

РАСЧЕТНО-АНАЛИТИЧЕСКАЯ РАБОТА

По дисциплине: Анализ данных

Москва, 2021

Оглавление

1. Загрузка исходных данных.....	4
2. Вычисление дополнительных признаков для каждой компании.....	4
3. Исследование цен и объемов торгов.....	5
4. Исследование логарифмических доходностей (логдоходностей) акций	10
4.1. Исследование до удаления выбросов.....	10
4.2. Удаление выбросов.....	13
5.2. Интервальные оценки параметров логарифмических доходностей (в предположении нормального распределения)	15
5.3. Проверка гипотезы о нормальности логдоходности	17
5.3.1. Проверка гипотезы по критерию Пирсона (хи-квадрат)	17
5.3.2. Проверка гипотезы по критерию Колмогорова-Смирнова	20
5.3.3. Сравнение результатов проверки по двум критериям	22
5.4. Построение диаграммы, содержащей гистограмму эмпирической плотности и график теоретической плотности	22
5.5. Построение графиков эмпирической функции распределения и график теоретической функции распределения.	23
6. Проверка гипотезы о равенстве нулю математического ожидания недельной логдоходности при альтернативной гипотезе о том, что оно больше нуля.	24
7. Проверка гипотезы о равенстве дисперсий логарифмической доходности за последний и предпоследний годы (двусторонний F-тест)	26
8. Проверка гипотезы о равенстве средних значений логарифмической доходности за последний и предпоследний годы (двусторонний t-тест)	27
9. Сравнительный анализ логдоходностей трех компаний.	28
9.1. Критерий Бартлетта	28

9.2. Гипотеза о значимости зависимости значений логдоходности акций компаний «ДВМП» от номера года.....	29
9.3. Гипотеза о значимости зависимости значений логдоходности акций компаний от названия компании.	30
10. Графики временных рядов логдоходностей трех компаний	31
11. Исследование тесноты связи между логдоходностями	33

1. Загрузка исходных данных

Для выполнения расчетно-аналитической работы были отобраны акции трех эмитентов: Акрон, ДЭК ао, Сбербанк.

На финансовом портале MFD.ru в разделе «Мосбиржа Акции и ПИФы» был произведен экспорт данных о котировках трех выбранных эмитентов в MS Excel. Временной интервал был выбран в соответствии с рекомендациями учебного пособия по расчетно-аналитической работе – с 1 января 2015 года до 12 февраля 2020 года. Результат данного шага представлен на рисунке 1.

<TICKER>;	<PER>;	<DATE>;	<TIME>;	<CLOSE>;	<VOL>
Акрон;	W;	05/01/15;	000000;	1856;	13070
Акрон;	W;	12/01/15;	000000;	2061;	30901
Акрон;	W;	19/01/15;	000000;	2258;	62687
Акрон;	W;	26/01/15;	000000;	2378;	52687
Акрон;	W;	02/02/15;	000000;	2306;	159429
Акрон;	W;	09/02/15;	000000;	2477;	166579
Акрон;	W;	16/02/15;	000000;	2530;	76318
Акрон;	W;	23/02/15;	000000;	2435;	19637
Акрон;	W;	02/03/15;	000000;	2512;	85522
Акрон;	W;	09/03/15;	000000;	2445;	46250
Акрон;	W;	16/03/15;	000000;	2300;	191669
Акрон;	W;	23/03/15;	000000;	2193;	56583
Акрон;	W;	30/03/15;	000000;	2325;	19192
Акрон;	W;	06/04/15;	000000;	2250;	46829
Акрон;	W;	13/04/15;	000000;	2145;	94395
Акрон;	W;	20/04/15;	000000;	2225;	158297
Акрон;	W;	27/04/15;	000000;	2262;	21948
Акрон;	W;	04/05/15;	000000;	2336;	17728
Акрон;	W;	11/05/15;	000000;	2257;	18297
Акрон;	W;	18/05/15;	000000;	2262;	12193
Акрон;	W;	25/05/15;	000000;	2216;	52886
Акрон;	W;	01/06/15;	000000;	2262;	40656

Рисунок 1. Первичные исходные данные

2. Вычисление дополнительных признаков для каждой компании

На следующем шаге были удалены столбцы периода и времени. Вместе с этим были сформированы столбцы цен закрытия и объемов торгов таким образом, что для каждого тиккера и каждого параметра был образован персональный столбец. Далее на основе данных о ценах и объёмах были рассчитаны показатели недельной доходности, логарифмической доходности, логарифма цены и логарифма объёма. Данные показатели были рассчитаны

для всех пяти акций. Получилась довольно объёмная таблица, более подробно с которой можно ознакомиться в файле Excel на листе «Данные с выбросами».

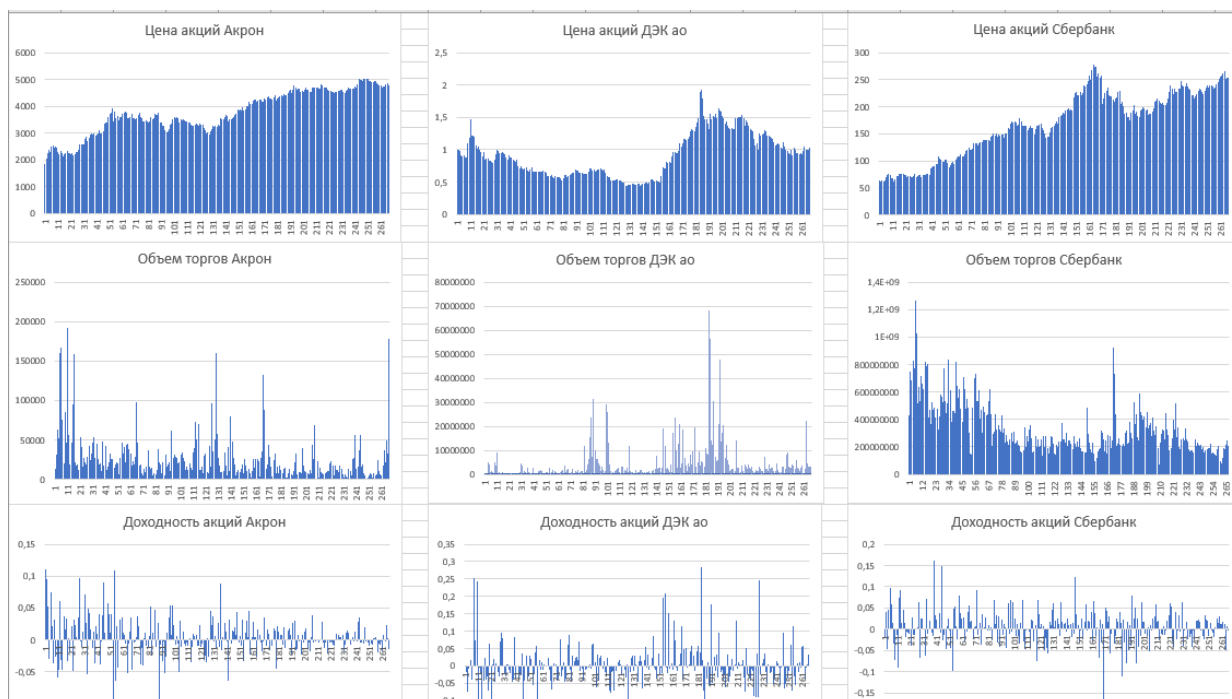
Акрон						ДЭК ао					
Цена	Объем	Лог.цен	Лог.объема	Доходности	Логдоходности	Цена	Объем	Лог.цен	Лог.объема	Доходности	Логдоходности
1856	13070	7,526179	9,47807481			0,999	2000	-0,001	7,60090246		
2061	30901	7,630947	10,3385438	0,11045259	0,104767668	0,997	624000	-0,003	13,34390565	-0,002002	-0,002004009
2258	62687	7,722235	11,0459094	0,09558467	0,091288164	0,98	266000	-0,0202	12,49125159	-0,01705115	-0,017198198
2378	52687	7,774015	10,872124	0,05314438	0,051780333	0,908	5221000	-0,09651	15,46819951	-0,07346939	-0,076308193
2306	159429	7,74327	11,979354	-0,03027754	-0,030745376	0,9	4321000	-0,10536	15,27899741	-0,00881057	-0,008849615
2477	166579	7,814803	12,023225	0,07415438	0,071533729	0,914	793000	-0,08992	13,5835785	0,015555556	0,015435808
2530	76318	7,835975	11,2426641	0,02139685	0,021171152	0,879	1275000	-0,12897	14,05845674	-0,03829322	-0,039045674
2435	19637	7,797702	9,88517082	-0,03754941	-0,038272546	0,879	902000	-0,12897	13,7123698	0	0
2512	85522	7,828835	11,3565289	0,03162218	0,031132492	1,1	5274000	0,09531	15,47829965	0,251422071	0,224280561
2445	46250	7,8018	10,7418167	-0,02667197	-0,027034126	1,181	3804000	0,166362	15,1515637	0,073636364	0,071051357
2300	191669	7,740664	12,1635252	-0,0593047	-0,061136	1,468	9283000	0,383901	16,04369533	0,243014395	0,217539393
2193	56583	7,693026	10,9434639	-0,04652174	-0,047638653	1,226	454000	0,203757	13,02585248	-0,16485014	-0,180144093
2325	19192	7,751475	9,8622488	0,06019152	0,05844957	1,22	859000	0,198851	13,6635242	-0,00489396	-0,004905979
2250	46829	7,718685	10,7542579	-0,03225806	-0,032789823	1,194	263000	0,177309	12,47990931	-0,02131148	-0,021541844
2145	94395	7,670895	11,4552434	-0,04666667	-0,047790664	1,066	728000	0,063913	13,49805633	-0,10720268	-0,113395689
2225	158297	7,707512	11,9722283	0,03729604	0,036617363	1,019	934000	0,018822	13,74723172	-0,04409006	-0,045091572
2262	21948	7,724005	9,9964313	0,01662921	0,016492462	1,042	116000	0,041142	11,66134547	0,022571148	0,022320189
2336	17728	7,756195	9,78290059	0,03271441	0,032190687	1	81000	0	11,30220443	-0,0403071	-0,041141943
2257	18297	7,721792	9,81449239	-0,03381849	-0,034403567	0,97	216000	-0,03046	12,28303369	-0,03	-0,030459207

Рисунок 2. Сформированный лист исходных данных

3. Исследование цен и объемов торгов

Для получения первоначального представления о поведении исследуемых величин используем средства визуализации, в данном случае воспользуемся сервисом MS Excel под названием «Вставка» для построения графиков.

