Глава 6

Основы теории–Суммарные страховые выплаты

Модели риска

§1 Введение

Определение и свойства производящей функции моментов нам должны быть известны. Свойство о том, что два различных распределения не могут иметь одну и ту же производящую функцию моментов, будет использоваться позже в этой главе. Пусть X – случайная величина, $M(t)$ – её производящая функция моментов. Известное нам свойство, что для любого натурального n:

$$\frac{d^n}{dt^n} M(t)\,|_{t=0} = E[X^n] \tag{6.1}$$

Менее известное свойство, но не менее широко применимое, то, что для $n = 2$ и $n = 3$:

$$\frac{d^n}{dt^n} \log M(t)\,|_{t=0} = E[(X - E[X])^n] \tag{6.2}$$

$\log M(t)$ — есть производящая функция семиинвариантов случайной величины X. Преимущество формулы (6.2) перед (6.1) в том, что центральные моменты представляют для нас больший интерес, чем нецентральные моменты. А для $n = 2$ и $n = 3$ эти центральные моменты дает нам как раз формула (6.2).

В этой главе будут использоваться свертки функций распределения. Предположим, что $\{X_i\}_{i=1}^n$ — независимые одинаково распределенные случайные величины с общей функцией распределения $F(x)$. Тогда функция распределения суммы $\sum\limits_{i=1}^{n} X_i$, обозначенная как $F^{n^*}(x)$:

$$F^{n^*}(x) = P(X_1 + X_2 + ... + X_n \leq x)$$

§2 Модели для краткосрочного страхования

2.1 Основная модель

Многие формы страхования, не связанные с жизнью (*non-life insurance*), например автострахование, могут рассматриваться как краткосрочное страхование, также как и некоторые формы страхования жизни (*life insurance*), например групповое страхование жизни или одногодичный страховой полис.

Краткосрочное страхование можно определить следующими его свойствами:

- Полис действует в течение фиксированного, относительно короткого, периода времени, обычно в течение года.
- Страховая компания получает премии от страховщиков – полисодержателей (*policyholder*).
- В свою очередь, страховщик оплачивает все иски согласно купленным полисам.

В конце периода действия полиса страхователь может продлить или не продолжать срок действия; в случае продления срока, премии, выплачиваемые страхователем, могут быть такой же величины или другой.

Страховщик может передать часть премий перестраховщику; в свою очередь, перестраховщик покрывает часть исков в течение времени действия полиса согласно некоторому установленному соглашению.

Важная особенность договора краткосрочного страхования состоит в том, что величина премий устанавливается таким образом, чтобы только покрыть иски, возникающие в течение действия полиса. Такой вид договора вместе с полисами страхования жизни, где коэффициент смертности растет с увеличением возраста, означает, что годичный страховой взнос в ранние годы более чем достаточен, чтобы покрыть ожидаемые иски в этом возрасте. Величина превышения затем откладывается в качестве резервов, используемых в более позднем возрасте, когда величины премий уже не будет достаточно для того, чтобы покрыть ожидаемые иски.

Теперь более подробно, рассмотрим страхование риска по договору краткосрочного страхования. Риск включает в себя отдельно каждый полис или определенную группу полисов. Для простоты будем считать, что срок действия договора равен одному году, но любой другой короткий период, например шесть месяцев, также может подойти. Случайная величина S показывает суммарные выплаты страховщика за год относительно данного риска. Все модели будут построены для этой случайной величины. В следующих двух параграфах будет изучена модель коллективного риска. Позже, в 4 параграфа, идея модели коллективного риска распространится и на модель индивидуального риска. Первым пунктом в построении модели коллективного риска является запись величины S для числа исков за указанный год, эту случайную величину обозначим за N, и для величины индивидуального иска. Пусть X_i обозначает величину i-го иска. Тогда:

$$S = \sum_{i=1}^{N} X_i$$

Поставленная задача состоит в следующем — определить моменты и установить распределение величины S через моменты и распределения N и X_i. Поставленную задачу будем рассматривать в условиях перестрахования и без. Для перестраховщика также определим моменты и распределение величины суммарных выплат за год относительно данного риска.

2.2 Рассмотрение упрощений в основной модели

Модель краткосрочного страхования, описанная в предыдущем параграфе, содержит несколько упрощений относительно реального процесса страхования. Во-первых, обычно предполагается, что моменты, и иногда распределения, величин N и X_i точно известны. На практике их обычно оценивают по существующим данным, используя методы, уже изученные в I Части.

Во-вторых, предполагается, по крайней мере неявно, что размер иска не изменеется, по крайней мере после того, как страховой случай, вызвавший иск, произошел. То есть, например, прибыль страховщика к концу года точно известна. На практике требуется, как минимум, кокроткий промежуток времени на урегулирование размера требования, а в некотороых случаях на урегулирование уходят годы. Это может произойти, если размер ущерба сложно определить, например, если возникают судебные разбирательства.

Эта модель вообще не включает в себя рассмотрение расходов. Предполагается, что премии идут на покрытие ущербов и содержат нагрузку для прибыли страховой компании. На самом деле, премии, выплачиваемые страховщиком, также содержат нагрузку на издержки. Учет издержек можно включить в модель простым способом.

Важным моментом в моделях долгосрочного страхования является понятие процентной ставки (*interest rate*), т.к. (как было сказано выше) доход от нагруженной премии может быть инвестирован на создание резервов. Сам процент (*interest*) менее важен, но все же остается значимым в краткосрочном страховании. В моделях для краткосрочного страхования возможно включение процента, но обычно его не учитывают, по крайней мере в самых простых моделях.

2.3 Замечания и предположения

В продолжение всей главы будут сделаны следующие два важных предположения:

- случайные величины $\{X_i\}_{i=1}^{\infty}$ – независимые и одинаково распределеные
- случайная величина N не зависит от $\{X_i\}_{i=1}^{\infty}$

Иначе говоря, эти предположения означают:

1. число исков не влияет на размер индивидуального иска
2. величина одного индивидуального иска не влияет на величину любого другого индивидуального иска
3. распределение индивидуальных исков не изменяется в течение всего (короткого) промежутка времени действия полиса

В течение всей главы будем предполагать, что все иски принимают неотрицательные значения, так что $P(X_i \leq x) = 0$, для $x < 0$. Многие формулы в этой главе будут получены, используя производящие функции моментов S, N и X_i. Эти функции будем обозначать $M_S(t), M_N(t)$ и $M_X(t)$ соответственно, а также будем предполагать их существование для некоторого положительного значения фиктивной переменной t. Существование производящей функции моментов неотрицательной случайной величины для положительного значения t, вообще говоря, не может считаться доказанным фактом. Например, производящие функции моментов Парето и логнормального распределений не существуют для какого-либо положительного значения t. Однако, все формулы, полученные в этой главе с помощью производящих функций моментов, могут быть получены (хотя и менее просто) без предположения о существовании производящей функции моментов для $t > 0$.

$G(x)$ и $F(x)$ будут обозначать функции распределения S и X_i соответственно, так что:

$$G(x) = P(S \leq x) \quad \text{и} \quad F(x) = P(X_i \leq x)$$

Для удобства будем считать, что плотность распределения $F(x)$ существует, и будем обозначать её $f(x)$. В случае, когда плотность не будет существовать, т.е. когда X_i имеет дискретное распределение или смешанное непрерывное/дискретное распределение, выражение вида

$$\int\limits_0^\infty x f(x) dx$$

будем интерпретировать надлежащим образом. Значение будет определяться в контексте.

k-ый центральный момент X_i, $k = 1, 2, 3, ...$ будем обозначать m_k.

§3 Модель коллективного риска

3.1 Модель коллективного риска

Вернемся к параграфу 2.1, где S представляется как сумма N случайных величин X_i, где X_i обозначает размер i-го иска. Таким образом:

$$S = X_1 + X_2 + ... + X_N$$

и $S = 0$ если $N = 0$.
Отметим, что число N исков из рискового портфеля – общее для всей группы (в отличии от числа исков в случае индивидуального полиса), что дает такое название модели — "Модель коллективного риска". В рамках этой модели могут быть выведены выражения для функции распределения, математического ожидания, дисперсии и производящей функции моментов S.

Выражение для $G(x)$, функции распределения S, может быть получено, если рассмотреть событие $\{S \le x\}$. Заметим, что если это событие произошло, то возможен один, и только один, из следующих вариантов:

$$
\begin{array}{lll}
 & \{S \le x \text{ и } N = 0\} & (\text{т.е. нет ни одного иска}) \\
\text{или} & \{S \le x \text{ и } N = 1\} & (\text{т.е. один иск размера } \le x) \\
\text{или} & \{S \le x \text{ и } N = 2\} & (\text{т.е. два иска общего размера } \le x) \\
 & \vdots & \\
\text{или} & \{S \le x \text{ и } N = r\} & (\text{т.е. } r \text{ исков общего размера } \le x) \\
 & \vdots &
\end{array}
$$

и так далее. Эти события взаимоисключающие и исчерпывающие. Таким образом

$$\{S \le x\} = \bigcup_{n=0}^{\infty} \{S \le x \text{ и } N = n\}$$

и следовательно

$$P(S \le x) = \sum_{n=0}^{\infty} P(S \le x \text{ и } N = n) = \sum_{n=0}^{\infty} P(N = n) P(S \le x | N = n)$$

Теперь заметим, что если $N = n$, то S – сумма фиксированного числа n случайных величин, $\{X_i\}_{i=1}^{n}$, и значит:

$$P(S \le x | N = n) = F^{n^*}(x)$$

где $F^{n^*}(x)$ – n-кратная свертка распределения $F(x)$. (Отметим, что $F^{1^*}(x)$—есть $F(x)$ и, для простоты, пусть $F^{0^*}(x)$ равняется 1 для всех неотрицательных x. В остальных случаях $F^{0^*}(x) = 0$.) Таким образом:

$$G(x) = \sum_{n=0}^{\infty} P(N = n) F^{n^*}(x) \tag{6.3}$$

Формула (3.1) является в общем виде выражением для функции распределения S. Ни распределение N ни X_i не были определены.

Заметим, что когда X_i принимает целочисленные положительные значения, $P(S = x)$ можно легко найти для $x = 1, 2, 3, ...$, так как

$$P(S = x) = G(x) - G(x - 1) =$$

$$= \sum_{n=1}^{\infty} P(N = n)\{F^{n^*}(x) - F^{n^*}(x - 1)\}$$

то есть

$$P(S = x) = \sum_{n=1}^{\infty} P(N = n) f_x^{n^*} \tag{6.3}$$

где $f_x^{n^*} = F^{n^*}(x) - F^{n^*}(x - 1)$ – плотность распределения $\sum_{i=1}^{n} X_i$. Как и в случае непрерывного распределения X_i, $P(S = 0) = P(N = 0)$.

Ранее уже обсуждались существование и методы аппроксимации для оценивания $G(x)$. Метод аппроксимации требует знание моментов S. Обсудим это более детально.

Для вычисления моментов S используется условное математическое ожидание при условии N. Для того, чтобы найти $E[S]$, воспользуемся равенством:

$$E[S] = E[E[S|N]]$$

Тогда $E[S|N = n] = \sum_{i=1}^{n} E[X_i] = nm_1$. Следовательно $E[S|N] = Nm_1$.
И тогда

$$E[S] = E[Nm_1] = E[N]m_1 \tag{6.3}$$

Формула (3.1) имеет весьма понятную интерпретацию. Она означает, что ожидаемая величина суммарных исков — есть результат ожидаемого числа исков и ожидаемой величины индивидуального иска.

Для выражения $Var[S]$ воспользуемся равенством:

$$Var[S] = E[Var[S|N]] + Var[E[S|N]]$$

$Var[S|N]$ можно найти, используя предположение о независимости величин индивидуальных исков:

$$Var[S|N=n] = Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var[X_i] = n(m_2 - m_1^2)$$

таким образом $Var[S|N] = N(m_2 - m_1^2)$. Следовательно:

$$Var[S|N] = E[N(m_2 - m_1^2)] + Var[Nm_1]$$

то есть

$$Var[S|N] = E[N](m_2 - m_1^2) + Var[N]m_1 \tag{6.3}$$

В отличии от выражения для $E[S]$, формула (3.1) не имеет очевидной интерпретации. Дисперсия S выражается через среднее и дисперсию и случайной величины N, и случайной величины X_i.

Производящая функция моментов S может быть также найдена, используя условное математическое ожидание. По определению, $M_S(t) = E[\exp\{tS\}]$. Таким образом:

$$M_S(t) = E[E[\exp\{tS\}|N]] \tag{6.3}$$

$E[\exp\{tS\}|N=n] = E[\exp\{tX_1 + tX_2 + ... + tX_n\}]$, и так как $\{X_i\}_{i=1}^n$ независимые случайные величины, то:

$$E[\exp\{tX_1 + tX_2 + ... + tX_n\}] = \prod_{i=1}^{n} E[\exp\{tX_i\}]$$

Так как $\{X_i\}_{i=1}^n$ одинаково распределены — они имеют одинаковые производящие функции моментов, $M_X(t)$, поэтому:

$$\prod_{i=1}^{n} E[\exp\{tX_i\}] = \prod_{i=1}^{n} M_X(t) = [M_X(t)]^n$$

Следовательно:

$$E[\exp\{tS\}|N] = [M_X(t)]^N \tag{6.3}$$

Подставляя (3.1) в (3.1), получим:

$$M_S(t) = E[(M_X(t))^n] = E[\exp\{N \log M_X(t)\}] = M_N(\log M_X(t)) \tag{6.3}$$

Таким образом, производящая функция моментов S выражается через производящие функции моментов N и X_i. Как в предыдущем результате, распределения N и X_i не указаны точно.

