МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
АЛТАЙСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Т.В. Саженкова, И.В. Пономарёв, С.П. Пронь

МЕТОДЫ АНАЛИЗА ВРЕМЕННЫХ РЯДОВ

Учебно-методическое пособие

Барнаул
Издательство
Алтайского государственного университета
2020

УДК 338.27
ББК 65.05
С 147

Рецензент – профессор кафедры математического анализа АлтГУ,
д. ф. – м. н. ***Е.Д. Родионов***

Саженкова, Т.В.

С 147 Методы анализа временных рядов [Текст]: учебно-методическое пособие / Т.В. Саженкова, И.В. Пономарёв, С.П. Пронь. – Барнаул: Изд-во Алт. ун-та. – 2020. – 60 с.

Учебно-методическое пособие предназначено студентам II курса магистратуры института математики и информационных технологий направления подготовки 02.04.01 "Математика и компьютерные науки", впервые знакомящимся с методами анализа временных рядов, но изучавшим ранее теорию вероятностей и математическую статистику. Объем материала соответствует семестровому курсу при 36-ти часовой аудиторной нагрузке. В пособии приведен необходимый теоретический материал, а также контрольные вопросы и практические задания по всем изучаемым разделам дисциплины.

УДК 338.27
ББК 65.05

Оглавление

ВВЕДЕНИЕ

Пособие «Методы анализа временных рядов» включает в себя разделы, касающиеся основных этапов моделирования и прогнозирования динамики экономических показателей, социально-экономических явлений и процессов, представленных временными рядами. Важнейшая задача прогнозирования явлений и процессов – это выявление закономерностей и установление основных тенденций развития. Для анализа общих тенденций не целесообразно рассматривать каждый случай в отдельности. Чем больше по числу единиц статистическая совокупность, тем качественнее проявляется закономерность, присущая изучаемому явлению или процессу. Для достоверности прогноза, полученного с использованием математической модели, необходимо уметь определять пределы возможного изменения прогнозируемого показателя, так как именно границы моделируемого интервала характеризуют амплитуду отклонений фактических данных от прогнозируемых.

Целью учебного курса является ознакомление обучаемых с теоретическими основами анализа временных рядов. Курс направлен на изучение методов математического моделирования и прогнозирования социально-экономических явлений и процессов. А также предполагается формирование практических навыков по экономико-статистическому анализу состояния и моделирования конкретных социально-экономических явлений и процессов на основе построения адекватных, и, в достаточной степени аппроксимирующих реальные явления и процессы, математических моделей. На основе полученных моделей возможна выработка конкретных предложений, рекомендаций и путей их прикладного использования. Моделирование и прогнозирование динамики социально-экономических явлений позволяют управлять экономическими явлениями и процессами и предвидеть их развитие.

Методы анализа временных рядов и прогнозирования находят широкое применение во многих областях знаний, как на макро, так и на микроуровнях исследования процессов экономического развития, оценке их эффективности, финансового состояния и финансовой устойчивости, деловой активности сегментов различных рынков и организационно-правовых структур.

Конечно, наиболее эффективным и целесообразным в решении задач анализа и прогнозирования временных рядов является широкое использование прикладного программного обеспечения, что существенно ускоряет проведение расчетов. В этой связи при решении задач исследования конкретных социально-экономических явлений и процессов распространены следующие программные продукты, такие как стандартные пакеты прикладных программ MS Excel, STATISTIKA, ОЛИМП, SPSS, STATGRAPHICS, Statistica Neural Networks и другие.

Относительно использования этих пакетов и прикладных программ следует обратиться к соответствующей литературе и источникам, в некотором количестве представленной в библиографическом списке данного пособия.

1. Понятие временного ряда. Классификация. Правила построения

Предметом моделирования и прогнозирования в сфере экономики является система, воспроизводящая объект исследования так, что на ее основе могут быть изучены структура и размещение социально-экономических явлений, их изменения во времени, связи и зависимости, следовательно, **объектом моделирования** и прогнозирования являются достаточно объёмные по численности статистические совокупности.

Анализ многих экономических процессов осуществляется через использование так называемых временных рядов. Во временном ряде содержится информация об особенностях и закономерностях протекания процесса. Анализ временного ряда позволяет выявить закономерности и использовать их для оценки характеристик процесса в будущем.

Временной ряд – это последовательность упорядоченных во времени числовых показателей, характеризующих уровень развития изучаемого явления.

Всякий ряд динамики включает два обязательных элемента: время и конкретное значение показателя – уровень ряда. В качестве показателя времени в рядах динамики могут указываться либо определенные моменты времени (даты), либо отдельные периоды (сутки, месяцы, кварталы, полугодия, годы и т.д.). Под длиной временного ряда понимают количество входящих в него уровней n. Общепринятое обозначение временного ряда Y_t или y_t, где $t = 1, ..., n$.

Используя ограниченное количество информации – временной ряд конечной длины – исследования позволяют выяснить механизм, порождающий ряд, проанализировать структуру ряда, описать характерные особенности ряда и предсказать будущее на основании знания прошлого.

Временные ряды классифицируются по четырём, далее перечисленным признакам:

1). В зависимости от характера временного параметра ряды делятся на **моментные и интервальные ряды**.

В интервальном ряду уровни ряда выражают величину явления за определенные интервалы времени, (например, за сутки, месяц, год и т.д.) Пример интервального ряда динамики, характеризующего финансовые результаты деятельности организаций:

Показатель	январь	февраль	март	апрель	май
финансовый результат (прибыль минус убыток), млн. руб.	342,1	315,5	417,1	308,5	431,9

Моментные ряды динамики отображают состояние изучаемого явления на определенные даты (моменты) времени. Примерами моментных рядов могут быть последовательности показателей численности населения на начало года, величины запаса какого-либо материала на начало периода, котировок акций на определенную дату и т. д. Примером моментного ряда абсолютных величин может служить временной ряд, характеризующий данные об остатках задолженности предприятия по кредиту:

Показатель на начало месяца	01.01.	01.02.	01.03.	01.04.	01.05.	01.06.
тыс. руб.	150	140	135	130	120	100

Важное аналитическое отличие моментных рядов от интервальных состоит в том, что сумма уровней интервального ряда дает вполне реальный показатель. Например, общий выпуск продукции за год, общие затраты рабочего времени, общий объем продаж акций и т. д., сумма же уровней моментного ряда реального содержания, как правило, не имеет, так как отдельные уровни моментного ряда содержат элементы повторного счета.

2). По форме представления уровней различают ряды **абсолютных, относительных и средних величин**.

Ряд динамики величин, характеризующих финансовые результаты деятельности организаций, является рядом абсолютных величин. Далее представлен ряд динамики относительных величин. Это ряд вложений банков в ценные бумаги, в % к активам банковского сектора:

Показатель	01.01. 2002	01.01. 2003	01.01. 2004	01.01. 2005	01.01. 2006	01.01. 2007
Вложения банков в ценные бумаги, в % к активам банковского сектора	17,8	18,8	17,9	15,2	15,8	14

В следующей таблице приведён ряд динамики средних величин – среднедушевые номинальные денежные доходы населения России в месяц (Российский статистический ежегодник. 2006):

Показатель	2000	2001	2002	2003	2004	2005
Среднедушевые денежные доходы в месяц, руб.	2281,1	3062,0	3947,2	5170,4	6410,3	8023,2

3). По расстоянию между датами или интервалам времени ряды динамики подразделяются на ряды динамики **с равноотстоящими уровнями** (уровни представлены через равные, следующие друг за другом интервалы (моменты) времени) и **неравноотстоящими уровнями**. Рядами с неравноотстоящими уровнями являются, например, ряды с равноотстоящими уровнями, в которых имеются пропуски значений уровней ряда, или ряды с неравномерными интервалами (моментами) времени.

4). Временные ряды подразделяются на **стационарные и нестационарные** ряды. Случайные процессы, протекающие во времени приблизительно однородно и имеющие вид непрерывных случайных колебаний вокруг некоторого среднего значения, причем ни средняя амплитуда, ни характеристика этих колебаний не обнаруживают существенных изменений с течением времени, в математической статистике называются стационарными. Всякий стационарный процесс можно рассматривать как процесс, неопределенно долго продолжающийся во времени. В связи с этим при проведении исследования в качестве начала отсчета можно выбрать любой момент времени. При этом на любом интервале времени должны быть получены одни и те же характеристики. В экономической практике в большинстве случаев приходится иметь дело со случайными процессами, имеющими вполне определенную тенденцию развития во времени. Такие процессы называются нестационарными, и временные ряды также называются нестационарными. Характеристики нестационарного случайного процесса меняются во времени, то есть зависят от начала отсчета.

Таким образом, исходя из выше сказанного, получаем итоговую таблицу.

Классификация временных рядов

Признак классификации	Виды временного ряда
1. В зависимости от того, как уровни выражают состояния явлений во времени	1. Интервальные ряды 2. Моментные ряды
2. В зависимости от качественной особенности изучаемого явления.	1. Абсолютных величин 2. Относительных величин 3. Средних величин
3. В зависимости от расстояния между уровнями.	1. С равноотстоящими уровнями по времени 2. С не равноотстоящими уровнями по времени
4. В зависимости от наличия тенденции изучаемого процесса.	1. Стационарные ряды 2. Нестационарные ряды

Успешность статистического анализа развития процессов во времени во многом зависит от правильного построения рядов динамики. Поэтому к составлению рядов динамики предъявляются определенные требования. Далее перечислены основные из этих требований.

1. Периодизация развития, т. е. расчленение его во времени на однородные этапы, в пределах которых показатель подчиняется одному закону развития. Вопрос о том, какие этапы развития прошло то или иное явление за определенный исторический отрезок времени, решается теорией той науки, к области которой относится изучаемая совокупность явления. При этом обращают внимание на значимые даты и события. А именно, время принятия управленческих решений по данному показателю, смену хозяйственного механизма, смену руководства, войну и т.д.

2. Сопоставимость. Вторым важным условием правильного построения ряда динамики является сопоставимость всех входящих в него уровней по территории, кругу охватываемых объектов, единицам измерения, времени регистрации, ценам, методологии расчета. Проблема сопоставимости данных особенно остро стоит в рядах динамики, потому что они охватывают значительные периоды времени, за которые могли произойти изменения и привести к несопоставимости статистических данных. Данное условие решается либо в процессе сбора и обработки данных, либо путем их пересчета. Выделяют следующие основные виды сопоставимости, играющие основополагающее значение в исследованиях.

Сопоставимость по территории предполагает одни и те же границы территории, изменение границ влияет на численность населения и объем продукции.

Сопоставимость по кругу охватываемых объектов означает сравнение совокупностей с равным числом элементов. Например, при характеристике динамики численности студентов вузов нельзя в одни годы учитывать только студентов дневных отделений, а в другие – численность студентов всех видов обучения.

Сопоставимость по времени регистрации для интервальных рядов динамики обеспечивается равенством периодов времени, за которые приводятся данные. Для моментных рядов показатели следуют приводить на одну и ту же дату.

Сопоставимость по ценам. На практике количество продукции, произведенной в разные периоды, оценивают в ценах одного и того же базисного периода, которые называют неизменными или сопоставимыми ценами.

Сопоставимость по методологии расчета. При определении уровней динамического ряда необходимо использовать единую методологию расчета.

Таким образом, прежде чем анализировать динамический ряд, надо, исходя из цели исследования, убедиться в сопоставимости уровней ряда. Если данные несопоставимые, они могут быть приведены к сопоставимому виду дополнительными расчетами.

3. Соответствие величины временных интервалов интенсивности изучаемых процессов. Чем больше вариация уровней во времени, тем чаще следует делать замеры. Соответственно для стабильных процессов интервалы можно увеличить. Так, переписи населения достаточно проводить один раз в десять лет, учет национального дохода, урожая – раз в год. Ежедневно регистрируются курсы покупки и продажи валют, ежечасно температура воздуха и т.д.

4. Упорядоченность числовых уровней рядов динамики во времени. Не допускается анализ рядов с пропусками отдельных уровней; если же такие пропуски неизбежны, то их восполняют условными расчетными значениями.

Применение к временным рядам определенного математического аппарата, а именно методов математической статистики, предъявляет к исходным данным ряд дополнительных требований: **однородность** данных, **устойчивость** тенденции, **полнота** данных.

5. Однородность данных означает отсутствие сильных изломов тенденций, а также аномальных, т.е. резко выделяющихся, нетипичных для данного ряда наблюдений. **Аномальные наблюдения** проявляются в виде сильного изменения уровня – скачка или спада, с последующим приблизительным восстановлением предыдущего уровня. Наличие аномалии резко искажает результаты моделирования, поэтому анализ временных рядов должен также начинаться с выявления и устранения аномальных (нехарактерных) значений уровней ряда. Обычно **аномальные значения** можно обнаружить визуально, при помощи графического представления временных рядов, но, прежде чем «подправить» обнаруженные таким образом значения ряда, их необходимо подвергнуть дальнейшему количественному и качественному анализу.

Нехарактерные уровни во временном ряду можно подразделить на три группы:

значения, отражающие объективное развитие процесса, но сильно отличающиеся от общей тенденции, так как они проявляют свои экстремальные воздействия крайне редко;

значения, возникающие вследствие изменений методики расчета;

значения, возникающие вследствие ошибок при измерении показателя, при записи и передаче информации, а также значения, связанные с различными катастрофическими явлениями, не влияющими на дальнейший ход развития явления, агрегировании и дезагрегировании показателей и т. д.

Аномальные значения первой группы не всегда должны исключаться из временного ряда и могут даже оказаться полезными на этапе исследования причинно-следственного механизма развития явления. Наличие нехарактерных пиковых значений для одного и того же момента времени в различных временных рядах свидетельствует, как правило, о причинных связях между соответствующими показателями.

Нехарактерные значения второй группы не должны исключаться из рассмотрения, а приниматься за «повторные» (пороговые), начиная с которых, должны быть пересчитаны по новой методике все предыдущие значения временного ряда.

Аномальные значения третьей группы (ошибки первого рода) должны быть исключены из рассмотрения в любом случае, так как они искажают представление о характере развития явления и могут оказать существенное влияние на выводы, полученные в результате анализа ряда, содержащего такую искаженную информацию. Для выявления и замены аномальных значений третьей группы существует ряд аналитических методов. Правда большинство из них разрабатывалось для статистических совокупностей, содержащих независимые и случайные наблюдения, что, вообще говоря, не является справедливым для экономических временных рядов.

6. Устойчивость тенденции характеризуется преобладанием закономерности над случайностью в изменении уровней ряда. На графиках устойчивых временных рядов закономерность прослеживается визуально, на графиках неустойчивых рядов изменения последовательных уровней представляются хаотичными, поэтому поиск закономерностей в формировании значений уровней таких рядов лишен смысла.

7. Полнота данных – требование, обусловленное тем, что закономерность может обнаружиться лишь при наличии минимально допустимого объема наблюдений.

2. Структура временного ряда: систематические и случайная компоненты

В практике моделирования и прогнозирования экономических процессов принято считать, что значения уровней временных рядов могут содержать следующие **структурообразующие элементы**:

- тренд, который будем обозначать $U_t\,(u_1, u_2, ..., u_n)$;
- сезонную компоненту – V_t;
- циклическую компоненту – C_t;
- случайную компоненту – E_t.

Эти компоненты временного ряда не наблюдаемы, они являются теоретическими величинами.

Под трендом понимают изменение, определяющее общее направление развития, основную тенденцию временного ряда. Это систематическая составляющая долговременного действия.

