МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«УФИМСКИЙ ГОСУДАРСТВЕННЫЙ АВИАЦИОННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Кафедра экономики предпринимательства

МЕТОДИЧЕСКИЕ УКАЗАНИЯ К ЛАБОРАТОРНЫМ ЗАНЯТИЯМ

УЧЕБНОЙ ДИСЦИПЛИНЫ

«*ЭКОНОМЕТРИКА (ПРОДВИНУТЫЙ УРОВЕНЬ)*»

Уровень подготовки
высшее образование - магистратура
(высшее образование - бакалавриат; высшее образование – специалитет, магистратура)

Направление подготовки (специальность)
38.04.01 Экономика
(код и наименование направления подготовки, специальности)

Направленность подготовки (профиль, специализация)
Экономика и финансы фирмы
(наименование профиля подготовки, специализации)

Квалификация (степень) выпускника
_____магистр_____

Составители: Орлова Е. В.

Лабораторная работа по теме 3. Построение моделей авторегрессии, авторегрессии и скользящего среднего в экономических временных рядах.

Краткие теоретические сведения

Ряд X_t, t=1…n называется **строго стационарным**, если его характеристики – среднее значение, дисперсия и ковариация не изменяются при изменении начала отсчета времени. Будем обозначать: μ, σ^2 – средняя и дисперсия временного ряда X_t. **Коррелограмма** стационарных рядов является четной функцией, т.е. $\rho(\tau)=\rho(-\tau)$.

Исследование **стохастических** временных рядов (значения которых являются случайными величинами) сводится к построению одной из следующих моделей, которые можно использовать для описания развития во времени реальных социально-экономических систем.

Модель белого шума ε_t. Процессом белого шума называется стационарный временной ряд X, для которого справедливы следующие условия: $\mu=0$, $\sigma^2>0$, $\rho(\tau)=0$ при $\tau\neq 0$, где μ – средняя величина случайного процесса. Т.е. при $t\neq s$ случайные величины x_t и x_s, соответствующие наблюдениям белого шума в моменты времени t и s, некоррелированы.

Модель авторегрессии порядка p – AR(p) определяется как:

$$x_t=\alpha_1 x_{t-1}+\alpha_2 x_{t-2}+\ldots+\alpha_p x_{t-p}+\varepsilon_t,$$

где ε_t – процесс белого шума, x_0 – некоторая случайная величина, $\alpha_p\neq 0$, $\operatorname{cov}(x_{t-s},\varepsilon_t)=0$ для всех $s>0$. То есть наблюдаемое значение ряда получается путем линейной комбинации p предшествующих значений этого ряда плюс некоррелированная с этими предшествующими значениями случайная составляющая ε_t, отражающая обновляемую (инновационную) информацию о состоянии исследуемой системы на момент времени t, влияющую на значение x_t.

Процесс скользящего среднего порядка q $MA(q)$:

$$x_t=\varepsilon_t+b_1\varepsilon_{t-1}+b_2\varepsilon_{t-2}+\ldots+b_q\varepsilon_{t-q},\quad b_q\neq 0,$$

где ε_t - процесс белого шума.

Смешанный процесс авторегрессии – скользящего среднего ARMA (p,q):

$$X_t=\sum_{j=1}^{p}\alpha_j x_{t-j}+\sum_{j=0}^{q}b_j\varepsilon_{t-j},\ \alpha_p\neq 0,\ \ b_q\neq 0,$$

где ε_t – процесс белого шума, b_0=1.

Модель авторегрессии-проинтегрированного скользящего среднего ARIMA(p,k,q) предназначена для описания нестационарных временных рядов X_t, может быть записана в виде:

$$x_t^k=\alpha_1 x_{t-1}^k+\alpha_2 x_{t-2}^k+\ldots+\alpha_p x_{t-p}^k+\varepsilon-b_1\varepsilon_{t-1}-b_2\varepsilon_{t-2}-\ldots-b_q\varepsilon_{t-q},$$

где $x_t^k=\Delta^k x_t$, k – порядок разности модели; p – число авторегрессионых членов; q – число членов скользящего среднего; ε – случайная компонента; b, α – параметры скользящего среднего и авторегрессии.

Идентификация модели временного ряда означает определение параметров этой модели. Если временной является нестационарным, то сначала его необходимо привести к стационарному виду, например путем взятия разностей, логарифмического преобразования и др. приемов. Число разностей, которые необходимо взять чтобы достижения стационарности, определяются параметром k. Если имеется медленное убывание выборочных коэффициентов автокорреляции в зависимости от лага, обычно берут разность первого порядка.

