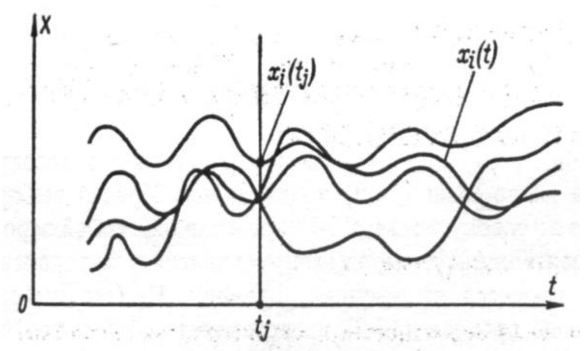
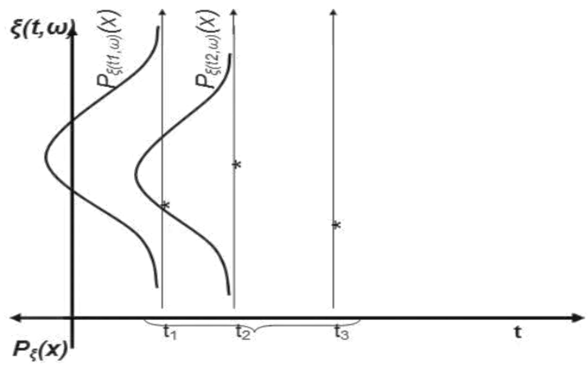
Лабораторная работа по анализу временных рядов

Динамический ряд – совокупность наблюдений некоторого явления (показателя), упорядоченная в зависимости от последовательности значений другого явления (признака).

Временной ряд (ВР) – это динамический ряд, у которого в качестве признака используется время. Временной ряд представляется набором чисел, привязанных к последовательным, обычно равностоящим моментам времени. Числа временного ряда (элементы) называют уровнями временного ряда y_t, t=1,…, n. Количество входящих в него уровней n называют длиной ряда.

ВР $y_1, y_2, ... y_n$ обычно интерпретируется как одна из возможных реализаций последовательности случайных величин $Y_1, Y_2, ... Y_n$, представляющей собой случайный (стохастический) процесс с дискретным временем $Y(t)$.





Характеристики случайного процесса $Y(t)$ в общем случае являются функциями от времени t:

математическое ожидание

$$\mu_t = M[Y_t] = \mu(t);$$

дисперсия

$$\sigma_t^{\ 2} = D[Y_t] \ = M[(Y_t - \mu_t)^2] = \sigma^2(t),$$

а автоковариация

$$\operatorname{cov}(Y_{t1}, Y_{t2}) = M\left[(Y_{t1} - \mu_{t1})(Y_{t2} - \mu_{t2})\right]$$

зависит от t_1 и t_2 .

Замечание: *важной особенностью временных рядов по сравнению с данными наблюдений, относящихся к одному периоду времени, является, как правило, наличие связи между последовательными уровнями ряда, вызванное действием каких-либо долговременных причин, что приводит к наличию таких составляющих ряда, как долговременная тенденция и периодическая составляющая.*

Корреляционная зависимость между последовательными уровнями временного ряда называется **автокорреляцией уровней временного ряда**.

Степень тесноты автокорреляционной связи между уровнями ряда может быть определена с помощью **коэффициентов автокорреляции**, т. е. коэффициентов линейной корреляции между уровнями исходного временного ряда и уровнями ряда, сдвинутыми на несколько шагов назад во времени. **Величина сдвига** называется **лагом**, обозначим ее через τ .

Например, имеется временной ряд за n лет, пусть лаг составляет $\tau = 2$ периода, тогда расположим уровни временного ряда следующим образом:

y_3	$y_1 = y_{3-2} = y_{3-\tau}$
y_4	$y_2 = y_{4-\tau}$
y_5	$y_3 = y_{5-\tau}$
…	…
y_n	$y_{n-2} = y_{n-\tau}$
$\bar{y}_{1\tau}$	$\bar{y}_{2\tau}$

$$r_2 = \frac{\sum\limits_{t=3=\tau+1}^{n}(y_t - \bar{y}_{12})(y_{t-2} - \bar{y}_{22})}{\sqrt{\sum\limits_{t=3}^{n}(y_t - \bar{y}_{12})^2 \sum\limits_{t=3}^{n}(y_{t-2} - \bar{y}_{22})^2}}$$ - коэффициент автокорреляции 2-порядка.

В общем случае

$$r_\tau = \frac{\sum\limits_{t=\tau+1}^{n}(y_t - \bar{y}_{1\tau})(y_{t-\tau} - \bar{y}_{2\tau})}{\sqrt{\sum\limits_{t=\tau+1}^{n}(y_t - \bar{y}_{1\tau})^2 \sum\limits_{t=\tau+1}^{n}(y_{t-\tau} - \bar{y}_{2\tau})^2}}$$ – коэффициент автокорреляции порядка τ .

Здесь $\bar{y}_{1\tau} = \frac{\sum\limits_{t=\tau+1}^{n} y_t}{n-\tau}$, $\bar{y}_{2\tau} = \frac{\sum\limits_{t=\tau+1}^{n} y_{t-\tau}}{n-\tau}$ – выборочные средние.

Функцию $r(\tau) = r_\tau$ называют **автокорреляционной функцией временного ряда** (АКФ или ACF), а ее график – **коррелограммой**.

Анализ автокорреляционной функции и коррелограммы позволяет выявить структуру ряда, т. е. определить присутствие в ряде той или иной компоненты.

Так, если наиболее высоким оказался коэффициент автокорреляции первого порядка, то исследуемый ряд содержит только тенденцию.

