



Лекция 2

Анализ временных рядов и прогнозирование

Курс лекций:

«Прогнозирование и моделирование в бизнесе»

(с использованием Microsoft Excel)

Подготовил:

к.э.н. Литвин Ю.В.

Цель и основные понятия

Цель прогнозирования: предсказание событий, которые произойдут в будущем, чтобы учесть их при разработке планов и стратегии развития компании

Временной ряд: набор числовых данных, полученных в течение последовательных периодов времени

Основное предположение: факторы, влияющие на исследуемый объект в прошлом, сохраняют свое влияние и в будущем

Классическая аддитивная модель временного ряда

Значение в будущем

=

Тренд

+

Сезонный компонент

+

Циклический компонент

+

Случайный компонент



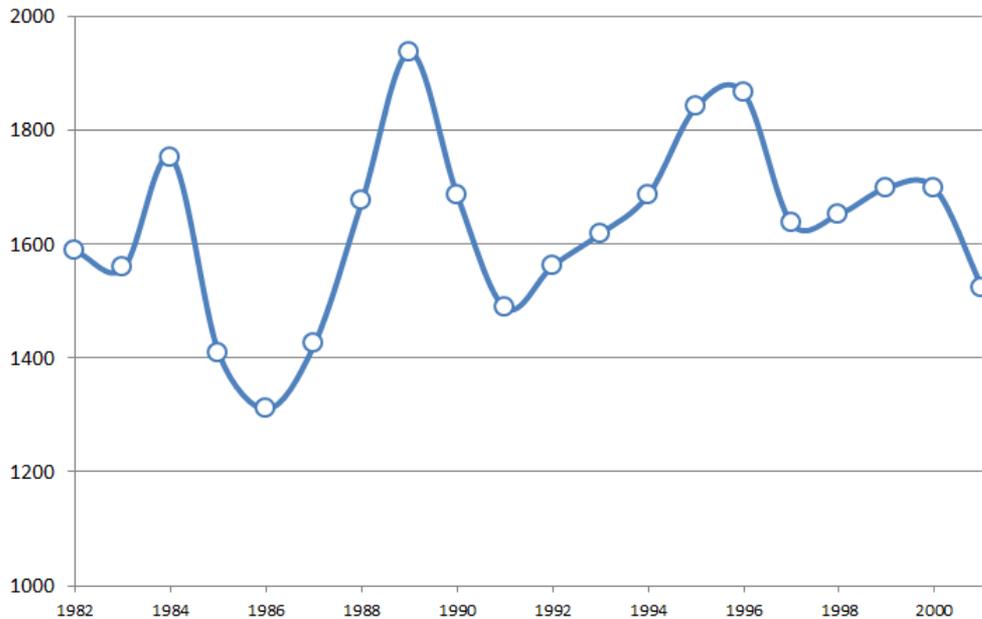
*Долговременное возрастание или убывание данных
(изменение среднего с течением времени)*

«Сглаживание» годовых временных рядов

Цель «сглаживания» :

- 1) выявление тренда в имеющихся данных;
- 2) построение прогнозов в условиях отсутствия тренда (сохраняющегося среднего значения);
- 3) использование для более сложных моделей прогнозирования.

Доходы компании *Cabot Corporation* в период 1982-2001 г. , млн долл
(специализируется на продаже химикатов, строительных материалов)

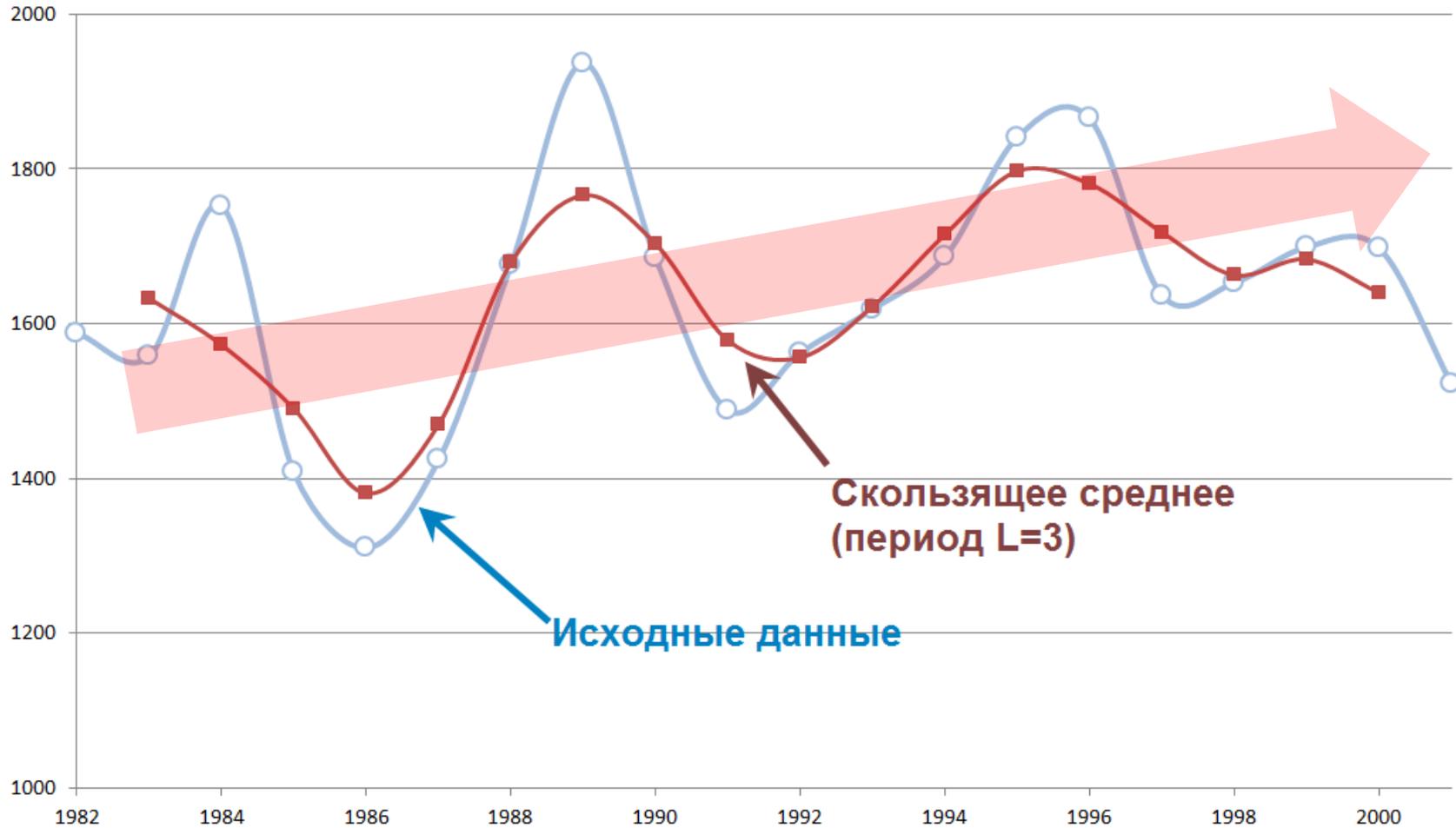


Вопросы:

- 1 Какова существующая долговременная тенденция в изменении доходов компании (рост / падение)?
- 2 Каков предполагаемый уровень доходов в 2002 году?

