 3 табл. 5.8);

b) разделив полученные суммы на 4, найдем скользящие средние (гр. 4 табл.5.8). Отметим, что полученные таким образом выравненные значения уже не содержат сезонной компоненты;

c) приведем эти значения в соответствие с фактическими моментами времени, для чего найдем средние значения из двух последовательных скользящих средних - центрированные скользящие средние (гр. 5 табл. 8).

Таблица 8

Расчет оценок сезонной компоненты в аддитивной модели

№ квартала, t	Потребление электроэнергии, y_t	Итого за четыре квартала	Скользящая средняя за четыре квартала	Централизованная скользящая средняя	Оценка сезонной компоненты
1	2	3	4	5	6
1	6,0	-	-	-	-
2	4,4	24,4	6,10	-	-
3	5,0	25,6	6,40	6,250	-1,250
4	9,0	26,0	6,50	6,450	2,550
5	7,2	27,0	6,75	6,625	0,575
6	4,8	28,0	7,00	6,875	-2,075
7	6,0	28,8	7,20	7,100	-1,100
8	10,0	29,6	7,40	7,300	2,700
9	8,0	30,0	7,50	7,450	0,550
10	5,6	31,0	7,75	7,625	-2,025
11	6,4	32,0	8,00	7,875	-1,475
12	11,0	33,0	8,25	8,125	2,875
13	9,0	33,6	8,40	8,325	0,675
14	6,6	33,4	8,35	8,375	-1,775
15	7,0	-	-	-	-
16	10,8	-	-	-	-

Шаг 2. Найдем оценки сезонной компоненты как разность между фактическими уровнями ряда и центрированными скользящими средними (гр. 6 табл. 5.8). Используем эти оценки для расчета значений сезонной компоненты S(табл. 5.9). Для этого найдем средние за каждый квартал (по всем годам) оценки сезонной компоненты S. В моделях с сезонной компонентой обычно предполагается, что сезонные воздействия за период взаимопогашаются. В аддитивной модели это выражается в том, что сумма значений сезонной компоненты по всем кварталам должна быть равна нулю.

Таблица 9

Расчет значений сезонной компоненты в аддитивной модели

Показатели	Годы	№ квартала, *i*			
		I	II	III	IV
	1	-	-	-1,250	2,550
	2	0,575	-2,075	-1,100	2,700
	3	0,550	-2,025	-1,475	2,875
	4	0,675	-1,775	-	-
Итого за i-й квартал (за все годы)		1,800	-5,875	-3,825	8.125
Средняя оценка сезонной компоненты для i-го квартала, S',		0,600	-1,958	-1,275	2,708
Скорректированная сезонная компонента, S_t		0,581	-1,977	-1,294	2,690

Для данной модели имеем:

0,6 - 1,958 - 1,275 + 2,708 = 0,075.

Определим корректирующий коэффициент:

к - 0,075 / 4 = 0,01875.

Рассчитаем скорректированные значения сезонной компоненты как разность между ее средней оценкой и корректирующим коэффициентом к:

Проверим условие равенства нулю суммы значений сезонной компоненты:

$$0{,}581 - 1{,}977 - 1{,}294 + 2{,}690 = 0.$$

Таким образом, получены следующие значения сезонной компоненты:

I квартал: $S_1 = 0{,}581$;;

II квартал: $S_2 = -1{,}979$;

IIIквартал: $S_3 = -1{,}294$;

IV квартал: $S_4 = 2{,}690$.

Занесем полученные значения в табл. 5.9 для соответствующих кварталов каждого года (стр. 3).

Шаг 3. Элиминируем влияние сезонной компоненты, вычитая ее значение из каждого уровня исходного временного ряда. Получим величины $T+E = Y - S$ (гр. 4 табл.10). Эти значения рассчитываются за каждый момент времени и содержат только тенденцию и случайную компоненту.

Таблица 10

Расчет выравненных значений Т н ошибок Е в аддитивной модели

t	y_t	S_t	$T+E=$ $= y_t - S_t$	T	$T+S$	$E=y_t -$ $-(T+S)$	E^2
1	2	3	4	5	6	7	8
1	6,0	0,581	0,419	5,902	6,483	-0,483	0,2333
2	4,4	-1,977	6,337	6,088	4,111	0,289	0,0835
3	5,0	-1,294	6,294	6,275	4,981	0,019	0,0004
4	9,0	2,690	6,310	6,461	9,151	-0,151	0,0228
5	7,2	0,581	6,619	6,648	7,229	-0,029	0,0008
6	4,8	-1,977	6,777	6,834	4,857	-0,057	0,0032
7	6,0	-1,294	7,294	7,020	5,727	0,273	0,0745
8	10,0	2,690	7,310	7,207	9,896	0,104	0,0108
9	8,0	0,581	7,419	7,393	7,974	0,026	0,0007
10	5,6	-1977	7,577	7,580	5,603	-0,030	0,0009
11	6,4	-1,294	7,694	7,766	6,472	-0,072	0,0052
12	11,0	2,690	8,310	7,952	10.642	0,358	0,1282
13	9,0	0,581	8,419	8,139	8,720	0,280	0,0784
14	6,6	-1,977	8,577	8,325	6,348	0,252	0,0635
15	7,0	-1,294	8,294	8,519	7,218	-0,218	0,0475
16	10,8	2,690	8,110	8,698	11,388	-0,588	0,3457

Определим компоненту Т данной модели. Для этого проведем аналитическое выравнивание ряда (Т + Е) с помощью линейного тренда. Результаты аналитического выравнивания следующие:

Константа 5,715416

Коэффициент регрессии 0,186421

Стандартная ошибка коэффициента регрессии 0,015188

Л-квадрат 0,914971

Число наблюдений 16

Число степеней свободы 14

Таким образом, имеем следующий линейный тренд:

$$T = 5{,}715 + 0{,}186 \cdot t .$$

Подставляя в это уравнение значения t= 1, ..., 16, найдем уровни Т для каждого момента времени (гр. 5 табл. 5.10).

Шаг 5. Найдем значения уровней ряда, полученные по аддитивной модели. Для этого прибавим к уровням Т значения сезон ной компоненты для соответствующих кварталов.

Шаг 6. В соответствии с методикой построения аддитивной модели расчет ошибки производится по формуле

$$E = Y - (T + S)$$

Это абсолютная ошибка. Численные значения абсолютных ошибок приведены в гр. 7 табл.10.

По аналогии с моделью регрессии для оценки качества построения модели или для выбора наилучшей модели можно применять сумму квадратов полученных абсолютных ошибок. Для данной аддитивной модели сумма квадратов абсолютных ошибок равна 1,10. По отношению к общей сумме квадратов отклонений уровней ряда от его среднего уровня, равной 71,59, эта величина составляет чуть более 1,5%:

$$(1 - 1{,}10 \ / \ 71{,}59) \cdot 100 = 1{,}536\%.$$

Следовательно, можно сказать, что аддитивная модель объясняет 98,5% общей вариации уровней временного ряда потребления электроэнергии за последние 16 кварталов.

Пример 5. Построение мультипликативной модели временного ряда.

Пусть имеются поквартальные данные о прибыли компании за последние четыре года (табл.11).

Таблица 11

Прибыль компаний, тыс. долл. США

Квартал / Год	I	II	III	IV
1	72	100	90	64
2	70	92	80	58
3	62	.80	68	48
4	52	60	50	30

График данного временного ряда (рис. 6) свидетельствует о наличии сезонных колебаний (период колебаний равен 4) и общей убывающей тенденции уровней ряда. Прибыль компании в весенне-летний период выше, чем в осенне-зимний период. Поскольку амплитуда сезонных колебаний уменьшается, можно предположить существование мультипликативной модели.

Определим ее компоненты.

