КОМПЬЮТЕРНАЯ МАТЕМАТИКА

УДК 519.8; 368; 65.0

Б.В. НОРКИН

ОБ ИДЕНТИФИКАЦИИ МОДЕЛЕЙ ДИНАМИЧЕСКОГО ФИНАНСОВОГО АНАЛИЗА СТРАХОВОЙ КОМПАНИИ*

Рассматриваются вопросы идентификации параметров процесса риска, моделирующего динамику капитала страховой компании. Трудность состоит в том, что моменты прихода отдельных требований и их величины являются внутренней информацией компании. Предложена регрессионная модель страховых выплат, связывающая выплаты с премиями и учитывающая задержки выплат. Проведена идентификация модели на реальных данных квартальной отчетности. Показано, что задержки выплат существенно влияют на вероятность разорения компании.

Введение. Ключевым вопросом практического использования математических моделей для целей управления сложными системами является идентификация моделей на реальных данных. В страховой математике такими моделями являются стохастические процессы риска, имитирующие стохастическую эволюцию капитала страховой компании [1, 2]. Например, в классической модели Крамера – Лундберга, с одной стороны, капитал компании монотонно возрастает с течением времени за счет непрерывно поступающих премий, а с другой – в случайные моменты времени (прихода страховых требований) он убывает на случайную величину (требования). Процесс останавливается, если капитал становится меньше нуля, т. е. компания становится неплатежеспособной. Обычно предполагается, что моменты прихода требований и величины требований являются независимыми случайными величинами с не зависящими от времени параметрами. Таким образом, предполагается, что поток требований является пуассоновским с некоторой интенсивностью, а распределение требований фиксировано. В смешанном пуассоновском потоке интенсивность является случайной. Задача состоит в идентификации

* Работа выполнена при поддержке гранта Президента Украины для молодых ученых GP/F49/121.

параметров интенсивности такого пуассоновского потока и параметров распределения величины требований.

Кроме классической модели мы рассматриваем модель процесса риска в дискретном времени. Это оправдано при больших масштабах бизнеса, когда ежедневно регистрируются приход многих премий и требований. Для этой модели предполагается, что выплаты в каждый период времени зависят от премий, полученных не только в данный период времени, но и в предыдущие периоды, со случайным коэффициентом пропорциональности.

Трудность идентификации параметров процесса риска состоит в том, что моменты прихода отдельных требований и их величины являются внутренней информацией компании, не доступны независимому стороннему наблюдателю; кроме того, страховые случаи и выплаты по ним могут быть значительно разнесены по времени и их сопоставление требует развитой системы электронного документооборота. Контролирующим органам, конкурентам, рейтинговым агентствам и общественности доступна лишь официальная финансовая отчетность, в которой фиксируются агрегированные величины, например, суммарный объем премий и требований за определенную единицу времени (квартал или год). Таким образом, независимую идентификацию параметров случайного процесса риска приходится проводить опосредовано через наблюдаемые величины.

Идентификация может быть проведена стандартными методами математической статистики, например, методом наименьших квадратов или наименьших модулей. Для этого необходимо вычислить средние значения наблюдаемых величин как функции параметров процесса риска. В настоящей работе процедура идентификации осуществлена на реальных данных процесса риска в дискретном времени. Новым элементом модели является регрессионная модель страховых выплат, связывающая выплаты с премиями, полученными в текущий и предшествующие периоды времени.

Классическая модель процесса риска. Классическая математическая модель стохастической эволюции резервов x^t страховой компании (с вычитанием постоянных дивидендов d) имеет вид [1– 4]:

$$x^t = u + (c-d)t - \sum\nolimits_{k=1}^{N_t} z_k, \quad 0 \le t \le T, \tag{1}$$

