МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра теории функций и стохастического анализа

МОДЕЛИРОВАНИЕ СОВОКУПНОГО УБЫТКА ПОРТФЕЛЯ РИСКОВ

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

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

механико-математического факультета
Хрипунова Константина Сергеевича

Научный руководитель
доцент, к. ф.-м. н. ______________ Л. В. Борисова

Заведующий кафедрой
д. ф.-м. н., доцент ______________ С. П. Сидоров

Саратов 2021

ВВЕДЕНИЕ

Актуальность темы. На протяжении нескольких десятков лет во многих видах деятельности являются актуальными вопросы, связанные с изучением страховых расчетов. В свою очередь при их рассмотрении требуются данные о характере страховых случаев (рисков) и процессе ущербов. Если имеющаяся в распоряжении база данных охватывает несколько страховых компаний, то зачастую данные изначально представлены в агрегированной форме.

Получается, что информация о страховой сумме каждого риска и размере каждого отдельного убытка отсутствует и доступны только суммарные годовые статистические показатели, такие как число рисков, их совокупная страховая сумма, а также число и суммарный размер убытков.

Таким образом в качестве данных имеем только одно значение совокупного убытка в каждом году. Из-за этого привлечение различных моделей распределения кажется бессмысленным, ведь параметры распределения нельзя оценить только по одному наблюдению. Решить эту проблему позволяет нормирование совокупного убытка на соответствующий объем с помощью распределения с небольшим количеством параметров.

Целью бакалаврской работы является анализ методов моделирования совокупного убытка портфеля рисков на конкретных примерах задач возникающих в страховых компаниях.

Объект исследования— роль и определение совокупного убытка.

Предмет исследования— алгоритмы моделирования и индентификации числа страховых рисков, совокупных убытков.

Для достижения поставленных целей в работе необходимо решить следующие **задачи**:

- рассмотреть две группы моделей, выявить их преимущества и недостатки и выделить общий принцип их построения;
- построить модель зависимости дисперсии от объема в случае однородных и неоднородных рисков;
- исследовать гамма-распределение, получить плотности при различных параметрах, написать алгоритм построения их графиков на языке про-

граммирования Python 3;

— построить модель совокупного убытка портфеля рисков в однородном и неоднородном случае;
— разработать численный метод, оценивающий параметры распределения;
— создать код на языке программирования C++, позволяющий смоделировать или идентифицировать совокупный страховой убыток, оценить параметры распределения с использованием приведенного численного метода.
— На языке Python 3 построить график, полученной модели и оценить насколько подходит гамма распределение при моделировании совокупного убытка.

Практическая значимость. Разработан способ моделирования совокупных убытков при различных страховых случаях и написан код, позволяющий смоделировать или идентифицировать убытки портфеля рисков по имеющимся статистическим данным, что позволит компании адекватно произвести страховой расчет в будущем.

Структура и содержание бакалаврской работы. Работа состоит из введения, шести разделов, заключения, списка использованных источников, содержащего 20 наименований, и четырех приложений. Общий объем работы составляет 42 страницы.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обосновывается актуальность темы работы, формулируется цель работы и решаемые задачи, вводятся основные понятия, отмечается практическая значимость полученных результатов.

В **первом** разделе даны определения риска, портфеля рисков и рассмотрены некоторые моменты при их постановке для страховой компании и приведен пример идеальных условий для портфеля, в котором поставлено понятие совокупного убытка.

Экономический риск — возможность случайного возникновения нежелательных убытков, измеряемых в денежном выражении.

Таким образом под риском можно понимать некоторый показатель опасности или совокупности опасностей, установленный в стоимостном выражении его ущерба за заданное время. В связи с этим на первый план выходит формирование сбалансированного страхового портфеля. Определим страховой портфель $S(n, c, \text{П}, \text{В}, \text{Р})$ — как те обязательства, которые взяла на себя страховая компания перед своими клиентами. Одновременно с обязательствами происходит и передача активов компании в размере того, сколько она должна будет выплатить при наступлении страхового случая. Здесь n – число договоров, c – собственный капитал компании, П – сумма поступивших страховых премий, то есть размер финансовых средств, которыми располагает страховщик, для ведения дела, В – сумма выплат, то бишь объем выполненных страховщиком обязательств, действительную платёжеспособность страховщика и Р – сумма страховых резервов — объем страхового фонда, для выполнения обязательств перед страхователями.

Пример. Рассмотрим идеальную ситуацию портфеля из I независимых одинаково распределенных рисков $R_i, \ \ i = 1, ..., I$. R_i представляет собой случайную величину совокупного убытка по риску i за определенный временной промежуток, например за один год. Под риском можно понимать мельчайшую единицу, которая в отдельности могла бы быть предметом договора страхования, так как полисе часто бывает указано несколько рисков. Соответственно квадрат коэффициента вариации $Vko(S)$ совокупного убытка

$$S = R_1 + ... + R_I$$

равен

$$(Vko(S))^2 = \frac{Var(S)}{(E(S))^2} = \frac{I \cdot Var(R_1)}{(I \cdot E(R_1))^2} = \frac{(Vko(R_1))^2}{I}$$

и с ростом объема портфеля I уменьшается. Таким образом, стандартное отклонение $Sta(S) = \sqrt{Var(S)}$ растет медленнее, чем математическое ожидание$E(S)$.

