DOI: https://doi.org/10.35681/1560-9189.2019.21.3.183724

УДК 004.942

Н. В. Кузнєцова
Інститут прикладного системного аналізу Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»
Проспект Перемоги, 37, 03056 Київ, Україна
e-mail: natalia-kpi@ukr.net

Динамічний метод оцінювання ризиків у системі фінансового менеджменту

Наведено основні визначення та формалізацію різних типів моделей виживання для оцінювання ризиків, зокрема, на основі пропорційних ризиків Кокса та їхньої модифікації, лінійних і непараметричних моделей. Запропоновано динамічний метод оцінювання ризиків, який дозволяє оцінити ступінь і рівень ризику та спрогнозувати момент переходу ризику з критичного до катастрофічного з використанням параметричних, напівпараметричних і непараметричних моделей на основі функцій виживання. Метод дозволяє застосовувати стратифікацію даних та окремо здійснювати моделювання різними функціями виживання для різних категорій даних. Автором розроблено два алгоритми, що дозволяють спрогнозувати такий момент часу на основі встановленої допустимої (критичної) імовірності настання ризику або обмеження щодо можливих економічних втрат, зокрема, момент переходу ризику від допустимого до критичного або катастрофічного (за визначеним обсягом критичних або катастрофічних втрат).

Ключові слова: *динамічний метод, фінансові ризики, моделі виживання, модель пропорційних ризиків Кокса, оцінка Каплан-Майєра.*

Вступ

Оцінювання ризиків є одним із основних засобів визначення стану компанії, підприємства, технічної, соціальної системи відносно впливу на них зовнішніх збурень, інших систем, конкурентів або випадкових впливів. Для технічних систем є характерним статичне оцінювання ризиків, тобто прогнозування реакції певної складової системи на певний тип збурення або впливу. Такий підхід є коректним у випадку, коли можна визначити всі типи збурень, які можуть виникати, і, відповідно, провести математичні розрахунки та технічні експерименти. Але визначення всіх видів таких впливів неможливо навіть для технічних систем, де здавалося би відоме і середовище функціонування, і параметри, і показники. Про це свідчать гучні «відкликання» автомобілів, мобільних телефонів, планшетів та інших пристроїв, які проходили доволі довгий період тестування, проте виявилися вразливими до появи певних ризиків або збурень, а тому зазнають значних порушень функціонування. Аналогічна ситуація спостерігається і в соціальній сфері: виникають спонтанні конфлікти, революції, протести, неочікувані та непрогнозовані зміни лідерів в електоральних уподобаннях у політичних перегонах. Економічна та фінансові системи випробовуються великою кількістю суттєвих коливань ринків, певними політичними домовленостями та рішеннями, технічними збоями та природними ризиками. Існуючий і простий підхід визначення і оцінювання ризиків статично, тобто одноразово, на початок дослідження, стає недостатнім. Виникає потреба в прогнозуванні і оцінюванні ризиків у динаміці та неперервному моніторингу і уточненню їхнього ступеню та рівня, визначенні завчасно моменту переходу ризику на більш високий ступінь або рівень. Стратегія, яку компанія обере в подальшому для своєї роботи, буде суттєво залежати від її фінансових можливостей і толерантності до ризику, тобто того, який рівень ризику і, відповідно, можливих втрат, вона може на себе прийняти.

Постановка задачі

Запропонувати метод, який дозволить оцінювати ризики в часі, визначати ймовірність настання певного ризику та можливі втрати. Запропонувати алгоритми прогнозування моменту переходу ризику на вищий рівень імовірності або втрат. Розробити динамічні моделі різних типів на основі теорії виживання та регресійних моделей і застосувати їх до вирішення практичних задач прогнозування динаміки зміни ризиків у часі.

Визначення та принцип динамічного оцінювання ризиків

Принцип динамічного оцінювання ризиків передбачає, що для оцінювання рівня втрат слід визначити очікувані моменти настання допустимого, критичного та катастрофічного ризиків, імовірність настання таких ризиків і розмір можливих втрат. Саме момент, в який переходить і змінюється рівень ризику від допустимого до катастрофічного, є найбільш важливим для оцінювання рівня втрат. Імовірність прояву ризиків також оцінюється та змінюється в часі, і момент, в який імовірність прояву ризику різко зростає, також може бути оцінений [1].

Визначення. *Динамічне оцінювання ризиків* (DE) — це оцінювання ризиків за ймовірністю, втратами та прогнозування часу (моменту) переходу ризику на більш високий (критичний) рівень з точки зору ймовірності або втрат $DE = < PR, Losses, t, S(t \mid x), \lambda(t \mid x) >$, де PR — ймовірність настання ризику; $Losses$ — рівень максимально можливих втрат; t — час; $S(t \mid x)$ — функція умовного виживання, тобто подальше функціонування фінансової системи навіть після прояву ризику; $\lambda(t \mid x)$ — умовний рівень небезпеки, тобто рівень втрат у момент часу t.

Динамічне оцінювання відрізняється від статичного передбаченою можливістю оцінювання ризиків у явному вигляді в динаміці, тобто прогнозування функції втрат і ймовірності ризику (переходу на вищий ступінь: критичний, катастрофічний) — як функцій часу. Формально це побудова функцій виживання [1].