Существует один специальный случай, который представляет особый интерес. Случай, когда все иски имеют одинаковый фиксированный размер.

Например, рассмотрим портфель одногодичного страхования с одинаковой величиной требования. Предполагая, что вероятность возникновения иска размера B равна 1 (т.е. $P(X_i = B) = 1$), получим $m_1 = B$ и $m_2 = B^2$. Тогда S может принимать следующие значения: $0, B, 2B, \ldots$. Фактически $S = BN$, значит:

$$P(S \leq Bx) = P(N \leq x)$$

и распределение S следует из распределения N. Формулы (3.1) и (3.1) дают среднее и дисперсию S, но так как $S = BN$ гораздо проще обратить внимание на то, что $E[S] = E[N]B$ и $Var[S] = Var[N]B^2$.

В следующих трех параграфах рассматривается обобщенное распределение, используя различные распределения для числа исков, N.

3.2 Обобщенное распределение Пуассона

Рассмотрим сначала величину суммарных исков, когда N имеет распределение Пуассона с математическим ожиданием λ. Обозначим $N \sim P(\lambda)$. S тогда будет иметь обобщенное распределение Пуассона с параметрами λ и $F(x)$. Для распределения Пуассона для N известно:

$$E[N] = Var[N] = \lambda$$

$$M_N(t) = \exp\{\lambda(e^t - 1)\}$$

Эти результаты представлены в таблице.

Эти равенства можно объединить с результатами, полученными в параграфе 3.1,

используя (3.1): $$\text{E[S]}=\lambda m_1 \tag{6.3}$$

используя (3.1): $$\text{Var[S]}=\lambda m_2 \tag{6.3}$$

используя (3.1): $$\mathrm{M}_S(t) = \exp\{\lambda(M_X(t) - 1)\} \tag{6.3}$$

Эти формулы для среднего значения и дисперсии имеют очень простой вид. Заметим, что дисперсия S выражается через второй центральный момент X_i, а не через дисперсию X_i.

Для того, чтобы показать, что центральный момент S равен λm_3, воспользуемся производящей функцией семиинвариантов (формула (6.2) для $n = 3$), то есть:

$$E[(S - \lambda m_1)^3] = \frac{d^3}{dt^3} \log M_S(t)|_{t=0}$$

Так как $\log M_S(t) = \lambda(M_X(t) - 1)$, то:

$$\frac{d^3}{dt^3} \log M_S(t)|_{t=0} = \lambda \frac{d^3}{dt^3} M_X(t)|_{t=0} = \lambda m_3$$

то есть $E[(S - \lambda m_1)^3] = \lambda m_3$
и коэффициент асимметрии равен $\lambda m_3 / (\lambda m_2)^{3/2}$.

Этот результат показывает, что распределение S имеет положительную асимметрию, т.к. m_3 – третий центральный момент X_i и, следовательно, больше нуля, т.к. X_i принимает неотрицательные значения. Заметим, что распределение S имеет положительную асимметрию даже тогда, когда распределение X_i имеет отрицательную асимметрию. Коэффициент асимметрии S равен $\lambda m_3 / (\lambda m_2)^{3/2}$, а следовательно, стремится к нулю при $\lambda \longrightarrow \infty$. Таким образом, для больших значений λ распределение S почти симметрично.

Рассмотрим следующее свойство обобщенного распределения Пуассона. Сумма независимых случайных величин, имеющих распределение Пуассона, является случайной величиной с обобщенным распределением Пуассона. Формальное определение этого свойства следующее:

Пусть $S_1, S_2, ..., S_n$ — независимые случайные величины. S_i имеет обобщенное распределение Пуассона с параметрами λ_i и $F_i(x)$. Определим следующую случайную величину $A = S_1 + S_2 + ... + S_n$. Тогда A будет иметь обобщенное распределение Пуассона с параметрами Λ и $F(x)$, где

$$\Lambda = \sum_{i=1}^{n} \lambda_i \quad \text{и} \quad F(x) = \frac{1}{\Lambda} \sum_{i=1}^{n} \lambda_i F_i(x)$$

Полученный результат представляет собой очень важное свойство. В дальнейшем оно нам пригодиться. Обозначим его как Результат 1.

Для доказательства этого свойства, прежде всего, заметим, что $F(x)$ — среднее взвешенное функций распределения и эти веса принимают положительные значения, а в сумме дают единицу. Это означает, что $F(x)$ — является функцией распределения и это распределение имеет производящую функцию моментов.

$$M(t) = \int\limits_0^\infty exp\{tx\}\frac{1}{\Lambda}\sum_{i=1}^{n}\lambda_i f_i(x)dx$$

где $f_i(x)$ – плотность $F_i(x)$. Следовательно:

$$M(t) = \frac{1}{\Lambda}\sum_{i=1}^{n}\lambda_i\int\limits_0^\infty exp\{tx\}f_i(x)dx = \frac{1}{\Lambda}\sum_{i=1}^{n}\lambda_i M_i(t) \tag{6.3}$$

где $M_i(t)$ – производящая функция моментов распределения $F_i(t)$.

Пусть $M_A(t)$ – производящая функция моментов A.

Тогда $M_A(t) = E[\exp\{tA\}] = E[\exp\{tS_1 + tS_2 + ... + tS_n\}]$.

Из независимости $\{S_i\}_{i=1}^n$:

$$M_A(t) = \prod_{i=1}^{n} E[\exp\{tS_i\}] = \exp\{\lambda_i(M_i(t) - 1)\}$$

Таким образом:

$$M_A(t) = \prod_{i=1}^{n}\exp\{\lambda_i(M_i(t) - 1)\} = \exp\{\prod_{i=1}^{n}\lambda_i(M_i(t) - 1)\}$$

То есть:

$$M_A(t) = \exp\{\Lambda(M(t) - 1)\} \tag{6.3}$$

где $\Lambda = \sum\limits_{i=1}^{n}\lambda_i$ и $M(t) = \frac{1}{\Lambda}\sum\limits_{i=1}^{n}\lambda_i M_i(t)$.

Согласно взаимно-однозначному соответствию между производящими функциями моментов, формула (3.2) показывает, что A имеет обобщенное распределение Пуассона с параметром Λ. Из (3.2) распределение величины индивидуального иска равно $F(t)$.

3.3 Обобщенное биномиальное распределение

В некоторых случаях в качестве распределения N целесообразно выбрать биномиальное распределение. Например, согласно группе полисов страхования n жизней, число смертей в год имеет биномиальное распределение, если предположить, что каждая застрахованная жизнь подчиняется одному закону смертности и все они независимы относительно смертности.

Для биномиального распределения N будем пользоваться обозначением $N \sim b(n,q)$. Основными результатами для этого распределения являются следующие формулы:

$$E[N] = nq$$

$$Var[N] = nq(1-q)$$

$$M_N(t) = (qe^t + 1 - q)^n$$

Эти формулы представлены в таблице.

Если N имеет биномиальное распределение, то S имеет обобщенное биномиальное распределение. Важным моментом относительно выбора биномиального распределения для N является то, что количество исков имеет верхний предел, n.
Запишем выражения для среднего, дисперсии и производящей функции через n, q, m_1, m_2 и $M_X(t)$, если $N \sim b(n,q)$.

С помощью формул (3.1) и (3.1) получим:

$$E[S] = nqm_1 \qquad (6.3)$$

$$Var[S] = nq(m_2 - m_1^2) + nq(1-q)m_1^2 = nqm_2 - nq^2m_1^2 \qquad (6.3)$$

Из формулы (3.1):

$$M_S(t) = (qM_X(t) + 1 - q)^n$$

Третий центральный момент найдём с помощью производящей функции семиинвариантов (используя формулу (6.2)).

$$\frac{d^3}{dt^3}\log M_S(t) = \frac{d^3}{dt^3} n\log(qM_X(t)+p) = \{\text{ где } p = 1-q\} =$$

$$= \frac{d^2}{dt^2}\left\{nq\left(\frac{d}{dt}M_X(t)\right)(qM_X(t)+p)^{-1}\right\} =$$

$$= \frac{d}{dt}\left\{nq\left(\frac{d^2}{dt^2}M_X(t)\right)(qM_X(t)+p)^{-1} - n\left(q\frac{d}{dt}M_X(t)\right)^2(qM_X(t)+p)^{-2}\right\} =$$

$$= nq\left(\frac{d^3}{dt^3}M_X(t)\right)(qM_X(t)+p)^{-1} -$$

$$-3nq^2\left(\frac{d^2}{dt^2}M_X(t)\right)(qM_X(t)+p)^{-2}\left(\frac{d}{dt}M_X(t)\right) +$$

$$+2n\left(q\frac{d}{dt}M_X(t)\right)^3(qM_X(t)+p)^{-3}$$

Положим $t = 0$:

$$E[(S - nqm_1)^3] = nqm_3 - 3nq^2m_2m_1 + 2nq^3m_1^3 \tag{6.3}$$

Коэффициент асимметрии, следовательно, равен:

$$\frac{nqm_3 - 3nq^2m_2m_1 + 2nq^3m_1^3}{(nqm_2 - nq^2m_1^2)^{3/2}}$$

Из формулы (3.3) следует, что обобщенное биномиальное распределение может иметь отрицательную асимметрию. Простой пример, иллюстрирующий этот факт, когда все иски имеют размер B. Тогда $S = BN$ и

$$E[(S - E[S])^3] = B^3E[(N - E[N])^3]$$

таким образом, коэффициент асимметрии S кратен соответствующему коэффициенту асимметрии N. Если $q > 0.5$, то биномиальное распределение N имеет отрицательную асимметрию.

3.4 Обобщенное отрицательное биномиальное распределение

Последнее распределение для N, которое мы рассмотрим, — отрицательное биномиальное распределение, которое имеет вероятностную функцию:

$$P(N = n) = \binom{k+n-1}{n}p^kq^n \quad \text{для } n = 0, 1, 2, ...$$

Параметры распределения $k > 0$ и p, причем $p + q = 1$ и $0 < p < 1$. Для этого распределения введем обозначение $NB(k,p)$. Если $N \sim NB(k,p)$:

$$E[N] = kq/p$$
$$Var[N] = kq/p^2$$
$$M_N(t) = p^k(1 - qe^t)^{-k}$$

Частный случай при $k = 1$ дает геометрическое распределение. Еще раз отметим, что эти результаты представлены в Таблице.

Отрицательное биномиальное распределение является альтернативным для распределения Пуассона для величины N. Но отрицательное биномиальное распределение имеет преимущество перед распределением Пуассона, т.к. дисперсия этого распределения больше математического ожидания. Эти две величины равны друг другу для распределения Пуассона. Таким образом, отрицательное биномиальное распределение дает лучшее соответствие набору данных, которые имеют выборочную дисперсию, превышающую выборочное среднее. На практике, такая ситуация встречается часто. В параграфе 6.2 приведена ситуация, когда применяется отрицательное биномиальное распределение N. Если N имеет отрицательное биномиальное распределение, то S имеет обобщенное отрицательное биномиальное распределение.

Выражения для среднего значения, дисперсии и производящей функции моментов S, если $N \sim NB(k,p)$, вытекают непосредственно из формул (3.1),(3.1) и (3.1):

$$E[S] = \frac{kq}{p}m_1$$
$$Var[S] = \frac{kq}{p}(m_2 - m_1^2) + \frac{kq}{p^2}m_1^2 = \frac{kq}{p}m_2 + \frac{kq^2}{p^2}m_1^2$$
$$M_S(t) = \frac{p^k}{(1 - qM_X(t))^k}$$

Как и раньше, третий центральный момент S может быть найден, используя производящую функцию семиинвариантов S, следующим образом:

$$\frac{d}{dt}\log M_S(t) = \frac{d}{dt}(k\log p - k\log(1 - qM_X(t))) =$$
$$= \frac{kq}{1 - qM_X(t)}\left(\frac{d}{dt}M_X(t)\right)$$

Тогда

$$\frac{d^2}{dt^2}\log M_S(t) = kq^2\left(\frac{d}{dt}M_X(t)\right)^2\frac{1}{(1-qM_X(t))^2}+\frac{kq}{1-qM_X(t)}\left(\frac{d^2}{dt^2}M_X(t)\right)$$

и

$$\frac{d^3}{dt^3}\log M_S(t) = 3kq^2\left(\frac{d}{dt}M_X(t)\right)\left(\frac{d^2}{dt^2}M_X(t)\right)\frac{1}{(1-qM_X(t))^2}+$$
$$+\frac{2kq^3}{(1-qM_X(t))^3}\left(\frac{d}{dt}M_X(t)\right)^3+\frac{kq}{1-qM_X(t)}\left(\frac{d^3}{dt^3}M_X(t)\right)$$

Положим $t = 0$ для третьей производной и получим:

$$E[(S-E[S])^3] = \frac{3kq^2m_1m_2}{p^2}+\frac{2kq^3m_1^3}{p^3}+\frac{kqm_3}{p} \tag{6.3}$$

Параметры k и p положительны, как и моменты $F(x)$. Тогда из формулы (3.4) следует, что обобщенное отрицательное биномиальное распределение имеет положительную асимметрию. Коэффициент асимметрии может быть найден по формуле $E[(S-E[S])^3]/(Var[S])^{3/2}$.