Смысл этого результата становится ясен из неравенства Чебышева:

$$P(|S - E(S)| > \varepsilon \cdot E(S)) \leq \frac{Var(S)}{\varepsilon^2 \cdot (E(S))^2} = \frac{(Vko(R_1))^2}{\varepsilon^2 \cdot I}$$

справедливого при любом $\varepsilon > 0$. Это означает, что вероятность превышения совокупным убытком S своего математического ожидания более чем на 100% уменьшается с ростом объема портфеля I. Абсолютный же разброс значений совокупного убытка вокруг среднего с ростом портфеля будет увеличиваться.

Во **втором** разделе рассказывается о двух группах моделей, а именно — индивидуальные и коллективные, а также о схеме построения моделей.

Обзор моделей совокупного убытка группы рисков

Существует две группы моделей: индивидуальные и коллективные. В индивидуальной модели распределение совокупного годового убытка группы рисков получается в результате свертки распределений годовых совокупных убытков отдельных рисков. В рисковом страховании узнать распределение убытка отдельного риска практически невозможно, поэтому индивидуальная модель имеет смысл только для однородных групп рисков. Она особенно важна для моделирования математических ожиданий совокупных убытков одновременно нескольких групп рисков.

В коллективной модели распределение совокупного убытка группы рисков строится на основе распределений числа убытков и размера убытка в одном страховом случае, при этом плоскость отдельных рисков не задействуется. Этот способ позволяет, в частности, реалистично смоделировать хвост распределения совокупного убытка и получить представление о вероятности любого размера годовых потерь. Таким образом, коллективная модель непосредственно решает важную практическую задачу, в то время как индивидуальная модель находит применение лишь опосредованно как составная часть других моделей.

Общие принципы построения

Решающее значение при моделировании имеет самый первый шаг — выбор модели. Следующим шагом будет оценка параметров модели, которая является только делом техники, поскольку в последние годы различные вспомогательные средства значительно усовершенствовались благодаря стремительному развитию персональных компьютеров. На третьем шаге проверяется соответствие модели данным (проверка гипотезы согласия). Только в случае удовлетворительного результата имеет смысл четвертый шаг — получение ответа на интересующий вопрос в рамках построенной модели. Наконец, на пятом шаге ответ необходимо утвердить, убедившись в его приемлемости и проверив чувствительность (влияние выбросов и т. д.). На этом этапе снова может возникнуть необходимость изменения модели и прохождения всех шагов заново.

Во **третьем** разделе приводится теорема об оптимальности линейной комбинации оценок, строится модель зависимости дисперсии ставки убытка от объема страховой суммы в двух случаях.

Моделирование зависимости дисперсии от объема

Приведем теорему об оптимальности линейной комбинации оценок, взвешенных в обратной пропорции со своими дисперсиями (неточностью), поскольку на нее можно будет ссылаться при построении оценок.

Теорема 1. *Пусть* $T_1, ..., T_I$ *— независимые несмещенные оценки величины* t*:* $E(t_i) = t, \quad 1 \leq i \leq I$ *и пусть* $T = T_1w_1 + ... + T_Iw_I$ *- линейная комбинация этих оценок, причем* $w_1 + ... + w_I = 1$*, то есть* $E(T) = t$*. Тогда среди всех возможных линейных комбинаций вида* T *минимальную дисперсию имеет комбинация с весами* w_i*, обратно пропорциональными* $Var(T_i)$ *:* $w_i = w \ / \ Var(T_i)$ *для всех* $i = 1, ..., I$*.*

Случай однородных рисков

Введем обозначения:

$$m = E(R_i), \ \ 1 \leq i \leq I, \ \ S^2 = Var(R_i), \ \ 1 \leq i \leq I, \tag{1}$$

где R_i—однородная группа рисков и для совокупного убытка группы рисков

$$S = \sum_{i=1}^{I} R_i, \tag{2}$$

Используем более наглядную нормированную на объем случайную величину, которая характеризует убыток на один страховой год:

$$Z = \frac{S}{I} \tag{3}$$

для нее справедливо

$$E(Z) = m, \ \ Var(Z) = \frac{s^2}{I} \tag{4}$$

При известных реализациях r_i величин R_i значения

$$\hat{m} = \frac{1}{I}\sum_{i=1}^{I} r_i, \ \ \hat{s^2} = \frac{1}{I-1}\sum_{i=1}^{I}(r_i - \hat{m})^2 \tag{5}$$

представляют собой несмещенные оценки для m и S^2.

Но данные по отдельным рискам часто бывают недоступны, и в распоряжении имеются только суммарные годовые показатели числа рисков I_j и очищенного от инфляции убытка на один страховой год z_j случайной величины Z_j портфеля за несколько лет $j = 1, ..., J$. Если данные очищены от инфляции, и изменения в объеме покрытия или структуре убытка пренебрежимо малы, то распределение величины R_i, значения параметров m и s^2 для всех лет можно считать одинаковыми: $E(Z_j) = m$ и $Var(Z_j) = \frac{s^2}{I_j}$. Тогда несмещенные оценки для m и s^2 находятся по формулам используя теорему 1.

$$\hat{m} = \frac{\sum_{j=1}^{J} I_j z_j}{\sum_{j=1}^{J} I_j}, \ \ \hat{s^2} = \frac{1}{J-1}\sum_{j=1}^{J} I_j(z_j - \hat{m})^2 \tag{6}$$

Случай неоднородных рисков

В большинстве видов имущественного страхования для рисков одной тарифной группы используется одинаковая ставка премии, которая умножается на соответствующие страховые суммы. Тем самым подразумевается пропорциональность

$$E(R_i) = m \cdot u_i, \ \ i = 1, ..., I \tag{7}$$

(m не совпадает с использованным выше). Принимаем следующие равенства

$$E(R_i) = E(R_1)u_i/u_1, \ \ Var(R_i) = Var(R_1)u_i/u_1 \tag{8}$$

для всех $i = 1, ..., I$.

