

DOI: 10.20535/kpissn.2021.3.251681

УДК 519.711.3

Л.Б. Левенчук, В.Г. Гуськова, П.І. Бідюк*

КПІ ім. Ігоря Сікорського, Київ, Україна

*corresponding author: pbidyuke_00@ukr.net

ЙМОВІРНІСНЕ МОДЕЛЮВАННЯ ОПЕРАЦІЙНИХ РИЗИКІВ

Received: 23 Jul. 2021. Accepted: 7 Dec. 2021.

Проблематика. Операційні ризики (ОР) притаманні всім видам людської діяльності, включно з промисловістю, транспортом, фінансовими послугами та ін. Ризики такого типу характеризуються множиною невизначеностей, неповнотою та низькою якістю даних, що ускладнює прогнозування й оцінювання. Для моделювання процесів, пов'язаних з ОР, необхідно виконати коректну обробку даних, а також ідентифікацію та врахування можливих невизначеностей. Для розв'язання таких задач придатний ймовірнісний підхід.

Мега дослідження. Зробити короткий огляд методів ймовірнісного аналізу даних, призначених для побудови математичних моделей ОР. Розробити нову ймовірнісну модель у вигляді мережі Баєса для формального опису операційного ризику шахрайства, пов'язаного з актуарними процесами.

Методика реалізації. Для обробки даних й експертних оцінок використовуються методи баєсового аналізу даних, призначені для побудови ймовірнісних моделей у формі баєсових мереж.

Результати дослідження. Описано методи дослідження ОР і їх модифікації. Запропонований метод моделювання застосовано для побудови моделі ОР страхового шахрайства. Для побудови моделі проаналізовано задачу, вибрано множину змінних й оцінено необхідні апріорні умовні ймовірності. Остаточна модель побудована з використанням інструментарію моделювання GeNIe. Функціонування моделі продемонстровано на прикладі.

Висновки. Показано, що задача моделювання, оцінювання та прогнозування ОР може бути розв'язана за допомогою ймовірнісного підходу, зокрема баєсової методології, яка дає змогу ідентифікувати та врахувати можливі невизначеності даних й експертних оцінок.

Ключові слова: операційний ризик; ймовірнісне моделювання; баєсова методологія; експертні оцінки; актуарні процеси.

Вступ

Останніми десятиліттями зростає інтерес наукових і ділових кіл до управління ризиками, причому не тільки у фінансовому секторі, а й в інших сферах діяльності: керуванні підприємствами, управлінні проектами, менеджменті ресурсів різних видів та ін. На відміну від суто фінансових ризиків, які визначаються змінами на фінансовому ринку та в економіці і характеризуються розрахунковими властивостями обраних моделей, ОР розглядають з точки зору недостатньо якісного функціонування систем (особливо керування), незадовільної реалізації виробничих процесів і негативного впливу випадкових зовнішніх факторів, з урахуванням можливих людських помилок.

Категоріями ОР є [1]–[3].

- *Людські ризики:* ризик виникнення людських помилок з боку менеджменту, персоналу, контрагентів – будь-кого, хто має вплив на діяльність підприємства.

- *Процесуальні ризики:* ризики, пов'язані з процедурами та практиками виробничої діяльності підприємств і компаній.

- *Системні ризики:* ризики, пов'язані з технологічними (апаратними та програмними) засобами, що використовуються на підприємствах, а також засобами безпеки функціонування підприємств.

- *Ризики зовнішніх подій:* форс-мажорні обставини, випадкові перебої в постачанні та наданні послуг, зміни в структурах систем, швидкі зміни на ринку товарів, непередбачувані змі-

Рекомендуємо цитувати цю статтю так: Л.Б. Левенчук, В.Г. Гуськова, П.І. Бідюк, “Ймовірнісне моделювання операційних ризиків”, *Наукові вісті КПІ*, № 3, с. 26–37, 2021. doi: 10.20535/kpissn.2021.3.251681.

Please cite this article as: L.B. Levenchuk, V.H. Huskova, and P.I. Bidyuk, “Probabilistic modelling of operational risks”, *KPI Science News*, no. 3, pp. 26–37, 2021. doi: 10.20535/kpissn.2021.3.251681.

ни під час виробничої діяльності підприємств і компаній.

Процес управління ОР пов'язаний із концептуальними труднощами, зумовленими насамперед складністю їх аналізу, та методологічними труднощами кількісної оцінки їх рівня, зумовленими браком єдиного розуміння типологічних особливостей ОР. Тож розвиток та уточнення теоретико-методологічних основ і розроблення методик управління ОР є актуальними [3], [4].

Відповідно до загального принципу стандартів якості управління (процесного підходу), будь-який вид ризику, зокрема й ОР, можна представити у вигляді причинно-наслідкової моделі такого типу [4], [5]:

джерело → ризикова подія → наслідки.

Джерелами ОР може бути як внутрішнє середовище (системи, персонал, методичні помилки в бізнес-процесах), так і зовнішнє; подією ризику є порушення процедур реалізації бізнес-процесів; наслідками – операційні втрати.

Операційні ризики важко оцінити кількісно, оскільки певні події можуть бути нечастими та малодоступними для спостереження, але водночас дуже небезпечними з погляду втрат [4], [5]. Для моделювання таких ситуацій розроблено методи аналізу розподілів екстремальних значень, проте їх надмірна чутливість до конкретних вибірок даних є проблемою за недостатньої кількості інформації про втрати [6]. В такому разі дані про втрати слід доповнити інформацією з інших джерел, а статистичні методи мають бути придатними для ефективного поєднання даних із різномірних джерел.

Для моделювання ОР часто використовують анкети експертних оцінок ризику, бо вони можуть містити апріорні експертні судження, які можна застосувати для точнішого оцінювання ОР у баєсовому аналізі [7]. Такі анкети розробляють, зазвичай щороку, у відділах внутрішнього аудиту більшості фінансових установ. У базовому варіанті експерти оцінюють якість систем внутрішнього та зовнішнього контролю організації з огляду на власний досвід у певний часовий інтервал.

У складнішій процедурі експерти висловлюють прогнози частоти та величини втрат на основі подій, пов'язаних з ОР. Якісне обґрунтування таких оцінок передбачає безпосередню участь аналітиків ризику в інтерв'ю, де вони пояснюють, які альтернативні прогнози щодо вимірювання ризику мають на увазі. Подавши ці дані як відповідні апріорні розподіли та до-

повнивши їх даними спостережень про збитки, можна на основі побудованого апостеріорного розподілу отримати відносно точні показники ОР, наприклад, Value at Risk (VaR) [6]–[9].

