Федеральное государственное автономное   
образовательное учреждение высшего образования

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ**

**«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

**Факультет экономических наук**

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА**

**Моделирование статистических свойств доходностей индексов высокодоходных облигаций эмитентов США как случайных величин**

образовательная программа «Экономика»

|  |  |
| --- | --- |
|  |  |
|  | Выполнил:  Студент БЭК199  Митерев Артём Святославович |
|  |
|  | Руководитель:  Доцент  Столяров Андрей Иванович |
|  |  |

Москва

2023

**Оглавление**

**Введение**

**Глава 1.** **Описание рынка, данных и методологии моделирования**

1.1. Рынок высокодоходных облигаций в США

1.2. Обзор литературы

1.3. Описание данных

1.4. Применяемые модели и статистические инструменты

1.4.1.Безусловные гомоскедастичные распределения

1.4.2. Условные гетероскедастичные модели

1.4.3. Безусловные гомоскедастичные модели для экстремальных значений

1.5. Оценка качества применяемых моделей

1.6. Гипотезы исследования

**Глава 2. Применение моделей**

2.1. Статистический обзор данных

2.2. Оценка параметров моделей

2.2.1. Безусловные гомоскедастичные распределения

2.2.2. Условные гетероскедастичные модели

2.2.3. Безусловные гомоскедастичные модели для экстремальных значений

2.3. Выводы по исходным гипотезам

2.4. Возможные практические применения

**Заключение**

**Введение**

Современный финансовый мир, включая коммерческие банки, управляющие активами, хедж-фонды, фонды прямых инвестиций и другие финансовые институты, на протяжении десятилетий в значительной степени полагался на несколько статистических моделей. Во-первых, с тех пор, как в 1952 году Гарри Марковиц представил Современную теорию портфеля (Modern Portfolio Theory, MPT), ее предпосылки и непосредственная модель стали ключевыми инструментами инвестиционных банкиров (Markowitz, 1952). Во-вторых, после того, как в 2004 году был введен стандарт Базель II, для банков и некоторых типов финансовых учреждений стала обязательной оценка риска с помощью показателя Value-at-Risk (VaR). В-третьих, оценка эффективности хедж-фондов и других типов альтернативных инвестиционных фондов часто основывается на таких показателях, как коэффициент Шарпа, который нацелен «усмирить» зависимость между риском инструмента и/или портфеля и его доходностью, таким образом «скорректировав» доходность на уровень риска (Sharpe, 1966). Общим фактором трех моделей в приведенных примерах является то, что базовое предположение о нормальности доходности активов в каком-либо виде часто делается или даже требуется моделью.

Удивителен тот факт, что вышеупомянутая нормальность распределения доходности финансовых активов была опровергнута еще в середине двадцатого века, и Бенуа Мандельброт был одним из первопроходцев в этой области (Mandelbrot, 1963).

Позже многие исследователи подтвердили его выводы и расширили тему распределения доходности активов с помощью различных классов теоретических распределений. Однако очевидно, что и по сей день предпосылки вышеупомянутых моделей все еще используются финансистами. Это служит свидетельством того, что альтернативные теоретические основы еще недостаточно полны, чтобы их мог легко применять широкий круг финансистов, инвесторов и финансовых учреждений, и недостаточно развиты, чтобы охватить все типы активов, доступные на рынке. В частности, пробел находится в сегменте облигаций на финансовом рынке, при этом существует мало исследований, изучающих статистическую природу распределения доходности различных типов облигаций, как это было сделано исследователями рынка акций.

В данной статье предпринимается попытка частично восполнить этот пробел, сосредоточив внимание на сегменте высокодоходных облигаций, который демонстрирует более высокую волатильность и более высокую доходность по сравнению с сегментом облигаций инвестиционного уровня и, следовательно, по своей природе ближе к должным образом исследованному рынку акций. Для этого в работе исследуется 3 статистические модели.

Во-первых, рассматривается принадлежность логарифомов доходностей индексов ВДО к безусловным гомоскедастичным распределениям – нормальному, Стьюдента и стабильному распределению Парето путем оценки параметров методом ММП и проведению двух тестов на качество подгонки (goodness-of-fit).

Во-вторых, путем оценки параметров и проведения статистических тестов проверяется примененимость модели ARMA-GARCH, служащей дополнением первой модели. В данной модели за ключевую случайную величину принимаются инновации авторегрессионного процесса, после чего также тестируется их принадлежность к трем описанным i.i.id. распределениям.

В-третьих, применяется инструментарий Теории экстремальных значений (EVT) для моделирования максимальных и минимальных дневных наблюдений за 3-, 6- и 12-месячные непересекающиеся окна. Для этого аргументируется выбор одного из трех распределений данной сферы статистики (Gumbel, Frechet, Weibull) и проверяется принадлежность выборки к выбранному распределению. Результирующие хвостовые индексы сравниваются между минимальными и максимальными значениями для выявления размера ассиметрии между этими распределениями краевых наблюдений.

Наконец, описываются способы и сферы применения полученных результатов для практических задач риск-менеджмента, включая расчет риск-метрик и стресс-тестирование моделей финансовых организаций. В частности, приводится и обсуждается пример расчета VaR и ES на основе стабильного распределения Парето.

**Глава 1. Описание рынка, данных и методологии моделирования**

В данной главе приводится описания рынка ВДО в США и используемых выборок данных, а также изучается сравнимая научная литература, в результате чего формируется теоретический инструментарий, который будет применяться в работе.

* 1. **Рынок высокодоходных облигаций в США**

Рынок высокодоходных облигаций в США имеет богатую историю, уходящую корнями в 1970-е годы. Еще раньше рынок существовал В это время инвестиционные банки начали выпускать долговые ценные бумаги для финансирования выкупа компаний с привлечением заемных средств (leveraged buyout, LBO). Этот тип финансирования позволял компаниям приобретать другие компании, используя долг, а не собственный капитал, увеличивая таким образом доходность на собственный вложенный капитал. Высокодоходные облигации, также известные как бросовые облигации, были привлекательным вариантом и для инвесторов, поскольку они предлагали высокую доходность в обмен на более высокий риск, связанный с кредитоспособностью заемщика.

Многие из самых известных корпораций США изначально финансировались именно за счет такого типа привлечения капитала. Например, в 1910-м году Уильям Дюрант, основатель General Motors, согласился выплатить кредиторам 7% за рефинансирование, которое спасло его компанию после череды приобретений, сдержавшим денежный поток и угрожавшему автомобильному гиганту банкротством. Буквально незадолго до этого, основатель J.P Morgan говорил, что ставка 5% сравнима с рисованием денег на Луне, проще говоря, инвестировать в инструменты с такой доходностью и выше, то же самое, что отправлять деньги в никуда, а значит и не получать их обратно. Видимо, не всегда это верно. Очень похожая ситуация в этом же году случилась с Чарли Флинтом, который не смог продать акции своей компании Computer Tabulating Recording. Вместо этого он разместил 6-ти процентных высокодоходных облигаций на сумму в 7 миллионов долларов. Через 14 лет, а именно в 1924 он сменил название своей компании на International Business Machine (IBM).

Рынок высокодоходных облигаций быстро рос в 1980-х годах, поскольку все больше компаний стремились финансировать LBO и другие сделки с использованием долга. За это время инвестиционные банки, такие как Drexel Burnham Lambert во главе с Майклом Милкеном, стали крупными игроками на данном рынке. В частности, Милкен помог этому банку запустить направление андеррайтинга подобных облигаций. Среди прочих, одно из главных достижений Майкла Милкена, что внесло его в огромное число рейтингов, среди которых “Пять человек, изменившие десятилетие 1980-х” или “Десять бизнес-лидеров, сформировавших столетие”, это демонстрация им факта того, что спреды между надежными и “мусорными” облигациями начали и продолжали увеличиваться с 1950-х годов, а значит выгоды от портфеля высокодоходных облигаций намного превосходят риски, вероятно, это знание и дало ему тот успех, которого он добился.

Причины такого роста весьма объективны. Конец 70-х, начало 80-х годов – это период связанный с резким ростов и взлетом стартапов в экономике США, так как стартапы зачастую не обладали доступом к более дешевым и традиционным способам финансирования, им приходилось выпускать высокодоходный долг. Относительное доверие к такого рода ценным бумагам возникало по конкретной причине. Как было упомянуто ранее, в этот момент рынок стартапов переживал хорошие времена, что приводило к тому, что немало таких компаний в итоге оказывались сверхприбыльными, а значит и платежеспособными. В конечном счете, такие успешные компании отказывались от такого рода финансирования в виду увеличения своего кредитного рейтинга и доверия со стороны кредиторов, но несмотря на это, именно широкий рост объемов рынка высокодоходный облигаций привел их к изначальному успеху (Yago, 1991).

Однако рынок высокодоходных облигаций пережил значительный спад в конце 1980-х годов, когда волна банкротств и дефолтов среди эмитентов бросовых облигаций привела к снижению доверия инвесторов. В частности, с 1980 по 1994 год было ликвидировано 1617 банковских организаций, державших в сумме активов на около 206,2 млрд долларов. Этот спад произошел неслучайно, у него были как внешние причины, так и внутренние. Важным фактором здесь стала изначально неблагоприятная атмосфера для многих отраслей, вызванная нефтяным кризисом 1973 г., который сильно подорвал отдельные отрасли американской промышленности. Стоит учитывать тот факт, что на тот момент в США действовало законодательство, которое во многом ставило банки в уязвимое положение. Банки в массе своей должны были быть локальными и привязаны к отдельным отраслям, а значит крах отрасли нередко приводил и к краху банку. Подобного рода опасения для банка означали, что одним из наиболее реалистичных выходов для него было взятие на себя более высоких рисков, однако закономерно не все могли совладать с этими рисками и не все могли вовремя от них отказаться, что сильно ударило по банкам во время рецессии, вынуждая их банкротиться (FDIC, 1997).

Позже, в 2008 году, произошел знаменитый кризис, вызванный выросшей популярностью продукта под названием CBO (collateralized bond obligation), включавший в себя бросовые облигации, но позиционировшийся как пакет облигаций класса AAA. Такая ситуация дала инвесторам ложное представление о том, что активы, в которые они инвестируют являются надежными, хотя во многом они были представлены ненадежными ипотечными долгами, или же “мусорными” облигациями. В результате, подавляющее большинство этих ценных бумаг не вернули вложенные в них деньги, что, в целом, нормальная ситуация для облигаций такого класса, однако, как было сказано ранее, инвесторы были убеждены, что это облигации с самым высоким рейтингом надежности. Это неизбежно привело к банкротству крупного инвестиционного банка Lehman Brothers, который закономерно потянул за собой всю финансовую систему США, а значит и почти всего мира, ведь глобальные компании, начавшие терять огромные средства на, казалось бы, надежных инвестициях, начали выводить свои активы из других стран, что стало детонатором уже глобального кризиса.

Несмотря на эти прецеденты, сегодня высокодоходные облигации являются важным источником финансирования для компаний ввиду их доступности, а для инвестиционных фондов – ввиду их доходности. Высокодоходные облигации выпускаются во многих странах, помимо США, и в разных валютах, а рынок стал более разнообразным с точки зрения типов эмитентов и инвесторов, которые способны вкладываться в этот инструмент. Исследование статистических свойств данного рынка, как следствие, представляет интерес для автора данной работы.

**1.2. Обзор литературы**

В научном мире широко освещается тема распределения доходности активов. Начиная с середины двадцатого века исследователи уделяют значительное внимание статистическим свойствам доходностей активов, подавляющая часть которых имеет толстые хвосты. В данном разделе приведены наиболее релевантные статьи, которые составляют основу для дальнейшего анализа, с описанием применимых методов и полученных результатов, которые будут применяться в последующем исследовании.

Одной из основополагающих статей по этому вопросу считается «The Variation of Certain Speculative Prices» (Mandelbrot, 1963), в которой автор показывает, как понятие нормальности распределения доходности акций, введенное Луисом Башелье (Bachelier, 1900) (сноски неправильно , решительно отвергается после изучения данных, накопленных с 1900 года. Автор утверждает, что остроконечность распределений слишком высока для того, чтобы изменения цен акций и некоторых товаров, таких как хлопок, могли быть смоделированы распределением Гаусса. Статистически этот вывод проявляется значимо более высоким показателем четвертого момента, или «эксцесса» (kurtosis). Вместо этого доказано, что семейство «стабильных паретианских» (Stable Pareto) распределений лучше подходит для данных логарифмической доходности фондового рынка.

Однако это не единственная статистическая конструкция, которая может быть применена к такой случайной величине. Существует целый раздел статистики «Теория экстремальных значений» (Extreme value theory), которая изучает свойства максимумов и минимумов значений случайных величин за определенный срок. Ряд исследователей-основоположников в EVT показывают, что максимумы и минимумы могут быть описаны лишь одним из трех классов распределений, что будет описано подробно в следующих разделах этой работы (Fréchet, 1927; Fisher, Tippett, 1928; Mises, 1936; Gnedenko, 1943).

В широко цитируемой работе Франсуа Лонгин применяет подход EVT на практике и показывает, что распределение экстремальных доходностей акций отличается от распределения «обычных» доходностей (Longin, 1996). Экстремальные доходности подчиняются распределению Фреше, описанному выше. В данной работе базовой переменной выбран логарифм дневного изменения доходности индекса наиболее торгуемых акций Нью-Йоркской биржи (NYSE) за период с 1885 по 1990 год. Далее автор изучает свойства максимумов и минимумов данных доходностей за периоды от 23 дней (месяц) до 559 дней (2 года).

Третий же тип подхода к моделированию был применен в статье «An empirical examination of daily stock return distributions for US stocks» (Rachev et al, 2005), где исследователи представили структуру, которая охватывает как безусловные гомоскедастические (предположение i.i.d. переменных), так и условные гетероскедастические модели с применением к широкому спектру акций. Авторы пришли к выводу, что гетероскедастические модели по типу ARMA-GARCH лучше описывают данные, что неудивительно ввиду феномена кластеризованной волатильности, которая свойственна многим финансовым рынкам.

В применении же к рынку облигаций авторегрессионный подход был применен в статье «Are Corporate Bond Market Returns Predictable?» (Hong, Lin, Wu, 2012). Здесь авторы показали наличие более высокой автокорреляции для корпоративных облигаций по сравнению с акциями. Кроме того, показана прямая зависимость между доходностью рынка акций и рынка облигаций, причем примечательно, что эта зависимость сильнее между акциями и высокодоходными облигациями, чем облигациями инвестиционных рейтингов. Это служит еще одним подтверждением того, что высокодоходные облигации представляют собой промежуточный класс между акциями и обычными долговыми инструментами в терминах своих статистических свойств. В данной работе эти выводы получены с помощью моделей векторной авторегрессии, чтобы рассмотреть зависимость *между* отдельными случайными величинами. В настоящей работе будет использован другой инструментарий, однако выводы, описанные выше, полезны для анализа.

Большинство последующих работ других авторов были так или иначе направлены на подтверждение и уточнение выводов, полученных в трех приведенных выше статьях. Например, Харрис и Кючюкёзмен проверили применимость семейства EGB (экспоненциальное обобщенное бета-распределение) и SGT (скошенный обобщенный Стьюдент) и обнаружили, что ежедневная доходность акций в США и Великобритании лучше описывается обобщенным и скошенным распределением Стьюдента (Harris & Küçüközmen, 2001).

Позже аналогичные подходы были применены и к рынку облигаций. В частности, Габриэль и Лау предложили использовать SGT и Pareto Stable распределения для моделировани дневных доходностей рисковых государственных облигаций Европейского Монетарного Союза (Gabriel, Lau, 2014). Однако для (предположительно) безрисковых государственных облигаций распределения Гаусса оказывается достаточно для описания данных за 1999–2011 годы. Для оценки параметров в этой работе использовалось максимальное правдоподобие (ММП), а качество соответствия проверялось по тестам Хи-квадрат, Колмогорова-Смирнова (KS), Крамера-вон-Мизеса (CM) и Андерсона-Дарлинга (AD). Полученный вывод интересен, поскольку показывает, что остроконечность распределения может убывать с уменьшением рискованности актива, что показывает различие акций и облигаций не на количественном (разница в оцененных параметрах), а на качественном уровне (принадлежность к семейству нетолстохвостых распределений).

В более недавней статье Вукович, Лапшина и Майти показывают, что распределение Гаусса неприменимо для нескольких небезрисковых государственных облигаций, но также показывают, что коэффициент Шарпа не является постоянным во времени для рынка облигаций. Этот вывод показывает на практике наличие более сложной статистической связи между риском и доходностью, чем та связь, на которую полагаются широко используемые статистические финансовые модели (Вукович, Лапшина, Майти, 2019).