Пусть

$$\nu = \sum_{i=1}^{I} u_i \tag{9}$$

совокупная страховая сумма группы рисков. Используя обозначения

$$m = \frac{E(R_1)}{u_1}, \ \ s^2 = \frac{Var(R_1)}{u_1} \tag{10}$$

для совокупного убытка

$$S = \sum_{i=1}^{I} R_i \tag{11}$$

независимых рисков, получим

$$E(S) = m \cdot \nu, \ \ Var(S) = s^2 \cdot \nu. \tag{12}$$

Соответствующей нормированной на объем случайной величиной теперь выступает ставка убытка

$$Z = \frac{S}{\nu}, \tag{13}$$

имеющая (аналогично убытку на один страховой год) математическое ожидание

$$E(Z) = m \tag{14}$$

и дисперсию

$$Var(Z) = \frac{s^2}{\nu}, \tag{15}$$

то оправдывает применение одинаковых обозначений для ставки убытка и убытка на один страховой год (условие $u_i = 1$ приводит к рассмотренному выше однородному случаю).

Несмещенные оценки для m и s^2 на базе реализаций r_i величин R_i рассчитываются по формулам

$$\hat{m} = \sum_{i=1}^{I} \frac{r_i}{\nu} = \sum_{i=1}^{I} \frac{u_i}{\nu} \cdot \frac{r_i}{u_i}, \quad \hat{s}^2 = \frac{1}{I-1} \sum_{i=1}^{I} u_i \left(\frac{r_i}{u_i} - \hat{m} \right)^2 \tag{16}$$

Пусть $Z_j, 1 \leq j \leq J$ - очищенная от инфляции ставка убытка группы рисков в j-м году, и ν_j - соответствующий объем, выраженный совокупной страховой суммой. Тогда величины m и s^2, могут считаться постоянными в течение нескольких лет: $E(Z_j) = m$ и $Var(Z_j) = s^2/\nu_j$ Несмещенные оценки параметров m и s^2 на основе реализаций z_j величин Z_j (снова аналогично убытку на один страховой год) равны

$$\hat{m} = \frac{\sum_{j=1}^{J} \nu_j z_j}{\sum_{j=1}^{J} \nu_j}, \tag{17}$$

$$\hat{s}^2 = \frac{1}{J-1} \sum_{j=1}^{J} \nu_j (z_j - \hat{m})^2 = \frac{1}{J-1} \sum_{j=1}^{J} \nu_j z_j^2 - \frac{\nu_+}{J-1} \cdot \hat{m}^2, \tag{18}$$

В **четвертом** разделе показан способ моделирования числа убытков с помощью стандартного распределения Пуассона.

Моделирование числа убытков

Пусть N — случайное число страховых событий, происходящих в определенном фиксированном временном промежутке (например, в предстоящем календарном году) и относящихся к портфелю рисков (например, совокупности рисков страхования автогражданской ответственности). Тогда при выполнении довольно общих предпосылок N имеет распределение Пуассона:

$$P(N = n) = \frac{\theta^n e^{-\theta}}{n!}, \quad n = 0, 1, 2, ...,$$

с параметром $\theta = E(N) = Var(N)$. Перечислим эти предпосылки:

1. случайные величины числа убытков в двух непересекающихся подынтервалах рассматриваемого временного промежутка независимы,
2. одновременно не могут произойти два и более убытка (регулярность так называемой функции интенсивности),
3. убытки могут происходить в любые моменты времени (непрерывность функции интенсивности).

В **пятом** разделе выводится плотность гамма-распределения, строятся графики при различных значениях параметра α, далее рассматривается модель совокупного годового убытка Z (как однородной, так и неоднородной) группы рисков и приводится численный метод, позволяющий получить оценки параметров распределения.

Моделирование совокупного убытка портфеля рисков с помощью гамма-распределения

Известно, что основная масса вероятностей сосредоточена в точке $R_i = 0$, ведь практически во всех видах рискового страхования подавляющее большинство рисков (в отдельно взятом году) не порождает убытков. Значит, распределение случайной величины R_i не имеет непрерывной плотности и далеко от нормального распределения.

Для получения приемлемой аппроксимации распределения совокупного убытка S малой (как это характерно для практики) группы рисков аппроксимируем распределение совокупного убытка R_i отдельного риска i непрерывным распределением, допускающим явный расчет сверток. Для грубой аппроксимации распределения величины R_i достаточно знать, что основная масса вероятностей находится в нуле. В последующем процессе свертки неточность аппроксимации довольно быстро нивелируется и достигаются очень близкие к реальности модели совокупного убытка.

