Федеральное государственное образовательное бюджетное учреждение

высшего образования

«Финансовый университет при Правительстве Российской Федерации»

Курсовая работа

на тему:

**«Проверка гипотезы о совпадении генеральных средних логарифмической доходности для всех дней недели»**

Вид исследуемых данных:

Котировки акций компаний из индекса Московской Биржи «ГАЗПРОМ ао», «Сбербанк», «ЛУКОЙЛ», «FIVE-гдр», «МТС-ао», «Роснефть», «Росбанк ао», «ДетскийМир», «Polymetal»

Выполнила:

студент группы ПМ20-6

Куприянова А. С.

Научный руководитель:

к.ф.- м.н., доцент

Департамент анализа данных,

принятия решений и финансовых технологий

Савинов Е. А.

Москва – 2022

Оглавление

ВВЕДЕНИЕ3

1. Теоретические основы5

1.1 Математическая статистика5

1.2 Проверка гипотезы в математической статистике5

1.3 Р-значение7

1.4 Однофакторный дисперсионный анализ8

1.5 Апостериорные значения11

1.5 Критерий Колмогорова11

1.6 Логарифмическая доходность12

2. Предварительный анализ данных13

3. Практическая часть20

3.1 Проверка гипотез на смоделированных данных20

3.2 Расчет мощности критерия 21

3.3 Проверка гипотез на реальных данных 22

ЗАКЛЮЧЕНИЕ26

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ27

ПРИЛОЖЕНИЕ28

Приложение 128

Приложение 2 (модельные данные)28

Приложение 3 (реальные данные)29

**ВВЕДЕНИЕ**

Данная работа была написана для того, чтобы разобраться (проверить гипотезу), совпадают ли генеральные средние логарифмической доходности в каждый день недели. Было принято решение проверять данную гипотезу на котировках акций индекса Московской Биржи, а именно: «ГАЗПРОМ ао», «Сбербанк», «ЛУКОЙЛ», «FIVE-гдр», «МТС-ао», «Роснефть», «Росбанк ао», «ДетскийМир», «Polymetal». Временной интервал был выбран практически случайно и составил дни с 01.01.2015 по 01.11.2022. Гипотеза проверяется методом дисперсионного анализа по критерию Колмогорова.

В ходе работы были рассмотрены теоретические аспекты, затронутые в дальнейшем в работе. Затем был проведен предварительный анализ данных, в ходе которой мы получили начальное представление о наших котировках. Практическая часть работы включает проверку гипотезу для смоделированных и фактических данных, представляющие собой реальные котировки акций за определенный период.

Нельзя не упомянуть и о крайней актуальности данного исследования, так как анализ генеральных средних полезен инвесторам и финансистам. Помимо этого, совпадение генеральных средних может говорить о схожести (или даже одинаковом происхождении) данных, которые вошли в нашу выборку, а следовательно, и генеральных совокупностей, откуда были эти выборки произведены.

Таким образом:

Цель курсовой работы: проверить совпадают ли генеральные средние логарифмической доходности в каждый день недели, основываясь на данные выборок.

Задачи:

1. Изучить необходимые для исследования теоретические основы
2. Получить необходимые данные
3. Сделать предварительный анализ полученных данных
4. Исследовать выполняемость поставленной гипотезы на модельных данных
5. Исследовать выполняемость поставленной гипотезы на реальных данных
6. Сделать выводы

Объект исследования: выборки из котировок акций индекса Московской Биржи, а именно: «ГАЗПРОМ ао», «Сбербанк», «ЛУКОЙЛ», «FIVE-гдр», «МТС-ао», «Роснефть», «Росбанк ао», «ДетскийМир», «Polymetal».

Предмет исследования: совпадение генеральных средние логарифмических доходностей.

1. **Теоретические основы**
   1. Математическая статистика

Математическая статистика – это наука об анализе данных, полученных в ходе проведения исследований или экспериментов, путем применения математических методов. Математическая статистика неразрывно связана с теорией вероятности, так как только так можно понять, насколько достоверными и надежными являются выводы, сделанные на ограниченном материале.

Таким образом, можно сказать, что в качестве цели данной науки рассматривается сбор и анализ-обработка данных (статистических), выводы которого можно использовать в дальнейшем. Например, в предсказывании результатов для других наборов данных, близких по происхождению к первоначальным.

* 1. Проверка гипотез в математической статистике

Стоит начать с того, что такое гипотеза вообще. Итак, чаще всего под этим понятием подразумевается предположение о каком-либо свойстве исследуемого объекта (например, набора данных, как в нашем случае). Здесь стоит отметить, что если мы делаем это предположение исходя из какой-то части/выборки большого массива данных (теоретическим языком, генеральной совокупности), то к слову гипотеза добавляется определение – статистическая.

Чаще всего работа производится с двумя вариантами гипотез:

1. Н0 – нулевая гипотеза, представляющая собой основное выдвигаемое предположение;
2. Нr (r = 1, 2, 3…) – антагонистические гипотезы, противоречащие нулевой.[[1]](#footnote-1) Иногда им присваивают название конкурирующих, а еще иногда альтернативных. На практике чаще всего рассматривают лишь одно такое предположение, называя его альтернативной гипотезой H1.

