Глава 2

Основы теории–Распределения ущерба

§1 Введение

1.1 Страховое множество

Суммарное количество исков в отдельный период времени представляет фундаментальную важность для правильного управления страховой компанией. Ключевым допущением во всех моделях, изучаемых здесь, является то, что вероятность предъявления иска и величину(сумму) иска можно изучать отдельно. Таким образом, вероятность предъявления иска вычисляется в соответствии с некоторой простой моделью событий, попавших в определенный промежуток времени, и затем величина иска выбирается из распределения, описывающего эту величину.

1.2 Статистический фон

Для описания распределения случайных величин может быть использован ряд статистических методов. Конечная цель состоит в том, чтобы описать колебания величин исков с помощью нахождения распределения ущерба, которое в достаточной мере описывает иски, имеющиеся в действительности. Как правило, это осуществляется в два этапа.

На первом этапе, можно предположить, что иски происходят, как реализации известного распределения. Например, можно допустить, что логарифм величины иска следует, в умеренном приближении, нормальному распределению с известным значением и известным стандартным отклонением. Завершив процесс нахождения величины иска, мы могли бы переключить свое внимание на его результаты, нужные в страховании. Например, иски выше определенного уровня могут инициировать некоторые меры для перестрахования, в то время как иски ниже определенного уровня могут быть никогда не предъявлены, если франшиза будет в силе.

На практике же, точное распределение, описывающее иски, вряд ли когда-нибудь станет известно. На втором этапе, типичным методом действия является предположение, что исковое распределение - это член некоторого семейства. Параметры этого семейства должны быть оценены с использованием величины иска, записанной с помощью соответствующего метода, такого как метод максимального правдоподобия. Правда, если крупные иски будут ограничены (перестрахование) или некоторые незначительные иски не будут предъявлены (франшиза), при вычислениях могут возникнуть трудности. Можно провести множество исследований характера распределения, которое используется, чтобы описать переменную величины иска. И основным заключением является то, что у исковых распределений имеется тенденция быть совершенно асимметричными и с "тяжелыми хвостами".

1.3 Приблизительная оценка и критерий согласия

Методы максимального правдоподобия, моментов и процентилей могут использоваться в соответствующих распределениях для множества разных видов информации. Приемлемость распределения можно проверить формально, используя критерий χ^2. Метод процентилей описывается в секции 2.3; другие методы и критерий χ^2 содержатся в теме С1. Формуляр для плотностей, моментов производящих функций (если они существуют) для распределений, обсуждаемых в этой части, приводится в *Формуляре и таблицах для актуарных исследований.*

Экспоненциальное распределение

Случайная величина X имеет экспоненциальное распределение с параметром λ, если

$$F(x) = 1 - e^{-\lambda x}$$

и, можно записать, $X \sim exp(\lambda)$.

Для оценки параметра экспоненциального распределения можно использовать метод максимального правдоподобия (ММП) или метод моментов.

Распределение Парето

Случайная величина X имеет Парето-распределение с параметрами α и λ, если

$$F(x) = 1 - \left[\frac{\lambda}{\lambda + x}\right]^{\alpha}$$

и, можно записать, $X \sim Pa(\alpha, \lambda)$.

Легко проверить, что распределение Парето имеет функцию плотности

$$f(x) = \frac{\alpha\lambda^{\alpha}}{(\lambda + x)^{\alpha+1}}, x > 0, \alpha > 0, \lambda > 0.$$

Метод моментов очень легко применить в случае Парето-распределений, но оценки, полученные таким способом, будут содержать довольно много стандартных ошибок, главным образом из-за S^2, выборочной дисперсии, имеющей очень большое отклонение. Однако, этот метод обеспечивает начальные оценки для дальнейшего использования более эффективных методов, которые не так просты в применении. Например, ММП, где для решения могут понадобиться численные методы.

Распределение Вейбулла

Парето-распределение — это распределение с верхней хвостовой частью, которая стремится к 0, как степень x. Это дает распределение с более тяжелой хвостовой частью, чем экспоненциальное. Выражения для верхних хвостов экспоненциального и Парето-распределений таковы

