ARIMA-процессы Бокса-Дженкинса

семейство линейных статистических моделей, основанных на нормальном распределении, которые позволяют имитировать поведение множества различных реальных временных рядов путем комбинирования процессов авторегрессии, процессов интегрирования и процессов скользящего среднего.

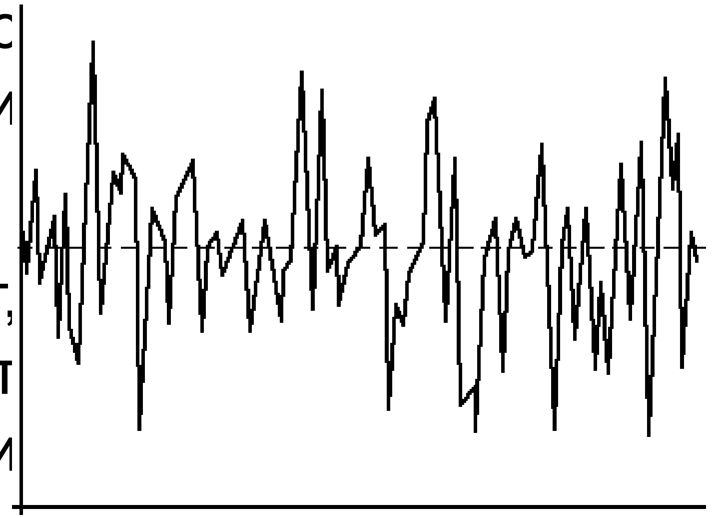
ARIMA – **A**utoregressive **I**ntegrated **M**oving **A**verage

Моделирование циклического поведения с помощью ARIMA-процессов Бокса-Дженкинса

Процесс случайного шума не обладает памятью: отправная точка

Процесс случайного шума состоит из случайной выборки (независимых наблюдений) из нормального распределения с постоянным средним и стандартным отклонением.

$$Y_t = \mu + \xi_t$$



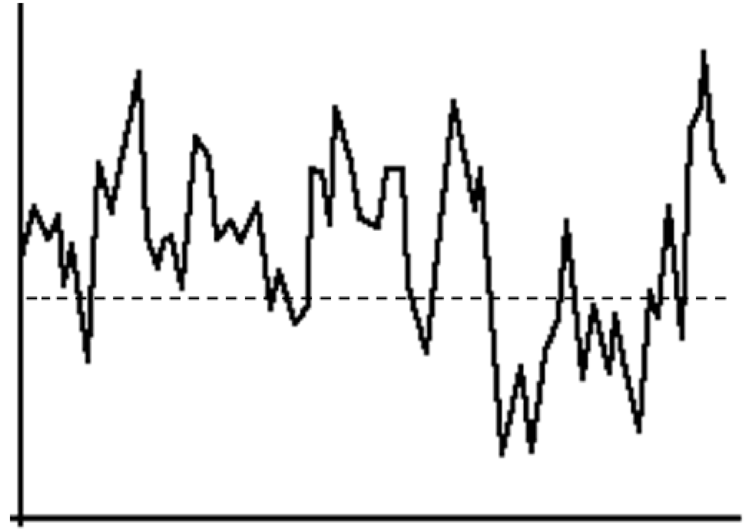
Тенденции (тренды) отсутствуют, наблюдения не содержат информации о прошлом поведении ряда.

Моделирование циклического поведения с помощью ARIMA-процессов Бокса-Дженкинса

Процесс авторегрессии (AR) обладает памятью о своем прошлом

Любое наблюдение процесса авторегрессии (AR) – линейная функция от предыдущего наблюдения плюс случайный шум.

$$Y_t = \delta + \varphi Y_{t-1} + \xi_t$$



Процесс авторегрессии «помнит» о своем предыдущем состоянии и использует эту информацию для определения своего дальнейшего поведения.

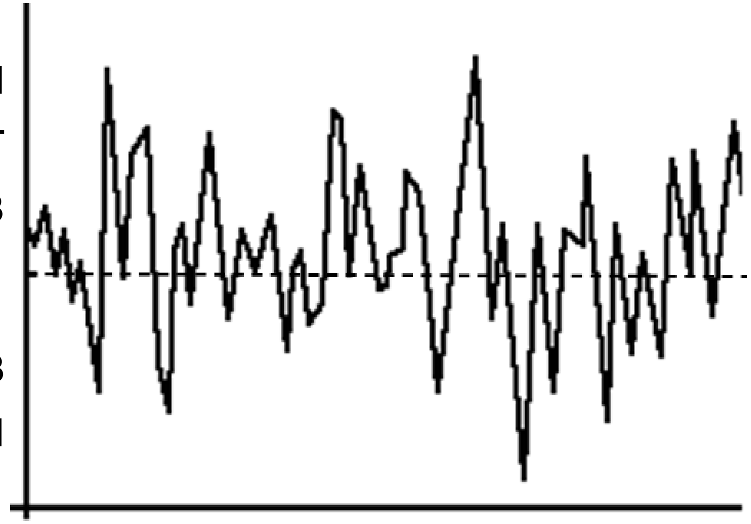
Моделирование циклического поведения с помощью ARIMA-процессов Бокса-Дженкинса

Процесс скользящего среднего (MA) имеет ограниченную память

Любое наблюдение MA состоит из константы (долгосрочное среднее значение процесса) плюс независимый случайный шум минус часть предыдущего случайного шума.

$$Y_t = \mu + \xi_t - \theta \xi_{t-1}$$

Процесс MA «не помнит» в точности своего прошлого, но помнит компонент случайного шума того состояния, в котором он (процесс) находился.



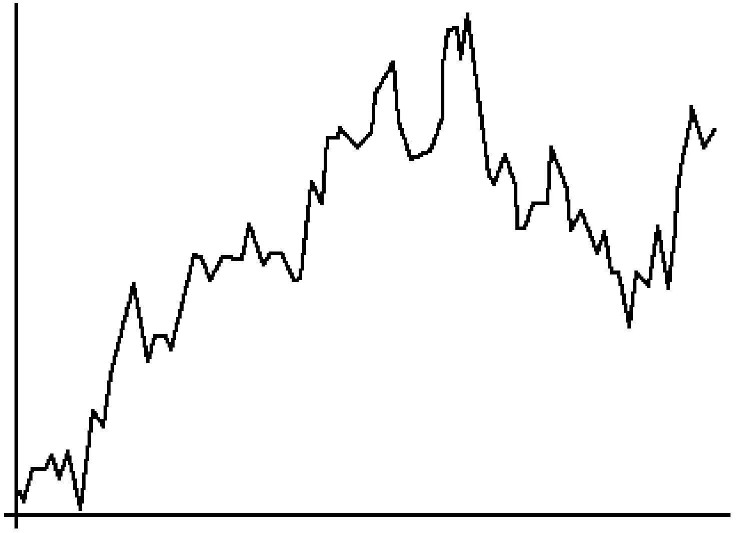
Память ограничена одним шагом в будущее; за пределами этого шага для процесса все начинается заново.

Моделирование циклического поведения с помощью ARIMA-процессов Бокса-Дженкинса

Чистый интегрированный (I) процесс помнит, где он находился, и затем движется случайно

$$Y_t = \delta + Y_{t-1} + \xi_t$$

Каждое наблюдение чистого интегрированного (I) процесса (pure integrated (I) process), называемого также случайным блужданием, заключается в случайном шаге в сторону от текущего наблюдения. Процесс знает, где он находится, но забыл, как он попал туда.

