

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерної інженерії та управління

Кафедра Комп'ютерних інтелектуальних технологій та систем

КВАЛІФІКАЦІЙНА РОБОТА

Пояснювальна записка

рівень вищої освіти

другий (магістерський)

Аналіз і прогнозування актуарних процесів
з використанням інтелектуального аналізу даних

Виконав:

студент 2 курсу, групи КІТм-22-1

Владислав ЦВЕК

Спеціальність 123 Комп'ютерна інженерія

Тип програми освітньо-професійна

Освітня програма Комп'ютерні інтелектуальні
технології

Керівник професор Микола КОРАБЛЬОВ

Допускається до захисту

Зав. кафедри

Олег РУДЕНКО

(підпис)

2023 р.

Харківський національний університет радіоелектроніки

Факультет	Комп'ютерної інженерії та управління
Кафедра	Комп'ютерних інтелектуальних технологій та систем
Рівень вищої освіти	другий (магістерський)
Спеціальність	123 Комп'ютерна інженерія
Тип програми	освітньо-професійна
Освітня програма	Комп'ютерні інтелектуальні технології

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« _____ » _____ 2024 р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові _____ Цвеку Владиславу Сергійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Аналіз і прогнозування актуарних процесів з використанням інтелектуального аналізу даних

затверджена наказом по університету від “ 03 ” листопада 2023 р. № 1290Ст

2. Термін подання студентом роботи до екзаменаційної комісії 17.01.2024

3. Вхідні дані до роботи _____

- 1) методи оцінювання фінансових ризиків
- 2) методи оцінювання кредитоспроможності осіб
- 3) інструментальні засоби оцінювання фінансових ризиків
- 4) середовище моделювання – Eviews
- 5) мова програмування – R-package

4. Перелік питань, що потрібно опрацювати в роботі _____

- 1) Аналіз предметної області і постановка задачі дослідження
- 2) Аналіз методів і моделей оцінювання фінансових ризиків у страхуванні
- 3) Аналіз типів даних і невизначеностей в актуарному моделюванні
- 4) Побудова математичних моделей для аналізу та прогнозування актуарних процесів
- 5) Визначення структури математичної моделі актуарних процесів
- 6) Прведення експериментальних досліджень

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій Плакати – 21 арк. ф. А4

6. Консультанти розділів роботи и (п.6 включається до завдання за наявності консультантів згідно до наказу, зазначеному у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Видача та узгодження теми проєкту	06.11.2023	Виконано
2	Огляд стану проблеми та постановка задачі	06.11 -17.11	Виконано
3	Аналіз літератури за напрямком кваліфікаційної роботи	18.11 - 01.12	Виконано
4	Аналіз методів і моделей оцінювання фінансових ризиків	02.12 - 09.12	Виконано
5	Побудова математичних моделей для аналізу та прогнозування актуарних процесів	10.12 – .17.12	Виконано
6	Проведення експериментальних досліджень	18.12.23 - 07.01.24	Виконано
7	Підготовка пояснювальної записки	08.01 - 15.01	Виконано
8	Підготовка графічного матеріалу	16.01 – 23.01	Виконано
9	Захист проєкту	24.01.2024	Виконано

Дата видачі завдання 06 листопада 2023 р.

Студент _____
(підпис)

Керівник роботи _____ проф. Корабльов М.М.
(підпис) (посада, ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 90 с., 11 рис., 4 табл., 2 дод., 31 джерело.

СТРАХОВА ДІЯЛЬНІСТЬ, АКТУАРНІ ПРОЦЕСИ, ФІНАНСОВІ РИЗИКИ, МАТЕМАТИЧНА МОДЕЛЬ, АНАЛІЗ, ПРОГНОЗУВАННЯ, ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ, РЕГРЕСІЙНІ МОДЕЛІ, ЕКСПЕРТНЕ ОЦІНЮВАННЯ

Метою кваліфікаційної роботи є побудова математичних моделей аналізу і прогнозування актуарних процесів на основі інтелектуального аналізу даних, використання яких дозволить отримати менші ЗЗЗ ризики у роботі страхових компаній.

Об'єктом дослідження є актуарні процеси у страховій діяльності.

Предметом дослідження є методи і моделі аналізу і прогнозування актуарних процесів з використанням інтелектуального аналізу даних.