Метод «скользящего среднего» (1)

Доходы компании *Cabot Corporation* в период 1982-2001 г. , млн долл
(специализируется на продаже химикатов, строительных материалов)

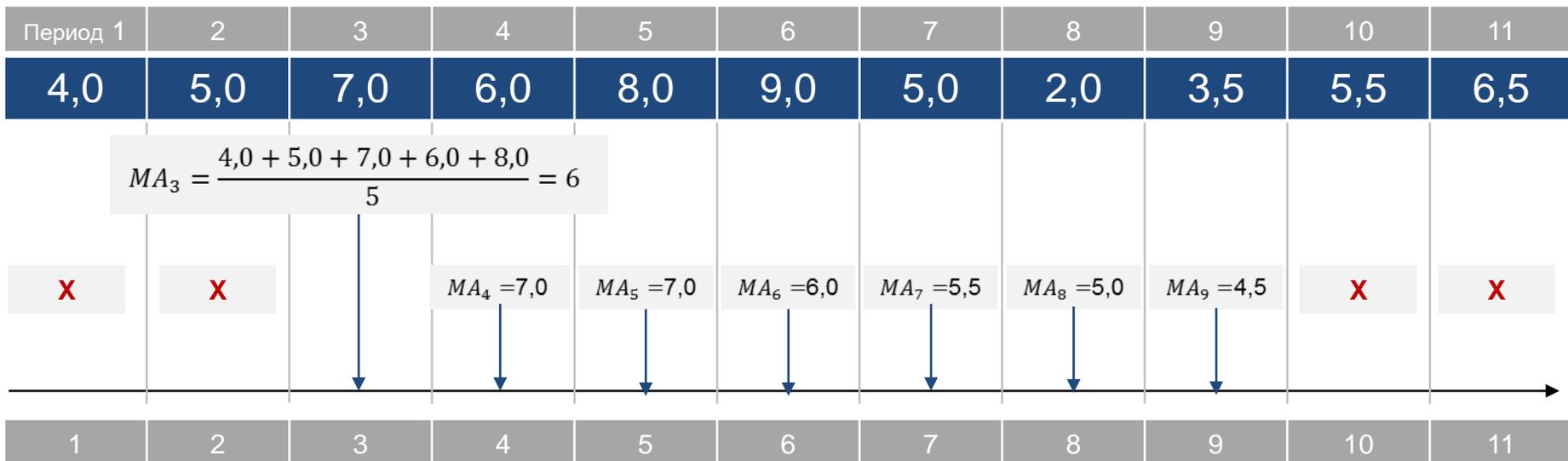


Метод «скользящего среднего» (2)

$$MA_i(L) = \frac{Y_1 + Y_2 + \dots + Y_L}{L}$$

где $MA_i(L)$ – скользящее среднее за L периодов, рассчитанное для периода i ;
 Y_j – значение показателя временного ряда в период j .

Например (требуется вычислить «скользящее» среднее за $L=5$ периодов)



Задача (1)

Вычислите первые 3 «скользящих средних» ($MA(L)$) для доходов компании Cabot Corporation (в млн долл.) за $L=3$ -х и $L=7$ -ми летний периоды. К каким периодам они будут отнесены?

Самостоятельно

Год	Значение
1982	1588
1983	1558
1984	1752
1985	1408
1986	1310
1987	1424
1988	1676
1989	1937
1990	1685
1991	1488
1992	1562
1993	1619
1994	1687
1995	1841
1996	1865
1997	1637
1998	1653
1999	1699
2000	1698
2001	1523

Самостоятельно

Метод экспоненциального сглаживания (1)

$$E_1 = Y_1$$

$$E_i = W \cdot Y_i + (1 - W) \cdot E_{i-1}$$

Прогноз

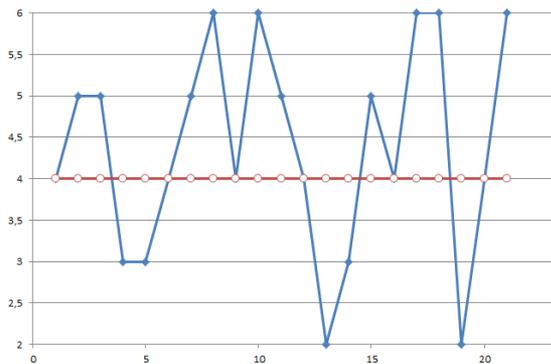
$$\widehat{E}_{N+q} = E_N$$

где E_i – значение экспоненциально сглаженного ряда для периода $i \in [1, N]$, N – число периодов;
 Y_i – фактическое значение показателя временного ряда в период i ;
 W – субъективный вес («сглаживающий» коэффициент), $[0, 1]$;
 q – период прогнозирования;

Главная проблема – определение W

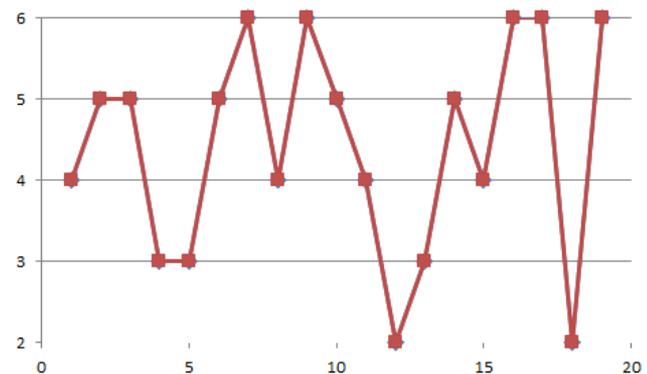
$W \rightarrow 0$: в большей мере учитывает долгосрочные тенденции

$$W = 0 \rightarrow E_i = E_{i-1}$$



$W \rightarrow 1$: с наибольшим весом учитываются последние значения

$$W = 1 \rightarrow E_i = Y_i$$



Метод экспоненциального сглаживания (2)

Например, требуется вычислить экспоненциально сглаженное значение для $W=0,5$

1	2	3	4	5	6	7	8	9	10	11
4,0	5,0	7,0	6,0	8,0	9,0	5,0	2,0	3,5	5,5	6,5
	$E_2 = 0,5 \cdot 4 + (1-0,5) \cdot 5 = 4,50$									
$E_1 = 4,00$		$E_3 = 5,75$	$E_4 = 5,88$	$E_5 = 6,94$	$E_6 = 7,97$	$E_7 = 6,48$	$E_8 = 4,24$	$E_9 = 3,87$	$E_{10} = 4,69$	$E_{11} = 5,59$
1	2	3	4	5	6	7	8	9	10	11

Чему равен прогноз на 12-15 периоды?

Задача (2)

Вычислите первые 3 экспоненциально сглаженных значений для доходов компании *Cabot Corporation* (в млн долл.) при $W=0,5$ и $W=0,25$. Чему равно значение дохода в 2002 г? Можно ли в этих условиях построить прогноз на 2003-2005 года?

Самостоятельно

Год	Значение
1982	1588
1983	1558
1984	1752
1985	1408
1986	1310
1987	1424
1988	1676
1989	1937
1990	1685
1991	1488
1992	1562
1993	1619
1994	1687
1995	1841
1996	1865
1997	1637
1998	1653
1999	1699
2000	1698
2001	1523

Самостоятельно

Задача (3)

Приведенные ниже данные представляют количество сотрудников (в тыс. чел.) нефтедобывающей компании в период 1983-2002 г.

Год	Кол-во	Год	Кол-во
1983	1,45	1993	1,73
1984	1,55	1994	1,77
1985	1,61	1995	1,90
1986	1,60	1996	1,82
1987	1,74	1997	1,65
1988	1,92	1998	1,73
1989	1,95	1999	1,88
1990	2,04	2000	2,00
1991	2,06	2001	2,08
1992	1,80	2002	1,88

Вопросы:

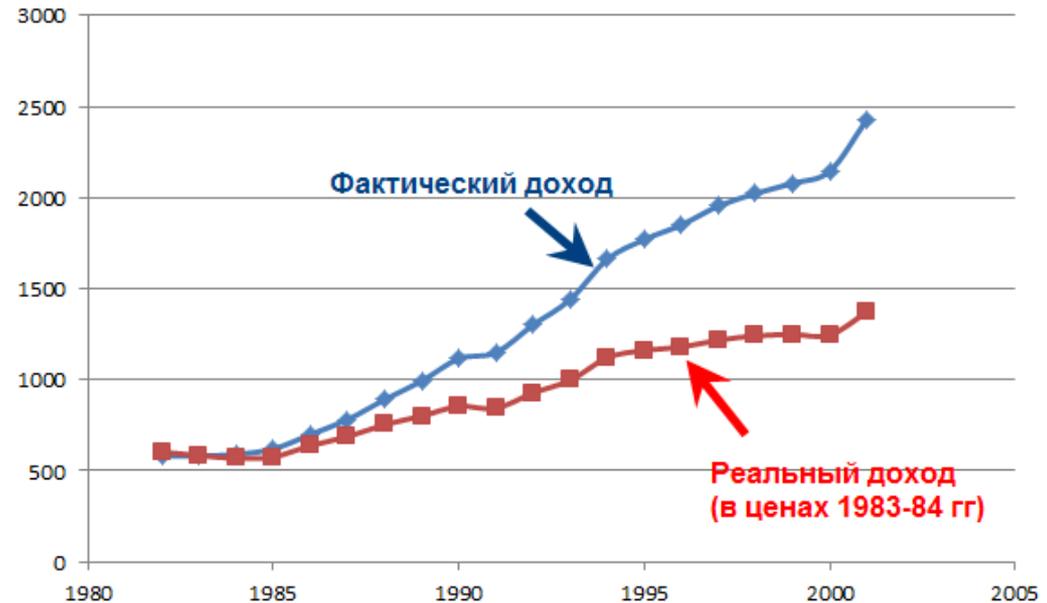
1. Постройте график временного ряда
2. Вычислите пятилетнее скользящее среднее **MA(5)** и построьте его график
3. Примените экспоненциальное сглаживание **E** с **W=0,5** и постройте его график
4. Сделайте выводы по проведенным исследованиям

Использование метода наименьших квадратов в вычислении трендов (1)

Доходы Wrigley Company, млн долл

Год	Фактический доход	Индекс цен (1983-84)	Реальный доход
1982	581,5	96,5	602,6
1983	581,7	99,6	584,0
1984	590,5	103,9	568,3
1985	620,3	107,6	576,5
1986	699	109,6	637,8
1987	781,1	113,6	687,6
1988	891,4	118,3	753,5
1989	992,9	124	800,7
1990	1116	130,7	853,9
1991	1148,9	136,2	843,5
1992	1301,3	140,3	927,5
1993	1440,4	144,5	996,8
1994	1661,3	148,2	1121,0
1995	1769,7	152,4	1161,2
1996	1850,6	156,9	1179,5
1997	1954,2	160,5	1217,6
1998	2023,4	163	1241,3
1999	2079,2	166,6	1248,0
2000	2145,7	172,2	1246,1
2001	2429,6	177,1	1371,9

Доходы Wrigley Company, млн долл



Использование метода наименьших квадратов в вычислении трендов (2)

Модель линейного тренда

$$Y = b_0 + b_1 \cdot x$$

где

Y – значение показателя временного ряда в период ;

b_1 – выборочный наклон; b_0 – «сдвиг»;

x – номер наблюдения, причем $x=0$ соответствует первому наблюдению

Использование метода наименьших квадратов в вычислении трендов (3)

Доходы Wrigley Company, млн долл

Реальный доход	Период (X)
602,6	0
584,0	1
568,3	2
576,5	3
637,8	4
687,6	5
753,5	6
800,7	7
853,9	8
843,5	9
927,5	10
996,8	11
1121,0	12
1161,2	13
1179,5	14
1217,6	15
1241,3	16
1248,0	17
1246,1	18
1371,9	19

Метод наименьших квадратов

Модель линейного тренда

$$Y = 498,95 + 45,49 \cdot x$$

где

498,95 – «сдвиг» (доходы в 0 период) ;

45,49 – среднее ежегодное увеличение доходов в период 0-19;

x – значение периода прогнозирования [0; 19]

$$r^2 = 0,968$$

Использование метода наименьших квадратов в вычислении трендов (4)

Доходы Wrigley Company, млн долл

Период (X)	Реальный доход (Y)	Предсказанное значение (Y')
0	602,6	499,0
1	584,0	544,4
2	568,3	589,9
3	576,5	635,4
4	637,8	680,9
5	687,6	726,3
6	753,5	771,8
7	800,7	817,3
8	853,9	862,8
9	843,5	908,2
10	927,5	953,7
11	996,8	999,2
12	1121,0	1044,7
13	1161,2	1090,1
14	1179,5	1135,6
15	1217,6	1181,1
16	1241,3	1226,6
17	1248,0	1272,0
18	1246,1	1317,5
19	1371,9	1363,0

Модель линейного тренда



Спрогнозируйте значение доходов в 20 и 21 периодах. Каким годам они соответствуют?

Использование метода наименьших квадратов в вычислении трендов (5)

Модель квадратичного тренда

$$Y = b_0 + b_1 \cdot x + b_2 \cdot x^2$$

где

Y – значение показателя временного ряда;

b_0 – «сдвиг»; b_1 – линейный эффект; b_2 – квадратичный эффект;

x – номер наблюдения, причем $x=0$ соответствует первому наблюдению

Использование метода наименьших квадратов в вычислении трендов (6)

Доходы Wrigley Company, млн долл

Реальный доход (Y)	Период (X)
602,6	0
584,0	1
568,3	2
576,5	3
637,8	4
687,6	5
753,5	6
800,7	7
853,9	8
843,5	9
927,5	10
996,8	11
1121,0	12
1161,2	13
1179,5	14
1217,6	15
1241,3	16
1248,0	17
1246,1	18
1371,9	19

Метод наименьших квадратов

Модель квадратичного тренда

$$Y = 512,93 + 40,815 \cdot x + 0,2453 \cdot x^2$$

где

512,93 – «сдвиг» (доходы в 0 период) ;

40,815 – среднее ежегодное линейное увеличение доходов в период 0-19;

0,2453 – среднее ежегодное квадратичное увеличение доходов в период 0-19;

x – значение периода прогнозирования [0; 19]

$$r^2 = 0,965$$

Использование метода наименьших квадратов в вычислении трендов (7)

Доходы Wrigley Company, млн долл

Период (X)	Реальный доход (Y)	Предсказанное значение (Y')
0	602,6	512,9
1	584,0	554,0
2	568,3	595,5
3	576,5	637,6
4	637,8	680,1
5	687,6	723,1
6	753,5	766,7
7	800,7	810,7
8	853,9	855,1
9	843,5	900,1
10	927,5	945,6
11	996,8	991,6
12	1121,0	1038,0
13	1161,2	1085,0
14	1179,5	1132,4
15	1217,6	1180,3
16	1241,3	1228,8
17	1248,0	1277,7
18	1246,1	1327,1
19	1371,9	1377,0

Модель квадратичного тренда



Спрогнозируйте значение доходов в 20 и 21 периодах. Каким годам они соответствуют?

Использование метода наименьших квадратов в вычислении трендов (8)

Модель экспоненциального тренда

$$Y = b_0 \cdot e^{b_1 \cdot x}$$

где

Y – значение показателя временного ряда;

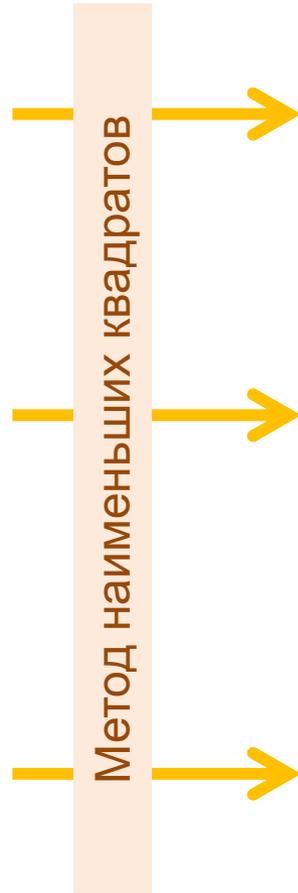
b_0 – сдвиг; e – константа = 2,71; b_1 – константа;

x – номер наблюдения, причем $x=0$ соответствует первому наблюдению

Использование метода наименьших квадратов в вычислении трендов (9)

Доходы Wrigley Company, млн долл

Реальный доход (Y)	Период (X)
602,6	0
584,0	1
568,3	2
576,5	3
637,8	4
687,6	5
753,5	6
800,7	7
853,9	8
843,5	9
927,5	10
996,8	11
1121,0	12
1161,2	13
1179,5	14
1217,6	15
1241,3	16
1248,0	17
1246,1	18
1371,9	19