Первый этап – построение гистограмм для исследуемых величин для каждой компании. Все построенные гистограммы можно найти на листе «Гистограммы» в прилагаемом файле.



Поскольку цены акций выбранных компаний отличаются на несколько порядков, их изображение на одном графике нецелесообразно, поэтому воспользуемся инструментом относительных цен, которые определяются по формуле:

$$t = \frac{x - x_{min}}{x_{max} - x_{min}}$$

где

x_{min} и x_{max} - соответственно минимальное и максимальное значения цены.

Рассчитанные данные об относительных ценах можно найти на листе Excel под названием «Изменение цен+опис.стат.». Далее по ним был построен линейный график, иллюстрирующий динамику относительных цен компаний, который изображен на рисунке 3.



Рисунок 3. Динамика относительных цен

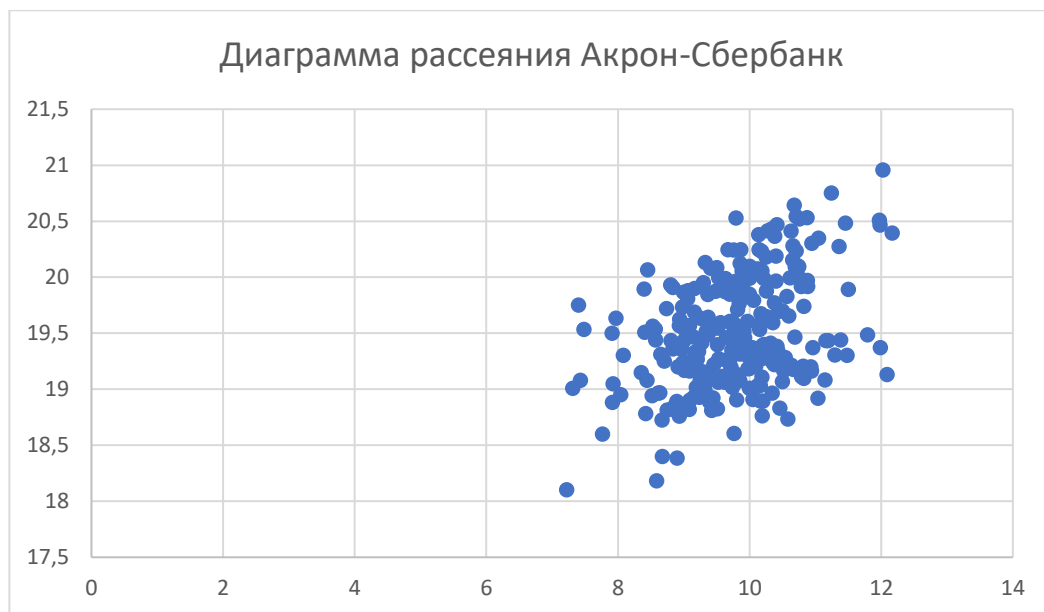
На нем мы можем наблюдать рост акций «Акрон» за весь период, резкий рост акций «Сбербанка» с конца 2017 года, а также стремительный рост акций «ДЭК ао», после – спад.

Также была сформирована описательная статистика и корреляционная матрица для цен трех компаний (рисунок ниже).

	Акрон	ДЭК ао	Сбербанк		
Акрон	1				
ДЭК ао	0,490246	1			
Сбербанк	0,853323	0,300597509	1		
Акрон		ДЭК ао		Сбербанк	
Среднее	3774,408	Среднее	0,917363	Среднее	168,4008
Стандартная ошибка	49,00238	Стандартная ошибка	0,020652	Стандартная оши	3,696917
Медиана	3675	Медиана	0,888	Медиана	173,25
Мода	3600	Мода	0,95	Мода	70,6
Стандартное отклонен	800,7055	Стандартное отклонен	0,33746	Стандартное откл	60,40812
Дисперсия выборки	641129,4	Дисперсия выборки	0,113879	Дисперсия выбор	3649,142
Эксцесс	-0,83344	Эксцесс	-0,53364	Эксцесс	-1,12395
Асимметричность	-0,31149	Асимметричность	0,56267	Асимметричность	-0,26435
Интервал	3184	Интервал	1,495	Интервал	216,49
Минимум	1856	Минимум	0,44	Минимум	61
Максимум	5040	Максимум	1,935	Максимум	277,49
Сумма	1007767	Сумма	244,936	Сумма	44963,01
Счет	267	Счет	267	Счет	267

Далее для изучения корреляции между объема торгов акций трех компаний были построены графики рассеяния и сформирована корреляционная матрица с помощью сервиса «Анализ данных». Данные графики можно найти, обратившись к листу «Логарифмы объемов+опис.стат.» в расчетном файле Excel.





Рисунки 4. Диаграмма рассеяния значений признаков «логарифмов объемов торгов» акций.

	<i>Акрон</i>	<i>ДЭК ао</i>	<i>Сбербанк</i>
<i>Акрон</i>	1		
<i>ДЭК ао</i>	-0,12865	1	
<i>Сбербанк</i>	0,401577	-0,21668	1

Рисунок 5. Корреляционная матрица значений признаков «логарифмов объемов торгов»

Проанализировав три графика рассеивания логарифмов объемов и обратившись к корреляционной матрице, можно сказать, что какая-либо связь между логарифмами объемов торгов компаний практически отсутствует. Также была сформирована описательная статистика для логарифмов объемов.

Описательная статистика признаков «логарифмов объемов торгов» акций компаний					
Акрон		ДЭК ао		Сбербанк	
Среднее	9,778740142	Среднее	14,53121039	Среднее	19,50056181
Стандартная оши	0,055161974	Стандартная ошибк	0,083819037	Стандартная ошиб	0,030550469
Медиана	9,803667217	Медиана	14,60850307	Медиана	19,42347275
Мода	9,857077055	Мода	12,84792653	Мода	#Н/Д
Стандартное откл	0,901354083	Стандартное отклон	1,369614354	Стандартное откл	0,499198774
Дисперсия выбор	0,812439183	Дисперсия выборки	1,875843477	Дисперсия выборк	0,249199416
Эксцесс	0,366714372	Эксцесс	2,826242349	Эксцесс	-0,21874732
Асимметричность	-0,077169304	Асимметричность	-0,65576059	Асимметричность	0,294555054
Интервал	4,946815718	Интервал	10,43433637	Интервал	2,856146683
Минимум	7,216709487	Минимум	7,60090246	Минимум	18,10367353
Максимум	12,1635252	Максимум	18,03523883	Максимум	20,95982021
Сумма	2610,923618	Сумма	3879,833174	Сумма	5206,650004
Счет	267	Счет	267	Счет	267

4. Исследование логарифмических доходностей (логдоходностей) акций

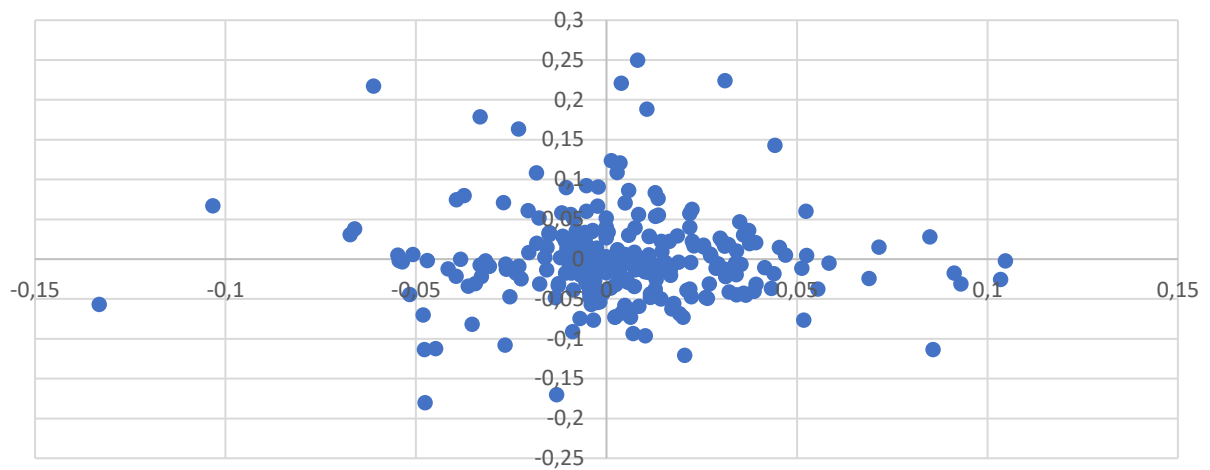
4.1. Исследование до удаления выбросов

Сгенерируем корреляционную матрицу логарифмических доходностей компаний до удаления выбросов с помощью сервиса «Анализ данных». Полученные расчеты можно увидеть, обратившись к листу «Логарифм доходности+опис.стат.» в расчетном документе Excel.

