РАСЧЕТ ОБЩЕГО ЧИСЛА СТРАХОВЫХ ВЫПЛАТ И ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Всеволод Малиновский

Теория риска, иначе называемая страховой или актуарной математикой, имеет целью выработку рекомендаций, основанных на анализе статистической информации при проведении и регулировании рисковых предприятий.

Одна из традиционных задач актуария заключается в выработке удовлетворительной тарифной системы, что в значительной мере сводится к предсказанию величины страховых выплат, ожидаемых по определенному виду, или видам, страхования, проводящимся страховой компанией.

Прогноз величины страховых выплат базируются на опыте компании, который заключается в поступившей за предшествующий период ее деятельности статистике страховых случаев. Не говоря о том, что эта статистика должна быть достоверной и достаточно полной, что составляет предмет отдельного обсуждения, качество прогноза в решающей степени зависит от того, какая методика применяется актуарием для обработки этой статистики. Другими словами, выбор методики обработки статистической информации решающим образом влияет на качество его рекомендаций руководству страховой компании.

Теория вероятностей, как математическая наука, позволяющая по вероятности одних случайных событий находить вероятности других случайных событий, связанных каким–либо образом с первыми, является естественной и общепринятой основой актуарных прогнозов и рекомендаций.

Однако все результаты этой теории имеют четко очерченные границы применимости. Поэтому перед обращением к какому–либо теоретическому результату актуарий обязан сознательно и полно проверить возможность его использования в рассматриваемой им конкретной ситуации.

Более того, и это является основным замечанием настоящей статьи, в определенных случаях актуарной практики даже корректное использование таких результатов теории вероятностей, как центральная предельная теорема или закон больших чисел, приводит к удивительно неточным выводам. Сразу заметим, что это ни в коей мере не противоречит известной мысли академика А. Н. Колмогорова, говорившего, что познавательная ценность теории вероятностей раскрывается только предельными теоремами. Существо кажущегося противоречия теории с практикой здесь в следующем: выводы, предлагаемые предельными теоремами, непременно должны сопровождаться оценкой погрешности той аппроксимации, которую они предлагают.

Основная задача данной заметки состоит в том, чтобы наглядно продемонстрировать приведенное выше соображение. Для этого вниманию читателя предлагается пример расчета общего числа страховых выплат по модельному страховому портфелю.

Перед тем, как переходить к существу примера, заметим, что рассматриваемые ниже расчеты общего числа страховых выплат в так называемой индивидуальной модели теории риска (individual risk model[1]) составляли экзаменационную работу по курсу «Aggregate claims distributions», предложенную автором в осеннем семестре 1993 года студентам Лаборатории Актуарной Математики в Университете Копенгагена (Дания). Задачи такого сорта, предполагающие написание программы расчета на ЭВМ, составляют обязательную часть образования современного актуария.

Перейдем к формулировке задачи, носящей модельный, заведомо упрощенный характер. Итак, пусть рассматривается портфель по страхованию от однородного риска лиц 30, 50 и 70 лет, которые во всем прочем предполагаются одинаковыми. Далее, пусть вероятности наступления страхового случая внутри данной возрастной группы известны (в экзаменационной задаче предлагалось находить эти вероятности, пользуясь таблицами смертности в Дании, G82M). Наступления страхового случая предполагаются независимыми событиями.

В момент заключения контракта портфель состоит из N страховых договоров, период действия которых фиксирован (например, равен 1 году) и размер страховых выплат по которым составляет 1, 5 и 10 единиц (например, миллионов крон). Во всем прочем эти договоры не различаются. Распределение по возрастам и страховым выплатам внутри портфеля определяется следующей таблицей 1:

Таблица 1. Структура страхового портфеля объема N: распределение договоров по размерам страховых возмещений и по возрасту страхуемых

структура портфеля объема N			
	выплаты		
возраст	$\alpha = 1$	$\alpha = 5$	$\alpha = 10$
$a = 30$	$N/10$	$N/5$	$N/30$
$a = 50$	$N/10$	$N/5$	$N/30$
$a = 70$	$N/10$	$N/5$	$N/30$

Задача, поставленная перед студентами, состояла в определении вероятностей общего числа страховых выплат S_N по указанному портфелю на период окончания действия договоров. При этом полный ответ предполагал

А) вычисление вероятностей $P(S_N = k)$, $k = 0, 1, \ldots,$

Б) определение среднего, дисперсии, третьего центрального момента величины S_N,

В) 95% и 5%; 85% и 15%; 75% и 25% квантилей распределения S_N.

