

Федеральное государственное образовательное бюджетное учреждение
высшего образования
**«Финансовый университет при Правительстве Российской
Федерации»
(Финуниверситет)**

Факультет информационных технологий и анализа больших данных
Департамент анализа данных и машинного обучения

Выпускная квалификационная работа
на тему **«Оценка рисков с помощью функции копула»**
Направление подготовки 01.03.02 «Прикладная математика и
информатика»

Профиль «Анализ данных и принятие решений в экономике и финансах»

Выполнил студент учебной группы ПМ17-3

Бачаев Умар Аптиеви

Руководитель: к.э.н., доцент, доцент
ДАДиМО

Гринева Наталья Владимировна

**ВКР соответствует предъявляемым
требованиям**

Руководитель Департамента д.э.н.,
профессор

_____ В.И.Соловьев
« _____ » _____ 2021г.

СОДЕРЖАНИЕ

| | |
|--|----|
| ВВЕДЕНИЕ | 3 |
| ГЛАВА 1. КОПУЛА-ФУНКЦИЯ: ОПРЕДЕЛЕНИЕ, СВОЙСТВА И РАЗНОВИДНОСТИ | |
| 1.1. Определение и основные свойства копул | 7 |
| 1.2. Классификация копула-функций: эллиптические и архимедовы копулы | 13 |
| 1.3. Меры зависимости случайных величин | 21 |
| ГЛАВА 2. СТАТИСТИЧЕСКИЕ МЕТОДЫ ОЦЕНКИ ПАРАМЕТРОВ КОПУЛА-ФУНКЦИЙ | |
| 2.1. Параметрические методы оценивания копула-функций | 26 |
| 2.2. Полупараметрические и непараметрические методы оценивания копула-функций | 28 |
| 2.3. Процедура выбора наилучшей копула-функции | 31 |
| ГЛАВА 3. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ РИСКОВ С ПОМОЩЬЮ КОПУЛА-ФУНКЦИЙ | |
| 3.1. Меры риска портфеля. Модель оптимального портфеля ценных бумаг | 35 |
| 3.2. Модель прогнозирования курса валюты | 49 |
| ЗАКЛЮЧЕНИЕ | 54 |
| СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ | 56 |
| ПРИЛОЖЕНИЕ 1 | 59 |
| ПРИЛОЖЕНИЕ 2..... | 61 |
| ПРИЛОЖЕНИЕ 3..... | 69 |

ВВЕДЕНИЕ

Одним из основных направлений применения математических моделей в экономике является анализ финансового рынка, в частности фондовой и валютной биржи. Однако в век цифровизации построить правильную модель поведения финансовых активов становится все сложнее и сложнее. Данные подчиняются нецентральному, скошенному, тяжелохвостному распределению и прогнозы стандартных математических моделей невозможно назвать адекватными.

Модели поведения ценных бумаг и оптимизации портфеля ценных бумаг, такие как модель Марковица, Тобина, предполагают нормальное распределение исходных данных, что чаще всего не соответствует реальности. Кроме того, базовые экономико-математические модели, работавшие качественно на финансовых рынках 10 лет тому назад, предполагают анализ частных распределений каждого актива, входящего в портфель. А в качестве меры взаимосвязи берут стандартный коэффициент корреляции, объясняющий функциональную линейную связь.

Так, актуальность выпускной квалификационной работы заключается в том, что подход, рассматриваемый в исследовании, сильно отличается от вышеописанных моделей. Моделирование поведения ценных бумаг на основе копула-функций подразумевает многомерный анализ, то есть описываются взаимосвязи ценных бумаг в совокупности и в качестве показателя тесноты связи используется не базовый коэффициент корреляции, а более сложные показатели, описывающие многомерную связь.

Целью исследования является моделирование поведения данных на валютном рынке и прогнозирование курсов на будущие периоды на основе разработанного метода прогнозирования риска и доходности портфеля, основанного на многомерных распределениях, таких как копула-функции при помощи созданного программного продукта по реализации

предложенных методов и выработка рекомендаций по их применению на практике.

Для достижения цели были поставлены и решены следующие задачи:

1. Проведен анализ современной литературы по теории копула функций.
2. Сформирована теоретическая база для дальнейшего исследования.
3. Исследованы статистические методы оценки копула функций.
4. Разработаны критерии, позволяющие выбрать наилучшую из рассмотренных копула функций.
5. Определены условия применимости копула функций для фондового рынка.
6. Проведено моделирование портфелей ценных бумаг, основанного на принципе минимизации риска.
7. Проведен анализ результатов моделирования с целью выбора наилучшего критерия и временного горизонта.
8. Разработано программное обеспечение для генерации портфелей и выбора оптимального из них.
9. Выработаны рекомендации по построению оптимального портфеля и выбора временного горизонта прогнозирования.

Объектами исследования в работе являются российская фондовая биржа и мировой валютный рынок. В первом разделе практической части исследования анализируются данные пяти российских акционерных обществ, которые выбраны из принципа диверсифицированного портфеля. В последней части исследования рассматриваются данные мирового валютного рынка, в частности курсы четырех мировых валют к рублю: евро, американский и канадский доллар и британские фунты стерлингов.

Предметом исследования выступают методы математического моделирования, основанные на многомерном анализе. Первый из которых является методом минимизации портфельного риска на основе копула-

функции и многомерного распределения данных, а второй – методом прогнозного моделирования, основанным, как и в первом случае, на копула-функции, а также на эконометрической модели временного ряда $AR(1)$ - $GARCH(1,1)$.

Методологическую базу исследования с области копула-функций составили труды зарубежных авторов Skral A., Clayton D.G., Frank M. J., Gumbel E.J., Aas K., Fantazzini D., McNeil A., Nelsen R., основой для моделирования портфеля ценных бумаг и прогнозирования курса валют составили труды отечественных авторов Алексеева В.В., Соложенцева Е.Д., Маркиной И.А., Антонова И.Н., Ульяновой М.К., а также труды зарубежных исследователей Fantazzini D., Ang A., Hennessy D., McNeil A., Patton A., Yau S., Nelsen R.

Теоретическая значимость исследования заключается в разработке математического аппарата на основе копула-функций для прогнозирования доходности и риска диверсифицированного портфеля на фондовом рынке.

Практическая значимость исследования состоит в возможности использования разработанных моделей реальными компаниями с целью долгосрочного и краткосрочного инвестирования в ценные бумаги. Также прогнозная модель, построенная во втором разделе практической части, может быть применена для прогнозирования курсов валют с целью минимизации валютного риска. Кроме того, в рамках работы для обеих моделей разработаны программные продукты на языке R, которые могут быть внедрены в компанию.

Апробация исследований. С докладом на тему «Моделирование оптимального портфеля ценных бумаг с помощью многомерных копула-функций» участвовал в XII Международном научном студенческом конгрессе, на которой получил призовое место. По результатам конференции в сборнике была опубликована статья по теме, соответствующей теме доклада, указанной ранее. Результаты исследования опубликованы в журнале из списка ВАК «Проблемы экономики и

юридической практики» и рекомендованного Финансовым университетом для публикации.

Выпускная квалификационная работа состоит из введения, трех глав, заключения и приложений. Работа содержит 17 таблиц, 15 рисунков и 97 формул. Список использованной литературы включает 30 источников. Общий объем работы составляет 71 страницу.

ГЛАВА 1. КОПУЛА-ФУНКЦИЯ: ОПРЕДЕЛЕНИЕ, СВОЙСТВА И РАЗНОВИДНОСТИ

1.1. Определение и основные свойства копул

Часто при анализе данных, встречается проблема установления закономерностей между случайными величинами, зная лишь многомерное распределение, и наоборот, зная лишь маргинальные распределения, построения многомерного распределения. Одним из наиболее эффективных методов устранения такой проблемы является копула-функция. Копула-функция позволяет получить информацию не только о самих частных функциях распределения, но и о структуре их зависимостей.

Копулы широко применяются для решения многих экономических задач, они используются как в эконометрических моделях, так и в моделях машинного обучения, в частности в задачах классификации. Также копула-функции позволяют достаточно точно идентифицировать риски (страховые, кредитные, портфельные).

Определение 1. [9]

Пусть некоторая функция $C(u_1, u_2, \dots, u_n)$, зависящая от n переменных, определена на множестве $I^n = [0; 1]^n$, то есть на единичном кубе (то есть $u_i \in [0; 1], i = \overline{0, n}$). Такая функция является **копулой**, если для нее выполняются свойства:

а) область значений функции расположена в интервале $[0; 1]$;

б) если хотя бы одно $u_i = 0, i = \overline{0, n}$, то $C(u_1, u_2, \dots, u_n) = 0$;

в) $C(1, 1, \dots, u_i, \dots, 1) = 1 \forall u_i$

г) функция $C(u_1, u_2, \dots, u_n)$ является n -возрастающей, то есть для всех $(\alpha_1, \dots, \alpha_n), (\beta_1, \dots, \beta_n) \in [0; 1]^n$, где $\alpha_i \leq \beta_i$ выполняется неравенство:

$$\sum_{i_1=1}^2 \dots \sum_{i_n=1}^2 (-1)^{i_1+\dots+i_n} \cdot C(u_{1i_1}, \dots, u_{ni_n}) \geq 0, \quad (1)$$

где $u_{j1} = \alpha_j, u_{j2} = \beta_j \forall j \in \{1, \dots, n\}$.

Основной теоремой в теории копула-функций является теорема Склара, сформулированная в 1959 году. Теорема показывает роль копул в описании многомерных распределений.

Теорема 1. (теорема Склара) [24, 28-29]

Пусть $H(x_1, \dots, x_n)$ – n -мерная функция распределения с некоторыми маргинальными распределениями $F_1(x_1), \dots, F_n(x_n)$. Тогда существует копула функция $C(\cdot)$ размерности n , такая, что $\forall x_i \in \mathbb{R}$:

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)). \quad (2)$$

В случае, когда все маргинальные распределения являются непрерывными, копула определяется однозначно единственным образом, в противном случае она может быть определена однозначно лишь на области определения $\text{Ran}(F_1(x_1)) \times \dots \times \text{Ran}(F_n(x_n))$, где $\text{Ran}(F_i(x_i))$ – область определения маргинальных распределений. Тогда можно однозначно определить и обратное, если $C(\cdot)$ – некоторая копула-функция, то функция $H(x_1, \dots, x_n)$ (определяется по правилу (2)) – многомерная функция распределения с маргинальными функциями $F_1(x_1), \dots, F_n(x_n)$.

В моделировании многомерных распределений с помощью копул наибольшее распространение получило следствие из теоремы Склара, опубликованное Нельсеном позже, в 1999 году.

Следствие. [24]

Пусть $F_1^{-1}(y_1), \dots, F_n^{-1}(y_n)$ – обратные маргинальные функции распределения, тогда для каждого (y_1, \dots, y_n) из n -мерного единичного куба существует копула-функция и она единственная, причем

$$C(y_1, \dots, y_n) = H(F_1^{-1}(y_1), \dots, F_n^{-1}(y_n)). \quad (3)$$

Хоть изначально копула была определена выражением (2), основным теоретическим инструментом для построения копула-функции является соотношение (3).

Стоит также отметить, что для дискретных распределений, в отличие от непрерывных, копула-функция имеет очень сложную форму, поэтому будем рассматривать копулы непрерывных законов статистики.

Для лучшей идентификации копула-функций необходимо выделить другое важное свойство, присущее всем копула-функциям – инвариантность относительно строго возрастающих преобразований случайных величин.

Теорема 2. (свойство инвариантности) [9]

Рассмотрим n случайных величин ξ_1, \dots, ξ_n . Предположим, что зависимость между этими случайными величинами определяется некоторой копула-функцией $C(\cdot)$. Если преобразования $T_j: \mathbb{R} \rightarrow \mathbb{R}$, $j = 1, \dots, n$ определяются строго возрастающими функциями, то структура зависимости случайных величин $T_1(\xi_1), \dots, T_n(\xi_n)$ определяется той же копулой $C(\cdot)$.

Ограниченность – это еще одно свойства копула-функцией. Любая копула ограничена сверху и снизу границами Фреше-Хёффдинга:

$$W(y_1, \dots, y_n) \leq C(y_1, \dots, y_n) \leq M(y_1, \dots, y_n), \quad (4)$$

где

$$W(y_1, \dots, y_n) = \max(y_1 + \dots + y_n - n + 1, 0); \quad (4a)$$

$$M(y_1, \dots, y_n) = \min(y_1, \dots, y_n). \quad (4б)$$

Рассмотрим границы Фреше-Хёффдинга на примере стандартного равномерного распределения. Пусть $Y_1 \sim Y_2 \sim U(0,1)$, если $Y_1 = Y_2$, то эти две случайные величины имеют копулу следующего вида:

$$\begin{aligned} C(y_1, y_2) &= P(Y_1 \leq y_1, Y_2 \leq y_2) = P(Y_1 \leq y_1, Y_1 \leq y_2) = \\ &= \min(y_1, y_2) = M(y_1, y_2). \end{aligned} \quad (5a)$$

Копула функция двух тождественно равных стандартных равномерных распределений соответствует верхней границе Фреше-Хёффдинга. Отобразим эту границу на графике (см. рис. 1), используя прикладной пакет «fCopulae» программного языка R.

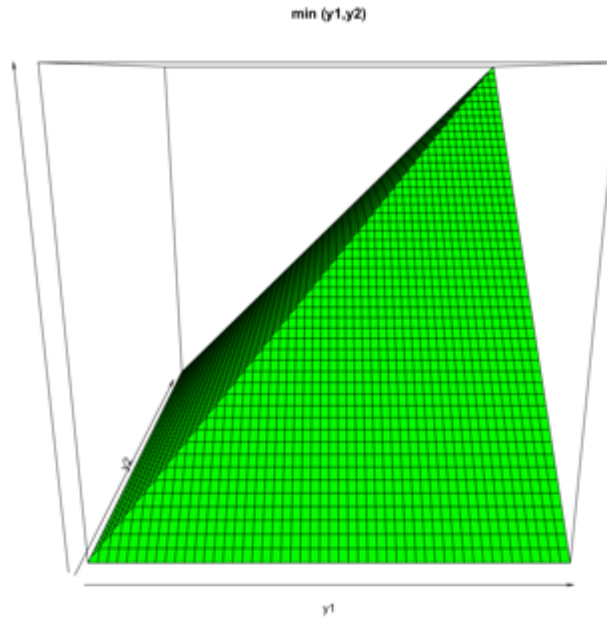


Рисунок 1. Верхняя граница Фреше-Хёфдинга стандартного равномерного распределения

Далее, предположим, что $Y_2 = 1 - Y_1$, тогда копула-функция будет иметь вид:

$$\begin{aligned}
 C(y_1, y_2) &= P(Y_1 < y_1, Y_2 < y_2) = P(Y_1 < y_1, 1 - Y_1 < y_2) = \\
 &= P(Y_1 < y_1, Y_1 \geq 1 - y_2) = P(1 - y_2 \leq Y_1 \leq y_1) = \\
 &= \begin{cases} y_1 + y_2 - 1, & y_1 + y_2 > 1 \\ 0, & y_1 + y_2 \leq 1 \end{cases} = \max(y_1 + y_2 - 1, 0) = W(y_1, y_2). \quad (56)
 \end{aligned}$$

Таким образом, копула стандартных равномерных распределений со взаимосвязью $Y_2 = 1 - Y_1$ соответствует нижней границе Фреше-Хёфдинга, изобразим ее (см. рис. 2).

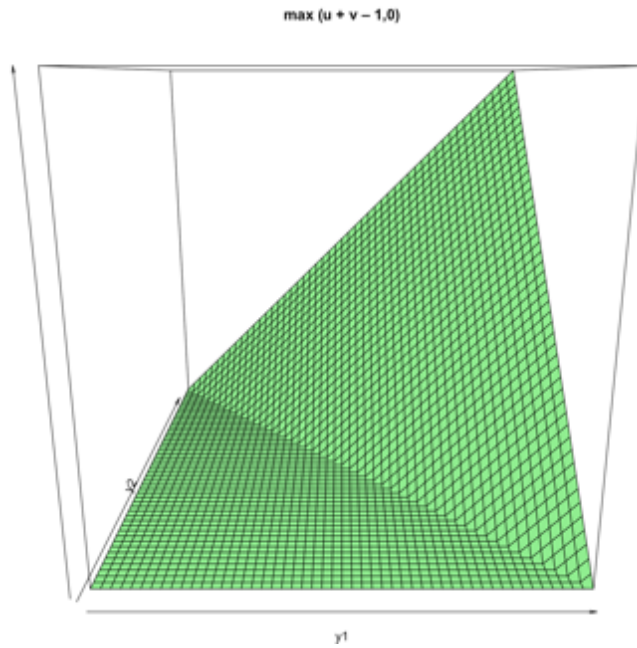


Рисунок 2. Нижняя граница Фреше-Хёфдинга стандартного равномерного распределения

И наконец рассмотрим копулу в случае независимости Y_1 от Y_2 .

$$\begin{aligned} C(y_1, y_2) &= P(Y_1 < y_1, Y_2 < y_2) = P(Y_1 < y_1) \cdot P(Y_2 < y_2) = \\ &= y_1 \cdot y_2 = P(y_1, y_2). \end{aligned} \quad (5в)$$

В третьем случае получаем самую тривиальную копула-функцию – мультипликативную. Ее график можно увидеть на рисунке 3.