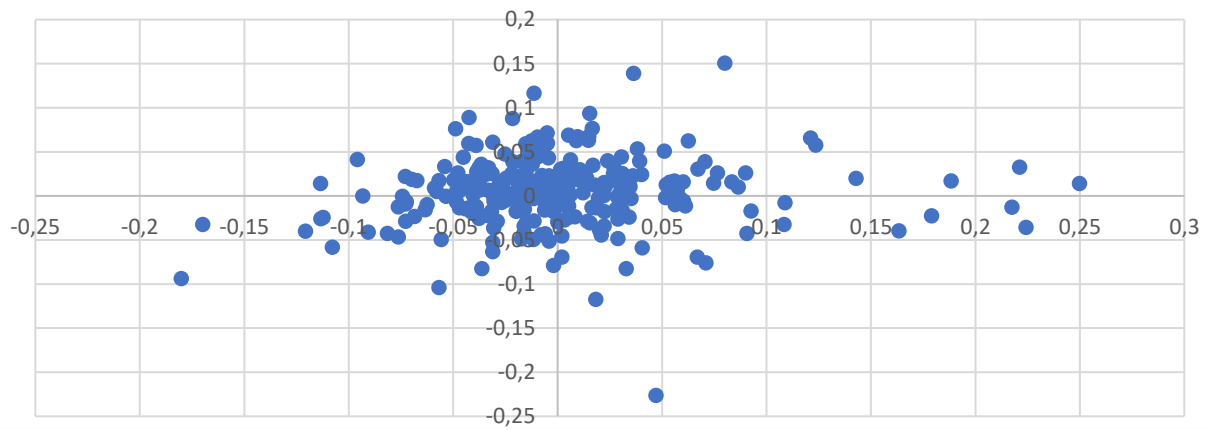
	Акрон	ДЭК ао	Сбербанк
Акрон	1		
ДЭК ао	-0,03451234	1	
Сбербанк	0,116651727	0,085018065	1

Рисунок 8. Корреляционная матрица значений признаков
«логарифмические доходности»

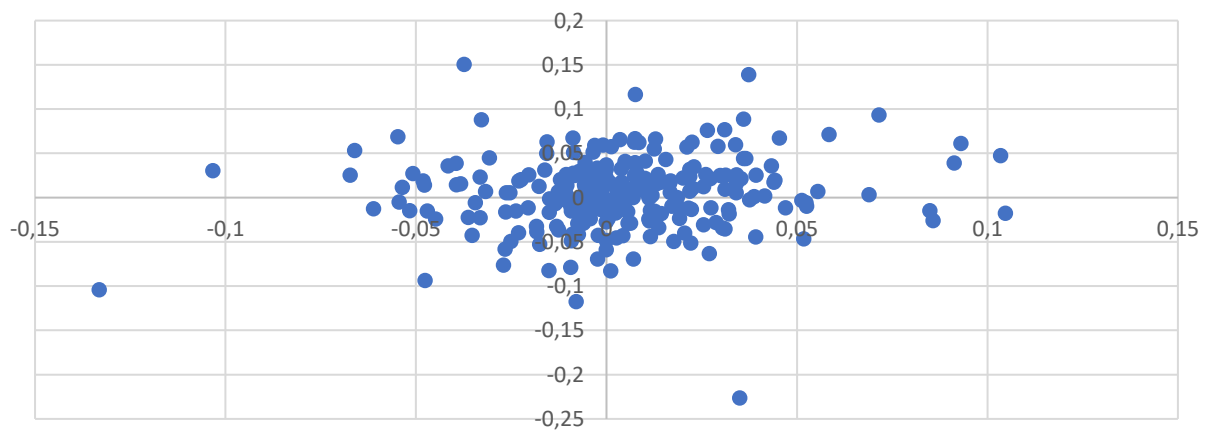
Акрон-ДЭК ао

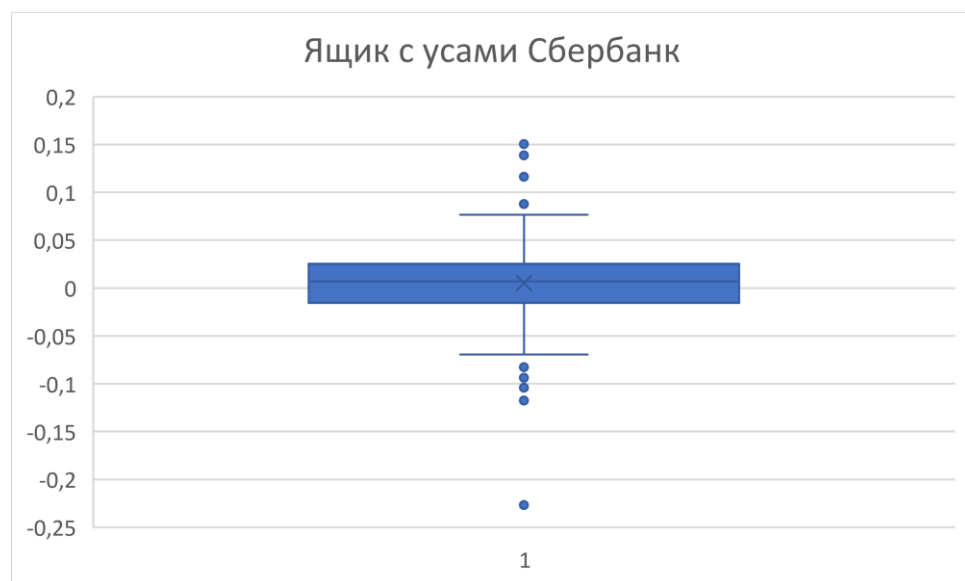
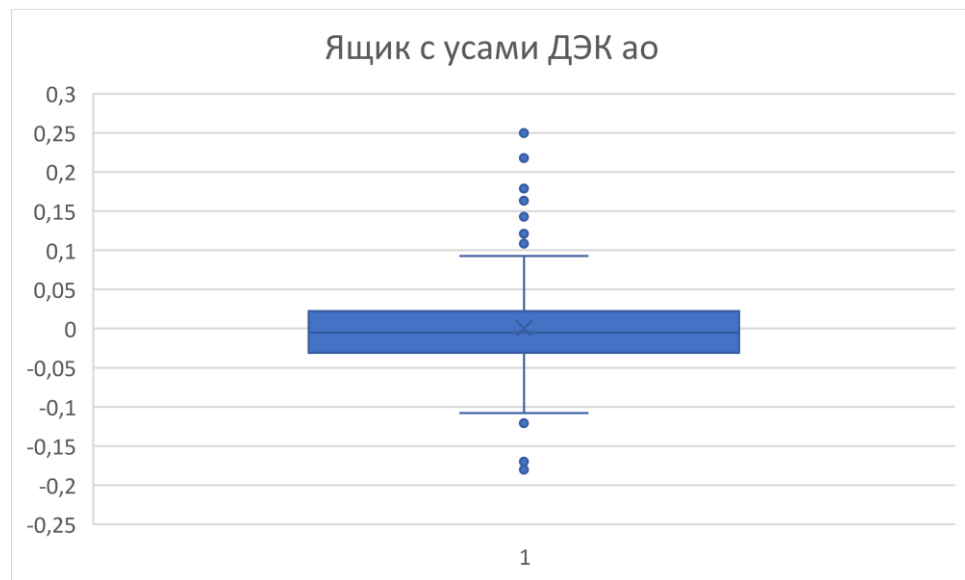
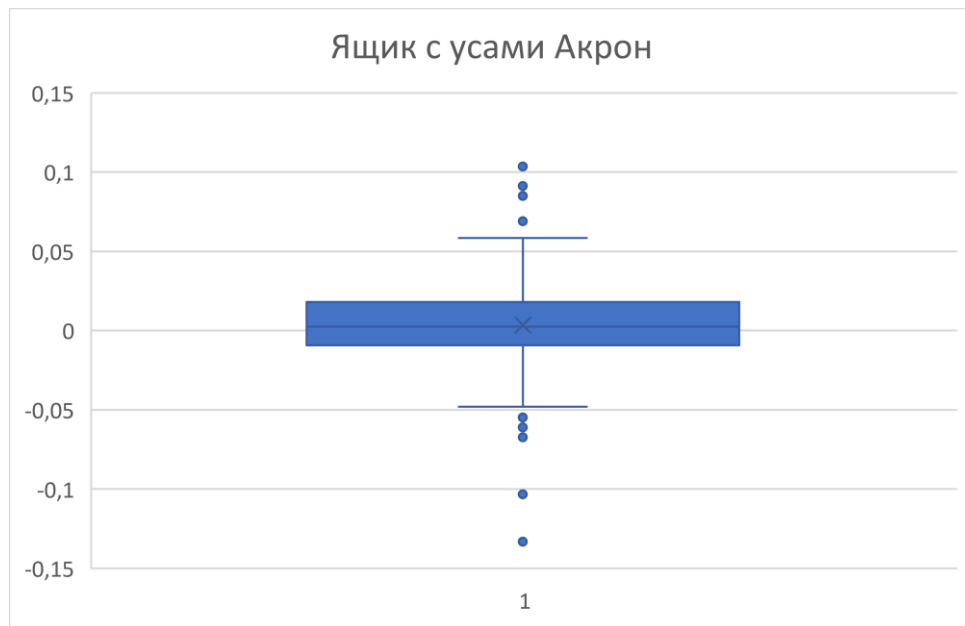


ДЭК ао-Сбербанк



Акрон-Сбербанк





<i>Акрон</i>		<i>ДЭК ао</i>		<i>Сбербанк</i>	
Среднее	0,003556455	Среднее	0,000125818	Среднее	0,005234
Стандартная оши	0,00181198	Стандартная ои	0,003486342	Стандартная оши	0,002458
Медиана	0,002394601	Медиана	-0,00523077	Медиана	0,00709
Мода	0	Мода	0	Мода	#Н/Д
Стандартное откл	0,029552499	Стандартное о	0,056860512	Стандартное откл	0,040096
Дисперсия выбор	0,00087335	Дисперсия выб	0,003233118	Дисперсия выбор	0,001608
Эксцесс	3,091485711	Эксцесс	4,11157857	Эксцесс	4,775319
Асимметричность	-0,05201731	Асимметрично	1,109944681	Асимметричность	-0,58237
Интервал	0,237935622	Интервал	0,429955891	Интервал	0,376907
Минимум	-0,13316795	Минимум	-0,18014409	Минимум	-0,2266
Максимум	0,104767668	Максимум	0,249811798	Максимум	0,150311
Сумма	0,946016912	Сумма	0,03346769	Сумма	1,39218
Счет	266	Счет	266	Счет	266

На рисунках выше представлены диаграммы рассеяния и диаграммы «Ящик с усами» для логарифмических доходностей компаний до удаления выбросов, а также описательная статистика.