Вычисления должны были быть проведены с использованием:

[1]См. например, Bühlmann H. *Mathematical Methods in Risk Theory*. Springer, 1970, или Gerber H. *An Introduction to Mathematical Risk Theory*. Huebner Foundation, Monograph Ser., N 8, 1979.

а) точной рекурсивной процедуры вычисления распределения S_N, предложенной Де Прилом (De Pril)[2],
б) нормальной аппроксимации распределения S_N, основанной на использовании центральной предельной теоремы (ЦПТ),
в) нормальной аппроксимации в ЦПТ с уточнениями Эджворта (используя третьи моменты).

От студентов требовалось написание пакета вычислительных программ, обеспечивающего решение вопросов А), Б) и В) в автоматическом режиме и при разнообразных усложнениях модельного портфеля. Подчеркнем однако, что «сверхзадача» исследования состоит в том, чтобы сделать соответствующие выводы из сравнения результатов, полученных точными вычислениями пункта а), и аппроксимации пунктов б) и в), при различных значениях N.

Математическая модель, соответствующая изложенному выше, состоит в том, что случайная величина S_N представима в виде

$$S_N = \sum_{a\in\{30,50,70\}} \sum_{\alpha\in\{1,5,10\}} \sum_{i=1}^{f(\alpha)} X_i^{(a,\alpha)} \tag{1}$$

где a обозначает возраст, α обозначает сумму страхового возмещения, (a,α) обозначает тип договора (полностью определенного возрастом застрахованного и суммой страхового возмещения по договору) и $f(\alpha)$ обозначает число застрахованных по договору типа (a,α). Наконец, через $X_i^{(a,\alpha)}$ обозначаются случайные величины, составляющие число страховых выплат по i–ому полису типа (a,α). Эти случайные величины независимы, но одинаково распределены только при фиксированных a, α и различных $i = 1,\dots,f(\alpha)$: легко проверить, что распределение $X_i^{(a,\alpha)}$ имеет вид

$$\mathbf{P}(X_i^{(a,\alpha)} = x) = \begin{cases} q(a), & \text{если } x = \alpha, \\ p(a), & \text{если } x = 0, \\ 0, & \text{в противном случае,} \end{cases}$$

где $q(a)$ обозначает вероятность наступления страхового случая для лица возраста а в рассматриваемый (одногодичный) период, и $p(a)$ — вероятность его ненаступления, т.е. $p(a) + q(a) = 1$. Для заинтересованного читателя заметим, что, согласно таблицам G82M,

$$q(30) = 0,001593144,$$
$$q(50) = 0,006773987,$$
$$q(70) = 0,036068784.$$

Представление (1) случайной величины S_N в виде суммы независимых слагаемых сразу подсказывает мысль воспользоваться центральной предельной теоремой теории вероятностей[3], согласно которой для $k = 0,1,2,\dots$

$$\mathbf{P}(S_N = k) \approx (2\pi B_N)^{-1/2} \exp\Big\{ -\frac{(k-M_N)^2}{2B_N}\Big\}. \tag{2}$$

[2]De Pril N. (1989) The aggregate claims distribution in the individual model with arbitrary positive claims. – ASTIN Bull. **19**, p. 9–24.

[3]Феллер В. Введение в теорию вероятностей и ее приложения, т.2, Мир, М., 1967.

Условия этой теоремы, как легко проверить, выполнены. Здесь

$$
\begin{aligned}
M_N &= \sum_{a\in\{30,50,70\}} \sum_{\alpha\in\{1,5,10\}} \sum_{i=1}^{f(\alpha)} \mathbf{E}X_i^{(a,\alpha)} = \sum_{a\in\{30,50,70\}} \sum_{\alpha\in\{1,5,10\}} \sum_{i=1}^{f(\alpha)} f(\alpha)\alpha q(\alpha) \\
&= N \times \frac{43}{30} \sum_{a\in\{30,50,70\}} q(\alpha) \approx N \times 0,063691478.
\end{aligned}
$$

$$
\begin{aligned}
B_N &= \sum_{a\in\{30,50,70\}} \sum_{\alpha\in\{1,5,10\}} \sum_{i=1}^{f(\alpha)} \mathbf{D}X_i^{(a,\alpha)} \\
&= \sum_{a\in\{30,50,70\}} \sum_{\alpha\in\{1,5,10\}} \sum_{i=1}^{f(\alpha)} f(\alpha)\alpha^2 q(\alpha)(1-q(\alpha)) \\
&= N \times \frac{235}{30} \sum_{a\in\{30,50,70\}} q(\alpha)(1-q(\alpha)) \approx N \times 0,363363093.
\end{aligned}
$$

Теория утверждает, что при достаточно больших N приближение (2) имеет место для любых $k = 0, 1, \ldots$. Дальнейшее уточнение аппроксимации (2) может быть получено обращением к так называемым асимптотическим разложениям, использующим не только моменты первого и второго порядка M_N и B_N , но и моменты высших порядков.