3.5 Распределение суммарного иска согласно договору перестрахования эксцедента убытка и пропорциональному договору перестрахования

Пропорциональное перестрахование

Распределение числа исков в отношении перестраховщика такое же, какое и распределение числа исков страховщика, так как каждый платит определенную пропорцию каждого иска. Для уровня удержания, равного α $(0 \preceq \alpha \preceq 1)$, величина индивидуального иска для страховщика распределена как αX_i и для перестраховщика — как $(1-\alpha)X_i$. Величины суммарного иска распределены как αS и $(1-\alpha)S$ соответственно.

Перестрахование эксцедента убытка

Величина выплат страховщика по каждому i-му иску, согласно договору перестрахования эксцедента убытка с уровнем удержания M, равна $Y_i = \min(X_i, M)$.

Величина выплат перестраховщика равна $Z_i = \max(0, X_i - M)$.

Таким образом, величина суммарного нетто-иска страховщика может быть представлена следующим образом:

$$S_I = Y_1 + Y_2 + ...Y_N$$

и суммарного иска перестраховщика:

$$S_R = Z_1 + Z_2 + ...Z_N$$

Если, например, $N \sim P(\lambda)$, то S_I имеет обобщенное распределение Пуассона с параметром Пуассона λ, а величины индивидуальных исков распределены как Y_i. Аналогично, S_R имеет обобщенное распределение Пуассона с параметром Пуассона, равным λ, и величины индивидуальных исков распределены как Z_i. Отметим, однако, если $F(M) > 0$, что будет часто встречаться, то вероятность (то есть $F(M)$) того, что $Z_i = 0$, отлична от нуля. Другими словами, 0 считается возможным иском для перестраховщика. С практической точки зрения, такое определение S_R весьма искусственно. Страховщик знает наблюдаемую величину числа исков N, но перестраховщик знает только о тех исках, величина которых превысила уровень M, т.к. страховщик может уведомить перестраховщика только о наличии исков, превышающих M.

Пример 6.16 Величина ежегодного суммарного иска имеет обобщенное распределение Пуассона с параметром 10. Величины индивидуальных исков равномерно распределены на (0,2000). Действует договор перестрахования эксцедента убытка с уровнем превышения 1,600. Вычислите среднее значение, дисперсию и коэффициент асимметрии суммарного иска для страховщика и для перестраховщика, согласно действующему договору.

Решение Пусть S_I и S_R — величины, определенные выше. Для того, что найти $E[S_I]$, вычислим $E[Y_i]$:

$$E[Y_i] = \int_0^M x f(x)dx + MP(X_i > M)$$

где $f(x) = 0.0005$ — плотность распределения $U(0, 2000)$ и $M = 1,600$. Учитывая это, получим:

$$E[Y_i] = \frac{0.0005x^2}{2} \Big|_0^M + 0.2M = 960$$

и значит

$$E[S_I] = 10E[Y_i] = 9,600$$

Чтобы найти $Var[S_I]$, вычислим $E[Y_i^2]$.

$$E[Y_i^2] = \int_0^M x^2 f(x)dx + M^2 P(X_i > M) =$$

$$= \frac{0.0005x^3}{3} \Big|_0^M + 0.2M^2 = 1,194,666.7$$

и значит

$$Var[S_I] = 10E[Y_i^2] = 11,946,667$$

Чтобы найти коэффициент асимметрии величины исков страховщика, вычислим $E[Y_i^3]$:

$$E[Y_i^3] = \int\limits_0^M x^3 f(x)dx + M^3 P(X_i > M) =$$

$$= \frac{0.0005x^4}{4} \Big|_0^M + 0.2M^3 = 1,638,400,000$$

Таким образом:

$$E[(S_I - E[S_I])^3] = 10E[Y_i^3] = 16,384,000,000$$

и коэффициент асимметрии равен

$$16,384,000,000/(11,946,667)^{3/2} = 0.397$$

Для того, чтобы найти $E[S_R]$, заметим, что ожидаемая величина суммарного иска за весь риск равна 10,000. Тогда

$$E[S_R] = 10,000 - E[S_I] = 400$$

Чтобы найти $Var[S_R]$, вычислим $E[Z_i^2]$:

$$E[Z_i^2] = \int\limits_M^{2000} (x - M)^2 f(x)dx = \{y = x - M\} = \int\limits_0^{2000-M} 0.0005y^2 dy =$$

$$= \frac{0.0005y^3}{3} \Big|_0^{200-M} = 10,666.7$$

Таким образом

$$Var[S_R] = 10E[Z_i^2] = 106,667$$

Для того, чтобы найти коэффициент асимметрии величины исков для перестраховщика, вычислим $E[Z_i^3]$:

$$E[Z_i^3] = \int\limits_M^{2000} (x - M)^3 f(x)dx = \{y = x - M\} =$$

$$= \int\limits_0^{2000-M} 0.0005y^3 dy = 3,200,000$$

и следовательно

$$E[(S_R - E[S_R])^3] = 10E[Z_i^3] = 32,000,000$$

коэффициент асимметрии равен

$$32,000,000/(106,667)^{3/2} = 0.92$$

Суммарные иски перестраховщика могут быть также представлены следующим образом:

$$S_R = W_1 + W_2 + ...W_{NR} \tag{6.3}$$

где случайная величина NR обозначает число действительных (ненулевых) выплат, производимых перестраховщиком.

Например, предположим, что по риску из прошлого примера за конкретный год было предъявлено восемь требований, следующих размеров:

403 1490 1948 443 1866 1704 1221 823

Тогда в формуле (**??**) наблюдаемая величина N равна 8, третий, пятый и шестой иски приводят к выплатам перестраховщика следующих размеров 348, 266 и 104 соответственно. "Выплаты"перестраховщика по остальным пяти искам равны 0.

В формуле (3.5) наблюдаемая величина NR равна 3, а величины W_1, W_2 и W_3 равны 348, 266 и 104 соответственно. Отметим, что величина S_R одинакова (то есть равна 718) для двух представлений.

Во II Части было показано, что W_i имеет следующую плотность распределения:

$$p(w) = \frac{f(w+M)}{1-F(M)}, \quad w > 0$$

Для распределения S_R виде (3.5) необходимо знать распределение величины NR. Его можно найти следующим образом. Определим

$$NR = I_1 + I_2 + ...I_N$$

где N обозначает число исков по данному риску (как и обычно). I_j — индикаторная случайная величина, которая принимает значение 1, если перестраховщик совершает (ненулевые) выплаты по j-му иску, и значение 0 в другом случае. Таким образом, NR дает число выплат, совершаемых перестраховщиком. Так как I_j принимает значение 1 только если $X_j > M$, то:

$$P(I_j = 1) = P(X_j > M) = \pi \text{ (так обозначим), и}$$

$$P(I_j = 0) = 1 - \pi$$

Кроме того, I_j имеет производящую функцию моментов:

$$M_I(t) = \pi \exp\{t\} + 1 - \pi$$

и по формуле (3.1) NR имеет производящую функцию моментов:

$$M_{NR}(t) = M_N(\log M_I(t))$$

Пример 6.17 Продолжая предыдущий пример и используя формулу (3.5) в качестве представления случайной величины S_R, можно увидеть, что S_R имеет обобщенное распределение Пуассона с параметром Пуассона, равным 0.2*10=2. Индивидуальные иски, W_i, имеют следующую плотность распределения:

$$p(w) = \frac{f(w+M)}{1-F(M)} = 0.0005/0.2 = 0.0025, \quad \text{для} \quad 0 < w < 400$$

т.е. W_i равномерно распределены на (0,400). $E[W_i] = 200$, $E[W_i^2]53,333.33$ и $E[W_i^3] = 16,000,000$ дают такие же результаты, что и раньше.
Таким образом, существует два способа задавать и устанавливать распределение S_R.

§4 Точные и приближенные вычисления $G(x)$ в модели коллективного риска

4.1 Введение

В этом параграфе будут проведены вычисления и приближенное представление $G(x)$, функции распределения суммарного иска для модели коллективного риска. В некоторых случаях возможно довольно легко найти распределение функции $G(x)$, например, если все иски одного размера, но предположения, сделанные в этих случаях, будут слишком ограничивающими, чтобы представлять интерес с практической точки зрения. В параграфе 4.2 будет выведена рекурсивная формула для вычисления $G(x)$. В параграфе 4.3 функция $G(x)$ будет аппроксимирована нормальным распределением. И наконец, в параграфе 4.4 $G(x)$ будет аппроксимирована (смещенным) гамма-распределением. В параграфе 4.2 предполагается, что распределение числа исков и распределение величины исков известны; в параграфах 4.3 и 4.4 предполагается, что только первые два или три момента этих распределений известны.

4.2 Рекурсивная формула для $G(x)$

В этом параграфе будем предполагать, что распределение индивидуальных исков, $F(x)$, дискретное распределение, принимающее положительные целые значения. Это означает, что индивидуальные иски могут принимать следующие значения: 1,2,3,... и, следовательно, возможные значения суммарного иска: 0,1,2,3,... . Кроме того, функция распределения для величины индивидуальных исков не имеет плотности распределения (т.к. каждая величина X_i – дискретная случайная величина).

Следующие обозначения будут использоваться для вероятностных функций величины индивидуального иска и суммарного иска соответственно.

$$f_k = P(X_i = k) \qquad k = 1, 2, 3, ...$$

$$g_k = P(S = k) \qquad k = 0, 1, 2, ...$$

Если распределение индивидуальных исков не является дискретным с положительными целыми значениями (что часто встречается), то оно может быть всегда аппроксимировано подходящим распределением.

Задача вычисления $G(x)$ согласно сделанным предположениям сводится к вычислению g_k для $k \leq x$. Предполагается, что:

- распределение числа исков известно
- распределение индивидуальных исков (т.е. f_k) известно

Перед доказательством рекурсивной формулы, или даже до формулировки этой формулы, необходимо сделать одно предположение относительно распределения N, числа исков. Обозначим $P(N = r)$ через p_r и предположим, что существуют такие a и b, что:

$$p_r = (a + b/r)p_{r-1} \qquad \text{для } r = 1, 2, 3, ... \tag{6.3}$$

Все три распределения числа исков, рассмотренные в главе 3, удовлетворяют Предположению (4.2).
Формула для g_r:

$$g_0 = p_0 \tag{6.3}$$

$$g_r = \sum_{j=1}^{r}(a + bj/r) f_j g_{r-j} \qquad \text{для } r = 1, 2, ... \tag{6.3}$$

Формула (4.2) следует из того факта, что минимальный размер иска равен 1. Суммарный иск равен нулю только в том случае, если не было предъявлено ни одного иска.

Для доказательства формулы (4.2) будем использовать следующие три формулы, для $n = 2, 3, ...$:

$$E\left[X_1 | \sum_{i=1}^{n} X_i = r\right] = r/n \tag{6.3}$$

$$E\left[X_1 | \sum_{i=1}^{n} X_i = r\right] = \sum_{j=1}^{r} j\ f_j\ f_{r-j}^{(n-1)*}/f_r^{n*} \tag{6.3}$$

$$p_n f_r^{n*} = \sum_{j=1}^{r-1}(a + bj/r)\ f_j\ p_{n-1}\ f_{r-j}^{(n-1)*} \tag{6.3}$$

Формулы (4.2) и (4.2) справедливы для любых значений r, для которых f_r^{n*} не равно нулю; (4.2) справедлива для $r = 1, 2, ...$, в любом случае.
Формула (4.2) следует из того, что $X_1, X_2, ..., X_n$ одинаково распределены, и если их сумма равна r, то ожидаемое значение любого из них должно быть равно r/n.

Чтобы увидеть, почему справедлива формула (4.2), заметим, что $f_j\ f_{r-j}^{(n-1)*}/f_r^{n*}$ — (условная) вероятность того, что X_1 равно j при условии, что $\sum\limits_{j=1}^{n} X_j = r$. (В (4.2) предполагается, что вероятность того, что $\sum\limits_{i=1}^{n} X_i = r$, т.е. f_r^{n*}, не равна нулю.) Учитывая, что $\sum\limits_{j=1}^{n} X_j = r$, значение X_1 не может быть больше r. Следовательно, правая часть (4.2) — есть сумма по каждому значению X_1 этого значения, умноженного на вероятность того, что X_1 принимает это значение, условие равенства r величины $\sum\limits_{j=1}^{n} X_j$. Эта сумма равняется левой части (4.2).