Самым известным распределением на интервале $(0; \infty)$, позволяющим рассчитать свертки в явном виде, является гамма-распределение с плотностью

$$g(x) = e^{-x\alpha\,/\,\mu} \cdot x^{\alpha-1}\frac{(\alpha\,/\,\mu)^{\alpha}}{\Gamma(\alpha)}, \;\; x > 0. \tag{19}$$

Отступая от обычного представления плотности гамма-распределения, выбрали параметризацию с участием математического ожидания μ. Дисперсия гамма-распределения равна $\mu^2\,/\,\alpha$, коэффициент вариации $1\,/\,\sqrt{\alpha}$, асимметрия $2\,/\,\Gamma(\alpha)$.

Получение плотностей γ - распределений

Используя формулу 19 гамма-распределения, необходимо получить плотности данного распределения при различных значения параметра α и постоянном математическом ожидании равном 1, чтобы построить их графики и более подробно понять как ведет себя функция.

$$g(x) = e^{-x\cdot\alpha} \cdot x^{\alpha-1}\frac{(\alpha)^{\alpha}}{\Gamma(\alpha)}, \;\; x > 0. \tag{20}$$

Итак имеем следущую картину: при $\alpha = 1$ получается убывающее по экспоненте из точки $g(0) = 1$ экспоненциальное распределение; при $\alpha < 1$ график выглядит еще более асимметричным и еще круче убывает из точки $g(0) = \infty$; при $\alpha > 1$ справедливо равенство $g(0) = 0$; единственная мода имеет значение $\mu - \mu/\alpha$, а вид плотности с ростом α становится все симметричнее и более похож на нормальное распределение. (Плотности при $\alpha = 0,04$ и $\alpha = 0,2$ пересекаются слева и справа за пределами рисунка 1)

Важно, чтобы аппроксимирующее распределение величины R_i имело как можно больший вероятностный вес вблизи нулевой точки, потому что, как показывает пракика, основная масса вероятностей сосредоточена именно в этой точке, поэтому принимаем в рассмотрение только гамма-распределение с параметром формы $\alpha < 1$. Например, при $\alpha = 0,02$ 89% вероятностной массы сосредоточено в области ниже значения $\mu\,/\,10$.

Однородный случай

Сначала предположим однородную группу рисков и аппроксимируем неизвестное распределение величины R_i гамма-распределением с параметрами

$$\mu = m \;\text{ и }\; \alpha = m^2\,/\,s^2, \tag{21}$$

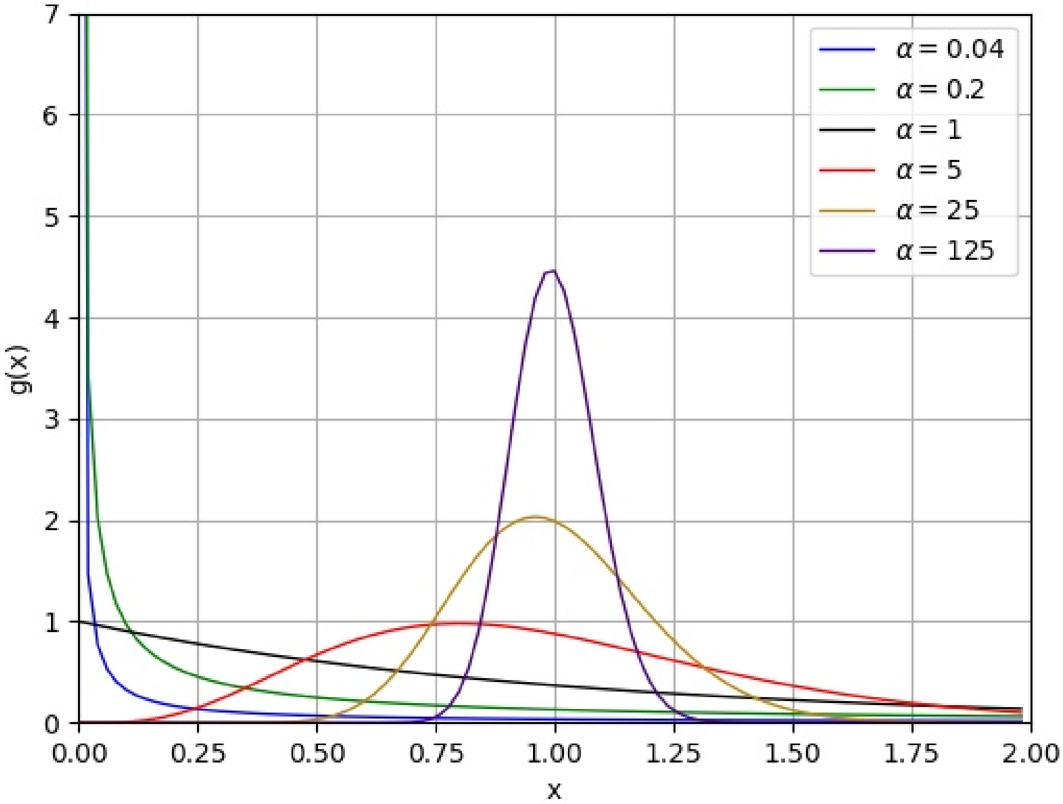

Рисунок 1 – Плотности гамма-распределений с математическим ожиданием 1

где $m = E(R_i)$ и $s^2 = Var(R_i)$.

Параметры распределения можно найти из условий равенства соответствующих теоретических и эмпирических моментов. Тогда в результате I-кратной свертки гамма распределений с параметрами μ и α получим распределение совокупного убытка группы I независимых рисков R_i.

