Спецкурс “Теория риска” (для 409 гр.)

Проф. Екатерина Вадимовна Булинская

(МГУ имени М.В.Ломоносова)

Лекция 5

Москва, 7 октября 2020 г.

С НОВЫМ УЧЕБНЫМ ГОДОМ!

План лекции

- Выпуклый порядок
- Достаточные условия стоп-лосс порядка
- Предположение о равенстве средних не является серьезным ограничением для $<_{sl}$
- Свойства инвариантности порядка стоп-лосс
- Сравнение биномиальной, пуассоновской и отрицательно биномиальной моделей
- Порядки более слабые, чем стоп-лосс

Стоп-лосс порядок

Пусть имеются два риска X и Y. Тогда X меньше Y в смысле стоп-лосс ($X <_{sl} Y$), если $\mathrm{E}(X-d)^+ \leq \mathrm{E}(Y-d)^+$ для любого $d \in R^+$. Иными словами, стоп-лосс премии упорядочены соответствующим образом для любого размера (вычитаемой) франшизы d.

В теории вероятностей стоп-лосс порядок обычно называют возрастающим выпуклым порядком, поскольку $X <_{sl} Y$ эквивалентно тому, что $\mathrm{E}f(X) \leq \mathrm{E}f(Y)$ для любых неубывающих выпуклых функций f, для которых соответствующие математические ожидания существуют.

Таким образом, ущерб X меньше чем Y, поскольку все лица с неубывающей вогнутой функцией полезности, т.е. стремящиеся к прибыли, но не склонные к риску, предпочтут X, а не Y. Следовательно, порядок $<_{sl}$ отражает предпочтение всех "рациональных" перестраховщиков.

Выпуклый порядок

Обычно стоп-лосс порядок усиливается с помощью дополнительного требования о равенстве средних сравниваемых рисков X и Y ($\mathsf{E}X = \mathsf{E}Y$). Более того, в этом случае говорят, что X меньше Y в смысле выпуклого порядка ($X <_{cx} Y$ или иногда в актуарной литературе это обозначается $X <_{sl,=} Y$).

Термин "выпуклый" используется, поскольку $X <_{cx} Y$ эквивалентно тому, что $\mathsf{E}f(X) \le \mathsf{E}f(Y)$ для всех выпуклых функций f, для которых указанные математические ожидания существуют.

Лемма

Если X - это случайная величина, заданная на $(\Omega, \mathcal{A})$, с $\mathsf{E}|X| < \infty$, и $\mathcal{A}_0$ - σ-подалгебра ($\mathcal{A}_0 \subseteq \mathcal{A}$), то $\mathsf{E}[X|\mathcal{A}_0] <_{cx} X$.

Доказательство следует из неравенства Йенсена для условных математических ожиданий, утверждающего, что для любой выпуклой функции f, для которой соответствующие математические ожидания существуют, верно

$$\mathsf{E}f(X) = \mathsf{E}(\mathsf{E}[f(X)|\mathcal{A}_0]) \ge \mathsf{E}f(\mathsf{E}[X|\mathcal{A}_0]). \square$$

Теорема, полезная для перестрахования

Теорема

Пусть $X_1, X_2, \dots, X_n$ *- n независимых одинаково распределенных рисков и пусть* Φ *- измеримое отображение* $\Phi : R_n^+ \to R$. *Тогда существует симметрическая функция* $\Phi^* : R_n^+ \to R$ *такая, что*

$$\Phi^*(X_1, X_2, \dots, X_n) <_{cx} \Phi(X_1, X_2, \dots, X_n).$$

Функция Φ^* *имеет вид*

$$\Phi^*(x_1, x_2, \dots, x_n) = \frac{1}{n!} \sum_{1 \le i_1 \ne i_2 \ne \dots \ne i_n \le n} \Phi(x_{i_1}, x_{i_2}, \dots, x_{i_n}).$$

Доказательство.

Пусть $\sigma(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ - σ-подалгебра, порожденная вариационным рядом $(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$, связанным с рисками $X_1, X_2, \ldots, X_n$. В силу предыдущей леммы

$$\mathsf{E}[\Phi(X_1, X_2, \ldots, X_n)|\sigma(X_{(1)}, X_{(2)}, \ldots, X_{(n)})] <_{cx} \Phi(X_1, X_2, \ldots, X_n).$$

Тем самым доказательство закончено, поскольку

$$\mathsf{E}[\Phi(X_1, X_2, \ldots, X_n)|\sigma(X_{(1)}, X_{(2)}, \ldots, X_{(n)})]$$

$$= \frac{1}{n!} \sum_{1 \le i_1 \ne i_2 \ne \ldots \ne i_n \le n} \Phi(X_{(i_1)}, X_{(i_2)}, \ldots, X_{(i_n)})$$

$$= \frac{1}{n!} \sum_{1 \le i_1 \ne i_2 \ne \ldots \ne i_n \le n} \Phi(X_{i_1}, X_{i_2}, \ldots, X_{i_n}) = \Phi^*(X_1, X_2, \ldots, X_n).\square$$

Заметим, что для справедливости этой теоремы достаточно требовать вместо независимости рисков X_1, $X_2, \ldots, X_n$ их перестановочность. В перестраховании особенно часто используются функции $\Phi(x_1, x_2, \ldots, x_n)$ вида $\sum_{i=1}^n \alpha_i x_i$ и $\sum_{i=1}^n \phi_i(x_i)$.