Ключові моменти, які потребували розробки та доопрацювання для статичного та динамічного оцінювання, наведено в таблиці.

Статичне та динамічне оцінювання ризиків

Статичне оцінювання	
Проблеми та обмеження	**Методи та способи подолання**
Неповнота вхідних даних	Адаптація для коротких вибірок, комбінований метод обробки неповних даних, урахування інформаційних ризиків
Неструктурованість вхідних даних	Критерії для формування структури моделі R^2, χ^2, IV, WOE
Недостатня ефективність існуючих методів інтелектуального аналізу даних	Оцінювання ризиків інтегрованими та комбінованими моделями (нейрон-нечіткі методи, дерева рішень, регресійні та байєсові моделі)
Визначення міри ступеня ризику	Скорингові карти
Критерії якості	$GINI, CA, BS$ та критерії ефективності
Динамічне оцінювання	
Момент настання ризику	Визначення часу з використанням параметричних, напівпараметричних та непараметричних моделей
Визначення ймовірності настання фінансового ризику	Використання функції ризику (hazard function)
Оцінювання втрат на конкретний момент часу	1) функції виживання; 2) імовірнісно-статистичний метод оцінювання втрат

Установлення рівня небезпеки та ключових моментів часу, які характеризують допустимий, критичний і катастрофічний рівні ризику, є задачею системного аналізу, яку необхідно вирішувати в рамках кожного виду ризику незалежно від типу ризику та галузі, в якій він спостерігається. Автором запропоновано підхід, що базується на визначенні втрат компанії як допустимих $\lambda(t_1 \mid x) = c_1$, критичних $\lambda(t_2 \mid x) = c_2$ та катастрофічних $\lambda(t_3 \mid x) = c_3$, де c_1, c_2, c_3 — певні константи, які визначаються компанією залежно від її фінансових обертів, потужностей, тощо (наприклад, обсягу власного капіталу).

Для визначення моментів часу t_1, t_2, t_3 запропоновано метод динамічного оцінювання ризиків на основі динамічних моделей виживання та алгоритми визначення моментів часу на основі допустимих втрат і ймовірності, які детально описані далі.

Динамічні моделі прогнозування ризиків

Модель пропорційних ризиків Кокса

Відома модель Кокса, запропонована в 1972 році [2], інтенсивно досліджується та використовується в різних областях діяльності [3–5], особливо в медицині та страхуванні, для оцінки умовного ризику захворювання при заданих значеннях вихідних ознак, а робота Д. Кокса на сьогодні є найбільш цитованою публікацією за статистикою в історії. Модель Кокса заснована на припущенні, що функцію ризику можна факторизувати, тобто представити у вигляді добутку двох функцій [3]:

$$h_i(t) = h_0(t) \cdot \psi(X_{i1}, ..., X_{ik}),$$

де $h_0(t)$ — базова функція інтенсивності, що включає фактор часу, але не включає коваріанти, а $\psi(X_{i1},...X_{ik})$ — лінійна функція досліджуваних ознак, яка не включає фактор часу.

Досить часто модель записують у наступному вигляді [3]:

$$h_i(t) = h_0(t) \cdot e^{\{\beta_1 x_{i1} + ... + \beta_k x_{ik}\}},$$

$$\ln h_i(t) = \ln h_0(t) + \beta_1 x_{i1} + ... + \beta_k x_{ik},$$

де $\beta_1, ..., \beta_k$ — невідомі параметри.

Модель пропорційних ризиків Кокса у вигляді функції умовного виживання $S(t \mid x)$ передбачає оцінку сукупної умовної функції ризику $L(t \mid x)$ з використанням максимальної правдоподібності. Метою є розробка умовної моделі для індивідуального $S(t \mid x)$, яка визначена в термінах $L(t \mid x)$. Для того, щоб описати $P\hat{D}^{PHM}$ (ймовірність дефолту при моделі пропорційних ризиків) визначимо наступні вирази по відношенню до теорії регресії Кокса [2].

Емпірична функція виживання

Усі спостереження вважаються завершеними та відсортованими за довжиною часових відрізків, починаючи з найменшого. Оцінка функції виживання обраховується за наступною формулою:

$$S_n(t) = \frac{\sum_{i=1}^{n} c_i(t)}{n}, \tag{1}$$

де $c_i(t) = \begin{cases} 1, t_i > t, \\ 0, t_i \le t, \end{cases}$ n — кількість спостережень.

Крива функції виживання $S_n(t)$ є ступінчастою функцією, кожна точка розриву якої відповідає смерті одного чи декількох суб'єктів. Оцінка математичного сподівання у такому випадку обраховується за формулою:

$$\hat{p} = \int_0^\infty S_n(t) dt. \tag{2}$$

Моделювання ризиків на основі припущення щодо їхньої пропорційності наразі є найбільш поширеним у кредитній галузі. Для оцінювання ризиків проводилося моделювання інших видів залежностей, зокрема представлення ймовірності настання ризику у вигляді узагальненої лінійної моделі або параметричної моделі, які будуть описані далі.

