**СОДЕРЖАНИЕ**

с.

ВВЕДЕНИЕ……………………………………………………………………...

1 Случайный процесс с тяжелыми хвостами………………………………….

1.1 Логарифмическая доходность……………………………………….

1.2 Распределение доходности валютного курса……………………….

2 Альтернативные модели………………………………………………………

2.1 Обобщенное гиперболическое распределение……………………..

2.2 Нормальное обратное гауссовское распределение…………………

2.3 Гиперболическое распределение…………………………………….

2.4 Гиперболический секанс……………………………………………..

2.5 Смесь гауссиан………………………………………………………..

2.5.1 Определение смеси…………………………………………..

2.5.2 Свойства смеси……………………………………………….

2.5.3 Статистическое оценивание параметров смешивающего распределения……………………………………………………...

ЗАКЛЮЧЕНИЕ………………………………………………………………….

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ…………………………...

**ВВЕДЕНИЕ**

Известно, что в окружающем нас мире часто встречается нормальное распределение. Многие случайные величины хорошо описываются нормальным распределением: параметры организма человека (рост, вес человека), измерительные погрешности, случайная продолжительность процессов (время заплыва на соревнованиях по плаванью или забега на соревнованиях по бегу) и многое другое.

Однако данные случайных процессов в финансовой математике значительно реже соотносятся с нормальным распределением. Замечено, что распределение доходности большинства финансовых индексов (ценных бумаг, валютных курсов) характеризуются тяжелыми, по сравнению с нормальным распределением, хвостами, асимметрией и положительным эксцессом.

Вместе с тем, в инвестиционной среде часто игнорируются эти особенности и при моделировании разброса значений финансовых возвратов делается допущение об их «нормальности». Распространенность такого подхода объясняется простотой построения модели, для которой требуется лишь два параметра, и наличием большого количества подходящих статистических тестов.

Возможно ли говорить об адекватной оценке рисков, предполагая, что данные распределены по нормальному закону? Какими проблемами может обернуться такой подход?

В данной работе исследуется состоятельность нормального распределения для моделирования доходности валютного курса EUR/USD и описываются альтернативные модели для приближения и оценки эмпирических данных: обобщенное гиперболическое распределение, обратное гауссовское распределение, гиперболическое-секанс распределение и смесь гауссовских распределений (EM-алгоритм).

Актуальность данной работы объясняется тем, что существует запрос на надежные и гибкие модель для достоверной оценки возможных значений доходности валютных курсов.

Данные для исследования представляют собой логарифмические доходности значений курса EUR/USD. Котировки EUR/USD за день/час в период с 05.05.2003 по 29.04.2020 и за минуту с 29.04.2019 по 29.04.2020 взяты с сайта швейцарского банка Dukascopy Bank.

1. **Случайный процесс с тяжелыми хвостами**

**1.1 Логарифмическая доходность**

Динамика финансовых индексов имеет стохастическую природу. На примере наблюдений изменения валютного курса EUR/USD это хорошо заметно. В анализе изменений финансовых индексов распространённым приёмом является переход от наблюдения за исходными значениями к наблюдению за изменением доходности, рассчитанной на основе исходных значений. Это удобно и наглядно, потому что главным образом интерес состоит в изучении изменения прибыли в фиксированные промежутки времени, а также в вычислении некоторых финансовых показателей (размах накопленных сумм, отклонения от размаха накопленных сумм). Кроме того, в сравнение с графиком изменения самой цены, график изменения доходности ведет себя более “однородно”.

Рассмотрим два подхода к вычислению доходности на примере валютного курса EUR/USD: процентная доходность и логарифмическая доходность. Положим - значение курса валют в момент времени , а - значение курса в предыдущий момент времени. Тогда процентная доходность в период с по выражается формулой:

Для вычисления логарифмической доходности применяется формула:

Чаще предпочтение отдается именно логарифмической доходности, чему есть ряд причин. Во-первых, распределение такой доходности будет иметь одинаковый (симметричный) вид как для значений евро в долларах, так и для значений доллара в евро. Во-вторых, логарифмическая доходность обладает аддитивностью по времени, что позволяет легко разложить доходность за 3 года, к примеру, на сумму доходностей за каждый год по отдельности в виде:

В-третьих, в отличие от самой стоимости валюты, для её логарифмической доходности можно задавать уровень, относительно которого происходят изменения.

Далее в работе под доходностью валютного курса следует понимать логарифмическую доходность.

