

# Курсовая работа на тему: Проверка гипотезы о равенстве дисперсий логарифмической доходности индекса фондового рынка и входящих в его состав акций

ВЫПОЛНИЛА: СТУДЕНТКА ГРУППЫ ПМ19-1 БАШМАКОВА А.А.

НАУЧНЫЙ РУКОВОДИТЕЛЬ:

ДОЦЕНТ, К.Э.Н.

ГРИНЕВА Н. В.

# Цель работы

Проверка гипотезы о равенстве дисперсий логарифмической доходности фондового рынка и входящих в его состав акций на реальных данных с использованием критерия Фишера как относительно простого, но находящего широкое распространение в дисперсионном анализе.

Исследуемые данные: котировки акций компаний, входящих в индекс ММВБ потребительского сектора (MOEXCN)

# Критерий Фишера

Пусть имеются две независимые выборки из нормальных распределений:

$$X_1, \dots, X_m \sim N(\mu_x, \sigma_x^2),$$

$$Y_1, \dots, Y_n \sim N(\mu_y, \sigma_y^2).$$

Будем считать, что параметры  $\mu_x, \sigma_x^2, \mu_y, \sigma_y^2$  известны. В качестве основной гипотезы примем  $H_0: \sigma_x^2 = \sigma_y^2$ , а в качестве дополнительной одну из трёх гипотез:

- 1)  $H_1: \sigma_x^2 > \sigma_y^2$ ;
- 2)  $H_1: \sigma_x^2 < \sigma_y^2$ ;
- 3)  $H_1: \sigma_x^2 \neq \sigma_y^2$ .

Теорема о построении критериев для проверки гипотезы с известным уравнением значимости  $\alpha$ :

Если верна  $H_0$ , то

$$\frac{s_x^2}{s_y^2} \sim F(m-1, n-1),$$

где  $s_x^2 = \sum_{i=1}^m (X_i - \bar{X})^2$ ,  $s_y^2 = \sum_{i=1}^n (Y_i - \bar{Y})^2$ , а  $F(m-1, n-1)$  – распределение Фишера с  $m-1$  и  $n-1$  степенями свободы.

# Данные

КОМПАНИИ ИНДЕКСА МОЕХСН

Тикер	Название компании
AGRO	ПАО «РУСАГРО»
<u>PJPq</u>	X5 <u>Retail Group</u>
AQUA	ПАО «Русская аквакультура»
DSKY	ПАО «Детский мир»
MGNT	ПАО «Магнит»
<u>LNTAq</u>	ПАО "ЛЕНТА"
MDMGDR	МГП «Мать и Дитя»
MVID	ПАО «М.Видео»
BELU	ПАО «Белуга Групп»
APTK	ПАО «Аптечная сеть 36,6»
SVAV	ПАО « <u>Соллерс</u> »
ORUP	OR GROUP

# Предварительный анализ:

КОЛИЧЕСТВО ТОРГОВЫХ ДНЕЙ

	A	B	C	D	E	F	G	H
1		Ticker	2015	2016	2017	2018	2019	2020
2	0	AGRO	229	252	252	254	252	250
3	1	PJPq	0	0	0	234	252	250
4	2	AQUA	151	231	228	223	251	250
5	3	DSKY	0	0	224	254	252	250
6	4	MGNT	250	252	252	254	252	250
7	5	LNTAq	250	252	252	254	252	250
8	6	MDMGDR	0	0	0	0	0	38
9	7	MVID	250	252	252	254	252	250
10	8	BELU	0	0	104	222	252	250
11	9	APTK	250	252	252	254	252	250
12	10	SVAV	250	252	252	254	252	250
13	11	ORUP	0	0	50	252	252	250

