

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра теории функций и стохастического анализа

**РОЛЬ И МЕТОДЫ ВЫБОРА ТАРИФНЫХ ФАКТОРОВ В
АВТОСТРАХОВАНИИ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

Студентки 4 курса 412 группы
направления 01.03.02 — Прикладная математика и информатика

механико-математического факультета
Егоровой Екатерины Вячеславовны

Научный руководитель

доцент, к. ф.-м. н., доцент

Л. В. Борисова

Заведующий кафедрой

д. ф.-м. н., доцент

С. П. Сидоров

Саратов 2023

Введение. В странах с рыночной экономикой физические и юридические лица получают определенный комплекс гарантий: по поводу возмещения ущерба, получения в определенных случаях некоторой денежной суммы и т.п. Такие гарантии предоставляются и обеспечиваются страхованием. На его основе становится возможна защита общественных и личных интересов, возникающих во всех сферах экономики.

Основная задача страхования состоит именно в защите имущественных интересов как юридических, так и физических лиц. Объективная необходимость страхования обуславливается существованием понятия риск как случайного события, которое может привести к ущербу. Последний может быть не только имущественным, связанным с убытками в материальной сфере, но и физическим вследствие нанесения ущерба людям, например в результате несчастного случая.

Несмотря на то, что в России страхование находится лишь на этапе своего развития, возникло оно достаточно давно и с тех пор развивалось. Существуют различные виды страхования: гражданской ответственности, недвижимости, страхование здоровья, жизни, социальное, медицинское, автомобильное, авиационное, космическое и многие другие. В данной работе рассматривается автострахование (ОСАГО и КАСКО).

Страхованием автомобилей и их владельцев занимается большинство страховых компаний России. Тем не менее большинство людей плохо осведомлены в теме тарифных факторов и их влияния на ставки автострахования. Основа страхового дела — это правильные актуарные расчеты страховых премий, ведь именно страховые взносы формируют базу для страховых выплат, страховые резервы. Задача любого страховщика — это сделать клиенту максимально выгодное предложение, дать ему возможность максимально сэкономить, но в то же время не занижить страховые ставки из-за угрозы обанкротиться.

Это свидетельствует об **актуальности** темы выпускной квалификационной работы, а также определяет ее **практическую значимость**.

Целью бакалаврской работы является анализ статистики выплат по автострахованию с помощью некоторых статистических методов для построения тарифных множителей. Объект исследования — рынок автострахования

в России.

Предметом исследования данной работы является совокупность показателей, характеризующих убыточность портфеля договоров автострахования и факторы, определяющие тарифообразование добровольного автострахования. В соответствии с целью поставлены и решены следующие **задачи**:

- проанализировать развитие и современное состояние рынка страхования ОСАГО и КАСКО в России;
- изучить методы построения классов значений факторов риска, а также методы отбора самих тарифных факторов (наиболее важных факторов риска, на основе которых строится страховой тариф);
- построить тарифных множителей для договоров добровольного автострахования, в зависимости от региона использования транспортного средства.

В первом разделе работы проведен экономико-статистический анализ рынка автострахования в России, рассмотрены основные тенденции и динамики рынка, такие как страховые премии, выплаты, а также количество страховых компаний на рынке. Определены основные тарифные факторы автомобильного страхования. Рассмотрены сходства и различия обязательного и добровольного автострахования, а так же проанализировано развитие и современное состояние рынка страхования ОСАГО и КАСКО в России.

Второй раздел представляет собой обзор методов построения классов значений факторов риска, эти методы необходимо применять для уменьшения количества классов значений с целью облегчения дальнейшего анализа этих факторов, построения множителей для выбранных факторов и т.д.

В третьем разделе представлены методы отбора тарифных факторов с помощью дисперсионного анализа, метода отношения правдоподобий и перекрёстной параметризации при наличии наблюдений за один год.

Четвертый раздел посвящен построению тарифных множителей для добровольного автомобильного страхования в зависимости от региона использования транспортного средства. Построение множителя произведено на основе данных о страховых выплатах по добровольному страхованию и количестве заключённых договоров в разрезе на субъекты РФ. Для анализа данных была реализована программное обеспечение на языке программиро-

вания Python, реализующее агломеративный кластер-метод с максимизацией функции правдоподобия.

Информационной базой исследования стала официальная статистика центрального банка Российской Федерации по страховым выплатам в сегменте добровольного автострахования во всех субъектах РФ.

