

УДК 336.02

Протасов Ю.М., Юров В.М.*Московский государственный областной университет*

МОДЕЛИРОВАНИЕ СЕЗОННЫХ И ЦИКЛИЧЕСКИХ КОЛЕБАНИЙ ОБЪЕМОВ ПРОДАЖ КОМПАНИИ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ГАРМОНИЧЕСКОГО АНАЛИЗА В MS EXCEL

Аннотация. В статье рассматривается компьютерная технология моделирования временных рядов с сезонными и циклическими колебаниями с использованием инструментов Excel. Для моделирования используются модель тренда и модели гармонического анализа. В качестве примера рассматривается временной ряд квартальных объемов продаж компании за восемь лет. При моделировании использовались все необходимые статистические процедуры, требуемые для идентификации и оценки параметров модели и проверки ее адекватности и точности.

Ключевые слова: временной ряд, коэффициенты Фурье, гармонический анализ, периодограмма, программа MS Excel.

Y. Protasov , V. Yurov*Moscow State Regional University*

MODELING SEASONAL AND CYCLICAL FLUCTUATIONS OF COMPANY'S SALES BY USING MS EXCEL HARMONIC ANALYSIS

Abstract. The article considers a computer technology of modeling time series with seasonal and cyclical fluctuations with the usage of MS EXCEL. For this procedure a pattern of harmonic analysis and a model of trend are used. As an example the time series of quarterly sales of the company are considered for the period of eight years. In modeling all necessary statistical procedures required for identification and evaluation of the model parameters as well as verification of its adequacy and accuracy have been carried out.

Key words: time series, Fourier coefficients, harmonic analysis, periodogram, trend, seasonality, cyclical component, significance of the model coefficients.

Для анализа поведения временных рядов с периодическими колебаниями и построения математических моделей, описывающих это поведение, широко используют гармонический (спектральный) анализ [5]. Для его применения в практической деятельности широким кругом экономистов, а также преподавателями и студентами в учебном процессе, необходимо

представлять процесс построения математической модели сезонных и циклических колебаний с использованием методов гармонического анализа в программе Excel.

В общем случае временной ряд Y_t можно представить состоящим из четырех различных компонент: тренда T_t , сезонной компоненты S_t , циклической компоненты C_t и случайной, нерегулярной компоненты ε_t . Здесь t -

момент времени. Тренд – компонента, представляющая основной рост (или спад) во временном ряду. Сезонная компонента обычно присутствует в квартальных, месячных или недельных данных. Она имеет периодические изменения уровней ряда, почти повторяющиеся из года в год. Сезонные колебания обусловлены влиянием погоды или повторением календарно-зависимых событий, таких как праздники или школьные каникулы. Циклическая компонента является более плавной и описывает длительные периоды относительного подъема и спада, состоит из циклов, меняющихся по амплитуде и протяженности. Циклические колебания чаще всего являются следствием периодических изменений в экономике – последовательных расширений и сокращений производства, обычно называемых бизнес-циклом.

Конкретные модели функциональных взаимосвязей между этими компонентами могут иметь самый разный вид. Однако можно выделить два их основных вида:

- аддитивная модель $Y_t = T + C_t + S_t + \varepsilon_t$;
- мультипликативная модель $Y_t = T \cdot C_t \cdot S_t \cdot \varepsilon_t$.

Аддитивная модель применяется, если амплитуда сезонных колебаний остается примерно постоянной. Мультипликативная модель применяется, если амплитуда сезонных колебаний меняется в зависимости от изменения общего уровня значений ряда. При построении модели необходимо выделить эти составляющие.

Для построения моделей сезонности и цикличности используются различные методы: метод применения сезонных фиктивных переменных, метод авторегрессии и скользящего

среднего, метод гармонического анализа (использовании одномерных рядов Фурье) и др. В большинстве методов длина периодической компоненты обычно известна (или предполагается) заранее или выявляется визуально по графику ряда и затем включается в теоретическую модель. В отличие от этих методов гармонический анализ позволяет распознать периодические колебания различной длины и на этой основе построить периодическую модель ряда. В результате успешного анализа можно обнаружить всего несколько повторяющихся циклов различной длины во временном ряду, который, на первый взгляд, выглядят как случайный шум.