Подводя итог, можно сказать, что тема распределений доходности акций широко освещена, но по-прежнему побуждает к созданию новых исследований, поскольку индустрия финансов и управления активами по-прежнему в значительной степени зависит от таких статистических моделей как MPT, коэффициент Шарпа и прочих структур, предполагающих нормальность распределения доходности активов. Однако вопрос распределения доходности облигаций является значительно менее изученной темой. Лишь несколько статей на эту тему проверяют применимость нескольких семейств распределений с предположением i.i.d. к долговому рынку, и в основном применяют эти распределения к государственным облигациям. Насколько известно автору, в научной сфере нет работ, в которых такой анализ применялся бы к индексам корпоративных облигаций в США (как наиболее зрелом рынке), и нет работ, в которых доходность корпоративных облигаций подгонялась бы к условным гетероскедастичным моделям и инструментарию Теории экстремальной значений.

Надо объяснить почему таких работ нет

**1.3. Описание данных**

В этой работе сочетаются несколько подходов, использованных авторами, указанными выше, для построения анализа в применении к рынку высокодоходных облигаций США.

Высокодоходный класс корпоративных облигаций был выбран, поскольку он служит промежуточным звеном между рискованными акциями, с одной стороны, и менее рискованными облигациями инвестиционного уровня и государственными облигациями, с другой стороны. Это позволяет более точно сравнить результаты данной работы с выводами соответствующих работ по акциям и государственным облигациям.

География изучаемых облигаций также выбрана не случайно. Важно понимать, что сама природа случайных величин с толстыми хвостами требует наличие большого числа наблюдений для получения сколь угодно репрезентативных выводов. Кроме того, стоит ожидать, что при повторении эксперимента через несколько лет полученные численные выводы (например, параметры распределений) будут отличаться от тех, что были получены в данной работе. Об этих свойствах тяжелохвостых величин предупреждает Нассим Талеб, иллюстрируя крайне медленную (или отсутствующую) сходимость метрик волатильности второго и четвертого моментов к некоторой константе в финансовых данных (Taleb, 2019).

По описанным выше причинам, для анализа был выбран рынок США, как представляющий наиболее развитый рынок высокодоходных облигаций, и имеющий длительную историю. Данный рынок, ввиду своей развитости, наименее подвержен структурным изменениям, свойственным развиавющихся странам, и посему статистические выводы, полученные в ходе анализа, могут быть более воспроизводимыми.

Данные представлены ежедневными и ежемесячными доходностями индексов высокодоходных облигаций США с различными кредитными рейтингами, предоставленными Bank of America (BB, Single-B, и CCC & Lower) с января 1997 г. по январь 2023 г. Используется несколько групп кредитных рейтингов, чтобы проверить, согласуются ли полученные выводы между рейтингами, и можно ли говорить о значимых различиях между ними.

В качестве переменных выбраны логарифмы дневных и месячных доходностей данных рейтингов:



где S(t) – значение индекса на момент t. Таким образом, переменная может принимать значения от –inf до +inf, при этом к каждому ее значению можно отнести доходность в классическом понимании S(t)/S(t-1)-1, которая не опустится ниже –100% при любом значении *x*.

**1.4. Применяемые модели и статистические инструменты**

По итогу рассмотрения описанных выше работ был сформулирован набор статистических инструментов, который будет использован для анализа доходностей высокодоходных облигаций в США.

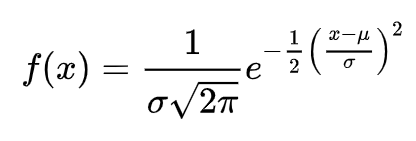
**1.4.1. Безусловные гомоскедастичные распределения**

Важно отметить, что в основе любой из рассматриваемых моделей лежит безусловное распределение, предполагающее одинаковость и независимость распределения случайных величин (independent and identically distributed variables, i.i.d.). Как отмечено в статьи Рэйчева и его коллег (Rachev et al, 2005), подобное распределение само по себе не может содержать все нюансы, которые бесспорно присутствуют на рынке, включая кластеризацию волатильности (эффекты ARCH), зависимость переменных во времени (эффекты ARMA) и зависимость самих параметров распределений от выбранного временного окна выборки данных. Тем не менее, наблюдения Мандельброта, в частности обозначенные им i.i.d. распределения, служат «строительными блоками», на которых могут быть построены дальнейшие, более сложные модели, которые будут описаны в следующих разделах.

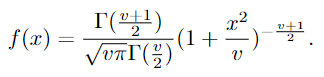
В данной работе в рамках первого подхода тестируется применимость распределений Гаусса (нормальное), Стьюдента и стабильного распределения Парето (также называемого распределением Levy Stable).

Для оценки параметров данных распределений применяется метод максимального правдоподобия (ММП), использованный в подавляющем большинстве упомянутых работ ввиду эффективности получаемых оценок.

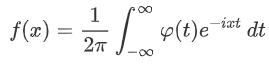
В случае нормального распределения оцениваются параметры положения и размерности (location & scale) функции распределения:

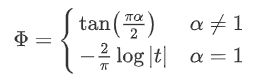


В случае распределения Стьюдента, помимо данных параметров также оценивается число степеней свободы (df, или *v*). Чем меньше данный параметр, тем тяжелее хвосты распределения. Функция распределения вероятности задается следующим образом:



Однако стабильное распределение Парето не может быть выражено в закрытой форме – его плотность выражена в виде следующей системы уравнений:



В данном случае помимо параметров положения и размерности оцениваются параметры альфа и бета, где альфа отвечает за тяжесть хвостов, а бета – за скошенность распределения.

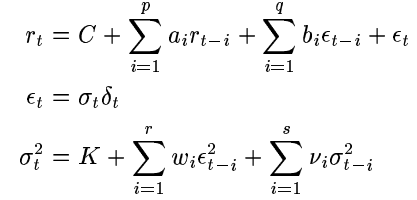
Стабильное распределение Парето представляет собой целый класс распределений, обладающих главным свойством: линейная комбинация двух величин, подчиняющихся стабильному распределению, также является стабильным распределением. Как подчеркивают Мандельброт, Рэйчев и Миттник (Mandelbrot, 1963; Rachev & Mittnik, 2000), это свойство крайне желательно для величин из финансового мира – например, для построения портфелей и моделей риск-менеджмента, обыкновенно состоящих из более чем одного инструмента, каждый из которых обладает своими статистическими свойствами (например, портфель из акций и облигаций). Примечательно, что кривая Гаусса является частным случаем этого распределения (при *alpha* = 2). Это означает, что в случае, если научный мир примет использование стабильного распределения Парето на практике, то значимую часть теории и инструментов, полагающиеся на нормальное распределение, можно будет перевести на новые рельсы без трудных изменений в остальных предпосылках.

Также важно, что по центральной предельной теореме сумма нормированных случайных величин с конечной дисперсией будет сходиться к нормальному распределению с увеличением числа наблюдений. Если же дисперсия не конечна, то в пределе подобная сумма будет являться стабильным, но не нормальным распределением. Для финансовых же данных не стоит поспешно делать предпосылку о конечности дисперсии. Как будет показано в Главе 3, значимая часть дисперсий финансовых величин определяется одним краевым наблюдением, а сходимость второго и четвертого моментов во времени работает слишком медленно в реальном мире (Taleb, 2019).

Упомянутые выше факторы делают данный класс распределений Парето весьма желательным для рассмотрения ввиду своей применимости и универсальности. Это также один из факторов в пользу выбора Парето-распределения вместо распределения Стьюдента при невозможности отвергнуть обе нулевые гипотезы (Tokat, Rachev, Schwartz, 2003).

**1.4.2. Условные гетероскедастичные модели**

Далее для каждого из наборов данных применяется авторегрессионная модель ARMA-GARCH, аналогично той, что была использована в статье Рэйчева и его коллег (Rachev et al, 2005):



С помощью ММП оцениваются параметры ARIMA (верхнее уравнение), а затем на остатках данной модели оцениваются параметры GARCH (нижнее уравнение). Стандартизированные ошибки , получаемые в результате, являются множителями из второго уравнения в указанной выше системе, и называются «инновациями» процесса ARMA-GARCH. По предположению модели, данные инновации сами по себе должны подчиняться некоторому i.i.d. распределению, поэтому в качестве следующего шага автором проводятся тесты на принадлежность остатков к тем же распределениям (нормальному, стьюдента, стабильного Парето).

Таким образом, в данной модели происходит изоляция эффектов временной волатильности рынков и автокорреляций от основополагающего распределения инноваций .

Для оценки качества условных моделей используются критерий AIC, равный



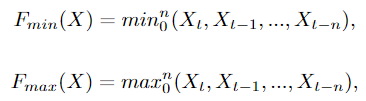
где L\* – значение максимизированной функции праводоподобия и *k* – число оцениваемых параметров, и критерий BIC, равный



где также учитывается размер выборки *n.* Наименьшее значение данных критериев соответствует наилучшей спецификации модели.

**1.4.3. Безусловные гомоскедастичные модели для экстремальных значений**

В рамках третьего подхода за исследуемые переменные берутся максимальные и минимальные значения дневной доходности индексов разных кредитных рейтингов за временные окна размером 3, 6 и 12 месяцев (используются только рабочие дни):



где *n={66, 132, 264}* дней

Далее для данных 9 векторов (3 индекса, соответствующих рейтингам BB, B и CCC и ниже) оцениваются параметры распределения Фреше (Frechet), имеющего следующую функцию распределения вероятности:

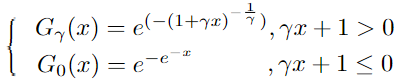


Где *alpha* – хвостовой параметр, а *m* и *s –* параметрыположения и размерности, как и в остальных распределениях (location & scale).

В дальнейшем в работе также потребуется определить кумулятивную функцию распределения Фреше:



В Теории экстремальных значений (EVT, extreme value theory), основные положения которой введены работами Frechet, Fisher, Tippett, Mises и Gnedenko (Fréchet, 1927; Fisher, Tippett, 1928; Mises, 1936; Gnedenko, 1943), основным результатом является теорма Фишера-Типетта-Гнеденко. Она утверждает, что максимум или минимум из выборки независимых одинаково распределенных переменных после должной нормализации в пределе может сходиться лишь к одному из трех распределений: Gumbel, Weibull и Frechet (иначе называемое обратным Weibull, т.е. лишь с заменой знака перед *x*). При группировке данные распределения можно записать в виде одной функции (Hans et al, 2007), называемой обобщенной функцией экстремальных значений (generalized extreme value distribution, GEV):



Примечательно, что максимумы значений, полученных из нормального распределения, сходятся к распределению Gumbel; для распределения Стьюдента и стабильного распределения Парето имеем сходимость к распределению Frechet (Longin, 1996). Таким образом, выводы касательно принадлежности Fmin(X) и Fmax(X) к одному из распределений EVT могут служить подтверждением принадлежности к некоторому распределению изначальной случайной величины. Более того, EVT применима не только для величин i.i.d. – ранее исследователями было показано, что если величина X подчиняется процессу ARCH, то Fmax(X) подчиняется распределению Frechet (De Haan et al, 1989). Как итог, для того, чтобы чрезмерно не расширять рамки работы, для анализа выбрано именно распределение Фреше, как соответствующее выводам Лонгина и соответствующее распределениям Стьюдента и Парето для i.i.d. величин. Также известно, что оценки параметров через ММП будут эффективными в нашем случае, то есть, когда все три параметра распределения Frechet неизвестны (Kotz, Nadarajah, 2000).

**1.5. Оценка качества применяемых моделей**

Для оценки качества трех описанных моделей применяются следующие методы:

1. Визуальное наблюдение

Простым визуальным сравнением гистограммы эмпирического распределения и функции распределения вероятности (pdf) предполагаемого теоретического распределения возможно увидеть несоответствие данных, например, нормальному распределению ввиду существенно более высокого пика. Однако визуально подходящие распределения могут отвергаться следующими методами ввиду того, что краевые значения, играющие большую роль, не всегда заметны на графиках.

1. Тест Колмогорова-Смирнова (KS-test)

KS-тест используется для определения того, относится ли выборка данных к некоторому распределению. Статистика считается по следующей формуле:

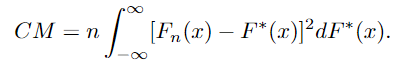


Где – эмпирическая кумулятивная функция распределения, и – кумулятивная функция проверяемого распределения. Другими словами, указанная статистика замеряет *максимальное расстояние* на графике между двумя этими функциями, то есть лишь между двумя точками. Важное следствие этого заключается в следующем: ввиду того, что, по определению, большее число наблюдений сосредоточено в центральной части распределения, наибольшее расстояние между и следует ожидать именно в этой части. Как итог, KS-тест высокочувствителен к различиям именно в значениях функций распределения в точках, близких к медиане, что делает его менее применимым в случае, когда исследователя интересуют распределения с тяжелыми хвостами. Как следствие, KS-тест в данной работе применяется индикативно – в дополнение к CM-тесту, описанному ниже.

Тем не менее, обозначим, что p-value в данном тесте представляет собой вероятность получения тестовой статистики по крайней мере такой же экстремальной, как и наблюдаемая тестовая статистика, при условии, что нулевая гипотеза (то есть гипотеза, что данные выборки получены из указанного распределения) верна. Таким образом, если p-value оказывается меньше выбранного уровня значимости, то нулевая гипотеза *H0* отклоняется и делается вывод в пользу *H1* (переменная не относится к указанному распределению). P-value в работе рассчитывается на основе алгоритма, предложенного Дурбином (Durbin, 1968) и реализованном в библиотеке scipy на языке программирования Python.

1. Тест Крамера-вон-Мизеса (CM-test)

В CM-тесте высчитывается интегрированное квадратичное отклонение между эмпирическим распределением Fn(x) и гипотетическим распределением F\*(x):



Данный тест был предложен в 1928–1931 годах (Cramer, 1928; Von Mises, 1931). P-value теста в работе рассчитывается на основе формулы приближения, указанной в статье «The Exact and Asymptotic Distribution of Cramér-von Mises Statistics» (Csorgo, Faraway, 1996) и реализованного в библиотеке scipy на языке программирования Python. Ввиду учета всего распределения, а не только наибольшего отклонения, а также большей чувствительности к хвостовым значениям, чем у KS-теста, CM-тест используется как первостепенный в данной работе.

Стоит отметить, что у перечисленных тестов есть проблема, заключающаяся в том, что критические значения их статистик изначально предназначались для случая, когда теоретическое распределение полностью определено (то есть его параметры известны). Однако на практике, включая случай данного исследования, параметры распределения оцениваются на тех же данных, на которых потом проводятся тесты. Это делает тесты чрезмерно консервативными (то есть они будут отвергать нулевую гипотезу реже, чем следует). Для учета этого фактора множество исследователей разрабатывали специальные таблицы и формулы для корректировки тестовых статистик для случая неполной спецификации распределения. Например, для AD-теста, частного случая CM-теста (Anderson, Darling, 1954), Стефенс, автор книг по оценке качества подгонки распределений (goodness-of-fit), приводит следующую скорректированную статистику для проверки распределения на нормальность (Stephens, 1974):

A\*2­ = A2 \*(1 + 4/n – 25/n2)

и следующую формулу для проверки экспоненциального распределения:

A\*2­ = A2 \*(1 + 0,6/n)

Таким образом, в случае данного исследования для дневной выборки коэффициент корректировки составил бы 1,00059 и 1,00009, а для месячной выборки – 1,0125 и 1,0019 для нормального и экспоненциального распределений, соответственно. Это крайне низкие значения корректировок. Примечательно, что описанная проблема даже не упоминается в большинстве описанной ранее литературы. По мнению автора, причина тому – чрезвычайно быстрая сходимости статистики *A\*2 к A2* при увеличении *n*. Для сохранения последовательности с научной литературой, описанной ранее, а также ввиду крайне низкого эффекта, в исследовании не применяются подобные корректировки.

