

Федеральное государственное образовательное бюджетное учреждение
высшего образования

**«ФИНАНСОВЫЙ УНИВЕРСИТЕТ
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)**

Департамент анализа данных и машинного обучения

**Расчётно - аналитическая работа по
предмету "Анализ данных"**

Вид исследуемых данных:

Выполнил:

студент группы ПИ20-5

Руководитель:

к.ф.-м.н., доцент Пяткина Д.А.

Москва

2022

Оглавление

1. Загрузка исходных данных	3
2. Вычисление дополнительных признаков для каждой компании.....	4
3. Исследование изменения цен и объемов торгов акций.....	4
4. Исследование логарифмических доходностей (логдоходностей) акций (удаление выбросов)	14
5. Интервальные оценки параметров логдоходностей (в предположении нормального распределения), доверительные интервалы	16
6. Проверка гипотезы о нормальности логдоходности по критерию Пирсона (хи-квадрат)	17
7. Проверка гипотезы о нормальности логдоходности по критерию Колмогорова-Смирнова	19
8. Проверка гипотезы о равенстве нулю математического ожидания недельной логдоходности при альтернативное гипотезе о том, что оно больше нуля	23
9. Проверка гипотезы о равенстве дисперсий за последний и предпоследний годы (двусторонний F-тест).....	23
10. Проверка гипотезы о равенстве средних значений логарифмической доходности за последний и предпоследний годы (двусторонний t-тест с информацией о равенстве дисперсий)	24
11. Сравнительный анализ логдоходностей трех компаний.....	24
12. Исследование тесноты связи между логдоходностями	29

1. Загрузка исходных данных

Вариант 102

102	70	10	58	14	51	Борисов Никита Алексеевич
-----	----	----	----	----	----	---------------------------

Тикеры для моего варианта:

- ВТБ ао
- ДатСБ ао
- РОСИНТЕРАо
- Соллерс
- ФСК ЕЭС ао

Для работы взял первые 3

Выгрузили данные по всем 5 Тикерам с сайта mfd.ru за период с 2015 по 2019 год.

```
<TICKER>;<PER>;<DATE>;<TIME>;<CLOSE>;<VOL>
ВТБ ао;W;05/01/15;000000;0,06531;63232270000
ВТБ ао;W;12/01/15;000000;0,06522;127227400000
ВТБ ао;W;19/01/15;000000;0,0638;124305460000
ВТБ ао;W;26/01/15;000000;0,06894;157334020000
ВТБ ао;W;02/02/15;000000;0,06826;86119860000
ВТБ ао;W;09/02/15;000000;0,0694;111104070000
ВТБ ао;W;16/02/15;000000;0,0666;101749690000
ВТБ ао;W;23/02/15;000000;0,068;136836940000
ВТБ ао;W;02/03/15;000000;0,06472;77666160000
ВТБ ао;W;09/03/15;000000;0,06332;85965190000
ВТБ ао;W;16/03/15;000000;0,0634;92591020000
ВТБ ао;W;23/03/15;000000;0,0595;78153790000
ВТБ ао;W;30/03/15;000000;0,06162;96640950000
ВТБ ао;W;06/04/15;000000;0,0587;250275670000
ВТБ ао;W;13/04/15;000000;0,05805;97081010000
ВТБ ао;W;20/04/15;000000;0,05984;122653670000
ВТБ ао;W;27/04/15;000000;0,0655;150612500000
ВТБ ао;W;04/05/15;000000;0,0675;146761730000
ВТБ ао;W;11/05/15;000000;0,06843;222016180000
ВТБ ао;W;18/05/15;000000;0,07777;306404170000
ВТБ ао;W;25/05/15;000000;0,08025;370794590000
ВТБ ао;W;01/06/15;000000;0,08107;205476980000
ВТБ ао;W;08/06/15;000000;0,0804;185527450000
```

Далее загрузили данные в Excel

<TICKER>	<DATE>	<CLOSE>	<VOL>
ВТБ ао	05.01.2015	0,06531	63232270000
ВТБ ао	12.01.2015	0,06522	1,27227E+11
ВТБ ао	19.01.2015	0,0638	1,24305E+11
ВТБ ао	26.01.2015	0,06894	1,57334E+11
ВТБ ао	02.02.2015	0,06826	86119860000
ВТБ ао	09.02.2015	0,0694	1,11104E+11
ВТБ ао	16.02.2015	0,0666	1,0175E+11
ВТБ ао	23.02.2015	0,068	1,36837E+11
ВТБ ао	02.03.2015	0,06472	77666160000
ВТБ ао	09.03.2015	0,06332	85965190000

Отредактировали данные по тикерам и расположили следующим образом

	A	B	C	D	E	F	G
	Дата	Цена			Объем торгов		
		ВТБ ао	ДгаСб ао	РОСИНТЕРао	ВТБ ао	ДгаСб ао	РОСИНТЕРао
3	05.01.2015	0,06531	0,0729	92,1	63232270000	42090000	2020
4	12.01.2015	0,06522	0,051	80	1,27227E+11	23570000	14790
5	19.01.2015	0,0638	0,08	81	1,24305E+11	118680000	36690
5	26.01.2015	0,06894	0,079	72,7	1,57334E+11	28310000	10030
7	02.02.2015	0,06826	0,05699	74	86119860000	23470000	960
8	09.02.2015	0,0694	0,0556	86	1,11104E+11	63100000	7070
9	16.02.2015	0,0666	0,05802	84,5	1,0175E+11	31760000	5780

2. Вычисление дополнительных признаков для каждой компании

Далее дополнили столбцами Логарифм цены, доходность, логарифм доходности, логарифм объема

	H	I	J	K	L	M	N	O	P	Q	R	S
	Логарифм цен			Доходности			Лог доходности			Логарифм объема торгов		
	ВТБ ао	ДгаСб ао	РОСИНТЕРао	ВТБ ао	ДгаСб ао	РОСИНТЕРао	ВТБ ао	ДгаСб ао	РОСИНТЕРао	ВТБ ао	ДгаСб ао	РОСИНТЕРао
0	Область построения	0,066664	4,522874943							24,87008061	17,55532074	7,61085279
0	-2,72998911	-2,97592965	4,382026635	-0,00137804	-0,30041152	-0,131378936	-0,00137899	-0,35726301	-0,140848309	25,56924187	16,97548527	9,601706556
0	-2,75200209	-2,52572864	4,394449155	-0,02177246	0,568627451	0,0125	-0,02201298	0,450201002	0,01242252	25,54600776	18,59194135	10,51025952
0	-2,67451872	-2,53830743	4,286341385	0,080564263	-0,0125	-0,102469136	0,077483371	-0,01257878	-0,10810777	25,7816369	17,15872566	9,213335881
0	-2,68443134	-2,86487947	4,304065093	-0,00986365	-0,27860759	0,017881706	-0,00991262	-0,32657204	0,017723709	25,17900588	16,97123357	6,866933284
0	-2,66786841	-2,88957208	4,454347296	0,01670085	-0,02439024	0,162162162	0,016562924	-0,02469261	0,150282203	25,43373317	17,96023133	8,863615759
0	-2,7090507	-2,8469675	4,436751534	-0,04034582	0,04352518	-0,01744186	-0,04118229	0,042604577	-0,017595762	25,34578161	17,27371819	8,662158962
0	-2,68824757	-2,71055333	4,406719247	0,021021021	0,146156498	-0,029585799	0,020803128	0,136414169	-0,030032287	25,64205583	16,40252243	8,330863613
0	-2,73768501	-2,4831482	4,430816799	-0,04823529	0,255338346	0,024390244	-0,04943743	0,227405135	0,024097552	25,07568548	17,09227786	9,788637711
0	-2,75955404	-2,30258509	4,49980967	-0,02163164	0,197891711	0,071428571	-0,02186904	0,180563104	0,068992871	25,17720828	16,81224233	8,503094267
0	-2,75829142	-2,30078671	4,54648119	0,001263424	0,0018	0,047777778	0,001262626	0,001798382	0,046671519	25,251458	17,15801894	9,326344048
0	-2,82177897	-2,33541825	4,48863637	-0,0615142	-0,03403873	-0,056203606	-0,06348755	-0,03463154	-0,05784482	25,08194439	16,63748451	9,360483303
0	-2,78676879	-2,27302629	4,540098189	0,035630252	0,064379456	0,052808989	0,035010181	0,062391959	0,05146182	25,2942684	16,30954212	8,788745882
0	-2,83531555	-2,16317104	4,536891345	-0,04738721	0,116116505	-0,003201708	-0,04854677	0,109855254	-0,003206844	26,24582883	16,43654938	8,527143522
0	-2,84645057	-2,28818921	4,382026635	-0,01107325	-0,11751914	-0,143468951	-0,01113502	-0,12501818	-0,154864711	25,29881162	16,70643772	9,027618735
0	-2,81608095	-2,22562405	4,494238625	0,030835487	0,064563825	0,11875	0,030369626	0,062565161	0,112211991	25,53263053	15,53290561	8,701180028
0	-2,72570514	-2,13724016	4,498697941	0,094585561	0,092407407	0,004469274	0,090375809	0,088383891	0,004459316	25,73797615	16,17636456	9,927692241
0	-2,6956768	-2,14558134	4,454347296	0,030524351	-0,00830649	-0,033381535	0,030077455	-0,00834118	-0,04350645	25,71207627	15,73378796	6,396929655

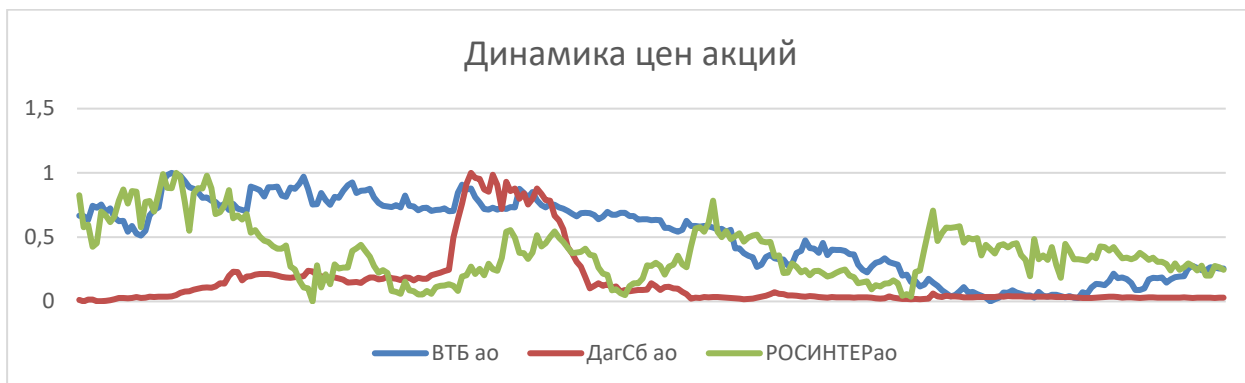
3. Исследование изменения цен и объемов торгов акций

Для получения представления о поведении различных цен акций создадим дополнительный столбец, в котором вычислим относительные цены по формуле

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

	T	U	V
	Относительные цены		
ао	ВТБ ао	ДгаСб ао	РОСИНТЕРао
79	0,666243117	0,01168	0,826446281
56	0,664337145	0	0,576446281
52	0,634265142	0,015466667	0,597107438
81	0,743117323	0,014933333	0,425619835
84	0,728716645	0,003194667	0,452479339
59	0,752858958	0,002453333	0,700413223
62	0,69356205	0,003744	0,669421488
13	0,723210504	0,008266667	0,617768595
11	0,653748412	0,017322667	0,659090909
67	0,624099958	0,026133333	0,783057851
48	0,625794155	0,026229333	0,871900826

Теперь мы можем построить график динамики цен всех трех акций сразу.



Здесь можем наблюдать, что цены акций ВТБ ао постепенно снижались, как и РОСИНТЕР ао, а цены ДагСБ ао в 2018 году аномально подскочили.

Для исследования изменения значений показателей удобно будет использовать описательную статистику.

На следующем рисунке представлена описательная статистика по цене акций. Здесь видим, что акции ВТБ самые дешевые, а акции РОСИНТЕР в разы дороже остальных.

Цена					
ВТБ ао		ДагСб ао		РОСИНТЕР ао	
Среднее	0,057810479	Среднее	0,341622337	Среднее	70,48965517
Стандартная ошибка	0,000859014	Стандартная ошибка	0,027109468	Стандартная ошибка	0,648568012
Медиана	0,0616	Медиана	0,12555	Медиана	69
Мода	0,06843	Мода	0,108	Мода	80
Стандартное отклонение	0,013877803	Стандартное отклонение	0,437966864	Стандартное отклонение	10,47793691
Дисперсия выборки	0,000192593	Дисперсия выборки	0,191814974	Дисперсия выборки	109,7871618
Эксцесс	-1,363248797	Эксцесс	4,700920798	Эксцесс	0,192511436
Асимметричность	-0,269231399	Асимметричность	2,392570475	Асимметричность	0,759760504
Интервал	0,04722	Интервал	1,875	Интервал	48,4
Минимум	0,03385	Минимум	0,051	Минимум	52,1
Максимум	0,08107	Максимум	1,926	Максимум	100,5
Сумма	15,088535	Сумма	89,16343	Сумма	18397,8
Счет	261	Счет	261	Счет	261

Далее описательная статистика по объему торгов

Объем торгов					
ВТБ ао		ДагСб ао		РОСИНТЕР ао	
Среднее	93057310307	Среднее	63913831,42	Среднее	17834,67433
Стандартная ошибка	3755631879	Стандартная ошибка	8239789,389	Стандартная ошибка	3007,121025
Медиана	80003560000	Медиана	25800000	Медиана	7240
Мода	#Н/Д	Мода	28310000	Мода	2020
Стандартное отклонение	60674089877	Стандартное отклонение	133117871,5	Стандартное отклонение	48581,52694
Дисперсия выборки	3,68135E+21	Дисперсия выборки	1,77204E+16	Дисперсия выборки	2360164760
Эксцесс	6,116243566	Эксцесс	35,18158736	Эксцесс	122,0781055
Асимметричность	2,141027709	Асимметричность	5,257474363	Асимметричность	9,892632583
Интервал	3,67826E+11	Интервал	1240220000	Интервал	662350
Минимум	11859280000	Минимум	220000	Минимум	40
Максимум	3,79686E+11	Максимум	1240440000	Максимум	662390
Сумма	2,4288E+13	Сумма	16681510000	Сумма	4654850
Счет	261	Счет	261	Счет	261

Наибольший объем торгов у самых дешевых акций (ВТБ).

Далее описательная статистика логарифма цен. Видим, что наибольший разброс значений имеет ДагСБ ао.

Логарифм цен					
ВТБ ао		ДагСб ао		РОСИНТЕРао	
Среднее	-2,882125181	Среднее	-1,562385945	Среднее	4,244954195
Стандартная ошибка	0,01595543	Стандартная ошибка	0,054867973	Стандартная ошибка	0,008905171
Медиана	-2,787093408	Медиана	-2,075051193	Медиана	4,234106505
Мода	-2,681943954	Мода	-2,225624052	Мода	4,382026635
Стандартное отклонение	0,257767862	Стандартное отклонение	0,886419232	Стандартное отклонение	0,143867439
Дисперсия выборки	0,066444271	Дисперсия выборки	0,785739055	Дисперсия выборки	0,02069784
Эксцесс	-1,218187731	Эксцесс	0,143255324	Эксцесс	-0,282002265
Асимметричность	-0,482945127	Асимметричность	1,021312904	Асимметричность	0,425151663
Интервал	0,87337398	Интервал	3,63137496	Интервал	0,656992779
Минимум	-3,38581628	Минимум	-2,975929646	Минимум	3,953164949
Максимум	-2,5124423	Максимум	0,655445313	Максимум	4,610157727
Сумма	-752,2346723	Сумма	-407,7827317	Сумма	1107,933045
Счет	261	Счет	261	Счет	261

Далее описательная статистика доходности.