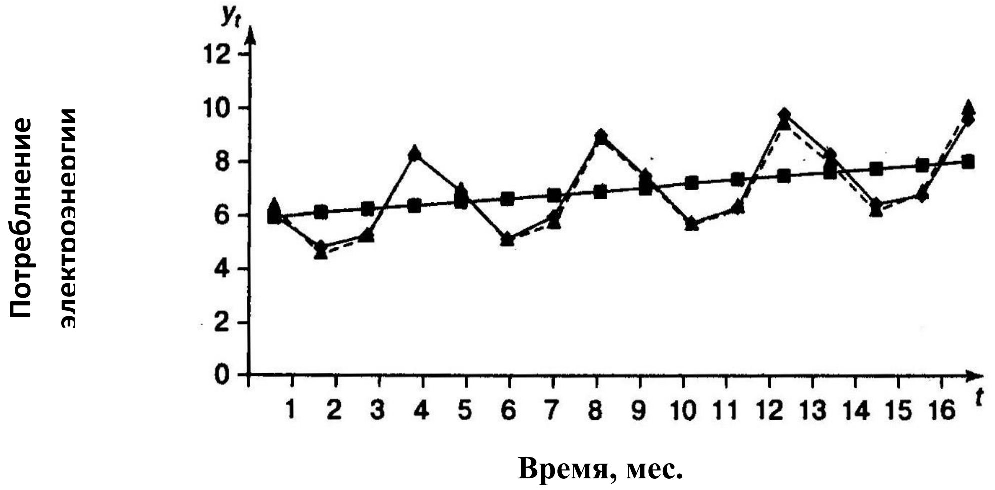


Рис. 6. Прибыль компании

Выявление и устранение сезонного эффекта (в некоторых источниках применяется термин «десезонализация уровней ряда») используются в двух направлениях. Во-первых, воздействие сезонных колебаний следует устранять на этапе предварительной обработки исходных данных при изучении взаимосвязи нескольких временных рядов. Поэтому в российских и международных статистических сборниках часто публикуются данные, в которых устранено влияние сезонной компоненты (если это помесячная или поквартальная статистика), например показатели объемов производства в отдельных отраслях промышленности, уровня безработицы и т.д. Во-вторых, это анализ структуры одномерных временных рядов с целью прогнозирования уровней ряда в будущие моменты времени. Пример 5.6. Прогнозирование по аддитивной модели.

Предположим, по данным примера 5.4 требуется дать прогноз потребления электроэнергии жителями района в течение первого полугодия ближайшего следующего года.

Прогнозное значение F, уровня временного ряда в аддитивной модели в соответствии с соотношением (5.5) есть сумма трендовой и сезонной компонент.

Объем электроэнергии, потребленной в течение первого полугодия ближайшего следующего, т. е. пятого, года, рассчитывается как сумма объемов потребления электроэнергии в I и во II кварталах пятого года, соответственно F_1и F_l%. Для определения трендовой компоненты воспользуемся уравнением тренда

$$T= 5,715 + 0,186 \cdot t.$$

Получим:

$$T_{17} = 5,715+ 0,186- 17 = 8,877;$$

$$T_{18} = 5,715 + 0,186-18 = 9,063,$$

Значения сезонной компоненты равны: Sv= 0,581 (I квартал); S2= -1,977 (II квартал).

Таким образом,

$$F17 = T,, + 5, = 8,877 + 0,581 = 9,458;$$

$$F18= Г,, + 52 = 9,063 - 1,977 = 7,086.$$

Прогноз объема потребления электроэнергии на первое полугодие ближайшего следующего (пятого) года составит:

(9,458 + 7,086) = 16,544 млн. кВт • ч.

Пример 5.7. Прогнозирование по мультипликативной модели.

Предположим, по данным примера 5 необходимо сделать прогноз ожидаемой прибыли компании за первое полугодие ближайшего следующего года.

Прогнозное значение F_t уровня временного ряда в мультипликативной модели в соответствии с соотношением (5.6) есть произведение трендовой и сезонной компонент. Для определения трендовой компоненты за каждый квартал воспользуемся уравнением тренда

$$T = 90{,}59 - 2{,}773 \cdot t.$$

Получим:

$$T_{17} = 90{,}59 - 2{,}773 \cdot 17 = 43{,}401;$$

$$T_{18} = 90{,}59 - 2{,}773 -18 = 40{,}626.$$

Значения сезонной компоненты равны S{= 0,913 (I квартал);

S_2= 1,202 (II квартал)

Список литературы

Основная литература:	1. Christopher Dougherty. Introduction to Econometrics. Oxford University Press, 2011. – 573 p. 2. Gujarati D.N. Basic Econometrics. McGraw-Hill, 5th edition, 2009. – 922 p. 3.Абдуллаев О.М., Ходиев Б.Ю., Ишназаров А.И. Эконометрика. Учебник. –Т.: Fan va texnologiya. 2007. – 612 с. 4. Абдуллаев О.М., Жамалов М.С. Эконометрическое моделирование. Учебник. –Т.: Fan va texnologiya. 2010. – 612 с
Дополнительная литература	1. Greene W.H.Econometric Analysis. Prentice Hall. 7th edition, 2011. -1232p. 2.Валентинов В,А.Эконометрика: Учебник. –М.:ИТК»Дашков и К»,2009–367с. 3.Кремер Н.Ш.Эконометрика:Учебник–М.:ЮНИТИ-ДАНА,2008. –562с. 4.Айвазян С.А. Прикладная статистика и основы эконометрика:Учебник. –М.ЮНИТИ,2007. –345с. 5.Елисеева И.И., Курышева С.В.и др. Эконометрика:Учебник–М.:Финансы и статистика, 2007. –260с. 6.Habibullayev I Iqtisodiy matematik usullsr va modellar: o‘quv qo‘llamna / O‘zbekiston Respublikasi Oliy va o‘rta maxsus ta`lim vazirligi/–Toshkent:“Tafakkur-Bo‘stoni”, 2012.112 b.

Интернет ресурсы:	www.ziyonet.uz **www.lex.uz** www.stat.uz **www.mehnat.uz** **www.mf.uz** www.tfi.uz www.mfer.uz www.conomical.uz **www.cisstat.com** **www.gks.ru** **www.unstats.un.org**

8-ТЕМА
Прикладные эконометрические модели
ПЛАН:
8.1. Функция полезности. Задача потребительского выбора
8.2. Решение задачи потребительского выбора и его свойства
8.3. Общая модель потребительского выбора
8.4. Взаимозаменяемость товаров. Эффекты компенсации

8.1. Функция полезности. Задача потребительского выбора

Пусть потребитель располагает доходом I, который он полностью тратит на приобретение товаров (продуктов), т.е. величина I — это не доход, а расход данного потребителя.

Учитывая структуру цен, доход и собственные предпочтения, потребитель приобретает определенные количества товаров, и математическая модель такого его поведения называется *моделью потребительского выбора.*

Рассмотрим потребительские наборы из двух товаров.

Потребительский набор (для краткости *набор*) — это вектор (x_1, x_2), координата x_1 которого равна количеству единиц первого товара, а координата x_2 равна количеству единиц второго товара.

***Выбор* п**отребителя характеризуется отношением предпочтения, суть которого состоит в следующем. Считается, что потребитель про каждые 2 набора может сказать, что либо один из них более желателен, чем другой, либо потребитель не видит между ними разницы. Отношение предпочтения транзитивно, т.е. если набор $A=(a_1,\ a_2)$ предпочтительнее набора $B=(b_1,\ b_2)$, а набор B предпочтительнее набора $C=(c_1,\ c_2)$, то набор A предпочтительнее набора C.

Функцией полезности потребителя называется функция $u(x_1,\ x_2)$, определенная на множестве потребительских наборов $(x_1,\ x_2)$, значение $u(x_1,\ x_2)$ которой на потребительском наборе $(x_1,\ x_2)$ равно потребительской оценке потребителя для этого набора.

Потребительскую оценку $u(x_1,\ x_2)$ набора $(x_1,\ x_2)$ принято называть *уровнем* (или *степенью)* удовлетворения потребностей потребителя, если он приобретает или потребляет данный набор $(x_1,\ x_2)$. Каждый потребитель имеет, вообще говоря, свою функцию полезности. Если набор A предпочтительнее набора B, то $u(A)>u(B)$.

Функция полезности удовлетворяет следующим свойствам:

1. Возрастание потребления одного продукта при постоянном потреблении другого продукта ведет к росту потребительской оценки, т.е.