Теперь получим (4.2). Во-первых, заметим, что (4.2) справедлива, если f_r^{n*} равно нулю, т.к. в этом случае для любых значений $j = 1, 2, ..., r$, либо одна из величин f_j или $f_{r-j}^{(n-1)*}$ равна нулю, либо обе должны быть равны нулю. Следовательно, если f_r^{n*} равна нулю, обе части (4.2) равны нулю. Теперь предположим, что f_r^{n*} не равно нулю. Тогда

используя (4.2) $\qquad p_n f_r^{n*} = p_{n-1}(a + b/n) f_r^{n*} =$

используя (4.2) $\qquad = p_{n-1} E[a + bX_1/r | \sum_{i=1}^{n} X_i = r] f_r^{n*} =$

используя (4.2) $\qquad = p_{n-1} \sum_{j=1}^{r} (a + bj/r) f_j f_{r-j}^{(n-1)*} =$

$$= p_{n-1} \sum_{j=1}^{r-1} (a + bj/r) f_j f_{r-j}^{(n-1)*} \qquad (\text{т.к. } f_0^{(n-1)*} = 0)$$

Наконец, теперь можем получить (4.2). Для $r = 1, 2, ...$

используя (3.1) $\qquad g_r = \sum_{n=1}^{\infty} p_n f_r^{n*} =$

$$= p_1 f_r + \sum_{n=1}^{\infty} p_{n+1} f_r^{(n+1)*} =$$

используя (4.2) $\qquad = (a + b) p_0 f_r + \sum_{n=1}^{\infty} \sum_{j=1}^{r-1} (a + bj/r) f_j p_n f_{r-j}^{n*} =$

$$= (a + b) g_0 f_r + \sum_{j=1}^{r-1} (a + bj/r) f_j \sum_{n=1}^{\infty} p_n f_{r-j}^{n*} =$$

используя (3.1) $\qquad = (a + b) g_0 f_r + \sum_{j=1}^{r-1} (a + bj/r) f_j g_{r-j} =$

$$= \sum_{j=1}^{r} (a + bj/r) f_j g_{r-j}$$

Формула (4.2) доказана.

В случая, когда N имеет распределение Пуассона ($a = 0$ и $b = \lambda$), формулы имеют следующий вид:

$$g_0 = e^{-\lambda}$$

$$g_r = \frac{\lambda}{r} \sum_{j=1}^{r} j \ f_j \ g_{r-j}$$

4.3 Аппроксимация $G(x)$ нормальным распределением

Рекурсивная формула для вычисления $G(x)$, доказанная в параграфе 4.2, является очень полезным "инструментом но она имеет некоторые недостатки. Во-первых, она может потребовать значительного времени для вычисления значений для $G(x)$ на компьютере. Во-вторых, её нельзя использовать, если распределения N и X_i не известны, или, по крайней мере, не могут быть оценены довольно точно.
В этом параграфе предполагается, что всё, что известно (или может быть довольно точно оценено), насчет S — это математическое ожидание и дисперсия этой случайной величины. Так как довольно много различных распределений имеют одинаковые средние значения и дисперсии, $G(x)$ нельзя найти только по этой информации. Один из способов аппроксимировать $G(x)$ в этой ситуации — предположить, что S аппроксимировано нормальным распределением.

Более формально, пусть $\Phi(z)$ – функция распределения нормально распределенной случайной величины со средним значением 0 и дисперсией 1. Таким образом:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{z} \exp\{-x^2/2\} dx$$

Теперь пусть μ и σ^2 обозначают среднее и дисперсию S. По предположению в этом параграфе, что S аппроксимирована нормальным распределением со средним значением μ и дисперсией σ^2, получим, что для любого x:

$$G(x) = P(S \leq x) = P((S-\mu)/\sigma \leq (x-\mu)/\sigma) \approx \Phi((x-\mu)/\sigma)$$

Для нормального распределения легко получить значения вероятностей.

S — сумма случайного числа независимых одинаково распределенных случайных величин. По Центральной Предельной Теореме можно использовать аппроксимацию нормальным распределением. Чем больше (ожидаемое) значение N (число случайных величин), тем лучше будет это приближение.

4.4 Аппроксимация $G(x)$ смещенным гамма-распределением

Теперь предположим, что нам известны, или могут быть достаточно точно оценены, первые три момента S. Другой способ аппроксимации распределения S — приближение смещенным гамма-распределением. Пусть μ, σ^2 и β обозначают среднее, дисперсию и коэффициент асимметрии S соответственно. Для аппроксимации смещенным гамма-распределением предположим, что S имеет приближенно то же самое распределение, что и случайная величина $Y + k$, где k – некоторая константа и Y имеет гамма-распределение с параметрами α и δ. Параметры k,α и δ подобраны так, чтобы величина $k + Y$ имела такие же первые три момента, что и S. Отметим, что $k + Y$ — это случайная величина Y, которая имеет простое гамма-распределение, смещенная на положительную или отрицательную величину k. Одна из причин, почему смещенное гамма-распределение дает приближение лучше, чем нормальная аппроксимация, — это положительная асимметрия у смещенного гамма-распределения, которая часто встречается на практике.

Коэффициент асимметрии, β, гамма-распределения с параметрами α и δ равен $2/\sqrt{\alpha}$.

Приравнивая коэффициенты асимметрии, средние значения и дисперсии S и $k + Y$, получим следующие формулы:

$$\beta = 2/\sqrt{\alpha}$$

$$\sigma^2 = \alpha/\delta^2$$

$$\mu = k + \alpha/\delta$$

из которых α, δ и затем k могут быть найдены через известные величины: β, σ^2, μ.

Основание для аппроксимации распределения S смещенным гамма, или нормальным, распределением — это то, что значения вероятностей таких как $P(a < k+Y < b)$ могут быть легче найдены, чем $P(a < S < b)$. Вероятности для гамма-распределения легко получить с помощью многих статистических компьютерных пакетов.

В некоторых простейших случаях возможно оценить вероятности для гамма-распределения из таблиц вероятностей для распределения χ^2. Если Y имеет $gamma(\alpha, \delta)$–распределение и 2α–целое число, то $2\delta Y$

имеет $\chi^2_{2\alpha}$–распределение. Это свойство может использоваться, даже если 2α не является целым числом, т.к. можно интерполировать между соседними целыми числами.

§5 Модель индивидуального риска

Согласно этой модели, рассматривается портфель, состоящий из фиксированного числа рисков. Это предполагает следующее:

- эти риски независимы
- величины исков по данным рискам не являются одинаково распределенными случайными величинами
- число рисков не изменяется в течение всего периода страхования.

Как и прежде, суммарных иск по данному портфелю обозначим за S. Таким образом:

$$S = Y_1 + Y_2 + ... + Y_n$$

где Y_j обозначает величину иска по j-му риску и n обозначает число рисков. Возможно, что по некоторым искам не предъявлено ни одного иска. Следовательно, некоторые наблюдаемые величины из $\{Y_j\}_{j=1}^n$ могут быть равны нулю.

Для каждого риска сделаем следующие предположения:

– количество исков по j-му риску, N_j, равно 0 или 1 (6.3)

– вероятность возникновения иска по j-му риску равна q_j (6.3)

Если по j-му риску предъявлен иск, то величину иска обозначим за X_j. Пусть $F_j(x), \mu_j$ и σ_j^2 обозначают функцию распределения, среднее и дисперсию X_j соответственно.
Предположение (§5) весьма ограничивающее. Это означает, что по каждому риску может быть предъявлен максимум один иск. Эта модель включает в себя такие риски, как одногодичное страхование, но не включает многие другие распространенные типа страхования. Например, при страхование жилищного имущества нет ограничений на число предъявленных исков за год.

Между моделью коллективного риска и моделью индивидуального риска существуют три важных различия:

(1) Количество рисков в портфеле определено по-разному. В модели коллективного риска нет необходимости определять это число, так как оно фиксировано в течение всего периода страхования (но не тогда, когда предполагается, что $N \sim b(n, q)$).

(2) Число исков по каждому индивидуальному риску ограничено. В модели индивидуального риска это число неограничено.

(3) Предполагается, что индивидуальные риски независимы. В модели коллективного риска независимыми являются величины индивидуальных исков.

Из предположений (§5) и (§5) следует, что $N_j \sim b(1, q_j)$. Таким образом, распределение Y_j — обобщенное биномиальное, в котором индивидуальные иски распределены как X_j. Из формул (3.3) и (3.3) следует:

$$E[Y_j] = q_j \mu_j \tag{6.3}$$

$$Var[Y_j] = q_j \sigma_j^2 + q_j(1 - q_j)\mu_j^2 \tag{6.3}$$

S — сумма n независимых случайных величин, имеющих обобщенное биномиальное распределение. В параграфе 3.4 было замечено, что нет общего результат для распределения такой суммы. Это распределение может быть установлено только тогда, когда обобщенные биномиальные величины одинаково распределены, а также независимы. Можно, хотя и затруднительно, найти функцию распределения S по определенным условиям. Однако, легко найти среднее и дисперсию S.

$$E[S] = E[\sum_{j=1}^{n} Y_j] = \sum_{j=1}^{n} E[Y_j] = \sum_{j=1}^{n} q_j \mu_j \tag{6.3}$$

Предположение о независимости индивидуальных рисков необходимо для:

$$Var[S] = Var[\sum_{j=1}^{n} Y_j] = \sum_{j=1}^{n} Var[Y_j] = \sum_{j=1}^{n} (q_j \sigma_j^2 + q_j(1 - q_j)\mu_j^2) \tag{6.3}$$

В случае, когда $\{Y_j\}_{j=1}^{n}$ — последовательность одинаково распределенных, а также независимых, случайных величин, тогда для каждого полиса величины q_j, μ_j и σ_j^2 одинаковы, обозначим их за q, μ и σ^2. Кроме того, $F_j(x)$ не зависит от j, обозначим за $F(x)$. Следовательно, S имеет обобщенное биномиальное распределение с биномиальными параметрами n и q, индивидуальные иски имеют функцию распределения $F(x)$. Этот случай сводится к модели коллективного риска и из (§5), (§5) видно, что:

$$E[S] = nq\mu$$

$$Var[S] = nq\sigma^2 + nq(1-q)\mu^2$$

которые соответствуют (3.3) и (3.3) соответственно.

§6 Параметр изменчивость/неопределенность

6.1 Введение

До сих пор рисковые модели изучались, предполагая, что параметры (а именно моменты и в некоторых случаях даже распределения) числа исков и величины индивидуального иска точно известны. Вообще говоря, эти параметры не известны, но должны быть оценены с помощью подходящих наборов данных. В этом параграфе мы рассмотрим как вводятся модели до того, как их можно будет доопределить, учитывая параметр изменчивости/разброса. Для этого рассмотрим несколько примеров. Большинство, но не все, из этих примеров будут рассматривать разброс в распределении числа исков, т.к. ему уделяется большее внимание в актуарной литературе, чем распределению индивидуальных исков. Все рассматриваемые примеры основываются на том, что число исков имеет распределение Пуассона.

6.2 Неопределенность в неоднородном портфеле

Рассмотрим портфель, содержащий n независимых полисов. Суммарные иски по i-му полису обозначим за S_i, где S_i имеет обобщенное распределение Пуассона с параметрами λ_i и $F(x)$. Будем считать для простоты, что распределение индивидуальных исков, $F(x)$, одинаково для всех полисов. В этом примере распределение индивидуальных исков, т.е. $F(x)$, предполагается известным, но значения параметров Пуассона, λ_i, не известны. Случайные величины λ_i предполагаются независимыми с одинаковым (известным) распределением. Другими словами, $\{\lambda_i\}_{i=1}^n$ трактуется как набор независимых одинаково распределенных случайных величин с известным распределением. Это означает, что если полис выбран из портфеля случайным образом, то параметр Пуассона для этого полиса не известен, но можно сделать вероятностные предположения о нем. Например, "вероятность 50%, что значение параметра Пуассона лежит между 3 и 5". Это важно для того, чтобы понять, что параметр Пуассона для выбранного из портфеля полиса – фиксированная величина; проблема состоит в том, что это величина нам не известна.

Пример 6.18 Предположим, что параметры Пуассона для портфеля полисов страхования не известны, но с равной вероятностью могут принимать два значения: 0.1 или 0.3.

(i) Найти среднее и дисперсию (через m_1 и m_2) суммарного иска по выбранному случайным образом из портфеля полису.

(ii) Найти среднее и дисперсию (через m_1, m_2 и n) суммарного иска по всему портфелю.

Пусть данная модель описывает автострахование. Все полисы по всему портфелю были разбиты на группы с учетом таких показателей как: "возраст водителя "тип автомобиля"и даже "статистика предъявленных исков в прошлом". Полисы каждой группы из портфеля имеют идентичные значения по каждому из этих показателей. Однако, существуют некоторые показатели, такие как "способность к вождению которые сложно измерить и поэтому они не могут быть явно учтены. Предполагается, что одни страховщики из данной группы полисов "умелые"водители, а другие — "неумелые"водители. Распределение величины индивидуальных исков для всех водителей из группы одинаковое, но "умелые"водители предъявляют меньше исков (в среднем 0.1 ежегодно), чем "неумелые"водители (в среднем 0.3 в год). Предполагается, что по каким-то данных известно, что страховщик из данной группы с равной вероятностью может быть как "умелым так и "неумелым"водителем, но не известно точно, является ли отдельный страховщик "умелым"или "неумелым"водителем.

