**Моделирование временных рядов**

1 Временной ряд и его компоненты3

2 Стационарные временные ряды и автокорреляционная функция5

3 Проверка наличия и нахождение тренда9

4 Моделирование сезонной компоненты15

4.1 Проверка наличия сезонной компоненты15

4.2 Выделение сезонной компоненты18

5 Экспоненциальное сглаживание и адаптивный прогноз26

6 ARIMA модели временных рядов31

**1 Временной ряд и его компоненты**

Один из важных разделов эконометрики посвящен **временным рядам.** **Временной ряд (динамический ряд, или ряд динамики)** – это последовательность наблюдений некоторого признака (случайной величины) **У** в последовательные моменты времени. Отдельные наблюдения называются **уровнями ряда**, которые будем обозначать  **(t = 1,2,..., n)**, где **n** – число уровней.

Показатели значений временного ряда образуют серию на­блюдений, проведенных через равные промежутки времени (день, неделя, месяц, квартал и т.д.). Например, временными рядами бу­дут серия ежедневных наблюдений в течение некоторого периода за ценами товара при закрытии торгов на бирже, дневные объемы выпуска товара, месячные показатели инфляции или индекса по­требительских цен, ежеквартальные оценки валового националь­ного продукта или средних зарплат, ежегодные данные об объеме, выручке и прибыли компании. По специфике своих значений по­казатели временного ряда, как правило, отличаются от простран­ственных данных. Особенностью данных временного ряда явля­ется наличие корреляции между ними.

**Основное положение, на котором базируется использование временных рядов для прогнозирования, гласит:** факторы, влияю­щие на отклик изучаемой системы, действовали некоторым обра­зом в прошлом и настоящем, и ожидается, что они будут действо­вать схожим образом и в не слишком далеком будущем.

**Основная цель анализа временных рядов** – оценка и вы­членение этих влияющих факторов с целью прогноза дальнейшего поведения системы и выработки рациональных управленческих решений.

Временной ряд рассматривается как **случайный процесс**, т.е. последовательность случайных величин

**{Y1, Y2,...,} = {}, t = 1,…, n**  **(1.1)**

зависящих от показателя времени t (через равные промежутки).

Конкретные показатели ряда его уровней,

**{} (1.2)**

рассматриваются как реализация **случайного процесса.**

В отличие от показателей пространственных данных, когда соответствующие случайные величины предполагаются одинаково распределенными и независимыми в совокупности, для временного ряда случайные величины (1.1) не предполагаются одинаково распределенными и статистически независимыми.

**Спецификация модели временного ряда обычно включает следующие составляющие:**

**Тренд (Т)**– плавно меняющаяся с течением времени составляющая, описывающая влияние долговременных факторов, эффект которых сказывается постепенно (например, рост населения, рост потребления, изменения структуры потребления, экономическое развитие).

**Сезонная компонента (s)** описывает регулярно изменяющиеся наблюдения. Она определяет короткопериодические колебания, связанные с изменениями внутригодовой активности, и повторяющиеся через более или менее продолжительные моменты времени (недели, месяцы, кварталы, сезоны и т.д.).

**Циклическая составляющая ()** описывает длительные периоды относительного подъема и спада. Она состоит из циклов, которые меняются по амплитуде и протяженности. На циклическую составляющую оказывают влияние такие факторы, как рост и исто­щение ресурсов, продолжительные периоды погодных условий, циклы солнечной активности, изменения в финансовой и налоговой политике и т.д. Факторы, влияющие на циклическую состав­ляющую, трудно идентифицируются формальными методами. В связи с такими особенностями мы исключаем циклические составляющие из рассмотрения.

**Случайная компонента** **(е)** вызывает отклонения от хода отклика, определяемого трендовой, циклической и сезонной составляющими. В терминах статистики она считается случайным остатком, на нее накладываются соответствующие ограничения.

Если временной ряд представляется в виде суммы компонент

**= + + , t = 1,…, n,**

то модель ряда называется **аддитивной.**

Представление ряда в виде произведения компонент

**= × × , t = 1,…, n,**

носит название **мультипликативной** модели.

Модель **смешанного** типа имеет следующий вид:

**= × + , t = 1,…, n.**

Задачи анализа и прогнозирования временных рядов начина­ются, как правило, с построения графика. На этом этапе можно исследовать компонентный состав ряда и провести выбор модели.

**2 Стационарные временные ряды и автокорреляционная функция**

**Стационарные временные ряды** – ряды, у которых вероятно­стные свойства не изменяются с течением времени. На интуитив­ном уровне можно дать следующее определение стационарного временного ряда: это такой ряд, который имеет **постоянную** **сред­нюю**, а значения ряда колеблются вокруг этой средней с некото­рой **постоянной дисперсией**. Хотя на практике временные ряды часто не являются стацио­нарными, для их исследования используются стационарные вре­менные ряды.

Временной ряд **{} t = 1,…, n** – называется **стационарным** (в широком смысле), если его среднее значение и дисперсия посто­янны (не зависят от времени **t**)*,* а его ковариации между значени­ями и с лагом **ῖ** зависят только от величины лага **ῖ** и не зависят от времени:

**М () = m,**

**D() = М(-m)2 = ,**

**= М [(-m)(** - **m)] = ρ(ῖ). (2.1)**

Примером **стационарного временного ряда** является белый шум ряд случайных остатков классической линейной модели регрессии, удовлетворяющий условиям Гаусса-Маркова. Для та­кого ряда среднее и ковариации равны нулю, а дисперсия посто­янна.

Классическим примером **нестационарного временного ряда** яв­ляется **случайное блуждание**, модель которого представляет­ся в виде:

**= + ,**

где – белый шум с нулевым средним и постоянной дисперсией.

Для стационарности временного ряда необходимо (но недостаточно) отсутствие в его структуре тренда и сезонной составляющей, а также условие гомоскедастичности случайного остатка.

Стационарный временной ряд представляется в виде:

**= m + ,**

где **m** – ожидаемое значение (математическое ожидание) уровней ряда;

– случайный остаток с нулевым средним, постоянной дисперсией и ковариацией, зависящей от лага.

Термин **автоковариация** для величины **γ(ῖ)** применяется в связи с тем, что характеризует статистическую зависимость между уровнями одного и того же ряда, отстоящими на ῖ лагов. На практике для стационарных временных рядов вместо коэффициента автоковариации чаще используется безразмерная величина – **коэффициент автокорреляции**:

= **cov {, } = γ(ῖ)**

**γ(0) (2.2)**

Статистическая значимость коэффициентов автокорреляции проверяется на основе (Q-статистики Бокса-Пиерса):

где **n –** объем выборки; **m –** длина лага. Проверяется совместная гипотеза о том, что все коэффициенты автокорреляции временно­го ряда равны нулю.

Для больших выборок Q-статистика имеет **ᵪ** **–** квадрат распре­деление с **m** степенями свободы. Если расчетное значение Q-статистики превосходит критическое значение **ᵪ** **–** квадрат заданного уровня значимости, то принимается гипотеза о том, что хотя бы одно отлично от нуля.

Другая Q-статистика Льюинга-Бокса (которая более предпоч­тительна при малых выборках) рассчитывается по формуле:

Изменение коэффициента автокорреляции в зависимости от временного лага ῖ называется **автокорреляционной функцией (АКФ).** Оценка этой функции по реализации временного ряда на­зывается **выборочной автокорреляционной функцией (ВАКФ).**

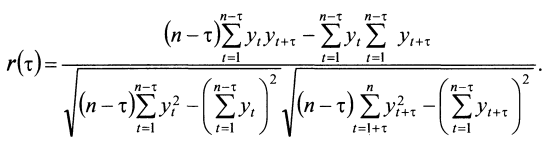
Ее значение вычисляется по формуле:

**, (2.3)**

где **= ) × ( - ) , = ,**

= = **) 2 , y =**  **.**

**Статистической оценкой ρ(ῖ) является выборочный коэффициент автокорреляции г(ῖ), определяемый по формуле:**

** (2.4)**

При расчете **г(ῖ)** следует помнить, что с увеличением **ῖ** число **n** - **ῖ** пар наблюдений уменьшается, поэтому лаг **ῖ** должен быть таким, чтобы число **n** - **ῖ** было достаточным для определения **г(ῖ)**. Обычно ориентируются на соотношение **ῖ ≤** .

Используя матрицу **R** парных коэффициентов автокорреля­ции временного ряда с различными лагами, можно построить **выборочную частную автокорреляционную функцию (ВЧАКФ) и ее график.** В дальнейшем под **стационарностью временного ряда** будем понимать его слабую стационарность. Проверка стационарности временного ряда проводится графическим анализом, **тестами корреллограммы** и **тестом единичного корня.**

Примеры графиков АКФ и ЧАКФ стационарного временного ряда, приведены на рисунке 2.1. На рисунке 2.2 приведены графики АКФ и ЧАКФ для процесса случайного блуждания, который не являет­ся стационарным.

Популярным тестом проверки стационарности является **тест единичного корня.** Исходной моделью для этого теста служит сто­хастический процесс единичного корня:

**= ρ × + , -1 ≤ ρ ≤ 1 , (2.5)**

где – белый шум. При **ρ = 1** эта модель является случайным блужданием (нестационарным случайным процессом). Идея тес­та единичного корня основана на построении регрессии по значениям и проверке статистической значимости условия **ρ = 1**. Если это условие выполнено, то – нестационарный.

