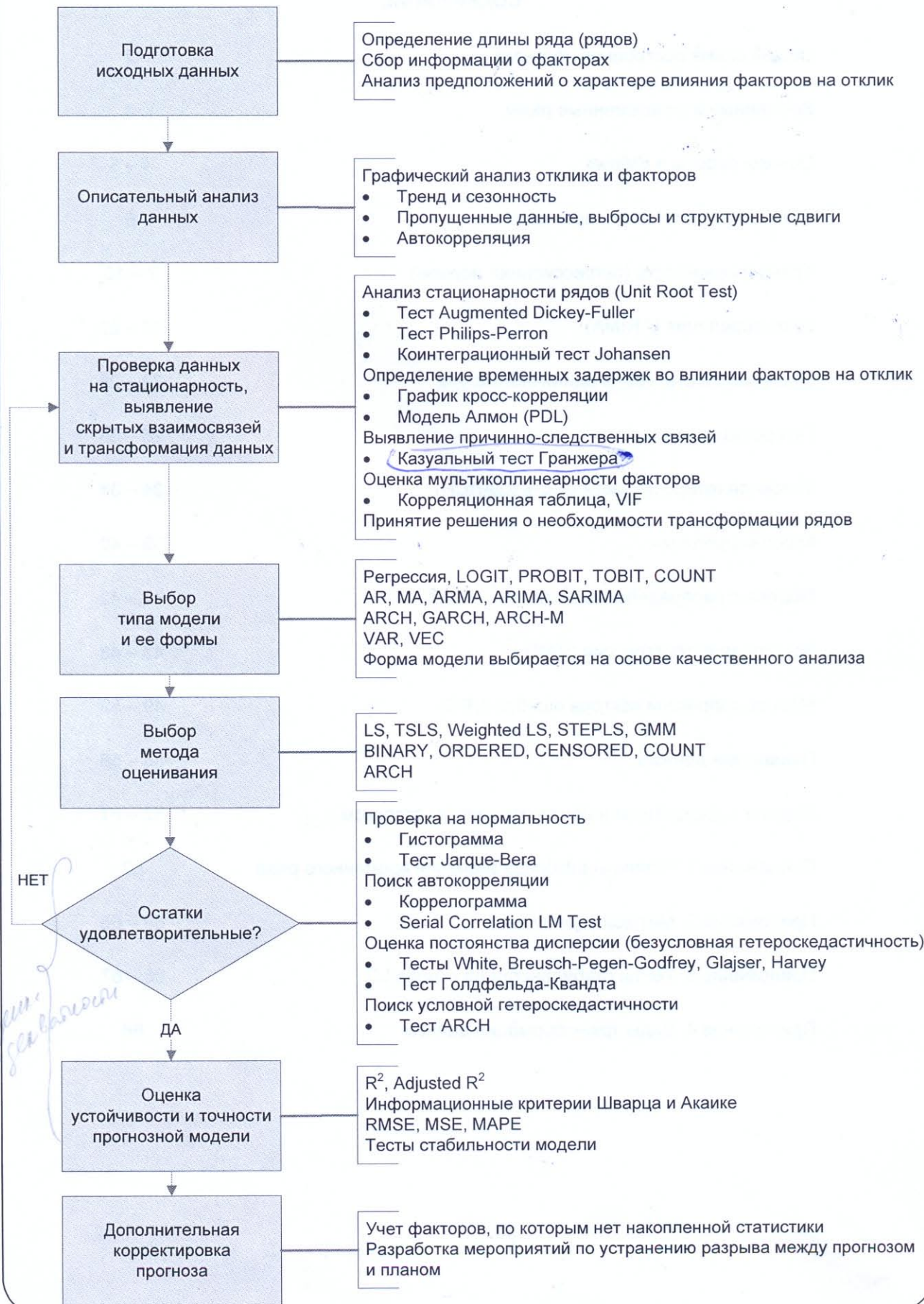




ОБЩАЯ СХЕМА ПОСТРОЕНИЯ ПРОГНОЗА

Общая схема построения прогноза





Временные и не временные ряды



Временной ряд – выборка значений, собранных и записанных в хронологическом порядке. Интервалы времени между любыми соседними значениями (*уровнями*) должны быть одинаковы.

Цели обработки временных рядов:

- Идентифицировать временной ряд
- Спрогнозировать уровень ряда на следующий момент времени на основе анализа тенденций прошлого и настоящего

Не временной ряд – выборка значений, собранная в определенный момент времени.

Цели обработки не временных рядов:

- Выявить закономерности, скрытые в данных
- Распространить выявленные закономерности на выборку большего объема (т.е. выполнить прогноз)

ПРИМЕРЫ ПОСТАНОВОК ЗАДАЧ ПРОГНОЗИРОВАНИЯ

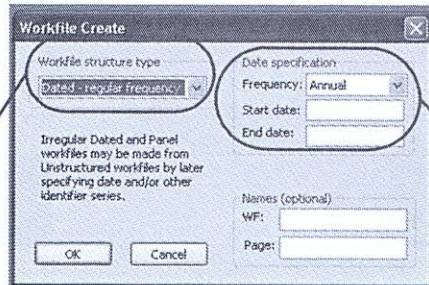
- Определить уровень расхода электроэнергии на предприятии на ближайший квартал.
- Построить прогноз продаж компании на неделю, месяц, квартал, год. При этом известно, что компания существует на рынке более 3 лет.
- Спрогнозировать измерение стоимости квартиры через месяц, если будут продолжаться тенденции повышения цен на объекты недвижимости, наблюдавшиеся за последние 2 года.
- Спрогнозировать изменения доли рынка некой брендовой торговой марки на последующие 2 года.
- Спрогнозировать численность населения на 10 лет вперед.
- Планируется строительство нового торгового центра. Требуется спрогнозировать «проходимость» секций будущего торгового центра с целью обоснования ставки арендной платы и оптимальной площади помещений.
- На основе риэлтерской базы данных по реализованным объектам недвижимости построить прогноз стоимости квартиры с учетом площади, возраста объекта, типа дома и других факторов.
- Спрогнозировать долю рынка новой торговой марки определенной группы товаров.
- Спрогнозировать стоимость автомобиля через три года эксплуатации.
- Какова будет ожидаемая сумма покупки, при условии, что покупатель – мужчина, который затратил 20 минут на выбор необходимых ему товаров?



Оснoвы работы в EViews

Создание нового файла EViews

File – New – Workfile...