Проведено аналіз методів і моделей, придатних для прогнозування та прийняття рішень на їх основі. Встановлено, що для розв'язання задачі аналізу та оцінювання фінансових ризиків можна скористатись регресійними моделями та мережами Байєса. Досліджено процес побудови математичних моделей для актуарних процесів, який складається з виконання відповідних етапів.

Обґрунтовано вибір середовища моделювання та обробки фінансових статистичних даних з використанням економетричного пакету Eviews та середовища програмування R-package. Побудована модель має широкий спектр застосування, а саме: для фінансового аналізу з метою вчасного запобігання настанню банкрутства страхових компаній, а також для аналізу ефективності виконання умов договорів та виплати страхових платежів.

ABSTRACT

Master's thesis:: 90 pages, 11 figures, 4 tables, 2 appendices, 31 sources.

INSURANCE ACTIVITY, ACTUARIAL PROCESSES, FINANCIAL RISKS, MATHEMATICAL MODELING, ANALYSIS, FORECASTING, INTELLECTUAL DATA ANALYSIS, REGRESSION MODELS, EXPERT VALUATION

The purpose of the qualification work is to build mathematical models of analysis and forecasting of actuarial processes based on intellectual data analysis, the use of which will allow to obtain lower risks in the work of insurance companies.

The object of research is actuarial processes in insurance activity.

The subject of research is methods and models of analysis and forecasting of actuarial processes using intelligent data analysis.

An analysis of methods and models suitable for forecasting and decision-making based on them was carried out. It has been established that regression models and Bayesian networks can be used to solve the problem of financial risk analysis and assessment. The process of building mathematical models for actuarial processes, which consists of the implementation of the corresponding stages, has been studied.

The choice of the environment for modeling and processing financial statistical data using the Eviews econometric package and the R-package programming environment is substantiated. The built model has a wide range of applications, namely: for financial analysis with the aim of timely preventing the onset of bankruptcy of insurance companies, as well as for the analysis of the effectiveness of the fulfillment of the terms of contracts and the payment of insurance payments.

АНОТАЦІЯ

Цвек В.С.. Аналіз і прогнозування актуарних процесів з використанням інтелектуального аналізу даних. – Магістерська кваліфікаційна робота.

У магістерській кваліфікаційній роботі розв'язано актуальну задачу побудови адаптивних математичних моделей для прогнозування та оцінювання ризиків актуарних процесів за допомогою узагальнених лінійних моделей; підвищення адекватності моделей актуарних процесів та якості оцінок прогнозів виплат внаслідок настання страхових випадків та операційних втрат страхових компаній.

Метою даної роботи є побудова математичних моделей аналізу і прогнозування актуарних процесів на основі інтелектуального аналізу даних, використання яких дозволить отримати менші ризики у роботі страхових компаній.

Об'єктом дослідження є актуарні процеси у страховій діяльності.

Предметом дослідження є методи і моделі аналізу і прогнозування актуарних процесів з використанням інтелектуального аналізу даних.

Методи дослідження - теорія узагальнених лінійних моделей та екстремальних значень, байєсівські мережі і теорія формування імовірнісного висновку, методи оцінювання невідомих параметрів математичних моделей (Монте-Карло і метод максимальної правдоподібності).

Наукова новизна полягає у побудові та дослідженні узагальненої лінійної моделі для формального опису актуарних процесів, а також побудовою адаптивних моделей актуарних процесів, які відрізняється запропонованим методом моделювання в умовах виродженості вибірки, методом оцінювання екстремальних значень і забезпечують підвищення адекватності моделей та якості оцінок прогнозів.

Практична цінність отриманих результатів полягає у побудові нових узагальнених лінійних моделей для обраних актуарних процесів, які забезпечують оцінювання короткострокових прогнозів страхових виплат високої точності. Відносні похибки оцінок прогнозів є незначними і коливаються в межах 5 – 12% для побудованих моделей.

У першому розділі розглянуто актуальність задачі побудови адаптивних математичних моделей для прогнозування та оцінювання ризиків актуарних процесів та проведено аналіз протікання фінансових процесів у страхуванні. Виконано аналіз методів і моделей оцінювання фінансових ризиків у страхуванні, зокрема, розглянуті моделі, що засновані на аналізі наслідків; моделі, що засновані на аналізі факторів ризику, метод функціональних кореляцій, регресійні моделі, та методи експертного оцінювання.