Основное содержание работы. В первом разделе «**Экономическая теория и статистический анализ рынка автомобильного страхования в России**» рассматриваются такие понятия как тарифные факторы, тарифные группы, тарифная политика страховщиков, её принципы и особенности построения тарифов, приводятся примеры тарифных факторов автострахования. Рассмотрены основные тенденции и динамики рынков обязательного и добровольного автострахования, такие, как страховые премии, выплаты, а также количество страховых компаний на рынке.

Все объекты, подлежащие страхованию, делятся на однородные с точки зрения риска категории — тарифные группы.

Под тарифной политикой понимается целенаправленная деятельность страховой организации по разработке, установлению, уточнению и упорядочению страховых тарифов. Цель тарифной политики — успешное и безубыточное развитие страховой организации.

Тарифные факторы — набор независимых наиболее значимых факторов, которые в дальнейшем будут учитываться при расчете тарифов.

К основным тарифным факторам в автостраховании относятся:

- Тип (категория) и назначение транспортного средства;
- Территория преимущественного использования транспортного средства;
- Количество лиц, допущенных к управлению транспортным средством
- Возраст и стаж водителя, допущенного к управлению транспортным средством;
- Мощность двигателя.

Основные различия ОСАГО и КАСКО.

- Полис ОСАГО защищает автогражданскую ответственность, а КАСКО — автомобиль;
- Стоимость полиса ОСАГО регулирует ЦБ, стоимость КАСКО зависит от условий договора страховой компании;

- Условия по ОСАГО всегда одинаковые, КАСКО может продаваться с разными наборами рисков и опций;
- ОСАГО нужен каждому водителю по закону, КАСКО — по желанию;
- Чтобы застраховаться от любых неожиданных расходов, лучше иметь оба полиса.

По итогам 2022 года общие сборы Российских страховщиков в сегменте КАСКО увеличились на 7,6% - до 222,8 млрд рублей, выплаты выросли на 11,2% - до 121,5 млрд рублей по сравнению с 2021 годом, следует из статистики Банка России. Средняя выплата составила 126,6 тыс. рублей, увеличившись на 14,3%. Сборы страхования КАСКО во II квартале 2022 года выросли на 6,4%, до 52 млрд руб., несмотря на спад продаж новых автомобилей на 81%. Заметный рост выплат сказался на величине коэффициента выплат, который вернулся к высокому допандемийному уровню (0,63). Квартальные темпы роста сборов премий по КАСКО снижались, но оставались в положительной области, т. е. рынок рос, но все медленнее от квартала к кварталу: 15,7% (I) и 6,4% (II). Количество проданных полисов сократилось на 8% и закрепилось на отметке 1,3 млн ед. Страхование каско физических лиц составляло 11% в сборах премий рынка конечных потребителей при величине сборов премий по виду, равной 29 млрд руб. Доля страхования каско юридических лиц достигала 15% портфеля рынка корпоративных клиентов, объем рынка по этому виду оценивался в 21 млрд руб. Объем рынка страхования каско физических лиц во II квартале был больше аналогичного для корпоративных клиентов в 1,4 раза в денежном измерении.

Средний размер премии по ОСАГО в 2022 году составил 6,9 тыс. руб. Среднемесячный темп прироста стоимости полиса в рассматриваемом периоде был равен 2%, а рост за январь — декабрь достиг 27%. Среди лидеров по количеству проданных страховок остаются Московский регион, Свердловская область, Самарская область и Санкт-Петербург. Самый высокий рост средней премии в группе рассматриваемых регионов был в Санкт-Петербурге и области (45%), а самый низкий — в Свердловской агломерации (12%). Доля водителей с высокими значениями КБМ выше в столичных регионах, что вполне ожидаемо.

На основе статистики можно сделать вывод, что рынок автострахова-

ния постепенно развивается и увеличивается. При сохранении положительной динамики предполагается дальнейший рост общих сборов Российских страховщиков, объёма выплат и средних выплат.

Второй раздел **«Методы построения классов значений факторов риска»** представляет собой обзор методов построения классов значений факторов риска, эти методы необходимо применять для уменьшения количества классов значений с целью облегчения дальнейшего анализа этих факторов для выявления тарифных факторов, построения множителей для выбранных факторов и т.д.

В разделе были рассмотрены два различных кластер-метода, оба они хорошо подходят для выявления классов в совокупности рисков.

Кластер-метод на основе критерия равенства математических ожиданий (2.1) применим только к логнормальному распределению, а агломеративный кластер-метод с максимизацией функции правдоподобия (2.2) может применяться к любому определенному виду распределения.

Оба метода могут применяться только при наличии наблюдений за несколько (желательно $j \geq 5$) лет, что является их недостатком.

Еще одна проблема представленных методов заключается в том, что их применение возможно лишь к определенной модели распределения, а реальные данные не всегда соответствуют какому-либо типу распределения.

