STUDIA UNIVERSITATIS MOLDAVIAE, 2017, nr.7(107)
Seria "Științe exacte și economice" *ISSN 1857-2073* *ISSN online 2345-1033* *p.72-77*

CZU: 336:368.8

РАСПРЕДЕЛЕНИЕ ПОТЕРЬ СТРАХОВЫХ ОПЕРАЦИЙ

Юлия КАПРИЯН, Андрей МУЛИК, Кирилл ГАВЛИЦКИЙ

Молдавский государственный университет

Исследуются проблемы распределения финансовых потерь при осуществлении страховых операций. В данном контексте рассматриваются аналитические методы расчёта необходимого распределения финансовых потерь. В частности, изучены особенности использования логарифмического распределения, экспоненциального распределения и распределения Парето для расчёта финансовых потерь при проведении тех или иных страховых операций.

Ключевые слова: *страховые операции, финансовые претензии, страховые политики, распределение финансовых потерь, страховые данные, эмпирическая функция распределения, аналитические методы, логарифмическое распределение, экспоненциальное распределение, распределение Парето.*

DISTRIBUTION OF LOSSES OF INSURANCE OPERATIONS

In the presented material, the problems of distribution of financial losses in the course of insurance operations are investigated. In this context, analytical methods for calculating the necessary distribution of financial losses are considered. In particular, the features of using the logarithmic distribution, exponential distribution and Pareto distribution for the calculation of financial losses in the course of conducting certain insurance operations are studied.

Keywords: *insurance operations, financial claims, insurance policies, financial loss distribution, insurance data, empirical function of distributions, analytical methods, logarithmic distribution, exponential distribution, Pareto distribution.*

Введение

Определение потерь страховых операций является довольно непростой задачей. Обычно страховщики хранят файлы данных, содержащих подробную информацию о страховой политике и возмещениях, используемую для учета и расчета страховых ставок. Однако распределение размеров финансовых претензий и другие данные, необходимые для проведения теоретических исследований с точки зрения риска, могут быть получены обычно только после дополнительной предварительной обработки исходных данных. Более того, статистика финансовых претензий зачастую ограничена.

Файлы данных, содержащие подробную информацию о некоторых страховых политиках и финансовых претензиях, могут отсутствовать. Может также возникнуть ситуация, когда предварительные или ретроспективные данные недоступны вообще, например, когда вводится новый вид страхования или когда застрахованы очень большие или особые виды рисков. В этом случае распределение должно основываться на знании подобных рисков или на экстраполяции меньших рисков.

Существуют три основных подхода к определению распределения финансовых потерь, которые наиболее часто используются в специализированной литературе по актуариату. Данные подходы сводятся к поиску аналитического выражения, которое хорошо подходит к наблюдаемым данным и легко обрабатывается. В данном случае особый интерес представляют основные характеристики и вопросы оценки наиболее популярных и полезных распределений потерь.

Исходя их этого, следует обращать внимание на то, что иногда может оказаться полезным разбить диапазон распределения размера претензии на интервалы, для которых используются различные методы. Например, требования к малым и средним размерам могут быть описаны с помощью эмпирического распределения размеров заявки, в то время как большие претензии, для которых нехватка данных исключает использование эмпирического подхода, – посредством аналитического распределения потерь.

В некоторых приложениях точная форма распределения потерь не требуется. Можно использовать подход, основанный на моменте, состоящем в оценке только самых низких характеристик (моментов) распределения, таких как среднее и дисперсия. Однако следует иметь в виду, что даже самые низкие три или четыре момента не полностью определяют форму распределения, и поэтому соответствие полученным данным может быть неудовлетворительным. Более подробная информация об основанных

STUDIA UNIVERSITATIS MOLDAVIAE, 2017, nr.7(107)
Seria "Ştiinţe exacte şi economice" ISSN 1857-2073 ISSN online 2345-1033 p.72-77

на моменте подходах может быть найдена, например, в исследованиях таких ученых, как Дайкин, Пентикайнен и Песонен (1994)*.