При применении гармонического анализа ряд раскладывается на функции синусов и косинусов различных частот [1]:

$$Y_t = T_t + \sum_{k=1}^{\frac{N}{2}} a_k \cos(\omega kt) + \sum_{k=1}^{\frac{N}{2}} b_k \sin(\omega kt). \quad (1)$$

где $\omega = \frac{2 \cdot \pi}{N}$ – основная частота колебаний;

N – число наблюдений; a_k и b_k – коэффициенты Фурье, определяемые по формулам:

$$\begin{aligned} a_0 &= \frac{1}{N} \sum Y_t; \\ a_k &= \frac{2}{N} \sum Y_t \cos(\omega kt); \\ b_k &= \frac{2}{N} \sum Y_t \sin(\omega kt) \end{aligned} \quad (2)$$

- при $k = 1 \dots (N/2 - 1)$

$$a_k = \frac{1}{N} \sum Y_t \cos(\omega kt) \quad b_k = 0$$

- при $k = N/2$.

Для распознавания периодических колебаний различной частоты используют периодограмму [4]. Значения периодограммы представляют собой дисперсию данных на соответствующей частоте. Значения периодограммы изображаются в зависимости от номера гармоники, частот или периодов. Дисперсия, учитываемая одной гармоникой равна [3]:

$$\sigma_k^2 = \frac{C_k^2}{2}, \quad (3)$$

где $C_k^2 = a_k^2 + b_k^2$.

Так как никакие две гармоники в силу их ортогональности не коррелируют между собой, то они не будут учитывать одну и ту же часть общей дисперсии, т. е. дисперсии, учитываемые различными гармониками, складываются и в результате их сумма должна равняться общей дисперсии ряда. Для оценки роли каждой гармоники в общей вариации уровней временного ряда рассчитывают удельный вес гармоники с номером k : $dk = \sigma_k^2 / \sigma^2$, где σ^2 – общая дисперсия временного ряда.

Вид периодограммы тесно связан со структурой ряда и является хорошим средством для выявления скрытых периодичностей. Наличие сезонных и циклических колебаний проявляет себя в виде острых узких пиков в периодограмме на соответствующих частотах (а при несимметричной форме сезонной или циклической волны – и на соответствующих кратных частотах), хотя подобные пики могут появиться и случайным образом.

Определение коэффициентов Фурье по формулам (2) достаточно трудоем-

ко. В Excel имеются более эффективные инструменты для этого. Первый из них – использование быстрого преобразования Фурье с помощью функции «Анализ Фурье». Существенным недостатком алгоритма быстрого преобразования Фурье является то, что число исходных данных обязательно должно быть равно степени числа 2, т. е. 4, 8, 16, и т. д. Если же это не так, то для построения модели целесообразно использовать второй способ (решение задачи линейной множественной регрессии), в которой зависимая переменная – наблюдаемый временной ряд, а независимые переменные – функции синусов и косинусов всех возможных (дискретных) частот. Для построения модели линейной множественной регрессии в Excel используется функция «Регрессия» надстройки «Анализ данных».

При построении модели периодических колебаний полезно вычестить среднее из значений ряда и удалить тренд, т. к. оба этих эффекта могут заслонить периодичности в данных.

Рассмотрим компьютерную технологию выполнения расчетов при построении модели методами гармонического анализа инструментами Excel. Процесс построения модели рассмотрим на примере временного ряда (см. график на рис. 1) поквартальных данных объема продаж компании за последние 8 лет [2]. Визуальный анализ графика показывает, что в ряду четко прослеживаются сезонные колебания с периодичностью в четыре квартала, остальные компоненты не прослеживаются. Поскольку амплитуда колебаний примерно одинакова, то будем строить аддитивную модель ряда.

Прежде всего определим тренд (или среднее значение уровня ряда)