Доходность					
ВТБ ао		ДагСб ао		РОСИНТЕРао	
Среднее	-0,00077222	Среднее	0,009657374	Среднее	0,000254938
Стандартная ошибка	0,002136282	Стандартная ошибка	0,008441261	Стандартная ошибка	0,003620259
Медиана	-0,00311396	Медиана	-0,003827611	Медиана	-0,001054853
Мода	0	Мода	0	Мода	0
Стандартное отклонение	0,034446515	Стандартное отклонение	0,136111244	Стандартное отклонение	0,058374916
Дисперсия выборки	0,001186562	Дисперсия выборки	0,018526271	Дисперсия выборки	0,003407631
Эксцесс	2,19754044	Эксцесс	13,93458612	Эксцесс	2,978568086
Асимметричность	0,718065546	Асимметричность	2,152862417	Асимметричность	0,869280189
Интервал	0,248451248	Интервал	1,35945246	Интервал	0,404505419
Минимум	-0,111961404	Минимум	-0,433962264	Минимум	-0,143468951
Максимум	0,136489844	Максимум	0,925490196	Максимум	0,261036468
Сумма	-0,200777261	Сумма	2,510917338	Сумма	0,066283842
Счет	260	Счет	260	Счет	260

Здесь можно наблюдать то, что видели изначально на динамике цен, доходность акций ВТБ ао на самом деле является убытком, то есть отрицательна в среднем.

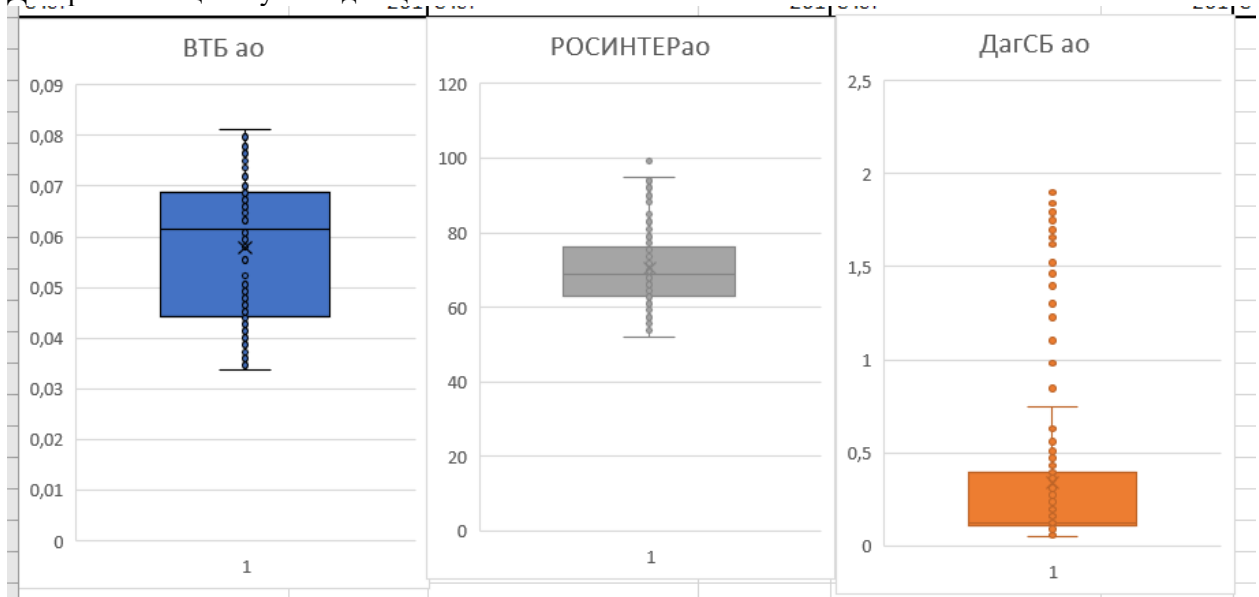
Более ясное понимание дает описательная статистика по логарифму доходности.

Логарифм доходности					
ВТБ ао		ДагСб ао		РОСИНТЕРао	
Среднее	-0,001356462	Среднее	0,001366547	Среднее	-0,001399969
Стандартная ошибка	0,00211694	Стандартная ошибка	0,007886702	Стандартная ошибка	0,00355601
Медиана	-0,003118821	Медиана	-0,00383496	Медиана	-0,00105541
Мода	0	Мода	0	Мода	0
Стандартное отклонение	0,034134624	Стандартное отклонение	0,127169248	Стандартное отклонение	0,057338945
Дисперсия выборки	0,001165173	Дисперсия выборки	0,016172018	Дисперсия выборки	0,003287755
Эксцесс	1,936354985	Эксцесс	7,239663046	Эксцесс	2,201441899
Асимметричность	0,532922861	Асимметричность	0,325042208	Асимметричность	0,526376365
Интервал	0,246684501	Интервал	1,224275115	Интервал	0,386798687
Минимум	-0,118740073	Минимум	-0,569094532	Минимум	-0,154864711
Максимум	0,127944428	Максимум	0,655180583	Максимум	0,231933977
Сумма	-0,352680047	Сумма	0,35530226	Сумма	-0,36399186
Счет	260	Счет	260	Счет	260

Далее логарифм объема торгов. Самый постоянный объем торгов у акций ВТБ, как и самый большой объем.

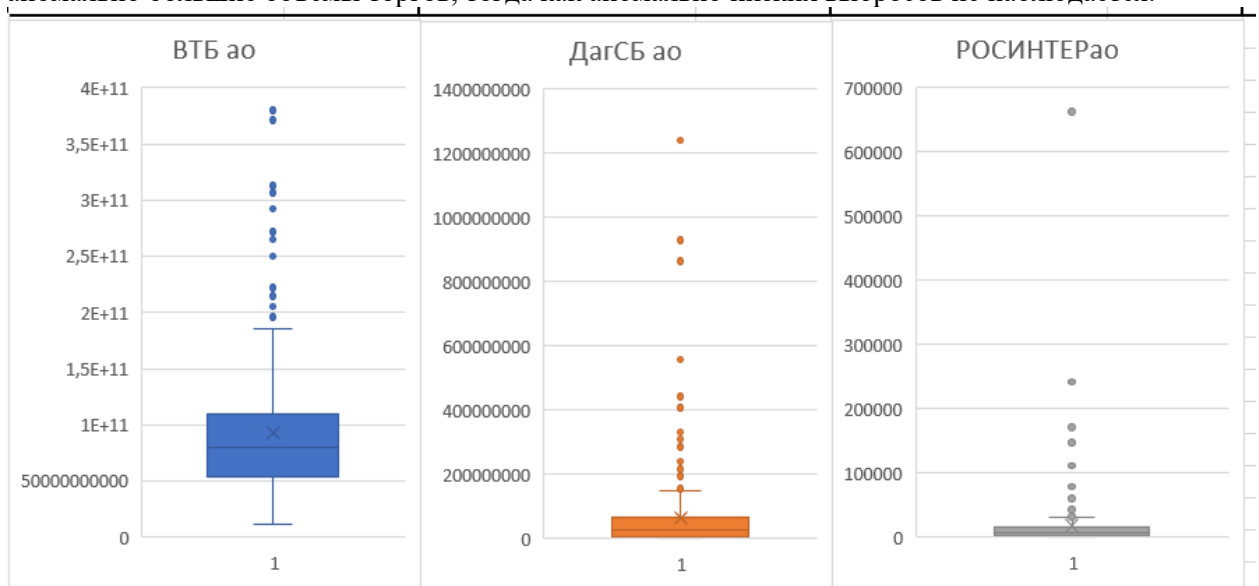
	AE	AF	AG	AH	AI	AJ
	Логарифм объема торгов					
	ВТБ ao		ДагСБ ao		РОСИНТЕРao	
9	Среднее	25,08526992	Среднее	16,76482594	Среднее	8,894745907
1	Стандартная ошибка	0,036143264	Стандартная ошибка	0,105466856	Стандартная ошибка	0,077966964
1	Медиана	25,10533697	Медиана	17,06588505	Медиана	8,887376485
0	Мода	#Н/Д	Мода	17,15872566	Мода	7,61085279
5	Стандартное отклонение	0,583912293	Стандартное отклонение	1,703869209	Стандартное отклонение	1,259594856
5	Дисперсия выборки	0,340953566	Дисперсия выборки	2,90317028	Дисперсия выборки	1,5865792
9	Эксцесс	0,433433885	Эксцесс	-0,392651718	Эксцесс	1,370457916
5	Асимметричность	-0,008959928	Асимметричность	-0,261003891	Асимметричность	0,01716892
7	Интервал	3,466232832	Интервал	8,637349167	Интервал	9,714730331
1	Минимум	23,19637652	Минимум	12,30138283	Минимум	3,688879454
7	Максимум	26,66260935	Максимум	20,93873199	Максимум	13,40360979
6	Сумма	6547,255449	Сумма	4375,619571	Сумма	2321,528682
0	Счет	261	Счет	261	Счет	261

Диаграммы Ящик с усами для цен

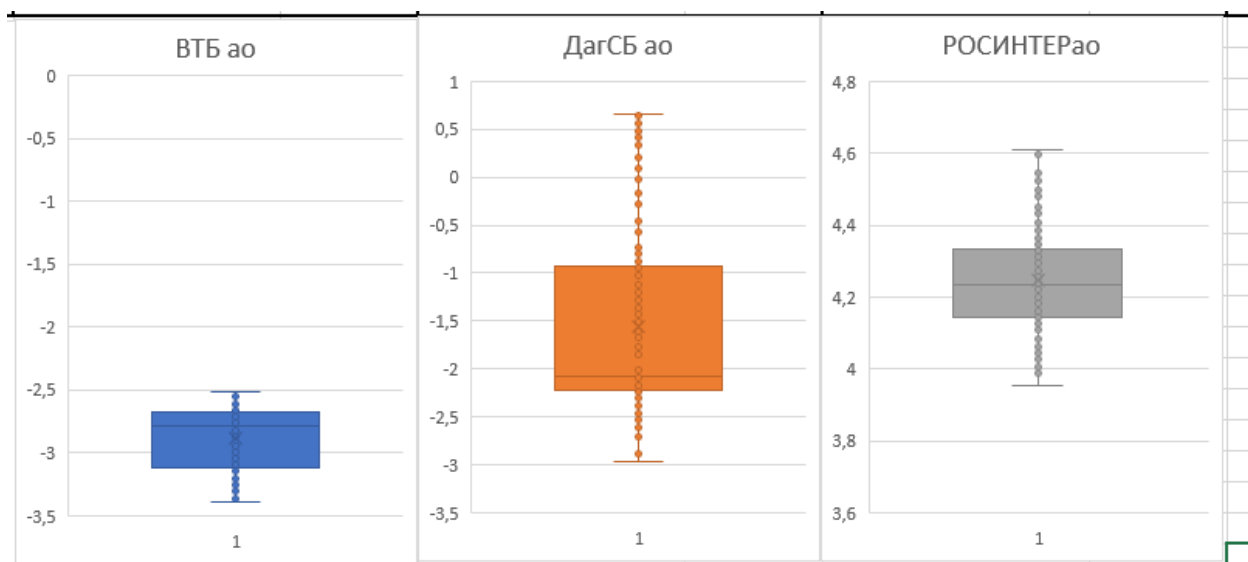


Как можно заметить, цены Акций ВТБ не имеют выбросов и распределены более равномерно. Росинтер имеет незначительное количество выбросов, а вот ДагСБ имеет целый период anomalously больших цен – выбросов.

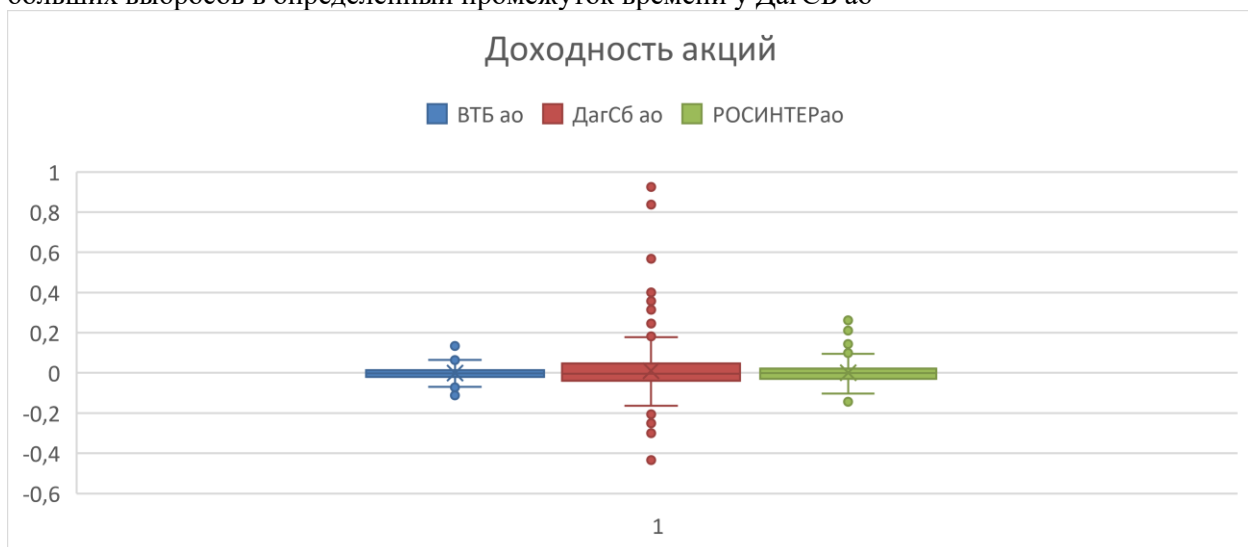
Из следующих диаграмм, построенных для объемов торгов, видно, что все три компании имели anomalously большие объемы торгов, тогда как anomalously низких выбросов не наблюдается.