При большом наборе распределений, для выбора необходимо сузить его до одной модели и уникальной оценки параметров. Тип распределения объективных потерь можно легко выбрать, сравнивая формы эмпирических и теоретических средних избыточных функций. Подтверждение пригодности можно проверить, построив соответствующие функции ограниченного ожидаемого значения. Наконец, гипотеза о том, что моделированное случайное событие устанавливается определенным распределением потерь, может быть проверена статистически.

Данные инструменты могут быть применены для моделирования реальных страховых данных. Например, в международном масштабе анализ может проводиться для двух наборов данных: (i) набор данных PCS (Property Claim Services), покрывающих убытки, вызванные катастрофическими событиями в США, которые произошли в период с 1990 по 1999 год или с 2006 по 2015 годы, и (ii) набора данных о среднеевропейских потерях от чрезвычайных событий в периоды с 1990 по 2015 годы.

Естественной оценкой распределения потерь является наблюдаемое (эмпирическое) распределение размеров претензий. Однако если в течение периода наблюдения были внесены изменения в денежные значения, следует использовать данные с исправленной инфляцией. Для выборки наблюдений {xi,…...xn} эмпирическая функция распределения (ЭФР) определяется как

$$F_n(x) = \frac{1}{n}\#\{i : x_i \leq x\}.$$

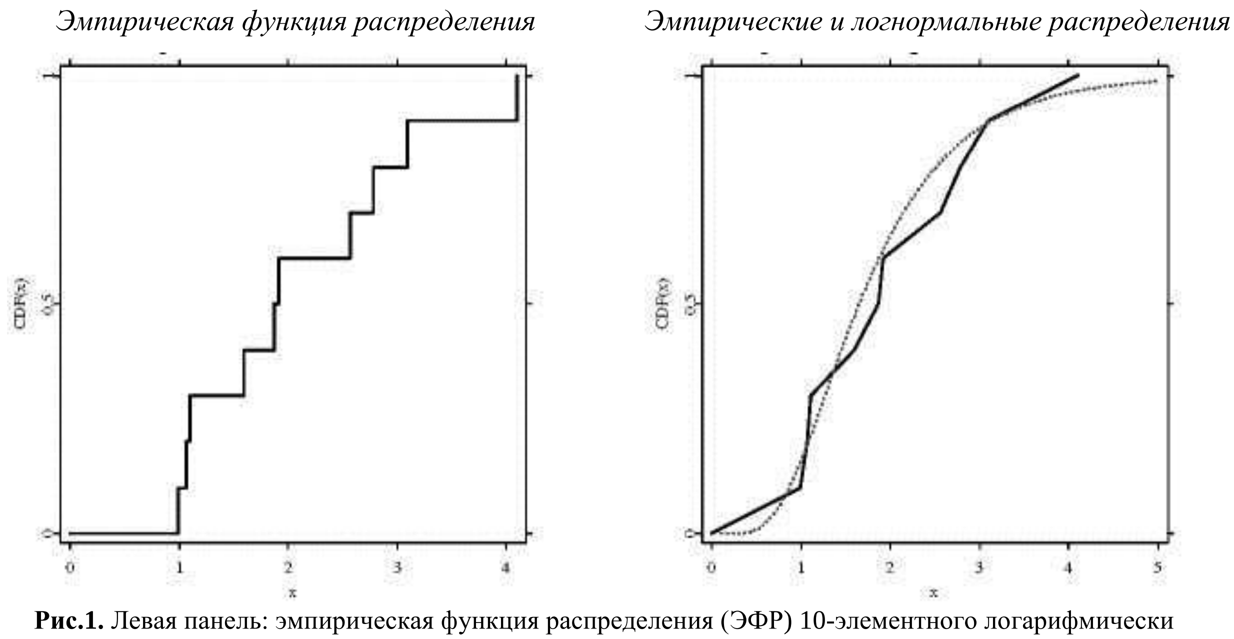

Рис.1. Левая панель: эмпирическая функция распределения (ЭФР) 10-элементного логарифмически нормально распределенного образца с параметрами μ = 0,5 и σ = 0,5, см. Правая панель: приближение ЭФР непрерывной частичнолинейной функцией (черная сплошная линия) и теоретическая функция распределения (пунктирная линия).