Условие Карлина-Новикова

Критерий пересечений Карлина-Новикова (1963г.) - достаточное условие для предпочтительности в смысле порядка $<_{cx}$.

Теорема

Пусть $\mathrm{E}X_k = m$, $k = 1, 2$, *и существует такой интервал* $[x_1, x_2]$, $x_1 \leq x_2$, *где эти функции распределения совпадают, в то время как* $F_1(x) \leq F_2(x)$ *при* $x < x_1$ *и* $F_1(x) \geq F_2(x)$ *при* $x > x_2$, *тогда* $F_1 <_{sl} F_2$.

Доказательство. Поскольку $\mathrm{E}\max(X_k, x) = x + \int_x^\infty (1 - F_k(t))\,dt$, то при $x \geq x_1$ имеем $\mathrm{E}\max(X_1, x) \leq \mathrm{E}\max(X_2, x)$. Аналогично, при $x \leq x_2$ интегрирование по частям дает

$$\mathrm{E}\min(X_1, x) \geq \mathrm{E}\min(X_2, x).$$

Поэтому, используя равенство $\mathrm{E}\min(X_k, x) + \mathrm{E}\max(X_k, x) = m + x$, получим, что при всех x

$$\mathrm{E}\max(X_1, x) \leq \mathrm{E}\max(X_2, x),$$

что, как было доказано, означает $F_1 <_{sl} F_2$. $\square$

Достаточные условия

Целый ряд теорем о "пересечении" можно получить, используя свойства преобразования стоп-лосс, при этом следующий результат не требует предположения о равенстве средних.

Теорема

Пусть риски X и Y удовлетворяют условиям: $\mathsf{E}X \le \mathsf{E}Y$ *и существует* $c \ge 0$ *такое, что* $F_X(t) \le F_Y(t)$ *при* $t < c$ *и* $F_X(t) \ge F_Y(t)$ *при* $t \ge c$*, тогда* $X <_{sl} Y$*.*

Доказательство. Введем стоп-лосс преобразование

$$m_X(d) = \mathsf{E}(X-d)^+ = \int_d^\infty (1 - F_X(u))\,du$$

и положим $\Delta(d) = m_Y(d) - m_X(d) = \int_d^\infty (F_X(u) - F_Y(u))\,du$.
Пусть $d_1 < d_2 < c$, тогда $\Delta(d_1) - \Delta(d_2) = \int_{d_1}^{d_2} (F_X(u) - F_Y(u))\,du \le 0$, т.е. $\Delta(d)$ возрастает при $d < c$. Наоборот, при $d \ge c$ функция $\Delta(d)$ убывает. Так как $m_X(0) = \mathsf{E}X$, то $\Delta(0) \ge 0$, а $\Delta(\infty) = 0$. Значит, $m_X(d) \le m_Y(d)$ при любом d, т.е. $X <_{sl} Y$. $\Box$

Теорема

Если $EX = EY$ и существуют три непересекающиеся интервала I_0, I_1, I_2, такие, что $0 \in I_0$, I_2 бесконечен вправо и $[0, \infty) = I_0 \cup I_1 \cup I_2$, при этом $dF_X(x) \le dF_Y(x)$ при $x \in I_0 \cup I_2$, а при $x \in I_1$, наоборот, $dF_X(x) \ge dF_Y(x)$, то $X <_{sl} Y$.

Доказательство. Пусть $\delta(x) = F_X(x) - F_Y(x)$, тогда $\delta(0) = \delta(\infty) = 0$.

Далее, $\delta(x) = \int_0^x dF_X(t) - \int_0^x dF_Y(t) \le 0$ при $x \in I_0$ и убывает.

Аналогично,

$$\delta(x) = (1 - F_Y(x)) - (1 - F_X(x)) = \int_x^\infty dF_Y(u) - \int_x^\infty dF_X(u) \ge 0$$

и убывает при $x \in I_2$, в то время как при $x \in I_1$ функция $\delta(x)$ возрастает.

Таким образом, существует такое $c \ge 0$, что сначала $\delta(x) \le 0$ при $x < c$, затем $\delta(x) \ge 0$ при $x \ge c$. Иначе говоря, выполнены условия предыдущей теоремы, обеспечивающие $X <_{sl} Y$. $\square$

Покажем, что при равенстве средних и дисперсий из стоп-лосс порядка вытекает совпадение распределений.

Лемма

Если $X <_{sl} Y$ *и* $\mathrm{E}X = \mathrm{E}Y$, $\mathrm{D}X = \mathrm{D}Y$, *то* $X \stackrel{d}{=} Y$.

Доказательство. Из условий леммы $\mathrm{E}X^2 = \mathrm{E}Y^2$.

Как известно,

$\mathrm{E}X^2 = \int_0^\infty x^2 dF_X(x) = -\int_0^\infty x^2 d(1 - F_X(x)) =$
$2\int_0^\infty x(1 - F_X(x))\,dx = -2\int_0^\infty x\,dm_X(x) = 2\int_0^\infty m_X(x)\,dx.$

Значит, $\int_0^\infty (m_Y(x) - m_X(x))\,dx = 0$, а поскольку $X <_{sl} Y$, подинтегральная функция неотрицательна.