Be monthly

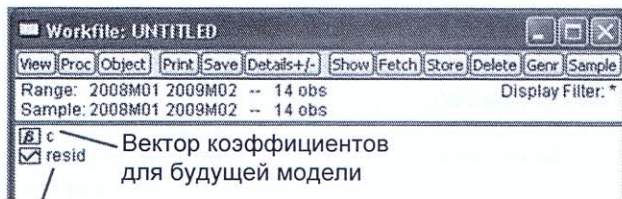
EViews позволяет строить модели для:

- **Unstructured / Undated** – не временных рядов
- **Dated-regular frequency** – временных рядов
- **Balanced Panel** – панельных данных

Для временных рядов и панельных данных возможна следующая периодизация:

- **Annual** – годовая (2001 2009)
- **Semi-annual** – полугодовая (2001s1 2009s2)
- **Quarterly** – квартальная (2001q1 2009q4)
- **Monthly** – месячная (2001m1 2009m12)
- **Weekly** – недельная (2001w1 2009w54)
- **Daily – 5 day week** – дни с понедельника по пятницу (2001d1 2009d365)
- **Daily – 7 day week** – все дни указанного периода (2001d1 2009d365)
- **Integer date** – пронумерованные по порядку наблюдения

После задания периодизации будет открыт новый рабочий файл.



Вектор коэффициентов для будущей модели

Остатки для будущей модели

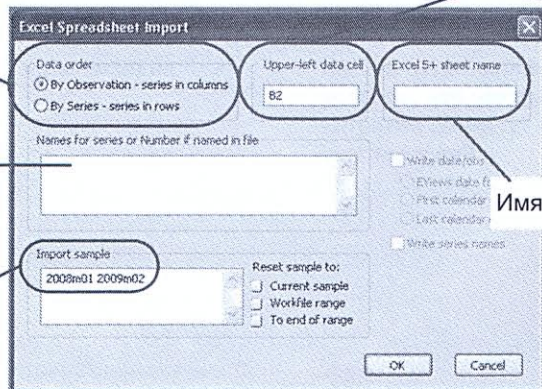
Импорт данных Microsoft Excel в новый файл EViews

File (Proc) – Import – Read Text-Lotus-Excel... В диалоговом окне **Open** выбрать таблицу с данными формата Excel.

Данные в Excel организованы в столбцах или строках

Количество импортируемых рядов (столбцов или строк)

Диапазон дат, указанный при создании рабочего файла EViews



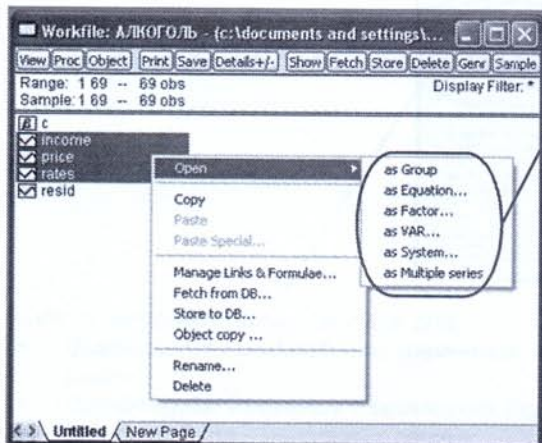
Верхняя левая ячейка таблицы Excel, содержащая данные

Имя листа с данными Excel



Основы работы в EViews (продолжение)

Создание объектов
для выделенных данных

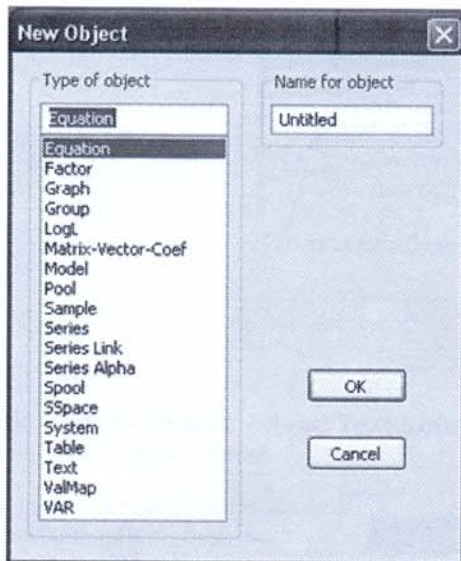


С участием выделенных переменных можно создать следующие объекты:

- **Group** – открыть выделенные переменные как группу (в одной таблице)
- **Equation** – открыть выделенные переменные как уравнение (по умолчанию – линейная регрессия на основе LS), причем первая из выделенных переменных будет считаться откликом
- **Factor** – «свернуть» выделенные переменные в меньшее число компонент с применением факторного анализа
- **VAR** – открыть выделенные переменные как VAR систему
- **System** – открыть выделенные переменные как систему
- **Multiply Series** – открыть все выделенные переменные, но каждую в своем окне

Создание новых объектов
без предварительного выделения данных

Object – New Object...



Без предварительного выделения переменных можно создать следующие объекты:

- **Equation** – создать новое уравнение (по умолчанию – линейная регрессия на основе LS)
- **Factor** – создать объект с результатами факторного анализа для указанных далее переменных
- **Graph** – построить график для указанных далее переменных
- **Group** – открыть указанные далее переменные как группу (в одной таблице)
- **LogL** – создать уравнение на основе метода максимального правдоподобия
- **Matrix-Vector-Coef** – создать матрицу или вектор
- **Model** – задать модель путем записи уравнения
- **Pool** – создать набор панельных данных
- **Sample** – создать выборку другого объема на основе имеющейся
- **Series Link** – создать ссылку на ряд
- **Series Alpha** – создать качественную переменную с текстовыми или цифровыми значениями
- **Spool** – создать сводный отчет с результатами моделирования
- **SSpace** – создать уравнение или систему для прогноза динамического (еще изменяющегося) состояния показателя
- **System** – создать систему уравнений
- **Table** – создать пустую таблицу
- **Text** – создать пустой текстовый комментарий
- **ValMap** – задать метки для качественных переменных
- **VAR** – создать VAR систему с участием указанных далее переменных