**1.2 Распределение доходности валютного курса**

Рассмотрим данные валютного курса EUR/USD за день в период с 05.05.2003 по 29.04.2020 (Рисунок 1.2.1).

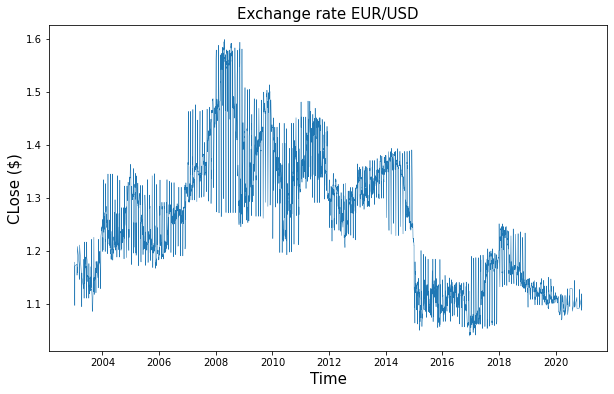


Рисунок 1.2.1 Дневные значения валютного курса EUR/USD с 05.05.2003 по 29.04.2020

Преобразуем данные, вычислив доходность за каждый день, и получим обновленное представление о курсе:

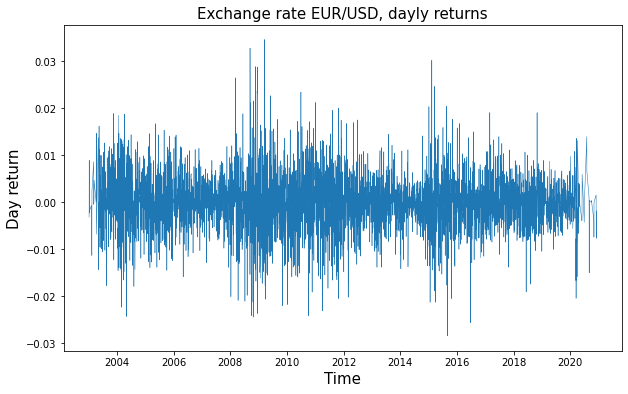


Рисунок 1.2.2 Дневные значения доходности для валютного курса EUR/USD с 05.05.2003 по 29.04.2020

На Рисунке 1.2.2 заметно, что данный временной ряд имеет высокую волатильность в кризисные 2008 и 2015 годы.

Перейдем к изучению распределения доходности (Рисунок 1.2.3). На первый взгляд кажется, что мы имеем дело с нормальным распределением. Построим нормальное распределение, вычислив выборочное среднее и стандартное отклонение для имеющихся данных:

,

Заметно, что гистограмма распределения имеет сильно вытянутую вершину и тяжёлые хвосты относительно модели, основанной на предположении о “нормальности” изучаемого случайного процесса.

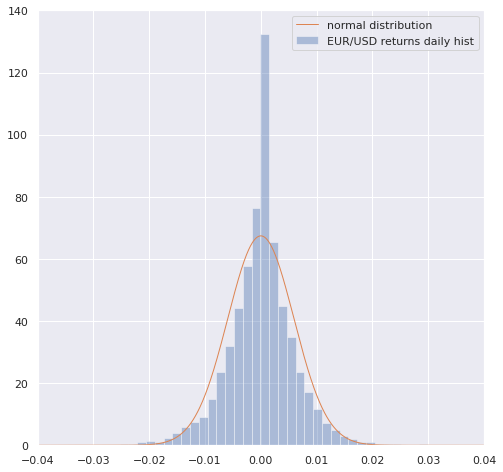


Рисунок 1.2.3 Нормальное распределение доходности

Эти особенности проявляются в большей степени при рассмотрении логарифма плотности распределения (Рисунок 1.2.4). На данном рисунке хорошо заметно, что в действительности хвосты реальных данных находятся выше нормального распределения. Получается, что частота появлений экстремальных событий на валютном рынке недооценивается. Значительное падение доходности или же её значительный рост на самом деле происходят чаще, чем предполагается нормальной моделью.