Проведено також аналіз типів даних і невизначеностей в актуарному моделюванні. Це свідчить про актуальність використання методів інтелектуального аналізу даних для розв'язання задач цього класу. На основі виконаного дослідження класів сучасних математичних моделей, придатних для прогнозування та прийняття рішень на їх основі, встановлено, що для розв'язання задачі аналізу та оцінювання фінансових ризиків можна скористатись універсальними логічними моделями. Моделі такого типу дають можливість враховувати не лише безпосередні залежності рівня ризику від факторів ризику, а й залежності між факторами ризику.

Досліджено типи даних і розподілів в актуарному моделюванні, що включені до узагальненого класу розподілів. Серед відомих типів розподілів виділено такі: нормальний розподіл, розподіл Пуассона, біноміальний розподіл, гамма розподіл. Саме ці розподіли пропонується використовувати у подальшому дослідженні.

У другому розділі були досліджені методика побудови математичних моделей для актуарних процесів довільної природи, базуючись на математичному інструменті узагальнених лінійних моделей, теорії екстремальних значень, проаналізовано описові та прогнозні статистики. Виявлені характеристики ефективності математичних моделей при їх застосуванні до аналізу актуарних процесів, серед яких такі: відображати об'єктивні закономірності процесів; правильно відтворювати функцію та/або структуру реального процесу; відповідати певним математичним умовам.

Виявлено характеристики ефективності математичних моделей при їх застосуванні до аналізу актуарних процесів, серед яких такі: відображати об'єктивні закономірності процесів; правильно відтворювати функцію та/або структуру

реального процесу (до уваги беруться актуарні процеси); відповідати певним математичним умовам (мати розв'язок, узгоджені розмірності тощо).

Результати досліджень будь-якої моделі повинні мати практичну цінність, тобто, якщо модель адекватна процесу або явищу, яке вивчається, то відповідно вона точно відтворює реальну ситуацію. Тому процес побудови математичної моделі складається з таких етапів: формулювання предмету та мети дослідження; виокремлення структурних чи функціональних елементів, що відповідають поставленій меті, та визначення найважливіших якісних характеристик цих елементів; встановлення та опис взаємозв'язків між елементами моделі; виконуються розрахунки за математичною моделлю, оцінюється прогнозне значення та аналізуються отримані результати шляхом оцінювання статистичних характеристик побудованої моделі, визначається її адекватність, виконується візуальне представлення результатів та оцінюється якість прогнозу.

Обґрунтовано вибір середовища моделювання та обробки фінансових статистичних даних з використанням економетричного пакету *Eviews* та середовища програмування *R-package*. Розглянуто їх основні функції стосовно аналізу фінансових даних.

У третьому розділі весь комплекс розроблених методів і алгоритмів використано для виконання експериментальних досліджень. Запропонована методика аналізу та оцінювання актуарних процесів задовольняє таким основним вимогам:

- забезпечує виявлення екстремальних значень в актуальних даних на початковому етапі аналізу статистичної вибірки;
- забезпечує обробку екстремальних значень шляхом наближення до одного з розподілів із класу узагальнених розподілів екстремальних значень;
- мета застосування методики – подолання проблеми виродженості початкової вибірки для подальшої побудови універсальної логічної моделі.

Експериментально доведено ефективність застосування комплексної моделі обробки екстремальних статистичних даних, завдяки якій вдалося успішно розв'язати проблему невиродженості даних у статистичній вибірці із застосуванням

теорії екстремальних значень. Залучення новітніх комбінованих методів до розв'язання задачі обробки екстремальних даних та оцінювання невідомих параметрів, вибору кращої моделі на основі алгоритмів зменшення порогів викидів відкриває нові можливості для дослідження особливостей методів математичного моделювання.

Створена система надає можливість актуаріям працювати з виродженими масивами даних та виконувати фінансовий аналіз страхової діяльності компаній у реальному часі з використанням фактичних даних стосовно страхових виплат за відповідними страховими договорами і страховими випадками.