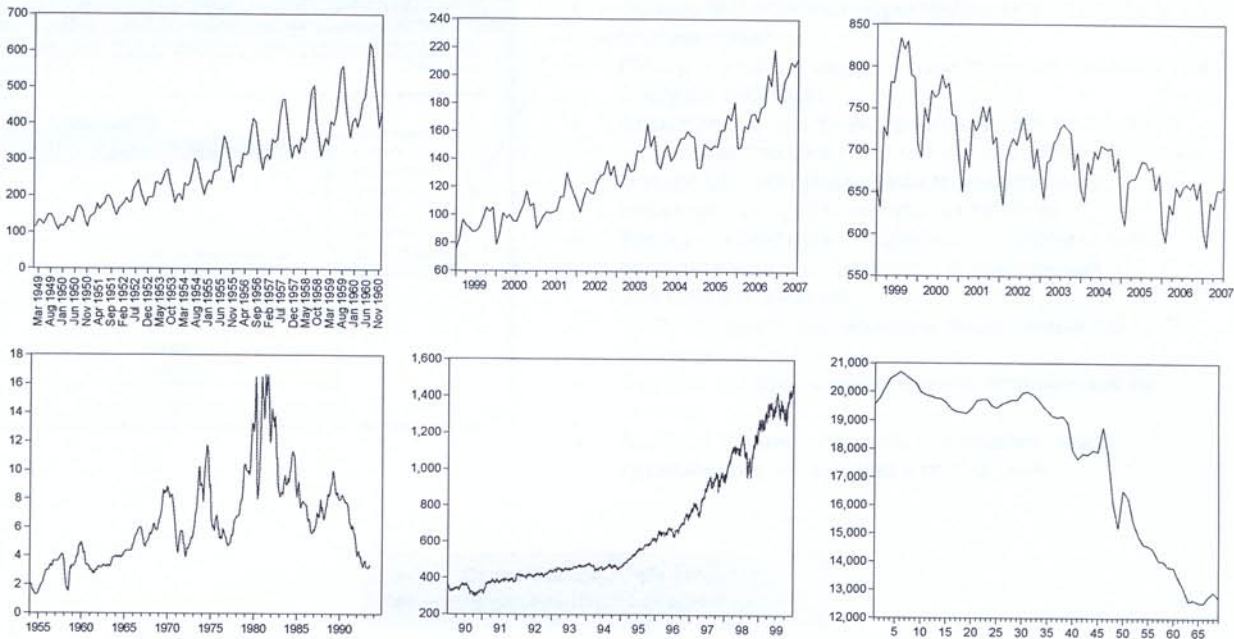
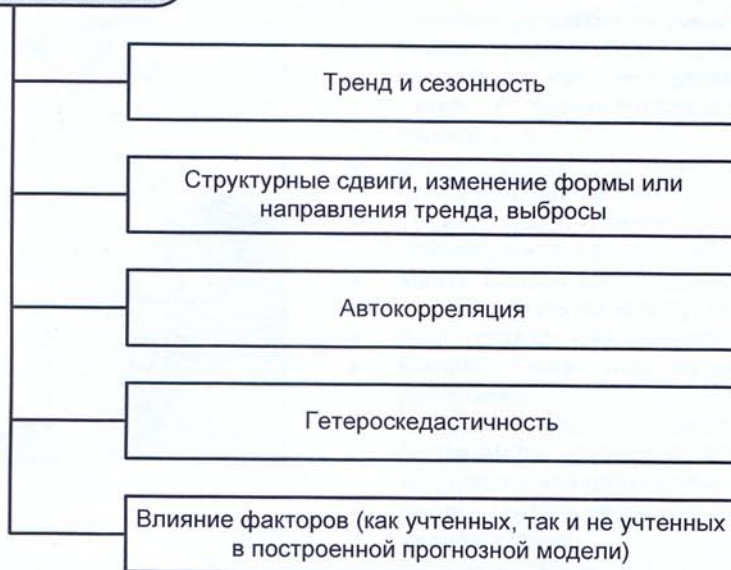
ЧТО ВЛИЯЕТ НА ТОЧНОСТЬ ПРОГНОЗА?

Что влияет на точность прогноза?

Цель: построить модель временного ряда и получить точный прогноз на будущие периоды.

Точность прогноза напрямую зависит от понимания причин вариации ряда.

ПРИМЕРЫ ВРЕМЕННЫХ РЯДОВ С РАЗЛИЧНОЙ ВАРИАЦИЕЙ

Причины вариации
временного ряда

Учет в модели всех этих причин позволяет добиться высокой точности прогноза.

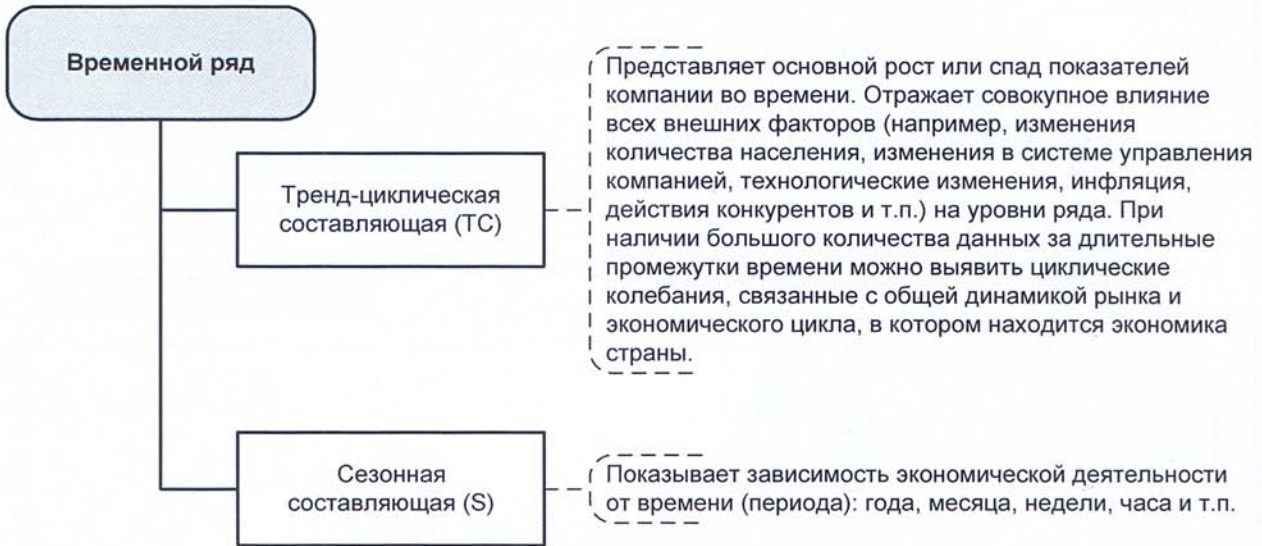
Далее рассмотрим суть каждой проблемы, способы обнаружения и учета в модели.



Тренд и сезонность временного ряда

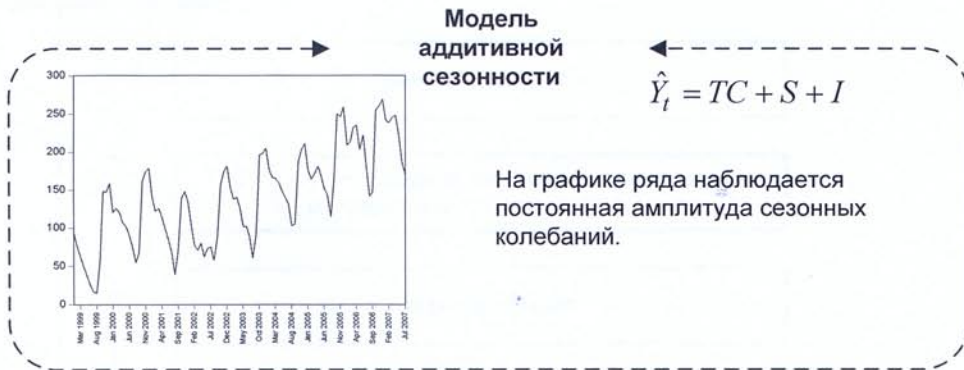
Каждый уровень временного ряда формируется под воздействием большого числа факторов, среди которых, в частности, можно выделить:

- Факторы, формирующие тенденцию временного ряда и глобальные циклические колебания (тренд-циклическая составляющая).
- Факторы, формирующие сезонные колебания временного ряда (сезонная составляющая).