где t – временной параметр; $x^0 = u \ge 0$ – начальный капитал (страховой резерв); z_k – случайные требования с функцией распределения $F(\cdot)$ и средним значением μ; N_t – число поступивших к моменту t случайных требований (пуассоновский поток с интенсивностью λ); c – агрегированная страховая премия в единицу времени; d – отчисления от премий в единицу времени, не связанные с формированием резервов (текущие расходы, дивиденды и пр.). Здесь $S_t = \sum\nolimits_{k=1}^{N_t} z_k$ – агрегированные случайные страховые требования за время t, причем их среднее значение $\bar{S}_t = \bar{N}_t\mu = \mu\lambda t,$ где $\bar{N}_t$ – среднее количество страховых случаев за время $t.$ Классическая состоятельная оценка λ_t^* для интенсивности λ имеет вид [1, (3.136)]: $\lambda_t^* = N_t/t$. Для смешанного пуассоновского потока сама интенсивность прихода требований λ является случайной [1, разд. III, § 4], тогда идентификации подлежит не только среднее значение, но и все ее распределение.

Можно предположить, что интенсивность потока требований λ и интенсивность потока премий c связаны: чем больше поток премий, тем больше договоров страхования и тем чаще приходят требования. Обычно предполагается, что $c = (1+\rho)\mu\lambda$, где ρ – фиксированный коэффициент, называемый страховой нагрузкой (safety loading) [1, разд. III].

Идентификация функции распределения требований $F(\cdot)$ требует более детального анализа страховой статистики, например, вычисления второго и высших моментов распределения требований. Состоятельная эмпирическая оценка $F_t^*(z)$ функции распределения требований $F(z)$ имеет вид [1, (3.138)]:

$$F_t^*(z) = \{\text{число } z_k \text{ таких, что } k \le N_t \text{ и } z_k \le z\} / N_t.$$

При наличии богатой статистики по моментам и величинам страховых требований $\{\tau_k, z_k, k = 1,\dots,K\}$ классическая модель может быть целиком реализована на основе этого статистического материала без идентификации параметров. Во-первых, для данной страховой нагрузки ρ нужно вычислить премии $\tilde{c} = (1+\rho)\sum_{k=1}^{K} z_k \Big/ (\tau_K - \tau_1)$. Затем нужно последовательно случайно выбирать пары $(\tau_{\tilde{k}}, z_{\tilde{k}})$ из статистики и вычислять приращения процесса (1):

$$t_{s+1} = t_s + (\tau_{\tilde{k}+1} - \tau_{\tilde{k}}), \quad x^{t_{s+1}} = x^{t_s} + \tilde{c}(\tau_{\tilde{k}+1} - \tau_{\tilde{k}}) - z_{\tilde{k}}, \quad t_1 = 0, \quad x^{t_1} = u.$$

Если предположить, что распределение требований экспоненциально, то $F(z) = 1 - e^{-z/\mu}$. В таком случае модель полностью идентифицирована и вероятность разорения компании $\psi(u, c, d, \mu, \lambda)$ для классического процесса риска (на бесконечном интервале времени) как функция параметров $(u, \rho, d, \lambda, \mu)$ известна в явном виде [1, (3.17)]:

$$\psi(u,\rho,d,\mu,\lambda) = \begin{cases} \dfrac{1}{1+r}\exp\left(-\dfrac{ru}{(1+r)\mu}\right), & r > 0, \\ 1, & r \le 0, \end{cases} \tag{2}$$

где $r = (c-d)/(\lambda\mu) - 1 = \big((1+\rho)\lambda\mu - d\big)/(\lambda\mu) - 1$. От формулы [1, (3.17)] выражение для вероятности разорения (2) отличается тем, что из величины премии c вычитаются постоянные выплаты d (дивиденды и пр.). Введение параметра d важно для исследования и тестирования модели при наличии постоянных выплат. Для смешанного процесса риска со случайной интенсивностью λ прихода требований, имеющей распределение $Q(\lambda)$, вероятность разорения задается интегралом [1, (3.89)]: $\Psi(u,c,d,\mu)=\int_0^{\infty}\psi(u,c,d,\mu,\lambda)\,dQ(\lambda)$, где $\psi(\cdot)$ описана в (2). Вероятность разорения на конечном интервале времени как функция начального капитала u может быть получена путем решения специального интегрального уравнения страховой математики [5].