В качестве примера, можно рассмотреть случай, когда мы предполагаем, что рост всех учеников в классе находится в пределах (160, 170) см. Это будет наша нулевая гипотеза. А в качестве альтернативной гипотезы может выступить как просто отрицание, что рост учеников не попадает в эти границы. А можно сделать еще бесконечное количество альтернативных гипотез, например, что рост учеников заключается в пределах (158, 180) или (165, 175) и т.д.

Если говорить о самом принципе статистической гипотезы, то можно с уверенностью сказать, что он заключается в том, чтобы выдвинуть какое-либо предположение (гипотезу) относительно неизвестных параметров распределения, называемые параметрическими, или свойств. Вернувшись к нашему небольшому исследованию, можно сказать, что у нас H0 говорит о том, что все генеральные средние логарифмических доходностей для всех дней недели за определенный-случайный период совпадают, а в H1 заключается предположение, что этого совпадения не происходит, т.е. есть группы, у которых генеральные средние не равны.

Существует еще одно понятие – статистический критерий. Под ним подразумевается правило, что если мы имеем какую-то выборку (x1, x2 … xn) и она целиком принадлежит K, то гипотеза считается принятой. В обратном же случае мы делаем вывод об отвержении данного предположения.

Более формально мы имеем неравенство:

K = {(x1 … xn) ∈ Rn : t > c} или K = {(x1 … xn) ∈ Rn : t < c} или K = {(x1 … xn) ∈ Rn : t < c1} ∪ {(x1 … xn) ∈ Rn : t > c2}, где c, c1, c2 (c1 < c2) – константы, а t = t(x1 … xn) – статистика критерия.[[2]](#footnote-2)

Естественно, не всегда все идет по плану. Возникновение ошибок – распространенная практика. Некоторые из них получили свои названия. Ошибка первого рода – это случай, когда мы сказали, что гипотеза Н0 неверна, хотя на самом деле это вообще не так. И логически понятно, что под ошибкой второго рода, мы рассматриваем случай, когда мы отвергли верную Н1[[3]](#footnote-3).

Математики придумали еще одно понятие. Мощность критерия. Что под этим подразумевается? Начнем с обозначений: α – это вероятность того, что мы совершим ошибку первого рода (или она сама всплывет). β отвечает за вероятность совершения ошибки второго рода, соответственно. А под термином мощности критерия подразумевается как раз величина, которую можно посчитать как 1 – β[[4]](#footnote-4).

Возвращаясь к статистической гипотезе, стоит еще сказать, что проверка статистической гипотезы заключается в сравнении рассматриваемого (получившегося) значения нашего параметра (критерия) и критического значения (который был выбран или определен для исследования). А дальше мы уже можем делать выводы и принимать решение опровергать или не отвергать ту или иную гипотезу. Тут стоит еще отметить, что критическое значение у каждого уровня значимости свое.

Стоит отметить, что уровень значимости, а именно вероятность того, что мы отвергнем правильную нулевую гипотезу, подбирается исследователем самостоятельно и зачастую случайно. Самыми распространенным являются значения 0.1, 0.05 и т.д. Тут может возникнуть логичное предположение, почему бы не понижать наше α максимально, ведь чем меньше вероятность такой ошибки – тем лучше. Однако это не так банально, так как при уменьшении вероятности α растет вероятность β, влекущая за собой совершение ошибки второго рода, а тем самым и понижение мощности критерия. Так что выбор (или лучше сказать подбор) α и β происходит путем оценивания последствий, которые будут вызваны данным соотношением.

* 1. Р-значение.

Еще одно очень важное понятие – это Р-значение (от англ. Probability value). Под этим понятием подразумевается вероятность получить для данной вероятностной модели распределения значений случайной величины такое же или более экстремальное значение статистики (среднего арифметического, медианы и др.), по сравнению с ранее наблюдаемым, при условии, что нулевая гипотеза верна[[5]](#footnote-5).

Следовательно, чем ниже наше р-значение, тем более абсурдной кажется наша нулевая гипотеза. Поэтому при нахождении р-значения (используя различные функции Python например stats.f\_oneway, ktest и др) мы сравниваем его с нашим критическим уровнем значимости. И если р-значение меньше данного уровня, то мы отвергаем нашу нулевую гипотезу H0.

* 1. Однофакторный дисперсионный анализ

Одним из самых часто используемых статистических методов является дисперсионный анализ (ANOVA)*.* Его суть заключается в том, что анализ данных основывается на сравнивании средних подмножеств. Существует несколько вариаций такого метода. Самым часто применяемым и простым в понимании является односторонний (еще называют однофакторный) дисперсионный анализ ANOVA. Если обобщить, то можно сказать, что он представляет собой некое расширение t-теста, когда мы работаем с независимыми выборками или группами в количестве больше 2.

Можно также сказать, что цель ANOVA – это проанализировать различия между группами (путем сравнения средних) и определить, являются ли они (различия) статистически значимыми[[6]](#footnote-6).

С математической точки зрения:

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 1. ANOVA.

Основная идея данного анализа в том, что изменчивость набора данных разделяется на категории. Как пример, на изменчивость необъяснимую (еще можно сказать случайную), и изменчивость, которая вызвана систематическими различиями между рассматриваемыми выборками или группами данных. Вернувшись к нашей нулевой гипотезе H0 об одинаковости средних значений, можно ожидать, что одинаковыми будут и внутренние с межгрупповыми отклонения. Здесь можно также предположить, что если отличия между выборками будет превосходить отличия внутри них, то это можно считать, что мы имеем дело с систематическими различиями. Как следстви мы может создать тест, который базируется на сравнении этих отличий. Такое сравнение-соотношение имеет название F-статистика. Критические значения могут быть взяты из таблицы Fd. Но тут еще важно обратиться к понятию степеней свободы.