Узагальнена лінійна модель

Для визначення «часу життя» до настання ризику можна сформувати таку узагальнену лінійну модель:

$$P(T \le t \mid X = x) = F_\theta(t \mid x) = g(\theta_0 + \theta_1 t + \theta^T x),$$

де $\theta = (\theta_2, \theta_3, .., \theta_{p+1})^T$; p — вимірний вектор; g — відома функція зв'язку, така як логістична чи пробіт-функція. Таким чином, ця модель характеризує умовний розподіл часу життя до настання ризику T [5] у термінах невідомих параметрів. Після оцінювання параметрів буде отримана оцінка функції умовного розподілу, $F_{\hat{\theta}}$ і, нарешті, оцінка ймовірності настання ризику може бути обчислена таким чином:

$$P\hat{R}^{GLM}(t \mid x) = \frac{F_{\hat{\theta}}(t+b \mid x) - F_{\hat{\theta}}(t \mid x)}{1 - F_{\hat{\theta}}(t \mid x)} = 1 - \frac{S_{\hat{\theta}}(t+b \mid x)}{S_{\hat{\theta}}(t \mid x)}, \tag{3}$$

де $\hat{\theta} = \hat{\theta}^{GML}$ є оцінкою максимальної правдоподібності вектора параметрів.

Розглянемо одновимірний випадок коваріантності. У такому випадку $\theta = \theta_2$ і умовний розподіл задається моделлю $F(t \mid x) = g(\theta_0 + \theta_1 t + \theta_2 x)$, зі щільністю $f(t \mid x) = \theta_1 g'(\theta_0 + \theta_1 t + \theta_2 x)$. Оскільки зазвичай задана випадкова цензурована справа вибірка, то умовна функція правдоподібності [5] представляє собою добуток членів, що включають умовну щільність, для нецензурованих даних та умовної функції виживання для цензурованих даних:

$$L(Y, X, \theta) = \prod_{i=1}^{n} f(Y_i \mid X_i)^{\delta_i} (1 - F(Y_i \mid X_i))^{1-\delta_i}, \tag{4}$$

де Y_i — строк обслуговування i-го клієнта до настання ризику, а δ_i є індикатором настання ризику для i-го клієнта.

Таким чином, логарифмічна функція правдоподібності визначається як:

$$l(\theta) = \ln(L(Y, X, \theta)) = \sum_{i=1}^{n} [\delta_i \ln(f(Y_i \mid X_i)) + (1-\delta_i)\ln(1 - F(Y_i \mid X_i))] =$$

$$= \sum_{i=1}^{n} [\delta_i \ln(\theta_i g'(\theta_0 + \theta_1 Y_i + \theta_2 X_i)) + (1-\delta_i)\ln(1 - g(\theta_0 + \theta_1 Y_i + \theta_2 X_i))] =$$

$$= \sum_{i=1}^{n} \delta_i [\ln(\theta_1) + \ln(g'(\theta_0 + \theta_1 Y_i + \theta_2 X_i))] + \sum_{i=1}^{n} (1-\delta_i)\ln(1 - g(\theta_0 + \theta_1 Y_i + \theta_2 X_i)).$$

І, нарешті, оцінка знаходиться як максимізація функції логарифмічної правдоподібності:

$$\hat{\theta}^{GML} = \arg\max_{\theta} l(\theta).$$

Таким чином, використання узагальненої лінійної регресії можливе для оцінювання та прогнозування часу настання ризику при вирішенні задачі максимізації функції логарифмічної правдоподібності.

Розвиток моделі пропорційних ризиків

За припущенням модель Кокса [2, 4] передбачала, що всі фактори в моделі мають один і той самий час виміру, тобто не потребують урахування лагових (попередніх) значень. Крім цього передбачається лінійна залежність моделі ризику від інших факторів і адитивність функції ризику, тобто спільний ефект від усіх предикторів є сумою ефектів окремих предикторів. У зальному випадку це не так, тому було розроблено моделі, що враховують і інші варіанти.

Модель Вейбулла

Модель є незначною модифікацією експоненційної моделі. Для неї зберігається припущення, що ε має стандартний розподіл екстремальних значень, проте послаблюється припущення, що $\sigma = 1$. Для $\sigma > 1$ ризик зменшується в часі, для $0,5 < \sigma < 1$ ризик зростає в часі зі швидкістю, що зменшується. Для $\sigma = 0,5$ функція ризику зростає лінійно, починаючи з 0. Модель називається моделлю Вейбулла (Weibull), тому що час T розподілений за розподілом Вейбулла, умовним для коваріант. Модель є досить популярною у біостатистиці через те, що функція виживання є доволі простою з математичної точки зору:

$$S_i(t) = \exp\{-[t_i e^{-\beta x_i}]^{\frac{1}{\sigma}}\},$$

де x_i — вектор значень коваріант; β — вектор коефіцієнтів.

Крім цього, модель є також моделлю пропорційних ризиків. Це означає, що її коефіцієнти можуть бути представлені у вигляді відповідного співвідношення.

