

УДК: 519.2

Научный руководитель: Албогачиева Мадина Магометовна

Физико-математический факультет

ФГБОУ ВО «Ингушский государственный университет»

Амерханов Магомед И., Мальсагов Амир М.

студенты 3-го курса специальности, «Математика»,

физико-математического факультета

ВЕРОЯТНОСТНЫЕ МОДЕЛИ В СТРАХОВАНИИ ИМУЩЕСТВА. АНАЛИЗ И МОДЕЛИРОВАНИЕ УБЫТКОВ С ПОМОЩЬЮ СОСТАВНЫХ РАСПРЕДЕЛЕНИЙ

***Аннотация:** В статье рассматривается применение теории составных (сложных) распределений к моделированию совокупного убытка в имущественном страховании. Дано формальное определение составного распределения $S = X_1, X_2, \dots, X_n$, где N — случайное число убытков за период наблюдения, а $X_1, \dots, X_n$ — независимые одинаково распределённые размеры отдельных убытков. Выведены общие формулы для производящей функции, математического ожидания и дисперсии суммы S , справедливые при произвольном распределении числа убытков. Рассмотрены три классических распределения частоты убытков — пуассоновское, биномиальное и отрицательное биномиальное — и получены соответствующие частные формулы для моментов S . На данных гипотетического портфеля договоров страхования имущества подобраны распределения частоты (отрицательное биномиальное) и тяжести (логнормальное) убытков, оценены их параметры методами моментов и максимального правдоподобия, проверено согласие моделей с данными с помощью критериев хи-квадрат и Колмогорова–Смирнова. Методом статистических испытаний (Монте-Карло) рассчитаны основные рисковые характеристики совокупного убытка: математическое ожидание, дисперсия, коэффициент вариации, квантили (VaR) и условная стоимость под риском ($CVaR$).*

***Ключевые слова:** составное распределение, страхование имущества, частота убытков, тяжесть убытков, распределение Пуассона, отрицательное биномиальное распределение, логнормальное распределение, рисковые характеристики, VaR , $CVaR$, актуарная математика*

Scientific adviser: Albogachieva Madina Magometovna

Faculty of Physics and Mathematics

FSBEI HE "Ingush State University"

Amerhanov Magomed I., Malsagov Amir M.

3rd year students of the specialty "Mathematics",

Faculty of Physics and Mathematics

Probabilistic Models in Property Insurance. Analysis and Modeling of Losses Using Compound Distributions

Abstract: *The article examines the application of compound (aggregate) distribution theory to modeling the total loss of an insurer in property insurance. A formal definition of the compound distribution $S = X_1, X_2, \dots, X_n$ is given, where N is the random number of claims during the observation period and $X_1, \dots, X_n$ are independent and identically distributed individual claim amounts. General formulas for the probability/moment generating function, the expected value and the variance of S are derived, valid for an arbitrary distribution of the claim count. Three classical claim-count distributions are considered — Poisson, binomial and negative binomial — and the corresponding closed-form expressions for the moments of S are obtained. Using data from a hypothetical property-insurance portfolio, frequency (negative binomial) and severity (lognormal) distributions are fitted, their parameters estimated by the method of moments and maximum likelihood, and goodness of fit checked with chi-square and Kolmogorov–Smirnov tests. Monte Carlo simulation is used to compute the main risk characteristics of the aggregate loss: mean, variance, coefficient of variation, Value-at-Risk (VaR) and Conditional VaR (CVaR).*

Keywords: *compound distribution, property insurance, claim frequency, claim severity, Poisson distribution, negative binomial distribution, lognormal distribution, risk measures, VaR, CVaR, actuarial mathematics*

Введение

В современной актуарной практике одной из центральных задач является количественная оценка совокупного (агрегированного) убытка страховщика по портфелю договоров за определённый период. В имущественном страховании совокупный убыток формируется как результат случайного числа страховых событий, каждое из которых порождает случайный по величине ущерб. Такая

структура риска естественно описывается аппаратом составных (сложных) распределений вероятностей.