Как мы можем видеть, корреляция между логдоходностями выбранных компаний почти отсутствует, либо она очень слабая, что может быть обусловлено разными финансовыми показателями, имеющимися у компаний за выбранный период.

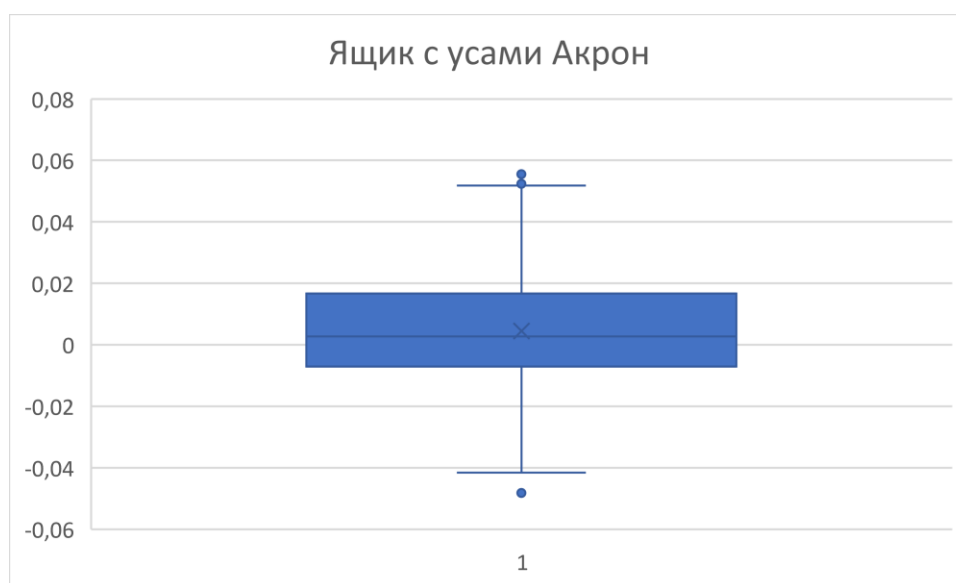
4.2. Удаление выбросов

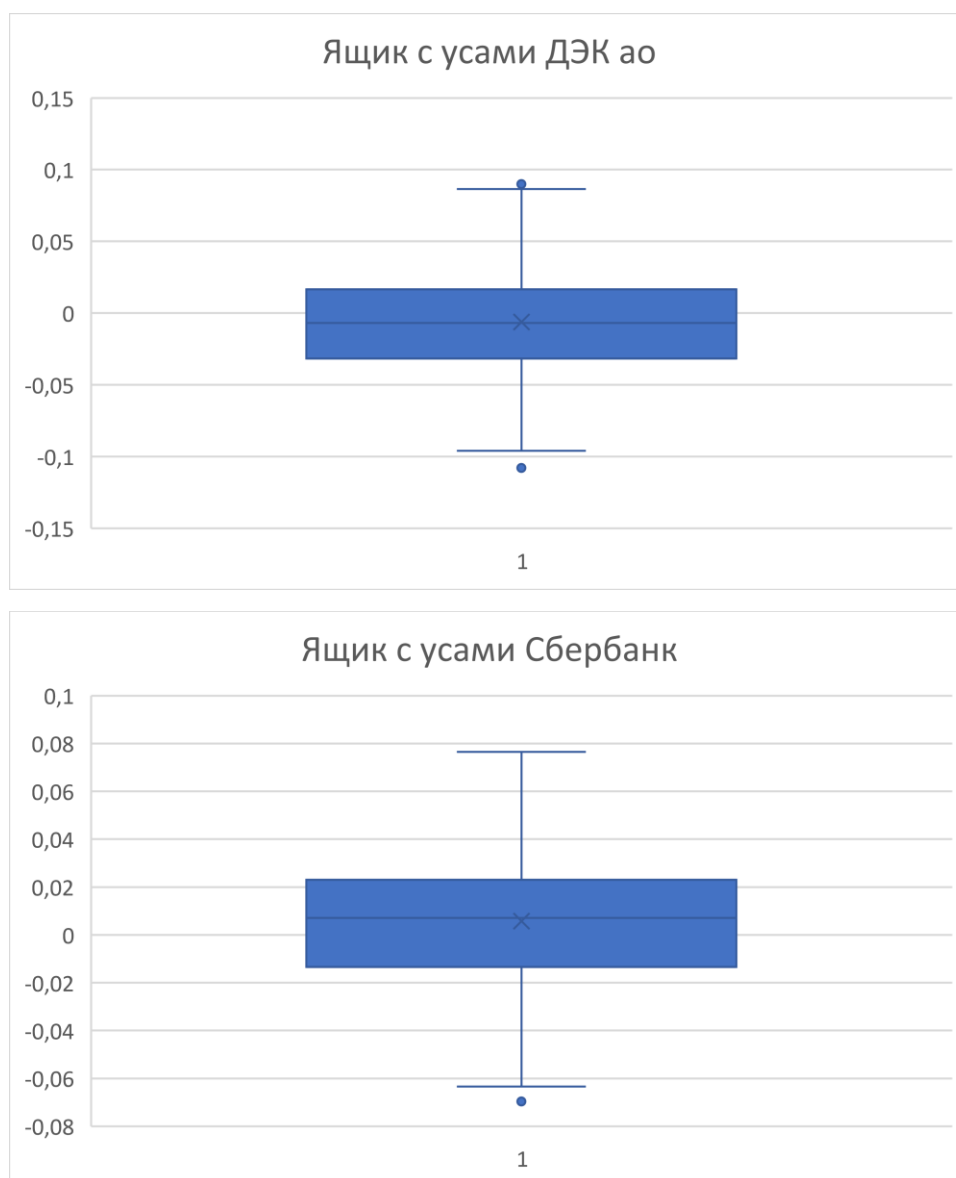
Следующим шагом было удаление выбросов из выборки. Были рассчитаны нижние и верхние границы для логдоходностей каждой компании.

	Акрон	ДЭК ао	Сбербанк
1 квартиль	-0,00696	-0,03145	-0,01336
3 квартиль	0,016551	0,016467	0,022964
Кв. размах	0,023506	0,047912	0,036322
Нижняя граница	-0,04221	-0,10331	-0,06784
Верхняя граница	0,05181	0,088335	0,077447

Рисунок 10. Расчет границ для удаления выбросов

Далее с помощью сортировки были найдены выбросы и удалены из выборки. По данным без выбросов были построены требуемые диаграммы, а также сформированы таблицы с помощью сервиса «Описательная статистика». Все расчеты можно найти, обратившись к Excel.





5.2. Интервальные оценки параметров логарифмических доходностей (в предположении нормального распределения)

Для расчета доверительного интервала для оценки математического ожидания, дисперсии и квадратичного отклонения был использован сервис «Описательная статистика» для каждой компании, а также статистические формулы `ДОВЕРИТ.СТЮДЕНТ` и `ХИ2.ОБР`. Результаты расчетов представлены ниже, их также можно найти на листе «Доверительные интервалы» в рабочем файле Excel.

Акрон		ДЭК ао		Сбербанк	
Среднее	0,004518	Среднее	-0,00618	Среднее	0,005687
Стандартная ошибка	0,001358	Стандартная ошибка	0,002582	Стандартная ошибка	0,002009
Медиана	0,002775	Медиана	-0,0068	Медиана	0,00709
Мода	0	Мода	0	Мода	#Н/Д
Стандартное отклонение	0,020142	Стандартное отклонение	0,038299	Стандартное отклонение	0,029801
Дисперсия выборки	0,000406	Дисперсия выборки	0,001467	Дисперсия выборки	0,000888
Эксцесс	0,129933	Эксцесс	0,109671	Эксцесс	-0,0755
Асимметричность	0,10387	Асимметричность	0,154933	Асимметричность	-0,09594
Интервал	0,103625	Интервал	0,200446	Интервал	0,146045
Минимум	-0,04812	Минимум	-0,10781	Минимум	-0,06955
Максимум	0,055506	Максимум	0,092633	Максимум	0,076492
Сумма	0,993971	Сумма	-1,35923	Сумма	1,251131
Счет	220	Счет	220	Счет	220

Рисунок 12. Результат процедуры «Описательная статистика».

Инт. оц. мат. ожидания	Акрон	ДЭК ао	Сбербанк
Среднее =	0,004518049	-0,006178306	0,00568696
Отклонение =	0,020141555	0,038298911	0,02980111
Альфа =	0,05	0,05	0,05
Размер выборки =	220	220	220
Эпсилон =	0,00267631	0,005088969	0,00395982
Нижняя граница =	0,001841739	-0,011267275	0,00172713
Верхняя граница =	0,007194359	-0,001089336	0,00964678

Рисунок 13. Интервальное оценивание математического ожидания.

