

# Лекция 9

---

АНАЛИЗ ВРЕМЕННЫХ РЯДОВ (РЯДЫ  
ДИНАМИКИ)

# Ряды динамики и анализ данных



Классификация технологических методов ИАД [1]

# Ряды динамики. Основные определения

Ряд наблюдений за значениями определенного показателя, упорядоченный в зависимости от возрастающих или убывающих значений другого показателя, называют *динамическим рядом, временным рядом, рядом динамики*.

Отдельные наблюдения временного ряда называются *уровнями* этого ряда.

Временные ряды бывают моментные, интервальные и производные.

- *Моментные ряды* характеризуют значения показателя на определенные моменты времени.
- *Интервальные ряды* характеризуют значения показателя за определенные интервалы времени.
- *Производные ряды* получаются из средних или относительных величин показателя.

# Временные ряды. Условия прогнозирования

Уровни ряда могут иметь **детерминированные** или **случайные значения**.

*Ряд последовательных данных о количестве дней в месяце, квартале, годе являются примерами рядов с детерминированными значениями.*

**Прогнозированию подвергаются ряды со случайными значениями уровней.** Каждый показатель таких рядов может иметь дискретную или непрерывную величину.

- **Важное значение для прогнозирования имеет *выбор интервалов между соседними уровнями ряда*.**

*При слишком большом интервале времени могут быть упущены некоторые закономерности в динамике показателя. При слишком малом – увеличивается объем вычислений, могут появляться ненужные детали в динамике процесса.*

- **Важным условием правильного отражения временным рядом реального процесса развития является *сопоставимость уровней ряда*.**

*Несопоставимость чаще всего встречается в стоимостных характеристиках, изменениях цен, территориальных изменениях, при укрупнении предприятий и др.*

Для несопоставимых величин показателя неправомерно проводить прогнозирование.

# Временные ряды. Условия для успешного изучения

Для успешного изучения динамики процесса необходимо, чтобы информация была полной, временной ряд имел достаточную длину, отсутствовали пропущенные наблюдения.

- Уровни временных рядов могут иметь аномальные значения.

*Появление таких значений может быть вызвано ошибками при сборе, записи или передачи информации. Это ошибки технического порядка или **ошибки первого рода**.*

*Однако аномальные значения могут отражать реальные процессы, например, скачок курса доллара и др. Такие аномальные значения относятся к **ошибкам второго рода**, они устранению не подлежат.*

Для выявления аномальных уровней временных рядов можно использовать метод Ирвина.

# Временные ряды. Выявление аномалий

**Метод Ирвина** предполагает использование следующей формулы:

$$\lambda_t = \frac{|y_t - y_{t-1}|}{\sigma_y}, t = 1, 2, \dots, n$$

где  $\sigma_y$ , среднее квадратическое отклонение временного ряда  $y_1, y_2, \dots, y_i, y_n$ .

$$\sigma_y = \sqrt{\frac{\sum (y_t - \bar{y})^2}{n-1}}, \bar{y} = \frac{\sum y_t}{n}$$

Расчетные значения  $\lambda_2, \lambda_3, \dots$  сравниваются с табличными значениями критерия Ирвина  $\lambda_\alpha$ , и **если оказываются больше табличных, то соответствующее значение  $y_t$  уровня ряда считается аномальным.**

Значения критерия Ирвина для уровня значимости  $\alpha = 0,05$  приведены в табл. 1.

Таблица 1

n	2	3	10	20	30	50	100
$\lambda_\alpha$	2,8	2,3	1,5	1,3	1,2	1,1	1,0

# Временные ряды. Выявление тренда

Если во временном ряду проявляется длительная тенденция изменения показателя, то в этом случае говорят, что имеет место **тренд**.

Пусть дан временной ряд  $y_1, y_2, \dots, y_t, y_n, t = 1, 2, \dots, n$ .

Считают, что значения уровней временных рядов показателей складываются из следующих составляющих (компонент): **тренда, сезонной, циклической и случайной**.

Под **трендом** понимают изменение, определяющее общее направление развития или основную тенденцию временного ряда. Тренд относят к систематической составляющей долговременного действия. Во временных рядах часто происходят регулярные колебания, которые относятся к периодическим составляющим рядов экономических процессов.