Если наиболее высоким оказался коэффициент автокорреляции порядка τ, то ряд содержит циклические колебания с периодичностью в τ моментов времени.

Если ни один из коэффициентов автокорреляции не является значимым, то ряд не содержит тенденции и циклических колебаний.

Замечание: *линейные коэффициенты автокорреляции характеризуют тесноту* ***только линейной*** *связи текущего и предыдущих уровней ряда. Для проверки ряда на наличие нелинейной тенденции рекомендуется вычислить линейные коэффициенты автокорреляции для временного ряда, состоящего из логарифмов исходных уровней. Отличные от нуля значения коэффициентов автокорреляции будут свидетельствовать о наличии нелинейной тенденции.*

Еще одной важной характеристикой временного ряда является **частная автокорреляционная функция** (ЧАКФ или PACF), которая измеряет автокорреляцию, существующую между разделенными τ тактами времени членами ВР y_t и $y_{t-\tau}$

Различают **стационарные** и **нестационарные** случайные процессы (а, следовательно, и временные ряды).

Стохастический процесс называется **стационарным** процессом в **узком** (сильном) смысле, если совместное распределение вероятностей случайных величин $Y_{t_1}, Y_{t_2}, \ldots Y_{t_n}$ такое же, как и у случайных величин $Y_{t_1+\tau}, Y_{t_2+\tau}, \ldots Y_{t_n+\tau}$ при любых n, t и τ.

Интуитивно, «стационарность» означает, что поведение ВР в будущем «похоже» на его поведение в прошлом. Точнее, основные вероятностные характеристики ВР неизменны во времени. Именно «неизменность поведения» во времени позволяет строить прогнозы стационарных временных рядов на основе их предыстории.

Стохастический процесс называется **стационарным** процессом в **широком** (слабом) смысле, если математическое ожидание μ_t и дисперсия σ_t^2 не зависят от времени (одинаковы для всех Y_t), а автоковариация $\operatorname{cov}(Y_{t_1}, Y_{t_2})$ зависит только от величины лага $\tau = t_2 - t_1$, т. е.

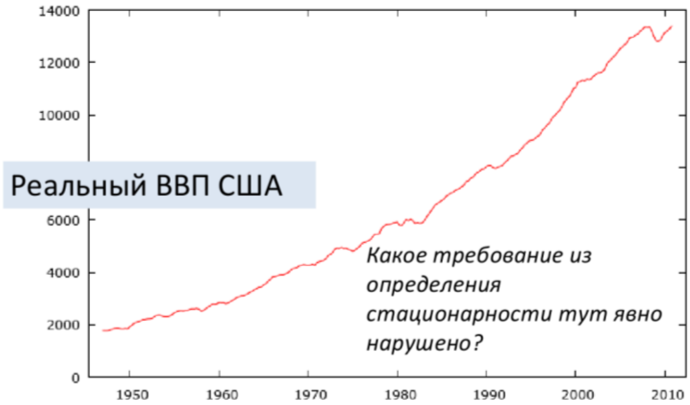
$$\mu_t = \mu = \text{const};$$

$$\sigma_t^{\,2} = \sigma^2 = \text{const};$$

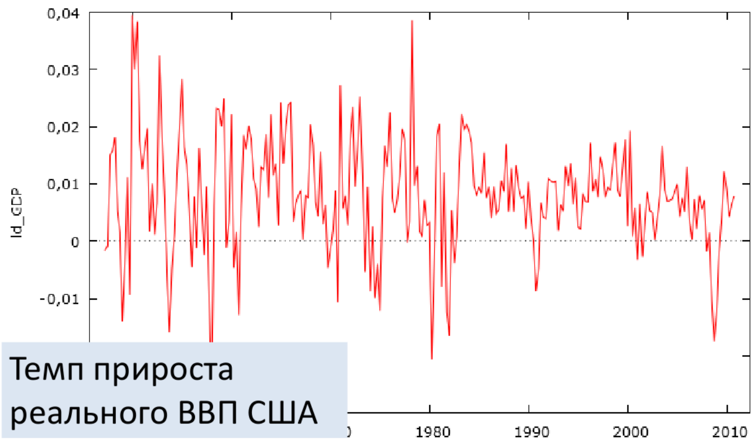
$$\operatorname{cov}(Y_{t_1}, Y_{t_2}) = \operatorname{cov}(Y_{t_1}, Y_{t_1+\tau}) = \gamma(\tau).$$

На практике признаками стационарности временного ряда являются отсутствие тенденции и периодической составляющей, а также систематических изменений размаха колебаний и систематически изменяющихся взаимозависимостей между элементами временного ряда.

Пример нестационарного ряда



Пример стационарного ряда



Для распознавания стационарности временных рядов могут использоваться следующие подходы:

– визуальный анализ графического представления временного ряда на наличие тенденции и периодической составляющей, на постоянство дисперсии и т. п.;
– анализ временного ряда на наличие автокорреляции;
– тесты на присутствие детерминистического тренда;
– тесты на постоянство статистических характеристик;
– тесты на наличие стохастического тренда, например, тесты на единичный корень (тесты Дики-Фуллера).

В случае стационарного ВР значения АКФ и ЧАКФ статистически значимы только на нескольких первых лагах, а далее функции быстро убывают. В случае нестационарности ВР значение на первом лаге автокорреляционной функции АКФ(1) близко к 1, а затем коррелограмма медленно убывает по угасающей экспоненте (синусоиде); значение ЧАКФ(1) близко к 1, а остальные значения статистически незначимы, то есть значения функции не выходят за пределы доверительного интервала.

Белый шум – это пример стационарного ВР.

Процессом белого шума будем называть последовательность независимых и одинаково распределенных случайных величин с ***нулевым математическим ожиданием и одинаковой дисперсией*** σ_{ε}^{2} и обозначать

$$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$$

Рассмотрим модели авторегрессии (AR), модели скользящего среднего (MA), смешанные модели ARMA и ARIMA.