экспоненциальное $\qquad P(X > x) = \exp(-\lambda x)$

Парето $\qquad P(X > x) = (\lambda/(\lambda + x))^{\alpha}.$

Воспользуемся дополнительной возможностью. Положим

$$P(X > x) = \exp(-\lambda x^{\gamma}), \gamma > 0.$$

Здесь у нас есть два случая. Если $\gamma < 1$, то получаем распределение с хвостовой частью, имеющей промежуточный вес между экспоненциальным и Парето-распределениями, если $\gamma > 1$, верхняя хвостовая часть будет легче, чем у экспоненциального (для экспоненциального распределения $\gamma = 1$). Распределение хвостовой части определяет распределение Вейбулла, очень гибкое распределение, которое может быть использовано как модель для ущерба в страховании, обычно с $\gamma < 1$. Случайная величина X имеет распределение Вейбулла с параметрами c и γ, если

$$F(x) = 1 - \exp(-cx^{\gamma})$$

и можно записать, что $X \sim W(c, \gamma)$. (Заметим, что изменения от γ к с описаны в *Таблицах для актуарных исследований*). Функция плотности для $W(c, \gamma)$

$$f(x) = c\gamma x^{\gamma-1} \exp(-cx^{\gamma}), x > 0, c > 0, \gamma > 0.$$

Ни метод моментов, ни метод максимального правдоподобия не могут применяться, если c и γ не известны (хотя на практике, если есть компьютер, эти уравнения довольно легко решаемы). В случае, когда γ — это известная величина γ^*, то достаточно простым окажется метод максимального правдоподобия.

Функция распределения $W(c, \gamma)$ — это элементарная функция, и на этом факте может базироваться простой метод оценки для c и γ. Метод основан на приравнивании подобранных выборочных процентилей к функции распределения. Например, приравнивание квартилей, 25-ого и 75-ого процентилей, к популяционным квартилям. Это соответствует способу, в котором выборочные моменты приравниваются к популяционным моментам в методе процентилей.

Первые два момента (в методе моментов) используются, если есть два неизвестных параметра, это интуитивно кажется очевидным (хотя теоретический базис для этого не так прост). Аналогично могла бы быть использована медиана, если бы был только один параметр для оценки. С двумя параметрами, оптимальная процедура менее проста, но пониженные и повышенные квартили кажутся вполне разумным выбором.

Пример 2.1 Оценить с и γ в распределении Вейбулла, используя метод процентилей, где первый выборочный квартиль равен 401 и третий квартиль равен 2836.75.

Решение Уравнения для с и γ

$$F(401) = 1 - \exp(-c * 401^{\gamma}) = 0.25$$

$$F(2836.75) = 1 - \exp(-c * 2836.75^{\gamma}) = 0.75,$$

которые могут быть переписаны в виде

$$-c * 401^{\gamma} = \log(3/4)$$

и

$$-c * 2836.75^{\gamma} = \log(1/4)$$

Делим одно уравнение на другое и получаем, что $\widetilde{\gamma} = 0.8038$, следовательно $\widetilde{c} = 0.002326$, где $\sim$ означает процентильную оценку. Заметим, что $\widetilde{\gamma} < 1$ дает более тяжелую хвостовую часть, чем у экспоненциального распределения.

Гамма-распределение

Случайная переменная X имеет гамма-распределение с параметрами α и λ, если

$$f(x) = \frac{\lambda^{\alpha}\Gamma(\alpha)}{x}^{\alpha-1} \exp(-\lambda x), x > 0$$

и можно записать, что $X \sim G(\alpha, \lambda)$. Среднее значение и вариация X

$$E(X) = \frac{\alpha}{\lambda}; Var(X) = \frac{\alpha}{\lambda^2}$$

Моменты имеют простую форму, поэтому метод моментов легко применим. Оценки ММП для гамма-распределения не могут быть получены в конечной форме (в терминах элементарных функций), но эти оценки могут использоваться, как исходные в поиске ММП-оценок.

Более удобно для получения ММП-оценок гамма-распределения использовать различную параметризацию. Положим $\mu = \alpha/\lambda$ и оценим параметры α и μ. Затем восстановим ММП-оценку λ, положив $\widetilde{\lambda} = \widetilde{\alpha}/\widetilde{\mu}$. Здесь мы используем свойство постоянства ММП-оценок.

Логнормальное распределение

Определение логнормального распределения очень простое: X имеет логноромальное распределение, если $\log X$ имеет нормальное распределение. Когда $\log X \sim N(\mu, \sigma^2), X \sim LogN(\mu, \sigma^2)$.

Оценка логнормального распределения является прямой до тех пор, пока μ и σ^2 могут быть оценены с помощью логорифма преобразованной информации. Пусть $x_1, x_2, \ldots, x_n$ будут наблюдаемыми переменными и пусть $y_i = \log x_i$. Оценки μ и σ^2 в ММП — это $\overline{y}$ и s_y^2, где нижний индекс y обозначает выборочную дисперсию, вычисленную на значениях y.

Смешанные распределения

Экспоненциальное распределение — это одна из простейших моделей для страховых потерь. Предположим, что каждый индивидуальный иск в большом страховом портфеле терпит убытки в соответствии с экспоненциальным распределением. Практическое изучение фактически любого страхового портфеля показывает, что значения этих различных распределений будут отличаться среди держателей полисов. Таким образом, описание ущербов для всего портфеля состоит в том, что каждый отдельный ущерб следует собственному экспоненциальному распределению, так как значения этих распределений различны.