Всего существует (в распространенной практике) 2 вида степеней свободы. Разница лишь в том, что 1 вид характеризуется различиями (изменчивостью) между группами. Из этого следует, что df = m - 1. А второй вид – внутри каждой выборки. Здесь df = n – m, где m – количество выборок или групп данных, n – количество наблюдений/испытаний. Расчет нашей F-статистики происходит следующим способом: сначала нужно посчитать количество элементов-наблюдений в каждой выборке (n). Затем найти средне выборочное , сумму элементов выборки (T) и сумму элементов выборки в квадрате (S). Затем суммируем по группам каждую описанную выше величину[[7]](#footnote-7).

Для удобства приведена таблица 1, где рассмотрен случай, когда k = n. А в таблице 5 приведены расчеты для дисперсионного анализа для F-статистики. P-значение получается путем сравнения F для k-1 и N-k степеней свободы.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Группа | 1 | 2 | … | k | Итоговое значение |
| Кол-во наблюдений | n1 | n2 | … | nk | N = |
| Сумма наблюдений | T1 | T2 | … | Tk | T = |
| Среднее значение наблюдений |  |  | … | ₖ | = T/N |
| Сумма квадратов наблюдений | S1 | S2 | … | Sk | S = |

Таблица 1. Расчеты по k группам.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Сумма квадратов (SS) | Степени свободы (df) | Среднее квадратов (MS) | F-statistic (Коэффициент дисперсии) |
| Между групп | / -T2/N | k-1 | SS/df | MSbetween/MSwithin |
| Внутри групп | S - / | N-k | SS/df |  |
| Итог | S – T2/N | N-1 |  |  |

Таблица 2. Таблица дисперсионного анализа для получения статистики F.

Стоит отметить, что на данный момент не разработан такой способ, который позволит вычислять идеальный уровень значимости. Именно по этой причине принято не сравнивать критические и фактические значения, а использовать P-значения. Тем самым, под ними подразумевается максимальный уровень значимости, для которого гипотеза все еще применяется. Обратно тоже возможно. То есть мы берем P-значение и смотрим, отвергается или нет гипотеза при каждом уровне значимости.

Теперь необходимо ответить еще на один вопрос, когда же мы можем вообще применять наш дисперсионный анализ. Выделяют несколько обязательных условий:

1. Данные должны иметь количественный непрерывный тип данных.
2. Выборки должны быть независимыми. Чаще всего под этим пунктом подразумевается, что ни один человек, ни один исследуемый объект (значение котировки акции в конкретный момент и т.д.) не может присутствовать одновременно в нескольких выборках.
3. Нормальность распределения признака в статистических совокупностях, из которых в дальнейшем извлекаются выборки.
4. Дисперсии изучаемого признака в статистических совокупностях должны быть равны.
   1. Апостериорные значения.

Если мы получили, что наш дисперсионный анализ показывает, что различия между группами статистически значимые (например, если p-значение меньше выбранного порога значимости), то мы получили лишь факт наличия этих различий. Для того чтобы определить, какие именно группы различаются принято проводить апостериорные сравнения. Другими словами, парные сравнения.

* 1. Критерий Колмогорова

Часто для проверок гипотез на различные утверждения применяется критерий Колмогорова.

Рассмотрим небольшой пример, где мы выдвинули гипотезу H0 (например, что наше распределение подчиняется определенному закону). Применение нашего критерия будет выглядеть подобным образом:

Пусть H0 – это гипотеза, что случайная величина X имеет заданный закон распределения. Критерий Колмогорова включает в себя шаги[[8]](#footnote-8):

Изображение выглядит как текст

Автоматически созданное описание

Рисунок 2. Шаги критерия Колмогорова

Применять критерий Колмогорова имеет смысл только при большом объеме выборок. Например, от 25 штук (в некоторых источниках – 30 или даже 50).

* 1. Логарифмическая доходность

Доходность можно считать по-разному. Для упрощения вычислений часто берут логарифмы от всех составляющих нашей выборки изменения цен и тем самым мы получаем способ подсчета логарифмической доходности: или , где t – рассматриваемый период, - цена акции в данный период, - цена акции за прошлый период. Это гораздо удобнее, чем считать проценты (для обычной процентной доходности), так как работа происходит с одной лишь операцией вычитания (чаще всего).

1. **Предварительный анализ данных**

В основу исследования легли котировки акций компаний индекса Московской Биржи с официального сайта Финам: <https://www.finam.ru/>. Был выбран временной интервал с 01.01.2015 – по 01.01.2022.

|  |  |
| --- | --- |
| **Тикер** | **Название** |
| GAZP | ГАЗПРОМ ао |
| SBER | Сбербанк |
| LKOH | ЛУКОЙЛ |
| FIVE | FIVE-гдр |
| MTSS | МТС-ао |
| ROSN | Роснефть |
| ROSB | Росбанк ао |
| DSKY | ДетскийМир |
| POLY | Polymetal |

Таблица 3. Исследуемые данные

Первое с чего стоит начать – это отобразить количество дней, когда акции продавались, по годам. Затем на глаз оценим их распределение.

