

Лабораторная работа #1

Анализ и обработка временных рядов

Цель: изучить основы анализа, обработки и прогнозирования временных рядов, приобрести навыки работы с методами анализа, обработки и прогнозирования временных рядов в системе STATISTICA StatSoft, осуществить обработку, анализ и прогнозирование ряда и интерпретацию результатов

1 Ход работы

- 1) изучить теоретические сведения
- 2) приобрести навыки анализа, обработки и прогнозирования временных рядов в системе STATISTICA StatSoft, реализуя приведенный ниже пример
- 3) на основе приобретенных практических навыков осуществить все этапы анализа, обработки и прогнозирования временного ряда (прогноз на 10% шагов от размера ряда, аналогично примеру) и интерпретацию результатов согласно варианту индивидуального задания, подобрав опытным путем НАИЛУЧШУЮ прогнозную модель
- 4) оформить отчет и подготовиться к защите лабораторной работы по полученным результатам и контрольным вопросам

2 Содержание отчета и требования к его оформлению

- 1) отчет оформляется в печатном виде
- 2) отчет содержит титульный лист, исходные данные, результаты выполнения этапов обработки данных в виде скриншотов и обязательных комментариев по ходу выполнения работы, выводы
- 3) к отчету прилагается финальный файл проекта в электронном виде с целью осуществления выборочного контроля

3 Варианты исходных данных

- исходные данные - расположены в папке DataSet, конкретные данные выбираются согласно нижеприведенной таблицы.

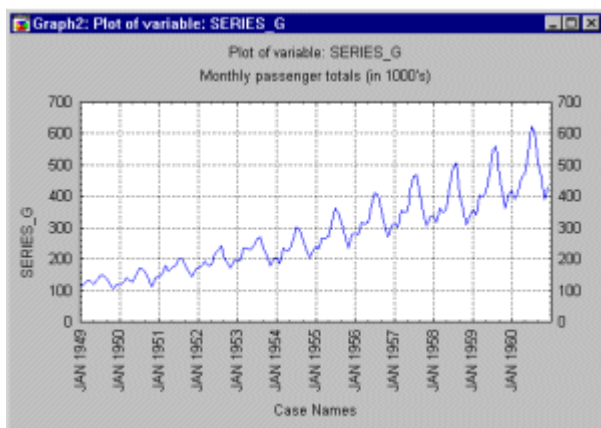
Вариант	Исходные данные (файл(-> столбец))
1	Brent crude oil spot price.xlsx
2	Canada Gas Production.xlsx
3	Cushing, OK WTI Spot Price FOB.xlsx
4	Henry Hub Natural Gas Spot Price.xlsx
5	Canada Crude Oil Production.xlsx
6	China_Electricity_Consumption.xlsx
7	Electricity_Net_Generation_From_Coal_Electric_Power_Sector_Monthly.xlsx
8	Monthly Mean Total Sunspot Number.xlsx
9	U.S. consumer price index.xlsx
10	US Natural Gas Storage Data.xlsx -> East Region
11	US Natural Gas Storage Data.xlsx -> Midwest Region
12	US Natural Gas Storage Data.xlsx -> Mountain Region
13	US Natural Gas Storage Data.xlsx -> Pacific Region
14	US Natural Gas Storage Data.xlsx -> South Central Region

4 Краткие теоретические сведения

Существуют две основные цели анализа временных рядов: 1) определение природы ряда и 2) прогнозирование (предсказание будущих значений временного ряда по настоящим и прошлым значениям). Обе эти цели требуют, чтобы модель ряда была идентифицирована и, более или менее, формально описана. Как только модель определена, вы можете с ее помощью интерпретировать рассматриваемые данные, не обращая внимания на глубину понимания и справедливость теории, вы можете экстраполировать затем ряд на основе найденной модели, т.е. предсказать его будущие значения.

Систематическая составляющая и случайный шум. Как и большинство других видов анализа, анализ временных рядов предполагает, что данные содержат систематическую составляющую (обычно включающую несколько компонент) и случайный шум (ошибку), который затрудняет обнаружение регулярных компонент. Большинство методов исследования временных рядов включает различные способы фильтрации шума, позволяющие увидеть регулярную составляющую более отчетливо.

Два общих типа компонент временных рядов. Большинство регулярных составляющих временных рядов принадлежит к двум классам: они являются либо трендом, либо сезонной составляющей. Тренд представляет собой общую систематическую линейную или нелинейную компоненту, которая может изменяться во времени. Сезонная составляющая - это периодически повторяющаяся компонента. Оба эти вида регулярных компонент часто присутствуют в ряде одновременно. Например, продажи компании могут возрастать из года в год, но они также содержат сезонную составляющую (как правило, 25% годовых продаж приходится на декабрь и только 4% на август).



Эту общую модель можно понять на "классическом" ряде - *Ряд G* (Бокс и Дженкинс, 1976, стр. 531), представляющем месячные международные авиаперевозки (в тысячах) в течение 12 лет с 1949 по 1960 (см. файл *Series_g.sta*). График месячных перевозок ясно показывает почти линейный тренд, т.е. имеется устойчивый рост перевозок из года в год (примерно в 4 раза больше пассажиров перевезено в 1960 году, чем в 1949). В то же время характер месячных перевозок повторяется, они имеют почти один и тот же характер в каждом годовом периоде (например, перевозок больше в отпускные периоды, чем в другие месяцы). Этот пример показывает довольно определенный тип модели временного ряда, в которой

амплитуда сезонных изменений увеличивается вместе с трендом. Такого рода модели называются моделями с *мультипликативной сезонностью*.

Анализ тренда. Не существует "автоматического" способа обнаружения тренда в временном ряде. Однако если тренд является монотонным (устойчиво возрастает или устойчиво убывает), то анализировать такой ряд обычно нетрудно. Если временные ряды содержат значительную ошибку, то первым шагом выделения тренда является сглаживание.

Сглаживание. Сглаживание всегда включает некоторый способ локального усреднения данных, при котором несистематические компоненты взаимно погашают друг друга. Самый общий метод сглаживания - *скользящее среднее*, в котором каждый член ряда заменяется простым или взвешенным средним n соседних членов, где n - ширина "окна". Вместо среднего можно использовать медиану значений, попавших в окно. Основное преимущество медианного сглаживания, в сравнении со сглаживанием скользящим средним, состоит в том, что результаты становятся более устойчивыми к выбросам (имеющимся внутри окна). Таким образом, если в данных имеются выбросы (связанные, например, с ошибками измерений), то сглаживание медианой обычно приводит к более гладким или, по крайней мере, более "надежным" кривым, по сравнению со скользящим средним с тем же самым окном. Основной недостаток медианного сглаживания в том, что при отсутствии явных выбросов, он приводит к более "зубчатым" кривым (чем сглаживание скользящим средним) и не позволяет использовать веса.

Относительно реже, когда ошибка измерения очень большая, используется метод *сглаживания методом наименьших квадратов, взвешенных относительно расстояния* или метод *отрицательного экспоненциально взвешенного сглаживания*. Все эти методы отфильтровывают шум и преобразуют данные в относительно гладкую. Ряды с относительно небольшим количеством наблюдений и систематическим расположением точек могут быть сглажены с помощью *бикубических сплайнов*.

Подгонка функции. Многие монотонные временные ряды можно хорошо приблизить линейной функцией. Если же имеется явная монотонная нелинейная компонента, то данные вначале следует преобразовать, чтобы устранить нелинейность. Обычно для этого используют логарифмическое, экспоненциальное или (менее часто) полиномиальное преобразование данных.

Анализ сезонности. Периодическая и сезонная зависимость (сезонность) представляет собой другой общий тип компонент временного ряда. Это понятие было проиллюстрировано ранее на примере авиаперевозок пассажиров. Можно легко видеть, что каждое наблюдение очень похоже на соседнее; дополнительно, имеется повторяющаяся сезонная составляющая, это означает, что каждое наблюдение также похоже на наблюдение, имевшееся в том же самом месяце год назад. В общем, периодическая зависимость может быть формально определена как корреляционная зависимость порядка k между каждым i -м элементом ряда и $(i-k)$ -м элементом (Kendall, 1976). Ее можно измерить с помощью автокорреляции (т.е. корреляции между самими членами ряда); k обычно называют *лагом* (иногда используют эквивалентные термины: сдвиг, запаздывание). Если ошибка измерения не слишком

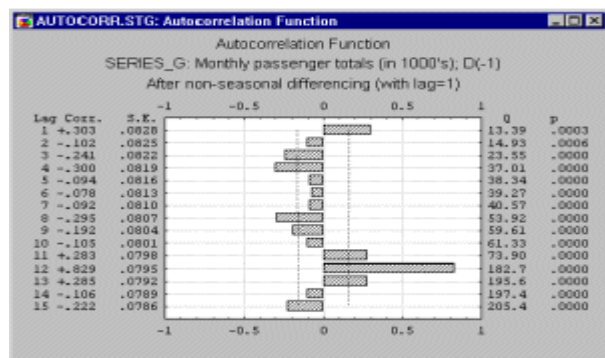
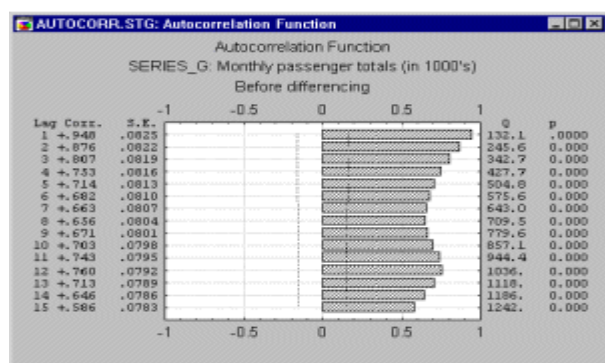
большая, то сезонность можно определить визуально, рассматривая поведение членов ряда через каждые k временных единиц.