Инт. оц. дисперсии	Акрон	ДЭК ао	Сбербанк
Среднее =	0,004518049	-0,006178306	0,00568696
Отклонение =	0,020141555	0,038298911	0,02980111
Альфа =	0,05	0,05	0,05
Размер выборки =	220	220	220
Хи^2 левая =	179,9068569	179,9068569	179,906857
Хи^2 правая =	261,8792591	261,8792591	261,879259
(n-1)*s^2 =	0,088844411	0,321230638	0,19449519
Нижняя граница =	0,000339257	0,001226636	0,00074269
Верхняя граница =	0,000493836	0,001785539	0,00108109

Рисунок 14. Интервальное оценивание дисперсии.

Инт. оц. квадрат. откл.	Акрон	ДЭК ао	Сбербанк
Среднее =	0,004518049	-0,006178306	0,00568696
Отклонение =	0,020141555	0,038298911	0,02980111
Альфа =	0,05	0,05	0,05
Размер выборки =	220	220	220
Хи^2 левая =	179,9068569	179,9068569	179,906857
Хи^2 правая =	261,8792591	261,8792591	261,879259
(n-1) =	219	219	219
Нижняя граница =	0,018418935	0,03502337	0,02725235
Верхняя граница =	0,022222412	0,042255634	0,03287991

Рисунок 15. Интервальная оценка квадратичного отклонения.

Данный анализ показывает, что верхняя граница логарифмической доходности компании «Сбербанк» выше, чем у остальных компаний. Также он показывает, что нижняя граница для «ДЭК ао» является отрицательной величиной, что свидетельствует о возможной вероятности убытков у данной компании.

5.3. Проверка гипотезы о нормальности логдоходности

5.3.1. Проверка гипотезы по критерию Пирсона (хи-квадрат)

Используем критерий согласия Пирсона (χ^2) для проверки гипотезы о соответствии эмпирического распределения данных о котировках акций предполагаемому теоретическому нормальному распределению $F(x)$ при объеме выборки $n > 100$.

H_0 : данные соответствуют нормальному распределению (логдоходность)

H_1 : данные не соответствуют нормальному распределению (логдоходность)

С помощью сервиса «Описательная статистика» и правила Стерджеса были рассчитаны количество интервалов и их длина. Далее были рассчитаны

границы сформированных интервалов, эмпирические и теоретические интервальные частоты. В итоге были получены значения критерия для каждой отдельной компании и произведено сравнение с критическими значениями. Все расчеты представлены ниже (с ними также можно ознакомиться на листе «критерий Пирсона» в расчетном документе Excel).

	Акрон	ДЭК ао	Сбербанк
Среднее	0,004518049	-0,006178306	0,005686957
Стандартная ошибка	0,001357943	0,002582112	0,00200919
Медиана	0,002774577	-0,006796739	0,007090111
Мода	0	0	#Н/Д
Стандартное отклонение	0,020141555	0,038298911	0,029801106
Дисперсия выборки	0,000405682	0,001466807	0,000888106
Эксцесс	0,129932684	0,109671076	-0,075504743
Асимметричность	0,103869652	0,154932833	-0,095938966
Интервал	0,103625486	0,200445665	0,146044638
Минимум	-0,048119248	-0,107812863	-0,069552283
Максимум	0,055506237	0,092632802	0,076492355
Сумма	0,993970757	-1,359227232	1,251130649
Счет	220	220	220
Правило Стерджеса	Акрон	ДЭК ао	Сбербанк
Число интервалов (l) =	8	8	8
Длина интервала (h) =	0,012953186	0,025055708	0,01825558

Рисунок 16. Расчет количества и длины интервалов

Интервалы	Акрон	ДЭК ао	Сбербанк
Первый интервал =	-0,035166063	-0,08275716	-0,0512967
Второй интервал =	-0,022212877	-0,05770145	-0,03304112
Третий интервал =	-0,009259691	-0,03264574	-0,01478554
Четвертый интервал =	0,003693495	-0,00759003	0,003470036
Пятый интервал =	0,01664668	0,017465678	0,021725616
Шестой интервал =	0,029599866	0,042521386	0,039981196
Седьмой интервал =	0,042553052	0,067577094	0,058236775
Восьмой интервал =	0,055506237	0,092632802	0,076492355
Эмп. инт. частоты	Акрон	ДЭК ао	Сбербанк
Первая частота =	8	4	7
	212	216	213
Вторая частота =	20	19	24
	200	201	196
Третья частота =	44	54	51
	176	166	169
Четвертая частота =	114	108	96
	106	112	124
Пятая частота =	165	167	161
	55	53	59
Шестая частота	192	198	195
	28	22	25
Седьмая частота =	211	212	207
	9	8	13
Восьмая частота =	220	220	219
	0	0	1

Рисунок 17. Расчет границ интервалов и эмпирических интервальных частот.

Эмп. инт. Частоты (ni)	Акрон	ДЭК ао	Сбербанк		Вероятность (pi)	Акрон	ДЭК ао	Сбербанк
Первая частота =	8	4	7		Первая =	0,024404	0,022777	0,027929605
Вторая частота =	20	19	24		Вторая =	0,092229	0,089266	0,096877311
Третья частота =	44	54	51		Третья =	0,246973	0,244759	0,24605044
Четвертая частота =	114	108	96		Четвертая =	0,483673	0,485298	0,470349804
Пятая частота =	165	167	161		Пятая =	0,726469	0,731499	0,704777075
Шестая частота =	192	198	195		Шестая =	0,893485	0,898237	0,875086703
Седьмая частота =	211	212	207		Седьмая =	0,970513	0,972934	0,96107937
Восьмая частота =	220	220	219		Восьмая =	0,994321	0,99506	0,991247654
Отд. теор. инт. частота (npi)	Акрон	ДЭК ао	Сбербанк		Отд. эмп. инт. Частоты (ni)	Акрон	ДЭК ао	Сбербанк
Первая частота =	5,368936862	5,010915899	6,144512994		Первая частота =	8	4	7
Вторая частота =	14,92145061	14,62756052	15,16849551		Вторая частота =	12	15	17
Третья частота =	34,04372056	34,20852437	32,81808828		Третья частота =	24	35	27
Четвертая частота =	52,07388022	52,91856659	49,34586014		Четвертая частота =	70	54	45
Пятая частота =	53,41527131	54,16428945	51,57399949		Пятая частота =	51	59	65
Шестая частота =	36,74344975	36,68223378	37,46811816		Шестая частота =	27	31	34
Седьмая частота =	16,94611379	16,43345007	18,91838685		Седьмая частота =	19	14	12
Восьмая частота =	5,2378149	4,867649093	6,637022394		Восьмая частота =	9	8	12
Значения	Акрон	ДЭК ао	Сбербанк		Отдельные значения	Акрон	ДЭК ао	Сбербанк
Первое =	1,28936015	0,203944942	0,119107571		Первое =	1,28936	0,203945	0,119107571
Второе =	0,004155903	0,020757829	0,338756648		Второе =	0,571987	0,009483	0,221143139
Третье =	1,965501833	0,000434727	0,181111554		Третье =	2,96314	0,018312	1,03144799
Четвертое =	0,541675894	0,014272615	0,540264098		Четвертое =	6,170959	0,0221	0,382737282
Пятое =	0,167676729	0,228960857	0,228254765		Пятое =	0,109211	0,431725	3,495123348
Шестое =	0,106095453	0,00076146	0,031970811		Шестое =	2,583721	0,880202	0,321015416
Седьмое =	0,029573306	0,019548348	0,093129494		Седьмое =	0,248933	0,360343	2,53002949
Восьмое =	0,007135547	0,005395547	0,003927925		Восьмое =	2,702279	2,01568	4,333498834
Значение критерия =	4,111174816	0,494076327	1,536522866		Значение критерия =	16,63959	3,94179	12,43410307

Рисунок 18. Расчет значений критерия согласия Пирсона для каждой компании

Критическое значение	Акрон	ДЭК ао	Сбербанк
Значение критерия (0.05) =	11,07049769	11,07049769	11,07049769
Гипотеза =	Отклоняется	Принимается	Отклоняется
Значение критерия (0.01) =	15,08627247	15,08627247	15,08627247
Гипотеза =	Отклоняется	Принимается	Принимается

Рисунок 19. Расчет критических значений критерия согласия Пирсона для каждой компании

Таким образом, мы получили, что для «ДЭК ао» на любом уровне принимается гипотеза H_0 о нормальном распределении. Для «Акрона» на любом уровне гипотезу H_0 о нормальном распределении отклоняем. Для «Сбербанка» H_0 принимается на уровне значимости равном 0.01, но отклоняется на уровне значимости 0.05.

5.3.2. Проверка гипотезы по критерию Колмогорова-Смирнова

Также в работе использовался критерий Колмогорова-Смирнова, который наряду с критерием Пирсона применяется для проверки гипотезы о том, что генеральная совокупность подчинена заданной непрерывной функции распределения $F_0(x)$.

H0: данные соответствуют нормальному распределению (логдоходность)

H1: данные не соответствуют нормальному распределению (логдоходность)

Все расчеты представлены в расчетном документе Excel на листах «Критерий Колмогорова» для каждой соответствующей компании. Ниже приведем пример расчета значений критерия для компании «Сбербанк».

n =	220		Критическое значение					
Xcp =	0,005687		a =	0,15	0,1	0,05	0,025	0,01
s =	0,029801		Sn кр =	0,775	0,819	0,895	0,995	1,035
Dn+	0,057467		Гипотеза	Принимается на любом уровне				
Dn-	0,037165							
Dn =	0,037165							
знач =	0,553001							

Рисунок 20. Расчет значений критерия Колмогорова-Смирнова для компании «Сбербанк».