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 3. Кол-во дней торгов для каждого года из выбранных

На данных, представленных на таблице 2, мы можем заметить нулевые значения. Это связано с более поздним присоединением компаний «ДетскийМир» и «FIVE-гдр» в индекс МосБиржы. За неимением достоверной информации мы исключим данные по этим компаниям из исследования. Также можем заметить, что количество торговых дней компании «Росбанк ао» насовсем равномерно. Также исключим данную компанию из дальнейшего анализа.

Наша таблица примет вид:

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 4. Итоговое количество торговых дней по годам

Следующий показатель, который необходимо проанализировать – это скачки цен на годовых временных интервалах. Ниже приведены таблицы максимальных скачков - изменений цен по годам вверх и вниз.

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 5. Максимальные изменения - скачки цен вверх

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 6. Максимальные скачки цен вниз.

Для более плотного анализа посмотрим на графики распределения цен и графики логарифмической доходности, которые построим при помощи библиотеки plotly.

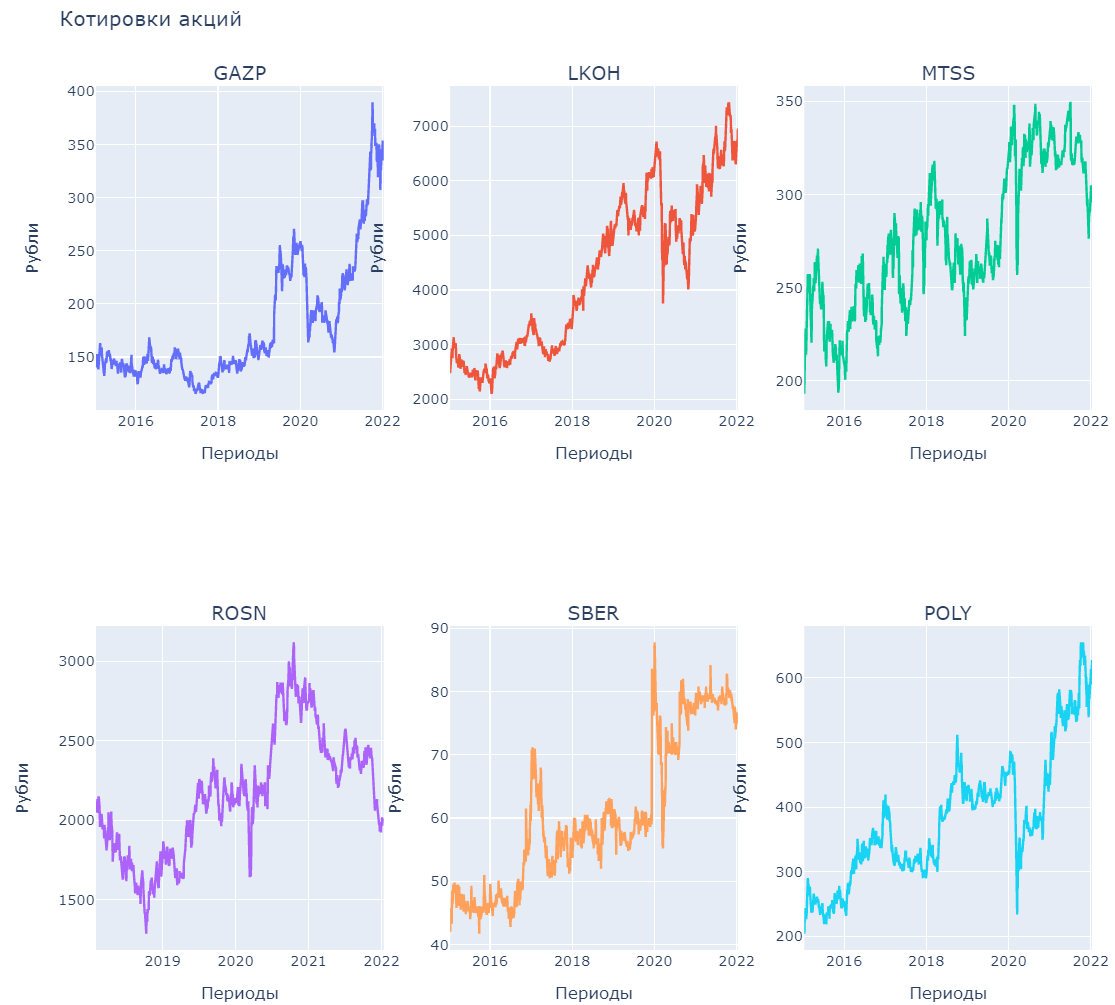


Рисунок 7. Цены по годам.