Решение Пусть $\lambda_i,\ i = 1, 2, ..., n$ — параметр Пуассона для i-го полиса из портфеля. $\{\lambda_i\}_{i=1}^{\infty}$ — набор независимых одинаково распределенных случайных величин, каждая из которых имеет следующее распределение:

$$P(\lambda_i = 0.1) = 0.5$$

$$P(\lambda_i = 0.3) = 0.5$$

Отсюда следует,что:

$$E[\lambda_i] = 0.2$$

$$Var[\lambda_i] = 0.01$$

(i) Моменты S_i могут быть найдены при условии λ_i. Т.к. $S_i|\lambda_i$ имеет прямое обобщенное распределение Пуассона, то, используя формулы (3.2) и (3.2), получим:

$$E[S_i] = E[E[S_i|\lambda_i]] = E[\lambda_i m_1] = 0.2m_1$$

$$Var[S_i] = E[Var[S_i|\lambda_i]] + Var[E[S_i|\lambda_i]] =$$
$$= E[\lambda_i m_2] + Var[\lambda_i m_1] = 0.2m_2 + 0.01m_1^2$$

(ii) Случайные величины $\{S_i\}_{i=1}^n$ – независимые и одинаково распределенные, каждая с распределением, данным в пункте (i). Следовательно, используя полученные результаты из пункта (i), получим:

$$E\left[\sum_{i=1}^n S_i\right] = nE[S_i] = 0.2nm_1$$

$$Var\left[\sum_{i=1}^n S_i\right] = nVar[S_i] = 0.2nm_2 + 0.01nm_1^2$$

Пример 6.19 Пусть параметры Пуассона имеют гамма-распределение с параметрами α и δ. Найти распределение числа исков по полису, выбранному из портфеля случайным образом.

Решение Пусть N_i обозначают число исков по i-му полису из страхового портфеля и λ_i–их параметры Пуассона. Тогда N_i имеет распределение Пуассона с параметром λ_i, но проблема в том, что (по предположению) значение λ_i нам не известно. А известно — распределение, по которому был выбран параметр λ_i. В итоге, задача может быть поставлена следующим образом:

Учитывая, что:

$$N_i|\lambda_i \sim P(\lambda_i) \quad \text{и} \quad \lambda_i \sim G(\alpha, \delta)$$

найдем безусловное распределение N_i.

Эта задача можно решить, избавляясь от условий стандартным способом:

Для $x = 0, 1, 2, ...$

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

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

Вычислим интеграл, сравнив подынтегральную функцию с плотностью гамма-распределения,и получим, что:

$$P(N_i = x) = \frac{\delta^\alpha}{\Gamma(\alpha)x!}\frac{\Gamma(x+\alpha)}{(\delta+1)^{x+\alpha}}$$

Последнее равенство показывает, что безусловное распределение N_i — отрицательное биномиальное с параметрами α и $\delta/(\delta+1)$.

6.3 Изменчивость в однородном портфеле

Теперь рассмотрим другой пример. Пусть, как и раньше, страховой портфель состоит из n полисов. Суммарный иск по каждому отдельному полису имеет обобщенное распределение Пуассона с параметрами λ и $F(x)$. Эти параметры одинаковы для всех полисов в портфеле. Если значение λ известно, то суммарные иски по различным полисам будут независимых величинами. Предположим, что значение λ не известно, возможно, потому что это значение изменяется от года к году, но есть некоторый вероятностный признак, что значение λ будет находиться в данном диапазоне. Так же как и в предыдущем примере, предположим для простоты, что нам известны моменты и распределение индивидуальных исков, т.е. $F(x)$. Неизвестность значения λ можно представить, как случайную величину λ с известным распределением.

Пример 6.20 Пусть параметр Пуассона, λ, принимает два значения 0.1 или 0.3 с приблизительно одинаковой вероятностью.

(i) Найти среднее и дисперсию (через m_1 и m_2) суммарного иска по выбранному случайным образом из портфеля полису.

(ii) Найти среднее и дисперсию (через m_1, m_2 и n) суммарного иска по всему портфелю.

Решение Используя те же обозначения, что и раньше, пусть S_i обозначает суммарный иск по i-му полису из портфеля. Тогда задача ставится следующим образом:
Случайные величины $\{S_i|\lambda\}_{i=1}^n$ независимые и одинаково распределенные, каждая из которых имеет обобщенное распределение Пуассона с параметрами λ и $F(x)$. Случайная величина λ имеет следующее распределение:

$$P(\lambda = 0.1) = 0.5$$

$$P(\lambda = 0.3) = 0.5$$

(i) При условии λ

$$E[S_i] = E[E[S_i|\lambda]] = E[\lambda m_1] = 0.2m_1$$

$$Var[S_i] = E[Var[S_i|\lambda]] + Var[E[S_i|\lambda]] =$$
$$= E[\lambda m_2] + Var[\lambda m_1] = 0.2m_2 + 0.01m_1^2$$

(ii)

$$E\left[\sum_{i=1}^{n} S_i\right] = nE[S_1] = 0.2nm_1 \qquad (\text{т.к. } \{S_i\}_{i=1}^{n} \text{ одинаково распределены})$$

$$Var\left[\sum_{i=1}^{n} S_i\right] = E\left[Var\left[\sum_{i=1}^{n} S_i|\lambda\right]\right] + Var\left[E\left[\sum_{i=1}^{n} S_i|\lambda\right]\right] =$$
$$= E[n\lambda m_2] + Var[n\lambda m_1] = 0.2nm_2 + 0.01n^2m_1^2$$

Полезно сравнить результаты с прошлыми результатами Примера 6.19. Значения среднего и дисперсии, когда рассматривается отдельно каждый полис (пункт (i)), во двух случаях одинаковы. Различия происходят, когда рассматривают более одного полиса. В этом случае второй пример дает большую дисперсию. Важно понять разницу (а также сходство) между двумя примерами. На практике второй пример может быть применен, если рассматривать полисы портфеля страхования зданий некоторой области. Число исков может зависеть, наряду с другими факторами, от погодных условий в течение года; необычно большое число ураганов может привести к большому ожидаемому числу исков (т.е. большое значение λ) для всех полисов вместе.

6.4 Изменчивость числа исков и величины исков и параметр неопределенности

Этот параграф содержит ещё два примера. Первый — весьма сложный пример, включающий в себя неопределенность как числа исков, так и величин исков.

Пример 6.21 Страховая компания моделирует иски, касающиеся компенсации ущерба при возникновения урагана, согласно полисам страхования имущества, используя следующие предположения.

Число ураганов ежегодно, K, по предположению имеет распределение Пуассона с параметром λ.

Число исков за i-ый ураган, N_i, $i = 1, 2, ..., K$, по предположению имеет распределение Пуассона с параметром Θ_i.

Пусть параметры Θ_i, $i = 1, 2, ..., K$, независимые и одинаково распределенные случайные величины, с $E[\Theta_i] = n$ и $Var[\Theta_i] = s_1^2$.

Величина j-го иска за i-ый ураган,X_{ij}, $j = 1, 2, ..., N_i$, имеет логнормальное распределение с параметрами μ_i и σ, где σ считается известной величиной. Средние значения величин исков,$\Lambda_i = \exp(\mu_i + \sigma^2/2)$, по предположению являются независимыми одинаково распределенными случайными величинами со средним p и дисперсией s_2^2.

Также предположим, что Θ_i и Λ_i независимы.

(i) Показать, что $E[X_{ij}] = p$ и $Var[X_{ij}] = \exp\{\sigma^2\}(p^2 + s_2^2) - p^2$.

(ii) Пусть S_i обозначает суммарный иск в следствии i-го урагана, так что случайная величина $S_i|\{\Theta_i, \Lambda_i\}$ имеет обобщенное распределение Пуассона.
Показать, что $E[S_i] = np$ и $Var[S_i] = (p^2+s_2^2)(n^2+s_1^2+n\exp\{\sigma^2\})-n^2p^2$.

(iii) Найти выражение для среднего и дисперсии ежегодного суммарного иска в следствии всех произошедших ураганов.

Решение

(i)

$$E[X_{ij}] = E[E[X_{ij}|\Lambda_i]] = E[\Lambda_i] = p$$

$$\begin{aligned} Var[X_{ij}] &= E[Var[X_{ij}|\Lambda_i]] + Var[E[X_{ij}|\Lambda_i]] = \\ &= E[\Lambda_i^2(\exp\{\sigma^2\} - 1)] + Var[\Lambda_i] = \\ &= (p^2 + s_2^2)(\exp\{\sigma^2\} - 1) + s_2^2 = \\ &= (p^2 + s_2^2)\exp\{\sigma^2\} - p^2 \end{aligned}$$

(ii)

$$E[S_i] = E[E[S_i|\Theta_i, \Lambda_i]] = E[\Theta_i\Lambda_i] = np$$

т.к. Θ_i и Λ_i независимы.

Т.к. $S_i|\{\Theta_i, \Lambda_i\}$ имеет обобщенное распределение Пуассона, то:

$$Var[S_i|\Theta_i, \Lambda_i] = \Theta_i E[X_{ij}^2|\Lambda_i] = \Theta_i(\Lambda_i \exp\{\sigma^2\})$$

и тогда

$$E[Var[S_i|\Theta_i, \Lambda_i]] = n(p^2 + s_2^2)\exp\{\sigma^2\}$$

А также

$$Var[E[S_i|\Theta_i,\Lambda_i]] = Var[\Theta_i\Lambda_i] = E[\Theta_i^2\Lambda_i^2] - n^2p^2 =$$

$$= (n^2+s_1^2)(p^2+s_2^2) - n^2p^2$$

Объединяя последние два результата, получим:

$$Var[S_i] = (n^2+s_1^2)(p^2+s_2^2) - n^2p^2 + n(p^2+s_2^2)\exp\{\sigma^2\}$$

(iii) Пусть R – случайная величина, обозначающая ежегодный суммарный иск в следствии всех произошедших ураганов. Тогда R может быть представлена следующим образом:

$$R = \sum_{i=1}^{K} S_i$$

где K имеет распределение Пуассона и случайные величины $\{S_i\}$ независимые и одинаково распределенные. Следовательно, R имеет распределение Пуассона и

$$E[R] = \lambda E[S_i] = \lambda np$$

$$Var[R] = \lambda E[S_i^2] = \lambda(Var[S_i] + E[S_i]^2) =$$

$$= \lambda(p^2+s_2^2)(n^2+s_1^2+n\exp\{\sigma^2\})$$

Пример 6.22 Каждый год страховая компания продает большое число полисов страхования жилищного имущества, ежегодная премия по каждому из которых равна £80. Ежегодный суммарный иск по каждому полису имеет обобщенное распределение Пуассона; параметр Пуассона равен 0.4 и величины индивидуальных исков имеют гамма-распределение с параметрами α и λ. Расходы, включенные в выплаты по иску, представляют собой случайную величину, равномерно распределенную между £50 и £b ($>$ £50). Величина расходов не зависит от величины иска, связанного с этими расходами. Случайная величина S представляет общий суммарный иск вместе с расходами за один год по данному портфелю страхования. Можно предположить, что S имеет приблизительно нормальное распределение.

(i) Положим, что

$$\alpha = 1; \quad \lambda = 0.01; \quad b = 100$$

Показать, что компания должна продать по крайней мере 884 полиса в год, чтобы быть хотя бы на 99% уверенной, что величина полученных премий превысит величину исков и расходов за год.

(ii) Предположим, что величины α, λ и b не известны точно, но могут принимать значения из заданных диапазонов:

$$0.95 \leq \alpha \leq 1.05; \quad 0.009 \leq \lambda \leq 0.011; \quad 90 \leq b \leq 110$$

Предполагая, что α, λ и b принимают наихудшие (для страховой компании) значения, найти число полисов, которое должна продать страховая компания, чтобы быть по крайней мере на 99% уверенной, что величина полученных премий превысит величину исков и расходов за год.

Решение Пусть X_i — величина i-го иска и Y_i —величина соответствующих расходов. Пусть N — общее число исков по всему страховому портфелю и пусть n — число полисов в этом портфеле. Тогда N имеет распределение Пуассона с параметром $0.4n$ и S можно записать так:

$$S = \sum_{i=1}^{N} (X_i + Y_i)$$

где $\{X_i + Y_i\}_{i=1}^{\infty}$ – последовательность независимых и одинаково распределенных случайных величин, независимых от N. Из этого можно заключить, что S имеет обобщенное распределение Пуассона, где $(X_i + Y_i)$ представляет "величину i-го индивидуального иска". Используя известные формулы, можем записать выражения для моментов величины S:

$$E[S] = 0.4nE[X_i + Y_i]$$

$$Var[S] = 0.4nE[(X_i + Y_i)^2] =$$
$$= 0.4n(E[X_i^2] + 2E[X_iY_i] + E[Y_i^2])$$

Записывая моменты X_i и Y_i через α, λ и b, получим:

$$E[X_i] = \alpha/\lambda \qquad E[Y_i] = (b + 50)/2$$

$$E[X_i^2] = \alpha(\alpha + 1)/\lambda^2 \qquad E[Y_i^2] = (b^2 + 50b + 2500)/3$$

$$E[X_iY_i] = E[X_i]E[Y_i]$$

Последнее равенство следует из независимости X_i и Y_i.