Актуальность исследования обусловлена тем, что корректная оценка распределения совокупного убытка лежит в основе тарификации, формирования страховых резервов и расчёта нормативов платёжеспособности страховщика. Использование неадекватной модели частоты или тяжести убытков приводит к недооценке риска и, как следствие, к финансовой неустойчивости страховой компании.

Цель работы — построение и анализ вероятностной модели совокупного убытка в имущественном страховании на основе составных распределений, включая вывод общих формул для его моментов и практическую оценку параметров модели по данным.

Задачи исследования:

1. Дать формальное определение составного распределения и вывести формулы для его производящей функции, математического ожидания и дисперсии;
2. Рассмотреть частные случаи распределения числа убытков;
3. Подобрать распределения частоты и тяжести убытков по данным и оценить их параметры;
4. Рассчитать основные рисковые характеристики совокупного убытка S .

Материалы и методы

Материалом исследования послужили данные гипотетического портфеля из 5000 договоров страхования имущества за один период наблюдения (год), сгенерированные на основе типичных для имущественного страхования распределений частоты и тяжести убытков. Такой подход, часто применяемый в учебных и методических целях, позволяет проверить работоспособность методов оценивания в условиях, когда истинные параметры генеральной совокупности известны.

Методологической основой являются положения теории риска и теории вероятностей: метод моментов и метод максимального правдоподобия — для оценки параметров распределений; критерии согласия хи-квадрат и Колмогорова–Смирнова — для проверки адекватности моделей; метод статистических испытаний (Монте-Карло) — для расчёта рисковых характеристик совокупного убытка.

1. Результаты и обсуждение

1.1 Составное распределение убытков: определение

Пусть страховой портфель наблюдается в течение фиксированного периода. Обозначим через N случайное число убытков (страховых случаев), произошедших за этот период, а через $X_1, X_2, \dots, X_n$ — размеры (тяжесть) отдельных убытков. Совокупный убыток портфеля определяется как случайная сумма:

$$S = X_1 + X_2 + \dots + X_N \quad (0)$$

При этом $N = 0$ означает $S = 0$. Предполагается:

- случайные величины $X_1, X_2, \dots$ независимы и одинаково распределены;
- величина N не зависит от величин $X_1, X_2, \dots$;
- при $N = 0$ полагается $S = 0$.

Распределение случайной величины S , построенной таким образом, называется составным (сложным) распределением.

1.2 Производящая функция и моменты составного распределения S

Обозначим через $GN(z) = E[z^N]$ производящую функцию распределения числа убытков N , а через $MX(t) = E[etX]$ — производящую функцию моментов размера убытка X . Тогда производящая функция моментов совокупного убытка S имеет вид:

$$M_S(t) = E[e^{tS}] = E[E(e^{tS} | N)] = E[(M_X(t))^N] = G_N(M_X(t)) \quad (1)$$

Формула (1) — ключевой результат теории составных распределений: она сводит распределение суммы случайного числа случайных слагаемых к композиции производящих функций N и X .

Математическое ожидание и дисперсия S находятся по формулам полного математического ожидания и полной дисперсии. По формуле полного математического ожидания:

$$E[S] = E[E(S | N)] = E[N \cdot E(X)] = E[N] \cdot E[X] \quad (2)$$

По формуле полной дисперсии:

$$Var[S] = E[Var(S | N)] + Var[E(S | N)] = E[N] \cdot Var(X) + Var(N) \cdot (E[X])^2 \quad (3)$$

Формулы (2)–(3) справедливы для произвольного распределения N при условии независимости N и X_i и существования соответствующих моментов.