а) если $x_1^2 > x_1^1$, то $u(x_1^2, x_2) > u(x_1^1, x_2)$ или, другими словами, $\dfrac{\partial u(x_1, x_2)}{\partial x_1} = u_1' > 0$;

б) если $x_2^2 > x_2^1$, то $u(x_1, x_2^2) > u(x_1, x_2^1)$ или, другими словами, $\dfrac{\partial u(x_1, x_2)}{\partial x_2} = u_2' > 0$.

Первые частные производные называются *предельными полезностями* продуктов: u_1' называется предельной полезностью первого продукта, а u_2' — предельной полезностью второго продукта.

2. Предельная полезность каждого продукта уменьшается, если объем его потребления растёт (это свойство предельной полезности называется *законом убывания предельной полезности*), т.е.

$$\frac{\partial^2 u}{\partial x_1^2} = u''_{11} < 0, \quad \frac{\partial^2 u}{\partial x_2^2} = u''_{22} < 0.$$

3. Предельная полезность каждого продукта увеличивается, если растет количество другого продукта, т.е.

$$\frac{\partial^2 u}{\partial x_1 \partial x_2} = u''_{12} = \frac{\partial^2 u}{\partial x_2 \partial x_1} = u''_{21} > 0.$$

В этом случае продукт, количество которого фиксировано, оказывается относительно дефицитным. Поэтому дополнительная его единица приобретает большую ценность и может быть потреблена более эффективно. Данное свойство не столь очевидно, как свойства 1 и 2, и справедливо не для всех товаров: если товары могут полностью замещать друг друга в потреблении, свойство 3 не выполняется.

Примером функции полезности может служить функция

$$u(x_1, x_2) = a_1 \ln(x_1 - x_1^*) + a_2 \ln(x_2 - x_2^*),$$

где $a_1 > 0$, $a_2 > 0$, $x_1 > x_1^* \geq 0$, $x_2 > x_2^* \geq 0$.

Действительно, имеем

$$u'_1 = \frac{a_1}{x_1 - x_1^*} > 0, \quad u'_2 = \frac{a_2}{x_2 - x_2^*} > 0,$$

$$\frac{\partial^2 u}{\partial x_1^2} = -\frac{a_1}{(x_1 - x_1^*)^2} < 0, \quad \frac{\partial^2 u}{\partial x_2^2} = -\frac{a_2}{(x_2 - x_2^*)^2} < 0,$$

т.е. выполнены свойства 1 и 2 функции полезности. Свойство 3 не выполнено, так как смешанные вторые частные производные функции $u(x_1, x_2)$ равны нулю.

Задача потребительского выбора (задача рационального поведения потребителя на рынке) заключается в выборе такого потребительского набора

$(x_1^0,\ x_2^0)$, который максимизирует его функцию полезности при заданном бюджетном ограничении.

Бюджетное ограничение означает, что денежные расходы на продукты не могут превышать денежного дохода, т.е. $p_1x_1 + p_2x_2 \le I$, где p_1 и p_2 — рыночные цены одной единицы соответственно первого и второго продуктов, а I — доход потребителя, который он готов потратить на приобретение первого и второго продуктов. Величины p_1, p_2 и I заданы.

Формально задача потребительского выбора имеет вид:

$$u(x_1, x_2) \quad (\max)\ (\max)$$

при условиях

$$\begin{array}{c} p_1x_1 + p_2x_2 \le I \\ x_1 \ge 0 \quad x_2 \ge 0 \end{array},$$

$$x_1 \ge 0,\ x_2 \ge 0.$$

8.2.Решение задачи потребительского выбора и его свойства

Набор $(x_1^0,\ x_2^0)$, который является решением задачи потребительского выбора, принято называть *оптимальным* для потребителя или *локальным рыночным равновесием* потребителя.

В приведенной постановке задача потребительского выбора является задачей нелинейного программирования. Однако если на каком-то потребительском наборе $(x_1,\ x_2)$ бюджетное ограничение $p_1x_1 + p_2x_2 \le I$ будет выполняться в виде строгого неравенства, то мы можем увеличить потребление какого-либо из продуктов и тем самым увеличить функцию полезности. Следовательно, набор $(x_1^0,\ x_2^0)$, максимизирующий функцию полезности, должен обращать бюджетное ограничение в равенство, т.е. $p_1x_1^0 + p_2x_2^0 = I$.

Мы также будем считать, что в оптимальной точке $(x_1^0,\ x_2^0)$ условия $x_1 \ge 0$, $x_2 \ge 0$ выполняются автоматически, вытекая из свойств функции $u(x_1, x_2)$.

Как правило, это действительно так. В то же время, если условия неотрицательности переменных не включать в явном виде в условие задачи, то она становится существенно проще с математической точки зрения.

Итак, задачу потребительского выбора можно заменить задачей на условный экстремум (ибо решение $(x_1^0,\ x_2^0)$ этих двух задач одно и то же):

$$u(x_1, x_2)\ (\max)$$

при условии

$$p_1x_1 + p_2x_2 = I\,.$$

Для решения этой задачи на условный экстремум применим метод Лагранжа. Выписываем функцию Лагранжа

$$L(x_1, x_2, \lambda) = u(x_1, x_2) + \lambda(p_1x_1 + p_2x_2 - I),$$

находим ее первые частные производные по переменным x_1, x_2, λ и приравниваем эти частные производные к нулю:

$$\frac{\partial L}{\partial x_1} = u_1' - \lambda \cdot p_1 = 0,\ \frac{\partial L}{\partial x_2} = u_2' - \lambda \cdot p_2 = 0,\ \frac{\partial L}{\partial \lambda} = p_1x_1 - p_2x_2 - I = 0.$$

Исключив из полученной системы трех уравнений с тремя неизвестными неизвестную λ, получим систему двух уравнений с двумя неизвестными x_1, x_2

$$\frac{u_1'}{u_2'} = \frac{p_1}{p_2},$$

$$p_1x_1 + p_2x_2 = I\,,$$

из которой получим решение $(x_1^0,\ x_2^0)$ задачи потребительского выбора.

Координаты x_1^0 и x_2^0 решения $(x_1^0,\ x_2^0)$ задачи потребительского выбора есть функции параметров p_1, p_2 и I:

$$x_1^0 = x_1^0(p_1, p_2, I),$$

$$x_2^0 = x_2^0(p_1, p_2, I).$$

Полученные функции называются *функциями спроса* на первый и второй продукты. Важным свойством функций спроса является их однородность нулевой степени относительно цен и дохода, т.е. значения функций спроса инварианты по отношению к пропорциональным изменениям цен и дохода:

$$x_1^0(\alpha p_1, \alpha p_2, \alpha I) = x_1^0(p_1, p_2, I),$$

$$x_2^0(\alpha p_1, \alpha p_2, \alpha I) = x_2^0(p_1, p_2, I)$$

для любого числа $\alpha > 0$. Это означает, что если все цены и доход изменяется в одно и то же число раз, величина спроса на продукт (первый или второй — безразлично) останется неизменной.

Решим одну простую задачу потребительского выбора с двумя товарами. Пусть неизвестные количества этих товаров равны x_1 и x_2, а их рыночные цены — соответственно p_1 и p_2. Рассматриваемая задача имеет вид:

$$u(x_1, x_2) = x_1 \cdot x_2 \text{ (max)} \qquad (4.1)$$

$$p_1 x_1 + p_2 x_2 \le I, \qquad (4.2)$$

$$x_1 \ge 0, \ \ x_2 \ge 0. \qquad (4.3)$$

Как мы выяснили, бюджетное ограничение в оптимальной точке должно выполняться как равенство, и, поскольку оба товара жизненно необходимы (полезность равна нулю, если один из них отсутствует), требования неотрицательности переменных будут выполнены автоматически. Следовательно, решаемая задача математического программирования превращается в классическую задачу на условный экстремум. Записав необходимые условия экстремума (согласно которым, отношения предельных полезностей товаров должны равняться отношениям их рыночных цен, а бюджетное ограничение выполняется как равенство), получаем систему уравнений

$$\frac{x_2}{x_1} = \frac{p_1}{p_2},$$

$$p_1 x_1 + p_2 x_2 = I .$$

Здесь первое условие означает, что в рассматриваемой задаче количества денег, затрачиваемые на оба товара, должны быть одинаковыми, то есть $x_2 \cdot p_2 = x_1 \cdot p_1$. Это вытекает из равенства «весов», или показателей степени у переменных x_1 и x_2 в функции полезности. Итак, $x_2 \cdot p_2 = x_1 \cdot p_1 = \frac{I}{2}$ и функции спроса приобретают вид

$$x_1 = \frac{I}{2 \cdot p_1}; \quad x_2 = \frac{I}{2 \cdot p_2}. \qquad (4.4)$$

8.3.Общая модель потребительского выбора

Таким образом, расход на каждый товар составляет половину общего дохода потребителя, и, чтобы найти необходимое количество каждого товара, следует разделить расходуемую на него сумму на его цену.