КЛЮЧОВІ СЛОВА: СТРАХОВА ДІЯЛЬНІСТЬ, АКТУАРНІ ПРОЦЕСИ, ФІНАНСОВІ РИЗИКИ, МАТЕМАТИЧНА МОДЕЛЬ, АНАЛІЗ, ПРОГНОЗУВАННЯ, ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ, РЕГРЕСІЙНІ МОДЕЛІ, ЕКСПЕРТНЕ ОЦІНЮВАННЯ

Публікації здобувача за темою роботи:

1. Цвек В.С. Побудова математичної моделі для аналізу та прогнозування актуарних процесів // Матеріали IV Всеукраїнської студентської наукової конференції «Розвиток сучасної науки: актуальні питання теорії та практики», 17 листопада 2023, м. Львів, Україна.

Використані публікації керівника та співробітників кафедри, що становлять теоретичну базу роботи

1. Mykola Korablyov, Oleksandr Fomichov, Sergey Lutskyu, Ihor Ivanisenko, Danylo Antonov, Stanislav Dykyi. Hybrid stock analysis model for financial market forecasting // 18th International Scientific and Technical Conference on Computer Science and Information Technology, CSIT 2023, October 19-21, 2023, Lviv Ukraine.

ЗМІСТ

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ	13
ВСТУП	14
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	16
1.1 Аналіз протікання фінансових процесів у страхуванні	16
1.2 Аналіз методів і моделей оцінювання фінансових ризиків у страхуванні	19
1.2.1 Моделі, що засновані на аналізі наслідків.....	19
1.2.2 Моделі, що засновані на аналізі факторів ризику.....	22
1.2.3 Метод функціональних кореляцій.....	22
1.2.4 Регресійні моделі.....	24
1.2.5 Методи експертного оцінювання	24
1.3 Аналіз типів даних і невизначеностей в актуарному моделюванні	30
2 ПОБУДОВА МАТЕМАТИЧНИХ МОДЕЛЕЙ ДЛЯ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ АКТУАРНИХ ПРОЦЕСІВ	33
2.1 Аналіз методів і моделей побудови математичних моделей актуарних процесів	33
2.2 Визначення структури математичної моделі актуарних процесів	34
2.2.1 Системний аналіз актуарного процесу	36
2.2.2 Попередня обробка експериментальних даних.....	37
2.2.3 Визначення наявності нелінійностей	38
2.2.4 Обробка екстремальних значень	40
2.2.5 Формування структури моделі	43
2.2.6 Оцінювання параметрів моделей.....	44
2.2.6 Діагностика моделей та вибір кращої з оцінених кандидатів	47
2.3 Обробка екстремальних значень	49
3 ЕКСПЕРЕМЕНТАЛЬНІ ДОСЛІДЖЕННЯ	54
3.1 Обґрунтування та вибір середовища для моделювання	54
3.2 Опис та структура статистичних даних	57
3.3 Попередня обробка даних	58

3.4 Прогнозування актуарних процесів за допомогою побудованих моделей	64
3.5 Результати експериментальних досліджень	65
3.6 Вибір та обґрунтування кращої моделі	65
ВИСНОВКИ	68
ПЕРЕЛІК ДЖЕРЕЛ ПОСЧИЛАННЯ	70
ДОДАТОК	А
	О

шибка! Закладка не определена.

Графічний матеріал	атестаційної	роботи
		О

шибка! Закладка не определена.

ДОДАТОК	Б
	О

шибка! Закладка не определена.

Наукова	публікація
	О

шибка! Закладка не определена.

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

ІАД – інтелектуальний аналіз даних

СК – страхова компанія

СД – страхова діяльність

ВІА – метод базових показників (Basic Indicator Approach)

LDA – метод розподілу збитків (Loss Distribution Approach)

ІМА – метод внутрішніх вимірів (Internal Measurements Approach)

Sb-АМА – сценарний аналіз (Scenario-based Advanced Measurement Approach)

GEV – gev-розподіл (Generalized extreme value)

УЛМ – універсальна логічна модель

БМ – Байєсівська мережа

МГВА – метод групового врахування аргументів

ТЕЗ – теорія екстремальних значень

МНК – метод найменших квадратів

ММП – метод максимальної правдоподібності

МКМЛ – метод Монте-Карло для марковських ланцюгів

ВСТУП

В теперішній час має місце удосконалення технологій пошуку та обробки даних, які використовуються в різних сферах людської діяльності, що призводить до накопичення великих масивів даних різної природи. При цьому діяльність підприємств супроводжується записом і реєстрацією всіх подробиць про їх функціонування, що призводить до утворення великих масивів даних та ускладнення процесів обробки та пошуку відповідної інформації. Критерії обробки таких великих і різнорідних масивів даних мають бути ефективними, простими, конкретними, зручними у використанні та адаптованими під користувачів.