(i) Положим

$$\alpha = 1; \quad \lambda = 0.01; \quad b = 100$$

Тогда можно записать

$$E[S] = 70n \quad \text{и} \quad Var[S] = 127.80^2 n$$

Следовательно, S имеет приблизительно нормальное распределение со средним значением $70n$ и стандартным отклонением $127.80\sqrt{n}$. Величина полученных премий равна $80n$. Тогда наименьшее возможное значение n найдем из условия:

$$P(S < 80n) \geq 0.99$$

Приводя к нормальному распределению стандартным способом, получим:

$$P\left[\frac{(S-70n)}{127.80\sqrt{n}} < \frac{(80n-70n)}{127.80\sqrt{n}}\right] \geq 0.99$$

Для стандартного нормального распределения 99-я процентиль равна 2.326. Таким образом, получим выражение для n:

$$\frac{(80n-70n)}{127.80\sqrt{n}} \geq 2.326$$

следовательно

$$n \geq 883.7 \qquad \text{(или } n \geq 884\text{, округляя до целого числа сверху).}$$

(ii) Для страховой компании наихудшая комбинация значений для α, λ и b — это комбинация, при которой $E[S]$ и $Var[S]$ принимают максимальные значения. Для того, чтобы это увидеть, предположим, что μ и σ обозначают среднее и стандартное отклонение величины суммарного иска и расходов по каждому полису соответственно. Следовательно:

$$E[S] = n\mu \qquad \text{и} \qquad Var[S] = n\sigma^2$$

Аналогично следуя всем шагам из пункта (i), получим выражение для n:

$$\frac{(80-\mu)\sqrt{n}}{\sigma} \geq 2.326$$

Таким образом:

$$n \geq [2.326\sigma/(80-\mu)]^2$$

Следовательно, наибольшее значение n получается при максимальных значениях μ и σ (при условии, что максимальное значение μ меньше 80). Теперь заметим, что

$$\mu = 0.4E[X_i + Y_i] \qquad \text{и} \qquad \sigma^2 = 0.4E[(X_i + Y_i)^2]$$

Из формул (данных выше) для моментов X_i и Y_i, μ и σ принимают наибольшие значения, когда α и b принимают максимально возможные значения, а λ — минимально возможное, т.е. когда

$$\alpha = 1.05; \qquad \lambda = 0.009; \qquad b = 110$$

Этот набор значений дает следующие результаты для μ и σ:

$$\mu = 78.67 \qquad \text{и} \qquad \sigma = 144.14$$

Подставляя эти значения в выражение для n, получим, что если n будет равно не меньше, чем 63,546, то страховая компания будет уверена, по крайней мере на 99%, что величина полученных премий превысит величину исков и расходов за год.

§7 Вопросы студентам

В1 Модели индивидуального и коллективного риска кажутся очень похожими. Какие основные отличия между ними?

О1 Обе модели в основном похожи. Главным отличием является то, что модель индивидуального риска рассматривает каждый полис, тогда как модель коллективного риска рассматривает отдельно каждый иск.

В2 Вы получили общие формулы для среднего и дисперсии суммарного иска в модели коллективного риска. Существует ли похожая формула для величины асимметрии?

О2 Да. Вот она:

$$skew[S] = E[N]skew[X] + 3Var[N]E[X]Var[X] + skew[N](E[X])^3$$

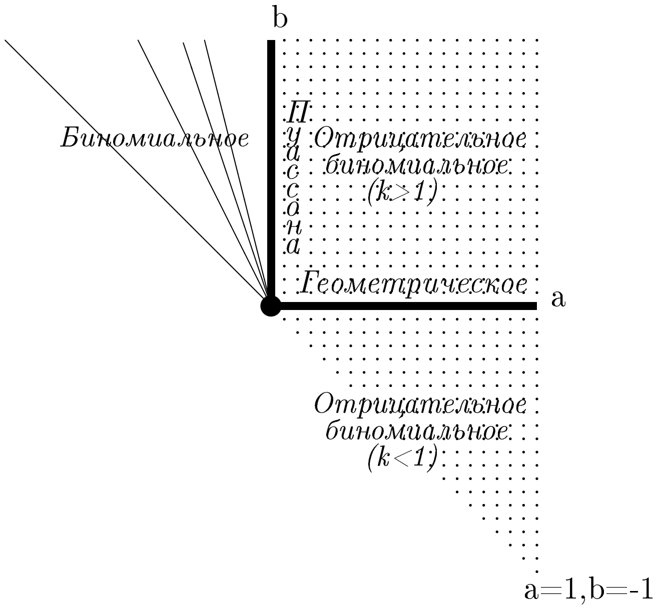
Но для экзамена вы не обязаны её знать.

В3 Есть ли ещё какие-нибудь распределения, которые можно вычислить с помощью итерационной формулы?

О3 Существует ещё несколько распределений, вероятностные функции которых могут быть выражены в итерационной форме. Однако, дискретные распределения, которые можно вычислить, используя итерационную формулу, приведенную здесь, должны быть распределением Пуассона, отрицательным биномиальным или биномиальным распределением. (Геометрическое распределение — это частный случай отрицательного биномиального распределения при $k = 1$.)

Диаграмма, приведенная ниже, показывает значения a и b, для которых распределения удовлетворяют итерационной формуле, вместе с соответствующими распределениями. Подходящие распределения обозначены черными линиями и заштрихованной областью. Начало координат соответствует "вырожденному распределению которое всегда принимает значение ноль. Незаштрихованные области соответствуют значениям a и b, которые не дают подходящую вероятностную функцию (например, когда формула дает отрицательное значение или когда сумма вероятностей не сходится). Сплошная линия с наклонными линиями соответствуют биномиальному распределению с $n = 1,\ n = 2, ...$

в зависимости от наклона. Чем больше n, тем ближе наклонные линии к вертикальной прямой, соответствующей распределению Пуассона.



В4 Применяются ли эти рекурсивные формулы на практике?

О4 Да, они применяются для вычисления вероятностей распределения суммарного иска на компьютерах. Однако, существует одна проблема, которая может возникнуть, когда размер индивидуального иска может принимать очень много различных значений. Рекурсивный тип формулы означает, что любые небольшие ошибки округления, которые возникают, имеют тенденцию к увеличению на каждом шаге вычисления. Это может привести к катастрофическому результату. Так что вы должны удостовериться, что арифметика, используемая вашим компьютером, достаточно точна (например, точность до 15 знаков), и вы должны проверять, что вычисляемые вероятности принимают корректные значения (в сумме дают единицу), и что среднее значение общего распределения принимает корректное значение.

В5 Как подобрать хи-квадрат вероятности для смещенного гамма-распределения, когда они не принимают значения,

близкие к 5%, 1% и т.д.?

O5 Если число степеней свободы достаточно велико, вы можете аппроксимировать хи-квадрат распределение нормальным распределением; т.е. $\chi_n^2 \sim N(n, 2n)$.

В6 Если мы должны использовать нормальную аппроксимацию в случае, описанном в предыдущем вопросе, почему мы сразу не применяем нормальное распределение?

О6 Подбор подходящего смещенного гамма-распределения в сущности растягивает и сжимает исходное распределение так, чтобы их асимметрия уменьшалась. Окончательное распределение является почти совсем симметричным, так что нормальная аппроксимация преобразованного распределения более точна.

В7 К какому распределению приближается смещенное гамма-распределение, когда асимметрия слишком мала?

О7 Если параметр k в смещенном гамма-распределении стремится к $-\infty$, при постоянных среднем и дисперсии исходного распределения, асимметрия стремится к нулю и распределение приближается к нормальному распределению.

В8 Действительно ли обобщенные распределения отличны от "исходных или они — это только сложные способ представить "исходные"распределения.

О8 Вообще говоря, обобщенное распределение — это распределение, отличающееся от любого стандартного распределения, т.е. не являются лишь сложным представлением простого распределения.

7.1 Подсказки для ответов на вопросы

1. Вопросы в этой теме главным образом алгебраические.

2. Отметим, что предполагается, что вы способны воспроизвести все доказательства из данной главы.

3. Значения величин, представленных в статистических экспериментах, часто представляют собой вещественные числа, которые имеют измерения, например, длины могут быть выражены в метрах. Как например в физических формулах, уравнения в статистике основываются на таких величинах, которые должны быть согласованы

с размерностью. Это правило иногда используется для проверки сложных выражений или для помощи в запоминании формул.

Например, если вы измеряете длину в единицах L (например, в метрах), то:

- все вероятности должны быть безразмерными, т.е. L^0
- среднее значение и среднеквадратичное отклонение распределения должны быть иметь ту же размерность, что и измеряемая величина, т.е. L
- дисперсия должна иметь размерность в квадрате, т.е. L^2

Пример 6.23 Если обобщенная случайная величина $S = X_1 + X_2 + ... + X_N$, то дисперсия равна

- (a) $Var[S] = E[N]Var[X] + Var[N](E[X])^2$ или
- (b) $Var[S] = (E[N])^2Var[X] + Var[N]E[X]$?

Решение N — это счетчик, поэтому должен быть безразмерным. Величины X и S выражены в денежных единицах, в £ скажем. Проверим размерность двух представленных альтернатив:

- (a) $£^2 = £^0£^2 + £^0£^2 \qquad \checkmark$
- (b) $£^2 = £^0£^2 + £^0£^1 \qquad \times$

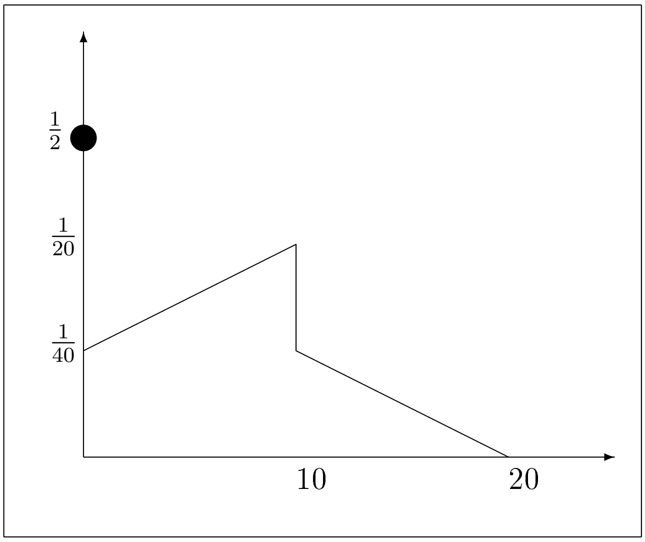
Таким образом, первая формула верная.

§8 Ответы на вопросы для самоподготовки

Решение 5.1

Если $N = 0$, то $S = 0$. Тогда имеет "точечную"вероятность того, что $S = 0$.
Если $N = 1$, то S имеет $U(0, 10)$–распределение. Это происходит с вероятностью $1/4$.
Если $N = 2$, то S — есть сумма двух независимых $U(0, 10)$–распределений. Это распределение имеет форму симметричного треугольника на интервале (0,20).
Объединяя все это, мы получим общее распределение, график которого

представлен ниже:



Решение 5.2

1. $E[t^S]$ — определение производящей функции вероятностей S.
2. Вместо S подставляем $X_1 + X_2 + ... + X_N$.
3. Следующий шаг доказательства получается, учитывая всевозможные значения N, т.е.:
$$\sum_{n=0}^{\infty} E[t^{X_1+X_2+...+X_N}|N=n]P(N=n)$$
что является суммой математических ожиданий при $N = n$, умноженных на вероятность $P(N = n)$.
4. Следующий шаг получается разложением на множители математических ожиданий, поскольку X_i не зависимы.
5. Далее, поскольку X_i одинаково распределены, получаем, что $E[t^{X_i}]$ – это производящая функция вероятностей случайной величины X, которая имеет то же распределение, что и X_i.
6. Следующее действие получается из определения $E[(G_X(t))^N]$, выписанного полностью:
$$\sum_{n=0}^{\infty}(G_X(t))^n P(N=n)$$

7. $E[(G_X(t))^N]$ — это определение производящей функции вероятностей величины N, но вместо t подставили $G_X(t)$. Следовательно, получили $G_N(G_X(t))$.

Решение 5.3

Используя вторую формулу:

$$M_S(t) = G_N(M_X(t)) = \exp(\lambda(M_X(t) - 1)) = \exp\left(\lambda\left[\left(\frac{\beta}{\beta - t}\right)^{\alpha} - 1\right]\right)$$

Решение 5.4

Мы имеем:

$$G_N(t) = \left(\frac{p}{1 - (1-p)t}\right)^k \quad \text{и} \quad G_X(t) = (1 - q + qt)^m$$

Объединяя это, получим:

$$G_S(t) = G_N(G_X(t)) = \left(\frac{p}{1 - (1-p)(1 - q + qt)^m}\right)^k$$

Будьте внимательны. Различайте правильно p и q.

Решение 5.5

При условии всех возможных значений N, получаем:

$$Var[S] = E[Var[S|N]] + Var[E[S|N]] =$$

$$= E[Var[X_1 + X_2 + ...X_N|N]] + Var[E[X_1 + X_2 + ...X_N|N]] =$$

$$= E[NVar[X]] + Var[NE[X]] =$$

$$= Var[X]E[N] + (E[X])^2 Var[N]$$

Последнее равенство справедливо в силу того, что $Var[X]$ и $E[X]$ — это константы в выражениях $E[NVar[X]]$ и $Var[NE[X]]$.