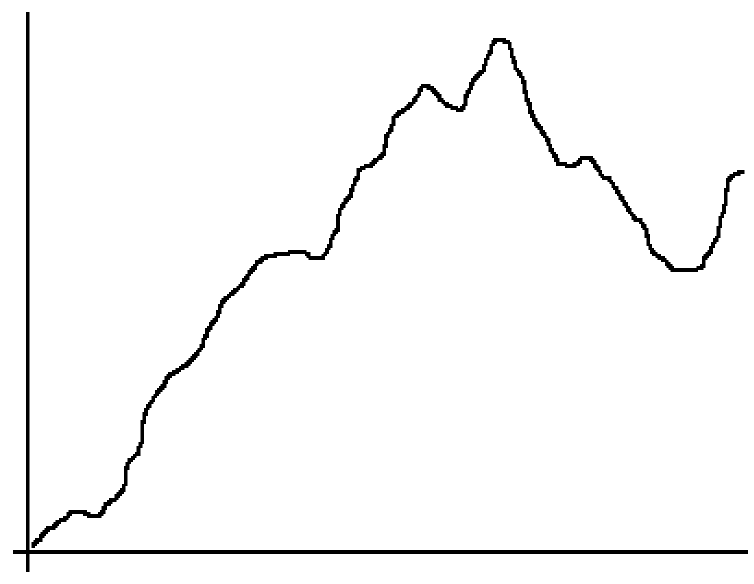


Моделирование циклического поведения с помощью ARIMA-процессов Бокса-Дженкинса

Процесс авторегрессионного интегрированного скользящего среднего (ARIMA)

$$Y_t - Y_{t-1} = \delta + \varphi(Y_{t-1} - Y_{t-2}) + \xi_t - \theta\xi_{t-1}$$

Состоит из линейной функции предыдущего изменения плюс независимый случайный шум минус определенная доля предыдущего случайного шума. Этот процесс знает, где он находится, помнит, как он попал в это состояние, и помнит даже часть предыдущего шумового компонента.

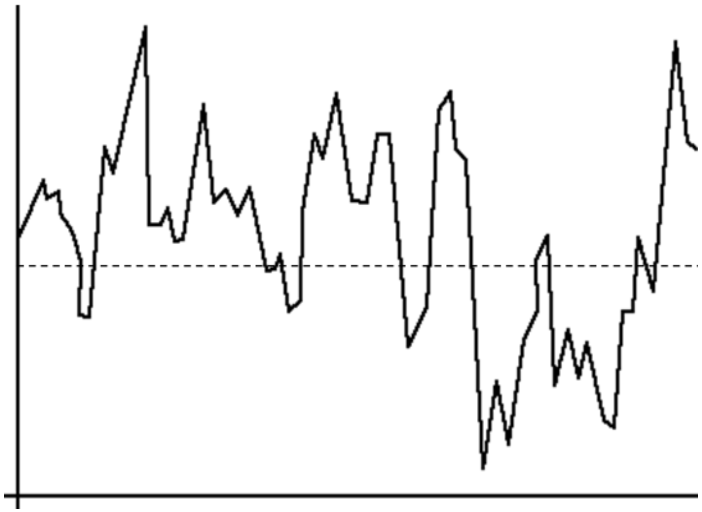


Моделирование циклического поведения с помощью ARIMA-процессов Бокса-Дженкинса

Процесс авторегрессии и скользящего среднего (ARMA)

Состоит из линейной функции от предыдущего наблюдения плюс независимый случайный шум минус некоторая доля предыдущего случайного шума.

$$Y_t = \delta + \varphi Y_{t-1} + \xi_t - \theta \xi_{t-1}$$



Запоминает как свое предыдущее состояние, так и компонент случайного шума предыдущего состояния.

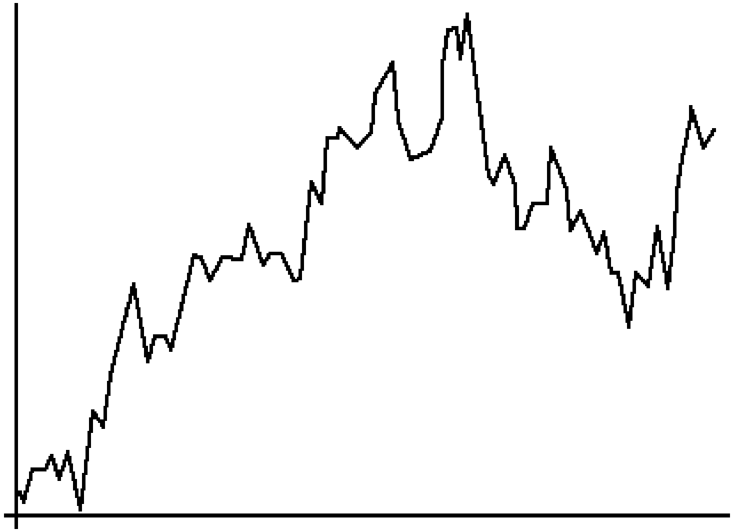
Сочетает в себе память процесса авторегрессии с памятью процесса скользящего среднего.

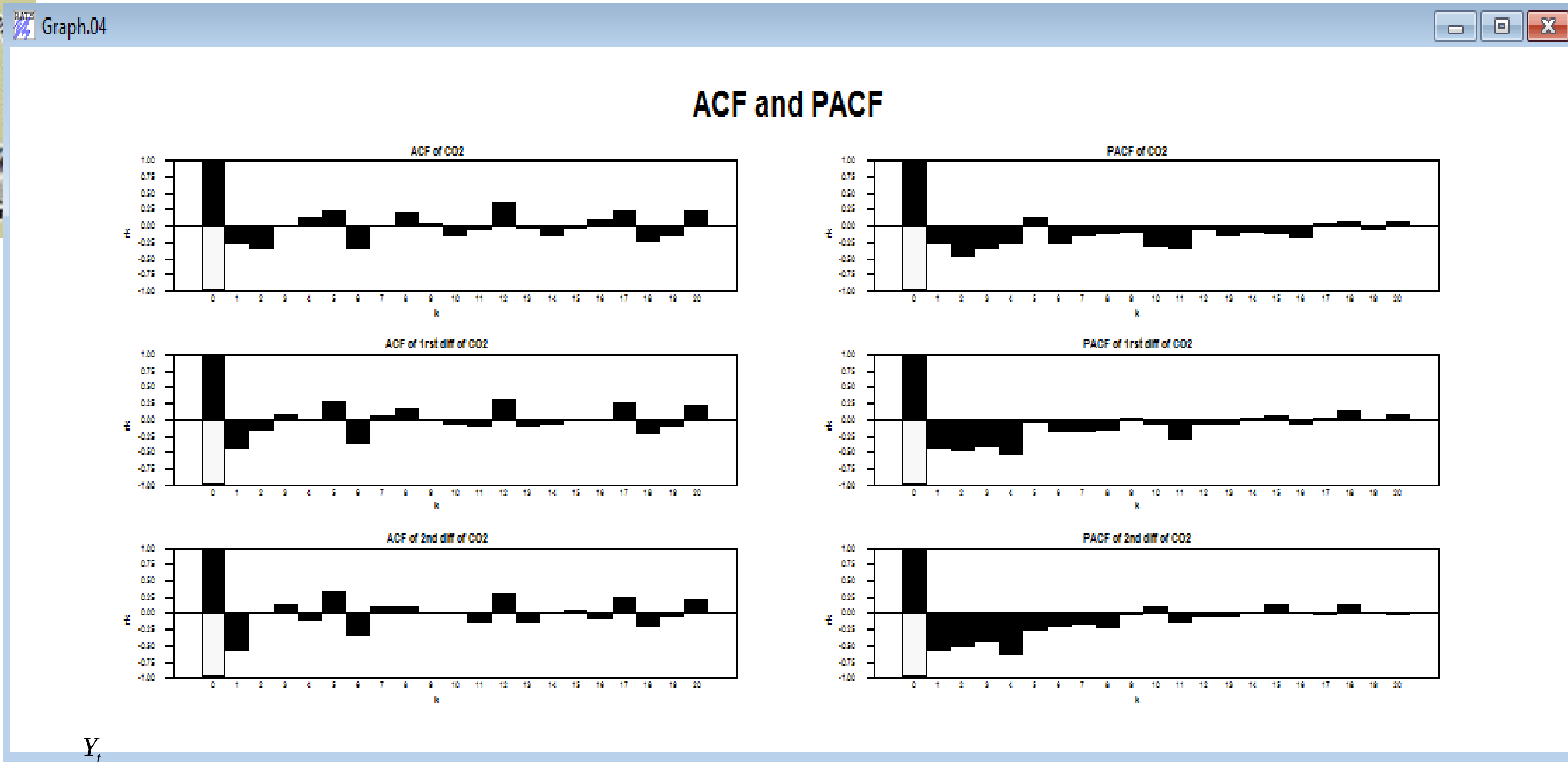
Моделирование циклического поведения с помощью ARIMA-процессов Бокса-Дженкинса

