Семинар 1. Введение в анализ временных рядов

Достовалов М.Ю.

25.01.2024

1. Введение данных «вручную». Численность населения РФ, млн.чел.

Создание вектора

y=c(146.8, 146.9, 146.8, 146.7, 146.2, 145.6)  
y

## [1] 146.8 146.9 146.8 146.7 146.2 145.6

Создание ts объекта (временного ряда)

year = ts(y, start=2017)  
year

## Time Series:  
## Start = 2017   
## End = 2022   
## Frequency = 1   
## [1] 146.8 146.9 146.8 146.7 146.2 145.6

year1 = ts(c(146.8, 146.9, 146.8, 146.7, 146.2, 145.6), start=2017)  
year1

## Time Series:  
## Start = 2017   
## End = 2022   
## Frequency = 1   
## [1] 146.8 146.9 146.8 146.7 146.2 145.6

Удаление вектора/объекта

remove(year, year1)

Зададим переменные времени

year = 2017:2022  
year

## [1] 2017 2018 2019 2020 2021 2022

t = 1:6  
t

## [1] 1 2 3 4 5 6

Обьединим переменные в одно множество с помощью команды column bind (cbind)

cbind(t,year, y)

## t year y  
## [1,] 1 2017 146.8  
## [2,] 2 2018 146.9  
## [3,] 3 2019 146.8  
## [4,] 4 2020 146.7  
## [5,] 5 2021 146.2  
## [6,] 6 2022 145.6

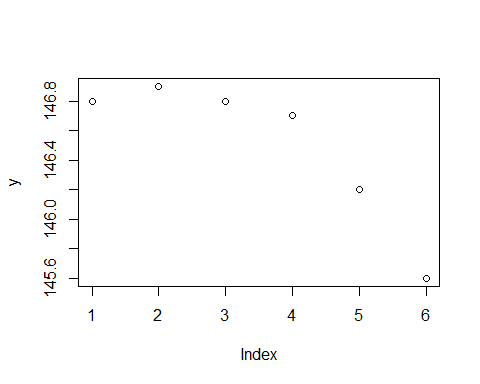
Создание квартального временного ряда начиная с III квартала 2017 года

q = ts(y, start=c(2017,3), frequency=4)   
q

## Qtr1 Qtr2 Qtr3 Qtr4  
## 2017 146.8 146.9  
## 2018 146.8 146.7 146.2 145.6

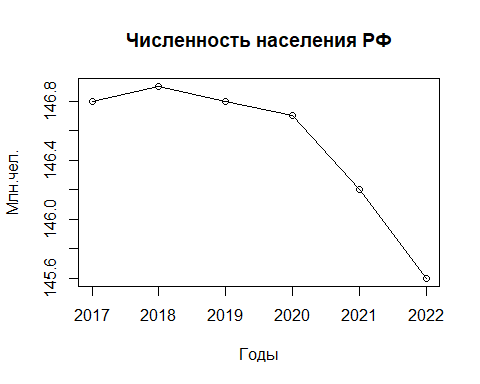
Построение графика ВР

plot(y)



Добавим заголовок и подписи на график

plot(year, y,   
 main = "Численность населения РФ", # Заголовок графика  
 xlab = "Годы", # Подпись оси X  
 ylab = "Млн.чел.", # Подпись оси Y  
 type = "o") # l-линия, р-точки, o-оба типа и пр.



1. Различные типы ВР, анализ составляющих ВР.

Устанавим пакеты и подгрузим библиотеки для работы с данными

# install.packages("tidyverse")  
library(tidyverse)

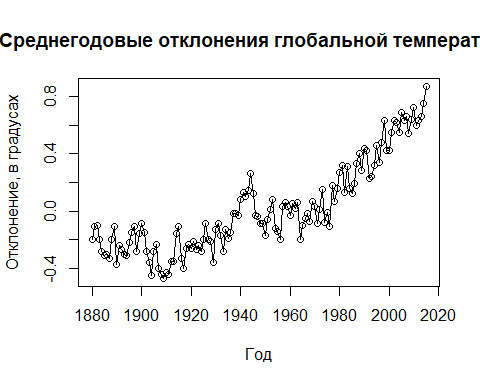
Пример 1. Глобальное потепление (Global Warming): импорт данных в R.

Постройте график и проанализируйте динамику основного тренда.

library(haven) # импортирует и экспортирует файлы SPSS, Stata

## Warning: пакет 'haven' был собран под R версии 4.3.2

# file.choose()  
globtemp = read\_dta("globtemp.dta") # укажите свой путь, где лежит файл  
plot(globtemp$year, globtemp$g\_temp,  
 main="Среднегодовые отклонения глобальной температуры",  
 xlab="Год",  
 ylab="Отклонение, в градусах",   
 type="o")



Пример 2. Объём авиаперевозок в ЕС: ВР с сезонностью, импорт данных.