- Если период колебаний не превышает 1 года, то их называют *сезонными*, более 1 года – *циклическими составляющими*.
- Чаще всего причиной сезонных колебаний являются природные, климатические условия, циклических – демографические циклы др.
- Тренд, сезонная и циклическая составляющая называются *регулярными*, или *систематическими компонентами временного ряда*.
- Если из временного ряда удалить регулярную компоненту, то останется *случайная компонента*.

# Временные ряды. Выявление тренда

Если временной ряд представлен в виде суммы составляющих компонент, то модель называется *аддитивной*, если в виде произведения, то *мультипликативной* или *смешанного* типа.

- $y_t = u_t + s_t + v_t + e_t$  — аддитивная форма,
- $y_t = u_t s_t v_t e_t$  — мультипликативная форма,
- $y_t = u_t s_t v_t + e_t$  — смешанная форма,

где  $y_t$  — уровни временного ряда;

$u_t$  — временной тренд;

$s_t$  — сезонная компонента;

$v_t$  — циклическая составляющая;

$e_t$  — случайная компонента.



# Временные ряды. Критерий восходящих серий

Прогнозирование временных рядов целесообразно начинать с построения графика исследуемого показателя. Однако в нем не всегда прослеживается присутствие тренда. Поэтому в этих случаях необходимо выяснить, существует ли тенденция во временном ряду или она отсутствует.

Дан временной ряд:  $y_1, y_2, \dots, y_i, y_n, t = 1, 2, \dots, n$ .

Рассмотрим *критерий «восходящих и нисходящих» серий*, согласно которому тенденция определяется по следующему алгоритму.

1. Для исследуемого временного ряда определяется последовательность знаков, исходя из условий

+, если  $y_{t+1} - y_t > 0$ ,

-, если  $y_{t+1} - y_t < 0$ .

# Временные ряды. Критерий восходящих серий

При этом, если последующее наблюдение равно предыдущему, то учитывается только одно наблюдение.

2. Подсчитывается число серий  $v(n)$ . Под серией понимается последовательность подряд расположенных плюсов или минусов, причем один плюс или один минус считается серией.
3. Определяется протяженность самой длинной серии  $l_{\max}(n)$
4. По табл. 2 находится значение  $l(n)$ .

Таблица 2

Длина ряда (n)	$n \leq 26$	$26 < n \leq 153$	$153 < n \leq 170$
Значение $l(n)$	5	6	7

5. Если нарушается хотя бы одно из следующих неравенств, то гипотеза об отсутствии тренда отвергается с доверительной вероятностью 0,95.

$$v(n) > \left[ (2n - 1)/3 - 1,96\sqrt{16n - 29/90} \right], \quad l_{\max}(n) \leq l(n)$$

# Временные ряды. Критерий восходящих серий

При этом, если последующее наблюдение равно предыдущему, то учитывается только одно наблюдение.

2. Подсчитывается число серий  $v(n)$ . Под серией понимается последовательность подряд расположенных плюсов или минусов, причем один плюс или один минус считается серией.
3. Определяется протяженность самой длинной серии  $l_{\max}(n)$
4. По табл. 2 находится значение  $l(n)$ .

Таблица 2

Длина ряда (n)	$n \leq 26$	$26 < n \leq 153$	$153 < n \leq 170$
Значение $l(n)$	5	6	7

5. Если нарушается хотя бы одно из следующих неравенств, то гипотеза об отсутствии тренда отвергается с доверительной вероятностью 0,95.

$$v(n) > \left[ (2n - 1)/3 - 1,96\sqrt{16n - 29/90} \right], \quad l_{\max}(n) \leq l(n)$$

# Анализ временных рядов. Общие положения

*Анализ временных рядов объединяет методы изучения временных рядов, как пытающиеся понять природу точек данных (откуда они взялись? что их породило?), так и пытающиеся построить прогноз.*

**Прогнозирование временных рядов** заключается в построении модели для предсказания будущих событий основываясь на известных событиях прошлого, предсказания будущих данных до того как они будут измерены. *Типичный пример — предсказание цены открытия биржи основываясь на предыдущей её деятельности.*

Понятие **анализ временных рядов** используется для того, чтобы

- отделить эту задачу от в первую очередь от более простых задач анализа данных (когда нет естественного порядка поступления наблюдений)
- и, во-вторых, от анализа пространственных данных, в котором наблюдения зачастую связаны с географическим положением.