<u>Модель авторегрессии</u>

Авторегрессионным процессом порядка p (обозначается AR(p)) называется стохастический процесс Y_t, определяемый соотношением

$$y_t = c + a_1 y_{t-1} + a_2 y_{t-2} + \ldots a_p y_{t-p} + \varepsilon_t, \tag{1}$$

где ε_t – процесс типа «белый шум». Свободный член c часто приравнивается нулю (т. е. рассматриваются центрированные процессы, средний уровень которых равен нулю).

Идея авторегрессии следующая: *если в какой-то момент времени t значение показателя составляло y_t, то в последующие моменты времени t+1, t+2, ... ,t+τ значения показателя так или иначе будут зависеть от полученного на наблюдении t. Типичный пример такой зависимости — рождаемость в стране: на то, сколько родится младенцев в 2023 г., влияет количество младенцев, рожденных 23—25 лет назад.*

Введем в рассмотрение лаговый оператор

$$B^p y_t = y_{t-p},$$

тогда уравнение (1) можно записать следующим образом

$$y_t = c + a_1 B^1 y_t + a_2 B^2 y_t + ... a_p B^p y_t + \varepsilon_t,$$

перенесем все, что связано с y_t в левую часть, получим

$$y_t - a_1 B^1 y_t - a_2 B^2 y_t - ... - a_p B^p y_t = c + \varepsilon_t \qquad (2)$$

$$\left(1 - a_1 B^1 - a_2 B^2 - ... - a_p B^p\right) y_t = c + \varepsilon_t$$

Выражение $\left(1 - a_1 B^1 - a_2 B^2 - ... - a_p B^p\right)$ обозначим как функцию от оператора сдвига

$$\varphi_p(B) = 1 - a_1 B^1 - a_2 B^2 - ... - a_p B^p.$$

Тогда уравнение (2) запишется кратко

$$\varphi_p(B) y_t = c + \varepsilon_t$$

AR-процесс является стационарным тогда и только тогда, когда комплексные решения (корни) его характеристического уравнения

$$1 - a_1 x^1 - a_2 x^2 - ... - a_p x^p = 0$$

лежат вне единичного круга, т. е. $| x | > 1$ (x — комплексное число, | | -модуль комплексного числа).

Процессы, у которых $|x| = 1$, называются процессами единичного корня и являются нестационарными.

Модели скользящего среднего MA(q)

Идея о том, что исследуемая величина может зависеть от своих же значений в прошлом, получила дальнейшее развитие. Так, предполагая, что при генерации y_t всегда существует некоторая ошибка (которая, конечно же, распределена нормально, что указывает на влияние множества мелких неучтенных факторов), появилась идея о том, что будущие значения y_t могут зависеть не только от прошлых значений ряда, но и от случайной ошибки на предыдущих наблюдениях. Так появилась **модель скользящей средней порядка *q***, MA(q), которая обычно записывается в виде

$$y_t = c_1\varepsilon_{t-1} + c_2\varepsilon_{t-2} + \dots c_q\varepsilon_{t-q} + \varepsilon_t \qquad (3)$$

где ε_t – процесс типа «белый шум».

Используя введенный нами ранее лаговый оператор B, формулу (3) можно переписать в виде

$$y_t = c_1B\varepsilon_t + c_2B^2\varepsilon_t + \dots c_qB^q\varepsilon_t + \varepsilon_t. \qquad (4)$$

Вынося в (4) за скобки ε_t, получим

$$y_t = \left(1 + c_1B + c_2B^2 + \dots + c_qB^q\right)\varepsilon_t = \theta_q(B)\varepsilon_t. \qquad (5)$$

Модели авторегрессии-скользящего среднего (ARMA(p,q))

Комбинация процессов авторегрессии и скользящего среднего порядков p и q соответственно называется авторегрессионным процессом скользящего среднего (ARMA(p,q))

$$y_t = c + a_1y_{t-1} + a_2y_{t-2} + \dots a_py_{t-p} + c_1\varepsilon_{t-1} + c_2\varepsilon_{t-2} + \dots c_q\varepsilon_{t-q} + \varepsilon_t$$

или

$$\varphi_p(B)y_t = \theta_q(B)\varepsilon_t + c.$$

Например, модель ARMA(2,1) будет иметь вид:

$\varphi_2(B)y_t = \theta_1(B)\varepsilon_t + c$ или $\left(1 - a_1B - a_2B^2\right)y_t = (1 + c_1B)\varepsilon_t + c$ или

$$y_t = c + a_1y_{t-1} + a_2y_{t-2} + c_1\varepsilon_{t-1} + \varepsilon_t$$

Использование ARMA-процессов позволяет строить более компактные модели реальных временных рядов по сравнению со схожими по поведению AR- или MA-процессами.

Модели ARIMA(p, d , q)

Если наблюдаемый временной ряд имеет признаки нестационарности (например, какие-либо детерминированные тренды – линейный, полиномиальный и т.п.), то необходимо использовать модификации модели ARMA – ARIMA.

Обозначим через $\Delta y_t = y_t - y_{t-1}$ – первые разности,

$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = y_t - y_{t-1} - (y_{t-1} - y_{t-2})$ – вторые разности,

…,

$\Delta^d y_t = \Delta^{d-1} y_t - \Delta^{d-1} y_{t-1}$ – разности порядка d.

Если процесс y_t стационарен, то он называется **процессом, интегрированным нулевого порядка.** Обозначение I(0).

Если процесс y_t не стационарен, а его первые разности Δy_t стационарны, то он называется **процессом, интегрированным 1-го порядка:** I(1).