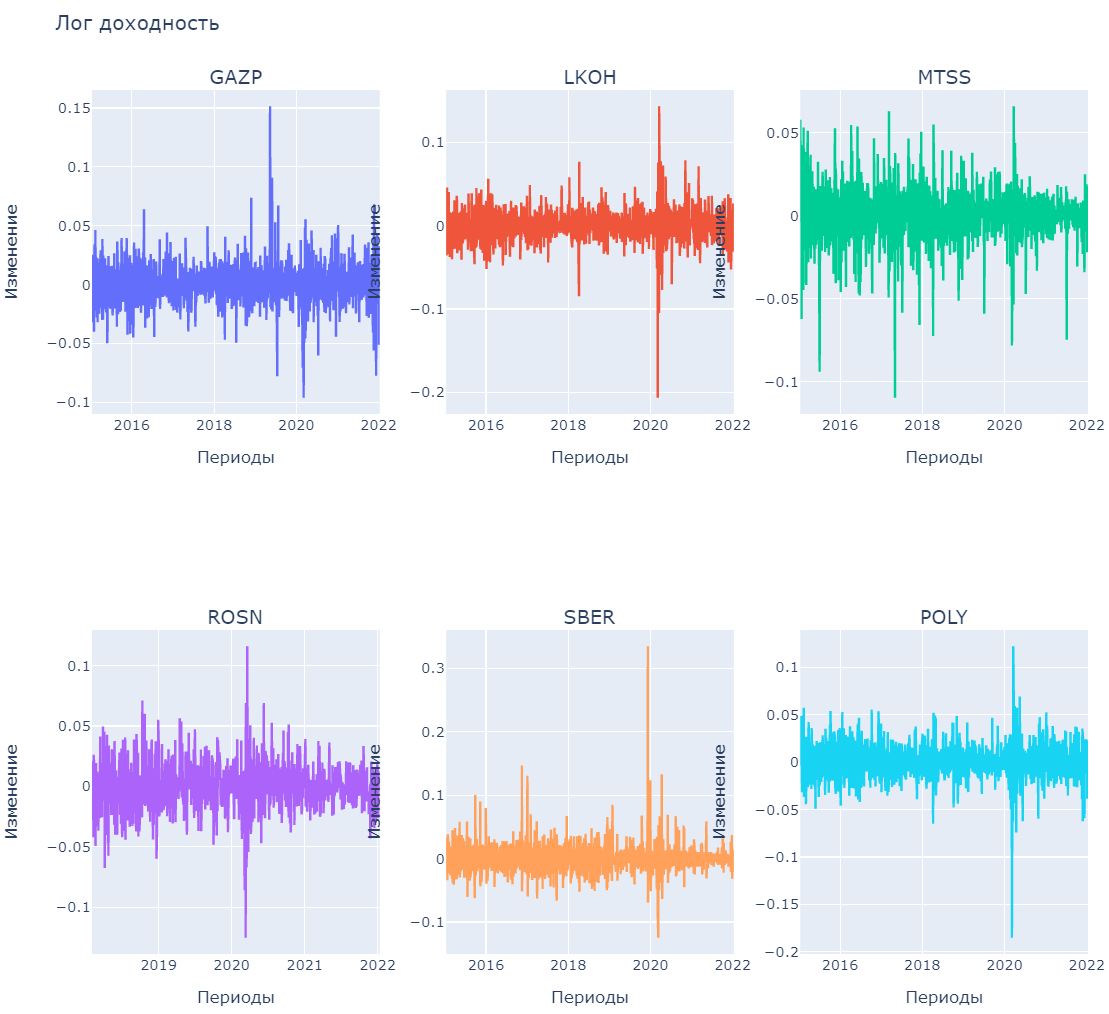


Рисунок 8. Логарифмическая доходность по годам.

На рисунках 5 и 6 видно, что максимальный скачок вверх составил 16% (компания Газпром в 2019 году), вниз – 18,6% (компания Лукойл в 2020 году).

Мы также видим, что 2020 год оказался очень богат на всевозможные выбросы, что можно связать с экономическими проблемами на фоне обострившейся ситуации с коронавирусом.

Однако выбросы находятся в пределах нормы, потому исключать какие-либо компании из рассмотрения не имеет смысла.

Итак, в основном этапе исследования примут участие 6 компаний:

|  |  |
| --- | --- |
| **Тикер** | **Название** |
| GAZP | ГАЗПРОМ ао |
| SBER | Сбербанк |
| LKOH | ЛУКОЙЛ |
| MTSS | МТС-ао |
| ROSN | Роснефть |
| POLY | Polymetal |

Таблица 4. Котировки акций, прошедшие в основное исследование

Также необходимо проверить нормальность распределения нашего результативного признака. Наш результативные признак – это логарифмическая доходность. Следовательно, мы должны проверить распределена ли наша логарифмическая доходность по нормальному распределению или нет для каждой компании. Сначала предлагаю прикинуть. То есть оценить чисто визуально. Для этого создадим столбец наших логарифмических доходностей в таблице каждой компании. Затем построим частотные гистограммы (для каждой компании):