Таким образом, $m_Y(x) = m_X(x)$, а по функции $m(x)$ однозначно восстанавливается функция распределения. $\square$

Теорема

Пусть $X <_v Y$, *тогда существуют случайные величины* Z_1 *и* Z_2 *такие, что* $Y \stackrel{d}{=} X + Z_1 + Z_2$, $\mathrm{E}(Z_1|X) = 0$ *и* $\mathrm{P}(Z_2 \geq 0) = 1$.

Доказательство. В силу определения порядка $<_v$, если $X <_v Y$, тогда существует случайная величина Z такая, что $Y \stackrel{d}{=} X + Z$ и $\mathrm{E}(Z|X) \geq 0$. Обозначим $Y_1 = X + Z$, тогда $Y_1 \stackrel{d}{=} Y$ и $\mathrm{E}(Y_1|X) \geq X$. У.м.о. определены с точностью до множества меры нуль, поэтому можно считать, что $\mathrm{E}(Y_1|X = x) \geq x$ для всех x.

Введем новую случайную величину $Y' = Y_1 I$, где

$$\mathrm{P}(I = 1|X = x, Y_1 = y) = x/\mathrm{E}(Y_1|X = x) = 1 - P(I = 0|X = x, Y_1 = y).$$

Очевидно, что $Y' \leq Y_1$, тогда $Z_2 = Y_1 - Y' \geq 0$, а случайная величина $Z_1 = Y' - X$ такова, что $\mathrm{E}(Z_1|X) \geq 0$. Действительно,

$$\mathrm{E}(Y'|X = x, Y_1 = y) = xy/\mathrm{E}(Y_1|X = x),$$

а это значит, что $\mathrm{E}(Y'|X, Y_1) = XY_1/\mathrm{E}(Y_1|X)$, откуда следует требуемое соотношение

$$\mathrm{E}(Y'|X) = \frac{X\mathrm{E}(Y_1|X)}{\mathrm{E}(Y_1|X)} = X. \square$$

Замечание

В заключение заметим, что предположение о равенстве средних у двух рисков не является существенным ограничением. В самом деле, было установлено, что порядки $<_{sl}$ и $<_{v}$ эквивалентны.

А, как показывает только что доказанная теорема, если $X <_v Y$ и $\mathsf{E}X < \mathsf{E}Y$, то существует риск Y' такой, что $\mathsf{E}Y' = \mathsf{E}X$ и

$$X <_v Y' <_1 Y_1 \stackrel{d}{=} Y.$$

Свойства инвариантности стоп-лосс порядка

1. Слабая сходимость сохраняет стоп-лосс порядок.

Теорема

Пусть последовательность $F_{k,n}$, $k = 1, 2$, слабо сходится к F_k при $n \to \infty$, кроме того,

$$\lim_{n\to\infty} \int_0^\infty \bar{F}_{k,n}(t)dt = \int_0^\infty \bar{F}_k(t)dt < \infty, \; k = 1, 2.$$

Тогда, если $F_{1,n} <_{sl} F_{2,n}$ при всех n, то $F_1 <_{sl} F_2$.

Доказательство. Так как в предположениях теоремы при всех x выполнено

$$\lim_{n\to\infty} \int_x^\infty \bar{F}_{k,n}(t)\,dt = \int_x^\infty \bar{F}_k(t)\,dt,$$

то

$$\int_x^\infty \bar{F}_1(t)\,dt \le \int_x^\infty \bar{F}_2(t)\,dt,$$

т.е. $F_1 <_{sl} F_2$. $\square$

Лемма

Порядок стоп-лосс сохраняется при суммировании независимых случайных величин.

Доказательство. Порядок $<_{sl}$ порождается классом $\mathcal{F}_{sl} = \{e_x(t)\}_{x \in R}$, здесь $e_x(t) = \int_{-\infty}^{t} \Theta_x(u)\,du = (t-x)^+$. Поскольку это семейство функций инвариантно относительно сдвига, то согласно доказанной ранее теореме, отношение $<_{sl}$ обладает свойством (3°), т.е. сохраняется при свертке.

Таким образом, если $X_i <_{sl} Y_i$, $i = \overline{1, n}$, а $X_1 \ldots, X_n$ независимы и $Y_1, \ldots, Y_n$ независимы, то $\sum_{i=1}^{n} X_i <_{sl} \sum_{i=1}^{n} Y_i$. $\square$

3. Смеси распределений

Пусть функции распределения соответственно X_i и Y_i, $i \geq 1$, обозначены F_i и G_i, а неотрицательные числа p_i, $i \geq 1$, в сумме равны 1.

Лемма

Если $F_i <_{sl} G_i$, $i \geq 1$, то $\sum_i p_i F_i <_{sl} \sum_i p_i G_i$.

Доказательство. Обозначим $X = \sum_i I_i X_i$, где $\mathrm{P}(I_i = 1) = p_i = 1 - \mathrm{P}(I_i = 0)$, $\sum_i I_i = 1$ и случайные величины $\{I_i\}$ не зависят от $\{X_i\}$. Тогда F, функция распределения X, равна смеси с весами p_i распределений F_i. Аналогично вводится случайная величина Y, распределение которой является смесью G_i. Следовательно,

$$E(X-d)^+ = \int_d^\infty (x-d)dF(x) = \sum_i p_i \int_d^\infty (x-d)dF_i(x)$$

$$= \sum_i p_i \mathrm{E}(X_i - d)^+ \leq \sum_i p_i \mathrm{E}(Y_i - d)^+ = \mathrm{E}(Y-d)^+.$$

Таким образом, $X <_{sl} Y$. $\square$

Лемма

Пусть $\{X_i\}$ и $\{Y_i\}$, $i \geq 1$, - две последовательности независимых одинаково распределенных случайных величин, причем $X_i <_{sl} Y_i$, а целочисленная случайная величина N от них также не зависит. Тогда $\sum_{i=1}^{N} X_i <_{sl} \sum_{i=1}^{N} Y_i$.