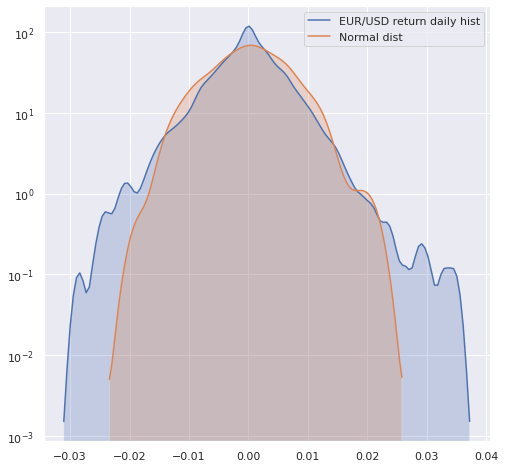


Рисунок 1.2.4 Плотность распределения доходности с логарифмическим масштабом оси ординат

Для большей убедительности рассмотрим квантильный график. На квантильном графике (QQ-plot) отображены квантили двух распределений. Точки нормального QQ-plot дают представление о нормальности рассматриваемого набора данных. Если данные распределены нормально, точки выстроятся на базовой линии, проходящей под углом 45 градусов. Если данные не распределены нормально, точки отклоняться от базовой линии. Значения на оси абсцисс - квантили нормального распределения, значения по оси ординат - квантили наблюдаемых значений. На Рисунке 1.2.5 видно, что большая часть наблюдений по центру попадает на базовую линию, хотя стоит отметить, что у квантилей доходности заметен s-образный “перегиб”. Возможно, имеет место бимодальность изучаемого распределения. На концах же заметны существенные отклонения.

Таким образом, подтверждается предположение о том, что у валютного курса EUR/USD наблюдается больше событий с сильным падением или ростом цен, чем ожидается при нормальном распределении.

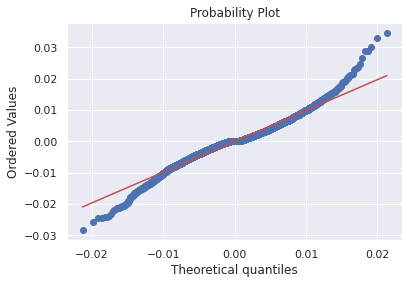
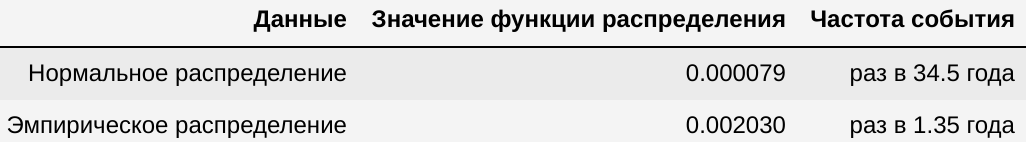


Рисунок 1.2.5 Квантильный график доходности

Таблица 1.2.1 Сравнение хвостов распределений.



Возьмем несколько точек, отклоняющихся от хвоста нормального распределения. Вычислив значения для нормальной и эмпирической функций распределения на левом хвосте, например, в точке -0.022356, легко заметить, насколько существенно нормальная модель недооценивает частоту экстремальных событий (Таблица 1.2.1). Такой подход совершенно не годится и для анализа, и моделирования стохастических процессов с тяжелыми хвостами.

Необходимость в точной оценке хвостов и вершины распределения значений доходности, приводит нас к поиску более состоятельных моделей, которые будут учитывать параметры эксцесса и асимметрии данных.

**2 Альтернативные модели**

**2.1 Обобщенное гиперболическое распределение**

Обобщенное гиперболическое распределение (GH) было описано в контексте анализа и моделирования форм песчаных дюн в 1977 г. Barndorff-Nielsen[1]. Данное распределение обладает свойствами, подходящими для описания процессов с тяжелыми хвостами:

* хвосты GH-распределения тяжелее, чем у нормального распределения [2];
* GH-распределение учитывает асимметрию распределения.

Благодаря своей гибкости, модели на основе обобщенного гиперболического распределения стали популярны областях теоретической и практической статистик и получили особое распространение в геоморфологии, теории турбулентности и финансовой математике. В 1995 г. Eberlein и Keller попробовали применить GH-распределение к анализу доходности [3].

Обобщенное гиперболическое распределение описывается пятью параметрами и имеет вид:

, (1)

где и - параметры положения и масштаба;

- параметр асимметрии;

- параметр устойчивости;

- параметр, который определяет подкласс обобщенных гиперболических распределений.

Функция - модифицированная функция Бесселя 3 типа с индексом , также известная как функция МакДональда. Она задается следующим образом:

Функция GH-распределения является безгранично делимой и для нее определены все моменты, а её плотность унимодальна. Форма плотности же позволяет моделировать распределения с положительным эксцессом (лептокуртические), положительным или отрицательным коэффициентом асимметрии и тяжелыми хвостами. С точки зрения поведения “хвостов”, это распределение занимает промежуточное положение между устойчивыми распределениями с индексом и гауссовским распределением : их хвосты убывают быстрее, чем у устойчивых распределений, но медленнее гауссовских.