Теперь рассмотрим свойства задачи потребительского выбора с произвольным числом товаров и целевой функцией общего вида.

Пусть заданы целевая функция полезности потребителя $u(x_1, x_2, \dots, x_n)$ (x_i — количество *i*-го товара), вектор цен $\bar{p} = (p_1, p_2, \dots, p_n)$ и доход I. Записав бюджетное ограничение и ограничения на неотрицательность, получаем задачу

$$u(\bar{x}) \text{ (max)}$$

при условиях

$$\bar{p}\,\bar{x} \le I,$$

$$\bar{x} \ge 0,$$

где $\bar{x} = (x_1, \dots, x_n)$, $\bar{p} = (p_1, \dots, p_n)$, $\bar{p}\,\bar{x} = p_1 x_1 + \dots + p_n x_n$.

Будем, как и ранее, считать, что неотрицательность переменных обеспечивается свойствами целевой функции и бюджетного ограничения. В этом случае можно записать функцию Лагранжа и исследовать ее на безусловный экстремум. Необходимые условия экстремума функции Лагранжа $L(\bar{x},\lambda)=u(\bar{x})+\lambda(\bar{p}\,\bar{x}-I)$ — равенство нулю частных производных:

$$L_i' = u_i' + \lambda p_i = 0 \text{ для всех } i=\overline{1,n} \qquad (4.5)$$

и

$$L_\lambda' = \bar{p}\,\bar{x} - I = 0.$$

Отсюда вытекает, что для всех *i*, *j* в точке $\bar{x}^0$ локального рыночного равновесия выполняется равенство

$$\frac{u_i'}{u_j'} = \frac{p_i}{p_j}, \qquad (4.6)$$

которое получается после перенесения в правую часть вторых слагаемых в условиях (4.5) и деления *i*-го равенства на *j*-ое. Итак, в точке оптимума отношение предельных полезностей любых двух товаров равно отношению их рыночных цен. Равенство (4.6) можно переписать и в другой форме:

$$\frac{u_i'}{p_i} = \frac{u_j'}{p_j}.$$

Последнее означает, что дополнительная полезность, приходящая на дополнительную единицу денежных затрат, в точке оптимума одинакова по всем видам товаров. Если бы это было не так, то по крайней мере одну денежную единицу можно было бы перераспределить так, чтобы выросло благосостояние (значение функции полезности) потребителя. Если бы для некоторых *i*, *j* имело бы место $\frac{u_i'}{p_i} > \frac{u_j'}{p_j}$, то некоторое количество денег можно было бы перераспределить от *i*-го товара к *j*-му, увеличив уровень благосостояния.

8.4.Взаимозаменяемость товаров. Эффекты компенсации

Если функция спроса имеет вид $x_i = \dfrac{I}{np_i}$, то спрос на *i*-й товар не зависит от цены на любой *j*-й товар. Вообще говоря, перекрестные функции спроса от цен характеризуют такие свойства товаров, как взаимозаменяемость и взаимодополняемость. Если при росте цены и снижении спроса на *i*-й товар растет спрос на *j*-й товар — эти товары взаимозаменяемы. Наоборот, если спрос на *j*-й товар также падает, — они взаимодополняемы.

Заметим, что реальная взаимозаменяемость может искажаться общим снижением благосостояния при росте цены *i*-го товара: *j*-й товар может заменять *i*-й в потреблении, но спрос на него может не расти, поскольку снизилось общее благосостояние потребителя. Для снятия этого искажения используются понятие *компенсированного изменения цены*, то есть такого, которое сопровождается увеличением дохода потребителя, позволяющим ему поддерживать прежний уровень благосостояния.

Для формального анализа компенсационных эффектов рассмотрим две задачи.

Сначала решим задачу (4.1) – (4.3) с ценами товаров $p_1 = 10$, $p_2 = 2$ и с доходом потребителя $I = 60$. Тогда, согласно формуле (4.4), $x_1 = \dfrac{60}{2 \cdot 10} = 3$,

$x_2 = \dfrac{60}{2 \cdot 2} = 15$ и $u(x_1, x_2) = 45$.

Пусть теперь p_2 меняется с 2 до 7. Каков необходимый размер компенсации? Чтобы приобрести прежний оптимальный набор, потребителю необходимо дополнительно $(7-2) \cdot 15 = 75$ денежных единиц. Однако прежняя структура потребления не будет оптимальной при новых ценах, так как в этом случае

$x_1 = \dfrac{60+75}{2 \cdot 10} = 6{,}75$, $x_2 = \dfrac{60+75}{2 \cdot 7} \approx 9{,}64$ и $u(x_1, x_2) \approx 65$.

Пусть для поддержания прежнего уровня благосостояния потребитель получает дополнительно M денежных единиц. Тогда при новых ценах его спрос на первый и второй товар будет равен соответственно $x_1 = \frac{60+M}{2\cdot 10}$ и $x_2 = \frac{60+M}{2\cdot 7}$. Целевая функция $x_1 \cdot x_2$ будет равна $\frac{(60+M)^2}{10\cdot 7\cdot 4}$, и это выражение должно равняется начальному $u(x_1, x_2) = 45$. Отсюда $M \approx 52{,}25$, что существенно меньше, чем 75.

Теперь решим задачу (4.1) – (4.3) в более общем виде. Очевидно, что

$$x_i = \frac{I}{2p_i};\ \ \frac{\partial x_i}{\partial p_i} = -\frac{I}{2p_i^2};\ \ \frac{\partial x_i}{\partial I} = \frac{1}{2p_i};\ \ \frac{\partial x_i}{\partial p_j} = 0 \quad (i, j = \overline{1, 2}).$$

Пусть теперь цена p_1 выросла в z раз ($z > 1$), и при этом потребитель получает необходимую компенсацию. Новый размер дохода обозначим через I^*, а спроса — x_1^* и x_2^*.

Очевидно, что

$$x_1^* = \frac{I^*}{2zp_1};\ \ x_2^* = \frac{I^*}{2p_2},$$

а условие компенсации имеет вид

$$\frac{(I^*)^2}{4zp_1p_2} = \frac{I^2}{4p_1p_2},$$

откуда

$$I^* = \sqrt{z}\cdot I;\ \ x_1^* = \frac{x_1}{\sqrt{z}};\ \ x_2^* = x_2\sqrt{z}.$$

Итак, спрос на первый товар в случае с компенсацией сократится в $\sqrt{z}$ раз (а не z раз, как без нее), а спрос на второй товар в $\sqrt{z}$ раз вырастет. В случае роста цены второго товара ситуация будет полностью симметричной.

Таким образом, $\left(\dfrac{\partial x_i}{\partial p_j}\right)_{comp} > 0$ при $i = 1$, $j = 2$ или при $i = 2$, $j = 1$.