I-кратная свертка гамма-распределений снова дает гамма-распределение, но с параметром среднего $I \cdot \mu$ и параметром формы $I \cdot \alpha$. Переход от совокупного убытка S к нормированной на объем величине $Z = S \;/\; I$ (убыток на один страховой год) не изменяет тип распределения, а значит Z будет иметь гамма-распределение с параметром среднего μ и параметром формы $I \cdot \alpha$.

Построенное на основе приемлемых аппроксимаций распределений отдельных рисков, гамма-распределение может считаться вполне реалистичной моделью для совокупного и нормированного убытков группы одинаково распределенных независимых рисков.

Неоднородный случай

В неоднородном случае, т.е. когда совокупный убыток группы рисков имеет разные страховые суммы, будем пользоваться очень выгодным свойством для гамма-распределения.

Свойство. Сумма независимых гамма-распределенных рисков R_i имеет гамма-распределение и в том случае, когда параметры μ и α не одинаковы для всех рисков i, но отношение этих параметров постоянно, т.е. $\mu_i \ / \ \alpha_i = const$.

Таким образом снова отталкиваемся от распределений отдельных рисков, что гарантирует наибольшую правдоподобность модели совокупного убытка. Предположим для совокупного убытка R_i отдельного риска со страховой суммой u_i гамма-распределение с параметрами

$$\mu_i = m \cdot u_i \ \text{ и } \ \alpha_i = m^2 u_i \ / \ s^2 \tag{22}$$

Тогда $\mu_i \ / \ \alpha_i$ действительно для всех рисков одинаково, и $S = R1 + ... + R_I$ тоже имеет гамма-распределение, но с параметром математического ожидания

$$\mu_1 + ... + \mu_I = m \cdot \nu \tag{23}$$

и с параметром формы

$$\alpha_1 + ... + \alpha_I = m^2 \cdot \nu \ / \ s^2, \tag{24}$$

где $\nu = u_1 + ... + u_I$.

Как видим, для получения распределения совокупного убытка будет снова достаточно просто сложить параметры. Ставка убытка $Z = S \ / \ \nu$ будет иметь гамма-распределение с параметрами m и $m^2 \cdot \nu \ / \ s^2$.

Соответственно, как в однородном, так и в неоднородном случае можно моделировать нормированную на объем величину убытка Z гамма-распределением с параметрами

μ и $\nu \cdot \alpha$,

где

$\mu = m$ – среднее значение,

ν – известный объем (число страховых лет или совокупная страховая сумма),

$\alpha = m^2 \; / \; s^2$.

Оценка параметров распределения

В построенной модели нормированный на известный объем ν_j совокупный убыток $Z_j = S_j \; / \; \nu_j$ j-го года имеет гамма-распределение с параметром математического ожидания μ и параметром формы $\nu_j \alpha$ (дисперсия равна $\mu^2 \; / \; (\nu_j \alpha)$). Два неизвестных параметра μ и α оцениваются привычными методами на основании годовых реализаций z_j, $1 \leq j \leq J$. Используя несмещенные оценки для m и s^2 (17) и (18), а также соотношения

$$\mu = m, \;\; \alpha = m^2 \; / \; s^2, \tag{25}$$

получаем оценки по методу моментов

$$\hat{\mu} = \frac{\sum\limits_{j=1}^{J} \nu_j z_j}{\sum\limits_{j=1}^{J} \nu_j}, \tag{26}$$

$$\hat{\alpha} = \frac{\hat{\mu}^2 (J-1)}{\sum\limits_{j=1}^{J} \nu_j (z_j - \hat{\mu})^2}. \tag{27}$$

Значение $\hat{\mu}$ одновременно является и оценкой максимального правдоподобия для μ. Оценка правдоподобия для α удовлетворяет уравнению

$$\sum_{j=1}^{J} \nu_j \big(\ln(\hat{\alpha}\nu_j z_j \;/\; \hat{\mu}) - \Psi(\hat{\alpha}\nu_j)\big) = 0 \tag{28}$$

где

$$\Psi(x) = \frac{d(\ln(\Gamma(x)))}{dx} = \Gamma^{'}(x) \;/\; \Gamma(x) \tag{29}$$

есть дигамма-функция. Это уравнение решается методом последовательных приближений с помощью аппроксимации

$$\Psi(x) \approx \ln(x) - x^{-1} \;/\; 2 - x^{-2} \;/\; 12 + (x^{-4} - 0,46x^{-6}) \;/\; 120, \tag{30}$$

имеющей при $x > 5$ точность до 8 десятичных знаков, а также рекурсии

$$\Psi(x) = \Psi(x+1) - 1 \;/\; x. \tag{31}$$

Заменив $\Psi(\alpha\nu_j)$ на

$$\Psi(\alpha\nu_j) = \Psi(\alpha\nu_j + 5) - \frac{1}{\alpha\nu_j + 4} - \frac{1}{\alpha\nu_j + 3} - \frac{1}{\alpha\nu_j + 2} - \frac{1}{\alpha\nu_j + 1} - \frac{1}{\alpha\nu_j} \tag{32}$$

и выразив α из последнего слагаемого правой части, получаем итерационную форму уравнения

$$\alpha = J \Bigg/ \sum_{j=1}^{J} \nu_j \left(\Psi(\alpha\nu_j + 5) - \frac{1}{\alpha\nu_j + 4} - \frac{1}{\alpha\nu_j + 3} - \right.$$
$$\left. - \frac{1}{\alpha\nu_j + 2} - \frac{1}{\alpha\nu_j + 1} \right) - \ln\left(\frac{\alpha\nu_j z_j}{\hat{\mu}} \right). \qquad (33)$$

Для запуска итерации используем вместо $\Psi(\alpha\nu_j + 5)$ формулу **??** и подставляем в правую часть в качестве стартового значения α его оценку по методу моментов.