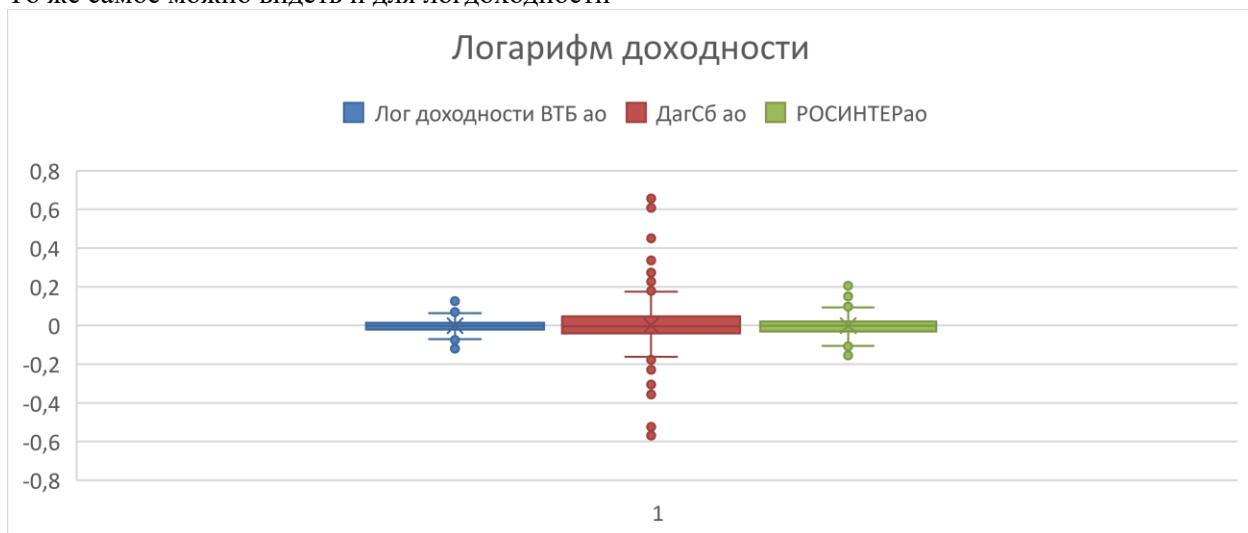
Далее логарифм цен



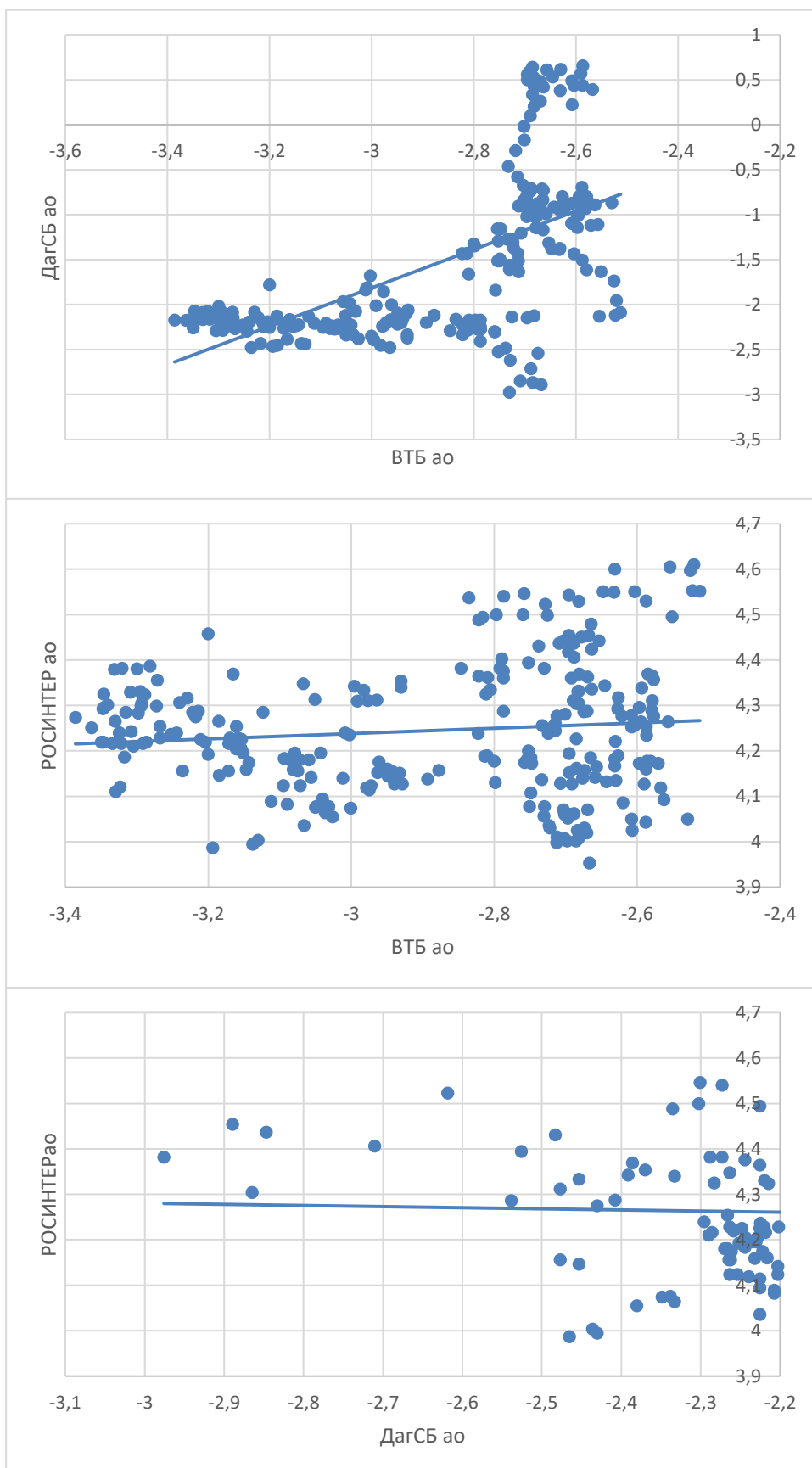
О доходности акций можно сказать, что она почти у всех одинаковая, за исключением anomalно больших выбросов в определенный промежуток времени у ДагСБ ао



То же самое можно видеть и для логдоходности



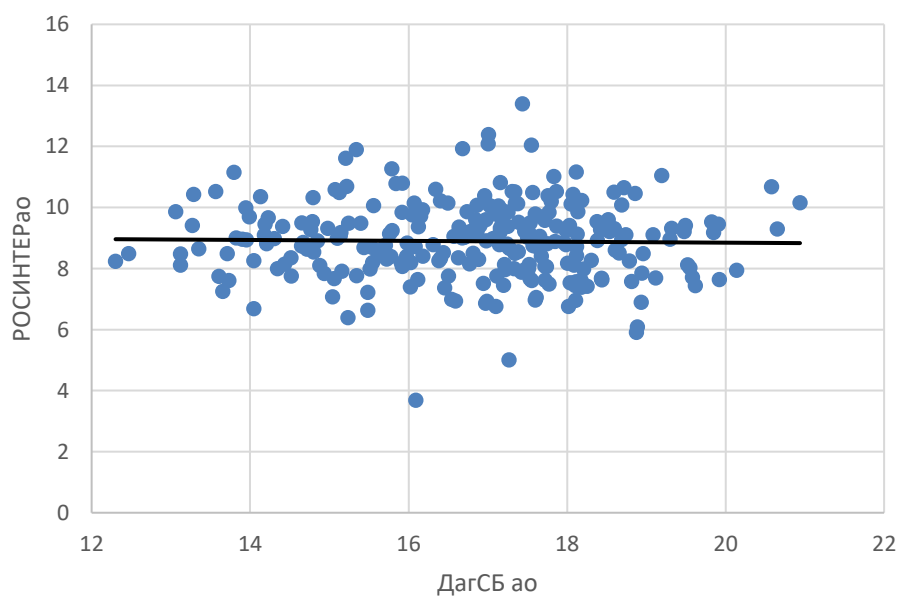
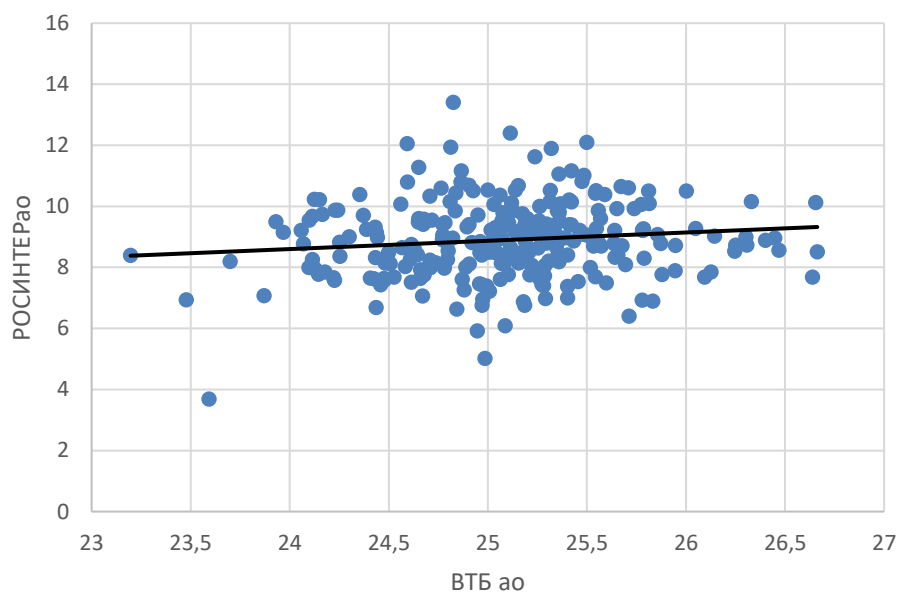
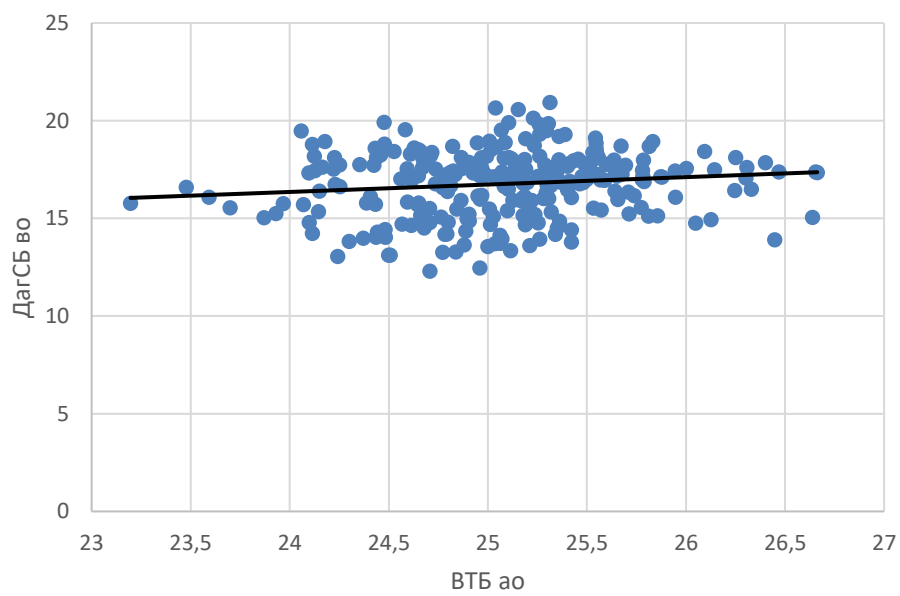
Диаграммы рассеяния для логарифма цен



И корреляции по ним

	<i>ВТБ ао</i>	<i>ДагСБ ао</i>	<i>РОСИНТЕР ао</i>
<i>ВТБ ао</i>	1		
<i>ДагСБ ао</i>	0,620660633	1	
<i>РОСИНТЕР ао</i>	0,104895655	-0,15194361	1

Здесь видим, что *ВТБ ао* и *ДагСБ ао* имеют более сильную зависимость между собой, чем в остальных случаях. При увеличении цен на акции *ВТБ ао* росли и цены на акции *ДагСБ ал*
 Диаграммы рассеяния для логарифма объема торгов

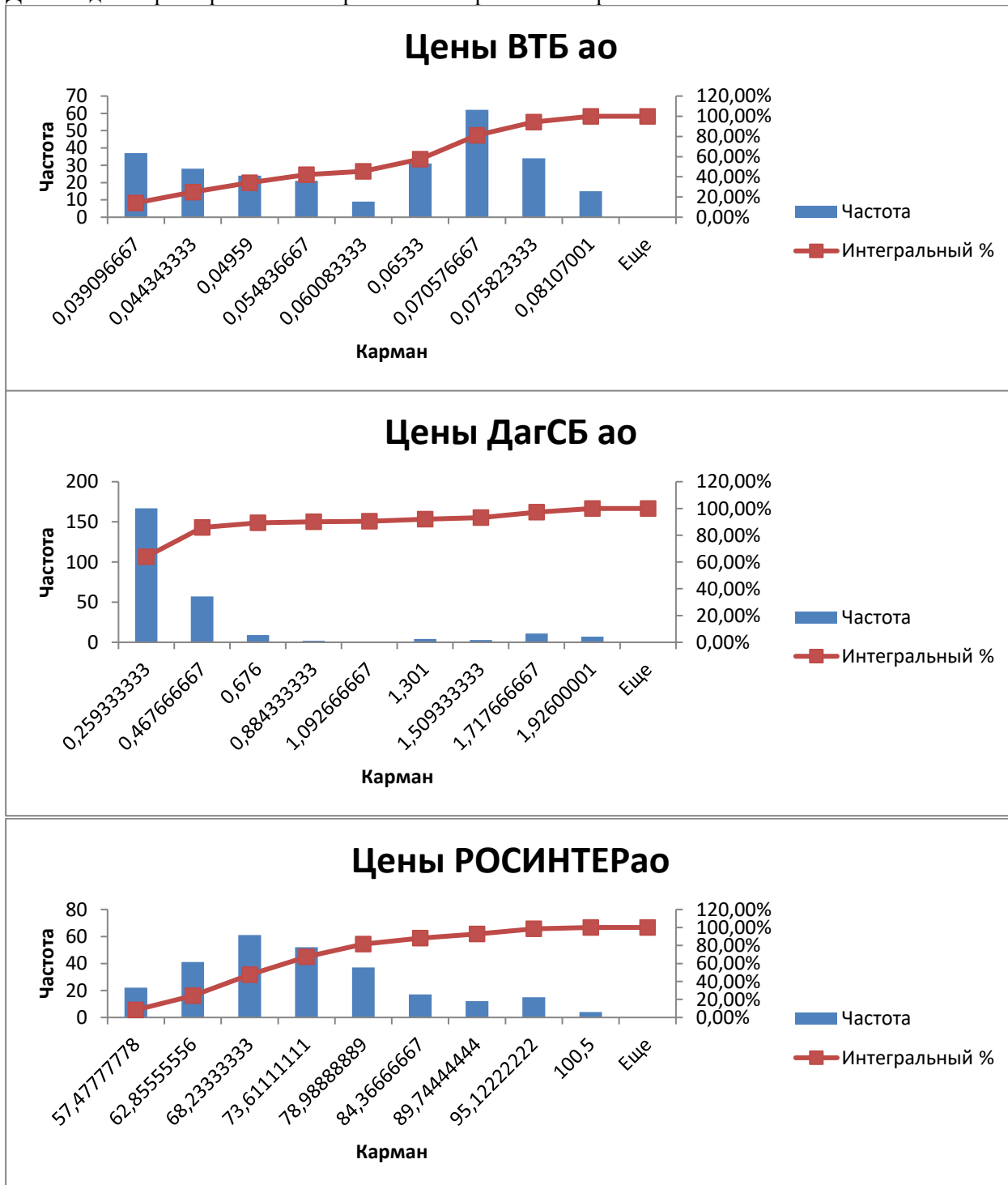


Логарифмы объемов торгов

	ВТБ ао	ДагСБ ао	РОСИНТЕРао
ВТБ ао	1		
ДагСБ ао	0,129967331	1	
РОСИНТЕРао	0,12583388	-0,019977193	1