Загрузим данные из файла

EU = read\_dta("EU.dta") # укажите свой путь, где лежит файл

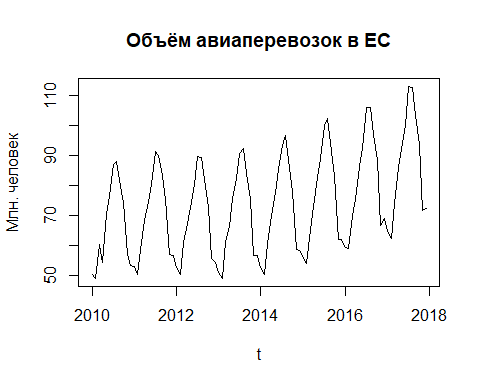
Cоздадим ts обьект и зададим время

eu1 = ts(EU$eu, start=c(2010,1), frequency=12)   
eu1

## Jan Feb Mar Apr May Jun Jul  
## 2010 50.39257 49.19072 60.47453 54.39661 70.78214 76.56784 87.23588  
## 2011 53.30106 50.52824 61.08048 68.55109 74.79131 80.47412 91.54129  
## 2012 52.79370 50.62262 61.00895 67.46822 72.91467 80.38313 89.85667  
## 2013 51.17889 49.29883 61.45054 66.42345 75.77541 82.42004 90.86588  
## 2014 52.77043 50.54291 60.94522 70.12556 77.67204 85.34586 93.35410  
## 2015 56.01254 54.11756 65.41499 73.14019 82.65354 89.20668 99.96532  
## 2016 59.34610 59.27203 69.99723 75.41469 86.34116 92.33430 106.11618  
## 2017 64.96925 62.51253 74.09251 85.97771 92.08812 99.93534 113.10291  
## Aug Sep Oct Nov Dec  
## 2010 88.12096 79.70252 74.52489 57.10109 53.58322  
## 2011 90.24802 83.60719 75.61700 57.30278 56.76479  
## 2012 89.47761 82.20876 72.88120 55.91630 54.46218  
## 2013 92.41832 83.99825 75.48750 56.96476 56.66147  
## 2014 96.86635 85.79675 78.05756 58.92865 58.56807  
## 2015 102.49474 91.54339 83.46172 62.26496 61.99648  
## 2016 106.04417 96.97013 88.56368 66.73495 68.98740  
## 2017 112.84474 103.44159 93.69891 71.71396 72.29410

Построим график ВР

plot.ts(eu1,  
 main="Объём авиаперевозок в ЕС",  
 ylab="Млн. человек",   
 xlab="t")



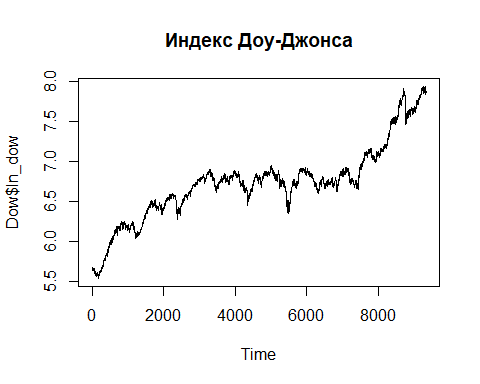
Пример 3. Индекс Доу-Джонса: высокочастотные ежедневные данные. \*19 октября 1987. Чёрный понедельник: индекс Доу-Джонса пережил самое большое падение в истории — на 22,6 %.

Загрузим данные из файла

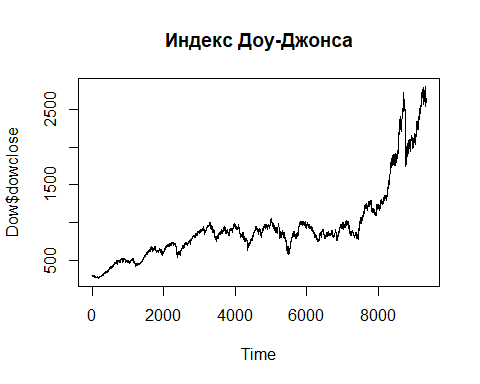
Dow = read\_dta("dow1.dta")

Построим график ВР

plot.ts(Dow$ln\_dow, main="Индекс Доу-Джонса")



plot.ts(Dow$dowclose, main="Индекс Доу-Джонса")



Рассчитаем первую разность и для нее построим график

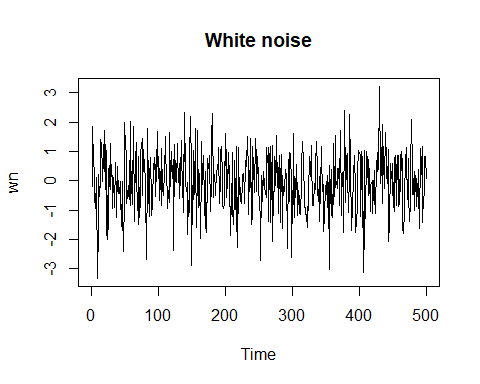
d\_dow = diff(Dow$ln\_dow)  
plot.ts(d\_dow, main="DJ Доходность")



1. Генерирование случайных процессов в R.