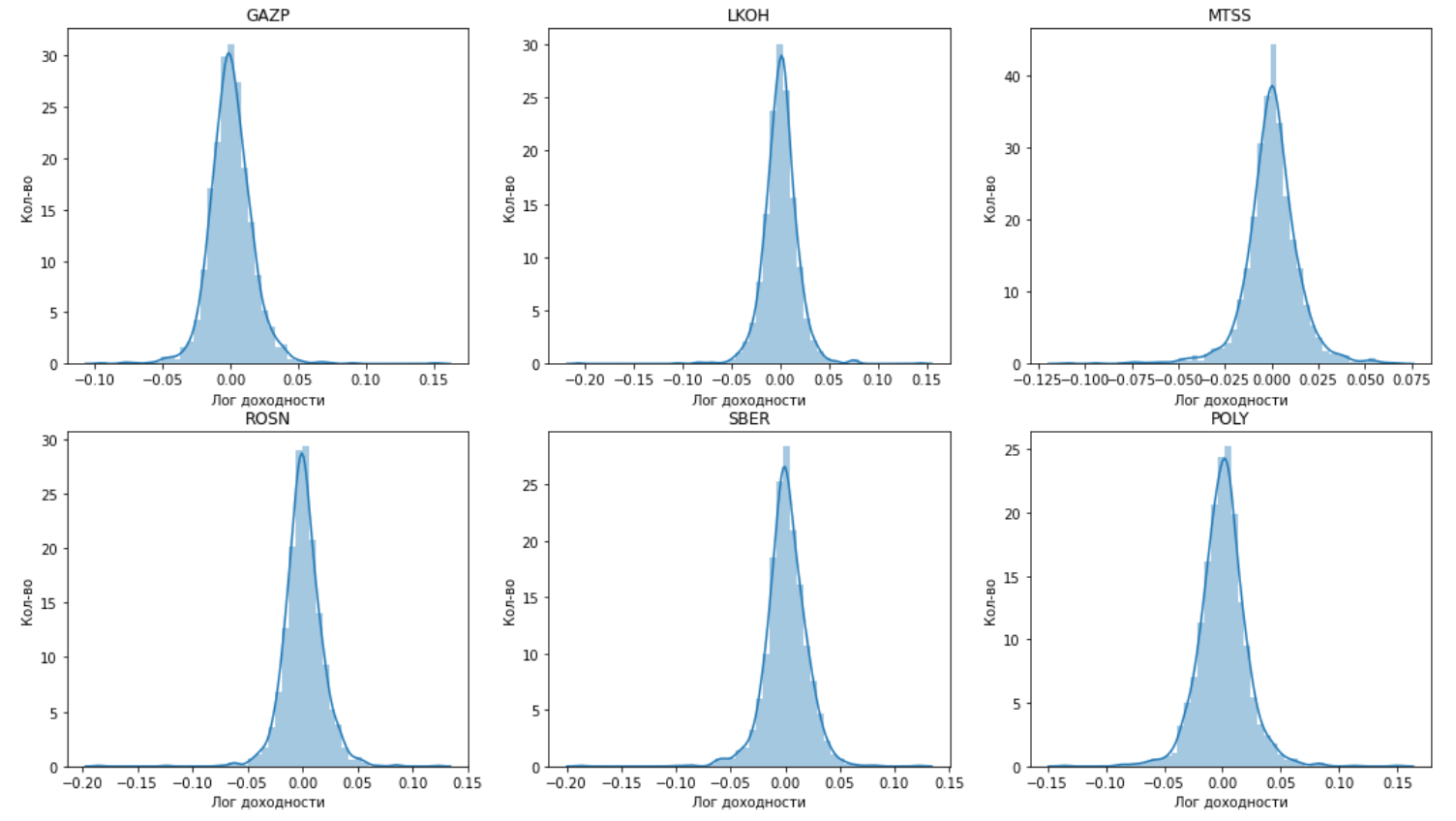


Рисунок 9. Распределение лог доходностей.

Действительно, все графики максимально приближены к графикам нормального распределения. Теперь проверим нормальность распределения функцией stats.normaltest из библиотеки scipy.stats. Получим результаты, подтверждающие наше предположение:

Изображение выглядит как текст

Автоматически созданное описание

Рисунок 10. Результат проверки нормальности распределения библиотекой Python.

Следовательно, и любые выборки, которые мы будем формировать из этих совокупностей, будут нормально распределены.

Следующий пункт, который необходимо проверить – это совпадение дисперсий у логарифмической доходности для каждого из дней недели. Для этого сгруппируем логарифмические доходности по дням недели для каждой компании и воспользуемся функцией stats.levene из библиотеки scipy.stats.

Изображение выглядит как текст

Автоматически созданное описание

Рисунок 11. P-значение проверки равенства дисперсий.

Наши p-значения довольно большие (больше 0.05), а следовательно, можно сделать вывод, что дисперсии равны.

Следовательно:

|  |  |
| --- | --- |
| Непрерывный тип данных выборок | Да. Следует из таблиц числа торговых дней. |
| Независимость выборок | Да. Выборки не пересекаются. |
| Нормальность распределения | Да. Следует из рисунков 8 и 9. |
| Гомогенность дисперсий | Да. Следует из рисунка 10. |

Таблица 5. Проверка условий.

Из таблицы 5 следует, что мы можем применять однофакторный дисперсионный анализ ANOVA.

1. **Практическая часть**
   1. . Проверка гипотез на смоделированных данных

Для того, чтобы проверить правильность работы нашей модели используем метод Монте-Карло (или еще можно сказать проведем эксперимент большое количество раз. В нашем случае 10000). Сам эксперимент заключается в том, что каждый раз (из 10000) генерируется выборка из нормального распределения (с параметрами 1 – среднее значение и 0 – среднеквадратичное отклонение) объемом 340 (примерно такое количество элементов исследуется в наших реальные данных) и находится р-значение. H0 гипотезой является наша гипотеза, что генеральные средние логарифмической доходности равны для каждой выборки. Альтернативной гипотезой является предположение, что среди всех выборок найдется хотя бы одна пара, у которой эти средние не равны. Для проверки данной гипотезы используется функция stats.f\_oneway из библиотеки scipy.stats.