Индекс *comp* означает, что перекрестная частная производная спроса рассчитывается при необходимой для поддержания прежнего уровня благосостояния компенсации дохода. Условие компенсации снимает «эффект дохода», оставляя лишь «эффект замены», что позволяет более точно определить понятие взаимозаменяемости и взаимодополняемости товаров и оценивать эти характеристики.

i-й и *j*-й товары называются *взаимозаменяемыми*, если $\left(\dfrac{\partial x_i}{\partial p_j}\right)_{comp} > 0$ и $\left(\dfrac{\partial x_j}{\partial p_i}\right)_{comp} > 0$ (эти два условия равносильны), и *взаимодополняемыми*, если $\left(\dfrac{\partial x_i}{\partial p_j}\right)_{comp} < 0$ и $\left(\dfrac{\partial x_j}{\partial p_i}\right)_{comp} < 0$.

Рассчитаем теперь эти частные производные для рассматриваемой задачи, когда p_1 растет в z раз. В этом случае приращения имеют вид

$$\Delta x_1 = \frac{x_1}{\sqrt{z}} - x_1;\ \Delta x_2 = x_2\sqrt{z} - x_2;\ \Delta p_1 = zp_1 - p_1.$$

Отсюда

$$\left(\frac{\partial x_1}{\partial p_1}\right)_{comp} = \lim_{z\to 1}\frac{x_1(1-\sqrt{z})}{p_1\sqrt{z}\cdot(z-1)} = \lim_{z\to 1}\left(-\frac{x_1}{p_1\sqrt{z}(1+\sqrt{z})}\right) =$$

$$= -\frac{x_1}{2}p_1 = -\frac{I}{4p_1^2},$$

$$\left(\frac{\partial x_2}{\partial p_1}\right)_{comp} = \lim_{z\to 1}\frac{x_2(\sqrt{z}-1)}{p_1\sqrt{z}\cdot(z-1)} = \lim_{z\to 1}\frac{x_2}{p_1\sqrt{z}(\sqrt{z}+1)} = \frac{x_2}{2p_2} = \frac{I}{4p_1p_2}.$$

Последняя величина положительна, что свидетельствует о взаимозаменяемости товаров в рассматриваемой задаче.

Список литературы

Основная литература:	1. Christopher Dougherty. Introduction to Econometrics. Oxford University Press, 2011. – 573 p. 2. Gujarati D.N. Basic Econometrics. McGraw-Hill, 5^{th} edition, 2009. – 922 p. 3.Абдуллаев О.М., Ходиев Б.Ю., Ишназаров А.И. Эконометрика. Учебник. –Т.: Fan va texnologiya. 2007. – 612 с. 4. Абдуллаев О.М., Жамалов М.С. Эконометрическое моделирование. Учебник. –Т.: Fan va texnologiya. 2010. – 612 с
Дополнительная литература	1. Greene W.H.Econometric Analysis. Prentice Hall. 7^{th} edition, 2011. -1232p. 2.Валентинов В,А.Эконометрика: Учебник. –М.:ИТК»Дашков и К»,2009–367с. 3.Кремер Н.Ш.Эконометрика:Учебник–М.:ЮНИТИ-ДАНА,2008. –562с. 4.Айвазян С.А. Прикладная статистика и основы эконометрика:Учебник. –М.ЮНИТИ,2007. –345с. 5.Елисеева И.И., Курышева С.В.и др. Эконометрика:Учебник–М.:Финансы и статистика, 2007. –260с. 6.Habibullayev I Iqtisodiy matematik usullsr va modellar: oʻquv qoʻllamna / Oʻzbekiston Respublikasi Oliy va oʻrta maxsus ta`lim vazirligi/–Toshkent:“Tafakkur-Boʻstoni”, 2012.112 b.
Интернет ресурсы:	www.ziyonet.uz **www.lex.uz** www.stat.uz **www.mehnat.uz** **www.mf.uz** www.tfi.uz www.mfer.uz www.conomical.uz **www.cisstat.com** **www.gks.ru** **www.unstats.un.org**

9-ТЕМА

ИСПОЛЬЗОВАНИЕ ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ ПРИ ПРОГНОЗИРОВАНИИ ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ

ПЛАН:

9.1.	**Специфика статистической оценки взаимосвязи двух временных рядов**
9.2.	**Включение в модель регрессии фактора времени**
9.3.	**Оценивание параметров уравнения регрессии при наличии автокорреляции в остатках**

9.1.Специфика статистической оценки взаимосвязи двух временных рядов

Изучение причинно-следственных зависимостей переменных, представленных в форме временных рядов, является одной из самых сложных задач эконометрического моделирования. Применение в этих целях традиционных методов корреляционно-регрессионного анализа, может привести к ряду серьезных проблем, возникающих как на этапе построения, так и на этапе анализа эконометрических моделей. В первую очередь эти проблемы связаны со спецификой временных рядов как источника данных в эконометрическом моделировании. В главе 5 было показано, что каждый уровень временного ряда содержит три основные компоненты: тенденцию, циклические или сезонные колебания и случайную компоненту. Рассмотрим подробнее, каким образом наличие этих компонент сказывается на результатах корреляционно-регрессионного анализа временных рядов данных.

Предварительный этап такого анализа заключается в выявлении структуры изучаемых временных рядов. Если на этом этапе было выявлено, что временные ряды содержат сезонные или циклические колебания, то перед проведением дальнейшего исследования взаимосвязи необходимо устранить сезонную или циклическую компоненту из уровней каждого ряда, поскольку ее наличие приведет к завышению истинных показателей силы и тесноты связи изучаемых временных рядов в случае, если оба ряда содержат циклические колебания одинаковой периодичности, либо к занижению этих показателей в случае, если сезонные или циклические колебания содержат только один из рядов или периодичность колебаний в рассматриваемых временных рядах различна.

Устранение сезонной компоненты из уровней временных рядов можно проводить в соответствии с методикой построения аддитивной и мультипликативной моделей, рассмотренной в п. 5.4. При дальнейшем изложении методов анализа взаимосвязей в этой главе мы примем предположение, что изучаемые временные ряды не содержат периодических колебаний. Предположим, изучается зависимость между рядами x и y. Для количественной характеристики этой зависимости используется линейный коэффициент корреляции. Если рассматриваемые временные ряды имеют тенденцию, коэффициент корреляции по абсолютной величине будет высоким (положительным в случае совпадения и отрицательным в случае противоположной направленности тенденций рядов x и y). Однако из этого еще нельзя делать вывод о том, что x причина y или наоборот. Высокий коэффициент корреляции в данном случае есть результат того, что x и y зависят от времени, или содержат тенденцию. При этом одинаковую или противоположную тенденцию могут иметь ряды, совершенно не связанные друг с другом причинно-следственной зависимостью. Это, естественно, не означает, что увеличение количества домов отдыха способствует росту числа выпускников вузов или увеличение числа последних стимулирует спрос на дома отдыха.

Для того чтобы получить коэффициенты корреляции, характеризующие причинно-следственную связь между изучаемыми рядами, следует избавиться от так называемой ложной корреляции, вызванной наличием тенденции в каждом ряде. Обычно это осуществляют с помощью одного из методов исключения тенденции, которые будут рассмотрены в п. 6.2.

Предположим, что по двум временным рядам x, и y, строится уравнение парной линейной регрессии вида

$$y_t = a + b \cdot x_t + \varepsilon_t \qquad (1)$$

Наличие тенденции в каждом из этих временных рядов означает, что на зависимую y, и независимую x, переменные модели оказывает воздействие фактор времени, который непосредственно в модели не учтен. Влияние фактора времени будет выражено в корреляционной зависимости между значениями остатков e, за текущий и предыдущие моменты времени, которая получила название «автокорреляция в остатках».

Автокорреляция в остатках есть нарушение одной из основных предпосылок МНК - предпосылки о случайности остатков, полученных по уравнению регрессии. Один из возможных путей решения этой проблемы состоит в применении к оценке параметров модели обобщенного МНК. При построении уравнения множественной регрессии по временным рядам данных, помимо двух вышеназванных проблем, возникает также проблема мультиколлинеарности факторов, входящих в уравнение регрессии, в случае если эти факторы содержат тенденцию.