Модель процесса риска с дискретным временем. На практике финансовое состояние компании регистрируется в дискретные моменты времени, например, поквартально. В этом случае математическая модель стохастической эволюции резервов x_t страховой компании может быть записана в дискретном времени:

$$x^t = x^{t-1} + c^t - d^t - s^t, \tag{3}$$

где $t=1,\dots,T$ – дискретный временной параметр; $x^0 = u \ge 0$ – начальный капитал (резерв); x^t – текущий страховой резерв в момент t; c^t, d^t – детерминированные управляемые параметры – соответственно суммарные квартальные премии и дивидендные выплаты в период времени t; $s^t = s^t(c^t, c^{t-1},\dots,\xi)$ – случайные страховые выплаты за период времени с индексом t, зависящие от премий в данный (и, возможно, в предшествующие) периоды времени, а также от случайных и неучтенных факторов ξ.

Пропорциональная модель выплат. В качестве агрегированной модели страховых s^t выплат в (3) в работе [6] предложено взять соотношение:

$$s^t = c^t\xi^t, \tag{4}$$

где $\{\xi^t\}$ – реализация некоторого случайного вектора ξ, который называется уровнем выплат, и распределение которого находится из страховой статистики. Например, пусть известна историческая статистика $\{c^\tau, s^\tau, \tau=1,\dots,m\}$, тогда в качестве эмпирического распределения вектора ξ можно взять равновероятный ряд $\{\xi^\tau = s^\tau / c^\tau,\ \tau=1,\dots,m\}$. Такой метод тем более точен, чем более агрегированной по времени является модель. Неточность такого метода идентификации может проистекать из того, что часть выплат (например, доля α) по страховым случаям в период t может выплачиваться не в данный период времени, а переноситься на последующий период $(t+1)$, поэтому более точная идентификация состоит в построении равновероятного ряда $\{\xi^\tau = s^\tau \big/ ((1-\alpha)c^\tau + \alpha c^{\tau-1}),$ $\tau = 2,\dots,m\}$. В этом случае страховые выплаты имеют вид $s^t = (1-\alpha)\xi^t c^t + \alpha\xi^{t-1}c^{t-1}$, а модель эволюции резервов приобретает вид:

$$x^t = x^{t-1} + c^t - d^t - \left((1-\alpha)\xi^t c^t + \alpha\xi^{t-1}c^{t-1}\right), \quad x^0 = u, \quad t = 1,2,\dots . \tag{5}$$

Имея стохастическую модель эволюции резервов, (параллельным) методом Монте-Карло можно оценить любые характеристики работы страховой компании и их зависимость от параметров модели [7].

Регрессионная модель выплат. Естественно предположить, что в модели (3) страховые выплаты s^t в данный период времени t связаны не только с премиями c^t за этот период, но и с премиями $c^{t-1},...,c^{t-k}$ в предшествующие k периодов времени, т. е.:

$$s^t = \alpha_0 \xi^t c^t + \alpha_1 \xi^{t-1} c^{t-1} + ... + \alpha_k \xi^{t-k} c^{t-k}, \tag{6}$$

где k – параметр максимальной задержки выплат (в кварталах); $\alpha_0, \alpha_1, ...,$ $\alpha_k \geq 0$ – детерминированные коэффициенты; $\{\xi^\tau\}_{\tau=1}^t$ – реализации некоторого случайного вектора ξ, который аккумулирует случайные и неучтенные факторы модели. Выплаты s^t связаны с договорами, по которым получены премии в размере $\alpha_0 c^t + \alpha_1 c^{t-1} + ... + \alpha_k c^{t-k}$. При этом выплаты, связанные с премиями $\alpha_0 c^t$, происходят в период t; выплаты, соответствующие премиями $\alpha_1 c^t$, происходят в период $t+1$, и т. д. Наконец, выплаты, связанные с долей премий $\alpha_k c^t$, происходят в период $t+k$. Для того, чтобы правильно разделить текущие премии c^t для установления соответствия с будущими выплатами по договорам, за которые получены премии c^t, необходимо потребовать $\sum_{i=0}^k \alpha_i = 1$.