Аналіз ОР – важлива галузь застосування системного аналізу. Метою проведення аналізу є побудова моделі розподілу втрат для розрахунку їх очікуваного значення. Баєсові моделі, що можуть бути адекватним модельно-вимірювальним інструментарієм для створення та використання баєсових методик аналізу ОР, включають статичні, динамічні, об'єктноорієнтовані баєсові мережі, а також діаграми впливів [7]–[9]. Метою моделювання операційних ризиків є побудова моделі, що відображає показники поточної діяльності, використовуючи міру VaR із певним рівнем довіри, що зазвичай становить 95 або 99 %.

Постановка задачі

Метою дослідження є: розглянути особливості певних методів моделювання ОР із використанням статистичних/експериментальних даних, які забезпечують обчислення оптимальних оцінок станів досліджуваних елементів і короткострокове прогнозування їхнього розвитку. Проаналізувати та навести приклади програмної реалізації баєсового аналізу даних, призначеної для побудови ймовірнісних моделей у формі баєсових мереж.

Методи моделювання операційних ризиків

Метод аналізу ієрархії (Analytic Hierarchy Process, AHP)

Згідно з методом аналізу ієрархії, обробка експертних оцінок, відбувається так:

1) Нехай R – комплексна вигода, яка включає ряд аспектів, наприклад, економічну вигоду R_1 , енергоефективність R_2 , соціальну вигоду R_3 : $R = \{R_1, R_2, R_3\}$. Вага вигоди R дорівнює $W = \{W_1, W_2, W_3\}$. W_1, W_2, W_3 – вагові коефіцієнти R_1, R_2, R_3 . Кожен аспект комплексної вигоди можна розділити на індекси другого рівня. Їхня вага дорівнює $w_1 = \{w_{11}, w_{12}, \dots, w_{1n}\}$, де нижні індекси відносяться до кількості рішень, що приймаються, число яких дорівнює n . Інші індекси можна додати так само, водночас кількість вкладень не має обмежень.

2) Ступінь тяжкості події може бути визначений попарним порівнянням. Для демонстрації важливості кожного показника використовують

так звану шкалу попарних порівнянь, приклад якої подається в табл. 1.

Таблиця 1. Шкала попарних порівнянь

Значення	Пояснення
1	Мало важливо
2	Досвід і судження щодо одного показника трохи переважають над іншим
3	Досвід і судження стосовно одного показника помітно переважають над іншим
4	Один показник суттєво кращий за інший
5	Докази, що віддають перевагу одному показнику перед іншим, мають найбільш можливе підтвердження

3) Матриця суджень будується попарним порівнянням показників. Наприклад, для індексу першого рівня $R = \{R_1, R_2, R_3\}$ матриця суджень має вигляд:

$$Q = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix},$$

де q_{12} – результат порівняння двох показників.

Якщо перший показник важливіший за другий, то значення q_{12} вище 1, в іншому випадку воно менше 1. Очевидно, що $q_{11} = q_{22} = q_{33} = 1$.

4) Для визначення ваги кожного індексу використовується формула:

$$\beta^T = \lambda_{\max} W$$

де λ_{\max} – найбільше власне значення матриці Q , W – вектор ознак λ_{\max} ; β^T – вектор вагових коефіцієнтів.

5) Для перевірки узгодженості обчислюється індекс узгодженості CR :

$$CR = \frac{CI}{RI},$$

де

$$CI = \frac{\lambda_{\max} - n}{n - 1},$$

а значення RI змінюється залежно від порядку матриці суджень Q , як показано в табл. 2.

Таблиця 2. Значення RI

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0,58	0,9	1,12	1,24	1,32	1,41	1,46	1,49

Коли $CR = 0$, узгодженість матриці суджень повністю стохастична.

Коли $CR > 0,1$, узгодженість матриці суджень не відповідає вимогам узгодженості, матрицю суджень слід додатково модифікувати.

Коли $CR < 0,1$, вимоги до узгодженості виконані.

Гібридні динамічні баєсові мережі (Hybrid Dynamic Bayesian Networks, HDBNs)

Гібридні динамічні баєсові мережі застосовуються для моделювання причинних ефектів і зв'язків між елементами управління у разі відмов [10], [11]. Ця модель апроксимує безперервний розподіл втрат й агрегує їх за всіма типам втрат. Узагальнена модель HDBNs ОР складається з трьох шарів:

- модель події, пов'язаної з втратами;
- модель тяжкості втрати;
- агрегована модель збитків.

Кожен шар, залежно від обставин, можна представити різними баєсовими, динамічними баєсовими та гібридними динамічними баєсовими мережами, з інтерфейсними зв'язками між ними, що містять загальні параметри. Використання гібридних баєсових мереж також було запропоновано для моделювання залежностей між класами ОР.

Модель події, пов'язаної з втратами. У цій моделі дані про втрати ОР описують за допомогою згортки двох розподілів: Пуассона, що моделює частоту втрат, і Парето, що моделює їх величину [11], [12]. Априорна модель подій втрат показує потенційні події втрат і те, як вони динамічно розвиваються в часі під впливом елементів управління бізнес-процесу. Продуктивність кожного елемента управління моделюється залежно від набору режимів відмов у роботі. На ці режими впливають причинно-наслідкові фактори, що ініціюють операційну відмову. Залежність між режимами експлуатаційних відмов для різних органів управління моделюють за допомогою коефіцієнтів залежності. Ці коефіцієнти обумовлені певним робочим режимом відмови, специфічним для управління, та впливають на ймовірність виникнення режиму відмови у вторинному управлінні. Режим експлуатаційної відмови може бути змодельований як функція причинних факторів і факторів залежності. Експлуатаційні відмови, що не залежать від інших режимів відмов, моделюються як функція причинно-наслідкових факторів. Далі бізнес-процес представляється послідовністю

дискретних залежних від часу подій, тож маємо динамічну баєсову мережу, що є графічним представленням повного спільного розподілу.