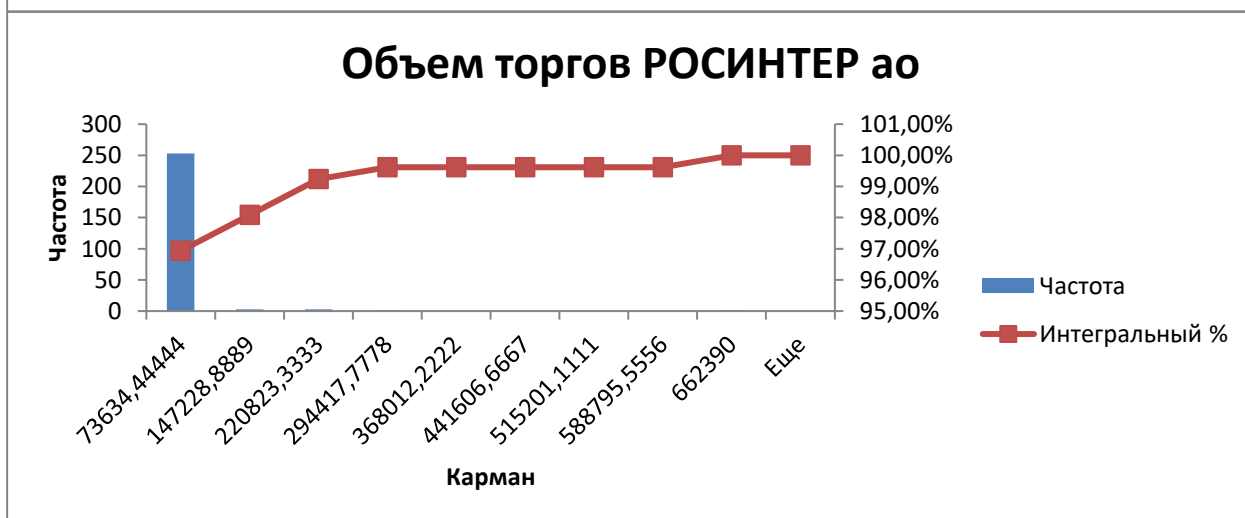
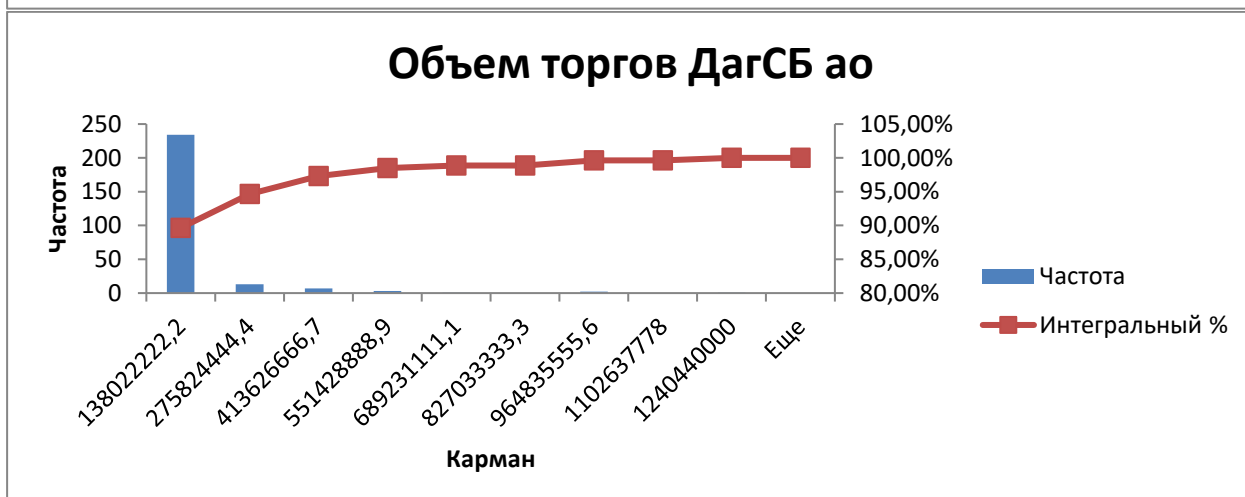
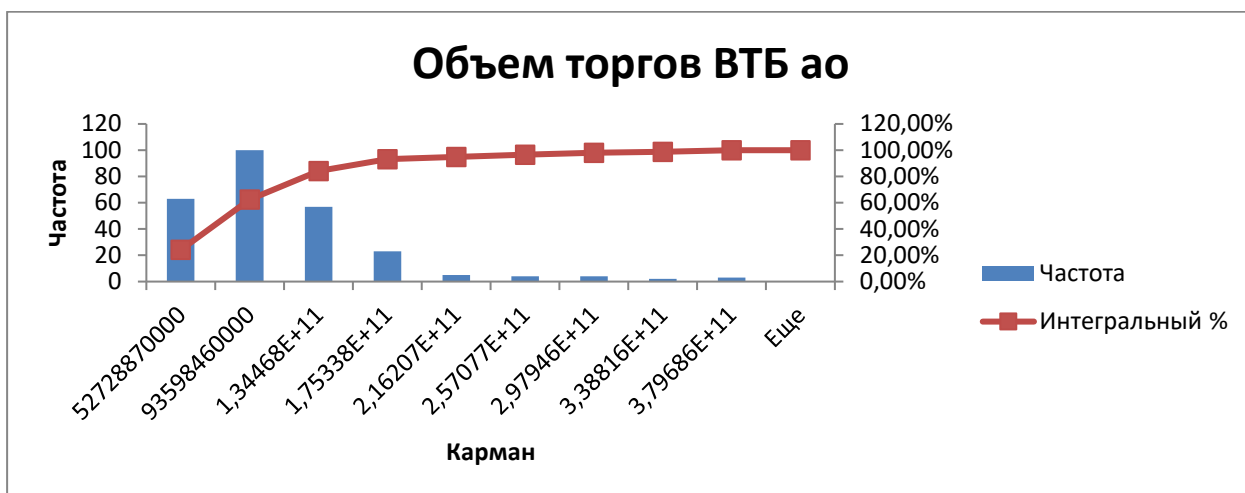
Здесь сложно проследить какую-либо связь, характеристики друг от друга слабо зависят.

Для каждой характеристики построили гистограммы интервальных частот



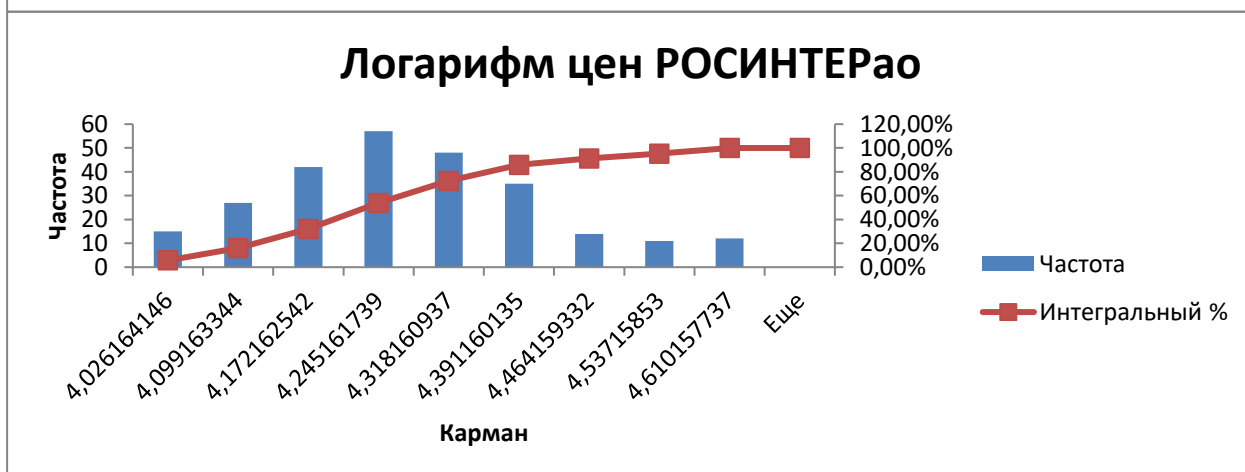
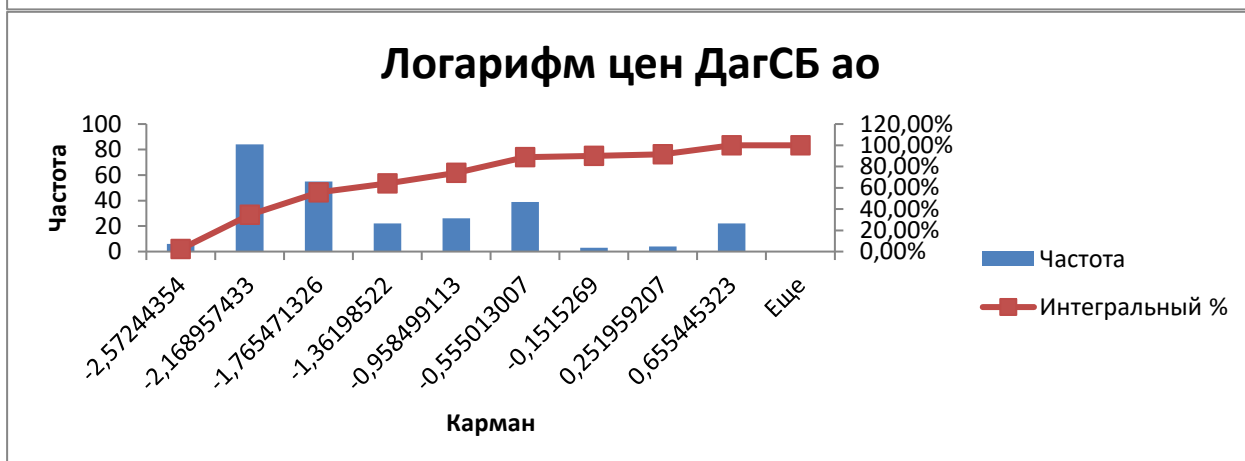
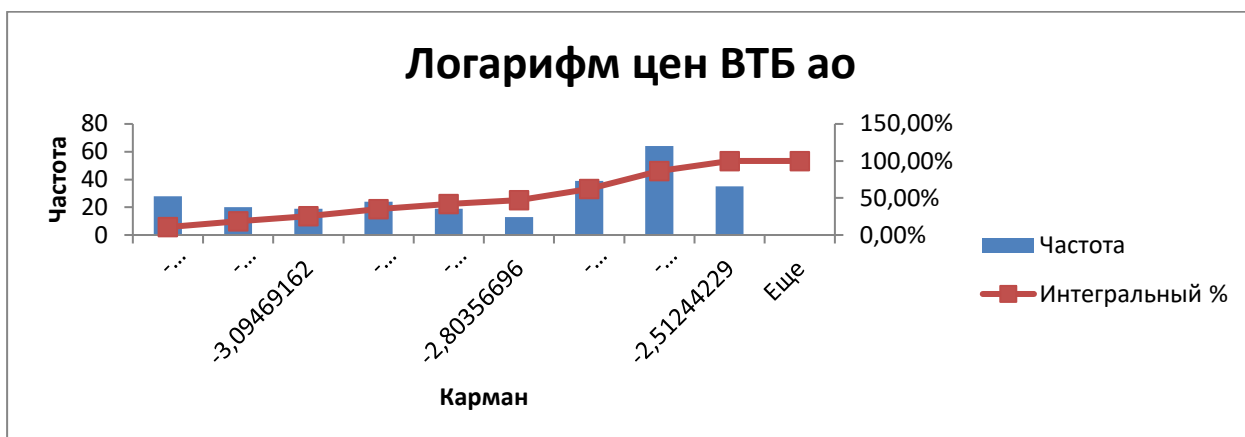
Распределение цен РОСИНТЕРао больше всего похоже на нормальное.

Далее графики для объема

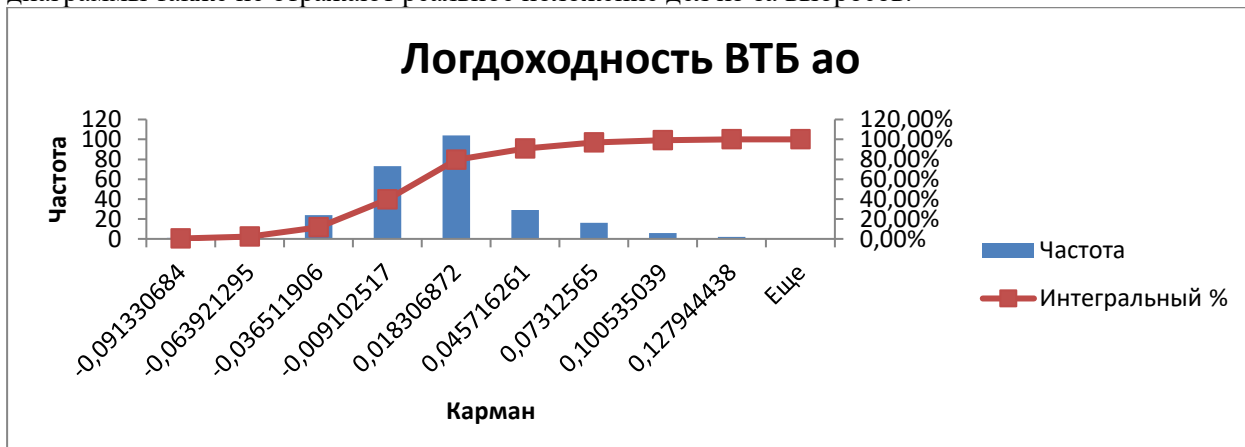


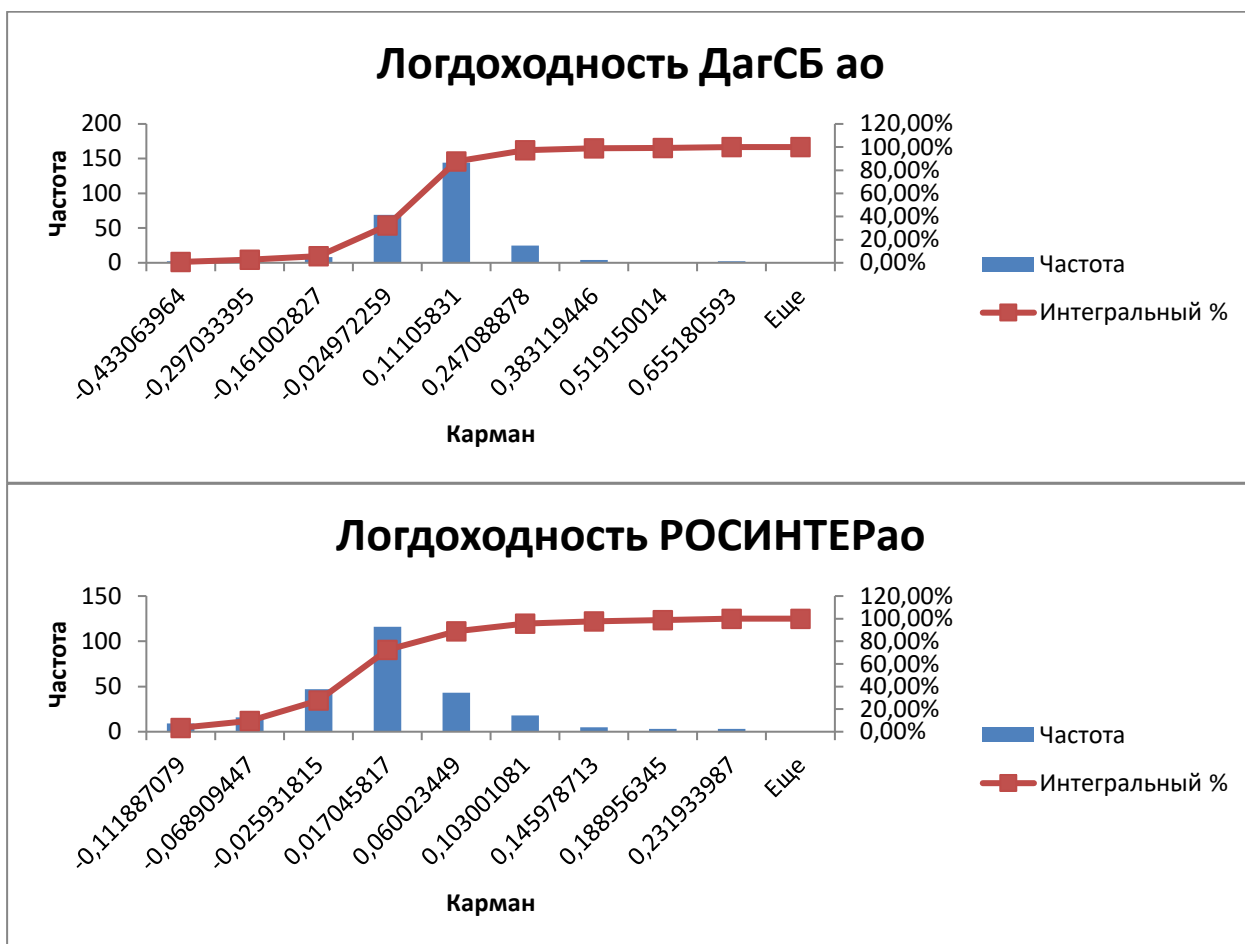
Данные диаграммы не отображают действительности вещей, потому что сильно искажены выбросами, ввиду чего имеем один интервал практически со всеми значениями, и интервалы с единичными аномальными значениями.