Автокорреляционная коррелограмма. Сезонные составляющие временного ряда могут быть найдены с помощью коррелограммы. Коррелограмма (автокоррелограмма) показывает численно и графически автокорреляционную функцию (АКФ), иными словами коэффициенты автокорреляции (и их стандартные ошибки) для последовательности лагов из определенного диапазона (например, от 1 до 30). На коррелограмме обычно отмечается диапазон в размере двух стандартных ошибок на каждом лаге, однако обычно величина автокорреляции более интересна, чем ее надежность, потому что интерес в основном представляют очень сильные (а, следовательно, высоко значимые) автокорреляции.

!!! Корреляция представляет собой меру зависимости переменных. Наиболее известна корреляция Пирсона. Коэффициенты корреляции изменяются в пределах от -1.00 до +1.00. Обратите внимание на крайние значения коэффициента корреляции. Значение -1.00 означает, что переменные имеют строгую отрицательную корреляцию. Значение +1.00 означает, что переменные имеют строгую положительную корреляцию. Отметим, что значение 0.00 означает отсутствие корреляции. Наиболее часто используемый коэффициент корреляции Пирсона r называется также линейной корреляцией, т.к. измеряет степень линейных связей между переменными.

!!! Корреляционную зависимость между последовательными уровнями временного ряда называют автокорреляцией уровней ряда. Количественно ее можно измерить с помощью линейного коэффициента корреляции между уровнями исходного временного ряда и уровнями этого ряда, сдвинутыми на несколько шагов во времени.

Исследование коррелограмм. При изучении коррелограмм следует помнить, что автокорреляции последовательных лагов формально зависимы между собой. Рассмотрим следующий пример. Если первый член ряда тесно связан со вторым, а второй с третьим, то первый элемент должен также каким-то образом зависеть от третьего и т.д. Это приводит к тому, что периодическая зависимость может существенно измениться после удаления автокорреляций первого порядка, т.е. после взятия разности с лагом 1).



Частные автокорреляции. Другой полезный метод исследования периодичности состоит в исследовании частной автокорреляционной функции (ЧАКФ), представляющей собой углубление понятия обычной автокорреляционной функции. В ЧАКФ устраняется зависимость между промежуточными наблюдениями (наблюдениями *внутри* лага). Другими словами, частная автокорреляция на данном лаге аналогична обычной автокорреляции, за исключением того, что при вычислении из нее удаляется влияние автокорреляций с меньшими лагами (см. Бокс и Дженкинс, 1976). На лаге 1 (когда нет промежуточных элементов внутри лага), частная автокорреляция равна, очевидно, обычной автокорреляции. На самом деле, частная автокорреляция дает более "чистую" картину периодических зависимостей.

Удаление периодической зависимости. Как отмечалось выше, периодическая составляющая для данного лага k может быть удалена взятием разности соответствующего порядка. Это означает, что из каждого i -го элемента ряда вычитается $(i-k)$ -й элемент. Имеются два довода в пользу таких преобразований.

Во-первых, таким образом можно определить скрытые периодические составляющие ряда. Напомним, что автокорреляции на последовательных лагах зависимы. Поэтому удаление некоторых автокорреляций изменит другие автокорреляции, которые, возможно, подавляли их, и сделает некоторые другие сезонные составляющие более заметными.

Во-вторых, удаление сезонных составляющих делает ряд *стационарным*, что необходимо для применения АРПСС и других методов, например, спектрального анализа.

4.1 АРПСС (ARIMA, модель авторегрессии и скользящего среднего)

Два основных процесса

4.1.1 Процесс авторегрессии. Большинство временных рядов содержат элементы, которые последовательно зависят друг от друга. Такую зависимость можно выразить следующим уравнением:

$$x_t = \xi + \phi_1 * x_{(t-1)} + \phi_2 * x_{(t-2)} + \phi_3 * x_{(t-3)} + \dots + \varepsilon$$

Здесь:

ξ - константа (свободный член),
 ϕ_1, ϕ_2, ϕ_3 - параметры авторегрессии.

Вы видите, что каждое наблюдение есть сумма случайной компоненты (случайное воздействие, ε) и линейной комбинации предыдущих наблюдений.

Требование стационарности. Заметим, что процесс авторегрессии будет стационарным только, если его параметры лежат в определенном диапазоне. Например, если имеется только один параметр, то он должен находиться в интервале $-1 < \phi < +1$. В противном случае, предыдущие значения будут накапливаться и значения последующих x_t могут быть неограниченными, следовательно, ряд не будет стационарным. Если имеется несколько параметров авторегрессии, то можно определить аналогичные условия, обеспечивающие стационарность.

4.1.2 Процесс скользящего среднего. В отличие от процесса авторегрессии, в процессе скользящего среднего каждый элемент ряда подвержен суммарному воздействию предыдущих ошибок. В общем виде это можно записать следующим образом:

$$x_t = \mu + \varepsilon_t - \theta_1 * \varepsilon_{(t-1)} - \theta_2 * \varepsilon_{(t-2)} - \theta_3 * \varepsilon_{(t-3)} - \dots$$

Здесь:

μ - константа,
 $\theta_1, \theta_2, \theta_3$ - параметры скользящего среднего.

Другими словами, текущее наблюдение ряда представляет собой сумму случайной компоненты (случайное воздействие, ε) в данный момент и линейной комбинации случайных воздействий в предыдущие моменты времени.

Обратимость. Не вдаваясь в детали, отметим, что существует "двойственность" между процессами скользящего среднего и авторегрессии. Это означает, что приведенное выше уравнение скользящего среднего можно переписать (*обратить*) в виде уравнения авторегрессии (неограниченного порядка), и наоборот. Это так называемое свойство обратимости. Имеются условия, аналогичные приведенным выше условиям стационарности, обеспечивающие обратимость модели.

4.1.3 Модель АРПСС (ARIMA)

Модель авторегрессии и скользящего среднего. Общая модель, предложенная Боксом и Дженкинсом (1976) включает как параметры авторегрессии, так и параметры скользящего среднего. Именно, имеется три типа параметров модели: параметры авторегрессии (p), порядок разности (d), параметры скользящего среднего (q). В обозначениях Бокса и Дженкинса модель записывается как ARIMA (p, d, q). Например, модель (0, 1, 2) содержит 0 (нуль) параметров авторегрессии (p) и 2 параметра скользящего среднего (q), которые вычисляются для ряда после взятия разности с лагом 1.

Идентификация. Как отмечено ранее, для модели ARIMA необходимо, чтобы ряд был *стационарным*, это означает, что его среднее постоянно, а выборочные дисперсия и автокорреляция не меняются во времени. Поэтому обычно необходимо брать разности ряда до тех пор, пока он не станет стационарным (часто также применяют логарифмическое преобразование для стабилизации дисперсии). Число разностей, которые были взяты, чтобы достичь стационарности, определяются параметром d (см. выше). Для того чтобы определить необходимый порядок разности, нужно исследовать график ряда и автокоррелограмму. Сильные изменения уровня (сильные скачки вверх или вниз) обычно требуют взятия несезонной разности первого порядка (лаг=1). Сильные изменения наклона требуют взятия разности второго порядка. Сезонная составляющая требует взятия соответствующей сезонной разности (см. ниже). Если имеется медленное убывание выборочных коэффициентов автокорреляции в зависимости от лага, обычно берут разность первого порядка. Однако следует помнить, что для некоторых временных рядов нужно брать разности небольшого порядка или вовсе не брать их. Заметим, что *чрезмерное количество взятых разностей* приводит к менее стабильным оценкам коэффициентов.

На этом этапе (который обычно называют *идентификацией* порядка модели) вы также должны решить, как много параметров авторегрессии (p) и скользящего среднего (q) должно присутствовать в эффективной и *экономной* модели процесса. (*Экономность* модели означает, что в ней имеется наименьшее число параметров и наибольшее число степеней свободы среди всех моделей, которые подгоняются к данным). На практике очень редко бывает, что число параметров p или q больше 2.

Оценивание и прогноз. Следующий, после идентификации, шаг (*Оценивание*) состоит в оценивании параметров модели. Полученные оценки параметров используются на последнем этапе (*Прогноз*) для того, чтобы вычислить новые значения ряда и построить доверительный интервал для прогноза. Процесс оценивания проводится по преобразованным данным (подвергнутым применению разностного оператора). До построения прогноза нужно выполнить обратную операцию (*интегрировать* данные). Таким образом, прогноз методологии будет сравниваться с соответствующими исходными данными. На интегрирование данных указывает буква *I* в общем названии модели (ARIMA = Авторегрессионное Проинтегрированное Скользящее Среднее).

Константа в моделях ARIMA. Дополнительно модели ARIMA могут содержать константу, интерпретация которой зависит от подгоняемой модели. Именно, если (1) в модели нет параметров авторегрессии, то константа μ есть среднее значение ряда, если (2) параметры авторегрессии имеются, то константа представляет собой свободный член. Если бралась разность ряда, то константа представляет собой среднее или свободный член преобразованного ряда. Например, если бралась первая разность (разность первого порядка), а параметров авторегрессии в модели нет, то константа представляет собой среднее значение преобразованного ряда и, следовательно, *коэффициент наклона линейного тренда* исходного.

Идентификация

Число оцениваемых параметров. Конечно, до того, как начать оценивание, вам необходимо решить, какой тип модели будет подбираться к данным, и какое количество параметров присутствует в модели, иными словами, нужно идентифицировать модель ARIMA. Основными инструментами идентификации порядка модели являются графики, автокорреляционная функция (АКФ), частная автокорреляционная функция (ЧАКФ). Это решение не является простым и требуется основательно поэкспериментировать с альтернативными моделями. Тем не менее, большинство встречающихся на практике временных рядов можно с достаточной степенью точности аппроксимировать одной из 5 основных моделей (см. ниже), которые можно идентифицировать по виду автокорреляционной (АКФ) и частной автокорреляционной функции (ЧАКФ). Ниже дается список этих моделей.

1. *Один параметр (p):* АКФ - экспоненциально убывает; ЧАКФ - имеет резко выделяющееся значение для лага 1, нет корреляций на других лагах.
2. *Два параметра авторегрессии (p):* АКФ имеет форму синусоиды или экспоненциально убывает; ЧАКФ имеет резко выделяющиеся значения на лагах 1, 2, нет корреляций на других лагах.
3. *Один параметр скользящего среднего (q):* АКФ имеет резко выделяющееся значение на лаге 1, нет корреляций на других лагах. ЧАКФ экспоненциально убывает.