Модель тяжкості втрати. В цій моделі для прогнозування загальних втрат за класами серйозності використовують ймовірності, що генеруються моделлю подій втрат, з урахуванням розподілу тяжкості та міри обсягу для масштабування втрат [11], [12]. Припускається, що це підмножина станів, які спостерігаються, зокрема тих, які призводять до втрат або ж є небезпечними. Для кожного дискретного кроку може бути більше однієї події втрат. Завдяки цьому можна моделювати загальні втрати за допомогою моделі умовної залежності, поданої як динамічна баєсова мережа, де загальна кількість подій, пов'язаних із втратами в межах кожного часового кроку, має біноміальний розподіл. При цьому враховують величину обсягу операцій, наприклад, загальну кількість угод, що укладаються щотижня, щомісяця, щороку тощо. До того ж припускається, що наслідки втрат, які виявляються пізніше, серйозніші, оскільки зазвичай призводять до великих очікуваних втрат або мають більшу невизначеність.

Агрегована модель збитків. Маючи набір змінних загальних втрат, для кожної події втрат можна обчислити загальні агреговані втрати як суму загальних втрат, пов'язаних із кожною подією в кожному проміжку часу. Ця агрегована сума є детермінованою функцією, яку можна розрахувати згорткою, семплінгом, або ж динамічною дискретизацією. За допомогою маржиналізації отримують загальний агрегований розподіл збитків, а на його основі розраховують VaR [13].

Підхід на основі оцінювання розподілу втрат (Loss Distribution Approach, LDA)

Однією зі складностей оцінювання ризиків є те, що як кількість подій, які призводять до певних втрат, так і розмір цих втрат є випадковими величинами. Тож побудова моделі ОР із використанням підходу розподілу втрат зводиться до [12], [15]:

- побудови моделі частоти випадків появи подій, що призвели до втрат (спрацювання ризиків) – quality model;
- побудови моделі оцінювання впливу факторів ризиків, що спрацювали, залежно від типу подій, які відбулися чи можуть відбутися – модель важливості факторів впливу;
- поєднання двох побудованих моделей для моделювання ОР загалом.

Побудова моделі частоти випадків появи подій, що призвели до втрат. Для розподілу частоти випадків появи подій приймається розподіл Пуассона з параметром α :

$$N \sim \text{Poisson}(\alpha); E(N) = \alpha; \quad \text{var}(N) = \alpha.$$

Слід зазначити, що частотна характеристика випадків подій є дискретною величиною, тоді як вплив описується неперервною величиною:

$$L = e^{\mu + \sigma Z}; Z \sim N(0, 1); \mu = \ln\left(\frac{E(L)}{\sqrt{1 + w}}\right);$$

$$\sigma^2 = \ln(1 + w); w = \frac{\text{var}(L)}{E(L)^2}.$$

де L – експонента, що характеризує розподіл; μ , σ – параметри розподілу; Z – позначення розподілу; N – нормальний; w – ваговий коефіцієнт для дисперсії; $\text{var}(L)$ – дисперсія L ; $E(L)^2$ – квадрат математичного сподівання для L .

Модель обробки даних із використанням анкетування й експертних оцінок

У цій моделі ОР дані про втрати описують за допомогою згортки двох розподілів: розподілу Пуассона, який моделює частоту втрат, і розподілу Парето, який моделює величину втрат [12], [13]. Апріорні розподіли можуть бути отримані для параметрів обох розподілів на основі простого для інтерпретації моменту, що збігається з частотними розподілами, отриманими шляхом експертного оцінювання. За допомогою теореми Баєса можна оновити ці апріорні розподіли даними про спостережувані втрати та визначити апостеріорний розподіл для параметрів розподілу частоти та впливу, а потім, шляхом згортки, і прогнозний розподіл втрат. Так, на основі оціненого прогностичного розподілу втрат можна розрахувати сумарні показники ризику: VaR і ймовірність. У порівнянні з класичною моделлю, застосування баєсової моделі зменшує VaR, а отже знижує й необхідний резервний обсяг капіталу, що є важливою фінансовою перевагою для прийняття рішень. Ще однією перевагою використання баєсового підходу є можливість об'єднання даних про ці втрати, що обов'язково мають бути "ретроспективними", з даними експертних оцінок, які можуть містити апріорні аспекти та прогнози.

Формальна побудова моделі втрат

Загальний обсяг втрат унаслідок появи певної події i визначається за формулою [12]:

$$L_{it} = X_{i1} + \dots + X_{in_i},$$

де $i = 1, \dots, I$ – подія; $t = 1, \dots, T$ – проміжок часу, причому T – кількість доступних періодів; $X_{i1} + \dots + X_{in_i}$ – втрати в кожен момент часу; а n_i – невідома частота появи певної події.

Кожен перетин події i з періодом t утворює матрицю M . Для кожного перетину i для кожного періоду загальна втрата може бути виражена формулою:

$$L_{it} = s_{it} * n_{it},$$

де s_{it} – втрати за цей період.

Припускається, що для кожного перетину i з моментом часу t , розподіл n_{it} не залежить від s_{it} , а втрати $L_{it} = s_{it} * n_{it}$ не залежать одна від одної для кожного $i = 1, \dots, I$ та $t = 1, \dots, T$.

Як і в попередній моделі, припускається, що розподіл ймовірностей $f(n_t | \eta)$ частоти появи подій ($n_t, t = 1, \dots, T$ – для кожної події i) – дискретний (тут η – вектор параметрів розподілу частот), а розподіл $f(s_t | \theta)$, що характеризує вплив величин s_t , ($s_t, t = 1, \dots, T$ – для кожної події i) – неперервний (тут θ – вектор параметрів розподілу впливу). З урахуванням цих припущень маємо [16]:

$$L(s, n | \theta, \eta) = \prod_{t=1}^T f(n_t | \eta) f(s_t | \theta).$$

де n_t – частота появи кожної події; T – момент появи події n_t ; s_t – втрати за період з частотою n_t ; L – інтегровані втрати за вибраний період.

Для розподілу частот втрат використовують розподіл Пуассона з параметром η , який має апіорний гамма-розподіл із параметрами a і b , що можуть бути встановлені порівнянням із розподілом експертних оцінок частот, який відповідає частоті появи ризикової події. Розподіл експертних оцінок частот можна природно інтерпретувати в термінах спрощеного математичного сподівання $E(\eta)$, що відповідає апіорному математичному сподіванню частоти. Якщо взяти гамма-пов'язаний апіорний частотний розподіл, то згідно з теоремою Баєса, розподіл параметра η можна отримати множенням функції правдоподібності на апіорний розподіл і нормуванням результату. Оскільки ймовірність

буде добутком розподілів Пуассона, пов'язаних з апіорним гамма-розподілом, то апостеріорний розподіл частотного параметра знову буде гамма-розподілом:

$$\pi(\eta | n) \sim \Gamma \left(a + \sum_{t=1}^T n_t, b + T \right).$$

де Γ – позначення гамма-розподілу; a, b – параметри розподілу; T – часовий інтервал, на якому досліджується процес; n_t – частота появи кожної події; $\pi(\eta | n)$ – апіорний розподіл для η ; n – загальна кількість подій.