Основными инструментами идентификации порядка модели являются графики автокорреляционной функции (АКФ), частной автокорреляционной функции (ЧАКФ). На практике большинство встречающихся временных рядов можно с достаточной степенью точности аппроксимировать одной из пяти основных моделей:

1) один параметр авторегрессии (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.

Мультипликативная сезонная *ARIMA* представляет собой обобщение обычной модели *ARIMA* на ряды, в которых имеется периодическая сезонная компонента. В дополнении 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 (P), сезонная разность (K), сезонное скользящее среднее (Q). Полная сезонная *ARIMA* может быть записана как: *ARIMA* (p,$к$,q)(P,K,Q). 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.

Применение методов эконометрического моделирования временных рядов в программе Statistica

Для анализа рядов динамики в программе *Statistica* предусмотрена функция *Statistics / Advanced Liner/Nonlinear Models / Time Series/Forecasting*. Общее назначение модуля – построить модель, описывающую временной ряд, сгладить его, спрогнозировать будущие его значения. Экранная форма модуля представлена на рис. 1.

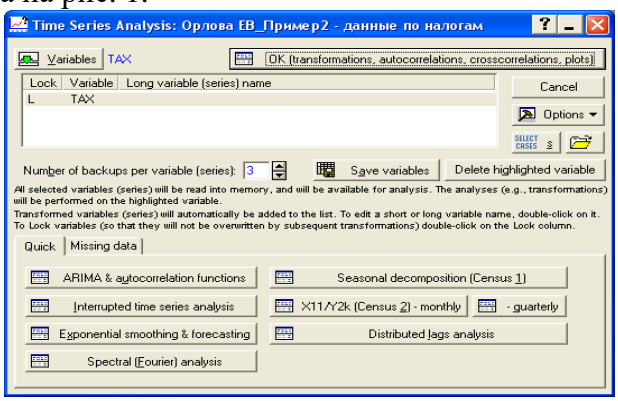

Рис.1. Основное окно модуля *Time Series/Forecasting*

Рассмотрим пример моделирования налоговых поступлений в бюджет [2] с помощью динамических методов эконометрического моделирования, в которых уровни ряда Y налоговых поступлений рассматриваются как функция тенденции T, периодических сезонных S и случайных E колебаний.

Эконометрическое моделирование и прогнозирование временного ряда осуществляется по следующим этапам: 1) идентификация модели, описывающей наблюдаемый временной ряд; 2) оценка параметров модели; 3) исследование адекватности модели; 4) прогнозирование; 5) оценка точности модели.

В результате предварительного анализа различных статистических моделей по критерию точности результатов прогнозирования выбрана одна из моделей в классе адаптивных методов – модель АРПСС (*ARIMA*).

1. *Идентификация модели*. Модель *ARIMA* $(p, d, q)(P_s, d_s, Q_s)$ имеет следующие параметры: d – порядок разности, p – порядок АР, q–порядок СС; сезонные параметры: сезонный параметр АР – P_s, сезонная разность – d_s, сезонный параметр СС – Q_s. Идентифицировать модель – значит определить эти параметры. Основным критерием идентификации является поведение автокорреляционный функции и частной автокорреляционной функции ряда.

Сначала необходимо отобразить исходный ряд данных и провести его графический анализ. Для этого в поле *Variables* выбрать исследуемая переменная, затем выбрать функцию *ARIMA & autocorrelation function.* Для просмотр ряда используется вкладка *Review series,* для построения автокорреляционных функций – *вкладка Autocorrelation.*

Наличие тренда, который хорошо виден на графике исходного ряда и отсутствие тенденции к затуханию, рис.2, 3, свидетельства нестационарности ряда.