# Предварительный анализ:

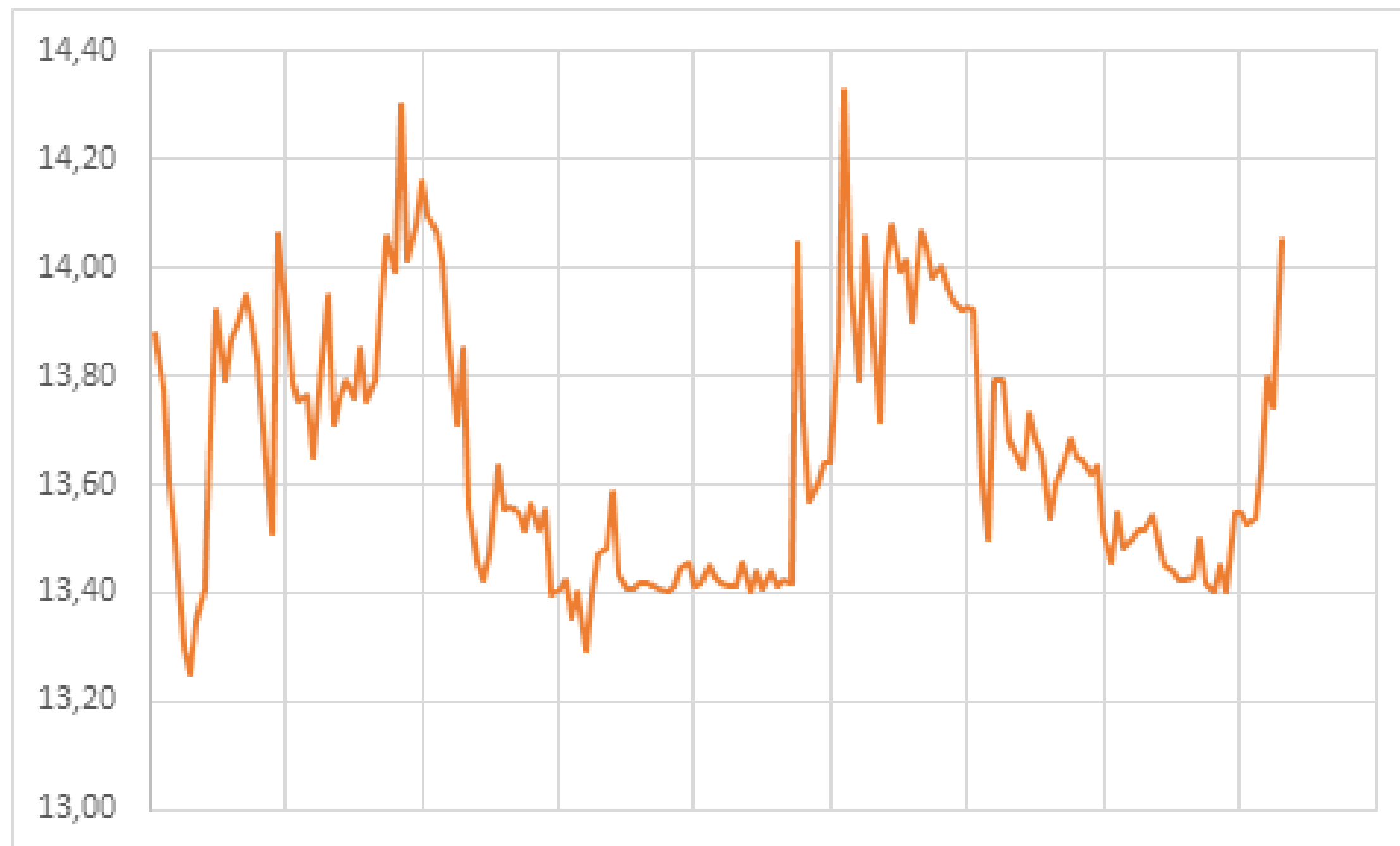
МАКСИМАЛЬНЫЕ ДНЕВНЫЕ ОТНОСИТЕЛЬНЫЕ СКАЧКИ ЦЕН