Тренд-циклическая составляющая (тренд) может иметь разное направление (возрастающий, убывающий) и разную форму (линейный, квадратичный, экспоненциальный, логарифмический, затухающий и т.д.). Временной ряд может не содержать тренда.

Сезонность может включаться в модель двумя способами: аддитивным и мультипликативным. Как правило, тип модели можно определить по построенному графику ряда. Временной ряд может не содержать сезонную составляющую.



Выявление тренда и сезонности временного ряда

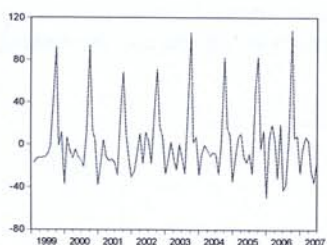
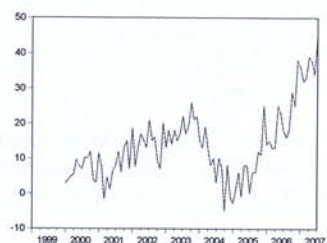
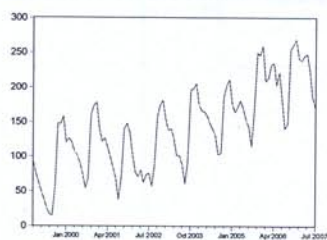
Способы выявления тренда и сезонности:

- Визуальный анализ графика ряда
- Анализ коррелограммы (графика автокорреляции)

Автокорреляция порядка p – это корреляция уровней ряда с этими же уровнями, но сдвинутыми на p периодов (лагов).

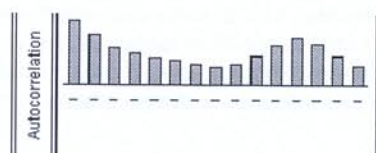
Автокорреляция обычно анализируется по графику, который называется **коррелограммой**. В EViews ее можно построить через: **View – Correlogram...**

График ряда



Коррелограмма и причины автокорреляции

1



Тренд и сезонность

2

Только тренд («ниспадающий забор»),
сильная автокорреляция

3



Только сезонность

(всплески на «сезонных» лагах:

на 12, 24 и т.д. – если данные собраны по месяцам,
на 4, 8, 12 и т.д. – если данные собраны по кварталам
и т.п.)

В случае, когда ряд содержит и тренд и сезонность, сначала строят коррелограмму самого ряда (рис. 1), а затем ряда, очищенного от тренда, для того, чтобы подтвердить наличие сезонности (рис.3).

Замечание

- В EViews ряд, очищенный от тренда, можно создать следующим образом:

```
genr t_data=d(data)
```

Это ряд первых разностей уровней исходного ряда data.



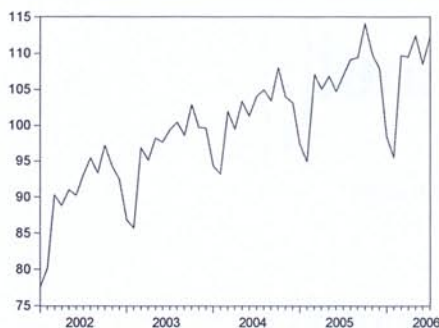
Постановка задачи

Требуется спрогнозировать объем железнодорожных перевозок (млн. тонн) до декабря 2008 г. Исходные данные собирались с января 2002 г. по месяцам. Дополнительная информация о внешних факторах, влияющих на объем железнодорожных перевозок, отсутствует.

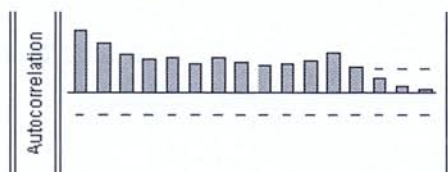
Реализация в EViews

Шаг 1. Анализ отклика

Исходный временной ряд (*COUNT*) имеет линейный тренд и аддитивную сезонность.

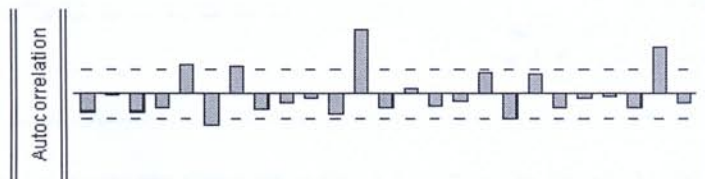


Коррелограмма исходного ряда.



Наличие существенной автокорреляции вызвано присутствием тренда и сезонности.

Коррелограмма ряда, очищенного от тренда путем взятия первой разности.



Присутствие сезонности очевидно.

Будем строить регрессионную модель с учетом линейного тренда и DUMMY переменными для сезонности, т.е. регрессию вида:

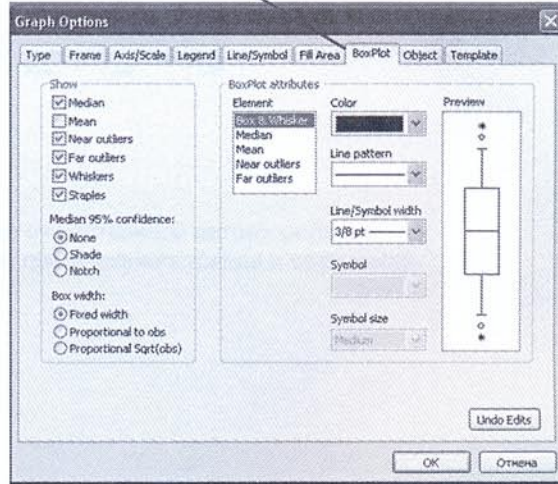
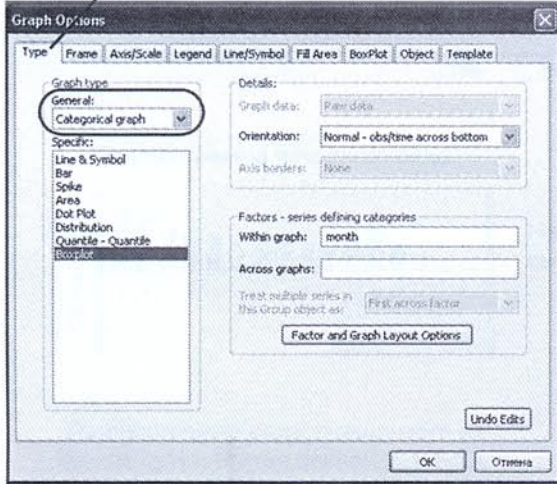
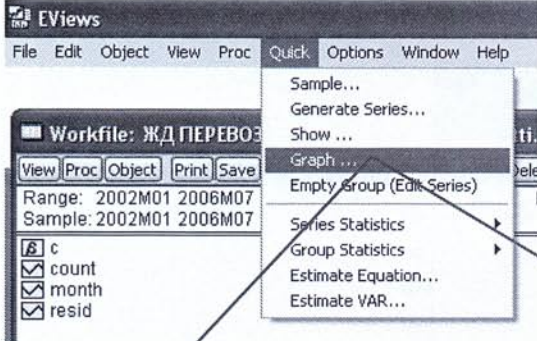
$$COUNT = const + A \cdot TREND + \sum_{i=1}^{11} B_i \cdot D_i$$