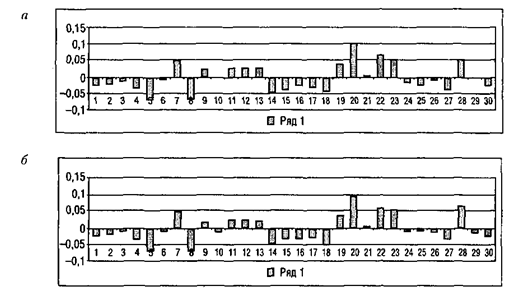
Формулу (2.5) можно переписать в виде:

Δ **= δ × + ,** **(2.6)**

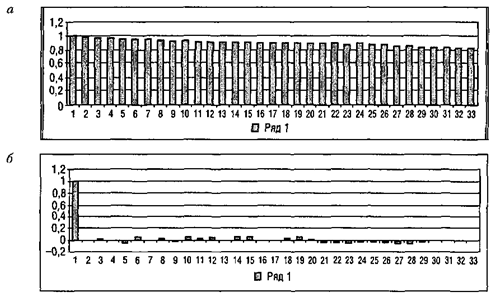
где **δ** = **ρ** - 1, а Δ ‒ оператор первой разности.

Оценивая коэффициент модели (2.6), можно проверить гипоте­зу о равенстве **δ = 0.** Если **δ = 0**, мы принимаем  нестационар­ным, если же оценка отрицательна, мы принимаем гипотезу о стационарности *.* При выполнении нулевой гипотезы **δ = 0** t-значение оценки коэффициента **δ** даже для больших выборок не следует асимптотически нормальному распределению. Дики и Фуллер построили ῖ-статистику для оценки t-значения коэффи­циента **δ** и рассчитали критические значения этой статистики методом Монте-Карло. Тест Дики-Фуллера и его обобщение на случай автокоррелированных остатков входит в стандартные ИПП (например, ЕViews). Нестационарные временные ряды, которые после выделения детерминированного тренда (например, с помощью МНК) приво­дятся к стационарному ряду остатков, называются **стационарны­ми относительно тренда.**

На основании значений функции автокорреляции и ее графика (коррелограммы) можно делать заключение о выборе вида модели ряда. Если колебания коррелограммы периодические и их ампли­туда практически постоянна, то принимается аддитивная модель. Если же амплитуда колебаний коррелограммы возрастает (убыва­ет), то принимается мультипликативная модель.

****

**Рис. 2.1 ‒ Примеры графиков АКФ (а) и ЧАКФ (б) стационарного временного ряда**

**

**Рис. 2.2 ‒ Примеры графиков АКФ (а) и ЧАКФ (б) для нестационарного случайного процесса**

Большинство экономических временных рядов являются неста­ционарными, поскольку их вероятностные характеристики зави­сят от времени. Если математическое ожидание временного ряда изменяется во времени в соответствии с некоторым детерминиро­ванным или стохастическим трендом, то **ряд не стационарен по среднему значению.** **Для описания таких рядов используются:**

**‒ модели с детерминированным трендом,** т.е. модели с трендом в виде детерминированной функции времени;

**‒ модели авторегрессии интегрированного скользящего среднего** (ARIMA-модели).

Разностно-стационарные процессы имеют стационарные ряды первых (или последующих) разностей. Тренд-стационарные процессы будут стационарными относительно линии тренда (детерминированного или стохастического).

**3 Проверка наличия и нахождение тренда**

Одной из важнейших задач исследования показателей экономического временного ряда является выявление основной тенденции изучаемого процесса, выраженной неслучайной составляющей **ƒ(t)** (тренда либо тренда с циклической и (или) сезонной компонентой). Для этого необходимо выбрать вид функции **ƒ(t)** (линейная, полиномиальная, экспоненциальная, логистическая, Гомперца и т.д.). Выбор той или иной функции является наиболее важным этапом анализа временного ряда. Неправильный выбор приводит к ошибкам при прогнозировании уровней ряда. Из нескольких вариантов предпочтение обычно отдается той функции, при которой меньше сумма квадратов отклонений фактических данных от расчетных на основе этих функций.

Для выявления основной тенденции чаще всего используется МНК. Значения временного ряда рассматриваются как зависимая переменная, а время **t** ‒ как объясняющая:

**= ƒ(t) + . (3.1)**

Согласно методу наименьших квадратов (МНК) в случае линейной зависимости параметры прямой  **=**  **ƒ(t) = +** находятся из системы нормальных уравнений, в которой в качестве берем **t**.

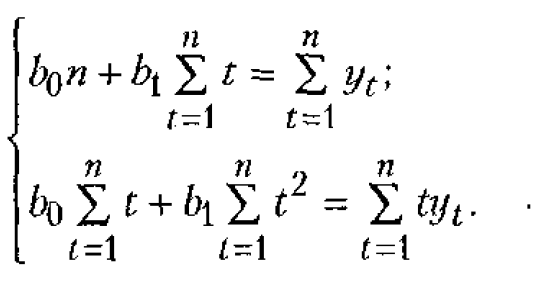
**Пример 3.1.** В таблице 3.1 приведены статистические данные по численности населения в Республике Беларусь за 1990 ‒ 2006 гг. Найти уравнение неслучайной составляющей (тренда) для временного ряда полагая тренд линейным.

**Таблица 3.1**

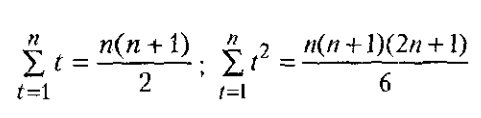
|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | 1 | 2 | 3 | 4 | 1 | 6 | 7 | 8 | 9 |
| Год | 1990 | 1995 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 |
| Численность насе­ления (млн. чел.) | 10,190 | 10,177 | 9,990 | 9,951 | 9,899 | 9,849 | 9,800 | 9,751 | 9,714 |

**Решение:**

Система нормальных уравнений для оценки параметров тренда имеет следующий вид:



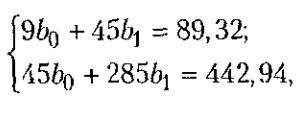
Для нахождения коэффициентов системы нормальных уравне­ний используем известные формулы:

,

(первая из них является суммой арифметической прогрессии, а вторая легко проверяется по индукции). По этим формулам будем иметь:

Далее:

Система нормальных уравнений имеет вид:



откуда = 10,58; = - 0,117 и уравнение тренда = 10,58 – 0,117**t** , т.е. численность населения ежегодно уменьшается в среднем на117 тыс. чел.

Остановимся подробнее на примерах моделей тренда.

**Полиномиальный тренд** имеет вид:

**=**  +  **+ 2  +** .... + **n  .**

Характерной особенностью линейного тренда является то, что приросты ординат остаются постоянными. Для полинома второй степени приросты ординат описываются линейной зависимостью  **- =**  - + **2t** , а приросты второго порядка постоянны.

Для выбора порядка полинома используются последовательные разности значений ряда. Нужная степень полинома соответствует случаю, когда разности постоянны (точнее их минимальный поря­док).

**Экспоненциальный тренд** имеет вид

**= abt .**

Характерной особенностью этой модели являются постоянные темпы роста и прироста, при этом приросты уровней ряда зависят от величины самой функции.

**Логистическая модель** тренда выглядит следующим образом:

**ƒ(t) =**

**Критерий Кокса-Стьюарта**

Одним из критериев проверки наличия тренда служит знако­вый критерий **Кокса-Стьюарта.** Для временного ряда, **t** *=* 1,…, **n***,* проверяется гипотеза об отсутствии тренда:

**: E {} = a = const,**

при альтернативной гипотезе:

**: E {} - E {} = 0 .**

**Гипотеза проверяется но следующему алгоритму.**

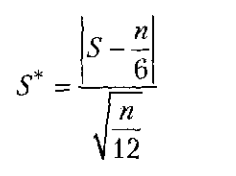
1. Временной ряд разбивается на три части: **n = + +** та­ких, что *= .*
2. Вычисляются разности всех наблюдений с одинаковыми но­мерами из третьей и первой частей и фиксируются их знаки.

Пусть  **‒** число положительных разностей; ‒ число отрица­тельных разностей.

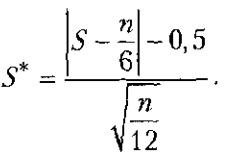
**S = max {}.**

Если > , то принимается решение о том, что тренд воз­растающий, а если , то - убывающий.

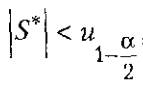
1. При **n** **> 30** статистика **S** аппроксимируется нормальным распределением со средним и дисперсией , а значит статистика



имеет стандартное нормальное распределение. При малых объ­емах выборки *(***n** *≤* **30** ) с учетом поправки на непрерывность



1. Гипотеза об отсутствии тренда принимается на уровне **ɑ**, если

, где  ‒ квантиль уровня  стандартно нормального распределения.

Особенностью теста Кокса-Стьюарта является тот факт, что его применение не требует предположений о распределении вероятностей уровней ряда. Такие тесты называются **непараметрическими**.

**Метод скользящего среднего**

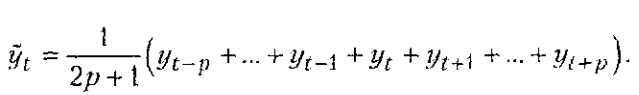
Для длинных временных рядов невозможно указать подходящую аппроксимирующую параметрическую кривую (функцию). Одним из методов выделения тренда в этом случае является **непа­раметрический метод скользящего среднего.** Он основан па пере­ходе от начальных значений членов ряда к их средним значениям на интервале времени, длина которого определена заранее. При этом сам выбранный интервал времени «скользит» вдоль ряда.