Чистый интегрированный (I) процесс помнит, где он находился, и затем движется случайно

$$Y_t = \delta + Y_{t-1} + \xi_t$$

Каждое наблюдение чистого интегрированного (I) процесса (pure integrated (I) process), называемого также случайным блужданием, заключается в случайном шаге в сторону от текущего наблюдения. Этот процесс знает, где он находится, но забыл, как он попал туда.

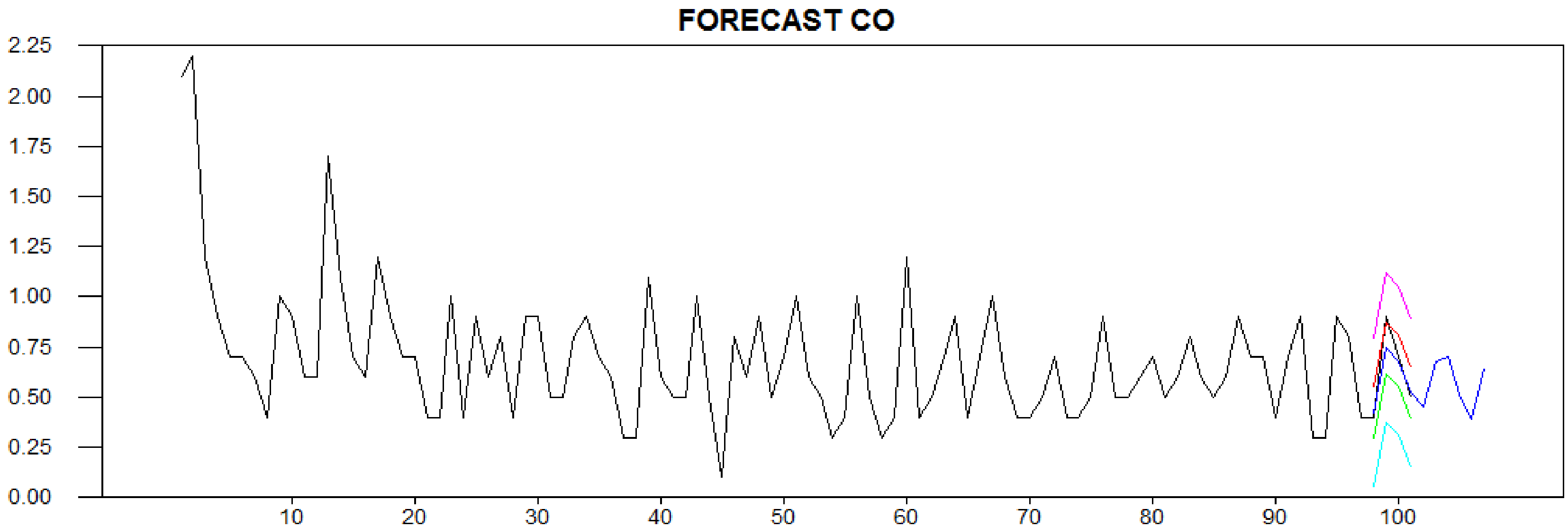


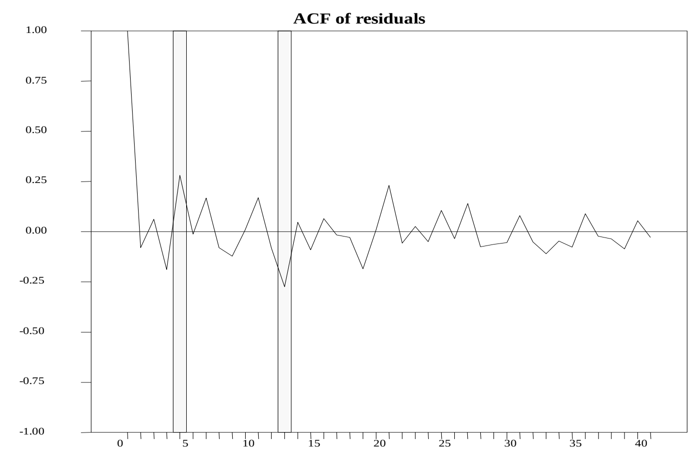


Y_t

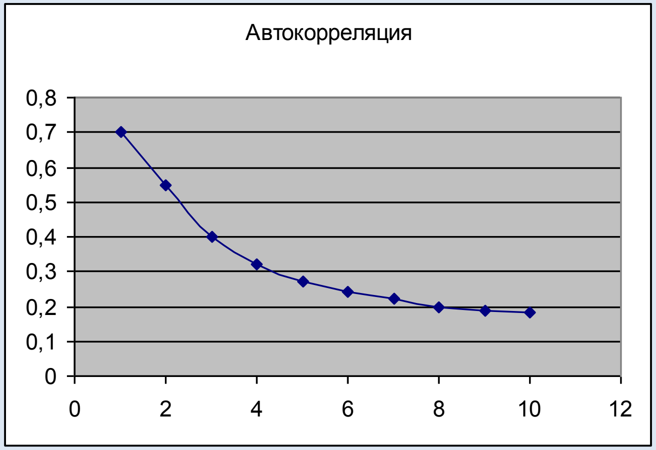
Ряд	Разность	Вид модели	Идентификация для Y_t
А	0	Смешанная модель авторегрессии первого порядка и скользящего среднего первого порядка	(1,0,1)
	2	Смешанная модель авторегрессии первого порядка и скользящего среднего первого порядка	(1,2,1)

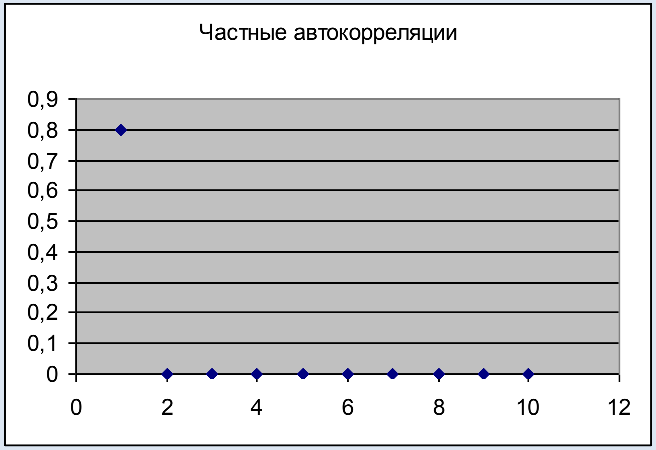
98	$\nabla^2 y_i - 0.354\nabla^2 y_{i-1} + 0.506\nabla^2 y_{i-4} - 0.186\nabla^2 y_{i-11} = \varepsilon_i - 0.785\varepsilon$ $(0.099) \quad (0.095) \quad (0.095) \quad (0.070)$	50.2009	40