Логарифмированные цены уже более точно отражают картину, но все еще искажены выбросами, рисунки ниже



Далее логдоходности. Более всего на нормальное распределение похожи ВТБ и РОСИНТЕР, но диаграммы также не отражают реальное положение дел из-за выбросов.





Далее необходимо избавиться от выбросов для Логдоходностей.

Для примера построим корреляционную матрицу признаков Логдоходности до удаления выбросов

	<i>ВТБ ао</i>	<i>ДагСб ао</i>	<i>РОСИНТЕР ао</i>
<i>ВТБ ао</i>	1		
<i>ДагСб ао</i>	0,006452271	1	
<i>РОСИНТЕР ао</i>	0,020434483	0,129605031	1

4. Исследование логарифмических доходностей (логдоходностей) акций (удаление выбросов)

Далее проведем поиск и удаление выбросов.

Для каждой организации посчитаем

q1, q3 первый и третий Квартиль, найдем размах IQR и вычислим нижнюю границу x0 и верхнюю x1/

q1	-0,020637137	-0,04006892	-0,030070459
q3	0,013093118	0,046520016	0,020421539
IQR	0,033730255	0,086588935	0,050491998
x0	-0,071232521	-0,169952322	-0,105808457
x1	0,063688501	0,176403418	0,096159536

Далее все значения – строки, находящиеся вне этой границы, удалим.

Итого, после удаления, остается 199 строк (было удалено 62 значения).

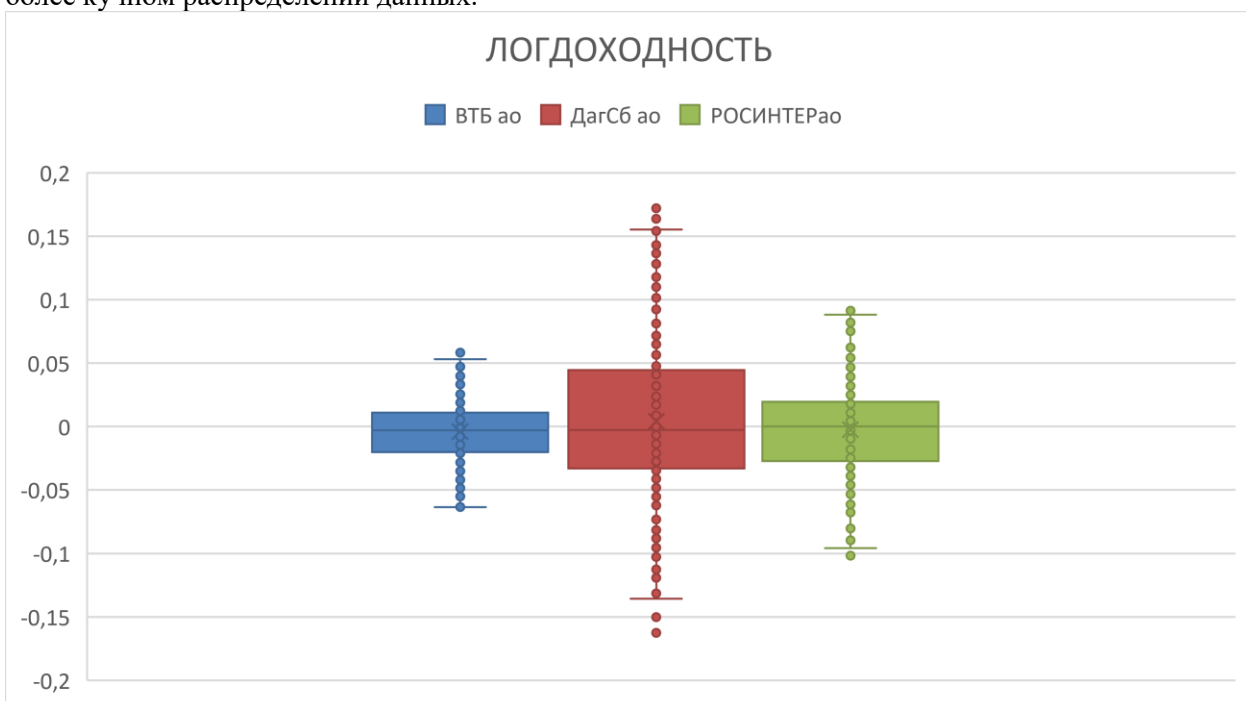
Теперь проверим корреляцию после удаления выбросов

	<i>ВТБ ао</i>	<i>ДагСб ао</i>	<i>РОСИНТЕР ао</i>
<i>ВТБ ао</i>	1		
<i>ДагСб ао</i>	0,050013471	1	
<i>РОСИНТЕР ао</i>	0,159906747	0,104673575	1

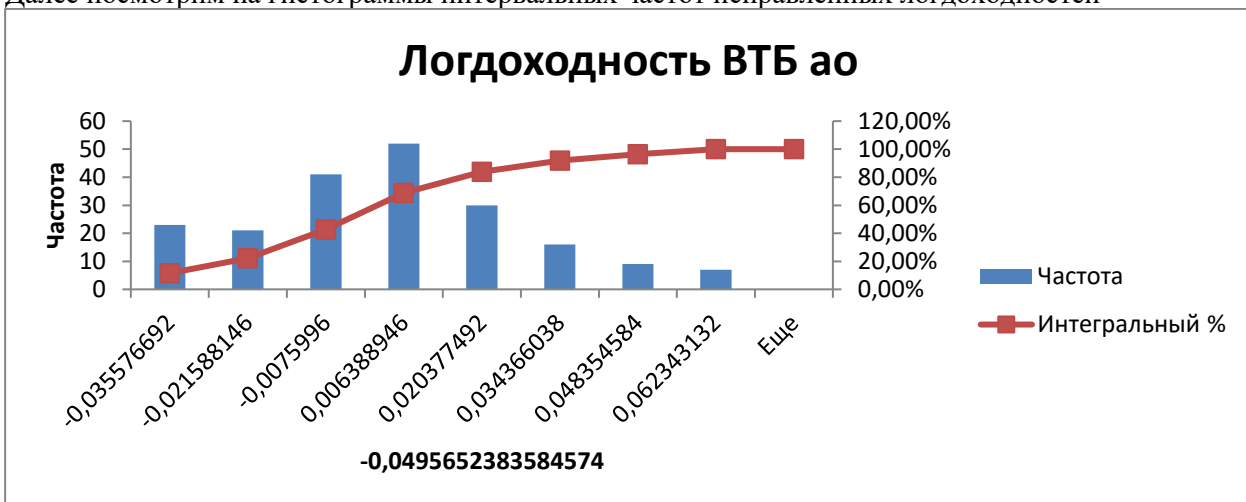
Теперь корреляция между признаками приняла более реальный вид.
Также можно увидеть разницу описательной статистики

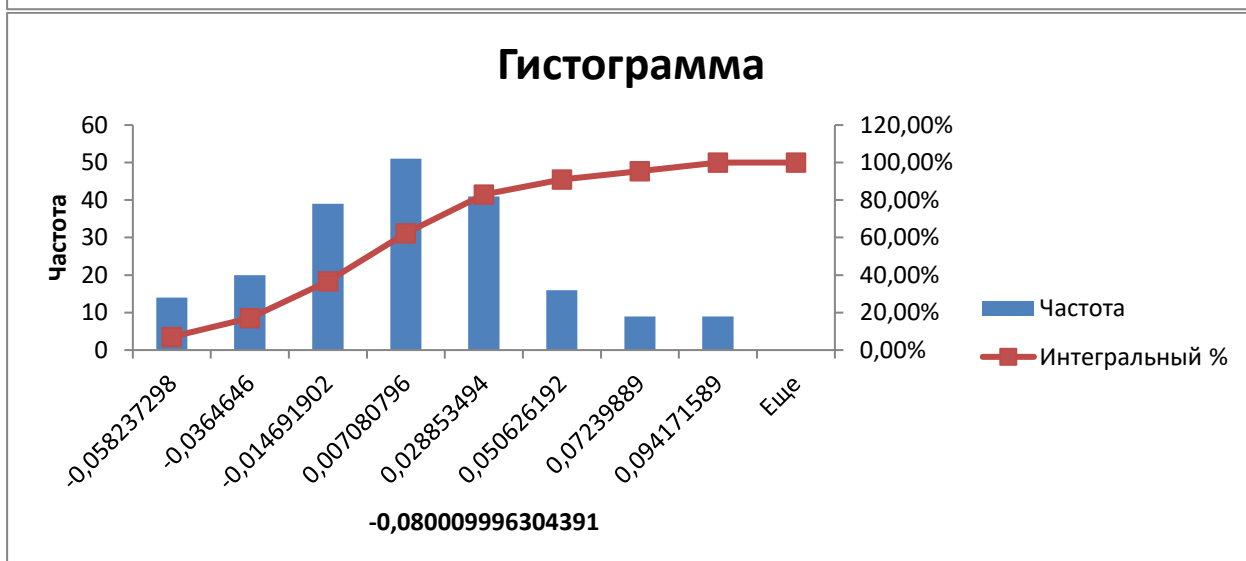
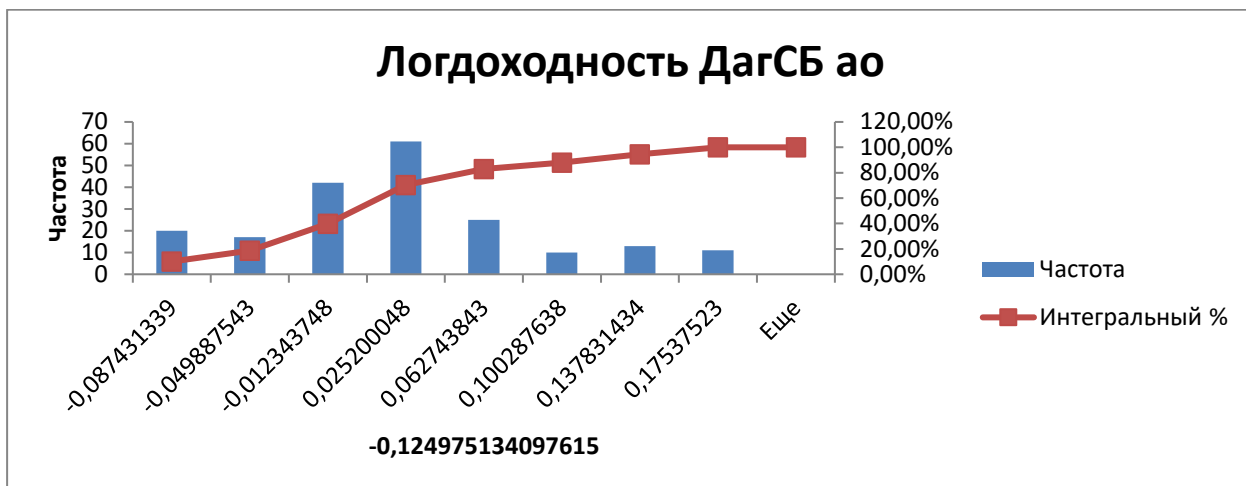
ВТБ ао		ДарСб ао		РОСИНТЕРао	
Среднее	-0,003877037	Среднее	0,00417262	Среднее	-0,002343035
Стандартная	0,001842022	Стандартная	0,004996106	Стандартная	0,002764207
Медиана	-0,003044142	Медиана	-0,002768442	Медиана	0
Мода	0	Мода	0	Мода	0
Стандартное	0,025984914	Стандартное	0,070478746	Стандартное	0,038993933
Дисперсия в	0,000675216	Дисперсия в	0,004967254	Дисперсия в	0,001520527
Эксцесс	0,114707985	Эксцесс	0,153476856	Эксцесс	0,354073631
Асимметрич	0,146735976	Асимметрич	0,307388142	Асимметрич	0,020095774
Интервал	0,125896915	Интервал	0,337894159	Интервал	0,195954282
Минимум	-0,063553784	Минимум	-0,162518929	Минимум	-0,101782694
Максимум	0,062343131	Максимум	0,175375229	Максимум	0,094171588
Сумма	-0,771530297	Сумма	0,830351419	Сумма	-0,466263986
Счет	199	Счет	199	Счет	199

Асимметрия и Эксцесс выборок уменьшились, также, как и стандартное отклонение, что говорит о более кучном распределении данных.



На диаграмме ящик с усами видно, что относительно очищенных данных появились новые, но уже не такие существенные выбросы, относительно которых мы уже можем дать более состоятельную оценку. Далее посмотрим на гистограммы интервальных частот исправленных логдоходностей

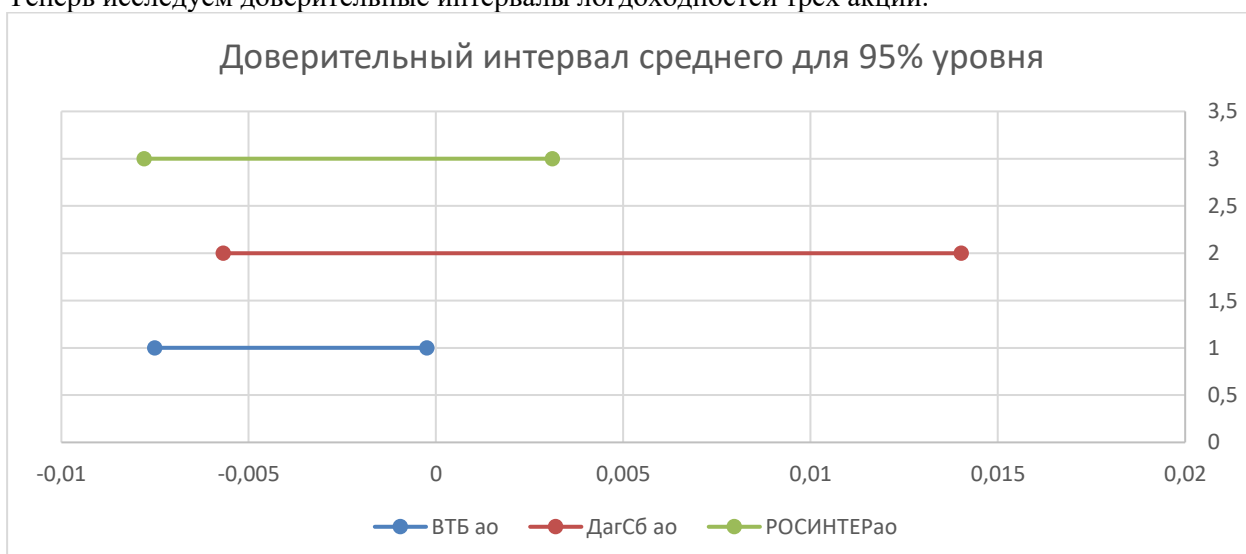




Здесь уже можно явно заметить, что логдоходности без выбросов для компаний ВТБ ао и РОСИНТЕРао вероятно имеют нормальное распределение.