	A	B	C	D	E	F	G	H
1		Ticker	2015	2016	2017	2018	2019	2020
2	0	AGRO	-0,112	-0,051	-0,047	-0,055	-0,032	-0,066
3	1	MGNT	-0,095	-0,073	-0,112	-0,102	-0,035	-0,136
4	2	LNTAq	-0,075	-0,055	-0,043	-0,077	-0,035	-0,134
5	3	MVID	-0,118	-0,077	-0,057	-0,055	-0,069	-0,064
6	4	APTK	-0,193	-0,104	-0,088	-0,065	-0,071	-0,109
7	5	SVAV	-0,064	-0,059	-0,056	-0,064	-0,047	-0,074

	A	B	C	D	E	F	G	H
1		Ticker	2015	2016	2017	2018	2019	2020
2	0	AGRO	0,209	0,055	0,051	0,058	0,043	0,104
3	1	MGNT	0,082	0,078	0,059	0,079	0,047	0,07
4	2	LNTAq	0,1	0,042	0,051	0,081	0,043	0,088
5	3	MVID	0,095	0,077	0,111	0,054	0,065	0,122
6	4	APTK	0,556	0,189	0,129	0,124	0,455	0,162
7	5	SVAV	0,167	0,124	0,077	0,049	0,04	0,092

# Предварительный анализ:

МАКСИМАЛЬНЫЕ ДНЕВНЫЕ ОТНОСИТЕЛЬНЫЕ СКАЧКИ ЦЕН



# Выводы по предваритель- ному анализу данных

01

Отметим, что у котировок PJPr данные с 2018 года, MDMGDR с 2020, у DSKY и BELU с 2017, у ORUP с конца 2017, а у AQUA с середины 2015, поэтому их анализировать дальше не будем. .

02

Оставшиеся 6 компании имеют более 229 торговых дней в каждом из рассматриваемых годов, что обеспечивает достаточное количество информации для дальнейшего анализа и обработки.

03

Максимальный скачок цен вверх у компании АРТК ("Аптеки 36,6") равный 0,555

$0,136 > 0,5$

Минимальный скачок цен вверх аналогично у компании АРТК ("Аптеки 36,6") равный -0,19