Тренд – некоторая аналитическая функция, которая связывает единым «законом движения» все последовательные уровни временного ряда. Тренд описывает общую тенденцию на базе лишь одного фактора – фактора времени t. Следовательно, не полностью описывает характер тенденции развития и не может рассматриваться как закон развития явления.

Наряду с долговременными тенденциями во временных рядах часто возникают более или менее регулярные колебания – периодические составляющие рядов динамики.

Сезонные колебания – это колебания, связанные со сменой времен года и имеющие выраженную годовую периодичность. Сезонные колебания наиболее отчетливо проявляются в сельском хозяйстве, в отраслях непосредственно занимающихся переработкой сельхозпродукции, в транспорте, в строительстве, в бытовом обслуживании, в торговле и потреблении.

При периоде колебания больше года считают, что во временных рядах имеет место **циклическая** составляющая. Примерами могут служить циклы деловой активности, демографические, инвестиционные и другие циклы. В экономических временных рядах редко предоставляется возможность для выделения и дальнейшего анализа циклической компоненты, так как ряды динамики экономических показателей часто оказываются слишком «короткими» для проведения такого исследования. Хорошо изучены циклические составляющие в рядах динамики, относящихся к естественным наукам. Например, по многолетним наблюдениям установлена цикличность солнечной активности (с периодом колебаний примерно в 11 лет).

Если из временного ряда удалить тренд и периодические составляющие, то останется нерегулярная компонента или случайная. **Случайная компонента** является действием большого числа относительно слабых второстепенных факторов. Влияние каждого из второстепенных факторов незначительно, но ощущается их суммарное воздействие.

Первые три компоненты (тренд, сезонная и циклическая компоненты) называют **систематическими**, детерминированными или регулярными, так как они вычисляются по определенному правилу как функция от времени t. В процессе формирования значений уровней определенного временного ряда не обязательно присутствуют все систематические компоненты. Случайная компонента является обязательной составляющей любого временного ряда, отображающего развитие экономического процесса, так как элемент случайности непременно присутствует в экономике. Если систематические компоненты временного ряда выделены правильно, то остающаяся после выделения из временного ряда этих компонент остаточная компонента будет случайной величиной.

Чаще всего встречаются временные ряды, в которых можно установить тенденцию и случайный компонент, особенно при использовании годовых данных, где влияние сезонности не отражается. Отметим, что в дальнейшем мы не будем рассматривать циклическую компоненту временных рядов; укажем только, что для моделирования и прогнозирования сезонных и циклических экономических процессов используются специальные методы (индексный и спектральный анализ, выравнивание по ряду Фурье и др.).

В зависимости от связи компонент ряда между собой может быть построена аддитивная или мультипликативная модель временного ряда.

В **аддитивной модели** результирующая компонента получается суммированием тренда и сезонной компоненты: $Y=U+V+E$.

В **мультипликативной модели** результирующая компонента временного ряда получается умножением тренда и сезонной компоненты: $Y=U \cdot V+E$. Данную модель называют еще и смешанной.

Все компоненты временного ряда взаимосвязаны между собой и являются теоретическими понятиями. С этой точки зрения разделение временных рядов на компоненты – это теоретическая абстракция, так как данное разделение является чисто математической процедурой и осуществляется на базе статистических методов. Но, несмотря на условность такого расчленения фактических уровней рядов, такой прием может оказаться довольно полезным для решения разных проблем анализа и прогнозирования на базе временных рядов. По поводу расчленения временных рядов на компоненты известный русский ученый Четвериков Н. С. отмечал, что «расчленению подвергается динамика, а не само явление, участвующее не раздельно во всем сложном движении».

Тип связи между компонентами можно также определить по динамике отклонений эмпирических значений уровней временного ряда от теоретических уровней, полученных по уравнению тренда. Если абсолютные отклонения имеют тенденцию к росту, а относительные варьируют приблизительно на одинаковом уровне, то это свидетельствует о мультипликативной связи тренда и сезонного компонента.

На практике выделить компоненты сложно, так как отдельные последующие значения уровней временных рядов зависят от предыдущих. Поэтому неверно допускать, что факторы, влияющие на колебания уровней, независимы. Кроме того, статистическая совокупность, изучаемая в течение длительного периода, перестает быть такой же самой совокупностью, так как могут измениться основные факторы, влияющие на ее формирование.

На первой стадии анализа для разложения рядов на составные компоненты и устранения влияния систематических компонент на изменение случайного компонента можно применить различные методы определения и установления отдельных неслучайных компонент временных рядов. Прежде чем перейти к выделению основной тенденции развития явления, следует проверить гипотезу о том, существует ли она вообще. Отсутствие тенденции означает неизменность среднего уровня ряда во времени.

Экономическим явлениям свойственны элементы вероятностного характера. Наличие случайного в социально-экономических явлениях объясняется сложным переплетением параметров экономической системы, влиянием на них большого числа взаимосвязанных факторов, действующих в разных направлениях. Это ведет к вариации показателей уровней временного ряда.

Итак, ввиду концепции о наличии вероятностных элементов в динамике процессов, уровни временного ряда могут рассматриваться как сумма детерминированного и случайного компонентов. Детерминированный компонент выражается некоторой функцией и определяется уравнением основной тенденции или тренда. Проявление случайного компонента оценивается с некоторой вероятностью. Отклонения фактических уровней временного ряда от тренда рассматривается как стационарный случайный процесс.

Контрольные вопросы и задания

1. Дайте определение временного ряда.
2. Назовите основные признаки классификации временных рядов. Приведите примеры временных рядов различных видов.
3. Какие требования предъявляются к составлению временных рядов?
4. В чём заключается сопоставимость уровней временного ряда?
5. Что такое аномальный уровень временного ряда?
6. Какие структурообразующие элементы (компоненты) выделяют при моделировании экономических процессов?
7. Что понимается под термином – тренд временного ряда?
8. Дайте характеристику сезонной и циклической компонент временного ряда.
9. Что означает термин – систематические компоненты временного ряда?
10. Какая модель временного ряда называется аддитивной (мультипликативной)?

3. Предварительный анализ и сглаживание временных рядов

Соответствие исходной информации всем указанным выше требованиям проверяется на этапе предварительного анализа временных рядов. Лишь после этого переходят к расчету и анализу основных показателей динамики развития, построению моделей прогнозирования, получению прогнозных оценок.

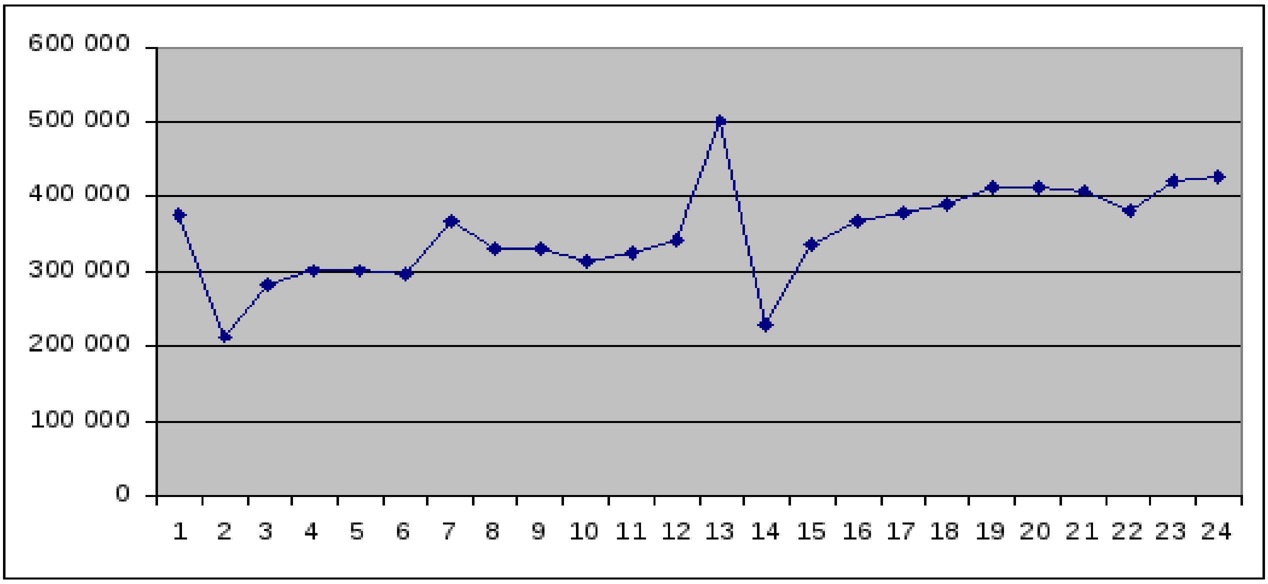
В ходе предварительного анализа определяют, соответствуют ли имеющиеся данные требованиям, предъявляемым к ним математическими методами, строят график временного ряда, рассчитывают основные динамические характеристики, проверяют наличие тренда, сглаживают временной ряд.

График временного ряда дает общее представление о динамике изучаемого процесса. Современные программные средства предоставляют исследователю большие возможности по построению различных графиков, в том числе и временных рядов. Графически ряды динамики изображаются в основном либо линейными, либо столбиковыми диаграммами.

По графику временного ряда можно сделать много выводов, в дальнейшем эти выводы могут быть проверены с помощью расчетов.

Визуальный анализ графиков временного ряда может позволить определить наличие тренда и его характер, наличие сезонных и циклических компонент. Выявление сезонных колебаний удобно производить по графику временного ряда, построенного методом наложения. В этом случае по оси абсцисс откладывается интервал времени предполагаемого периода колебаний, например, год, с разбивкой по месяцам. А по оси ординат – значения уровней ряда за несколько лет. Если во временном ряду имеются периодические изменения, то на графике наблюдаются пики или впадины в определенный период времени.

Пример линейного графика динамики поступления единого социального налога (ЕСН) в некий региональный бюджет за некоторые 24 месяца (в тыс. руб.):



месяц	ЕСН, тыс. руб.	месяц	ЕСН, тыс. руб.
1	376842	13	501145
2	213854	14	229470
3	282036	15	336705
4	302301	16	368367
5	300327	17	378257
6	296497	18	388235
7	369010	19	413939
8	330873	20	414 123
9	330538	21	408 899
10	314843	22	382 538
11	324831	23	421 251
12	340315	24	427 532

Необходимой процедурой предварительного анализа временных рядов экономических показателей является выявление и **устранение аномальных уровней** ряда. В предварительный анализ входит также сглаживание (выравнивание) временных рядов, так как очень часто уровни экономических рядов динамики колеблются, при этом тенденция развития экономического явления во времени скрыта случайными отклонениями уровней в ту или иную сторону. Часто аномальные значения уровней временного ряда можно обнаружить визуально по графику ряда. С целью более четко выявить тенденцию развития исследуемого процесса, в том числе для дальнейшего применения методов прогнозирования на основе трендовых моделей на предварительной стадии производится **сглаживание** (выравнивание) временных рядов.

Для выявления аномальных уровней временных рядов аналитически используются методы, рассчитанные для статистических совокупностей.

Метод Ирвина, например, предполагает использование следующей формулы:

$$\lambda_t = \frac{|y_t - y_{t-1}|}{\sigma_y}, \quad t = 2, ..., n,$$

где среднеквадратическое отклонение σ_y рассчитывается с использованием формул:

$$\sigma_y = \sqrt{\frac{\sum_{t=1}^{n}\left(y_t - \bar{y}\right)^2}{n-1}}, \quad \bar{y} = \frac{\sum_{t=1}^{n} y_t}{n}.$$

Расчетные значения λ_t сравниваются с табличным значением критерия Ирвина $\lambda_{кр}$, значение которого зависит от количества членов временного ряда n и уровня значимости α. Если расчетное значение λ_t больше табличного значения $\lambda_{кр}$, соответствующее значение уровня ряда y_t считается аномальным.

n	2	3	10	20	30	50	100
α=0,05	2,8	2,2	1,5	1,3	1,2	1,1	1,0
α=0,01	3,7	2,9	2,0	1,8	1,7	1,6	1,5

В приведённом фрагменте справочной таблицы значений $\lambda_{кр}$ критерия Ирвина представлены значения для уровней значимости, соответствующих 5%-ной и 1%-ной ошибкам.

После выявления аномальных значений уровней временного ряда необходимо установить причины их возникновения. Аномальные наблюдения первого рода подлежат устранению. Наиболее простой подход заключается в замене аномального наблюдения средней арифметической простой из двух соседних уровней ряда.

Перейдем к вопросу сглаживания временных рядов экономических показателей.

Методы сглаживания временных рядов делятся на две основные группы:

1) механическое выравнивание отдельных уровней временного ряда с использованием фактических значений соседних уровней;

2) аналитическое выравнивание с использованием кривой, проведенной между конкретными уровнями ряда так, чтобы она отображала тенденцию, присущую ряду, и одновременно освобождала его от незначительных колебаний. Эти методы рассматриваются позже.

Суть методов механического сглаживания заключается в следующем. Берется несколько первых уровней временного ряда, образующих интервал сглаживания. Для них подбирается полином, степень которого должна быть меньше числа уровней, входящих в интервал сглаживания; с помощью полинома определяются новые, выровненные значения уровней в середине интервала сглаживания. Далее интервал сглаживания сдвигается на один уровень ряда вправо, вычисляется следующее сглаженное значение и т.д.

Самым простым методом механического сглаживания является **метод простой скользящей средней**. Сначала для временного ряда y_1, y_2, y_3,...,y_n определяется интервал сглаживания m ($m<n$). Если необходимо сгладить мелкие беспорядочные колебания, то интервал сглаживания берут по возможности большим; интервал сглаживания уменьшают, если нужно сохранить более мелкие колебания. При прочих равных условиях интервал сглаживания рекомендуется брать нечетной длины. Для первых m уровней временного ряда вычисляется их среднее арифметическое; это будет сглаженное значение уровня ряда, находящегося в середине интервала сглаживания. Затем интервал сглаживания сдвигается на один уровень вправо, повторяется вычисление среднего арифметического и т.д. Таким образом, для вычисления сглаженных уровней ряда применяется формула:

$$\bar{y}_t = \frac{\sum_{k=t-p}^{t+p} y_k}{m}, \qquad t > p,$$

где $p = \frac{m-1}{2}$ (при нечетном m); для четных m формула усложняется.

В результате такой процедуры получаются n-m+1 сглаженных значений уровней ряда. При этом первые p и последние p уровней ряда теряются (не сглаживаются). Другой недостаток метода в том, что он применим лишь для рядов, имеющих линейную тенденцию.

Когда тренд выравниваемого ряда имеет изгибы и к тому же желательно сохранить мелкие волны, использовать для сглаживания метод простой скользящей средней нецелесообразно. Такой вывод целесообразен, поскольку при этом: выравниваются и выпуклые и вогнутые линии, происходит сдвиг волны вдоль ряда, изменяется знак волны, т.е. на кривой, соединяющей сглаженные уровни, вместо выпуклого участка образуется вогнутый и наоборот. Последнее имеет место в тех случаях, когда интервал сглаживания в полтора раза превышает длину волны.

Если временной ряд содержит периодические колебания, то для устранения периодических колебаний интервал сглаживания должен быть равен периоду колебаний. Тогда для сглаживания сезонных колебаний с годовой или квартальной периодичностью рекомендуется использовать центрированную скользящую среднюю (тринадцати- или пятичленную):

$$\tilde{y}_t = \frac{\frac{1}{2}y_{t-6} + y_{t-5} + y_{t-4} + y_{t-3} + y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2} + y_{t+3} + y_{t+4} + y_{t+5} + \frac{1}{2}y_{t+6}}{12},$$

$$\tilde{y}_t = \frac{\frac{1}{2}y_{t-2} + y_{t-1} + y_t + y_{t+1} + \frac{1}{2}y_{t+2}}{4}.$$

При нелинейном характере развития более надежным является использование других методов сглаживания. Например, разумно использование метода взвешенной скользящей средней.