Для вычисления самой функции правдоподобия требуется значение $\ln((x))$. При $x > 3$ удобно воспользоваться аппроксимацией

$$\ln(\Gamma(\mathrm{x})) \approx \frac{1}{2}\ln\left(\frac{2\pi}{x}\right) + x \cdot \ln(x) - x +$$
$$+\left(\left(\left(1 - \frac{2}{3x^2}\right)\frac{2}{7x^2} - 1\right)\frac{1}{30x^2} + 1\right)\frac{1}{12x}, \qquad (34)$$

имеющей точность до 7 десятичных знаков, а при $0 < x \leq 3$ - рекурсией

$$\ln(\Gamma(x)) = \ln(\Gamma(x+1)) - \ln(x). \qquad (35)$$

Шестой раздел посвящен вычислительному эксперименту.

Вычислительный эксперимент

Рассмотрим задачу следующего вида: в результате наблюдения за группой из N_0 однородных объектов на некотором интервале времени T каждому из одинаковых по продолжительности интервалов времени $\Delta t = t_j - t_{j-1} (j = 1, 2, ..., J)$ поставлены в соответствие число страховых событий ν_j и значения страховых ущербов по множеству страховых случаев $z_j^{(\nu)} (\nu = 1, 2, ..., m_i)$.

При страховании N объектов (где $N \neq N_0$) удалось найти вероятности $p_0^{(0)}, p_1^{(0)}, ..., p_k^{(0)}, ...$ того, что на годовом интервале времени произошло $0, 1, 2, 3, ..., k, ...$ страховых событий. Тогда если $F_0(x)$ есть функция распределения страхового ущерба в единичном страховом событии, то функция распределения суммы k независимых одинаково распределенных случайных величин, каждая из которых описывается функцией распределения $F_0(x)$, выражается k-кратной сверткой этой функции

$$P\{W < x|k\} = F^{(k)}(x) = (*)^k F_0(x)(k = 2, 3, ...). \tag{36}$$

Так как рапределение вероятностей подчиняется условию нормированности $\sum_{k=0}^{\infty} p_k^{(0)} = 1$, то для вычисления функции распределения совокупного страхового ущерба на годовом интервале времени можно применить формулу полной вероятности:

$$R_0(x) = \sum_{k=0}^{\infty} p_k^{(0)} P\{W < x|k\} = \sum_{k=0}^{\infty} p_k^{(0)} F^k(x). \tag{37}$$

В качестве аппроксимации $F_0(x)$ можно взять гамма-распределение с плотностью, соответствующей формуле 19, поскольку позволит избежать необходимости определения сверток численными методами. С помощью данного распределения можно вычислить эти свертки в аналитической форме.

Используя свойство замкнутости по операции свертки гамма-распределения формулу 37 можно переписать в виде:

$$R(x) = p_0 + \sum_{k=1}^{\infty} p_k F_\gamma(x|k\alpha, \mu), \tag{38}$$

где $F_\gamma(x|k\alpha, \mu)$ функция распределения совокупного страхового убытка в результате наступления k страховых событий.

Для нахождения функции распределения совокупного страхового убытка по формуле 38 нужно найти параметры $F_\gamma(x|k\alpha, \mu)$ и функции распределения числа страховых событий p_k. Таким образом, в соответствии с таблицей 1 используем следующие статистические данные:

Далее по формулам 28 и 29 нужно решить уравнение максимального правдоподобия (для нахождения α), используя численный метод, приведен-

Таблица 1 – Статистические данные о страховых убытках по месяцам

j	ν_j	$z_j \times 10^{-6}$(руб.)
1	3	15,6
2	4	79,4
3	5	44,6
4	8	51,5
5	12	19,4
6	13	38,6
7	24	133,3
8	13	38,6
9	17	40,5
10	25	20,4
11	25	21,8
12	42	87,2

ный в предыдущей главе, причем в качестве начального значения α для запуска итерации необходимо взять оценки, полученные по методу моментов.

$$\sum_{j=1}^{J} \nu_j \Bigg(\ln\left(\frac{\hat{\alpha}\nu_j z_j}{\hat{\mu}}\right) - \Psi(\hat{\alpha}\nu_j) \Bigg) = 0 \tag{39}$$

где

$$\Psi(x) = \frac{d\Big(\ln\left(\Gamma(x)\right)\Big)}{dx} = \frac{\Gamma'(x)}{\Gamma(x)} \tag{40}$$

откуда

$$\alpha_k = J \Bigg/ \sum_{j=1}^{J} \nu_j \Bigg(\Psi(\alpha_{k-1}\nu_j + 5) - \frac{1}{\alpha_{k-1}\nu_j + 4} - \frac{1}{\alpha_{k-1}\nu_j + 3} -$$

$$- \frac{1}{\alpha_{k-1}\nu_j + 2} - \frac{1}{\alpha_{k-1}\nu_j + 1} \Bigg) - \ln\Bigg(\frac{\alpha_{k-1}\nu_j z_j}{\hat{\mu}} \Bigg),$$

где k — число итераций, а

$$\alpha_0 = \frac{\hat{\mu}^2(J-1)}{\sum\limits_{j=1}^{J} \nu_j (z_j - \hat{\mu})^2}. \tag{41}$$

Итерации проводим до тех пор, пока разница между предыдущем и полученным α будет больше ε, где ε — заданная точность.