**1.6. Гипотезы исследования**

**Гипотеза 1:** распределение логарифма дневной и месячной доходностей индексов высокодоходных облигаций (далее по тексту – «переменные») демонстрирует более высокий эксцесс по сравнению с нормальным распределением и не может считаться нормальным согласно статистическим тестам;

**Гипотеза 2:** дневныеи месячныепеременные могут, согласно статистическим тестам, быть смоделированы распределением Стьюдента или стабильным распределением Парето, как универсальными распределениями с повышенным эксцессом;

**Гипотеза 3:** инновациимодели ARMA-GARCH в применении к дневными месячнымпеременным могут, согласно статистическим тестам, быть смоделированы распределением Стьюдента или стабильным распределением Парето;

**Гипотеза 4:** инновации процесса при применении модели ARMA-GARCH демонстрируют менее толстые хвосты, чем исходные переменные, что отражается в более высоком числе степеней свободы в распределении Стьюдента и более высокой степени показателя *alpha* в стабильном распределении Парето (при оценке параметров с помощью ММП). В случае получения данного вывода будет получено обоснование использования условных гетероскедастичных моделей, в том числе для моделирования переменных в применении к VaR и ES;

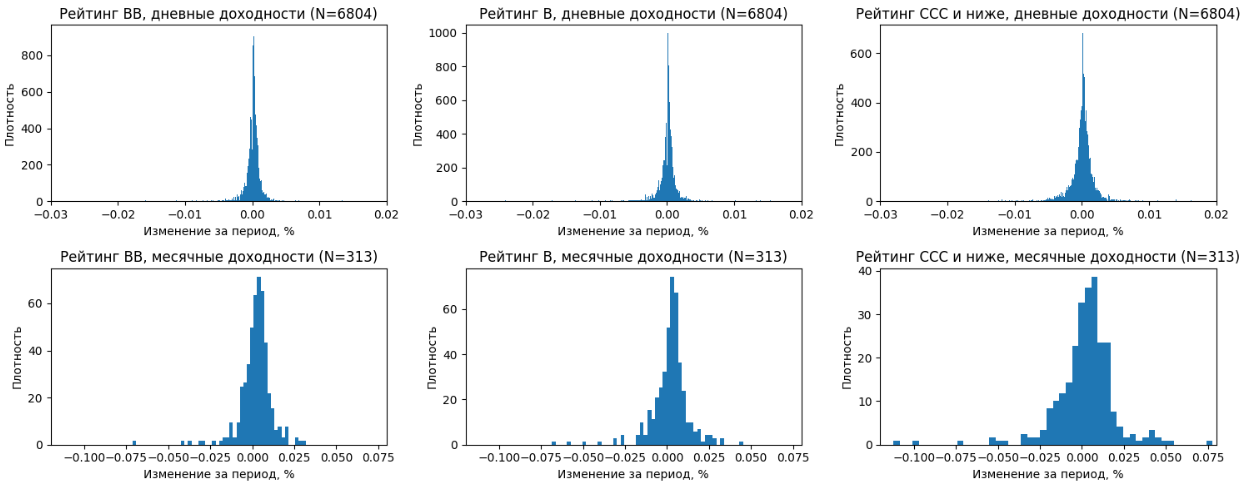
**Гипотеза 5:** экстремальные значения логарифмов доходностей за окна определенных размеров могут, согласно статистическим тестам, быть смоделированы одним из трех распределений Теории экстремальных значений.

**Глава 2. Применение моделей**

В данной главе приводится более подробное статистическое описание данных, и далее применяются описанные ранее методики по оценке параметров моделей и оценке их качества. Наконец, описываются возможные способы применения полученных результатов.

**2.1. Статистический обзор данных**

Для моделирования обозначенных переменных крайне важно изучить их статистические свойства. Выборка данных составляет по 6804 наблюдения для дневных и по 313 наблюдения для месячных доходностей для каждого из трех индексов. Визуальное рассмотрение эмпирических распределений логарифмов данных наблюдений показывает наличие экстремальных наблюдений, а также скошенность данных наблюдений влево, то есть в отрицательную область доходности.

*Рисунок 1: Эмпирические распределения логарифмов доходностей разных рейтингов*

Примечательно, что 15 из 18 крайних дневных наблюдений относится на 2008 и 2020 год, т.е. на года финансового кризиса и разгара пандемии COVID, из которых 9 наблюдения относится на 2008 год, и 6 – на 2020 год. Для месячных данных – лишь 4 из 14 крайних наблюдений относится на 2008, и также лишь 4 наблюдения – на 2009 год.

Также видно, что у индекса рейтинга CCC и ниже существенно больше отклонений, чем у индексов более высоких рейтингов. Например, у индекса рейтинга BB за промежуток 1996–2023 не было ни одного месячного наблюдения, превосходящего 10%, и было лишь одно наблюдение ниже, чем -10%. Таким образом, ВДО наиболее низкого рейтинга действительно имеют наивысший реализовавшийся риск в терминах разовых доходностей.

*Таблица 1. Крупнейшие дневные и месячные изменения индексов ВДО разных кредитных рейтингов. Для месячных данных – показаны все изменения больше 10,0% по модулю; для дневных – больше 3,0% по модулю.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Месяц** | **Рейтинг BB** | **Рейтинг B** | **Рейтинг CCC**  **и ниже** |  | **День** | **Рейтинг BB** | **Рейтинг B** | **Рейтинг CCC**  **и ниже** |
| 10.2008 | -15,1% | -14,6% | -22,8% |  | 10.10.2008 | -3,6% | -5,4% | -5,9% |
| 03.2020 | -9,3% | -12,5% | -20,2% |  | 18.03.2020 | -3,2% | -3,5% | -4,5% |
| 06.2002 | -8,1% | -6,1% | -10,8% |  | 16.03.2020 | -3,1% | -3,9% | -4,1% |
| 09.2008 | -7,1% | -8,6% | -10,1% |  | 12.03.2020 | -3,0% | -3,4% | -4,3% |
| 09.2001 | -5,0% | -7,1% | -11,7% |  | 09.03.2020 | -2,9% | -4,0% | -5,3% |
| 11.2008 | -3,5% | -10,7% | -15,0% |  | 15.10.2008 | -2,0% | -2,1% | -3,1% |
| 08.1998 | -2,3% | -6,1% | -11,3% |  | 14.09.2001 | -1,9% | -3,1% | -4,8% |
| 03.2016 | 3,3% | 4,0% | 10,2% |  | 20.11.2008 | -1,9% | -2,7% | -3,1% |
| 04.2003 | 3,9% | 5,6% | 11,5% |  | 02.10.2008 | -1,4% | -2,7% | -3,4% |
| 09.2009 | 4,4% | 4,9% | 10,1% |  | 06.10.2008 | -1,2% | -2,3% | -3,1% |
| 07.2009 | 4,6% | 4,7% | 10,8% |  | 12.12.2008 | -0,3% | -0,9% | -3,2% |
| 05.2009 | 4,9% | 6,1% | 12,7% |  | 26.12.2008 | 0,1% | 0,3% | 4,2% |
| 12.2008 | 6,4% | 6,7% | 11,5% |  | 30.04.2009 | 0,9% | 1,4% | 3,3% |
| 04.2009 | 7,6% | 11,0% | 19,6% |  | 30.12.2008 | 1,0% | 2,5% | 3,8% |
|  |  |  |  |  | 14.10.2008 | 1,6% | 3,1% | 3,4% |
|  |  |  |  |  | 06.01.2009 | 2,5% | 2,8% | 3,3% |
|  |  |  |  |  | 09.04.2020 | 3,0% | 3,3% | 3,8% |
|  |  |  |  |  | 26.03.2020 | 3,1% | 3,6% | 3,8% |

Средний показатель доходности схож между рейтингами и не позволяет говорить о наличии объяснимой зависимости – так, доходность рейтинга BB выше, чем рейтинга «CCC и ниже», несмотря на более высокий риск последнего. Дисперсия, однако, возрастает с понижением рейтинга (см. Таблицу 1).

При рассмотрении метрик более высоких моментов, однако, важно отметить, что данные метрики крайне непостоянны во времени. Как отмечает Нассим Талеб в книге «Statistical Consequences of Fat Tails», 79% эксцесса дневных логарифмических доходностей индекса SP500 за 56 лет пришлось на лишь одно экстремальное наблюдение. Для 15 рассмотренных Талебом финансовых инструментов (среди которых нет долговых инструментов) данный процент не опускается ниже 20%, при том, что для гауссовой случайной величины данный показатель был бы равен приблизительно 0,8% +- 0,28% (Taleb, 2019).

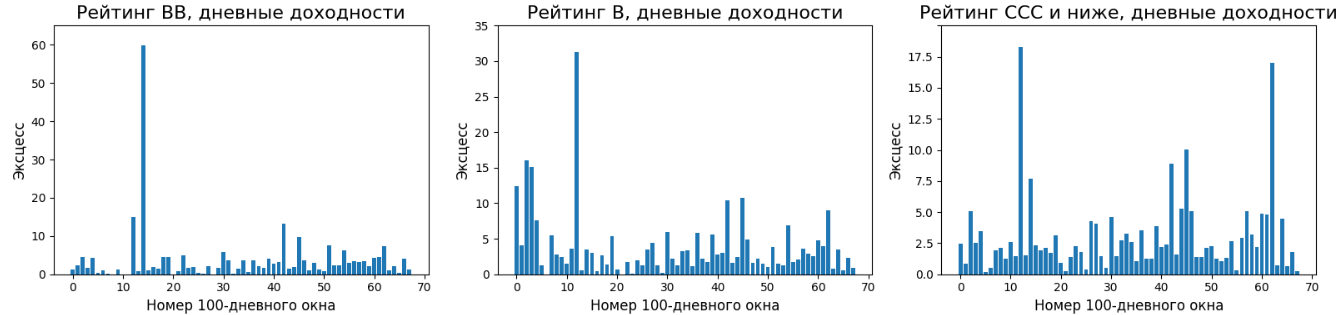
Повторяя эксперимент на рассматриваемых данных, можно увидеть, что удаление из выборки крайнего наблюдения понижает эксцесс на долю от 18% до 68% в зависимости от рейтинга (см. Таблицу 1). Если говорить интуитивно, четвертый момент (эксцесс), показывает стабильность второго момента (дисперсии). Тот факт, что от 18 до 68% эксцесса относится на одно наблюдение, показывает, что мы не можем ориентироваться на метрики второго момента для совершения количественных выводов ввиду их нестабильности (более того, как будет показано далее, в некоторых случаях для некоторых рассматриваемых переменных и вовсе нельзя говорить о *конечности* дисперсии согласно полученным параметрам распределений Фреше).

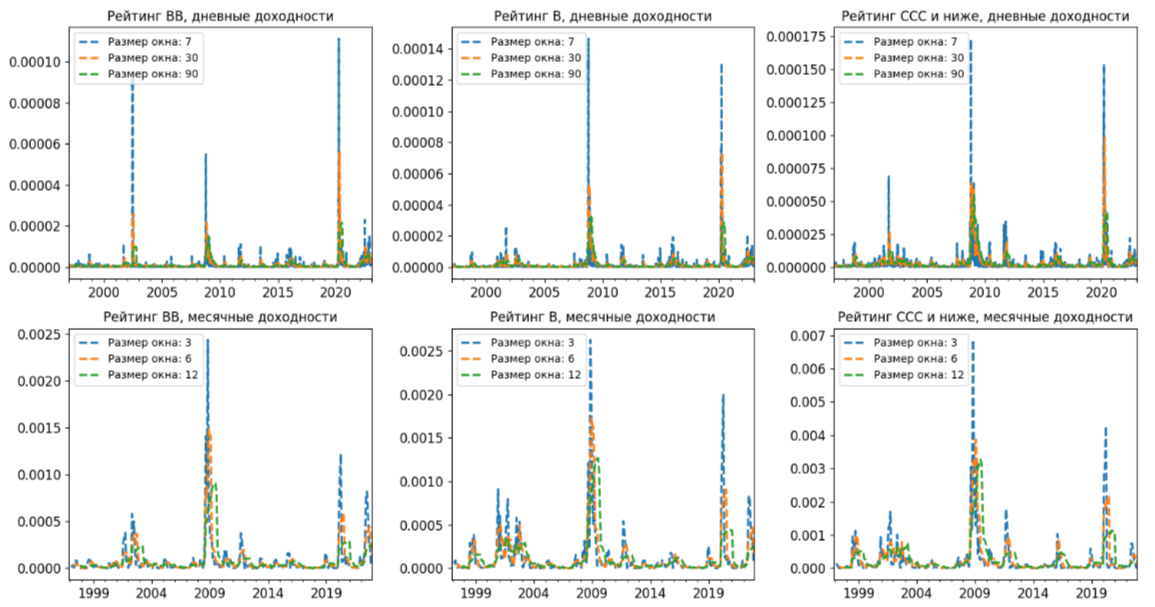
*Таблица 2. Описательные статистики логарифмов дневных и месячных доходностей*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Индекс** | **Период доходности** | **Среднее1** | **Дисперсия1** | **Скошенность2** | **Эксцесс2** | **Эксцесс, без крайнего наблюдения2** | **Доля крайнего наблюдения в эксцессе** |
| Рейтинг BB | Дневные | 0.10734 | 0.00143 | -0.00528 | 0.00013 | 0.00006 | 55,1 % |
| Рейтинг B | 0.08781 | 0.00201 | -0.00612 | 0.00017 | 0.00012 | 30,0 % |
| Рейтинг CCC и ниже | 0.09958 | 0.00387 | -0.01205 | 0.00041 | 0.00034 | 18,0 % |
| Рейтинг BB | Месячные | 2.25316 | 0.09058 | -1.77139 | 0.13617 | 0.04356 | 68,0 % |
| Рейтинг B | 1.85362 | 0.13286 | -2.21232 | 0.20120 | 0.12353 | 38,6 % |
| Рейтинг CCC и ниже | 2.06888 | 0.32992 | -8.08761 | 1.32244 | 0.78327 | 40,8 % |

1 Умножено на 10^3 для наглядности 2 Умножено на 10^6 для наглядности

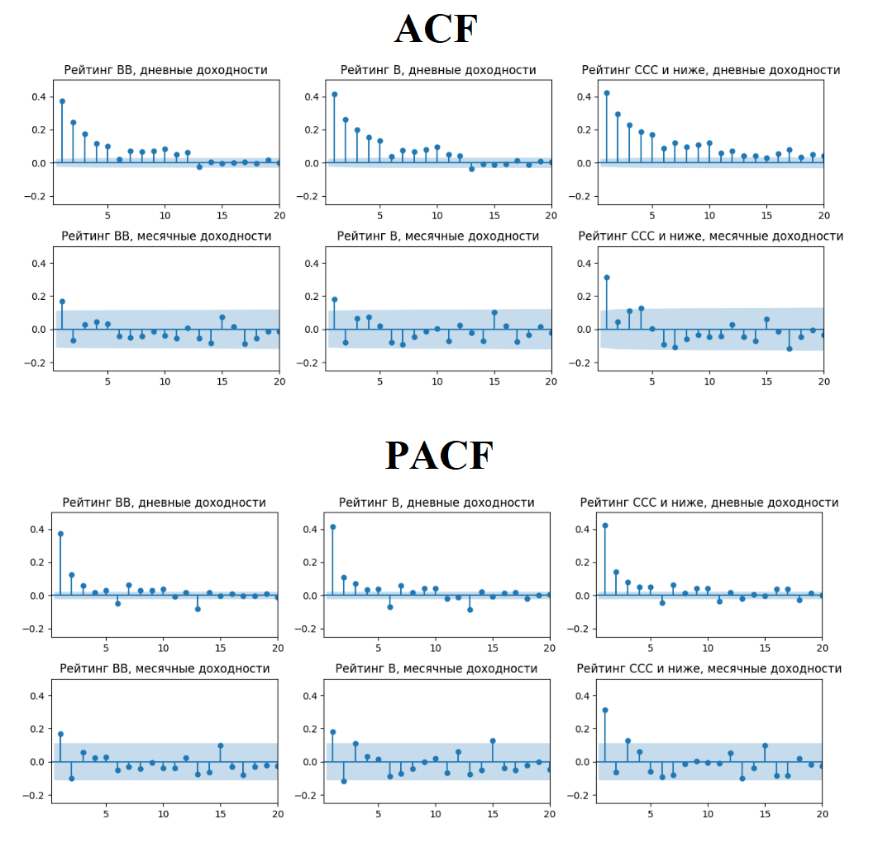
Аналогично, можно заметить, что эксцесс и дисперсия нестабильны во времени (см. Рисунки 2, 3), что показывает опасность использования исторических показателей второго момента для оценки будущих рисков. Эти выводы касаются и корреляций, при использовании которых современные теории также предполагают стабильность ковариаций *между* инструментами, что является еще более сильным допущением, чем стабильность дисперсии.