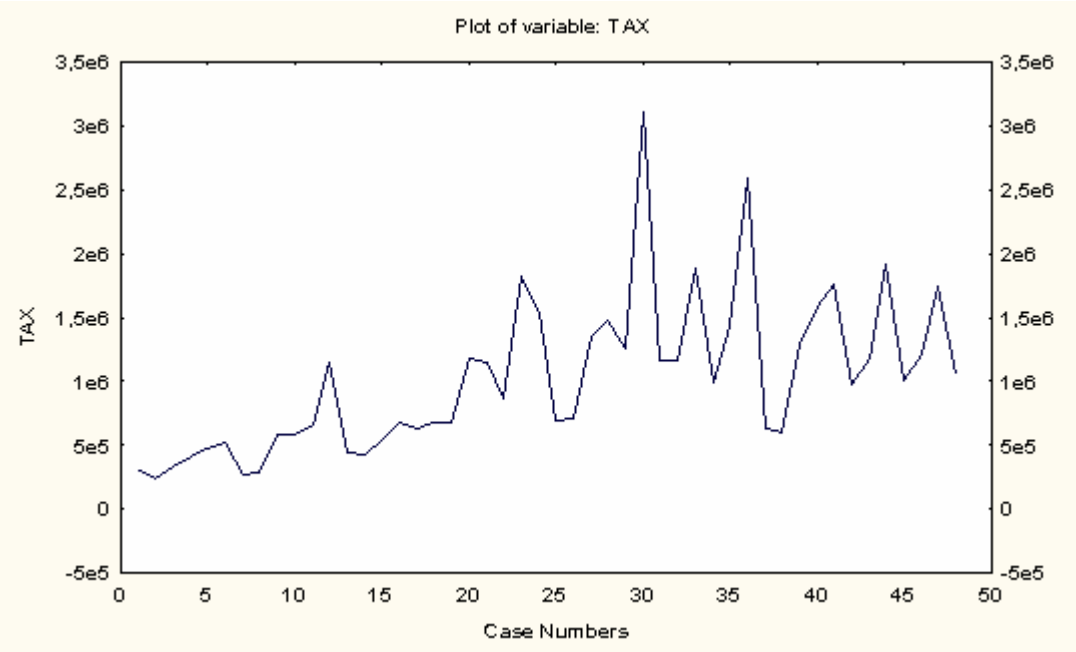

Рис. 2. Исходный ряд данных налоговых поступлений

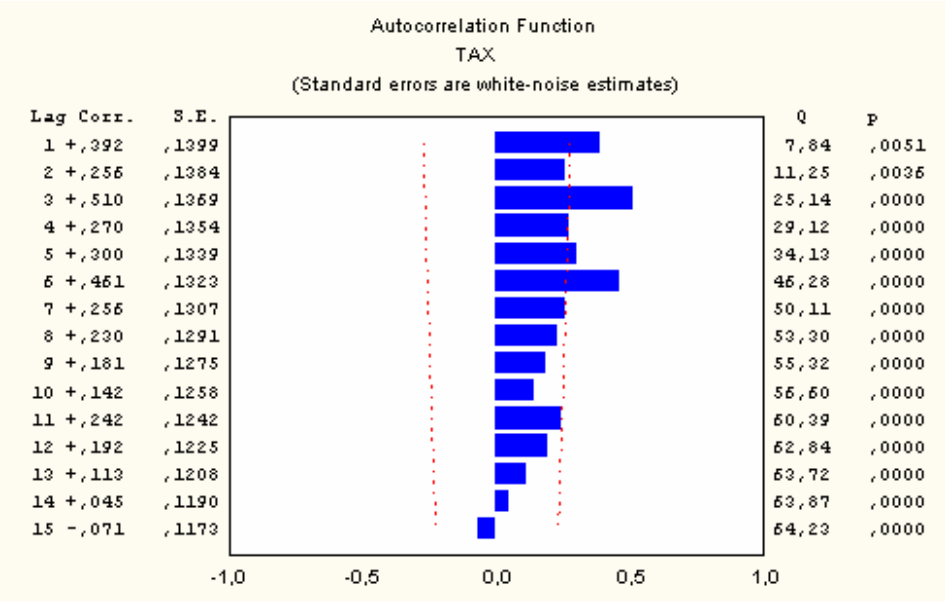

Рис. 3. Автокорреляционная функция исходного временного ряда

Для удаления тренда используем функцию взятия первых разностей, рис. 4. Для всех преобразований ряда используется функции *OK* (*transformations, autocorrelations…*) и *OK* (*transform selected series*).