Решение 5.6

Используя формулу из таблицы для производящей функции моментов обобщенного геометрического распределения, получим:

$$\frac{p}{1 - qM_X(t)} = \frac{p}{1 - q\beta/(\beta - t)}$$

Решение 5.7

Используем общие формулы:

$$E[S] = E[N]E[X] = mpm_1$$

и

$$Var[S] = E[N]Var[X] + Var[N](E[X])^2 =$$

$$= mp(m_2 - m_1^2) + mpqm_1^2 = mpm_2 - mp^2m_1^2$$

Решение 5.8

Производящая функция моментов обобщенного распределения Пуассона:

$$M_S(t) = e^{\lambda[M_X(t)-1]}$$

Тогда производящая функция семиинвариантов:

$$K_S(t) = \log M_S(t) = \lambda[M_X(t) - 1]$$

Дифференцируя по t, получим:

$$K'_S(t) = \lambda M'_X(t)$$

Дифференцируем ещё раз по t:

$$K''_S(t) = \lambda M''_X(t)$$

И ещё раз дифференцируем по t:

$$K'''_S(t) = \lambda M'''_X(t)$$

Положим $t = 0$:

$$skew[S] = K'''_S(0) = \lambda m_3$$

Решение 5.9

Производящая функция семиинвариантов:

$$K(t) = \log(q + pM(t))^m = m\log(q + pM(t))$$

Дифференцируем это выражение по t:

$$K'(t) = mpM'(t)[q + pM(t)]^{-1}$$

Дифференцируем снова:

$$K''(t) = mpM''(t)[q + pM(t)]^{-1} - mp^2[M'(t)]^2[q + pM(t)]^{-2}$$

И снова:

$$K'''(t) = mpM'''(t)[q + pM(t)]^{-1} - mp^2M'(t)M''(t)[q + pM(t)]^{-2}+$$
$$+2mp^3[M'(t)]^3[q + pM(t)]^{-3} - 2mp^2M'(t)M''(t)[q + pM(t)]^{-2}$$

Положим $t = 0$:

$$K'''(0) = mpm_3 - 3mp^2m_1m_2 + 2mp^3m_1^3$$

Следовательно, коэффициент асимметрии равен:

$$\frac{mpm_3 - 3mp^2m_1m_2 + 2mp^3m_1^3}{[mpm_2 - mp^2m_1^2]^{3/2}}$$

Решение 5.10

Да, при условии, что параметр p в распределении числа исков одинаковый для двух распределений, и что распределения величин индивидуальных исков одинаковы. Если эти условия выполнены, то обобщенное биномиальное распределение обладает свойством аддитивности, т.е. если S_1 имеет обобщенное биномиальное распределение с параметрами m и p, и распределение величины иска имеет производящую функцию моментов $M_X(t)$, и S_2 имеет обобщенное биномиальное распределение с параметрами n и p, и распределение величины иска имеет производящую функцию моментов $M_X(t)$, и S_1 и S_2 независимы, тогда $S_1 + S_2$ также имеет обобщенное биномиальное распределение с параметрами $m+n$ и p, и величина иска имеет распределение с производящей функцией моментов $M_X(t)$. (Вы можете доказать это, используя производящие функции моментов).

Решение 5.11

Пусть $S = S_1+S_2$. Тогда S имеет обобщенное распределение Пуассона с параметрами 300 и $F(x)$, где

$$F(x) = \frac{1}{3}F_1(x) + \frac{2}{3}F_2(x) = 1 - \frac{1}{3}\exp(-x/\alpha) - \frac{2}{3}\exp(-x/\beta)$$

$F(x)$ — пример смешанного экспоненциального распределения. Мы можем интерпретировать распределение S следующим образом. Число исков имеет распределение Пуассона с параметром 300. С вероятностью 1/3 величина индивидуального иска получается из экспоненциального распределения с параметром α, с вероятностью $2/3$ величина индивидуального иска получается из экспоненциального распределения с параметром β.

Решение 5.12

Итак, мы имеем 5 полисов ("индивидуальные риски"). Тогда $n = 5$.

Компании A, C и D не предъявляли исков. Следовательно: $X_1 = X_3 = X_4 = 0$
Компания B предъявила иск в размере £$2,500$. Следовательно: $X_2 = 2,500$
Компания E предъявила два иска общего размера £$20,000$. Следовательно: $X_5 = 20,000$

Решение 5.13

Вероятность того, что произойдет ровно 5 смертных случаев среди штатных сотрудников, подчиняется биномиальному закону:

$$\binom{500}{5} * 0.005^5 * (1 - 0.005)^{500-5} =$$

$$= \frac{500 * 499 * 498 * 497 * 496}{5 * 4 * 3 * 2 * 1}(0.005)^5(0.995)^{495} =$$
$$= 0.0667$$

и вероятность того, что произойдет ровно 20 смертей среди числа рабочих, есть:

$$\binom{2,500}{20} * 0.010^{20} * (1 - 0.010)^{2,500-20} =$$
$$= \frac{2,500 * 2,499 * ... * 2,481}{20!}(0.010)^{20}(0.990)^{2,480}$$

Таким образом, учитывая, что случаи смерти независимы, общая вероятность равна 0.0035, т.е. 0.35%.

Решение 5.14

X_i имеет следующую функцию вероятностей:

$$X_i = \begin{cases} b & \text{с вероятностью} \quad q \\ 0 & \text{с вероятностью} \quad p \end{cases}$$

Таким образом:

$$E[X_i] = q * b + p * 0 = qb$$

А также:

$$E[X_i^2] = q * b^2 + p * 0^2 = qb^2$$

Тогда:

$$Var[X_i] = E[X_i^2] - (E[X_i])^2 = b^2 q - (bq)^2 = b^2 q(1-q)$$

Решение 5.15

Аналогично:

$$E[X_i^3] = q * b^3 + p * 0^3 = qb^3$$

Таким образом, получаем:

$$\begin{aligned} skew[X_i] &= E[X_i^3] - 3E[X_i^2]E[X_i] + 2(E[X_i])^3 = \\ &= b^3 q - 3(b^2 q)(bq) + 2(bq)^3 = \\ &= b^3 q(1-q)(1-2q) \end{aligned}$$

(В качестве альтернативы вы можете вычислить $E[(X_i - \mu)^3]$ непосредственно).

Решение 5.16

Мы имеем:

$$\mu = \frac{5}{0.002} = 2,500$$

$$\sigma^2 = \frac{5}{0.002^2} = 1,250,000$$

Тогда среднее значение и дисперсия величины суммарного иска будут равны:

$$E[S] = 1000 * 2500 * 0.004 = 10,000$$

$$Var[S] = 1000 * [1,250,000 * 0.004 + 2,500^2 * 0.004 * 0.996] = (£5,468)^2$$

Решение 5.17

Мы рассматриваем три иска. Тогда $N = 3$.

Первый иск (предъявленный компанией E в марте) был в размере £8,000. Следовательно: $X_1 = 8,000$
Второй иск (предъявленный компанией B в сентябре) был в размере £2,500. Следовательно: $X_2 = 2,500$
Третий иск (предъявленный компанией E в ноябре) был в размере £12,000. Следовательно: $X_3 = 12,000$

Решение 5.18

Согласно такому договору, если брутто-величина индивидуального иска равна X, тогда нетто-выплаты прямого страховщика равны $D = kX$. Следовательно, производящая функция моментов будет равна:

$$M_D(t) = E[e^{tD}] = E[e^{tkX}] = E[e^{(kt)X}] = M_X(kt)$$

Таким образом, производящая функция моментов величины суммарного иска равна:

$$M_{Snet}(t) = M_N[\log M_D(t)] = e^{\lambda[e^{\log M_D(t)}-1]} = e^{\lambda[M_X(kt)-1]}$$

Решение 5.19

Пусть S обозначает величину суммарного иска до перестрахования. Тогда:

$$S = X_1 + X_2 + ... + X_N$$

где X имеет распределение $Pareto(\alpha, \lambda)$ и $\alpha = 3,\ \lambda = 1000$. Таким образом:

$$E[X] = \frac{\lambda}{\alpha - 1} = 500 \quad \text{и} \quad Var[X] = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)} = 750,000$$

Тогда, если параметр Пуассона (значение которого точно не известно) обозначить за μ, то получим:

$$E[S] = 500\mu \quad \text{и} \quad Var[S] = \mu E[X^2] = \mu(750,000 + 500^2) = 1,000,000\mu$$

(i) Ожидаемая прибыль страховщика без перестрахования равна разности полученных премий и ожидаемых исков. Но если используется коэффициент нагрузки на премии, равный 0.2, то суммарная премия будет равна $1.2 * 500\mu$, и ожидаемая прибыль страховщика равна $600\mu - 500\mu = 100\mu$.

Теперь рассмотрим результат применения перестрахования. Для каждой случайной величины X_i имеем $X_i = Y_i + Z_i$, где

$$\begin{array}{lll} Y_i = X_i & \text{если } X_i < 1,000 & \text{(распределение} \\ Y_i = 1,000 & \text{если } X_i \geq 1,000 & \text{исков страховщика)} \end{array}$$

и

$$\begin{array}{lll} Z_i = 0 & \text{если } X_i < 1,000 & \text{(распределение} \\ Z_i = X_i - 1,000 & \text{если } X_i \geq 1,000 & \text{исков перестраховщика)} \end{array}$$

Таким образом, для перестраховщика общий суммарный иск представляет следующую случайную величину: $S_R = Z_1 + Z_2 + ... + Z_N$ с обобщенным распределением Пуассона, в котором каждая величина Z_i имеет распределение, данное выше.

Премии для перестраховщика равны $1.3E[S_R]$, где

$$E[S_R] = E[Z]E[N] = \mu E[Z]$$

$$E[Z] = \int\limits_{1,000}^{\infty} (x - 1,000) \frac{3 * 1,000^3}{(1,000 + x)^4} dx$$

Положим $u = x - 1,000$ в этом интеграле:

$$E[Z] = \int\limits_{0}^{\infty} u \frac{3 * 1,000^3}{(2,000 + u)^4} du = \left(\frac{1,000}{2,000}\right)^3 \int\limits_{0}^{\infty} u \frac{3 * 2,000^3}{(2,000 + u)^4} du$$

Последний интеграл — есть математическое ожидание $Pareto(3, 2, 000)$ распределения. Следовательно, можем записать:

$$E[Z] = \left(\frac{1}{2}\right)^3 * \frac{2000}{3 - 1} = 125$$

Тогда:

$$E[S_R] = 125\mu$$

Следовательно, премии перестраховщика равны 162.5μ. Итак, ожидаемая прибыль страховщика равна $600\mu - 162.5\mu - E[S - S_R]$, т.к. $S - S_R$ — это величина выплат страховщика по искам в результате перестрахования. Но

$$E[S - S_R] = E[S] - E[S_R] = 375\mu$$

Окончательно получим, что ожидаемая прибыль страховщика равна 62.5μ, и уменьшение (в процентном соотношении) ожидаемой прибыли без применения перестрахования (которая равна 100μ) равно 37.5%.

(ii) Теперь рассмотрим дисперсию прибыли, сначала без перестрахования. Прибыль страховщика равна разности полученных премий с учетом нагрузки и выплат по искам, так что дисперсия прибыли (до перестрахования) равна:

$$Var[S] = 1,000,000\mu$$

Тогда среднеквадратичное отклонение величины прибыли равно $1,000\sqrt{\mu}$.

С учетом перестрахования, прибыль страховщика равна разности нагруженных премий и нетто-выплат по искам. Тогда, если величина суммарных нетто-выплат по искам равна S_I, то дисперсия прибыли равна дисперсии S_I (т.к. другие выражения — константы). Значит:

$$S_I = Y_1 + Y_2 + ... + Y_N$$

имеет ещё одно распределение Пуассона.

Дисперсия S_I равна $Var[S_I] = \mu m_2$, где $m_2 = E[Y^2]$ и

$$E[Y^2] = \int\limits_0^{1,000} x^2 \frac{3 * 1,000^3}{(1,000 + x)^4} dx + \int\limits_{1,000}^{\infty} 1,000^2 \frac{3 * 1,000^3}{(1,000 + x)^4} dx =$$

$$= 3 * 1,000^3 \int\limits_0^{1,000} \frac{x^2}{(1,000 + x)^4} dx + 3 * 1,000^5 \int\limits_{1,000}^{1,000} \frac{1}{(1,000 + x)^4} dx$$

Второй интеграл равен:

$$\left.\frac{(1,000 + x)^{-3}}{-3}\right|_{1,000}^{\infty} = \frac{1}{3 * 2,000^3}$$

Для первого интеграла положим u=1,000+x и получим:

$$\int\limits_{1,000}^{2,000} \frac{(u - 1,000)^2}{u^4} du = -\frac{1}{u} + \frac{1,000}{u^2} - \left.\frac{1,000,000}{3u^3}\right|_{1,000}^{2,000} = \frac{1}{24,000}$$

Отсюда получаем $E[Y^2]$:

$$E[Y^2] = \frac{3*1,000^3}{24,000} + \frac{3*1,000^5}{3*2,000^3} = 250,000$$

Окончательный результат: $Var[S_I] = 250,000\mu$, среднеквадратичное отклонение равно $500\sqrt{\mu}$. Таким образом, уменьшение (в процентном соотношении) среднеквадратичного отклонения равно 50%.