Модель экспоненциального тренда

$$Y = 550,38 \cdot e^{0,0508 \cdot x}$$

где

550,38 – «сдвиг» (доходы в 0 период) ;

0,0508 – ежегодное увеличение доходов в период 0-19;

x – значение периода прогнозирования.

$$r^2 = 0,962$$

Использование метода наименьших квадратов в вычислении трендов (10)

Доходы Wrigley Company, млн долл

Период (X)	Реальный доход (Y)	Предсказанное значение (Y')
0	602,6	550,4
1	584,0	579,1
2	568,3	609,2
3	576,5	641,0
4	637,8	674,4
5	687,6	709,5
6	753,5	746,5
7	800,7	785,4
8	853,9	826,3
9	843,5	869,4
10	927,5	914,7
11	996,8	962,4
12	1121,0	1012,5
13	1161,2	1065,3
14	1179,5	1120,8
15	1217,6	1179,2
16	1241,3	1240,7
17	1248,0	1305,3
18	1246,1	1373,3
19	1371,9	1444,9

Модель экспоненциального тренда



Спрогнозируйте значение доходов в 20 и 21 периодах. Каким годам они соответствуют?

Приведенные ниже данные представляют объем продаж ресторана быстрого питания в период 1992-2001 г.

Год	Объем продаж, тыс. ед.
1992	1,45
1993	1,55
1994	1,61
1995	1,60
1996	1,74
1997	1,92
1998	1,95
1999	2,04
2000	2,06
2001	1,80

Вопросы:

1. Постройте график временного ряда
2. Используя скользящее среднее ($MA(3)$), сделайте выводы о будущих тенденциях объема продаж
3. Используя метод экспоненциального сглаживания (E), сделайте выводы относительно будущих тенденций объема продаж ($W=0,5$).
4. Постройте уравнение линейного тренда и сделайте выводы;
5. Постройте уравнение квадратичного тренда и сделайте выводы;
6. Постройте уравнение экспоненциального тренда и сделайте выводы;
7. Сделайте выводы о предполагаемом значении объема продаж в 2002 г.

Вычисление тренда при помощи авторегрессии (1)

Часто при построении прогнозных моделей можно заметить, что значения временного ряда достаточно сильно коррелируют с предшествующими и последующими значениями

Автокорреляция 1-ого порядка – зависимость между последовательными значениями временного ряда

$$Y_t = A_0 + A_1 \cdot Y_{t-1}$$

Пример

31	34	37	40	43	46	49	?
----	----	----	----	----	----	----	---

$$Y_t = 3 + 1 \cdot Y_{t-1}$$

Автокорреляция 2-ого порядка – зависимость между последовательными значениями временного ряда, разделенными 2 интервалами

$$Y_t = A_0 + A_1 \cdot Y_{t-1} + A_2 \cdot Y_{t-2}$$

Автокорреляция p-ого порядка – зависимость между последовательными значениями временного ряда, разделенными p интервалами

$$Y_t = A_0 + A_1 \cdot Y_{t-1} \dots + A_p \cdot Y_{t-p}$$

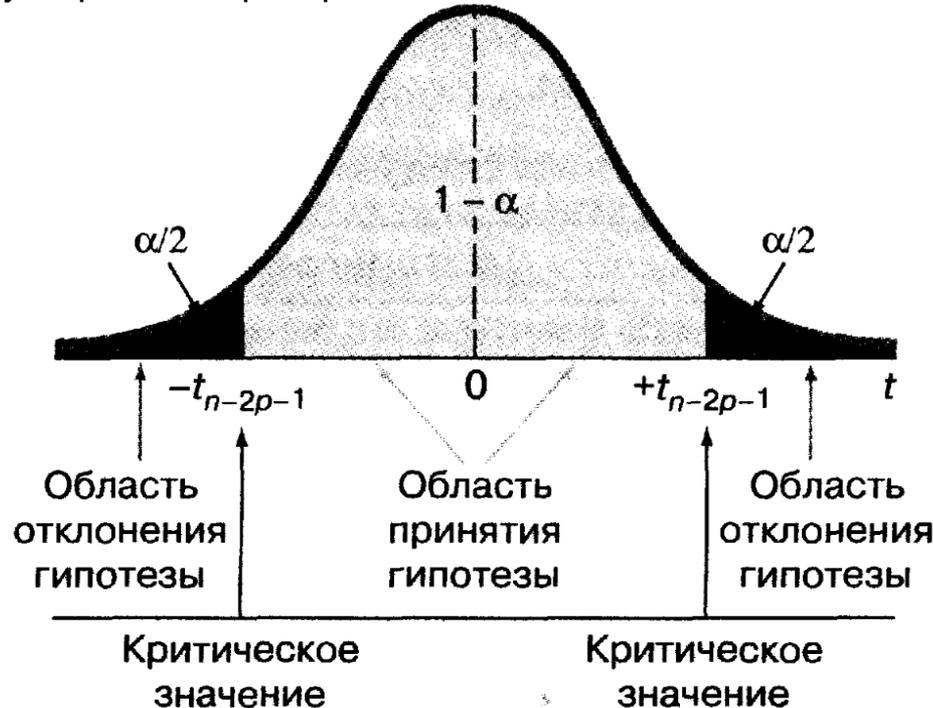
Сколько порядков (прошлых периодов) следует использовать в авторегрессионной модели?

$$H_0: A_p = 0$$
$$H_1: A_p \neq 0$$

$$t = \frac{a_p - A_p}{S_{a_p}}$$

где A_p – гипотетическое значение ($A_p = 0$),
 a_p – оценка параметра авторегрессии A_p , S_{a_p} – стандартная ошибка оценки

Области принятия и отклонения гипотезы для двустороннего критерия значимости



Алгоритм

- 1 Выберите наивысший порядок коэффициента p в регрессионной модели
- 2 Оцените число «степеней свободы» как $n-2p-1$ (n – число измерений) и значимость параметра A_p , имеющий наивысший приоритет по t -критерию
- 3 Если нулевая гипотеза принимается, удалите из модели a_p и вернитесь к шагу 1 с коэффициентом $p-1$.

В противном случае расчет завершен!

Пример построения авторегрессионной модели (1)

Исходные данные

Год	Реальный доход	Лаг t-1	Лаг t-2	Лаг t-3
1982	602,59			
1983	584,04	602,59		
1984	568,33	584,04	602,59	
1985	576,49	568,33	584,04	602,59
1986	637,77	576,49	568,33	584,04
1987	687,59	637,77	576,49	568,33
1988	753,51	687,59	637,77	576,49
1989	800,73	753,51	687,59	637,77
1990	849,73	800,73	753,51	687,59
1991	843,54	849,73	800,73	753,51
1992	927,51	843,54	849,73	800,73
1993	996,82	927,51	843,54	849,73
1994	1120,99	996,82	927,51	843,54
1995	1161,22	1120,99	996,82	927,51
1996	1179,48	1161,22	1120,99	996,82
1997	1217,57	1179,48	1161,22	1120,99
1998	1241,35	1217,57	1179,48	1161,22
1999	1248,02	1241,35	1217,57	1179,48
2000	1246,05	1248,02	1241,35	1217,57
2001	1371,08	1246,05	1248,02	1241,35

Исключение из анализа

$$Y_t = 54,763 + 1,0687 \cdot Y_{t-1} - 0,0730 \cdot Y_{t-2} - 0,0073 \cdot Y_{t-3}$$

$$r^2 = 0,975$$

Переменная	Стандартная ошибка оценки (S_{a_i})
a_{t-1}	0,330
a_{t-2}	0,541
a_{t-3}	0,325

Пример построения авторегрессионной модели (2)

$$Y_t = 54,763 + 1,0687 \cdot Y_{t-1} - 0,0730 \cdot Y_{t-2} - 0,0073 \cdot Y_{t-3}$$

1 Задача: анализ значимости a_{t-3}

2 Вычисляем t -статистику:

$$t = \frac{a_{t-3} - A_{t-3}}{s_{a_{t-3}}} = \frac{-0,0073 - 0}{0,325} = -0,022$$

$$H_0: A_{t-3} = 0$$

$$H_1: A_{t-3} \neq 0$$

Вычисляем число степеней свободы:

$$n - (2 \cdot p) - 1 = 20 - (2 \cdot 3) - 1 = 13$$

По таблице оцениваем критические значения двустороннего t -критерия при уровне значимости $\alpha = 0,05$: $t_{крит} = \pm 2,1604$.