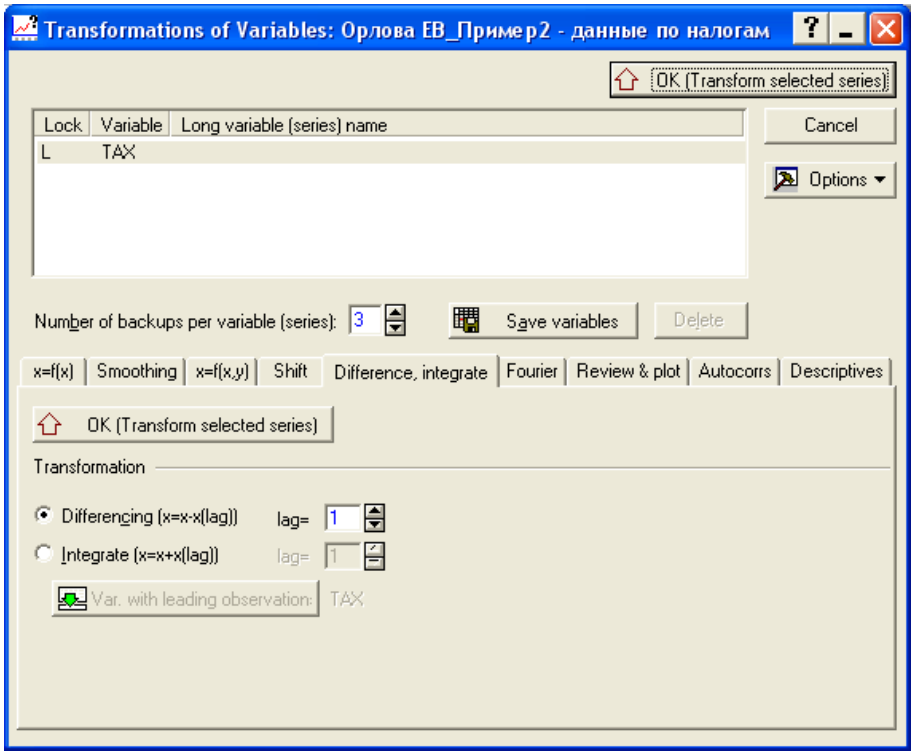

Рис. 4. Реализация функции «взятие разности»

Опять применив к полученному ряду критерия стационарности, можно прийти к заключению, что ряд первых разностей – стационарный, рис 5.

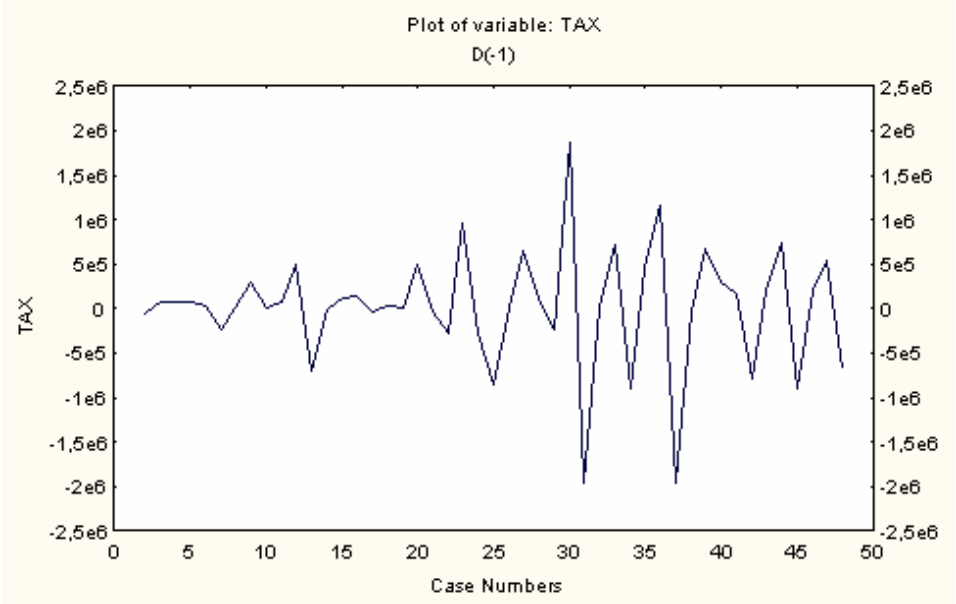

Рис. 5. Ряд первых разностей

На рис. 6 показаны автокорреляционные и частные автокорреляционные функции преобразованного ряда. В соответствии с критериями, описанными в теоретической части методических указаний, можно идентифицировать исследуемый ряд как процесс скользящего среднего СС(2) – p=0, q=2.