Если процесс y_t не стационарен, его первые разности Δy_t не стационарны, а вторые $\Delta^2 y_t$ стационарны, то процесс называется интегрированным 2-го порядка.

Аналогично определяется **процесс интегрированный k-го порядка I(k).**

Если

– процесс является интегрированным d-го порядка,
– и его разность d-го порядка описывается процессом ARMA(p,q),

то исходный процесс называется интегрированным процессом авторегрессии со скользящим средним в остатках ARIMA(p,d,q).

ARIMA позволяет свести нестационарный ряд к стационарному, путем взятия разностей некоторого порядка от уровней исходного временного ряда.

Характерная запись модели ARIMA(p,d,q) имеет следующий вид:

$$\Delta^d y_t = \sum_{i=1}^{p} a_i \cdot \Delta^d y_{t-i} + \sum_{j=1}^{q-1} c_j \cdot \varepsilon_{t-j} + \varepsilon_t$$

В основе прогнозирования с использованием моделей класса ARIMA лежит методология Бокса-Дженкинса.

Методология Бокса — Дженкинса

Шаг 1. Определение порядка интегрированности ряда d и переход к стационарным разностям;

Шаг 2. Анализ автокорреляционной функции АКФ и частной автокорреляционной функции ЧАКР для определения параметров p и q;

Разработаны правила по определению p и q по АКФ и ЧАКФ:

Модель	Поведение АКФ и ЧАКФ
AR(1)	АКФ экспоненциально затухает, ЧАКФ имеет статистически значимый выброс на лаге 1
AR(2)	АКФ имеет форму затухающей синусоиды или экспоненциально затухает; ЧАКФ статистически значимы только для сдвигов 1 и 2
MA(1)	АКФ имеет выброс на лаге 1 (остальные значения нулевые), ЧАКФ экспоненциально затухает монотонно либо осциллируя, то есть меняя знак
MA(2)	АКФ имеет выбросы на лаге 1 и 2 (остальные значения нулевые), ЧАКФ экспоненциально затухает или изменяется синусоидально
ARMA(1,1)	АКФ и ЧАКФ экспоненциально затухают, начиная с первого лага (первое значение не нулевое), затухание может быть монотонное или колебательное

Шаг 3. Оценивание методом максимального правдоподобия параметров модели и проверка ее адекватности (статистической значимости коэффициентов модели, и того, что остатки модели должны иметь свойства белого шума). В случае, если несколько моделей адекватны, выбирается модель с наименьшим количеством параметров и наилучшими статистическими характеристиками качества подгонки модели (например, с наименьшими значениями критерия Акаике (AIC) и Шварца (BIC));

Шаг 4. Прогнозирование.

Реализация ARIMA в R

Загрузим в R таблицу с данными по количеству арендованных велосипедов по датам за два года.

day <- read.csv("~/R/examples/time_series/day.csv")

Некоторые из столбцов:

- dteday – дата наблюдения.
- year – закодированный год наблюдения (0 = 2011, 1 = 2012).
- cnt – общее число арендованных велосипедов за этот день.

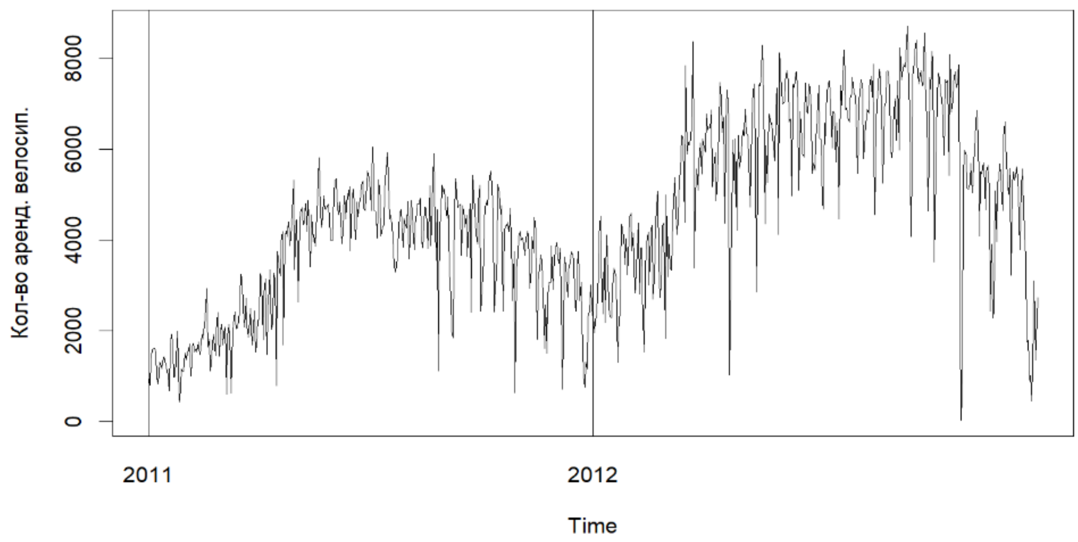
Для преобразования числового массива во временной ряд применяется функция **ts().** Если уровни ряда измерены на интервалах, меньших чем год (например, по месяцам или кварталами), можно указать количество измерений, сделанных за год, с помощью параметра frequence. Для временных рядов по месяцам frequence=12, а для данных по кварталам frequence=4. Также можно указать первый год, в который собирались данные, и первый интервал в году с помощью параметра start. Например, если первая точка данных соответствует второму кварталу 2000 года, то start=c(2000, 2).