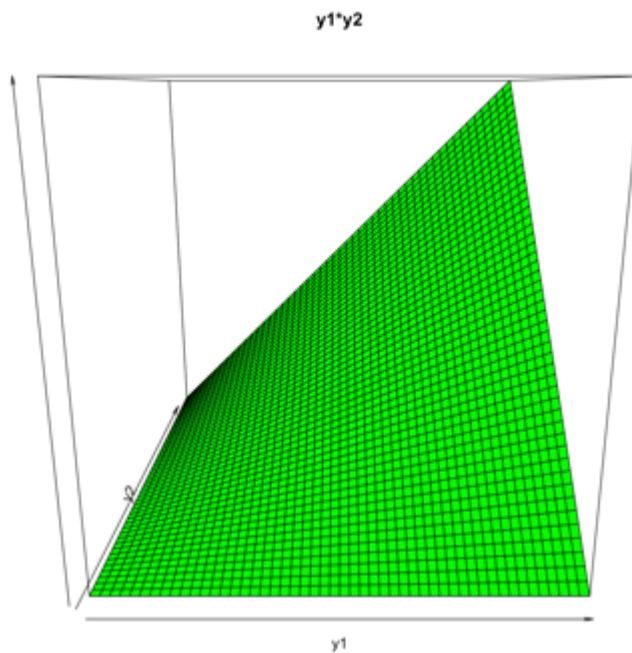


Рисунок 3. Копула независимых стандартных равномерных распределений

Для визуализации построим линии уровня для $M(y_1, y_2), W(y_1, y_2), P(y_1, y_2)$.

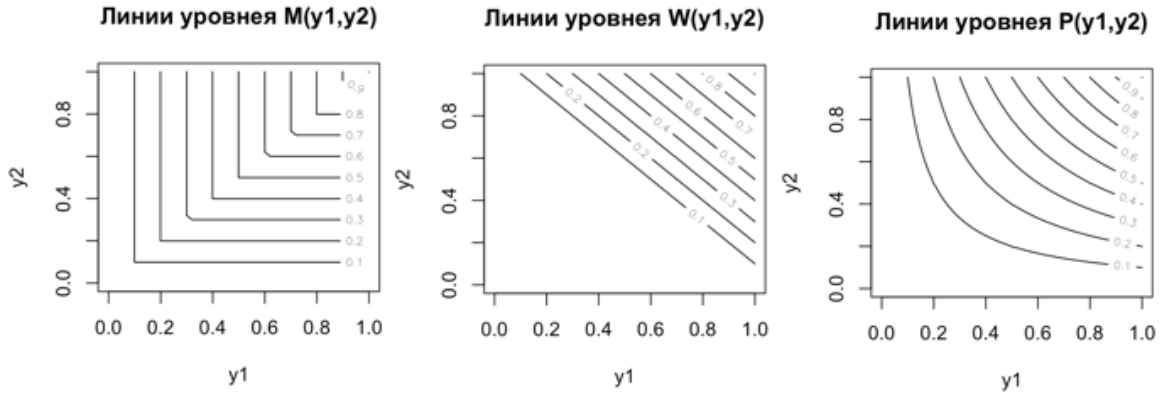


Рисунок 4. Линии уровня для границ Фреше-Хёфдинга и мультипликативной копулы

Стоит также отметить, что в общем случае, в отличие от рассмотренного примера с равномерным распределением, границы Фреше-Хёфдинга не являются копула-функциями.

Для дифференцируемой копула-функций, так же и для функций распределения, определена плотность.

$$c(y_1, \dots, y_n) = \frac{\partial^n C(y_1, \dots, y_n)}{\partial y_1 \dots y_n}, \quad (6)$$

здесь $c(y_1, \dots, y_n)$ – плотность копула-функции $C(y_1, \dots, y_n)$.

Плотность копулы используется для определения абсолютно непрерывной A_c и сингулярной S_c компоненты.

$$A_c(y_1, \dots, y_n) = \int_0^{y_1} \dots \int_0^{y_n} \frac{\partial^n C(t_1, \dots, t_n)}{\partial t_1 \dots t_n} dt_1 \dots dt_n \quad (7)$$

$$S_c(y_1, \dots, y_n) = C(y_1, \dots, y_n) - A_c(y_1, \dots, y_n) \quad (8)$$

В случае, когда $A_c(y_1, \dots, y_n) = C(y_1, \dots, y_n)$, копула-функция называется абсолютно непрерывной, а когда равенство не выполняется $A_c(y_1, \dots, y_n) \neq C(y_1, \dots, y_n)$ – сингулярной.

Плотность копула-функции $c(F_1(x_1), \dots, F_n(x_n))$ может быть представлена через плотность многомерного распределения:

$$\begin{aligned}
f_H(x_1, \dots, x_n) &= \frac{\partial^n (C(F_1(x_1), \dots, F_n(x_n)))}{\partial F_1(x_1) \dots \partial F_n(x_n)} \prod_{i=1}^n f(x_i) = \\
&= c(F_1(x_1), \dots, F_n(x_n)) \prod_{i=1}^n f(x_i),
\end{aligned} \tag{9}$$

тогда

$$c(F_1(x_1), \dots, F_n(x_n)) = \frac{f_H(x_1, \dots, x_n)}{\prod_{i=1}^n f(x_i)}. \tag{9a}$$

1.2. Классификация копула-функций: эллиптические и архимедовы копулы

Определение 2. (распределения эллиптического типа) [15]

Пусть \mathbf{Y} – некоторый случайный вектор размерности n , а $\sigma \in \mathbb{R}^{n \times n}$ – симметричная, неотрицательно определённая квадратная матрица. Если существует некоторое $\mu \in \mathbb{R}^n$ и функция $\psi: \mathbb{R} \rightarrow \mathbb{R}^+$ такие, что характеристическая функция случайного вектора $\mathbf{Y} - \mu$ имеет вид:

$$\phi_{\mathbf{Y}-\mu}(z) = e^{iz^T \mu} \psi\left(-\frac{z^T \sigma z}{2}\right) \quad \forall z \in \mathbb{R}^n, \tag{10}$$

то \mathbf{Y} называется случайным вектором, имеющим эллиптическое распределения с параметрами μ и σ .

В случае, когда функция является копулой некоторого многомерного распределения эллиптического типа, копула-функция является эллиптической. К эллиптическим статистическим распределениям относятся симметричные распределения, например распределение Гаусса (нормальной) или распределение Коши. Эллиптические распределения широко используются при актуарном моделировании, в моделях страхования и кредитования и т.д. С помощью эллиптических копул моделируются многомерные экстремальные события, так как такие эллиптические распределения, как распределение Стьюдента, позволяют описывать тяжелые хвосты распределения. Плотность многомерного эллиптического распределения в случае, когда она существует имеет следующий вид [15]

$$f(\mathbf{y}) = \frac{\gamma}{\sqrt{|\sigma|}} g((\mathbf{y} - \mu)^T \sigma^{-1} (\mathbf{y} - \mu)), \mathbf{y} \in \mathbb{R}^n \quad (11)$$

$$\gamma = \frac{1}{2\pi \int_0^{+\infty} g(u) du} \quad (12)$$

для функции $g(\cdot)$ такой, что $g: \mathbb{R} \rightarrow \mathbb{R}^+$ и $\int_0^{+\infty} g(u) du$ сходится. Здесь функция $g(\cdot)$ является оператором плотности эллиптического распределения, а γ нормирующей константой. Ниже, в таблице 1 можно увидеть список операторов плотности и нормирующих констант для некоторых эллиптических распределений.

Таблица 1. Функция $g(t)$ и γ для распределений эллиптического типа

| Распределение | Генератор | Константа |
|------------------|--|----------------------------------|
| Нормальное | $e^{-t/2}$ | $(2\pi)^{-1}$ |
| Коши | $(1+2t)^{-3/2}$ | $(2\pi)^{-1}$ |
| Стьюдента | $\left(1 + \frac{2t}{\nu}\right)^{\frac{-\nu-2}{2}}$ | $(2\pi)^{-1}$ |
| Логистическое | $e^t(1+e^{-t})^{-2}$ | π^{-1} |
| Лапласа | $e^{-\sqrt{t}}$ | $(2\pi)^{-1}$ |
| Котца | $e^{-rt/2}$ | $r(2\pi)^{-1}$ |
| Экспоненциальное | $e^{-r(t/2)^s}$ | $sr^{1/s}(2\pi\Gamma(1/s))^{-1}$ |

Используя выражение (3) из следствия теоремы Склара, определим копулу двумерного эллиптического распределения

$$C(y_1, y_2) = \frac{\gamma}{\sqrt{1-\rho^2}} \int_{-\infty}^{F^{-1}(y_1)} \int_{-\infty}^{F^{-1}(y_2)} g\left(\frac{u^2 - 2\rho uv + v^2}{1-\rho^2}\right) du dv, \quad (13)$$

в двумерного случае матрица σ определяется как корреляционная, то есть $\sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, а $|\sigma| = 1 - \rho^2$.

Рассмотрим подробнее копула-функцию нормального распределения. Для простоты визуализации и вычислений рассмотрим двумерный случай, в многомерном случае все делается аналогично. Копула двумерного

гауссовского распределения может быть получена из выражения (13) при

$$g(t) = e^{-\frac{t}{2}} \text{ и } \gamma = \frac{1}{2\pi}:$$

$$C_N(y_1, y_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{-1}(y_1)} \int_{-\infty}^{\Phi^{-1}(y_2)} e^{-\frac{u^2-2\rho uv+v^2}{2(1-\rho^2)}} du dv, \quad (14)$$

тогда запишем это в виде

$$C(y_1, y_2) = \Phi_2(\Phi^{-1}(y_1), \Phi^{-1}(y_2)), \quad (15)$$

где $\Phi_k(\cdot)$ - функция k -мерного нормального распределения с $\text{Var}(Y_1) = \text{Var}(Y_2) = 1$ и $\mathbb{E}(Y_1) = \mathbb{E}(Y_2)$ и коэффициентом корреляции ρ . В общем случае копула многомерного гауссовского распределения определяется так:

$$C(y_1, \dots, y_n) = \Phi_n(\Phi^{-1}(y_1), \dots, \Phi^{-1}(y_n)). \quad (16)$$

Плотность многомерной нормальной копула функции определяется из соотношения (9)

$$\begin{aligned} c(\Phi(y_1), \dots, \Phi(y_n)) &= \frac{f^N(y_1, \dots, y_n)}{\prod_{i=1}^n f_i^N(x_i)} = \\ &= \frac{(2\pi)^{-\frac{n}{2}} |\sigma|^{-\frac{1}{2}} e^{-\frac{1}{2} \mathbf{y}^T \sigma^{-1} \mathbf{y}}}{\prod_{i=1}^n (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2} y_i^2}} = \frac{1}{|\sigma|^{\frac{1}{2}}} e^{-\frac{1}{2} \eta^T (\sigma^{-1} - I) \eta}, \end{aligned} \quad (17)$$

где $\eta = (\Phi^{-1}(x_1), \dots, \Phi^{-1}(x_n))^T$ - вектор, элементами которого являются значения функции распределения для стандартного нормального распределения в точках $x_i = \Phi(y_i)$, а I - единичная матрица.

Гауссовская копула часто используется при анализе финансовых данных в стабильной экономике (иначе присутствуют тяжелые хвосты), так как она позволяет моделировать портфели большой размерности и достаточно легко может быть оценена. На рисунке 5 изобразим графики плотности двумерной нормальной копула-функции при различных значениях коэффициента корреляции.

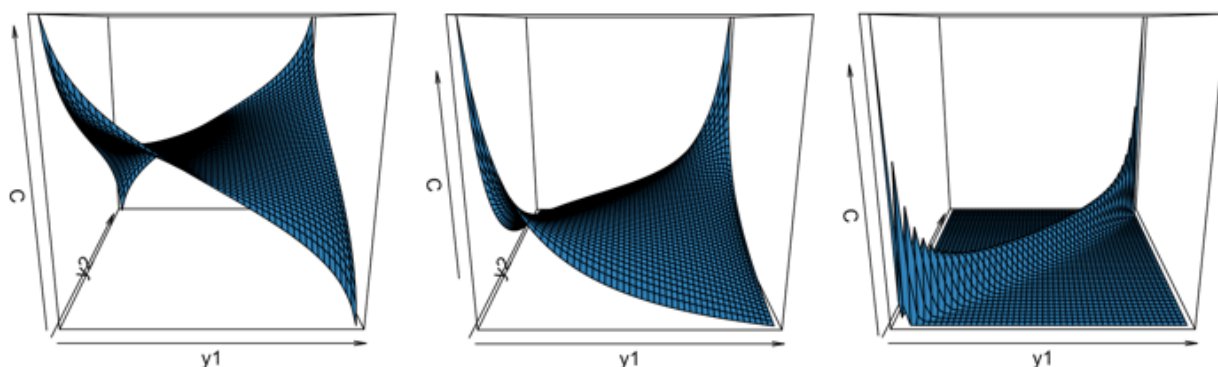


Рисунок 5. Плотность копулы нормального распределения при $\rho = 0,01; 0,5; 0,99$ соответственно

Изобразим также линии уровня в каждом из случаев.

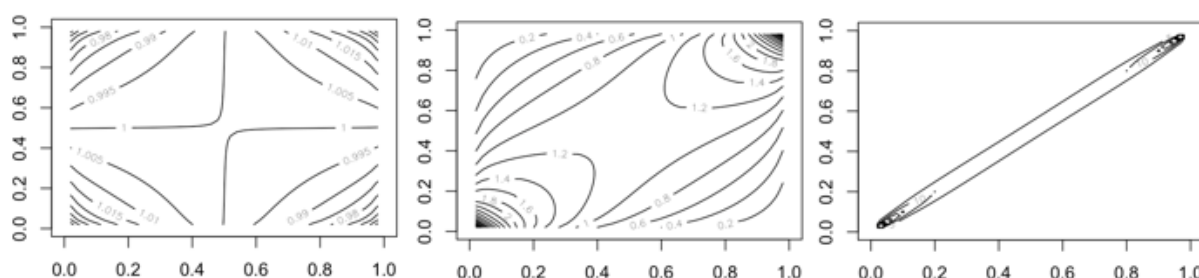


Рисунок 6. Линии уровня копулы нормального распределения при $\rho = 0,01; 0,5; 0,99$ соответственно

Другим большим классом копула-функций является класс архимедовых копул. Определим двумерные архимедовы копулы, а далее обобщим их для многомерного случая.

Определение 3. (архимедова копула) [23]

Рассмотрим некоторую непрерывную, убывающую (строго) и выпуклую функцию $\varphi(u)$, принимающую неотрицательные значения, определенную на $u \in [0; 1]$ и $\varphi(1) = 0$. Пусть $\varphi^{[-1]}(u)$ – псевдообратная функция, то есть:

$$\varphi^{[-1]}(u) = \begin{cases} \varphi^{-1}(u), & 0 \leq u \leq \varphi(0) \\ 0, & u > \varphi(0) \end{cases}, \quad (18)$$

где $\varphi^{-1}(u)$ - функция, обратная к $\varphi(u)$.

Функция $C: [0; 1]^2 \rightarrow [0; 1]$, определенная следующим образом:

$$C(y_1, y_2) = \varphi^{[-1]}(\varphi(y_1) + \varphi(y_2)), \quad (19)$$

называется **архимедовой копулой** с генератором $\varphi(u)$. Кроме того, если $\varphi(0) \rightarrow \infty$, то $\varphi^{[-1]}(u) = \varphi^{-1}(u)$. В таком случае $\varphi(u)$ называется **строгим генератором**, а $C(y_1, y_2)$ **строгой копула-функцией**.

Проиллюстрируем вышесказанное наглядно графиками строгого и нестрогого генератора, а также графиками функций обратных к ним (см. рис. 7).

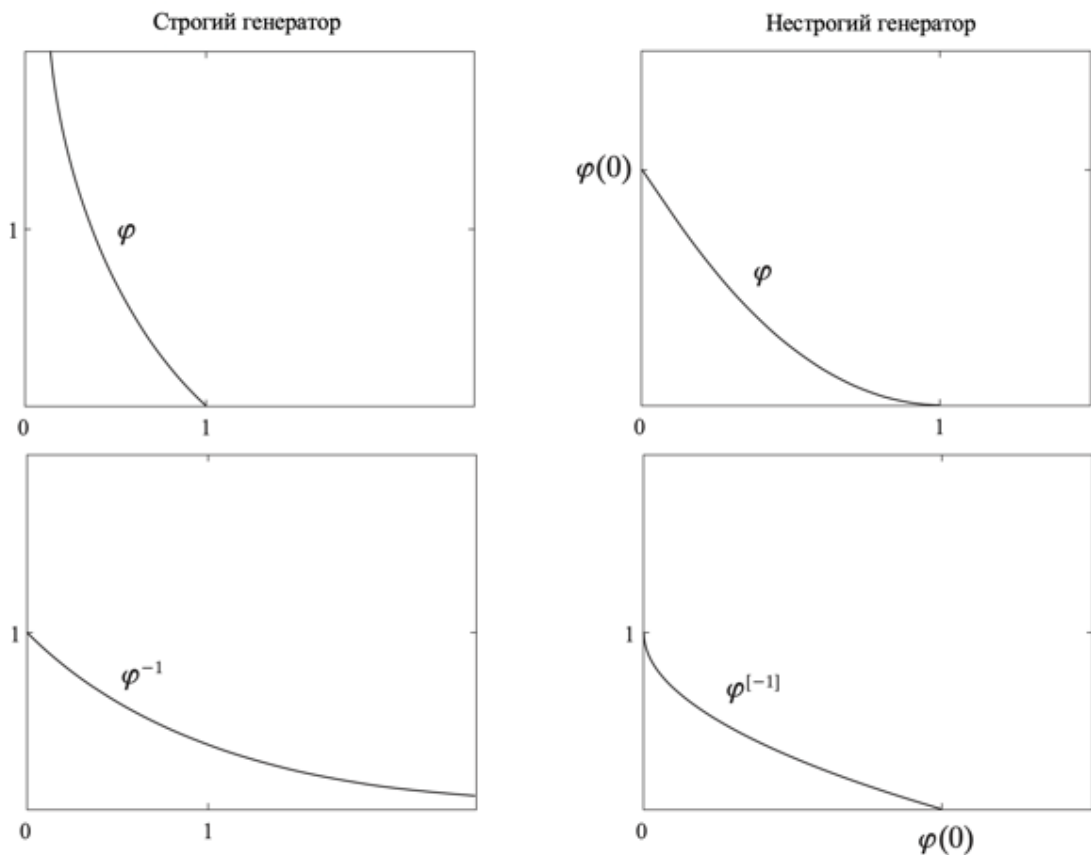


Рисунок 7. Графики строгого и нестрогого генераторов и обратных к ним функций

Плотность архимедовой копулы определяется по общему правилу (6), однако она может быть представлена в более развернутом виде, если функция-генератор дважды дифференцируема:

$$c_{\varphi}(y_1, y_2) = \frac{\left(\varphi^{[-1]}(\varphi(y_1) + \varphi(y_2))\right)''}{\left(\varphi^{[-1]}(\varphi(y_1))\right)' \cdot \left(\varphi^{[-1]}(\varphi(y_2))\right)'}. \quad (20)$$

В таблице 2 приведем список архимедовых копул с их генерирующими функциями и диапазоном значений параметра α [24].