Оптимальные коэффициенты регрессии $\alpha_0^*, \alpha_1^*, ..., \alpha_k^*$ могут быть найдены методами математической статистики, а в качестве эмпирического распределения вектора ξ могут быть взяты равновероятные векторы:

$$\left\{\xi^\tau = s^\tau \Big/ \left(\alpha_0^* c^\tau + \alpha_1^* c^{\tau-1} + ... + \alpha_k^* c^{\tau-k}\right),\ \ \tau = k+1, ..., t\right\}.$$

Тогда динамика модели описывается соотношением:

$$x^t = x^{t-1} + c^t - d^t - \left(\alpha_0^* \xi^t c^t + \alpha_1^* \xi^{t-1} c^{t-1} + ... + \alpha_k^* \xi^{t-k} c^{t-k}\right),$$

$$x^0 = u, \qquad t = 1, 2, ...$$

При постоянных $c^t = c$ и $d^t = d$ эта модель приобретает вид:

$$x^t = x^{t-1} + \left(1 - \alpha_0^* \xi^t - \alpha_1^* \xi^{t-1} - ... - \alpha_k^* \xi^{t-k}\right) c - d, \quad x^0 = u, \ t = 1, 2, ... \tag{7}$$

Идентификация параметров методом наименьших квадратов и наименьших модулей. Параметры $\alpha_0, \alpha_1, ..., \alpha_k$ в предыдущей модели (6), связанные с задержками выплат, могут быть известны менеджменту страховой компании, в таком случае они задаются экспертно. Эти параметры могут быть также идентифицированы и методами математической статистики. Пусть имеется ряд наблюдений $\{(c^\tau, s^\tau),\ \tau = 1, ..., m\}$. Предположим, что выплаты связаны с премиями соотношением (6), где $\alpha = (\alpha_0, ..., \alpha_k)$ – неизвестный векторный параметр, $\{\xi^\tau\}$ – независимые реализации некоторой случайной величины ξ с неизвестным средним y. Заметим, что среднее значение $\overline{s}^\tau$ выплат в период τ задается выражением:

$$\overline{s}^\tau(\alpha_0, ..., \alpha_k; y; c^\tau, ..., c^{\tau-k}) = \left(\alpha_0 c^\tau + \alpha_1 c^{\tau-1} + ... + \alpha_k c^{\tau-k}\right) y.$$

Параметры $(\alpha_0, \alpha_1, ..., \alpha_k, y)$ могут быть найдены методом наименьших квадратов $(\beta = 2)$ или наименьших модулей $(\beta = 1)$:

$$\sum_{\tau=k+1}^{t} \left| \frac{s^\tau}{\alpha_0 c^\tau + \alpha_1 c^{\tau-1} + ... + \alpha_k c^{\tau-k}} - y \right|^\beta \to \min_{\alpha_0,...,\alpha_k, y}, \qquad (8)$$

где $y \ge 0$, $\alpha_0 \ge 0, ..., \alpha_k \ge 0$, $\sum_{i=0}^{k} \alpha_i = 1$. Содержательно, параметр y является средним уровнем выплат, которые, обычно, меньше премий, поэтому естественно добавить условие, что $0 \le y \le 1$. Метод наименьших модулей может быть полезен для идентификации параметров, поскольку он уменьшает роль экстремальных выбросов наблюдений, а такие выбросы встречаются в страховой статистике. Задача (8) – задача нелинейной глобальной оптимизации и поэтому должна решаться соответствующими методами. Для небольших значений k можно задействовать полный перебор на сетке значений параметров.

Пусть найдено оптимальное решение $\alpha^* = (\alpha_0^*, ..., \alpha_k^*)$ и y^*. В качестве эмпирического распределения случайного вектора ξ из (6) можно взять равновероятный ряд: $\left\{ \xi^\tau = s^\tau \Big/ \left(\alpha_0^* c^\tau + \alpha_1^* c^{\tau-1} + ... + \alpha_k^* c^{\tau-k} \right), \ \tau = k+1, ..., t \right\}$.

Данные. В табл. 1 представлены данные официальной годовой отчетности одной известной украинской страховой компании.