Рассмотрим некоторые подклассы обобщенного гиперболического распределения.

**2.2 Нормальное обратное гауссовское распределение**

Согласно данным статистического анализа приращений цен на датских и немецких биржах, проведенного Барндорфф-Нильсеном и его коллегами, распределение логарифмов приращений биржевых цен хорошо аппроксимируется, так называемым, нормальным обратным гауссовским распределением [4].

Семейство GH-распределения включает в себя широкий спектр распределений с разными свойствами. Одним из наиболее привлекательных для нашей задачи является нормальное обратное гауссовское распределение (NIG-распределение). Если взять для уравнения (1), получаем NIG-распределение:

где и .

Привлекательность именно этого подкласса GH-распределения определяется его уникальным свойством: замкнутостью относительно свертки, в том смысле, что для - независимых случайных величин из NIG-распределения с одинаковыми параметрами асимметрии и устойчивости и любыми и справедлива сумма , которая принадлежит, опять же, NIG-распределению с такими же параметрами асимметрии и устойчивости. При этом параметры положения и масштаба для данного распределения будут представлять собой суммы вида:

и соответственно.

Квантильный график (Рисунок 2.2.1) хорошо иллюстрирует большой потенциал данной модели. Наблюдаются сильные отклонение на хвостах и вогнутости по центру, схожие с реальными данными.

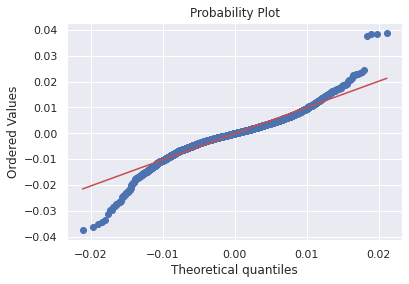


Рисунок 2.2.1 Квантильный график NIG-распределения

**2.3 Гиперболическое распределение**

В ряде работ, посвященных анализу доходностей финансовых индексов ([3], [2]), наблюдается большая схожесть эмпирических данных с моделью, основанной на гиперболическом распределении.

Гиперболическое распределение (HYP-распределение) является подклассом семейства GH-распределений и определяется параметром . Плотность HYP-распределения:

**2.4 Гиперболический секанс**

Гиперболическое-секанс распределение - это симметричное распределение похожее на нормальное. Данный закон распределения случайной величины также определяется двумя параметрами: средним и стандартным отклонением, но имеет более вытянутую и острую вершину.

Плотность распределения задается в форме:

Так как гиперболическое-секанс распределение является лептокуртическим, его хвосты более “тяжелые”, чем у нормального распределения (Рисунок 2.4.1). Данные качества делают эту модель неплохой альтернативой для анализа и моделирования стохастических процессов с тяжелыми хвостами, хотя симметричность накладывает некоторые ограничения на эмпирические данные, которые подлежат анализу.

Состоятельность HypSech-распределения возможна на данных, чей коэффициент асимметрии будет стремиться к нулю.

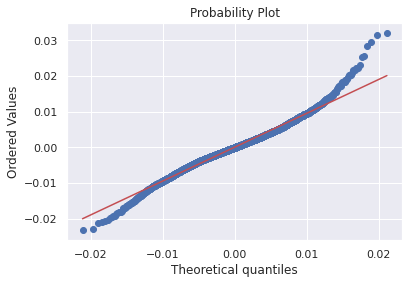


Рисунок 2.4.1 Квантильный график HypSech-распределения

**2.5 Смеси гауссиан**

**2.5.1 Определение смеси**

Рассмотрим специальную функцию распределения на , где 𝕐 – это подмножество *n*-мерного евклидова пространства и с борелевской -алгеброй . Если мы фиксируем , то будет функцией распределения величины , если же мы фиксируем , наблюдается обратное - будет функцией распределения величины . Тогда смесь функции представляется следующим образом:

, (7)

где - вероятностная мера на вероятностном пространстве ().

Распределение называется смешиваемым, в то время как мера задает смешивающее распределение. – смесь функции по . Возьмём **Y** *m*-мерную случайную величину в пространстве () и перепишем функцию распределения:

Для функции плотность распределения , а для смеси , соответственно:

. (9)

В дискретном случае смесь функции распределения приобретает вид:

где - дискретная случайная величина, значения , , … которой возникают с вероятностями , , …, соответственно;

– компоненты смеси;

– веса компонент.