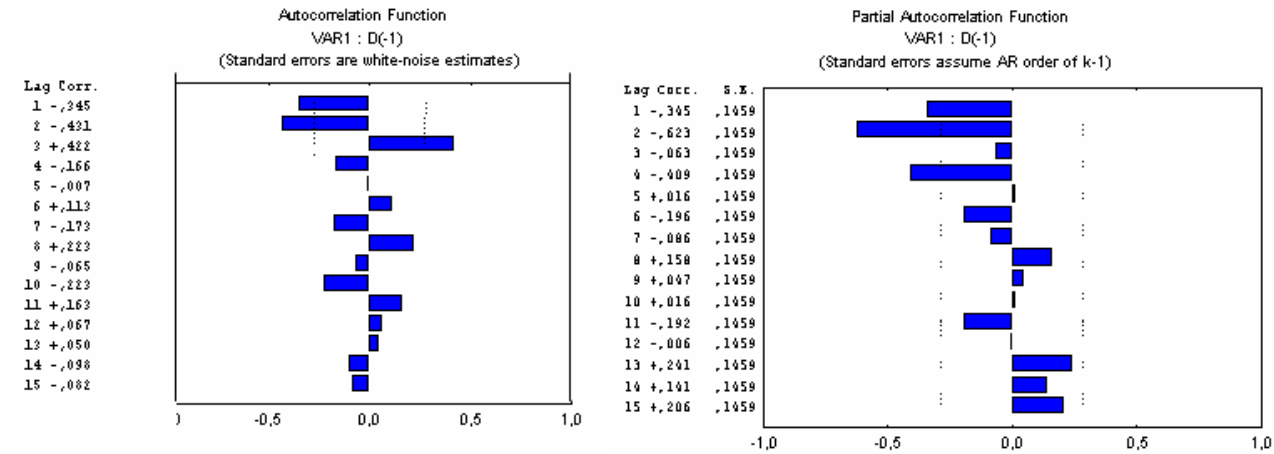

Рис. 6. Автокорреляционная и частная автокорреляционная функции

Определение наличия сезонной компоненты проводится с использованием спектрального анализа, для этого используется функция – *Time Series Analysis / Spectral (Fourier) analysis / Single Fourier Analysis / Periodogram.* Для выявления сезонной компоненты исследуются пики спектральной плотности на периодограмме. На рис. 7 видно, что в исходном временном ряде имеется трехмесячный цикл.

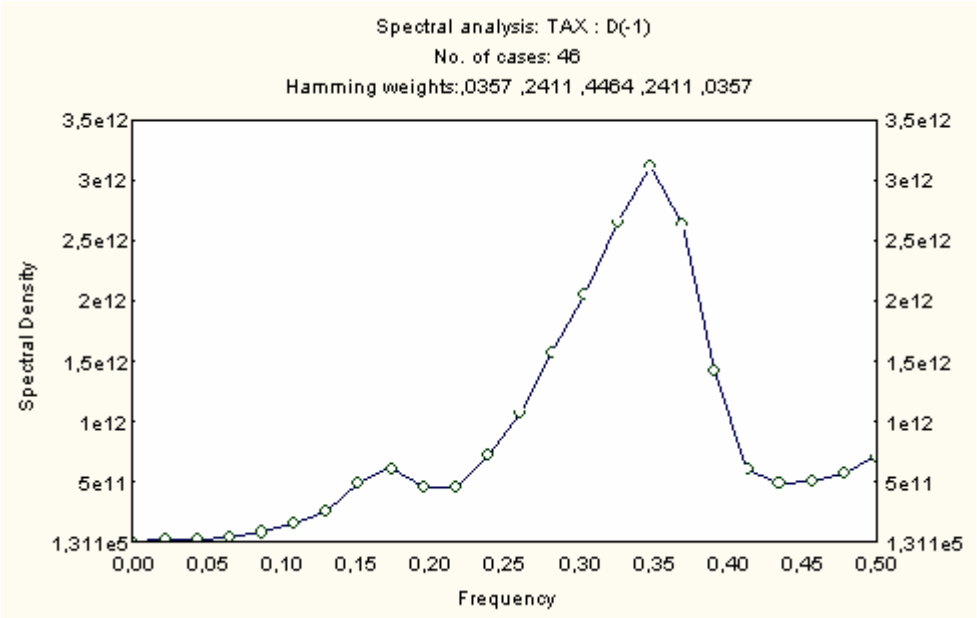

Рис. 7. График оценки спектральной плотности

В точке со значением частоты равной 0,33 спектральная плотность максимальна. Следовательно, период = 1/частота = 3. Наличие сезонной составляющей с данным периодом связано с квартальными налоговыми платежами. Для удаления сезонной составляющей необходимо взять разность третьего порядка.

2. *Определение параметров* P_s, Q_s аналогично определению параметров p, q. Получены следующие значения параметров: d=1, p = 0, q = 2, d_s=1, Q_s = 1, P_s = 0. Таким образом, модель представляется в виде *ARIMA* (0,1,2)(0,1,1).