Таблица 2. Список архимедовых копул

| Номер | $C_\alpha(u, v)$ | Генератор $\varphi_\alpha(t)$ | Диапазон параметра α |
|-------|---|--|-------------------------------------|
| 1 | $\max([u^{-\alpha} + v^{-\alpha} - 1]^{-1/\alpha}, 0)$ | $(t^{-\alpha} - 1) / \alpha$ | $[-1, \infty) \setminus \{0\}$ |
| 2 | $\max(1 - [(1-u)^\alpha + (1-v)^\alpha]^{1/\alpha}, 0)$ | $(1-t)^\alpha$ | $[1, \infty)$ |
| 3 | $\frac{uv}{1 - \alpha(1-u)(1-v)}$ | $\ln \frac{1 - \alpha(1-t)}{t}$ | $[-1, 1)$ |
| 4 | $\exp(-[(\ln u)^\alpha + (\ln v)^\alpha]^{1/\alpha})$ | $(-\ln t)^\alpha$ | $[1, \infty)$ |
| 5 | $-\frac{1}{\alpha} \ln \left(1 + \frac{(e^{-\alpha u} - 1)(e^{-\alpha v} - 1)}{e^{-\alpha} - 1} \right)$ | $-\ln \frac{e^{-\alpha t} - 1}{e^{-\alpha} - 1}$ | $(-\infty, \infty) \setminus \{0\}$ |
| 6 | $1 - [(1-u)^\alpha + (1-v)^\alpha - (1-u)^\alpha(1-v)^\alpha]^{1/\alpha}$ | $-\ln[1 - (1-t)^\alpha]$ | $[1, \infty)$ |
| 7 | $\max(\alpha uv + (1-\alpha)(u+v-1), 0)$ | $-\ln[\alpha t + (1-\alpha)]$ | $(0, 1]$ |
| 8 | $\max\left(\frac{\alpha^2 uv - (1-u)(1-v)}{\alpha^2 - (\alpha-1)^2(1-u)(1-v)}, 0\right)$ | $\frac{1-t}{1 + (\alpha-1)t}$ | $[1, \infty)$ |
| 9 | $uv \exp(-\alpha \ln u \ln v)$ | $\ln(1 - \alpha \ln t)$ | $(0, 1]$ |
| 10 | $uv[1 + (1-u^\alpha)(1-v^\alpha)]^{-1/\alpha}$ | $\ln(2t^{-\alpha} - 1)$ | $(0, 1]$ |
| 11 | $\max([u^\alpha v^\alpha - 2(1-u)^\alpha(1-v)^\alpha]^{1/\alpha}, 0)$ | $\ln(2-t^\alpha)$ | $(0, \frac{1}{2}]$ |
| 12 | $(1 + [(u^{-1}-1)^\alpha + (v^{-1}-1)^\alpha]^{1/\alpha})^{-1}$ | $(t^{-1} - 1)^\alpha$ | $[1, \infty)$ |
| 13 | $\exp(1 - [(1-\ln u)^\alpha + (1-\ln v)^\alpha - 1]^{1/\alpha})$ | $(1 - \ln t)^\alpha - 1$ | $(0, \infty)$ |
| 14 | $(1 + [(u^{-1/\alpha} - 1)^\alpha + (v^{-1/\alpha} - 1)^\alpha]^{1/\alpha})^{-\alpha}$ | $(t^{-1/\alpha} - 1)^\alpha$ | $[1, \infty)$ |
| 15 | $\max\{[1 - [(1-u^{1/\alpha})^\alpha + (1-v^{1/\alpha})^\alpha]^{1/\alpha}], 0\}$ | $(1 - t^{1/\alpha})^\alpha$ | $[1, \infty)$ |
| 16 | $\frac{1}{2}(S + \sqrt{S^2 + 4\alpha}), S = u + v - 1 - \alpha(u^{-1} + v^{-1} - 1)$ | $(\alpha t^{-1} + 1)(1-t)$ | $[0, \infty)$ |
| 17 | $\left(1 + \frac{[(1+u)^{-\alpha} - 1][(1+v)^{-\alpha} - 1]}{2^{-\alpha} - 1}\right)^{-1/\alpha} - 1$ | $-\ln \frac{(1+t)^{-\alpha} - 1}{2^{-\alpha} - 1}$ | $(-\infty, \infty) \setminus \{0\}$ |
| 18 | $\max(1 + \alpha / \ln[e^{\alpha/(u-1)} + e^{\alpha/(v-1)}], 0)$ | $e^{\alpha/(u-1)}$ | $[2, \infty)$ |
| 19 | $\alpha / \ln(e^{\alpha/u} + e^{\alpha/v} - e^\alpha)$ | $e^{\alpha/t} - e^\alpha$ | $[0, \infty)$ |
| 20 | $[\ln(\exp(u^{-\alpha}) + \exp(v^{-\alpha}) - e)]^{-1/\alpha}$ | $\exp(t^{-\alpha}) - e$ | $(0, \infty)$ |
| 21 | $1 - (1 - \{\max([1 - (1-u)^\alpha]^{1/\alpha} + [1 - (1-v)^\alpha]^{1/\alpha} - 1, 0)\}^{1/\alpha})$ | $1 - [1 - (1-t)^\alpha]^{1/\alpha}$ | $[1, \infty)$ |
| 22 | $\max\left(\left[1 - (1-u^\alpha)\sqrt{1 - (1-u^\alpha)^2} - (1-v^\alpha)\sqrt{1 - (1-v^\alpha)^2}\right]^{1/\alpha}, 0\right)$ | $\arcsin(1 - t^\alpha)$ | $(0, 1]$ |

Копулы, в которых $\varphi_\alpha(0) \rightarrow \infty$ являются строгими. Такими являются функции № 3, 4, 5, 6, 9, 10, 12, 13, 14, 17, 19, 20, 22. Стоит выделить наиболее широко применяемые архимедовы копулы: 1, 4, 5 - копулы Клейтона, Гумбея и Франка соответственно. Рассмотрим их подробнее.

Определение 4. (копула-функция Клейтона) [14]

Пусть генерирующая функция $\varphi(u)$ имеет вид:

$$\varphi(u) = \frac{u^{-\alpha} - 1}{\alpha}, \alpha \in [-1; +\infty) \setminus \{0\}, \quad (21)$$

а обратная функция

$$\varphi^{-1}(u) = (1 + u)^{-\frac{1}{\alpha}}.$$

Тогда используя соотношение (19), получаем:

$$C_{clayton}(y_1, y_2) = \max\left((y_1^{-\alpha} + y_2^{-\alpha} - 1)^{-\frac{1}{\alpha}}, 0\right). \quad (22)$$

Функция вида (22) называется **копула-функцией Клейтона**.

В случае, когда $\alpha > 0$, генерирующая функция в нуле будет бесконечно большой, то есть $\varphi(0) \rightarrow +\infty$, и выражение (22) можно представить в виде:

$$C(y_1, y_2) = (y_1^{-\alpha} + y_2^{-\alpha} - 1)^{-\frac{1}{\alpha}}, \quad (23)$$

тогда плотность копулы Клейтона для $\alpha > 0$ будет следующая:

$$c(y_1, y_2) = (1 + \alpha)(y_1 y_2)^{-1-\alpha} (y_1^{-\alpha} + y_2^{-\alpha} - 1)^{-2-\frac{1}{\alpha}}. \quad (24)$$

Стоит отметить, что при $\alpha = -1$ копула функция Клейтона соответствует нижней границе Фреше-Хёфдинга, при $\alpha \rightarrow 0$ случаю независимости и при $\alpha \rightarrow +\infty$ случаю комонотонности (абсолютная зависимость).

Определение 5. (копула-функция Гумбеля) [18]

Копула функция относится к классу копул Гумбеля, если функция-генератор имеет вид:

$$\varphi(u) = (-\ln u)^\alpha, \alpha \geq 1, \quad (25)$$

обратная функция

$$\varphi^{-1}(u) = e^{-u^{\frac{1}{\alpha}}}, \quad (26)$$

а сама копула-функция

$$C_{Gumbel}(y_1, y_2) = e^{-((- \ln y_1)^\alpha + (- \ln y_2)^\alpha)^{\frac{1}{\alpha}}}. \quad (27)$$

Функцию плотности копула-функции Гумбеля можно определить, как:

$$c(y_1, y_2) = e^{-((- \ln y_1)^\alpha + (- \ln y_2)^\alpha)^{\frac{1}{\alpha}}} \cdot (y_1 y_2)^{-1} ((- \ln y_1)^\alpha + (- \ln y_2)^\alpha)^{\frac{1}{\alpha}-2}.$$

$$\cdot (\ln y_1 \cdot \ln y_2)^{\alpha-1} \cdot \left(((-\ln y_1)^\alpha + (-\ln y_2)^\alpha)^{-\frac{1}{\alpha}} + \alpha - 1 \right) \quad (28)$$

При $\alpha = 1$ получаем копулу функцию двух независимых случайных величин Y_1 и Y_2 , а при $\alpha \rightarrow +\infty$ для комонотонных случайных величин (абсолютно зависимых). В таком случае мы делаем вывод, что копула Гумбеля интерполирует промежуточную структуру зависимости случайных величин Y_1 и Y_2 между независимостью и абсолютной зависимостью.

Определение 6. (копула-функция Франка) [17]

Пусть функция-генератор имеет вид:

$$\varphi(u) = \ln \left(\frac{e^{-\alpha} - 1}{e^{-\alpha u} - 1} \right), \quad (29)$$

а функция, обратная генератору

$$\varphi^{-1}(u) = -\frac{1}{\alpha} \ln(1 + e^u(e^{-\alpha} - 1)). \quad (30)$$

Архимедова копула с таким генератором называется **копулой Франка**:

$$C(y_1, y_2) = -\frac{1}{\alpha} \ln \left(1 + \frac{(e^{-\alpha y_1} - 1)(e^{-\alpha y_2} - 1)}{e^{-\alpha} - 1} \right). \quad (31)$$

Дважды продифференцировав по y_1 и y_2 , получим плотность копула-функции Франка.

$$c(y_1, y_2) = \frac{-\alpha(e^{-\alpha} - 1)e^{-\alpha(y_1+y_2)}}{\left((e^{-\alpha y_1} - 1)(e^{-\alpha y_2} - 1) + e^{-\alpha} - 1 \right)^2}, \alpha \neq 0 \quad (32)$$

При значении $\alpha = 0$ копула Франка совпадает с копулой независимых случайных величин, а при $\alpha \rightarrow \pm\infty$ с верхней и нижней границей Фреше-Хёффдинга соответственно.

Чтобы показать разницу между этими тремя видами копул, используя прикладной пакет «copula» в среде R, построим графики каждого.

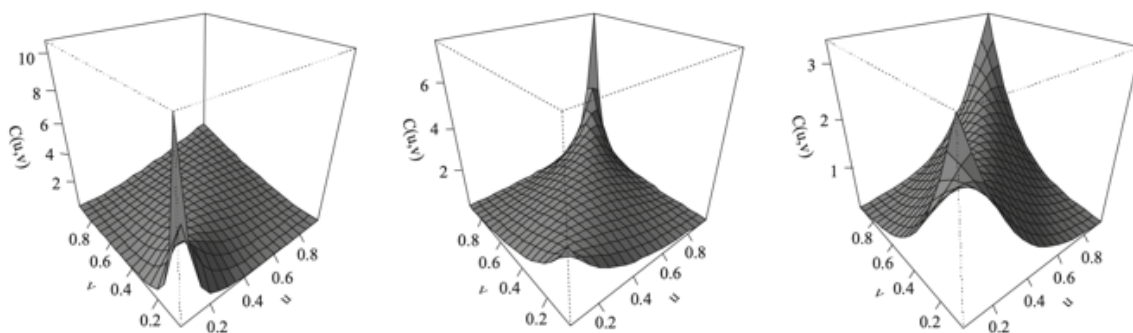


Рисунок 8. Плотности копула-функций Клейтона, Гумбеля и Франка
соответственно

Анализ графиков показал, что копула-функции Клейтона и Гумбеля описывают хвостовую зависимость, так как в точках $(0;0)$ и $(1;1)$ плотность резко растет. Копула Гумбеля описывает верхнюю хвостовую зависимость, Клейтона - нижнюю, а копула-функция Франка вовсе не имеет хвостовой зависимости.

Архимедовы копулы можно обобщить и для n -мерного случая. Тогда в общем виде копула будет иметь вид:

$$C(y_1, \dots, y_n) = \varphi^{[-1]}(\varphi(y_1) + \dots + \varphi(y_n)). \quad (33)$$

Функция вида (33) называется серийно-итеративной. Для нее выполняется еще одно важное свойство:

$$C(y_1, \dots, y_n) = C(C(y_1, \dots, y_{n-1}), y_n). \quad (34)$$

1.3. Меры зависимости случайных величин

Меры зависимости - инструмент, описывающий структуру зависимости между случайными величинами. Одной из самых распространенных мер является коэффициент корреляции. И, действительно, иногда он очень хорошо описывает структурные взаимосвязи случайных величин, однако главным недостатком коэффициента корреляции является то, что он описывает только функциональную линейную связь. Часто бывает так, что случайные величины бывают связаны нелинейно, и тогда коэффициент корреляции не совсем хороший показатель точности. Для выделения нелинейных зависимостей используются другие метрики, основанные на копула-

функциях. Такими мерами являются коэффициент ранговой корреляции ρ -Спирмена и τ -Кендалла.

Коэффициенты ρ -Спирмена и τ -Кендалла – метрики, основанные на согласованности и рассогласованности случайных величин.

Определение 7. (Коэффициент ρ -Спирмена) [1]

Для двух непрерывных случайных величин Y_1 и Y_2 коэффициент ρ -Спирмена определяется как:

$$\begin{aligned}\rho_S(Y_1, Y_2) &= 3Q(C, P) = 12 \int \int y_1 y_2 dC(y_1, y_2) - 3 = \\ &= 12 \int \int C(y_1, y_2) dy_1 dy_2 - 3,\end{aligned}\quad (35)$$

где $y_1 = F_{Y_1}(x_1)$, $y_2 = F_{Y_2}(x_2)$, $Q(\cdot)$ – разность вероятности согласованности и рассогласованности случайных величин.

Иными словами, если выразить коэффициент корреляции ρ -Спирмена через обычный коэффициент корреляции, получим:

$$\rho_S(Y_1, Y_2) = \rho(F(Y_1), F(Y_2)), \quad (36)$$

где $F(Y_1)$ и $F(Y_2)$ – маргинальные функции распределения.

В многомерном случае для коэффициента корреляции ρ -Спирмена также можно определить корреляционную матрицу $\rho_S(Y_1, \dots, Y_n)$. Такая матрица будет симметричной, по ее главной диагонали будут стоять единицы, а элементы определяться по следующему правилу:

$$\rho_S(Y_i, Y_j) = \rho(F(Y_i), F(Y_j)). \quad (37)$$

Рассмотрим случай, когда $Y_1 \sim Y_2 \sim N(0,1)$. В таком случае коэффициент корреляции ρ -Спирмена можно выразить через обычный коэффициент корреляции:

$$\rho_S = \frac{6}{\pi} \arcsin\left(\frac{\rho}{2}\right). \quad (38)$$

Несмещенная оценка такой характеристики имеет вид:

$$\hat{\rho}_S = \frac{6}{k(k-1)(k-2)} \sum_{1 \leq i \leq j \leq l \leq k} 3 \operatorname{sign}(Y_{1i} - Y_{1j}) \operatorname{sign}(Y_{2i} - Y_{2l}), \quad (39)$$

где (Y_1, Y_2) – выборка объема k , элементами которой являются Y_{1i}, Y_{2i} .

Определение 8. (Коэффициент τ -Кендалла) [2]

Пусть (Y_1, Z_1) – случайный вектор с совместной функцией распределения $H(y, z)$, а (Y_2, Z_2) – случайный вектор, такой, что Y_2 и Z_2 независимы, а распределения Y_2 и Y_1 , Z_2 и Z_1 совпадают. **Коэффициент τ -Кендалла** для вектора (Y, Z) определяется как разница между вероятностью согласованности и рассогласованности:

$$\begin{aligned}\tau(Y, Z) &= P((Y_1 - Y_2)(Z_1 - Z_2) > 0) - P((Y_1 - Y_2)(Z_1 - Z_2) < 0) \\ &= P(Y_1 > Y_2, Z_1 > Z_2) + P(Y_2 > Y_1, Z_2 > Z_1) - \\ &\quad - P(Y_1 < Y_2, Z_1 > Z_2) - P(Y_1 > Y_2, Z_1 < Z_2)\end{aligned}\quad (40)$$

В n -мерном случае коэффициент τ -Кендалла будет иметь вид симметричной корреляционной матрицы, элементами которой будут являться значения:

$$\rho_\tau(\mathbf{Y}, \mathbf{Z})_{ij} = \mathbb{Cov}(\text{sign}(Y_i - Z_i), \text{sign}(Y_j - Z_j)). \quad (41)$$

Коэффициент τ -Кендалла может быть выражен через копула-функцию. Сформулируем теорему об определении коэффициента в терминах копула-функции.