Метод взвешенной скользящей средней отличается от предыдущего метода сглаживания тем, что уровни, входящие в интервал сглаживания, суммируются с разными весами. Это связано с тем, что аппроксимация ряда в пределах интервала сглаживания осуществляется с использованием полинома не первой степени, как в предыдущем случае, а степени, начиная со второй. Используется формула средней арифметической взвешенной:

$$\bar{y}_t = \frac{\sum_{k=t-p}^{t+p} \rho_k y_k}{\sum_{k=t-p}^{t+p} \rho_k},$$

причем веса ρ_k определяются с помощью метода наименьших квадратов. Эти веса рассчитаны для различных степеней аппроксимирующего полинома и различных интервалов сглаживания. Так, для полиномов второго и третьего порядков числовая последовательность весов при интервале сглаживания m=5 имеет вид:{-3; 12; 17; 12; -3}, а при m=7 имеет вид: {-2; 3; 6; 7; 6; 3; -2}. Для полиномов четвертой и пятой степеней и при интервале сглаживания m=7 последовательность весов выглядит следующим образом: {5; -30; 75; 131; 75; -30; 5}.

К этой же группе методов выравнивания временных рядов примыкает **метод экспоненциального сглаживания**. Его особенность заключается в том, что в процедуре нахождения сглаженного уровня используются значения только предшествующих уровней ряда, взятые с определенным весом, причем вес наблюдения уменьшается по мере удаления его от момента времени, для которого определяется сглаженное значение уровня ряда.

Если для исходного временного ряда соответствующие сглаженные значения уровней обозначить через S_t, t = 1, 2, ..., n, то экспоненциальное сглаживание осуществляется по формуле:

$$S_t = \alpha y_t + (1\text{-}\alpha)\, S_{t\text{-}1}\ ,$$

где α – **параметр сглаживания** ($0<\alpha<1$); величина β = 1-α называется **коэффициентом дисконтирования**.

Используя приведенное выше рекуррентное соотношение для всех уровней ряда, начиная с первого и кончая моментом времени t, можно получить, что экспоненциальная средняя, т.е сглаженное данным методом значение уровня ряда, является взвешенной средней всех предшествующих уровней:

$$S_t=\alpha\sum_{i=0}^{t-1}(1-\alpha)^i y_{t-i} + (1-\alpha)^t S_0,$$

здесь S_0 – величина, характеризующая начальные условия.

Важную роль в методе экспоненциального сглаживания играет выбор оптимального параметра сглаживания α, так как именно он определяет оценки коэффициентов модели, а, следовательно, и результаты прогноза

В зависимости от величины параметра прогнозные оценки по-разному учитывают влияние исходного ряда наблюдений: чем больше α, тем больше вклад последних наблюдений в формирование тренда, а влияние начальных условий быстро убывает. При малом α прогнозные оценки учитывают все наблюдения, при этом уменьшение влияния более «старой» информации происходит медленно. Известны два основных соотношения, позволяющие найти приближенную оценку α. Первое соотношение Р. Брауна, выведенное из условия равенства скользящей и экспоненциальной средней $\alpha = \frac{2}{n+1}$. Вторым является соотношение Мейера $\alpha \approx \frac{\sigma_n}{\sigma_\varepsilon}$, где σ_n – среднеквадратическая ошибка модели; σ_ε – среднеквадратическая ошибка исходного ряда.

Однако использование последнего соотношения затруднено тем, что достоверно определить σ_n и σ_ε из исходной информации очень сложно.

В практических задачах обработки экономических временных рядов рекомендуется (необоснованно) выбирать величину параметра сглаживания в интервале от 0,1 до 0,3. Если при подходе к правому концу временного ряда сглаженные этим методом значения при выбранном параметре α начинают значительно отличаться от соответствующих значений исходного ряда, необходимо перейти на другой параметр сглаживания. Других точных рекомендаций для выбора оптимальной величины параметра α пока нет.

Что касается начального параметра S_0, то в конкретных задачах его берут или равным значению первого уровня ряда, или равным среднему арифметическому нескольких первых членов ряда, например, членов y_1, y_2, y_3 :

$$s_0 = \frac{y_1 + y_2 + y_3}{3}.$$

Этот порядок выбора величины S_0 обеспечивает хорошее согласование сглаженного и исходного рядов для первых уровней.

Заметим, что при этом методе сглаживания не теряются ни начальные, ни конечные уровни сглаживаемого временного ряда.

Контрольные вопросы и задания

1. Какие графики используются для изображения временного ряда?

2. Какие процедуры выполняют в ходе предварительного анализа временного ряда?

3. Какие методы выявления аномальных уровней временного ряда существуют?

4. (*a*). Даны сведения о выручке предприятия от реализации продукции по месяцам, в тыс. руб.:

1	2	3	4	5	6	7	8	9	10	11	12
456	476	566	522	662	687	702	743	721	752	784	767

(*b*). В следующей таблице приведены годовые данные о трудоемкости производства 1т цемента (в нормо-сменах) за 10 лет:

1	2	3	4	5	6	7	8	9	10
7,9	8,3	7,5	6,9	7,2	6,5	5,8	4, 9	5,1	4,4

Определите наличие аномальных уровней для данных временных рядов методом Ирвина.

5. Поясните сущность методов сглаживания.

6. Перечислите основные шаги алгоритма сглаживания по простой скользящей средней.

7. Как проводится сглаживание временного ряда с периодическими колебаниями?

8. В чем состоит отличие алгоритма метода взвешенной скользящей средней от простой?

9. Сколько значений теряется при использовании скользящей средней с длиной интервала сглаживания $m=7$?

10. В чем состоит особенность метода экспоненциального сглаживания?

11. Объясните, как выбирается сглаживающий параметр α при реализации процедуры экспоненциального сглаживания.

12. Сгладить временные ряды, данные в задании 4, используя:

а) пятичленную простую скользящую среднюю для ряда (*a*);

б) пятичленную взвешенную скользящую среднююдля ряда (*b*);

в) методом экспоненциального сглаживания, приняв сглаживающий параметр *α=0,4* для ряда (*a*) и *α=0,1* и *α=0,3* для ряда *(b)*.

Результаты сглаживания показать графически.

4. Расчет показателей динамики развития экономических процессов

При изучении явления во времени исследователь часто сталкивается с необходимостью описать интенсивность изменения и рассчитать средние показатели динамики. Эта проблема решается путем построения соответствующих показателей.

Для **количественной оценки динамики** социально-экономических явлений, т.е. для характеристики интенсивности изменения во времени применяются такие статистические показатели как:

- **абсолютный прирост;**
- **темпы роста;**
- **темпы прироста**;
- **абсолютное значение одного процента прироста**.

В основе расчета показателей рядов динамики лежит сравнение его уровней. В зависимости от применяемого способа сопоставления показатели динамики могут вычисляться на постоянной и переменной базах сравнения.

Для расчета показателей динамики на постоянной базе каждый уровень ряда сравнивается с одним и тем же **базисным уровнем** (чаще всего начальным уровнем). Исчисляемые при этом показатели называются базисными.

Для расчета показателей динамики на переменной базе каждый последующий уровень сравнивается с предыдущим. Вычисленные таким образом показатели динамики называются **цепными**.

Формулы расчета показателей динамики:

Показатель	Цепной	Базисный
Абсолютный прирост	$\Delta y_ц = y_i - y_{i-1}$	$\Delta y_б = y_i - y_1$
Темп роста	$T_ц = \frac{y_i}{y_{i-1}}$	$T_б = \frac{y_i}{y_1}$
Темп прироста	$\Delta T_ц = \frac{\Delta y_ц}{y_{i-1}} \cdot 100 = T_ц - 100$ в %.	$\Delta T_б = \frac{\Delta y_б}{y_1} \cdot 100 = T_б - 100$ в %.
Абсолютное значение одного процента прироста	$A = \frac{\Delta y_ц}{\Delta T_ц(\%)} = 0{,}01 \cdot y_{i-1}$	

Цепной абсолютный прирост – разность между сравниваемым уровнем и уровнем, который ему предшествует.
Базисный абсолютный прирост – это разность между сравниваемым уровнем и уровнем, принятым за постоянную базу сравнения.

Между базисными и цепными абсолютными приростами имеется связь: сумма абсолютных приростов равна базисному абсолютному приросту последнего ряда динамики.

Распространенным статистическим показателем является темп роста. Он характеризует отношение двух уровней ряда и может выражаться в виде коэффициентов или в процентах. Цепные темпы роста исчисляются делением сравниваемого уровня на предыдущий. Базисные темпы исчисляются делением сравниваемого уровня на уровень, принятый за постоянную базу сравнения. Если темп роста больше единицы (или 100 %), то это показывает увеличение изучаемого уровня по сравнению со сравниваемым. Темп роста равный единице (или 100%) показывает, что уровень изучаемого периода по сравнению со сравниваемым не изменился. Темп роста меньше единицы (или 100%) показывает на уменьшение изучаемого уровня по сравнению со сравниваемым.

Темп прироста характеризует абсолютный прирост в относительных величинах. Исчисленный в процентах темп прироста показывает, на сколько процентов изменился сравниваемый уровень с уровнем, принятым за базу сравнения.

При анализе динамики развития следует знать, какие абсолютные значения скрываются за темпами роста и прироста. При снижении темпов прироста абсолютный прирост не всегда уменьшается, в отдельных случаях он может возрастать. Поэтому, чтобы правильно оценить значение полученного темпа прироста его рассматривают в сопоставлении с показателем абсолютного прироста. Результат выражают показателем, который называют абсолютным значением одного процента прироста. Абсолютное значение одного процента прироста показывает, какое абсолютное значение скрывается за относительным показателем.

Между базисными и цепными темпами роста имеется взаимосвязь: произведение последовательных цепных коэффициентов (темпов роста) равно базисному темпу роста, а частное от деления последующего базисного коэффициента (темпа роста) на предыдущий равно соответствующему цепному темпу роста.

Следующая группа показателей анализа рядов динамики – **система средних показателей**, применяется для получения обобщающих показателей динамики социально-экономических явлений и включает в себя:

- **средний уровень ряда;**
- **средний абсолютный прирост;**
- **средний темп роста;**
- **средний темп прироста**.

Средний уровень ряда – это показатель, обобщающий итоги развития явления за единичный интервал или момент из имеющейся временной последовательности. Например, средний уровень урожайности за ряд лет лучше опишет урожайность, чем уровень одного года, значение которого формируется под действием множества случайных факторов. Средний уровень ряда определяется по-разному для моментных и интервальных рядов, при этом следует обратить внимание на то, какие – равноотстоящие или не равноотстоящие во времени – уровни наблюдаются в ряду динамики.

Для интервальных рядов с равными периодами времени средний уровень рассчитывается по формуле простой средней арифметической:

$$\overline{y} = \frac{\sum y_i}{n}$$

где n – длина временного ряда.

Если в интервальном ряду интервалы имеют неравную длительность, то средний уровень рассчитывается по формуле средней арифметической взвешенной, где в качестве весовых коэффициентов используется продолжительность интервалов времени между уровнями (число периодов времени, при которых значение уровня не изменяется):

$$\overline{y} = \frac{\sum y_i t_i}{\sum t_i}$$

Для моментного ряда с равноотстоящими датами средний уровень рассчитываем по формуле средней хронологической простой:

$$\overline{y} = \frac{\frac{1}{2}y_1 + y_2 + ... + y_{n-1} + \frac{1}{2}y_n}{n-1}$$

Для моментного ряда динамики с неравноотстоящими датами средний уровень определяется по формуле средней хронологической взвешенной:

$$\overline{y} = \frac{(y_1 + y_2)t_1 + (y_2 + y_3)t_2 + ... + (y_{n-1} + y_n)t_{n-1}}{2(t_1 + t_2 + ... + t_{n-1})},$$

где $t_1, t_2, ...t_{n-1}$ – продолжительность интервала времени между соседними уровнями.

Средний абсолютный прирост представляет собой обобщенную характеристику индивидуальных абсолютных приростов ряда динамики (для определения среднего абсолютного прироста сумма цепных абсолютных приростов делится на число приростов):

$$\Delta\overline{y} = \frac{\Sigma\Delta y_ц}{n-1}$$

Средний темп роста – обобщающая характеристика индивидуальных темпов роста динамики, рассчитывается по формуле средней геометрической простой:

$$\overline{T} = \sqrt[n]{T_{1ц} \cdot T_{2ц} \cdot \ldots \cdot T_{nц}}$$

Средний темп прироста (%) определяется по формуле $\Delta\overline{T} = \overline{T} - 100$.

Показатели динамических рядов имеют большое практическое значение и находят самое широкое применение в анализе общественных явлений и процессов. К недостаткам среднего прироста и среднего темпа роста следует отнести то, что они учитывают лишь конечный и начальный уровни ряда, исключая влияния промежуточных уровней. Тем не менее, эти показатели имеют весьма широкую область применения, что объясняется чрезвычайной простотой их вычисления. Они могут быть использованы как приближенные, простейшие способы прогнозирования, предшествующие более глубокому количественному и качественному анализу.

При статистическом анализе временных рядов часто возникает необходимость, кроме определения основных характеристик ряда, оценить **зависимость изучаемого показателя** y_t от его значений, рассматриваемых с некоторым запаздыванием во времени. В значительной части временных рядов между уровнями, особенно близко расположенными, существует взаимосвязь. Ее удобно представить в виде коррелированной зависимости между рядом y_1, $y_2, \ldots, y_n$ и этим же рядом, но сдвинутым относительно первоначального положения на τ моментов времени $y_{1+\tau}$, $y_{2+\tau}$, $\ldots, y_{n+\tau}$. Временное смещение τ называют сдвигом, а само явление – **автокорреляцией**.

Автокорреляционная зависимость особенно существенна между последующими и предшествующими уровнями временного ряда. Поскольку классические методы математической статистики применимы лишь в случае независимости отдельных членов ряда между собой, то при анализе временных рядов чрезвычайно важно установить наличие и степень их автокорреляции. Для определения уровня автокорреляции используется автокорреляционная функция, состоящая из множества коэффициентов корреляции, вычисленных относительно исходного ряда и этого же ряда, сдвинутого на τ моментов времени. Значение нормированной автокорреляционной функции вычисляется по формуле:

$$r_\tau = \frac{(n-\tau)\sum_{t=1}^{n-\tau} y_t y_{t+\tau} - \sum_{t=1}^{n-\tau} y_t \sum_{t=1}^{n-\tau} y_{t+\tau}}{\sqrt{\left((n-\tau)\sum_{t=1}^{n-\tau} y_t^2 - \left(\sum_{t=1}^{n-\tau} y_t\right)^2\right)\left((n-\tau)\sum_{t=1}^{n-\tau} y_{t+\tau}^2 - \left(\sum_{t=1}^{n-\tau} y_{t+\tau}\right)^2\right)}}.$$

Сдвиг, которому соответствует наибольший коэффициент автокорреляции, называется временным запаздыванием, или **временным лагом**. Совокупность коэффициентов $r_0, r_1, r_2, \ldots$ называют коррелограммой. Это же название используется для обозначения графика r_τ, как функции от τ.