Источник: Разработано авторами на основании [1].

Данный рисунок представляет собой частичнопостоянную функцию со скачками размера 1/n в точках Xi. Очень часто, особенно если образец достаточно большой, ЭФР аппроксимируется непрерывной частичнолинейной функцией с «точками перехода», связанными линейными функциями.

Принцип эмпирической функции распределения подходит только тогда, когда имеется достаточно большой объем входящих данных. Это редко бывает для хвоста распределения, особенно в ситуациях, когда возможны исключительно большие финансовые претензии. Часто целесообразно разделить

* Cížek P., Härdle W., Weron R. *Statistical Tools for Finance and Insurance.* Springer Berlin Heidelberg New York. ISBN 3-540-22189-1

диапазон соответствующих значений претензий на две части, обрабатывая размеры претензий до определенного предела на дискретной основе, тогда как хвост заменяется аналитической эмпирической функцией распределения.

Аналитические методы

Зачастую желательно найти явное аналитическое выражение для распределения потерь. Это особенно характерно, если статистика претензий слишком скудна для использования эмпирического подхода. Следует, однако, подчеркнуть, что многие стандартные статистические данные, такие как распределение Гаусса, не подходят для подгонки распределения размера претензии. Основной причиной этого является сильно искаженная нагрузка распределений потерь. Логарифмическое распределение, Парето, Бур, Weibull и Гамма распределения, являются типичными для распределения размеров претензий, рассматриваемых в данных исследованиях.

Логарифмическое распределение

Рассмотрим случайную величину X, имеющую нормальное распределение с плотностью

$$f_N(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\}, \qquad -\infty < x < \infty.$$

Пусть $Y = e^X$, так что $X = \log Y$. Тогда функция плотности вероятности Y задается формулой

$$f(y) = f_N(\log y)\frac{1}{y} = \frac{1}{\sqrt{2\pi}\sigma y} \exp\left\{-\frac{1}{2}\frac{(\log y-\mu)^2}{\sigma^2}\right\}, \qquad y > 0,$$

где б > 0 – шкала, -∞ < μ < ∞ – параметр местоположения. Распределение Y называется логнормальным, однако иногда его называют законом Кобба-Дугласа, особенно применительно к эконометрическим данным. Логарифмическая эмпирическая функция распределения задается формулой

$$F(y) = \Phi\left(\frac{\log y - \mu}{\sigma}\right), \qquad y > 0,$$

где Ф(-) является стандартной нормой (со средним значением 0 и дисперсией l) функции распределения. Необработанный момент mk логарифмически нормального значения можно легко получить, используя результаты для обычных случайных величин:

$$m_k = \mathrm{E}\left(Y^k\right) = \mathrm{E}\left(e^{kX}\right) = M_X(k) = \exp\left(\mu k + \frac{\sigma^2 k^2}{2}\right),$$

где Mx (k) – моментная функция нормального распределения. В частности, среднее и дисперсия –

$$\begin{aligned} \mathrm{E}(X) &= \exp\left(\mu + \frac{\sigma^2}{2}\right), \\ \mathrm{Var}(X) &= \left\{\exp\left(\sigma^2\right) - 1\right\}\exp\left(2\mu + \sigma^2\right), \end{aligned}$$

соответственно. Для обоих методов оценки стандартных параметров известны в закрытой форме. Метод оценок моментов может быть представлен следующей формулой:

$$\begin{aligned} \hat{\mu} &= 2\log\left(\frac{1}{n}\sum_{i=1}^n x_i\right) - \frac{1}{2}\log\left(\frac{1}{n}\sum_{i=1}^n x_i^2\right), \\ \hat{\sigma}^2 &= \log\left(\frac{1}{n}\sum_{i=1}^n x_i^2\right) - 2\log\left(\frac{1}{n}\sum_{i=1}^n x_i\right), \end{aligned}$$

тогда как оценки максимального правдоподобия

$$\begin{aligned} \hat{\mu} &= \frac{1}{n}\sum_{i=1}^n \log(x_i), \\ \hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^n \left\{\log(x_i) - \hat{\mu}\right\}^2. \end{aligned}$$

STUDIA UNIVERSITATIS MOLDAVIAE, 2017, nr.7(107)
Seria "Științe exacte și economice" ISSN 1857-2073 ISSN online 2345-1033 p.72-77

Следовательно, генерация логарифмически нормального изменения достаточно проста, так как необходимо использовать только показатель нормального отклонения.