Теорема 3. [10]

В случае двумерной случайной величины коэффициент τ -Кендалла можно определить следующим выражением:

$$\tau(Y_1, Y_2) = 4 \int_0^1 \int_0^1 C(y_1, y_2) dC(y_1, y_2) - 1. \quad (42)$$

Тогда для эллиптических распределений можно написать явный вид коэффициента τ -Кендалла через обычный парный коэффициент корреляции:

$$\tau(Y_1, Y_2) = \frac{2}{\pi} \arcsin \rho(Y_1, Y_2). \quad (43)$$

Для архимедовых копула-функций коэффициент оценивается через генерирующую функцию

$$\tau(Y_1, Y_2) = 1 + 4 \int_0^1 \frac{\varphi(u)}{\varphi'(u)} du = 1 + 4 \int_0^1 \frac{1}{(\ln \varphi(u))'} du \quad (44)$$

Определим коэффициент τ -Кендалла для копула-функции Клейтона и Гумбеля:

$$\tau_{clayton}(Y_1, Y_2) = \frac{\alpha}{\alpha + 2}; \quad (45)$$

$$\tau_{Gumbel}(Y_1, Y_2) = 1 - \frac{1}{\alpha}. \quad (46)$$

Несмещенная оценка коэффициента τ -Кендалла имеет вид:

$$\hat{\tau} = \frac{2}{k(k-1)} \sum_{1 \leq i < j \leq k} \text{sign}(Y_i - Y_j) \text{sign}(Z_i - Z_j). \quad (47)$$

Другими важными мерами зависимости являются коэффициенты верхней и нижней хвостовой зависимости. Такая мера широко используется при моделировании экстремальных событий, которые всегда имеют тяжелые хвосты.

Определение 9. (верхняя хвостовая зависимость) [10]

Пусть (Y_1, Y_2) – случайный вектор, у которого компоненты являются непрерывными случайными величинами с маргинальными функциями распределения F_{Y_1} и F_{Y_2} . Тогда коэффициент верхней хвостовой зависимости λ^U случайных Y_1 и Y_2 определяется выражением:

$$\lambda^U = \lim_{y \rightarrow 1} P\left(Y_2 > F_{Y_2}^{-1}(y) \mid Y_1 > F_{Y_1}^{-1}(y)\right) = \lim_{y \rightarrow 1} \frac{1 - 2y + C(y, y)}{1 - y}, \quad (48)$$

при условии, что пределы существуют.

Случайные величины Y_1 и Y_2 называют асимптотически зависимыми на верхнем хвосте, когда $\lambda^U \in (0, 1]$ и асимптотически независимыми, когда $\lambda^U = 0$. Также коэффициент λ^U можно интерпретировать как вероятность того, что одна переменная примет экстремальное значение, когда вторая также принимает экстремальное значение. Нижняя хвостовая зависимость определяется аналогично.

Определение 10. (нижняя хвостовая зависимость) [10]

Пусть (Y_1, Y_2) – случайный вектор, у которого компоненты являются непрерывными случайными величинами с маргинальными функциями

распределения F_{Y_1} и F_{Y_2} . Тогда коэффициент нижней хвостовой зависимости λ^L случайных Y_1 и Y_2 определяется выражением:

$$\lambda^L = \lim_{y \rightarrow 0} P(Y_2 \leq F_{Y_2}^{-1}(y) | Y_1 \leq F_{Y_1}^{-1}(y)) = \lim_{y \rightarrow 0} \frac{C(y, y)}{y}. \quad (49)$$

Случайные величины Y_1 и Y_2 называют асимптотически зависимыми на нижнем хвосте, когда $\lambda^L \in (0, 1]$ и асимптотически независимыми, когда $\lambda^L = 0$.

Для эллиптических распределений λ^L и λ^U совпадают, так как такие распределения являются симметричными. Нормальное распределение – единственное устойчивое распределение с легкими хвостами. Логично предположить, что λ^L и λ^U для него будут нулевыми, проверим это.

$$\lambda^L = \lambda^U = 2 \lim_{y \rightarrow \infty} (P(Y_1 > y | Y_2 = y)) = 2 \lim_{y \rightarrow \infty} \left(1 - \Phi \left(y \frac{\sqrt{1 - \rho}}{\sqrt{1 + \rho}} \right) \right) = 0 \quad (50)$$

Как и ожидалось, коэффициент хвостовой зависимости для нормального распределения равен 0.

В таком случае рассмотрим распределение с тяжелыми хвостами, например распределение Стьюдента. Определим для него коэффициент хвостовой зависимости.

$$\lambda^L = \lambda^U = 2 \lim_{y \rightarrow \infty} (P(Y_1 > y | Y_2 = y)) = 2 - 2t_{v+1} \left(\sqrt{v+1} \cdot \frac{\sqrt{1 - \rho}}{\sqrt{1 + \rho}} \right) \quad (51)$$

Таким образом, коэффициент хвостовой зависимости для распределения Стьюдента ненулевой при небольших степенях свободы, когда $v \rightarrow +\infty$ коэффициент хвостовой зависимости $\lambda^L = \lambda^U \rightarrow 0$. Это связано с тем, что при высоких степенях свободы распределение Стьюдента по вероятности сходится к стандартному нормальному распределению.

ГЛАВА 2. СТАТИСТИЧЕСКИЕ МЕТОДЫ ОЦЕНКИ ПАРАМЕТРОВ КОПУЛА-ФУНКЦИЙ

2.1. Параметрические методы оценивания копула-функций

Пусть $z_{1t}, \dots, z_{nt}, t = 1, \dots, k$ – некоторые многомерные наблюдения, n – размерность наблюдаемых значений, а k – число наблюдений. Пусть $f_H(\cdot)$ – плотность многомерного распределения (Z_1, \dots, Z_n) .

$$f_H(\tilde{z}_1, \dots, \tilde{z}_n, \alpha_1, \dots, \alpha_n, \gamma) = c(F_1(\tilde{z}_1, \alpha_1), \dots, F_n(\tilde{z}_n, \alpha_n), \gamma) \cdot \prod_{j=1}^n f_j(\tilde{z}_j, \alpha_j), \quad (52)$$

f_j – плотность маргинального распределения Z_j , $\tilde{z}_1, \dots, \tilde{z}_n$ – текущие значения случайных величин Z_1, \dots, Z_n , а $c(\cdot)$ – плотность копула-функции.

Пусть $\theta = (\alpha_1, \dots, \alpha_n, \gamma)$ – вектор параметров, подлежащих оцениванию, где α_j – параметры маргинальных распределений случайного вектора, γ – вектор параметров совместной копула-функции случайного вектора. Тогда логарифм функции правдоподобия будет иметь вид:

$$l(\theta) = \ln L(\theta) = \sum_{t=1}^k \ln(c(F_1(z_{1t}, \alpha_1), \dots, F_n(z_{nt}, \alpha_n), \gamma)) + \\ + \sum_{t=1}^k \sum_{j=1}^n \ln f_j(z_{jt}, \alpha_j). \quad (53)$$

Оценка максимального правдоподобия для вектора θ будет считаться по правилу:

$$\hat{\theta}_{\text{ОМП}} = \arg \max_{\theta} l(\theta). \quad (54)$$

Сформулируем теорему о статистических свойствах оценки максимального правдоподобия.

Теорема 4. [10]

Пусть $\hat{\theta}_{\text{ОМП}}$ – оценка максимального правдоподобия вектора параметров θ . Тогда при выполнении условия регулярности оценка обладает свойствами:

- $\hat{\theta}_{\text{ОМП}} \xrightarrow{P} \theta$ – состоятельность оценок

- $\sqrt{k}(\hat{\theta}_{\text{ОМП}} - \theta) \xrightarrow{d} N(0, I^{-1}(\theta))$ – асимптотическая нормальность ($I(\theta)$ – информационная матрица по Фишеру).

Другим параметрическим методом оценки параметров является модифицированный ММП, так называемый **двухшаговый метод максимального правдоподобия (ДММП)**. В этом методе по отдельности оценивается α_j и γ .

Первый шаг метода заключается в оценивании α_j по стандартному методу максимального правдоподобия одномерных случайных величин:

$$\hat{\alpha}_j = \arg \max_{\alpha} l_j(\alpha) = \arg \max_{\alpha} \sum_{t=1}^k \ln f_j(z_{jt}, \alpha), \quad (55)$$

где $l_j(\cdot)$ – логарифм функции правдоподобия маргинального распределения.

Второй шаг заключается в оценивании параметров копула-функции γ , используя оценки $\hat{\alpha}_j$, полученные на предыдущем шаге

$$\hat{\gamma}_{\text{О2МП}} = \arg \max_{\gamma} l_c(\gamma) = \arg \max_{\gamma} \sum_{t=1}^k \ln(c(F_1(z_{1t}, \hat{\alpha}_1), \dots, F_n(z_{nt}, \hat{\alpha}_n), \gamma)), \quad (56)$$

где $l_c(\cdot)$ – логарифм функции правдоподобия для копулы.

Сформулируем свойства оценок ДММП для двумерного случая для простоты понимания. В многомерном случае все свойства будут сохраняться.

Теорема 5. [10]

В случае, когда объем выборки большой, то есть $n \rightarrow \infty$, оценки параметров $\hat{\alpha}_1$ и $\hat{\alpha}_2$ являются состоятельными, то есть $\hat{\alpha}_1 \xrightarrow{P} \alpha_1$ и $\hat{\alpha}_2 \xrightarrow{P} \alpha_2$.

Теорема 6.

Пусть функция $\frac{1}{n} \sum_{t=1}^k \ln(c(F_1(z_{1t}, \hat{\alpha}_1), F_2(z_{2t}, \hat{\alpha}_2), \gamma))$ имеет единственный максимум в некоторой точке $\gamma' \in D$, где D – внутренность

множества допустимых значений γ . Тогда $\hat{\gamma}_{02МП} \xrightarrow{P} \gamma_0$ при $n \rightarrow \infty$, это значит, что оценка $\hat{\gamma}_{02МП}$ является состоятельной.

Теорема 7.

Оценки ДММП являются асимптотически нормальными, то есть $\sqrt{k}(\hat{\theta}_{02МП} - \theta) \xrightarrow{d} N(0, G)$, где G – это информационная матрица Годамби.

2.2. Полупараметрические и непараметрические методы оценивания копула-функций

Одним из наиболее распространенных полупараметрических методов оценивания копула-функция является **канонический метод максимального правдоподобия (КММП)**. В отличие от вышеописанных методов, здесь не делаются никакие предположения о маргинальных распределениях. КММП состоит из двух шагов.

1. Наблюдения z_{1t}, \dots, z_{nt} преобразуем:

$$\hat{y}_{it} = F_{ik}(z_{it}), \quad (57)$$

где $F_{ik}(z) = \frac{1}{k+1} \sum_{t=1}^k I_{\{x_{it} \leq z\}}$, $I_{\{x_{it} \leq z\}}$ – индикатор. Иначе говоря, $F_{ik}(z)$ – непараметрическая оценка маргинальной функции распределения.

2. Далее, оцениваем параметры копулы, максимизирую логарифм функции правдоподобия:

$$\hat{\gamma}_{КММП} = \arg \max_{\gamma} \sum_{t=1}^k \ln(c(F_{1k}(z_{1t}), \dots, F_{nk}(z_{nt}), \gamma)). \quad (58)$$

Опишем некоторые асимптотические свойства метода для двумерного случая. Введем обозначения: $l(y_1, y_2, \gamma) = \ln c(y_1, y_2, \gamma)$. Пусть

$$S_k = \frac{1}{k} \sum_{t=1}^k l'_y(F_{1k}(z_{1t}), F_{2k}(z_{2t}), \gamma), \quad (59)$$

$$H_k = \frac{1}{k} \sum_{t=1}^k l''_{yy}(F_{1k}(z_{1t}), F_{2k}(z_{2t}), \gamma). \quad (60)$$

Так как оценка $\hat{\gamma}_{\text{КММП}}$ является решением уравнения $S_k = 0$, разложив в ряд Тейлора в окрестности точки $\hat{\gamma}_{\text{КММП}}$ до двух слагаемых S_k получим:

$$0 = \frac{1}{k} \sum_{t=1}^k l'_y(F_{1k}(z_{1t}), F_{2k}(z_{2t}), \gamma) \cong S_k + H_k(\hat{\gamma}_{\text{КММП}} - \gamma). \quad (61)$$

Таким образом, получаем свойство оценок КММП:

$$\sqrt{k}(\hat{\gamma}_{\text{КММП}} - \gamma) \cong -\sqrt{k} \frac{S_k}{H_k}. \quad (62)$$

Выделим некоторые асимптотические свойства оценок КММП:

- $\hat{\gamma}_{\text{КММП}} \rightarrow \gamma$, при $k \rightarrow +\infty$;
- $\sqrt{k}(\hat{\gamma}_{\text{КММП}} - \gamma) \xrightarrow{d} N(0, \sigma^2 h^{-2})$, где

$$\sigma^2 = \mathbb{V}_{\text{ar}} \left(l'_y(F_{1k}(z_{1t}), F_{2k}(z_{2t}), \gamma) + W_1(z_{1t}) + W_2(z_{2t}) \right) \quad (63)$$

$$h = \mathbb{E} \left(l''_{yy}(F_{1k}(z_{1t}), F_{2k}(z_{2t}), \gamma) \right) \quad (64)$$

$$W_i(z_{it}) = - \int I_{\{F_i(z_{it}) \leq y_i\}} l''_{y_i \gamma}(y_{1t}, y_{2t}, \gamma) c(y_{1t}, y_{2t}, \gamma) dy_1 dy_2 \quad (65)$$

Также стоит отметить, что КММП хорошо работает тогда, когда случайные величины являются независимыми. В остальных случаях параметрические методы являются более эффективными.

Существует еще одна полупараметрическая процедура оценивания параметров копула-функции, однако она хорошо подходит для оценки копулы Стьюдента, поэтому не будем рассматривать ее, а перейдем к непараметрическим методам оценки копула-функций.

Для начала введем определение ядерных функций.

Определение 9. (ядерные функции) [10]

Ядерной функцией называется некоторая действительная функция $K(x)$, $-\infty < x < +\infty$, удовлетворяющая условиям:

- *Четность:* $K(x) = K(-x)$;
- $K_{\max} = K(0)$
- $\int_{-\infty}^{+\infty} K(x) dx = 1$
- *Неотрицательность:* $K(x) \geq 0$;

- Ограниченность: $|K(x)| \leq M$;
- $\int_{-\infty}^{+\infty} x^2 K(x) dx < \infty$

Возьмем некоторое n -мерное ядро (для простоты возьмем произведение одномерных ядер) и его кумулятивную функцию:

$$\kappa(\mathbf{t}) = \prod_{i=1}^n \kappa_i(t_i), K(\mathbf{T}) = \prod_{i=1}^n \int_{-\infty}^{t_i} \kappa_i(t) dt = \prod_{i=1}^n K_i(t_i), \quad (66)$$

также возьмем

$$\kappa(\mathbf{t}, \boldsymbol{\eta}) = \prod_{i=1}^n \kappa_i\left(\frac{t_i}{\eta_i(k)}\right), K(\mathbf{T}, \boldsymbol{\eta}) = \prod_{i=1}^n K_i\left(\frac{t_i}{\eta_i(k)}\right), \quad (67)$$

здесь $\boldsymbol{\eta}(k)$ – матрица с элементами $\eta_i(k)$ на главной диагонали и определителем $|\boldsymbol{\eta}(k)|$, называемая шириной окна ядерной функции.

Ядерная оценка \hat{f}_i плотности f_i случайной величины X_{it} в точке x_i определяется следующим образом:

$$\hat{f}_i(x_i) = \frac{1}{k \cdot \eta_i} \sum_{i=1}^k \kappa\left(\frac{x_i - X_{it}}{\eta_i(k)}\right), \quad (68)$$

а оценка плотности вектора \mathbf{X} в точке $\mathbf{x} = (x_1, \dots, x_n)^T$ равна:

$$\hat{f}(\mathbf{x}) = \frac{1}{k \cdot |\boldsymbol{\eta}(k)|} \sum_{i=1}^k \kappa(\mathbf{x} - \mathbf{X}_i, \boldsymbol{\eta}(k)). \quad (69)$$

Тогда оценка функции распределения будет иметь вид:

$$\hat{F}_i(x_i) = \int_{-\infty}^{x_i} \hat{f}_i(y) dy, \quad (70)$$

$$\hat{F}(\mathbf{x}) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} \hat{f}_i(\mathbf{y}) d\mathbf{y}. \quad (71)$$

Пользуясь следствием из теоремы Склара и выражением (3), запишем оценку копула-функции:

$$\hat{C}(\mathbf{y}) = \hat{H}(\hat{\xi}) = \hat{H}(\hat{\xi}_1, \dots, \hat{\xi}_n), \quad (72)$$

где $\hat{\xi}_i = \inf_{x \in \mathbb{R}} \{x: F_i(x) \geq y_i\}$ – ядерная оценка квантиля y_i -уровня для

маргинальной функции распределения X_{it} . Такая оценка является асимптотически нормальной и состоятельной.

2.3. Процедура выбора наилучшей копула-функции

При анализе многомерных распределений часто бывает так, что несколько различных копула-функций хорошо описывают исследуемые данные. В таком случае приходится выбирать среди всех моделей наилучшую. Для того, чтобы сделать это используются различные методы, рассмотрим наиболее популярные:

- информационный критерий Акаике (AIC);
- тесты отношения псевдоправдоподобия;
- байесовский метод выбора копулы.

Критерий Акаике – это критерий, с помощью которого можно выбрать копулу, отвечающую минимальному из следующих значений:

$$AIC_c(\hat{\alpha}_l) = -2 \sum_{i=1}^k \ln c_l(y_{1i}, \dots, y_{ni}, \hat{\alpha}_l) + 2q_l, l = 1, \dots, L, \quad (73)$$

здесь q_l – штраф за количество параметров в модели. Критерий хорош тем, что он достаточно простой, однако для стандартных (не вложенных) копула-функций теряется смысл ввода штрафной части в формуле (73) [10].