Сейчас будем искать описание отклонения отдельных средних значений. Один из способов это сделать — предположение, что экспоненциалы этих значений следуют распределению. В экспоненциальном случае, удобно сделать следующее предположение. Пусть $\lambda_i = 1/\theta_i$ описывает взаимодействие со средним значением ущерба для i-ого держателя полиса. Предположим, что вариация для λ_i может быть описана известным гамма-распределением $G(\alpha, \delta)$, то есть предположим, что $\lambda \sim G(\alpha, \delta)$, где

$$f(\lambda) = \frac{\delta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\delta\lambda), \lambda > 0.$$

Особо заметим, что это — функция плотности λ с известными значениями α и δ.

Такая формулировка имеет много общего с тем, что используется в оценке Байеса. Действительно, фундаментальная идея этой оценки состоит в том, что интересующий параметр (здесь, λ)может быть представлен как случайная величина с известным распределением. Заметим, однако, что целью здесь является не оценить отдельное λ_i, а описать совокупность ущерба для всего портфеля. Отдельное λ_i может быть представлено с помощью байесовской оценки, когда к $G(\alpha, \delta)$ распределению можно было бы относиться, как к основному распределению. В задаче описания ущерба для всего портфеля, $G(\alpha, \delta)$ распределение используется, чтобы усреднить экспоненциальные распределения; к нему относятся, как к смешивающему распределению, и , следовательно, в таких случаях распределение ущерба называют смешанным.

Предельное распределение X

$$f_X(x) = \int f_{X,\lambda}(x, \lambda) d\lambda$$

$$= \int f_\lambda(\lambda) f_{X|\lambda}(x \mid \lambda) d\lambda$$

$$= \int_0^\infty \frac{\delta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\delta\lambda) \times \lambda \exp(-\lambda x)) d\lambda$$

$$= \frac{\delta^\alpha}{\Gamma(\alpha)} \int_0^\infty \lambda^\alpha \exp\{-(x+\delta)\lambda\} d\lambda$$

$$= \frac{\delta^\alpha}{\Gamma(\alpha)} \times \frac{\Gamma(\alpha+1)}{(x+\delta)^{\alpha+1}} (G(\alpha+1, x+\delta) - \text{интеграл})$$

$$= \frac{\alpha\delta^\alpha}{(x+\delta)^{\alpha+1}}, \qquad x > 0$$

что является Парето-распределением $Pa(\alpha, \delta)$. Такой результат дает очень хорошую интерпретацию распределения Парето: $Pa(\alpha, \delta)$ возникает, когда экспоненциально распределенные ущербы усредняются $G(\alpha, \delta)$-смешивающим распределением.

Обобщения распределения Парето

Производящая функция распределения Парето $Pa(\alpha, \lambda)$

$$F(x) = 1 - \frac{\lambda^\alpha}{(\lambda + x)^\alpha}$$

Дополнительный параметр γ может быть представлен таким образом:

$$F(x) = 1 - \frac{\lambda^\alpha}{(\lambda + x^\gamma)^\alpha}$$

Такая производящая функция преобразует распределения Бурра и Парето. Дополнительный параметр дает особую гибкость, когда требуется подстроится к имеющейся информации. С того момента, как мы получим производящую функцию в конечном виде, станет возможным приблизить распределение Бурра к имеющейся информации, используя методы процентилей. ММП обычно требует использования компьютерных программ, которые позволяют нелинейную оптимизацию.

Второе обобщение распределения Парето задействуют идею смешанного распределения, обсужденного ранее. Если ущербы — это экспоненциалы со средним значением $1/\lambda$, и $\lambda \sim G(\alpha, \delta)$, тогда предельное распределение ущербов — это $Pa(\alpha, \delta)$. Можно сделать обобщение, если предположить, что потери — это $G(k, \lambda)$ и $\lambda \sim G(\alpha, \delta)$. В частном случае, если $k = 1$, то $Pa(\alpha, \delta)$ распределение получено также, как и ранее

$$\begin{aligned} f_X(x) &= \int f_\lambda(\lambda) f_{X|\lambda}(x \mid \lambda) d\lambda \\ &= \int_0^\infty \frac{\delta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\delta\lambda) \times \frac{\lambda^k}{\Gamma^(k)} x^{k-1} \exp(-\lambda x)) d\lambda \\ &= \frac{\delta^\alpha x^{k-1}}{\Gamma(\alpha)\Gamma(k)} \int_0^\infty \lambda^{\alpha+k-1} \exp\{-(x+\delta)\lambda\} d\lambda \\ &= \frac{\Gamma(\alpha+k)\delta^\alpha}{\Gamma(\alpha)\Gamma(k)} \frac{x^{k-1}}{(x+\delta)^{\alpha+k}}, \qquad x > 0, \end{aligned}$$

где конечный интеграл вычисляется от $G(\alpha + k, \delta + x)$ — интегрируемой функции. А, значит, мы нашли функцию плотности обобщенного Парето-распределения.