ТАБЛИЦА 1. Данные годовой отчетности страховой компании

Год	2012	2011	2010	2009	2008
Число страховых случаев, тыс.	29000	21000	25000	30000	25000
Суммарные требования, млн. грн.	188.8	209.40	264.80	245.3	198.7
Суммарные премии, млн. грн.	350.1	407.69	407.65	372.9	395.7
Уровни выплат, т. е. требования / премии, %	41.9	51.4	65.0	65.7	50.2
Страховая нагрузка	0.85	0.95	0.54	0.52	0.99

Из табл. 1 можно сделать вывод, что в день в среднем происходит около 70 страховых случаев со средней выплатой порядка 8500 грн. При этом уровень выплат (отношение суммарных страховых выплат к собранным премиям) незначительно колеблется около значения 55 %, что соответствует средней страховой нагрузке $\rho = 0.77$. При таком интенсивном потоке страховых случаев, при точности регистрации событий в один день, запаздывании процесса выплат по отношению к процессу страховых случаев, представляется, что классическая модель с пуассоновским потоком требований для компании в целом мало применима. Действительно, пуассоновский процесс – это модель появления редких событий. А для крупной компании страховой случай – это совсем не редкое событие. Классическая модель, возможно, применима для отдельных направлений страхового бизнеса компании.

На рис. 1 показано квартальную отчетность по выплатам (ряд 1) и премиям (ряд 2) той же страховой компании за 2005 – 2011 гг., где по вертикальной оси отложены выплаты и премии y в тысячах гривен (т. е. реальные величины равны $y \times 1000$ грн.), по горизонтальной оси – номера кварталов, уходящие в прошлое. На рис. 2 показан ряд из уровней выплат (отношение квартальных выплат к премиям). Из диаграмм можно сделать несколько выводов: первое – квартальные колебания уровней выплат могут быть весьма значительными, второе – выплаты и премии коррелируют, и третье – данные делятся на две группы (ближайшие 18 кварталов и более старые). Поэтому в дальнейшем мы будем работать только с данными из первой группы и искать зависимости в виде (4) или (6).

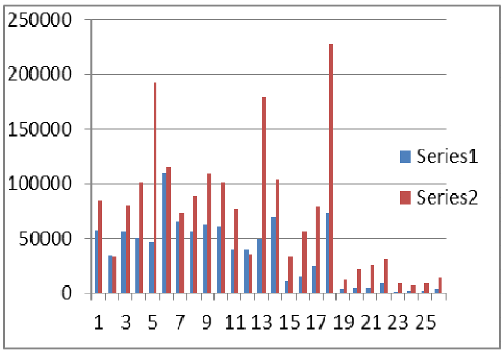

РИС. 1. Данные по выплатам и премиям

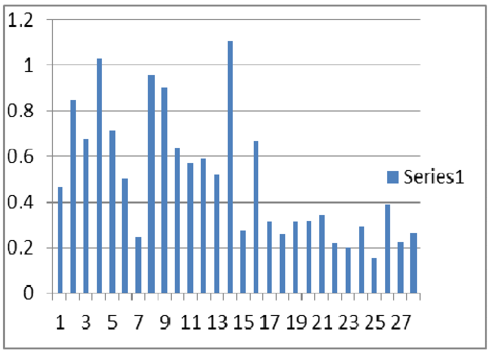

РИС. 2. Данные по уровням выплат

Результаты численных экспериментов. Вычислительные эксперименты показывают важность правильного выбора параметра задержки выплат. Например, на рис. 3 показаны примеры графиков зависимости вероятности разорения процесса (5) как функции параметра $\alpha \in [0,1]$ для $x^0 = u = 10$, $c^t = c = 1$ и двух различных значений $d = 0$ и $d = 0.15$ для статистики с рис. 1.