Тогда рассмотрим другой метод сравнения копула-функций – **тесты отношения псевдоправдоподобия**. Идея тестов – сравнение исследуемых моделей с точки зрения правдоподобия. Лучше та модель, которая оказалась наиболее правдоподобной. Рассмотрим нулевую гипотезу теста:

$$H_0: \max_{j=2, \dots, L} \mathbb{E} \left(\ln \frac{c_j(F_1^0(\delta_{1t}), \dots, F_n^0(\delta_{nt}), \alpha_j^*)}{c_1(F_1^0(\delta_{1t}), \dots, F_n^0(\delta_{nt}), \alpha_1^*)} \right) \leq 0. \quad (74)$$

Пусть $C_1(\cdot)$ – наилучшая из рассматриваемых моделей, модели $C_l(\cdot), l = 2, \dots, L$ – остальные, а $C_0(\cdot)$ – истинная, не наблюденная копула, описывающая зависимости на самом деле. Тогда при постановке гипотезы (74) используется расстояние Кульбака-Лейблера, а затем проверяется является ли модель $C_1(\cdot)$ наилучшей среди всех остальных. Тогда очевидно, что вид альтернативной гипотезы будет следующий:

$$H_1: \max_{j=2,\dots,L} \mathbb{E} \left(\ln \frac{c_j(F_1^0(\delta_{1t}), \dots, F_n^0(\delta_{nt}), \alpha_j^*)}{c_1(F_1^0(\delta_{1t}), \dots, F_n^0(\delta_{nt}), \alpha_1^*)} \right) > 0. \quad (75)$$

При принятии альтернативной гипотезы делается вывод, что среди прочих моделей есть копула, описывающая исходные данные лучше, чем $C_1(\cdot)$.

Статистика теста рассчитывается по следующему правилу:

$$LR_n(F_{1k}, \dots, F_{nk}, \tilde{\alpha}_j, \tilde{\alpha}_1) = \frac{1}{k} \sum_{t=1}^k \left(\ln \frac{c_j(F_{1k}(\delta_{1t}), \dots, F_{nk}(\delta_{nt}), \tilde{\alpha}_j)}{c_1(F_{1k}(\delta_{1t}), \dots, F_{nk}(\delta_{nt}), \tilde{\alpha}_1)} \right), \quad (76)$$

Здесь знак « \sim » над α означает, что параметры $\tilde{\alpha}_j$ и $\tilde{\alpha}_1$ оценены каноническим методом максимального правдоподобия.

Для принятия или отклонения основной гипотезы, в отличие от других статистических тестов, тут недостаточно сравнить статистику с некоторой критической точкой. Здесь необходимо проделать нетривиальную работу с использованием метода Монте-Карло. Однако, при всей сложности реализации теста, он дает хорошие результаты.

Другим интересным подходом сравнения копула-функций является **байесовский**. В этом подходе параметры копулы рассматриваются как случайные величины, а значит они не подлежат оцениванию. Параметризация плотности копула-функций в этом методе рассматривается в терминах τ -Кендалла. Априорные распределения параметров заменяются априорными распределениями τ -Кендалла, и они берутся одними и теми же для всех тестируемых копул и используется как база при сравнении.

Пусть $C_q(\cdot)$, $q = 1, \dots, Q$ – множество копул, подлежащих сравнению. Тогда на первом шаге необходимо протестировать следующие Q гипотез:

$$H_q: \text{данные извлечены из копулы } C_q(\cdot), q = 1, \dots, Q.$$

Для проверки гипотез необходимо вычислить вероятности реализации гипотезы H_q для заданных наблюдений Y , которые представляют собой k независимых пар (y_{1t}, y_{2t}) , $t = 1, \dots, k$ с компонентами, имеющими равномерное распределение. Применяя теорему Байеса, получаем:

$$P(H_q|X, Inf) = \frac{P(X|H_q, Inf)P(H_q|Inf)}{P(X|Inf)}, \quad (77)$$

где $P(X|H_q, Inf)$ – функция правдоподобия, $P(H_q|Inf)$ – априорное распределение для копулы, $P(X|Inf)$ – нормирующая константа, Inf – дополнительная информация. Наилучше копулой признается та, у которой максимальная апостериорная вероятность $P(H_q|X, Inf)$.

Если записать это выражение в терминах τ -Кендалла получим:

$$\begin{aligned} P(H_q|X, Inf) &= \int_{-1}^1 P(H_q, \tau|X, Inf) d\tau = \\ &= \int_{-1}^1 \frac{P(X|H_q, \tau, Inf)P(H_q|\tau, Inf)P(\tau|Inf)}{P(X|Inf)} d\tau, \end{aligned} \quad (78)$$

где $P(H_q|\tau, Inf)$ – априорная вероятность гипотезы H_q для копулы, $P(\tau|Inf)$ – априорная плотность τ -Кендалла, правдоподобие $P(X|H_q, \tau, Inf)$, зависящее теперь еще от τ , вычисляется по правилу:

$$P(X|H_q, \tau, Inf) = \prod_{t=1}^k P(y_{1t}, y_{2t}|\tau, q, Inf) = \prod_{t=1}^k c_q(y_{1t}, y_{2t}|g_q^{-1}(\tau)), \quad (79)$$

где $c_q(y_{1t}, y_{2t}|g_q^{-1}(\tau))$ – плотность q -ой копулы, параметризированной в терминах τ -Кендалла.

Вероятность $P(\tau|Inf)$ можно вычислить по правилу:

$$P(\tau|Inf) = \begin{cases} \frac{1}{\lambda(\Delta)}, & \tau \in \Delta, \\ 0, & \tau \notin \Delta \end{cases}, \quad (80)$$

где Δ – множество, которое хранит информация о τ (то есть, если известно, что $\tau \geq 0$, то $\Delta = [0, 1]$, если о τ ничего не известно, то $\Delta = [-1, 1]$), $\lambda(\Delta)$ – длина интервала (норма Лебега).

Для вероятности $P(H_q|\tau, Inf)$ выполняется следующее соотношение:

$$P(H_q|\tau, Inf) \propto I_{\{\tau \in \Omega_q\}}, \quad (81)$$

де Ω – область значений τ в условиях выполнения гипотезы H_q .

Тогда, если подставить (79), (80) и (81) в (78) получим:

$$\begin{aligned}
 P(H_q|X, Inf) &= \frac{1}{P(X|Inf)} \int_{-1}^1 \prod_{t=1}^k c_q(y_{1t}, y_{2t} | g_q^{-1}(\tau)) \cdot \frac{I_{\{\tau \in \Omega_q \cap \Delta\}}}{\lambda(\Delta)} d\tau = \\
 &= \frac{1}{P(X|Inf)\lambda(\Delta)} \int_{\Omega_q \cap \Delta} \prod_{t=1}^k c_q(y_{1t}, y_{2t} | g_q^{-1}(\tau)), \quad (82)
 \end{aligned}$$

где нормирующую константу $P(X|Inf)$ можно рассчитать по правилу:

$$P(X|Inf) = \sum_{q=1}^Q P(X|H_q, Inf)P(H_q|Inf). \quad (83)$$

Такой метод является очень интересным, так как он не требует предварительной оценки параметров копулы и дает хорошие результаты.

ГЛАВА 3. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ РИСКОВ С ПОМОЩЬЮ КОПУЛА-ФУНКЦИЙ

3.1. Меры риска портфеля. Модель оптимального портфеля ценных бумаг

Моделирование оптимального портфеля ценных бумаг – специфическая задача нахождения удовлетворяющего некоторой выбранной инвестиционной стратегии сочетания ценных бумаг. При выборе ценных бумаг и расчете оптимального портфеля основываются на их исторических котировках. Анализируя исторические данные и их поведение, инвесторы пытаются выявить некоторые тенденции, с помощью которых они вкладываются в те или иные бумаги, такой процесс называется прогнозированием динамики исследуемых котировок. На курс ценных бумаг влияет очень много факторов, как внутренних (в самой компании), так и внешних, поэтому прогнозирование – это процесс с неопределенной вероятностью. В связи с этим при построении портфеля перед инвестором встает выбор: инвестировать в ценные бумаги с наибольшей доходностью, но с большим риском разорения или, наоборот, инвестировать в безрисковые активы с небольшой доходностью. Чаще всего выбирают последний вариант.

Под риском в портфельном анализе чаще всего подразумевается среднеквадратическое отклонение (СКО) доходностей котировок, показывающее средний разброс, однако СКО не позволяет учесть скошенность асимметричных распределения доходностей. В связи с этим в конце прошлого столетия была предложена новая мера риска – стоимость под риском VaR (Value-at-Risk). Стоимостью под риском VaR называется величина убытков, которая не будет превышена с некоторой заданной вероятностью [16].

$$\begin{aligned} VaR(p) &= \inf(k \in \mathbb{R}, P(K > k) \leq 1 - p) = \inf(k \in \mathbb{R}, 1 - F(k) \leq 1 - p) = \\ &= \inf(k \in \mathbb{R}, F(k) \geq p) = x_p \end{aligned} \quad (84)$$

Однако показатель VaR также имеет некоторые недостатки, например, VaR сложно оптимизировать для дискретных распределений, поскольку в этом случае он не будет выпуклой и гладкой функцией и будет иметь множество локальных экстремумов. Также VaR не является когерентной мерой риска, то есть может возникнуть ситуация, когда диверсификация портфеля повлечет за собой увеличение величины VaR . Для преодоления вышеуказанных недостатков стоимости под риском, используется другой показатель – средние ожидаемые потери $CVaR$ (Conditional Value-at-Risk) [5].

$$CVaR(p) = E(K | K > VaR(p)) \quad (85)$$

Таким образом, целевой переменной в модели, подлежащей оптимизации, будет показатель $CVaR$.

Для построения модели рассмотрим 5 акций российских компаний. Выберем эмитенты так, чтобы портфель был диверсифицированным. Для этого рассмотрим ценные бумаги компаний из разных отраслей экономики: финансы – Сбербанк, энергетика – Русгидро, нефть и газ – Роснефть, добыча металлов – Алроса, транспорт – Аэрофлот.

Загрузим ежедневные значения цен закрытия акций компаний за годовой период с 15 апреля 2020 года по 15 апреля 2021 года (505 значения).

Таблица 3. Часть исходных данных

| Дата | SBER | HYDR | ROSN | ALRS | AFLT |
|----------|--------|--------|--------|-------|-------|
| 20190415 | 237,8 | 0,532 | 428,9 | 94,3 | 98,72 |
| 20190416 | 233,24 | 0,5314 | 435,15 | 95,4 | 99,12 |
| 20190417 | 234,3 | 0,533 | 436,5 | 96 | 98,62 |
| 20190418 | 232,79 | 0,529 | 435,65 | 95,8 | 98 |
| 20190419 | 232,6 | 0,5276 | 432 | 95,79 | 97,6 |
| 20190422 | 235,41 | 0,5465 | 439,6 | 95,96 | 97,6 |
| 20190423 | 235,67 | 0,542 | 441 | 96,3 | 96,8 |
| 20190424 | 236 | 0,546 | 441,3 | 95,58 | 96,32 |
| 20190425 | 227,8 | 0,537 | 436 | 95,2 | 96,34 |
| 20190426 | 223,18 | 0,5355 | 434,45 | 94 | 97,12 |

| | | | | | |
|----------|--------|--------|--------|-------|-------|
| 20190429 | 228,8 | 0,5374 | 433,7 | 94,49 | 96,98 |
| 20190430 | 225,17 | 0,54 | 428,8 | 94,23 | 96,32 |
| 20190502 | 229,1 | 0,5409 | 423,25 | 95,34 | 96,86 |
| 20190503 | 232,52 | 0,5425 | 423,55 | 95,6 | 96,7 |

Разделим выборку на тренировочную (475 элементов) и тестовую (30 элементов) части. Далее, для тренировочной выборки рассчитаем ежедневные значения доходностей акций. В работе будут использоваться арифметические доходности ценных бумаг, так как они лучше подходят для моделирования портфеля на небольших временных отрезках.

$$r_i = \frac{p_i - p_{i-1}}{p_{i-1}} \quad (86)$$

Так, получаем выборку из 474 объектов из ежедневных значений доходностей.

Таблица 4. Ежедневные доходности акций

| Дата | SBER | HYDR | ROSN | ALRS | AFLT |
|-------------|-------------|-------------|-------------|-------------|-------------|
| 20190416 | -1,92% | -0,11% | 1,46% | 1,17% | 0,41% |
| 20190417 | 0,45% | 0,30% | 0,31% | 0,63% | -0,50% |
| 20190418 | -0,64% | -0,75% | -0,19% | -0,21% | -0,63% |
| 20190419 | -0,08% | -0,26% | -0,84% | -0,01% | -0,41% |
| 20190422 | 1,21% | 3,58% | 1,76% | 0,18% | 0,00% |
| 20190423 | 0,11% | -0,82% | 0,32% | 0,35% | -0,82% |
| 20190424 | 0,14% | 0,74% | 0,07% | -0,75% | -0,50% |
| 20190425 | -3,47% | -1,65% | -1,20% | -0,40% | 0,02% |
| 20190426 | -2,03% | -0,28% | -0,36% | -1,26% | 0,81% |
| 20190429 | 2,52% | 0,35% | -0,17% | 0,52% | -0,14% |
| 20190430 | -1,59% | 0,48% | -1,13% | -0,28% | -0,68% |
| 20190502 | 1,75% | 0,17% | -1,29% | 1,18% | 0,56% |
| 20190503 | 1,49% | 0,30% | 0,07% | 0,27% | -0,17% |
| 20190506 | 0,12% | -1,35% | 0,34% | -0,42% | -1,82% |

Проведем первичный анализ данных. Он будет состоять из двух этапов: тестирование гипотезы о нормальном распределении данных, определение распределения с тяжелыми хвостами для исследуемых данных.

Тестировать гипотезу о нормальности будем в программной среде R с помощью статистического критерия Колмогорова-Смирнова. Критерий основывается на сравнении эмпирической и теоретической функций распределения. Проведем тест на 5% уровне значимости. В качестве результата теста выведем *p-значения* статистики (*pvalue*). В случае, если значение статистики больше 5%, гипотеза о нормальности принимается.

В результате теста Колмогорова-Смирнова отклоняются гипотезы о нормальном распределении для доходностей всех рассматриваемых эмитентов. В таблице 5 приведены *p-значения* статистики для каждой акции. Все значения *pvalue* оказались ниже уровня значимости. Таким образом, моделирование портфеля с применением модели Марковица может привести к получению неоптимальных результатов и плохих портфелей [20].

Таблица 5. Результаты теста Колмогорова-Смирнова

| Акция | SBER | HYDR | ROSN | ALRS | AFLT |
|---------------|-------------|--------------|--------------|----------|--------------|
| <i>pvalue</i> | 0.002467783 | 2.444741e-06 | 1.547181e-05 | 0.012624 | 2.124785e-05 |

Рассмотрим распределение хвостов доходностей. Это позволит визуально отвергнуть гипотезу о нормальном распределении. Для этого с помощью ядерных функций (см. определение 9) построим распределение данных. Для простоты реализации рассмотрим гауссовское (нормальное) ядро с фиксированной шириной интервала (рис. 9).

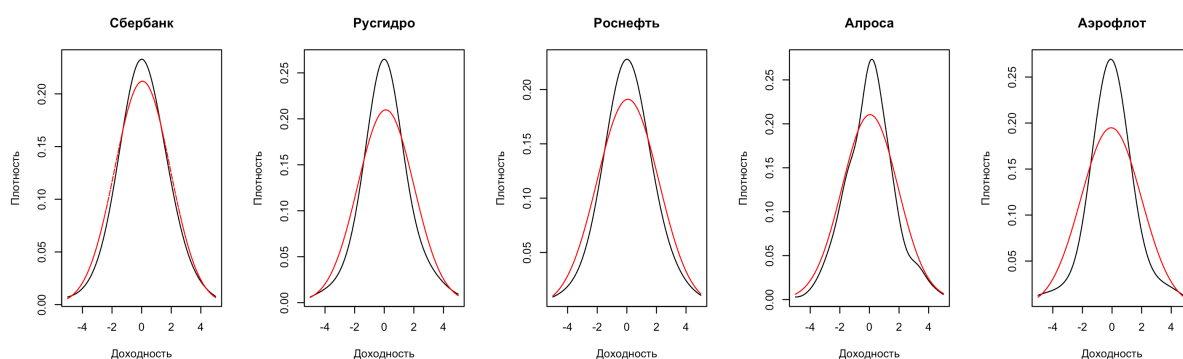


Рисунок 9. Ядерные оценки функции плотности по рассматриваемым акциям

Анализ рисунка 9 показывает, что распределения доходностей имеют тяжелые хвосты. Причем как правый хвост, так и левый. Наиболее толстые хвосты наблюдаются у Аэрофлота. Связано это с тем, что акции компании дважды потерпели резкий спад, связанный с первой и второй волной пандемии. Если ко второй волне коронавируса компании прочих отраслей научились существовать в условиях пандемии, то для авиакомпании вторая волна стала в очередной раз рычагом спада. В связи с этим особо четко прослеживаются тяжелые хвосты у этой компании. Так, мы приходим к выводу, что вероятность наступления экстремально негативных событий гораздо выше, чем экстремально позитивных. Именно из-за этого для моделирования портфеля используются такие несимметричные меры риска, как VaR и $CVaR$ [5].

Для нахождения оптимального портфеля используется метод имитационного моделирования Монте-Карло. Будет сгенерировано множество портфелей (долей), из которого будет выбран портфель исходя из двух принципов:

- портфель минимального риска $CVaR \rightarrow \min$;
- портфель максимального отношения доходности к риску (обратное значение коэффициента вариации) $\frac{r}{CVaR} \rightarrow \max$.

Выделим несколько предпосылок данного подхода нахождения оптимального портфеля:

- Веса ценных бумаг в портфеле положительные, то есть отсутствуют короткие продажи;
- Оценки плотности распределения данных будут неизменны во времени;
- В модели учитываются только такие показатели, как доходность и риск;
- Множество сгенерированных портфелей является полным, а значит описывает все возможные комбинации ценных бумаг.

Процесс построения оптимального портфеля состоит из трех этапов: генерация портфелей, оценка копула-функций, расчет портфеля исходя из двух принципов, описанных выше.