*Рисунок 2: Эксцесс дневных доходностей за 100 дней (не пересекающиеся окна)*

*Рисунок 3: Дисперсии доходностей во времени – сдвигающиеся окна*

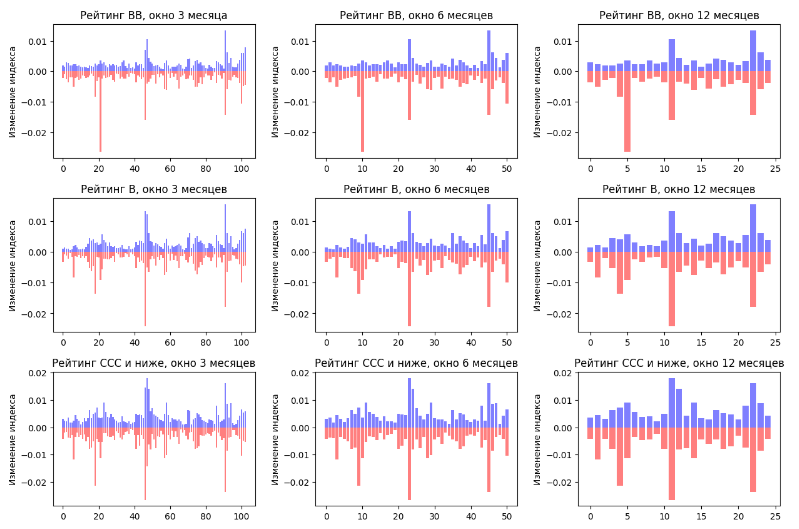
Далее, при рассмотрении переменных как временных рядов, очевидно наличие статистически значимых автокорреляций и частичных автокорреляций (см. Рисунок 4), что подтверждает необходимость рассмотрения авторегрессионной модели.

*Рисунок 4: Частичные автокорреляции (сверху) и автокорреляции (снизу) с указанием доверительных интервалов (синие поля)*



Для целей моделирования экстремальных значений отдельно с помощью распределения Frechet, в Рисунке 5 приведены временные ряды, состоящие из максимальных и минимальных доходностей за рассматриваемые временные промежутки. Несмотря на то, что каждое из данных значений по определению является крайним в довольно широкой подвыборке (66, 132 и 264 дня), очевидно наличие значимых выбросов.

*Рисунок 5. Максимальные (синие) и минимальные (красные) дневные доходности индексов за периоды 3, 6 и 12 месяцев*



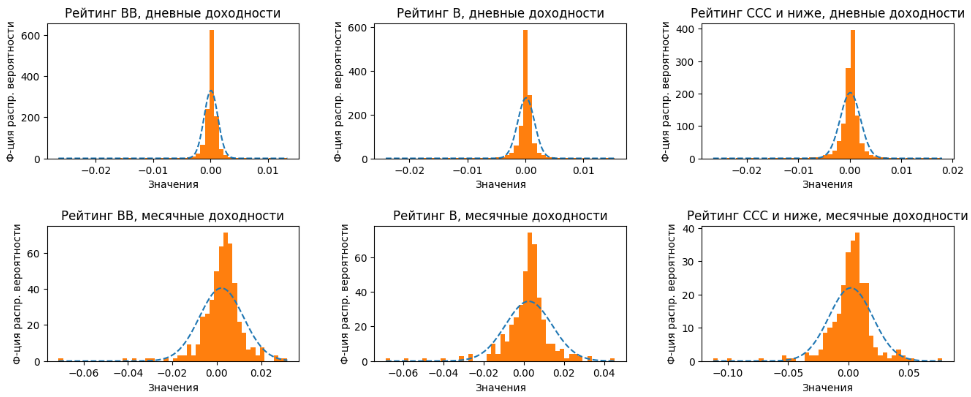
**2.2.1. Безусловные гомоскедастичные распределения**

Для начала, предположение о нормальности распределения отвегается тремя тестами для индексов всех рассматриваемых кредитных рейтингов на уровне значимости 0,1%. Это переносит выводы, полученные многими указанными исследователями (Mandelbrot, 1963; Rachev et al, 2005; Gabriel, Lau, 2014; Harris, Küçüközmen, 2001, Vukovic et al., 2019) на рынок высокодоходных облигаций, и является вполне ожидаемым результатом. Примечательно, что в работе Габриэля и Лау (Gabriel, Lau, 2014) не была отвергнута гипотеза нормальности распределения государственных облигаций, считавшихся близкими к безрисковым, однако она была отвергнута для рискованных облигаций. Таким образом, Таблицу 2 можно считать подтверждением наличия качественного различия между распределениями малорисковых государственных и высокорисковых высокодоходных облигаций.

*Таблица 3. Результаты тестов на нормальность: значения статистик и соответствующие p-значения (в скобках)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Индекс** | **Период доходности** | **KS-тест** | **CM-тест** | **Положение (loc)** | **Размерность**  **(scale)** |
| Рейтинг BB | Дневные | **0,14**  (0,00) | **58,95**  (0,00) | 0,0001 | 0,0012 |
| Рейтинг B | **0,16**  (0,00) | **76,04**  (0,00) | 0,0001 | 0,0014 |
| Рейтинг CCC  и ниже | **0,14**  (0,00) | **56,24**  (0,00) | 0,0000 | 0,0020 |
| Рейтинг BB | Месячные | **0,13**  (0,00) | **1,72**  (0,00) | 0,0021 | 0,0099 |
| Рейтинг B | **0,12**  (0,00) | **1,85**  (0,00) | 0,0024 | 0,0116 |
| Рейтинг CCC  и ниже | **0,14**  (0,00) | **1,78**  (0,00) | 0,0021 | 0,0182 |

*Рисунок 6. Эмпирические распределения логарифмов доходностей (оранжевые) и функции распределения вероятности нормального распределения, соответствующие параметрам, полученным с ММП (синие)*

**

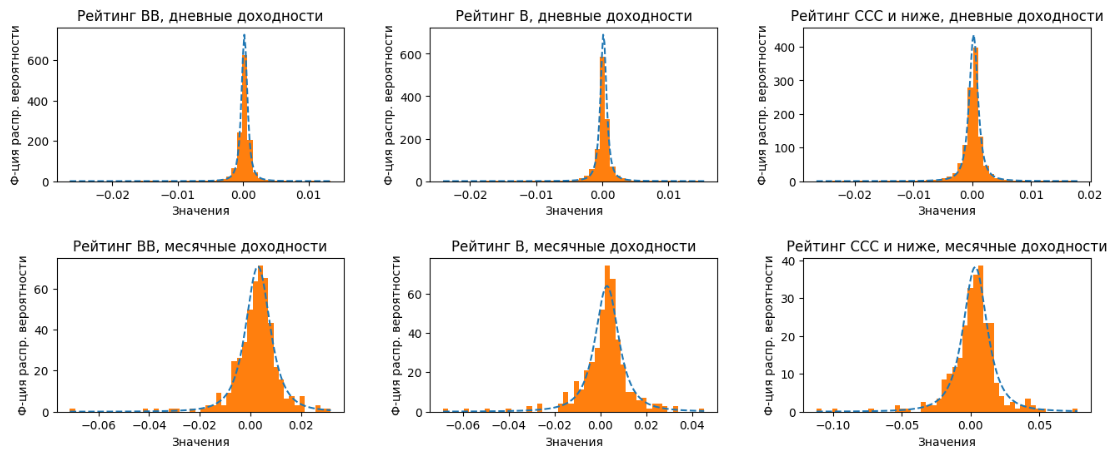
Далее, предположение о принадлежности переменных к распределению Стьюдента отвергается для дневных данных, но не отвергается для месячных данных. Это является весьма интересным наблюдением. Месячные доходности, поскольку считаются по упорядоченному временному ряду, фактически учитывают в себе эффекты временной зависимости (ARMA) и нестационарности волатильности (GARCH). Как следствие, месячная доходность представляет собой ансамбль из ~22 дневных доходностей, и вполне логично, что данный ансамбль будет иметь менее остроконечное распределение.

Также мы можем видеть, что рейтинг BB имеет наибольшее число степеней свободы (2,59), что означает наличие наименее толстых хвостов распределения. Для рейтингов B и CCC и ниже данный параметр довольно близок (2,09 и 2,30, соответственно).

*Таблица 4. Результаты тестов на соответствие распределению Стьюдента: значения статистик и соответствующие p-значения (в скобках)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Индекс** | **Период доходности** | **KS-тест** | **CM-тест** | **Ст. свободы (df)** | **Положение (loc)** | **Размерность**  **(scale)** |
| Рейтинг BB | Дневные | **0,02**  (0,00) | **1,15**  (0,00) | **1,89** | **0,0002** | **0,0005** |
| Рейтинг B | **0,03**  (0,00) | **2,26**  (0,00) | **1,80** | **0,0002** | **0,0005** |
| Рейтинг CCC и ниже | **0,03**  (0,00) | **2,01**  (0,00) | **1,86** | **0,0002** | **0,0008** |
| Рейтинг BB | Месячные | **0,04**  (0,78) | **0,08**  (0,70) | **2,59** | **0,0030** | **0,0051** |
| Рейтинг B | **0,04**  (0,65) | **0,13**  (0,47) | **2,09** | **0,0029** | **0,0056** |
| Рейтинг CCC и ниже | **0,04**  (0,73) | **0,07**  (0,77) | **2,30** | **0,0034** | **0,0094** |

*Рисунок 7. Эмпирические распределения логарифмов доходностей (оранжевые) и функции распределения вероятности распределения Стьюдента, соответствующие параметрам, полученным с ММП (синие)*



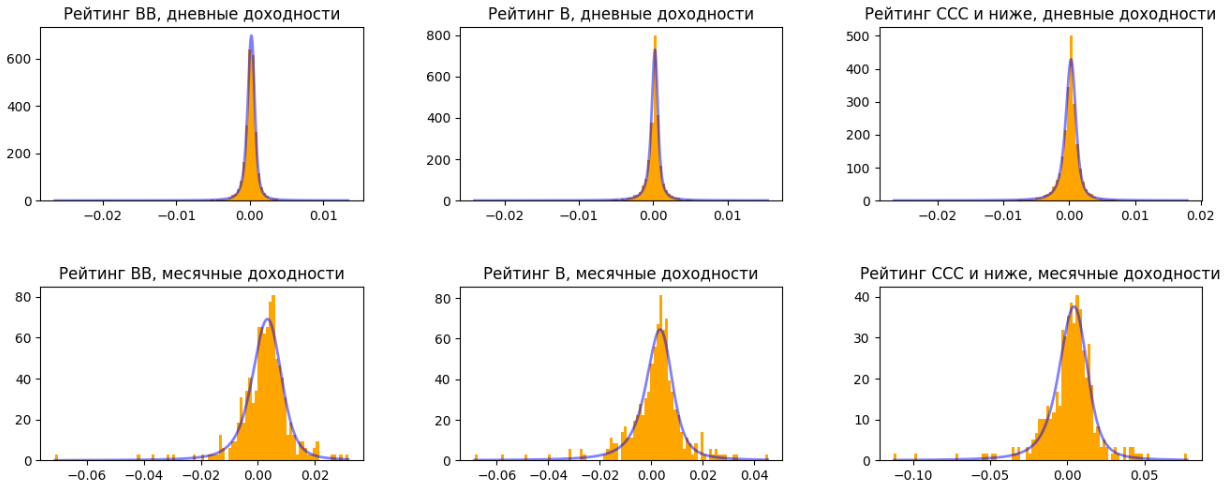
Наконец, при оценке параметров стабильного распределения Парето нулевая гипотеза не отвергается для месячных доходностей на всех разумных уровнях уверенности, причем p-value тестов выше, чем для распределения Стьюдента. Хвостовой параметр *alpha* варьируется от 1,33 до 1,50 в зависимости от рейтинга, а параметр *beta* –от -0,27 до -0,24, что свидетельствует о довольно высокой степени скошенности.

Однако для дневных доходностей гипотеза принадлежности к Парето отвергается на уровне уверенности 1% двумя тестами, помимо индекса рейтинга BB. Этот вывод расходится с результатом, полученным Габриэлем и Лау для *государственных облигаций Европейского Союза* на выборке с 1999 по 2011 год (около 12 лет), которые показали возможность принадлежности к Парето-распределению согласно тем же тестам. Для большинства рассматриваемых ими стран, однако, нулевая гипотеза не отвергалась лишь на уровне значимости 1% или 5%. Таким образом, настоящая работа тестирует эти выводы в приложении к ВДО, но на значимо более длительном промежутке (около 26 лет), чем и объясняется отвергание нулевой гипотезы в условиях отсутствия асимптотической сходимости выборки к предполагаемому распределению (что также наблюдается из графика, где видно наличие более высокого «горба» в эмпирическом распределении). Для того, чтобы удостовериться в этом объяснении, был также проведен аналогичный тест для последних 25% наблюдений выборки по рейтингу BB, в результате чего p-value CM-теста увеличился до 0,188.

*Таблица 5. Результаты тестов на соответствие стабильному распределению Парето: значения статистик и соответствующие p-значения (в скобках)*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Индекс** | **Период доходности** | **KS-тест** | **CM-тест** | **alpha** | **beta** | **loc** | **scale** |
| Рейтинг BB | Дневные | **0,027**  (0,00) | **0,688**  (0,014) | **1,366** | **-0,171** | **0,000088** | **0,000417** |
| Рейтинг B | **0,029**  (0,00) | **0,754**  (0,009) | **1,200** | **-0,146** | **0,000016** | **0,000410** |
| Рейтинг CCC и ниже | **0,031**  (0,00) | **1,075**  (0,002) | **1,309** | **-0,179** | **0,000038** | **0,000683** |
| Рейтинг BB | Месячные | **0,025**  (0,99) | **0,036**  (0,95) | **1,496** | **-0,244** | **0,0022** | **0,0042** |
| Рейтинг B | **0,036**  (0,81) | **0,046**  (0,90) | **1,329** | **-0,248** | **0,0012** | **0,0045** |
| Рейтинг CCC и ниже | **0,024**  (0,98) | **0,024**  (0,99) | **1,424** | **-0,272** | **0,0012** | **0,0077** |

*Рисунок 8. Эмпирические распределения логарифмов доходностей (оранжевые) и функции распределения вероятности стабильного распределения Парето, соответствующие параметрам, полученным с ММП (синие)*



Подводя итог для безусловных распределений, нельзя отвергнуть гипотезу принадлежности месячных доходностей к стабильному распределению Парето и распределению Стьюдента, но можно отвергнуть эту гипотезу для дневных доходностей. Выбирая из двух распределений, для месячных доходностей CM-тестом большее предпочтение отдается распределению Pareto Stable. Факт неотвергания данного распределения для месячных доходностей повторяет выводы, полученные исследователями рынка акций и товарных рынков (Mandelbrot, 1963; Rachev et al., 2005) и рынков облигаций (Gabriel & Lau, 2014), но в применении к рынку высокодоходных облигаций. Факт отвергания гипотезы для дневных доходностей аргументирует тезис Лонгина, утверждавшего в своей статье, что Парето-распределение не должно рассматриваться отдельно от других статистических моделей, поскольку не учитывает зависимость параметров во времени, нестабильность дисперсии, наличие автокорреляции и ряд других факторов, присущих финансовым переменным (Longin, 1996). В следующей главе рассматривается гетероскедастичная модель с целью учесть эти факторы.

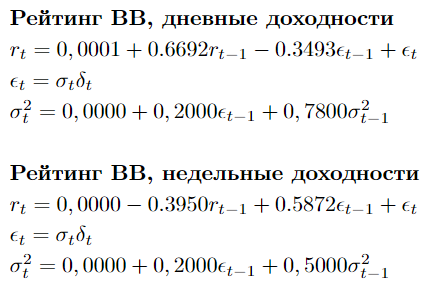
**2.2.2. Условные гетероскедастичные модели**

Далее применяется модель ARMA-GARCH, которая позволяет учесть наличие автокорреляции и непостоянность волатильности во времени (гетероскедастичность). Для выбора наилучшей спецификации модели сначала подбирается порядок авторегрессионной модели ARMA(p, q). Для этого для каждой из возможных комбинаций *p* и *q* (от 0 до 12, исходя из графиков ACF и PACF)рассчитываются метрики AIC и BIC, после чего для дальнейшего анализа выбирается пара *p\*, q\** с наименьшим значением данных показателей. Далее остатки модели ARMA(p\*, q\*) служат временным рядом, на основе которого аналогичным образом подбирается оптимальная спецификация модели GARCH(r\*, s\*). Наконец, проводятся статистические тесты на принадлежность остатков модели GARCH(r\*, s\*), иначе называемых «инновациями», к одному из трех рассматриваемых безусловных распределений.