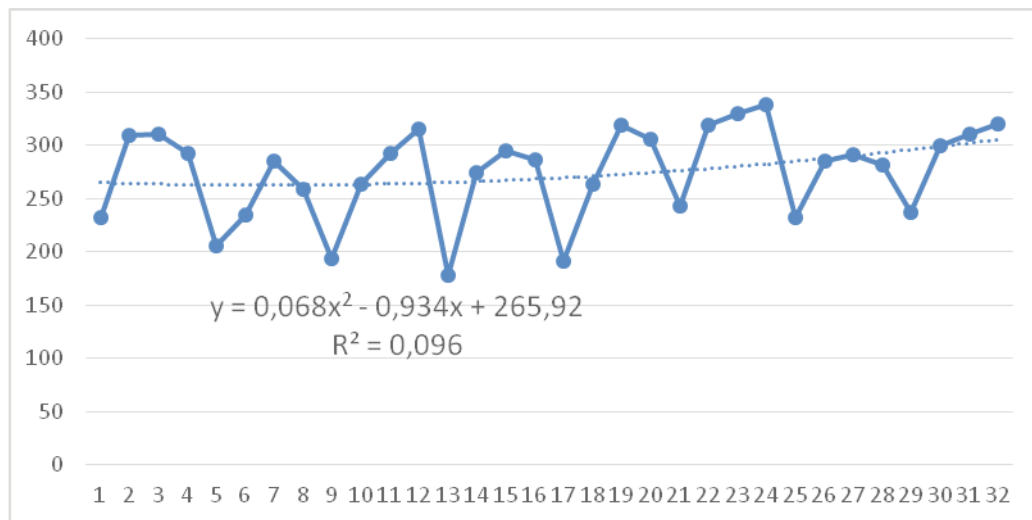


Рис. 1. График объемов продаж компании (тыс. долл)

Таблица 1

Результаты расчетов гармонических коэффициентов

Номер гармоники k	Комплексные числа Z_k	$\text{Re}(Z_k)$	$\text{Im}(Z_k)$	a_k	b_k
0	-8,0000000057821E-003	-0,008	0	-0,000	0
1	-32,950607783414+91,679066078454i	-32,9506	91,67907	-2,05941	-5,72994
2	57,3048822828274-169,16207224783i	57,30488	-169,162	3,581555	10,57263
3	292,356441434689-82,1623559738739i	292,3564	-82,1624	18,27228	5,135147
4	2,2278595974798-34,8583750084717i	2,22786	-34,8584	0,139241	2,178648
5	-40,9429284638622-60,7257877534981i	-40,9429	-60,7258	-2,55893	3,795362
6	-18,3816716833468-59,4477148842969i	-18,3817	-59,4477	-1,14885	3,715482
7	53,2161613256538-120,166632118465i	53,21616	-120,167	3,32601	7,510415
8	-702,628+126,452i	-702,628	126,452	-43,9143	-7,90325
9	121,397468486083+5,27202420699329i	121,3975	5,272024	7,587342	-0,3295
10	0,887399297488471-28,9664931323218i	0,887399	-28,9665	0,055462	1,810406
11	42,2383331919762+15,7334160911891i	42,23833	15,73342	2,639896	-0,98334
12	4,56414040252018-7,56237500847153i	4,56414	-7,56238	0,285259	0,472648
13	-118,500084532315-85,3846763244868i	-118,5	-85,3847	-7,40626	5,336542
14	6,36539010303101+79,5111495041452i	6,36539	79,51115	0,397837	-4,96945
15	55,5052163411896+2,62019820861156i	55,50522	2,620198	3,469076	-0,16376
16	-480,64	-480,64	0	-30,04	0

и исключим его из ряда. Для построения модели тренда воспользуемся инструментом «Подбор линии тренда» Мастера диаграмм Excel. Анализ трендовых моделей для данного ряда показывает, что можно выбрать полиномиальную модель: $T_t = 265,92 - 0,934t + 0,068t^2$. Для полученного детрендированного ряда вычислим коэффициенты Фурье в виде комплексных чисел с помощью функции «Анализ Фурье» надстройки «Анализ данных», а затем определим коэффициенты a_k и b_k (см. результаты расчетов в табл. 1).

Первое комплексное число соответствует нулевой гармонике, второе – первой гармонике и т. д. Во втором столбце таблицы с помощью инженерной функции «МНИМ.ВЕЩ» рассчитаны действительные части $\text{Re}(Z_k)$ комплексных чисел Z_k , а в третьем столбце с помощью инженерной функции «МНИМ.ЧАСТЬ» вычислены мнимые части $\text{Im}(Z_k)$ чисел Z_k . Действительные и мнимые части рассчитанных в режиме «Анализ Фурье» комплексных чисел связаны с гармоническими коэффициентами a_k и b_k следующими соотношениями:

$$\begin{aligned} a_0 &= \frac{\text{Re}(Z_0)}{N}; \\ a_k &= \frac{2 \cdot \text{Re}(Z_k)}{N}; \\ b_k &= \frac{2 \cdot \text{Im}(Z_k)}{N}. \end{aligned} \quad (4)$$