Т.к. $-0,022 \in [-2,1604; +2,1604]$, то H_0 отклонить нельзя

Таким образом, коэффициент третьего порядка a_{t-3} не имеет статистической значимости и должен быть исключен

Пример построения авторегрессионной модели (3)

Исходные данные

Год	Реальный доход	Лag t-1	Лag t-2
1982	602,59		
1983	584,04	602,59	
1984	568,33	584,04	602,59
1985	576,49	568,33	584,04
1986	637,77	576,49	568,33
1987	687,59	637,77	576,49
1988	753,51	687,59	637,77
1989	800,73	753,51	687,59
1990	849,73	800,73	753,51
1991	843,54	849,73	800,73
1992	927,51	843,54	849,73
1993	996,82	927,51	843,54
1994	1120,99	996,82	927,51
1995	1161,22	1120,99	996,82
1996	1179,48	1161,22	1120,99
1997	1217,57	1179,48	1161,22
1998	1241,35	1217,57	1179,48
1999	1248,02	1241,35	1217,57
2000	1246,05	1248,02	1241,35
2001	1371,08	1246,05	1248,02

Исключение из анализа

$$Y_t = 31,642 + 1,2119 \cdot Y_{t-1} - 0,206 \cdot Y_{t-2}$$

$$r^2 = 0,976$$

Переменная	Стандартная ошибка оценки (S_{a_i})
a_{t-1}	0,2725
a_{t-2}	0,2753

Пример построения авторегрессионной модели (4)

$$Y_t = 31,642 + 1,2119 \cdot Y_{t-1} - 0,206 \cdot Y_{t-2}$$

1 Задача: анализ значимости a_{t-2}

2 Вычисляем *t*-статистику:

$$t = \frac{a_{t-2} - A_{t-2}}{s_{a_{t-2}}} = \frac{-0,206 - 0}{0,2753} = -0,748$$

$$H_0: A_{t-2} = 0$$

$$H_1: A_{t-2} \neq 0$$

Вычисляем число степеней свободы:

$$n - 2p - 1 = 20 - (2 \cdot 2) - 1 = 15$$

По таблице оцениваем критические значения двустороннего *t*-критерия при уровне значимости $\alpha = 0,05$: $t_{крит} = \pm 2,1315$.

Т.к. $-0,748 \in [-2,1315; +2,1315]$, то H_0 отклонить нельзя

Таким образом, коэффициент второго порядка a_{t-2} не имеет статистической значимости и должен быть исключен

Пример построения авторегрессионной модели (5)

Исходные данные

Год	Реальный доход	Лag t-1
1982	602,59	
1983	584,04	602,59
1984	568,33	584,04
1985	576,49	568,33
1986	637,77	576,49
1987	687,59	637,77
1988	753,51	687,59
1989	800,73	753,51
1990	849,73	800,73
1991	843,54	849,73
1992	927,51	843,54
1993	996,82	927,51
1994	1120,99	996,82
1995	1161,22	1120,99
1996	1179,48	1161,22
1997	1217,57	1179,48
1998	1241,35	1217,57
1999	1248,02	1241,35
2000	1246,05	1248,02
2001	1371,08	1246,05

Исключение из анализа

$$Y_t = 18,4217 + 1,0242 \cdot Y_{t-1}$$

$$r^2 = 0,976$$

Переменная	Стандартная ошибка оценки (S_{a_i})
a_{t-1}	0,03874

Пример построения авторегрессионной модели (6)

$$Y_t = 18,4217 + 1,0242 \cdot Y_{t-1}$$

1 Задача: анализ значимости a_{t-1}

2 Вычисляем *t*-статистику:

$$t = \frac{a_{t-1} - A_{t-1}}{s_{a_{t-1}}} = \frac{+1,0242 - 0}{0,03874} = 26,437$$

$$H_0: A_{t-1} = 0$$

$$H_1: A_{t-1} \neq 0$$

Вычисляем число степеней свободы:

$$n - 2p - 1 = 20 - (2 \cdot 1) - 1 = 17$$

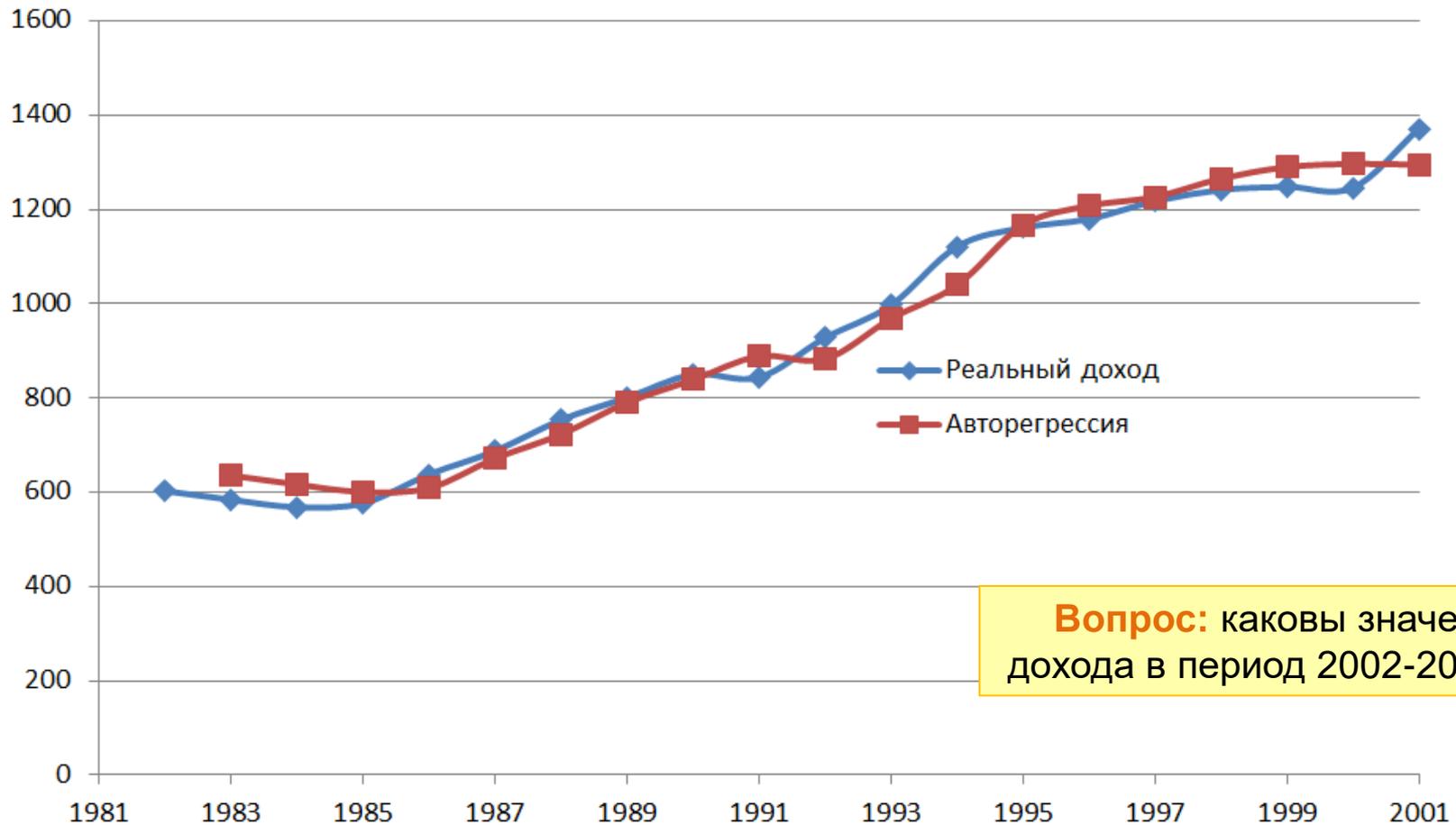
По таблице оцениваем критические значения двустороннего *t*-критерия при уровне значимости $\alpha = 0,05$: $t_{крит} = \pm 2,1098$.