1.3 Частные случаи распределения числа убытков N

а) Пуассоновское распределение. Для $N \sim \text{Pois}(\lambda)$:

$$P(N = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad G_N(z) = e^{\lambda(z-1)}, \quad E[N] = \text{Var}[N] = \lambda$$

Подставляя в (2)–(3) и учитывая, что $\text{Var}(X) + (E[X])^2 = E[X^2]$, получаем известную формулу составного пуассоновского распределения:

$$E[S] = \lambda \cdot E[X], \quad \text{Var}[S] = \lambda \cdot E[X^2] \quad (4)$$

б) Биномиальное распределение. Для $N \sim \text{Bin}(n, p)$:

$$P(N = k) = C_n^k p^k (1-p)^{n-k}, \quad G_N(z) = (1-p + pz)^n, \quad E[N] = np, \quad \text{Var}[N] = np(1-p)$$

$$E[S] = np \cdot E[X], \quad \text{Var}[S] = np \cdot [\text{Var}(X) + (1-p) \cdot (E[X])^2] \quad (5)$$

в) Отрицательное биномиальное распределение. Для $N \sim \text{NB}(r, p)$ (число «неудач» до r -го «успеха»):

$$P(N = k) = C_{k+r-1}^k p^r (1-p)^k, \quad G_N(z) = \left[\frac{p}{1 - (1-p)z} \right]^r$$

$$E[N] = \frac{r(1-p)}{p}, \quad \text{Var}[N] = \frac{r(1-p)}{p^2} = E[N]/p$$

$$E[S] = \frac{r(1-p)}{p} \cdot E[X], \quad \text{Var}[S] = \frac{r(1-p)}{p} \cdot \text{Var}(X) + \frac{r(1-p)}{p^2} \cdot (E[X])^2 \quad (6)$$

Поскольку $p \in (0,1)$, из $\text{Var}[N] = E[N]/p$ следует $\text{Var}[N] > E[N]$ — отрицательное биномиальное распределение допускает «избыточную дисперсию» (overdispersion) относительно пуассоновского, что часто наблюдается в реальных портфелях имущественного страхования, где число убытков варьируется от полиса к полису сильнее, чем предсказывает модель Пуассона.

1.4 Данные для анализа частоты и тяжести убытков

Для практической иллюстрации сформирована выборка по портфелю из 5000 договоров страхования имущества за один год наблюдения. Основные описательные статистики выборки приведены в таблице 1.

Показатель	Число убытков N	Размер убытка X, у.е.
Среднее	1,463	16 066
Стандартное отклонение	1,537	28 415
Медиана	1,000	7 968

Показатель	Число убытков N	Размер убытка X, у.е.
Индекс дисперсии Var/Mean	1,616	—
Объём выборки	5 000 полисов	7 313 убытков

Таблица 1 — Описательные статистики выборки

1.5 Подбор и оценка параметров распределения частоты убытков N

Оценка параметра пуассоновского распределения методом моментов: $\hat{\lambda} = \bar{x} = 1,463$.

Оценка параметров отрицательного биномиального распределения методом моментов (из соотношений $E[N] = \bar{x}$, $Var[N] = s^2$):

$$\hat{r} = \frac{\bar{x}^2}{s^2 - \bar{x}} = 2,375, \quad \hat{p} = \frac{\hat{r}}{\hat{r} + \bar{x}} = 0,619 \quad (7)$$

Проверка согласия распределений с эмпирическими данными по критерию хи-квадрат (группировка по $k = 0, 1, \dots, 7$ и «8 и более») дала следующие результаты (таблица 2).

Распределение	χ^2 (набл.)	Число степеней свободы	$\chi^2_{кр}$ ($\alpha=0,05$)	Вывод
Пуассон ($\hat{\lambda}=1,463$)	1555,9	7	14,1	отклоняется
Отриц. биномиальное ($\hat{r}=2,375$; $\hat{p}=0,619$)	7,5	6	12,6	не отклоняется

Таблица 2 — Критерий согласия χ^2 для распределений частоты убытков

Индекс дисперсии выборки ($Var/Mean \approx 1,62 > 1$) указывает на избыточную дисперсию числа убытков, что делает пуассоновскую модель неадекватной (гипотеза о согласии отклоняется на уровне значимости 0,05), тогда как отрицательное биномиальное распределение хорошо согласуется с данными.