Смоделируем множество портфелей. На генерируемые доли накладываются следующие ограничения:

$$\begin{cases} 0 \leq \omega_i \leq 1 \\ \sum \omega_i = 1, i = 1, \dots, N \end{cases} \quad (87)$$

Для начала необходимо смоделировать $N - 1$ стандартных равномерных случайных величин, то есть $U(0; 1)$, тогда получаем выборку (z_1, \dots, z_{N-1}) . Далее, необходимо упорядочить элементы выборки и добавить на концах интервала 0 и 1, то есть $(0, z_{(1)}, \dots, z_{(N-1)}, 1)$. Далее, необходимо вычесть из последующих элементов выборки предыдущие. Тогда получим значения долей портфеля, удовлетворяющие условию (87).

$$\begin{cases} \omega_1 = z_{(1)} \\ \omega_2 = z_{(2)} - z_{(1)} \\ \dots \\ \omega_{N-1} = z_{(N-1)} - z_{(N-2)} \\ \omega_N = 1 - z_{(N-1)} \end{cases} \quad (88)$$

Выполняя алгоритм M раз, получаем необходимое множество портфелей (в данной работе $M = 10000$). Таким образом, получаем матрицу долей размерности $M \times N$:

$$\Omega = \begin{pmatrix} \omega_{11} & \dots & \omega_{1N} \\ \vdots & \ddots & \vdots \\ \omega_{M1} & \dots & \omega_{MN} \end{pmatrix}. \quad (89)$$

В случае трех портфелей множество сгенерированных портфелей графически представлено на рисунке 10.

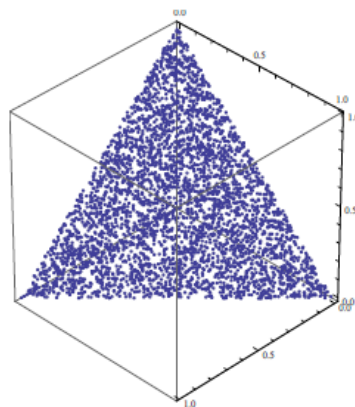


Рисунок 10. Сгенерированное множество портфелей в случае трех ценных бумаг

В рамках модели предполагается оценивание параметров четырех параметрических копул: Гаусса (нормальная), Стьюдента, Клейтона и Гумбеля. Для оценки копула-функций необходимо использовать эмпирическую функцию распределения (кумулятивную функцию распределения). Кумулятивная функция соответствует вероятности того, что доходность по данному активу окажется меньше или равна значению r_i .

После того, как получаем значения векторов кумулятивных функций генерируется копула-функций, оцениваются ее параметры методом максимального правдоподобия и выбирается наилучшая среди них. В качестве критерия выбора выступает логарифм правдоподобия. Наилучшей будем считать такую копулу, у которой $\ln(L)$ наибольший.

Таблица 6. Значения логарифма правдоподобия для копула-функций

| Копула | Нормальная | Стьюдента | Гумбеля | Клейтона |
|-------------------|------------|-----------|----------|----------|
| Значение $\ln(L)$ | 266,6947 | 314,7413 | 186,4849 | 252,6444 |

Анализ таблицы 6 показывает, что лучше всего исходные данные описывает копула Стьюдента. Для моделирования портфеля далее копула-функция Стьюдента и будет использоваться.

В этой работе будем рассматривать 3 горизонта для инвестирования: неделя ($t = 7$ дней), полмесяца ($t = 15$ дней), месяц ($t = 30$ дней). Используя копула-функцию, подобранную ранее, сгенерируем 7, 15 и 30 значений доходностей. В итоге получим три матрицы Ψ следующего вида:

$$\Psi = \begin{pmatrix} r_{11} & \cdots & r_{1N} \\ \vdots & \ddots & \vdots \\ r_{t1} & \cdots & r_{tN} \end{pmatrix}. \quad (90)$$

Далее, считаем кумулятивную (накопленную) доходность (недельную, полумесячную и месячную) на основе рассчитанных ранее доходностей.

$$r_{\text{накоп}} = \left(\prod_{j=1}^t (1 + r_{j1}) - 1 \right) \quad (91)$$

В результате получим вектор накопленных доходностей. Повторим итерации Z раз и получим матрицу накопленных доходностей (в работе $Z = 1000$).

$$\Psi_{\text{накоп}} = \begin{pmatrix} r_{\text{накоп } 11} & \cdots & r_{\text{накоп } 1N} \\ \vdots & \ddots & \vdots \\ r_{\text{накоп } Z1} & \cdots & r_{\text{накоп } ZN} \end{pmatrix} \quad (92)$$

И наконец, перейдем к матрице кумулятивных доходностей по каждому портфелю акций, для этого необходимо выполнить следующее матричное умножение:

$$\Psi_{\text{накоп}}^{\text{портф}} = \Omega \cdot (\Psi_{\text{накоп}})^T. \quad (93)$$

Теперь можно рассчитать ожидаемые значения доходностей и $CVaR$ для каждого портфеля акций. Для получения оценки доходностей необходимо найти среднее арифметическое всех значений доходностей портфеля, а для получения оценки $CVaR$ необходимо: [5]

- сортировать исходные данные по убыванию;
- рассчитать порядковый номер, соответствующий 5% квантилю для вычисления VaR ;
- вычислить $CVaR$ как среднее всех значений меньше VaR .

Получив значения доходностей и риска, можно построить достижимое множество для каждого горизонта инвестирования (рис. 11).

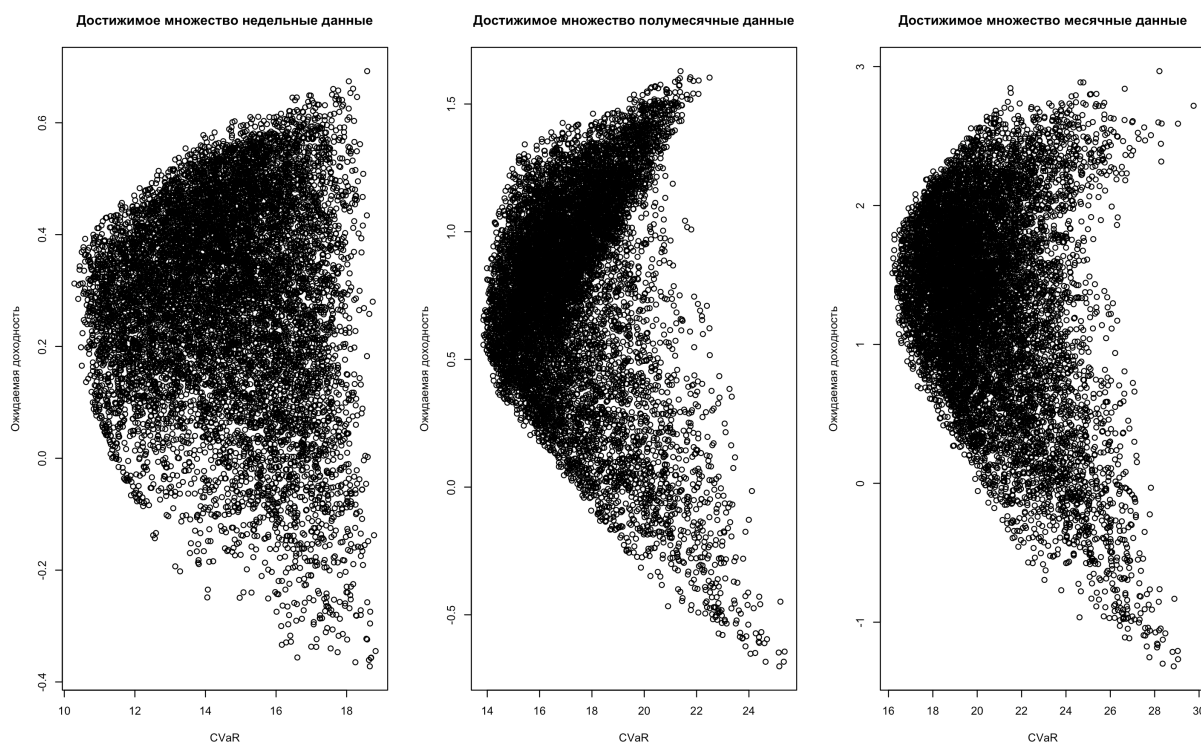


Рисунок 11. Достижимое множество для трех периодов инвестирования

И, наконец, решая оптимизационную задачу, получим по двум принципам для каждого горизонта инвестирования оптимальные портфели.

Таблица 7. Портфели для недельного горизонта инвестирования

| № п/п | Ценная бумага | $r/CVaR \rightarrow \max$ | $CVaR \rightarrow \min$ |
|-------|---------------|---------------------------|-------------------------|
| 1 | Сбербанк | 68,8% | 31,1% |
| 2 | Русгидро | 27,4% | 7,4% |
| 3 | Роснефть | 1,7% | 13,7% |
| 4 | Алроса | 1,4% | 41,3% |
| 5 | Аэрофлот | 0,8% | 6,5% |

Можно заметить, что портфель минимального риска является более диверсифицированным. Наибольшую долю в портфеле максимума отношения доходности к риску имеет Сбербанк, в портфеле минимального риска вклад Сбербанка также весьма значимый. Это связано с тем, что акции Сбербанка являются стабильными на рынке (голубые фишки).

Таблица 8. Портфели для полумесечного горизонта инвестирования

| № п/п | Ценная бумага | $r/CVaR \rightarrow \max$ | $CVaR \rightarrow \min$ |
|-------|---------------|---------------------------|-------------------------|
| 1 | Сбербанк | 48,8% | 36,3% |

| | | | |
|---|----------|-------|-------|
| 2 | Русгидро | 0,0% | 8,5% |
| 3 | Роснефть | 49,4% | 14,4% |
| 4 | Алроса | 1,4% | 20,4% |
| 5 | Аэрофлот | 0,4% | 20,4% |

В случае полумесячного инвестирования также основная доля портфеля приходится на голубые фишки: Сбербанк и Роснефть. Диверсификация портфеля – способ снижения портфельного риска, в связи с этим портфель минимального риска является более диверсифицированным.

Таблица 9. Портфели для месячного горизонта инвестирования

| № п/п | Ценная бумага | $r/CVaR \rightarrow \max$ | $CVaR \rightarrow \min$ |
|-------|---------------|---------------------------|-------------------------|
| 1 | Сбербанк | 2,1% | 30,4% |
| 2 | Русгидро | 59,7% | 18,7% |
| 3 | Роснефть | 36,8% | 24,2% |
| 4 | Алроса | 0,9% | 11,1% |
| 5 | Аэрофлот | 0,6% | 15,7% |

В таблице 9 представлены доли для месячного периода инвестирования.

Проверим результаты модели на тестовой выборке. Пусть $S = 1000000$ рублей – капитал для вложения в ценные бумаги. Тогда в соответствии с полученными долями вложим деньги в рассмотренные эмитенты на различные периоды (горизонты инвестирования).

Для начала рассмотрим принцип минимального риска. Предполагается покупка акций на всю сумму и продажа акций после истечения временного горизонта.

Таблица 10. Портфель минимального риска в случае недельного вложения

| Эмитент | SBER | HYDR | ROSN | ALRS | AFLT | Сумма |
|--------------------------|----------|----------|-----------|-----------|----------|---------|
| Сумма вложения в эмитент | 311473,9 | 73546,76 | 136668,97 | 413366,72 | 64943,65 | 1000000 |

| | | | | | | |
|-------------------------------------|----------|----------|----------|----------|----------|------------|
| Количество акций | 1133,66 | 92441,88 | 253,04 | 3989,26 | 935,25 | 98753,09 |
| Сумма после окончания периода | 331403,7 | 75210,7 | 144538,6 | 431637,5 | 65935,01 | 1048725,57 |

При недельном инвестировании в модели минимального риска доходность портфеля составляет 4,87%. В таблице 10 видно, что в рамках данного портфеля, ни один из эмитентов не пошел в убыток. Таким образом, модель, основанная на копула-функциях, в первом случае дала хороший результат. Проверим, действительно ли выбранный портфель является диверсифицированным. Для этого сравним коэффициент вариации портфеля с коэффициентами вариации отдельных ценных бумаг. Коэффициенты вариации ценной бумаги и портфеля считаются по следующим формулам:

$$W_i = \frac{\sigma_i}{r_i} \quad (94a)$$

$$W = \frac{\sigma}{r} = \frac{\sqrt{\sum \omega_i^2 \sigma_i^2 + 2 \sum \sum \omega_i \omega_j \rho_{ij} \sigma_i \sigma_j}}{\sum \omega_i r_i} \quad (94b)$$

Коэффициент вариации портфеля в случае недельного горизонта составил $W = 143,41$, а суммарный коэффициент корреляции бумаг составил $\sum W_i = 161,56$. Таким образом, делаем вывод, что портфель диверсифицирован, и добавление выбранных акций в портфель снижает риск.

Таблица 11. Портфель минимального риска в случае полумесячного вложения

| | | | | | | |
|--------------------------------|-----------|----------|-----------|-----------|-----------|---------|
| Эмитент | SBER | HYDR | ROSN | ALRS | AFLT | Сумма |
| Сумма вложения в эмитент | 362623,45 | 85184,52 | 144286,21 | 203982,00 | 203923,82 | 1000000 |

| | | | | | | |
|-------------------------------|-----------|-----------|-----------|-----------|-----------|------------|
| Количество акций | 1319,83 | 104700,74 | 252,60 | 1885,23 | 2892,54 | 111050,94 |
| Сумма после окончания периода | 384361,06 | 85111,23 | 147064,83 | 203322,17 | 196692,48 | 1016551,77 |

В случае полумесячного горизонта инвестирования модель также дала положительную доходность – 1,66%. В модели минимального риска часто случается так, что доходность от акции бывает невеликой, однако она не несет больших рисков. Коэффициент вариации портфеля составляет $W = 115,19 < 161,56$. Портфель также можно считать правильно подобранным.

Таблица 12. Портфель минимального риска в случае месячного вложения

| | | | | | | |
|-------------------------------|-----------|-----------|-----------|-----------|-----------|------------|
| Эмитент | SBER | HYDR | ROSN | ALRS | AFLT | Сумма |
| Сумма вложения в эмитент | 304302,30 | 186864,00 | 241568,60 | 110594,40 | 156670,70 | 1000000 |
| Количество акций | 1107,56 | 230184,77 | 420,38 | 1035,53 | 2368,77 | 235117,01 |
| Сумма после окончания периода | 319531,26 | 190731,10 | 233329,25 | 109900,60 | 157476,08 | 1010968,29 |

Для месячного периода инвестирования результат практически совпал с полумесячным, и доходность составляет 1,1%. Для всех периодов инвестирования по принципу минимального риска средняя доходность составляет 2,54%. Результаты моделирования для рассмотренного принципа можно считать положительными, действительно, имея небольшой риск, инвестор получил небольшую доходность. Коэффициент вариации портфеля $W = 50,23$. Так, в месячном случае диверсификация портфеля снижает риск.

Таблица 13. Портфель максимального отношения доходности к риску в случае недельного вложения

| Эмитент | SBER | HYDR | ROSN | ALRS | AFLT | Сумма |
|-------------------------------------|-----------|-----------|----------|----------|---------|------------|
| Сумма вложения в эмитент | 687539,83 | 273547,31 | 16985,32 | 13785,45 | 8142,11 | 1000000 |
| Количество акций | 2502,42 | 343825,17 | 31,45 | 133,04 | 117,25 | 346609,33 |
| Сумма после окончания периода | 731532,37 | 279736,16 | 17963,36 | 14394,76 | 8266,40 | 1051893,04 |

Для модели с максимизацией отношения доходности к риску при недельном горизонте доходность оказалась довольно большой – 5,19%, что свидетельствует о том, что риск был «оправданным», то есть максимизация совокупного показателя доходности и риска принесла хорошие результаты. Коэффициент вариации составил $W = 39,92$, что значительно меньше коэффициента вариации по отдельным бумагам.

Таблица 14. Портфель максимального отношения доходности к риску в случае полумесячного вложения

| Эмитент | SBER | HYDR | ROSN | ALRS | AFLT | Сумма |
|-------------------------------------|-----------|--------|-----------|----------|---------|------------|
| Сумма вложения в эмитент | 487670,11 | 164,33 | 494037,95 | 14423,37 | 3704,24 | 1000000 |
| Количество акций | 1774,96 | 206,54 | 914,72 | 139,19 | 53,34 | 3088,76 |
| Сумма после окончания периода | 516903,69 | 167,90 | 532547,48 | 15012,17 | 3627,43 | 1068258,66 |

В случае полумесячного горизонта модель также показала высокие результаты. Доходность в таком случае составляет 6,83%. В данном случае максимизация доходности также дала хорошие результаты. При полумесячном горизонте результаты также оказались хорошими, $W = 42,52$.

*Таблица 15. Портфель максимального отношения доходности к
риску в случае месячного вложения*

| Эмитент | SBER | HYDR | ROSN | ALRS | AFLT | Сумма |
|-------------------------------------|----------|-----------|-----------|---------|---------|------------|
| Сумма вложения в эмитент | 20814,85 | 596735,81 | 368275,48 | 8654,74 | 5519,12 | 1000000 |
| Количество акций | 75,76 | 733451,09 | 644,74 | 79,99 | 78,29 | 734329,86 |
| Сумма после окончания периода | 21856,54 | 607737,57 | 357862,93 | 8489,16 | 5204,41 | 1001150,62 |

Примером «неоправданного» риска является модель максимизации отношения доходности к риску при месячном горизонте инвестирования. Доходность хоть и оказалась положительной, но она значительно меньше других случаев, всего 0,12%. Это может быть связано с тем, что месяц – достаточно большой срок для точного прогноза с учетом того, что на цены акций влияет множество прочих факторов, невозможных отразить в математических моделях. С точки зрения диверсификации, портфель можно считать правильным, коэффициент вариации портфеля почти в 7 раз меньше суммарного коэффициента по отдельной бумаге.

Средняя доходность по принципу максимизации совокупного показателя доходности и риска составила 4,04%. Средняя доходность по всем горизонтам инвестирования и показателям составила 3,29%.

Таким образом, мы делаем вывод, что модель оптимального портфеля, основанная на многомерных копула-функциях, хорошо описывает исходные данные и дает надежные результаты. На основе анализа данных по 5 акциям различных отраслей за двухгодовой период можно сделать вывод, что наиболее результативно применять модель на средних периодах – неделя, полмесяца. Как было отмечено ранее, оценки плотности распределения и копула-функций неизменные, в связи с этим при моделировании на более длительные периоды результаты могут

ухудшиться. Это связано с тем, что статистическое распределение данных меняется со временем.