Реализация в Eviews (продолжение)

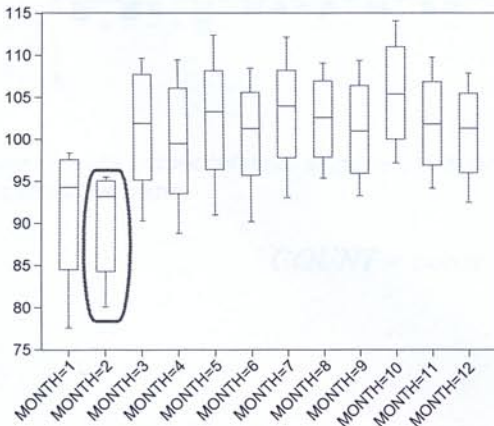
Шаг 2. Выбор базового месяца для сезонных DUMMY (для аддитивной сезонности)

Необходимо выбрать базовый сезон (месяц), по сравнению с которым будет оцениваться средний рост или падение значений отклика других сезонов (месяцев).
Графическим инструментом определения базового сезона является ящик с усами.



@ month - если нет в данных

COUNT by MONTH



Базовый сезон (месяц) определяется экспертно с учетом наименьшего разброса значений и/или отличной от других медиане.

В нашем случае возьмем MONTH=2 за базовый.

Количество DUMMY-переменных для учета сезонности всегда должно быть на одну меньше, чем количество сезонов.

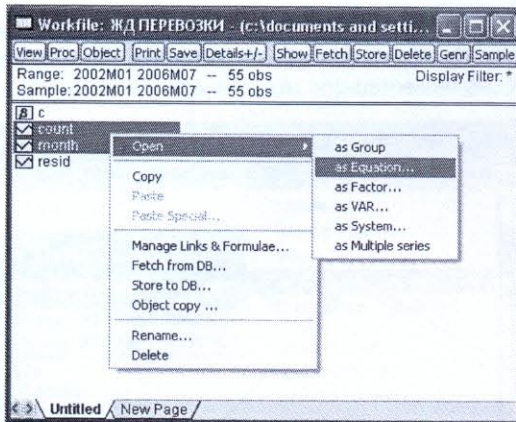
Для базового сезона DUMMY не создается.

В случае мультипликативной сезонности необходимо строить модель для логарифмированного (по основанию e) отклика.

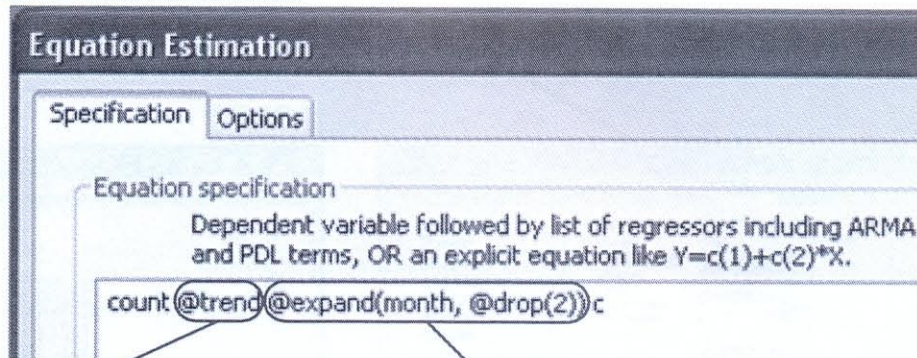


Реализация в Eviews (продолжение)

Шаг 3. Построение регрессионной модели



В силу отсутствия данных о влияющих на отклик факторах будем строить линейную регрессионную модель. В качестве факторов возьмем линейный тренд и DUMMY для аддитивной сезонности.

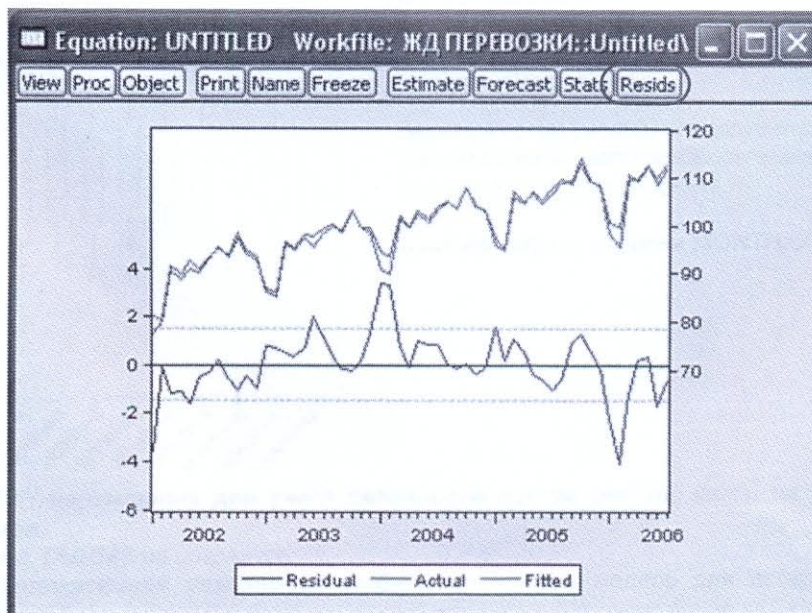


Переменная для линейного тренда
(номер периода, номер по порядку)

Создание 11 DUMMY-переменных для учета аддитивной сезонности. Второй месяц - базовый.

Шаг 4. Анализ остатков

ГРАФИК ПОДГОНКИ МОДЕЛИ И ГРАФИК ОСТАТКОВ





Реализация в Eviews (продолжение)

Шаг 4. Анализ остатков
(продолжение)

	Coefficient	Std. Error	t-Statistic	Prob.
Correlogram - Q-statistics			0	0.0000
Correlogram Squared Residuals			7	0.0000
Histogram - Normality Test			6	0.1391
Serial Correlation LM Test...			1	0.0000
Heteroskedasticity Tests...			1	0.0000
MONTH=4	8			
MONTH=5	1			
MONTH=6	8.962632	0.932681	9.609535	0.0000
MONTH=7	11.21829	0.933446	12.01815	0.0000
MONTH=8	12.57000	0.987815	12.72506	0.0000
MONTH=9	10.89066	0.987895	11.02411	0.0000
MONTH=10	14.86132	0.988135	15.03976	0.0000
MONTH=11	10.83197	0.988536	10.95759	0.0000
MONTH=12	9.277632	0.989097	9.379896	0.0000
R-squared	0.974024	Mean dependent var	99.72000	
Adjusted R-squared	0.966602	S.D. dependent var	8.057626	
S.E. of regression	1.472547	Akaike info criterion	3.814928	
Sum squared resid	91.07258	Schwarz criterion	4.289388	
Log likelihood	-91.91052	Hannan-Quinn criter.	3.998406	
F-statistic	131.2375	Durbin-Watson stat	0.802019	
Prob(F-statistic)	0.000000			