В нашем случае данные ежедневные, поэтому указываем следующие параметры при вызове функции ts()

```
# Создаем объект временной ряд
tsday<-ts(data=day[,16],frequency = 365, start=c(2011,1))

print(tsday, calendar = T)
```

Нарисуем график временного ряда

```
#рисуем график временного ряда
plot.ts(tsday, ylab="Кол-во аренд. велосип.",xaxt="n")
axis(1,at=seq(2011, 2012, by=1), tck=1)
```



Ниже показан код для перевода данных из ежедневных в ежемесячные, а затем ежеквартальные:

```
#переход к ежемесячным данным (сначала сгруппируем данные
# по годам (столбец yr) и по месяцам (столбец mnth)), а затем с
# помощью функции sum найдем суммарное количество арендованных
# велосипедов для каждого месяца
library(dplyr)
day12<-day %>%
  group_by(yr, mnth) %>%
  summarize(summa_cnt=sum(cnt))
tsday12<-ts(day12$summa_cnt, start=c(2011,1), frequency = 12)
tsday12
# переход от ежемесячным данным к поквартальным
tsday4 <- aggregate(tsday12, nfrequency = 4, FUN = "sum")
tsday4
```

Задание: нарисуйте графики ВР для ежемесячных и ежеквартальных данных.

Для исследования временного ряда используются автокорреляционная и частная автокорреляционные функции. В среде R реализованы функции acf() и Acf(), pacf() и Pacf() соответственно для построения АКФ и ЧАКФ. У этих функций одинаковые аргументы:

Аргумент	Назначение
x	временной ряд или числовой вектор
lag.max	максимальный лаг, который будет вычислен; по умолчанию максимальный лаг вычисляется по формуле 10*log10(N/m), где N – длина временного ряда, m – количество периодов

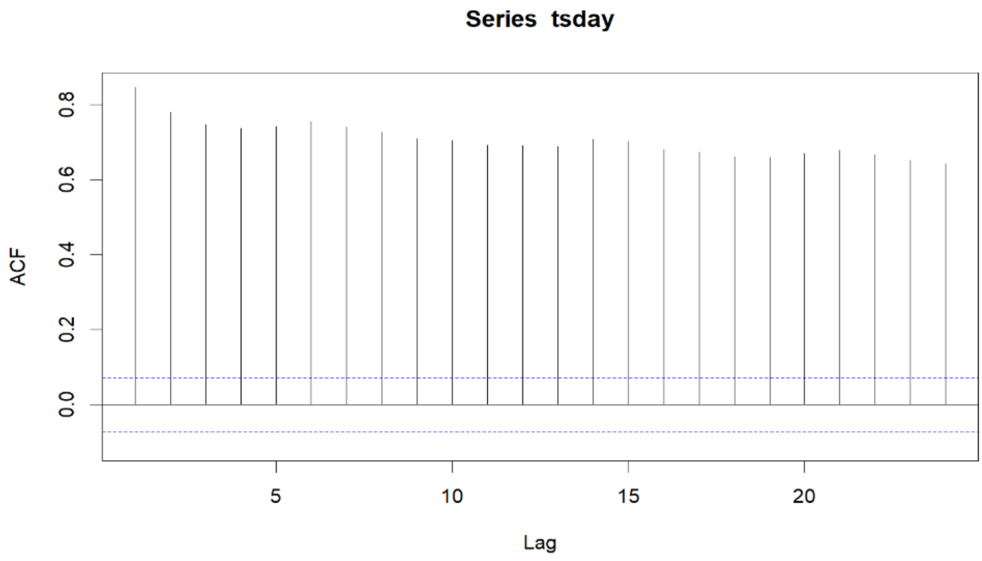
type	тип вычисляемой функции, по умолчанию равен “correlation” – автокорреляционная, “covariance” – автоковариационная, “partial” – частная автокорреляционная
plot	построение графика функции, по умолчанию равен TRUE
na.action	тип обработки пропущенных значений в исходном ВР. Если установлено na.fail (по умолчанию в функции acf()), то во ВР не допускаются пропущенные значения и выдается ошибка, в случае их наличия; если задано na.contiguous (по умолчанию в функции Acf()), то выбирается самый длинный фрагмент ВР между пропущенными значениями и по нему рассчитывается функция; na.interp – заполняются значения во ВР на основе линейной интерполяции

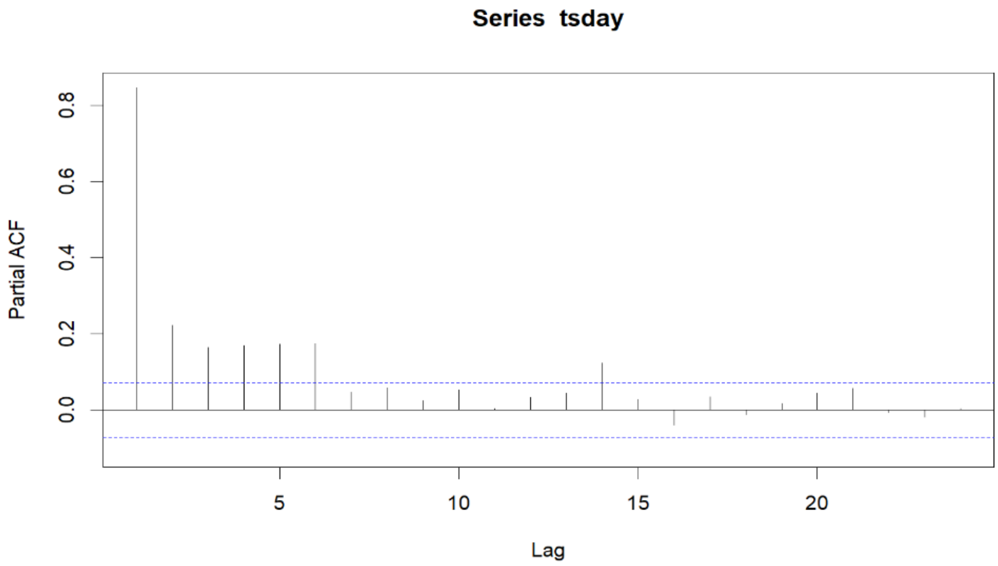
Основное отличие функций acf() и Acf() заключается в том, что Acf() не отображает значение коэффициента автокорреляции при лаге 0 (это избыточно, так как значение всегда равно единице) и по горизонтальной оси указаны лаги в единицах времени, а не в сезонных единицах.