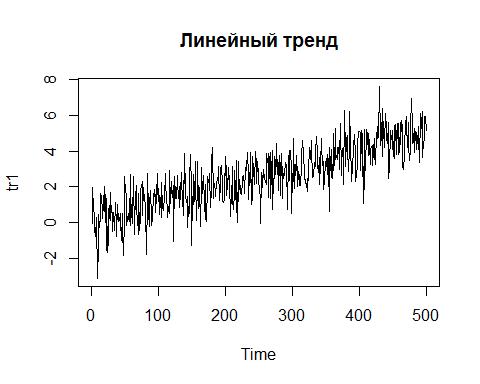
Сгенерируем белый шум (БМ) и построим график

wn = rnorm(500,0,1) # 500 наблюдений N(0,1)   
plot.ts(wn, main="White noise")



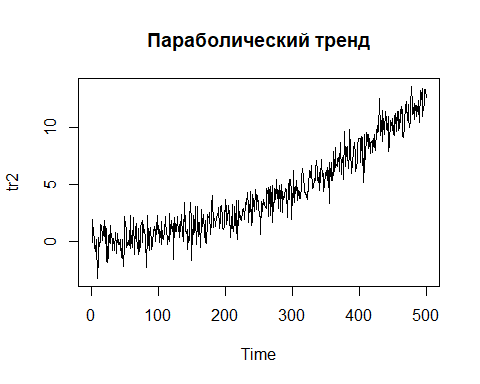
Сгенерируем линейный тренд + БШ и построим график

t = 1:500 # переменная времени от 1 до 500  
tr1=0.1+0.01\*t+wn  
plot.ts (tr1, main="Линейный тренд")



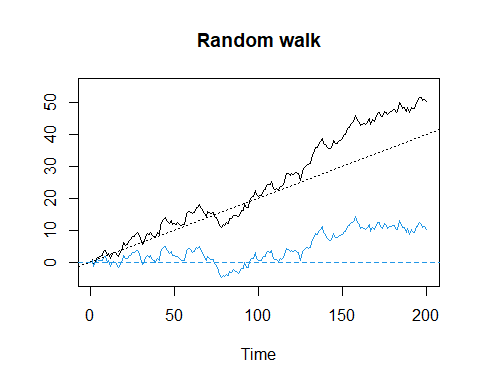
Сгенерируем параболический тренд + БШ и построим график

tr2=0.1+0.00005\*t^2+wn  
plot.ts (tr2, main="Параболический тренд")



Сгенерируем случайное блуждание/блуждание с дрейфом (2-е представление) и построим график

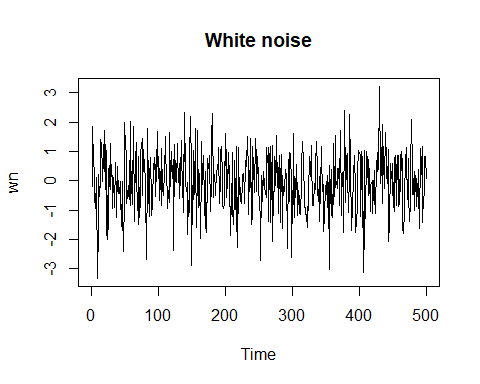
set.seed(154)   
w = rnorm(200); x = cumsum(w) # две команды в одной строке задаются через ";"  
wd = w +.2; xd = cumsum(wd)  
plot.ts(xd, ylim=c(-5,55), main="Random walk", ylab='')  
lines(x, col=4); abline(h=0, col=4, lty=2); abline(a=0, b=.2, lty=3)



1. Стационарность ВР.

Построим график и найдем характеристики для сгенерированного БШ

plot.ts(wn, main="White noise")



summary(wn)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -3.32342 -0.69892 -0.02768 -0.03376 0.71801 3.21982

Проверим его на стационарность

library(psych)

## Warning: пакет 'psych' был собран под R версии 4.3.2

##   
## Присоединяю пакет: 'psych'

## Следующие объекты скрыты от 'package:ggplot2':  
##   
## %+%, alpha

describe(wn)

## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 500 -0.03 1 -0.03 -0.01 1.09 -3.32 3.22 6.54 -0.21 0.18 0.04

describe(wn[1:250])

## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 250 0.02 1.01 0.01 0.05 1.04 -3.32 2.34 5.67 -0.26 0.06 0.06

describe(wn[251:500])

## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 250 -0.09 0.99 -0.09 -0.07 1.06 -3.12 3.22 6.34 -0.17 0.31 0.06

1. Показатели динамики ВР.

Рассчитаем лаги и разности

Ly=lag(y)  
Dy=y-Ly  
Ty= (y-Ly)/Ly\*100  
DLny=log(y/Ly)\*100  
cbind(y, Ly, Dy, Ty, DLny)

## y Ly Dy Ty DLny  
## [1,] 146.8 NA NA NA NA  
## [2,] 146.9 146.8 0.1 0.06811989 0.0680967  
## [3,] 146.8 146.9 -0.1 -0.06807352 -0.0680967  
## [4,] 146.7 146.8 -0.1 -0.06811989 -0.0681431  
## [5,] 146.2 146.7 -0.5 -0.34083163 -0.3414138  
## [6,] 145.6 146.2 -0.6 -0.41039672 -0.4112412

Построим столбиковую диаграмму

barplot(Ty)