1.6 Подбор и оценка параметров распределения тяжести убытков X

Рассмотрены два стандартных для актуарной практики распределения тяжести убытков — логнормальное и гамма-распределение. Параметры оценивались методом максимального правдоподобия.

Логнормальное распределение: $\hat{\mu} = 9,003$, $\hat{\sigma} = 1,151$, откуда $E[X] = 15\,755$, $Var[X] \approx 6,85 \cdot 10^8$.

Гамма-распределение: $\hat{\alpha}$ (форма) = 0,862, $\hat{\theta}$ (масштаб) = 18 629, откуда $E[X] = 16\,066$, $Var[X] \approx 2,99 \cdot 10^8$.

Проверка согласия по критерию Колмогорова–Смирнова (таблица 3) показывает, что логнормальное распределение хорошо описывает эмпирические данные ($p\text{-value} \gg 0,05$), тогда как гамма-распределение отклоняется — его более лёгкий правый хвост не воспроизводит наблюдаемую тяжесть крупных убытков.

Распределение	Параметры	Статистика D (K-S)	p-value	Вывод
Логнормальное	$\hat{\mu}=9,003$; $\hat{\sigma}=1,151$	0,008	0,706	не отклоняется
Гамма	$\hat{\alpha}=0,862$; $\hat{\theta}=18\,629$	0,093	<0,001	отклоняется

Таблица 3 — Критерий согласия Колмогорова–Смирнова для распределений тяжести убытков

1.7 Расчёт рисковых характеристик совокупного убытка S

На основании подобранных распределений — отрицательного биномиального для N и логнормального для X — по формулам (2)–(3) рассчитаны теоретические характеристики совокупного убытка:

$$E[N] = 1,463; \quad Var[N] = 2,364; \quad E[X] = 15\,755; \quad Var[X] \approx 6,85 \cdot 10^8$$

$$E[S] = E[N] \cdot E[X] \approx 23\,044 \text{ у. е.} \quad (8)$$

$$Var[S] = E[N] \cdot Var[X] + Var[N] \cdot (E[X])^2 \approx 1,59 \cdot 10^9, \quad \sigma(S) \approx 39\,850 \text{ у. е.} \quad (9)$$

Коэффициент вариации совокупного убытка $CV(S) = \sigma(S)/E[S] \approx 1,73$, что свидетельствует о значительной изменчивости и правосторонней асимметрии распределения S — типичной черте страховых портфелей.

Поскольку аналитическое выражение функции распределения S при отрицательном биномиальном N и логнормальном X не выражается в элементарных функциях, для расчёта квантилей использован метод Монте-Карло (200 000 траекторий портфеля). Результаты симуляции приведены в таблице 4.

Показатель	Значение, у.е.
$E[S]$ (симуляция)	22 985
$\sigma(S)$ (симуляция)	40 009
Асимметрия S	5,70
VaR 90%	63 197
VaR 95%	93 323
VaR 99%	180 415
VaR 99,5%	227 152
CVaR 95%	151 206
CVaR 99%	262 237
Рисковая надбавка ($VaR99\% - E[S]$)	157 371

Таблица 4 — Рисковые характеристики совокупного убытка S (Монте-Карло, 200 000 симуляций)

Результаты симуляции хорошо согласуются с теоретическими значениями (8)–(9), что подтверждает корректность выведенных формул. Высокая асимметрия ($\approx 5,7$) показывает, что распределение S существенно отличается от нормального: нормальное приближение дало бы оценку VaR 99% на уровне $E[S] + 2,33 \cdot \sigma(S) \approx 115\,894$ у.е., то есть занизило бы истинное значение (180 415 у.е., по Монте-Карло) примерно на треть. Это подчёркивает необходимость использования эмпирических или иных методов, учитывающих асимметрию распределения S, при расчёте капитала под риском.

2. Обсуждение

Результаты подтверждают, что выбор распределения частоты убытков существенно влияет на оценку риска: пуассоновская модель, не допускающая избыточной дисперсии, статистически значимо отклоняется данными ($\chi^2 = 1555,9$ при критическом значении 14,1), тогда как отрицательное биномиальное распределение адекватно описывает наблюдаемую вариацию числа убытков между полисами. Использование пуассоновской модели в подобных условиях привело бы к занижению дисперсии совокупного убытка и, соответственно, к недостаточному запасу платёжеспособности.