3.2. Модель прогнозирования курса валюты

Еще одним важным направлением применения копулярных функций является эконометрическое моделирование, в частности временные ряды. В работе рассматривается модель прогнозирования курса валюты. В модели копула-функции применяется для описания рискованной части, то есть случайной составляющей.

В качестве эконометрической модели в подходе выбирается модель AR(1)-GARCH(1,1). Модель задается в виде двух регрессионных уравнений, одно из которых (AR(1)) описывает поведение непосредственно курса валюты, а второе его условной дисперсии (GARCH(1,1)) [2].

$$p_t = \beta_0 + \beta_1 p_{t-1} + e_t \quad (95)$$

$$h_t = \gamma_0 + \gamma_1 e_{t-1}^2 + \gamma_2 h_{t-1}. \quad (96)$$

Рассмотрим 4 мировых валюты: канадский доллар (CAD), евро (EUR), британские фунты-стерлингов (GBP) и американский доллар (USD). Загрузим котировки за период с 15 апреля 2019 года по 15 апреля 2021 года.

Таблица 16. Курсы мировых валют

| CAD | EUR | GBP | USD |
|------------|---------|----------|---------|
| 48,0649973 | 72,6611 | 84,17321 | 64,2838 |
| 47,9695834 | 72,2771 | 83,59278 | 64,0705 |
| 47,8209481 | 72,0477 | 83,18739 | 63,8038 |
| 47,682504 | 71,7602 | 82,91138 | 63,832 |
| 47,7558116 | 71,962 | 83,12954 | 63,9901 |
| 47,8053206 | 72,042 | 83,20149 | 64,0307 |
| 47,7622418 | 71,972 | 83,12652 | 63,973 |
| 47,7890462 | 71,796 | 82,82371 | 63,8038 |
| 47,4265524 | 71,46 | 82,38182 | 63,694 |
| 47,8327765 | 71,8796 | 83,2989 | 64,5778 |

| | | | |
|------------|---------|----------|---------|
| 47,854701 | 71,878 | 83,24549 | 64,5464 |
| 48,0568685 | 72,2723 | 83,52711 | 64,6795 |
| 48,100472 | 72,346 | 83,63357 | 64,7469 |
| 48,255669 | 72,4292 | 83,98111 | 64,9908 |

Далее, отделим последние 30 значений выборки для проверки прогноза. Оценим модель AR(1)-GARCH(1,1) на тренировочной части выборки и получим значения остатков модели. Полученные остатки не являются стандартизированными, поэтому, поделив на значение среднеквадратического отклонения, получим нормированные оценки [2].

$$z_t = \frac{e_t}{\sqrt{h_t}} \quad (97)$$

Для ежедневных значений случайных ошибок курса валют оценим эмпирическую функцию распределения. По принципу модели, рассмотренной ранее, выберем копула-функцию, описывающую данные наилучшим образом.

Таблица 17. Значения логарифма правдоподобия для копула-функций

| Копула | Нормальная | Стьюдента | Гумбеля | Клейтона |
|-------------------|------------|-----------|----------|----------|
| Значение $\ln(L)$ | 1118,257 | 1293,409 | 1168,099 | 909,7846 |

Как и в предыдущем случае, наилучшей оказалась модель копулы, основанной на распределении Стьюдента, что было предсказуемо. Курсы валют, также как и курсы акций имеют тяжелые хвосты, более того, распределение валютных котировок является более непредсказуемым и валютные вложения более рисковыми.

Выбрав наилучшее многомерное распределение, необходимо сгенерировать значения функций распределений остатков. Для перехода от функции распределения к самим случайным остаткам модели, необходимо найти квантили распределения в точках, соответствующих функциям распределения. Для этого необходимо оценить параметры маргинальных распределений для каждой валюты.

Для отбора подходящих значений остатков используем модель скользящего среднего. Для случайных остатков каждой из валют построим скользящее среднее с интервалом сглаживания 2. Для выбора необходимого набора сгенерированных остатков наложим условие для каждой из валют: модуль отклонения сгенерированных остатков от оценок скользящего среднего не должен отличаться более, чем 0,5. Повторив итерацию 30 раз, получим месячные оценки случайных остатков. Просуммировав с прогнозом по модели авторегрессии (AR(1)-GARCH(1,1)), получим месячные прогнозы курсов валют.

Сравним значения прогнозов с фактическими значениями. Для этого построим графики фактического и прогнозного рядов. Рассмотрим канадский доллар.

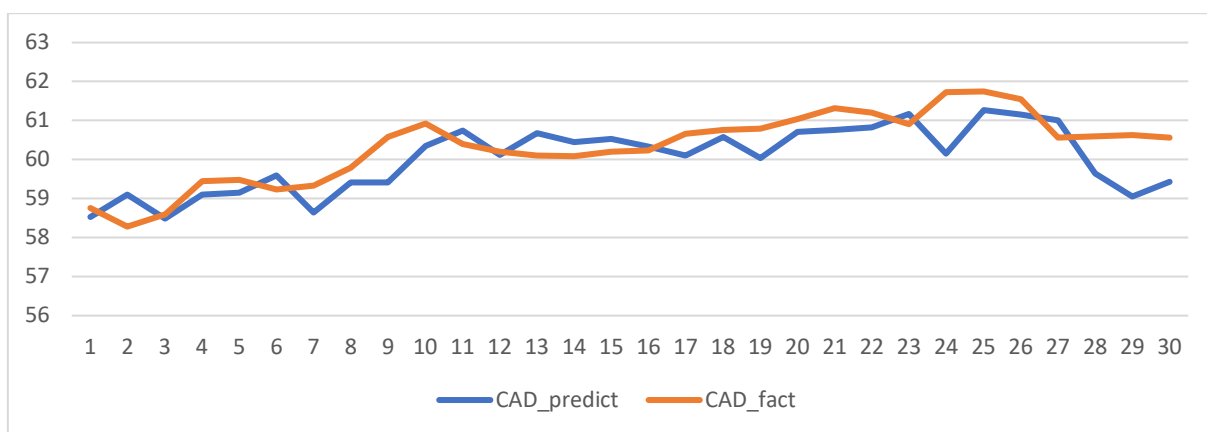


Рисунок 12. Фактические и прогнозные значения курса канадского доллара

Анализ рисунка 12 показывает, что полученные значения прогноза неплохо описывают тенденцию поведения курса валюты, особенно на начальных периодах прогноза. Некоторые резкие изменения курса модель определяет с небольшим лагом, это может быть связано с большим временным периодом – день.

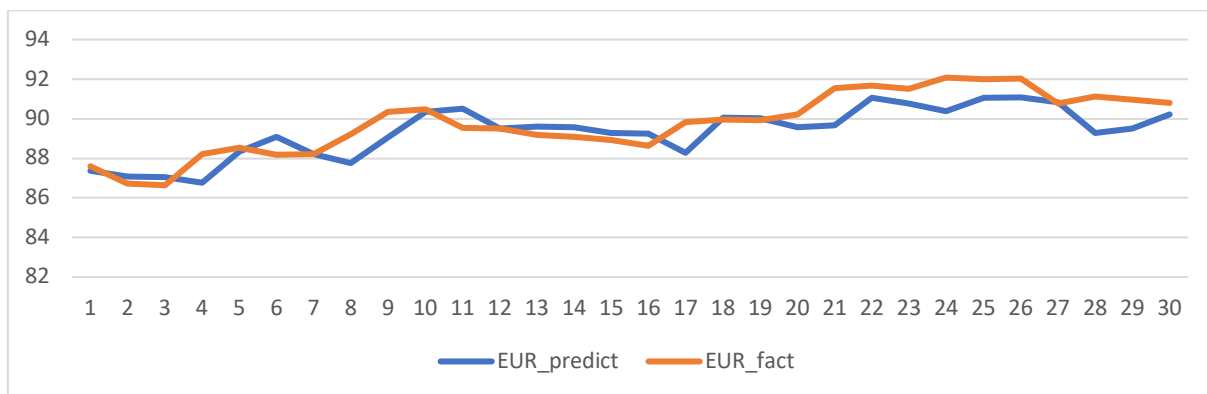


Рисунок 13. Фактические и прогнозные значения курса евро

Аналогичный результат модель показала и на курсе евро. Модель также хорошо описывает общую тенденцию поведения валюты и на некоторых интервалах дает неплохой прогноз, однако для более далеких прогнозных дней, модель перестает давать хорошие результаты.

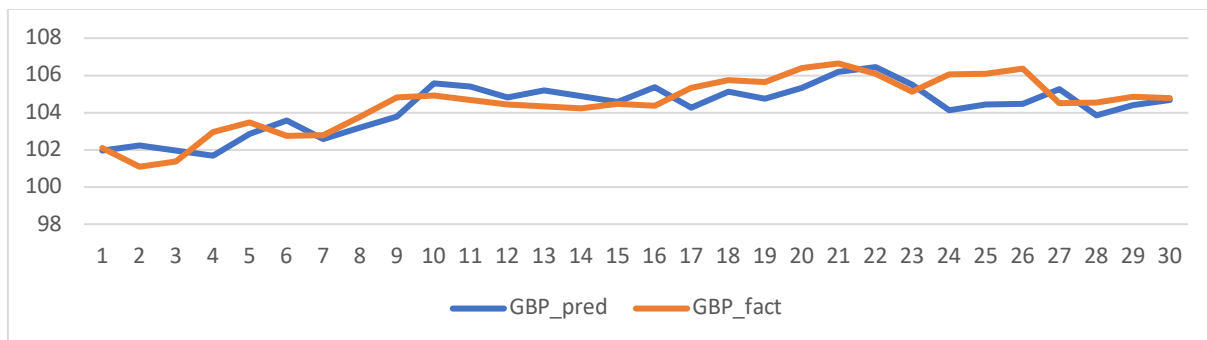


Рисунок 14. Фактические и прогнозные значения курса британского фунта стерлингов

Результат моделирования для британских фунтов стерлингов представлен на рисунке 14. Модель дала весьма неплохие результаты, полученные прогнозы показывают поведение валюты.

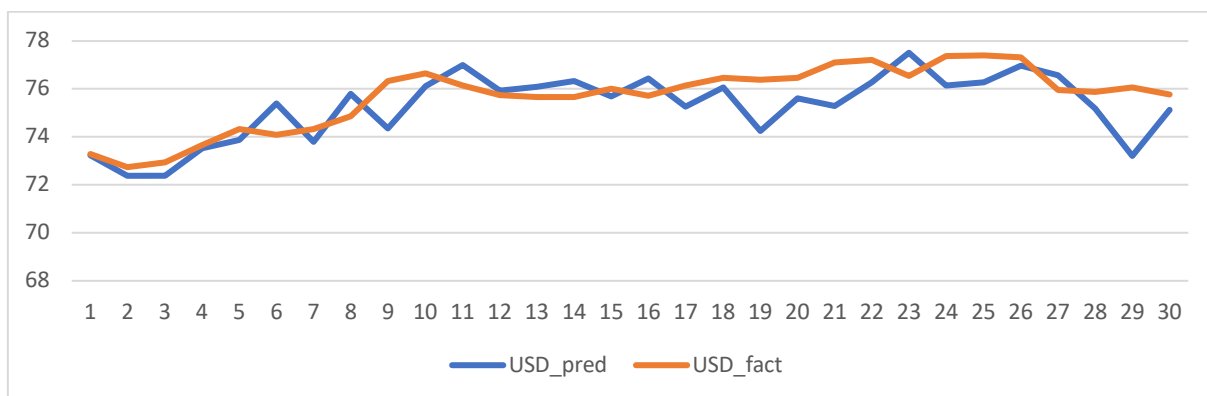


Рисунок 15. Фактические и прогнозные значения курса американского доллара

Наименее точный прогноз по модели получился у курса американского доллара. В прогнозах можно заметить резкие скачки цен. Такая ошибка прогноза может быть объяснена тем, что доллар является ключевой валютой в мировой экономике. На его курс влияет очень много различных факторов, невозможных отразить ни в одной математической модели.

Таким образом, мы изучили еще одно направление применения копула-функций. На основе многомерной связи между случайными величинами в модели генерируются остатки и, учитывая некоторые условия, они добавляются к прогнозу по модели. Как видно из анализа последних рисунков, модель хорошо подходит для рассмотрения тенденции поведения валют, а также для не долгосрочного прогнозирования. Также стоит отметить, что обе модели, рассмотренные в работе, предназначены для работы с котировками. Первая модель может быть использована для формирования портфеля валют, а вторая для прогнозирования курса акций. А использование этих моделей в совокупности может принести еще лучшие результаты.

ЗАКЛЮЧЕНИЕ

Исследование было посвящено анализу многомерных экономико-математических моделей, основанных на копула-функциях. В ходе выполнения работы принимались во внимание уже существующие труды, посвященные многомерному анализу. Было выяснено, что копула-функции широко используются в различных направлениях: страхование, кредитование, портфельный анализ. На копулах основываются множество эконометрических, актуарных моделей, а также моделей машинного обучения. Также стоит отметить, что большинство исследователей сходятся во мнении о том, что применение копула-функций при анализе рисков приводят к хорошим результатам за счет того, что копулы хорошо описывают данные на тяжелых хвостах.

В ходе исследования была сформирована теоретическая база для моделирования исследуемых данных. Были исследованы параметрические и непараметрические статистические методы оценки копула функций

На основе проведенного анализа теоретического материала по многомерным копулам и примеров моделирования с помощью копул были определены условия применимости копула функций для фондового рынка.

В практической части было проведено моделирование портфелей ценных бумаг, основанного на принципах минимизации риска и максимального отношения доходности к риску. В ходе проведенного анализа результатов моделирования, с целью выбора наилучшего критерия и временного горизонта, были рассчитаны значения доходностей. При использовании подхода минимизации риска было установлено, что портфель получался более диверсифицированным, однако доходность такого подхода получалась невысокой. Напротив, при максимизации отношения доходности к риску портфель получался менее диверсифицированными, а доходность более высокой. Например, средняя доходность по всем горизонтам инвестирования для принципа

минимизации риска составила 2,54%, а для принципа максимизации отношения доходности к риску 4,04%.

Также в работе была построена прогнозная модель для курса мировых валют, основанная на эконометрической модели временного ряда. Эта модель может быть использована как для валютного, так и для фондового рынка. Таким образом, мы делаем вывод, что для получения наиболее качественных результатов могут использоваться обе модели в совокупности.

Для получения вышеописанных результатов было разработано программное обеспечение для генерации портфелей и выбора оптимального из них, а также для построения прогноза курса валюты. Программы были написаны на языке R в программной среде RStudio. Программные коды для модели оптимального портфеля и курса валюты представлены в приложениях 2 и 3 соответственно.

Таким образом, в результате выполнения поставленных задач цель исследования достигнута.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИК

1. Айвазян С.А. (2010). Методы эконометрики. М.: Магистр.
2. Алексеев В.В., Шоколов В.В., Соложенцев Е.Д. (2006). Логико-вероятностное моделирование портфеля ценных бумаг с использованием копул. Управление финансовыми рисками. № 3. С. 272–283.
3. Антонов И.Н. (2016). Прогнозирование курсов валют на основе копулярных моделей. Бизнес. Образование. Право. Вестник Волгоградского института бизнеса, No 1 (34), С. 158-164.
4. Благовещенский Ю. Н. (2012). Основные элементы теории копул. Прикладная эконометрика, 26 (2), 113–130.
5. Крицкий О. Л., Ульянова М. К. (2007). Определение многомерного финансового риска портфеля акций. Прикладная эконометрика, 8 (4), 3–17.
6. Маркина И. А, Переверзев Е. В. (2013). Оценка рисков при внутрикорпоративных расчетах. Бизнес. Образование. Право. Вестник Волгоградского института бизнеса. No 4 (25). С. 38–41.
7. Пеникас Г. (2010). И. Модели «копула» в управлении валютным риском банка. Прикладная эконометрика. No 8. С. 62–87.
8. Фантаццини Д. (2008). Эконометрический анализ финансовых данных в задачах управления риском. Прикладная эконометрика. No 2 (10), С. 91 – 137.
9. Фантаццини Д. (2011a). Моделирование многомерных распределений с использованием копула - функций. I. Прикладная эконометрика, 22 (2), 98 – 134.
10. Фантаццини Д. (2011b). Моделирование многомерных распределений с использованием копула - функций. II. Прикладная эконометрика, 23 (3), 98 – 132.

11. Фантазини Д. (2011с). Моделирование многомерных распределений с использованием копула - функций. III. Прикладная эконометрика, 25 (4), 100 – 130.
12. Aas K., Czado C., Frigessi A., Bakken H. (2009). Pair-copula constructions of multiple dependence. In Insurance: Mathematics and Economics, 44 (2), 182 – 198.
13. Ang A., Chen J. (2002). Asymmetric correlations of equity portfolios. Journal of Financial Economics, 63 (3), 443 – 494.
14. Clayton D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika, 65, 141 – 151.
15. Fang K., Kotz S., Ng K. (1987). Symmetric multivariate and related distributions. London: Chapman Hall.
16. Fantazzini D. (2006). Dynamic copula modeling for Value at Risk. Frontiers in Finance and Economics, 2 (5), 1–38.
17. Frank M. J. (1979). On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. Aequationes Mathematicae, 19, 194 – 226.
18. Gumbel E. J. (1960). Bivariate exponential distributions. Journal of the American Statistical Association, 55, 698 – 707.
19. Hennessy D., Lapan H. (2002). The Use of Archimedean Copulas to Model Portfolio Allocations. Mathematical Finance. No 12. P. 143–154.
20. Markowitz H. (1952). Portfolio selection. Journal of Finance. No 7. P. 77–91.
21. McNeil A. (2008). Sampling nested Archimedean copulas. Journal of Statistical Computation and Simulation, 78 (6), 567 – 581.
22. McNeil A., Frey R., Embrechts P. (2005). Quantitative risk management: Concepts, techniques and tools. New Jersey: Princeton Series in Finance.
23. McNeil A., Nešlehová J. (2009). Multivariate Archimedean copulas, d-monotone functions and L1-norm symmetric distributions. Annals of Statistics, 37 (5), 3059 – 3097.