Скользящее среднее позволяет сгладить как случайные, так и периодические колебания, выявить имеющуюся тенденцию в раз­витии показателей и поэтому служит важным инструментом при фильтрации компонент временного ряда. Сглаживание по простой скользящей средней представляется в виде следующей последова­тельности шагов.

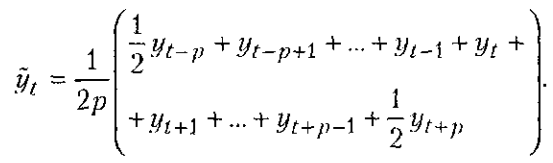
1. Сначала определяется длина интервала сглаживания **m**, включающего **m** последовательных уровней ряда. Чем сильнее колебания показателей, тем большим должен быть интервал сгла­живания.
2. Затем исходный ряд разбивается на участки по **m** уровней. Каждый последующий участок сдвигается по отношению к преды­дущему на один лаг времени.
3. Срединное значение каждого участка заменяется средней арифметической (или взвешенной средней) уровней из данного участка.

Интервал **m** удобнее брать из нечетного числа наблюдений **m = 2р +1**. Однако, если период колебаний ряда является четным, **m** можно брать также и четным. Если представить уровни участка сглаживания для **m = 2р +1** в виде

,…, , , ,…, , **то скользящая средняя определяется формулой**

** (3.2)**

При четном **m = 2p** скользящая средняя находится по формуле:

 **(3.3)**

Указанные формулы применяются, когда основная тенденция ряда является линейной. Когда же тренд сглаживаемого ряда имеет колебания, целесообразно использовать взвешенную скользящую среднюю. На каждом из участков слагаемые берутся со своими весовыми коэффициентами. Например, если сглажи­вание производится с помощью полинома, то на основе метода наименьших квадратов можно показать, что для интервала дли­ны **m** = **5** весовые коэффициенты в формуле (3.2) будут следующими: - , , , , - Для интервала длины **m** = **7**: - , , , , , - .

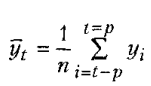
**Пример 3.2.** В таблице 3.2 приведены данные об урожайности пшеницы. Выполнить сглаживание уровней временного ряда при помощи простых скользящих средних. В качестве интервала сгла­живания принять значение **m = 5.**

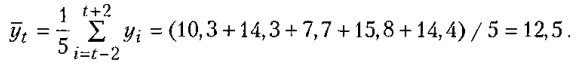
**Таблица 3.2**

|  |  |  |  |
| --- | --- | --- | --- |
| **t** |  | **t** |  |
| 1 | 10,3 | 9 | 17,1 |
| 2 | 14,3 | 10 | 7,7 |
| 3 | 7,7 | 11 | 15,3 |
| 4 | 15,8 | 12 | 16,3 |
| 5 | 14,4 | 13 | 19,9 |
| 6 | 16,7 | 14 | 14,4 |
| 7 | 15,3 | 15 | 18,7 |
| 8 | 20,2 | 16 | 20,7 |

**Решение:**

Для вычисления значений уровней ряда, сглаженных при по­мощи простых скользящих средних, воспользуемся формулой:

, в которой величине интервала сглаживания **m = 5** соответствует значение **р = 2***.* Таким образом, число членов сгла­женного ряда будет на четыре меньше числа членов исходного ря­да (теряются два первых и два последних наблюдения). Первое сглаженное значение соответствует значению **t** = **3** и вычисляется по данным, соответствующим **t** = **1**, ... , **t** = **5**.



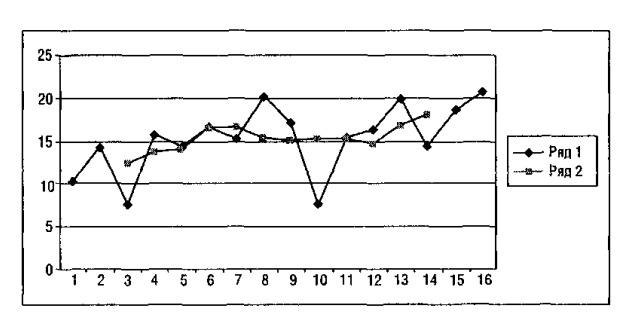
Результаты вычислений приведены в таблице 3.3.

**Таблица 3.3**

|  |  |  |
| --- | --- | --- |
| **t** |  |  |
| 1 | 10,3 |  |
| 2 | 14,3 |  |
| 3 | 7,7 | 12,5 |
| 4 | 15,8 | 13,8 |
| 5 | 14,4 | 14,0 |
| 6 | 16,7 | 16,5 |
| 7 | 15,3 | 16,7 |
| 8 | 20,2 | 15,4 |
| 9 | 17,1 | 15,1 |
| 10 | 7,7 | 15,3 |
| 11 | 15,3 | 15,3 |
| 12 | 16,3 | 14,7 |
| 13 | 19,9 | 16,9 |
| 14 | 14,4 | 18,0 |
| 15 | 18,7 |  |
| 16 | 20,7 |  |

На рисунке 3.1 представлены уровни исходного временного ряда (ряд 1) и уровни ряда сглаженного простой средней (ряд 2).

**Замечание.** Оказывается, что в результате сглаживания полу­ченный ряд, хотя и является более гладким, может содержать сис­тематические колебания, вызванные усреднением показателей. Этот факт впервые был выявлен в 1927 г. Е.Е. Слуцким (эффект Слуцкого-Юла).



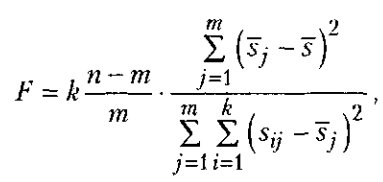
**Рис. 3.1 ‒ Графики исходного и сглаженного временных рядов**

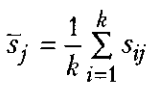
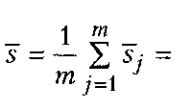
**4 Моделирование сезонной компоненты**

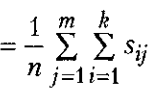
**4.1 Проверка наличия сезонной компоненты**

Приведем два критерия для обнаружения сезонной компонен­ты временного ряда. Первый из них ‒ дисперсионный, второй ‒ основан на вычислении коэффициента автокорреляции уровней ряда. Для применения **дисперсионного критерия** выдвигается ги­потеза о том, что во временном ряде, из которого отфильтрован тренд, отсутствуют сезонные колебания. В случае справедливости данной гипотезы F-статистика будет иметь F-распределение с

*(***m*-*1**) и *(***n-m***)* степенями свободы:

 **(4.1)**

где  ‒ среднее для периода с номером **j;  **

‒  общее среднее; ‒ значения ряда после выделенияиз него тренда; **i** = **1**,..., **k***,* **j** *=* **1**,..., **m**; **k**‒ число лет наблюдений; **m** ‒ период колебаний (**m**= **4** для квартальных данных**, m= 12**для месячных данных); **n = km** *‒* количество всех наблюдений.

Проверяемая гипотеза отвергается (т.е. сезонные колебания присутствуют) на уровне значимости **ɑ***,* если F-статистика (4.1) больше критического значения **F(ɑ, m *-*1, n – m)**.

Для обнаружения сезонных колебаний с использованием **коэффициента автокорреляции** его значения рассчитываются по формуле (2.2).

Расчетные значения коэффициентов автокорреляции сравниваются с табличными значениями с заданным уровнем значимости, В случае, когда расчетное значение больше табличного, соответ­ствующий коэффициент признается значимым, что может свиде­тельствовать о наличии сезонных колебаний. В таблице 4.1 приведеныкритические значения коэффициента корреляции **r(ῖ)** в зависимости от количества наблюдений. Точнее заключение о наличии или отсутствии сезонных колебаний можно сделать на основании коррелограммы. Если значения коэффициента корреляции **r(ῖ)** изменяются периодически, то этот период и будет периодом сезонных колебаний.

**Таблица 4.1**

|  |  |
| --- | --- |
| Число наблюдений, n | r(ῖ) |
| 10 | 0,360 |
| 15 | 0,328 |
| 20 | 0,300 |
| 25 | 0,276 |
| 30 | 0,257 |

**Пример 4.1.** Исследовать структуру временного ряда по квар­тальным данным потребления электроэнергии за 2001-2004 гг. (таблица 4.2). Оценить уровень и структуру потребления электро­энергии в 2005 г.

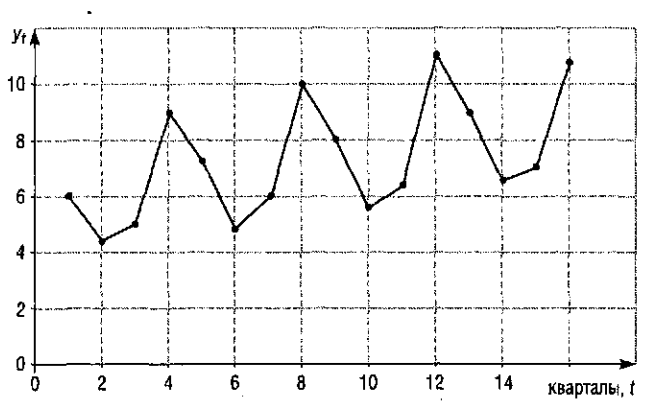
**Решение:**

В данной задаче в качестве зависимой переменной выступает потребление электроэнергии, в качестве независимой переменной ‒время **t (t = 1,16 )** . Проверим наличие сезонности в ряду.