Если временной ряд состоит из уровней, среднее значение которых равно нулю, то выражение автокорреляционной функции значительно упрощается:

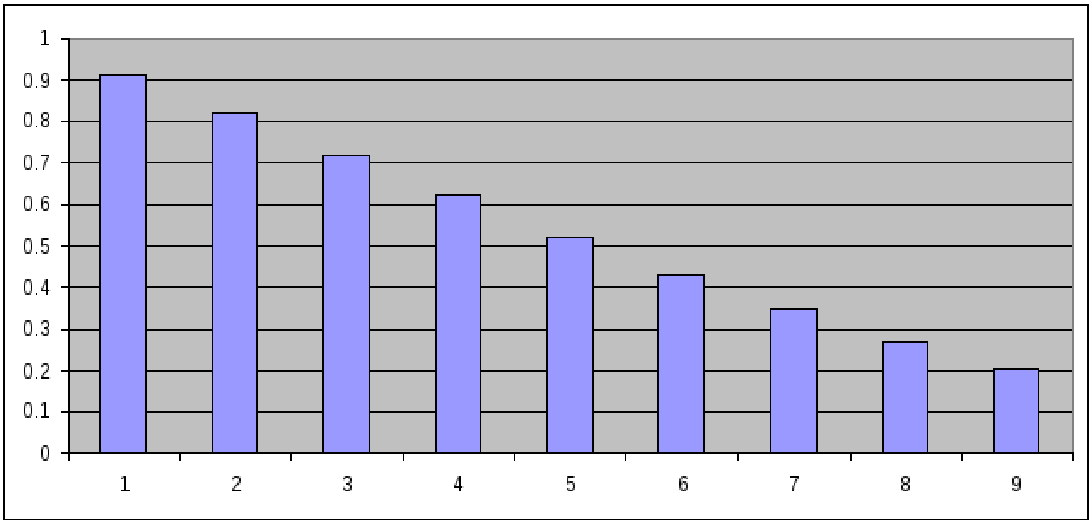
$$r_\tau = \frac{\sum_{t=1}^{n-\tau} y_t y_{t+\tau}}{\sqrt{\sum_{t=1}^{n-\tau} y_t^2 \sum_{t=1}^{n-\tau} y_{t+\tau}^2}}$$

В ряде случаев при больших значениях n используется упрощенная формула для вычисления коэффициента автокорреляции:

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

где $\bar{y}$ – средний уровень ряда.

Коррелограмма ряда «Денежная масса» для заданного ряда денежной массы на начало периода-месяца за три года (млрд.руб.):



Месяц (на начало месяца, млрд.руб)	1 год	2 год	3 год
Январь	130,4	187,8	266,6
Февраль	116,4	178	232,9
Март	120,4	180,8	242
Апрель	119,1	174,1	251,5
Май	128,6	195,2	279,1
Июнь	129,9	205,3	289,3
Июль	129,8	216,4	321,8
Август	129,3	218,2	334
Сентябрь	133,4	216,2	341,6
Октябрь	154,2	212,8	351
Ноябрь	166,4	222	349,7
Декабрь	167,3	219,3	358,3

Здесь при расчётах коэффициентов автокорреляции в Excel использовалась статистическая функция КОРРЕЛ, расчет производился по изначальной формуле для нормированной автокорреляционной функции.

В практических расчетах при построении коррелограмм количество членов автокорреляции r_τ ограничивают числом, содержащимся между $n/4$ и $n/3$. Анализ автокорреляционной функции и коррелограммы позволяет выявить структуру временного ряда. Если наиболее высоким оказался коэффициент автокорреляции первого порядка, то исследуемый ряд содержит только тенденцию. Если наиболее высоким оказался коэффициент автокорреляции порядка τ, то ряд содержит циклические колебания с периодичностью в τ моментов времени. Если ни один из коэффициентов автокорреляции не является значимым, то можно сделать одно из двух предположений относительно структуры этого ряда: либо ряд не содержит тенденции и циклических колебаний, либо ряд содержит сильную нелинейную тенденцию.

Контрольные вопросы и задания

1. С помощью, каких статистических показателей, описывается динамика экономических процессов?

2. Какая взаимосвязь существует между цепными и базисными коэффициентами роста?

3. Что такое автокорреляция уровней временного ряда?

4. Что такое автокорреляционная функция?

5. Что такое временной лаг?

6. В таблице приведены (*a*) данные о доходах от операций с коммерческими кредитами, на начало месяца в тыс. руб.

1	2	3	4	5	6	7	8	9	10	11	12
2380	2978	3270	3451	3633	3733	4008	4235	4467	4439	4590	5110

(*b*) годовые данные о трудоемкости производства 1т цемента (в нормо-сменах) за 10 лет:

1	2	3	4	5	6	7	8	9	10
7,9	8,3	7,5	6,9	7,2	6,5	5,8	4, 9	5,1	4,4

Определить средний уровень ряда, средний абсолютный прирост, средний темп роста и прироста для данных временных рядов

7. Для временных рядов, приведенных в предыдущем задании рассчитать три первых коэффициента автокорреляции, сделать вывод об их значимости. Построить графики автокорреляционных функций. Найти величину временного лага.

5. Методы выявления наличия тенденции во временном ряду

Временные ряды, отображающие развитие экономического процесса, имеют, как правило, тенденцию к возрастанию или убыванию. По графику временного ряда можно определить наличие тенденции, но не всегда тенденция прослеживается достаточно четко. Поэтому прежде, чем перейти к определению тенденции и выделению тренда, необходимо выяснить, существует ли вообще тенденция во временном ряду.

В настоящее время существует большое количество методов, позволяющих определить наличие тенденции во временном ряду. Основные подходы к решению данной задачи основаны на статистической проверке гипотезы о случайности ряда. Рассмотрим некоторые из них.

Метод проверки разности средних уровней. По этому методу временной ряд разбивают на две примерно равные по числу уровней части (в первой части n_1 первых уровней исходного ряда, во второй остальные $n_2=n-n_1$ уровней). Каждая из частей рассматривается как самостоятельная выборочная совокупность, имеющая нормальное распределение. Для каждой из этих частей вычисляются средние значения и дисперсии:

$$\overline{y_1} = \frac{\sum_{t=1}^{n_1} y_t}{n_1}, \quad \sigma_1^2 = \frac{\sum_{t=1}^{n_1} \left(y_t - \overline{y_1}\right)^2}{n_1 - 1},$$

$$\overline{y_2} = \frac{\sum_{t=n_1+1}^{n} y_t}{n_2}, \quad \sigma_2^2 = \frac{\sum_{t=n_1+1}^{n} \left(y_t - \overline{y_2}\right)^2}{n_2 - 1}.$$

Если временной ряд не имеет тенденции, то средние, вычисленные для каждой совокупности, не должны существенно различаться между собой. Таким образом, проверка наличия тренда в исследуемом ряду сводится к проверке гипотезы о равенстве средних двух нормально распределенных совокупностей.

Предварительно проверяется дополнительная гипотеза равенства (однородности) дисперсий обеих частей ряда с помощью F-критерия Фишера, которая основана на сравнении расчетного значения этого критерия:

$$F = \begin{cases} \sigma_1^2 / \sigma_2^2, \text{ если } \sigma_1^2 > \sigma_2^2 \\ \sigma_2^2 / \sigma_1^2, \text{ если } \sigma_1^2 < \sigma_2^2 \end{cases}$$

с табличным (критическим) значением критерия Фишера F_α с заданным уровнем значимости (уровнем ошибки) α. В качестве α чаще всего берут значения 0,1 (10%-ная ошибка), 0,05 (5%-ная ошибка), 0,01 (1%-ная ошибка). Величина 1-α называется доверительной вероятностью.

Если расчетное значение F меньше табличного F_α, то гипотеза о равенстве дисперсий принимается, и переходят к следующему этапу метода. Если F больше или равно F_α, гипотеза о равенстве дисперсий отклоняется и делается вывод, что данный метод для определения наличия тренда ответа не дает.

На следующем этапе проверяется основная гипотеза об отсутствии тренда с использованием t-критерия Стьюдента. Для этого определяется расчетное значение критерия Стьюдента по формуле:

$$t = \frac{|\bar{y}_1 - \bar{y}_2|}{\sigma \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

где σ – среднеквадратическое отклонение разности средних:

$$\sigma = \sqrt{\frac{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}{n_1 + n_2 - 2}}\ .$$

Если расчетное значение t меньше табличного значения статистики Стьюдента t_α с заданным уровнем значимости α, гипотеза принимается, т.е. тренда нет, в противном случае тренд есть. Заметим, что в данном случае табличное значение t_α берется для числа степеней свободы, равного n_1+n_2-2, при этом данный метод применим только для рядов с монотонной тенденцией.

Метод Фостера-Стьюарта. Этот метод дает более надежные результаты и обладает большими возможностями по сравнению с рассмотренными ранее. Кроме определения наличия тенденции процесса, метод Фостера-Стьюарта позволяет обнаружить тренд дисперсии временного ряда (если тренда дисперсии нет, то разброс уровней ряда постоянен, если дисперсия увеличивается, то ряд «раскачивается»), что важно знать при анализе и прогнозировании экономических процессов.

Реализация метода содержит *четыре этапа*.

На *первом этапе* производится сравнение каждого уровня исходного временного ряда, начиная со второго уровня, со всеми предыдущими, при этом определяются две числовые последовательности:

$$k_t = \begin{cases} 1, & \text{если } y_t \text{ больше всех предыдущих уровней,} \\ 0, & \text{в противном случае,} \end{cases}$$

$$l_t = \begin{cases} 1, & \text{если } y_t \text{ меньше всех предыдущих уровней,} \\ 0, & \text{в противном случае,} \end{cases}$$

t = 2, 3, ..., n.

Второй этап. По этому методу проверка гипотезы об отсутствии во временном ряду тренда осуществляется при помощи двух вспомогательных величин d и s.

$$d = \sum_{2}^{n} (k_t - l_t),$$

$$s = \sum_{2}^{n} (k_t + l_t).$$

Нетрудно заметить, что величина s, характеризующая изменение временного ряда, принимает значения от 0 (все уровни ряда равны между собой) до (n-1) (ряд монотонный).

Величина d характеризует изменение дисперсии уровней временного ряда и изменяется от -(n-1) (ряд монотонно убывает) до (n-1) (ряд монотонно возрастает).

Обе величины (s и d) асимптотически нормальны и имеют независимые распределения. Показатель d применяется для обнаружения тенденций в средней, s – для обнаружения тенденций изменения дисперсии.

Третий этап. Проверка гипотезы о том, можно ли считать случайными разности (d-0) и (s-μ), проводится с использованием t-критерия Стьюдента:

$$t_d = \frac{|d - 0|}{\sigma_2}, \quad \sigma_2 = \sqrt{2\sum_{t=2}^{n}\frac{1}{t}} \approx \sqrt{2\ln n - 0{,}8456},$$

$$t_s = \frac{|s - \mu|}{\sigma_1}, \quad \sigma_1 = \sqrt{2\sum_{t=2}^{n}\frac{1}{t} - 4\sum_{t=2}^{n}\frac{1}{t^2}} \approx \sqrt{2\ln n - 3{,}4253},$$

где μ – математическое ожидание величины s, определенной для ряда, в котором уровни расположены случайным образом;
σ_1 – среднеквадратическое отклонение для величины s;
σ_2 – среднеквадратическое отклонение для величины d.

Значения величин μ, σ_1, σ_2 табулированы:

n	μ	σ_1	σ_2	n	μ	σ_1	σ_2
10	3,858	1,288	1,964	60	7,360	2,201	2,713
15	4,636	1,521	2,153	65	7,519	2,236	2,742
20	5,191	1,677	2,279	70	7,666	2,268	2,769
25	5,632	1,791	2,373	75	7,803	2,297	2,793
30	5,990	1,882	2,447	80	7,931	2,324	2,816
35	6,294	1,956	2,509	85	8,051	2,349	2,837
40	6,557	2,019	2561	90	8,165	2,371	2,857
45	6,790	2,072	2,606	95	8,273	2,395	2,876
50	6,998	2,121	2,645	100	8,375	2,416	2,894
55	7,187	2,163	2,681				

Четвёртый этап. Расчетные значения t_d и t_s сравниваются с табличным значением t-критерия Стьюдента с заданным уровнем значимости t_α. Если расчетное значение меньше табличного, то гипотеза об отсутствии тренда в среднем и дисперсии принимается. Если t_s больше табличного значения t_α, а t_d меньше t_α, то для данного временного ряда имеется тренд в среднем, а тренда дисперсии уровней ряда нет.

Контрольные вопросы и задания

1. С помощью, каких методов, определяется наличие тенденции во временном ряду?
2. Перечислите основные этапы определения наличия тренда.
3. Для временных рядов (*a*) и (*b*):

(*a*)

1	2	3	4	5	6	7	8	9	10
7,9	8,3	7,5	6,9	7,2	6,5	5,8	4, 9	5,1	4,4

(*b*)

1	2	3	4	5	6	7	8	9	10
43	47	50	48	54	57	61	59	65	62

определить наличие тренда во временном ряду, используя:
а) метод проверки разности средних уровней;
б) метод Фостера-Стьюарта.

6. Трендовые модели экономических процессов на основе кривых роста

Для моделирования тенденции развития экономических процессов и явлений широко используются кривые роста. Аналитически кривые роста задаются в виде простых математических функций одной переменной (времени). Изменение рассматриваемого показателя связывают с течением времени, и определяется функция поведения показателя в предыдущих периодах, чтобы выявить возможное поведение изучаемого показателя в будущем. Найденная функция времени (кривая роста) позволяет получить теоретические значения уровней временного ряда, т.е. те уровни, которые наблюдались бы, если бы тенденция процесса полностью совпадала с кривой. Эта функция далее применяется для прогнозирования.

Решение задачи о выборе типа кривой предполагает знакомство с основными видами используемых кривых роста и их свойствами, главным образом с характером изменения приростов и некоторыми их преобразованиями.

В настоящее время насчитывается большое количество типов кривых роста, применяемых для моделирования экономических процессов. Эти кривые могут быть разделены на три класса в зависимости от того, какой тип динамики развития они хорошо описывают.

К первому классу относятся функции, используемые для описания процессов с монотонным характером тенденции развития и без предела роста. К этому классу функций относятся полиномиальные кривые и простая экспоненциальная кривая.

Ко второму классу относятся функции, используемые для описания процессов с пределом роста без точки перегиба. Такие процессы чаще всего демографические, хотя встречаются и в исследовании экономических процессов на промышленных предприятиях. Функции, относящиеся ко второму классу, называются кривыми с насыщением. К этому классу функций относится модифицированная экспонента.

К третьему классу относятся функции, используемые для описания процессов с пределом роста и имеющие точку перегиба, то есть функции насыщения с точками перегиба. Такие кривые называются S -образными кривыми. К этому классу функций относятся кривая Гомперца и логистическая кривая.

Из полиномиальных кривых в экономических исследованиях обычно применяются полиномы первого, второго и третьего порядка. Использовать для выделения тенденции полиномы высоких степеней нецелесообразно, поскольку полученные таким образом аппроксимирующие функции будут отражать случайные отклонения, что противоречит смыслу выделения тенденции.

Полином первой степени

$$y_t = a_0 + a_1 t$$

на графике изображается прямой линией и используется для описания процессов, развивающихся во времени равномерно. Если для полинома первой степени рассчитать первые приросты по формуле

$$u_t = y_t - y_{t-1},\ t = 2,3,\ldots,n,$$

то они будут постоянной величиной и равны a_1.