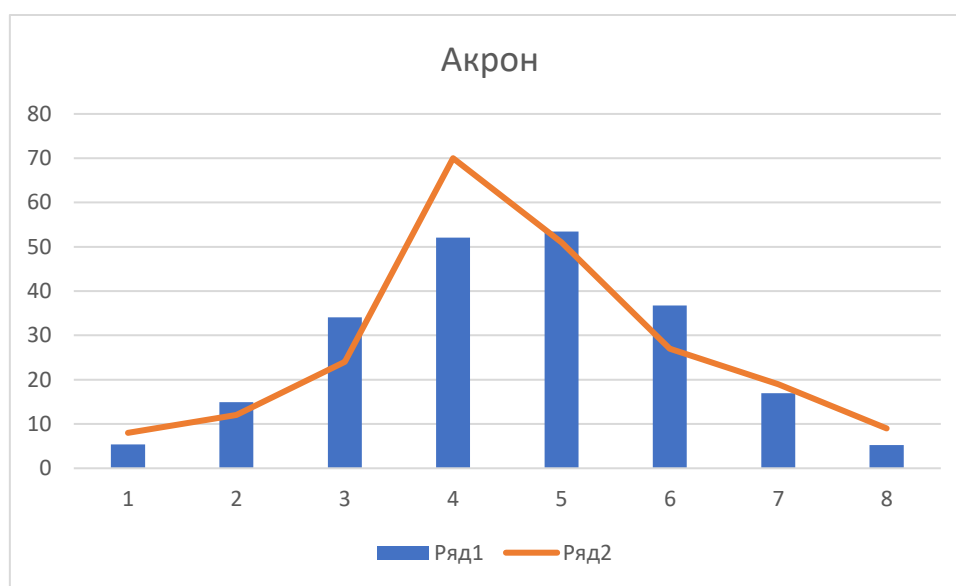
Расчет показателей для остальных компаний показал аналогичный результат. Рассчитанные значения критерия оказались ниже критических, кроме компании «Акрон» (на уровнях значимости 0.15 и 0.1 отклоняется), поэтому мы можем сделать вывод, что у нас нет оснований отвергнуть ранее выдвинутую гипотезу.

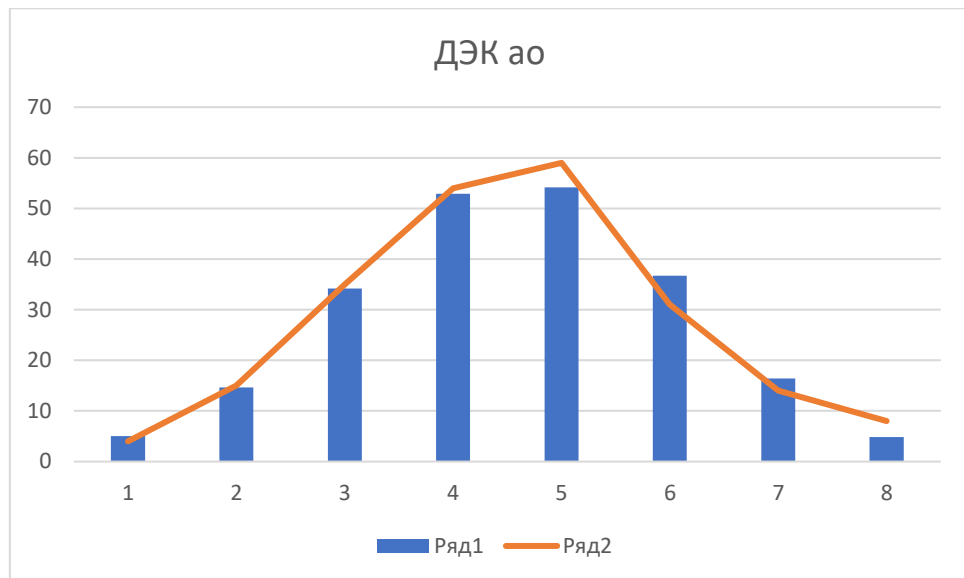
n =	220		Критическое значение					
Xcp =	0,004518		a =	0,15	0,1	0,05	0,025	0,01
s =	0,020142		Sn кр =	0,775	0,819	0,895	0,995	1,035
Dn+	0,061471		Гипотеза	Принимается на уровне 0,05 0,025 0,01				
Dn-	0,059457							
Dn =	0,059457							
знач =	0,884701							

5.3.3. Сравнение результатов проверки по двум критериям

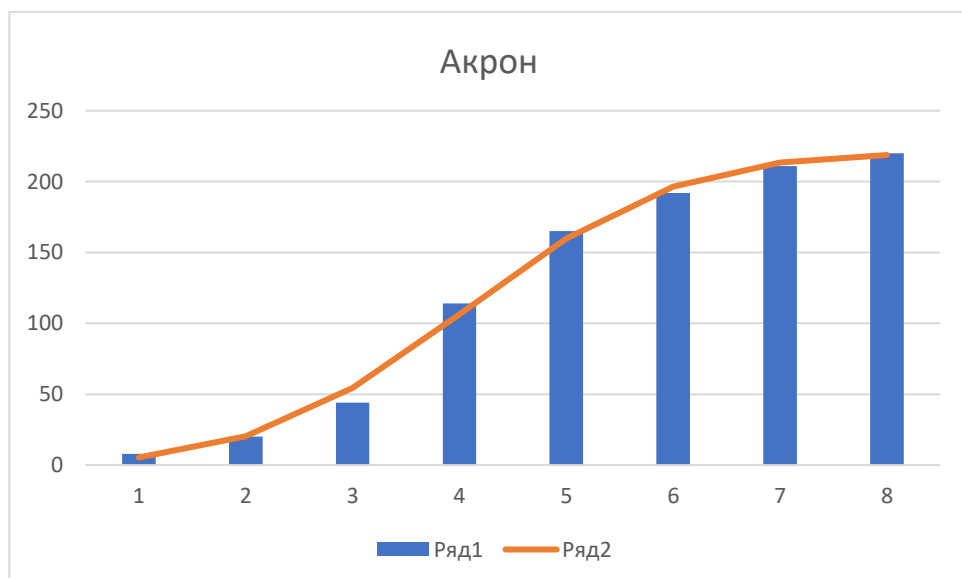
Таким образом, расчеты критерия согласия Пирсона и критерия Колмогорова-Смирнова дали одинаковые результаты для компании «ДЭК ао», что позволяет принять гипотезу о принадлежности выборки логарифмических доходностей трех компаний к нормальному распределению. Но результаты расчетов для «Акрон» и «Сбербанк» различаются, что говорит о неоднозначности, и не дает уверенности отнести распределения к нормальному.

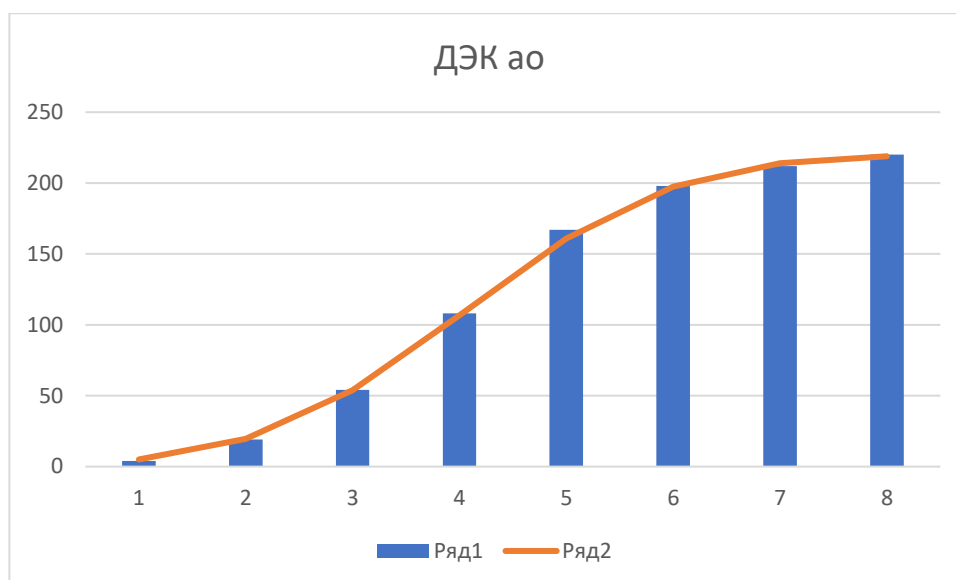
5.4. Построение диаграммы, содержащей гистограмму эмпирической плотности и график теоретической плотности





5.5. Построение графиков эмпирической функции распределения и график теоретической функции распределения.





6. Проверка гипотезы о равенстве нулю математического ожидания недельной логдоходности при альтернативной гипотезе о том, что оно больше нуля.

Далее выдвигается гипотеза о равенстве нулю математического ожидания недельной логдоходности при альтернативной гипотезе о том, что оно больше нуля для каждой выбранной компании.

H_0 : математическое ожидание равно нулю для каждой компании

H_1 : математическое ожидание больше нуля для каждой компании

В расчетном документе Excel был создан новый лист «Z-test». На этот лист скопированы данные о логдоходностях акций трех компаний, оставшиеся после удаления выбросов.

Для вычисления необходимых значений были использованы функции СРЗНАЧ и СТЬЮДЕНТ.ОБР.

Акрон			
альфа =	0,05	альфа =	0,01
Среднее =	0,004518049	Объем =	220
Откл =	0,020141555	t =	3,327126
t кр (0.05) =	1,970855367	t кр (0.01) =	2,598465
Zтест =	0,000438733	Отклоняем 0,1	
ДЭК ао			
альфа =	0,05	альфа =	0,01
Среднее =	-0,006178306	Объем =	220
Откл =	0,038298911	t =	2,392733
t кр (0.05) =	1,970855367	t кр (0.01) =	2,598465
Zтест =	0,991638305	Принимаем	
Сбербанк			
альфа =	0,05	альфа =	0,01
Среднее =	0,005686957	Объем =	220
Откл =	0,029801106	t =	2,830473
t кр (0.05) =	1,970855367	t кр (0.01) =	2,598465
Zтест =	0,002323965	Отклоняем 0,1	

Поскольку вычисленные значения критерия T для «ДЭК ао» оказались менее критических значений, как при $\alpha = 0,01$, так и при $\alpha = 0,05$ (см. Рисунок 19), то нет оснований отвергнуть гипотезу H_0 о равенстве нулю математического ожидания недельной логдоходности данного актива. Для компаний «Акрон» и «Сбербанк» мы вынуждены отклонить гипотезу H_0 на уровне значимости 0.01.