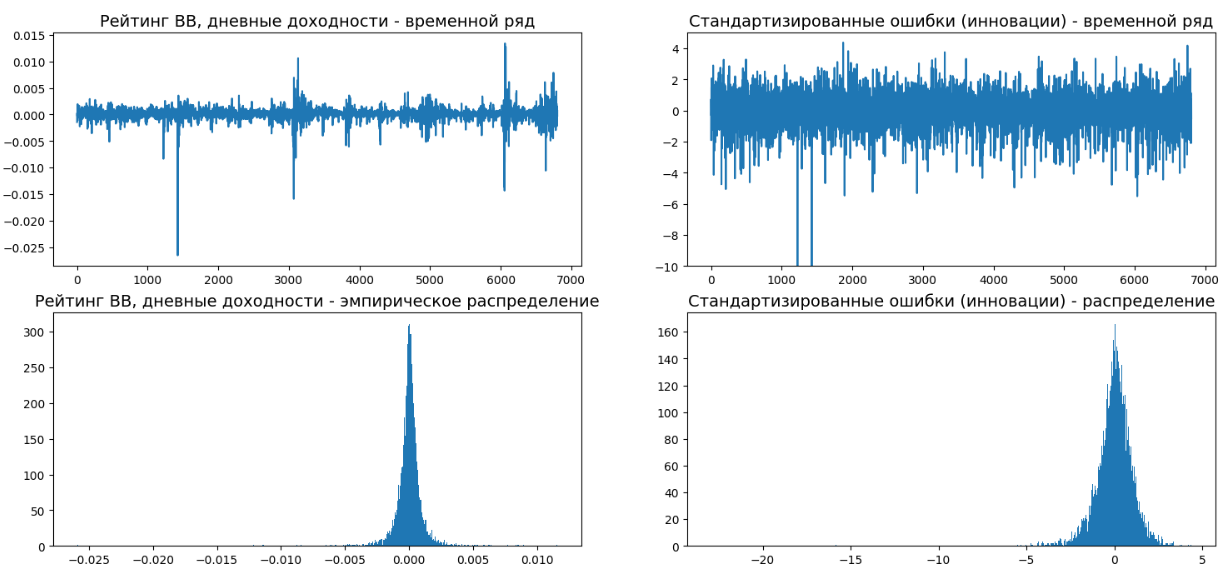
По итогу применения подобного алгоритма оптимальной спецификацией была выявлена комбинация ARMA(7,10)-GARCH(2,1). Однако разница в значениях AIC и BIC между этой спецификацией и моделью ARMA(1,1)-GARCH(1,1) оказалась крайне мала, а в компоненте GARCH комбинация (1,1) и вовсе является вторым по приоритету выбором. Как следствие, по мнению автора, не имеет значимого смысла использовать модель с большим числом параметров при отсутствии значимого выигрыша в качестве модели. Как было описано ранее, метрики второго момента, включая корреляции, являются крайне нестабильными в финансовых данных, из-за чего выбор спецификации модели может быть подвержен конкретному временному окну, выбранному исследователем. Для того, чтобы минимизировать данный эффект, дальнейший анализ будет проводиться с моделью ARMA(1,1)-GARCH(1,1). Аналогичный подход применяется в нескольких статьях (Rachev & Mittnik, 2000; Rachev et al, 2005).

Как видно из Рисунков 8 и 9, применение данной модели позволяет перейти от необходимости моделировать переменную с толстыми хвостами и очевидными выбросами в виде экстремальных значений к моделированию величины с меньшей волатильностью и более тонкими хвостами.

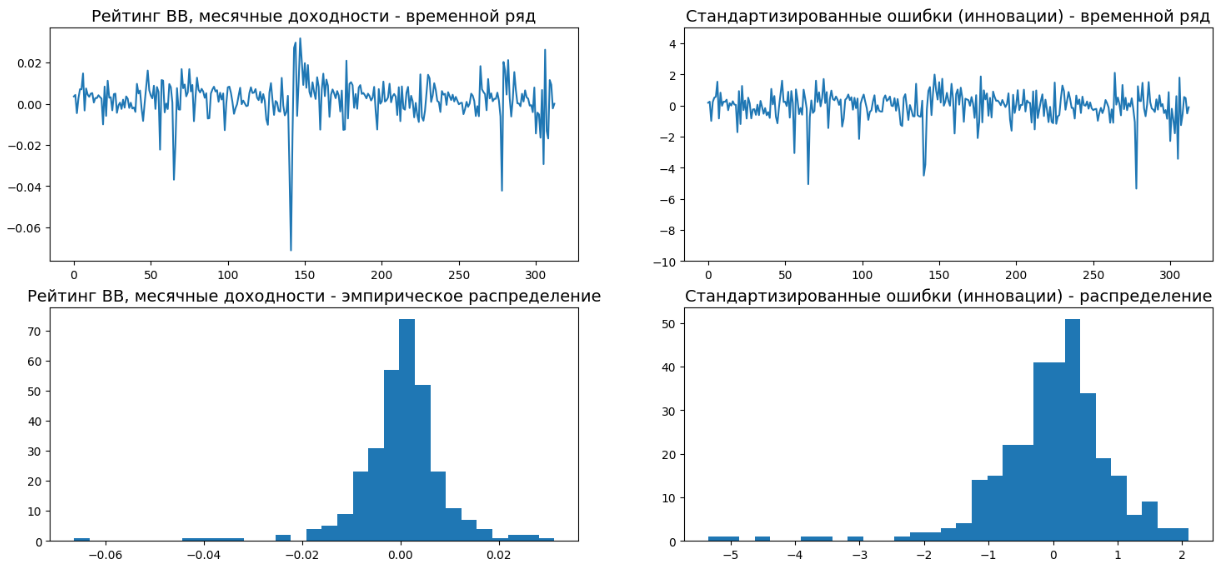
Поскольку выводы практически не различаются между кредитными рейтингами, ниже приведены результаты только для индекса рейтинга BB. Итоговые формулы для доходностей выглядят следующим образом:



*Рисунок 9. Результат применения модели ARMA(1,1)-GARCH(1,1): исходная дневная переменная, относящаяся к рейтингу BB (слева) и стандартизированные ошибки автокорреляционной гетероскедастичной модели (справа). Границы правого верхнего графика были усечены для наглядности*



*Рисунок 10. Результат применения модели ARMA(1,1)-GARCH(1,1): исходная месячная переменная, относящаяся к рейтингу BB (слева) и стандартизированные ошибки автокорреляционной гетероскедастичной модели (справа)*

****

Далее для данных инноваций необходимо оценить параметры безусловных распределений. Нормальность остатков отвергается тремя тестами на любых разумных уровнях уверенности в случае дневных доходностей. Для месячных данных нормальность также отвергается CM-тестом на уровне уверенности 1%, что демонстрируется в Таблице 6, но отвергается лишь на уровне 2% KS-тестом для индекса рейтинга CCC и ниже. Это, вероятно, служит проявлением упомянутой ранее особенности KS-теста, уделяющего меньшее внимание хвостам распределения.

*Таблица 6. Результаты тестов на соответствие остатков модели ARMA(1,1)-GARCH(1,1) нормальному распределению: значения статистик и соответствующие p-значения (в скобках)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Индекс** | **Период доходности** | **KS-тест** | **CM-тест** | **Положение (loc)** | **Размерность**  **(scale)** |
| Рейтинг BB | Дневные | **0,07**  (0,00) | **11,11**  (0,00) | -0,0052 | 1,052 |
| Рейтинг B | **0,08**  (0,00) | **13,91**  (0,00) | 0,0024 | 0,9862 |
| Рейтинг CCC  и ниже | **0,06**  (0,00) | **10,55**  (0,00) | 0,0004 | 1,0293 |
| Рейтинг BB | Месячные | **0,10**  (0,002) | **0,87**  (0,004) | -0,0285 | 0,9480 |
| Рейтинг B | **0,11**  (0,00) | **1,13**  (0,00) | -0,1020 | 1,0194 |
| Рейтинг CCC  и ниже | **0,09**  (0,016) | **0,88**  (0,005) | -0,1130 | 1,0545 |

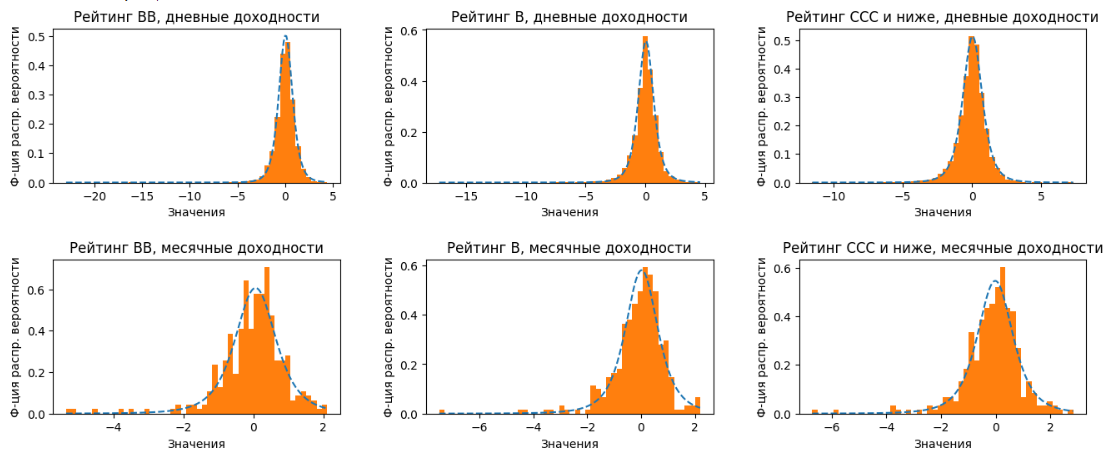
Однако нельзя отвергнуть гипотезу о соответствии месячных инноваций распределению Стьюдента: p-value по результатам CM-теста для всех рейтингов составляет не менее 0,48. Стоит отметить, что число степеней свободы по оценке через ММП у инноваций выше, чем у исходных переменных, не пропущенных через модель ARMA-GARCH (3,17–3,38 для инноваций против 2,17–2,84 для исходных переменных). Это служит аргументом в пользу гипотезы №3 о том, что применение условной гетероскедастичной модели позволяет существенно снизить степень тяжести хвостов распределения «корневой» переменной.

Для дневных же данных распределение Стьюдента не отвергается CM-тестом на уровне уверенности 5% для рейтинга BB и CCC и ниже; для рейтинга B p-value составляет 3,6%. Можно заметить, что на Рисунке 9 функция распределения с оцененными параметрами имеет хорошую подгонку с визуальной точки зрения – как для дневных, так и для месячных данных. Значимо более низкие p-value для дневых данных, по мнению автора, объясняется тем, что при кратном повышении числа наблюдений значительно понижается порог отвергания нулевой гипотезы. Другими словами, чем больше выборка – тем точнее должна быть оценка параметров теоретического распределения и тем качественнее данные должны подходить под это распределение для неотвергания нулевой гипотезы. Вместе с тем, в остатках модели ARMA-GARCH сохраняется незначительная часть автокорреляционного эффекта, которая тем не менее не позволяет говорить об асимптотической сходимости распределения случайной величины (инноваций) к i.i.d. распределению Стьюдента.

*Таблица 7. Результаты тестов на соответствие остатков модели ARMA(1,1)-GARCH(1,1) распределению Стьюдента: значения статистик и соответствующие p-значения (в скобках)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Индекс** | **Период доходности** | **KS-тест** | **CM-тест** | **Степень свободы *(DF)*** | **Положение**  ***(m)*** | **Размерность**  ***(scale)*** |
| Рейтинг BB | Дневные | **0,015**  (0,089) | **0,301**  (0,135) | 4,25 | 0,0420 | 0,7492 |
| Рейтинг B | **0,002**  (0,008) | **0,51**8  (0,036) | 3,49 | 0,0586 | 0,6669 |
| Рейтинг CCC и ниже | **0,015**  (0,093) | **0,416**  (0,066) | 3,79 | 0,0479 | 0,7277 |
| Рейтинг BB | Месячные | **0,04**  (0,68) | **0,09**  (0,65) | 3,38 | 0.0623 | 0.6115 |
| Рейтинг B | **0,05**  (0,41) | **0,12**  (0,48) | 3,17 | 0,0229 | 0,6356 |
| Рейтинг CCC и ниже | **0,039**  (0,71) | **0,097**  (0,60) | 3,35 | -0,0170 | 0,6797 |

*Рисунок 11. Эмпирические распределения инноваций модели ARMA-GARCH (оранжевые) и функции распределения вероятности распределения Стьюдента, соответствующие параметрам, полученным с ММП (синие)*



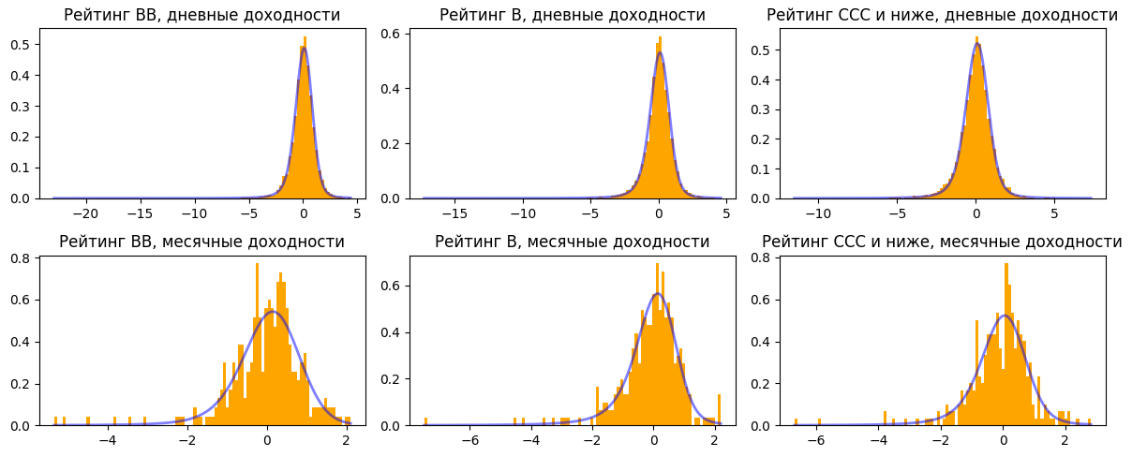
Повторяя эксперимент для стабильного распределения Парето, нулевая гипотеза не отвергается для месячных данных на любых разумных степенях уверенности. Очевидно наличие существенно отрицательной скошенности, которая проявляется в параметре *beta* от -0,44 для рейтинга CCC и ниже до -0,70 для рейтинга B (напомним, что параметр может принимать значения от -1 до 1). Изучив Рисунок 9, можно увидеть проявление этой особенности в наличии большого числа наблюдений в левой части графика.

Для дневных же доходности гипотеза принадлежности не отвергается только на уровне уверенности 10% для рейтинга BB и 1% для двух других индексов. Этот результат аналогичен тому, что был получен для распределения Стьюдента ранее – очевидно отсутствие асимптотической сходимости выборки к предполагаемому распределению при увеличении числа наблюдений, что также проявляется в хорошей визуальной подгонке (Рисунок 10), за исключением наличия некоторого «горба» у гистограммы.

*Таблица 8. Результаты тестов на соответствие остатков модели ARMA(1,1)-GARCH(1,1) стабильному распределению Парето: значения статистик и соответствующие p-значения (в скобках)*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Индекс** | **Период доходности** | **KS-тест** | **CM-тест** | **alpha** | **beta** | **loc** | **scale** |
| Рейтинг BB | Дневные | **0,013**  (0,17) | **0,294**  (0,14) | **1,693** | **-0,323** | **-0,014** | **0,582** |
| Рейтинг B | **0,02**  (0,01) | **0,587**  (0,02) | **1,654** | **-0,342** | **-0,012** | **0,537** |
| Рейтинг CCC и ниже | **0,022**  (0,00) | **0,573**  (0,03) | **1,659** | **-0,211** | **0,020** | **0,545** |
| Рейтинг BB | Месячные | **0,05**  (0,41) | **0,076**  (0,72) | **1,762** | **-0,703** | **-0,021** | **0,520** |
| Рейтинг B | **0,022**  (0,99) | **0,019**  (0,99) | **1,592** | **-0,607** | **-0,134** | **0,502** |
| Рейтинг CCC и ниже | **0,037**  (0,77) | **0,036**  (0,95) | **1,644** | **-0,435** | **-0,117** | **0,543** |

*Рисунок 12. Эмпирические распределения инноваций модели ARMA-GARCH (оранжевые) и функции распределения вероятности стабильного распределения Парето, соответствующие параметрам, полученным с ММП (синие)*

****

Подводя итог рассмотрения условной гетероскедастичной модели, с большой степенью уверенности нельзя отвергать принадлежность месячных данных к распределениям Стьюдента и стабильному распределению Парето, что касается и случая i.i.d. модели, и случая рассмотрения инноваций модели ARMA-GARCH. Использование авторегрессионной модели более предпочтительно, поскольку ее инновации обладают более высокими показателями *alpha* и *df* для Стьюдента и Парето соответственно, что означает более тонкие хвосты. Выбирая между двумя распределениями для остатков ARMA-GARCH – предпочтение стоит отдать Парето ввиду более высоких p-value CM-теста, а также, что более важно – ввиду возможности учесть наличие отрицательной скошенности в показателе *beta.*

Что касается дневных данных – принадлежность к Стьюденту и Парето не отвергается лишь на степени уверенности 1% в большинстве рассматриваемых случаев, что объясняется отсутствием асимптотической сходимости выборки к одному из этих двух распределений с рассматриваемыми параметрами при увеличении ее размера. Это в свою очередь может объясняться изменением параметров во времени, а также сохранения части автокорреляционного эффекта (в отличие от месячных данных, где последний эффект несколько сглаживается за счет рассмотрения ансамбля из 5 упорядоченных наблюдений временного ряда).