Доказательство. Очевидно, что в силу формулы полной вероятности и условий независимости

$$\mathrm{P}(\sum_{i=1}^{N} X_i \leq x) = \sum_{n=0}^{\infty} \mathrm{P}(\sum_{i=0}^{n} X_i \leq x | N = n)\mathrm{P}(N = n)$$

$$= \sum_{n=0}^{\infty} F^{n*}(x) p_n, \qquad (1)$$

здесь $p_n = \mathrm{P}(N = n)$ и F - функция распределения X_i.
Распределение $\sum_{i=1}^{N} Y_i$ аналогичным образом выражается через функцию распределения G случайных величин Y_i. В силу двух предыдущих лемм требуемый результат получен. $\square$

Теорема (В)

Пусть риски X_i, $i \geq 1$, независимы и одинаково распределены, а целочисленные случайные величины N_j, $j = 1, 2$, от них не зависят и $N_1 <_{sl} N_2$, тогда $S_{N_1} <_{sl} S_{N_2}$, где $S_n = \sum_{i=1}^{n} X_i$.

Доказательство. Фиксируем $d \geq 0$ и рассмотрим функцию $f(n) = \mathrm{E}(S_n - d)^+$. Если проверить, что (при любом d) справедливо $\mathrm{E}f(N_1) \leq \mathrm{E}f(N_2)$, то утверждение теоремы будет доказано, т.е. достаточно установить, что $f(n)$ - неубывающая выпуклая функция n. Так как $X_i \geq 0$, то $S_{n-1} \leq S_n$ и, значит, $f(n-1) \leq f(n)$, т.е. функция неубывающая.

Для доказательства выпуклости проверим, что $f(n) - f(n-1) \geq f(n-1) - f(n-2)$. Пусть $x, y, z \geq 0$, тогда $(x + y + z - d)^+ + (x - d)^+ \geq (x + y - d)^+ + (x + z - d)^+$. Это легко проверяется, если расписать, чему равны левая и правая части приведенного неравенства при $0 < d < x$, $x < d < x + y$, $x + y < d < x + z$ и $x + z < d < x + y + z$, так как y и z входят в правую и левую части неравенства симметрично.

Возьмем $x = S_{n-2}$, $y = X_{n-1}$, $z = X_n$. Поскольку

$S_{n-2} + X_{n-1} + X_n = S_n$, $S_{n-2} + X_{n-1} = S_{n-1}$, а $S_{n-2} + X_n \stackrel{d}{=} S_{n-1}$, из предыдущего неравенства мы получим $f(n) + f(n-2) \geq 2f(n-1)$, а тем самым выпуклость функции $f(\cdot)$. $\square$

Следствие

Пусть $\{X_i\}$ *и* $\{Y_i\}$ *две последовательности независимых одинаково распределенных случайных величин,* $X_i <_{sl} Y_i$, $\forall i$ *, а целочисленные случайные величины* N_1 *и* N_2 *не зависят соответственно от* $\{X_i\}$ *и* $\{Y_i\}$*, кроме того* $N_1 <_{sl} N_2$*, тогда*

$$\sum_{i=1}^{N_1} X_i <_{sl} \sum_{i=1}^{N_2} Y_i.$$

Доказательство очевидным образом следует из цепочки неравенств

$$\sum_{i=1}^{N_1} X_i <_{sl} \sum_{i=1}^{N_1} Y_i <_{sl} \sum_{i=1}^{N_2} Y_i,$$

справедливых в силу доказанных леммы и теоремы.

Как известно, составным пуассоновским распределением называется распределение суммы S_N пуассоновского числа N независимых одинаково распределенных слагаемых X_i, $i \geq 1$, не зависящих от N.

Очевидно, что функция распределения S_N получается подстановкой $p_n = e^{-\lambda}\lambda^n/n!$ в формулу (1), где λ - параметр пуассоновского распределения.

Таким образом, только что полученное следствие показывает, когда можно упорядочить составные пуассоновские распределения в смысле порядка стоп-лосс.

Распределения Кокса

Предположим, что существует (структурная) случайная величина Λ такая, что при условии $\Lambda = \lambda$ распределение N является пуассоновским с параметром λ, тогда N имеет распределение Кокса (или распределение Пуассона, взвешенное по параметру).

Если $W(\lambda)$ - это функция распределения Λ, то

$$\mathrm{P}(N = n) = \int_0^\infty e^{-\lambda} \frac{\lambda^n}{n!} dW(\lambda). \tag{2}$$

Напомним, что производящая функция моментов пуассоновского распределения равна

$$g_N(t) = \mathrm{E}e^{tN} = e^{\lambda(e^t - 1)},$$

откуда легко получить, что $\mathrm{E}N = \mathrm{D}N = \lambda$.

Докажем теперь, что стоп-лосс порядок сохраняется также при суммировании случайного числа (имеющего распределение Кокса) случайных слагаемых.