Завдяки урахуванню спостережуваних фактичних даних і використанню теореми Баєса експертні оцінки частоти оновлюються та коригуються. З наведеної вище формули видно, що вплив первинних думок швидко зменшується зі збільшенням загальної частоти даних стосовно втрат $\sum_{t=1}^T n_t$ і кількості періодів T . У зв'язку

з цим також зрозуміло, що a може тлумачитись як апіорна загальна частота, а b – як апіорна тривалість часового інтервалу. Іншими словами, апіорна оцінка може бути кількісно виражена в термінах даних апіорного досвіду в тій же шкалі, що й спостережувані дані, тому її відносну вагу також можна виразити кількісно. Якщо для спрощення встановити $b = 1$, то a можна отримати як апіорне сподівання значення частотного розподілу оцінок експертів: $a = E(\eta)$.

Для моделювання ступеня тяжкості застосовують розподіл Парето, який характеризується параметрами θ та s_m , причому перший задає форму, а другий – пороговий параметр. Кумулятивна функція розподілу Парето (типу I) задається так:

$$F(x) = 1 - \left(\frac{s_m}{x} \right)^\theta, \text{ для } x > s_m, \text{ та } F(0) = 0.$$

де F – кумулятивна функція розподілу Парето; θ і s_m – параметри розподілу; x – значення змінної розподілу.

У загальному випадку пороговий параметр s_m можна обмежити відомим мінімальним порогом, за якого подія не реєструється, коли приносить незначний збиток (припущення, яке часто роблять у фінансових компаніях для цілей звітності та бухгалтерського обліку). Тож залишається θ – параметр форми розподілу Парето. Для інтерпретації цього параметра використовують індекс Джині, що набуває значення від 0

до 1: високе значення індексу Джині (близьке до 1) означає концентрований розподіл втрат, а низьке значення (близьке до 0) означає дуже мінливий розподіл втрат. Завдяки використанню розподілу Парето індекс Джині (G) можна виразити через параметр θ :

$$G = \frac{1}{2\theta - 1}.$$

Апріорним для параметра θ розподілу Парето є гамма-розподіл із гіперпараметрами (c, d). Гіперпараметри $\frac{\theta s_m}{\theta - 1}$ та d можна встановити порівнянням із розподілом величини втрат згідно з експертними оцінками. Однак, на відміну від параметра η , θ не є очікуваним значенням величини втрат, а дорівнює $\frac{\theta s_m}{\theta - 1}$. Так, отримання з розподілу експертних оцінок апіорної оцінки θ є значно складнішою задачею, ніж отримання з розподілу частоти самооцінки апіорної інформації про η . Для розв'язання цієї проблеми використовують теорема Баєса.

Якщо прийняти апіорний гамма-розподіл для розподілу ступеня тяжкості, то, згідно з теоремою Баєса, апостеріорний розподіл параметра θ також є гамма-розподілом із параметрами [15], [16]:

$$\pi(\theta | s) \sim \Gamma \left(c + T, d + \sum_{t=1}^T \ln \left(\frac{s_t}{s_m} \right) \right).$$

Завдяки теоремі Баєса експертні оцінки параметра θ оновлюються і коригуються спостережуваними фактичними даними.

З наведеної вище формули зрозуміло, що так само як і для частотного параметра, важливість експертних оцінок зменшується зі збільшенням числа періодів T , а сума логарифмованих впливів втрат $\sum_{t=1}^T \ln \left(\frac{s_t}{s_m} \right)$ збільшується.

Це допомагає отримувати апіорні ймовірності з анкет, оскільки гіперпараметр c може бути інтерпретований (аналогічно b) як апіорна тривалість періоду та (аналогічно a) як апіорна сума логарифмічних впливів відносних втрат. Слід зазначити, що роль двох гіперпараметрів змінюється на протилежну, оскільки d , що відповідає зведеним даним про спостережувані втрати

$\sum_{t=1}^T \ln \left(\frac{s_t}{s_m} \right)$, з'являється в знаменнику, а не в чи-

сельнику апіорного математичного сподівання:

$$E(\theta) = \frac{c}{d}.$$

Щоб полегшити розв'язання цієї задачі, запропоновано взяти $c = 1$, так що апіорне математичне сподівання $E(\theta)$, отримане з розподілу ступеня тяжкості втрат з анкет, може бути використано для встановлення однозначного значення

$d = \frac{1}{E(\theta)}$. Так само як попереднє очіку-

ване значення $E(\eta)$, отримане з розподілу частот під час опитування, може бути використано для однозначного задавання $a = E(\eta)$.

Для обчислення математичного сподівання $E(\theta)$ кожна шкала розподілу опитувань може бути виражена в термінах конкретного порядку величин за відношенням до порогового s_m : наприклад, “низькі”, “середні” та “високі” значення ступеня тяжкості можуть бути віднесені до $s_m e^1$, $s_m e^2$, $s_m e^3$, тобто це альтернативи, що їх експерти мали обирати під час опитувань.

$E(\theta) = 1p_1 + 2p_2 + 3p_3$, де p_i – спостережувані частоти самооцінки трьох пунктів шкали ($i = 1, 2, 3$). Коефіцієнт Джині записують так:

$$G = \frac{E(\theta)}{E(\theta) - 2} = \frac{d}{2 - d}.$$

Позначимо через y частоту втрат у майбутньому інтервалі з функцією щільності $f(y | \eta)$. Прогностичну щільність y , враховуючи спостережувані дані $n = (n_t, t = 1, \dots, T)$, можна знайти за виразом:

$$f(y | n) = \int_H f(y | \eta) \pi(\eta | n) d\eta.$$

Позначимо через z ступені тяжкості майбутніх втрат із функцією щільності $f(z | s)$. Прогностична щільність z , із урахуванням спостережуваних даних, є такою:

$$f(y | n) = \int_{\Theta} f(z | \theta) \pi(\theta | s) d\theta.$$

Ці два прогностичні розподіли можна використати для оцінювання прогностичного розподілу втрат, застосувавши процес згортки ОР в повну баєсову парадигму. Фактичний механізм згортки на основі імітаційного моделювання можна звести до таких кроків:

1) для кожного прогнозованого періоду генерують n випадкових спостережень із прогнозного частотного розподілу;

2) для тих самих періодів генерують величину витрат з прогнозованого розподілу впливу, що відповідає частотному спостереженню, зробленому на першому кроці (тобто, якщо моделювана частота подій для періоду k дорівнює n_k , то моделюють вплив із прогнозного розподілу);

3) для кожного періоду розраховуються збитки, отримані на другому етапі; отримуючи спостереження за збитками за період із граничного розподілу, отримують одне спостереження за збитками за кожен період;

4) використовуючи спостереження за збитками, отримані на третьому кроці, оцінюють прогнозний розподіл збитків і отримують VaR-зведені показники ризику, що дають змогу встановити, наскільки великий ризик капіталу.