Методы исключения тенденций. Сущность всех методов исключения тенденции заключается в том, чтобы устранить или зафиксировать воздействие фактора времени на формирование уровней ряда. Основные методы исключения тенденции можно разделить на две группы:

• методы, основанные на преобразовании уровней исходного ряда в новые переменные, не содержащие тенденции. Полученные переменные используются далее для анализа взаимосвязи изучаемых временных рядов. Эти методы предполагают непосредственное устранение трендовой компоненты T из каждого уровня временного ряда. Два основных метода в данной группе - это метод последовательных разностей и метод отклонений от трендов;

• методы, основанные на изучении взаимосвязи исходных уровней временных рядов при элиминировании воздействия фактора времени на зависимую и независимые переменные модели. В первую очередь это метод включения в модель регрессии по временным рядам фактора времени.

Рассмотрим подробнее методику применения, преимущества и недостатки каждого из перечисленных выше методов.

Метод отклонений от тренда

Пусть имеются два временных ряда x, и y,, каждый из которых содержит трендовую компоненту T и случайную компоненту e. Проведение аналитического выравнивания по каждому из этих рядов позволяет найти параметры соответствующих уравнений трендов и определить расчетные по тренду уровни x, и y, соответственно. Эти расчетные значения можно принять за оценку трендовой компоненты T каждого ряда. Поэтому влияние тенденции можно устранить путем вычитания расчетных значений уровней ряда из фактических. Эту процедуру проделывают для каждого временного ряда в модели. Дальнейший анализ взаимосвязи рядов проводят с использованием не исходных уровней, а отклонений от тренда xt - St,и y, - y, при условии, что последние не содержат тенденции.

Пример 1. Измерение взаимосвязи расходов на конечное потребление и совокупного дохода.

Вернемся к примеру 1. Пусть помимо расходов на конечное потребление имеются данные о совокупном доходе (д. е). Исходные данные за 8 лет представлены в табл. 1. Требуется охарактеризовать тесноту и силу связи между временными рядами совокупного дохода xf и расходов на конечное потребление yг

Таблица 1

Расходы на конечное потребление и совокупный доход (усл. ед.)

Год	1	2	3	4	5	6	7	8
Расходы на конечное потребление, y_t	7	8	8	10	11	12	14	16
Совокупный доход, x_t	10	12	11	12	14	15	17	20

Корреляционно-регрессионный анализ, проведенный по исходным данным рядов, приводит к следующим результатам:

Уравнение регрессии $\hat{y}_t = -2{,}05 + 0{,}92 \cdot x_t$,

Коэффициент корреляции $r_{xy} = 0{,}982$,

Коэффициент ковариации $r_{xy}^2 = 0{,}965$.

Как было показано в примере 1, коэффициент автокорреляции первого порядка по ряду расходов на конечное потребление r}| = 0,976. Аналогично можно рассчитать, что коэффициент автокорреляции первого порядка временного ряда совокупного дохода г* = 0,880. Можно предположить, что полученные результаты содержат ложную корреляцию ввиду наличия в каждом из рядов линейной или близкой к линейной тенденции. Применим метод устранения тенденции по отклонениям от тренда. Результаты расчета линейных трендов по каждому из рядов представлены в табл. 2.

Таблица 2

Результаты расчета параметров линейных трендов расходов на конечное потребление и совокупного дохода

Показатели	Расходы на конечное потребление	Совокупный доход
константа	5,071428	8,035714
Коэффициент регрессии	1,261904	1,297619
Стандартная ошибка коэффициента регрессии	0.101946	0,179889
R-квадрат	0,962315	0,896611
Число наблюдений	8	8
Число степеней свободы	6	6

трендам $\hat{y}_t = 5,07 + 1,26 \cdot t$ и $\hat{x}_t = 8,04 + 1,3 \cdot t$ определим расчетные значения $\hat{y}_t$ и $\hat{x}_t$, и отклонения от трендов $y_t - \hat{y}_t$, и $x_t - \hat{x}_t$, (табл. 3).

Таблица 3

t, вакт	y_t	x_t	$\hat{y}_t$	$\hat{x}_t$	$y_t - \hat{y}_t$	$x_t - \hat{x}_t$
1	7	10	6,33	9,34	0,67	0,66
2	8	12	7,59	10,64	0,41	1,36
3	8	11	8,85	11,94	-0,85	-0,94
4	10	12	10,11	13,24	-0,11	-1,24
5	11	14	11,37	14,54	-0,37	-0,54
6	12	15	12,63	15,84	-0,63	-0,84
7	14	17	13,89	17,14	0,11	-0,14
8	16	20	15,15	18,44	0,85	1,56

Проверим полученные отклонения от трендов на автокорреляцию. Коэффициенты автокорреляции первого порядка по отклонениям от трендов составляют:

$$r_1^{\Delta x_t} = 0,254, \quad r_1^{\Delta y_t} = 0,129 .$$

Следовательно, временные ряды отклонений от трендов можно использовать для получения количественной характеристики тесноты связи исходных временных рядов расходов на конечное потребление и общего дохода. Коэффициент корреляции по отклонениям от трендов $r_{\Delta x \Delta y} = 0,860$ (сравните это значение с коэффициентом корреляции по исходным уровням рядов $r_{xy} = 0,982$). Связь между расходами на конечное потребление и совокупным доходом прямая и тесная.

Результаты построения модели регрессии по отклонениям от трендов следующие:

Константа	0,017313
Коэффициент регрессии	0,487553
Стандартная ошибка коэффициента	0,117946

регрессии	
R-квадрат	0,740116
Число наблюдений	8
Число степеней свободы	6

Содержательная интерпретация параметров этой модели затруднительна, однако ее можно использовать для прогнозирования. Для этого необходимо определить трендовое значение факторного признака $\hat{x}_t$ (и с помощью одного из методов оценить величину предполагаемого отклонения фактического значения от трендового. Далее по уравнению тренда для результативного признака определяют трендовое значение $\hat{x}_t$ а по уравнению регрессии по отклонениям от трендов находят величину отклонения $y_t - \hat{y}_t$. Затем находят точечный прогноз фактического значения у, по формуле

$y_t - \hat{y}_t$

Метод последовательных разностей. В ряде случаев вместо аналитического выравнивания временного ряда с целью устранения тенденции можно применить более простой метод - метод последовательных разностей.

Если временной ряд содержит ярко выраженную линейную тенденцию, ее можно устранить путем замены исходных уровней ряда цепными абсолютными приростами (первыми разностями).

Пусть

$$y_t = \hat{y}_t + \varepsilon_t \qquad (2)$$

где ε_t— случайная ошибка;

$$\hat{y}_t = a + b \cdot t \qquad (3)$$

Тогда

$$\Delta_t = y_t - y_{t-1} = a + b \cdot t + \varepsilon_t - (a + b \cdot (t-1) + \varepsilon_{t-1}) = b + (\varepsilon_t - \varepsilon_{t-1}). \qquad (4)$$

Коэффициент b— константа, которая не зависит от времени. При наличии сильной линейной тенденции остатки е, достаточно малы и в соответствии с предпосылками МНК носят случайный характер. Поэтому первые разности уровней ряда Д, не зависят от переменной времени, их можно использовать для дальнейшего анализа.

Если временной ряд содержит тенденцию в форме параболы второго порядка, то для ее устранения можно заменить исходные уровни ряда на вторые разности.

Пусть имеет место соотношение (2), однако

$$\hat{y}_t = a + b_1 \cdot t + b_2 \cdot t^2 \qquad (5)$$

Тогда

$$\begin{aligned}\Delta_t &= y_t - y_{t-1} = a + b_1 \cdot t + b_2 \cdot t^2 + \varepsilon_t - (a + b_1 \cdot (t-1) + b_2 \cdot (t-1)^2 + \varepsilon_{t-1}) = \\ &= b_1 - b_2 + 2 \cdot b_2 \cdot t + (\varepsilon_t - \varepsilon_{t-1}).\end{aligned} \qquad (6)$$

Как показывает это соотношение, первые разности Δ_t , непосредственно зависят от фактора времени (t) и, следовательно, содержат тенденцию.