5. Интервальные оценки параметров логдоходностей (в предположении нормального распределения), доверительные интервалы

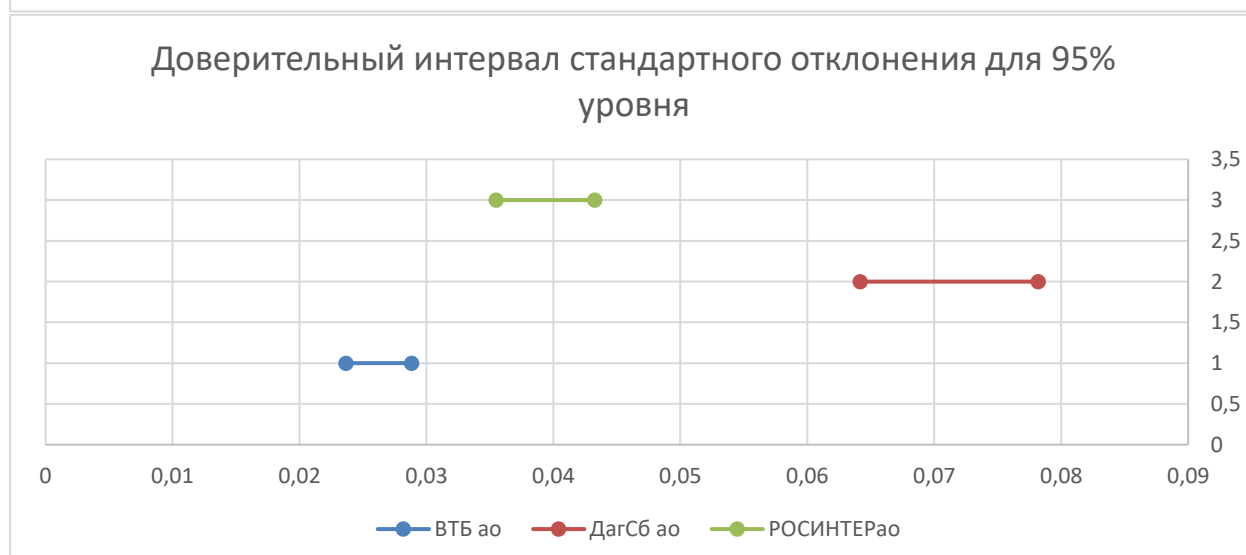
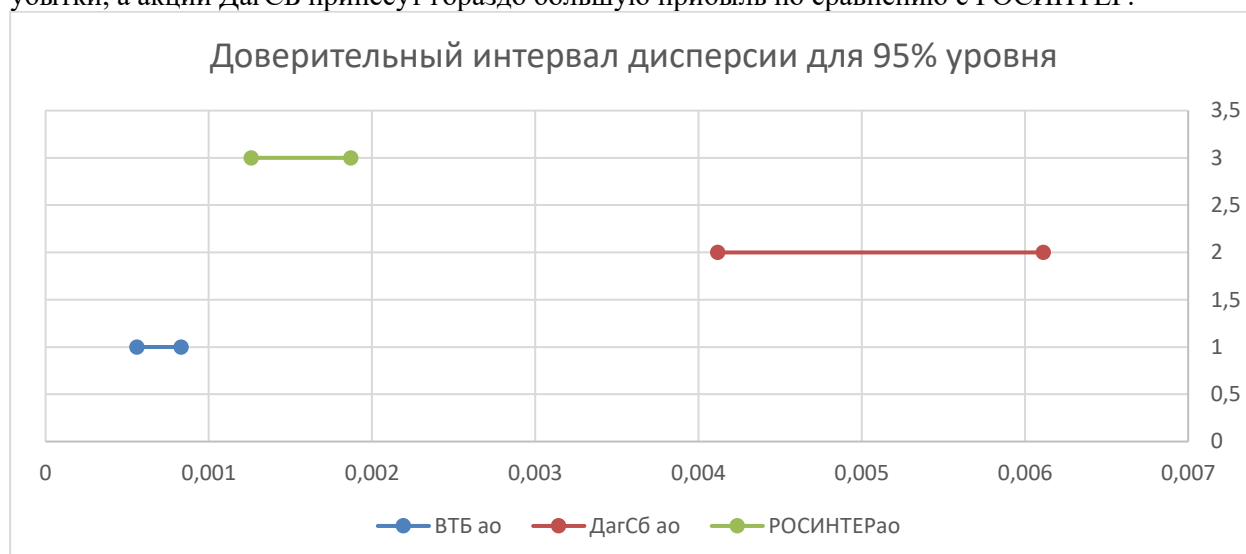
Теперь исследуем доверительные интервалы логдоходностей трех акций.



Здесь мы видим, что нижние границы интервалов почти совпадают у всех трех акций, но у ДагСБ она немного ближе к 0, верхняя же граница у ДагСБ значительно выше, нежели у ВТБ и РОСИНТЕР. Следовательно инвестиции в акции ВТБ и РОСИНТЕР в худшем случае принесут одинаковые убытки, но акции ДагСБ чуть меньше убытков. А вот в лучшем случае акции ДагСБ принесут гораздо больше

прибыли, нежели акции двух других компаний. Причем акции ВТБ вообще не принесут прибыли, а только убытки.

Следовательно с 95% ной вероятностью можно утверждать, что акции ВТБ в среднем принесут только убытки, а акции ДагСБ принесут гораздо большую прибыль по сравнению с РОСИНТЕР.



Здесь мы можем видеть, что разброс логдоходности акций ВТБ минимален, в целом, то же самое можно сказать и о акциях РОСИНТЕР, хотя по сравнению с акциями ВТБ он все же больше, а вот разброс логдоходности ДагСБ ао существенно отличается от ВТБ и РОСИНТЕР. Следовательно ДагСБ ао имеет менее стабильное колебание акций и с большей вероятностью может отклониться от среднего значения.

6. Проверка гипотезы о нормальности логдоходности по критерию Пирсона (хи-квадрат)

Для компании ВТБ ао:

Н₀: Нормальность присутствует

Н₁: распределение данных отлично от нормального

Альфа = 0,05

Сначала мы вычислили все необходимые параметры для выборки и разбили ее на 9 интервалов, посчитав для каждого частоту. Потом по методу моментов нашли необходимые параметры для нормального распределения, рассчитав выборочное среднее и выборочное стандартное отклонение, далее по формуле

$$\chi^2 = \sum_{j=1}^l \frac{(n_j - np_j)^2}{np_j},$$

Нашли сумму, то есть наблюдаемое значение 12,046.

Вычисляем критическую область для 95% вероятности по формуле $\chi^2_{0,95;9-1-2}$, так как 9

[illegible]

Для компании ДагСБ ао:

H1: распределение данных отлично от нормального

[illegible]

Следующая компания РОСИНТЕР ао.

H0: Нормальность присутствует

H1: распределение данных отлично от нормального

[illegible]

Здесь аналогично с предыдущей ситуацией, должно было быть бы 9 интервалов, но 1ый интервал содержал меньше 5 значений и слился со 2 интервалом в один. Рассчитанное значение 11,97 немного больше критического 11,075. То есть принимаем H1: распределение данных отлично от нормального с вероятностью 95%, но вот для 99% уровня наблюдаемое значение уже не попадает в критическую область, сл-но отвергнуть гипотезу с вероятностью 99% мы не можем.

7. Проверка гипотезы о нормальности логдоходности по критерию Колмогорова-Смирнова

Уровень значимости α	0,15	0,10	0,05	0,025	0,01
Порядок p квантиля	0,85	0,90	0,95	0,975	0,99
Квантили порядка p для $\sqrt{n}D_n$	1,138	1,224	1,358	1,480	1,626
Квантили порядка p для $\sqrt{n}D_n(\theta^*)$	0,775	0,819	0,895	0,995	1,035

Для компании ВТБ ао:

H0: Нормальность присутствует

H1: распределение данных отлично от нормального

Альфа = 0,05

Берем уже вычисленные в предыдущем примере левую, правую границу, середину интервала и частоту. Далее по частоте находим интегральный процент и переводим в доли – столбец функция распределения.

$$F_n(x) = \frac{1}{n} \sum_{X_k < x} k.$$

Считаем плотность по формуле

$$p_n(x) = \frac{ni}{n(x(\text{прав}) - x(\text{лев}))}$$

Далее находим НОРМ.РАСП для плотности точечная, для функции распределения – интегральная.

Находим модуль разницы функции распределения эмпирической и функции распределения нормированной, и берем супремум из найденных значений и умножаем на корень объема выборки.

$$\sqrt{n}D_n(\theta^*) = \sqrt{n} \sup_x |F_n(x) - F(x, \theta^*)|.$$

Найденное значение сравниваем с критическим, для 95% уровня критическое 0,895. Наблюдаемое 1,95, критическая область правосторонняя, следовательно принимаем гипотезу H1: распределение данных отлично от нормального.

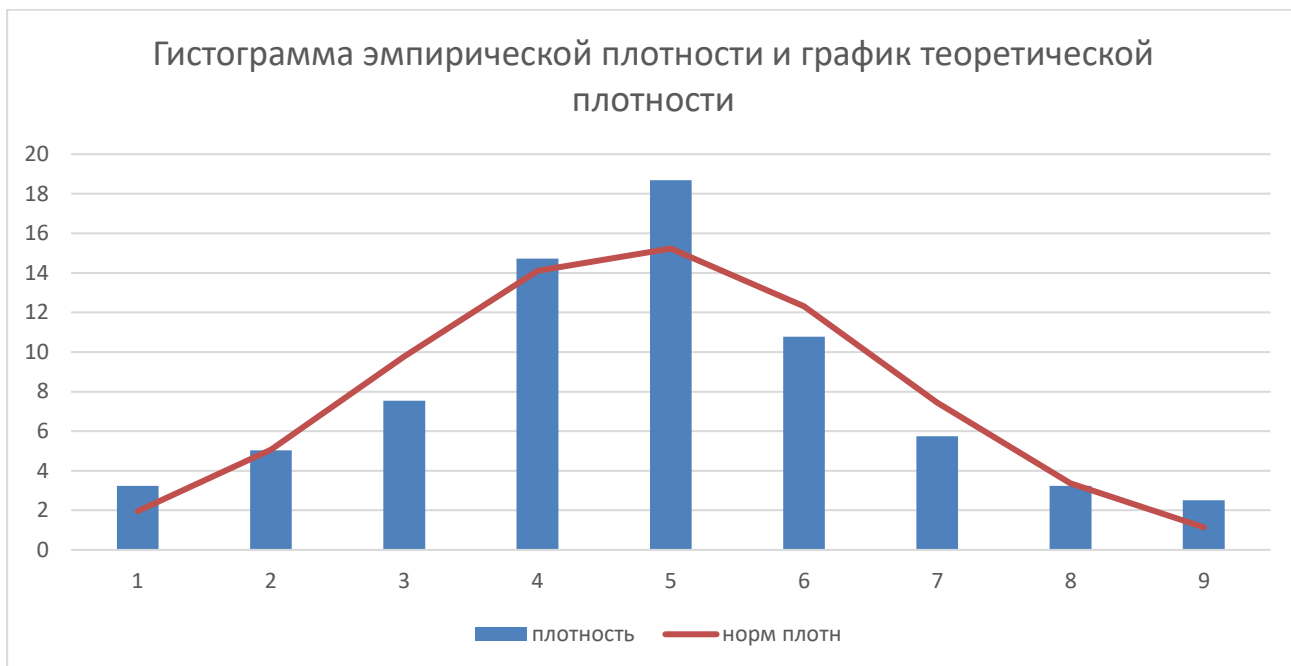
КОЛМОГОРОВА-СМИРНОВА										
левая	правая	середина	частота	Ф-ция распр	плотности	норм плотн	Ф-ция распр Норм	F_n(x)-F(x,θ^*)	sup=	
-0,06355	-0,04957	-0,05656	9	0,045226131	3,233083	1,966143013	0,021309464	0,023916666	0,138345	
-0,04957	-0,03558	-0,04257	14	0,115577889	5,02924	5,06626099	0,068231569	0,04734632		
-0,03558	-0,02159	-0,02858	21	0,221105528	7,54386	9,770128052	0,170863449	0,050242078	$\sqrt{n} D_n(\theta^*)$	
-0,02159	-0,0076	-0,01459	41	0,427135678	14,72849	14,10110659	0,340013857	0,087121822	1,95159	
-0,0076	0,006389	-0,00061	52	0,688442211	18,68004	15,23162985	0,550097642	0,138344569		
0,006389	0,020377	0,013383	30	0,83919598	10,77694	12,31345143	0,74673209	0,09246389	альфа 0,05	
0,020377	0,034366	0,027372	16	0,91959799	5,747703	7,449951892	0,885429539	0,034168451	0,895	
0,034366	0,048355	0,04136	9	0,964824121	3,233083	3,373396403	0,959150157	0,005673964		
0,048355	0,062343	0,055349	7	1	2,51462	1,143198193	0,988673642	0,011326358		
								1,951590308	>	0,895
Крит область правосторонняя, сл-но принимаем H1 с вер 95%										
ГИПОТЕЗА О НОРМАЛЬНОМ НЕ ПОДТВЕРДИЛАСЬ										

Также построены диаграммы:

- гистограмма эмпирической плотности и график теоретической плотности

- гистограмма эмпирической ФР и график теоретической ФР

По ним можно заметить, что распределение близко к теоретическому нормальному.