24. Nelsen R. (1999). An introduction to copulas. Lecture Notes in Statistics. New York: Springer-Verlag.
25. Nelsen R. B. (2005). Some properties of Schur-constant survival models and their copulas. Brazilian Journal of Probability and Statistics, 19, 179 – 190.
26. Nelsen R. B. (2006). An introduction to copulas. Lecture Notes in Statistics, 2nd Edition. New York: Springer-Verlag.
27. Patton A. (2012). A review of copula models for economic time series. Journal of Multivariate Analysis, 110, 4–18.
28. Sklar A. (1959). Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statis. Univ. Paris, 8, 229 – 231.
29. Sklar A. (1996). Random variables, distribution functions, and copulas: Personal look backward and forward. Lecture notes. Monograph series, 28, 1 – 14.
30. Yau S, Kwon RH, Rogers JS, Wu D. (2011). Financial and operational decisions in the electricity sector: contract portfolio optimization with the conditional value-at-risk criterion. International Journal of Production Economics; 134: 67–77.

ПРИЛОЖЕНИЕ 1

Программный код на языке R для иллюстраций в главах 1 и 2

```
library (fCopulae)
library(scatterplot3d)
library(copula)
# Generate Grid:
N = 50; uv = grid2d (x = (0:N)/N);
u = uv$x; v = uv$y
# Compute Frechet and Product Copulae:
W = matrix (apply (cbind (u+v-1, 0), 1, max), ncol =
N+1)
Pi = matrix (u*v, ncol = N+1)
M = matrix (apply (cbind (u, v), 1, min), ncol = N+1)
persp(M, col='green', xlab = 'y1', ylab='y2', main =
'min (y1,y2)')
persp(W, col='lightgreen', xlab = 'y1', ylab='y2',
main = 'max (u + v - 1,0)')
persp(Pi, col='green', xlab = 'y1', ylab='y2', main =
'y1*y2')
contour (M, xlab = "y1", ylab = "y2", main = "Линии
уровня M(y1,y2)")
contour (W, xlab = "y1", ylab = "y2", main = "Линии
уровня W(y1,y2)")
contour (Pi, xlab = "y1", ylab = "y2", main = "Линии
уровня P(y1,y2)")
# Compute the Normal Copula Density:
c1= dellipticalCopula (u, v, rho = 0.01, type =
"norm", output = "list")
c2= dellipticalCopula (u, v, rho = 0.5, type =
"norm", output = "list")
```

```

c3= dellipticalCopula (u, v, rho = 0.99, type =
"norm", output = "list")
# Create a Perspective Plot:
persp (c1,xlab = 'y1', ylab='y2', zlab='C', col =
"steelblue", main = "Копула нормального распределения
(rho=0.01)", cex = 0.5)
persp (c2,xlab = 'y1', ylab='y2', zlab='C', col =
"steelblue", main = "Копула нормального распределения
(rho=0.5)", cex = 0.5)
persp (c3,xlab = 'y1', ylab='y2', zlab='C', col =
"steelblue", main = "Копула нормального распределения
(rho=0.99)", cex = 0.5)
contour (c1, nlevels = 20, main = "Линии уровня
нормальной копулы (rho=0,01)", cex = 0.5)
contour (c2, nlevels = 20, main = "Линии уровня
нормальной копулы (rho=0,05)", cex = 0.5)
contour (c3, nlevels = 20, main = "Линии уровня
нормальной копулы (rho=0,99)", cex = 0.5)

```

ПРИЛОЖЕНИЕ 2

Программный код на языке R для модели оптимального портфеля

```
library(readxl)
data=read_excel("~/Desktop/данные.xlsx", sheet =
"Лист2")
data=data*100 #переход к %
num_of_shares=ncol(data) #количество цб
train_data=data[1:474,] #тренировочное множество
test_data=data[475:504,] #тестовое множество
names_shares=c("Сбербанк", "Русгидро", "Роснефть", "Алроса",
"Аэрофлот")

#Проверка гипотезы о нормальном распределении
library(stats)
alpha=0.05 #уровень значимости
pv=matrix(nrow=1,ncol=num_of_shares)
for (i in 1:num_of_shares) {

kolmogorov=ks.test(data[,i],"pnorm",mean=mean(data[,i
]),sd=sd(data[,i]))
  pv[i]=kolmogorov$p.value #добавляем значения pvalue в
таблицу
}

#Подбор ядерного распределения
library('np')
x=seq(-5,5,length=10000)
dens=matrix(ncol=num_of_shares,nrow=10000)
old_params=par(mfrow=c(2,5))
for (i in 1:num_of_shares) {
```

```

    y_dat=data[,i]
    y_plot=npudens(tdat=y_dat, edat=x,
ckertype="gaussian",bwtype="fixed")
    dens[,i]=y_plot$dens
    plot(x, dens[,i],
type="l",main=names_shares[i],xlab="Доходность",ylab="
Плотность")
    lines(x,dnorm(x,mean=mean(data[,i]),
sd=sd(data[,i])), lty="dashed", col="red")
}

library (copula)
cdf_func=pobs(train_data) #эмпирическая функция
распределения
#Создание копул
norm_copula=normalCopula(dim=num_of_shares,dispstr="e
x") #Копула Гаусса
t_copula=tCopula(dim=num_of_shares,dispstr="ex")
#Копула Стьюдента
gumbel_copula=gumbelCopula(dim=num_of_shares,param=2)
#Копула Гумбеля
clayton_copula=claytonCopula(dim=num_of_shares,param=
2) #Копула Клейтона
#Подгонка
norm_fit=fitCopula(cdf_func,copula=norm_copula)
t_fit=fitCopula(cdf_func,copula=t_copula)
gumbel_fit=fitCopula(cdf_func,copula=gumbel_copula)
clayton_fit=fitCopula(cdf_func,copula=clayton_copula)
#Выбор оптимальной копулы
norm_fit@loglik

```

```

t_fit@loglik
gumbel_fit@loglik
clayton_fit@loglik

#Генерация портфелей
M=10000
X=matrix(ncol=num_of_shares,nrow=M)
for (k in 1:M) {
  p=c(0,1,runif(n=num_of_shares-1,min=0,max=1))
  q=t(matrix(sort(p)))
  n=ncol(q)
  w=c(rep(0,times=n-1))
  for (j in 1:n-1) {
    w[j]=q[j+1]-q[j]
  }
  X[k,]=w
}

```

```

#Моделирование доходностей с многомерным распределением на
основе лучшей копулы
Z=1000 #количество симуляций
r.week=matrix(nrow=Z,ncol=M)
r.halfmonth=matrix(nrow=Z,ncol=M)
r.month=matrix(nrow=Z,ncol=M)
r_c.week=matrix(nrow=Z,ncol=num_of_shares)
r_c.halfmonth=matrix(nrow=Z,ncol=num_of_shares)
r_c.month=matrix(nrow=Z,ncol=num_of_shares)
for (l in 1:Z) {
  D=30
  copula_gen=rCopula(n=D,copula=t_fit@copula)
  r_matrix=matrix(nrow=D,ncol=num_of_shares)

```

```

r_matrix_c.week=matrix(nrow=1,ncol=num_of_shares)

r_matrix_c.halfmonth=matrix(nrow=1,ncol=num_of_shares
)
r_matrix_c.month=matrix(nrow=1,ncol=num_of_shares)
for (j in 1:num_of_shares) {
  fact_cdf=ecdf(train_data[,j])
  r_matrix[,j]=quantile(fact_cdf,t(copula_gen[,j]))
  r_matrix_c.week[j]=(prod(1+r_matrix[1:7,j]/100)-
1)*100

r_matrix_c.halfmonth[j]=(prod(1+r_matrix[1:15,j]/100)
-1)*100

r_matrix_c.month[j]=(prod(1+r_matrix[1:30,j]/100)-
1)*100
}
#Считаем CVaR и ожидаемую доходность
r_portfolio=matrix(ncol=M,nrow=D)
for (i in 1:M) {
  ZZ=0
  for (j in 1:num_of_shares){
    ZZ=ZZ+r_matrix[,j]*X[i,j]
    r_portfolio[,i]=ZZ
  }
}
r_portfolio_week=matrix(ncol=M,nrow=1)
r_portfolio_halfmonth=matrix(ncol=M,nrow=1)
r_portfolio_month=matrix(ncol=M,nrow=1)
for (i in 1:M) {

```



```
r_portfolio_week[,i]=(prod((r_portfolio[1:7,i]/100)+1)-1)*100
```

```
r_portfolio_halfmonth[,i]=(prod((r_portfolio[1:15,i]/100)+1)-1)*100
```

```
r_portfolio_month[,i]=(prod((r_portfolio[1:30,i]/100)+1)-1)*100
```

```

    }
    r.week[l,]=r_portfolio_week
    r.halfmonth[l,]=r_portfolio_halfmonth
    r.month[l,]=r_portfolio_month
    r_c.week[l,]=r_matrix_c.week
    r_c.halfmonth[l,]=r_matrix_c.halfmonth
    r_c.month[l,]=r_matrix_c.month
}

```

```

mean_r_week=matrix(ncol=M,nrow=1)
mean_r_halfmonth=matrix(ncol=M,nrow=1)
mean_r_month=matrix(ncol=M,nrow=1)
CVaR_week=matrix(ncol=M,nrow=1)
CVaR_halfmonth=matrix(ncol=M,nrow=1)
CVaR_month=matrix(ncol=M,nrow=1)
VaR_week=matrix(ncol=M,nrow=1)
VaR_halfmonth=matrix(ncol=M,nrow=1)
VaR_month=matrix(ncol=M,nrow=1)
gamma_=0.95
VaR_bound=(1-gamma_)*D
for (i in 1:M) {
    mean_r_week[,i]=mean(r.week[,i])

```

```

mean_r_halfmonth[,i]=mean(r_halfmonth[,i])
mean_r_month[,i]=mean(r_month[,i])
sort_r_week=sort(r_week[,i]*(-1),decreasing=TRUE)
sort_r_halfmonth=sort(r_halfmonth[,i]*(-
1),decreasing=TRUE)
sort_r_month=sort(r_month[,i]*(-1),decreasing=TRUE)
VaR_week[i]=sort_r_week[VaR_bound]
VaR_halfmonth[i]=sort_r_halfmonth[VaR_bound]
VaR_month[i]=sort_r_month[VaR_bound]
CVaR_week[i]=mean(sort_r_week[1:VaR_bound])

CVaR_halfmonth[i]=mean(sort_r_halfmonth[1:VaR_bound])
#значения квартального CVAR
CVaR_month[i]=mean(sort_r_month[1:VaR_bound])
}
a.par=par(mfrow=c(1,3))
plot(CVaR_week, mean_r_week,
xlab="CVaR",ylab="Ожидаемая доходность",
main="Достижимое множество недельные данные")
plot(CVaR_halfmonth, mean_r_halfmonth,
xlab="CVaR",ylab="Ожидаемая доходность",
main="Достижимое множество полумесячные данные")
plot(CVaR_month, mean_r_month,
xlab="CVaR",ylab="Ожидаемая доходность",
main="Достижимое множество месячные данные")

#Поиск оптимального портфеля по максимуму r/CVaR
coeff_week=mean_r_week/CVaR_week
max_coeff_week=order(coeff_week,decreasing = TRUE)
max_coeff_week_portfolio=X[max_coeff_week[1],]

```

```

char_max_coeff_week=c(mean_r_week[max_coeff_week[1]],
CVaR_week [max_coeff_week[1]],
VaR_week[max_coeff_week[1]],
coeff_week[max_coeff_week[1]])
names(char_max_coeff_week)=c("Доходность", "CVaR ", "VaR
", "Коэффициент")

```

```

coeff_halfmonth=mean_r_halfmonth/CVaR_halfmonth
max_coeff_halfmonth=order(coeff_halfmonth,decreasing
= TRUE)
max_coeff_halfmonth_portfolio=X[max_coeff_halfmonth[1
],]
char_max_coeff_halfmonth=c(mean_r_halfmonth[max_coeff
_halfmonth[1]],
CVaR_halfmonth[max_coeff_halfmonth[1]],
VaR_halfmonth[max_coeff_halfmonth[1]],
coeff_halfmonth[max_coeff_halfmonth[1]])
names(char_max_coeff_halfmonth)=c("Доходность", "CVaR
", "VaR ", "Коэффициент")

```

```

coeff_month=mean_r_month/CVaR_month
max_coeff_month=order(coeff_month,decreasing = TRUE)
max_coeff_month_portfolio=X[max_coeff_month[1],]
char_max_coeff_month=c(mean_r_month[max_coeff_month[1
]], CVaR_month[max_coeff_month[1]],
VaR_month[max_coeff_month[1]],
coeff_month[max_coeff_month[1]])
names(char_max_coeff_month)=c("Доходность", "CVaR
", "VaR ", "Коэффициент")

```

```

# Портфель минимального риска по CVaR

```

```

min_CVaR_week=order(CVaR_week, decreasing = FALSE)
min_CVaR_week_portfolio=X[min_CVaR_week[1],]
char_min_CVaR_week=c(mean_r_week[min_CVaR_week[1]],CV
aR_week[min_CVaR_week[1]],VaR_week[min_CVaR_week[1]],
coeff_week[min_CVaR_week[1]])
names(char_min_CVaR_week)=c("Доходность", "CVaR ", "VaR
", "Коэффициент")

```

```

min_CVaR_halfmonth=order(CVaR_halfmonth, decreasing =
FALSE)
min_CVaR_halfmonth_portfolio=X[min_CVaR_halfmonth[1],
]
char_min_CVaR_halfmonth=c(mean_r_halfmonth[min_CVaR_h
alfmonth[1]],CVaR_halfmonth[min_CVaR_halfmonth[1]],Va
R_halfmonth[min_CVaR_halfmonth[1]],
coeff_halfmonth[min_CVaR_halfmonth[1]])
names(char_min_CVaR_halfmonth)=c("Доходность", "CVaR
", "VaR ", "Коэффициент")

```

```

min_CVaR_month=order(CVaR_month, decreasing = FALSE)
min_CVaR_month_portfolio=X[min_CVaR_month[1],]
char_min_CVaR_month=c(mean_r_month[min_CVaR_month[1]]
,CVaR_month[min_CVaR_month[1]],VaR_month[min_CVaR_mon
th[1]], coeff_month[min_CVaR_month[1]])
names(char_min_CVaR_month)=c("Доходность", "CVaR ", "VaR
", "Коэффициент")

```

ПРИЛОЖЕНИЕ 3

Программный код на языке R для модели прогнозирования курса валюты

```
library(fBasics)
library(fGarch)
library(copula)
library(MASS)
library(tseries)
data1=data[1:611,]
data2=data[612:641,]
cur1_fit=garchFit(~arma(1,0)+garch(1,1),
data=data1$CAD, cond.dist = "std",trace = F)
cur2_fit=garchFit(~arma(1,0)+garch(1,1),
data=data1$EUR, cond.dist = "std",trace = F)
cur3_fit=garchFit(~arma(1,0)+garch(1,1),
data=data1$GBP, cond.dist = "std",trace = F)
cur4_fit=garchFit(~arma(1,0)+garch(1,1),
data=data1$USD, cond.dist = "std",trace = F)

cur1_res=cur1_fit@fit$series$z
cur2_res=cur2_fit@fit$series$z
cur3_res=cur3_fit@fit$series$z
cur4_res=cur4_fit@fit$series$z

cur_result=data.frame(cur1_res)
cur_result$'EUR'=data.frame(cur2_res)
cur_result$'GBP'=data.frame(cur3_res)
cur_result$'USD'=data.frame(cur4_res)

cdf_func=pobs(cur_result)
```

```

norm_copula=normalCopula(dim=4,dispstr="ex") #Копула
Гаусса
t_copula=tCopula(dim=4,dispstr="ex") #Копула Стьюдента
gumbel_copula=gumbelCopula(dim=4,param=2) #Копула
Гумбеля
clayton_copula=claytonCopula(dim=4,param=2) #Копула
Клейтона
#Подгонка
norm_fit=fitCopula(cdf_func,copula=norm_copula)
t_fit=fitCopula(cdf_func,copula=t_copula)
gumbel_fit=fitCopula(cdf_func,copula=gumbel_copula)
clayton_fit=fitCopula(cdf_func,copula=clayton_copula)
#Выбор оптимальной копулы
norm_fit@loglik
t_fit@loglik
gumbel_fit@loglik
clayton_fit@loglik

tfit_cad=fitdistr(cur_result$curl_res, "t")
tfit_eur=fitdistr(ts(cur_result$EUR), "t")
tfit_gbp=fitdistr(ts(cur_result$GBP), "t")
tfit_usd=fitdistr(ts(cur_result$USD), "t")

for (i in 1:1000){
  copula_gen=rCopula(n=1,copula=t_fit@copula)
  znach_cad=qt(copula_gen[,1],
df=tfit_cad$estimate[3])
  znach_eur=qt(copula_gen[,2],
df=tfit_eur$estimate[3])

```

```

    znach_gbp=qt(copula_gen[,3],
df=tfit_gbp$estimate[3])
    znach_usd=qt(copula_gen[,4],
df=tfit_usd$estimate[3])
    if (abs(znach_cad-mean(curl1_res[(length(curl1_res)-
1):length(curl1_res)]))<0.5
        & abs(znach_eur-
mean(cur2_res[(length(cur2_res)-
1):length(cur2_res)]))<0.5
        & abs(znach_gbp-
mean(cur3_res[(length(cur3_res)-
1):length(cur3_res)]))<0.5
        & abs(znach_eur-
mean(cur4_res[(length(cur4_res)-
1):length(cur4_res)]))<0.5){
        curl1_res[length(curl1_res)+1]=znach_cad
        cur2_res[length(cur2_res)+1]=znach_eur
        cur3_res[length(cur3_res)+1]=znach_gbp
        cur4_res[length(cur4_res)+1]=znach_usd
    }
}

```