Полином второй степени

$$y_t = a_0 + a_1 t + a_2 t^2$$

применяется в тех случаях, когда процесс развивается равноускоренно. Первые конечные разности этой параболы (приросты первого порядка) линейно зависят от времени, вторые конечные разности постоянны:

$$u_t = y_t - y_{t-1} = a_1 + 2a_2 t - a_2 = (a_1 - a_2) + 2a_2 t,$$

$$u_t^{(2)} = u_t - u_{t-1} = 2a_2.$$

Полином третьей степени

$$y_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3$$

может дважды изменять знак прироста ординат.

Параметры полиномов невысоких степеней могут иметь конкретную интерпретацию в зависимости от содержания временного ряда. Например, их можно трактовать как оценки скорости роста (параметр a_1), ускорения роста (параметр a_2), изменения ускорения (параметр a_3), уровень ряда при $t = 0$ (параметр a_0).

Преимуществами полиномов низкой степени являются, как большая гладкость, так и простота толкования их записи. Но в случае, если степень полинома недостаточна, возникает смещение в оценке тренда. Недостатком полиномов высоких степеней является неустойчивость параметров по отношению к выборочным ошибкам.

Для экспоненциальных кривых характерна зависимость приростов от значений самой функции. Эти кривые хорошо описывают процессы, имеющие лавинообразный характер, когда прирост зависит от достигнутого уровня функции.

Простая экспонента имеет вид: $y_t = ab^t$

где *a* и *b* – положительные числа, при этом если *b* больше единицы, то функция возрастает с ростом времени *t,* если *b* меньше единицы – функция убывает. Параметр*a* характеризует начальные условия развития, а параметр *b* – постоянный темп роста. Прологарифмировав выражение для данной функции по любому основанию, получим:

$$\log y_t = \log a + t \log b,$$

то есть логарифм ординаты простой экспоненты линейно зависит от времени.

Простая экспонента, так же как и полиномы, используется для описания монотонно возрастающих или убывающих процессов без насыщения, их рост ничем не ограничен.

В случае, когда процесс характеризуется насыщением, его следует описывать при помощи кривой, имеющей асимптоту, отличную от нуля. К этому классу кривых относится **модифицированная экспонента**:

$$y_t = k + ab^t.$$

Логарифмы приростов ординат модифицированной экспоненты линейно зависят от времени, отношение последовательных приростов данной кривой для равноотстоящих во времени уровней равны между собой, т.е.

$$\frac{u_2}{u_1} = \frac{u_3}{u_2} = ... = \frac{u_t}{u_{t-1}} = b = const \, .$$

Кривую Гомперца и логистическую кривую используют для моделирования процессов, которые сначала растут медленно, затем ускоряются, а затем снова замедляют свой рост, стремясь к какому-либо пределу. В качестве примера можно привести процесс ввода некоторого объекта в промышленную эксплуатацию, процесс изменения спроса на товары, обладающие способностью достигать некоторого уровня насыщения, и др.

Кривая Гомперца имеет вид:

$$y_t = ka^{b^t} \, ,$$

где *a, b* – положительные параметры, причем b<1; k – асимптота функции. Логарифм данной функции представляет собой модифицированную экспоненту:

$$\log y_t = \log k + b^t \log a \, .$$

Логистическая кривая или кривая Перла–Рида имеет вид:

$$\frac{1}{y_t} = k + ab^t \, .$$

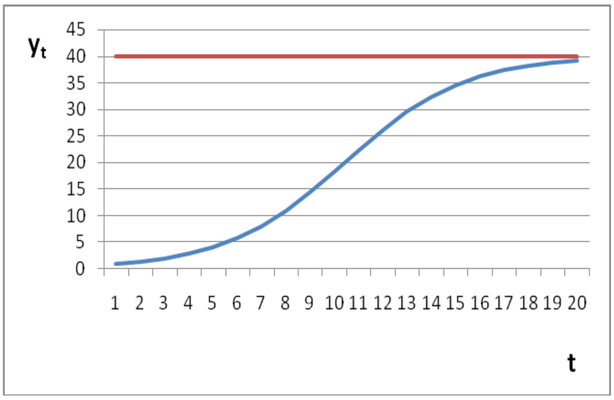
Наиболее распространенная форма записи уравнения логистической кривой следующая:

$$y_t = \frac{k}{1 + be^{-at}} \, ,$$

где a и b – параметры; k – верхний предел функции при бесконечном возрастании времени, т.е. при $t \to +\infty$ функция стремится к асимптоте. Логистическая кривая симметрична относительно точки перегиба, координатами которой являются:

$$t = \frac{\ln b}{a} \, ; \; y_t = \frac{k}{2} \, .$$

С помощью логистической функции хорошо описывается развитие новой отрасли и нового производства. Сначала технические методы производства не достаточно разработаны, издержки производства высоки и спрос на данный товар на рынке еще очень мал, поэтому производство развивается медленно. В дальнейшем, благодаря усовершенствованию технологии изготовления, переходу к массовому производству и увеличению емкости рынка производство растет быстрее. Затем наступает период насыщения рынка, рост производства все более замедляется и наконец, почти прекращается. Наступает стабилизация производства на определенном уровне. В результате конфигурация логистической кривой напоминает латинскую букву *S*. Пример логистической кривой:



Рассмотренные особенности и свойства кривых позволяют правильно выбрать кривую роста.

Приемлемым методом, применяемым для выбора вида полиномиальной кривой роста, является **метод конечных разностей (метод Тинтнера)**. Этот метод может быть использован для предварительного выбора полиномиальной кривой при выполнении следующих предположений:

1) уровни временного ряда состоят только из двух компонент: систематической составляющей (тренда) и случайной компоненты

$$y_t = u_t + \varepsilon_t ,$$

где u_t – систематическая составляющая (тренд), ε_t – случайная компонента;

2) тренд является достаточно гладким и поэтому может быть исключен либо уменьшен путем образования последовательных разностей;

3) случайная компонента ε_t подчиняется нормальному закону распределения с математическим ожиданием равным нулю и постоянной дисперсией.

Выбор кривой роста с помощью метода конечных разностей состоит в следующем:

1. Вычисляются разности k-го порядка с использованием рекурсивных отношений

$$u_t = y_t - y_{t-1};$$
$$u_t^{(2)} = u_t - u_{t-1};$$
$$\dots\dots\dots\dots\dots$$
$$u_t^{(k)} = u_t^{(k-1)} - u_{t-1}^{(k-1)};$$

2. Вычисляются дисперсии для исходного ряда и для каждого разностного ряда по следующим формулам:
для исходного ряда по обычной формуле

$$\sigma_0^2 = \frac{\sum_{t=1}^{n} y_t^2 - \frac{1}{n}\left(\sum_{t=1}^{n} y_t\right)^2}{n-1}$$

для k-го разностного ряда по формуле

$$\sigma_k^2 = \frac{\sum_{t=k+1}^{n}\left(u_t^k\right)^2}{(n-k)C_{2k}^k},$$

где C_{2k}^k – биномиальный коэффициент, значения которого приведены в таблице

k	0	1	2	3	4	5	6
C_{2k}^k	1	2	6	20	70	252	924

3. Вычисленное очередное значение σ_k^2 сравнивается с предыдущим *(k*-1*)* значением. Процесс заканчивается при выполнении условия:

$$\left|\sigma_k^2 - \sigma_{k-1}^2\right| \le \beta,$$

где β – заранее заданная малая величина. Таким образом, расчет ведется до тех пор, пока разности не будут примерно равны друг другу. В качестве аппроксимирующего полинома принимается полином, имеющий k-1 порядок.

Более универсальным является **метод характеристик прироста**, который основывается на сравнении характеристик изменения приростов исследуемого временного ряда с соответствующими характеристиками кривых роста. При рассмотрении свойств кривых роста были найдены различные преобразования приростов u_t, причем для каждой кривой было найдено такое преобразование u_t, которое характеризуется линейным уравнением относительно времени. Для эмпирических рядов можно определить аналогичные характеристики, взяв вместо прироста u_t значение среднего прироста $\bar{u}_t$. Если какая-либо из найденных по наблюдениям характеристик показывает развитие во времени близкое к линейному, то тенденция развития процесса может быть описана с помощью соответствующей кривой.

Процесс выбора кривых с использованием этого метода включает следующие этапы:

сглаживание ряда с помощью метода скользящей средней;

определение средних приростов;

вычисление производных характеристик прироста;

выбор наилучшей кривой.

Предварительное сглаживание методом простой скользящей средней осуществляется, например, для интервала сглаживания *m*=3. Сглаженные уровни рассчитываются по формулам:

$$\bar{y}_t = \frac{y_{t-1} + y_t + y_{t+1}}{3}; \quad \bar{y}_1 = \frac{5y_1 + 2y_2 - y_3}{6};$$

$$\bar{y}_n = \frac{-y_{n-2} + 2y_{n-1} + 5y_n}{6}, \quad t = 2,\dots,n-1.$$

Затем вычисляются первые средние приросты:

$$\bar{u}_t = \frac{\bar{y}_{t+1} - \bar{y}_{t-1}}{2}, \quad t = 2,3,\dots,n-1;$$

вторые средние приросты: $\bar{u}_t^{(2)} = \frac{\bar{u}_{t+1} - \bar{u}_{t-1}}{2}$,

а также ряд производных величин, связанных с вычисленными средними приростами и сглаженными уровнями ряда:

$$\frac{\bar{u}_t}{\bar{y}_t}; \ \log \bar{u}_t; \ \log \frac{\bar{u}_t}{\bar{y}_t}; \ \log \frac{\bar{u}_t}{\bar{y}_t^2}.$$

Полученные характеристики кривых роста сведены в таблице:

Показатель	Характер изменения показателя во времени	Вид кривой роста
Первый средний прирост $\bar{u}_t$	Примерно одинаковы	Полином первого порядка
То же	Изменяются линейно	Полином второго порядка
Второй средний прирост $\bar{u}_t^{(2)}$	Изменяются линейно	Полином третьего порядка

$\frac{\bar{u}_t}{\bar{y}_t}$	Примерно одинаковы	Простая экспонента
$\log \bar{u}_t$	Изменяются линейно	Модифицированная экспонента
$\log \frac{\bar{u}_t}{\bar{y}_t}$	Изменяются линейно	Кривая Гомперца
$\log \frac{\bar{u}_t}{\bar{y}_t^2}$	Изменяются линейно	Логистическая кривая

Приведенные в таблице характеристики изменения показателей временных рядов исследуются на линейность и постоянство уровня. Кривые роста, характеристики которых отвечают этим условиям, выбираются в качестве аппроксимирующих для исследуемого эмпирического ряда.

На практике при предварительном выборе отбирают обычно две-три кривые роста для дальнейшего исследования и построения трендовой модели данного временного ряда.

Оценки параметров большинства кривых роста находятся с помощью **метода наименьших квадратов** (МНК). Суть МНК состоит в определении таких коэффициентов (параметров модели), при которых сумма квадратов отклонений фактических уровней ряда от соответствующих расчетных значений по кривой роста была бы наименьшей:

$$\sum_{t=1}^{n} (y_t - \hat{y}_t)^2 \to \min ,$$

где y_t – фактическое значение уровня ряда; $\hat{y}_t$ – расчетное значение;
n – длина временного ряда.

При оценке параметров полиномиальных кривых в результате минимизации выражения получается система нормальных уравнений.

Если полином имеет невысокую степень, системы для оценивания параметров выглядят достаточно просто. Нормальные уравнения для оценивания полинома первого порядка $y_t = a_0 + a_1 t$ имеют вид:

$$\sum_{t=1}^{n} y_t = a_0 n + a_1 \sum_{t=1}^{n} t$$

$$\sum_{t=1}^{n} y_t t = a_0 \sum_{t=1}^{n} t + a_1 \sum_{t=1}^{n} t^2$$

Для оценивания параметров полинома второго порядка

$$y_t = a_0 + a_1 t + a_2 t^2$$

используют систему нормальных уравнений:

$$\sum_{t=1}^{n} y_t = a_0 n + a_1 \sum_{t=1}^{n} t + a_2 \sum_{t=1}^{n} t^2;$$

$$\sum_{t=1}^{n} y_t t = a_0 \sum_{t=1}^{n} t + a_1 \sum_{t=1}^{n} t^2 + a_2 \sum_{t=1}^{n} t^3;$$

$$\sum_{t=1}^{n} y_t t^2 = a_0 \sum_{t=1}^{n} t^2 + a_1 \sum_{t=1}^{n} t^3 + a_2 \sum_{t=1}^{n} t^4.$$

При оценивании параметров экспоненциальной кривой, уравнение после логарифмирования приводят к линейному виду: $log y_t = log a + t\ log b.$

Таким образом, составляют на основе метода наименьших квадратов систему нормальных уравнений, аналогичную системе для полинома первой степени. Решая эту систему, находят логарифмы параметров, а затем и сами параметры модели с помощью потенцирования, то есть получается экспоненциальная функция, аппроксимирующая временной ряд.

При определении параметров кривых роста, имеющих асимптоты (модифицированная экспонента, кривая Гомперца, логистическая кривая), различают два случая. Если значение асимптоты k известно заранее, то путем несложной модификации формулы и последующего логарифмирования определение параметров сводят к решению системы нормальных уравнений , неизвестными которой являются логарифмы параметров кривой. Если значение асимптоты заранее неизвестно, то для нахождения параметров указанных выше кривых роста используются приближенные методы: метод трех точек, метод трех сумм и др.

Таким образом, при моделировании экономической динамики, заданной временным рядом, путем сглаживания исходного ряда , определения наличия тренда, отбора одной или нескольких кривых роста и определения их параметров, в случае наличия тренда, получают одну или несколько трендовых моделей для исходного временного ряда. Встает вопрос, насколько эти модели близки к экономической реальности, отраженной во временном ряду, насколько обосновано применение этих моделей для анализа и прогнозирования изучаемого экономического явления.

Контрольные вопросы и задания

1. Из каких предположений исходят при экстраполировании на основе кривых роста?

2. Охарактеризуйте основные типы кривых роста, применяемых для моделирования экономических процессов.

3. Назовите методы предварительного выбора кривых роста.

4. Какие предположения должны выполняться при предварительном выборе полиномиальной кривой методом конечных разностей (методом Тинтнера)?

5. В чем суть метода конечных разностей (метода Тинтнера)?

6. Перечислите основные этапы выбора кривой на основе метода характеристик прироста.

7. Какой метод используется для оценки параметров кривых роста?

8. Для временного ряда:

1	2	3	4	5	6	7	8	9	10
43	47	50	48	54	57	61	59	65	62

а) постройте линейную модель, определив её параметры методом наименьших квадратов; *б*) осуществите выбор наилучшей кривой роста методом характеристик прироста.

9. В таблице приведены данные о ценах на сырье (руб. за единицу объёма)

Период	1	2	3	4	5	6	7	8	9	10	11	12
Цена, руб.	203	220	231	240	248	260	273	282	279	270	268	260

а) выберите степень полинома, аппроксимирующего данный временной ряд, методом конечных разностей; *б*) постройте выбранную полиномиальную модель.

7. Оценка адекватности и точности трендовых моделей

Вопрос о возможности применения модели может быть решен только после установления её адекватности, т.е. соответствия модели исследуемому процессу или объекту. Так как полного соответствия модели реальному процессу или объекту быть не может, **адекватность** в какой-то мере условное понятие. При моделировании имеется в виду адекватность не вообще, а по тем свойствам модели, которые считаются существенными для исследования.