**Таблица 4.2**

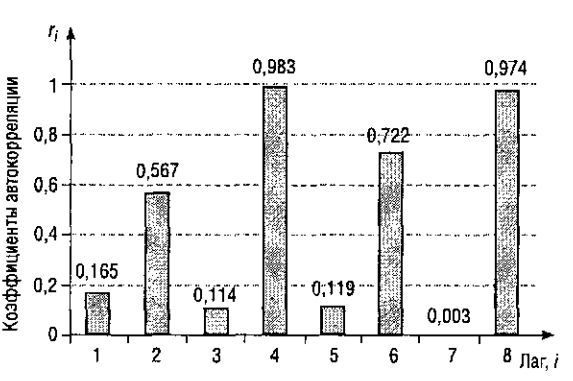
|  |  |
| --- | --- |
| Квартал | Потребление электроэнергии (млн кВт/ч) |
| 1-й квартал 2001 г.  2-й квартал 2001 г.  3-й квартал 2001 г.  4-й квартал 2001 г. | 6  4,4  5  9 |
| 1-й квартал 2002 г.  2-й квартал 2002 г.  3-й квартал 2002 г.  4-й квартал 2002 г. | 7,2  4,8  6  10 |
| 1-й квартал 2003 г.  2-й квартал 2003 г.  3-й квартал 2003 г.  4-й квартал 2003 г. | 8  5,6  6,4  11 |
| 1-й квартал 2004 г.  2-й квартал 2004 г.  3-й квартал 2004 г.  4-й квартал 2004 г. | 9  6,6  7  10,8 |

Первоначально изобразим ряд графически (рисунок 4.1). Из графика видно, что в 4-м квартале потребление электро­энергии ежегодно возрастает. Поэтому есть подозрение на нали­чие сезонной компоненты в ряде. Визуально также видно, что амплитуда сезонных колебаний постоянна. Это позволяет предположить аддитивную структуру временного ряда **у = Т + S + Е.**

**

**Рис. 4.1** *‒*  **График потребления электроэнергии за период 1-й квартал 2001 г. ‒ 4-й квартал 2004 г.**

Чтобы формально проверить наличие сезонности, рассчитаем автокорреляционную функцию временного ряда. Для этого рас­считаем коэффициенты автокорреляции первого, второго, ..., вось­мого порядков (рисунок 4.2).



**Рис. 4.2 ‒ Коррелограмма для данных примера 4.1**

Анализ автокорреляционной функции позволяет сделать вывод о сезонных колебаниях периодичностью 4 квартала.

**4.2 Выделение сезонной компоненты**

**1 способ (сглаживанием).** Для расчета сезонных компо­нент воспользуемся методом скользящей средней.

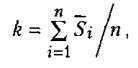
Просуммируем уровни ряда последовательно за каждые четыре квартала со сдвигом на один момент времени и определим услов­ные годовые объемы потребления электроэнергии (таблица 4.3, графа 3). Полученные суммы разделим на длину периода (в нашем случае на 4) и найдем скользящие средние, которые уже не зависят от се­зонности (таблица 4.3, графа 4). Чтобы привести эти значения в соот­ветствие с фактическими моментами времени, найдем средние значения из каждых двух соседних скользящих средних (таблица 4.3, графа 5). Оценку сезонной компоненты найдем, вычитая из факти­ческого значения уровня ряда центрированную скользящую среднюю (таблица 4.3, графа 6). На следующем этапе подготовим вто­рую вспомогательную таблицу 4.4. Занесем в нее оценки сезонных компонент, распределив их по кварталам. За каждый квартал найдем среднюю оценку сезонной компоненты.

Например, для 1-го квартала: = (0,575+ 0,55 +0,675)/3 = 0,6.

**Таблица 4.3**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № периода, **t** | Потребление электроэнергии, | Итого за 4 квартала | Скользящая средняя за 4 квартала | Центрированная скользящая средняя | Оценка сезонной компоненты |
| 1 | 6 | ‒ | ‒ | ‒ | ‒ |
| 2 | 4,4 | 24,4 | 6,1 | ‒ | ‒ |
| 3 | 5 | 25,6 | 6,4 | 6,25 | -1,25 |
| 4 | 9 | 26 | 6,5 | 6,45 | 2,55 |
| 5 | 7,2 | 27 | 6,75 | 6,625 | 0,575 |
| 6 | 4,8 | 28 | 7 | 6,875 | -2,075 |
| 7 | 6 | 28,8 | 7,2 | 7,1 | -1,1 |
| 8 | 10 | 29,6 | 7,4 | 7,3 | 2,7 |
| 9 | 8 | 30 | 7,5 | 7,45 | 0,55 |
| 10 | 5,6 | 31 | 7,75 | 7,625 | -2,025 |
| 11 | 6,4 | 32 | 8 | 7,875 | -1,475 |
| 12 | 11 | 33 | 8,25 | 8,125 | 2,875 |
| 13 | 9 | 33,6 | 8,4 | 8,325 | 0,674 |
| 14 | 6,6 | 33,4 | 8,35 | 8,375 | -1,775 |
| 15 | 7 | ‒ | ‒ | ‒ | ‒ |
| 16 | 10,8 | ‒ | ‒ |

Сезонные воздействия за период должны взаимопогашаться. В аддитивной модели это выражается в том, что сумма всех сезон­ных компонент за период должна быть равна нулю. **Рассчитаем корректирующий коэффициент по формуле:**



где **n** ‒ длина периода.

Для нашего примера **k** = (0, б — 1,958 — 1,275 + 2,708)/4 = 0,01875.

Скорректированные значения сезонной компоненты рассчиты­ваем как разность между средним значением сезонной компонен­ты и корректирующим коэффициентом: = – **k , i = 1,n .**

**Таблица 4.4**

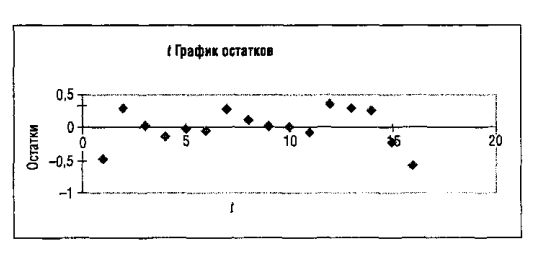
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № квартала | 2001 г. | 2002 г. | 2003 г. | 2004 г. | Средняя оценка сезонной компоненты для i-го квартала, | Скорректированная сезонная компонента, |
| 1 |  | 0,575 | 0,55 | 0,675 | 0,6 | 0,6 |
| 2 |  | -2,075 | -2,025 | -1,775 | -1,958 | -2,0 |
| 3 | -1,25 | -1,1 | -1475 |  | -1,275 | -1,3 |
| 4 | 2,55 | 2,7 | 2,875 |  | 2,708 | 2,7 |
| Корректирующий коэффициент | | | | | 0,01875 |  |

Исключим сезонную компоненту из исходного ряда, т.е. рассчи­таем **у - S.**С этой целью заполним рабочую таблицу (таблицу 4.5). Далее в преобразованном ряду **у - S** выделим линейный тренд: **T** = **5,7+ 0,2t,R2 = 0,91, DW = 1,29 t*-*ст*.* (38,9), (12,3).**

**Таблица 4.5**

|  |  |  |  |
| --- | --- | --- | --- |
| № периода, t | Исходный ряд, y | Сезонная компонента, S | Преобразованный ряд, у - S |
| 1 | 6 | 0,6 | 5,4 |
| 2 | 4,4 | -2,0 | 6,4 |
| 3 | 5 | -1,3 | 6,3 |
| 4 | 9 | 2,7 | 6,3 |
| 5 | 7,2 | 0,6 | 6,6 |
| 6 | 4,8 | -2,0 | 6,8 |
| 7 | 6 | -1,3 | 7,3 |
| 8 | 10 | 2,7 | 7,3 |
| 9 | 8 | 0,6 | 7,4 |
| 10 | 5,6 | -2,0 | 7,6 |
| 11 | 6,4 | -1,3 | 7,7 |
| 12 | 11 | 2,7 | 8,3 |
| 13 | 9 | 0,6 | 8,4 |
| 14 | 6,6 | -2,0 | 8,6 |
| 15 | 7 | -1,3 | 8,3 |
| 16 | 10,8 | 2,7 | 8,1 |

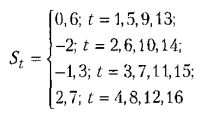
График остатков позволяет предположить гомоскедастичность остатков (рисунок, 4.3).



**Рисунок 4.3 ‒ График зависимости остатков от факторов времени**

В данном случае статистика **DW** лежит в области неопределенности (при уровне значимости 5 % = **1,37 ,**  = **1,1**). Поэтому мы не можем ни отклонить, ни принять нулевую гипотезу относительно автокорреляции остатков. Если условно принять допущение об отсутствии статической значимости автокорреляции остатков, то использованием последней модели найдем потребление электроэнергии в каждом квартале.

Зная значения сезонных компонент



и тренд **Т = 5,7 + 0,2t** , можно прогнозировать потребление электроэнергии в последующих кварталах с использованием модели **= 5,7 + 0,2t +**  :

= **5,7 + 0,2 × 17 + 0,6 = 9,7 (млн кВт/ч);**

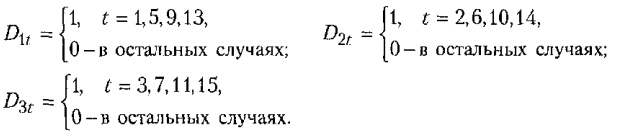
= **5,7 + 0,2 × 18 - 2 = 7,3 (млн кВт/ч);**

= **5,7 + 0,2 × 19 – 1,3 = 8,2 (млн кВт/ч);**

= **5,7 + 0,2 × 20 + 2,7 = 12,4 (млн кВт/ч).**

В целом году: 9,7 + 7,3 + 8,2 + 12,4 = 37,6 (млн кВт/ч).