Моменты обобщенного Парето-распределения могут быть получены либо непосредственным вычислением $\int xf(x)dx$, либо использованием условного аргумента математического ожидания. Что касается оценки, так как производящая функция моментов не определена, то метод процентилей использовать нельзя. ММП может быть применен, но, опять же, необходимо подходящее программное обеспечение; метод моментов может обеспечить начальные оценки для любой итерационной схемы.

§2 Формулы

Гамма-распределение

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \qquad x > 0$$

Логнормальное распределение

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp[-\frac{1}{2}(\frac{\log x - \mu}{\sigma})^2], \qquad x > 0$$

Парето-распределение

$$f_X(x) = \frac{\alpha\lambda^\alpha}{(\lambda + x)^{\alpha+1}}, \qquad x > 0 \qquad \text{для двух параметров}$$

$$f_X(x) = \frac{\Gamma(\alpha + k)\lambda^\alpha x^{k-1}}{\Gamma(\alpha)\Gamma(k)(\lambda + x)^{\alpha+k}}, \qquad x > 0 \qquad \text{общий вид}$$

Распределение Бурра

$$f_X(x) = \frac{\alpha\gamma\lambda^\alpha x^{\gamma-1}}{(\lambda + x^\gamma)^{\alpha+1}}, \qquad x > 0$$

Исковая частота

$$\text{Исковая частота} = \frac{\text{количество исков}}{\text{среднее количество полисов}}$$

Рисковая премия

Рисковая премия=Ожидаемая частота иска×Средняя величина иска

§3 Вопросы студентам

В1 Почему рассматриваемые распределения называются распределениями "ущербов а не "исков"?

О1 Ущерб — это полная стоимость восстановления ущерба, тогда как исковая величина — это только действительная величина оплачиваемой суммы. В следующей части мы увидим, что страховщик не всегда обязан возмещать все потери, например, если применяется франшиза, если ущерб превышает сумму, указанную в полисе или если использовалось еще какое-либо страхование и стоимость разделена. Поэтому величина ущерба и величина иска не всегда одинаковы.

В2 Все ли основные страховые иски оплачиваются на основе компенсации ущерба?

О2 Подавляющее большинство — да. Однако, есть несколько ситуаций, когда дело обстоит иначе. Например:

- В страховании имущества, покрытие может обеспечиваться на основе системы *новое-за-старое*, которая означает, что все предметы будут заменены новыми эквивалентами(превышающими стоимость старых предметов, которые были утеряны, украдены или повреждены).
- В страховании от несчастных случаев выплачиваются заранее заданные суммы, если застрахованный человек получил определенные травмы, например, потерю конечности или глаза.
- Травмированные люди могут получить компенсационные выплаты, превышающие их действительную материальную стоимость.

В3 Верно ли, что, если дисперсия распределения равна бесконечности, то асимметрия также будет бесконечной?

О3 Да. Для типов распределения ущербов, которые мы будем использовать, обнаружится, что, когда мы меняем параметры, высшие моменты будут "идти" впереди, то есть сначала асимметрия станет бесконечной, затем дисперсия, затем среднее значение.

В4 Есть ли смысл в том, чтобы стандартное отклонение было больше, чем среднее значение?

О4 Да. Распределения, которые мы рассматриваем здесь, очень несимметричные, и нет причин, по которым стандартное отклонение не могло бы быть больше среднего значения.

В5 Должны ли мы принимать во внимание реальные периоды воздействия на примере частоты иска?

О5 Да. Чтобы не усложнять ситуацию, можно предположить, что профиль держателя полиса останется постоянным на протяжении всего периода.

В6 Как сильно отличаются премии, реально присутствующие в общем страховании, с теоретической значением премии?

О6 На практике, общие страховые премии находятся под сильным влиянием конкурентного давления других компаний. И действительные премии могут значительно отличаться от теоретических премий. Тем не менее, важно, что страхователи осознают, как реальные премии сопоставляются с теоретическими, поэтому они могут разделять ущербы по областям.

3.1 Подсказки для ответов на вопросы

1. Вопросы к этой теме обычно очень простые. Однако, вам нужно с легкостью уметь интегрировать стандартные функции, интегрировать по частям или с помощью замены переменных.

2. Параметры, используемые в Парето, Бурра и Вейбулла распределениях, не являются стандартными. Поэтому возьмите на заметку конкретные функции плотности, которые заданы в задачах, использующих эти распределения.