Аналогично предыдущему вычисляем наблюдаемое и критическое значение для компании ДагСБ ао. Для компании ДагСБ ао:

H_0 : Нормальность присутствует

H_1 : распределение данных отлично от нормального

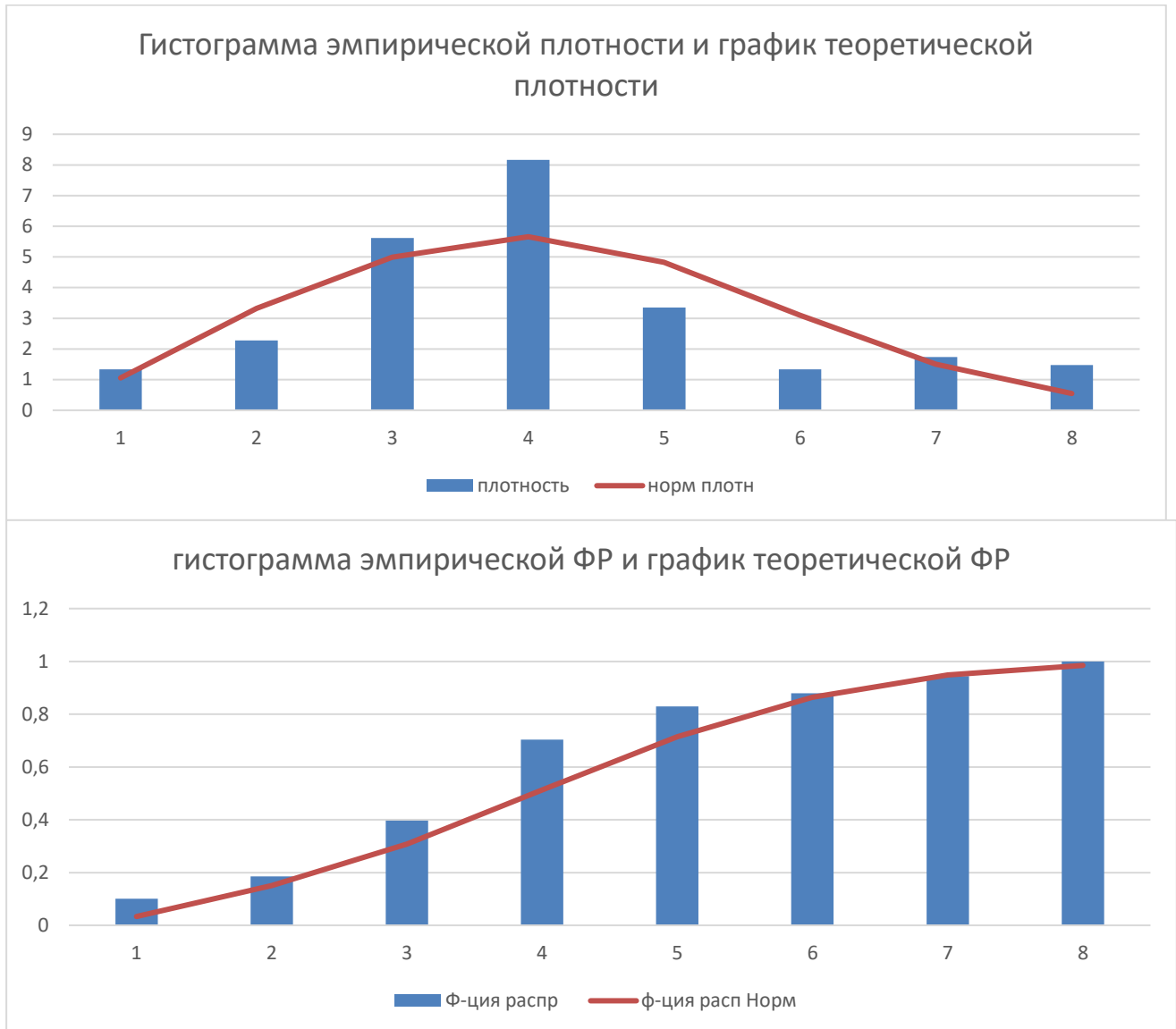
Альфа = 0,05

КОЛМОГОРОВА-СМИРНОВА									
левая	правая	середина	частота	Ф-ция рас	плотности	норм пло	Ф-ция рас	$ F_n(x) - F(x, \theta^*) $	sup=
-0,16252	-0,08743	-0,12498	20	0,100503	1,33847	1,056114	0,033443	0,067059	0,190752
-0,08743	-0,04989	-0,06866	17	0,18593	2,2754	3,318649	0,150711	0,035219	
-0,04989	-0,01234	-0,03112	42	0,396985	5,621575	4,993607	0,308293	0,088692	$\sqrt{n} D_n(\theta^*)$
-0,01234	0,0252	0,006428	61	0,703518	8,164669	5,657564	0,512765	0,190752	2,690894
0,0252	0,062744	0,043972	25	0,829146	3,346176	4,826216	0,713861	0,115285	
0,062744	0,100288	0,081516	10	0,879397	1,33847	3,099889	0,863766	0,015631	альфа 0,05
0,100288	0,137831	0,11906	13	0,944724	1,740011	1,499158	0,948459	0,003735	0,895
0,137831	0,175375	0,156603	11	1	1,472317	0,545897	0,984721	0,015279	
								2,690894	> 0,895
Крит область правосторонняя, сл-но принимаем H_1 с вер 95%									
ГИПОТЕЗА О НОРМАЛЬНОМ НЕ ПОДТВЕРДИЛАСЬ									

Здесь аналогично предыдущему принимаем H_1 : распределение данных отлично от нормального с вероятностью 95%.

Также построены диаграммы:

- гистограмма эмпирической плотности и график теоретической плотности
- гистограмма эмпирической ФР и график теоретической ФР



Здесь уже более заметно расхождение эмпирической функции и теоретической, больше отличается от нормального распределения.

Аналогично предыдущему вычисляем наблюдаемое и критическое значение для компании РОСИНТЕР ао.

Для компании РОСИНТЕР ао:

H0: Нормальность присутствует

H1: распределение данных отлично от нормального

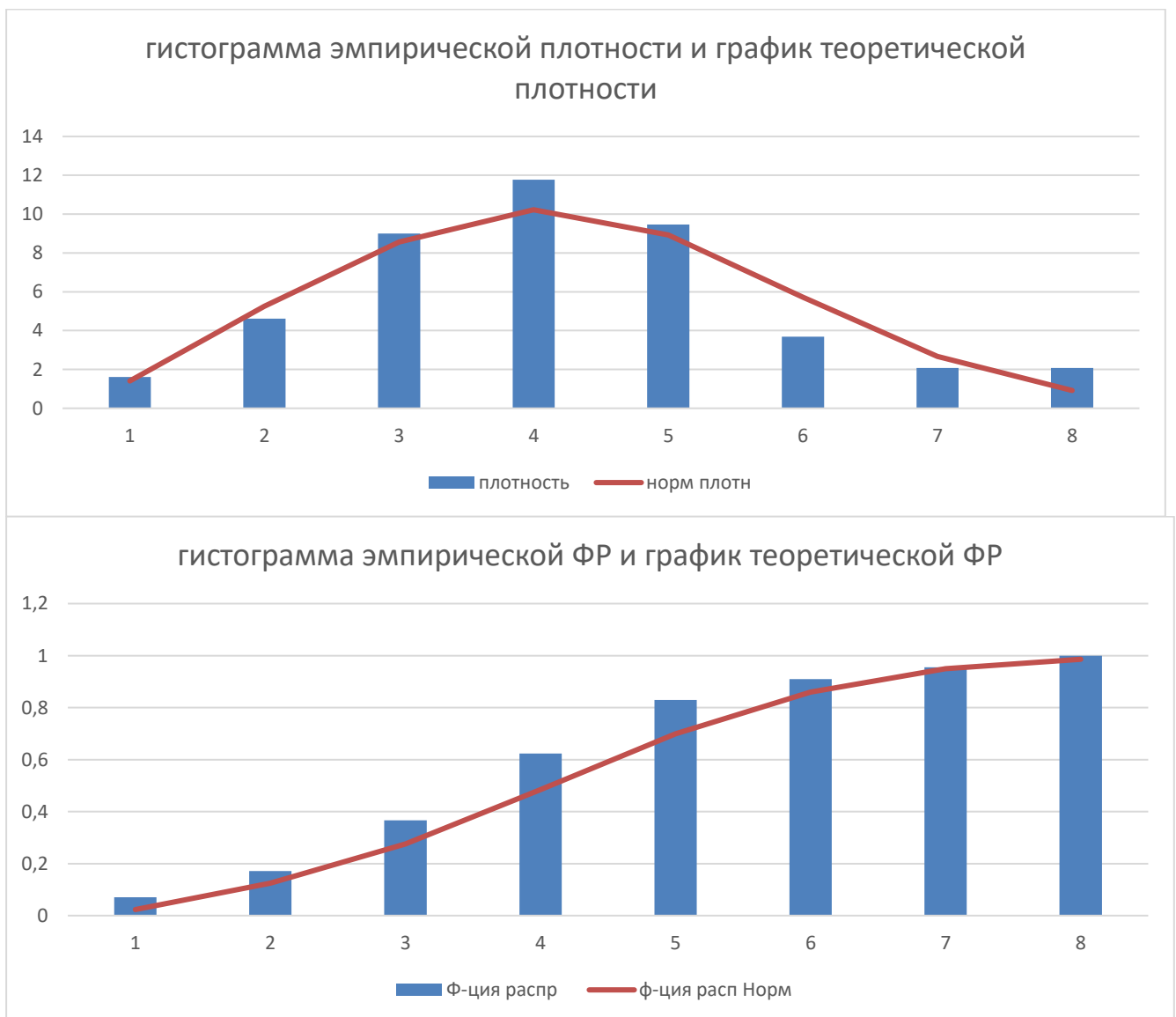
Альфа = 0,05

КОЛМОГОРОВА-СМИРНОВА									
левая	правая	середина	частота	Ф-ция рас	плотности	норм пло	Ф-ция рас	$ F_n(x) - F(x, \theta^*) $	sup=
-0,10178	-0,05824	-0,08001	14	0,070352	1,615596	1,407529	0,023198	0,047154	0,138075
-0,05824	-0,03646	-0,04735	20	0,170854	4,615988	5,255572	0,124203	0,046651	
-0,03646	-0,01469	-0,02558	39	0,366834	9,001177	8,566692	0,275632	0,091202	$\sqrt{n} D_n(\theta^*)$
-0,01469	0,007081	-0,00381	51	0,623116	11,77077	10,22369	0,485041	0,138075	1,947786
0,007081	0,028853	0,017967	41	0,829146	9,462775	8,933121	0,698766	0,13038	
0,028853	0,050626	0,03974	16	0,909548	3,69279	5,714787	0,859754	0,049793	альфа 0,05
0,050626	0,072399	0,061513	9	0,954774	2,077195	2,67669	0,949245	0,005529	0,895
0,072399	0,094172	0,083285	9	1	2,077195	0,917904	0,985952	0,014048	
								1,947786	> 0,895
									Крит область правосторонняя, сл-но принимаем Н1 с вер 95%
									ГИПОТЕЗА О НОРМАЛЬНОМ НЕ ПОДТВЕРДИЛАСЬ

Здесь аналогично предыдущему принимаем Н1: распределение данных отлично от нормального с вероятностью 95%.

Также построены диаграммы:

- гистограмма эмпирической плотности и график теоретической плотности
- гистограмма эмпирической ФР и график теоретической ФР



Распределение близко к нормальному.

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

Расчет теоретического значения производится по формуле

$$T = \frac{(\bar{X} - 0)\sqrt{n}}{S}$$

находится критическая точка распределения Стьюдента $t_{\alpha}(n-1)$, соответствующая уровню значимости α и числу степеней свободы $n-1$. Нулевая гипотеза принимается, если

$$T \leq t_{\alpha}(n-1)$$

и отклоняется при

$$T > t_{\alpha}(n-1).$$

Гипотезы сформулированы следующим образом

H0 : m=0

H1: m>0

В результате проверки получаем следующую таблицу.

Проверка равенства 0 матожидания			
	ВТБ ао	ДагСБ ао	РОСИНТЕР ао
наблюд	-2,104772484	0,8351745	-0,847633856
крит 5%	1,652585784	1,652585784	1,652585784
крит 1%	2,345328349	2,345328349	2,345328349

Для всех значений наблюденное меньше критического, следовательно везде принимаем H0: m=0.

Также данную проверку можно сделать при помощи Z.ТЕСТ

Z.ТЕСТ	0,982344441	0,201809733	0,801679043
--------	-------------	-------------	-------------

Все значения >0.1 , следовательно как и было уже сказано везде принимаем H0.

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

Гипотезы

H0: дисперсии по годам равны

H1: дисперсии по годам различаются

В диапазонах F-теста указываем 2018 и 2019 год – логдоходность.

Он возвращает значение p-value, которое мы сравниваем с $\alpha/2$. В результате сравнения получаем, что p-value для ВТБ и РОСИНТЕР больше $\alpha/2$, а вот для ДагСБ меньше $\alpha/2$. Следовательно для 1 и 3 компании принимаем H0 дисперсии равны по годам, а для ДагСБ H1: с вер 99% - дисперсии не равны по годам.

Равенство дисперсий за последний и предпоследний годы			
	ВТБ ао	ДагСБ ао	РОСИНТЕР ао
F.Тест	0,297894694	0,000001	0,780779613
крит 1%	0,005	0,005	0,005
крит 5%	0,025	0,025	0,025

На основе предположений о равенстве дисперсий проверяем равенство средних.

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

Гипотезы:

H0: средние по годам равны

H1: средние по годам не равны

Для 1 и 3 компании берем ТТЕСТ с типом 2 – дисперсии равны, а для ДагСб берем с типом 3 – дисперсии не равны на основании предыдущего пункта, получаем следующие p-value.

1	Равенство средних за последний и предпоследний годы			
2		ВТБ ао	ДагСб ао	РОСИНТЕРПао
3	ТТЕСТ	0,218802382	0,377516782	0,524906972

Все p-value > 0.1, поэтому принимаем H0: средние по годам равны. Во всех трех случаях.

11. Сравнительный анализ логдоходностей трех компаний

Будет применен однофакторный дисперсионный анализ.

Для компании ВТБ ао:

Гипотезы:

H0: m1=m2=m3=m4=m5

H1: средние не равны по годам

Для начала разделим данные по годам

	2015	2016	2017	2018	2019
	-0,04118	-0,05852	-0,04181	0,059082	0,016188
	0,020803	-0,03521	-0,02453	0,015026	0,052997
	0,001263	-0,02635	-0,01191	-0,02604	-0,00378
	-0,06349	-0,00457	0,018389	-0,00606	0,025511
	0,03501	0,034411	-0,0059	-0,00304	-0,02092
	-0,04855	0,026492	-0,01249	-0,04426	-0,01438
	0,030077	0,013761	-0,00733	0,013296	-0,01431
	0,013684	-0,05146	-0,01727	0,012673	-0,00139
	0,031391	0,011203	0,018279	-0,05103	-0,03788
	0,010166	0,002011	-0,01001	-0,00375	-0,00991
	-0.0083	0.008271	-0.02313	-0.03088	0.01876

Затем проведем дисперсионный анализ

Однофакторный дисперсионный анализ						
ИТОГИ						
	Группы	Счет	Сумма	Среднее	Дисперсия	
2015		31	-0,146229447	-0,004717079	0,000733742	
2016		41	-0,264538675	-0,006452163	0,000679559	
2017		38	-0,251447505	-0,00661704	0,000249772	
2018		42	-0,220299168	-0,005245218	0,000976259	
2019		47	0,110984498	0,002361372	0,000711807	
Дисперсионный анализ						
	Источник вариации	SS	df	MS	F	P-Значение F критическое
	Между группами	0,002486802	4	0,000621701	0,919241357	0,453786578 2,418202375
	Внутри групп	0,131205921	194	0,000676319		
	Итого	0,133692723	198			

Для обоснованности выводов по дисперсионному анализу сначала необходимо рассчитать критерий Бартлетта.