Існує співвідношення між лог-формою моделі функції виживання

$$\log h(t) = \alpha \log t + \beta_o^* + \beta_1^* x_1 + \ldots + \beta_k^* x_k$$

та лог-моделлю для часу

$$\log T_i = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \sigma\varepsilon .$$

Для моделі Вейбулла відношення між параметрами порівняно з експоненційною моделлю визначається наступним чином:

$$\beta_j^* = \frac{-\beta_j}{\sigma} \text{ для } j = 1,..,k \text{ та } \alpha = (1/\sigma) - 1 .$$

Оскільки $\beta_j = 0$ тільки тоді, коли $\beta_j^* = 0$, і тест на нуль-гіпотезу, що коефіцієнт дорівнює 0, буде таким самим, незважаючи, яка форма використовується. З іншого боку, стандартна похибка і довірчі інтервали для коефіцієнтів у логарифмічній моделі часу виживання не так легко представляються у логарифмічному форматі ризику.

Функція ризику з часовою залежністю предикторів

$$h(t) = h_0(t) e^{\{\beta_1 X_1 + \beta_2 (X_1 \cdot f(t))\}}, \tag{5}$$

де $\beta_1 X_1$ — частина регресорів, що не залежить від часу; $\beta_2 (X_1 \cdot f(t))$ — частина регресорів, залежних від часу [3].

$$\ln h_i(t) = \alpha(t) + \beta_1 x_{i1} + \ldots + \beta_2 x_{i2}(t) . \tag{6}$$

Моделювання різними функціями ризику для різних страт

У моделі Кокса припускалося, що всі індивіди розподілені за однаковим законом, який і потрібно знайти. Проте в реальних практичних задачах оцінювання ризиків виникає необхідність моделювання різних груп клієнтів або класів, побудови для них різних моделей і скорингових карт залежно від віку, доходу, типу клієнта, типу позики і т.д. Для того щоб врахувати різницю у поведінці таких груп (страт) і було запропоновано використання стратифікації і моделювання їхніми різними функціями [3]:

$$h_{ig}(t) = h_{0g}(t)e^{\{\beta_1 x_{i1} + \ldots + \beta_k x_{ik}\}}, \tag{7}$$

причому $h_{0g}(t)$ буде різною базовою функцією для кожної страти, а параметри $x_{i1},\ldots,x_{ik}$ будуть по-різному обраховуватися для кожної страти. Таким чином, будуть формуватися різні функції виживання після реалізації ризику для кожної групи або страти.

Тоді

$$h_g(t) = h_{0g}(t)e^{\{\beta_1 X_1 + \beta_2(gX_1)\}} \text{ або } h_g(t) = h_{0g}(t)e^{\{\beta_{g1} X_1\}}, \tag{8}$$

$$L = \prod_{i=1}^{g} L_i,$$

де L_i — часткова функція правдоподібності страти i; g — номер страти.

Для оцінювання параметрів моделі методом часткової правдоподібності необхідно виконати наступні кроки [3]:

1) побудувати різні функції виживання для кожної групи або страти;

2) знайти добуток цих функцій;

3) обрати значення параметрів β, які максимізують функцію часткової правдоподібності.

Непараметрична модель оцінювання фінансових ризиків

Непараметрична модель, що описана у [5], передбачає застосування запропонованого Бераном [6] оцінювача для умовної функції виживання:

$$\hat{S}_h(t \mid x) = \prod_{i=1}^{n} \left(1 - \frac{1_{\{Y_i \le t, \delta_i = 1\}} B_{ni}(x)}{1 - \sum_{j=1}^{n} 1_{\{Y_j < Y_i\}} B_{nj}(x)} \right),$$

де Y_i — спостережуваний час для i-го клієнта; δ_i — це індикатор, що показує настання ризику для i-го клієнта; X_i — вектор пояснюючих коваріант для i-го клієнта. Терми B_{ni} — це непараметричні ваги Надарая-Вотсона [3]:

$$B_{ni}(x) = \frac{K((x - X_i)/h)}{\sum_{j=1}^{n} K((x - X_j)/h)}, \quad 1 \le i \le n,$$

і $h \equiv h_n$ є згладжуючим параметром, який наближається до нуля, якщо розмір популяції наближається до нескінченності.

Для оцінювання ймовірності настання ризику в момент *t* для заданого набору параметрів виконують підстановку, тоді теоретичне значення умовної функції виживання [3] за його оцінкою $\hat{S}_h$ буде:

$$P(\hat{R}^{NPM}(t \mid x)) = \frac{\hat{F}_h(t+b \mid x) - \hat{F}_h(t \mid x)}{1 - \hat{F}_h(t \mid x)} = 1 - \frac{\hat{S}_h(t+b \mid x)}{\hat{S}_h(t \mid x)}.$$

Оцінки для порівняння динамічних моделей

Оцінка Каплан-Майєра (КМ) є узагальненням емпіричної функції виживання і враховує відцензуровані спостереження. Формула Каплан-Майєра [7] для ймовірності виживання у певний час t_j обмежується добутком характеристик, що відповідають особам, що залишились живими після настання часу t_j. Тому часто таку оцінку також називають Product-limit estimator [8].