Математична модель операційного ризику у формі басової мережі

Розглянемо застосування цієї моделі для моделювання шахрайських вимог — операційної події, загальної для страхових компаній. У цьому прикладі моделюємо вартість шахрайських вимог, що виникають протягом тижня, з портфеля комерційного страхування від пожежі з річною валовою премією в 1 млн грн. Шахрайські вимоги можуть включати випадки підпалу майна самими застрахованими чи штучно завищені суми претензій.

Основна гіпотеза полягає в тому, що рівень контролю під час андеррайтингу нових полісів разом зі стадією економічного циклу відповідає за пояснення випадків шахрайства. Рівень контролю в андеррайтингу визначатиме кількість страхувальників із вищою схильністю подавати шахрайські вимоги, прийняті на бухгалтерський облік страховика. Під час економічного спаду ці страхувальники можуть вдатися до підпалів або штучного завищення суми вимог для отримання економічної вигоди. Політика, що пропонує відшкодування втрачених доходів, як і багато інших комерційних політик пожежної безпеки, особливо вразлива до таких зловживань.

Рівень контролю за андеррайтингу залежить від таких факторів:

1) досвід андеррайтера: старші андеррайтери краще оцінюють, які пропозиції ймовірніше призведуть до шахрайського позову;

2) розмір бізнесу: великий розмір бізнесу тисне на всю систему страхування та знижує якість андеррайтингу;

3) опора на філію: андеррайтинг для менш складних ризиків може бути частково переданий

на аутсорсинг персоналу філії для зниження витрат і підвищення ефективності, однак це також збільшує ризик шахрайства через менший контроль якості андеррайтингу на рівні філії.

Очевидно, що відділ обробки претензій зобов'язаний виявляти шахрайські претензії. Ймовірність виявлення залежить від рівня контролю претензій. Якщо відділ дійсно виявляє шахрайство, то можна вжити заходів щодо зменшення збитків, почавши розслідування задля виявлення доказів, які можуть бути використані для відмови чи зменшення платежу. Ці заходи є дорогими та не завжди можуть забезпечити усунення шахрайських виплат за претензіями.

Рівень контролю в управлінні претензіями залежить від таких факторів:

1) досвід оцінювача вимог: досвідченіший оцінювач вимог здатний виявляти шахрайські вимоги з більшою ймовірністю;

2) користування послугами оцінювачів збитків, котрі спеціалізуються на оцінюванні суми збитків, спричинених пожежею: це допомагає зменшити штучно завищені позови; також оцінювачі можуть встановити причини пожежі, особливо якщо можливий підпал;

3) вибіркові перевірки: компанія може проводити вибіркові перевірки претензій для їх детальнішого оцінювання до затвердження платежу.

Нарешті, вартість страховки залежатиме як від рівня розповсюдженості шахрайства, так і від ступеня його виявлення. Навіть у випадку виявлення всіх випадків шахрайства неможливо повністю позбутися витрат через неможливість спростувати всі підозрілі твердження і необхідність витрат на розслідування шахрайських вимог. Випадкові змінні, використані в моделі, наведено в табл. 3.

Вузол “Ціна шахрайства” має задані значення: 0, 50 000, 100 000 або 150 000 грн; але це не означає, що розмір збитків може бути тільки таким. Ці значення встановлені тому, що потім потрібно застосовувати лінійну інтерполяцію для підрахунку величини Capital at Risk (CaR) для різних рівнів довіри. Значення вузла “Ціна шахрайства” потрібно розуміти як “втраати не перевищують 50 000 грн”, “втраати близькі до нуля” тощо.

Після побудови загальної структури басової мережі потрібно задати ймовірнісні розподіли для вузлів змінних. Усі ймовірнісні розподіли є результатом експертного оцінювання, бо статистичних даних стосовно операційних витрат у страхових компаніях немає у відкритому доступі.

Таблиця 3. Вершини баєсової мережі та їх можливі значення

Назва вершини	Можливі значення	Назва вершини у мережі
Досвід андеррайтера	Senior, Junior	Underwriter experience
Опора на філію	Yes, No	Branch reliance
Розмір бізнесу	High, Low	Business volume
Досвід оцінювача вимог	Senior, Junior	Claims assessor experience
Вибіркові перевірки	Yes, No	Random checks
Користування послугами оцінювачів збитків	Yes, No	Engage loss adjuster
Андеррайтинговий контроль	High, Low	Underwriting control
Контроль вимог	High, Low	Claims control
Економічний цикл	Up, Down	Economic cycle
Шахрайська вимога	Yes, No	Fraudulent claim
Виявлення шахрайства	Yes, No	Fraud detected
Ціна шахрайства	0, 50 000, 100 000, 150 000 (грн)	Cost of fraud

Наведемо таблиці умовних ймовірностей для вузлів баєсової мережі: “Досвід андеррайтера”, “Опора на філію”, “Розмір бізнесу”, “Досвід оцінювача вимог” (табл. 4), “Вибіркові перевірки”,

“Користування послугами оцінювачів збитків” (табл. 5).