Построим АКФ и ЧАКФ по нашему временному ряду tsday для ежедневных данных с помощью функций Acf() и Pacf()

```
install.packages("forecast")
library(forecast)
# Построение АКФ и ЧАКФ
tsday_acf<-Acf(tsday, lag.max = 24)
tsday_pacf<-Pacf(tsday, lag.max = 24)
```

Графики этих функций изображены на рисунках:



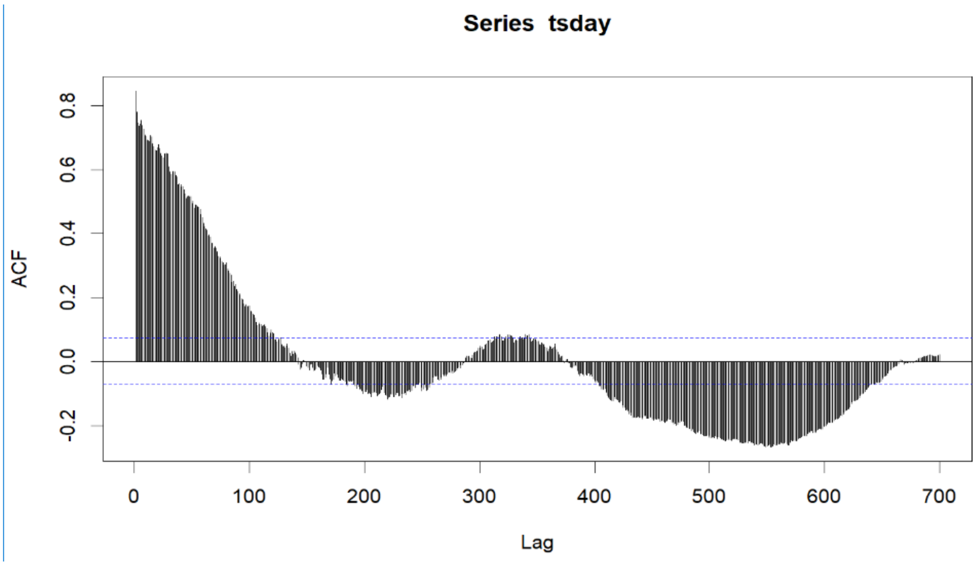


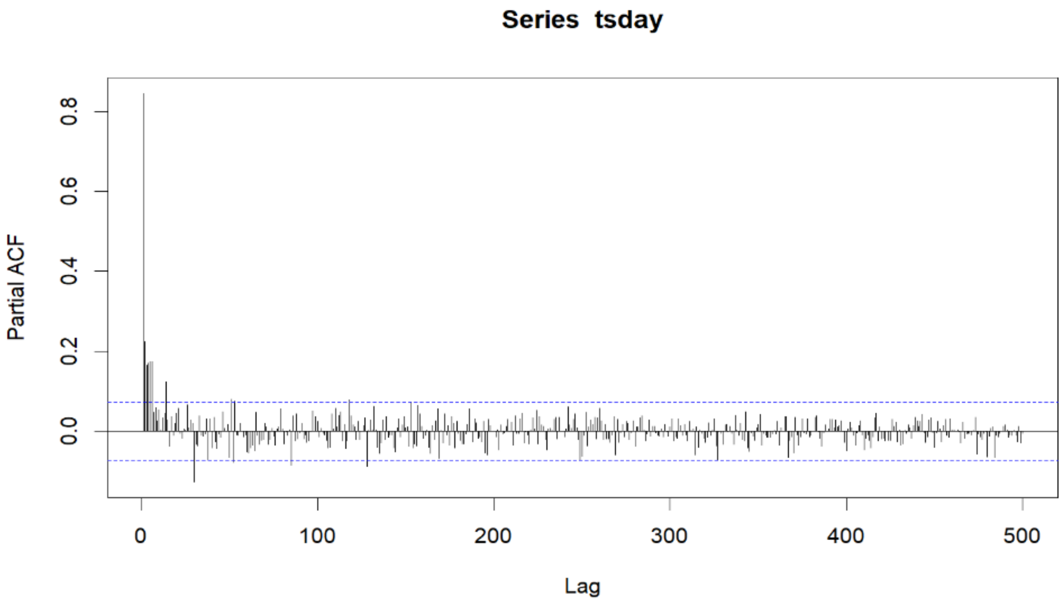
Задание: постройте АКФ и ЧАКФ для ежемесячных данных с максимальным лагом 12 и 24.

Построим также графики АКФ и ЧАКФ с максимальным лагом равным 700

```
# Построение АКФ и ЧАКФ
tsday_acf<-Acf(tsday, lag.max = 700)
tsday_pacf<-Pacf(tsday, lag.max = 700)
```

Графики приведены на рисунках





Из рисунков видно, что значения функции АКФ очень медленно убывают с увеличением лага, а значения функции ЧАКФ значимы только при первых нескольких лагах, затем лежат внутри доверительного интервала для значения равного 0. Все это говорит о нестационарности ВР и о наличии тренда.