Логарифмически нормальное распределение применимо наилучшим образом при моделировании размеров страховых претензий. Оно правильно искажено, имеет густой хвост и хорошо подходит для многих ситуаций; при малой б похоже на нормальное распределение (см. левую панель на рисунке 2), хотя подобное не всегда желательно. Оно бесконечно делится и закрыто по масштабам и силовым преобразованиям. Однако существуют также некоторые недостатки. В частности, преобразование Лапласа не имеет замкнутого представления формы, а функция генерации момента не существует.

Экспоненциальное распределение

Рассмотрим случайную величину со значениями плотности и распределения соответственно:

$$f(x) = \beta e^{-\beta x}, \qquad x > 0,$$
$$F(x) = 1 - e^{-\beta x}, \qquad x > 0.$$

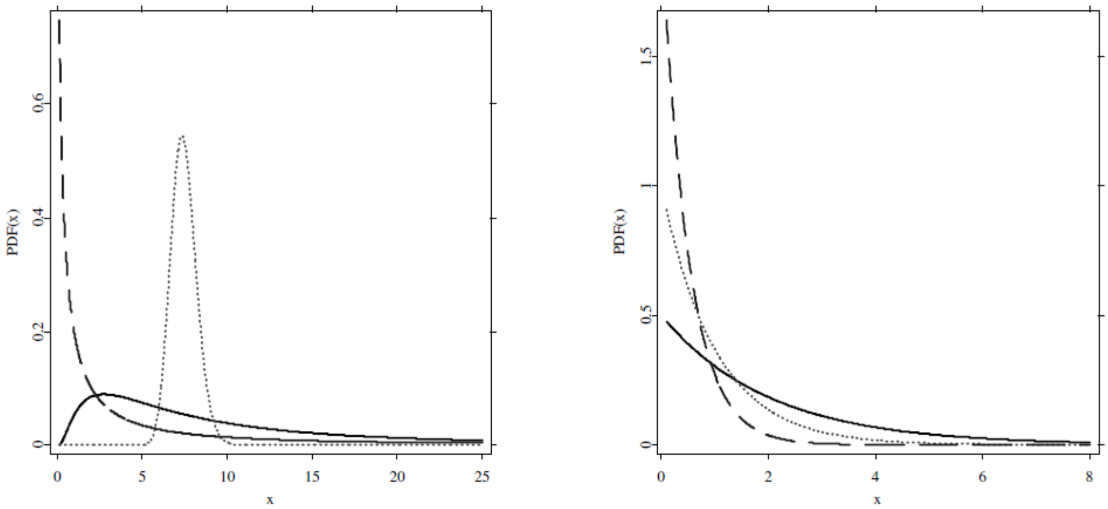

Рис.2. Левая панель: функции логарифмически нормальной плотности с параметрами μ = 2 и б = 1 (черная сплошная линия), μ = 2 и б = 0,1 (штриховая пунктирная линия), p = 0,5, a = 2 (точечная пунктирная линия). Правая панель: Экспоненциальные функции логарифмически нормальной плотности с параметром β = 0,5 (черная сплошная линия), β = 1 (штрихпунктирная линия) и β = 5 (точечная пунктирная линия).

Это распределение называется экспоненциальным распределением с параметром (или интенсивностью) β > 0. Преобразование Лапласа трансформируется в

$$L(t) \stackrel{\text{def}}{=} \int_0^\infty e^{-tx} f(x) dx = \frac{\beta}{\beta + t}, \qquad t > -\beta.$$

Общую формулу для k-ого необработанного момента можно представить в следующем виде:

$$L(t) \stackrel{\text{def}}{=} \int_0^\infty e^{-tx} f(x) dx = \frac{\beta}{\beta + t}, \qquad t > -\beta.$$

Таким образом, среднее и дисперсия составляют, соответственно, β^{-1} и β^{-2}. Оценщик максимального правдоподобия (равного методу оценки моментов) для β задается формулой

$$\hat{\beta} = \frac{1}{\hat{m}_1},$$

где $\hat{m}_k = \frac{1}{n} \sum_{i=1}^{n} x_i^k$ является образцом К-го необработанного момента.