Посмотрим на распределение р-значений, которые у нас получились. Напомню, что равномерность распределения р-значения говорит о том, что нулевая гипотеза верна.



Рисунок 12. Р-значение на модельных данных.

На рисунке видно, что наше р-значение распределено равномерно. Для статистического подтверждения этому посчитаем р-значение критерием Колмогорова.



Рисунок 13. Р-значение р-значений модельных данных – проверка на равномерность распределения.

Получаем р-значение больше 0.1 (и даже больше 0.05), а следовательно р-значения действительно имеют равномерное распределение. Это говорит о том, что наша нулевая гипотеза не отвергается, что и следовало ожидать, так как все наши выборки имели одинаковое математическое ожидание (1) и дисперсию (0). Следовательно, наша модель работает правильно.

* 1. . Расчет мощности критерия.

Для того, чтобы рассчитать мощность нашего критерия (ANOVA) воспользуемся алгоритмом:

1. Построим таблицу квантилей модельных данных (999 строк)

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 14. Таблица квантилей.

1. Найдем вероятность ошибки второго рода (H1) = β
2. По формуле 1- β найдем мощность критерия
3. Произведем этот анализ для нескольких уровней значимости (Верхняя строка – уровень значимости, нижняя – оценка мощности критерия).

Изображение выглядит как текст

Автоматически созданное описание

Рисунок 15. Оценка мощности критерия.

Интересно отметить, что на данном рисунке видно, как проявляется свойство несмещённости критерия, т. е. мощность во всех рассмотренных случаях не меньше ее верхней границы α.

* 1. Проверка гипотез на реальных данных

Для того, чтобы провести однофакторный дисперсионный анализ ANOVA воспользуемся функцией stats.f\_oneway из библиотеки scipy.stats (Напомню, что в качестве аргументов мы берем данные для каждого для недели, для каждой компании. То есть мы вычисляем нашу функцию 6 раз (так как в исследовании на данном этапе участвуют 6 компаний) и каждый раз подставляем в функцию 5 аргументов (списки логарифмических доходностей за понедельник, вторник, среду, четверг и пятницу). Получим следующие результаты:

Изображение выглядит как текст

Автоматически созданное описание

Рисунок 16. P-значения ANOVA для каждой компании.

Можно сразу заметить, что лишь 5 из 6 компаний прошли порог значимости 0.05. То есть можно сделать вывод, что генеральные средние у компаний с тикерами LKOH, MTSS, ROSN, SBER и POLY совпадают. А у компании с тикером GAZP – нет.

Для того чтобы проверить, в каких именно группах существенна между средними запустим пост-hoc тест, а именно тест Tukey HSD (Honest Significant Difference). Для этого импортируем pairwise\_tukeyhsd и MultiComparison из statsmodels.stats.multicomp. Для компании GASP получим следующие результаты:

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 17. Апостериорные сравнения GASP.

Мы видим, что критический порог в 0.05 не прошла пара Вторника и Четверга. Правда, стоит отметить, что до 0.05 не хватает всего 0.0136.

Для того чтобы убедиться в том, что подобной ситуации не возникло в других компаниях (и тем самым подтвердить результаты дисперсионного анализа), проведем апостериорные сравнения для остальных 5 компаний:

Изображение выглядит как стол

Автоматически созданное описание Изображение выглядит как стол

Автоматически созданное описание

Рисунок 18. Апостериорные сравнения для компаний LKOH, MTSS, ROSN, SBER и POLY.

Действительно, существенных различий найдено не было. Это подтверждает результаты дисперсионного анализа на рисунке 16.

**ЗАКЛЮЧЕНИЕ**

Проведя анализ наших данных, мы приходим к выводу, что нулевая гипотеза H0, говорящая о том, что генеральные средние логарифмических доходностей равны для каждого для недели не отвергается для всех компаний (GASP, LKOH, MTSS, ROSN, SBER и POLY), кроме пары четверга и вторника у компании “Газпром ао” (ей соответствует тиккер GASP).

Однако, если мы понизим критический порог до 0.01, то гипотеза будет принята для всех компаний и всех пар дней недели.

Если посмотреть на ситуацию бытовым взглядом, то наш вывод кажется довольно правдоподобным. Ведь если бы данные различия в генеральных средних логарифмических доходностей существовали, то умные инвесторы уже давно бы воспользовались этим фактом, тем самым вызвав изменения на рынке, и сгладили бы данные различия повышением (или понижением) спроса на те или иные акции в определенные дни недели, что привело бы к стабилизации рыночный цены акции и тем самым сглаживанию логарифмической доходности.

Цели достигнуты. Задачи выполнены.

**СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

**Литература**

1. Браилов А.В. Лекции по математической статистике. – М.: Финакадемия, 2007.
2. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. – М.: ФИЗМАТЛИТ, 2006.
3. Гмурман В.Е. Теория вероятностей и математическая статистика. Учеб. пособие для вузов. Изд. 9-е, стер. – М.: Высш. шк., 2003.

**Интернет-ресурсы**

1. Сайт Википедия: https://ru.wikipedia.org/wiki/P-значение

**ПРИЛОЖЕНИЕ**

Приложение 1.