4. *Два параметра скользящего среднего (q):* АКФ имеет резко выделяющиеся значения на лагах 1, 2, нет корреляций на других лагах. ЧАКФ имеет форму синусоиды или экспоненциально убывает.
5. *Один параметр авторегрессии (p) и один параметр скользящего среднего (q):* АКФ экспоненциально убывает с лага 1; ЧАКФ - экспоненциально убывает с лага 1.

Сезонные модели. Мультипликативная сезонная ARIMA представляет естественное развитие и обобщение обычной модели ARIMA на ряды, в которых имеется периодическая сезонная компонента. В дополнение к несезонным параметрам, в модель вводятся сезонные параметры для определенного лага (устанавливаемого на этапе идентификации порядка модели). Аналогично параметрам простой модели ARIMA, эти параметры называются: сезонная авторегрессия (ps), сезонная разность (ds) и сезонное скользящее среднее (qs). Таким образом, полная сезонная ARIMA может быть записана как ARIMA $(p,d,q)(ps,ds,qs)$. Например, модель $(0,1,2)(0,1,1)$ включает 0 регулярных параметров авторегрессии, 2 регулярных параметра скользящего среднего и 1 параметр сезонного скользящего среднего. Эти параметры вычисляются для рядов, получаемых после взятия одной разности с лагом 1 и далее сезонной разности. Сезонный лаг, используемый для сезонных параметров, определяется на этапе идентификации порядка модели.

Общие рекомендации относительно выбора обычных параметров (с помощью АКФ и ЧАКФ) полностью применимы к сезонным моделям. Основное отличие состоит в том, что в сезонных рядах АКФ и ЧАКФ имеют существенные значения на лагах, кратных сезонному лагу (в дополнении к характерному поведению этих функций, описывающих регулярную (несезонную) компоненту ARIMA).

4.2 Оценивание параметров

Существуют различные методы оценивания параметров, которые дают очень похожие оценки, но для данной модели одни оценки могут быть более эффективны, а другие менее эффективны. В общем, во время оценивания порядка модели используется так называемый *квазиньютоновский алгоритм* максимизации правдоподобия (вероятности) наблюдения значений ряда по значениям параметров. Практически это требует вычисления (условных) сумм квадратов (SS) остатков модели. Имеются различные способы вычисления суммы квадратов остатков SS ; вы можете выбрать: (1) приближенный метод максимального правдоподобия МакЛеода и Сейлза (1983), (2) приближенный метод максимального правдоподобия с итерациями назад, (3) точный метод максимального правдоподобия по Меларду (1984).

Стандартные ошибки оценок. Для всех оценок параметров вычисляются так называемые *асимптотические стандартные ошибки*, для вычисления которых используется матрица частных производных второго порядка, аппроксимируемая конечными разностями.

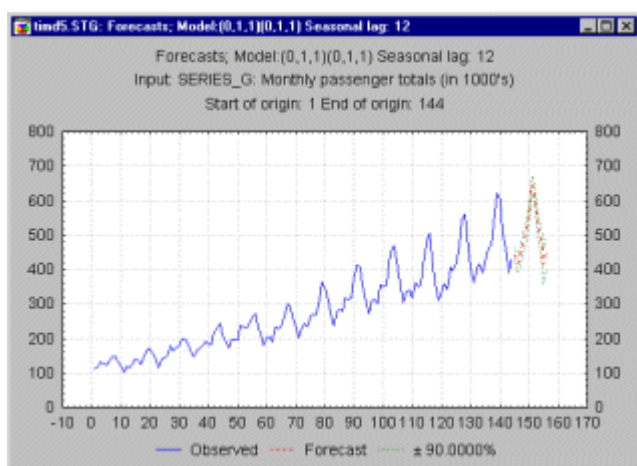
Штраф. Процедура оценивания минимизирует (условную) сумму квадратов остатков модели. Если модель не является адекватной, может случиться так, что оценки параметров на каком-то шаге станут неприемлемыми - очень большими (например, не удовлетворяют условию стационарности). В таком случае, SS будет

приписано очень большое значение (*штрафное значение*). Обычно это "заставляет" итерационный процесс удалить параметры из недопустимой области. Однако в некоторых случаях и эта стратегия может оказаться неудачной, и вы все равно увидите на экране (во время *процедуры оценивания*) очень большие значения SS на серии итераций. В таких случаях следует с осторожностью оценивать пригодность модели. Если модель содержит много параметров и, возможно, имеется интервенция, то следует несколько раз испытать процесс оценивания с различными начальными. Если модель содержит много параметров и, возможно, интервенцию, вам следует повторить процедуру с различными начальными значениями параметров.

4.3 Оценивание модели

Оценки параметров. Если значения вычисляемой t статистики не значимы, соответствующие параметры в большинстве случаев удаляются из модели без ущерба подгонки.

Другой критерий качества. Другой обычной мерой надежности модели является сравнение прогноза, построенного по урезанному ряду с "известными (исходными) данными".



Однако качественная модель должна не только давать достаточно точный прогноз, но быть экономной и иметь независимые остатки, содержащие только шум без систематических компонент (в частности, АКФ остатков не должна иметь какой-либо периодичности). Поэтому необходим всесторонний анализ остатков. Хорошей проверкой модели являются: (а) график остатков и изучение их трендов, (б) проверка АКФ остатков (на графике АКФ обычно отчетливо видна периодичность).

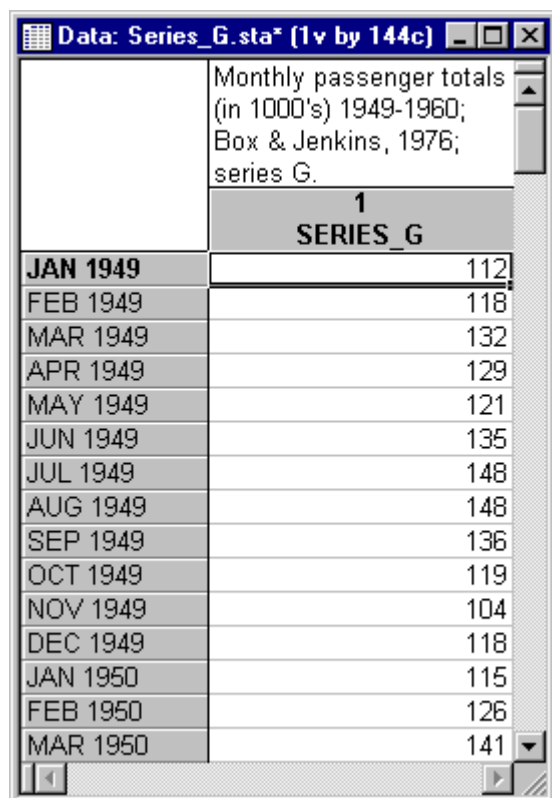
Анализ остатков. Если остатки систематически распределены (например, отрицательны в первой части ряда и примерно равны нулю во второй) или включают некоторую периодическую компоненту, то это свидетельствует о неадекватности модели. Анализ остатков чрезвычайно важен и необходим при анализе временных рядов. Процедура оценивания предполагает, что остатки не коррелированы и нормально распределены.

Ограничения. Следует напомнить, что модель ARIMA является подходящей только для рядов, которые являются *стационарными* (среднее, дисперсия и автокорреляция примерно постоянны во времени); для нестационарных рядов

следует брать разности. Рекомендуется иметь, как минимум, 50 наблюдений в файле исходных данных. Также предполагается, что параметры модели постоянны, т.е. не меняются во времени.

5 Пример реализации методов обработки, анализа и прогнозирования временных рядов в системе STATISTICA StatSoft

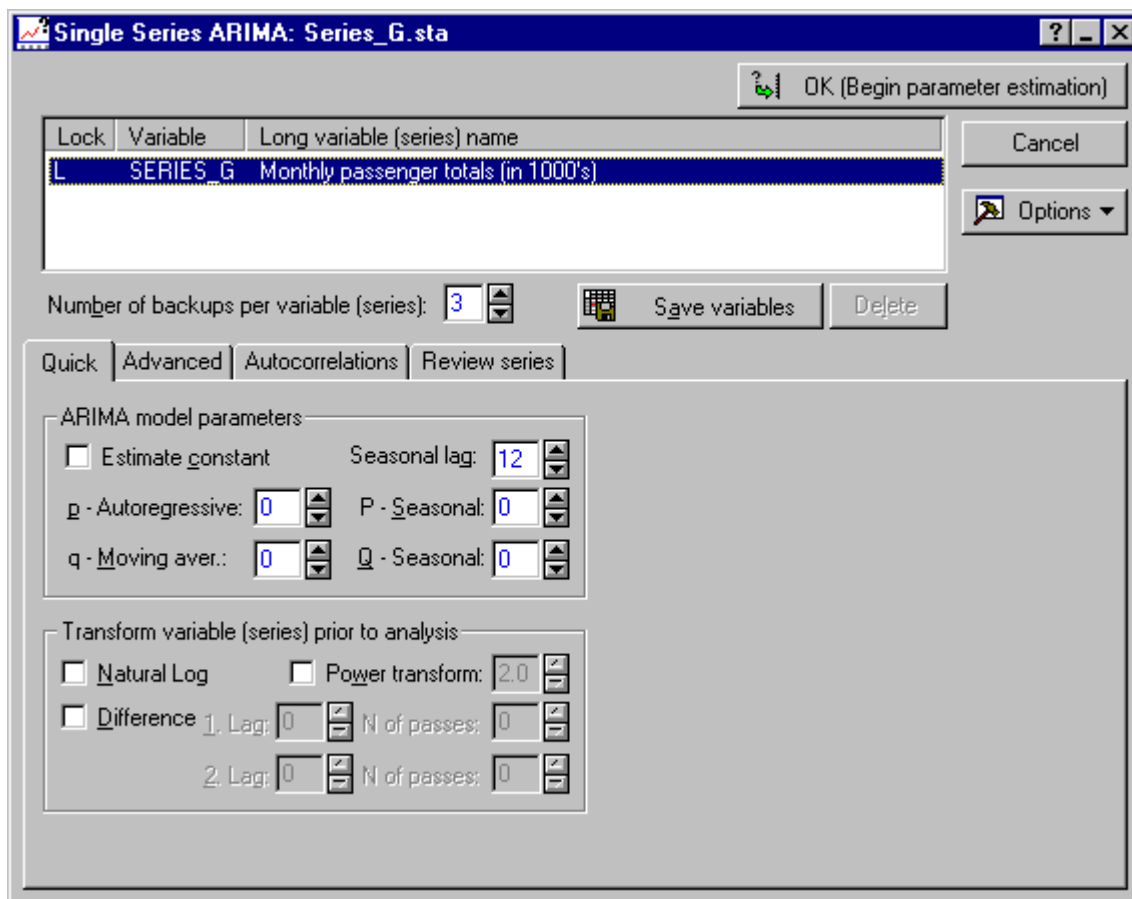
Общие сведения. Исходные данные примера представляют собой ежемесячное общее количество пассажиров (измеряемое в тысячах) международных авиаперелетов за двенадцать лет подряд: 1949-1960 гг. (см. Box and Jenkins, 1976, «Series G»). Данные частично приведены ниже; они также включены в STATISTICA StatSoft в пример файла данных Series_G.sta. Откройте этот файл данных через меню Файл - Открыть примеры (File - Open Examples); он находится в папке Datasets.