Чтобы создать экспоненциальную случайную величину X с интенсивностью β, можно использовать метод обратного преобразования (L'Ecuyer, 2004; Ross, 2002). Метод состоит в том, что случайное число U равномерно распределено на интервале (0,1), полагая X = F^{-1} (U), где F^{-1} (x) = - log (1 - x) является обратным экспоненциальной ЭФР. Действительно, мы можем представить X = - log U, поскольку 1 - U имеет такое же распределение, что и U.

Экспоненциальное распределение имеет много интересных особенностей. Например, оно обладает следующим свойством: P (X> x + y\X> y) = P (X> x). Оно также проявляется как время возникновения событий в пуассоновском процессе. N-й корень преобразования Лапласа равен

$$L(t) = \left(\frac{\beta}{\beta + t}\right)^{\frac{1}{n}},$$

являясь преобразованием гамма-вариации Лапласа. Таким образом, экспоненциальное распределение делится бесконечно.

Экспоненциальное распределение часто используется при разработке моделей страховых рисков. Его полезность в значительной степени связана с его многочисленными и разнообразными прикладными математическими свойствами. Однако недостатком экспоненциального распределения является то, что его плотность монотонно уменьшается (см. правую панель на рис.2), что может оказаться неприемлемым в некоторых практических ситуациях.

Распределение Парето

Предположим, что переменная X имеет (условно β) экспоненциальное распределение со средним β^{-1}. Далее предположим, что β имеет гамма-распределение. Безусловное распределение X является помесью и называется распределением Парето. Более того, можно представить, что если X – экспоненциальная случайная величина, а Y – гамма-случайная величина, то X / Y – случайная величина Парето.

Функции плотности и распределения вариации Парето представляются следующим выражением:

$$\begin{aligned} f(x) &= \frac{\alpha\lambda^\alpha}{(\lambda + x)^{\alpha+1}}, \qquad x > 0, \\ F(x) &= 1 - \left(\frac{\lambda}{\lambda + x}\right)^\alpha, \qquad x > 0, \end{aligned}$$

соответственно. Явствует, что параметр формы α и масштабный параметр λ – оба положительны. К-ый необработанный момент,

$$m_k = \lambda^k k! \frac{\Gamma(\alpha - k)}{\Gamma(\alpha)},$$

существует только при к < α. В приведенной выше формуле

$$\Gamma(a) \stackrel{\text{def}}{=} \int_0^\infty y^{a-1} e^{-y} dy$$

стандартная гамма-функция. Среднее значение и отклонение, соответственно,

$$\begin{aligned} \mathrm{E}(X) &= \frac{\lambda}{\alpha - 1}, \\ \mathrm{Var}(X) &= \frac{\alpha\lambda^2}{(\alpha - 1)^2(\alpha - 2)}, \end{aligned}$$

Заметим, что среднее существует только при α > 1 и дисперсия только при α > 2. Следовательно, распределение Парето имеет очень толстые (или тяжелые) хвосты, см. Рис.3. Метод оценок моментов представлен следующими формулами:

$$\begin{aligned} \hat{\alpha} &= 2\frac{\hat{m}_2 - \hat{m}_1^2}{\hat{m}_2 - 2\hat{m}_1^2}, \\ \hat{\lambda} &= \frac{\hat{m}_1\hat{m}_2}{\hat{m}_2 - 2\hat{m}_1^2}, \end{aligned}$$

где, как и раньше, mk – образец К-го необработанного момента. Заметим, что оценки хорошо определены только при $\hat{m}_2 - 2\hat{m}_1^2 > 0$.