Определим вторые разности:

$$\begin{aligned}&\Delta_t^2 = \Delta_t - \Delta_{t-1} = b_1 - b_2 + 2 \cdot b_2 \cdot t + (\varepsilon_1 - \varepsilon_2) - \\ &- (b_1 - b_2 + 2 \cdot b_2 \cdot (t-1) + (\varepsilon_1 - \varepsilon_2)) = \\ &= 2 \cdot b_1 + (\varepsilon_t - 2 \cdot \varepsilon_{t-1} + \varepsilon_{t-2})\end{aligned} \qquad (7)$$

Очевидно, что вторые разности A_2,не содержат тенденции, поэтому при наличии в исходных уровнях тренда в форме параболы второго порядка их можно использовать для дальнейшего анализа. Если тенденции временного ряда соответствует экспоненциальный или степенной тренд, метод последовательных разностей следует применять не к исходным уровням ряда, а к их логарифмам.

Пример 2. Изучение зависимости расходов на конечное потребление от совокупного дохода по первым разностям.

Обратимся вновь к данным о расходах на конечное потребление у, и совокупном доходе х, (табл. 1). Проанализируем зависимость между этими рядами, используя для этого первые разности (табл. 4)

Первые разности временных рядов расходов на конечное потребление и совокупного дохода

t	y_t	x_t	$\Delta_t y$	$\Delta_t x$
1	7	10	-	-
2	8	12	1	2
3	8	11	0	-1
4	10	12	2	1
5	11	14	1	2
6	12	15	1	1
7	14	17	2	2
8	16	20	2	3
Коэффициенты автокорреляции первого порядка			-0,109	-0,156

Результаты проверки временных рядов первых разностей на автокорреляцию приведены в последней строке табл. 4. Поскольку полученные ряды не содержат автокорреляции, будем использовать их вместо исходных данных для измерения зависимости между расходами на конечное потребление и совокупным доходом. Коэффициент корреляции этих рядов по первым разностям составляет $r_{\Delta_t x \Delta_t y} = 0{,}717$. Это подтверждает вывод о наличии тесной прямой связи между расходами на конечное потребление и совокупным доходом, приведенный в примере 6.1.

Построение уравнения регрессии зависимости расходов на конечное потребление от совокупного дохода по первыми разностям привело к следующим результатам:

Константа	0,676471
Коэффициент регрессии	0,426471
Стандартная ошибка коэффициента регрессии	0,184967
R-квадрат	0,515219
Число наблюдений	7
Число степеней свободы	5

Таким образом, уравнение регрессии имеет вид:

$$\hat{\Delta}_t y = 0{,}68 + 0{,}43 \cdot \Delta_t x; \quad R^2 = 0{,}515.$$

В отличие от уравнения регрессии по отклонениям от тренда, параметрам данного уравнения легко дать интерпретацию. При изменении прироста дохода на 1 д. е. прирост потребления изменяется в среднем на 0,43 д. е. в ту же сторону. При всей 270 своей простоте метод последовательных разностей имеет два существенных недостатка. Во-первых, его применение связано с сокращением числа пар наблюдений, по которым строится уравнение регрессии, и, следовательно, с потерей числа степеней свободы. Во-вторых, использование вместо исходных уровней временных рядов их приростов или ускорений приводит к потере информации, содержащейся в исходных данных

.

9.2.Включение в модель регрессии фактора времени

В корреляционно-регрессионном анализе устранить воздействие какого-либо фактора можно, если зафиксировать воздействие этого фактора на результат и другие включенные в модель факторы. Этот прием широко используется в анализе временных рядов, когда тенденция фиксируется через включение фактора времени в модель в качестве независимой переменной.

Модель вида

$$y_t = a + b_1 \cdot x_t + b_2 \cdot t + \varepsilon_t \qquad (8)$$

относится к группе моделей, включающих фактор времени. Очевидно, что число независимых переменных в такой модели может быть больше единицы. Кроме того, это могут быть не только текущие, но и лаговые значения независимой переменной, а также лаговые значения результативной переменной.

Преимущество данной модели по сравнению с методами отклонений от трендов и последовательных разностей в том, что она позволяет учесть всю информацию, содержащуюся в исходных данных, поскольку значения y, и x, есть уровни исходных временных рядов. Кроме того, модель строится по всей совокупности данных за рассматриваемый период в отличие от метода последовательных разностей, который приводит к потере числа наблюдений. Параметры a и Ь модели с включением фактора времени определяются обычным МНК. Расчет и интерпретацию параметров покажем на примере.

Пример 3. Построение модели регрессии с включением фактора времени.

Вернемся к данным табл. 1. Построим уравнение регрессии, описывающее зависимость расходов на конечное потребление x_t от совокупного дохода x_t и фактора времени. Для расчета параметров уравнения регрессии (8) воспользуемся обычным МНК. Система нормальных уравнений имеет вид:

$$\begin{cases} n \cdot a + b_1 \cdot \sum x_t + b_2 \cdot \sum t = \sum y_t, \\ a \cdot \sum x_t + b_1 \cdot \sum x_t^2 + b_2 \cdot \sum t \cdot x_t = \sum x_t \cdot y_t, \\ a \cdot \sum t + b_1 \cdot \sum t \cdot x_t + b_2 \cdot \sum x_t^2 = \sum t \cdot y_t \end{cases} \quad (9)$$

Рассчитав по исходным данным необходимые величины, получим:

$$\begin{cases} 8 \cdot a + 111 \cdot b_1 + 36 \cdot b_2 = 86, \\ 111a + 1619 \cdot b_1 + 554 \cdot b_2 = 1266, \\ 36 \cdot a + 554 \cdot b_1 + 204 \cdot b_2 = 440 \end{cases}$$

Решив эту систему относительно a, b_1 и b_2, находим: a = 1,15; b_x = 0,49; b_2 = 0,63. Следовательно, уравнение регрессии имеет вид:

$$y_t = 1{,}15 + 0{,}49 \cdot x_t + 0{,}63 \cdot t + \varepsilon_t$$

Интерпретация параметров этого уравнения следующая. Параметр b_x=0,49 характеризует, что при увеличении совокупного дохода на 1 д. е. расходы на конечное потребление возрастут в среднем на 0,49 д. е. в условиях существования неизменной тенденции. Параметр b_2 =0,63 означает, что воздействие всех факторов, кроме совокупного дохода, на расходы на конечное потребление приведет к его среднегодовому абсолютному приросту на 0,63 д. е.

9.3.Оценивание параметров уравнения регрессии при наличии автокорреляции в остатках

Обратимся вновь к уравнению регрессии (1). Примем некоторые допущения относительно этого уравнения:

• пусть y_t и x_t, не содержат тенденции, например, представляют собой отклонения выровненных по трендам значений от исходных уровней временных рядов;

• пусть оценки a и b параметров уравнения регрессии найдены обычным МНК;

• пусть критерий Дарбина - Уотсона показал наличие автокорреляции в остатках первого порядка.

Чтобы понять, каковы последствия автокорреляции в остатках для оценок параметров модели регрессии, найденных обычным МНК, построим формальную модель, описывающую автокорреляцию в остатках. Автокорреляция в остатках первого порядка предполагает, что каждый следующий уровень остатков e, зависит от предыдущего уровня e. Следовательно, существует модель регрессии вида

$$\varepsilon_t = c + d \cdot \varepsilon_{t-1} + u_t$$

где c и d- параметры уравнения регрессии.

В соответствии с рабочими формулами МНК имеем

$$\varepsilon_t = \overline{\varepsilon_t} - d \cdot \overline{\varepsilon_{t-1}}; \quad d = \frac{\overline{\varepsilon_t \varepsilon_{t-1}} - \overline{\varepsilon_t} \cdot \overline{\varepsilon_{t-1}}}{\overline{\varepsilon_{t-1}^2} - \overline{\varepsilon}_{t-1}^2} \quad (24)$$

С учетом соотношений (16) и (17) получим:

$$c = 0; \quad d = \frac{\overline{\varepsilon_t \varepsilon_{t-1}}}{\overline{\varepsilon_{t-1}^2}} = \frac{\sum_{t=2}^{n} \varepsilon_t \varepsilon_{t-1}}{\sum_{t=2}^{n} \varepsilon_{t-1}^2} \approx r_1^{\varepsilon} \quad (25)$$

Таким образом, имеем:

$$\varepsilon_t = r_1^{\varepsilon} \cdot \varepsilon_{t-1} + u_t \quad (26)$$

где и, - случайная ошибка.