Логнормальное распределение оказалось предпочтительным для описания тяжести убытков по сравнению с гамма-распределением: критерий

Колмогорова–Смирнова не отклоняет гипотезу о логнормальности, тогда как для гамма-распределения гипотеза отклоняется на высоком уровне значимости. Это согласуется с общей практикой актуарного моделирования имущественных рисков, где тяжесть убытков часто имеет более «тяжёлый» правый хвост, чем допускает гамма-распределение с малым параметром формы.

Совместное действие случайной частоты и тяжёлого хвоста тяжести убытков приводит к сильной правосторонней асимметрии совокупного убытка S (асимметрия $\approx 5,7$), что делает нормальное приближение непригодным для расчёта капитала под риском. В связи с этим для практических актуарных расчётов рекомендуется использовать эмпирические квантили, полученные методом Монте-Карло, либо аналитические аппроксимации, учитывающие асимметрию (например, трансформированное гамма- или NP2-приближение).

Заключение

В ходе проведённого исследования построена и проанализирована вероятностная модель совокупного убытка в имущественном страховании на основе аппарата составных распределений. Установлено, что предложенный подход позволяет формально связать характеристики частоты и тяжести убытков с рисковыми характеристиками портфеля в целом.

Основные выводы:

1. Составное распределение $S = X_1 + X_2 + \dots + X_N$ является универсальным инструментом описания совокупного убытка: производящая функция, математическое ожидание и дисперсия S выражаются через соответствующие характеристики N и X по формулам (1)–(3) независимо от конкретного вида распределения числа убытков.

2. Выбор распределения частоты убытков существенно влияет на точность модели. Данные портфеля продемонстрировали избыточную дисперсию ($\text{Var}/\text{Mean} \approx 1,62$), из-за чего пуассоновская модель была статистически отклонена, а отрицательное биномиальное распределение показало хорошее согласие с данными.

3. Тяжесть убытков в рассмотренном портфеле адекватно описывается логнормальным распределением; гамма-распределение с оценёнными параметрами не проходит проверку по критерию Колмогорова–Смирнова из-за более лёгкого правого хвоста.

4. Совокупный убыток S характеризуется значительной правосторонней асимметрией ($\approx 5,7$), вследствие чего величины VaR и CVaR, рассчитанные

методом Монте-Карло, существенно превышают оценки, которые дало бы нормальное приближение — расхождение между VaR 99% и математическим ожиданием составило около 157 тыс. у.е., то есть почти 7,8 математического ожидания убытка.

5. Перспективы дальнейших исследований связаны с учётом зависимости между частотой и тяжестью убытков (например, с помощью копул), применением распределений с более тяжёлыми хвостами (Парето, Вейбулла) для моделирования крупных убытков, а также с оценкой влияния перестрахования на распределение совокупного убытка.

Список литературы

1. Фалин Г.И. Математические основы теории страхования жизни и пенсионных схем. — М.: Анкил, 2002.
2. Корнилов И.А. Основы страховой математики. — М.: ЮНИТИ-ДАНА, 2004.
3. Вентцель Е.С. Теория вероятностей. — М.: КНОРУС, 2018.
4. Кремер Н.Ш. Теория вероятностей и математическая статистика. — М.: ЮНИТИ-ДАНА, 2019.
5. Bowers N.L., Gerber H.U., Hickman J.C., Jones D.A., Nesbitt C.J. Actuarial Mathematics. — Schaumburg: Society of Actuaries, 1997.
6. Klugman S.A., Panjer H.H., Willmot G.E. Loss Models: From Data to Decisions. — Hoboken: Wiley, 2019.
7. Kaas R., Goovaerts M., Dhaene J., Denuit M. Modern Actuarial Risk Theory: Using R. — Berlin: Springer, 2008.