3. На следующем этапе *оценки параметров* идентифицированной модели получают точечные оценки неизвестных параметров $q(1)$, $q(2)$, $Q_s(1)$, стандартная ошибка, верхняя и нижняя границы доверительного интервала. Для этого используется функция *ARIMA & autocorrelation function,* где выставляются порядки всех параметров, рис. 8.

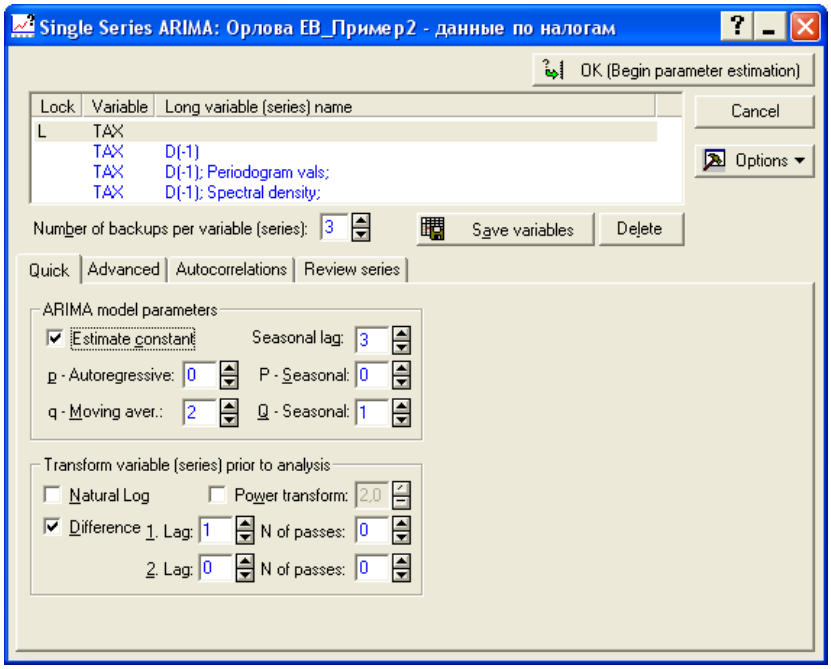

Рис. 8. Окно для оценки параметров модели

Ширина интервала между нижней и верхней границами 95 % доверительного интервала мала, так же как и стандартная ошибка, что является показателем качества оценки параметров модели, табл. 1.

Таблица 1

Результаты оценки параметров модели

Параметр	Значение параметра	Asympt. Std.Err.	Lower 95 % Conf	Upper 95 % Conf
q(1)	0,823273	0,000014	0,823245	0,823302
q(2)	0,176725	0,000003	0,176719	0,176731
Qs(1)	0,534129	0,175346	0,180011	0,888246

4. Для исследования адекватности модели исследуются остатки, представляющие собой разности фактических значений налоговых поступлений и значений, рассчитанных с помощью построенной модели. По автокорреляционный функции и частной автокорреляционной функции остатков можно сделать вывод, что ряд остатков почти не отличается от белого шума. Кроме этого, распределение остатков подчиняется нормальному закону распределения, рис. 9. Таким образом, модель адекватна и вполне пригодна для прогнозирования.

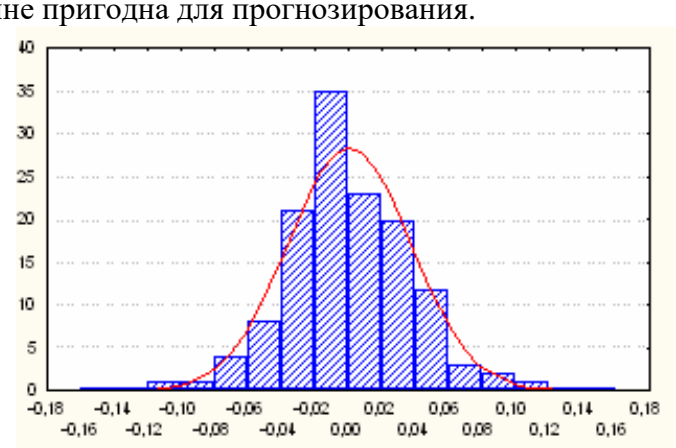

Рис. 9. График распределения остатков регрессионной модели

5. Результатом прогнозирования является как точечный, так и интервальный прогноз с доверительной вероятностью 0,95, рис. 10.