**2 способ (включением фиктивной сезонной компоненты).** В качестве факторов в модель включают показатели: время **t** и три фиктивные сезонные компоненты **D1***,***D2**,**D3**, описывающие сезонную компоненту соответственно 1-го, 2-го и 3-го кварталов:



Для 4-го квартала значения всех сезонных компонент равны нулю.

**Таблица 4.6**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **t** | **D1t** | **D2t** | **D3t** | **yt** |
| 1 | 1 | 0 | 0 | 6 |
| 2 | 0 | 1 | 0 | 4,4 |
| 3 | 0 | 0 | 1 | 5 |
| 4 | 0 | 0 | 0 | 9 |
| 5 | 1 | 0 | 0 | 7,2 |
| 6 | 0 | 1 | 0 | 4,8 |
| 7 | 0 | 0 | 1 | 6 |
| 8 | 0 | 0 | 0 | 10 |
| 9 | 1 | 0 | 0 | 8 |
| 10 | 0 | 1 | 0 | 5,6 |
| 11 | 0 | 0 | 1 | 6,4 |
| 12 | 0 | 0 | 0 | 11 |
| 13 | 1 | 0 | 0 | 9 |
| 14 | 0 | 1 | 0 | 6,6 |
| 15 | 0 | 0 | 1 | 7 |
| 16 | 0 | 0 | 0 | 10,8 |

Включение в модель всех четырех сезонных компонент может привести к эффекту мультиколлинеарности. Формируем таблицу входных данных для построения модели.

Уравнение тренда с включением фиктивных сезонных компо­нент выглядит следующим образом:

**yt = 8,3 + 0,19t – 2,1D1 – 4,5D2 – 3,9D3 ,**

**R2 = 0,98, F = 180, DW = 1,2 ,**

**t-ст. (36,6) (11,1) (-9,5) (-20,6) (-18,2) .**

Статистические характеристики свидетельствуют о значимости уравнения в целом и его коэффициентов.

Как и в предыдущем способе построения модели, в данном слу­чае **DW** статистика лежит в области неопределенности (при уров­не значимости 5 % = 1,93, 0,74). Поэтому мы не можем ни отклонить, ни принять нулевую гипотезу относительно авто­корреляции остатков, Если условно принять допущение об от­сутствии статистически значимой автокорреляции остатков, то с использованием последней модели найдем потребление электро­энергии в каждом квартале:

= **8,3 + 0,19 × 17 – 2,1 = 9,43 (млн кВт/ч);**

= **8,3 + 0,19 × 18 – 4,5 = 7,22 (млн кВт/ч);**

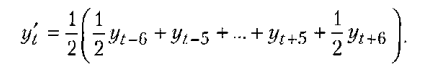
= **8,3 + 0,19 × 19 – 3,9 = 8,01 (млн кВт/ч);**

= **8,3 + 0,19 × 20 = 12,1 (млн кВт/ч).**

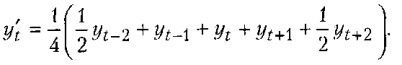
В целом по году: 9,43 + 7,22 + 8,01 + 12,1 = 36,76 (млн кВт/ч).

Рассмотрим алгоритм расчета для случая **мультипликативной сезонности:**

1 Для описания тенденции воспользуемся процедурой скользящей средней при четной длине интервала сглаживания **l = 2p**. Тогда для временных рядов месячной динамики скользящая сред­няя при **l = 12** на каждом активном участке будет определяться выражением:

 **(4.2)**

Для рядов квартальной динамики при **l = 4** можно использо­вать выражение

 **(4.3)**

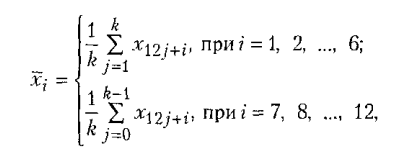
2 Рассчитаем отношение фактических значений **yt** к уровням сглаженного ряда , полученным на предыдущем шаге



Уровни вновь полученного ряда **xt** отражают эффект сезонности и случайности.

3 Для исключения влияния случайных факторов определим предварительные значения сезонной составляющей как средние значения из уровней **xt** для одноименных месяцев (кварталов).

**Например, для временных рядов месячной динамики процедура усреднения может быть описана выражением:**



где **k** ‒ число целых периодов (циклов) во временном ряду, полученном на втором шаге.

Разные пределы суммирования объясняются тем, что при ис­пользовании скользящей средней с четным значением длины ин­тервала сглаживания **(l = 2p) p** первых и **p** последних уровней ряда будут потеряны.

4 Проведем корректировку первоначальных значений сезон­ной составляющей, вызванную тем, что суммарное воздействие се­зонности на динамику предполагается нейтральным. Это свойство для сезонных колебаний в мультипликативной форме выражается в том, что средняя арифметическая из значений коэффициентов сезонности для полного сезонного цикла должна быть равна еди­нице. Поэтому окончательные оценки коэффициентов сезонности получим с помощью следующего выражения:

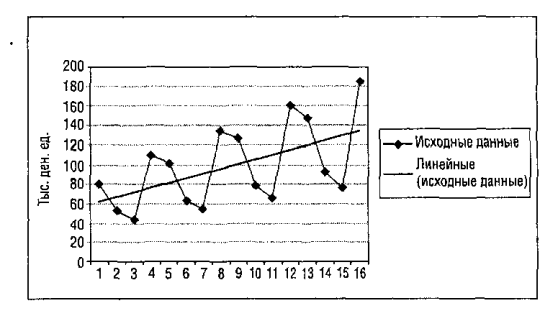
**Si = Xi × k (I = 1,2,…, m) ,**

где **k** *=* ; **m ‒** число фаз в полном сезонном цикле (как правило, **m** *=* **12** ‒ для рядов месячной динамики и **m** *=* **12** ‒ для квар­тальных данных).

**Пример 4.2**. В таблице 4.7 представлены квартальные данные об из­менении прибыли фирмы за последние четыре года, Рассчитать про­гнозную оценку уровня прибыли в 1-м полугодии следующего года.

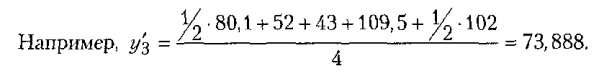
**Решение:**

Графический анализ исходного временного ряда свидетель­ствует о наличии трендовой компоненты: в анализируемом перио­де наглядно выражена тенденция роста прибыли фирмы, причем характер тенденции близок к линейному развитию. На рисунке 4.4 от­четливо видны сезонные колебания. Наблюдается устойчиво по­вторяющееся увеличение прибыли в 1-м и 4-м кварталах по срав­нению с 3-м и 2-м кварталами, причем наиболее существенные всплески в динамике показателя отмечаются в 4-м квартале. Так как амплитуда сезонных колебаний изменяется с течением време­ни, увеличивается с ростом тренда, то для описания и прогнозиро­вания динамики временного ряда можно предложить модель с мультипликативной сезонностью.



**Рис. 4.4 ‒ Квартальная динамика прибыли фирмы.**

Поясним последовательность вычислений. В графе 4 таблицы 4.7 показаны результаты сглаживания исходного временного ряда с помощью скользящей средней в соответствии с выражением (4.3).



В графе 5 таблицы 4.7 показаны уровни полученные делением исходного временного ряда на значения сглаженного ряда. Уров­ни отражают влияние случайных факторов и сезонности.

**Таблица 4.7**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № квартала | **t** | Прибыль, , тыс. ден. ед. | Скользящая средняя, **y’** | **xt = yt / yt’** | Десезонализированный ряд, **yt** | Расчетные уровни прибыли, **yt** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1  2  3  4 | 1  2  3  4 | 80,1  52  43  109,5 | 73,888  78,063 | 0,582  1,403 | 63,400  69,217  73,393  78,245 | 80,074  51,982  43,008  109,544 |
| 1  2  3  4 | 5  6  7  8 | 102  63,5  53  134,1 | 80,75  85,075  91,15  95,975 | 1,263  0,746  0,581  1,397 | 80,782  7,987  90,461  95,824 | 101,967  63,478  53,01  134,153 |
| 1  2  3  4 | 9  10  11  12 | 126  78,1  65,2  159,6 | 99,325  103,038  109,913  114,338 | 1,269  0,751  0,593  1,396 | 99,730  103,959  111,284  114,045 | 125,96  78,07  65,212  159,664 |
| 1  2  3  4 | 13  14  15  16 | 147,5  92  75,9  184 | 117,413  121,8 | 1,256  0,755 | 116,747  122,462  129,546  131,481 | 147,452  91,968  75,914  184,073 |

Предварительную оценку сезонности получим усреднением уровней временного ряда **xt** для одноименных кварталов.

Например, значение сезонной составляющей для 1-го квартала:

**X1 = = 1,263 .**

Аналогичным образом вычислены **X2 = 0,75, X3 = 0,586 , X4  = 1,399 .** Так как *=* **3,998** (достаточно близка к 4), то найденные оценки сезонности можно было бы оставить в неизменном виде, ведь очевидно, что процедура корректировки приведет к их незначительному изменению.

