

Зимняя Школа по Эконометрике 2023

Домашнее задание по применению bootstrap

Галина Писарева

1. Идея: для чего будем применять бутстрэп

Пусть мы имеем данные о дневных доходностях акций компании МТС. Задача состоит в том, чтобы оценить среднюю доходность акций и их риск, причем нас интересуют как точечные оценки, так и доверительные интервалы. Кроме того, мы хотим понять, насколько чувствительны акции МТС к рыночному индексу МОЕХ. Используем для этих целей bootstrap.

2. Работа с данными о доходностях и их дисперсии

Вся работа производится в Google Colab. Загружаем и считываем excel-файл, создав для его содержимого таблицу df. Построим график ряда доходности акций МТС (рис. 1):

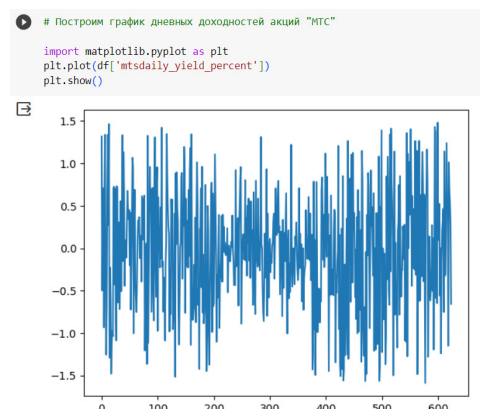


Рис. 1: График дневных доходностей акций МТС

Для того чтобы убрать большие скачки, данные были предварительно «очищены»: удалены по 10% самых больших и самых маленьких значений. Несмотря на это,



как видно из рис. 1, дисперсия все равно «ползает», хоть и едва заметно. Для упрощения решения не будем брать это в расчет (хотя, по-хорошему, нужно применять GARCH-модели).

Вычислим точечные оценки для математического ожидания и для дисперсии ряда (получим среднюю доходность и средний риск). С помощью команды `np.mean()`. Оно оказалось равным **-0,036**, что близко к 0, как и ожидалось.

А теперь применим бутстрэп, чтобы задать доверительный интервал. Мы хотим узнать, в каких пределах следует ожидать доходность акций МТС с 95% уверенностью (рис. 2, 3). Этот интервал покрывает 0.

```
# Запустим генератор (псевдо)случайных чисел и "размножим" оценки
rng = np.random.default_rng(121123)
n_boot = 10000
mu_hat_star = [np.mean(rng.choice(x, size = len(x))) for i in range (n_boot)]
mu_hat_star [1:10]
```

Рис. 2: Применение bootstrap для получения массива точечных оценок

```
# Построим доверительный интервал для математического ожидания доходности акций с использованием бутстрэп (5% ошибка)
boot_x = IIDbootstrap(x, seed = 121123)
boot_x.conf_int(np.mean, method = 'basic', reps = 10000, size = 0.95)

array([[0.09052387],
       [0.01786174]])
```

Рис. 3: Построение доверительного интервала для средней доходности акций МТС

Теперь работаем с риском. С помощью команды `np.var()` нам удалось получить оценку дисперсии доходности: **0,476**. Снова используем bootstrap для задания доверительного интервала (рис. 4).

```
[20] # Точечная оценка для риска акций (дисперсии их доходности)
np.var(x)

0.4759331642745556

[21] # Интервал
boot_x.conf_int(np.var, method = 'basic', reps = 10000, size = 0.95)

array([[0.43084821],
       [0.52120595]])
```

Рис. 4: Точечная оценка и доверительный интервал для разброса доходности

Наконец, нам бы хотелось понять, как сильно коррелирована доходность акций МТС с рынком (представлен МОЕХ). Чтобы это устроить, используем возможности bootstrap: вычислим точечную оценку корреляции и построим доверительный интервал для нее. В процессе нам понадобится использовать `obs id`, чтобы гарантировать «привязку» пар значений индекса и акций (рис. 5, 6).



```
[30] def corr(x, y):
      corr_mat = np.corrcoef(x, y)
      return corr_mat[0, 1]

[31] # Точечная оценка корреляции
      corr(x,y)

0.7624225428827404
```

Рис. 5: Точечная оценка корреляции доходности акций МТС и рыночного индекса

```
[32] obs_id = rng.choice(range(len(x)), size = len(x))
      obs_id[1:10]

array([ 18, 497,  90, 179,  14,  79, 290, 476,  74])

[33] # Доверительный интервал для корреляции
      boot_xy = IIDBootstrap(x, y, seed = 121123)
      boot_xy.conf_int(corr, method = 'basic', reps = 10000, size = 0.95)

array([[0.72244597],
       [0.80728873]])
```

Рис. 6: Доверительный интервал для корреляции (95%)

Из рис. 5 видно, что точечная оценка корреляции составила около **0,76**, что означает довольно сильную корреляцию. Доверительный же интервал располагается в границах от **0,72** до **0,81**.

3. Заключительные ремарки

Проведенный анализ показывает, что доходность акций МТС близка к 0 в среднем с небольшим разбросом вокруг среднего. С вероятностью 95% доходность акций компании находится в границах интервала от -0,09 до 0,02.

Акции МТС оказываются довольно сильно коррелированными с доходностью рыночного индекса МОЕХ.

У использованного метода существуют свои ограничения. В частности, мы предполагаем, что ряд доходности является слабо стационарным, хотя в реальности это может быть не так (дисперсия может быть непостоянной во времени, и тогда полученные оценки не будут достаточно точными, чтобы рассчитывать на них).