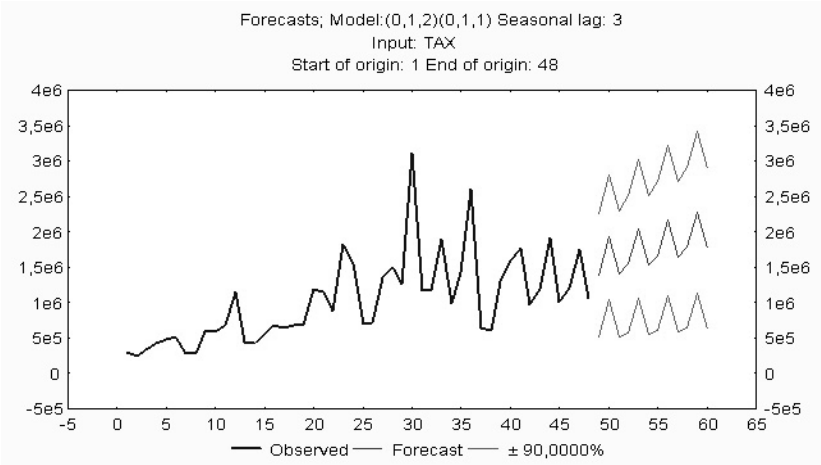

Рис.10. График прогноза налоговых поступлений в бюджет

Примеры формализованного представления моделей ARIMA

Общий вид модели *ARIMA*(1,1,0) по первым разностям можно представить как $\Delta^1 y_t = \alpha\Delta^1 y_{t-1}$, где $\Delta^1 y_t = y_t - y_{t-1}$, $\Delta^1 y_{t-1} = y_{t-1} - y_{t-2}$, …, $\Delta^1 y_n = y_n - y_{n-1}$ – первые разности.

Модель *ARIMA*(1,1,0), где $\alpha = -0{,}31$ имеет вид: $\Delta^1 y_t = -0{,}31\Delta^1 y_{t-1}$, тогда $y_t - y_{t-1} = -0{,}31(y_{t-1} - y_{t-2})$, откуда $y_t = 0{,}69\, y_{t-1} - 0{,}31\, y_{t-2}$ – искомая модель регрессии.

Модель *ARIMA*(1,1,1), где $\alpha = 0{,}02, \beta = 0{,}79$ имеет вид: $\Delta^1 y_t = 0{,}02\Delta^1 y_{t-1} + \varepsilon_t - 0{,}79\varepsilon_{t-1}$, тогда $y_t - y_{t-1} = 0{,}02(y_{t-1} - y_{t-2}) + \varepsilon_t - 0{,}79\varepsilon_{t-1}$,

откуда $y_t = 1{,}02\, y_{t-1} - 0{,}02\, y_{t-2} + \varepsilon_t - 0{,}79\varepsilon_{t-1}$ – искомая модель регрессии.

Модель *ARIMA*(0,1,1), где $\beta = -0{,}79$ имеет вид: $\Delta^1 y_t = \varepsilon_t + 0{,}79\varepsilon_{t-1}$, тогда $y_t - y_{t-1} = \varepsilon_t - 0{,}79\varepsilon_{t-1}$, откуда $y_t = y_{t-1} + \varepsilon_t - 0{,}79\varepsilon_{t-1}$ – искомая модель регрессии.

Порядок выполнения лабораторной работы

В соответствии с вариантом задания провести эконометрический анализ и моделирование временного ряда, включающий следующие задачи.

1. Определить тип модели временного ряда (белый шум, стационарный, нестационарный).
2. На основе анализа автокорреляционных и частных автокорреляционных функций провести идентификацию регрессионной модели.
3. Определить параметры порядка модели тренда.
4. Определение наличия сезонной компоненты.
5. Оценить параметры идентифицированной модели временного рядя.
6. Провести анализ адекватности модели.
7. Осуществить прогнозирование временного ряда на 2-3 периода вперед по построенной модели.

Задания для лабораторной работы

Задание. По статистическим данным, приведенным в табл.3, в соответствии с вариантом задания, провести компонентный анализ временного ряда, построить аппроксимирующую его модель, оценить адекватность построенной модели, получить прогноз на *n* шагов вперед. Варианты заданий приведены в табл. 2.