При правильном отображении моделью систематических компонент изменение **остаточной компоненты** не будет связано с изменением времени, т.е. будет носить **случайный характер**. Поэтому принято считать, что модель адекватна изучаемому процессу, если остаточная компонента $\varepsilon_t = y_t - \hat{y}_t$ представляет собой случайную компоненту временного ряда. Это требование будет выполняться, если остаточная компонента удовлетворяет свойствам:

случайности колебаний уровней остаточной последовательности;

соответствию распределения остаточной компоненты нормальному закону распределения;

равенству математического ожидания остаточной компоненты нулю;

независимости значений уровней остаточной компоненты между собой.

Рассмотрим проверку перечисленных свойств подробнее.

Проверка случайности колебаний уровней остаточной последовательности. Проверку случайности колебаний уровней остаточной компоненты проводят, используя критерии серий, основанный на медиане или критерий поворотных точек (критерий пиков).

Критерий серий. Для исследования случайности отклонений от тренда мы располагаем набором разностей $\varepsilon_t = y_t - \bar{\hat{y}}_t$, (t = 1, 2, ..., n). Ряд из величин ошибки располагают в порядке возрастания их значений и находят медиану ε_m полученного вариационного ряда, т.е. срединное значение при нечетном n или среднюю арифметическую из двух срединных значений при n четном. Возвращаясь к исходной последовательности ε_t и сравнивая значения этой последовательности с ε_m, будем ставить знак «плюс», если значение ε_t превосходит медиану, и знак «минус», если оно меньше медианы; в случае равенства сравниваемых величин соответствующее значение ε_t опускается. Таким образом, получается последовательность, состоящая из плюсов и минусов, общее число которых не превосходит n. Последовательность подряд идущих плюсов или минусов называется серией. Для того чтобы последовательность ε_t была случайной выборкой, протяженность самой длинной серии не должна быть слишком большой, а общее число серий – слишком малым.

Обозначим протяженность самой длинной серии через $K_{\max}$, а общее число серий – через v. Выборка признается случайной, если выполняются следующие неравенства для 5%-ного уровня значимости:

$$v(n) > \left[\frac{1}{2}(n + 1 - 1{,}96\sqrt{n - 1})\right],$$

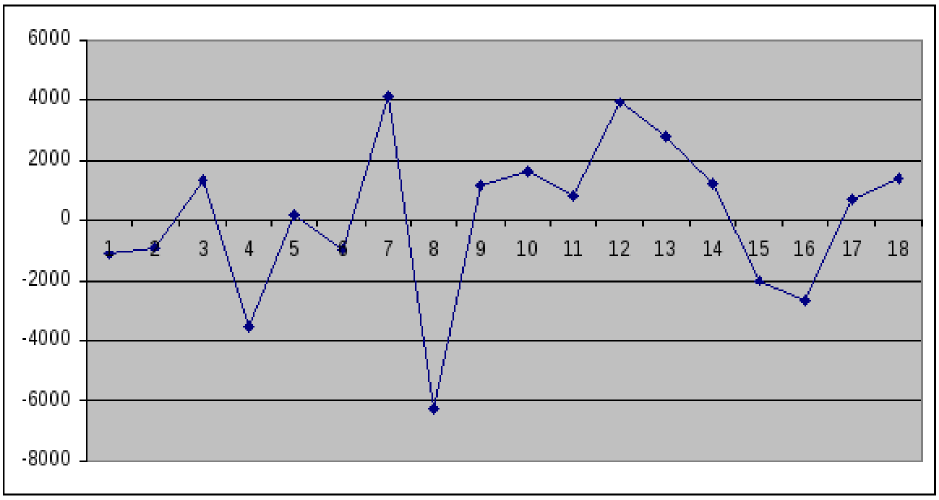
$$K_{max} < [3{,}3(lgn + 1)]\ ,$$

где квадратные скобки, как обычно, означают целую часть числа.

Если хотя бы одно из этих неравенств нарушается, то гипотеза о случайном характере отклонений уровней временного ряда от тренда отвергается и, следовательно, трендовая модель признается неадекватной.

Критерий поворотных точек (критерий пиков). Точка считается поворотной, если она одновременно больше или меньше двух рядом стоящих точек, т.е. $\varepsilon_{t-1} < \varepsilon_t > \varepsilon_{t+1}$ или $\varepsilon_{t-1} > \varepsilon_t < \varepsilon_{t+1}$. Наиболее просто определить поворотные точки, если ряд остатков изобразить графически.

Начальное значение также как и последнее значение нельзя считать поворотной точкой. На представленном далее графике ряда остатков, хорошо видно, что в ряду число поворотных точек равно 10.



Общее число поворотных точек для ряда остатков обозначим через p. В случайной выборке математическое ожидание числа точек поворота и дисперсия выражается формулами:

$$\bar{p} = \frac{2}{3}(n-2)\,,\ \sigma_p^2 = \frac{16n-29}{90}\,.$$

Ряд является случайным с 5% уровнем значимости, если выполняется неравенство:

$$p > \left[\bar{p} - 1{,}96\sigma_p^2\right],$$

где квадратные скобки – целая часть числа. Если это неравенство не выполняется, то ряд остатков нельзя считать случайным, т.е. он содержит регулярную компоненту, модель не является неадекватной.

Проверка соответствия распределения случайной компоненты нормальному закону распределения может быть произведена лишь приближенно с помощью исследования показателей асимметрии (γ_1) и эксцесса (γ_2), так как временные ряды, как правило, не очень велики. При нормальном распределении показатели асимметрии и эксцесса некоторой генеральной совокупности равны нулю. Мы предполагаем, что отклонения от тренда представляют собой выборку из генеральной совокупности, поэтому можно определить только выборочные характеристики асимметрии и эксцесса и их ошибки:

$$\gamma_1 = \frac{\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^3}{\sqrt{\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2\right)^3}}\ ;\ \sigma_{\gamma_1} = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}}\ ;$$

$$\gamma_2 = \frac{\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^4}{\sqrt{\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2\right)^2}}\ ;\ \sigma_{\gamma_2} = \sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}}\,.$$

В этих формулах γ_1 – выборочная характеристика асимметрии; γ_2 – выборочная характеристика эксцесса; σ_{γ_1} и σ_{γ_2} – соответствующие среднеквадратические ошибки.

Если одновременно выполняются следующие неравенства:

$$|\gamma_1| < 1{,}5\sigma_{\gamma_1}\ ;\ \left|\gamma_2 + \frac{6}{n+1}\right| < 1{,}5\sigma_{\gamma_2}\ ,$$

то гипотеза о нормальном характере распределения остаточной компоненты принимается.

Если выполняется хотя бы одно из неравенств

$$|\gamma_1| \geq 2\sigma_{\gamma_1}\ ;\ \left|\gamma_2 + \frac{6}{n+1}\right| \geq 2\sigma_{\gamma_2}\ ,$$

то гипотеза о нормальном характере распределения отвергается, модель признается неадекватной. Другие случаи требуют дополнительной проверки с помощью более сложных критериев.

Кроме рассмотренного метода известен ряд других методов проверки выполнения нормального закона распределения случайной величины. Это метод Вестергарда, *RS*-критерий и т.д. Рассмотрим наиболее простой из них, основанный на *RS*-критерии. Этот критерий численно равен отношению размаха вариации случайной величины R к стандартному отклонению S.

В нашем случае $R=\varepsilon_{\max}-\varepsilon_{\min}$, , а

$$S = S_{\hat{y}} = \sqrt{\frac{\sum \varepsilon_t^2}{n-1}}\ .$$

Далее вычисленное значение *RS*-критерия сравнивается с табличными (критическими) нижней и верхней границами данного отношения. Если это значение не попадает в интервал между критическими границами, то с заданным уровнем значимости гипотеза о нормальности распределения отвергается; в противном случае эта гипотеза принимается. Критические границы *RS*-критерия табулированы. Для иллюстрации приведем несколько пар значений критических границ *RS*-критерия для уровня значимости $\alpha=0{,}05$: при $n=10$ нижняя граница равна 2,67, а верхняя равна 3,685; при $n=20$ эти числа составляют соответственно 3,18 и 4 ,49; при $n=30$ они равны 3,47 и 4,89.

Проверка равенства математического ожидания остаточной компоненты нулю осуществляется с помощью *t*-критерия Стьюдента. Такая проверка осуществляется, если остаточная компонента распределена по нормальному закону. Расчетное значение *t*-критерия Стьюдента определяется по формуле:

$$t = \frac{\bar{\varepsilon} - 0}{S_\varepsilon}\sqrt{n}\ ,$$

где $\bar{\varepsilon}$ – среднее арифметическое значение уровней остаточной последовательности, S_ε – среднеквадратическое отклонение для последовательности остатков.

Если расчетное значение t-критерия меньше табличного значения t_α статистики Стьюдента с заданным уровнем значимости α и числом степеней свободы n-1, то гипотеза о равенстве нулю математического ожидания уровней ряда остатков принимается; в противном случае эта гипотеза отвергается и модель считается неадекватной.

Независимость значений уровней случайной компоненты. Если систематические компоненты подобраны неправильно, то последовательные значения остаточной компоненты могут коррелировать между собой, в этом случае говорят, что существует автокорреляция остатков. Существуют различные критерии проверки отсутствия существенной автокорреляции в остаточной последовательности. Наиболее распространенным является d-критерий Дарбина–Уотсона. При установлении наличия автокорреляции по критерию Дарбина–Уотсона вычисляется d-значение:

$$d = \frac{\sum_{t=2}^{n}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n}\varepsilon_t^2}$$

Заметим, что расчетное значение критерия Дарбина–Уотсона в интервале от 2 до 4 свидетельствует об отрицательной связи. В этом случае его надо преобразовать по формуле $d'=4-d$ и в дальнейшем использовать значение d'.

Расчетное значение критерия d (или d') сравнивается с верхним d_2 и нижним d_1 критическими значениями статистики Дарбина–Уотсона.

n	k=1		k=2		k=3	
	d_1	d_2	d_1	d_2	d_1	d_2
15	1,08	1,36	0,95	1,54	0,82	1,75
20	1,20	1,41	1,10	1,54	1,00	1,68
30	1,35	1,49	1,28	1,57	1,21	1,65

Здесь представлен фрагмент табличных значений для различного числа уровней ряда n и числа определяемых параметров модели k (уровень значимости 5%)

Если расчетное значение критерия d больше верхнего табличного значения d_2 , то гипотеза о независимости уровней остаточной последовательности, т.е. об отсутствии в ней автокорреляции, принимается. Если значение d меньше нижнего табличного значения d_1, то эта гипотеза отвергается и модель неадекватна. Если значение d находится между значениями d_1 и d_2, включая сами эти значения, то считается, что нет достаточных оснований сделать тот или иной вывод и необходимы дальнейшие исследования, например, по большему числу наблюдений.

Если **все четыре гипотезы** о свойствах остаточной компоненты подтверждаются, то исследуемая **модель адекватна,** и ее можно использовать для анализа и построения прогнозных оценок.

Если модель неадекватная необходимо тщательно проанализировать методику моделирования и, возможно, осуществить выбор и оценку другой математической модели.

Оценка точности модели. Как правило, о точности модели и прогноза судят по величине погрешности (ошибки). Ошибка прогноза это расхождение между фактическим и прогнозируемым значением исследуемого показателя. Использование данного подхода к оценке точности возможно только в том случае, когда период упреждения закончился, и исследователи имеют фактические значения на период упреждения или когда разрабатывается ретро прогноз.

Ретроспективное прогнозирование разрабатывается для некоторого момента времени в прошлом, для которого имеются фактические данные. В этом случае имеющаяся информация делится на две части. Первая часть, включающая более ранние данные, используется для подбора математической модели. По построенной математической модели дается прогноз на последующий оставшийся период времени. Прогнозные качества модели оцениваются по более поздним данным второй части ряда. Полученные ошибки прогноза в какой-то мере характеризуют точность подобранных моделей и могут использоваться при сопоставлении различных моделей прогнозирования. В то же время при использовании ошибки ретроспективного прогноза в качестве меры точности необходимо учитывать, что она получена при использовании только части имеющихся данных. При использовании полного объема имеющихся данных трансформируется вид подобранной модели, и изменяются значения критериев точности и качества.

Наличие данных о реализации прогнозов дает возможность оценить качество прогнозов величиной: $\gamma = p/(p+q)$,

где p – число прогнозов, подтвержденных фактическими данными (фактическая реализация охвачена интервальным прогнозом); q – число прогнозов, не подтвержденных фактическими данными. Использование коэффициентов γ для разных моделей имеет смысл в том случае, если доверительные вероятности прогнозов приняты одинаковыми. В том случае, если прогноз дается в виде точечной оценки, в качестве показателей точности прогноза могут использоваться такие статистические характеристики как средняя абсолютная и среднеквадратическая ошибка прогноза.

Г. Тейлом предложен в качестве меры качества прогноза коэффициент расхождения (или коэффициент несоответствия):

$$\upsilon = \frac{\sqrt{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}}{\sqrt{\sum_{t=1}^{n} y_t^2}},$$

где участвуют предсказанное и фактическое значения уровней ряда . Этот коэффициент равен нулю, когда предсказанное и фактическое значения совпадают (случай совершенного прогнозирования). И коэффициент равен единице, когда экстраполяция строится исходя из неизменности приростов. Если коэффициент больше единицы, то прогноз дает худшие результаты, чем прогноз методом простой экстраполяции.

В условиях отсутствия информации о результатах эксплуатации модели оценивается так называемая априорная или предполагаемая точность. Исследуя априорную точность модели, характеризуют только точность аппроксимации. Чаще всего в качестве показателей точности применяются следующие показатели: абсолютная ошибка $\varepsilon_t = y_t - \hat{y}_t$; средняя абсолютная ошибка $\bar{\varepsilon}$; средняя квадратическая ошибка

$$S_{\hat{y}} = \sqrt{\frac{\sum \varepsilon_t^2}{n-k}},$$

где n — количество уровней ряда y_t, k – число оцениваемых параметров модели; относительная ошибка

$$e_{отн} = \frac{y_t - \hat{y}_t}{y_t} \cdot 100;$$

средняя относительная ошибка

$$\bar{e}_{отн} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \cdot 100.$$

Последний показатель чаще других используется при сравнении точности прогнозов, осуществляемых по различным методикам. Обычно лучшим признается тот прогноз, который имеет меньшее значение этого показателя. Принято считать, что если значение средней относительной ошибки менее 3-5%, то точность высокая; если значение средней относительной ошибки не превышает 10%, то точность хорошая; от 10% до 15% точность удовлетворительная.

Применяются также коэффициент сходимости и коэффициент детерминации:

$$\phi^2 = \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y}_t)^2},$$

$$R^2 = 1 - \phi^2.$$

Чем меньше значение коэффициента сходимости, тем лучше точность модели; чем больше значение коэффициента детерминации, тем лучше точность модели.

Контрольные вопросы и задания

1. Когда модель считается адекватной экономическому процессу?
2. Какие свойства остаточной компоненты проверяются при оценке адекватности модели? Какие статистические критерии при этом используются?
3. Поясните суть ретроспективного прогнозирования.
4. Какие статистические характеристики используются для оценки точности модели?
5. Для линейной модели временного ряда:

1	2	3	4	5	6	7	8	9	10
43	47	50	48	54	57	61	59	65	62

а) оцените адекватность модели (принять t_α = 1,09 при доверительной вероятности 0,7; критические уровни критерия Дарбина–Уотсона d_1 = 1,08, d_2 = 1,36); *б*) осуществите оценку точности модели, используя показатели среднего квадратического отклонения и средней относительной ошибки аппроксимации.

8. Прогнозирование экономической динамики на основе трендовых моделей

Прогнозирование временных рядов состоит в формировании на основе математической модели точечного и интервального прогнозов исследуемого показателя и выдачи пользователю степени доверия к полученным результатам.