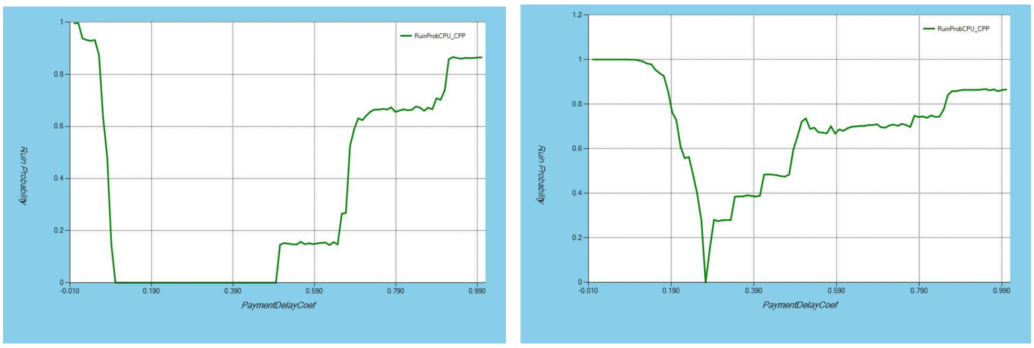

РИС. 3. Зависимость вероятности разорения как функция α при $d = 0$ и $d = 0.15$

Идентификация модели (3), (6) на реальных данных. В табл. 2, 3 приведены результаты идентификации параметров регрессионной модели страховых выплат (6) методом наименьших квадратов и наименьших модулей. Оба метода дают близкие результаты.

ТАБЛИЦА 2. Оценка параметров модели выплат методом наименьших квадратов

	Функция	Среднее ξ	α_0	α_1	α_2	α_3
$k = 0$	0.068	0.589	1			
$k = 1$	0.052	0.573	0.844	0.156		
$k = 2$	0.047	0.576	0.786	0.102	0.112	
$k = 3$	0.04	0.595	0.745	0.05	0.19	0.015

ТАБЛИЦА 3. Оценка параметров модели выплат методом наименьших модулей

	Функция	Среднее ξ	α_0	α_1	α_2	α_3
$k = 0$	0.208	0.572	1			
$k = 1$	0.181	0.610	0.794	0.206		
$k = 2$	0.167	0.608	0.744	0.164	0.092	
$k = 3$	0.141	0.615	0.7	0.175	0.06	0.065

На рис. 4, 5 показаны графики вероятности разорения процесса (7) для случая $k = 0$ ($\alpha_0 = 1$) и $k = 1$ ($\alpha_0 = 0.794$, $\alpha_1 = 0.206$), полученные с помощью системы страхового моделирования [7]. Параметры модели были следующими: $t = 40$, $c = 100$, $u = 100$, $d = 25$. Графики существенно различаются. Дальнейшее увеличение k не дает существенных изменений данных зависимостей.

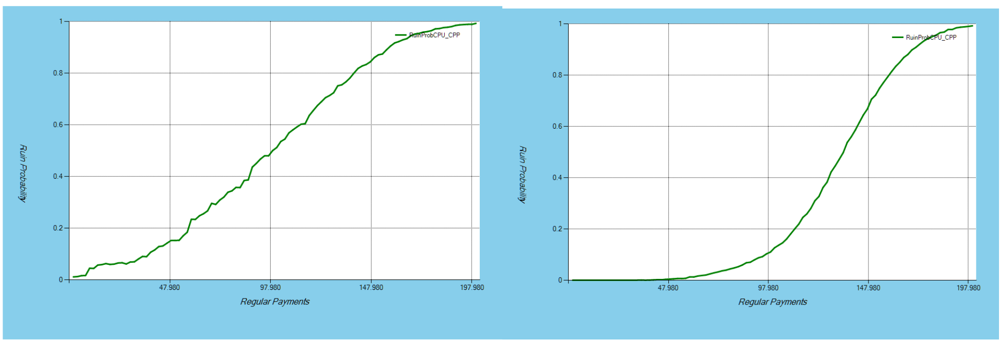

РИС. 4. Графики зависимости вероятности разорения как функции постоянных платежей d для $k = 0,1$