Таблица 2

Распределение заданий по вариантам

Номер варианта	Компания
1	Аэрофлот

2	Газпром
3	Камаз
4	МТС
5	НМК
6	Норникель

Таблица 3

Динамика цен акции компаний

Дата	Аэрофлот	Газпром	Камаз	МТС	НМК	Норникель
10.12.2009	44,05	166,09	68,70	224,25	91,54	4 112,51
09.12.2009	44,30	167,01	68,25	224,90	94,19	4 117,82
08.12.2009	43,39	167,20	68,18	225,50	95,20	4 178,00
07.12.2009	44,10	166,00	69,01	226,00	97,30	4 122,00
04.12.2009	44,64	168,55	68,72	229,50	97,51	4 227,13
03.12.2009	44,00	163,28	68,72	225,00	94,02	4 177,00
02.12.2009	45,00	167,50	68,24	229,37	91,20	4 228,06
01.12.2009	45,13	168,59	68,02	228,00	89,34	4 090,00
30.11.2009	43,70	166,49	67,50	219,30	87,60	4 049,83
27.11.2009	43,95	169,00	68,89	218,83	89,02	3 980,00
26.11.2009	43,70	167,19	70,32	215,93	86,31	3 870,00
25.11.2009	45,74	171,99	73,10	222,21	89,50	3 985,00
24.11.2009	46,20	175,91	74,50	223,88	90,11	3 981,82
23.11.2009	46,50	179,20	70,47	229,55	92,05	4 120,12
20.11.2009	46,01	177,99	69,99	226,51	87,59	4 107,00
19.11.2009	46,50	180,70	70,39	223,44	84,75	4 106,96
18.11.2009	45,52	184,66	73,10	230,89	86,27	4 250,92
17.11.2009	44,15	183,15	72,64	224,30	85,70	4 195,85
16.11.2009	45,02	186,39	66,90	225,67	85,00	4 240,00
13.11.2009	42,97	178,90	65,36	220,80	80,47	4 070,00
12.11.2009	41,80	179,35	66,01	220,30	81,36	4 074,00
11.11.2009	42,90	181,20	66,25	217,68	82,33	4 115,00
10.11.2009	42,17	183,31	66,85	215,58	81,80	4 155,00
09.11.2009	42,00	181,82	63,20	214,98	80,01	4 091,40
06.11.2009	40,89	174,95	62,49	204,11	76,19	3 898,97
05.11.2009	40,20	180,00	64,40	205,50	76,60	3 965,00
03.11.2009	39,35	174,18	62,22	199,50	74,98	3 826,15
02.11.2009	40,67	180,99	64,94	206,00	80,00	3 947,17
30.10.2009	39,88	175,90	70,23	203,49	76,40	3 861,10
29.10.2009	42,40	184,10	62,21	208,78	78,35	4 042,03
28.10.2009	42,29	178,59	60,89	204,00	73,32	3 931,36
27.10.2009	43,00	189,73	63,28	216,50	81,60	4 230,77
26.10.2009	44,65	193,80	64,10	224,88	84,90	4 245,00
23.10.2009	45,20	193,50	64,66	230,49	85,03	4 298,00
22.10.2009	45,35	192,23	62,90	231,60	85,50	4 301,00
21.10.2009	46,10	197,00	63,40	222,57	85,67	4 297,13
20.10.2009	46,40	196,49	64,00	218,00	88,15	4 166,00
19.10.2009	44,52	195,82	64,30	217,19	87,38	4 119,92
16.10.2009	42,85	191,50	62,90	212,74	83,15	3 923,00
15.10.2009	44,60	196,01	64,72	212,40	83,97	4 050,00
14.10.2009	45,55	200,45	67,20	215,80	81,30	4 111,22
13.10.2009	43,53	190,50	66,78	210,00	78,00	3 915,00

Дата	Аэрофлот	Газпром	Камаз	МТС	НМК	Норникель
12.10.2009	43,15	192,94	69,00	212,12	79,99	4 096,96
09.10.2009	43,30	181,61	69,93	208,30	79,00	3 882,00
08.10.2009	43,45	178,81	69,90	210,50	80,85	3 858,88
07.10.2009	43,79	174,15	67,37	207,51	80,00	3 680,01
06.10.2009	43,91	175,10	68,99	207,70	82,00	3 704,90
05.10.2009	43,40	170,50	68,68	204,50	79,12	3 578,01
02.10.2009	44,01	171,01	68,05	197,00	78,15	3 583,05
01.10.2009	44,96	175,97	70,00	201,98	80,00	3 783,36
30.09.2009	45,50	175,00	70,50	203,00	76,22	3 732,50
29.09.2009	44,10	179,49	68,47	202,79	75,00	3 764,70
28.09.2009	44,00	180,59	69,60	202,60	74,46	3 801,99
25.09.2009	42,49	177,80	61,18	200,50	74,41	3 766,79
24.09.2009	40,90	177,10	62,00	199,89	74,10	3 651,00
23.09.2009	40,00	180,18	65,10	200,03	73,52	3 788,00