Теорема

Пусть $Z_j = \sum_{i=1}^{N_j} X_i$, *где* X_i, $i \geq 1$, *независимые одинаково распределенные случайные величины и* N_j, $j = 1, 2$, *имеют структурные функции* Λ_j *такие, что* $\Lambda_1 <_{sl} \Lambda_2$, *тогда* $Z_1 <_{sl} Z_2$.

Доказательство. Согласно теореме В достаточно проверить, что $N_1 <_{sl} N_2$.

Для этого сначала подсчитаем $\mathrm{E}(N_j - d)^+$ при $k \leq d < k + 1$.

$$\mathrm{E}(N_j - d)^+ = \sum_{n=k+1}^{\infty} (n - k - d + k)\mathrm{P}(N_j = n)$$

$$= \mathrm{E}(N_j - k)^+ - (d - k)\mathrm{P}(N_j \geq k + 1),$$

т.е. как функция d это математическое ожидание убывает линейно на указанном промежутке. Поэтому далее мы сравниваем лишь $\mathrm{E}(N_j - k)^+$, $j = 1, 2$, $k \geq 0$.

Используя соотношение (2) и вид гамма-распределения, получим

$$\mathrm{E}(N_j - k)^+ = \sum_{n=k+1}^{\infty} (n-k)\mathrm{P}(N_j = n)$$

$$= \int_0^{\infty} \sum_{n=k+1}^{\infty} (n-k)\frac{\lambda^n}{n!}e^{-\lambda}dW_j(\lambda)$$

$$= \int_0^{\infty} \int_0^{\lambda} (\lambda - u)\frac{u^{k-1}}{(k-1)!}e^{-u}du dW_j(\lambda)$$

$$= \int_0^{\infty} \frac{u^{k-1}}{(k-1)!}e^{-u} \left[\int_u^{\infty} (\lambda - u)dW_j(\lambda)\right] du.$$

Поскольку внутренний интеграл равен $\mathrm{E}(\Lambda_j - u)^+$, то утверждение $N_1 <_{sl} N_2$ следует из условия $\Lambda_1 <_{sl} \Lambda_2$. $\square$

Следствие

Отрицательно биномиальное распределение рискованнее, чем пуассоновское (с тем же самым математическим ожиданием).

Доказательство. Воспользуемся предыдущей теоремой. Пусть N_1 имеет пуассоновское распределение с параметром λ. Очевидно, что это усреднение по структурной случайной величине Λ_1, равной λ с вероятностью 1. Было показано, что отрицательно биномиальное распределение получается усреднением по случайной величине Λ_2, имеющей гамма-распределение с плотностью

$$p_{(\alpha,q)}(\lambda) = \frac{\lambda^{\alpha-1}}{q^{\alpha}\Gamma(\alpha)} e^{-\lambda/q}.$$

Далее, $\mathsf{E}N_2 = \alpha q$, $\mathsf{D}N_2 = \alpha q(1+q)$. Если $\alpha q = \lambda$, то $N_1 <_{sl} N_2$ согласно указанной теореме. $\square$

Лемма

Разно распределенные испытания менее опасны в смысле порядка стоп-лосс, чем одинаково.

Доказательство. Положим $S_n = I_{A_1} + \ldots + I_{A_n}$, где I_{A_j} - независимые бернуллиевские случайные величины, $\mathrm{P}(I_{A_j} = 1) = q_j = 1 - \mathrm{P}(I_{A_j} = 0)$.

Иначе говоря, S_n - это число успехов в n независимых испытаниях с вероятностью успеха q_j в j-м испытании.

Обозначим $\bar{q} = n^{-1} \sum_{i=1}^{n} q_i$ и пусть случайная величина B_n будет числом успехов в n независимых испытаниях с вероятностью успеха $\bar{q}$ в каждом испытании.

Очевидно, $\mathrm{E}B_n = n\bar{q} = \sum_{i=1}^{n} q_i = \mathrm{E}S_n$.

Докажем по индукции, что $S_n <_{sl} B_n$. При $n = 1$ утверждение верно, так как $S_1 \stackrel{d}{=} B_1$.

Если все q_i равны, то они совпадают с $\bar{q}$, а $S_n \stackrel{d}{=} B_n$. Если не все q_i одинаковы, то существуют по крайней мере две вероятности, которые между собой различны. Предположим для определенности, что $q_1 < \bar{q} < q_2$ и что $S_{n-1} <_{sl} B_{n-1}$.

Рассмотрим случайные величины I_{C_1} и I_{C_2} независимые между собой и не зависящие от введенных ранее случайных величин. Пусть
$\mathrm{P}(I_{C_1} = 1) = \bar{q} = 1 - \mathrm{P}(I_{C_1} = 0)$, а
$\mathrm{P}(I_{C_2} = 1) = q_1 + q_2 - \bar{q} = 1 - \mathrm{P}(I_{C_2} = 0)$.

Нетрудно проверить (упражнение), используя критерий пересечений, что $I_{A_1} + I_{A_2} <_{sl} I_{C_1} + I_{C_2}$.
Значит, $S_n = I_{A_1} + \ldots + I_{A_n} <_{sl} I_{C_1} + I_{C_2} + I_{A_3} + \ldots + I_{A_n}$ (в силу сохранения порядка стоп-лосс при свертке).
Но $q_1 + q_2 - \bar{q} + q_3 + \ldots + q_n = n\bar{q} - \bar{q} = (n-1)\bar{q}$, следовательно, по предположению индукции $I_{C_2} + I_{A_3} + \ldots + I_{A_n} <_{sl} B_{n-1}$. В свою очередь, $I_{C_1} + B_{n-1} \stackrel{d}{=} B_n$, что заканчивает индукцию.