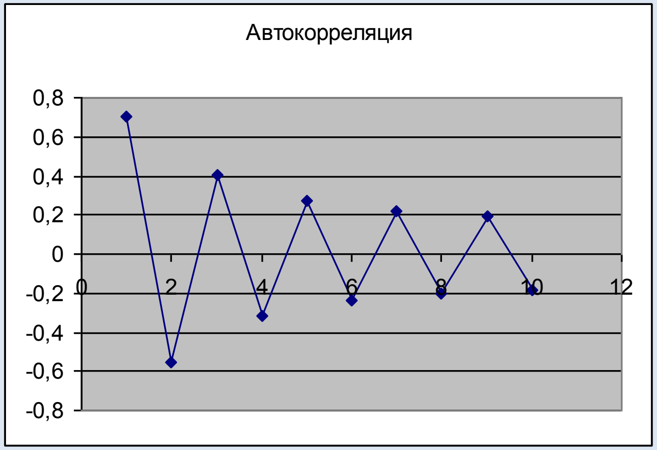


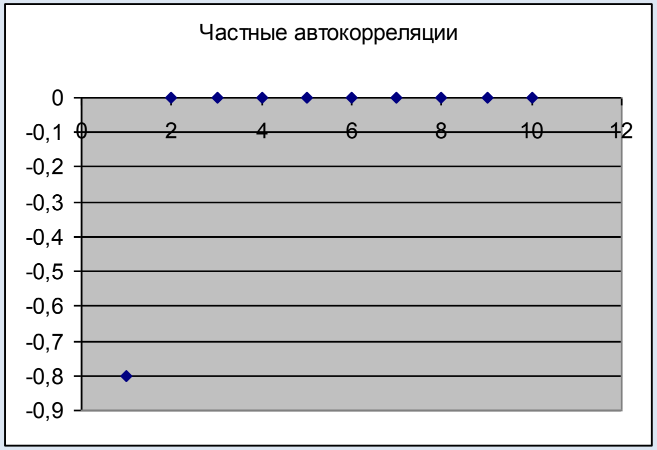


AR(1)

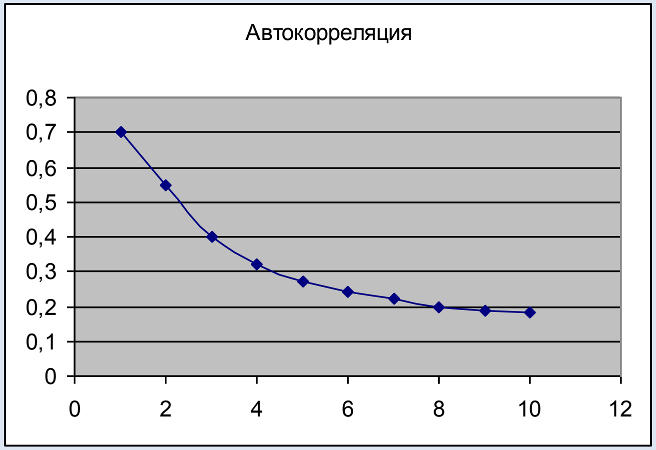

Автокорреляция
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0
2
4
6
8
10
12


Частные автокорреляции
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0
2
4
6
8
10
12


Автокорреляция
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
0
2
4
6
8
10
12

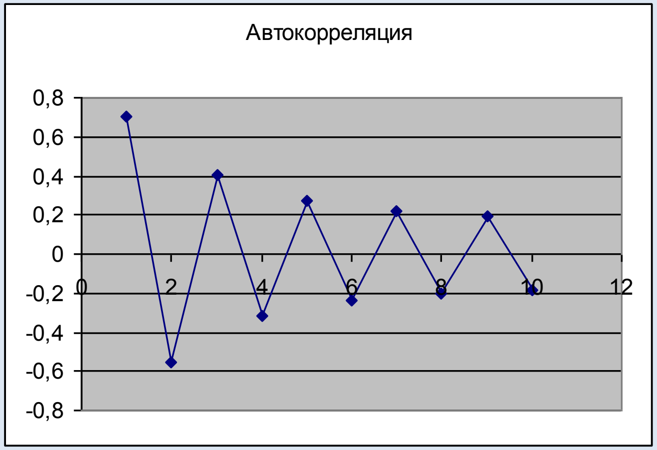

Частные автокорреляции
0
-0,1
-0,2
-0,3
-0,4
-0,5
-0,6
-0,7
-0,8
-0,9
0
2
4
6
8
10
12

AR(2)

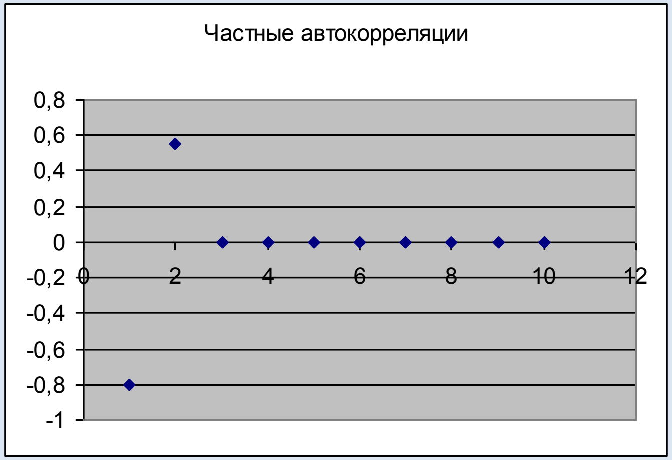


Автокорреляция
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0
2
4
6
8
10
12

Частные автокорреляции
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0
2
4
6
8
10
12

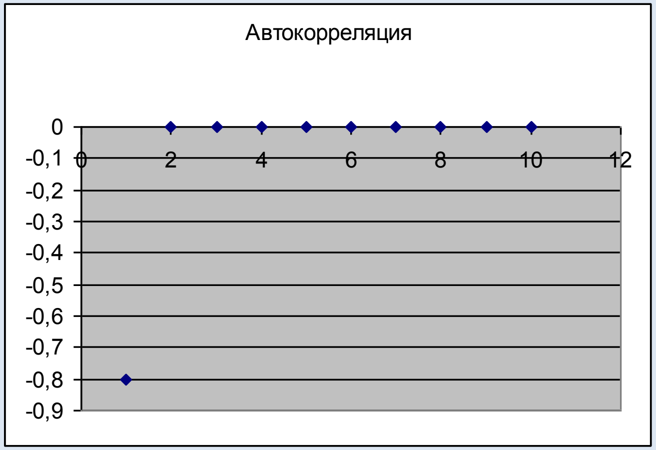


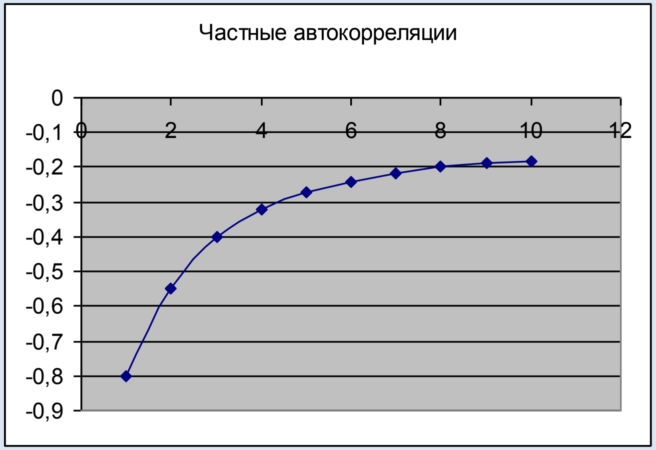
Автокорреляция
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
0
2
4
6
8
10
12