Не вдаваясь здесь в математические детали, обратимся теперь к точным вычислениям распределения $\mathbf{P}(S_N = k)$. Одним из подходов здесь является рекуррентная процедура, предложенная в [2]. Перед прямыми вычислениями вероятностей $\mathbf{P}(S_N = k)$ по формуле свертки эта процедура имеет преимущество в скорости проведения вычислений и в объеме вспомогательной информации, запоминаемой на каждом шаге. Итак, обозначая для простоты $p(k) = \mathbf{P}(S_N = k)$, имеем

$$
P_S(0) = \prod_{a\in\{30,50,70\}} (p(a))^{n_a}, \tag{3}
$$

$$
sP_S(s) = \sum_{i=1}^{\min(s,10)} \sum_{k=1}^{[s/i]} a(i,k) P_S(s-ki), \quad s = 1, 2, \ldots, m, \tag{4}
$$

где

$$
m = \sum_{a\in\{30,50,70\}} \sum_{\alpha\in\{1,5,10\}} f(\alpha)\alpha = 4,3 \times N
$$

обозначает максимально возможные выплаты по портфелю,

$$
n_\alpha = \sum_{\alpha\in\{1,5,10\}} f(\alpha) = \frac{1}{3}
$$

и

$$
a(i,k) = (-1)^{k+1} i \sum_{j\in\{30,50,70\}} f(j) \Big(\frac{1-p(j)}{p(j)}\Big)^k
$$

Подчеркнем, что указанная процедура точных расчетов (3), (4) — это всего лишь одна из многих известных процедур численного вычисления вероятностей $\mathbf{P}(S_N = k)$.

Результаты расчетов $\mathbf{P}(S_N = k)$ приведены на графиках 1–5 для портфелей объема 60, 300, 600, 1200, 3000. Очевидно, что аппроксимация (2) распределения $\mathbf{P}(S_N = k)$ знаменитой кривой Гаусса, как и уточнения этой аппроксимации с помощью асимптотических разложений, практически не дают информации о существе дела в первых трех случаях. Приближения начинают удовлетворительно «работать» только при объемах портфеля N между 1200 и 3000.

Заметим также, что знание среднего M_N и дисперсии B_N величины S_N, нахождением которых некоторые актуарии часто ограничиваются, не дает удовлетворительной картины явления при «малых» и «средних» выборках: значения $\mathbf{P}(S_N = k)$ сильно флуктуируют, что вполне объяснимо после некоторого размышления о природе задачи. В самом деле, ясно, например, что событие $S_N = 4$, для реализации которого необходимо иметь 4 страховых случая с размером выплат в 1 единицу, менее вероятно, чем событие $S_N = 5$, которое, кроме реализации 5 страховых случаев с размером выплат в 1 единицу, может быть результатом одного страхового случая с размером выплаты в 5 единиц. Эти вероятности различаются тем больше, чем меньше N.

Для справедливости заметим, что отмеченные трудности особенно отчетливы при использовании локальных предельных теорем теории вероятностей. Для приближений вероятностей вида $\mathbf{P}(S_N < k)$, т.е. при использовании интегральных предельных теорем, описанные трудности могут быть менее отчетливыми (см. графики 6–10).

В заключение хочется надеяться, что приведенные расчеты предостерегут актуариев от излишнего доверия к фразам типа «из центральной предельной теоремы известно, что при достаточно большом числе N договоров страхования (не менее нескольких десятков) нормальная аппроксимация $\mathbf{P}(S_N = k)$ является точной».

Список литературы

[1] Феллер В. Введение в теорию вероятностей и ее приложения, т. 2, Мир, М., 1967.

[2] De Pril N. (1989) The aggregate claims distribution in the individual model with arbitrary positive claims. — ASTIN Bull. v. 19, p. 9–24.

[3] Buhlmann H. Mathematical Methods in Risk Theory. Springer, 1970.

[4] Gerber H. An Introduction to Mathematical Risk Theory. Huebner Foundation, Monograph Ser., № 8, 1979.