	1 SERIES_G
JAN 1949	112
FEB 1949	118
MAR 1949	132
APR 1949	129
MAY 1949	121
JUN 1949	135
JUL 1949	148
AUG 1949	148
SEP 1949	136
OCT 1949	119
NOV 1949	104
DEC 1949	118
JAN 1950	115
FEB 1950	126
MAR 1950	141

Ряд показывает четкую тенденцию к росту с годами, но в то же время в данных присутствует сильная сезонность (например, показатели за март обычно выше, чем за февраль и апрель).

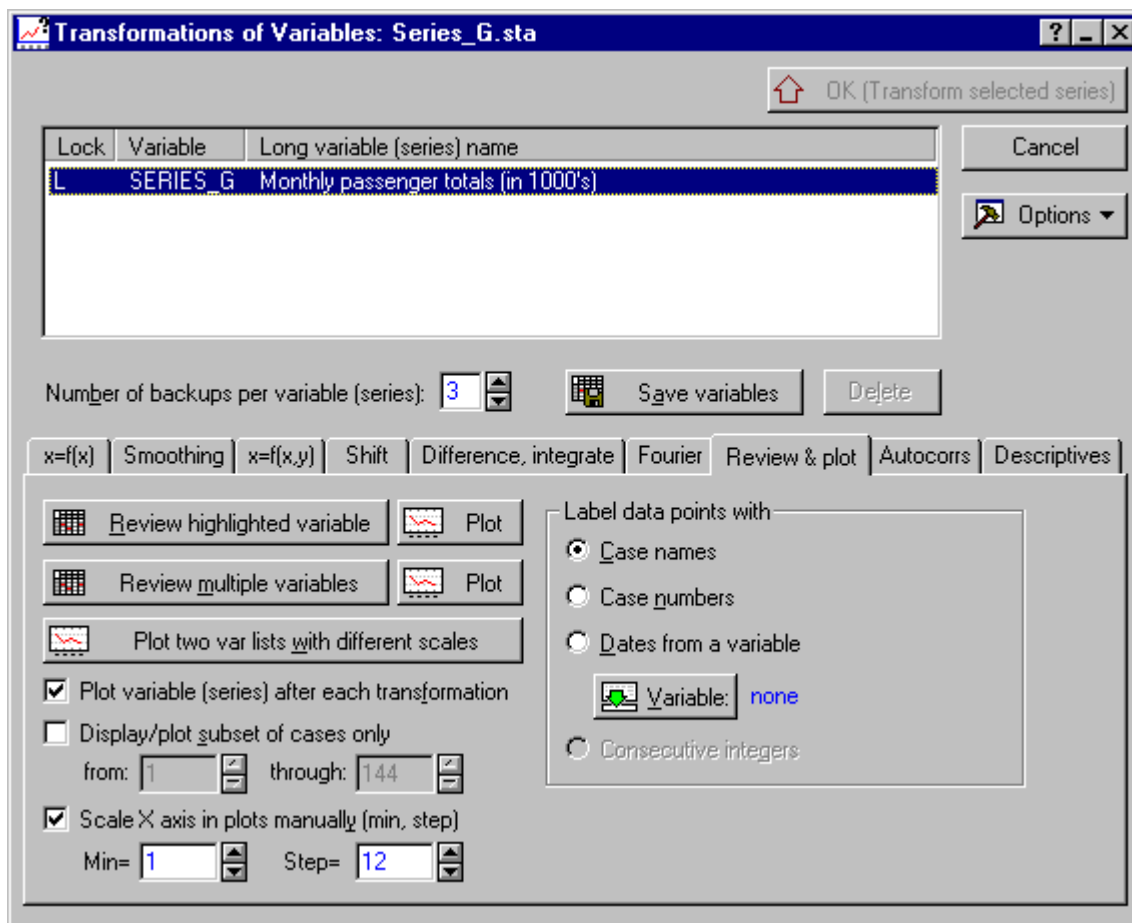
Определение анализа. Чтобы начать анализ, выберите «Временные ряды / прогнозирование» (Time Series/Forecasting) в меню «Статистика» - «Расширенные линейные / нелинейные модели» (Statistics - Advanced Linear/Nonlinear Models), чтобы отобразить панель запуска анализа временных рядов (Time Series Analysis). Затем нажмите кнопку «Переменные» (Variables), чтобы отобразить стандартный диалог выбора переменных. Здесь выберите переменную и нажмите кнопку ОК. Теперь на вкладке «Быстрая» (Quick tab) нажмите кнопку «Функции ARIMA и автокорреляции» (ARIMA & autocorrelation functions). После считывания всех 144 точек данных отображается диалоговое окно ARIMA для отдельных рядов (Single Series ARIMA).



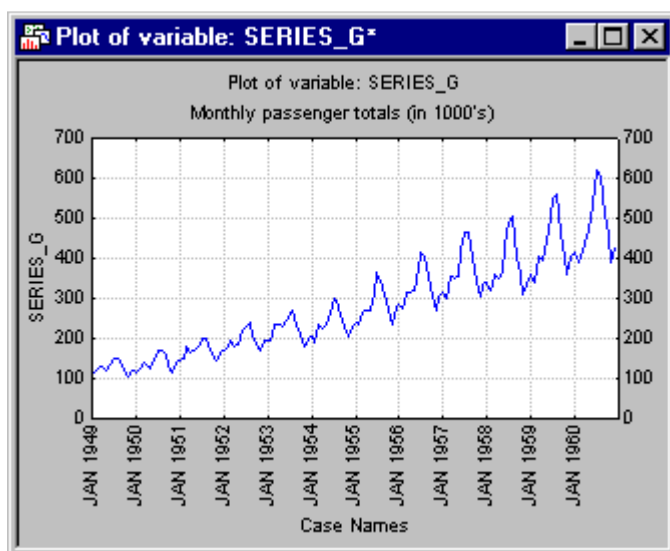
Фаза идентификации. Прежде чем вы сможете указать параметры для оценки в ARIMA, сначала необходимо идентифицировать модель. Хотя большинство необходимых параметров, таких как автокорреляция или частичная автокорреляция (Autocorrelations or Partial Autocorrelations), доступны в этом диалоговом окне, параметры преобразования будут использоваться, чтобы проиллюстрировать, как можно идентифицировать природу процесса ARIMA. Поэтому нажмите кнопку «Другие преобразования и графики» (Other transformations & plots) на вкладке «Дополнительно» (Advanced), чтобы отобразить диалоговое окно «Преобразования переменных» (Transformations of Variables).

В этом диалоговом окне сначала выберите более подходящий масштаб для горизонтальной оси x на последующих линейных графиках. На вкладке «Просмотр и построение» (Review & plot) установите флажок «Масштабировать ось X на графиках вручную» (Scale X axis in plots manually), а затем введите 1 в качестве минимума и 12 в качестве шага (12 месяцев в году). Кроме того, файл данных Series_g.sta содержит названия случаев с соответствующими датами для каждого наблюдения. Вы можете использовать эти метки для обозначения разных лет на последующих линейных графиках; Таким образом, выберите опцию «Имена наблюдений» (Case names) в разделе «Обозначить точки данных» (Label data points). Эти варианты маркировки и масштабирования горизонтальной оси x на линейных графиках являются «общими для модуля»; то есть они будут использоваться во всем модуле временных рядов всякий раз, когда запрашиваются линейные графики рядов.

Обратитесь к описанию вкладки «Просмотр и построение» (Review & plot) для подробного обсуждения того, как линейные графики могут быть масштабированы и помечены в модуле «Временные ряды» (Time Series). Вкладка Transformations of Variables - Review and Plot теперь будет выглядеть следующим образом.



Чтобы построить график временного ряда, нажмите кнопку «График» (Plot) рядом с кнопкой «Просмотр выделенной переменной» (Review highlighted variable)



И тренд, и сезонность ряда очевидны. Чтобы определить конкретные лаги для разности ARIMA, будут использоваться автокорреляции и частичные автокорреляции.

Мультипликативная сезонность. Из графика ряда также видно, что амплитуда сезонных изменений со временем увеличивается (т.е. есть свидетельства мультипликативной сезонности), что может смещать значения автокорреляций. Для стабилизации этой изменчивости будет выполнено преобразование данных в натуральный логарифм.

Логарифмическое преобразование. Щелкните вкладку $x = f(x)$ и выберите кнопку параметра Natural log ($x = \ln(x)$).