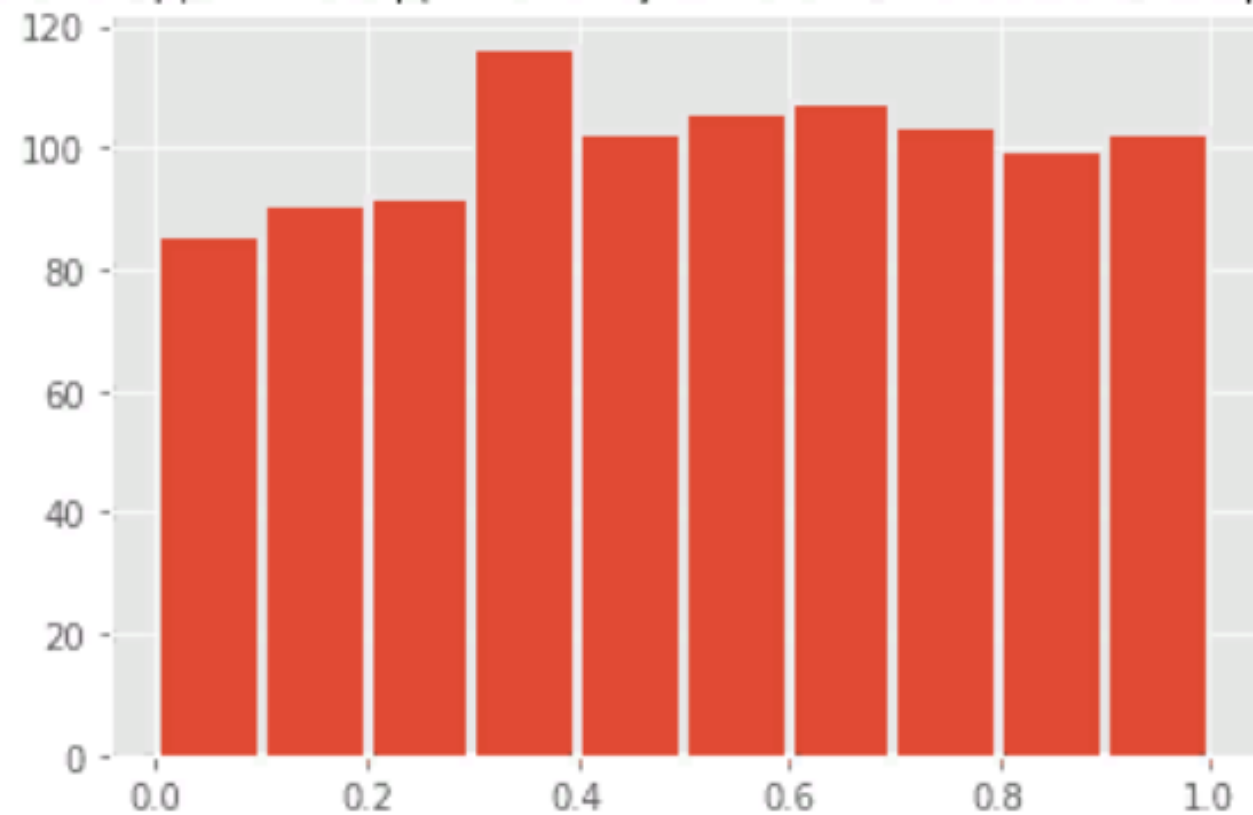
$-0,193 > -0,5$

Приходим к выводу, что компания АРТК ут  
пригодны для дальнейшей проверки  
распределения их логарифмических доходностей  
на нормальность.