Однако для реализации обсуждаемого методического подхода в полном объеме проведем корректировку ранее полученных оценок сезонности. Поправочный коэффициент **k = 4 /**  3,998 = 1,0006 . Тогда окончательная оценка сезонности для 1-го квартала:

**S1 = 1,0006 × 1,263 = 1,263**. Аналогичным образом вычислены **S2 = 0,751, S3 = 0,586, S4 =1,4**.

В графе 6 таблицы 4.7 представлен десезонализированный времен­ной ряд, полученный делением исходного временного ряда на со­ответствующие коэффициенты сезонности (коэффициенты сезон­ности принимаются одинаковыми для каждого анализируемого года).

Для описания тенденции воспользуемся моделью линейного тренда, так как это согласуется с результатами графического ана­лиза динамики показателя. **Модель для десезонализированного временного ряда имеет вид:**

**yt(1) = 59,026+ 4,563t.**

На заключительном этапе расчетные уровни прибыли (графа 7 таблица 4.7) были вычислены умножением значений, полученных по трендовой модели, на соответствующие коэффициенты сезон­ности.

Прогнозная оценка уровня прибыли в 1-м полугодии следую­щего года состоит из суммы оценок прибыли в 17-м квартале и в 18-м, т.е. суммы следующих значений:

**y17** = **(59,026 + 4,563 × 17) × 1,263 = 172,52;**

**y18 = (59,026 + 4,563 × 18) × 0,751 = 106,01** .

Таким образом, по этой модели ожидаемый уровень прибыли в 1-м полугодии следующего года составит 278,53 тыс. ден. ед.

**5 Экспоненциальное сглаживание и адаптивный прогноз**

**Модель экспоненциального сглаживания** является основой для построения адаптивных методов прогнозирования. **Адаптивные методы** прогнозирования позволяют строить самокорректирующи­еся экономико-математические модели, которые учитывают изме­нения путем учета результата прогноза, сделанного на предыдущем шаге. Один из первых методов адаптивного прогнозирования, осно­ванный на экспоненциальном сглаживании предложил Р. Браун. Существует несколько модификаций и развитий метода Брауна: модель Ч. Хольта, модель Дж. Бокса и Г. Дженкинса и др. Известны модели P. Брауна нулевого, первого и второго порядков, описывающие процессы, соответственно не имеющие тенденции разви­тия, имеющие линейную и параболическую тенденции развития. Мы рассмотрим **линейную адаптивную модель Р. Брауна.** Предпо­ложим, что временной ряд представляется в виде:

**yt = a0 + a1 × t + Ɛt ,**

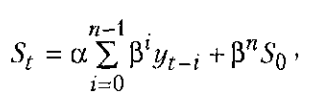
где **a0 , a1** ‒ параметры; **Ɛt** ‒ случайные неавтокоррелированные остатки с нулевым математическим ожиданием и дисперсии **𝛔2** (белый шум).

Модель экспоненциального сглаживания ряда описывается следующей рекуррентной формулой:

**St = ɑ × yt + β × St – 1 ,** **(5.1)**

где **St** ‒ значение экспоненциальной средней в момент **t**; **ɑ**‒ па­раметр сглаживания; **ɑ = соnst,** **0** < **ɑ** < **1**, **β** = **1 - ɑ**.

Если последовательно использовать формулу (5.1), то экспо­ненциальную среднюю можно выразить через предшествую­щие значения уровней временного ряда в виде:

 **(5.2)**

где **n** ‒ длина ряда; **S0** ‒ начальное значение экспоненциальной средней.

Из формулы (5.2) видно, что величина **St** является взвешен­ной суммой всех членов ряда. При этом веса уровней ряда убыва­ют по мере их удаления в прошлое соответственно экспоненци­альной функции. Данное свойство явилось основой для названия экспоненциального сглаживания.

Например, для **ɑ = 0,1** вес текущего наблюдения **yt** будет ра­вен 0,1, вес предыдущего уровня **yt - 1** ‒ 0,09, уровня **yt - 1** ‒ 0,081 и т.д. На первом шаге должна быть определена начальная величина **S0**. Часто на практике в качестве значения **S0** берется среднее всех уровней ряда. Для длинных рядов выбор не играет суще­ственной роли.

При значении **ɑ**, близком к единице, дисперсия экспоненци­альной средней незначительно отличается от дисперсии ряда.

С уменьшением **ɑ** **(0 < ɑ < 1)** дисперсия экспоненциальной сред­ней уменьшается. Таким образом, экспоненциальная средняя иг­рает роль фильтра, поглощающего колебания временного ряда. Поиск компромиссного значения параметра **ɑ** составляет важный этап исследования.

**Опишем этапы построения линейной адаптивной модели Р. Брауна.**

1 По первым пяти точкам временного ряда оцениваем значе­ния параметров **а0** и **а1** методом наименьших квадратов. Их оцен­ки **а0** и **а1** являются коэффициентами уравнения регрессии:

**yt = а0 + а1 × t .**

2 С использованием параметров **а0** и **а1** , которые соответ­ствуют моменту времени **t** = **0**, вычисляют начальные условия экспоненциальных средних:

**S0(1) = a0(0) - × a1(0) ;**

**S0(2) = a0(0) - × a1(0) .**

3 С учетом выбранного значения параметра сглаживания **ɑ** или коэффициента дисконтирования **β (ɑ + β = 1)** вычисляют зна­чения экспоненциальных средних:

**St(1) = ɑyt + ΒSt – 1(1) ;**

**St(2) = ɑSt(1) + ΒSt – 1(2) .**

4 Корректируют параметры модели **а0(t)** и **а1(t)** по следую­щим формулам:

**а0(t) = 2St(1) – St(2) ;**

**а1(t) = (St(1) - St(2)) .**

5 По модели со скорректированными параметрами **а0(t)** и **а1(t)** находят прогноз на следующий момент времени:

**yt(ῖ) = а0(t) + а1(t)ῖ = y(t + 1) = а0(t) + а1(t) .**

6 Возврат к пункту 3, если **t** *<* **n** *.* Если **t** *<* **n**, то построенную модель можно использовать для прогнозирования на будущее. То­чечный прогноз рассчитывают по формуле:

**y(n + ῖ) = а0(n) + а1(n)ῖ , ῖ** **= 1, 2, ... .**

**Пример 5***.***1** По имеющейся информации **(yt)** в таблице 5.1 об объемах экспорта (млрд дсн. ед.) построить адаптивную модель Брауна с линейной тенденцией.

Построить прогноз на три шага вперед, используя оптимальное значение параметра сглаживания. Результаты моделирования и прогнозирования привести на графике.

**Таблица 5.1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **t** | **yi** | **t** | **yi** | **t** | **yi** |
| 1 | 7,44 | 11 | 21,56 | 21 | 36,01 |
| 2 | 8,91 | 12 | 22,95 | 22 | 37,42 |
| 3 | 10,38 | 13 | 24,43 | 23 | 38,76 |
| 4 | 11,82 | 14 | 25,87 | 24 | 40,24 |
| 5 | 13,17 | 15 | 27,25 | 25 | 41,73 |
| 6 | 14,55 | 16 | 28,73 | 26 | 43,1 |
| 7 | 16,01 | 17 | 30,29 | 27 | 44,51 |
| 8 | 17,34 | 28 | 31,7 | 28 | 46 |
| 9 | 18,68 | 19 | 33,07 | 29 | 47,48 |
| 10 | 20,09 | 20 | 34,59 | 30 | 48,84 |

**Решение:**

**1** По первым пяти точкам временного ряда оцениваем значе­ния **b0** и **b1** параметров модели с помощью МНК для линейной аппроксимации **yt = b0 + b1t .** Получаем начальные значения параметров модели **b0** **= 5,96** и **b1** *=* **1,44**, которые соответствуют моменту времени **t = 0**. Все рас­четы (таблица 5.2) показаны для **ɑ** = **0,3, β = 0,7.**