При $k=16$

$$\begin{aligned} a_k &= \frac{\text{Re}(Z_k)}{N}; \\ b_k &= \frac{\text{Im}(Z_k)}{N} \end{aligned}$$

Для построения периодограммы для каждой гармоники определим период как N/k , частоту, как величину, обратную периоду, и дисперсию по формуле (3). По результатам расчетов построим периодограмму (табл. 2 и рис. 2). Периодограмма указывает на наличие сезонных эффектов при периоде, равном четырем кварталам, а также на наличие циклических колебаний с периодом примерно в три года (10,67 кварталов). Эти периоды соответствуют 8 и 3 гармоникам и частотам 0,25 и 0,09375. Кратными этим частотам будут частоты 0,5 и 0,28125 (16 и 9 гармоник), поэтому 16 гармоника может быть включена в модель сезонности, а 9-я – в модель циклической компоненты. Однако 16 гармоника является вырожденной (ее значения равны -30,0), и поэтому ее включение в модель не уменьшит остаточную дисперсию, но сделает сезонную волну несимметричной. В рассматриваемой задаче в этом нет необходимости.

Для оценки роли и значимости каждой гармоники рассчитываются удельный вес гармоник d_k и статистика F_k , имеющая распределение Фишера с 2 и $N-2q+1$ степенями свободы [1]:

$$F_k = \frac{\sigma_k^2}{\frac{\sigma_{\text{ост}}^2}{N}}$$

$$\text{где } = \left[\sum_{k=1}^{\frac{N}{2}} \sigma_k^2 - \sum_j \frac{[\sigma_{k(j)}^2]^N}{[N - (2q + 1)]^1} \right]^{-1}$$

– остаточная дисперсия;

j – номера учитываемых в модели гармоник;

q – число учитываемых в модели гармоник.

Таблица 2

Результаты расчетов дисперсии данных на соответствующей частоте

k	Период, кв.	Частота	Дисперсия σ_k^2	Доли дисперсии, $d_{k\%}$	Статистика F_k	Значимость F_k
1	32	0,03125	18,53671	1,136591	0,890145	0,42321322
2	16	0,0625	62,30402	3,820215	2,991879	0,06840279
3	10,666667	0,09375	180,1229	11,04436	8,649619	0,00139667
4	8	0,125	2,382949	0,146112	0,114431	0,89233812
5	6,4	0,15625	10,47645	0,642371	0,503086	0,61065337
6	5,333333	0,1875	7,562337	0,46369	0,363148	0,69909177
7	4,571429	0,21875	33,73433	2,068445	1,619944	0,21800318
8	4	0,25	995,4614	61,03742	47,8027	2,8653E-09
9	3,555556	0,28125	28,83816	1,768232	1,384827	0,26891758
10	3,2	0,3125	1,640323	0,100578	0,078769	0,92448151
11	2,909091	0,34375	3,968002	0,243301	0,190546	0,82769685
12	2,666667	0,375	0,152385	0,009344	0,007318	0,99271121
13	2,461538	0,40625	41,66565	2,554759	2,000812	0,15630444
14	2,285714	0,4375	12,42684	0,76196	0,596745	0,55825555
15	2,133333	0,46875	6,030653	0,369774	0,289596	0,75104315
16	2	0,5	225,6004	13,83285	10,83348	0,00040891