Прогнозирование на основе кривых роста основано на экстраполяции, т.е. на продлении в будущее тенденции, наблюдавшейся в прошлом. При этом предполагается, что во временном ряде присутствует тренд и характер развития показателя обладает свойствами инерционности, т.е. изменение сложившейся тенденции маловероятно.

Процедура прогнозирования сводится к выполнению следующих этапов:

1. Предварительный анализ данных.
2. Выбор одной или нескольких кривых, форма которых соответствует характеру изменения временного ряда.

3. Оценка параметров выбранных кривых.
4. Проверка адекватности выбранных кривых прогнозируемому процессу.
5. Оценка точности моделей.
6. Окончательный выбор кривой роста.
7. Расчет точечного и интервального прогноза.
8. Верификация прогноза.

В данной главе рассматриваются два заключительных этапа этой достаточно длительной процедуры.

Прогноз на основании трендовых моделей (кривых роста) содержит два элемента: точечный и интервальный прогнозы.

Точечный прогноз – это прогноз, которым называется единственное значение прогнозируемого показателя. Это значение определяется подстановкой в уравнение выбранной кривой роста величины времени t, соответствующей периоду упреждения: $t=n+1$; $t=n+2$ и т.д. Такой прогноз называется точечным, так как на графике его можно изобразить в виде точки.

Очевидно, что точное совпадение фактических данных в будущем и прогностических точечных оценок маловероятно. Поэтому точечный прогноз должен сопровождаться двусторонними границами, т.е. указанием интервала значений, в котором с достаточной долей уверенности можно ожидать появления прогнозируемой величины. Установление такого интервала называется **интервальным прогнозом**.

Интервальный прогноз на базе трендовых моделей осуществляется путем расчета доверительного интервала – такого интервала, в котором с определенной вероятностью можно ожидать появления фактического значения прогнозируемого экономического показателя. Расчет доверительных интервалов при прогнозировании с использованием кривых роста опирается на выводы и формулы теории регрессий. Перенесение выводов теории регрессий на временные экономические ряды не совсем правомерно, так как динамические ряды, как выше уже отмечалось, отличаются от статистических совокупностей. Поэтому к оцениванию доверительных интервалов для кривых роста следует подходить с известной долей осторожности.

Методы, разработанные для статистических совокупностей, позволяют определить доверительный интервал, зависящий от стандартной ошибки оценки прогнозируемого показателя, от времени упреждения прогноза, от количества уровней во временном ряду и от уровня значимости (ошибки) прогноза. Стандартная (среднеквадратическая) ошибка оценки прогнозируемого показателя $S_{\hat{y}}$ определяется по известной формуле:

$$S_{\hat{y}} = \sqrt{\frac{\sum \varepsilon_t^2}{n-k}} \, ,$$

где n – количество уровней ряда y_t, k – число параметров модели, абсолютная ошибка $\varepsilon_t = y_t - \hat{y}_t$, $\hat{y}_t$ – соответствующее значение уровня по модели.

В случае прямолинейного тренда для расчета доверительного интервала можно использовать аналогичную формулу для парной регрессии, таким образом доверительный интервал прогноза U_y в этом случае будет иметь вид

$$U_y = \hat{y}_{t+L} \pm t_\alpha S_{\hat{y}} \sqrt{1 + \frac{1}{n} + \frac{3(n + 2L - 1)^2}{n(n^2 - 1)}},$$

где L – период упреждения, $\hat{y}_{n+L}$ – точечный прогноз по модели на (n+L)-й момент времени, $S_{\hat{y}}$ рассчитана для числа параметров модели, равного двум, t_α – табличное значение критерия Стьюдента для уровня значимости α и для числа степеней свободы, равного n-2.

Значения величины $K = t_\alpha \sqrt{1 + \frac{1}{n} + \frac{3(n+2L-1)^2}{n(n^2-1)}}$ табулированы.

Далее представлена таблица значений K для уровня значимости α=0,2 :

Число уровней в ряду n	Период упреждения L					
	1	2	3	4	5	6
7	1,932	2,106	2,300	2,510	2,733	2,965
10	1,692	1,774	1,865	1,964	2,069	2,180
13	1,581	1,629	1,682	1,738	1,799	1,863
15	1,536	1,572	1,611	1,653	1,697	1,745

Иногда для расчета доверительных интервалов прогноза относительно линейного тренда применяют приведенную выше формулу в несколько преобразованном виде:

$$U_y = \hat{y}_{t+L} \pm t_\alpha S_{\hat{y}} \sqrt{1 + \frac{1}{n} + \frac{(t_L - \bar{t})^2}{\sum(t - \bar{t})^2}}.$$

Здесь t – порядковый номер уровня ряда (t=1,2 ,...,n); t_L=n+L – время, для которого делается прогноз; $\bar{t}$ – время, соответствующее середине периода наблюдений для исходного ряда, например, $\bar{t}$=(n+1)/2; суммирование ведется по всем наблюдениям.

Эту формулу можно упростить, если, как часто делается на практике, перенести начало отсчета времени на середину периода наблюдений ($\bar{t}$ = 0):

$$U_y = \hat{y}_{t+L} \pm t_\alpha S_{\hat{y}} \sqrt{1 + \frac{1}{n} + \frac{{t_L}^2}{\sum t^2}}.$$

Формула для расчета доверительных интервалов прогноза относительно тренда, имеющего вид полинома второго или третьего порядка, выглядит следующим образом:

$$U_y = \hat{y}_{t+L} \pm t_\alpha S_{\hat{y}} \sqrt{1 + \frac{1}{n} + \frac{{t_L}^2}{\sum t^2} + \frac{\sum t^4 - 2{t_L}^2 \sum t^2 + n{t_L}^4}{n\sum t^4 - (\sum t^2)^2}}.$$

Аналогично вычисляются доверительные интервалы для экспоненциальной кривой роста и кривых роста, имеющих асимптоту (модифицированная экспонента, кривая Гомперца и логистическая).

Не следует обольщаться технической простотой процедуры экстраполяции и пытаться заглянуть слишком далеко, это неизбежно приведет к грубым ошибкам. Оптимальная длина периода упреждения определяется отдельно для каждого экономического явления. Она, как правило, не превышает для рядов годовых наблюдений одной трети объема данных, а для квартальных и месячных рядов – двух лет.

При выравнивании временных рядов с использованием кривых роста приходится решать вопрос о том, какой длины должен быть ряд, выбираемый для прогнозирования. Очевидно, что если период ряда экономической динамики слишком короткий, можно не обнаружить тенденцию его развития. С другой стороны, очень длительный временной ряд может охватывать периоды с различными трендами и его описание с помощью одной кривой роста не даст положительных результатов. Поэтому рекомендуется поступать следующим образом. Если нет никаких соображений качественного порядка, следует брать возможно больший промежуток времени. Если развитие обнаруживает циклический характер, следует брать период от середины первого до середины последнего периода цикла. Если ряд охватывает периоды с разными трендами, лучше сократить ряд, отбросив наиболее ранние уровни, которые относятся к периоду с иной тенденцией развития.

При экстраполяционном прогнозировании экономической динамики с использованием трендовых моделей весьма важным является заключительный этап – **верификация** прогноза. Верификация прогнозов по трендовым моделям, сводится к сопоставлению расчетных результатов по модели с соответствующими данными действительности – массовыми фактами и закономерностями экономического развития. Верификация прогнозной модели представляет собой совокупность критериев, способов и процедур, позволяющих на основе многостороннего анализа оценивать качество получаемого прогноза. Однако чаще всего на этапе верификации в большей степени осуществляется оценка метода прогнозирования, с помощью которого был получен результат, чем оценка качества самого результата. Это связано с тем, что до сих пор не найдено эффективного подхода к оценке качества прогноза до его реализации.

Даже в тех случаях, когда прогноз не оправдался, нельзя категорически утверждать, что он был бесполезен, поскольку пользователь, если он хотя бы частично контролирует ход событий и может воздействовать на экономический процесс, может использовать прогнозную информацию желаемым для себя образом. Так, получив прогноз событий, определяющих нежелательное направление перспективного развития, пользователь может принять меры, чтобы прогноз не оправдался; такой прогноз называется **самодеструктивным**. Если прогноз предсказал ход событий, устраивающий пользователя, то он может использовать свои возможности для увеличения вероятности правильного прогноза; подобный прогноз называется **саморегулирующим**. Таким образом, показателем ценности прогноза является не только его достоверность, но и полезность для пользователей.

Контрольные вопросы и задания

1. Перечислите основные этапы прогнозирования экономической динамики на основе одномерных временных рядов с использованием трендовых моделей.
2. Опишите порядок получения точечного и интервального прогноза экономического показателя на основе трендовых моделей.
3. От каких факторов зависит ширина доверительного интервала прогноза?
4. Для моделей, являющихся результатом исследования временных рядов:

1	2	3	4	5	6	7	8	9	10
43	47	50	48	54	57	61	59	65	62

1	2	3	4	5	6	7	8	9	10	11	12
203	220	231	240	248	260	273	282	279	270	268	260

заданий 8 и 9 главы VI, построить точечный и интервальный прогнозы на два шага вперед (t_{α}=1,09). Результаты прогнозирования отобразите графически.

9. Адаптивные модели прогнозирования

При краткосрочном прогнозировании, а также при прогнозировании в ситуации изменения внешних условий, когда наиболее важными являются последние реализации исследуемого процесса, наиболее эффективными оказываются адаптивные методы, учитывающие неравноценность уровней временного ряда.

Адаптивные модели прогнозирования – это модели дисконтирования данных, способные быстро приспосабливать свою структуру и параметры к изменению условий. Инструментом прогноза в адаптивных моделях, как и в кривых роста, является математическая модель с единственным фактором «время».

При оценке параметров адаптивных моделей в отличие от рассматриваемых ранее моделей «кривых роста» наблюдения (уровням ряда) присваиваются различные веса в зависимости от того, насколько сильным признается их влияние на текущий уровень. Это позволяет учитывать изменения в тенденции, а также любые колебания, в которых прослеживается закономерность . Все адаптивные модели базируются на двух схемах: **скользящего среднего** (СС-модели) и **авторегрессии** (АР-модели).

Согласно схеме скользящего среднего, оценкой текущего уровня является взвешенное среднее всех предшествующих уровней, причем веса при наблюдениях убывают по мере удаления от последнего уровня, т.е. информационная ценность наблюдений признается тем большей, чем ближе они к концу интервала наблюдений. Такие модели хорошо отражают изменения, происходящие в тенденции, но в чистом виде не позволяют отражать колебания.

Реакция на ошибку прогноза и дисконтирование уровней временного ряда в моделях, базирующихся на схеме СС, определяется с помощью параметров сглаживания (адаптации), значения которых могут изменяться от нуля до единицы. Высокое значение этих параметров (свыше 0,5) означает придание большего веса последним уровням ряда, а низкое (менее 0,5) – предшествующим наблюдениям. Первый случай соответствует быстро изменяющимся динамичным процессам, второй – более стабильным.

В авторегрессионной схеме оценкой текущего уровня служит взвешенная сумма не всех, а нескольких предшествующих уровней, при этом весовые коэффициенты при наблюдениях не ранжированы. Информационная ценность наблюдений определяется не их близостью к моделируемому уровню, а теснотой связи между ними.

Общая схема построения адаптивных моделей может быть представлена следующим образом. По нескольким первым уровням ряда оцениваются значения параметров модели. По имеющейся модели строится прогноз на один шаг вперед, причем его отклонение от фактических уровней ряда расценивается как ошибка прогнозирования, которая учитывается в соответствии с принятой схемой корректировки модели. Далее по модели со скорректированными параметрами рассчитывается прогнозная оценка на следующий момент времени и т.д. Таким образом, модель постоянно «впитывает» новую информацию и к концу периода обучения отражает тенденцию развития процесса, существующую в данный момент.

В практике статистического прогнозирования наиболее часто используются две базовые СС-модели – Брауна и Хольта, первая из них является частным случаем второй. Эти модели представляют процесс развития как линейную тенденцию с постоянно изменяющимися параметрами.

Модель Брауна (модель экспоненциального сглаживания). Модель Брауна может отображать развитие не только в виде линейной тенденции, но также в виде случайного процесса, не имеющего тенденции, а также в виде изменяющейся параболической тенденции. Соответственно различают модели Брауна:

• нулевого порядка, описывающая процессы, не имеющие тенденции развития. Она имеет один параметр A_0 (оценка текущего уровня). Прогноз развития на k шагов вперед осуществляется согласно формуле $Y(t+k)=A_0$. Такая модель также называется «наивной» («будет, как было»);

• первого порядка ($Y(t+k)=A_0+A_1k$). Коэффициент A_0 – значение , близкое к последнему уровню, и представляет как бы закономерную составляющую этого уровня. Коэффициент A_1 определяет прирост, сформировавшийся в основном к концу периода наблюдений, но отражающий также (правда, в меньшей степени) скорость роста на более ранних этапах;

• второго порядка, отражающей развитие в виде параболической тенденции с изменяющимися «скоростью» и «ускорением». Она имеет три параметра (A_2 – оценка текущего прироста или «ускорение»). Прогноз осуществляется по формуле: $Y(t+k)=A_0+A_1\mathrm{k}+A_2k^2$.

Порядок модели обычно определяют либо априорно на основе визуального анализа графика процесса, либо методом проб и сравнения.

Рассмотрим **этапы построения линейной адаптивной модели Брауна**.

Этап 1. По первым пяти точкам временного ряда оцениваются начальные значения A_0 и A_1 параметров модели с помощью метода наименьших квадратов для линейной аппроксимации: $Y_p(t)=A_0+A_1t$, (t= 1, 2, 3, 4, 5).

Этап 2. С использованием параметров A_0 и A_1 по модели Брауна находим прогноз на один шаг(k=1): $Y_p(t,k)=A_0(t)+A_1(t)k$.

Этап 3. Расчетное значение $Y_\mathrm{p}(t,k)$ экономического показателя сравнивают с фактическим $Y(t)$ и вычисляется величина их расхождения (ошибки). При k=1 имеем: $e(t+1)=Y(t+1) - Y_p(t,1)$.

Этап 4. В соответствии с этой величиной корректируются параметры модели. В модели Брауна модификация осуществляется следующим образом:

$$A_0(t) = A_0(t\text{-}1) + A_1(t\text{-}1) + (1\text{-}\beta^2)\ e(t),$$

$$A_1(t) = A_1(t\text{-}1) + (1\text{-}\beta)^2\ e(t),$$

где β – **коэффициент дисконтирования** данных, изменяющийся в пределах от 0 до 1 ($\alpha + \beta$ =1), характеризующий обесценение данных за единицу времени и отражающий степень доверия более поздним наблюдениям. Оптимальное значение β находится итеративным путем, т.е. многократным построением модели при разных р и выбором наилучшей, или по формуле: $\beta = (n\text{-}3)/(n\text{-}1)$, где n – длина временного ряда, α – **параметр сглаживания**, $e(t)$ – ошибка прогнозирования уровня $Y(t)$, вычисленная в момент времени (t-1) на один шаг вперёд.

Этап 5. По модели со скорректированными параметрами рассчитывается прогноз на следующий момент времени. Возврат на этап 3, если $t<n$. Если $t=n$, то построенную модель, проверив на адекватность, можно использовать для прогнозирования на будущее.

Этап 6. Интервальный прогноз строится как для линейной модели кривой роста.