§4 Ответы на вопросы для самоподготовки

Решение 1.1

Функция плотности для $N(0,1)$-распределения $f(x) = \frac{1}{\sqrt{2\pi}} e^{-1/2x^2}$. Тогда производящая функция моментов:

$$M(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tX} \frac{1}{\sqrt{2\pi}} e^{-1/2x^2} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-1/2(x^2-2tx)} dx$$

$$= e^{1/2t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-1/2(x-t)^2} dx$$

Но последний интеграл – это интеграл по всей числовой прямой функции распределения $N(t,1)$, и потому он равен 1. Тогда производящая функция моментов стандартного нормального распределения равна $e^{1/2t^2}$.

Решение 1.2

Если X имеет $LogGamma(\alpha, \lambda)$ распределение, то $Y = \log X$ имеет $Gamma(\alpha, \lambda)$ распределение. Тогда, используя функцию плотности $Gamma(\alpha, \lambda)$ распределения, найдем математическое ожидание величины иска:

$$E(X) = E(e^Y) = \int_0^{\infty} e^y \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy$$

Простейший способ вычислить этот интеграл — это подобрать константы таким образом, чтобы сделать его похожим на другое гамма-распределение (в данном случае, $Gamma(\alpha, \lambda - 1)$-распределение). Что дает (обеспечивая $\lambda > 1$):

$$E(X) = \frac{\lambda^\alpha}{(\lambda-1)^\alpha} \int_0^{\infty} \frac{(\lambda-1)^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-(\lambda-1)y} dy$$

$$= \left[\frac{\lambda}{\lambda-1}\right]^\alpha P[0 < Gamma(\alpha, \lambda-1) < \infty] = \left[\frac{\lambda}{\lambda-1}\right]^\alpha$$

(Заметим, что вы можете применять эту формулу для нахождения среднего значения логнормального и логгамма-распределений, используя следующий метод:

Если X имеет $LogN(\mu, \sigma^2)$ распределение, то, по определению, $Y = \log X$ имеет $N(\mu, \sigma^2)$ распределение. Тогда можно использовать формулу для производящей функции моментов нормального распределения:

$$E(X) = E(e^Y) = M_Y(1) = e^{\mu + \frac{1}{2}\sigma^2}$$

Аналогично, если X имеет $LogGamma(\alpha, \lambda)$ распределение, то, по определению, $Y = \log X$ имеет $Gamma(\alpha, \lambda)$ распределение:

$$E(X) = E(e^Y) = M_Y(1) = (1 - 1/\lambda)^{-\alpha}$$

Хотя этот метод кажется довольно кратким, он не может рассматриваться в качестве доказательства основных принципов, так как предполагает известной формулу для производящей функции моментов.)

Решение 1.3

Среднее значение Парето-распределения: $\int\limits_0^\infty x\alpha\lambda^\alpha(\lambda + x)^{-\alpha-1}dx$

Существуют два возможных способа вычисления этого интеграла:

(a) Воспользуемся заменой $t = \lambda + x$:

$$\int\limits_0^\infty x\alpha\lambda^\alpha(\lambda + x)^{-\alpha-1}dx = \int\limits_\lambda^\infty (t - \lambda)\alpha\lambda^\alpha t^{-\alpha-1}dx$$

$$= \alpha\lambda^\alpha \int\limits_\lambda^\infty t^{-\alpha}dt - \alpha\lambda^{\alpha+1} \int\limits_\lambda^\infty t^{-\alpha-1}dt$$
$$= \alpha\lambda^\alpha \left[\frac{t^{-\alpha+1}}{-\alpha+1}\right]_\lambda^\infty - \alpha\lambda^{\alpha+1}\left[\frac{t^{-\alpha}}{-\alpha}\right]_\lambda^\infty$$
$$= \frac{\alpha\lambda}{\alpha - 1} - \lambda$$
$$= \frac{\lambda}{\alpha - 1}$$

Эти вычисления возможны только в том случае, когда интеграл сходится на бесконечности. Такое условие требует, чтобы выражения, находящиеся в квадратных скобках были отрицательны, что будет выполнятся при $\alpha > 1$.

(б) Запишем начальное x в интеграле, как $(\lambda+x)-\lambda$ и разобьем его на два интеграла, соответствующих распределению Парето:

$$\int_0^\infty x\alpha\lambda^\alpha(\lambda+x)^{-\alpha-1}dx = \int_0^\infty [(\lambda+x)-\lambda]\alpha\lambda^\alpha(\lambda+x)^{-\alpha-1}dx$$

$$\frac{\alpha\lambda}{\alpha-1}\int_0^\infty (\alpha-1)\lambda^{\alpha-1}(\lambda+x)^{-\alpha}dx - \lambda\int_0^\infty \alpha\lambda^\alpha(\lambda+x)^{-\alpha-1}dx$$

$$\frac{\alpha\lambda}{\alpha-1}P[0<Pareto(\alpha-1,\lambda)<\infty] - \lambda P[0<Pareto(\alpha,\lambda)<\infty]$$

$$= \frac{\alpha\lambda}{\alpha-1} - \lambda$$

$$= \frac{\lambda}{\alpha-1}$$

Как альтернативный вариант вы можете использовать интегрирование по частям, или записать этот интеграл, как функцию плотности обобщенного распределения Парето.