Т.к. $26,437 \notin [-2,1315; +2,1315]$, то H_0 отклоняется

Таким образом, коэффициент первого порядка a_{t-1} обладает статистической значимостью

Пример построения авторегрессионной модели (7)

$$Y_t = 18,4217 + 1,0242 \cdot Y_{t-1}$$



Вопрос: каковы значения дохода в период 2002-2005 гг.?

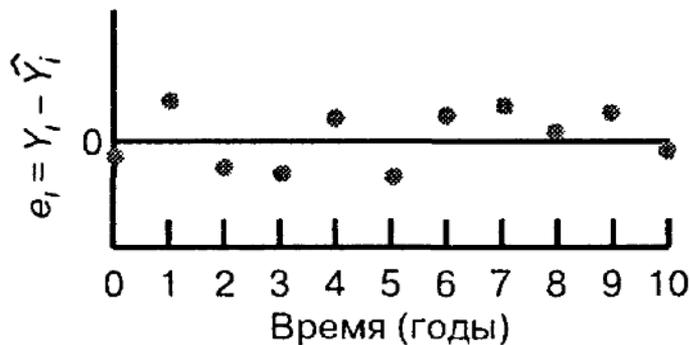
Приведенные ниже данные представляют объем продаж ресторана быстрого питания в период 1992-2001 г.

Год	Объем продаж, ед.
1992	1,45
1993	1,55
1994	1,61
1995	1,60
1996	1,74
1997	1,92
1998	1,95
1999	2,04
2000	2,06
2001	1,80

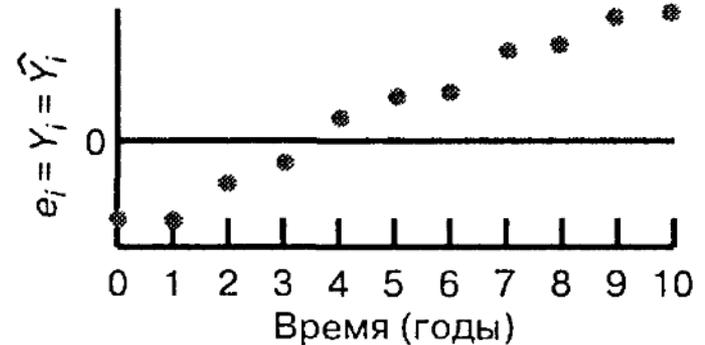
Вопросы:

1. Используя авторегрессию, начиная с $p=2$, постройте соответствующее уравнение, а также сделайте выводы относительно значимости коэффициентов в авторегрессионной модели
2. Постройте исходный временной ряд, а также авторегрессионную модель, полученную в п.1
3. Сделайте выводы о предполагаемом значении объема продаж в 2002-2005 г.

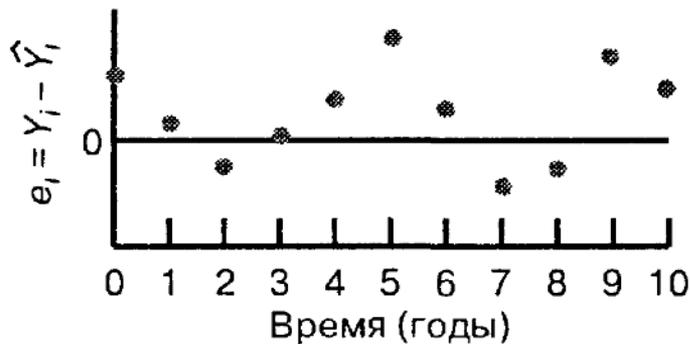
Анализ остатков



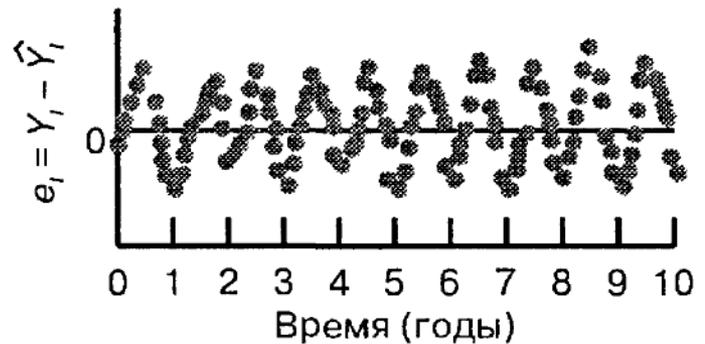
Панель А
Случайно распределенные
ошибки прогнозирования



Панель Б
Не учитывается тренд



Панель В
Не учитывается циклический компонент



Панель Г
Не учитывается сезонный компонент

Измерение абсолютной и относительной погрешности прогнозирования

1 Среднее абсолютное отклонение (MAD)

$$MAD = \frac{\sum_{i=1}^n |Y_i - \hat{Y}_i|}{n}$$

2 Среднеквадратичная ошибка оценки (S_{YX})

$$S_{YX} = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - 2}}$$

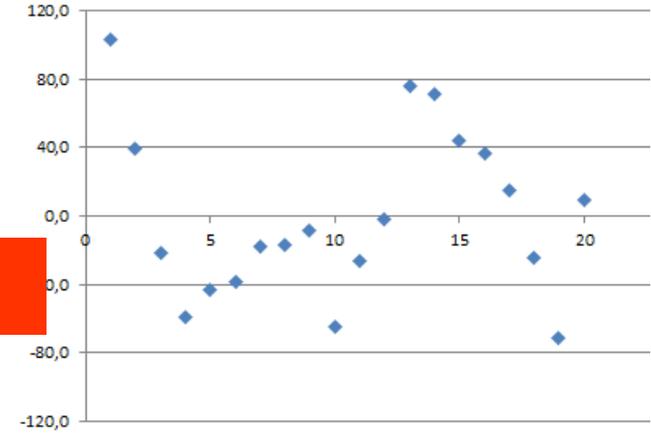
3 Принцип экономии

При равенстве (близких) значений MAD и S_{YX} следует выбирать наиболее простую модель

Сравнение 4-х моделей прогнозирования по уровню ошибки (1)

Модель простой линейной регрессии

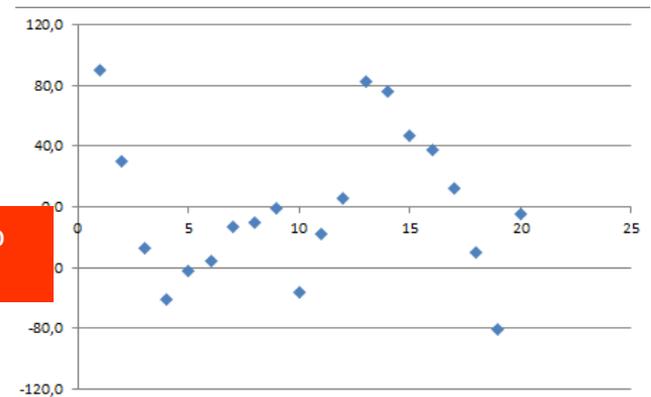
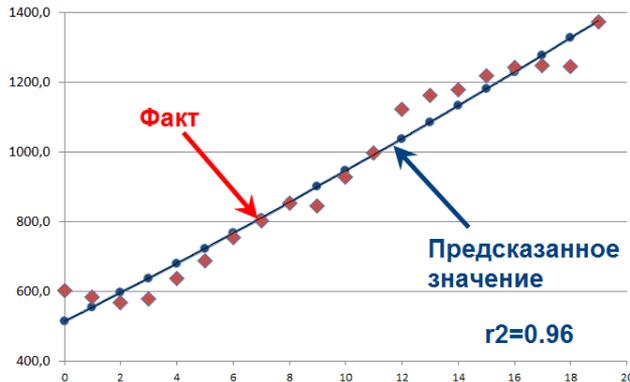
$$Y_1 = 498,95 + 45,49 \cdot x$$