Таблиця 4. Умовні розподіли вузлів (змінних)

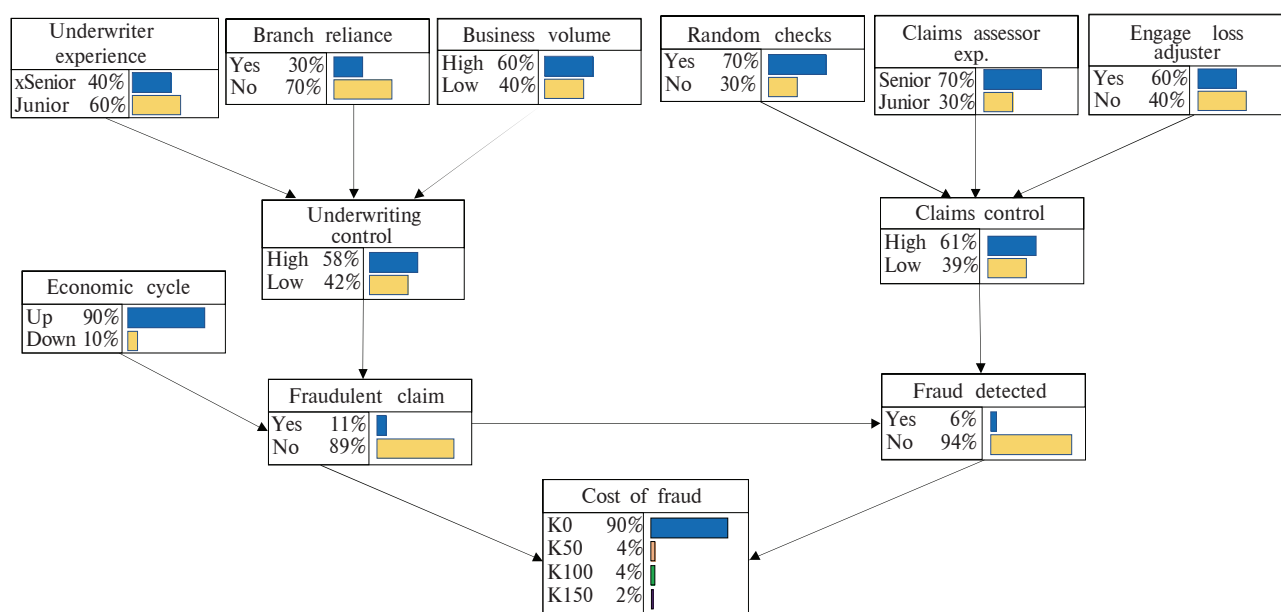
Назва вузла (змінної)							
Досвід андеррайтера		Опора на філію		Розмір бізнесу		Досвід оцінювача вимог	
Senior	Junior	Yes	No	High	Low	Senior	Junior
0,4	0,6	0,3	0,7	0,6	0,4	0,7	0,3

Таблиця 5. Ймовірності для вузлів: вибіркові перевірки і користування послугами оцінювачів

Назва вузла (змінної)			
Вибіркові перевірки		Користування послугами оцінювачів збитків	
Yes	No	Yes	No
0,7	0,3	0,6	0,4

Будуємо структуру моделі (рис. 1) у вигляді баєсової мережі, яка може бути використана для стрес-тестування, причинно-наслідкового моделювання та розподілу капіталу. Для побудови моделі використано програму GeNie – графічний інтерфейс користувача (GUI) для “SMILE Engine”, що дає змогу створювати та вивчати інтерактивні моделі.

Використовуючи GeNie розраховано умовні (маргіальні) ймовірності (рис. 2).

**Рис. 1.** Маргіальні значення вузлів баєсової мережі

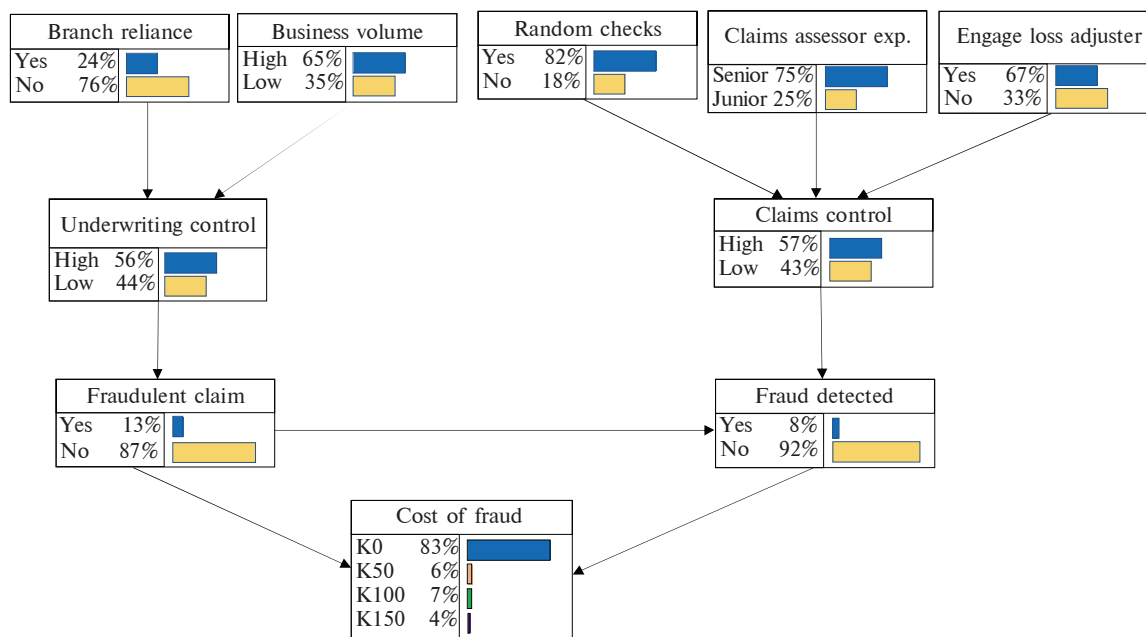


Рис. 2. Умовні ймовірності вузлів баєсової мережі

Одне з основних застосувань баєсових мереж – це моделювання різних сценаріїв подій. Створивши модель, страхова компанія може ухвалити рішення про розмір капіталу, який слід спрямувати на захист компанії від усіх випадків шахрайських вимог, окрім екстремальних. На цьому етапі модель може бути налаштована так, щоб відображати фактично відомий стан певних змінних, щоб показати фактичний стан компанії. Змоделюємо варіант ситуації, коли страхова компанія не має фактичних даних і може спиратися лише на експертні оцінки.

Розрахуємо капітал для покриття збитків, спричинених ОР CaR, за різних довірчих інтервалів: 90, 95 і 99 %. Такі рівні довіри є стандартними для розрахунку CaR у фінансових установах (табл. 6).