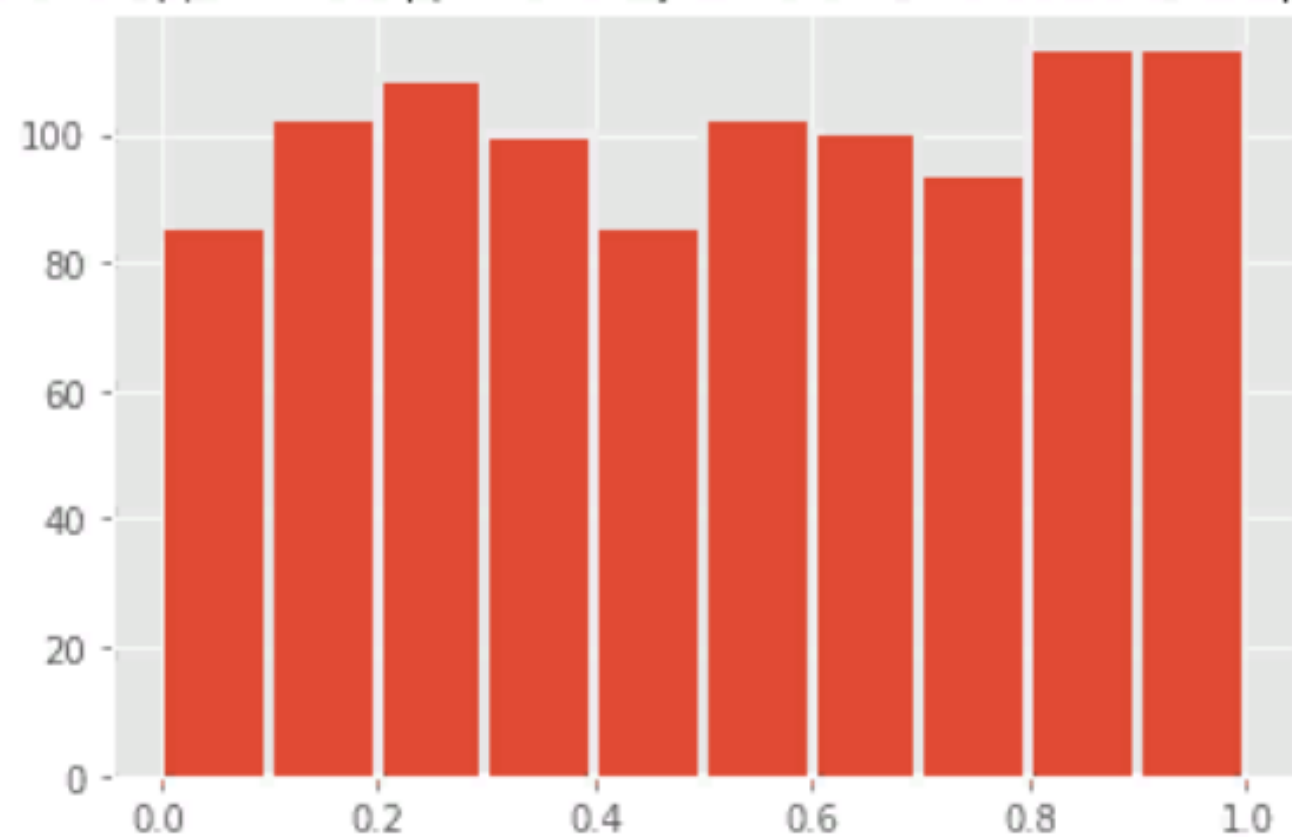


# Модельные данные

Гистограмма Р-значений модельных данных 1, вычисленных с помощью критерия Колмогорова



Гистограмма Р-значений модельных данных 2, вычисленных с помощью критерия Колмогорова