Решение 5.20

Для того, чтобы найти вероятности выигрышей за один год, воспользуемся решением предыдущего примера. Единственное отличие состоит в том, что "параметр выигрыша"увеличится в 12 раз, так что: $\lambda = 12,000/15,000$.
Следовательно:

$$p_S(0) = e^{-\lambda} = e^{-12,000/15,000} = 0.4493$$

$$p_S(50) = \lambda p_X(50)p_S(0) = (12,000/15,000)*(15/16)*0.4493 = 0.3370$$

$$p_S(100) = \frac{\lambda}{2}[p_X(50)p_S(50) + 2p_X(100)p_S(0)] = 0.1488$$

$$p_S(150) = \frac{\lambda}{3}[p_X(50)p_S(100) + 2p_X(100)p_S(50) + 0] = 0.0484$$

Тогда: $P(S \geq 200) = 1 - 0.4493 - 0.3370 - 0.1488 - 0.0484 = 0.0165$.

Итак, вероятности равны:

(a) 44.9% (b) 33.7% (c) 14.9% (d) 4.8% (e) 1.7%

Решение 5.21

Для распределения Пуассона:

$$\frac{p_N(n)}{p_N(n-1)} = \frac{\lambda}{n}$$

Таким образом, $a = 0, \quad b = \lambda$.

Для геометрического распределения (которое является частным случаем биномиального распределения, когда $k = 1$):

$$\frac{p_N(n)}{p_N(n-1)} = pq^n / pq^{n-1} = q$$

Таким образом, $a = q, \quad b = 0$.

Для биномиального распределения:

$$\frac{p_N(n)}{p_N(n-1)} = \frac{\binom{m}{n} p^n q^{m-n}}{\binom{m}{n-1} p^{n-1} q^{m-n+1}} =$$

$$= \frac{m!}{n!(m-n)!} \frac{(n-1)!(m-n+1)!}{m!} \frac{p}{q} = \left(\frac{m-n+1}{n}\right) \frac{p}{q} =$$

$$= \left(-1 + \frac{m+1}{n}\right) \frac{p}{q}$$

Таким образом, $a = -\frac{p}{q}, \quad b = \frac{(m+1)p}{q}$.

Решение 5.22

Отношение вероятностных функций на двух соседних значениях (для $n = 1, 2, ...$) равно:

$$\frac{p_N(n)}{p_N(n-1)} = \frac{1}{2^{3n+1/2}} \frac{(2n)!}{n!^2} \Bigg/ \frac{1}{2^{3n-2\frac{1}{2}}} \frac{(2n-2)!}{(n-1)!^2} = \frac{2^{3n-2\frac{1}{2}}}{2^{3n+1/2}} \frac{(2n)!}{(2n-2)!} \frac{(n-1)!^2}{n!^2} =$$

$$= \frac{1}{8} * 2n(2n-1) * \frac{1}{n^2} = \frac{1}{2} - \frac{1}{4n}$$

Получили выражение вида $a + \frac{b}{n}$, где $a = 1/2, \quad b = -1/4$.

Сравнивая эти значения со значениями из таблицы, мы видим, что данное распределение соответствует отрицательному биномиальному распределению (единственное распределение, которое имеет положительное значение a и ненулевое b) с параметрами $p = q = 1/2, \quad k = 1/2$.

(Чтобы проверить, действительно ли это отрицательное биномиальное распределение, мы должны проверить значение $p_N(0)$ и сравнить его с нормирующей постоянной.)

Решение 5.23

Для распределения Пуассона $a = 0, \quad b = \lambda$.
Следовательно:

$$E[N] = \frac{a+b}{1-a} = \frac{0+\lambda}{1-0} = \lambda \quad \checkmark$$

Для отрицательного биномиального распределения $a = q$, $b = (k-1)q$.
Следовательно:

$$E[N] = \frac{a+b}{1-a} = \frac{q+(k-1)q}{1-q} = \frac{kq}{p} \quad \checkmark$$

Для геометрического распределения $a = q, \quad b = 0$.
Следовательно:

$$E[N] = \frac{a+b}{1-a} = \frac{q+0}{1-q} = \frac{q}{p} \quad \checkmark$$

Для биномиального распределения $a = -\frac{p}{q}, \quad b = \frac{(m+1)p}{q}$.
Следовательно:

$$E[N] = \frac{a+b}{1-a} = \left(-\frac{p}{q} + \frac{(m+1)p}{q}\right) \Bigg/ \left(1+\frac{p}{q}\right) = \frac{mp}{q} \Bigg/ \left(\frac{q+p}{q}\right) = mp \quad \checkmark$$

Решение 5.24

Начнем с того, что

$$G'_N(t) = \left(\frac{a+b}{1-at}\right) G_N(t)$$

Дифференцируя по t, получим:

$$G''_N(t) = \frac{a(a+b)}{(1-at)^2} G_N(t) + \left(\frac{a+b}{1-at}\right) G'_N(t)$$

Положим $t = 1$:

$$G''_N(1) = \frac{a(a+b)}{(1-a)^2} + \left(\frac{a+b}{1-a}\right)^2$$

Используя формулу для выражения дисперсии через производящую функцию вероятностей, упрощая, получим:

$$Var[N] = G''_N(1) + G'_N(1) - [G'_N(1)]^2 =$$

$$= \frac{a(a+b)}{(1-a)^2} + \left(\frac{a+b}{1-a}\right)^2 + \left(\frac{a+b}{1-a}\right) - \left(\frac{a+b}{1-a}\right)^2 = \frac{a+b}{(1-a)^2}$$

Решение 5.25

Используя рекурсивную формулу, получим:

$$p_S(0) = p^k = 0.4^2 = 0.16$$

$$p_S(1) = q\left(1 + \frac{1}{1}\right) p_X(1)p_S(0) = 0.6 * 2 * 0.4 * 0.16 = 0.0768$$

$$p_S(2) = q\left(1 + \frac{1}{2}\right) p_X(1)p_S(1) + q\left(1 + \frac{2}{2}\right) p_X(2)p_S(0) =$$

$$= 0.6 * \frac{3}{2} * 0.4 * 0.0768 + 0.6 * 2 * 0.6 * 0.16 = 0.1428$$

Решение 5.26

Для распределения Парето с параметрами 4 и 3 получим $m_1 = 1$ и $m_2 = 3$.

(i) Если $\lambda = 10$, используя стандартные формулы, среднее и дисперсия S равны 10 и 30 соответственно. Тогда

$$P(S \leq x) \doteqdot P(N(10, 30) \leq x) = P\left(N(0,1) \leq \frac{x - 10}{\sqrt{30}}\right) = \Phi\left(\frac{x - 10}{5.477}\right)$$

(a) Из Таблицы получаем значение $\Phi(1.645) = 0.95$. Таким образом:

$$\frac{x - 10}{5.477} = 1.645$$

Получаем $x = 19.01$.

(b) Т.к. $\Phi(2.326) = 0.99$, аналогично получаем:

$$\frac{x - 10}{5.477} = 2.326$$

Находим $x = 22.74$.

(ii) Для $\lambda = 50$ используем тот же самый способ для вычисления значения x. Среднее и дисперсия S равны 50 и 150 соответственно.

(a)

$$\frac{x - 50}{\sqrt{150}} = 1.645 \quad \Rightarrow \quad x = 70.15$$

(b)

$$\frac{x - 50}{\sqrt{150}} = 2.326 \quad \Rightarrow \quad x = 78.49$$

Решение 5.27

k-ый момент $Gamma(\alpha, \lambda)$-распределения равен:

$$E[X^k] = \int_0^\infty x^k \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx =$$

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

$$= \frac{1}{\lambda^k} \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} P(0 < Gamma(\alpha+k, \lambda) < \infty) = \frac{1}{\lambda^k} \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}$$

Следовательно:

$$skew[X_i] = E[X_i^3] - 3E[X_i^2]E[X_i] + 2E[X_i]^3 =$$

$$= \frac{1}{\lambda^3} \frac{\Gamma(\alpha+3)}{\Gamma(\alpha)} - 3\frac{1}{\lambda^2} \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \frac{1}{\lambda} \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} + 2\left[\frac{1}{\lambda} \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}\right]^3 =$$

$$= \frac{(\alpha+2)(\alpha+1)\alpha}{\lambda^3} - 3\frac{(\alpha+1)\alpha}{\lambda^2} \frac{\alpha}{\lambda} + 2\left(\frac{\alpha}{\lambda}\right)^3 =$$

$$= \frac{\alpha^3 + 3\alpha^2 + 2\alpha - 3\alpha^3 - 3\alpha^2 + 2\alpha^3}{\lambda^3} = \frac{2\alpha}{\lambda^3}$$

Решение 5.28

Сначала вычислим m_3, третий центральный момент величины X_i. Он вычисляется по формуле:

$$E[X^3] = \int_0^\infty x^3 \frac{4 * 3^4}{(3+x)^5} dx$$

Рассмотрим этот интеграл в качестве производящей функции вероятностей обобщенного распределения Парето с параметрами $k = 4,\ \lambda = 3$ и $\alpha = 1$. А также учтем, что:

$$\int_0^\infty \frac{\Gamma(5)}{\Gamma(1)\Gamma(4)} \frac{3x^3}{(3+x)^5} dx = 1$$

Таким образом:

$$E[X^3] = \frac{\Gamma(1)\Gamma(4)}{\Gamma(5) * 3} * 4 * 3^4 = 27$$

(i) Если $\lambda = 10$, то уравнения для вычисления параметров смещенного гамма-распределения имеют вид:

$$\frac{2\alpha}{\lambda^3} = 270 \qquad \frac{\alpha}{\lambda^2} = 30 \qquad \frac{\alpha}{\lambda} + k = 10$$

Решая систему из трех уравнений, получим:

$$\alpha = 1.481 \qquad \lambda = 0.222 \qquad k = 3.333$$

Т.к. $2\alpha = 2.962 \doteqdot 3$, то $2\lambda(S - k)$ имеет приближенно χ^2_3–распределение. Тогда

$$P(S \le x) = P(2\lambda(S-k) \le 2\lambda(x-k)) \doteqdot$$

$$\doteqdot P(\chi^2_3 \le 0.444(x - 3.333))$$

(a) Из Таблицы получим $P(\chi^2_3 \le 7.815) = 0.95$, тогда $0.444(x - 3.333) = 7.815$. Из этого выражения находим значение $x = 20.93$.

(b) Аналогично:

$$P(\chi^2_3 < 11.34) = 0.99$$

Следовательно, $0.444(x - 3.333) = 11.34$ и $x = 28.87$.

(ii) Если $\lambda = 50$, уравнения для вычисления параметров смещенного гамма-распределения имеют вид:

$$\frac{2\alpha}{\lambda^3} = 1350 \qquad \frac{\alpha}{\lambda^2} = 150 \qquad \frac{\alpha}{\lambda} + k = 50$$

Решая систему из трех уравнений, получим:

$$\alpha = 7.407 \qquad \lambda = 0.222 \qquad k = 16.67$$

В этом случае $2\alpha = 14.814$, и значит, распределение $2\lambda(S-k)$ расположено между χ^2_{14}–распределением и χ^2_{15}–распределением. Из Таблицы получим значения:

$$P(\chi^2_{14} \le 23.68) = 0.95 \qquad P(\chi^2_{14} \le 29.14) = 0.99$$

$$P(\chi^2_{15} \le 25.00) = 0.95 \qquad P(\chi^2_{15} \le 30.31) = 0.99$$

Используя линейную интерполяцию:

$$P(2\lambda(S-k) \le 24.75) \doteqdot 0.95 \qquad P(2\lambda(S-k) \le 30.31) \doteqdot 0.99$$

т.к. $P(S \leq x) \doteqdot P(2\lambda(S-k) \leq 2\lambda(x-k))$, получим:

$$2\lambda(x-k) = 24.75 \quad \Rightarrow \quad x = 72.41$$

и

$$2\lambda(x-k) = 30.31 \quad \Rightarrow \quad x = 84.94$$

Заключение части III

Обобщенное распределение — есть сумма случайного числа независимых одинаково распределенных случайных величин. Конкретные примеры обобщенных распределений включают в себя обобщенное распределение Пуассона, обобщенные биномиальное, отрицательное биномиальное и геометрическое распределения. Для обобщенного распределения можно найти выражения для моментов производящей функции моментов.

Величина суммарного иска для страхового портфеля может быть описана, используя модель индивидуального риска или модель коллективного риска. Модель индивидуального риска рассматривает выплаты по каждому полису (риску). А модель коллективного риска рассматривает выплаты по каждому иску.

Если величина исков принимает дискретные значения, для нахождения вероятностной функции суммарного иска может быть использована рекурсивная формула.

Метод моментов используется для аппроксимации распределения величины суммарного иска нормальным распределением или смещенным гамма-распределением.

Часть IV