Проверим гипотезу о наличии единичного корня в исходном ВР (о нестационарности ВР) с помощью расширенного теста Дики-Фуллера:

```
#Расширенный тест Дики-Фуллера
install.packages("tseries")
library(tseries)

adf.test(tsday)
```

Результаты применения расширенного теста Дики-Фуллера:

```
> adf.test(tsday)

        Augmented Dickey-Fuller Test

data:  tsday
Dickey-Fuller = -1.6351, Lag order = 9, p-value = 0.7327
alternative hypothesis: stationary
```

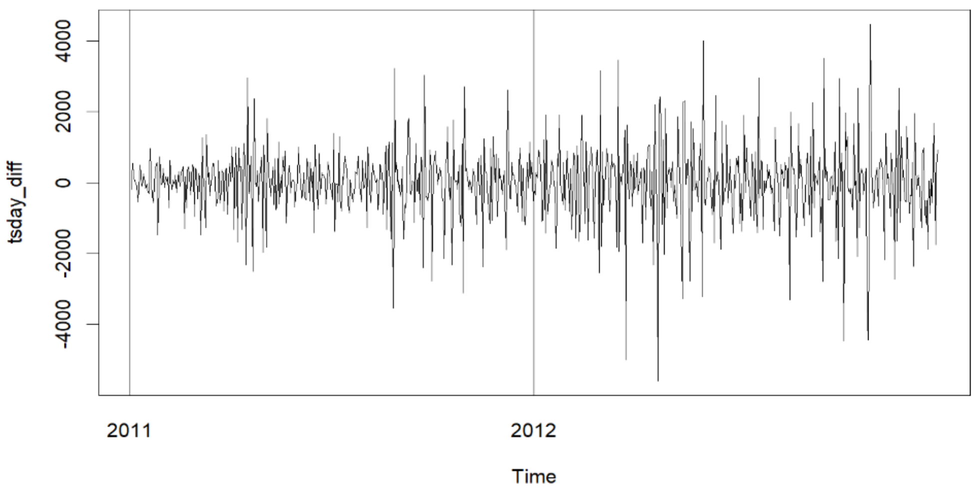
Здесь нулевая гипотеза состоит в том, что ряд нестационарный и видно, что для первых 9 лагов эта гипотеза подтверждается с вероятностью 0,7327>0.05.

Далее выполним преобразование ВР для его сведения к стационарному виду с помощью разностного оператора Δ, реализованного функцией diff().

У этой функции аргумент lag отвечает за величину лага. Например, если lag=1, то будет вычислена разница между первым и вторым элементом вектора, затем между вторым и третьим элементом и так далее. Аргумент differences – это порядок разностного оператора d

```
tsday_diff<-diff(tsday, lag=1, differences=1)
plot.ts(tsday_diff, xaxt="n")
axis(1,at=seq(2011, 2012, by=1), tck=1)
tsday_diff
adf.test(tsday_diff)
```

График ВР первых разностей изображен на рисунке



На рисунке видно, что ряд первых разностей не имеет статистически значимого тренда.

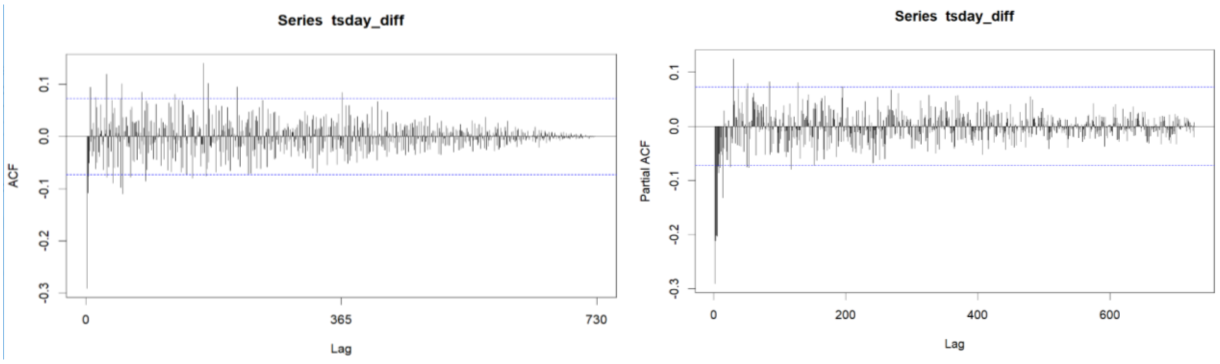
Результаты теста Дики-Фуллера:

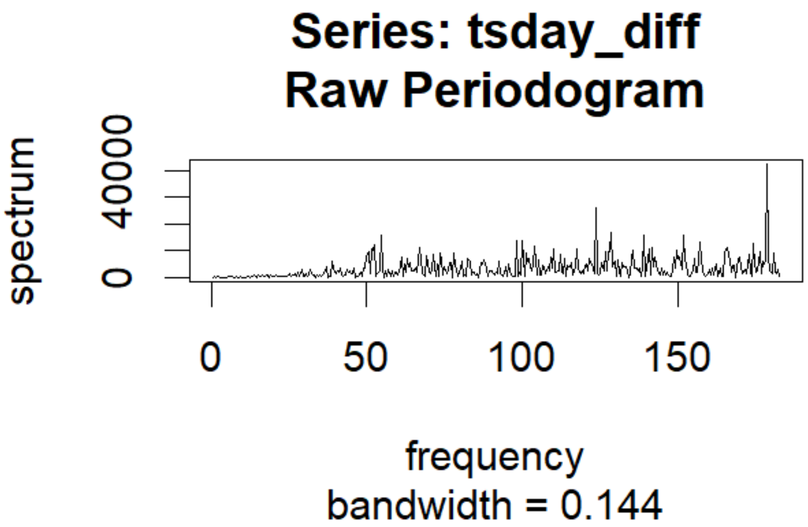
```
> adf.test(tsday_diff)

        Augmented Dickey-Fuller Test

data:  tsday_diff
Dickey-Fuller = -13.798, Lag order = 8, p-value = 0.01
alternative hypothesis: stationary
```

Гипотеза о нестационарности ВР первых разностей отвергается на уровне значимости меньшем 0,01.