Теорема

Биномиальное распределение менее рискованное в смысле $<_{sl}$, чем пуассоновское (с тем же самым средним).

Доказательство. Пусть N_1 - это пуассоновская случайная величина с параметром λ.

Распределение B_n (биномиальное $Bi(n, \bar{q})$) можно рассмотреть как S_{n+1} с $q_i = \bar{q}$, $i = \overline{1, n}$, и $q_{n+1} = 0$. По доказанному оно менее рискованное, чем $Bi(n+1), \sum_{i=1}^{n+1} q_i/(n+1)) = Bi(n+1, n\bar{q}/(n+1))$. Если обозначить $n\bar{q} = \lambda$ и воспользоваться теоремой о сохранении порядка при слабой сходимости и теоремой Пуассона о сходимости (при $n \to \infty$) биномиального распределения $Bi(n, \lambda/n)$ к пуассоновскому с параметром λ, мы получим требуемый результат. $\square$

Сравнение моделей

Встает вопрос, каким образом перейти от индивидуальной модели риска к "соответствующей" коллективной модели. Иными словами, как выбрать распределение числа страховых случаев N и размера отдельного требования X, чтобы получившаяся модель давала адекватное описание рассматриваемого портфеля рисков (или была более рискованной). Наиболее распространенными являются "биномиальная", "пуассоновская" и "отрицательно биномиальная" модели.

Прежде всего заметим, что если все контракты "одинаковы", т.е. все V_i, $i = \overline{1, n}$, имеют одну и ту же функцию распределения F, то $N^* = \sum_{i=1}^{n} 1_i$ имеет биномиальное распределение с параметрами n и $q = \mathrm{P}(V_1 > 0) = \mathrm{P}(1_i = 1)$. Возьмем $N \stackrel{d}{=} N^*$ и пусть функция распределения всех X_i, $i \geq 1$, равна $G(x) = q^{-1}[F(x) - (1-q)]$, $x \geq 0$, тогда $S^{ind} \stackrel{d}{=} S^{col}$.

В общем случае, когда распределения V_i различны, мы также можем получить биномиальную модель. Для этого положим

$$q = n^{-1}\sum_{i=1}^{n} q_i, \text{ а } G(x) = (nq)^{-1}\sum_{i=1}^{n} q_i G_i(x), \qquad (3)$$

как и ранее, здесь обозначено

$$q_i = \mathrm{P}(V_i > 0),\ G_i(x) = q_i^{-1}[F_i(x) - (1 - q_i)],\ F_i(x) = \mathrm{P}(V_i \le x).$$

Отметим, что при таким выборе $G(x)$ мы будем иметь

$$\mathrm{E}X_1^k = \frac{1}{nq}\sum_{i=1}^{n} \mathrm{E}V_i^k,\ k \ge 1.$$

Далее, пусть N имеет биномиальное распределение $Bi(n, q)$ с параметрами n и q, тогда $\mathrm{E}N = nq$, следовательно, в силу д.з.1

$$\mathrm{E}S^{col} = \mathrm{E}N\mathrm{E}X_1 = nq(nq)^{-1}\sum_{i=1}^{n} \mathrm{E}V_i = \sum_{i=1}^{n} \mathrm{E}V_i = \mathrm{E}S^{ind}.$$

Таким образом чистые премии, подсчитанные для индивидуальной и коллективной моделей, совпадают. Отметим также, что $N^* <_{sl} N$ (N доминирует N^* в смысле порядка стоп-лосс), как следует из леммы.

Перейдем к построению пуассоновской модели коллективного риска. В этом случае S^{col} будет иметь сложное (или составное) пуассоновское распределение. Точнее, мы снова предполагаем, что X_i, $i \geq 1$, независимы и распределены одинаково с функцией распределения G, введенной выше. А число страховых случаев (или число поступлений требований) N имеет распределение Пуассона $Poi(\lambda)$ с параметром $\lambda = nq$. В этом случае $\mathsf{E}N = \mathsf{D}N = nq$, поэтому также $\mathsf{E}S^{col} = \mathsf{E}S^{ind}$.

Теперь рассмотрим третью модель - отрицательно биномиальную. Это означает, что распределение размера отдельного требования как и в прежних случаях G, а N имеет отрицательно биномиальное распределение $NB(n, q)$. Тогда снова $\mathsf{E}N = nq$, т.е. $\mathsf{E}S^{col} = \mathsf{E}S^{ind}$.