Не учитывает циклическую компоненту

Модель квадратичной регрессии

$$Y_t = 31,642 + 1,2119 \cdot Y_{t-1} - 0,206 \cdot Y_{t-2}$$



Не учитывает циклическую компоненту

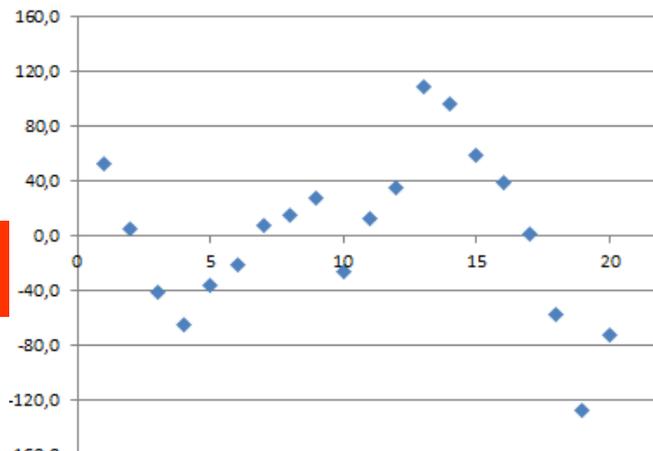
Сравнение 4-х моделей прогнозирования по уровню ошибки (2)

Экспоненциальная регрессионная модель

$$Y_1 = 550,38 \cdot e^{0,0508 \cdot x}$$

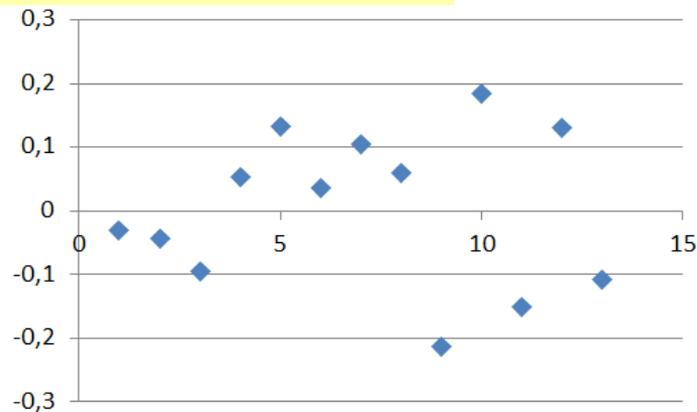
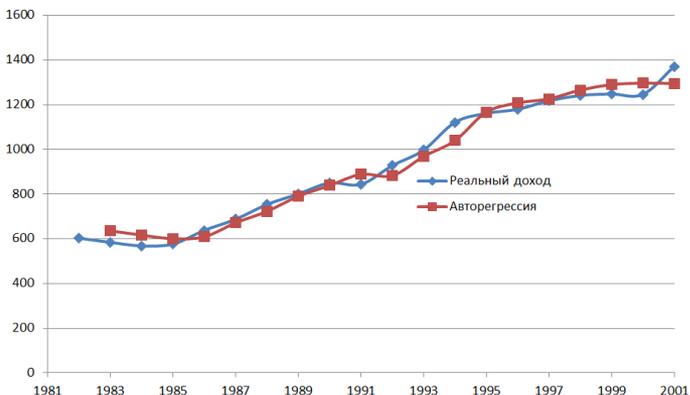


Не учитывает циклическую компоненту



Авторегрессия 1-ого порядка

$$Y_t = 18,4217 + 1,0242 \cdot Y_{t-1}$$



Задача (4)

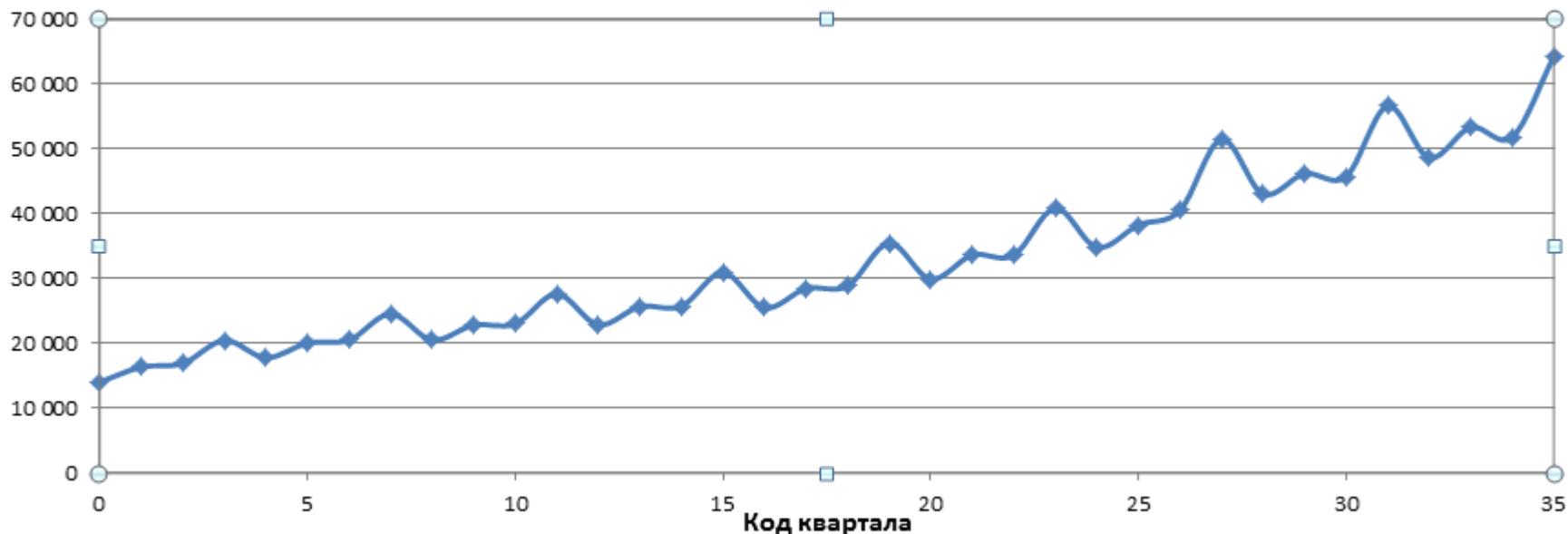
Используя построенные на предыдущих слайдах модели, оцените качество построенных зависимостей по показателям MAD и S_{yx} .

На основе проведенного анализа выберите лучшую прогностическую модель

Анализ временных рядов основе квартальных данных

График квартальных доходов компании Wall Mart, млн долл. за период 1994-2002 г.