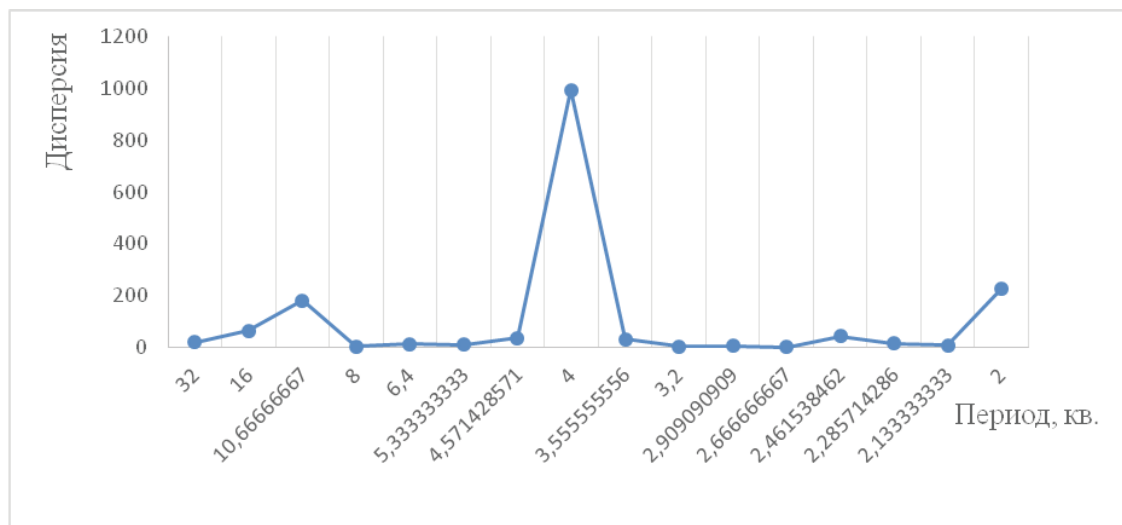


Рис. 2. Периодограмма для детрендированного ряда

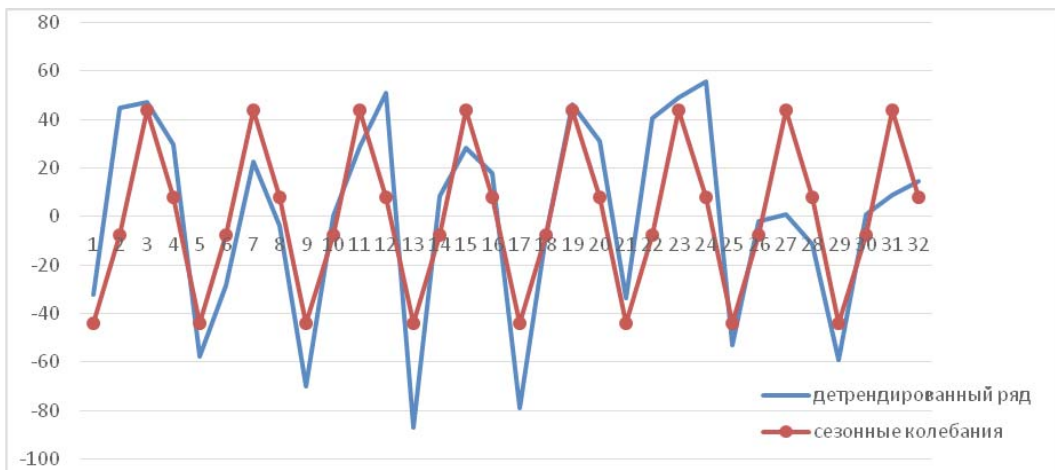


Рис. 3. Графики детрендрованного ряда и сезонной компоненты

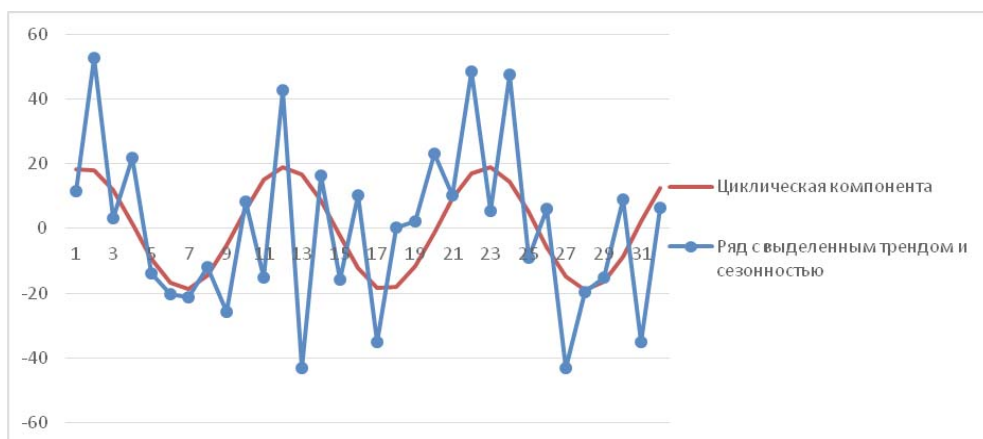


Рис. 4. Графики ряда с выделенным трендом и сезонностью и циклической компоненты

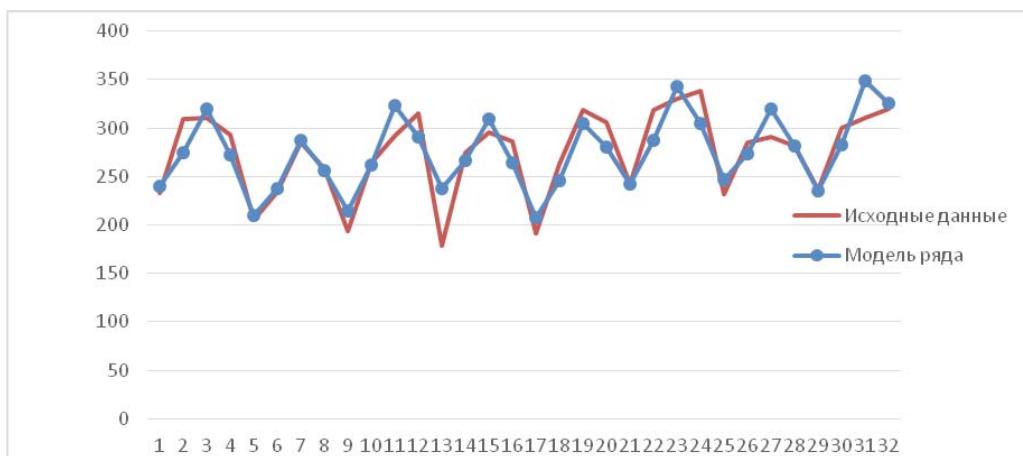


Рис. 5. Графики исходного и модельного рядов

При учете в модели 3, 8 и 9 гармоник $\sigma_{\text{ост}}^2 = 333,19$.

Рассчитанные значения удельных весов d_k , статистики F_k и значимости F представлены в 5, 6 и 7 столбцах табл. 2. Значимость F_k находилась с помощью статистической функции Excel «F.РАСП» при данном F_k , числе степеней свободы 2 и 27. Анализ этих данных показывает, что 9 гармоника является незначимой, следовательно, в модели необходимо оставить только 3 и 8 гармоники. В результате получаем следующие модели сезонной и циклической компонент:

$$S_t = -43,9143 \cos(8\omega t) - 7,90325 \sin(8\omega t);$$

$$C_t = 18,27222 \cos(3\omega t) + 5,1351 \sin(3\omega t).$$

Сезонные и циклические волны в соответствии с моделями являются симметричными с постоянными амплитудами 44,6 и 20,0 тыс. долл. Показатели точности полученной аддитивной модели временного ряда: коэффициент детерминации $R^2 = 0,748$; средняя относительная ошибка аппроксимации $\bar{\epsilon}_{\text{отн}} = 6,26$ %; процентная ошибка модели – 7,76 %. В заключение представим графики исходных и модельных значений уровней рассматриваемого временного ряда (рис. 3–5). Представленные данные свидетельствуют о достаточно высокой точности построенной модели и ее пригодности для практического применения.