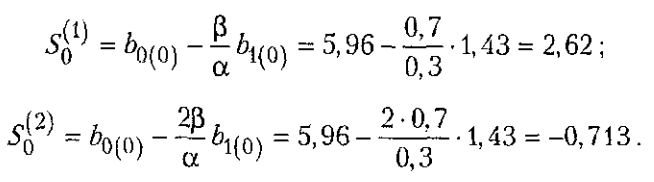
**Таблица 5.2**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **t** | **yi** | **b0** | **b1** | **St(1)** | **St(2)** | **yt** | **et = yt - yt** |
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
| 0 |  | 5,96 | 1,43 | 2,623333 | - 0,71333 |  |  |
| 1 | 7,44 | 6,702167 | 1,4345 | 4,068333 | 0,721167 | 7,39 | 0,05 |
| 2 | 8,91 | 9,601767 | 1,4399 | 5,520833 | 2,161067 | 8,136667 | 0,773333 |
| 3 | 10,38 | 12,51191 | 1,445255 | 6,978583 | 3,606322 | 11,04167 | - 0,66167 |
| 4 | 11,82 | 15,41461 | 1,447406 | 8,431008 | 5,053728 | 13,95717 | - 2,13717 |
| 5 | 13,17 | 18,26572 | 1,43969345 | 9,852706 | 6,493421 | 16,86202 | - 3,69202 |
| 6 | 14,55 | 21,09325 | 1,43054189 | 11,26189 | 7,923963 | 19,70541 | - 5,15541 |
| 7 | 16,01 | 23,94394 | 1,428708856 | 12,68633 | 9,352672 | 22,52379 | - 6,51379 |
| 8 | 17,34 | 26,74593 | 1,418926872 | 14,08243 | 10,7716 | 25,37265 | - 8,03265 |
| 9 | 18,68 | 29,51673 | 1,407030281 | 15,4617 | 12,17863 | 28,16486 | - 9,48486 |
| 10 | 20,09 | 32,29891 | 1,401468226 | 16,85019 | 13,5801 | 30,9234 | - 10,8334 |
| 11 | 21,56 | 35,12135 | 1,404910679 | 18,26313 | 14,98501 | 33,70038 | - 12,1404 |
| 12 | 22,95 | 37,93313 | 1,40525552 | 19,66919 | 16,39026 | 36,52627 | - 13,5763 |
| 13 | 24,43 | 40,78272 | 1,412151495 | 21,09744 | 17,80241 | 39,33839 | - 14,9084 |
| 14 | 25,87 | 43,64037 | 1,418036888 | 22,5292 | 19,22045 | 42,19487 | - 16,3249 |
| 15 | 27,25 | 46,47339 | 1,417497411 | 23,94544 | 20,63795 | 45,05841 | - 17,8084 |
| 16 | 28,73 | 49,33876 | 1,4228583 | 25,38081 | 22,06081 | 47,89089 | - 19,1609 |
| 17 | 30,29 | 52,26931 | 1,437827889 | 26,85357 | 23,49864 | 50,76162 | -20,4716 |
| 18 | 31,7 | 55,17234 | 1,442658477 | 28,3075 | 24,94129 | 53,70713 | - 22,0071 |
| 19 | 33,07 | 58,03401 | 1,438486203 | 29,73625 | 26,37978 | 56,61499 | - 23,545 |
| 20 | 34,59 | 60,94097 | 1,44377803 | 31,19237 | 27,82356 | 59,4725 | - 24,8825 |
| 21 | 36,01 | 63,83109 | 1,444231003 | 32,63766 | 29,26779 | 62,38475 | - 26,3747 |
| 22 | 37,42 | 66,70335 | 1,441372169 | 34,07236 | 30,70916 | 65,27532 | - 27,8553 |
| 23 | 38,76 | 69,52646 | 1,430847845 | 35,47865 | 32,14001 | 68,14473 | - 29,3847 |

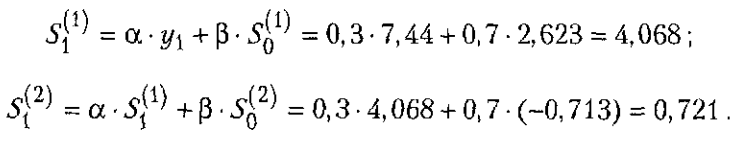
**Окончание таблицы 5.2**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
| 24 | 40,24 | 72,384 | 1,430114621 | 36,90706 | 33,57012 | 70,95731 | - 30,7173 |
| 25 | 41,73 | 75,27274 | 1,435145025 | 38,35394 | 35,00527 | 73,81412 | - 32,0841 |
| 26 | 43,1 | 78,12377 | 1,43174867 | 39,77776 | 36,43702 | 76,70788 | - 33,6079 |
| 27 | 44,51 | 80,96674 | 1,428124556 | 41,19743 | 37,86514 | 79,55552 | - 35,0455 |
| 28 | 46 | 83,84448 | 1,431918413 | 42,6382 | 39,29706 | 82,39486 | - 36,3949 |
| 29 | 47,48 | 86,74338 | 1,438104745 | 44,09074 | 40,73516 | 85,2764 | - 37,7964 |
| 30 | 48,84 | 89,59693 | 1,434106621 | 45,51552 | 42,16927 | 88,18148 | - 39,3415 |

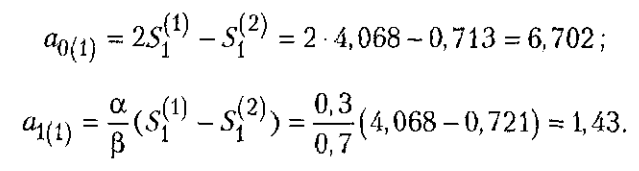
1. Вычисляем начальные условия экспоненциальных средних:



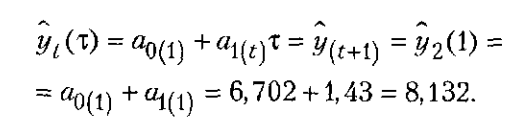
**3** С учетом выбранного значения параметра сглаживания **β (ɑ + β = 1)** вычисляем значения экспоненциальных средних **(t = 1):**



**4** Корректируем параметры модели **а0(t)** и **а1(t)** **(t = 1):**



**5** По модели со скорректированными параметрами **а0(t)** и **а1(t)** находим прогноз на следующий момент времени **ῖ = 1**, **t = 2:**

****

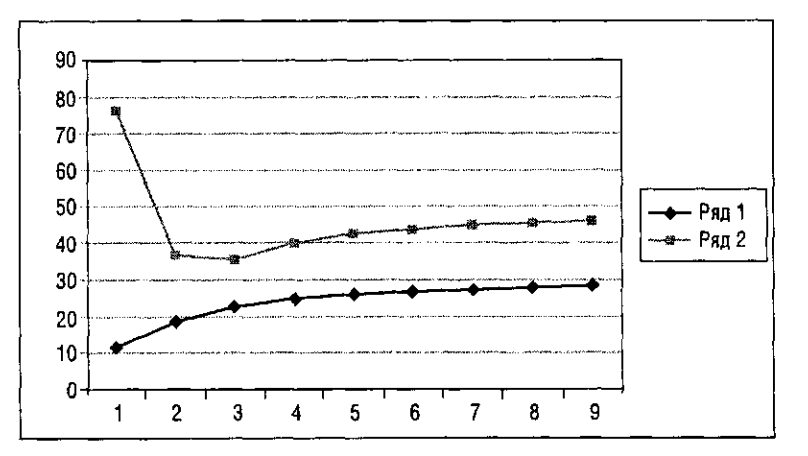
**6** Возвращаемся к пункту 3, вычисления повторяем до конца наблюдений.

Для выбора лучшей модели вычислены показатели точности, результаты по которым для различных параметров сглаживания приведены в таблице 5.3.

**Таблица 5.3**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Показатели точности модели, % | Параметры сглаживания **ɑ** | | | | | | | | |
| 0,1 | 0,2 | 0,3 | 0,4 | 0,5 | 0,6 | 0,7 | 0,8 | 0,9 |
| Среднеквадратичная ошибка | 11,93 | 18,79 | 22,67 | 24,73 | 25,99 | 26,85 | 27,46 | 27,92 | 28,28 |
| Средняя относительная ошибка | 76,21 | 36,59 | 35,82 | 40,07 | 42,51 | 43,98 | 44,97 | 45,68 | 46,22 |
| Средняя ошибка | -0,23 | -14,03 | -18,64 | -20,94 | -22,32 | -23,26 | -23,90 | -24,40 | -24,78 |
| Максимальная по модулю ошибка | 20,23 | 34,55 | 39,34 | 41,74 | 43,19 | 44,16 | 44,86 | 45,38 | 45,79 |

На рисунке 5.1 приведены графики средней относительной ошибки (Ряд 2) и среднеквадратичной ошибки (Ряд 1) для различных па­раметров сглаживания. Для построения прогноза выбрана модель с параметром сглаживания 0,3.



**Рис. 5.1 Графики средней относительной и среднеквадратичной ошибки**

**6 АRIМА-модели временных рядов**

Среди временных рядов особо выделяются стационарные вре­менные ряды. К рядам с такими свойствами могут быть приведе­ны нестационарные ряды после выделения тренда, фильтрации сезонной компоненты или взятия разностей. Ряд случайных ос­татков классической регрессионной модели является стационарным при выполнении условий Гаусса-Маркова. Для моделирования стационарных рядов часто применяются модели авторегрессии и скользящего среднего. **Авторегрессионная модель порядка ρ (AR(ρ)) представляется в виде:**

**= ɑ1 × yt – 1 + ɑ2 × yt – 2 +…+ ɑp × yt – p  + Ɛt . (6.1)**

В частности, для порядка **ρ** = **1 AR(1)** модель авторегрессии называется **марковским процессом:**

**= ɑ× yt – 1 + Ɛt , (6.2)**

где **|ɑ|** < **1**, а случайные остатки образуют белый шум (с постоянной дисперсией **𝛔02**). Для такого процесса математическое ожидание и дисперсия равны соответственно:

**Myt = 0; D(yt) = 𝛔02**

**1 - ɑ2 ,**

а автокорреляционная функция ‒ **ρ(ῖ) = ɑῖ .**

Таким образом, коэффициент **ɑ** является коэффициентом автокорреляции первого порядка: **ɑ** **= ρ(1)**. В силу равенства **ρ**(ῖ) = **ɑῖ (|ɑ| < 1 )** степень тесноты корреляционной связи между членами ряда и их смещенными значениями экспоненциально убывают.

Для ряда **AR(1)** существует отличительный признак: **значения его частной автокорреляционной функции равны пулю для лага ῖ > 1.** Указанные свойства автокорреляционной функции и частной автокорреляционной функции могут служить признаками процесса **AR(1)**. Выборочные коэффициенты автокорреляции частной автокорреляции могут, конечно, несколько отличаться от теоретических.

Для авторегрессионной модели порядка **p** частная автокорреляциониая функция обращается в нуль, начиная со значения аргу­мента **ῖ** *=* **p** *+***1**.

На практике часто встречаются модели **AR(2)** с остатками в виде белого шума:

**= ɑ1 × yt – 1 + ɑ2 × yt – 2 + Ɛt .**

**Они называются процессами Юла.**

**Модель скользящего среднего порядка** **q (MA(q))** имеет вид:

**= Ɛt + γ1  × Ɛt – 1 + γ2  × Ɛt – 2**  +**…** + **γq  × Ɛt – q .**

Характерной особенностью модели **MA(1)** является условие обращения в нуль автокорреляционной функции ряда  для ла­гов **ῖ** *>* **2**.