Решение 1.4

Формула, данная в *Таблицах* для среднего значения распределения Бурра:

$$\frac{\lambda^{1/\gamma}\Gamma(\alpha-1/\gamma)\Gamma(1+1/\gamma)}{\Gamma(\alpha)}$$

Когда $\gamma=1$, получаем:

$$\frac{\lambda\Gamma(\alpha-1)\Gamma(2)}{\Gamma(\alpha)} = \frac{\lambda\Gamma(\alpha-1)\times 1}{(\alpha-1)\Gamma(\alpha-1)} = \frac{\lambda}{\alpha-1},$$

что является формулой для среднего значения распределения Парето.

Решение 1.5

Среднее значение распределения Вейбулла:

$$E(X) = \int_0^\infty xc\gamma x^{\gamma-1}e^{-cx^\gamma}dx$$

Подставляя $u = cx^\gamma$, получаем, что $\frac{du}{dx} = c\gamma x^{\gamma-1}$, и, в итоге:

$$E(X) = \int\limits_0^\infty x e^{-u} du = \int\limits_0^\infty \left[\frac{u}{c}\right]^{1/\gamma} e^{-u} du$$

Попробуем представить этот интеграл в виде функции плотности для $Gamma(1 + 1/\gamma, 1)$ распределения:

$$E(X) = c^{-1/\gamma}\Gamma(1 + 1/\gamma) \int\limits_0^\infty \frac{1}{\Gamma(1 + 1/\gamma)} u^{1/\gamma} e^{-u} du$$

$$= \frac{\Gamma(1 + 1/\gamma)}{c^{1/\gamma}} P[0 < Gamma(1 + 1/\gamma, 1) < \infty] = \frac{\Gamma(1 + 1/\gamma)}{c^{1/\gamma}}$$

Это соответствует формуле, приведенной в *Таблицах*, для среднего значения распределения Вейбулла.

Решение 1.6

Формула для дисперсии распределения Вейбулла, данная в *Таблицах*:

$$\frac{\Gamma(1 + 2/\gamma)}{c^{2/\gamma}} - \left[\frac{\Gamma(1 + 1/\gamma)}{c^{1/\gamma}}\right]^2$$

При $\gamma = 1$:

$$\frac{\Gamma(3)}{c^2} - \left[\frac{\Gamma(2)}{c}\right]^2 = \frac{2}{c^2} - \left[\frac{1}{c}\right]^2 = \frac{1}{c^2},$$

что является формулой дисперсии экспоненциального распределения с $\lambda = c$.

Решение 1.7

Производящая функция для распределения Вейбулла:

$$F_X(x) = \int\limits_0^x c\gamma x^{\gamma-1} e^{-cx\gamma} dx = \left[-e^{-cx\gamma}\right]_{0^x} = 1 - e^{-cx^\gamma}$$

Тогда:

$$F_X(5000) = 1 - e^{-0.00001 \times 5000^{1.5}} = 1 - 0.0291 = 0.9709$$

$$F_X(2500) = 1 - e^{-0.00001 \times 2500^{1.5}} = 1 - 0.2865 = 0.7135$$

Получаем требуемую вероятность: $F_X(5000) - F_X(2500) = 0.9709 - 0.7135 = 0.2574$

Решение 1.8

При $k = 2$ функция плотности обобщенного распределения Парето:

$$f_X(x) = \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)\Gamma(2)}\lambda^\alpha x(\lambda+x)^{-\alpha-2} = (\alpha+1)\alpha\lambda^\alpha x(\lambda+x)^{-\alpha-2}$$

Воспользуемся подстановкой $t = \lambda + x$ и запишем вероятность иска, превышающего M:

$$P(X > M) = \int_M^\infty (\alpha+1)\alpha\lambda^\alpha x(\lambda+x)^{-\alpha-2}dx$$

$$= \int_{\lambda+M}^\infty (\alpha+1)\alpha\lambda^\alpha (t-\lambda)t^{-\alpha-2}dt$$

$$= (\alpha+1)\alpha\lambda^\alpha \int_{\lambda+M}^\infty t^{-\alpha-1}dt - (\alpha+1)\alpha\lambda^{\alpha+1} \int_{\lambda+M}^\infty t^{-\alpha-2}dt$$

$$= (\alpha+1)\alpha\lambda^\alpha \left[\frac{t^{-\alpha}}{-\alpha}\right]_{\lambda+M}^\infty - (\alpha+1)\alpha\lambda^{\alpha+1} \left[\frac{t^{-\alpha-1}}{-\alpha-1}\right]_{\lambda+M}^\infty$$

$$= (\alpha+1)\left[\frac{\lambda}{\lambda+M}\right]^\alpha - \alpha\left[\frac{\lambda}{\lambda+M}\right]^{\alpha+1}$$