Метод 2.1 имеет существенный недостаток по сравнению с методом 2.2. Каждый раз, когда создается новый класс значений, его компоненты теряют возможность быть объединенными с другими классами значений, среди которых могла бы найтись группа из большего числа компонентов, что привело бы к более сильному сокращению общего числа классов. Этот недостаток может исправить грамотная перестановка порядка классов значений, как в методе 2.2.

Несмотря на свои недостатки представленные методы показывают хорошие результаты и могут быть эффективно использованы на практике.

В третьем разделе **«Сравнение методов выбора тарифных факторов»** представлены схема пошагового отбора, методы отбора тарифных факторов с помощью дисперсионного анализа, метода отношения правдоподобий и перекрёстной параметризации при наличии наблюдений за один год.

Все излагаемые методы отбора тарифных факторов построены по одинаковой схеме, типичной для регрессионного анализа. Эта схема была описана в разделе 3.1. Методы отличаются друг от друга только в отношении применяемых статистик критерия (или, что одно и то же, в отношении моделей распределения). После соответствующего преобразования данных и статистики критерия отбор факторов в разделе 3.2 проводится с помощью дисперсионного анализа. В разделе 3.3 за основу берется другой стандартный метод проверки гипотез — метод отношения правдоподобий. И в первом, и во втором случае допускается детализация отбора, когда на значимость исследуется не весь фактор, а только отдельные комбинации его (классов) значений.

Все рассмотренные методы отбора тарифных факторов сводятся к проверке гипотезы равенства нескольких математических ожиданий. Не гарантируя достижения оптимума, последовательный отбор, тем не менее, позволяет легко выявить взаимную зависимость факторов риска. Пошаговая оптимизация является стандартным методом многомерной статистики.

Четвёртый раздел **«Вычислительный эксперимент»** посвящен построению тарифных множителей для добровольного автомобильного страхования в зависимости от региона использования транспортного средства.

Цель вычислительного эксперимента — разбиение базы данных на классы для построения системы тарифных множителей добровольного страхования автотранспорта в зависимости от региона использования транспортного средства.

В качестве информационной базы использована статистика центрального банка Российской Федерации о выплатах по страхованию в период с 2016 по 2022 год.

Задача вычислительного эксперимента - написание программного обеспечения, которое позволяет на основе наблюдений за несколько лет построить классы значений тарифного фактора таким образом, чтобы количество этих классов значений было наименьшим — такой подход является наиболее удобным для практического использования итоговой системы.

В качестве метода обработки информации выбран «агломеративный кластер-метод с максимизацией функции правдоподобия» из раздела 2. Выбор метода обусловлен простотой его реализации по сравнению с другими

аналогичными и более широким спектром применения, так как применяется к любому определённом виду распределения. На основе этого метода реализован алгоритм на языке программирования Python. Фрагмент кода представлен в приложении А.

Подадим на вход программы статистические данные за 7 лет о размере тарифных факторов, таких как: КТ (территориальный коэффициент), КБМ (коэффициент «бонус-малус», чем больше стаж безаварийной езды, тем выше скидка), КВС (коэффициент возраста и стажа), КО (коэффициент ограничения, влияет на стоимость полиса при покупке страховки с неограниченным числом водителей), КМ (коэффициент мощности двигателя), КС (коэффициент сезонности, зависит от того, сколько месяцев в году используется авто). С помощью «агломеративного кластер-метода с максимизацией функции правдоподобия» разделим данные на два класса. Представлен результат кластеризации, на основе которого можно сделать вывод, что территориальный коэффициент наиболее отличается от тарифных факторов, которые попали во второй кластер.

Далее в данном эксперименте рассматриваемым тарифным фактором является «регион использования транспортного средства». На входе в программу подаются суммы выплат по добровольному страхованию автомобильного транспорта с 2016 по 2022 год (Приложение Б) в регионах Российской Федерации и количество заключённых договоров в этих регионах за период с 2016 по 2022 год (Приложение В).

Загружаем файлы и производим анализ данных. Представлены сведения о пропусках. В данных наборах данных пропуски и дубликаты отсутствуют. Если в дальнейшем, при реализации данной программы на другом наборе данных, будут присутствовать пропуски, то заполняем их с помощью детерминированного подхода (deterministic approach), представленного в Приложении А. Он предполагает, что мы заполняем пропуски строго теми значениями, которые будут предсказаны линейной регрессией.