Таблиця 6. Капітал на покриття ОР із заданими рівнями довіри

Процентиль, %	90	95	99
CaR, грн	56 818	64 286	125 394

Логічно, що витрати на капітал (CaR) зростають разом із рівнем довіри. За “Базелем II”, фінансові установи, що використовують метод АМА (advanced management approach), повинні бути спроможними покрити витрати з 99,9 % рівнем довіри. Але фінансові установи, зокрема й СК, розраховують капітал для 95 % рівня довіри. За 95 % рівня довіри CaR дорівнює 64 286 грн. При використанні методу ВІА (basic indicator approach)

із річним валовим доходом в 1 млн дол. треба виділити 150 тис. грн для покриття ОР, що майже у 2,5 раза більше, ніж за АМА.

Сценарний аналіз

Баєсова мережа може бути використана для тестування різних сценаріїв, щоб допомогти менеджеру оптимізувати профіль ризику. Наприклад, якщо страхова компанія бажає скоротити витрати зниженням рівня кваліфікованості андеррайтера до “Низький”, то ефект цієї дії можна легко дослідити, встановивши свідомо Underwriter experience = “Низький” у мережу. Потім інформація поширюється мережею, й отримуємо нові умовні розподіли. Розглянемо один із можливих сценаріїв у страховій компанії. У таблиці 7 показано, які значення приймають деякі вузли.

Таблиця 7. Апостеріорні значення для першого сценарію

Змінні	Значення
Underwriter experience	Junior
Branch reliance	Yes
Business volume	High
Random checks	No
Claims assessor experience	Junior
Engage loss adjuster	No
Economic cycle	Down

Отримані результати зображено на рис. 3.

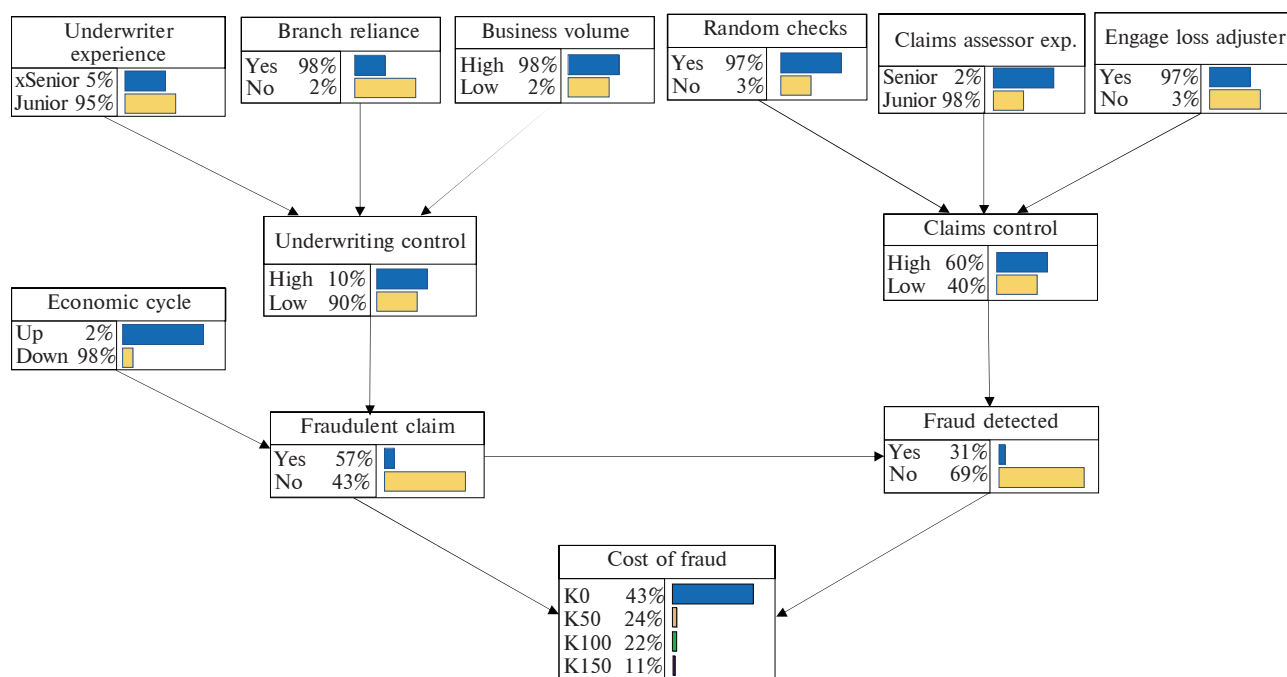


Рис. 3. Умовні розподіли ймовірностей за сценарієм

Розрахуємо капітал для покриття збитків, спричинених ОР CaR, за різних довірчих інтервалів: 90, 95 і 99 % (табл. 8).

Таблиця 8. Капітал на покриття операційного ризику із заданими рівнями довіри за сценарієм

Процентиль, %	90	95	99
CaR, грн	94 374	119 243	143 849

Можна зробити висновок, що компроміс між нижчими витратами – це збільшення потреби в капіталі на 55 тис. грн, водночас припускаємо, що в СК встановлений рівень довіри 95 %.

Висновки

Розглянуто причини виникнення ОР, який є надзвичайно притаманним промисловості,

фінансовим послугам, транспортуванню тощо. Різноманітність причин виникнення такого ризику вимагає комплексного підходу до його оцінювання та менеджменту. В роботі подано короткий огляд баєсових моделей аналізу, моделювання й оцінювання ОР. Показано, що ймовірнісні баєсові моделі працюють задовільно та забезпечують оцінювання необхідного ризикового капіталу (Capital at Risk). Побудована модель ОР шахрайства у сфері страхування ілюструє можливе практичне застосування розглянутої методики моделювання.

У подальших дослідженнях необхідно узагальнити розглянуту методику аналізу на ризики інших типів і створити спеціалізовану систему підтримки прийняття рішень для моделювання та прогнозування ризиків втрат із прийнятною точністю.

References

- [1] Basel Committee on Banking Supervision. *International Convergence of Capital Measurement and Capital Standards. A Revised Framework. Comprehensive Version*, Basel, Switzerland: Bank for International Settlements, 2006.
- [2] M. G. Cruz, *Modeling, Measuring and Hedging Operational Risk*, London: Wiley, 2002.
- [3] C. Alexander, *Operational Risk: Regulation, Analysis and Management*, Harlow, England: Pearson Education Ltd, 2003.
- [4] P. V. Shevchenko, *Modeling Operational Risk Using Bayesian Inference*, New York: Springer, 2011.
- [5] S. O. Dmytrov, *Modeling of operational risk assessment of a commercial bank*, Sumy, Ukraine: SHEI "Ukrainian Academy of Banking", 2010.