Также была использована статистическая функция Z.ТЕСТ. Рассчитанные с помощью нее значения оказались ниже уровней значимости, равных 0,05 и 0,01, для компаний «Акрон» и «Сбербанк». Для компании «ДЭК ао» проводимый тест показал результат выше уровней значимости.

Таким образом, можно сделать вывод, что нет оснований отвергнуть гипотезу H_0 о равенстве нулю математического ожидания недельной доходности активов только для компании «ДЭК ао».

7. Проверка гипотезы о равенстве дисперсий логарифмической доходности за последний и предпоследний годы (двусторонний F-тест)

Выдвигается гипотеза о равенстве дисперсий рассматриваемых генеральных совокупностей.

H_0 : дисперсии рассматриваемых генеральных совокупностей равны

H_1 : дисперсии рассматриваемых генеральных совокупностей не равны

Обратимся к статистической функции F.ТЕСТ, которая возвращает вероятность сходства дисперсий тестируемых выборок. Рассчитанные показатели оказались больше уровней значимости, кроме компании «Акрон». Это позволяет сделать вывод, что нет оснований отвергнуть гипотезу о равенстве дисперсий логарифмической доходности за 2018 и 2019 годы для «ДЭК ао» и «Сбербанка», но есть основания отвергнуть эту гипотезу для «Акрона».

Акрон			
альфа =	0,05	альфа =	0,01
альфа/2 =	0,025	альфа/2 =	0,005
F-тест =	0,003628031	Отклоняем 0,01	
ДЭК ао			
альфа =	0,05	альфа =	0,01
альфа/2 =	0,025	альфа/2 =	0,005
F-тест =	0,818489442	Принимаем	
Сбербанк			
альфа =	0,05	альфа =	0,01
альфа/2 =	0,025	альфа/2 =	0,005
F-тест =	0,064400915	Принимаем	

8. Проверка гипотезы о равенстве средних значений логарифмической доходности за последний и предпоследний годы (двусторонний t-тест)

Для решения данной применим критерий для проверки гипотезы о сравнении двух математических ожиданий.

Н₀: средние значения логдоходности за последний и предпоследний годы равны

Н₁: средние значения логдоходности за последний и предпоследний годы не равны

Рассмотрим решение указанной задачи на примере логдоходности акций выбранных компаний.

С помощью статистической формулы TТЕСТ рассчитаем показатель Т для каждой из компаний. Затем рассчитаем количество степеней свободы k с помощью следующей формулы:

$$k = \left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y} \right)^2 : \left(\left(\frac{s_x^2}{n_x} \right)^2 / (n_x - 1) + \left(\frac{s_y^2}{n_y} \right)^2 / (n_y - 1) \right)$$

Полученное значение используем в расчете критических значений критерия t с помощью функции СТЬЮДЕНТ.ОБР.

Акрон			
t =	0,218325153	n1 =	43
k =	70	n2 =	49
t кр (0.05) =	1,994437112	D1 =	0,000306
t кр (0.01) =	2,647904624	D2 =	0,000127
Принимается			
ДЭК ао			
t =	0,073727688	n1 =	43
k =	75	n2 =	49
t кр (0.05) =	1,992543495	D1 =	0,000261
t кр (0.01) =	2,643912872	D2 =	0,000133
Принимается			
Сбербанк			
t =	0,529619182	n1 =	43
k =	75	n2 =	49
t кр (0.05) =	1,992102154	D1 =	0,000268
t кр (0.01) =	2,642983067	D2 =	0,00014
Принимается			

Значения критериев ниже критических значений на уровнях значимости 0,05 и 0,01, значит, нет оснований отвергать гипотезу Н0 о равенстве средних значений логдоходностей за 2018 и 2019 годы для всех компаний.

9. Сравнительный анализ логдоходностей трех компаний.

9.1. Критерий Бартлетта

Перед проверкой следующих гипотез необходимо провести проверку на равенство дисперсий критерием Бартлетта.

Дисперсионный анализ					Акрон	ДЭК ао	Сбербанк
					0,05178	-0,076308193	-0,046852555
				Расчет Барклета	-0,03075	-0,008849615	0,044677965
				v	0,021171	-0,039045674	0,057119068
				0,000920198	-0,03827	0	0,015399216
				0,997974684	0,036617	-0,045091572	0,043851883
					0,016492	0,022320189	0,015067432
					0,032191	-0,041141943	-0,01837322
				84,85744155		-0,0344	-0,030459207
						0,002213	-0,072681551
						-0,02055	0,061276555
						0,032039	0,018540508
						0,022273	-0,047012182
						0,004396	0,005998818
Однофакторный дисперсионный анализ					0,034062	-0,019324273	0,005503436
ИТОГИ					-0,00232	0,066894235	-0,069494283
	Группы	Счет	Сумма	Среднее	Дисперсия	-0,01051	0,089948237
Акрон	220	0,993971	0,004518049	0,000405682		0,0136	0,076562714
ДЭК ао	220	-1,35923	-0,00617831	0,001466807		0,027009	-0,031090587
Сбербанк	220	1,251131	0,005686957	0,000888106			

Наблюдаемое числовое значение статистики Бартлетта оказалось больше критической точки, поэтому есть основания отвергнуть проверяемую гипотезу, таким образом, применение дисперсионного анализа необосновано. Это следует учитывать при дальнейшей проверке гипотез дисперсионным анализом.

9.2. Гипотеза о значимости зависимости значений логдоходности акций компании «ДВМП» от номера года.

Выдвигается гипотеза об отсутствии зависимости значений логдоходности акций компании.

H0: зависимость значений логдоходности акций компании отсутствует

H1: присутствует зависимость значений логдоходности акций компании

Чтобы сравнить логдоходности за несколько лет обратимся к сервису «Однофакторный дисперсионный анализ» в разделе «Анализ данных». Результаты процедуры представлены с помощью двух таблиц (см. Рисунок 29). Все расчеты можно найти на рабочем листе «Дисп анализ Акрон».

Однофакторный дисперсионный анализ					
ИТОГИ					
<i>Группы</i>	<i>Счет</i>	<i>Сумма</i>	<i>Среднее</i>	<i>Дисперсия</i>	
2015	33	0,4075959	0,012351389	0,00067494	
2016	44	0,2604008	0,005918201	0,00061409	
2017	45	0,1375359	0,003056354	0,00039891	
2018	43	0,1770775	0,004118082	0,00030565	
2019	49	0,0121971	0,00024892	0,00012713	
Дисперсионный анализ					
<i>Источник вариации</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-Значение</i>
Между группами	0,0031	4	0,0007759	1,91918151	0,108448168
Внутри групп	0,0845	209	0,000404287		
Итого	0,0876	213			
$F = 1,9191851 < 2,414848367 = F \text{ критическое}$ $P - \text{значение} = 0,108448 > 0,05 = \alpha$					
Гипотеза принимается					

Рисунок 29. Результаты процедуры «Однофакторный дисперсионный анализ» для компании «Акрон».

В результате можно сравнить полученные показатели. Критерий фактического F – распределения оказался меньше критического, а P -значение превысило заданный уровень значимости. Таким образом, отсутствуют основания отвергнуть выдвинутую гипотезу H_0 .

9.3. Гипотеза о значимости зависимости значений логдоходности акций компаний от названия компании.

Данная гипотеза также рассматривается с использованием сервиса «Однофакторный дисперсионный анализ».

Критерий фактического F – распределения оказался больше критического, а P -значение не превысило заданный уровень значимости. Таким образом, есть основания отвергнуть выдвинутую гипотезу (см. Рисунок 30). Расчеты представлены на листе «Дисп анализ 2» в расчетном файле Excel.

Однофакторный дисперсионный анализ							
ИТОГИ							
Группы	Счет	Сумма	Среднее	Дисперсия			
Акрон	96	0,200498997	0,002088531	0,0002119			
ДЭК ао	96	-0,80874573	-0,008424435	0,0015534			
Сбербанк	96	0,547035581	0,005698287	0,0008583			
Дисперсионный анализ							
Источник вариации	SS	df	MS	F	P-Значение	F критическое	
Между группами	0,01033613	2	0,005168065	5,9094012	0,003057503	3,027443265	
Внутри групп	0,249246667	285	0,00087455				
Итого	0,259582797	287					
F = 5,9094012 > 3,0274 = F критическое					Гипотеза отклоняется		
P - значение = 0,0030575 < 0,05 = альфа							

Рисунок 30. Результаты процедуры «Однофакторный дисперсионный анализ» для трех компаний.

10. Графики временных рядов логдоходностей трех компаний

Для визуализации имеющихся данных о динамике показателей логдоходности компаний используем сервис «Вставка» для построения графиком (см. Рисунок 31, 32 и 33). Данные графики представлены на рабочем листе «Зависимость ЛД от Т» в расчетном файле Excel.