**2.2.3. Безусловные гомоскедастичные модели для экстремальных значений**

Далее в качестве переменной рассматривается не дневная или месячная доходность, а их максимальные и минимальные значения за период 3, 6 или 12 месяцев (непересекающиеся окна). Этот подход позволяет применить обширный инструментарий Теории экстремальных значений, и в частности частично применить подход Франсуа Лонгина, примененный к рынку акций, на рынок ВДО. Кроме того, в риск-менеджменте для управленцев представляют интерес именно вероятности экстремальных событий, из-за чего на практике и применяются модели VaR и ES, которые учитывают только хвостовую часть распределения. В частности, теория экстремальных значений не рассматривает свойства центральной части распределения случайных величин, а фокусируется именно на краевых значениях, что и является ключевым аспектом в моделировании финансовых случайных величин.

Оценка параметров в данном подходе производится только для дневных данных, поскольку для месячных доходностей доступно лишь 313 наблюдений, и это число уменьшится кратно после выведения максимальных и минимальных доходностей за некоторый период – чего недостаточно для качественной оценки параметров распределений с тяжелыми хвостами.

Для выбора правильной модели для максимумов и минимумов применяется модель, описанная одним из основоположников EVT (Gumbel, 1958) и использованная Лонгином в описанной выше работе (Longin, 1996). Сначала путем упорядочивания ряда максимумов *(Yn,i)i=1,N* (в случае данной работы – за окна 3, 6 и 12 месяцев) получается ряд *(Y\*n,i)i=1,N,*где каждое следующее значение больше предыдущего. На основании получившихся значений рассчитывается кумулятивная функция распределения F(Y\*n,i) = i/(N+1), принимающая значение от 0 до 1 в зависимости порядкового номера наблюдения.

Вспомним, что общая функция экстремальных значений, GEV, в частном случае распределения Gumbel задается следующим образом (с добавлением параметров location *(mu)* и scale *(s)*):



Данная функция может быть преобразована в уравнение прямой линии путем взятия двойного логарифма:

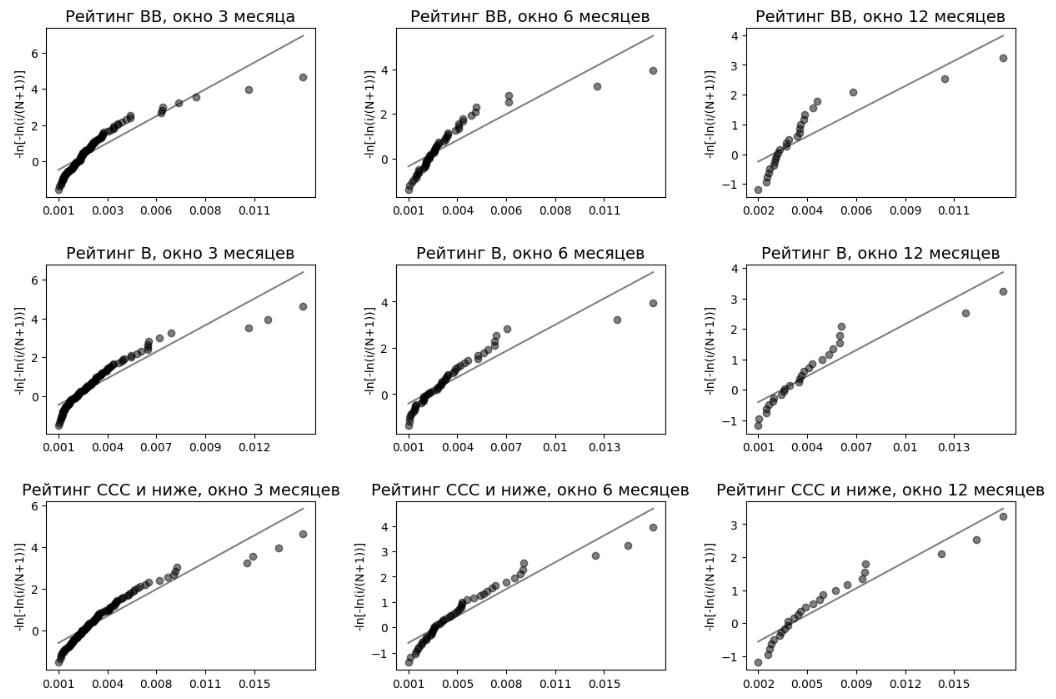


Аналогично, на основе рассчитанных ранее значений F(Y\*n,i) берется двойной логарифм для расчета метрики *–ln(–ln(i/(N+1)),* которая затем может быть использована для оценки параметров *alpha* и *beta* в следующей регрессии:

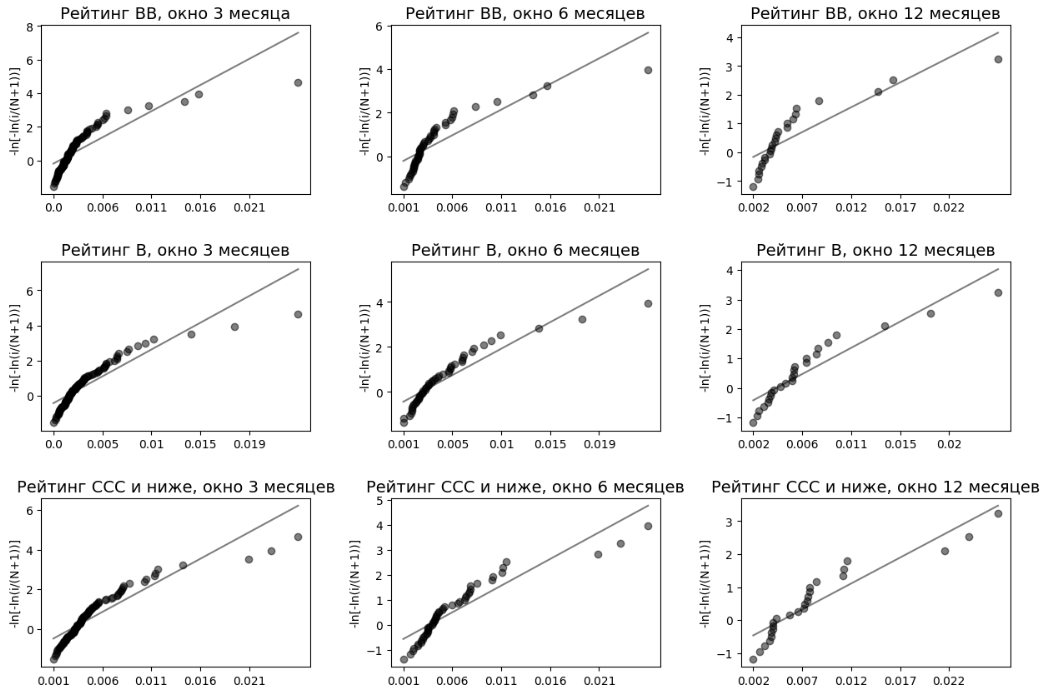


Получившиеся параметры не представляют большого интереса для данной работы, однако через сравнение получившихся линий регрессий и эмпирического распределения метрики *–ln(–ln(i/(N+1))* (см. Рисунки 13 и 14) можно визуально определить вероятную принадлежность исходной величины к одному из трех распределений EVT. Если точки (наблюдения) лежат на прямой, то это индикация распределения Gumbel. Если точки образуют выпуклую кривую, то это признак распределения Weibull, а если вогнутую кривую – то это признак принадлежности к Frechet. На Рисунках 13 и 14 явно прослеживается вогнутость кривых, и особенно заметно непопадание точек на прямую в самых левых частях графиков. Это верно и для максимальных, и для минимальных доходностей.

*Рисунок 13. Регрессии –ln(–ln(i/(N+1)) (ось Y)* на *(Y\*n,i)i=1,N (ось Х) (спецификация Gumbel) для максимальных доходностей*



*Рисунок 14. Регрессии –ln(–ln(i/(N+1)) (ось Y)* на *(Y\*n,i)i=1,N (ось Х) (спецификация Gumbel) для минимальных доходностей*



Если же сделать аналогичное преобразование для функции кумулятивного распределения Фреше,



то получим следующее уравнение:



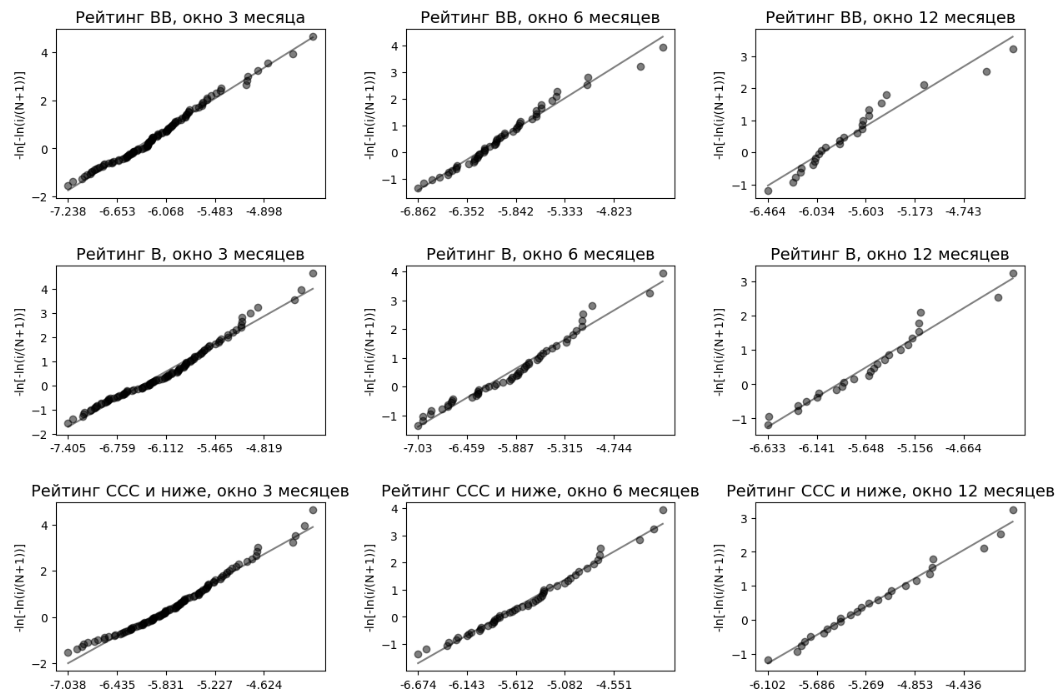
Поскольку *alpha* и *ln(s)* считаются константами, то, заменяя коэффициенты, возможно построить следующую регрессию теоретической функции распределения *–ln(–ln(i/(N+1))* нафактические значения Yn,i\*:



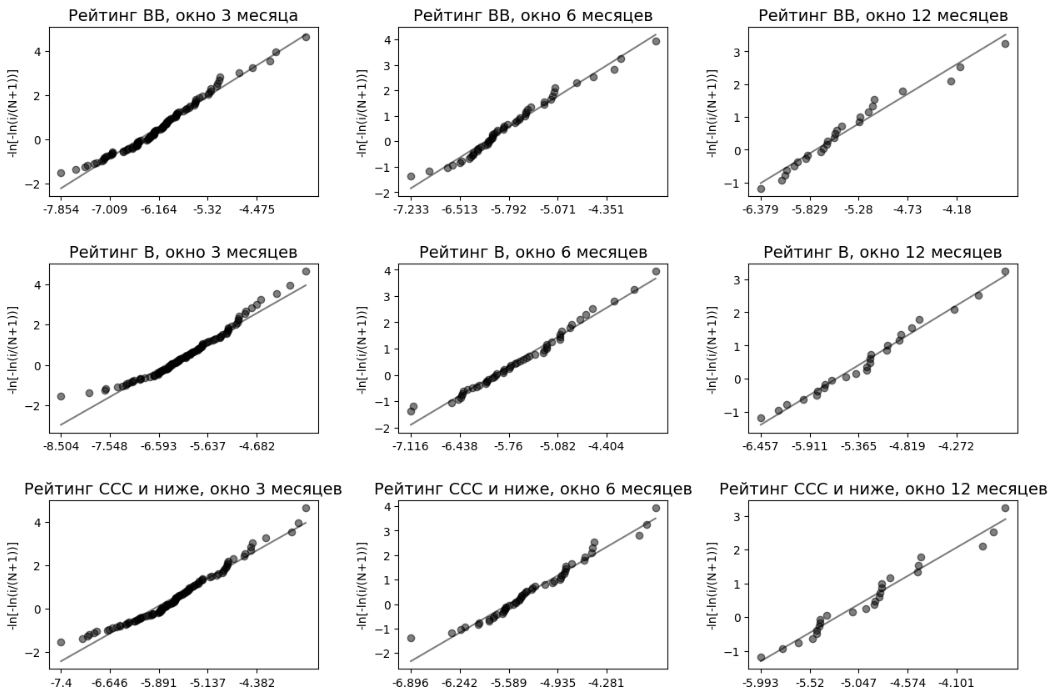
Оценивая эту регрессию с помощью метода наименьших квадратов, получаем, что значения *ln(Yn,i\*)* лежат крайне близко к прямой (см. Рисунки 15 и 16). Примечательно, что даже самые правые значения (то есть экстремальные значения *среди* экстремальных значений также находятся крайне близко к линии. Это означает, что яркие кризисные события, как некоторые дни 2008 и 2020 годов, не являются статистическими отклонениями как таковыми, если рассматривать их с точки зрения EVT, и конкретно распределения Фреше. Это является довольно интересным выводом.

Следовательно, целесообразно в дальнейшем анализе рассматривать именно распределение Frechet. Это наблюдение повторяет вывод Лонгина для рынка акций (Longin, 1996).

*Рисунок 15. Регрессии –ln(–ln(i/(N+1)) (ось Y)* на *ln[(Y\*n,i)i=1,N](ось Х) (спецификация Frechet) для максимальных доходностей*



*Рисунок 16. Регрессии –ln(–ln(i/(N+1)) (ось Y)* на *ln[(Y\*n,i)i=1,N](ось Х) (спецификация Gumbel) для минимальных доходностей*



После оценки параметров распределения Фреше для экстремальных значений дневных доходностей можно увидеть, что нулевая гипотеза принадлежности этих переменных к этому распределению не отвергается тестами для всех временных окон и для всех индексов (см. Таблицу 9). Данные выводы одинаковы для минимальных и максимальных значений.