- [6] A. Frachot, P. Georges, and T. Roncalli, "Loss distribution approach for operational risk," *SSRN Electron. J.*, pp. 56–99, 2001. doi: 10.2139/ssrn.1032523.
- [7] C. Alexander, "Bayesian methods for measuring operational risk," *SSRN Electron. J.*, pp. 45–65, 2000. doi: 10.2139/ssrn.248148.
- [8] X. Hao, "Operational risk control of commercial banks based on Bayesian network," *Proc. 2013 ICETIS*, 2013, pp. 913–918. doi: 10.2991/icetis-13.2013.209.
- [9] Yongqi Yang, Wanlei Xue, Chaoyuan Li, Xin Zhao, and Nan Xu, "An AHP-Bayesian model for operational risk evaluation of integrated energy systems," *IOP Conf. Ser. Materials Sci. Eng.*, vol. 484, pp. 1–8, 2019. doi: 10.1088/1757-899X/484/1/012052
- [10] Y. K. Yoon, "Modeling operational risk in financial institutions using Bayesian networks", *J. Risk Insur.*, UK, №74(4), pp. 795–827, 2007.
- [11] M. Neil, N. Fenton, and M. Taylor, "Using Bayesian networks to model expected and unexpected operational losses," *Risk Analysis*, vol. 25, no. 4, pp. 963–972, Aug. 2005. doi: 10.1111/j.1539-6924.2005.00641.x.
- [12] A. Stuart and K. Ord, *Kendalls Advanced Theory of Statistics*, vol. 1, Wiley, 1994.
- [13] S. Figini, L. Gao and P. Giudici, "Bayesian operational risk models," *J. Oper. Risk*, vol. 10, no. 2, pp. 45–60, 2015. doi: 10.21314/JOP.2015.155
- [14] C. K. Chow and C. N. Liu, "Approximating discrete probability distributions with dependence trees," *IEE Trans. Inf. Theory*, vol. 14, pp. 462–467, 1968.
- [15] T. Minka *et al.* (2014). *Infer.NET 2.6 Microsoft Research Cambridge* [Online]. Available: <http://research.microsoft.com/infernet>
- [16] M. Neil, D. Hager and L. B. Andersen, "Modeling operational risk in financial institutions using hybrid," *J. Oper. Risk*, vol. 4, no. 1, pp. 1–28, 2009.

Л.Б. Левенчук, В.Г. Гуськова, П.И. Бидюк

ВЕРОЯТНОСТНОЕ МОДЕЛИРОВАНИЕ ОПЕРАЦИОННЫХ РИСКОВ

Проблематика. Операционные риски присущи всем видам деятельности человека, включая промышленное производство, транспорт, финансовые услуги и т.п. Риски такого типа характеризуются множеством неопределенностей, неполнотой и низким качеством данных, что усложняет прогнозирование и оценивание. Для моделирования процессов, связанных с операционными рисками, необходимо выполнить корректную обработку данных, а также идентификацию и учет возможных неопределенностей. Для решения этих задач хорошо подходит вероятностный подход.

Цель исследования. Сделать краткий обзор методов вероятностного анализа данных, предназначенных для построения математических моделей операционных рисков. Разработать новую вероятностную модель в форме байесовской сети для формального описания операционного риска мошенничества, связанного с актуарными процессами.

Методика реализации. Для обработки данных и экспертных оценок используются методы байесовского анализа данных, предназначенные для построения вероятностных моделей в форме байесовских сетей.

Результаты исследования. Предложенный метод моделирования использован для построения модели операционного риска мошенничества, характерного для сферы страхования. Для построения модели выполнен анализ задачи, выбрано множество переменных и выполнено оценивание априорных условных вероятностей. Результирующая модель построена с помощью инструментария моделирования GeNIe. Функционирование модели продемонстрировано на примере.

Выводы. Показано, что важная практическая задача моделирования, оценивания и прогнозирования операционных рисков может быть решена с помощью вероятностного подхода, в частности, байесовской методологии, которая дает возможность идентифицировать и учесть имеющиеся неопределенности данных и экспертных оценок. Построенная с помощью этой методологии модель операционного риска демонстрирует возможность применения байесовского подхода для решения упомянутых задач.

Ключевые слова: операционные риски; вероятностное моделирование; байесовская методология; экспертные оценки; актуарные процессы.

L.B. Levenchuk, V.H. Huskova, P.I. Bidyuk

PROBABILISTIC MODELLING OF OPERATIONAL RISKS

Background. Operational risks are inherent to all types of human activity, including industrial production, transport, financial services, etc. Risks of this kind are characterized by many uncertainties, incompleteness and low data quality, that complicate prediction and assessment. To perform modelling of the process associated with operational risk it is necessary to carry out a proper data processing as well as identification, taking into consideration possible uncertainties. The probabilistic approach to modelling is very helpful in solving the problems.

Objective. The purpose of the paper is to make brief overview of probabilistic data analysis methods designed to build mathematical models of operational risks. To develop a new probabilistic model in the form of a Bayesian network to describe formally the operational risk of fraud associated with actuarial processes.

Methods. The basic methodology used for data and expert estimates processing are Bayesian data analysis techniques that help to construct probabilistic models in the form of Bayesian networks.

Results. The proposed modelling method was applied to constructing model of operational risk, more specifically risk of fraud in actuarial sphere. To construct the model the problem was analysed, a set of variables was selected, and prior estimates for conditional

probabilities were estimated. The final model was constructed using the modelling system GeNIe. The model functioning was demonstrated using illustrative example.

Conclusions. It was shown that modelling, estimating and forecasting financial and other types of risks is important practical problem that can be solved using probabilistic approach, namely Bayesian methodology that helps to identify and take into consideration possible uncertainties of data and expert estimates. The operational risk model constructed using the methodology illustrates the possibilities of application the Bayesian techniques to solving the problems mentioned.

Keywords: operational risks; probabilistic modelling; Bayesian methodology; expert estimates; actuarial processes.

Рекомендована Радою
Інституту прикладного системного аналізу
КПІ ім. Ігоря Сікорського

Надійшла до редакції
23 липня 2021 року

Прийнята до публікації
7 грудня 2021 року