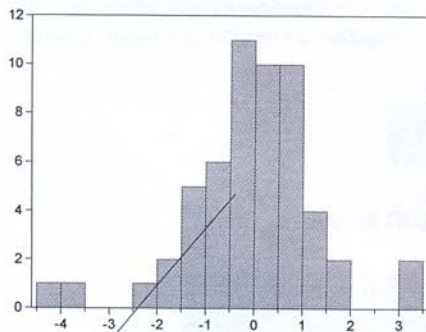
Коррелограмма (для оценки автокорреляции)

Гистограмма, тест на нормальность

Тест на автокорреляцию

НОРМАЛЬНОСТЬ ОСТАТКОВ

Residual Tests – Histogram – Normality Test



Series: Residuals	
Sample 2002M01 2006M07	
Observations 55	
Mean	1.06e-14
Median	0.023947
Maximum	3.420000
Minimum	-4.084211
Std. Dev.	1.298664
Skewness	-0.359122
Kurtosis	5.179583
Jarque-Bera	12.06896
Probability	0.002395

АНАЛИТИЧЕСКИ – Тест Жарка-Бера.

Распределение остатков противоречит нормальному закону, т.к. Probability = 0.002395 < 0.05.

ГРАФИЧЕСКИ – Гистограмма.

Распределение остатков достаточно симметрично.

Рекомендации

- Если тест Jarque-Bera утверждает нормальность остатков (Probability > 0.05), то гистограмму можно не анализировать.
- Если тест Jarque-Bera утверждает ненормальность остатков (Probability < 0.05), то следует проанализировать гистограмму. В случае, когда она достаточно симметрична, можно говорить о нормальности остатков, игнорируя результаты аналитического теста.

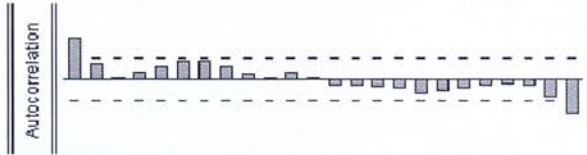


Реализация в Eviews (продолжение)

Шаг 4. Анализ остатков
(продолжение)

КОРРЕЛОГРАММА

Residual Tests – Correlogram – Q-statistics



Отсутствие регулярных сезонных всплесков и «ниспадающего забора» говорят о том, что тренд и сезонность учтены в модели. Хотя, если обратить внимание на 24 лаг, то можно сделать вывод о том, что существует модель лучше построенной.

Шаг 5. Оценка приемлемости модели
для прогнозирования

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	0.404342	0.012590	32.11610	0.0000
C	79.77145	0.729895	109.2817	0.0000
MONTH=1	1.404342	0.931406	1.507786	0.1391
MONTH=3	10.87566	0.931406	11.67681	0.0000
MONTH=4	8.891316	0.931661	9.547511	0.0000
MONTH=5	11.24697	0.932086	12.06645	0.0000
MONTH=6	8.962632	0.932681	9.609535	0.0000
MONTH=7	11.21829	0.933446	12.01815	0.0000
MONTH=8	12.57000	0.987815	12.72508	0.0000
MONTH=9	10.89066	0.987895	11.02411	0.0000
MONTH=10	14.86132	0.988735	15.03978	0.0000
MONTH=11	10.83197	0.996536	10.95759	0.0000
MONTH=12	9.277632	0.999097	9.379896	0.0000
R-squared	0.974024	Mean dependent var	99.72000	
Adjusted R-squared	0.966602	S.D. dependent var	8.057626	
S.E. of regression	1.472547	Akaike info criterion	3.814928	
Sum squared resid	91.07258	Schwarz criterion	4.289388	
Log likelihood	-91.91052	Hannan-Quinn criter	3.990406	
F-statistic	131.2375	Durbin-Watson Stat	0.802019	
Prob(F-statistic)	0.000000			

Коэффициент детерминации $R^2=0.974$, т.е. изменение отклика на 97.4% происходит под воздействием учтенных в модели факторов.

R^2 более 30%, значит прогнозировать по такой модели целесообразно.

При добавлении новых факторов R^2 увеличивается, что не обязательно означает улучшение качества регрессионной модели. Поэтому в случае большого количества факторов рекомендуется интерпретировать скорректированный R^2 (Adjusted R Squared).

Для F-статистики Prob.=0.000...<0.05, значит ошибка прогноза по построенной модели будет меньше, чем при «наивном» прогнозе (по среднему).



Реализация в EViews (продолжение)

Шаг 6. Анализ регрессионного уравнения и удаление факторов, не влияющих на отклик

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	0.404342	0.012590	32.11610	0.0000
C	79.77145	0.729895	109.2917	0.0000
MONTH=1	1.404342	0.931406	1.507766	0.1391
MONTH=3	10.87566	0.931406	11.67661	0.0000
MONTH=4	8.891316	0.931661	9.543511	0.0000
MONTH=5	11.24697	0.932086	12.06645	0.0000
MONTH=6	8.962632	0.932681	9.609535	0.0000
MONTH=7	11.21829	0.933446	12.01815	0.0000
MONTH=8	12.57000	0.987815	12.72506	0.0000
MONTH=9	10.89066	0.987895	11.02411	0.0000
MONTH=10	14.86132	0.988135	15.03976	0.0000
MONTH=11	10.83197	0.988536	10.95759	0.0000
MONTH=12	9.277632	0.989097	9.379896	0.0000

R-squared	0.974024	Mean dependent var	99.72000
Adjusted R-squared	0.966602	S.D. dependent var	8.057626
S.E. of regression	1.472547	Akaike info criterion	3.814928
Sum squared resid	91.07258	Schwarz criterion	4.289388
Log likelihood	-91.91052	Hannan-Quinn criter.	3.998406
F-statistic	131.2375	Durbin-Watson stat	0.802019
Prob(F-statistic)	0.000000		

Не влияющие DUMMY для сезонности нельзя исключать из модели. Их незначимость (Prob.>0.05) указывает на несущественное отличие данного сезона от базового.

Модель: _____

Интерпретация коэффициентов регрессии:

- Каждый месяц наблюдается рост ЖД перевозок в среднем на 0.4 тыс.тонн.
- ЖД перевозки в январе не существенно отличаются от числа ЖД перевозок в феврале.
- В марте наблюдается рост ЖД перевозок в среднем на 10.88 тыс.тонн по сравнению с февралем.
- В апреле наблюдается рост ЖД перевозок в среднем на 8.89 тыс.тонн по сравнению с февралем.
- В мае наблюдается рост ЖД перевозок в среднем на 11.25 тыс.тонн по сравнению с февралем.
- И т.д.