*Таблица 9. Результаты тестов на соответствие максимальных и минимальных значений за окна 3, 6 и 12 месяцев распределению Фреше: значения статистик и соответствующие p-значения (в скобках)*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Индекс** | **Тип значения** | **Размер окна** | **KS-тест** | **CM-тест** | ***alpha*** | **Положение**  **(m)** | **Размерность**  **(s)** |
| Рейтинг BB | Максимальное | 3 мес. | **0,096**  (0,29) | **0,186**  (0,30) | **2,13** | **0,0000** | **0,0016** |
| 6 мес. | **0,088**  (0,80) | **0,068**  (0,76) | **2,45** | **0,0000** | **0,0020** |
| 12 мес. | **0,092**  (0,97) | **0,037**  (0,95) | **2,84** | **0,0000** | **0,0026** |
| Минимальное | 3 мес. | **0,096**  (0,29) | **0,186**  (0,30) | **1,53** | **0,0000** | **0,0015** |
| 6 мес. | **0,088**  (0,80) | **0,086**  (0,76) | **1,74** | **0,0000** | **0,0022** |
| 12 мес | **0,085**  (0,99) | **0,028**  (0,98) | **2,10** | **0,0000** | **0,0033** |
| Рейтинг B | Максимальное | 3 мес. | **0,079**  **(0,52)** | **0,138**  (0,43) | **1,68** | **0,0000** | **0,0016** |
| 6 мес. | **0,113**  (0,50) | **0,094**  (0,62) | **1,76** | **0,0000** | **0,0019** |
| 12 мес. | **0,127**  (0,77) | **0,052**  (0,87) | **1,89** | **0,0000** | **0,0027** |
| Минимальное | 3 мес. | **xx** | **xx** | **1,36** | **0,0000** | **0,0019** |
| 6 мес. | **0,072**  (0,94) | **0,041**  (0,93**)** | **1,56** | **0,0000** | **0,0026** |
| 12 мес | **0,129**  (0,75) | **0,041**  (0,93) | **1,75** | **0,0000** | **0,0037** |
| Рейтинг CCC  и ниже | Максимальное | 3 мес. | **0,71**  (0,65) | **0,119**  (0,50) | **1,79** | **0,0000** | **0,0024** |
| 6 мес. | **0,062**  (0,98) | **0,031**  (0,97) | **1,90** | **0,0000** | **0,0031** |
| 12 мес. | **0,064**  (0,99) | **0,015**  (0,99) | **2,25** | **0,0000** | **0,0043** |
| Минимальное | 3 мес. | **0,119**  (0,10) | **0,25**  (0,19) | **1,48** | **0,0000** | **0,0025** |
| 6 мес. | **0,099**  (0,66) | **0,10**  (0,58) | **1,65** | **0,0000** | **0,0037** |
| 12 мес | **0,127**  (0,77) | **0,072**  (0,74) | **1,99** | **0,0000** | **0,0053** |

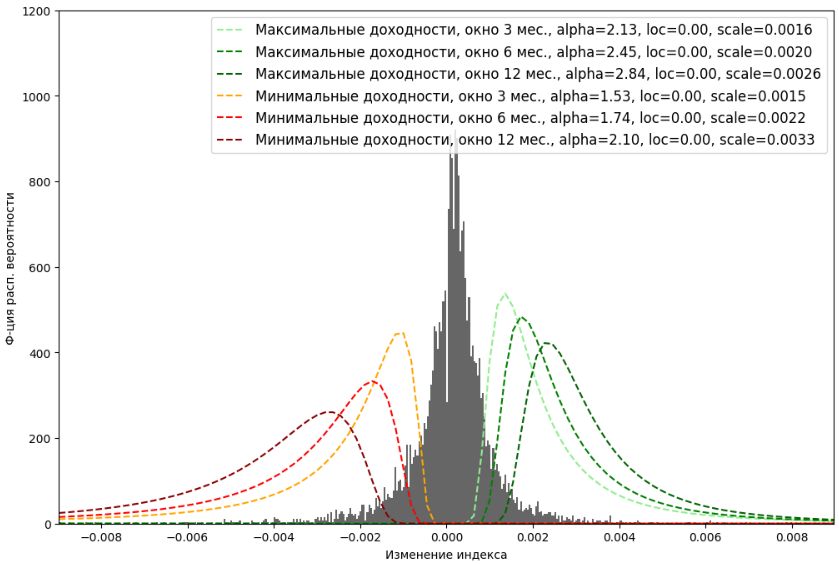
Рассматривая результаты на примере индекса BB, для окон 3, 6 и 12 месяцев были получены параметры alpha 2,13, 2,45 и 2,84 для максимальных и 1,53, 1,74 и 2,10 для минимальных доходностей. Во-первых, видно, что степень *alpha* увеличивается с увеличением окна, что весьма логично – за более длительный период выше вероятность возникновения более экстремального наблюдения. Эта закономерность повторяется для остальных индексов.

Во-вторых, в качестве более интересного вывода, для минимальных доходностей показатель alpha существенно ниже, чем для максимальных, что свидетельствует о том, что максимумы убытков по портфелю высокодоходных облигаций имеют более толстые хвосты, чем доходности (см. Рисунок 17).

Этот вывод подтверждается тестом LR (likelihood ratio). В рамках этого теста берется функцию логарифма правдоподобия *L(x, alpha, loc, scale)*, где *x* – рассматриваемая выборка, а alpha, loc, scale – параметры распределения Фреше. Рассчитывается значение функции *L1 =* *L(x, alpha\*, loc, scale)*, где *alpha\** – хвостовой параметр распределения Фреше, оцененный через ММП по выборке максимальных наблюдений, и значение *L2 =* *L(x, alpha’, loc, scale)*, где *alpha’* – тот же параметр, но оцененный через ММП по выборке минимальных значений.Далее рассчитывается значение тестовой статистики *LR = –2 \* (L1 – L2).* В случае корректности нулевой гипотезы, то есть равенства двух моделей, статистика сходится асимптотически к распределению Хи-квадрат с одной степенью свободы по теореме Самуэля Вилкса (Wilks, 1938). Одна степень свободы выбирается, поскольку в обеих моделях оценивается по три параметра, при этом два из них предполагаются равными (location, scale). Различие в последнем параметре *(alpha)* и создает одну степень свободы. Таким образом, в случае данной работы подобным образом сравниваются хвостовые индексы (tail indecies) 2,13 и 1,53 (окно 3 месяца), а также 2,45 и 1,74 (окно 6 месяцев) и пара 2,84 и 2,10 (окно 12 месяцев). Для окна 3 месяца имеем LR = 17,45 (p-value ~ 0,00003), для окна 6 месяцев имеем LR = 9,02 (p-value = 0,00268), для окна 12 месяцев – LR = 3,26640 (p-value = 0,07013). Как следствие, гипотеза равенства параметров *alpha* для минимальных и максимальных наблюдений отвергается на уровне уверенности 1% для окон 3 и 6 месяцев и на уровне 10% для окна 12 месяцев (вероятно, ввиду более низкого числа наблюдений для данного окна). Это означает, что распределение минимумов действительно имеет более тяжелый хвост, чем распределение максимумов. Данный результат повторяет качественный вывод, полученный Франсуа Лонгином для рынка акций (в его работе, однако, LR-тест конкретно для равенства показателей *alpha* не применялся).

В-третьих, для минимальных значений параметр *alpha* меньше двух, что означает, что второй момент распределения бесконечен. Другими словами, распределение Фреше с данными параметрами имеет бесконечную дисперсию. Как следствие, при рассмотрении такой случайной величины на практике не стоит применять понятие выборочной дисперсии, либо считать дисперсию равной некоторму значению, полученному параметрическими методами.

*Рисунок 17. Эмпирическое распределение логарифма дневной доходности (серым) и функции распределения вероятности, соответствующие распределению Frechet для максимальных и минимальных значений за окна 3, 6 и 12 месяцев с оцененными через ММП параметрами – на примере рейтинга BB*

**

**2.3. Возможные практические применения**

Далее описываются методы потенциального применения полученных результатов для практических задач портфельного управления и риск-менеджмента.

Экстремальные наблюдения долгового рынка могут объясняться целым рядом причин – например, в основе месячных и дневных отклонений 2008 и 2020 годов лежали совершенно различные фундаментальные факторы. Несмотря на это, в работе было показано, что максимумы и минимумы могут быть семплированы из широко изученного в статистическом мире распределения Фреше. Также было установлено, что с помощью модели ARMA-GARCH возможно «понизить» толщину хвостов доходности индексов ВДО, упростив задачу до стресс-тестирования остатков (инноваций) этого процесса. Эти два вывода можно применить на практике способами, описанными ниже.

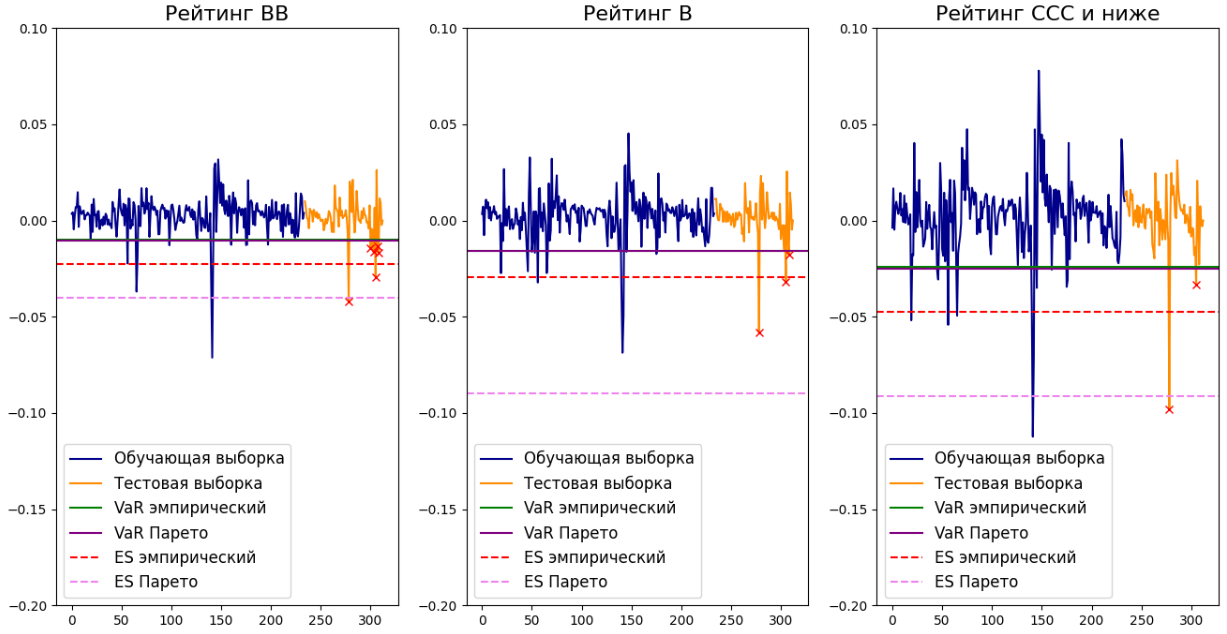
Во-первых, EVT (распределение Фреше) в применении к рынку ВДО можно применять для расчета лимитов маржинальной торговли для брокеров и трейдинговых компаний, использующих производные финансовые инструменты. Рынок деривативов известен своей повышенной зависимостью от значимых изменений в ценах инструментов, и особая опасность находится в краевых наблюдениях. Таким образом, брокеров и трейдинговых компаний в рамках их риск-моделей может интересовать вероятность и последствия краевых дневных и/или месячных изменений за определенное среднесрочное временное окно. Как показано в работе, эти краевые наблюдения не являются статистическими выбросами в фреймворке Теории экстремальных значений – другими словами, с использованием описанных выше регрессий и/или распределения Фреше с оцененными параметрами трейдинговая фирма может оценить, какое наибольшее дневное или месячное отклонение стоит ожидать за период из следующих N месяцев.

Во-вторых, аналогичным образом, результаты работы могут быть использованы для установки требований по капиталу для финансовых организаций, включая банки и пенсионные фонды. Это может достигаться с помощью использования того же распределения Фреше, как было описано выше, либо путем интеграции в риск-модели этих организаций условных гетероскедастичных моделей с остатками в форме распределений Стьюдента или Парето, и последующего стресс-тестирования и расчетов таких метрик как VaR и ES. Этот подход состоятелен с принципами управления кредитным риском, установленными Базельским комитетом еще в 1999 году (Basel committee on Banking Supervision, 1999), где помимо прочего подчеркивается необходимость проводить стресс-тестирование предпосылок по волатильности (то есть учитывать гетероскедастичность и/или отсутствия конечной дисперсии как таковой для фундаментальной случайной величины). Примечательно, что акцент на рынке ВДО, сделанный в этой работе, полезен для более рискофобных институциональных инвесторов, таких как пенсионные фонды, для которых ВДО может служить более привлекательным инструментом, чем рынок акций.

В качестве примера использования результатов работы для расчета указанных выше метрик на Рисунке 11 приведен расчет ежемесячных VaR и ES на основе i.i.d. модели и распределения Парето (как наиболее простой из моделей, подтвержденной тестами). Можно увидеть, что VaR практически идентичен для этих двух подходов, однако ES становится кратно больше при расчете через распределение Парето. Этот эффект достигается за счет того, что при использовании эмпирического распределения предполагается невозможность получения наблюдения более экстремального, чем исторический минимум, а также эмпирическое распределение дискретно, что не позволяет качественно оценить эмпирический ES в хвостах, где, по определению, исторических наблюдений доступно немного. Использование ES по распределению Парето решает эти проблемы, и, как можно видеть, на тестовой выборке единственное превышение реализовавшейся величины (среди тестовой части) над ESПарето в рейтингах BB и CCC и ниже было весьма небольшим, в то время как ESэмпирический превышался значительно.

Стоит отметить, что значимая опасность использования ES на основе распределения Парето заключается в возможности получить чрезмерно пессимистичный порог при уменьшении параметра *alpha.* Так, для трех индексов (75% каждой из выборок) оцененные параметры были 1.54, 1.28 и 1,42. На Рисунке 18 можно увидеть, что снижение alpha с 1,54 до 1,28 увеличивает ES более чем в 2 раза.[[1]](#footnote-1)

*Рисунок 18. Value-at-Risk и Expected shortfall, рассчитанный по 75% выборки по эмпирическому распределению (наиболее простой метод) и по стабильному распределению Парето*



Другим возможным, однако более амбициозным, применением может быть решение задачи аллокации активов в рамках инвестиционного портфеля. В книге «Safe Haven: Investing for Financial Storms» Марк Спитзнагел с помощью симуляций Монте-Карло показывает, как при предположении свойств распределения доходности активов (включая безрисковые, такие как денежные средства) инвестор может увеличить стоимость своего портфеля *в медианном случае* на горизонте 25 лет с помощью вложения части портфеля в актив с нулевой или даже отрицательнойожидаемойдоходностью (Spitznagel, 2021). Перенося аналогию на настоящую работу, полученные модели можно использовать для сбора инвестиционного портфеля путем максимизации его *долгосрочного темпа роста* стоимости (compound annual growth rate, CAGR) через совершение Монте-Карло симуляций. В данном случае аллокация портфеля может происходить только между рынком ВДО (например, ETF высокодоходных облигаций) и безрисковым активом (денежные средства) ввиду практического отсутствия корреляции между ними. Добавление же рынка акций и/или других финансовых активов в задачу оптимизации портфеля создает необходимость крайне точно оценивать устойчивые корреляции между инструментами, что являлось бы отдельной задачей, труднореализуемой на практике – о чем в частности заявляют Спитзнагел и Талеб (Spitznagel, 2021; Taleb, 2019). Этот подход является некоторой альтернативой Портфельной теории Марковица, однако требует значительных дополнительных исследований, выходящих за рамки данной работы.

**2.4. Выводы по исходным гипотезам**

По итогу проведенного анализа можно выделить следующие результаты, исходя из изначально поставленных гипотез:

**Гипотеза 1:** распределение логарифма дневной и месячной доходностей индексов высокодоходных облигаций (далее по тексту – «переменные») не может считаться нормальным согласно статистическим тестам на уровне значимости 0,1%. На практике это проявляется в виде наличия значимо более высокого эксцесса (более толстых хвостов), а также нестабильности четвертого момента, рассчитанного по выборке данных – от 18 до 68% эксцесса приходится на одно крайнее наблюдение, в зависимости от рассматриваемого рейтинга, при том что для нормального распределения этот показатель должен быть на уровне 0,8% +- 0,28%. Таким образом, гипотеза 1 подтверждается.

**Гипотеза 2:** принадлежность месячных величин t-распределению не отвергается статистическими тестами с p-value >0,25. Это означает, что месячныепеременные действительно могут быть смоделированы распределением Стьюдента со степенями свободы 2,59, 2,09 и 2,30 для рейтингов BB, B и CCC и ниже, соответственно. Для дневных величин, однако, эта гипотеза отвергается тестами на уровне уверенности 1%, несмотря на наличие хорошей визуальной схожести эмпирического распределения и предполагаемого распределения Стьюдента. Это является проявлением того вывода, что дневная переменная не имеет асимптотической сходимости к предполагаемому t-распределению, что может объясняться прежде всего нестабильностью параметров *(df, loc, scale)* во времени и наличием автокорреляции.

Что касается стабильного распределения Парето, гипотеза принадлежности месячной величины к нему также не отвергается, причем получены более низкие тестовые статистики в случае данного распределения, что говорит о более хорошем качестве подгонки. Для дневных доходностей гипотеза отвергается для рейтингов B и CCC и ниже, и для рейтинга BB не отвергается лишь на уровне 1% и только CM-тестом.

Обобщая, можно сказать, что гипотеза 2 не отвергается для месячных доходностей (т.е. подходят оба распределения), но отвергается для дневных доходностей.

**Гипотеза 3:** принадлежностьдневных и месячныхинновациймодели ARMA(1,1)-GARCH(1,1) к распределению Стьюдента и стабильному распределению Парето не отвергается на уровне уверенности 25% для месячных и 1-5% для дневных данных (см. Таблицу 10 для детализации до индексов разных кредитных рейтингов).

**Гипотеза 4:** применении модели ARMA-GARCH инновации процесса действительно демонстрируют менее тяжелые хвосты, чем исходные переменные. Это отражается в более высоком числе степеней свободы *df* в распределении Стьюдента (3,17–3,38 для недельных и 3,49–4,25 для дневных доходностей в ARMA-GARCH, но 2,09–2,59 и 1,80–1,89 для исходных переменных) и более высокой степени показателя *alpha* в стабильном распределении Парето (1,59–1,76 для инноваций ARMA-GARCH и 1,33-1,50 для исходных месячных переменных, что является большой разницей в рамках распределения Парето) при оценке параметров с помощью ММП. Это означает, что применение условной гетероскедастичной модели позволяет «объяснить» часть отклонений индекса и «упростить» задачу моделирования потенциальных изменений индекса. Таким образом, например, для расчета VaR и ES в данном случае будет необходимо считать соответствующий процентиль от распределения с существенно более тонкими хвостами. Гипотеза 4, следовательно, подтверждается.