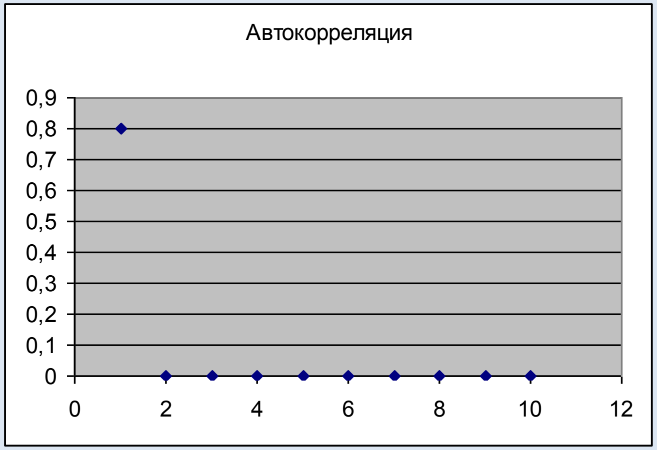


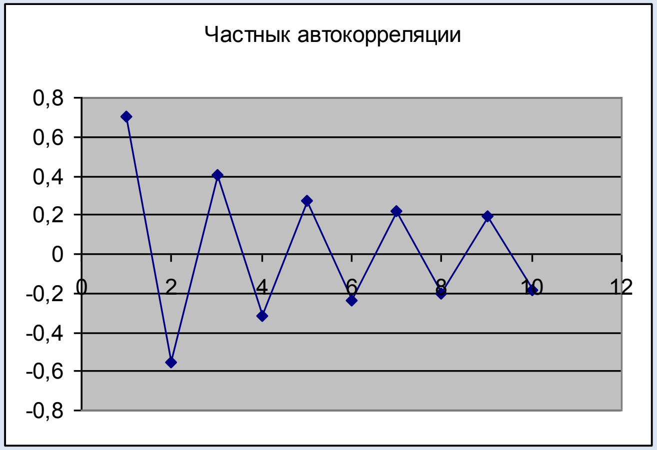
Частные автокорреляции
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
-1
0
2
4
6
8
10
12

MA(1)

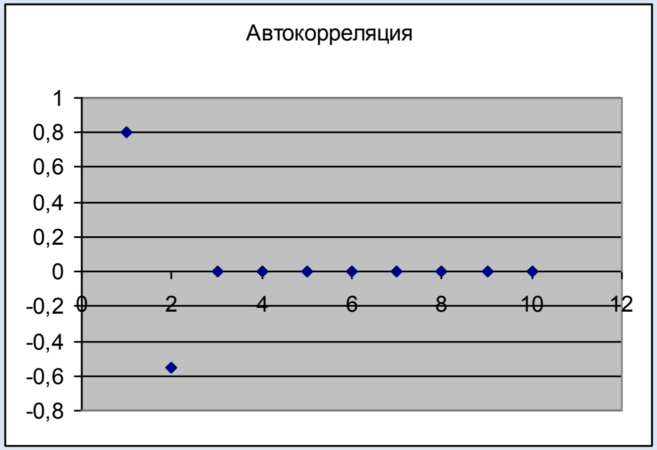

Автокорреляция
0
-0,1
-0,2
-0,3
-0,4
-0,5
-0,6
-0,7
-0,8
-0,9
0
2
4
6
8
10
12

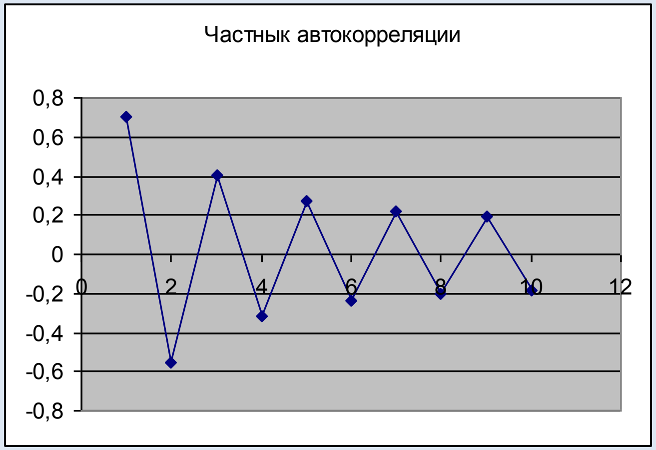

Частные автокорреляции
0
-0,1
-0,2
-0,3
-0,4
-0,5
-0,6
-0,7
-0,8
-0,9
0
2
4
6
8
10
12

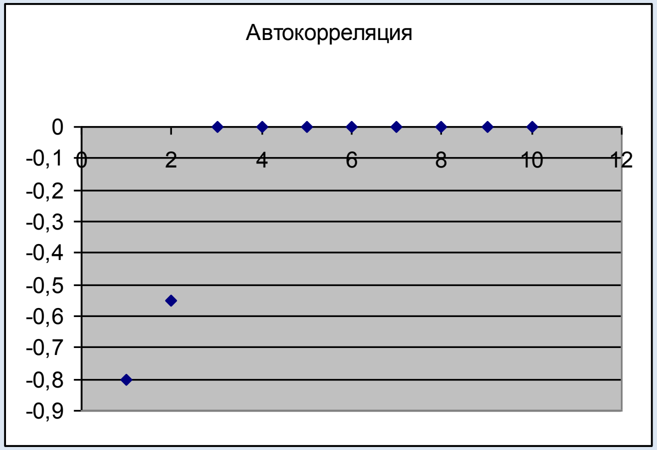

Автокорреляция
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0
2
4
6
8
10
12

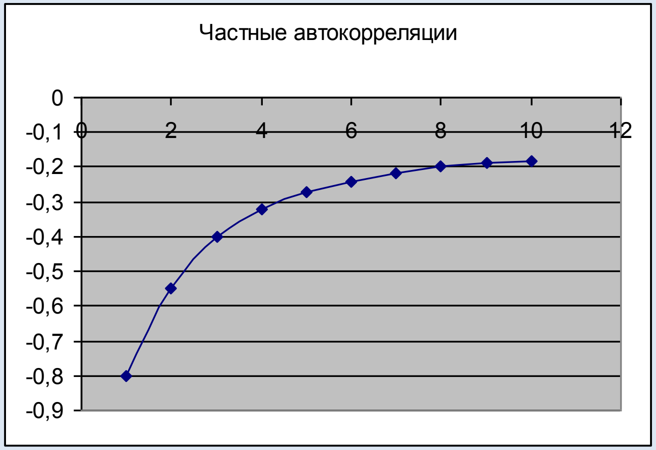

Частнык автокорреляции
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
0
2
4
6
8
10
12

MA(2)