Сравнение дисперсий суммарного ущерба

Поскольку в биномиальной модели $\mathrm{D}N = nq(1-q)$, а

$$\mathrm{D}X_1 = \mathrm{E}X_1^2 - (\mathrm{E}X_1)^2 = (nq)^{-1}\sum_{i=1}^{n} EV_i^2 - (nq)^{-2}(\sum_{i=1}^{n}\mathrm{E}V_i)^2,$$

то (с помощью задачи из з.1) получим

$$\mathrm{D}S^{col} = \sum_{i=1}^{n}\mathrm{E}V_i^2 - n^{-1}(\sum_{i=1}^{n}\mathrm{E}V_i)^2.$$

С другой стороны, в силу независимости V_i

$$\mathrm{D}S^{ind} = \sum_{i=1}^{n}\mathrm{D}V_i = \sum_{i=1}^{n}\mathrm{E}V_i^2 - \sum_{i=1}^{n}(\mathrm{E}V_i)^2,$$

поэтому

$$\Delta_{Bi} = \mathrm{D}S^{col} - \mathrm{D}S^{ind} = \sum_{i=1}^{n}(\mathrm{E}V_i)^2 - n^{-1}(\sum_{i=1}^{n}\mathrm{E}V_i)^2 \geq 0.$$

В пуассоновском случае

$$\mathsf{D}S^{col} = \sum_{i=1}^{n} \mathsf{E}V_i^2,$$

значит,

$$\Delta_{Po} = \mathsf{D}S^{col} - \mathsf{D}S^{ind} = \sum_{i=1}^{n} (\mathsf{E}V_i)^2 > \Delta_{Bi}.$$

Наконец, в отрицательно биномиальном случае мы имеем $\mathsf{D}N = nq(1+q)$, следовательно,

$$\mathsf{D}S^{col} = \sum_{i=1}^{n} \mathsf{E}V_i^2 + n^{-1}(\sum_{i=1}^{n} \mathsf{E}V_i)^2.$$

Таким образом,

$$\Delta_{NB} = \mathsf{D}S^{col} - \mathsf{D}S^{ind} = \sum_{i=1}^{n} (\mathsf{E}V_i)^2 + n^{-1}(\sum_{i=1}^{n} \mathsf{E}V_i)^2 > \Delta_{Po}.$$

Стоп-лосс порядок моделей

Заметим, что качественный результат, т.е. возрастание дисперсии $\mathrm{D}S^{col}$ при переходе от биномиальной модели к пауссоновской и от пуассоновской к отрицательно биномиальной следовало ожидать в силу доказанных свойств инвариантности стоп-лосс порядка. В самом деле, $S^{col} = \sum_{i=1}^{N} X_i$, а по доказанной теореме порядок стоп-лосс сохраняется при суммировании случайного числа случайных слагаемых, т.е. $S^{col}_{Bi} <_{sl} S^{col}_{Po} <_{sl} S^{col}_{NB}$. При равенстве средних ($\mathrm{E}S^{col} = nq$ для любой модели) это влечет за собой возрастание дисперсий.

Итак, каждая из трех моделей коллективного риска правильно задает среднее, т.е. чистую премию, но переоценивает дисперсию. Как уже отмечалось ранее, выбор более рискованной модели приводит к принятию более "осторожного" решения, в частности, к большей страховой нагрузке, если премия-брутто подсчитывается на основе принципа дисперсии или среднего квадратичного (разумеется, при использовании принципа среднего размер премии-брутто останется тем же самым).

Другой переход к составному пуассоновскому распределению

Интересно, что к тому же самому составному пуассоновскому распределению, которое возникает при рассмотрении пуассоновской модели коллективного риска, можно придти на основе других рассуждений.

А именно, пусть $V_1, \ldots, V_n$ - это n независимых рисков с функциями распределения $F_1, \ldots, F_n$. Введем независимые случайные величины Y_i, $i = \overline{1, n}$, такие, что $Y_i \overset{d}{=} \sum_{k=1}^{N_i} V_k^{(i)}$, где N_i имеет распределение Пуассона с параметром 1, величины $N_1, \ldots, N_n$ независимы, а $V_k^{(i)} \overset{d}{=} V_i$, $k \geq 1$, т.е. имеют функцию распределения F_i, независимы друг от друга и от N_i.

Иначе говоря, каждый риск V_i заменяется на другой риск Y_i, имеющий составное пуассоновское распределение. Покажем, что при этой замене мы получим для суммы Y_i такое же распределение, как в случае пауссоновской модели коллективного риска.

Лемма

При сделанных предположениях $\sum_{i=1}^{n} Y_i \stackrel{d}{=} S_{Po}^{col}$.

Доказательство. Очевидно, что производящая функция моментов для Y_i имеет вид

$$g_{Y_i}(t) = \mathsf{E}e^{Y_i t} = \exp\{g_{V_i}(t) - 1\},$$

откуда следует, что соответствующая функция для $\sum_{i=1}^{n} Y_i$ равна

$$\prod_{i=1}^{n} g_{Y_i}(t) = \exp\{\sum_{i=1}^{n}(g_{V_i}(t) - 1)\} = \exp\{n[n^{-1}\sum_{i=1}^{n} g_{V_i}(t) - 1]\}.$$

Нетрудно понять, что это производящая функция сложного пуассоновского распределения с параметром $\lambda = n$ и функцией распределения отдельного слагаемого $\hat{G}(x) = n^{-1}\sum_{i=1}^{n} F_i(x)$, которую можно переписать в виде $\hat{G}(x) = qG(x) + (1-q)I(x)$. Здесь $G(x)$ определено в (3), а через $I(x)$ обозначена функция распределения случайной величины равной нулю.

Значит, мы имеем

$$n^{-1}\sum_{i=1}^{n} g_{V_i}(t) = g_{\hat{G}}(t) = \int_0^\infty e^{tx}\,d\hat{G}(x) = q\int_0^\infty e^{tx}dG(x) + 1 - q,$$

откуда следует

$$\exp\{n[n^{-1}\sum_{i=1}^{n} g_{V_i}(t) - 1]\} = \exp\{n[1 - q + q\int_0^\infty e^{tx}dG(x) - 1]\}$$

$$= \exp\{nq(g_G(t) - 1)\},$$

здесь и выше $g_G(t)$ (соотв. $g_{\hat{G}}(t)$) означает производящую функцию моментов случайной величины, имеющей функцию распределения $G(x)$ (соотв. $\hat{G}(x)$). $\square$

Отсюда сразу можно получить следующее соотношение между индивидуальной и коллективной (пуассоновской) моделями.