Рассмотренная технология моделирования временного ряда в среде Excel позволила получить модель ряда, аналогичную модели, полученной с использованием профессионального статистического пакета STATISTIKA. При этом использовались все необходимые статистические процедуры, необходимые для идентификации и оценки параметров модели и проверки ее адекватности и точности. В отличие от статистических пакетов, моделирование в Excel доступно более широкому кругу специалистов и студентов, наглядно представляет процесс моделирования и дает широкую свободу для применения различных статистических инструментов, критериев и тестов для улучшения качества модели.

ЛИТЕРАТУРА:

1. Андерсен Т. Статистический анализ временных рядов. М.: Мир, 1976. 754 с.
2. Ханк Д.Э, Уичерн Д.У., Райтс А.Д. Бизнес-прогнозирование. М.: ИД «Вильямс», 2003. 652 с.
3. Харченко Л.П. Статистика: учебник. М.: Инфра-М, 2008. 445 с.
4. Электронный учебник по статистике / StatSoft [электронный ресурс]. URL: <http://www.statsoft.ru/home/textbook/default.htm> (дата обращения: 28.03.2015 г.)
5. Юров В.М. Моделирование нестационарных временных рядов с выраженными колебаниями с использованием инструментов Excel // Вестник Московского государственного областного университета. Серия: Экономика. 2015. № 1. С. 62–71.