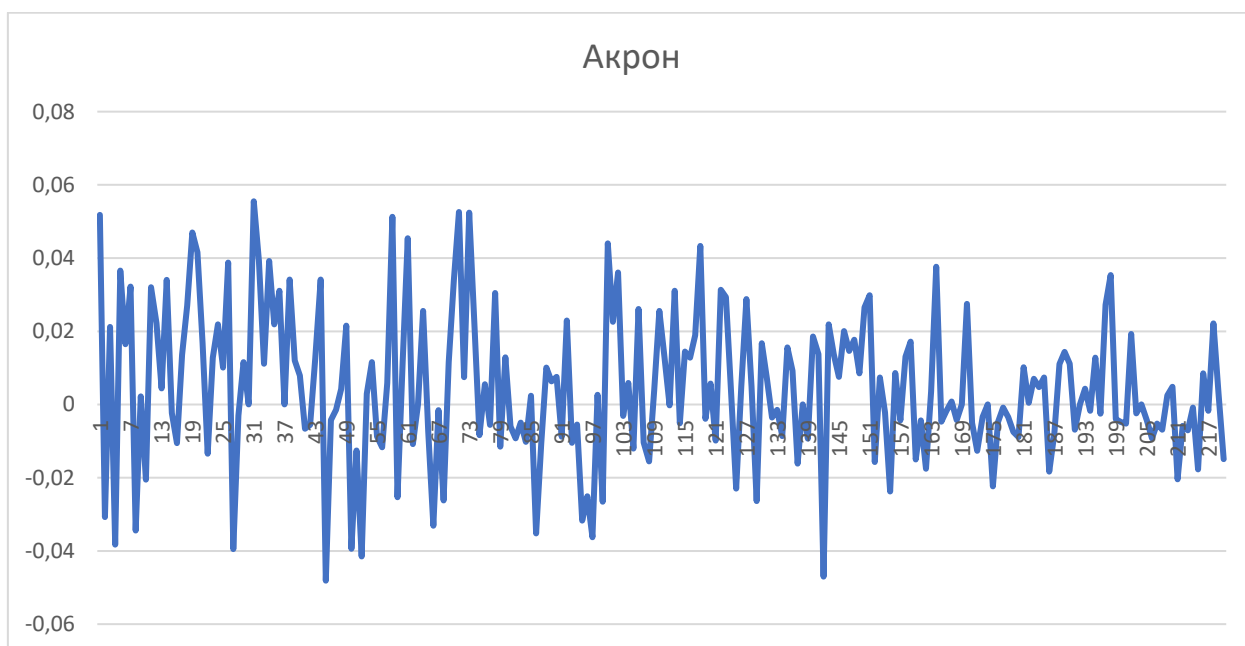


Рисунок 31. Зависимость показателей логдоходности компании «Акрон» от времени.

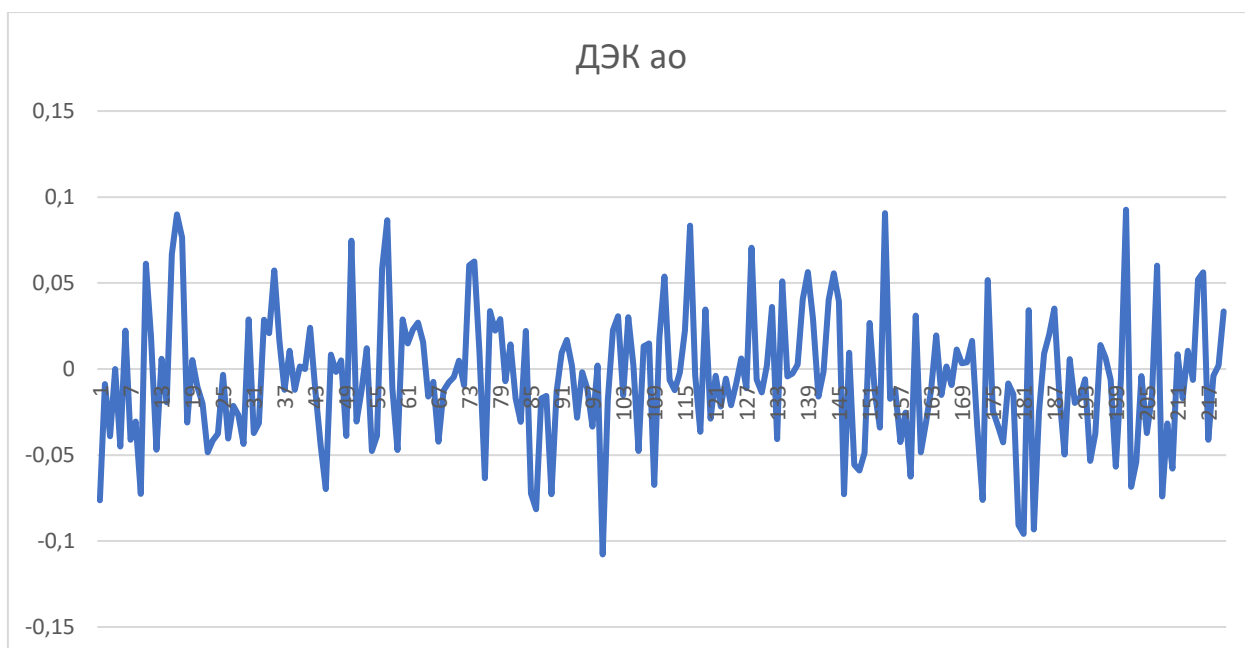


Рисунок 32. Зависимость показателей логдоходности компании «ДЭК ао» от времени.

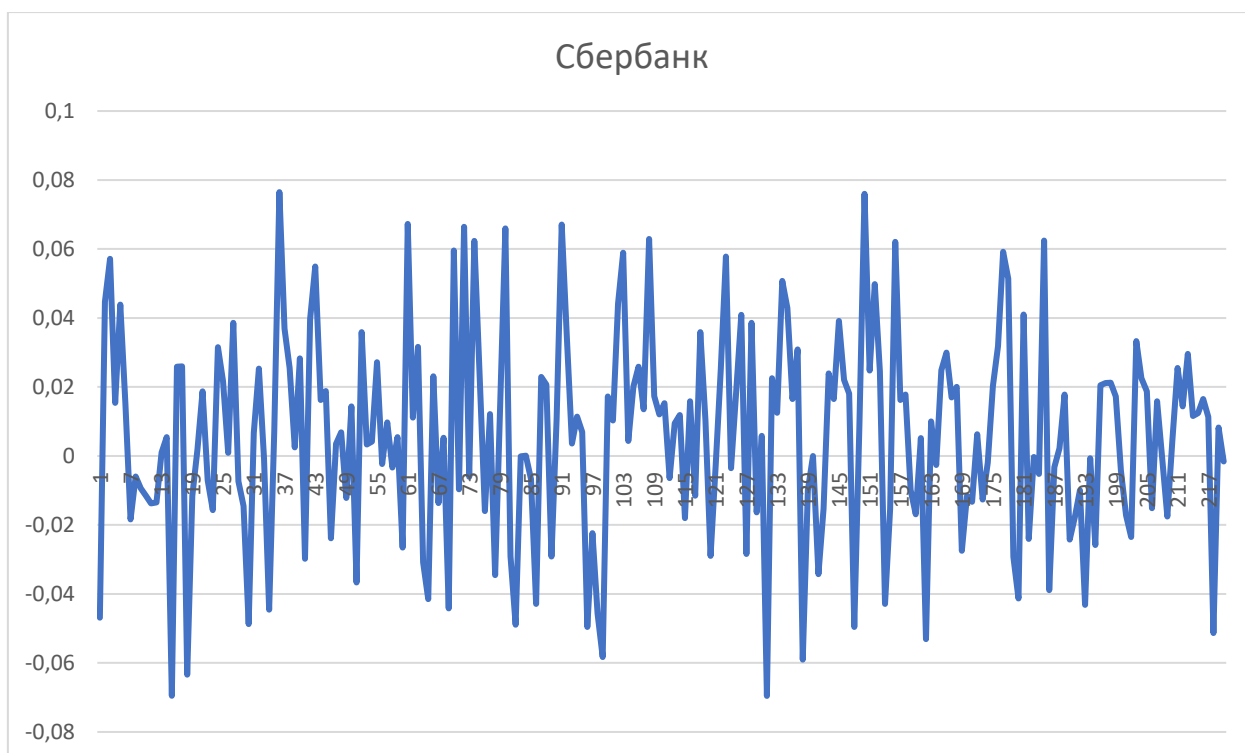


Рисунок 33. Зависимость показателей логдоходности компании «Сбербанк» от времени.

С помощью представленных графиком можно сделать вывод, что динамика логдоходностей компании «ДЭК ао» за рассматриваемый период является относительно «спокойным», а акции «Сбербанк» отличается более высокой волатильностью показателей логдоходности за весь период.

11. Исследование тесноты связи между логдоходностями

С помощью сервиса «Корреляция» в разделе «Анализ данных» сгенерируем корреляционную матрицу для логарифмических доходностей трех компаний.

	<i>Акрон</i>	<i>ДЭК ао</i>	<i>Сбербанк</i>
<i>Акрон</i>	1		
<i>ДЭК ао</i>	-0,006150168	1	
<i>Сбербанк</i>	0,072495529	0,083771685	1

Рисунок 34. Корреляционная матрица

Для каждого значения статистического коэффициента корреляции вычисляем значение:

$$t = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2},$$

далее находим критические значения распределения Стьюдента с помощью статистической функции СТЬЮДЕНТ.ОБР.2Х. Результаты вычислений приведены в следующей таблице:

r	t	ткр (0,05)	ткр (0,01)
-0,006150168	-0,087843626	1,971660889	2,600143757
0,072495529	1,038174979	1,971660889	2,600143757
0,083771685	1,200719546	1,971660889	2,600143757

Рисунок 35. Расчет фактического и критических значений критерия Т.

Из представленной таблицы следует, что гипотезу о наличии корреляции между логдоходностями компаний нет оснований отклонить на уровне значимости $\alpha = 0,05$ и на уровне значимости $\alpha = 0,01$.