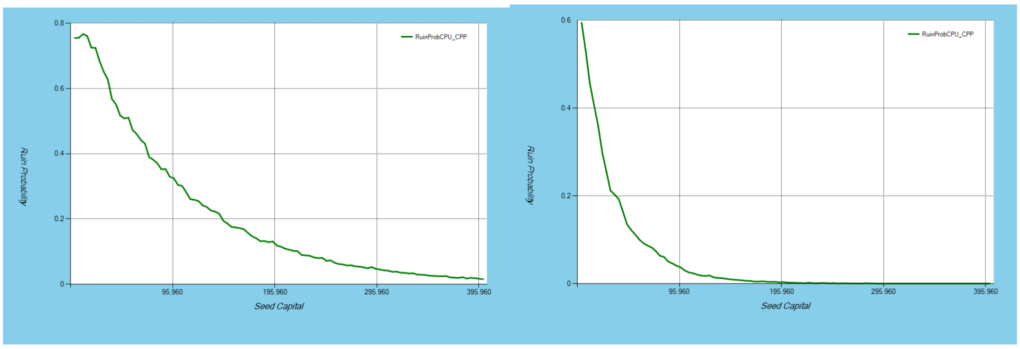

РИС. 5. Графики зависимости вероятности разорения от начального капитала u для k = 0,1

Выводы. В рамках модели эволюции страховых резервов в дискретном времени в работе предложена регрессионная модель страховых выплат, связывающая выплаты с премиями и учитывающая задержки выплат. Проведена идентификация модели на реальных данных квартальной отчетности. Показано, что задержки выплат существенно влияют на вероятность разорения компании.

Б.В. Норкін

ПРО ІДЕНТИФІКАЦІЮ МОДЕЛЕЙ ДИНАМІЧНОГО ФІНАНСОВОГО АНАЛІЗУ СТРАХОВОЇ КОМПАНІЇ

Розглядаються питання ідентифікації параметрів процесу ризику, що моделює динаміку капіталу страхової компанії. Труднощі полягають у тому, що моменти приходу окремих вимог і їх величини є внутрішньою інформацією компанії. Запропонована регресійна модель страхових виплат, що зв'язує виплати з преміями та враховує затримки виплат. Проведена ідентифікація моделі на реальних даних квартальної звітності. Показано, що затримки виплат суттєво впливають на ймовірність розорення компанії.

B.V. Norkin

ON IDENTIFICALION OF AN INSURANCE COMPANY DYNAMIC FINANTIAL ANALYSYS MODELS

We consider some aspects of insurance company assets dynamics modeling. The difficulty is that statistics for the moments of individual claim arrival and their quantities is an internal company information. The paper considers the re-insurance payments regression model that relates to the payment of premiums and takes into account the payment delay. It is shown that payments delay significantly affects ruin probability.

1. *Леоненко М.М., Мішура Ю.С., Пархоменко Я.М., Ядренко М.Й.* Теоретико-ймовірнісні та статистичні методи в економетриці та фінансовій математиці. – К.: Інформтехніка, 1995. – 380 с.
2. *de Finetti B.* Su un' impostazione alternativa della teoria collettiva del rischio // Transactions of the XV-th International Congress of Actuaries 2. – 1957. – P. 433 – 443.
3. *Kaufmann R., Gadmer A., Klett R.* Introduction to dynamic financial analysis // ASTIN Bulletin. – 2001. –Vol. 31, N 1. – P. 213 – 249.
4. *Норкин Б.В.* Математические модели оптимизации страхового дела // Кибернетика и системный анализ. – 2011. – № 1. – С. 128 – 145.
5. *Норкин Б.В.* Распараллеливание методов оценки риска банкротства страховой компании // Теорія оптимальних рішень. – 2010. – № 9. – С. 33 – 39.
6. *Норкин Б.В.* Об оптимизации портфеля страховых договоров // Прикладна статистика. Актуарна та фінансова математика (Донецкий нац. ун-т). – 2011. – № 1-2. – С. 197 – 203.
7. *Norkin B.V.* On performing actuarial calculations on GPU // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка. – К.: Век+, 2012. – № 56. – С. 113 – 119.

Получено 25.06.2013

Об авторе:

Норкин Богдан Владимирович,
кандидат физико-математических наук, докторант
Института кибернетики имени В.М. Глушкова НАН Украины.
E-mail: norkin@i.com.ua