**Гипотеза 5:** экстремальные значения логарифмов доходностей за окна 3, 6 и 12 месяцев действительно могут быть смоделированы распределением Фреше, как одним из трех классов распределений Теории экстремальных значений. Этот вывод получен теми же статистическими тестами с p-value >0,25 во всех случаях (однако с большей уверенностью для 6-месячных и 12-месячных окон). Примечательно, что распределение минимальных доходностей имеет существенно более тяжелые хвосты в соответствии со значениями параметра *alpha* (хвостовой индекс) распределения Фреше (1,53–1,74 для минимальных и 2,13–2,45 для максимальных значений – на примере логарифмов дневных доходностей индекса BB). Этот вывод, в частности, служит косвенным аргументом в пользу принадлежности дневной переменной к стабильному распределению Парето с *некоторыми* параметрами.

**Заключение**

В ходе данной работы были изучены статистические свойства логарифмов дневных и месячных доходностей индексов высокодоходных облигаций трех групп кредитных рейтингов эмитентов США (далее по тексту – «величин»), а также были предложены 3 метода моделирования этих случайных величин для целей риск-менеджмента.

Во-первых, была рассмотрена принадлежность данных величин к распределениям, предполагающие независимость наблюдений – нормальное, Стьюдента и стабильное распределение Парето (иначе называемое Levy Stable). Данная принадлежность не была отвергнута для месячных величин для распределений Стьюдента и Парето, но отвергнута для дневных данных. В ходе анализа были представлены аргументы неприменимости данного подхода для моделирования дневных величин, среди которых – наличие неучтенных эффектов, таких как непостоянство параметров во времени, наличие автокорреляции и гетероскедастичность.

Во-вторых, был проверен эффект применения модели ARMA-GARCH, которая призвана учесть часть описанных выше эффектов. В данной модели за «корневую» случайную величину принимаются инновации процесса. Было показано, что применение этой модели позволяет «снизить» толщину хвостов по сравнению с изначальной величиной, что упрощает стресс-тестирование и другие задачи риск-менеджмента. Также показано, что принадлежность инноваций к распределению Стьюдента и стабильному распределению Парето не отвергается для месячных и дневных данных на выборке существенного размера (~26 лет).

В-третьих, был применен инструментарий Теории экстремальных значений для рассмотрения краевых, то есть максимальных и минимальных, дневных наблюдений за 3-, 6- и 12-месячные окна. Применение логарифмических трансформаций исходных данных показало необходимость применения распределения Фреше (Frechet) как одного из трех распределений данной сферы статистики. Как итог, на высокой степени уверенности не была отвергнута принадлежность краевых наблюдений к данному распределению. Кроме того, была показана и количественно посчитана ассиметрия в виде наличия статистически значимо более тяжелых хвостов для минимальных наблюдений по сравнению с максимальными, а также отсутствие конечности дисперсии для оцененных распределений минимальных значений. Эти факторы служат важными статистическими свойствами на практике.

Наконец, были описаны сферы и способы применения полученных результатов для практических задач риск-менеджмента, включая расчет Value-at-Risk, Expected Shortfall и стресс-тестирования внутренних риск-моделей финансовых организаций. Также приведенный пример расчета VaR и ES на основе стабильного распределения Парето как наиболее простой модели показал важность изучения свойств популяции случайных финансовых величин и опасность использования эмпирических распределений для расчета риск-метрик.

Краткое количественное описание некоторых выводов, полученных в ходе работы, сформулировано в Таблицах 10 и 11.

*Таблица 10. Результаты рассмотрения моделей i.i.d.-распределений и ARMA(1,1)-GARCH(1,1)*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Рейтинг** | **Период доходности** | **Безусловные модели** | | | **Инновации ARMA-GARCH** | | |
|  |  | **Нормальное** | **Стьюдент** | **Парето** | **Нормальное** | **Стьюдент** | **Парето** |
| BB | Дневные | Отвергается | Отвергается | **alpha = 1,37\***  **beta = -0,17\*** | Отвергается | **df = 4,25\*\*** | **alpha = 1,69\*\***  **beta = -0,32\*\*** |
| B | Отвергается | Отвергается | Отвергается | Отвергается | **df = 3,49\*** | **alpha = 1,65\***  **beta = -0,34\*** |
| CCC и ниже | Отвергается | Отвергается | Отвергается | Отвергается | **df = 3,79\*\*** | **alpha = 1,66\***  **beta = -0,21\*** |
| BB | Месячные | Отвергается | **df = 2,59\*\*\*** | **alpha = 1,50\*\*\***  **beta = -0,24\*\*\*** | Отвергается | **df = 3,38\*\*\*** | **alpha = 1,76\*\*\***  **beta = -0,70\*\*\*** |
| B | Отвергается | **df = 2,09\*\*\*** | **alpha = 1,33\*\*\***  **beta = -0,25\*\*\*** | Отвергается | **df = 3,17\*\*\*** | **alpha = 1,59\*\*\***  **beta = -0,61\*\*\*** |
| CCC и ниже | Отвергается | **df = 2,30\*\*\*** | **alpha = 1,42\*\*\***  **beta = -0,27\*\*\*** | Отвергается | **df = 3,35\*\*\*** | **alpha = 1,64\*\*\***  **beta = -0,44\*\*\*** |

Легенда: «отвергается» - принадлежность выборки к распределению отвергается на уровне 1% тестом CM;   
\*\*\* p-value > 25%, CM-тест принадлежности выборки к обозначенному распределению с заданными параметрами  
\*\* p-value > 5%  
\* p-value > 1%

*Таблица 11. Результаты рассмотрения распределения Фреше, выбранного из трех распределений Теории экстремальных значений*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | ***Распределение Фреше*** | | |
| **Индекс** | **Тип значения** | **Размер окна** | ***alpha*** | **Положение**  **(m)** | **Размерность**  **(s)** |
| Рейтинг BB | Максимальное | 3 мес. | **2,13\*\*\*** | **0,0** | **0,0016** |
| 6 мес. | **2,45\*\*\*** | **0,0** | **0,0020** |
| 12 мес. | **2,84\*\*\*** | **0,0** | **0,0026** |
| Минимальное | 3 мес. | **1,53\*\*\*** | **0,0** | **0,0015** |
| 6 мес. | **1,74\*\*\*** | **0,0** | **0,0022** |
| 12 мес | **2,10\*\*\*** | **0,0** | **0,0033** |
| Рейтинг B | Максимальное | 3 мес. | **1,68\*\*\*** | **0,0** | **0,0016** |
| 6 мес. | **1,76\*\*\*** | **0,0** | **0,0019** |
| 12 мес. | **1,89\*\*\*** | **0,0** | **0,0027** |
| Минимальное | 3 мес. | **добавить** | **0,0** | **добавить** |
| 6 мес. | **1,56\*\*\*** | **0,0** | **0,0026** |
| 12 мес | **1,75\*\*\*** | **0,0** | **0,0037** |
| Рейтинг CCC  и ниже | Максимальное | 3 мес. | **1,79\*\*\*** | **0,0** | **0,0024** |
| 6 мес. | **1,90\*\*\*** | **0,0** | **0,0031** |
| 12 мес. | **2,25\*\*\*** | **0,0** | **0,0043** |
| Минимальное | 3 мес. | **1,48\*\*\*** | **0,0** | **0,0025** |
| 6 мес. | **1,65\*\*\*** | **0,0** | **0,0037** |
| 12 мес | **1,99\*\*\*** | **0,0** | **0,0053** |

Легенда: \*\*\* p-value > 25%, CM-тест принадлежности выборки к распределению Фреше с заданными параметрами

**Список литературы**

1. Markowitz, H. The Utility of Wealth. / H. Markowitz // [*Journal of Political Economy*](https://www.journals.uchicago.edu/journal/jpe)*.* – 1952 – Vol. 60, No. 2 – Pp. 151-158
2. Sharpe, W. Mutual Fund Performance / W. Sharpe // *The Journal of Business.* – Vol. 39, No. 1 – Pp. 119–138.
3. Mandelbrot, B. The Variation of Some Other Speculative Prices. / B. Mandelbrot // The Journal of Business. – 1963 – Vol. 40, No. 4 – Pp. 393–413. <http://www.jstor.org/stable/2351623> (1967??)
4. FDIC <https://www.fdic.gov/bank/historical/history/3_85.pdf>
5. [Glenn Yago](https://www.google.es/search?tbo=p&tbm=bks&q=inauthor:%22Glenn+Yago%22) Oxford University Press, 1991 – Junk Bonds: How High Yield Securities Restructured Corporate America
6. Bachelier L. The Theory of Speculation / L. Bachelier // Annales scientifiques de l’Ecole Normale Superieure – 1900 – Vol. 3, No. 17 – Pp. 21-86. <https://fermlab.hse.ru/data/2013/06/09/1283651835/Bachelier%20(en).pdf>
7. Fama, E. Mandelbrot and the Stable Paretian Hypothesis / E. Fama // The Journal of Business. – 1963 – Vol. 36, No. 4, Pp. 420–429. <http://www.jstor.org/stable/2350971>
8. Fama, E. The Behavior of Stock-Market Prices. / E. Fama // The Journal of Business, 1965 – Vol. 38 – No. 1, Pp. 34–105 <http://www.jstor.org/stable/2350752>
9. Longin, F. The Asymptotic Distribution of Extreme Stock Market Returns. / F. Longin // The Journal of Business – 1996 – Vol. 69 – No. 3 – Pp. 383–408. <http://www.jstor.org/stable/2353373>
10. Rachev, S.T., Stoyanov, S.V., Biglova, A., Fabozzi, F.J. (2005). An Empirical Examination of Daily Stock Return Distributions for U.S. Stocks. In: Baier, D., Decker, R., Schmidt-Thieme, L. (eds) Data Analysis and Decision Support. Studies in Classification, Data Analysis, and Knowledge Organization. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-28397-8\_30
11. Hong, Lin, Wu, 2012
12. Harris R., Küçüközmen C. The Empirical Distribution of UK and US Stock Returns / R. Harris, C. Küçüközmen // Journal of Business Finance – Vol. 28 – No. 5-6 – 2001 – Pp. 715-740
13. Gabriel C., Lau C. On the distribution of European government bond returns: Empirical evidence / C. Gabriel, C. Lau // 2014  
    <https://efmaefm.org/0efmameetings/EFMA%20ANNUAL%20MEETINGS/2012-Barcelona/papers/On_the_distribution_of_European_government_bond_returns_gabriel_Lau.pdf>
14. Vukovic D., Lapshina K., Maiti M. European Monetary Union bond market dynamics: Pre & post crisis / D. Vukovic, K. Lapshina, M. Maiti // Research in International Business and Finance – Vol. 50 – 2019 – Pp. 369-380 <https://www.sciencedirect.com/science/article/pii/S0275531918308730>
15. Taleb N. Statistical consequences of fat tails: Real world preasymptotics, epistemology, and applications / N. Taleb // arXiv preprint arXiv:2001.10488. – 2020.
16. FRED <https://fred.stlouisfed.org/series/BAMLH0A2HYBEY>
17. FRED <https://fred.stlouisfed.org/series/BAMLH0A1HYBBEY>
18. FRED <https://fred.stlouisfed.org/series/BAMLH0A3HYCEY>
19. Mittnik S., Paolella M., Rachev S. Diagnosing and treating the fat tails in financial returns data / S. Mittnik, M. Paolella, S. Rachev // Journal of Empirical Finance – 2000 – Vol. 7 – No. 3-4, Pp. 389-416 <https://www.sciencedirect.com/science/article/pii/S0927539800000190>
20. Tokat, Y., Rachev, S. T., Schwartz, E. S. The stable non-Gaussian asset allocation: a comparison with the classical Gaussian approach / // Journal of Economic Dynamics and Control – 2003 – Vol. 27 – No. 6, Pp. 937-969 <https://www.sciencedirect.com/science/article/pii/S0165188902000507>
21. Fréchet, M. (1927) Sur la loi de Probabilité de l'écart Maximum. Annales de la Société Polonaise de Mathematique, 6, 93-116.
22. Fisher, R.A. and Tippett, L.H.C. (1928) Limiting Forms of the Frequency Distribution of the Largest or Smallest Members of a Sample. Proceedings of the Cambridge Philosophical Society, 24, 180-190.  
    <http://dx.doi.org/10.1017/S0305004100015681>
23. Misés, R. D. (1936). Les lois de probabilité pour les fonctions statistiques. In Annales de l'institut Henri Poincaré (Vol. 6, No. 3-4, pp. 185-212). <http://www.numdam.org/item/AIHP_1936__6_3-4_185_0.pdf>
24. Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. Annals of mathematics, 423-453. <https://www.jstor.org/stable/1968974>
25. Hans et al, 2007
26. Dekkers, A. L., Einmahl, J. H., & De Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. The Annals of Statistics, 1833-1855. <https://www.jstor.org/stable/2241667>
27. Kotz, S., & Nadarajah, S. (2000). Extreme value distributions: theory and applications. world scientific .<https://books.google.com/books?hl=en&lr=&id=GwBqDQAAQBAJ&oi=fnd&pg=PR5&dq=28.%09Kotz,+Nadarajah,+2000&ots=DDjd8w1azL&sig=6dcrst4VmZ_T45-YM5_78-RWujE>
28. Massey, F. J. (1952). Distribution table for the deviation between two sample cumulatives. The annals of mathematical statistics, 23(3), 435-441.<https://www.jstor.org/stable/2236686>
29. Anderson, T.W. and Darling, D.A. (1954) A Test of Goodness of Fit. Journal of American Statistical Association, 49, 765-769.  
    <https://doi.org/10.1080/01621459.1954.10501232>
30. Arshad et al., 2003; <https://www.researchgate.net/publication/46027298_Anderson_Darling_and_Modified_Anderson_Darling_Tests_for_Generalized_Pareto_Distribution>
31. Razali, N. M., & Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. Journal of statistical modeling and analytics, 2(1), 21-33. <https://www.nbi.dk/~petersen/Teaching/Stat2017/Power_Comparisons_of_Shapiro-Wilk_Kolmogorov-Smirn.pdf>
32. Cramér, H. (1928). On the composition of elementary errors: First paper: Mathematical deductions. Scandinavian Actuarial Journal, 1928(1), 13-74. <https://www.tandfonline.com/doi/abs/10.1080/03461238.1928.10416862>
33. Von Mises, 1931
34. Farrell, Patrick & Rogers-Stewart, Katrina. (2006). Comprehensive study of tests for normality and symmetry: Extending the Spiegelhalter test. Journal of Statistical Computation and Simulation. 76. 803-816. 10.1080/10629360500109023. <https://www.researchgate.net/publication/233339727_Comprehensive_study_of_tests_for_normality_and_symmetry_Extending_the_Spiegelhalter_test>
35. Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. Journal of the American statistical Association, 69(347), 730-737. <https://www.tandfonline.com/doi/abs/10.1080/01621459.1974.10480196>
36. James, G. S. (1959). [Review of Statistics of Extremes., by E. J. Gumbel]. Journal of the Royal Statistical Society. Series A (General), 122(2), 243–245. <https://doi.org/10.2307/2342609>
37. Basel committee on Banking Supervision, 1999 <https://www.bis.org/publ/bcbs61.pdf>
38. Safe Haven: Investing for Financial Storms» Mark Spitznagel ISBN 10: 1119401798 ISBN 13: 9781119401797 <https://www.abebooks.com/9781119401797/Safe-Haven-Investing-Financial-Storms-1119401798/plp> Publisher: Wiley, 2021
39. [https://core.ac.uk/download/pdf/41363083.pdf Hong et al. 2012](https://core.ac.uk/download/pdf/41363083.pdf%20Hong%20et%20al.%202012)
40. [Wilks, Samuel S.](https://en.wikipedia.org/wiki/Samuel_S._Wilks) (1938). ["The large-sample distribution of the likelihood ratio for testing composite hypotheses"](https://doi.org/10.1214%2Faoms%2F1177732360). The Annals of Mathematical Statistics. 9 (1): 60–62. [doi](https://en.wikipedia.org/wiki/Doi_(identifier)):[10.1214/aoms/1177732360](https://doi.org/10.1214%2Faoms%2F1177732360).)

1. Индекс рейтинга CCC и ниже практически равен индексу рейтинга B в терминах ES ввиду более высокого параметра размерности, scale, что тоже можно наблюдать на графике в виде более широких вариаций месячных значений. [↑](#footnote-ref-1)