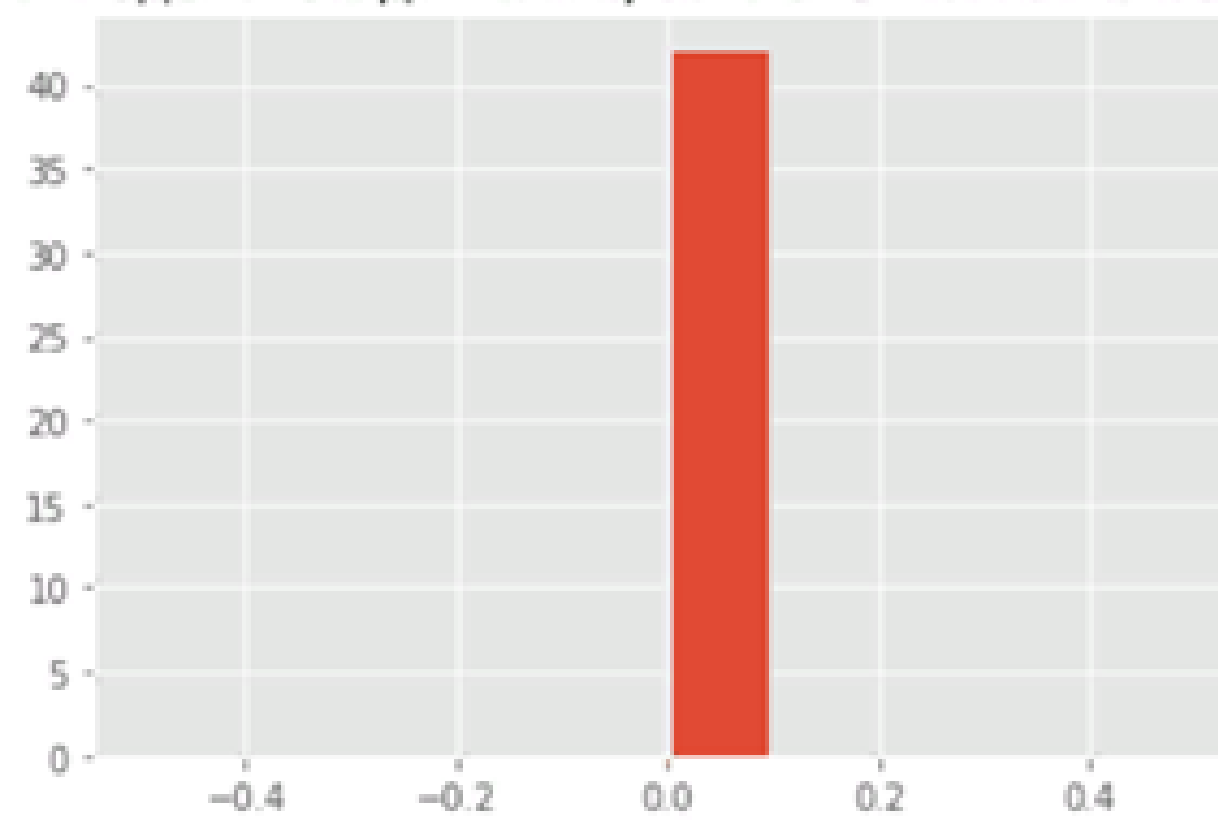
# Модельные данные

- Проверим равномерность распределения  $P$ -значений на отрезке  $[0,1]$
- Возьмем заведомо нормально распределенные случайные величины и произведем на них проверку.
- Из гистограммы видно, что  $P$ -значения распределены равномерно. Это подтверждается значением критерия Колмогорова. Следовательно, можно перейти к проверке по критерию Фишера на модельных данных.
- Проверим Критерий Фишера на модельных данных. Критерий Фишера равен 1.2487322294814869, значит, можно сделать вывод, что гипотеза принимается.

# Реальные данные

	Тикер	2015	2016	2017	2018	2019	2020
0	AGRO	0.0	0.0	0.0	0.0	0.0	0.0
1	MGNT	0.0	0.0	0.0	0.0	0.0	0.0
2	LNTAq	0.0	0.0	0.0	0.0	0.0	0.0
3	MVID	0.0	0.0	0.0	0.0	0.0	0.0
4	APTK	0.0	0.0	0.0	0.0	0.0	0.0
5	SVAV	0.0	0.0	0.0	0.0	0.0	0.0
6	MOEXCN	0.0	0.0	0.0	0.0	0.0	0.0

Гистограмма P-значений модельных данных 1, вычисленных с помощью критерия Колмогорова



# Значение дисперсии для реальных данных

	Тикер	2015	2016	2017	2018	2019	2020
0	AGRO	0.055917	0.055917	0.055917	0.055917	0.055917	0.055917
1	MGNT	0.234778	0.234778	0.234778	0.234778	0.234778	0.234778
2	LNTAq	0.136766	0.136766	0.136766	0.136766	0.136766	0.136766
3	MVID	0.111443	0.111443	0.111443	0.111443	0.111443	0.111443
4	SVAV	0.078697	0.078697	0.078697	0.078697	0.078697	0.078697
5	MOEXCN	0.008455	0.008455	0.008455	0.008455	0.008455	0.008455

# Проверка критерия Фишера

	A	B	C	
1	AGRO	43.73868939867883		
2	MGNT	771.065195439747		
3	LNTAq	261.6566468132922		
4	MVID	173.73253262154708		
5	SVAV	86.6344458183792		

# Модельные данные

- При анализе модельных данных, распределенных по нормальному закону, гипотеза равенстве дисперсий по критерию Фишера была полностью принята.
- Для реальных данных гипотеза была отвергнута.
- Выбранная гипотеза не верна для реальных данных, а значит дисперсии логарифмической доходности индекса фондового рынка и входящих в его состав акций не равны
- Критерий Фишера крайне чувствителен к отклонению от нормального распределения и достигает максимальной эффективности для выборок одинакового объёма, что крайне редко можно наблюдать в современных реалиях, а не на моделях.