Для обрахунку оцінки КМ усі спостереження сортують у порядку зростання часу їхнього життя. Перше входження починається з нуля. Ймовірність виживання для цього часу дорівнює 1. Подальші спостереження виключаються в момент часу їхньої загибелі (можливо, у результаті цензурування) [7]. Множина ризику, що позначається $R(t_j)$ — це кількість усіх індивідів, що дожили хоча б до часу t_j.

Основна ідея оцінки — представлення її у вигляді:

$$\hat{S}(t_i) = \prod_{j=1}^{i} \hat{P}(T > t_j \mid T \geq t_j). \tag{9}$$

Враховуючи те, що

$$\hat{P}(T > t_j \mid T \geq t_j) = \frac{n_j - d_j}{n_j}. \tag{10}$$

оцінку Каплан-Майєра представимо так:

$$\hat{S}(t) = \prod_{t_j < t} (1 - \frac{d_j}{n_j}), \tag{11}$$

де d_j — кількість зразків, що вибули під час дослідження (загибель у час t_j ($d_j = |D_j|$)); n_j — кількість зразків, які знаходяться під загрозою вибування (number at risk).

За обчисленою оцінкою Каплан-Майєра можливе порівняння функцій виживання для різних груп, що моделюються при оцінюванні фінансових ризиків. Тобто є реальний інструмент порівняння і визначення часу життя та часу переходу до більш високого ступеня ризику для різних страт, клієнтів, груп:

$$\textit{Оцінка LogRank} = \frac{(\sum_{j=1}^{r}(d_{1j} - e_{1j}))^2}{\mathrm{var}(\sum_{j=1}^{r}(d_{1j} - e_{1j})}, \tag{12}$$

де e_{1j} — кількість випадків 1-го виду на j-му часовому інтервалі; r — кількість часових проміжків [3].

Цей критерій подібний до χ^2 і використовується, коли функції ризику однакові, а рівень ризику є константою в часі.

Оцінка Вілкоксона (Wilcoxon) обчислюється за виразом:

$$Wilcoxon = \frac{(\sum_{j=1}^{r} n_j (d_{1j} - e_{1j}))^2}{\mathrm{var}(\sum_{j=1}^{r} n_j (d_{1j} - e_{1j})}. \tag{13}$$

Критерій порівняння функцій виживання, є чутливим до зразків, що цензуруються. Його рекомендовано [1, 3] використовувати на ранніх періодах дослідження.

Алгоритми прогнозування часу настання ризику

Візуальне порівняння кривих виживання для різних страт або груп, яке часто використовується, при прогнозуванні часу настання ризику є можливим, але не дуже зручним. Для задач, де зрізи даних надходять у часовому проміжку щоденно, погодинно, щохвилинно, і точність прогнозування часу є критичною, необхідно розробити алгоритми встановлення часу.

Для цього можна використати кілька різних підходів. Якщо функція ризику визначається у процесі моделювання через параметричний, непараметричний розподіл, то можливе обчислення часу через похідну функції ризику.

Алгоритм 1 розрахунку моменту переходу на вищий ступінь ризику.

Для визначення моменту часу t виконаємо наступні кроки.

1. Задамо вид початкової функції $\hat{\Lambda}_0(t)$. Нехай $\hat{\Lambda}_0(t) = \exp(-a \cdot t)$, де $a > 0$, але має невелике значення, щоб $\hat{\Lambda}_0(t) = \exp(-a \cdot t)$ не швидко спадала.

2. Підставивши базову функцію небезпеки для пропорційних ризиків, отримаємо: $\hat{\Lambda}(t) = \exp(x^T \cdot \beta^{PHM}) \cdot \exp(-a \cdot t)$.

3. Взявши похідну від функції небезпеки за часом, отримаємо: $\frac{\partial \Lambda(t \mid x)}{\partial t} = \exp(x^T \cdot \beta^{PHM}) \cdot \exp(-a \cdot t) \cdot (-a) = -a \cdot \exp(x^T \cdot \beta^{PHM} - a \cdot t)$.

Для можливості диференціювання функції необхідно існування її похідної, а це не завжди можна гарантувати. Тому можливо адаптувати алгоритм без прямого обрахунку похідної. Виходячи з визначення похідної як швидкості зміни певної функції, можна обчислити швидкість зміни ймовірності, тобто переходу її до критичної імовірності настання ризику $P_{крит}(t)$:

1) задається ймовірність, яка є критичною, $P_{крит}$;

2) розраховується значення $P_{поточн}(t) - P_{крит} = \Delta P$ як «запас по ймовірності»;

3) визначається момент переходу ризику до критичного як:

$$t_{крит} = \frac{\Delta P}{\frac{\partial P(t)}{\partial t}}.$$

Якщо встановити ймовірність ризику, яка є критичною, неможливо, тоді пропонується розробити алгоритм розрахунку часу через критичний (або катастрофічний) рівень ризику, тобто критичні (або катастрофічні) втрати. Це завжди є можливим, оскільки фінансова система або підприємство функціонує для отримання певного прибутку, а тому може визначити, які втрати за ризиками є більшими, ніж отриманий прибуток.