На практике чаще используется объединенная модель **ARMA:**

**= ɑ× yt – 1 + Ɛt + γ × Ɛt – 1 .**

Стационарность этого процесса выполняется при условии **|ɑ| < 1 .**

Условие же **|γ| < 1**  обеспечивает возможность выражения теку­щих значений в виде сходящегося ряда своих прошлых значе­ний **, , … .**

Для процесса **АRМА** значения автокорреляционной функ­ции будут экспоненциально убывать от значений **ρ(1)**, причем, если **ɑ** положительно ‒ то монотонно, если отрицательно - то знакопеременно.

Поведение частной автокорреляционной функции определяет­ся начальным значением **ρ(1)**, после которого функция экспо­ненциально убывает. Если **γ** положительно, то функция убывает монотонно, если отрицательно ‒ то знакопеременно.

Анализ поведения автокорреляционной функции и частной ав­токорреляционной функции оказывают существенную помощь в выборе вида модели.

**Нестационарным однородным** называется временной ряд, для которого случайный остаток, полученный вычитанием из исход­ного ряда его неслучайной составляющей, представляет собой ста­ционарный временной ряд.

Для описания таких рядов используется модель **авторегрессии ‒проинтегрированного скользящего среднего *(*ARIMA*-*модель)**, разработанная Боксом-Дженкинсом в начале 1970-х гг.

**Эта модель используется для описания временных рядов, обла­дающих свойствами:**

* ряд включает аддитивную неслучайную составляющую в ви­де полинома;
* ряд, полученный из исходного применением к нему последовательных разностей, может быть описан моделью **ARMA (р,q)**.

Характерным свойством степени используемого полинома может служить обращение в нуль разности соответствующего порядка. Если в нуль обращаются разности второго порядка, то берется полином первой степени, если третьего порядка, то берет­ся полином второго порядка и т.д.

Если порядок используемых разностей равен **d***,* модель обозна­чается **ARIMA (р,d,q)**.

Методология построения модели **ARIMA** состоит из следую­щей последовательности шагов.

**На первом шаге** строится стационарный ряд. Сначала вы­деляется неслучайная трендовая составляющая, которую можно обнаружить визуальным анализом графика или построением мо­дели регрессии. После удаления трендовой составляющей проводится анализ автокорреляционной и частной автокорреляционной функций остатков. Быстрое затухание этих функций является одним из критериев стационарности.

Для выявления стационарности временного ряда и порядка интегрированности **d** часто используется **тест единичного корня Дики-Фуллера**. Следует отметить, что тесты Дики-Фуллера нельзя применять к временным рядам, содержащим структурные и сезонные изменения, а также в случае гетероскедастичности случайных остатков.

**На втором шаге** после получения стационарного ряда ос­татков исследуются графики автокорреляционной и частной авто­корреляционной функций. На их основе выдвигаются гипотезы о значениях порядка **р** авторегрессии и порядка **q** скользящего среднего.

**На третьем шаге** оцениваются параметры модели.

**На четвертом шаге** проводится проверка адекватности мо­дели на основе анализа ее ряда случайных остатков.

**Пример 6.1** В таблице 6.1 приведены квартальные уровни ВВП России, выраженные в процентах к уровню ВВП за 1-й квар­тал 2000 г.

**Таблица 6.1**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Квартал | Год | | | | | |
| 2000 | 2001 | 2002 | 2003 | 2004 | 2005 |
| 1-й | 100 | 105 | 108 | 117 | 125 | 132 |
| 2-й | 105 | 110 | 115 | 124 | 133 | 143 |
| 3-й | 122 | 130 | 135 | 144 | 153 | 164 |
| 4-й | 115 | 120 | 125 | 137 | 149 |  |

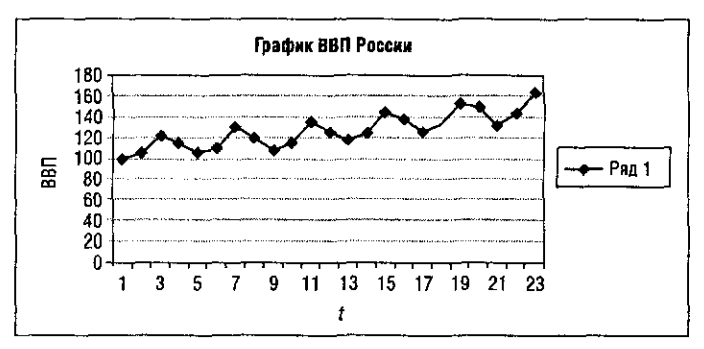
1. Построить график ряда и на основе его визуального анализа определить структуру модели ряда.
2. Вводя фиктивные переменные построить уравнение регрес­сии с учетом линейного тренда и сезонных компонент.
3. По значениям остатков модели проверить гипотезы о нали­чии автокорреляции и гетероскедастичности остатков.
4. Предполагая, что ряд остатков удовлетворяет условиям мо­дели авторегрессии первого порядка **AR(1): ut = ρ × ut – 1 + Ɛt,** най­ти оценку параметра **ρ** и проверить условие стационарности ряда остатков, используя графики АКФ и ЧЛКФ.

**Решение:**

Графический анализ исходного временного ряда свидетель­ствует о наличии трендовой компоненты: в анализируемом перио­де наглядно выражена тенденция роста ВВП, причем характер тенденции близок к линейному развитию. На рисунке 6.1 отчетливо видны сезонные колебания периодичностью в четыре квартала. Так как амплитуда сезонных колебаний не изменяется с течением времени, то для описания и прогнозирования динамики временно­го ряда можно предложить модель с аддитивной сезонностью.

После включения фиктивных переменных оставим 20 наблю­дений, для того чтобы данные содержали все годовые кварталы. В результате построения регрессии получим следующее уравне­ние регрессии:

**y = 107,22 – 1,83t – 12,14D1 – 8,13D2 – 9,43D3 , R2 = 0,98 , F = 223,7 .**

****

**Рисунок6*.*1 ‒ График ВВП России**

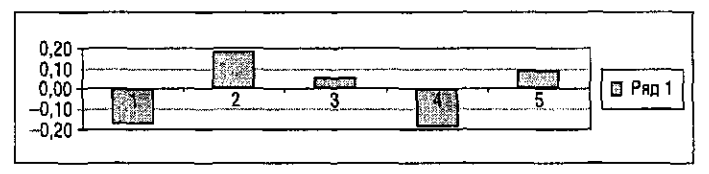
По значениям остатков модели проверяем гипотезы о наличии автокорреляции и гетероскедастичности остатков. В нашем случае статистика **DW** = **0,7** лежит в области положительной автокорре­ляции (при уровне значимости **5% DWu = 0,777, DWl =1,534**). По критерию Голдфелда-Квандта гипотеза о наличии гетероскеда­стичности остатков отклоняется, так как значение статистики равно **S3** / **S1 = 1,44**.

Строим модель авторегрессии первого порядка **АR(1)**. Для это­го определяем уточненную спецификацию модели стационарного временного ряда **ut**, соседние уровни которого **ut – 1** , **ut** коррели­руют с коэффициентом **ρ**. В результате проведенной регрессии получили = - **0,158 + 0,587** , **ρ = 0,59 ˂ 1 ,** свободный член, равный - **0,158**, незначим на уровне **0,05**. Для проверки свойства остатков этой регрессии рассчитаем матрицу парных коэффици­ентов корреляции (таблица 6.2).

**Таблица 6.2**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Столбец 1 | Столбец 2 | Столбец 3 | Столбец 4 | Столбец 5 | Столбец 6 |
| Столбец 1 | 1 |  |  |  |  |  |
| Столбец 2 | -0,17146 | 1 |  |  |  |  |
| Столбец 3 | 0,177273 | -0,123491902 | 1 |  |  |  |
| Столбец 4 | 0,056885 | 0,148072173 | -0,09201 | 1 |  |  |
| Столбец 5 | -0,18169 | 0,117943049 | 0,111294 | -0,04783 | 1 |  |
| Столбец 6 | 0,089418 | -0,139700928 | 0,294856 | 0,101962 | -0,04073 | 1 |

По данным столбца 1 таблицы 6.2 строим график АКФ (рисунок 6.2).



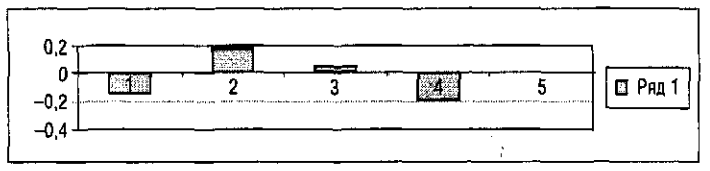
**Рис. 6.2 ‒ График АКФ**

Для построения графика ЧАКФ используем матрицу, обрат­ную матрице таблица 6.2. Вычислим частные коэффициенты корреляции (таблица 6.3).

**Таблица 6.3**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 | 2 | 3 | 4 | 5 |
| - 0,1358 | 0,181104 | 0,050023623 | -0,18462 | 0,005084 |

График ЧАКФ имеет следующий вид (рисунок 6.3).



**Рис. 6.3 ‒ График ЧАКФ**

Графики АКФ и ЧАКФ соответствуют графикам стационарно­го временного ряда. Стационарность ряда остат­ков подтверждается также условием **ρ < 1**.