Применяем к наборам данных агломеративную кластеризацию, используя евклидову метрику и метод Уорда для определения связей между регионами (создаем 2 кластера). Критерий минимальной дисперсии Уорда минимизирует общую дисперсию внутри кластера. Добавляем информацию о

принадлежности регионов к кластерам в исходный DataFrame. Выводим результат кластеризации. Отображаем дендрограмму для визуализации иерархии кластеров и связей между ними. Для удобства её анализа масштаб можно изменять.

Количество загружаемых файлов и кластеров можно изменять. Для подсчёта тарифных множителей для фактора «регион использования транспортного средства» на основе дендрограммы разделим базу данных средних выплат на классы значений. Для более точных значений множителей разделим данные на 6 кластеров (количество можно изменять). Применяем агломеративную кластеризацию, используя евклидову метрику и метод Уорда для определения связей между регионами (создаем 6 кластеров). Добавляем информацию о принадлежности регионов к кластерам в исходный DataFrame. Выводим результат кластеризации. Выводим результат подсчёта тарифных множителей для классов. Отображаем дендрограмму для визуализации иерархии кластеров и связей между ними.

На основе работы программы с использованием агломеративной кластеризации можно сделать следующие выводы:

1) Определение структуры данных: получено представление о структуре данных, которые включают информацию о тарифных факторах, выплатах по страхованию КАСКО с 2016 по 2022 год в регионах Российской Федерации и количество заключённых договоров в этих регионах за тот же период;

2) Кластеризация регионов: программа позволяет кластеризовать регионы на основе их страховых выплат и количества заключённых договоров, что может помочь в выявлении групп регионов со схожими характеристиками. Это может быть полезно для страховых компаний при разработке тарифов, определении рисков или планировании маркетинговых стратегий;

3) Визуализация результатов: дендрограмма позволяет визуально изучить иерархию кластеров и определить оптимальное количество кластеров для дальнейшего анализа. Можно экспериментировать с разным количеством кластеров и выбрать тот, который наилучшим образом описывает структуру данных;

4) Анализ полученных кластеров: на основе полученных кластеров можно определить, какие тарифные факторы или регионы входят в каждый кла-

стер. Это может дать информацию о схожести факторов или регионов в рамках каждого кластера и позволить выявить особенности каждой группы.

В целом, программа предоставляет полезную информацию о тарифных факторах и регионах Российской Федерации, которая может быть использована страховыми компаниями для анализа данных и принятия решений на основе этого анализа. Данная программа подходит для реализации не только представленного набора данных, но и других статистических наборов данных. Фрагмент программы представлен в Приложении А.

Таким образом, в результате вычислительного эксперимента были получены тарифные множители для субъектов РФ.

Заключение. Рост интенсивности дорожного движения, сопровождающийся количественным ростом дорожно-транспортных происшествий, объективно требует организации страхования гражданской ответственности владельцев автотранспортных средств и самих автотранспортных средств.

Страхование представляет собой рынок, который развивается или стагнирует в зависимости от изменения доходов населения и предприятий.

Ролью тарифных факторов в автостраховании является их влияние на ставку страхования.

В первом разделе работы был проведен экономико-статистический анализ рынка автострахования в России, рассмотрены основные тенденции и динамики рынка, такие как страховые премии, выплаты, а также количество страховых компаний на рынке. Определены основные тарифные факторы автомобильного страхования. Рассмотрены сходства и различия обязательного и добровольного автострахования, а так же проанализировано развитие и современное состояние рынка страхования ОСАГО и КАСКО в России.

Во втором разделе были изучены методы построения классов значений факторов риска, а именно — агломеративный кластер-метод на основе критерия равенства математических ожиданий и агломеративный кластер-метод с максимизацией функции правдоподобия, сделаны выводы о достоинствах и недостатках данных методов.

В третьем разделе было дано определение тарифных факторов, а также рассмотрены методы их отбора, а именно — схема пошагового отбора, на основе которой строятся большинство методов выбора тарифных факторов,

метод отбора тарифных факторов с помощью дисперсионного анализа, метод отношения правдоподобий, перекрёстная параметризация при наличии наблюдений за один год, а результатом третьего раздела стал вывод об эффективности данных методов при практическом применении.

В вычислительном эксперименте была разработана программа на языке Python, фрагмент которого представлен в приложении А. Она позволяет строить новые классы значений как для нескольких факторов риска, так и для конкретного фактора риска. Разработанное программное обеспечение имеет практическую значимость для построения классов значений факторов риска и может быть использовано для сокращения числа классов значений не только по отношению к КАСКО, но и для тарифных факторов других видов страхования. Программа предоставляет полезную информацию о регионах Российской Федерации, которая может быть использована страховыми компаниями для анализа данных и принятия решений на основе этого анализа.