Алгоритм 2 визначення моменту настання критичного (катастрофічного) рівня ризику за втратами.

1. Задати інтервал часу, на якому буде здійснюватися пошук критичного часу $\Delta T = (0, T)$.
2. Задати крок збільшення часу $t := \Delta t$.
3. $t := 0$.
4. Обчислити $Losses_{поточ}(t)$.
5. Якщо $Losses_{поточ}(t) \geq Losses_{крит}$, то $t_{крит} := t$ і STOP.
6. $t := t + \Delta t$.
7. Якщо $t \geq \Delta T$, то STOP, і в цьому інтервалі не відбудеться перехід ризику до критичного (катастрофічного).
8. Go to Step 4.

За визначеним на попередньому кроці t можна обчислити: $\lambda(t \mid x)$, $P(t \mid x)$ та очікувані втрати EL для: допустимого, критичного та катастрофічного рівнів ризику в моменти часу (t_1, t_2, t_3).

Розроблені алгоритми дозволяють визначити не лише ступінь і рівень ризику, як передбачалось у статичному оцінюванні, а й спрогнозувати момент часу, коли рівень або ступінь ризику різко змінюється.

Метод динамічного оцінювання ризиків

Для моделювання ризиків у динаміці автором було запропоновано метод динамічного оцінювання та прогнозування ризиків, який передбачає побудову динамічних моделей різних типів для окремих страт, визначення кращої з них, та використання такої моделі для прогнозування часу, рівню та ступеню ризику. Метод може бути представлений як послідовність описаних нижче кроків.

Крок 1. Визначення значущості характеристик фінансових ризиків-параметрів моделей на основі критеріїв:

— оцінки кореляції $R^2, \chi^2, ACF, PACF$ [9];

— ваги категорії змінної: $WOE_i = \ln\left(\dfrac{g_i}{b_i}\right)$, де $\sum_{i=1}^{k} g_i = 1$ — розподіл одиничних значень цільової змінної та $\sum_{i=1}^{k} b_i = 1$ — розподіл нульових значень цільової змінної [10];

— інформаційного значення (IV): $IV = \sum_{i=1}^{k}(g_i - b_i)\ln\left(\dfrac{g_i}{b_i}\right) = \sum_{i=1}^{k}(g_i - b_i)WoE_i$.

Крок 2. Розробка моделей виживання різного виду.

За можливістю може бути висунуте припущення щодо розподілу ризиків (пропорційні ризики, з часовими коваріантами, непараметричні моделі тощо). Тоді буде будуватися тільки один вид моделей виживання. Якщо явно визначити вид моделі неможливо, тоді будується множина обраних видів моделей і визначається краща з них.

2.1. Модель пропорційних ризиків.

2.1.1. Оцінка функції умовного рівня небезпеки визначається як

$$\hat{\lambda}(t \mid x) = \hat{\lambda}_0(t)\exp(x^T\hat{\beta}),$$

де $\hat{\lambda}_0(t)$ — оцінка базової функції рівня небезпеки $\lambda_0(t)$ [1, 5]; $\hat{\beta}$ — оцінка вектора параметрів β, а ймовірність настання ризику

$$P\hat{R}^{PHM}(t \mid x) = \frac{\hat{F}_{\hat{\beta}}(t+b \mid x) - \hat{F}_{\hat{\beta}}(t \mid x)}{1 - \hat{F}_{\hat{\beta}}(t \mid x)} = 1 - \frac{\hat{S}_{\hat{\beta}}(t+b \mid x)}{\hat{S}_{\hat{\beta}}(t \mid x)},$$

де $1 - \hat{F}_{\hat{\beta}}(t \mid x) = \hat{S}_{\hat{\beta}}(t \mid x) = \exp(-\hat{\Lambda}(t \mid x))$.

2.1.2. Оцінюється інтегральна функція базового ризику $\Lambda_0(t)$:

$$\hat{\Lambda}_0(t) = \sum_{i=1}^{n} \frac{1\{Y_i \le t, \delta_i = 1\}}{\sum_{j=1}^{n} 1\{Y_j \ge Y_i\}},$$

2.1.3. Параметр β оцінюється як $\hat{\beta}^{PHM} = \arg\max_{\beta} L(\beta)$, де часткова функція правдоподібності задається виразом [1]:

$$L(\beta) = \prod_{i=1}^{n} \frac{\exp(x_i^T\beta)}{\left(\sum_{j=1}^{n} 1_{\{Y_j > Y_i\}} \exp(x_j^T\beta)\right)}.$$

2.1.4. Оцінка умовної інтегральної функції ризику визначається за формулою [5]:

$$\hat{\Lambda}(t \mid x) = \int_0^t \hat{\lambda}(s \mid t)ds = \exp(x^T\hat{\beta}^{PHM})\hat{\Lambda}_0(t).$$

2.2. Побудова узагальненої лінійної моделі для оцінювання ризиків.

Припущення щодо пропорційності ризиків не завжди є допустимим, а тому можливе представлення функції ризику та функцій виживання через узагальнену лінійну регресію (3), (4) [9]. Оцінка параметрів визначається в результаті максимізації логарифмічної функції правдоподібності: $\hat{\theta}^{GML} = \arg\max_{\theta} l(\theta)$.