STUDIA UNIVERSITATIS MOLDAVIAE, 2017, nr.7(107)
Seria "Științe exacte și economice" ISSN 1857-2073 ISSN online 2345-1033 p.72-77

К сожалению, для оценок максимального правдоподобия нет выражений замкнутой формы, и их можно оценивать только численно.

Как и для многих других распределений, симуляция Парето вариации X может быть проведена методом обратного преобразования. Обратный метод для ЭФР имеет простую аналитическую форму F^{-1} (x) = λ $\{(1-\mathbf{x})^{-\frac{1}{\alpha}}$- 1}. Следовательно, можно представить множество X = λ ($U^{-\frac{1}{\alpha}}$ - 1), где U равномерно распределено на единичном интервале. Следует быть осторожными, если значение достаточно большое, но очень близкое к единице. Теоретическое среднее существует, но правый хвост очень тяжелый. Среднее значение выборки будет, как правило, значительно ниже E (X).

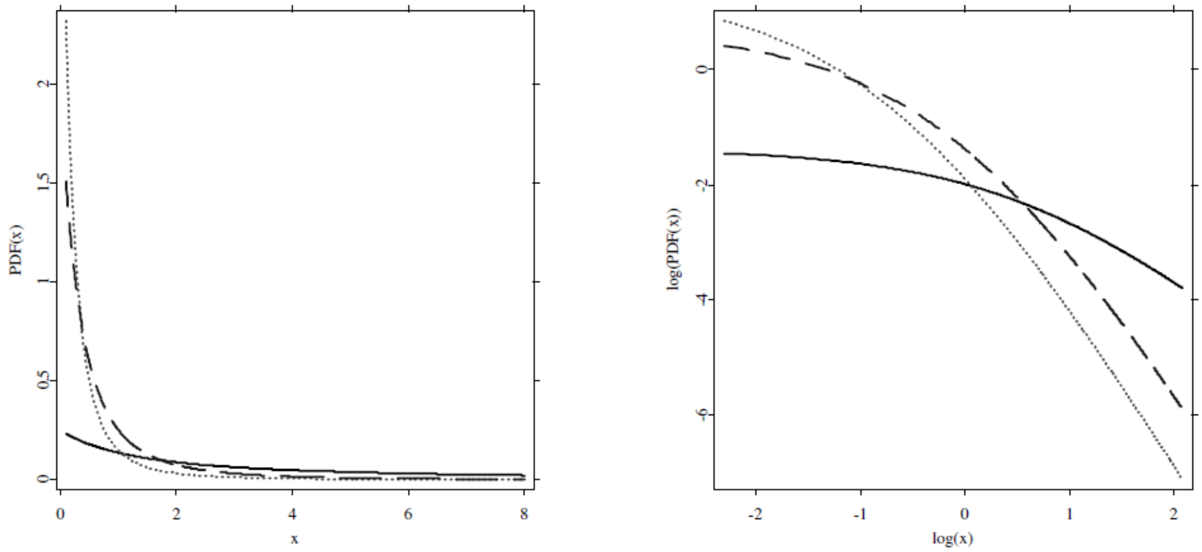

Рис.3. Левая панель: ЭФР Парето с параметрами α = 0,5 и λ = 2 (черная сплошная линия), α = 2 и λ = 0,5 (штрихпунктирная линия), a = 2 и A = 1 (точечная пунктирная линия). Правая панель: те же плотности Парето на двойном логарифмическом участке. Наглядно видны толстые степенные хвосты распределения Парето.

Выводы

Закон Парето очень полезен при моделировании размеров финансовых претензий в страховании. Его главный недостаток заключается в отсутствии математической приемлемости в некоторых ситуациях. Как и для логарифмически нормального распределения, преобразование Лапласа не имеет замкнутого представления формы, а функция генерации моментов не существует. Более того, как и экспоненциальная ЭФР, плотность Парето монотонно убывает, что может оказаться неприемлемым в некоторых практических финансовых ситуациях.

Литература:

1. CÍŽEK P., HÄRDLE, W., WERON, R. *Statistical Tools for Finance and Insurance.* Springer Berlin Heidelberg New York. ISBN 3-540-22189-1.

Prezentat la 06.09.2017