В модели Хольта участвуют **два параметра сглаживания**: α_1 *и* α_2. Корректировка параметров в этой модели осуществляется по формулам:

$$A_0(t) = A_0(t\text{-}1) + A_1(t\text{-}1) + \alpha_1 e(t),$$

$$A_1(t) = A_1(t\text{-}1) + \alpha_1\alpha_2 e(t),\ \text{где } 0< \alpha_1 , \alpha_2 <1.$$

Параметры α_1 *и* α_2 называют также **параметрами адаптации**.

В моделях Брауна и Хольта параметры сглаживания характеризуют степень адаптации модели к изменению ряда наблюдений. Они определяют скорость реакции модели на изменения, происходящие в развитии. Чем они больше, тем быстрее реагирует модель на изменения. Обычно для устойчивых рядов их величина большая, а для неустойчивых – маленькая. В различных методах прогнозирования используется различный подход к их определению. Их можно взять фиксированными, а наилучшее значение определить методом подбора, чтобы ошибка прогноза на один шаг вперед была наименьшей. При использовании компьютера это не представляет особого труда.

Альтернативу этому подходу составляет динамическое изменение параметров сглаживания. В методах эволюции и симплекс-планирования параметры адаптации постоянно меняются на каждом шаге. Для каждого параметра сглаживания формируется несколько значений.

Контрольные вопросы и задания

1. Поясните суть адаптивных методов прогнозирования.
2. Перечислите этапы построения адаптивной модели Брауна.
3. Чем отличается модель Хольта от модели Брауна?
4. Как влияют параметры сглаживания в моделях Хольта и Брауна на скорость адаптации моделей к изменениям в прогнозируемом процессе?
5. Для временного ряда

1	2	3	4	5	6	7	8	9	10
43	47	50	48	54	57	61	59	65	62

постройте модель Брауна с параметром сглаживания α=0,4 и α=0,7.

6. Оцените адекватность моделей, построенных в задании 5.
7. Постройте точечный и интервальный прогнозы на два шага вперед (t_α=1,09) по моделям, построенным в задании 5. Результаты прогнозирования отобразите графически.

10. Исследование сезонных колебаний в экономических процессах

При анализе временных рядов важным является выявление периодических колебаний. Если эти колебания происходят благодаря природно-климатическим факторам, связанными со сменой времен года и имеют выраженную годовую периодичность, то такие колебания называют сезонными. Сезонные колебания наиболее отчетливо проявляются в сельском хозяйстве, в отраслях непосредственно занимающихся переработкой сельхозпродукции, в транспорте, в строительстве, в бытовом обслуживании, в торговле и потреблении. Сезонные изменения социально-экономических процессов и явлений определяются не только климатическими факторами, но и социальными, экономическими, юридическими. Например, увеличение уровня безработных зимой, увеличение средней заработной платы и среднедушевого дохода в конце года, периодические денежные потоки налоговых платежей, отчислений в различные фонды, осуществления платежей за услуги.

Для определения наличия сезонности в исследуемом процессе часто бывает достаточно экономического анализа и графического отображения процесса за два-три года. Первое представление о возможном характере процесса и наличии в нем сезонных колебаний дает графическое представление временного ряда. Визуальный анализ графиков временного ряда позволяет определить наличие тренда и его характер, наличие сезонных и циклических компонент. Выявление сезонных колебаний удобно производить по графику временного ряда, построенного методом наложения. В этом случае по оси абсцисс откладывается интервал времени предполагаемого периода колебаний, например, год, с разбивкой по месяцам. А по оси ординат - значения уровней ряда за несколько лет. Если во временном ряду имеются периодические изменения, то на графике наблюдаются пики или впадины в определенный период времени.

При графическом изображении процесса сезонность часто бывает выражена настолько ярко, что нет необходимости ее доказывать численным способом. Но могут возникнуть какие ситуации, когда нет твердой уверенности, что колебания обусловлены сезонным фактором, а не каким-то иным случайным внешним воздействием. В этом случае необходимо использовать специальные статистические критерии.

Выявить наличие сезонных колебаний можно, проверив на случайность остаточный ряд после выделения тренда. Случайная компонента должна иметь математическое ожидание равное нулю, постоянную дисперсию и нулевую автокорреляцию между соседними уровнями ряда. Если эти условия не выполняются, то можно сделать предположение о наличии в остаточном ряду сезонной компоненты. Если процесс подвержен периодическим колебаниям, имеющим определенный и постоянный период, равный годовому промежутку, то мы имеем дело с так называемым тренд-сезонным временным рядом.

Почти всюду, где не оговорено специально, будем рассматривать тренд-сезонный временной ряд Y_t, $t = 1, 2, ..., n$, порождаемый аддитивным случайным процессом: $Y_t=U_t+V_t+\varepsilon_t$.

Сезонная компонента V_t имеет период T_0: $V_{t+T_0}= V_t$ (T_0=12 для ряда месячных данных; T_0=4 – для ряда квартальных данных). Кроме того T_0 нацело делит n, т.е. $n = m\times T_0$, m – целое число. Очевидно, если T_0 – число месяцев или кварталов в году, то m – число лет, представленных во временном ряду.

Часто исходные данные тренд-сезонного временного ряда представляются в виде матрицы размера $m\times T_0$. В этом случае выражение ряда перепишется с учетом введения двойной индексации:

$$Y_{ij} = U_{ij}+V_{ij}+\varepsilon_{ij}, \quad i = 1,...,m, \quad j = 1,...,T_0.$$

Соотношения, устанавливающие связь между индексами t и (i,j):

$i = \left[\frac{t}{T_0}\right] + 1$, $j = t - (i-1)T_0$, где [ ] – целая часть числа.

При исследовании сезонной волны V_t чаще всего предполагается, что она не изменяется год от года. На самом же деле такое предположение далеко от действительности, по крайней мере, для большинства экономических процессов. Для сезонной волны характерно изменение со временем, как ее размаха, так и формы. В результате возникает необходимость в анализе и предсказании изменений сезонной волны.

Задачи, которые возникают при исследовании сезонных временных рядов:

1) определение наличия во временном ряду тренда и определение степени его гладкости;

2) выявление наличия во временном ряду сезонных колебаний;

3) фильтрация компонент ряда;

4) анализ динамики сезонной волны;

5) исследование факторов, определяющих сезонные колебания;

6) прогнозирование тренд-сезонных процессов.

Под степенью гладкости тренда понимают минимальную степень полинома, адекватно сглаживающего компоненту U_t. Этот пункт используется в некоторых итерационных алгоритмах фильтрации при выделении из временного ряда его компонент U_t , V_t , ε_t.

Выявление наличия во временном ряду сезонных колебаний сводится к проверке на случайность остаточного ряда: $l_t = Y_t - U_t$.

Под фильтрацией компонент ряда понимается выделение из ряда его составляющих U_t , V_t , ε_t.

Анализ динамики, или эволюции, сезонной волны может рассматриваться как процесс решения трех взаимосвязанных задач:

1. Анализ динамики амплитуды сезонной волны в каждом месяце (квартале, неделе).

2. Анализ динамики точек экстремума сезонной волны.

3. Исследование изменений формы волны.

В настоящее время развиваются три основных направления фильтрации компонент временного ряда вида: регрессионные, спектральные и итерационные. Далее рассмотрим более подробно итерационные.

Итерационные методы фильтрации составляющих временного ряда появились в свое время как результат признания невозможности выделения компонент ряда прямыми методами. Основная идея итерационных процедур заключается в многократном применении скользящей средней:

$$Y_t^{'}=\frac{\frac{Y_{t-\frac{T_0}{2}}}{2}+Y_{t-\frac{T_0}{2}+1}+\ldots+Y_t+\ldots+Y_{t+\frac{T_0}{2}-1}+\frac{Y_{t+\frac{T_0}{2}}}{2}}{T_0}$$

и одновременной оценке сезонной компоненты в каждом цикле.

При этом переход от одного шага итерации к другому в разных методах может сопровождаться изменением параметров сглаживания, то есть производиться методом взвешенной скользящей. Итерационные методы отличает простота и удовлетворительная чистота фильтрации компонент, но их недостатком является дополнительная работа с $T_0/2$ членов ряда в начале и в его конце.

Далее рассмотрим **итерационный метод Четверикова**. **1.** Здесь эмпирический ряд Y_t выравнивается с помощью выше указанной скользящей средней, выпадающие при этом первые и последние уровни либо восстанавливаются экстраполированием выровненного ряда, либо в дальнейшем остаются в стороне.

Получается предварительная оценка тренда $Y_t^{'} = U_t^{'}$ и отклонения исходного временного ряда от выровненного ряда $l_t = Y_t - Y_t^{'}$, t=1, 2, ..., n или $l_{ij} = Y_{ij} - Y_{ij}^{'}$, i=1, ..., m, j=1, ..., T_0.

2. Для каждого года i вычисляется среднеквадратическое отклонение, на которое делятся затем отдельные месячные (квартальные) отклонения соответствующего года:

$$\tilde{l}_{ij} = \frac{l_{ij}}{\sigma_i},$$

$$\sigma_i = \sqrt{\frac{\sum_{j=1}^{T_0} l_{ij}^2 - \left(\sum_{j=1}^{T_0} l_{ij}\right)^2 / T_0}{T_0 - 1}}.$$

3. Из нормированных таким образом отклонений получается предварительная сезонная волна:

$$V_j^1 = \frac{\sum_{i=1}^{m} \tilde{l}_{ij}}{m}.$$

4. Средняя предварительная сезонная волна умножается на среднеквадратическое отклонение каждого года и вычитается из исходного временного ряда: $U_{ij}^1 = Y_{ij} - V_j^1 \sigma_i$.

5. Полученный ряд U_{ij}^1 вновь сглаживается скользящей средней. Для месячных данных по пяти или семи точкам. В результате получается новая оценка тренда $U_t^{(2)}$.

6. Отклонения эмпирического ряда Y_t от тренда $U_t^{(2)}$, обозначаемые $l_t^{(2)} = Y_t - U_t^{(2)}$, вновь подвергается обработке по пп. 2 и 3 для выявления окончательной средней сезонной волны.

7. Исключение окончательной сезонной волны производится после умножения средней сезонной волны на коэффициент напряжённости сезонной волны k_i :

$$k_i = \frac{\sum_{j=1}^{T_0} l_{ij}^{(2)} \varepsilon_{ij}}{\sum_{j=1}^{T_0} \varepsilon_{ij}^2},$$

где ε_{ij} – случайная компонента, $\varepsilon_{ij} = l_{ij}^{(2)} - V_j^{(2)}$.

Коэффициенты напряжённости логично вычислять для каждого года кроме первого и последнего, поскольку при неоднократном сглаживании в них остаётся не достаточное количество уровней, что влечёт искажение при расчётах средних характеристик ряда.

8. Окончательные значения сезонной волны: $V_{ij} = V_j^{(2)} \cdot k_i$.

В итерационном **методе Шискина-Эйзенпресса** на втором и последующих итерационных этапах используются более сложные пятнадцати- и двадцатиодноточечные скользящие Спенсера. Однако при применении вычислительной техники их использование особых трудностей не представляет.

Скользящая средняя с симметрично-равными весами, как в методе Четверикова, позволяет успешно выделить линейный тренд. Если тренд нелинеен, то такое сглаживание даёт искажённые его значения. Скользящая средняя Спенсера позволяет получать оценки тренда, выраженного полиномами до третьего порядка включительно.

Перейдём к рассмотрению статистических методов оценки уровня сезонности. Мультипликативная модель сезонного временного ряда имеет вид:

$$Y_{ij} = U_{ij} I_j + \varepsilon_{ij},$$

где I_j - постоянная пропорциональности для j-го месяца, не меняющаяся от года к году, называемая **индексом сезонности**.

Приближённые оценки индексов сезонности могут быть получены следующим образом:

$$I_j = \frac{\sum_{i=1}^{m} I_{ij}}{m}, \quad I_{ij} = \frac{Y_{ij}}{\overline{Y_i}}, \quad \overline{Y_i} = \frac{\sum_{j=1}^{T_0} Y_{ij}}{T_0}.$$

Индексы сезонности характеризуют степень отклонения уровня сезонного временного ряда от тренда. Чаще всего рассматривают не ряд относительных показателей, а ряд процентов:

$$I_j = \frac{\sum_{i=1}^{m} I_{ij}}{m} \cdot 100\%.$$

Так как производится усреднение индексов за анализируемый период, индексы сезонности дают представление о средней сезонной волне.

Графическое изображение индексов сезонности – это изображение средней сезонной волны, которая достаточно адекватно воспроизводит фактическое изменение процесса в течение года.

Контрольные вопросы и задания

1. Дайте определение сезонных колебаний.

2. Какие факторы вызывают сезонные изменения в экономических процессах?

3. В чем отличие циклической компоненты временного ряда от сезонной?

4. Как графически определить наличие сезонных колебаний во временном ряде?

5. Какие показатели используются для измерения сезонности?

6. Что такое сезонная волна?

7. Какими моделями возможно представление сезонной компоненты в тренд-сезонных экономических процессах?

8. В таблице приведены данные по выручке от реализации лако-красочной продукции, млн. руб.

Месяц	1 год	2 год	3 год
Январь	7 254	9 783	8 367
Февраль	7 373	8 943	7 880
Март	7 468	7 855	8 633
Апрель	8 157	7 934	9 748
Май	8 795	8 467	12 454
Июнь	9 589	9 796	14 875
Июль	10 995	11 892	16 893
Август	12 790	13 934	19 356
Сентябрь	13 562	14 563	20 767
Октябрь	13 457	14 985	17 986
Ноябрь	12 976	13 478	15 733
Декабрь	11 157	9 623	11 124

а) постройте график временного ряда, методом наложения сделайте вывод о наличии сезонных колебаний, выделите сезонную волну методом Четверикова;

б) рассчитайте индексы сезонности, постройте график сезонной волны и сделайте выводы.

9. Используя данные задания 8, составьте план продаж лако-красочной продукции на четвёртый год, если предполагается а) снижение спроса на 15%, б) повышение спроса на 7%.

Библиографический список

1. Афанасьев В.Н. Анализ временных рядов и прогнозирование: Учебник. / В.Н. Афанасьев, М.М. Юзбашев – М.: Финансы и статистика, 2001. – 228 с.

2. Лукашин, Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов: учеб. пособие / Ю.П. Лукашин. – М.: Финансы и статистика, 2003.

3. Орлова, И.В. Экономико-математические методы и модели: компьютерное моделирование / И.В. Орлова, В.А. Половникова. – М.: Вузовский учебник, 2007. – 365 с.

4. Садовникова, Н.А. Анализ временных рядов и прогнозирование. Вып. 2: Учебное пособие / Н.А. Садовникова, Р.А. Шмойлова. – М.: Московский государственный университет экономики, статистики и информатики, 2004. – 200 с.

5. Тихонов Э.Е. Методы прогнозирования в условиях рынка: Учебное пособие / Э.Е. Тихонов. – Невинномысск, 2006. – 221 с.

6. Федосеев В.В. Экономико-математические методы и прикладные модели: Учебное пособие. / В.В. Федосеев, А.Н. Гармаш, Д.М. Дайитбегов и др.; под ред. В.В. Федосеева. – М.: ЮНИТИ, 1999. – 391 с.

7. Шелобаев С.И. Анализ и прогнозирование финансовых процессов: Учебное пособие. / С.И. Шелобаев, И.С. Шелобаева. – Тула: Левша, 2009. – 265с.

Издательская лицензия JIP 020261 от 14.01.1997
Подписано в печать 05.03.2020
Формат 60x84 1/16. Бумага офсетная.
Усл.-печ. л. 3,49. Тираж 100. Заказ 89.

Издательство Алтайского государственного университета

Типография Алтайского государственного университета
656099 Барнаул, ул. Димитрова, 66