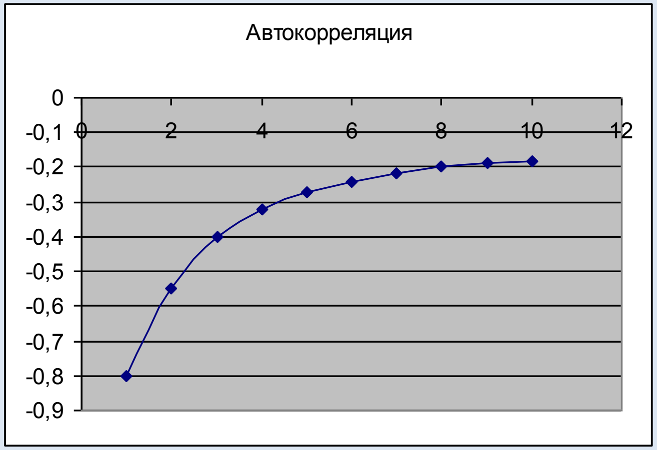

Автокорреляция
1
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
0
2
4
6
8
10
12

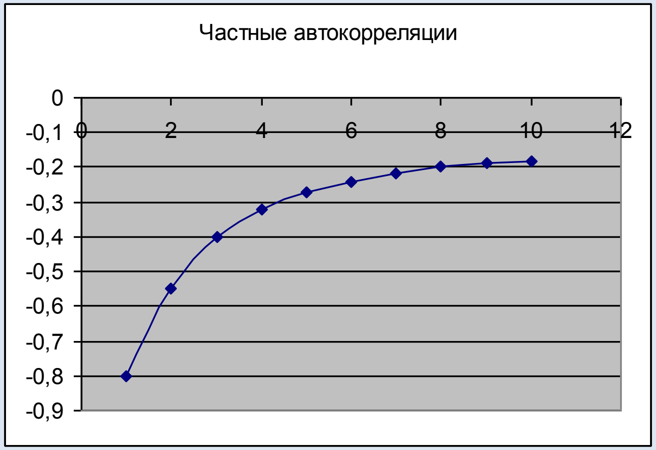

Частнык автокорреляции
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
0
2
4
6
8
10
12

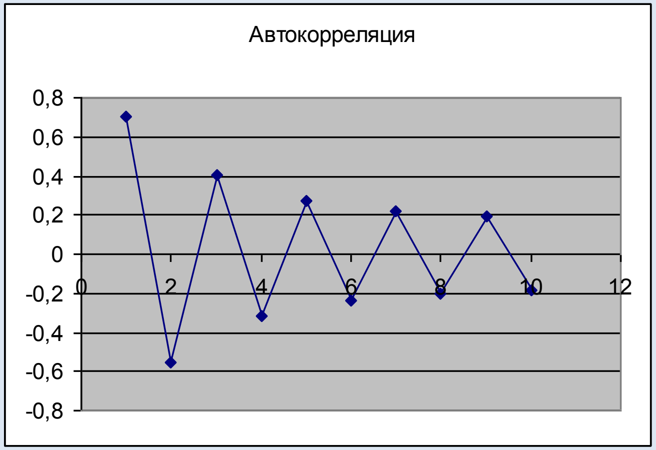

Автокорреляция
0
-0,1
-0,2
-0,3
-0,4
-0,5
-0,6
-0,7
-0,8
-0,9
0
2
4
6
8
10
12

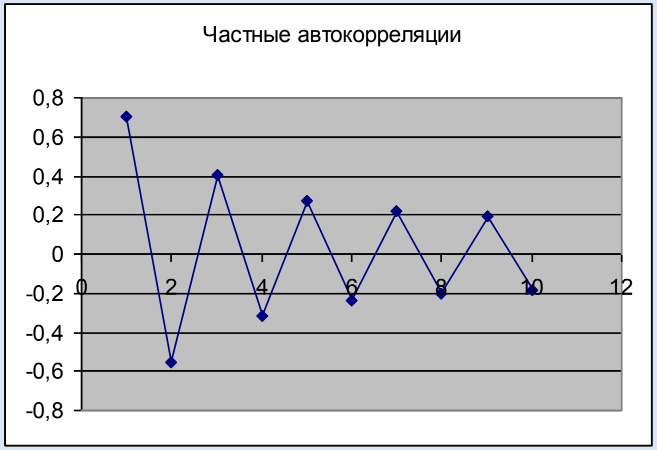

Частные автокорреляции
0
-0,1
-0,2
-0,3
-0,4
-0,5
-0,6
-0,7
-0,8
-0,9
0
2
4
6
8
10
12

ARMA(1,1)

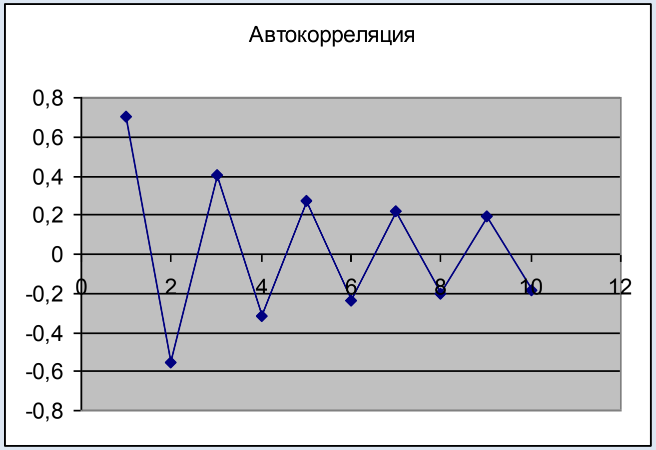

Автокорреляция
0
-0,1
-0,2
-0,3
-0,4
-0,5
-0,6
-0,7
-0,8
-0,9
0
2
4
6
8
10
12

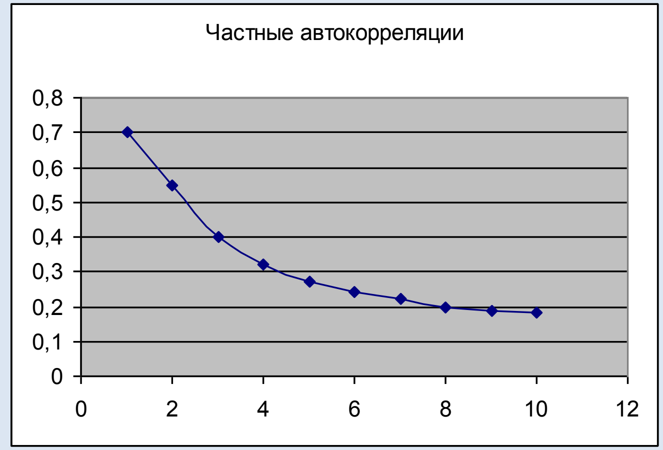

Частные автокорреляции
0
-0,1
-0,2
-0,3
-0,4
-0,5
-0,6
-0,7
-0,8
-0,9
0
2
4
6
8
10
12

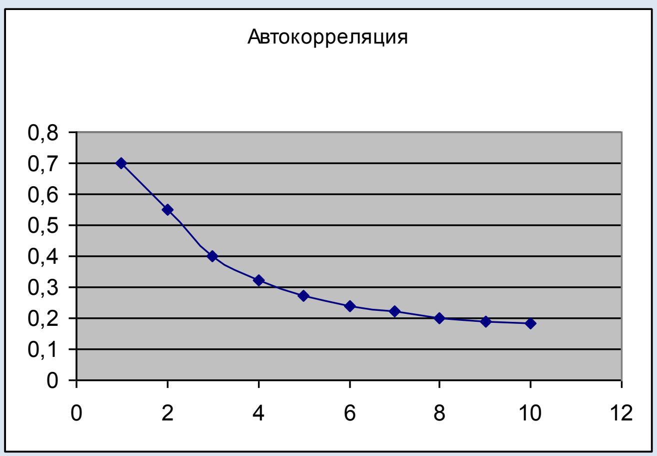

Автокорреляция
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
0
2
4
6
8
10
12