H0: дисперсии равны

H1: дисперсии не равны

ni	s^2i	ni-1	1/(ni-1)	(ni-1)*s^2	(ni-1)*s^2(ост)/si^2	общие	
31	0,000733742	30	0,033333333	0,022012257	-2,44476562	v-1=	4
41	0,000679559	40	0,025	0,027182359	-0,191155773	сумм(1/(ni-1))	0,13149
38	0,000249772	37	0,027027027	0,009241553	36,85635909	сумм(ni-1)	194
42	0,000976259	41	0,024390244	0,040026634	-15,04958957	q=	0,989582
47	0,000711807	46	0,02173913	0,032743119	-2,352512988	S^2(ост)	0,000676
						Φ=	16,64312
						крит 1%	13,2767
						крит 5%	9,487729

По формуле рассчитываем Φ

$$\Phi_{v-1} = q \left((n_1 - 1) \ln \frac{s_{ост}^2}{s_1^2} + (n_2 - 1) \ln \frac{s_{ост}^2}{s_2^2} + (n_3 - 1) \ln \frac{s_{ост}^2}{s_3^2} \right),$$

$$q = \frac{1}{1 + \frac{1}{3(v-1)} \left(\frac{1}{n_1 - 1} + \frac{1}{n_2 - 1} + \frac{1}{n_3 - 1} - \frac{1}{(n_1 - 1) + (n_2 - 1) + (n_3 - 1)} \right)},$$

$$s_{ост}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2}{(n_1 - 1) + (n_2 - 1) + (n_3 - 1)},$$

Находим Критическое значение Φ через ХИ2.ОБР. Получаем, что наблюдаемое попадает в критическую область (правосторонняя). Следовательно H1, гипотеза о равенстве дисперсий не подтверждается и Применение дисперсионного анализа необоснованно.

Но если его все же применить, то p-value>0.1, сл-но принимаем H0 о равенстве средних, то есть логдоходность не зависит от года.

Аналогично проверим для ДагСБ оа:

Гипотезы:

H0: m1=m2=m3=m4=m5

H1: средние не равны по годам

Данные по годам

	2015	2016	2017	2018	2019
	0,042605	0,171975	0,092373	0,110348	0,102279
	0,136414	-0,04402	-0,0482	0,155216	-0,05297
	0,001798	-0,0025	-0,05716	0,175375	0,0648
	-0,03463	-0,03304	-0,00656	-0,13162	-0,02261
	0,062392	-0,05306	-0,15635	-0,02477	-0,00201
	0,109855	-0,1154	-0,05535	-0,1304	-0,01376
	-0,00834	0,009132	-0,10717	-0,02166	-0,02517
	0,022815	0,012048	-0,11926	-0,09273	0,01781
	0,012526	-0,04595	-0,11778	0,056467	-0,00373
	0,028632	0,137442	0,154151	-0,04001	-0,00625
	0,132725	0,083819	0,117783	-0,06754	-0,01346
	0,101783	0,005013	-0,11778	0,081126	-0,02616
	0,067793	-0,07257	0,068993	-0,04426	0,014668
	0,04839	0,026527	-0,16252	0,004515	-0,00521
	0,013519	0,085265	0,057158	-0,01818	-0,01845
	-0,00277	-0,02598	0,120628	0,013668	-0,07643
	0,1692	-0,05129	-0,09531	-0,02288	-0,02716
	-0,02387	0,107574	0,04879	-0,00905	-0,00394

Дисперсионный анализ

Однофакторный дисперсионный анализ					
ИТОГИ					
Группы	Счет	Сумма	Среднее	Дисперсия	
2015	31	1,040041822	0,033549736	0,00323662	
2016	41	0,535281133	0,013055637	0,005978749	
2017	38	-0,166456358	-0,00438043	0,009005576	
2018	42	-0,514563644	-0,012251515	0,005302871	
2019	47	-0,063951533	-0,001360671	0,001111101	
Дисперсионный анализ					
Источник вариации	SS	df	MS	F	P-Значение F критическое
Между группами	0,045537176	4	0,011384294	2,354586749	0,055282587 2,418202375
Внутри групп	0,937979045	194	0,004834944		
Итого	0,983516221	198			

Критерий Бартлетта.

H0: дисперсии равны

H1: дисперсии не равны

ni	s ² i	ni-1	1/(ni-1)	(ni-1)*s ²	(ni-1)*s ² (ост)/si ²	общие	
31	0,00323662	30	0,033333333	0,097098602	12,04019556	v-1=	4
41	0,005978749	40	0,025	0,239149943	-8,493673031	сумм(1/(r	0,13149
38	0,009005576	37	0,027027027	0,333206303	-23,01305514	сумм(ni-1	194
42	0,005302871	41	0,024390244	0,217417718	-3,787537049	q=	0,989582
47	0,001111101	46	0,02173913	0,05110648	67,6475789	S ² (ост)	0,004835
						Ф=	43,93101
						крит 1%	13,2767
						крит 5%	9,487729
						Ф>крит	

Снова применение дисперсионного анализа не обосновано, но если применить, то p-value между 0,1 и 0,05, то есть отклоняем H0 с вероятностью 90%, и логдоходность зависит от года.

Аналогично проверим для РОСИНТЕРАо:

Гипотезы:

H0: m1=m2=m3=m4=m5

H1: средние не равны по годам

Разделим данные

2015	2016	2017	2018	2019
-0,0176	-0,00699	0,088193	-0,00402	0,041847
-0,03003	0,081917	-0,05609	0,001343	0,00545
0,046672	-0,06123	0,016349	-0,07524	-0,01506
-0,05784	-0,02299	0,031918	0,004332	0,015058
0,051462	0,006182	0,027116	-0,09993	0,00542
-0,00321	0	-0,03105	0,001591	-0,0613
0,04435	0,08124	0,02125	0,054151	0,0262

Далее дисперсионный анализ по этим данным

Однофакторный дисперсионный анализ						
ИТОГИ						
Группы	Счет	Сумма	Среднее	Дисперсия		
2015	31	-0,172257081	-0,005555668	0,001628124		
2016	41	0,100150448	0,002442694	0,001607153		
2017	38	0,120763181	0,003177978	0,001780159		
2018	42	-0,132253485	-0,003148893	0,001416604		
2019	47	-0,382667049	-0,008141852	0,001303536		
Дисперсионный анализ						
Источник вариации	SS	df	MS	F	P-Значение	F критическое
Между группами	0,004025195	4	0,001006299	0,657226407	0,622460429	2,418202375
Внутри групп	0,297039111	194	0,001531129			
Итого	0,301064305	198				

Определяем критерий Бартлетта

H0: дисперсии равны

H1: дисперсии не равны

ni	s ² i	ni-1	1/(ni-1)	(ni-1)*s ²	(ni-1)*s ² (ост)/si ²	общие	
31	0,001628124	30	0,033333333	0,048843724	-1,842685631	v-1=	4
41	0,001607153	40	0,025	0,06428611	-1,938339203	сумм(1/(r	0,13149
38	0,001780159	37	0,027027027	0,065865866	-5,575780445	сумм(ni-1	194
42	0,001416604	41	0,024390244	0,058080769	3,187467505	q=	0,989582
47	0,001303536	46	0,02173913	0,059962642	7,402564896	S ² (ост)	0,001531
						Ф=	1,220379
						крит 1%	13,2767
						крит 5%	9,487729
						Ф<крит	

В отличие от предыдущих компаний здесь наблюдаемое меньше критического, сл-но принимаем H0 и делаем вывод, что применение дисперсионного анализа обосновано.

По p-value>0.1 делаем вывод, что средние равны, принимаем H0, сл-но логдоходность не зависит от года.

Далее дисперсионный анализ логдоходности в зависимости от названия

Возьмем данные логдоходности по трем компаниям

H0: m1=m2=m3

H1: средние не равны по компаниям

ВТБ ао	ДагСб ао	РОСИНТЕР
-0,04118	0,042605	-0,0176
0,020803	0,136414	-0,03003
0,001263	0,001798	0,046672
-0,06349	-0,03463	-0,05784

Выполним для них дисперсионный анализ

Однофакторный дисперсионный анализ						
ИТОГИ						
Группы	Счет	Сумма	Среднее	Дисперсия		
ВТБ ао	199	-0,771530297	-0,003877037	0,000675216		
ДагСб ао	199	0,830351419	0,00417262	0,004967254		
РОСИНТЕРао	199	-0,466263986	-0,002343035	0,001520527		
Дисперсионный анализ						
Источник вариации	SS	df	MS	F	P-Значение	F критическое
Между группами	0,007270392	2	0,003635196	1,52248972	0,219018147	3,010891637
Внутри групп	1,418273249	594	0,002387665			
Итого	1,425543641	596				

Далее рассчитаем критерий Бартлетта

H0: дисперсии равны

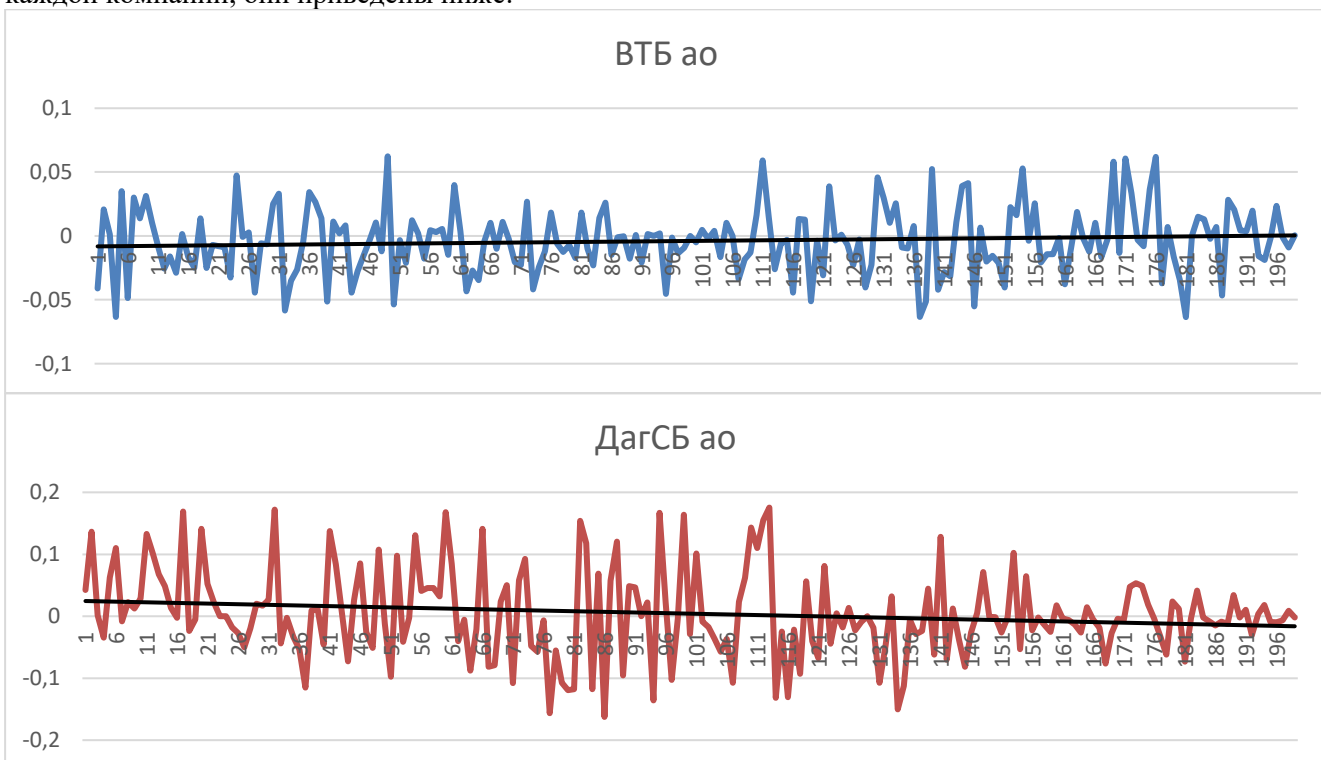
H1: дисперсии не равны

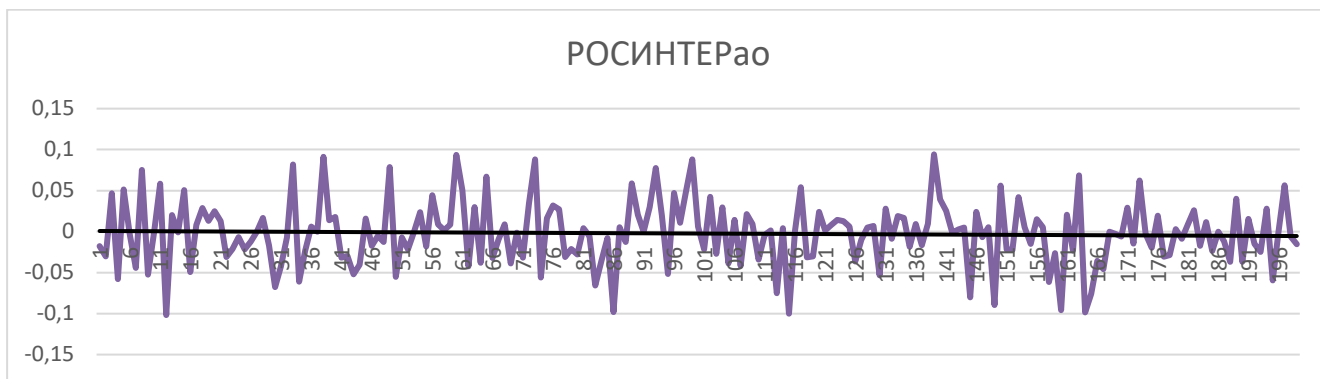
ni	s ² i	ni-1	1/(ni-1)	(ni-1)*s ²	(ni-1)*s ² (ост)/si ²	общие	
199	0,000675216	198	0,005050505	0,133692723	250,0817312	v-1=	2
199	0,004967254	198	0,005050505	0,983516221	-145,045104	сумм(1/(r	0,015152
199	0,001520527	198	0,005050505	0,301064305	89,34932543	сумм(ni-1	594
						q=	0,99776
						S ² (ост)	0,002388
						Ф=	193,9506
						крит 1%	9,21034
						крит 5%	5,991465
						Ф>крит	

Здесь принимаем H1: дисперсии не равны, сл-но применение дисперсионного анализа не обосновано, но если его применить, то:

p-value > 0.1, принимаем H0, средние равны, а значит логдоходность не зависит от компании.

Эти различия нагляднее будут отображены на графике зависимости логдоходности от времени для каждой компании, они приведены ниже:





12. Исследование тесноты связи между логдоходностями

Для этого исследования используем процедуру корреляции.

Выдвигается нулевая гипотеза

$H_0: r=0$,

при альтернативной гипотезе

$H_1: r \neq 0$.

Далее вычисляем t

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

где

t_{n-2} - распределение Стьюдента с $n-2$ степенями свободы.

Если $|t| > t_{кр}$, то гипотеза H_0 отклоняется, в противном случае – не отклоняется.

Найдем матрицу корреляции с помощью Корреляция-Анализ Данных

	<i>ВТБ ао</i>	<i>ДагСБ ао</i>	<i>РОСИНТЕРао</i>
<i>ВТБ ао</i>	1		
<i>ДагСБ ао</i>	0,050013471	1	
<i>РОСИНТЕР</i>	0,159906747	0,104673575	1

Далее найдем t для каждого и $t_{кр}$ для альфа 0,01 и 0,05

	r	t	t кр (0,05)	t кр (0,01)
ВТБ ао и ДагСБ ао	0,050013471	0,702852107	1,972079034	2,601015642
ВТБ ао и РОСИНТЕРао	0,159906747	2,273655258	1,972079034	2,601015642
ДагСБ ао и РОСИНТЕРао	0,104673575	1,477278877	1,972079034	2,601015642

Получаем, что для ВТБ и ДагСБ нет оснований отклонить гипотезу о наличии корреляции между логдоходностями компаний ни на уровне 0,05 ни на уровне 0,01

Для ВТБ и РОСИНТЕР есть основания отклонить гипотезу о наличии корреляции между логдоходностями на уровне значимости альфа = 0,05

Для ДагСБ и РОСИНТЕР нет оснований отклонить гипотезу о наличии корреляции между логдоходностями ни на уровне значимости альфа = 0,05 ни на альфа = 0,01