Технические характеристики компьютера, на котором выполнена курсовая работа:

* Процессор: Intel Core i5 8265U 1.60 ГГц
* Ядро процессора: Whiskey Lake
* Количество ядер процессора: 4
* Объем кэша L2: 1 МБ
* Объем кэша L3: 6 МБ
* Память: 8 ГБ LPDDR3 2133 МГц

Приложение 2 (модельные данные).

Изображение выглядит как текст

Автоматически созданное описание

Изображение выглядит как текст

Автоматически созданное описание

Приложение 3 (реальные данные).

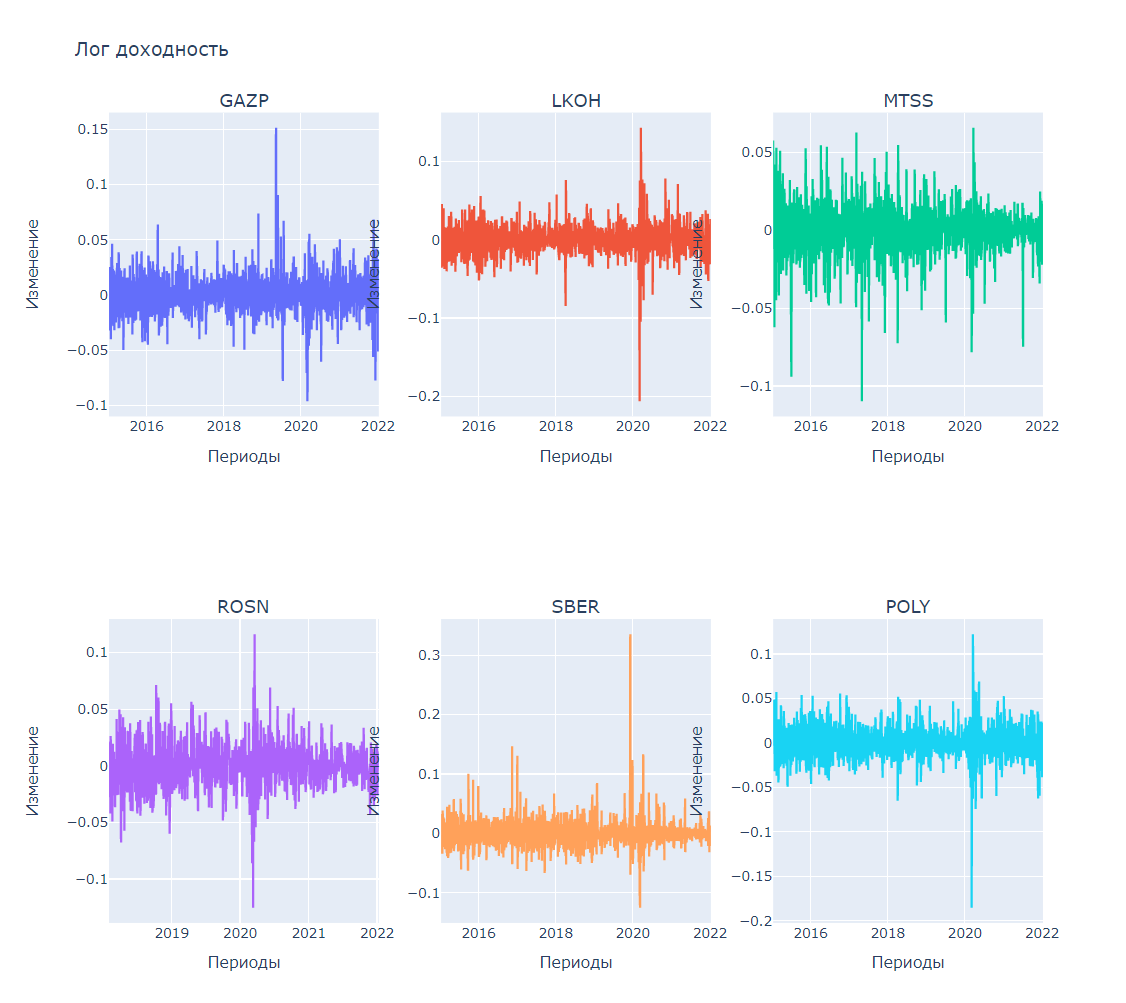
Изображение выглядит как стол

Автоматически созданное описание

Изображение выглядит как стол

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описаниеИзображение выглядит как стол

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описаниеИзображение выглядит как стол

Автоматически созданное описание

1. Браилов А.В. Лекции по математической статистике – С.80. [↑](#footnote-ref-1)
2. Браилов А.В. Лекции по математической статистике – С.81. [↑](#footnote-ref-2)
3. Браилов А.В. Лекции по математической статистике – С.82. [↑](#footnote-ref-3)
4. Браилов А.В. Лекции по математической статистике – С.82. [↑](#footnote-ref-4)
5. Сайт Википедия: https://ru.wikipedia.org/wiki/P-значение [↑](#footnote-ref-5)
6. Гмурман В.Е. Теория вероятностей и математическая статистика. Учеб. пособие для вузов. Изд. 9-е, стер. – М.: Высш. шк., 2003 – С.349. [↑](#footnote-ref-6)
7. Браилов А.В. Лекции по математической статистике – С.119. [↑](#footnote-ref-7)
8. Браилов А.В. Лекции по математической статистике – С.110. [↑](#footnote-ref-8)