2.3. Побудова непараметричної моделі оцінювання фінансових ризиків.

$$\hat{S}_h(t \mid x) = \prod_{i=1}^{n} \left(1 - \frac{1_{\{Y_i \le t, \delta_i = 1\}} B_{ni}(x)}{1 - \sum_{j=1}^{n} 1_{\{Y_j < Y_i\}} B_{nj}(x)} \right),$$

де

$$B_{ni}(x) = \frac{K((x - X_i)/h)}{\sum_{j=1}^{n} K((x - X_j)/h)}, 1 \le i \le n,$$

$$P(\hat{R}^{NPM}(t \mid x)) = \frac{\hat{F}_h(t+b \mid x) - \hat{F}_h(t \mid x)}{1 - \hat{F}_h(t \mid x)} = 1 - \frac{\hat{S}_h(t+b \mid x)}{\hat{S}_h(t \mid x)}.$$

Крок 3. Порівняння моделей та обрання кращої за множиною статистичних критеріїв.

3.1. Обрання кращої моделі на основі критеріїв GINI, AUC, неправильної класифікації (MR), Байєса-Шварца (BSC), Колмогорова-Смірнова (KS):

$$AUC = \int_{\tilde{y} \in \tilde{Y}} Se(\tilde{y}) dFPR(\tilde{y}), \; GINI = 2 \cdot AUC - 1, \; MR = \frac{FP + FN}{N},$$

$$BSC = N \ln \left(\sum_{k=1}^{N} e^2(k) \right) + n \ln(N), \; KS = \max_{x \in X} | F_B(x) - F_G(x) | \text{ тощо.}$$

3.2. Перевірка гіпотези про однаковий розподіл функцій ризику.

Застосовуються критерії: оцінка Каплан-Майєра, LogRank, Wilcoxon (11)–(13) для порівняння функцій виживання, що були побудовані для різних страт.

Крок 4. Визначення моменту настання ризику.

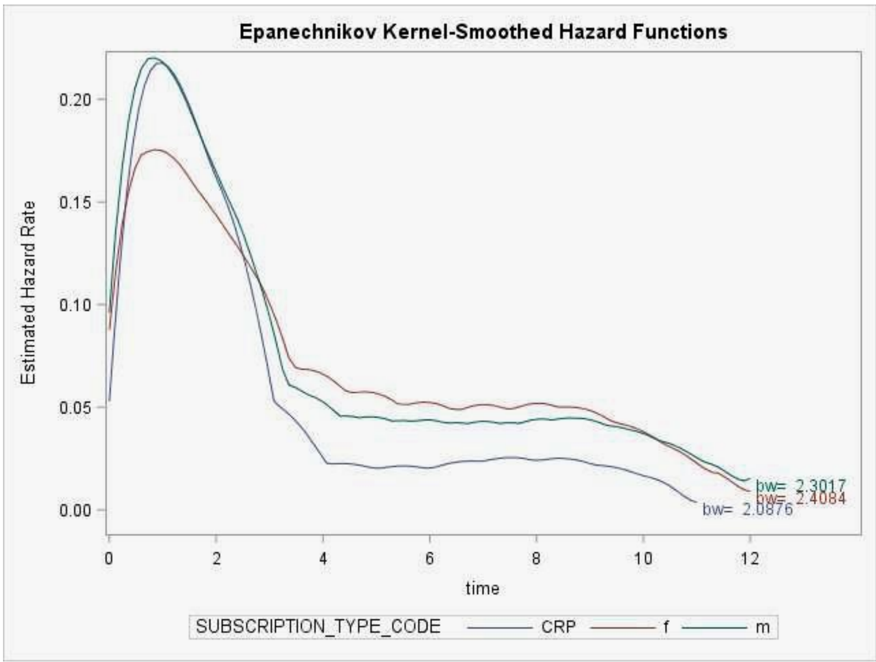
Залежно від постановки задачі, тобто необхідності визначення моменту переходу ризику на більш високий ступінь або рівень, використовується один з описаних алгоритмів.

Крок 5. Визначення ймовірності настання фінансового ризику та можливих втрат.

За визначеним на попередньому кроці t можна обчислити: $\lambda(t \mid x)$, $P(t \mid x)$ та очікувані втрати EL для допустимого, критичного та катастрофічного рівнів ризику в моменти часу (t_1, t_2, t_3) та відповідних функцій виживання.

Було проведене експериментальне дослідження застосування динамічного методу оцінювання ризиків для практичних задач моделювання кредитних ризиків і ризиків телекомунікаційної компанії. Результати моделювання наведено у роботах [1, 11, 12]. Обчислення настання моменту часу критичного рівня або ступеню ризику дозволило визначити для телекомунікаційних компаній критичних і схильних до відтоку клієнтів, обчислити можливі втрати. Для кожної групи клієнтів може бути оцінений рівень ризику залежно від фактору часу (див. рисунок) та обсяг можливих втрат. Залежно від фінансової ситуації компанії, можуть бути здійснені переоцінювання сукупного рівня ризику через можливі втрати. Для цього оцінюється сукупний рівень втрат без деталізації за типами клієнтів і, відповідно, сукупних втратах за моделлю Кокса та сукупною функцією ризику, або окремо будуються моделі збитковості за типами клієнтів.