Transformations of Variables: Series_G.sta

OK (Transform selected series)

Lock	Variable	Long variable (series) name
L	SERIES_G	Monthly passenger totals (in 1000's)

Number of backups per variable (series): 3

Save variables Delete

x=f(x) Smoothing x=f(x,y) Shift Difference, integrate Fourier Review & plot Autocorr Descriptives

OK (Transform selected series)

Transformation

☐ Add a constant ($x=x+C$) C= 0.0

☐ Power ($x=x^C$) C= 2.00

☐ Inverse power ($x=x^{1/C}$) C= 2.00

☒ Natural log ($x=\ln(x)$)

☐ Exponent ($x=\exp(x)$)

☐ Mean subtract ($x=x-M$) M= 0.0

☐ Standardize ($x=(x-M)/SD$) M= 0.0 S= 1.0

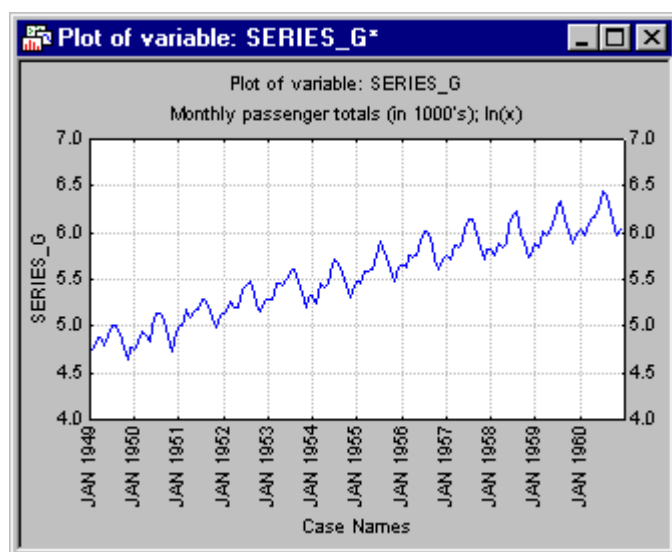
☒ Estimate mean & std. dev. from data

☐ Trend subtract ($x=x-(a+b*t)$) a= 0.0 b= 1.0

☐ Autocorr. ($x=x-(a+b*x(lag))$) a= 0.0 b= 1.0

☒ Estimate a/b from data lag= -1

Теперь нажмите кнопку ОК (Преобразовать выбранный ряд, Transform selected series). После преобразования всех наблюдений преобразованный ряд будет автоматически нанесен на график.



Обратите внимание, что желаемый эффект был достигнут, поскольку амплитуда изменений теперь достаточно стабильна, и ряд готов для дальнейшего анализа с автокорреляциями.

Автокорреляции. Теперь щелкните вкладку Autocorr. Измените значение по умолчанию для параметра «Число лагов» (Number of lags) в поле «Автокорреляции и взаимные корреляции» (Autocorrelations & crosscorrelations) с 15 на 25. Затем нажмите кнопку «Автокорреляции» (Autocorrelations), чтобы отобразить электронную таблицу с автокорреляциями и графиком автокорреляции.

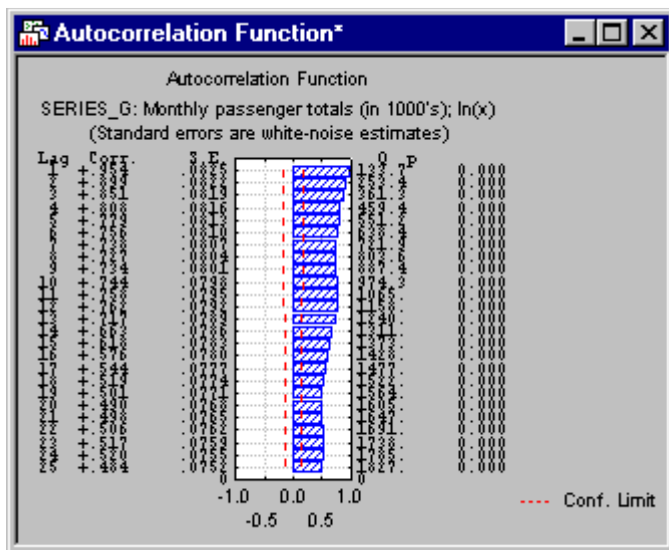
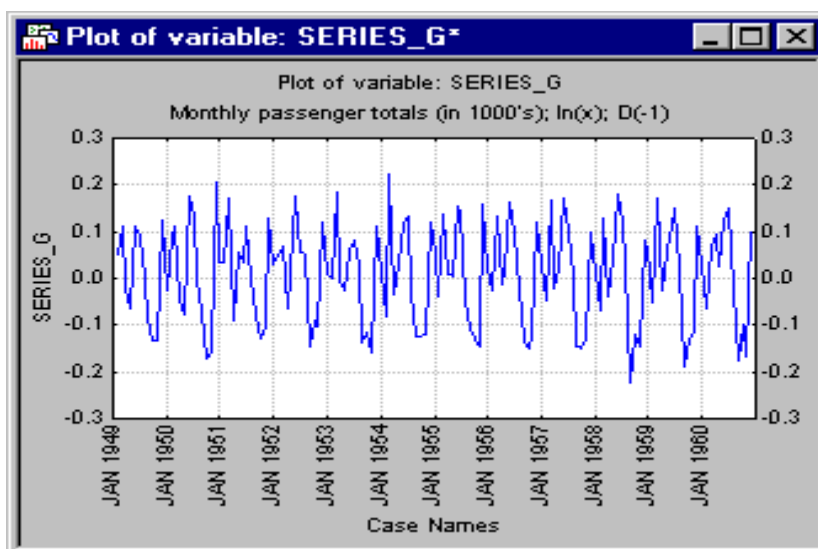


График показывает сильные последовательные зависимости для лагов от 1 до 12 с наивысшим значением автокорреляции для лага, равного единице.

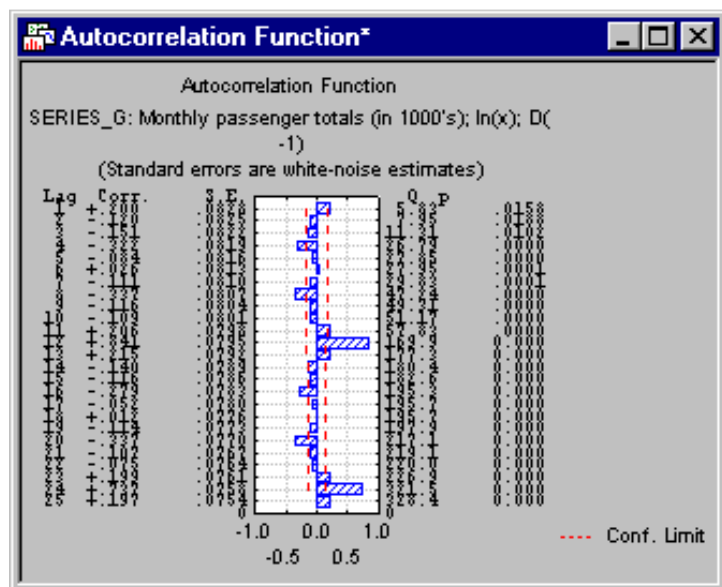
Разность. Чтобы удалить последовательную зависимость, сначала будет выполнено несезонное разностное преобразование для ряда, то есть разница с лагом, равным 1.

Обратите внимание, что преобразованный (записанный) ряд автоматически выбирается (выделяется) в активной рабочей области. Таким образом, просто щелкните вкладку «Разница, интегрировать» (Difference, integrate), выберите кнопку параметра «Разность» $x = x - x(\text{lag})$, (Differencing ($x = x - x(\text{lag})$)) (не изменяйте значение задержки по умолчанию, равное 1), и нажмите кнопку «ОК» (преобразовать выбранный ряд), ОК (Transform selected series). После преобразования всех наблюдений преобразованный (разностный) ряд снова (по умолчанию) будет построен.



Теперь каждый элемент преобразованного ряда представляет собой разницу между его исходным (т. е. предыдущим) значением и исходным значением соседнего элемента. Обратите внимание, что ряд теперь короче (на количество элементов, равное запаздыванию, т. е. 1), поскольку первый элемент ряда не может быть различен.

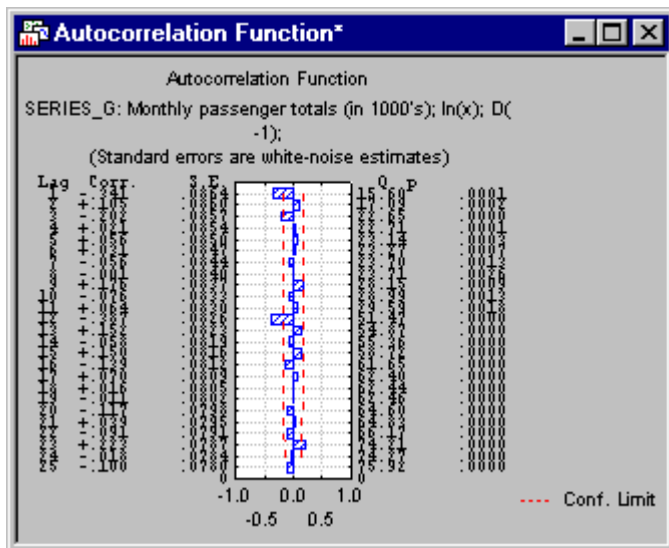
Вернитесь в диалоговое окно «Преобразования переменных» (Transformations of Variables), щелкните вкладку «Автокорреляции» (Autocorr) и снова нажмите кнопку «Автокорреляции» (Autocorrelations).



Вы видите, что после взятия разности исчезла корреляция не только на лаге 1, но также на большинстве других лагов (как объяснялось ранее, автокорреляции для последовательных лагов взаимозависимы).

Сезонность. Однако, как часто происходит, удаление зависимостей на малых лагах приводит к более отчетливой зависимости на лагах высокого порядка (в данном случае, это видно на лаге 12). Имеется также отчетливая (сезонная) зависимость на лаге 24 (и других лагах, кратных 12, таких как 36, 48 и т.д.). Это показывает сильную сезонную зависимость. Таким образом, в ряде авиаперевозок отчетливо видна сезонность.