Подставляя $\alpha = 5, \lambda = 200$ и $M = 300$ получаем искомую вероятность:

$$(\alpha+1)\left[\frac{\lambda}{\lambda+M}\right]^\alpha - \alpha\left[\frac{\lambda}{\lambda+M}\right]^{\alpha+1} = 6\left[\frac{200}{200+300}\right]^5 - 5\left[\frac{200}{200+300}\right]^6$$

$$= 6 \times 0.4^5 - 5 \times 0.4^6 = 0.041,$$

то есть 4.1% исков будут превышать £300000.

Решение 1.9

Дисперсия $Gamma(\alpha, \lambda/1.1)$ распределения $\frac{1.1^2\alpha}{\lambda^2}1.21 \times \frac{\alpha}{\lambda^2}$. То есть дисперсия распределения размера иска возрастет на 21%. Если сравнить это со стандартными отклонениями, то мы увидим, что стандартное отклонение в данном случае возрастет на 10%, в отличие от того, как мы могли бы ожидать.

Решение 1.10

Пусть $Y = (1+k)X$. Сделаем подстановку $y = (1+k)x$ в интеграле функции плотности:

$$\int_0^\infty \frac{\alpha\lambda^\alpha}{(\lambda+x)^{\alpha+1}}dx = \int_0^\infty \frac{\alpha\lambda^\alpha}{(\lambda+\frac{y}{1+k})^{\alpha+1}}\frac{dy}{1+k} = \int_0^\infty \frac{\alpha\lambda^\alpha(1+k)^\alpha}{(\lambda(1+k)+y)^{\alpha+1}}dy$$

Это функция плотности $Pareto(\alpha, \lambda(1+k))$ распределения. Значит, это распределение инфляционных размеров иска.

Решение 1.11

Если N — это общее число исков, то мы имеем:

$$P(N=n) = \int_0^\infty p_{N|\theta}(n) f_\theta(\theta) d\theta$$

$$= \int_0^\infty \frac{e^{-\theta}\theta^n}{n!} \times \frac{1}{\Gamma(\alpha)}\delta^\alpha\theta^{\alpha-1}e^{-\delta\theta}d\theta$$

Преобразуем это выражение к интегралу $Gamma(n+\alpha, \delta+1)$ распределения:

$$P(N=n) = \frac{\Gamma(n+\alpha)}{n!(\delta+1)^{n+\alpha}}\frac{\delta^\alpha}{\Gamma(\alpha)}\int_0^\infty \frac{1}{\Gamma(n+\alpha)}(\delta+1)^{n+\alpha}\theta^{n+\alpha-1}e^{-(\delta+1)\theta}d\theta$$

$$\frac{\Gamma(n+\alpha)}{n!\Gamma(\alpha)}\frac{\delta^\alpha}{(\delta+1)^{n+\alpha}}, \qquad n = 0, 1, 2, \ldots$$

Представим его в виде:

$$P(N=n) = \binom{n+\alpha-1}{n}\left[\frac{\delta}{\delta+1}\right]^\alpha\left[\frac{1}{\delta+1}\right]^n, \qquad n = 0, 1, 2, \ldots$$

и получим отрицательное биномиальное распределение с параметрами $p = \frac{\delta}{\delta+1}$ и $k = \alpha$.

Решение 1.12

Воспользуемся тем же способом:

$$P(X=x)=\int\limits_0^\infty f_{X|p}(x)f_p(p)dp$$

$$=\int\limits_0^\infty \binom{n}{x}p^x(1-p)^{n-x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}dp$$

Приведем это выражение к виду $Beta(x+\alpha,n-x+\beta)$ распределения:

$$P(X=x)=\frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)\Gamma(\alpha+\beta)n!}{\Gamma(n+\alpha+)\Gamma(\alpha)\Gamma(\beta)(n-x)!x!}\times$$

$$\times\int\limits_0^1 \frac{\Gamma(n+\alpha+\beta)}{\Gamma(x+\alpha)\Gamma(n-x+\beta)}p^{x+\alpha-1}(1-p)^{n-x+\beta-1}dp$$

Тогда общее распределение исков:

$$P(X=x)=\frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)\Gamma(\alpha+\beta)n!}{\Gamma(n+\alpha+)\Gamma(\alpha)\Gamma(\beta)(n-x)!x!},\qquad x=0,1,2,\ldots,n.$$

