

## Домашнее задание 2. Временной ряд: введение

Николаева В.Н. Достовалов М.Ю.

30.01.2024

```
# импорт данных
library(dplyr)

##
## Присоединяю пакет: 'dplyr'

## Следующие объекты скрыты от 'package:stats':
##
## filter, lag

## Следующие объекты скрыты от 'package:base':
##
## intersect, setdiff, setequal, union

data <- readxl::read_excel("WB_Russia.xls", sheet = 'Лист1')

# подготовка данных
names <- data$`Indicator Code`
data <- data %>% select(-1)
data <- t(data)
colnames(data) <- names
data <- data.frame(data)
```

### Показатель: индекс Джинни

Индекс Джини - это статистический показатель, который используется для оценки экономического неравенства в обществе. Он может варьироваться от 0 до 1, где 0 соответствует полному равенству (все имеют одинаковый доход), а 1 - абсолютному неравенству (весь доход имеет один человек, а все остальные - нулевой).

Наиболее часто в современных экономических расчётах в качестве изучаемого признака берётся уровень годового дохода. Коэффициент Джини можно определить как макроэкономический показатель, характеризующий дифференциацию денежных доходов населения в виде степени отклонения фактического распределения доходов от абсолютно равного их распределения между жителями страны.

**Страна:** Российская Федерация

**Период:** с 1997 по 2020 год

```
index <- na.omit(data$SI.POV.GINI)
```

```
# лаговые признаки
lag_1 <- lag(index)
lag_2 <- lag(index, 2)
```

```
lag_3 <- lag(index, 3)
lag_4 <- lag(index, 4)
lag_5 <- lag(index, 5)

# формулы расчёта показателей
increment <- index - lag_1
growth_rate <- increment / lag_1
logarithmic_diff <- log(index) - log(lag_1)
```

```
index_with_lags <- cbind(
  index,
  lag_1, lag_2, lag_3, lag_4, lag_5,
  increment, growth_rate, logarithmic_diff
)
round(index_with_lags, 3)
```

```
##      index lag_1 lag_2 lag_3 lag_4 lag_5 increment growth_rate
## [1,] 38.4    NA    NA    NA    NA    NA        NA         NA
## [2,] 38.1 38.4    NA    NA    NA    NA       -0.3      -0.008
## [3,] 37.4 38.1 38.4    NA    NA    NA       -0.7      -0.018
## [4,] 37.1 37.4 38.1 38.4    NA    NA       -0.3      -0.008
## [5,] 36.9 37.1 37.4 38.1 38.4    NA       -0.2      -0.005
## [6,] 37.3 36.9 37.1 37.4 38.1 38.4        0.4       0.011
## [7,] 40.0 37.3 36.9 37.1 37.4 38.1        2.7       0.072
## [8,] 40.3 40.0 37.3 36.9 37.1 37.4        0.3       0.007
## [9,] 41.3 40.3 40.0 37.3 36.9 37.1        1.0       0.025
## [10,] 41.0 41.3 40.3 40.0 37.3 36.9       -0.3      -0.007
## [11,] 42.3 41.0 41.3 40.3 40.0 37.3        1.3       0.032
## [12,] 41.6 42.3 41.0 41.3 40.3 40.0       -0.7      -0.017
## [13,] 39.8 41.6 42.3 41.0 41.3 40.3       -1.8      -0.043
## [14,] 39.5 39.8 41.6 42.3 41.0 41.3       -0.3      -0.008
## [15,] 39.7 39.5 39.8 41.6 42.3 41.0        0.2       0.005
## [16,] 40.7 39.7 39.5 39.8 41.6 42.3        1.0       0.025
## [17,] 40.9 40.7 39.7 39.5 39.8 41.6        0.2       0.005
## [18,] 39.9 40.9 40.7 39.7 39.5 39.8       -1.0      -0.024
## [19,] 37.7 39.9 40.9 40.7 39.7 39.5       -2.2      -0.055
## [20,] 36.8 37.7 39.9 40.9 40.7 39.7       -0.9      -0.024
## [21,] 37.2 36.8 37.7 39.9 40.9 40.7        0.4       0.011
## [22,] 37.5 37.2 36.8 37.7 39.9 40.9        0.3       0.008
## [23,] 37.7 37.5 37.2 36.8 37.7 39.9        0.2       0.005
## [24,] 36.0 37.7 37.5 37.2 36.8 37.7       -1.7      -0.045
##      logarithmic_diff
## [1,]                NA
## [2,]              -0.008
## [3,]              -0.019
## [4,]              -0.008
## [5,]              -0.005
## [6,]               0.011
## [7,]               0.070
## [8,]               0.007
## [9,]               0.025
## [10,]              -0.007
```

```
## [11,]      0.031
## [12,]     -0.017
## [13,]     -0.044
## [14,]     -0.008
## [15,]      0.005
## [16,]      0.025
## [17,]      0.005
## [18,]     -0.025
## [19,]     -0.057
## [20,]     -0.024
## [21,]      0.011
## [22,]      0.008
## [23,]      0.005
## [24,]     -0.046
```

```
plot(ts(index, start = 1997), main = 'Индекс Джини', ylab = '%')
```