Взятие сезонной разности. Сезонная разность с лагом 12 устранил эту зависимость. Вернитесь к вкладке «Разница, интеграция» (Difference, integrate). Снова выберите кнопку параметра «Разность» ($x = x - x(\text{lag})$) (*Differencing ($x = x - x(\text{lag})$)*), но измените значение лага на 12. Затем нажмите кнопку «ОК» (преобразовать выбранный ряд) (*OK (Transform selected series)*), чтобы выполнить разностное преобразование. Опять же, по умолчанию преобразованный ряд будет отображен в виде графика (снимите флажок «График (ряд) после каждого преобразования» (*the Plot variable (series) after each transformation*) на вкладке «Просмотр и построение» (Review & plot), если вы не хотите строить ряд после каждого преобразования). Как и прежде, верните вкладку «Автокорреляции» (Autocorr) и нажмите кнопку «Автокорреляции» (Autocorrelations).



Большинство сильных автокорреляций теперь удалено. Хотя еще остались автокорреляции, большие 2-х стандартных ошибок (показанных точечной линией на графике автокорреляций), не нужно брать еще разности ряда, т.к. они могут исключить эффект скользящего среднего.

Теперь нажмите кнопку "Частные автокорреляции" (т.е. автокорреляции, контролирующие все корреляции "в пределах лага") (Partial autocorrelations).

Параметры, подлежащие оценке. Коррелограмма выглядит хорошо, и теперь ряд готов для ARIMA. Основываясь на изучении природы ряда (т.е. на этапе идентификации ARIMA), можно прийти к выводу, что сезонная АРПСС (с лагом 12) и несезонная модель (с лагом 1) достаточно хорошо подходят к преобразованному ряду. Будут оцениваться два параметра скользящего среднего модели АРПСС: один сезонный (Q_s) и один несезонный (q). Параметры авторегрессии отсутствуют в модели.

Комплексные преобразования ARIMA. Ранее было выполнено логарифмическое преобразование данных и два типа разности (несезонная и сезонная) были взяты. Все эти преобразования уже выполнены и результаты просмотрены. Преобразованный ряд можно теперь непосредственно использовать в ARIMA. Однако в ситуациях, похожих на данную, рекомендуется анализировать исходный ряд и задать необходимые преобразования внутри ARIMA (эти преобразования будут частью спецификации ARIMA). Если вы захотите построить прогноз (после оценки параметров ARIMA), то он будет вычислен из проинтегрированных рядов ("интегрирование", более точно суммирование, в данном случае означает просто операцию, обратную взятию разностей с соответствующими лагами). Таким образом, проводя обратные преобразования, вы возвращаетесь к исходному ряду и прогноз соответствует исходным данным (что обеспечивает более легкую интерпретацию результатов).

Заметим, внутри ARIMA доступны только преобразования логарифм, возведение в степень и взятие сезонных/несезонных разностей. В некоторых случаях определенные преобразования рекомендуется выполнять до работы в ARIMA. Речь идет о преобразованиях (например, сглаживание), не изменяющих диапазон данных и к которым не нужно применять обратные преобразования.

Диалог спецификаций ARIMA. Теперь вернитесь в диалоговое окно ARIMA для отдельных рядов Single Series ARIMA, нажав кнопку «Отмена» Cancel в диалоговом окне «Преобразования переменных» Transformations of Variables. В этом диалоговом окне выделите исходную (непреобразованную) переменную Series_G. Диалоговое окно ARIMA для отдельных рядов позволяет указать параметры авторегрессии и скользящего среднего для их оценки (сезонные или несезонные). Вы не можете перейти к следующему шагу, пока не будет выбран хотя бы один из параметров авторегрессии или скользящего среднего, которые необходимо оценить (p, P, q или Q). Однако сначала вам нужно будет указать преобразования и разность.

В поле «Преобразовать переменную (ряд) до анализа» (Transform variable (series) prior to analysis) на вкладке «Быстрый» (Quick) установите флажки «Натуральный логарифм» (Natural Log) и «Разность» (Difference). Затем укажите 1. Лаг = 1 (1. Lag = 1), и число переходов равное 1 (N. of passes 1). Теперь указываете логарифмическое преобразование и несезонное простое разностное преобразование. Чтобы указать преобразование сезонной разности, укажите в поле «2. Лаг» лаг, равную 12, и снова установите для числа проходов значение 1 (N. of passes to 1.).

Параметры ARIMA. Вам по-прежнему необходимо указать параметры модели ARIMA. Вы не можете запустить ARIMA, пока не будет запрошен хотя бы один из параметров авторегрессии или скользящего среднего для оценки (p, P, q или Q). На этапе идентификации ARIMA было решено оценить два параметра скользящего среднего, один регулярный (q) и один сезонный (Q), и никаких авторегрессионных параметров. Ниже показано диалоговое окно ARIMA для отдельных рядов (Single Series ARIMA) со всеми необходимыми настройками.

Lock	Variable	Long variable (series) name
L	SERIES_G	Monthly passenger totals (in 1000's)
	SERIES_G	Monthly passenger totals (in 1000's); ln(x)
	SERIES_G	Monthly passenger totals (in 1000's); ln(x); D(-1)
	SERIES_G	Monthly passenger totals (in 1000's); ln(x); D(-1); D(-12)

Number of backups per variable (series): 3

Buttons: OK (Begin parameter estimation), Cancel, Options, Save variables, Delete

Tabs: Quick, Advanced, Autocorrelations, Review series

ARIMA model parameters

☐ Estimate constant Seasonal lag: 12

p - Autoregressive: 0 P - Seasonal: 0

q - Moving aver.: 1 Q - Seasonal: 1

Transform variable (series) prior to analysis

☒ Natural Log ☐ Power transform: 2.0

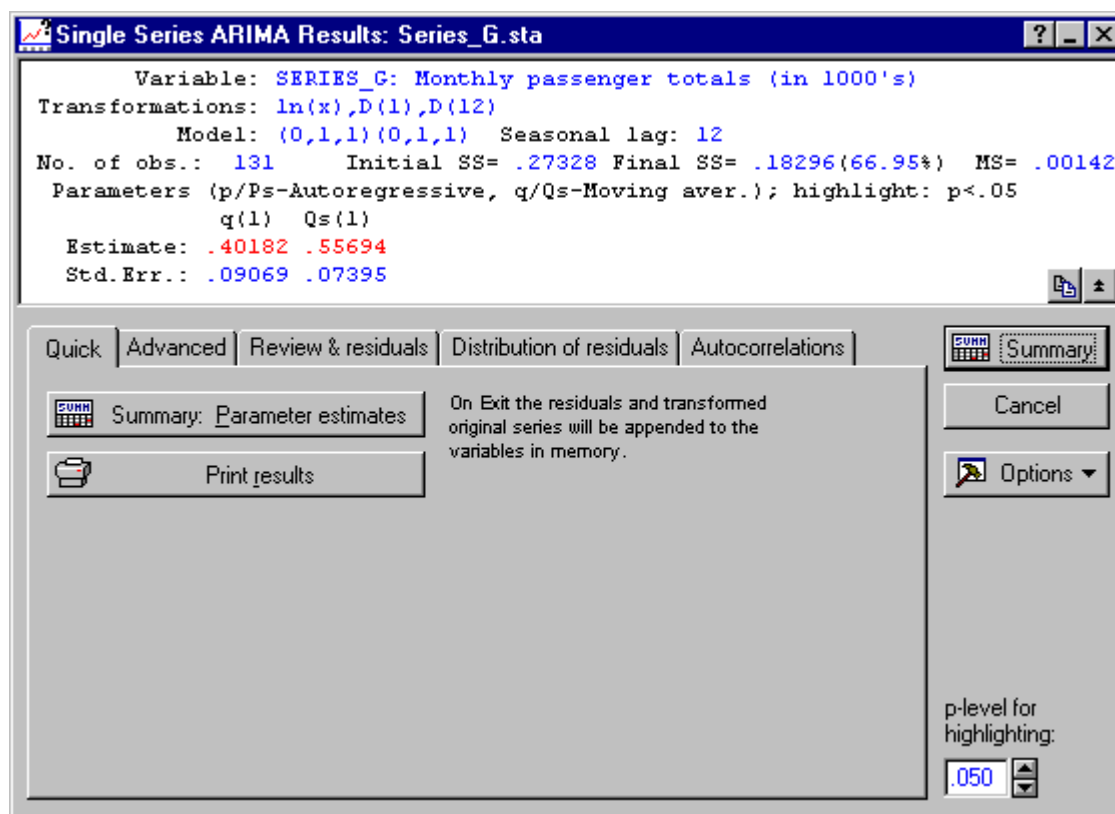
☒ Difference 1. Lag: 1 N of passes: 1

2. Lag: 12 N of passes: 1

Оценивание параметров. Параметры ARIMA оцениваются максимизацией функции правдоподобия. Доступны два метода максимизации функции правдоподобия: Приближенный (МакЛеода и Сейлза) и Точный (Меларда).

Далее нажмите ОК (Начать оценивание параметров) (OK (Begin parameter estimation)) и запустите итеративную процедуру оценивания.

Просмотр результатов. После того, как процедура оценивания сойдется, нажмите ОК и откройте диалоговое окно *Результаты одномерной ARIMA* (Single Series ARIMA Results).



Вывод ARIMA. Сводное поле в верхней части диалогового окна Результаты ARIMA для отдельных рядов (Single Series ARIMA Results) отражает спецификации ARIMA (модель ARIMA, преобразования), оценки параметров и их стандартные ошибки. Нажмите кнопку Сводка: оценки параметров (Summary: Parameter estimates), чтобы просмотреть оценки параметров в электронной таблице.