Так, наприклад у роботі [1], на початковому етапі визначено, що телекомунікаційна компанія вважає допустимим рівнем ризику такий, який за функцією ризику не більше $\lambda(t_1) = 0{,}15$, критичний не більший $\lambda(t_2) = 0{,}3$, а катастрофічний більший за $\lambda(t_3) = 0{,}4$. Як видно за побудованою функцією збитковості за окремими групами клієнтів (де *f* — жінки, *m* — чоловіки, а CRP — корпоративні клієнти), перехід ризику в зону критичного або катастрофічного ризику не відбувається зовсім (див. рисунок). Найбільший рівень втрат спостерігався серед груп чоловіків і корпоративних клієнтів через 1 місяць та 1,5 місяці. У разі критичного рівня ризику були задіяні завчасно напрацьовані запобіжні заходи для того, щоб уникнути такої фінансової небезпеки.



Графік функції збитковості для згрупованих даних

Під час моделювання кредитних ризиків [12] динамічний метод оцінювання дозволив установити групи клієнтів і проміжки часу, де клієнти мають прострочки в оплаті та потребують додаткових механізмів для запобігання фінансових втрат, що пов'язані з несплатою щомісячних платежів за кредитами.

Висновки

Динамічне оцінювання ризиків є сьогодні актуальним і необхідним підходом для технічних, економічних, соціальних сфер життя, його доцільно використовувати для банків і фінансових компаній, які працюють на ринку в умовах перехідного періоду та змушені швидко реагувати на зовнішні впливи та дії конкурентів. Такий підхід може бути рекомендовано для використання у Національному банку України при проведенні стрес-тестування українських банків, для оцінювання їхньої діяльності і своєчасного виявлення та реагування на економічні проблеми всередині таких банків. У статті наведено формалізацію динамічних моделей на основі пропорційних ризиків Кокса та їхню модифікацію, лінійних і непараметричних моделей. Запропоновано та розроблено метод динамічного оцінювання ризиків, який передбачає побудову різних типів моделей виживання і прогнозування часу переходу на вищий ступінь або рівень ризику, використовує переваги ідеологічно різних методів і дозволяє оцінювати можливі втрати у формі точкових оцінок і ймовірностей настання ризикових ситуацій. Запропоновано два алгоритми прогнозування часу: алгоритм, який дозволяє визначити момент переходу ризику від допустимого до критичного або катастрофічного (за визначеним обсягом критичних або катастрофічних втрат) та алгоритм розрахунку моменту переходу ризику на вищий ступінь за заданою для фінансової системи критичною ймовірністю ризику.

1. Кузнєцова Н.В., Бідюк П.І. Динамічне моделювання фінансових ризиків. *Індуктивне моделювання складних систем*. 2017. Вип. 9. С. 122–137.

2. Cox D.R. Regression Models and Life-Tables. *Journal of the Royal Statistical Society, Series B.* 1972. Vol. 34. No 2. P. 187–220.

3. Allison P.D. Survival Analysis Using SAS: A Practical Guide: Second Edition. Cary, NC: SAS Institute Inc., 2010. 324 p.

4. Cox D.R. The regression analysis of binary sequences (with discussion). *Journal of the Royal Statistical Society, Series B.* 1958. Vol. 20. P. 215–242.

5. Cao R., Vilar J.M., Devia A. Modelling consumer credit risk via survival analysis. *SORT*. January-June 2009. **33**(1). P. 3–30.

6. Beran J., Dja A.K. Credit risk modeling based on survival analysis with immunes. *Statistical Methodology*. 2007. Vol. 4. P. 251–276.

7. Kaplan E.L., Meier P. Non-parametric estimation for incomplete observations. *J. Am. Stat. Assoc*. 1958. No 53. P. 457–481. URL: http://www.jstor.org/stable/2281868 (Last accessed 25.09.2018).

8. Dabrowska D. Non-parametric regression with censored survival time data. Scandinavian *Journal of Statistics*. 1987. Vol. 14, No 3. P. 181–197.

9. Бідюк П.І., Романенко В.Д., Тимощук О.Л. Аналіз часових рядів: навч. посіб. — Київ: НТУУ «КПІ», 2013. 600 с.

10. Shutt R, O'Neil C. Doing Data Science. Straight Talk From the Frontline. O'Reilly, 2013. 408 p.

11. Kuznietsova N.V. Information Technologies for Clients' Database Analysis and Behaviour Forecasting. *CEUR Workshop Proceeding* (ISSN 1613-0073) 2017. Vol. 2067. P. 56–62 [Online]. URL: http://ceur-ws.org/Vol-2067/ (Last accessed 25.09.2018).

12. Kuznetsova N.V., Bidyuk P.I. Modeling of credit risks on the basis of the theory of survival. *Journal of Automation and Information Sciences*. 2017. Vol. 49. Isssue. 11. P. 11–24.

Надійшла до редакції 10.07.2019