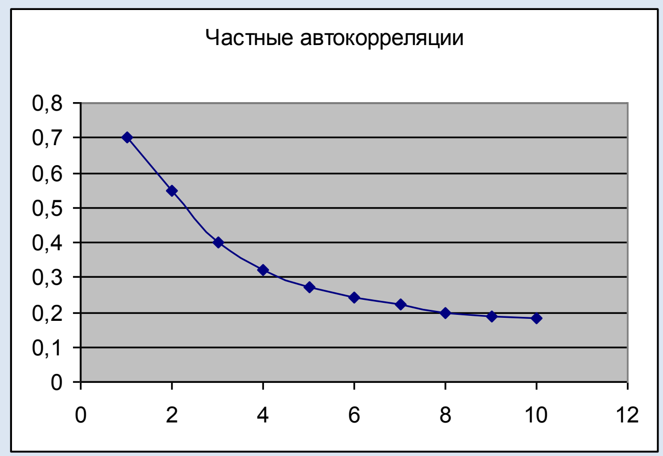

Частные автокорреляции
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
0
2
4
6
8
10
12

ARMA(1,1)


Автокорреляция
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
0
2
4
6
8
10
12


Частные автокорреляции
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0
2
4
6
8
10
12


Автокорреляция
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0
2
4
6
8
10
12


Частные автокорреляции
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0
2
4
6
8
10
12

Мультипликативная модель временных рядов

- Основное предположение: факторы, влияющие на исследуемый объект в настоящем и прошлом, будут влиять на него и в будущем.
- Таким образом, основные цели анализа временных рядов заключаются в идентификации и выделении факторов, имеющих значение для прогнозирования.

Компоненты мультипликативной модели

- Долговременная тенденция называется трендом (trend).
- Тренд является компонентой временного

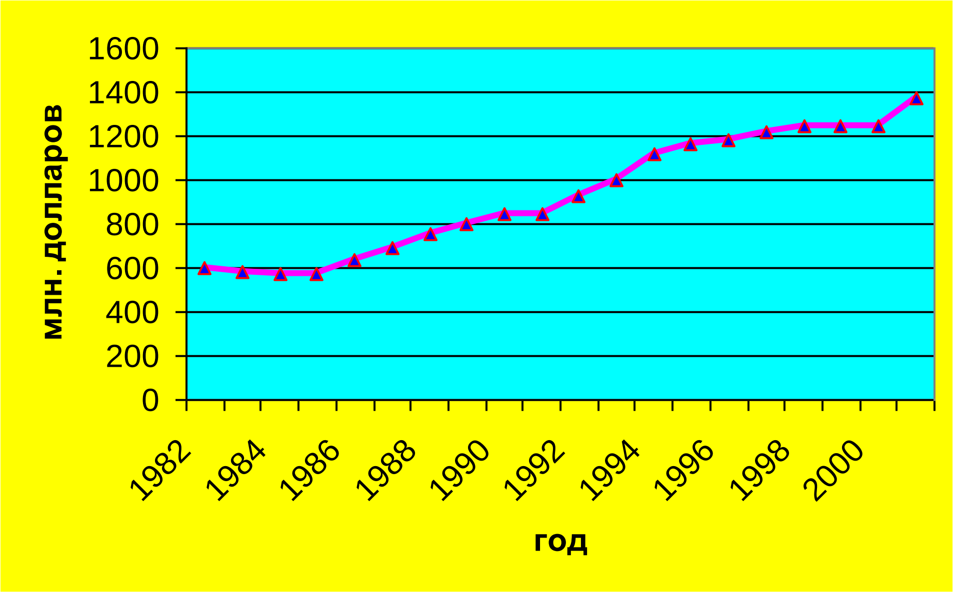
ряда.

- Циклический компонент (cyclical component) описывает колебание данных вверх и вниз. Его длина изменяется в интервале от 2 до 10 лет.

Компоненты мультипликативной модели

- Любые наблюдаемые данные, не лежащие на кривой тренда и не подчиняющиеся циклической зависимости, называются случайными компонентами (random component)

Фактический валовой доход компании WWC за период с 1982 по 2001 годы



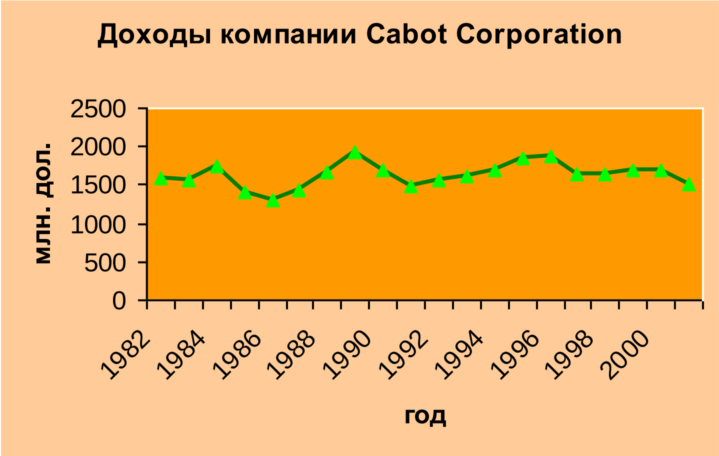
Классическая мультипликативная модель временного ряда для ежегодных данных

$$Y_i = T_i * C_i * I_i$$

- T_i – значение тренда;
- C_i – значение циклического компонента в i-том году
- I_i – значение случайного компонента в i-том году

Сглаживание годовых временных рядов:

скользящее среднее и экспоненциальное среднее



Скользящее среднее

Moving average

подвижная динамическая средняя, которая исчисляется по ряду при последовательном передвижении на один интервал, т.е. сначала вычисляют средний уровень из определенного числа первых по порядку уровней ряда, затем - средний уровень из такого же числа членов, начиная со второго.

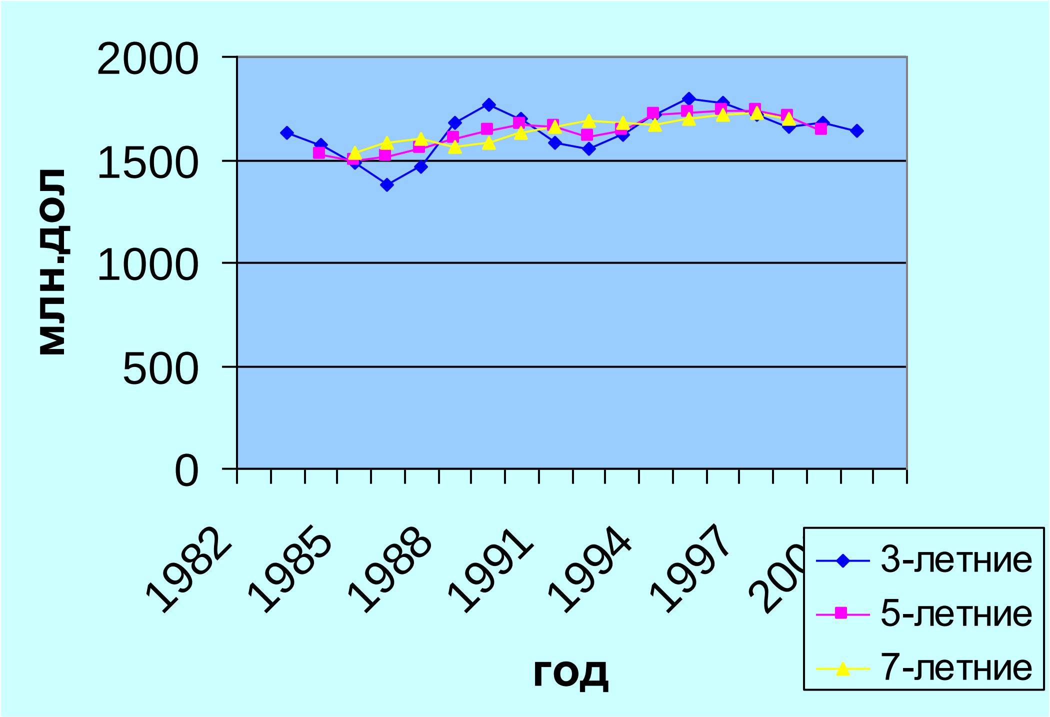
$$\bar{y}_1 = \frac{y_1 + y_2 + y_3}{3} \qquad \bar{y}_2 = \frac{y_2 + y_3 + y_4}{3}$$