Заметим, что $\left| r_1^{\varepsilon} \right| \prec 1$

Учитывая соотношение (26), уравнение (1) можно переписать в виде

$$\varepsilon_t = c + d \cdot \varepsilon_{t-1} + u_t$$

Найденные соотношения показывают, что текущий уровень ряда .у, зависит не только от факторной переменной но и от остатков предшествующего периода ε_{t-1}.

Допустим, мы не принимаем во внимание эту информацию и определяем оценки параметров a и b уравнения (1) обычным МНК. Тогда можно показать, что полученные оценки неэффективны, т. е. они не имеют минимальную дисперсию. Это приводит к увеличению стандартных ошибок, снижению фактических значений t-критерия и широким доверительным интервалам для коэффициента регрессии. На основе таких результатов можно сделать ошибочный вывод о незначимом влиянии исследуемого фактора на результат, в то время как на самом деле его влияние статистически значимо.

Отметим, что при соблюдении прочих предпосылок МНК автокорреляция остатков не влияет на свойства состоятельности и несмещенности оценок параметров уравнения регрессии обычным МНК, за исключением моделей авторегрессии. Применение МНК к моделям авторегрессии ведет к получению смещенных, несостоятельных и неэффективных оценок.

Рассмотрим основной подход к оценке параметров модели регрессии в случае, когда имеет место автокорреляция остатков. Для этого вновь обратимся к исходной модели (1). Для момента времени t-1 эта модель примет вид:

$$y_{t-1} = a + b \cdot x_{t-1} + \varepsilon_{t-1}$$

Умножим обе части уравнения (28) на r_1^{ε}

$$r_1^{\varepsilon} \cdot y_{t-1} = r_1^{\varepsilon} \cdot a + r_1^{\varepsilon} \cdot b \cdot x_{t-1} + r_1^{\varepsilon} \cdot \varepsilon_{t-1}$$

Вычтем почленно из уравнения (1) уравнение (29):

$$y_t - r_1^{\varepsilon} \cdot y_{t-1} = a - r_1^{\varepsilon} \cdot a + b \cdot x_t - r_1^{\varepsilon} \cdot b \cdot x_{t-1} + \varepsilon_t - r_1^{\varepsilon} \cdot \varepsilon_{t-1} \quad (30)$$

Проведя тождественные преобразования в (30), имеем:

$$y_t - r_1^{\varepsilon} \cdot y_{t-1} = a \cdot (1 - r_1^{\varepsilon}) + b \cdot (x_t - r_1^{\varepsilon} \cdot x_{t-1}) + \varepsilon_t - r_1^{\varepsilon} \cdot \varepsilon_{t-1} \quad (31)$$

Или

$$y_t^{'} = a^{'} + b \cdot x_t^{'} + u_t$$

В формуле (32)

$$y'_t = y_t - r_1^{\varepsilon} \cdot y_{t-1}$$
$$x'_t = x_t - r_1^{\varepsilon} \cdot x_{t-1}$$
$$u_t = \varepsilon_t - r_1^{\varepsilon} \cdot \varepsilon_{t-1}$$
$$a' = a \cdot (1 - r_1^{\varepsilon})$$

Поскольку u_t — случайная ошибка, то для оценки параметров уравнения (6.32) можно применять обычный МНК.

Итак, если остатки по исходному уравнению регрессии содержат автокорреляцию, то для оценки параметров уравнения используют обобщенный МНК. Для его реализации необходимо выполнять следующие условия.

1. Преобразовать исходные переменные y_{t-1} и x_{t-1} к виду (33) и (34).
2. Применив обычный МНК к уравнению (32), определить оценки параметров a' и b.
3. Рассчитать параметр а исходного уравнения из соотношения (36) как

$$a = a' / (1 - r_1^{\varepsilon}) \quad (37)$$

4. Выписать исходное уравнение (1).

Обобщенный метод наименьших квадратов аналогичен методу последовательных разностей. Однако мы вычитаем из y_t (или x_{t-1}) не все значение предыдущего уровня y_{t-1} (или x_{t-1}), а некоторую его долю -$r_1^{\varepsilon} \cdot y_{t-1}$ или $r_1^{\varepsilon} \cdot x_{t-1}$ Если $r_1^{\varepsilon} = 1$ данный метод есть просто метод первых разностей, так как

$$y'_t = y_t - y_{t-1} \quad (38)$$

и

$$x'_t = x_t - x_{t-1} \quad (39)$$

Поэтому в случае, если значение критерия Дарбина - Уотсона близко к нулю, применение метода первых разностей вполне обоснованно. Если $r_1^{\varepsilon} = -1$ т. е. в остатках наблюдается полная отрицательная автокорреляция, то изложенный выше метод модифицируется следующим образом:

$$y'_t = y_t - (-1) \cdot y_{t-1} = y_t + y_{t-1} \quad (40)$$

Аналогично

$$x_t^{'} = x_t - (-1) \cdot x_{t-1} = x_t + x_{t-1} \quad (41)$$

Поскольку

$$a^{'} = a \cdot (1 - r_1^{\varepsilon}) = 2 \cdot a \quad (42)$$

имеем:

$$y_t + y_{t+1} = 2 \cdot a + b \cdot (x_t + x_{t-1}) + u_t \quad (43)$$

Следовательно:

$$(y_t + y_{t+1}) / 2 = a + b \cdot (x_t + x_{t-1}) / 2 + u_t / 2 \quad (43)$$

В сущности, в модели (44) мы определяем средние за два периода уровни каждого ряда, а затем по полученным усредненным уровням обычным МНК рассчитываем параметры а и b. Данная модель называется моделью регрессии по скользящим средним.

Основная проблема, связанная с применением данного метода, заключается в том, как получить оценку r_1^2. Существует множество способов оценить численное значение коэффициента автокорреляции остатков первого порядка. Однако основными способами являются оценка этого коэффициента непосредственно по остаткам, полученным по исходному уравнению регрессии, и получение его приближенного значения из соотношения между коэффициентом автокорреляции остатков первого порядка и критерием Дарбина - Уотсона: $r_1^2 = 1 - d / 2$.

Список литературы

Основная литература:	1. Christopher Dougherty. Introduction to Econometrics. Oxford University Press, 2011. – 573 p. 7. Gujarati D.N. Basic Econometrics. McGraw-Hill, 5th edition, 2009. – 922 p. 2.Абдуллаев О.М., Ходиев Б.Ю., Ишназаров А.И. Эконометрика. Учебник. –Т.: Fan va texnologiya. 2007. – 612 с. 3. Абдуллаев О.М., Жамалов М.С. Эконометрическое моделирование. Учебник. –Т.: Fan va texnologiya. 2010. – 612 с

Дополнительная литература	1. Greene W.H.Econometric Analysis. Prentice Hall. 7th edition, 2011. -1232p. 2.Валентинов В,А.Эконометрика: Учебник. –М.:ИТК»Дашков и К»,2009–367с. 3.Кремер Н.Ш.Эконометрика:Учебник–М.:ЮНИТИ-ДАНА,2008. –562с. 4.Айвазян С.А. Прикладная статистика и основы эконометрика:Учебник. –М.ЮНИТИ,2007. –345с. 5.Елисеева И.И., Курышева С.В.и др. Эконометрика:Учебник–М.:Финансы и статистика, 2007. –260с. 6.Habibullayev I Iqtisodiy matematik usullsr va modellar: o‘quv qo‘llamna / O‘zbekiston Respublikasi Oliy va o‘rta maxsus ta`lim vazirligi/–Toshkent:“Tafakkur-Bo‘stoni”, 2012.112 b.
Интернет ресурсы:	www.ziyonet.uz www.lex.uz www.stat.uz www.mehnat.uz www.mf.uz www.tfi.uz www.mfer.uz www.conomical.uz www.cisstat.com www.gks.ru www.unstats.un.org