*Модель временного ряда в общем смысле отражает идею, что близкие во времени наблюдения будут теснее связаны, чем удалённые.*

Кроме того, **модели временных рядов зачастую используют однонаправленный порядок по времени** в том смысле, что значения в ряду выражаются в некотором виде через прошлые значения, а не через последующие

# Анализ временных рядов. Общие положения

Методы анализа временных рядов зачастую делят на два класса: **анализ в частотной области** и **анализ во временной области**.

- **Первый основывается на спектральном анализе** и с недавних пор вейвлетном анализе, и может рассматриваться в качестве не использующих модели методов анализа, хорошо подходящих для исследований на этапе разведки.
- **Методы анализа во временной области** также имеют безмодельное подмножество, состоящее из кросс-корреляционного анализа и автокорреляционного анализа, но именно здесь появляются частично и полностью определённые модели временных рядов.

# Анализ временных рядов. Общие положения

**Существует несколько методов анализа данных, применимых для временных рядов.**

## Общее исследование

- Визуальное изучение графических представлений временных рядов
- Автокорреляционный анализ для изучения зависимостей
- Спектральный анализ для изучения циклического поведения, не связанного с сезонностью

## Описание

- Разделение компонент: тренд, сезонность, медленно и быстро меняющиеся компоненты, циклическая нерегулярность
- Простейшие свойства частных распределений

## Прогнозирование и предсказание

- Полноценные статистические модели при стохастическом моделировании для создания альтернативных версий временных рядов, показывающих, что могло бы случиться на произвольных отрезках времени в будущем (предсказание)
- Упрощённые или полноценные статистические модели для описания вероятные значения временного ряда в ближайшем будущем при известных последних значениях (прогноз)

# Анализ временных рядов. Общие положения

## Модели временных рядов

Как показано Боксом и Дженкинсом, модели временных рядов могут иметь различные формы и представлять различные стохастические процессы.

При моделировании изменений уровня процесса можно выделить три широких класса имеющих практическую ценность:

***авторегрессионные модели, интегральные модели и модели скользящего среднего.***

Эти три класса линейно зависят от предшествующих данных. На их основе построены модели авторегрессионного скользящего среднего (Autoregressive Moving Average, ARMA) и авторегрессионного интегрированного скользящего среднего (Autoregressive Integrated Moving Average, ARIMA).

# Анализ временных рядов. Общие положения

## Модели временных рядов

- Эти модели в свою очередь обобщает модель авторегрессионного дробноинтегрированного скользящего среднего (autoregressive fractionally integrated moving average, ARFIMA).
- Расширения моделей на случаи, когда данные представляются не скалярно, а векторно, называют моделями многомерных временных рядов. Для таких моделей в сокращённых названиях появляется буква «v» от слова «vector».
- Существуют расширения моделей на случай, когда исследуемый временной ряд является ведомым для некоторого «вынуждающего» ряда (который, однако, может не быть причиной возникновения исследуемого ряда).
- Отличие от многомерного ряда заключается в том, что вынуждающий ряд может быть детерминированным или управляться исследователем, проводящим эксперимент.
- Для таких моделей в сокращении появляется буква «x» от «exogenous» (экзогенный, вызываемый внешними причинами).



# Анализ временных рядов. Общие положения

## Модели временных рядов

- Нелинейная зависимость уровня ряда от предыдущих точек интересна, отчасти из-за возможности генерации хаотических временных рядов. Но главным всё же является то, что опытные исследования указывают на превосходство прогнозов, полученных от нелинейных моделей, над прогнозами линейных моделей.
- Среди прочих типов нелинейных моделей временных рядов можно выделить модели, описывающие изменения дисперсии ряда со временем (гетероскедастичность).
- Такие модели называют моделями авторегрессионной условной гетероскедастичности (AutoRegressive Conditional Heteroscedasticity, ARCH). К ним относится большое количество моделей: GARCH, TARARCH, EGARCH, FIGARCH, CGARCH и др.

# Анализ временных рядов. Общие положения

## Модели временных рядов

- В этих моделях изменения дисперсии связывают с ближайшими предшествующими данными.
- Противовесом такому подходу является представление локально изменчивой дисперсии, при котором дисперсия может быть смоделирована зависящей от отдельного меняющегося со временем процесса, как в бистохастических моделях.
- В последнее время значительное внимание снискали исследования в области безмодельного анализа и методы, основанные на вейвлетных преобразованиях (например локально стационарные вейвлеты) в частности.
- Методы многомасштабного анализа разлагают заданный временной ряд на составные части, чтобы показать зависимость от времени с разным масштабом.