$$\bar{y}_3 = \frac{y_3 + y_4 + y_5}{3}$$

Скользящие средние с продолжительностью периода, равной 3

- Для того, чтобы исключить циклические колебания, длина периода должна быть целым числом, кратным средней длине цикла.
- Для трехлетнего периода невозможно выполнить вычисления для первого и последнего года, а при пятилетнем периоде сглаживания - первых двух и последних двух лет.

Скользящие средние для доходов Cabot Corporation



Экспоненциальное сглаживание

Exponential smoothing

Этот метод позволяет делать краткосрочные прогнозы (в рамках одного периода), когда наличие долговременных тенденций остается под вопросом.

При экспоненциальном сглаживании веса, присвоенные наблюдаемым значениям, убывают со временем, поэтому после выполнения вычислений наиболее часто встречающиеся значения получат наибольший вес, а редкие величины - наименьший

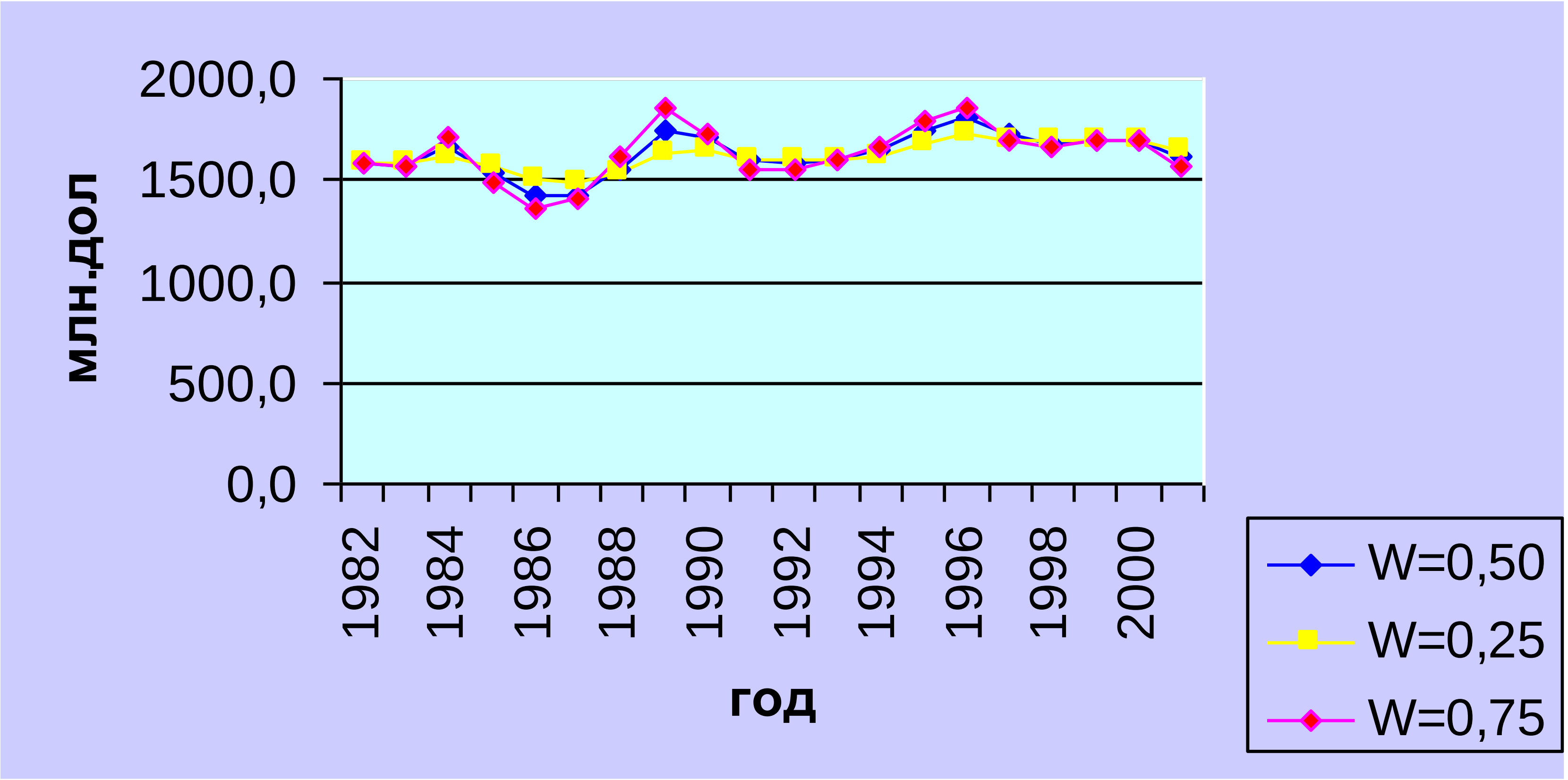
Вычисление экспоненциально сглаженного значения в i-том периоде времени

$$E_i = Y_i$$

$$E_i = WY_i + (1 - W)E_{i-1} \qquad i = 2,3,4.....$$

- Где E_i- значение экспоненциально сглаженного ряда, вычисленное для i-го периода;
- E_{i-1}- значение экспоненциально сглаженного ряда, вычисленное для предшествующего периода
- Y_i- наблюдаемое значение временного ряда в i-ом периоде;
- W – субъективный вес или сглаживающий коэффициент (0<W<1)

Экспоненциально сглаженный временной ряд для доходов Cabot Corporation



Экспоненциально сглаженный временной ряд для доходов Cabot Corporation

- Допустим, что коэффициент сглаживания равен 0,25. Первое наблюдаемое значение Y_{1982} = 1587,7 одновременно является первым сглаженным значением E_{1982} = 1587,7.
- Используя значение временного ряда для 1983 года получаем:

  $E_{1983} = WY_{1983}+(1-W)E_{1982} =$

  $= 0{,}25*1558{,}0+0{,}75*1587{,}7 = 1580{,}3$

Экспоненциально сглаженный временной ряд для доходов Cabot Corporation

- Сглаженное значение временного ряда для 1984 года:

  $E_{1984} = WY_{1984} + (1\text{-}W)E_{1983} =$

  $= 0{,}25*1752{,}5+0{,}75*1580{,}3 = 1623{,}3$ и т.д.

Этот процесс продолжается до тех пор пока не будут сглажены все значения вариационного ряда.

Прогнозирование значений для (i+1)-го интервала

$$\widehat{Y}_{i+1} = E_i$$

Для предсказания доходов компании Cabot Corporation в 2002 году можно использовать сглаженное значение, вычисленное для 2001 года