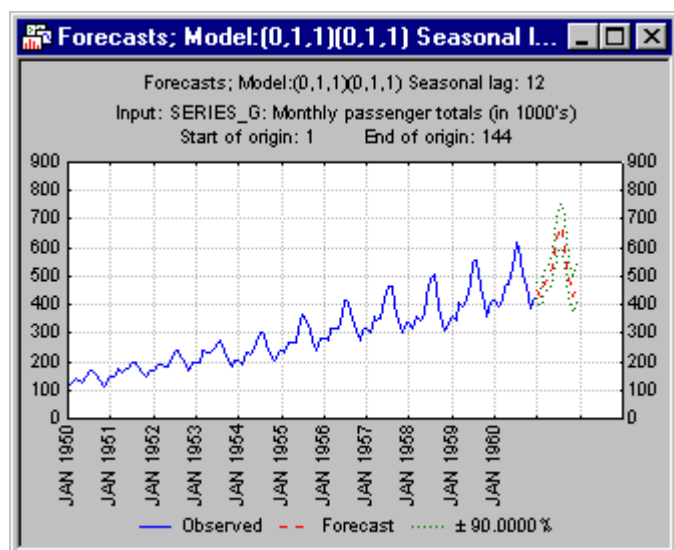
Data: Input: SERIES_G: Monthly passenger totals (in 1000's) (Series_G)*						
Input: SERIES_G: Monthly passenger totals (in 1000's) (Series_G.sta)						
Transformations: ln(x), D(1), D(12)						
Model: (0,1,1)(0,1,1) Seasonal lag: 12 MS Residual=.00142						
Paramet.	Param.	Asympt. Std.Err.	Asympt. t(129)	p	Lower 95% Conf	Upper 95% Conf
q(1)	0.401823	0.090686	4.430945	0.000020	0.222399	0.581247
Qs(1)	0.556937	0.073949	7.531376	0.000000	0.410627	0.703246

Обе оценки (сезонных и несезонных параметров) высоко значимы.

Параметры прогноза. Поле «Прогнозирование» (Forecasting) на вкладке «Дополнительно» (Advanced) содержит параметры для расчета прогнозов. По умолчанию STATISTICA будет вычислять прогнозы для одного полного сезонного цикла после последнего наблюдаемого значения, то есть после случая 144 (начиная со случая 145). Сначала взгляните на прогнозы в электронной таблице; нажмите кнопку «Прогнозировать случаи» (Forecast cases). Таблица будет содержать прогнозы и их доверительные интервалы; обратите внимание, что если бы вы запросили прогнозы для случаев, которые также наблюдались, в электронной таблице также были бы отображены наблюдаемые и остаточные значения.

Data: Forecasts; Model:(0,1,1)(0,1,1) Seasonal lag: 12 (Series_G)*			
Forecasts; Model:(0,1,1)(0,1,1) Seasonal lag: 12 (Series_G.sta)			
Input: SERIES_G: Monthly passenger totals (in 1000's)			
Start of origin: 1 End of origin: 144			
CaseNo.	Forecast	Lower 90.0000%	Upper 90.0000%
145	450.3939	423.1511	479.3906
146	425.6906	395.8395	457.7928
147	478.9716	441.3850	519.7590
148	492.3758	450.0683	538.6603
149	509.0141	461.8264	561.0232
150	583.2924	525.5728	647.3510
151	669.9440	599.7494	748.3542
152	667.0088	593.4766	749.6517
153	558.1345	493.7240	630.9478
154	497.1610	437.3513	565.1500
155	429.8310	376.1138	491.2202
156	477.1891	415.4215	548.1407

График прогнозов. Гораздо лучшую «картину» того, насколько хорошо прогнозы расширяют наблюдаемые ряды, можно получить, построив прогнозы вместе с наблюдаемыми значениями. Нажмите кнопку "Построить ряды и прогнозы" (Plot series & forecasts).

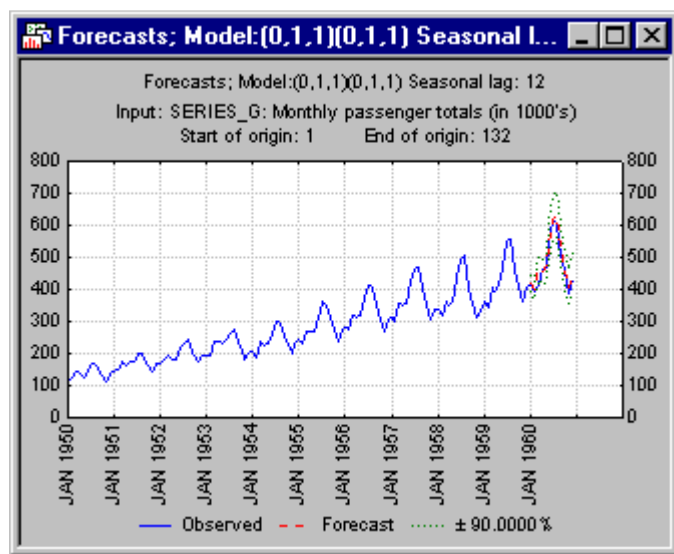


На графике отображаются исходный ряд, прогнозы и их доверительные интервалы (для вероятности, установленной в диалоговом окне «Результаты ARIMA для отдельных рядов» (Single Series ARIMA Results)); Напомним, что раньше вы

потребовали пометить точки на оси X именами наблюдений и использовали шаг 12, чтобы аккуратно отобразить последовательные годы. Просматривая график, вы видите, что построенная АРПСС довольно разумно прогнозирует ряд.

Вернитесь на вкладку «Результаты ARIMA для отдельных рядов - Дополнительно» (Single Series ARIMA Results - Advanced), чтобы увидеть, насколько хорошо текущая модель ARIMA предсказывает последние 12 наблюдаемых случаев ряда. Установите для поля «Начать со случая» (Start at case) значение 133 (т.е. $144 - 12 + 1$), а затем снова нажмите кнопку «Построить ряд и прогнозы» (Plot series & forecasts).

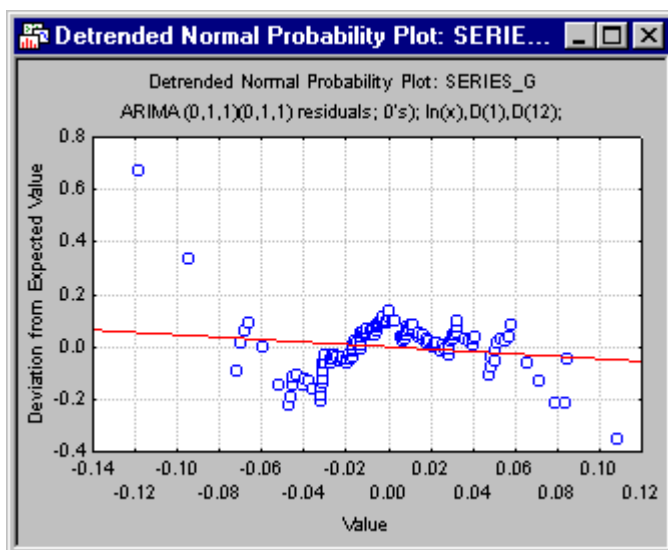
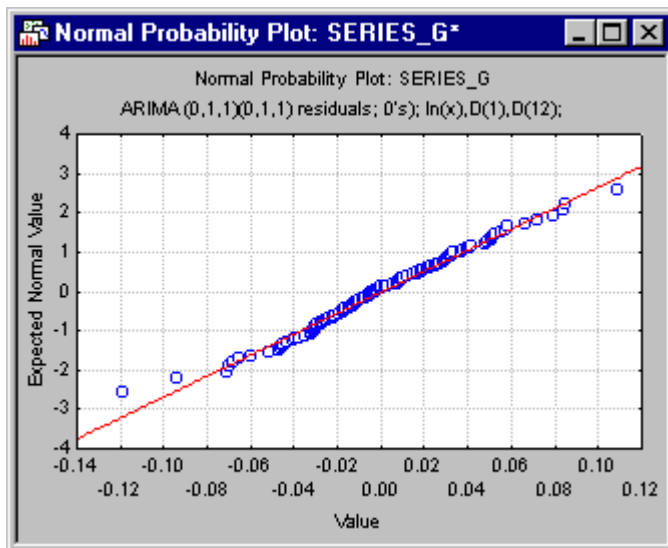
Вы видите, как STATISTICA построила прогноз.



Опять же, двухпараметрическая модель ARIMA, очевидно, очень хорошо соответствует ряду, и наблюдаемые значения находятся в пределах доверительного интервала предсказанных значений.

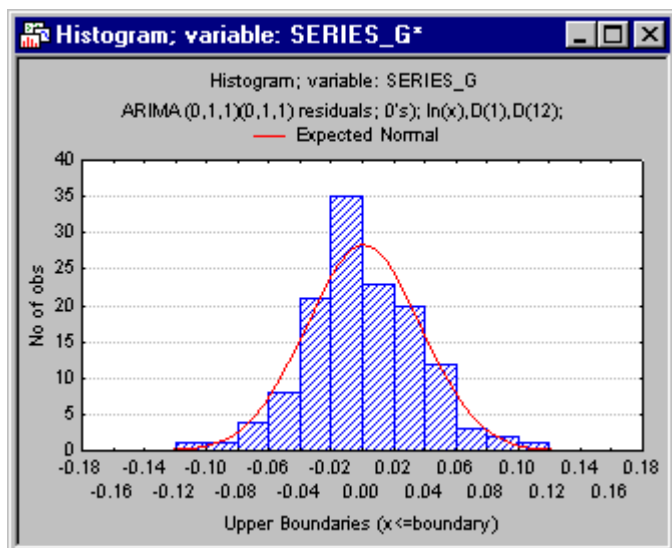
Анализ остатков. В общем, кажется, что модель достаточно адекватно подходит к данным. Однако имеются и другие важные способы оценки адекватности. Имеются два предположения модели АРПСС: 1) остатки (наблюдаемые минус оцененные значения) нормально распределены, 2) остатки независимы друг с другом, т.е. между ними нет остаточной корреляции. Если последнее условие не выполнено, то, вполне вероятно, что вы не заметили некоторый дополнительный параметр, влияющий на ряд.

Графики нормальной вероятности. Первое предположение - нормальное распределение остатков - можно проверить, исследуя графики нормальной вероятности остатков. Ниже показаны графики нормальной вероятности и нормальной вероятности без тренда (нажмите эту кнопку на вкладке «Распределение остатков» (Distribution of residuals)).