Доход, млн долл



	1994	1995	1996	1997	1998	1999	2000	2001	2002
1 кв	13 920	17 690	20 440	22 772	25 409	29 819	34 717	42 985	48 565
2 кв	16 237	19 942	22 723	25 587	28 366	33 521	38 170	46 112	53 269
3 кв	16 827	20 418	22 913	25 644	28 777	33 509	40 432	45 676	51 754
4 кв	20 361	24 448	27 550	30 856	35 386	40 785	51 394	56 556	64 211

Прогнозирование квартальных временных рядов

$$Y = \beta_0 \cdot \beta_1^x \cdot \beta_2^{Q_1} \cdot \beta_3^{Q_2} \cdot \beta_4^{Q_3}$$

где

Y – значение прогнозируемой переменной;

x – код квартала (0, 1, 2, 3, 4, 5, 6, ...), причем $x=0$ соответствует 1-ому кварталу

Q_1 – индекс (1 для первого кв. и 0 для остальных); Q_2 – индекс (1 для второго кв. и 0 для остальных);

Q_3 – индекс (1 для третьего кв. и 0 для остальных);

β_0 – «сдвиг» переменной Y ; β_1 – темп ежеквартального роста доходов; β_2 – изменение в первом кв по отношению к четвертому; β_3 – изменение в втором кв по отношению к четвертому; β_4 – изменение в третьем кв по отношению к четвертому

Модификация:

$$\log Y = b_0 + b_1 \cdot x + b_2 \cdot Q_1 + b_3 \cdot Q_2 + b_4 \cdot Q_3$$

где

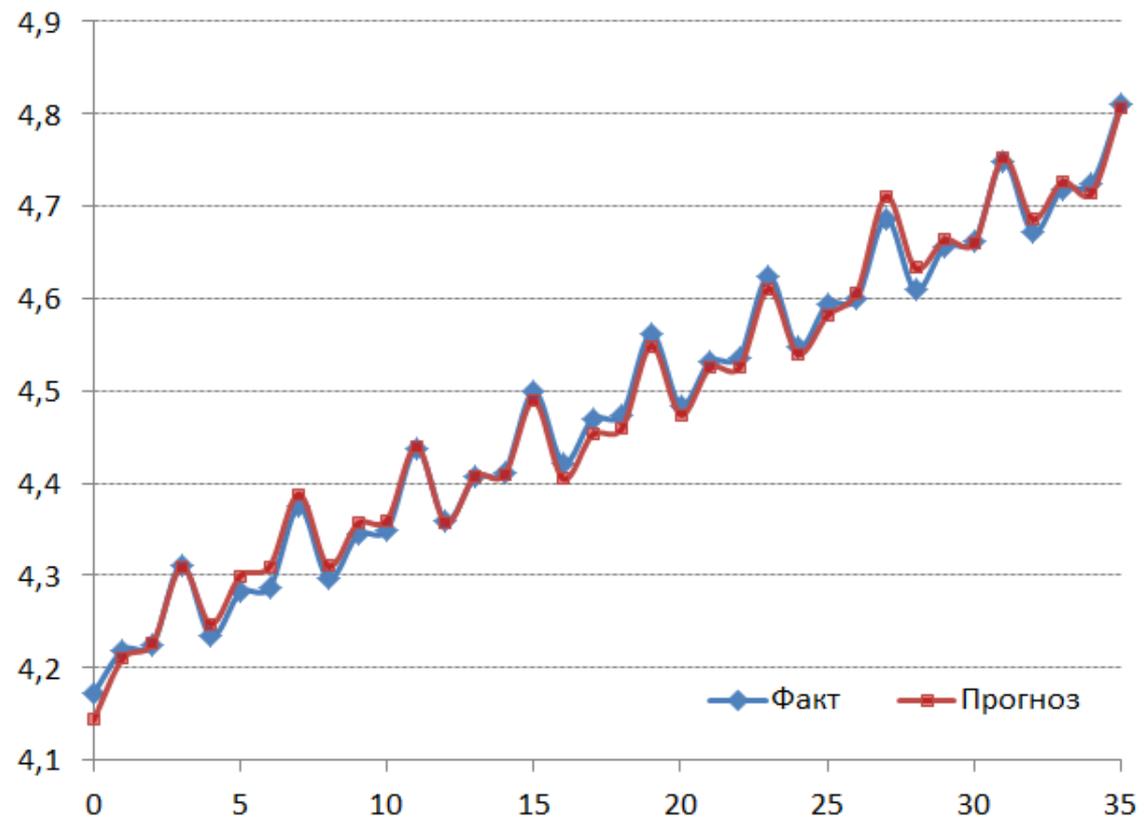
$b_1 = \log(\beta_1)$, $b_2 = \log(\beta_2)$, $b_3 = \log(\beta_3)$, т.е., например $10^{b_1} = \beta_1$

Пример прогнозирования квартальных временных рядов

$$\text{Log}Y = 4,2652 + 0,01561 \cdot x - 0,09293 \cdot Q_1 - 0,06121 \cdot Q_2 - 0,07185 \cdot Q_3$$

$r^2 = 0,994$

Период	Квартал	Значение	Log(значение)
0	1	13920	4,14
1	2	16237	4,21
2	3	16827	4,23
3	4	20361	4,31
4	1	17690	4,25
5	2	19942	4,30
6	3	20418	4,31
7	4	24448	4,39
8	1	20440	4,31
9	2	22723	4,36
10	3	22913	4,36
11	4	27550	4,44
12	1	22772	4,36
13	2	25587	4,41
14	3	25644	4,41
15	4	30856	4,49
16	1	25409	4,40
17	2	28366	4,45
18	3	28777	4,46
19	4	35386	4,55
20	1	29819	4,47



Используя полученное регрессионное уравнение спрогнозируйте значение доходов в 36-40 периодах. Каким годам они соответствуют?

тами (ВШУП)

Задача (5)

Ниже приведены индексы курсов акций, вычисленные компанией *Standard & Poor* в период 1994-2003 годы.

Год	Кв	Значение	Год	Кв	Значение
1994	1	445,77	1998	4	1229,23
1994	2	444,27	1999	1	1286,37
1994	3	462,69	1999	2	1372,71
1994	4	459,27	1999	3	1282,71
1995	1	500,71	1999	4	1469,25
1995	2	544,75	2000	1	1498,58
1995	3	584,41	2000	2	1454,6
1995	4	615,93	2000	3	1436,51
1996	1	645,5	2000	4	1320,28
1996	2	670,63	2001	1	1160,33
1996	3	687,31	2001	2	1224,38
1996	4	740,74	2001	3	1040,94
1997	1	757,12	2001	4	1148,08
1997	2	885,14	2002	1	1147,38
1997	3	947,28	2002	2	989,81
1997	4	970,43	2002	3	815,28
1998	1	1101,75	2002	4	879,28
1998	2	1133,84	2003	1	848,18
1998	3	1017,01			

Вопросы:

1. Постройте график временного ряда
2. Постройте уравнение экспоненциального тренда с учетом квартальных данных
3. Предскажите значение временного ряда в 2003 и в 2004 годах
4. Дайте качественную интерпретацию коэффициентов регрессионной модели (β_0 , β_1 , β_2 , β_3 , β_4)
5. Сделайте выводы по результатам построения прогнозной модели

Основное предположение: факторы, влияющие на исследуемый объект в прошлом, сохраняют свое влияние в будущем

СПАСИБО ЗА ВНИМАНИЕ!

Литвин Юрий Васильевич

Email: litvinj@simplecs.ru

Тел.: 8 (929) 906-90-33

Приложение

(t – статистика для различных степеней свободы)

Число средней свободы <i>df</i>	α			Число средней свободы <i>df</i>	α		
	0,10	0,05	0,01		0,10	0,05	0,01
1	6,3138	12,706	63,657	18	1,7341	2,1009	2,8784
2	2,9200	4,3027	9,9248	19	1,7291	2,0930	2,8609
3	2,3534	3,1825	5,8409	20	1,7247	2,0860	2,8453
4	2,1318	2,7764	4,6041	21	1,7207	2,0796	2,8314
5	2,0150	2,5706	4,0321	22	1,7171	2,0739	2,8188
6	1,9432	2,4469	3,7074	23	1,7139	2,0687	2,8073
7	1,8946	2,3646	3,4995	24	1,7109	2,0639	2,7969
8	1,8595	2,3060	3,3554	25	1,7081	2,0595	2,7874
9	1,8331	2,2622	3,2498	26	1,7056	2,0555	2,7787
10	1,8125	2,2281	3,1693	27	1,7033	2,0518	2,7707
11	1,7959	2,2010	3,1058	28	1,7011	2,0484	2,7633
12	1,7823	2,1788	3,0545	29	1,6991	2,0452	2,7564
13	1,7709	2,1604	3,0123	30	1,6973	2,0423	2,7500
14	1,7613	2,1448	2,9768	40	1,6839	2,0211	2,7045
15	1,7530	2,1315	2,9467	60	1,6707	2,0003	2,6603
16	1,7459	2,1199	2,9208	120	1,6577	1,9799	2,6174
17	1,7396	2,1098	2,8982	∞	1,6449	1,9600	2,5758