Плотность распределения дискретной смеси:

Вообще говоря, функции распределения в смеси могут представлять собой различные распределения. Для дискретной смеси , а само распределение и его плотность соответственно:

**2.5.2 Свойства смеси гауссиан**

Выше часто упоминалась островершинность как эмпирического распределения, так и теоретических. Для объяснения того, как у смеси распределений появляется такое важное для нас свойство, рассмотрим коэффициент эксцесса для смеси.

Пусть и – независимые случайные величины с конечными четвертыми моментами. Предполагая, что и , имеем:

где – коэффициент эксцесса;

при том (по неравенству Йенсена).

Получается, что правая часть всегда больше или равна . Положив случайной величиной из стандартного нормального распределения, а неотрицательной случайной величиной, независимой от , у которой , получим – случайную величину с одной вершиной и симметричную. Для имеем:

где – функция стандартного нормального распределения;

– плотность распределения случайной величины .

Если выполняется условие нормирования и выходит, что:

и смеси нормальных распределений имеют более тяжелые хвосты и острую вытянутую вершину, в сравнение со стандартным нормальным распределением: (коэффициент эксцесса для нормального распределения равен трем).

Таким образом, существуют все предпосылки для состоятельности данного распределения в качестве модели для анализа доходностей валютного курса EUR/USD.

**2.5.3 Статистическое оценивание смешивающего распределения**

Оценка смешивающего распределения сводится к поиску наиболее близкой к истинной дискретной смеси нормальных распределений. Необходимо найти оптимальное количество распределений, составляющих смесь, и оценить их параметры: их веса в смеси, средние и параметры масштаба.

Пусть – случайная величина, распределение которой является смесь распределений. Имеется конечная выборка размера *n* реализаций данной случайной величины. Предполагается, что смесь состоит из k компонент, являющихся гауссианами. Требуется отыскать такое k распределений, соответствующие им значения параметров и веса, чтобы получившаяся смесь согласовывалась с имеющейся выборкой самым лучшим образом.

То есть, задача сводится к оценке параметров смеси плотностей:

где – значение весов компонент смеси;

– параметры распределений в смеси;

– число распределений в смеси.

Для задач такого рода хороших, хоть и не всегда стабильных, результатов можно добиться применяя EM-алгоритм (Expectation Maximization). В дальнейшем, при приближении эмпирических данных смесью гауссовских распределений, именно этот алгоритм предпочтителен для отыскания оптимального количества компонент смеси, а также их весов и параметров.

**ЗАКЛЮЧЕНИЕ**

В ходе научно-исследовательской работы было проведено вычисление доходностей и исследование данных валютного курса EUR/USD, затем последовала проверка модели нормального распределения для анализа доходностей на состоятельность, после чего были выдвинуты и проанализированы на пригодность альтернативные модели для изучения стохастической природы данных.

Были решены следующие задачи:

– проведен ознакомительный статистический анализ с данными о доходности валютного курса EUR/USD;

– выполнена проверка модели нормального закона распределения на состоятельность;

– описаны особенности распределения с тяжелыми хвостами и риски при неверном выборе статистической модели;

– приведены альтернативные модели для анализа и моделирования стохастического процесса с положительным коэффициентом эксцесса и тяжелыми хвостами

– показана пригодность свойств альтернативных моделей для изучения доходностей валютного курса EUR/USD.

На исследуемых данных было показано, что модель на основе нормального распределения недооценивает вероятность событий на хвостах примерно в 25 раз, что отражено в Таблице 1.2.1 и на Рисунке 1.2.3, Рисунке 1.2.4 и Рисунке 1.2.5.

Модели основанные на NIG-распределении, гиперболическом секанс распределении и смеси гауссовских распределений обладают подходящими свойствами для описания распределения доходности валютных курсов и показали большой потенциал на Рисунке 2.2.1 и Рисунке 2.4.1. В дальнейшем следует провести вычисления параметров этих трех моделей на данных в разном временном масштабе и оценить их состоятельность с помощью квантильных графиков и статистических тестов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ:

1. Barndoff-Nielsen O. E. Exponentially decreasing distributions for the logarithm of particle size // Proc. of the Royal Society London.– 1977. – V. A353. – P. 401–419.
2. Истигечева Е. В. Оценивание параметров гиперболического и обратного гауссовского распределений.
3. Eberlein E., Keller U. Hyperbolic distributions in finance // Bernoulli. – 1995. – V. 1. – P. 281–299.
4. Королёв В. Ю. Вероятностно-статистический анализ хаотических процессов с помощью смешанных гауссовских моделей, 2008. — 65с.