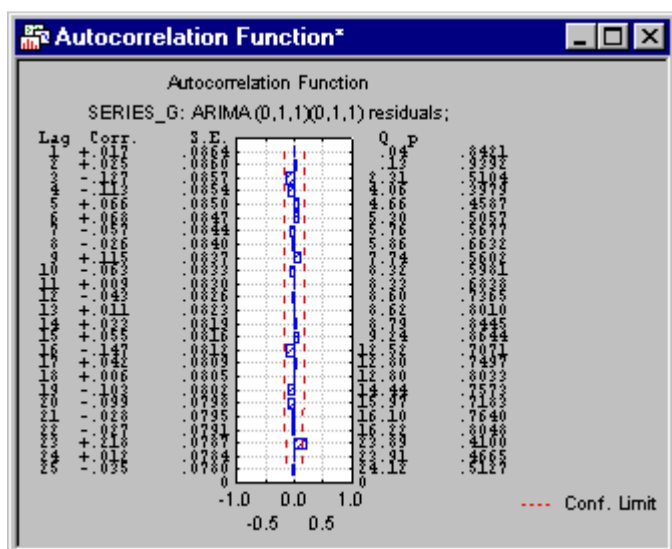


Стандартный нормальный график строится следующим образом. Сначала упорядочиваются остатки. Из этих рангов вы можете вычислить значения z (то есть стандартизованные значения нормального распределения), исходя из предположения, что остатки происходят из нормального распределения. Эти значения z отложены по оси ординат на графике. Если наблюдаемые значения (нанесенные на ось x) нормально распределены, то все точки должны попадать на прямую линию на графике, как в случае на графике, показанном выше.

График нормальной вероятности без тренда строится так же, как и график стандартной нормальной вероятности, за исключением того, что перед построением графика линейный тренд удаляется. Это часто «расширяет» график, тем самым позволяя вам легче выявлять закономерности отклонений. На графике выше большинство остатков группируются близко к горизонтальной линии, и снова похоже, что остатки действительно распределены нормально. Нажмите кнопку «Гистограмма» (Histogram), чтобы увидеть, насколько хорошо нормальное распределение соответствует фактическому распределению остатков.



Автокорреляция остатков. Теперь обратим ваше внимание на второе предположение ARIMA - что остатки независимы друг от друга. Это можно проверить, построив график функции автокорреляции (нажмите кнопку «Автокорреляции» (Autocorrelations) на вкладке «Автокорреляции» (Autocorrelations)).



На графике выше видно, что практически не остается остаточной автокорреляции после того, как вы подгоните текущую модель ARIMA к данным. Таким образом, вы можете быть удовлетворены тем, что условие независимых остатков также выполняется.

Дальнейшие анализы. Когда вы выходите из диалогового окна Результаты ARIMA для отдельных рядов (Single Series ARIMA Results), остатки ARIMA автоматически добавляются в активную рабочую область. Кроме того, если на вкладке «Дополнительно» (Advanced) установлен флажок «Добавлять прогнозы к исходным рядам при выходе» (Append forecasts to original series on Exit) (который установлен по умолчанию), в активную рабочую область будет добавлен другой ряд с исходными данными и прогнозами. Теперь выйдите из диалогового окна «Результаты ARIMA для отдельных рядов» (Single Series ARIMA Results), нажав кнопку «Отмена» (Cancel); снова отобразится диалоговое окно "ARIMA одного ряда" (Single Series ARIMA).

Single Series ARIMA: series_G.sta

OK (Begin parameter estimation)

Lock	Variable	Long variable (series) name
L	SERIES_G	Monthly passenger totals (in 1000's)
	SERIES_G	Monthly passenger totals (in 1000's); ln(x);D(1);D(12);
	SERIES_G	ARIMA (0,1,1)(0,1,1) residuals;
	SERIES_G	+forecasts, Model: (0,1,1)(0,1,1);

Number of backups per variable (series): 3

Save variables Delete

Quick Advanced Autocorrelations Review series

ARIMA model parameters

☐ Estimate constant Seasonal lag: 12

p - Autoregressive: 0 P - Seasonal: 0

q - Moving aver.: 1 Q - Seasonal: 1

Estimation method

☐ Approximate (McLeod & Sales)

Backcast cases: 0

☒ Exact (Melard)

Estimation options

Maximum number of iterations: 50

Convergence criterion (required accuracy): .0001

Max. no. of iterations for backcasting: 10

☐ User-defined start values

Transform variable (series) prior to analysis

☒ Natural Log ☐ Power transform: 2.0

☒ Difference 1. Lag: 1 N of passes: 1

2. Lag: 12 N of passes: 1

Other transformations & plots

Как видите, в активную рабочую область были добавлены и остатки, и прогнозы. Чтобы «завершить» анализ, взгляните на другой информативный сводный график. Постройте исходный ряд, прогнозы и остатки на одном графике. Такой график может выявить другие «проблемы» с вашей моделью ARIMA, например, если остатки особенно велики, а соответствие особенно плохо в одной области ряда (например, может быть трехлетний период, когда модель ARIMA предсказывает больше пассажиров международных авиакомпаний, чем наблюдалось). Поскольку значения остатков и наблюдаемых рядов (и прогнозов) несовместимы (помните, что остатки относятся к зарегистрированным рядам и рядам с двойной разностью, в то время как прогнозы генерируются в терминах немодифицированного ряда), вам необходимо щелкнуть "Постройте два списка переменных с разными масштабами" (Plot two var lists with different scales). Нажмите кнопку на вкладке «Обзор ряда» (Review series), чтобы отобразить диалоговое окно «Выбрать переменные для графика / электронной таблицы» (Select variables for plot/spreadsheet). Выберите в качестве первой переменной для построения график исходный ряд с добавленными прогнозами, а в качестве второй переменной выберите остатки ARIMA.

Select variables for plot/spreadsheet

Variables (series) to be plotted against left-Y axis

SERIES_G: Monthly passenger totals (in 1000's); ln(x);

SERIES_G: Monthly passenger totals (in 1000's); ln(x);

SERIES_G: Monthly passenger totals (in 1000's); ln(x);

SERIES_G: ARIMA (0,1,1)(0,1,1) residuals;

SERIES_G: +forecasts, Model: (0,1,1)(0,1,1);

Variables (series) to be plotted against right-Y axis

SERIES_G: Monthly passenger totals (in 1000's); ln(x);

SERIES_G: Monthly passenger totals (in 1000's); ln(x);

SERIES_G: Monthly passenger totals (in 1000's); ln(x);

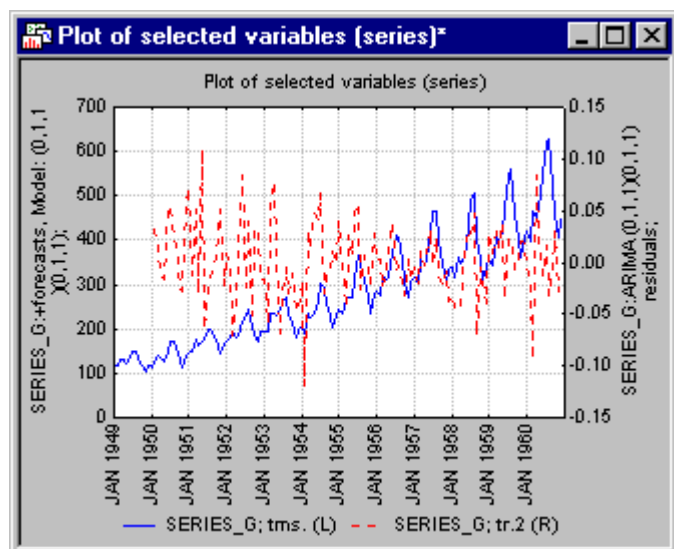
SERIES_G: ARIMA (0,1,1)(0,1,1) residuals;

SERIES_G: +forecasts, Model: (0,1,1)(0,1,1);

Select variables for the Spreadsheet/plot

OK Cancel

Теперь нажмите кнопку ОК, чтобы увидеть график.



Снова из графика видно, что подгонка модели АРПСС очень хорошая, т.к. остатки имеют примерно равную вариацию на всем протяжении ряда и нет очевидного тренда или сдвига в них.

5 References

1. Боровиков В.П. Популярное введение в современный анализ данных в системе STATISTICA. Учебное пособие для вузов. – М.: Горячая линия – Телеком, 2013. – 288 с.
2. <http://statsoft.ru/home/textbook/glossary/default.htm>
3. Box, G. E. P., & Jenkins, G. M. (1976). Time series analysis: Forecasting and control. San Francisco: Holden-Day.
4. McLeod, A. I., & Sales, P. R. H. (1983). An algorithm for approximate likelihood calculation of ARMA and seasonal ARMA models. Applied Statistics, 211-223 (Algorithm AS).
5. Melard, G. (1984). A fast algorithm for the exact likelihood of autoregressive-moving average models. Applied Statistics, 33, 104-119.

6 Контрольные вопросы

1. Цели анализа временных рядов. Систематическая составляющая и случайный шум временного ряда
2. Два общих типа компонент временных рядов
3. Понятие и анализ тренда временного ряда
4. Понятие и виды сглаживания временного ряда (уметь выполнить сглаживание на примере ряда)

5. Понятие и анализ сезонности. Понятие лага
6. Что такое автокорреляция уровней ряда. Диапазон значений автокорреляции
7. Автокорреляционная коррелограмма. Что дает исследование коррелограмм?
8. Как осуществляется удаление периодической зависимости?
9. Процесс авторегрессии. Дать определение
10. Процесс скользящего среднего. Дать определение
11. Дать определение модели авторегрессии и скользящего среднего АРПСС (ARIMA). Каковы параметры модели ARIMA (p, d, q)
12. Понятие разности ряда. Как привести ряд к стационарному виду?
13. Что определяется на этапе идентификации модели ARIMA?
14. Что выполняется на этапе оценивания модели ARIMA?
15. Какова цель анализа остатков модели ARIMA?