Найдем АКФ и ЧАКФ для ВР первых разностей, а также нарисуем спектрограмму

```
tsday_diff_acf<-Acf(tsday_diff)
tsday_diff_acf<-Pacf(tsday_diff)
s<-spec.pgram(tsday_diff, detrend=FALSE, log="no", fast=FALSE,
              plot=TRUE)
round(s$spec, 2)
```





В R реализована возможность построения моделей класса ARIMA (p,d,q) с автоматическим подбором структуры модели с помощью функции **auto.arima().** Функция возвращает наилучшую модель ARIMA в соответствии со значением выбранного информационного критерия AIC или

BIC. Функция перебирает все варианты структуры модели в пределах заданных максимальных порядков p, d и q для модели:

```
tsdayA1<-auto.arima(tsday, ic="aic", max.p=5, max.d = 3, max.q = 5,
                    max.P=0, max.D=0, max.Q=0)
```

Результаты выполнения функции:

```
Series: tsday
ARIMA(1,1,1)

Coefficients:
         ar1      ma1
      0.3584  -0.8896
s.e.  0.0423   0.0192

sigma^2 = 857769:  log likelihood = -6021.96
AIC=12049.91   AICc=12049.94   BIC=12063.69
```

Таким образом, получаем модель ARIMA(p=1, d=1, q=1)

$$y_t = \underset{(0{,}0423)}{0.3584}\, y_{t-1} - \underset{(0{,}0192)}{0.8896}\, \varepsilon_{t-1} + \varepsilon_t,$$

где в скобках указаны значения стандартных ошибок коэффициентов модели (они намного меньше значений параметров, что говорит о статистической их значимости).

Значение информационного критерия Акаике здесь самое маленькое из возможных с учетом ограничений на параметры. Значение логарифма функции правдоподобия равно -6021,96.

Параметры max.P, max.Q, max.D служат уже для подбора модели SARIMA (p,d,q) (P,D,Q), которая учитывает сезонность, циклические колебания.

В результате применения функции accuracy() получаем характеристики точности построенной модели

```
> accuracy(tsdayA1)
                   ME     RMSE      MAE       MPE     MAPE      MASE
Training set 12.87842 924.2558 647.6408 -44.46123 58.52916 0.2788367
                   ACF1
Training set 0.01117124
```

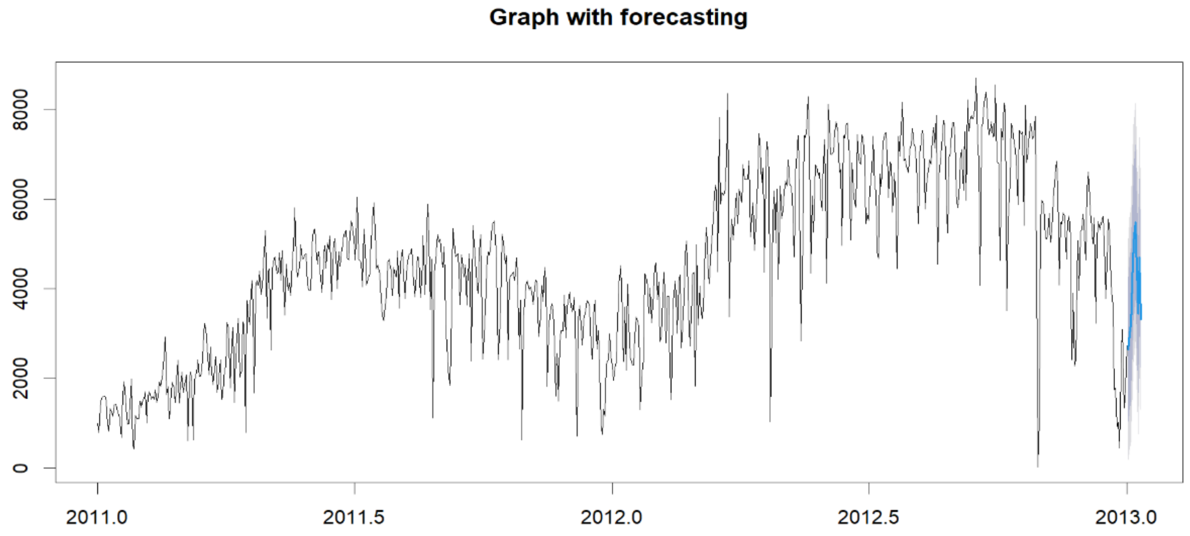
где

ME	$\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t$	средняя ошибка

MAE	$\frac{1}{n}\sum_{t=1}^{n}\lvert\varepsilon_t\rvert$	Средняя абсолютная ошибка
RMSE	$\sqrt{\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2}$	Квадратный корень из среднеквадратичной ошибки
MPE	$\frac{1}{n}\sum_{t=1}^{n}\frac{\varepsilon_t}{y_t}\cdot 100\%$	Средняя процентная ошибка
MAPE	$\frac{1}{n}\sum_{t=1}^{n}\frac{\lvert\varepsilon_t\rvert}{y_t}\cdot 100\%$	Средняя абсолютная процентная ошибка
ACF1		Коэффициент автокорреляции ошибки ε_t с лагом 1

С помощью функции forecast() сделаем прогноз на 10 дней вперед и нарисуем график

```
plot(forecast(tsdayA1, 10), main = "Graph with forecasting")
```



Задание

постройте прогноз на 30 дней вперед и нарисуйте график.