Таким образом, с помощью программы на языке программирования $C++$, используя численный метод, удалось получить значения параметра α с точностью до 8 знаков после запятой за 15 итераций. Математическое ожидание было посчитано из оценки по методу моментов по формуле, полученной в предыдущем разделе.

$$\hat{\mu} = \frac{\sum\limits_{j=1}^{J} \nu_j z_j}{\sum\limits_{j=1}^{J} \nu_j},$$

Откуда получаем $\alpha = 0,151006$, $\mu = 56,7586$.

Для выяснения распределения числа страховых случаев расчитываются выборочные математическое ожидание и дисперсия числа страховых случаев на месячном интервале времени при $J = 12$:

$$m = \frac{\sum\limits_{j=1}^{J} \nu_j}{J} \tag{42}$$

$$D = \frac{\sum\limits_{j=1}^{J} (\nu_j - m)^2}{J-1} \tag{43}$$

Соответственно получаем $m = 15,916$, $D = 130,447$.

Полученные значения удовлетворяют условию $m < D$, поэтому представляется возможным идентифицировать распределение числа страховых событий с помощью распределения Пуассона.

Таким образом для вероятностей $p_k(t)$ пуассоновского потока событий справедлива формула:

$$p_k(t) = e^{-\theta}\frac{\theta^k}{k!} \quad k = 0, 1, 2, ..., \tag{44}$$

где $\theta = \lambda N_t$, N_t — количество объектов; λ — параметр потока, который выражает интенсивность появления страховых событий в группе страхуемых объектов.

Оценка параметра λ определяющего распределение Пуассона, находится по формуле:

$$\lambda = \frac{\nu_c}{N_0 T}$$

где nu_c – число страховых случаев, зафиксированных в группе наблюдаемых объектов, T – интервал времени.

Таким образом с помощью программы на языке *Python* находим значение параметра $\lambda = 0.015916$ при $N = 1000$, $T = 12$ и $nu_c = \sum_{j=1}^{T} nu_j = 191$ и далее подсчитываем вероятности p_k для каждого $k = 0, ..., 20$ по формуле 44.

Далее, также на языке *Python*, можно построить график полученной модели, используя формулу 38. Причем функцию распределения $F_\gamma(x|k\alpha, \mu)$ можно найти, используя:

$$F_\gamma(x|\alpha, \mu) = \frac{1}{\Gamma(\alpha)}\gamma(\alpha, \alpha x \mu)$$

где γ – нижняя неполная гамма- функция, т.е. вычисляется по формуле:

$$\gamma(s, x) = \int_0^x t^{s-1}e^{-t}dt$$

Но вполне можно отказаться от данной формулы и разработки очередного численного метода, используя библиотеку *Python Scipy*, в котором уже есть эта функция.

Таким образом, в соответствии с рисунком 2 имеем следующий график для $R(x)$.