Шаг 7. Прогнозирование

Изменить длину ряда с учетом прогнозного периода

Workfile: ЖД ПЕРЕВОЗКИ - (c:\documents and setti...)

Range: 2002M01 2006M07 -- 55 obs

Sample: 2002M01 2006M07 -- 55 obs

Workfile structure

Workfile structure type: Dated - regular frequency

Date specification: Frequency: Monthly, Start date: 2002M01, End date: 2008M12

Два раза щелкнуть левой кнопкой мыши на **Range** и в открывшемся диалоговом окне указать конечную дату прогнозного периода (декабрь 2008 года).

Реализация в EViews (продолжение)

Шаг 7. Прогнозирование
(продолжение)Продлить значения переменной, отвечающей за
создание DUMMY для сезонности

Year	Month	Value
2006	M03	3.000000
2006	M04	4.000000
2006	M05	5.000000
2006	M06	6.000000
2006	M07	7.000000
2006	M08	NA
2006	M09	NA
2006	M10	NA
2006	M11	NA
2006	M12	NA
2007	M01	NA
2007	M02	NA
2007	M03	NA
2007	M04	NA

Щелчок на кнопке **Edit** позволит вставить пропущенные значения месяцев прогнозного периода. Также можно воспользоваться командой

```
genr month=@month
```

Сохранить прогнозные значения
и стандартную ошибку прогноза

Equation: UNTITLED Workfile: ЖД ПЕРЕВОЗКИ: Untitled

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: COUNT
Method: Least Squares
Date: 03/03/09 Time: 13:10
Sample (adjusted): 2002M01 2008M07
Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
----------	-------------	------------	-------------	-------

Forecast

Forecast of
Equation: UNTITLED Series: COUNT

Series names
Forecast name: forecast
S.E. (optional): SE_forecast
GARCH (optional):

Method
Static forecast (no dynamics in equation)
 Structural (ignore ARMA)
 Coef uncertainty in S.E. calc

Forecast sample
2002m01 2008m12

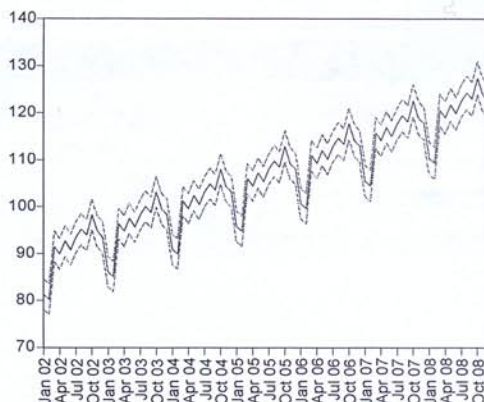
Output
 Forecast graph
 Forecast evaluation

Insert actuals for out-of-sample observations

OK Cancel

Сохранить границы 95% доверительного интервала прогноза можно путем ввода следующих команд:

```
genr low=forecast-2*se_forecast  
genr up=forecast+2*se_forecast
```



Forecast: FORECAST
Actual: COUNT
Forecast sample: 2002M01 2008M12
Included observations: 55

Root Mean Squared Error	1.286804
Mean Absolute Error	0.915914
Mean Abs. Percent Error	0.947962
Theil Inequality Coefficient	0.006432
Bias Proportion	0.000000
Variance Proportion	0.006580
Covariance Proportion	0.993420

— FORECAST
- - - ± 2 S.E.



АВТОКОРРЕЛЯЦИЯ

Суть автокорреляции

Автокорреляция порядка p – это корреляция уровней ряда с этими же уровнями, но сдвинутыми на p периодов (лагов).

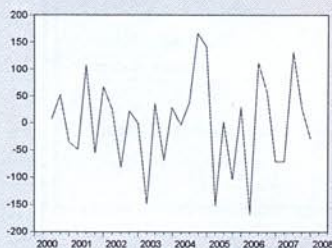
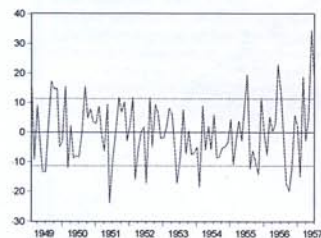
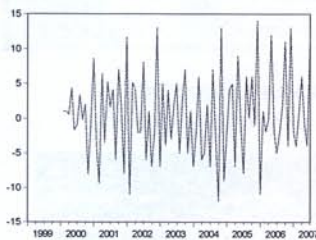
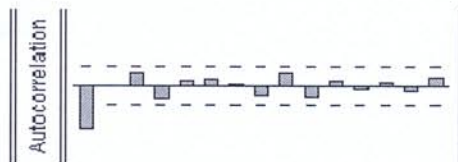
Причинами автокорреляции временного ряда или остатков могут быть:

- Тренд и сезонность (см. способы выявления тренда и сезонности на стр. 8)
- Влияние не учтенных в модели факторов
- Не верная форма модели
- Наличие собственной (остаточной) автокорреляции временного ряда (природа явления)

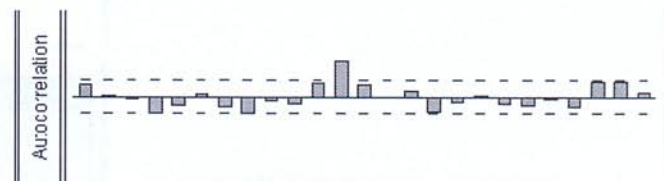
Остаточная автокорреляция временного ряда может быть выявлена следующими способами:

- После очищения ряда от тренда и сезонности
- После построения модели временного ряда и последующего анализа остатков на наличие автокорреляции

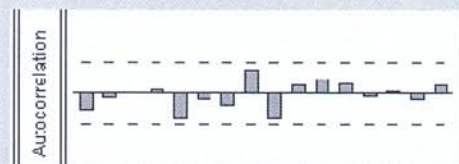
График ряда

Коррелограмма и
примеры остаточной автокорреляции

Остаточная автокорреляция 1-го порядка



Остаточная сезонная автокорреляция 1-го порядка



Автокорреляция отсутствует

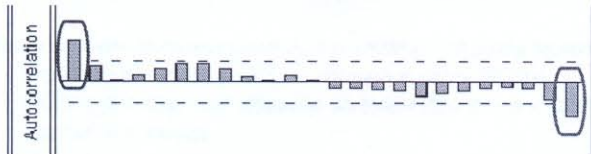


Выявление автокорреляции в анализе остатков

В примере на стр. 13 была построена коррелограмма остатков. Проанализируем ее с точки зрения наличия остаточной автокорреляции.

Графически

Residual Tests – Correlogram – Q-statistics



Коррелограмма показывает остаточную автокорреляцию на 1-ом лаге (указывает на возможность построения авторегрессионной модели вместо регрессионной) и, вероятно, остаточную сезонную автокорреляцию на 24-ом лаге (указывает на возможность учета в авторегрессионной модели сезонной автокорреляции).