Следствие

В смысле порядка стоп-лосс индивидуальная модель менее рискованная, т.е. $S^{ind} <_{sl} S^{col}_{Po}$.

Доказательство. Поскольку $S^{col}_{Po} \stackrel{d}{=} \sum_{i=1}^{n} Y_i$, достаточно проверить, что $S^{ind} <_{sl} \sum_{i=1}^{n} Y_i$. Последнее соотношение очевидным образом вытекает из свойств порядка стоп-лосс. В самом деле, $1 <_{sl} N$, если N имеет распределение Пуассона с параметром 1. Значит, $V_i <_{sl} Y_i$, $i = \overline{1, n}$, откуда $S^{ind} = \sum_{i=1}^{n} V_i <_{sl} \sum_{i=1}^{n} Y_i$. $\square$

Стоп-лосс порядок степени n

Определение

$X <_{(n)} Y$, если $\mathsf{E}X^k \le \mathsf{E}Y^k$ для $k = \overline{1, n-1}$ и $\mathsf{E}[(X-d)^+]^n \le \mathsf{E}[(Y-d)^+]^n$ при любом $d \ge 0$.

Теорема

Порядок $<_{(n)}$ - это общее предпочтение всех лиц, имеющих функцию полезности $u \in \mathcal{U}_n$, что означает u дифференцируема n раз, первые $(n-1)$ производных непрерывны и меняют знак, т.е. $(-1)^{k-1}u^{(k)}(x) \ge 0$, $k = \overline{1, n-1}$, а также $(-1)^{n-1}u^{(n)}(x) \ge 0$ и не возрастает по x.

Доказательство. 1. Пусть $\mathsf{E}u(-X) \ge \mathsf{E}u(-Y)$ для любой функции $u \in \mathcal{U}_n$, тогда $\mathsf{E}f(X) \le \mathsf{E}f(Y)$ для $f(x) = -u(-x)$. Если взять $f(x) = [(x-d)^+]^n$, то $f^{(k)}(x) = n(n-1)\ldots(n-k+1)[(x-d)^+]^{n-k} \ge 0$, $k = \overline{1, n}$. Для $u(x) = -f(-x)$ имеем $u^{(k)}(x) = (-1)^{k-1}f^{(k)}(-x)$, т.е. выполнены все условия, наложенные на функцию полезности. Следовательно, $\mathsf{E}[(X-d)^+]^n \le \mathsf{E}[(Y-d)^+]^n$. Аналогичным образом проверяется, что $\mathsf{E}X^k \le \mathsf{E}Y^k$ при $k = \overline{1, n-1}$.

2. Обратно, пусть $X <_{(n)} Y$, а $u(\cdot)$ удовлетворяет требованиям к функции полезности. Берем $f(x) = -u(-x)$, тогда $f^{(k)}(x) \geq 0$ при $k = \overline{1, n}$, а $f^{(n)}(x)$ не убывает по x. Рассмотрим разложение Тейлора

$$f(x) = f(0) + f'(0)x + \ldots + f^{(n)}(0)x^n/n!$$

$$+ \int_{0+}^{x} [(x-t)^n/n!] df^{(n)}(t). \tag{4}$$

Последнее слагаемое можно переписать в виде

$$\int_{0+}^{\infty} \{[(x-t)^+]^n/n!\} df^{(n)}(t).$$

Пользуясь формулой (4), из $X <_{(n)} Y$ получаем $\mathsf{E}f(X) \leq \mathsf{E}f(y)$, следовательно, $\mathsf{E}u(-X) \geq \mathsf{E}u(-Y)$. $\square$

Теперь докажем, что с ростом n порядок стоп-лосс (степени n) становится слабее.

Теорема

Если $m > n$, то из $X <_{(n)} Y$ следует $X <_{(m)} Y$.

Доказательство. Достаточно проверить, что $\mathcal{U}_m \subset \mathcal{U}_n$. В самом деле, если $u \in \mathcal{U}_m$, то $(-1)^{k-1}u^{(k)}(x) \geq 0$ при $k = \overline{1,m}$. Поскольку $n < m$, осталось проверить, что $v(x) = (-1)^{n-1}u^{(n)}(x)$ не возрастает по x. Но $v'(x) = (-1)^{n-1}u^{(n+1)}(x) \leq 0$ по определению класса $\mathcal{U}_m$. $\square$

Обратное утверждение неверно, как показывает нижеследующий пример. Пусть $\mathsf{P}(X = i) = 1/4$, $i = \overline{0,3}$, и $\mathsf{P}(Y = 0) = 1/3$, $\mathsf{P}(Y = 2) = 1/2$, $\mathsf{P}(Y = 3) = 1/6$. Можно проверить, что справедлив стоп-лосс порядка 2, но не выполнен порядка 1.

Заметим также, что стохастический порядок $<_{st}$ - это порядок стоп-лосс степени 0, т.е. $<_{(0)}$, в то время как $<_{(1)}$ это то, что ранее обозначалось $<_{sl}$.