Для проверки адекватности полученной модели необходимо построить

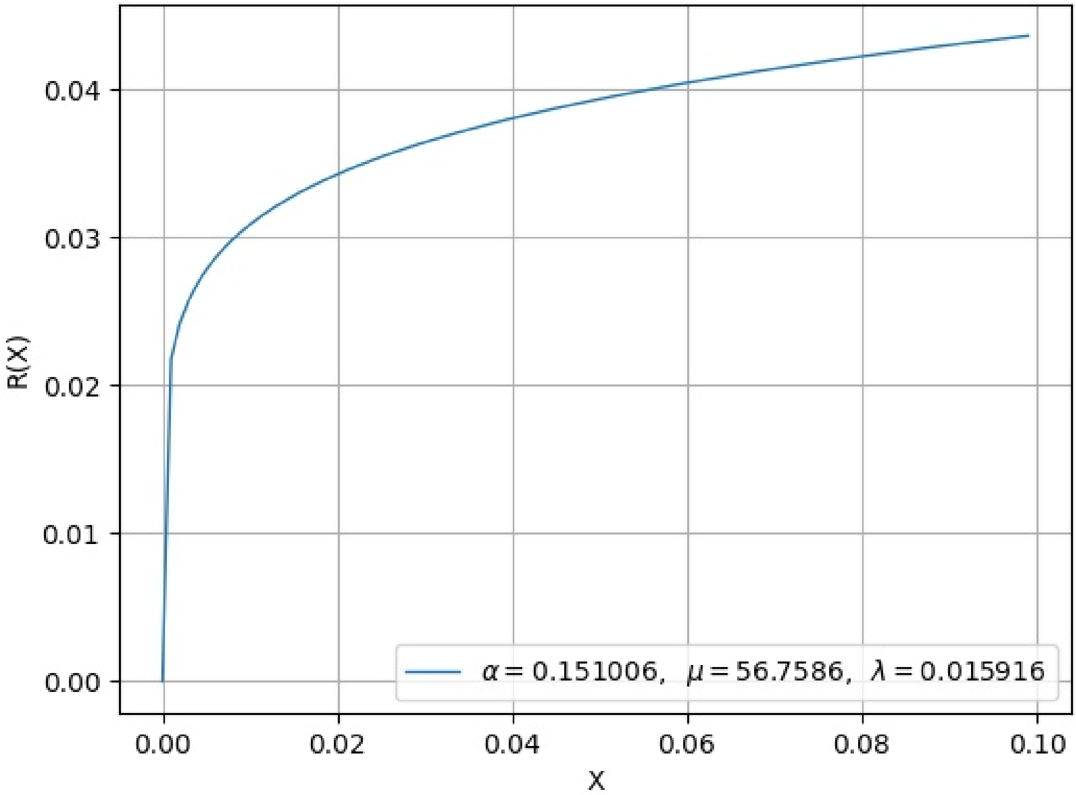

Рисунок 2 – График модели $R(x)$

график функции гамма-распределения. Таким образом, с помощью алгоритма на языке $Python$ в соответствии с рисунком 3 имеем следующий график:

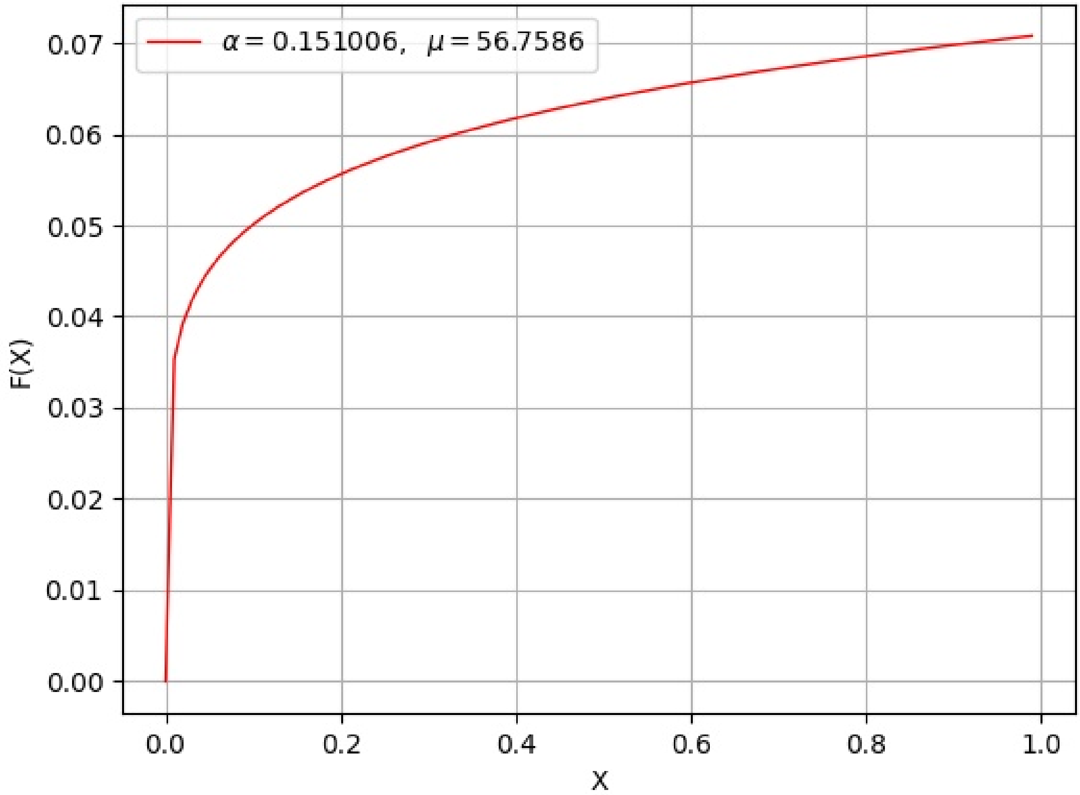

Рисунок 3 – График гамма-распределения

Соответственно на основе полученных параметров и графиков можно сделать вывод о том, что когда нормированный на объем ущерб (убыток) z_j зависит в большей степени от размера убытка, то использование гамма-распределения для данного случая является наилучшим вариантом.

В дайльнейшем, на основе таких данных у страховой компании есть возможность четко представлять характер рисков, а это в свою очередь для будущих клиентов позволит адекватно произвести страховой расчет.

В **заключении** приведены результаты бакалаврской работы.

Основные результаты

1. Рассмотрены две группы моделей, выявлены их преимущества и недостатки, выделен общий принцип их построения.
2. Построена модель зависимости дисперсии от объема в случае однородных и неоднородных рисков;

3. Исследованно гамма-распределение, получены плотности при различных параметрах, написан алгоритм построения их графиков на языке программирования Python 3;
4. Построена модель совокупного убытка портфеля рисков в однородном и неоднородном случае.
5. Разработан численный метод, оценивающий параметры распределения;
6. Создан код на языке программирования C++, позволяющий смоделировать или идентифицировать совокупный страховой убыток, оценить параметры распределения с использованием приведенного численного метода.
7. На языке Python 3 построен график полученной модели и сделан вывод о том, насколько подходит гамма распределение при моделировании совокупного убытка.