Аналитически

Residual Tests – Serial Correlation LM Test... (с 1 лагом)

Breusch-Godfrey Serial Correlation LM Test				
F-statistic	15.86022	Prob. F(1, 41)	0.0003	
Obs*R-squared	15.34134	Prob. Chi-Square(1)	0.0001	
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Date: 03/04/09 Time: 10:41				
Sample: 2002M01 2006M07				
Included observations: 55				
Presample missing value lagged residuals set to zero.				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	-0.000725	0.010822	-0.066994	0.9469
C	0.018125	0.627324	0.028893	0.9771
MONTH=1	-0.000725	0.800496	-0.000906	0.9993
MONTH=3	0.000725	0.800496	0.000906	0.9993
MONTH=4	0.001450	0.800716	0.001811	0.9986
MONTH=5	0.002175	0.801081	0.002715	0.9978
MONTH=6	0.002900	0.801593	0.003618	0.9971
MONTH=7	0.003625	0.802250	0.004519	0.9964
MONTH=8	-0.082651	0.849230	-0.097324	0.9229
MONTH=9	0.000725	0.849046	0.000854	0.9993
MONTH=10	0.001450	0.849253	0.001707	0.9986
MONTH=11	0.002175	0.849597	0.002560	0.9980
MONTH=12	0.002900	0.850090	0.003411	0.9972
RESID(-1)	0.529634	0.132991	3.982489	0.0003
R-squared	0.278933	Mean dependent var	-2.01E-14	
Adjusted R-squared	0.050303	S.D. dependent var	1.298664	
S.E. of regression	1.265580	Akaike info criterion	3.524268	
Sum squared resid	65.66939	Schwarz criterion	4.035225	
Log likelihood	-82.91736	Hannan-Quinn criter.	3.721859	
F-statistic	1.220017	Durbin-Watson stat	1.703329	
Prob(F-statistic)	0.300591			

Оценка существенности автокорреляции остатков вплоть до лага 1 – статистика Бреуша - Пагана.

Если Prob. < 0.05, то автокорреляция остатков вплоть до лага 1 присутствует.

Тестируемая модель имеет вид:

$$\text{Остаток}_t = \text{const} + \sum a_i \cdot \text{Фактор}_i + b_1 \cdot \text{Остаток}_{t-1}$$

В данном случае имеется автокорреляция остатков 1-го порядка, т.к. первый лаг является значимым (Prob. < 0.05).

Замечание

- Тест чувствителен к количеству лагов остатков, участвующих в тесте. Поэтому число лагов целесообразно выбирать согласно коррелограмме.



Учет автокорреляции: АВТОРЕГРЕССИЯ (AR)

В отличие от регрессионного анализа, где прогнозная модель строится на основе внешних факторов, авторегрессия учитывает ещё и внутренний фактор – влияние предыдущих p уровней исходного ряда на прогнозируемый уровень.

Авторегрессионная модель порядка p ($AR(p)$) имеет следующий вид:

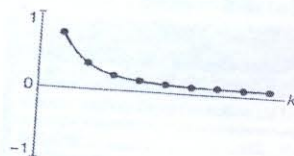
$$\hat{Y}_t = B_0 + \sum_{i=1}^p AR_i \cdot Y_{t-i} + \sum_{j=1}^k B_j \cdot X_j$$

где X_j - факторы, Y_t - отклик, AR_i, B_j - параметры (коэффициенты) авторегрессии.

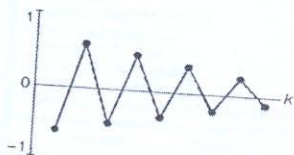
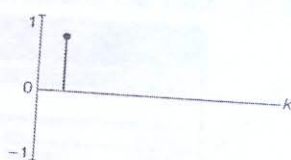
Математически суть авторегрессионного анализа сводится к нахождению параметров авторегрессии, проверке их значимости и оценке приемлемости всей построенной модели в целом.

Выбор порядка AR модели

Автокорреляция



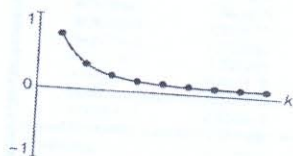
Частная корреляция



Авторегрессионная модель 1-го порядка:

$$\hat{Y}_t = B_0 + AR_1 \cdot Y_{t-1} + \sum_{j=1}^k B_j \cdot X_j$$

Автокорреляция



Частная корреляция



Авторегрессионная модель 2-го порядка:

$$\hat{Y}_t = B_0 + AR_1 \cdot Y_{t-1} + AR_2 \cdot Y_{t-2} + \sum_{j=1}^k B_j \cdot X_j$$

Замечания

- При включении AR компоненты в модель в данных могут быть пропущенные значения.
- В EViews есть возможность включать в модель AR составляющие не последовательно, а выборочно например, только AR(2).
- Учет остаточной сезонной автокорреляции происходит аналогичным образом, с поправкой на периодические лаги на графиках автокорреляции. Например, для данных, фиксируемых по месяцам, AR(12) – учет значения прошлого года.



Учет автокорреляции: СКОЛЬЗЯЩЕЕ СРЕДНЕЕ (МА)

В отличие от регрессионного анализа, где прогнозная модель строится на основе внешних факторов, и авторегрессии, модели скользящего среднего учитывают ещё и внутренний фактор – влияние предыдущих q остатков (т.е. ошибок прогнозирования) на прогнозируемый уровень.

Модель скользящего среднего порядка q ($AR(q)$) имеет следующий вид:

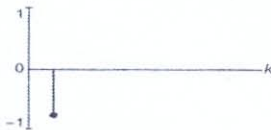
$$\hat{Y}_t = B_0 + \sum_{i=1}^q MA_i \cdot e_{t-i} + \sum_{j=1}^k B_j \cdot X_j$$

где X_j - факторы, Y_t - отклик, MA_i, B_j - параметры (коэффициенты) модели скользящего среднего.

Математически суть моделирования скользящим средним сводится к нахождению параметров уравнения, проверке их значимости и оценке приемлемости всей построенной модели в целом.

Выбор порядка МА модели

Автокорреляция



Частная корреляция



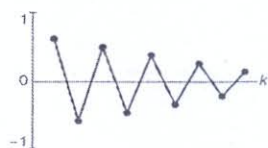
Авторегрессионная модель 1-го порядка:

$$\hat{Y}_t = B_0 + MA_1 \cdot e_{t-1} + \sum_{j=1}^k B_j \cdot X_j$$

Автокорреляция



Частная корреляция



Авторегрессионная модель 2-го порядка:

$$\hat{Y}_t = B_0 + MA_1 \cdot e_{t-1} + MA_2 \cdot e_{t-2} + \sum_{j=1}^k B_j \cdot X_j$$

Замечания

- При включении МА компоненты в модель в данных не должно быть пропущенных значений.
- В EViews есть возможность включать в модель МА составляющие не последовательно, а выборочно, например, только МА(2)
- Учет остаточной сезонной автокорреляции с помощью скользящего среднего происходит аналогичным образом, с поправкой на периодические лаги на графиках автокорреляции. Например, для данных, фиксируемых по месяцам, МА(12) – учет ошибки прошлого года.