Решение 1.13

Приравниваем формулы среднего значения и стандартного отклонения Парето-распределения к заданным величинам:

$$\frac{\lambda}{\alpha-1}=5000\text{и}\frac{\lambda}{\alpha-1}\sqrt{\frac{\alpha}{\alpha-2}}=7500$$

Делим второе равенство на первое:

$$\sqrt{\frac{\alpha}{\alpha-2}}=\frac{7500}{5000}=1.5$$

Тогда $\frac{\alpha}{\alpha-2}=1.5^2=2.25$

$$\alpha=2.25(\alpha-2)\qquad\Rightarrow\qquad\alpha=4.5/1.25=3.6$$

Можно найти λ:

$$\lambda=5000(3.6-1)=13000$$

Далее, процент исков, превышающих £25000:

$$P(X > 25000) = \int\limits_{25000}^{\infty} \alpha\lambda^{\alpha}(\lambda + x)^{-\alpha-1}dx$$

$$= [-\lambda^{\alpha}(\lambda + x)^{-\alpha}]_{25000}^{\infty}$$

$$= \left[\frac{\lambda}{\lambda + 25000}\right]^{\alpha}$$

$$= \left[\frac{13000}{13000 + 25000}\right]^{3.6} = 0.021,$$

то есть 2,1% исков будут превышать £25000.

Оба ответа одинаковые. Значит, £25000 — это критическая точка, в которой вероятности любых распределений ущерба равны.

Решение 1.14

Подобная функция:

$$L(\alpha, \lambda, \gamma) = \alpha^n \gamma^n \lambda^{n\alpha} \Pi x_i^{\gamma-1} \Pi(\lambda + x_i^{\gamma})^{-\alpha-1}$$

Возьмем логарифм:

$$\log L = n \log \alpha + n \log \gamma + n\alpha \log \lambda + (\gamma - 1) \sum \log x_i - (\alpha + 1) \sum \log(\lambda + x_i^{\gamma})$$

Дифференцируем по α:

$$\frac{\partial}{\partial \alpha} \log L = \frac{n}{\alpha} + n \log \lambda - \sum \log(\lambda + x_i^{\gamma})$$

Устремляем равенство к 0 и преобразуем:

$$\widetilde{\alpha} = \frac{n}{\sum \log(\lambda + x_i^{\gamma}) - n \log \lambda}$$

Подстановка известных значений $\lambda = 500$ и $\gamma = 2$ дает искомый результат.

После упрощений получим $\sum \log(500 + x_i^2) = 47.6245$. Подставляя данное значение, находим $\widetilde{\alpha} = 0.3021$

Решение 1.15

Ожидаемая частота иска:

$$\frac{315}{6200} = 0.050806$$

Средняя величина иска — это среднее значение $Gamma(50, 0.02)$ распределения, которая равна:

$$\frac{50}{0.02} = 2500$$

Тогда рисковая премия:

$$0.050806 \times 2500 = \pounds 127.02$$

Тогда значение премии P удовлетворяет равенству:

$$P = 127.02 + 50 \times 0.050806 + 0.35P$$

То есть значение премии равно £199.32.

Решение 1.16

Вот список возможных причин:

1. Текущее распределение размера иска может не подходить для исков, которые будут иметь место в будущем.

2. Необходимость конкуренции на рынке может означать, что теоретическая премия не соответствует действительной.

3. Частота иска может изменятся сверх ожидания. Если за последние несколько лет имело место 3150 исков, то можно предположить, что частота иска возрастет сверх ожидаемой нормы, в этом случае 0.050806

4. Законодательство может вводить ограничения (верхние или нижние) на премии.

5. Эти вычисления игнорируют временные денежные характеристики (таких как доля (капитала) в деле и инфляция)

Решение 1.17

Проверим, что:
H_0 : $N(\mu, \sigma^2)$ представляет собой удачную модель для этих исков.
H_1 : $N(\mu, \sigma^2)$ представляет собой неудачную модель.

Значение тестовой статистики:

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(120-104)^2}{104} + \ldots + \frac{(62-76)^2}{76} = 25.94$$

Сравнивая это значение с соответствующей точкой χ_2^2 распределения (5 элементов минус 2 оценочных параметра минус 1), мы получим вероятностное значение, гораздо меньшее, чем 0.005. Тогда для нас становится очевидной неоптимальность H_0, и можно сделать заключение, что нормальное распределение в данном случае не является подходящим.

Заключение части I

Размеры индивидуальных исков могут быть описаны с помощью распределений ущерба. Например, стандартных распределений, таких как логнормальное, Гамма, Парето, распределения Бура, Вейбулла.
Рисковая премия и теоретическая офисная премия могут быть вычислены, зная распределение ущерба и частоту исков.