- Видно, что в 2020 году (последний год, когда индекс публиковался), индекс принял самое низкое значение за всю историю наблюдений в РФ. Темп прироста его составил -4,6%, что говорит о том, что он стал ниже, чем в 2019 году на 4,6%.
- После 2000 года, индекс начал стремительно расти, к 2010 году индекс спустился до локального минимума, при этом оставшись довольно высоким. Затем, после роста в 2013 году начал стремительно падать.

```
library(glue)
```

```
mean_index <- mean(index)
var_index <- var(index)
```

```

len_index <- length(index)

# расчёт ACF
dev_from_mean <- index_with_lags[, 'index'] - mean_index
variance <- sum(dev_from_mean^2) / len_index

acf_vector <- c()
for (i in 1:5) {
  lag_dev_from_mean <- index_with_lags[, glue('lag_{i}')] - mean_index
  auto_cov <- sum((dev_from_mean * lag_dev_from_mean), na.rm = T) / (len_index -
i)
  auto_corr <- auto_cov / variance
  acf_vector[i] <- auto_corr
}

acf_vector

## [1] 0.80306343 0.56337975 0.30198341 0.12713150 -0.05773391

```

Видно, что ACF постепенно убывает с увеличением лагов

```

# расчёт PACF
pacf1 <- acf_vector[1]
pacf2 <- (acf_vector[2] - acf_vector[1]^2) / (1 - acf_vector[1]^2)
pacf3 <- (acf_vector[1]^3 + acf_vector[1] * acf_vector[2]^2 + acf_vector[3] -
(acf_vector[1]^2) * acf_vector[3] - 2 * acf_vector[1] * acf_vector[2]) / (1 + 2
* (acf_vector[1]^2) * acf_vector[2] - acf_vector[2]^2 - 2 * acf_vector[1]^2)

pacf_vector <- c(pacf1, pacf2, pacf3)

pacf_vector

## [1] 0.8030634 -0.2296075 -0.2079209

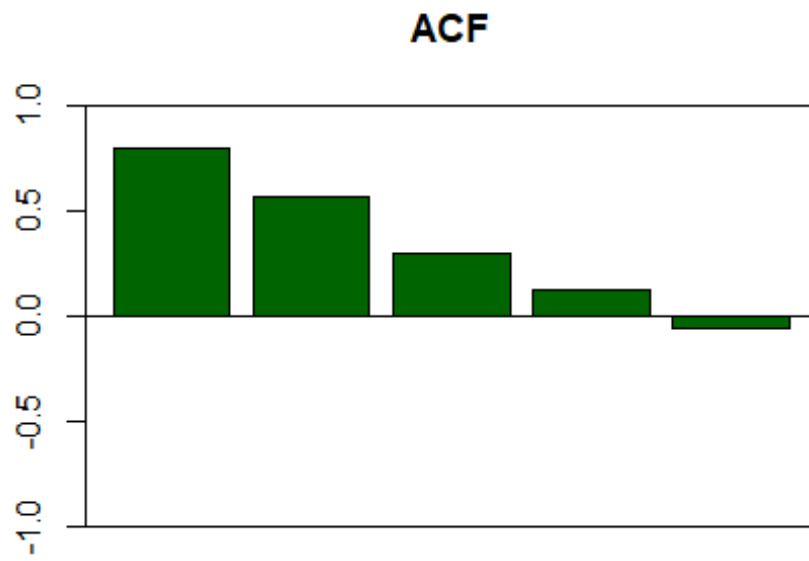
```

Видно, что высокое значение PACF принимает только в первом лаге, когда как на остальных лагах значения визуальнo куда ниже

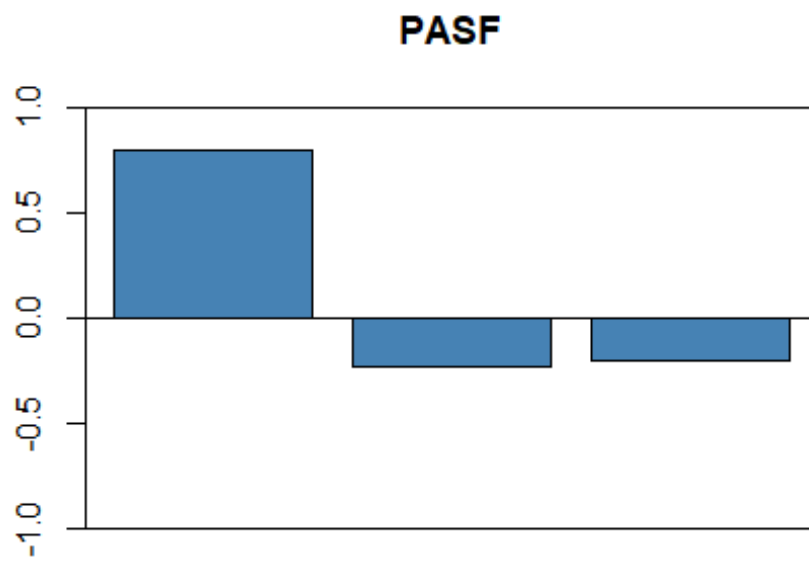
```

barplot(acf_vector, main = 'ACF', col = 'darkgreen', ylim = c(-1, 1))
abline(h = 0)
box()

```



```
barplot(pacf_vector, main = 'PASF', col = 'steelblue', ylim = c(-1, 1))  
abline(h = 0)  
box()
```



Исходя из визуального анализа автокорреляционной функции (ACF) и частной автокорреляционной функции (PACF), можно сделать вывод, что измерения показателя индекса Джини могут соответствовать авторегрессионной модели первого порядка  $AR(1)$ .