Обробка статистичних і експериментальних даних може виконуватись із застосовувати як класичних статистичних методів, так і сучасних методів інтелектуального аналізу даних (ІАД). Слід зазначити, що методи ІАД виникли та отримали розвиток на основі результатів методів штучного інтелекту, функціонального моделювання, прикладної статистики, розпізнавання образів, баз даних і знань тощо. Задача ІАД полягає у виявленні в отриманих даних тих чи інших знань, які необхідні для обґрунтованого прийняття рішень в тій чи іншій сфері діяльності.

До методів та алгоритмів ІАД можна віднести наступні: символні правила, нейронні мережі, нечітка логіка та нейро-нечіткі системи, дерева рішень, метод опорних векторів, байєсівські мережі та байєсівська регресія, методи ближчого сусіда, кластерний аналіз, методи пошуку асоціативних правил, метод обмеженого перебору, еволюційне програмування, штучні імунні системи, генетичні алгоритми, методи візуалізації даних та ін. Методи ІАД мають такі властивості як багатofункціональність, гнучкість, можливість подолання невизначеностей різної природи та типів, можливість створення комбінованих багатовимірних моделей, яким притаманні характеристики високої якості оцінок прогнозів та управління з їх використанням.

Методи ІАД широко застосовуються для розв'язання актуарних задач, тому

що його інструменти візуалізації дозволяють виконувати аналіз даних особами, які є фахівцями в предметній області, але можуть не володіти необхідними математичними знаннями. Тому ІАД важливий та надзвичайно користний для застосування керівниками та актуаріями у сфері страхування. Це дозволяє розв'язувати такі задачі: оцінювання фінансових ризиків страхових компаній, прогнозування страхових виплат у страхових випадках, виявлення шахрайства та ін.

Відомо, що страхова (актуарна) діяльність орієнтована на перерозподіл грошових коштів, а також на їх акумуляування для страхової діяльності, що дозволяє інвестувати ці кошти у різні галузі діяльності, і таким чином сприяти їх подальшому розвитку. У зв'язку з цим виникають задачі аналізу та менеджменту фінансових ризиків у страхуванні, розв'язання яких можливе з використанням методів інтелектуального аналізу даних, а саме сучасних методів прогнозування, математичного моделювання та інтелектуальної підтримки прийняття рішень.

Сучасний стан фінансів страхових компаній (СК) потребує впровадження нових економіко-математичних моделей для прогнозування і оцінювання їх подальшого розвитку за умов невизначеності, ризику реалізації багатьох процесів, а також великої конкуренції. Широкий спектр форм прояву ризиків, складність та частота наслідків його реалізації потребують поглибленого аналізу можливих ризиків та математичного обґрунтування фінансової політики діяльності СК. Тому задача аналізу і прогнозування актуарних процесів з метою підвищення рівня і фінансової стабільності СК на основі використання інтелектуального аналізу даних є актуальною.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

Міра впливу фінансових ризиків на результати функціонування підприємств різних видів діяльності та рівень їх фінансової стабільності істотно зростає. Зростання впливу фінансових ризиків на результати економічної діяльності суб'єктів господарювання обумовлене розширенням сфери фінансових стосунків, нестабільністю економічної ситуації в країні, зміною кон'юнктури фінансового ринку, наявністю нових інноваційних фінансових інструментів та ін.

Надійність учасників страхового ринку є головною умовою його ефективного функціонування. Підтримка здатності кожного страховика, що діє на ринку, у повному обсязі та своєчасно виконувати взяті на себе обов'язки, є важливим для фактичної реалізації функції страхування. Існуючий стан фінансів страхових компаній (СК) вказує на необхідність підвищення їх фінансової стабільності та конкурентоспроможності.

Широкий спектр форм прояву ризиків, складність та частота наслідків його реалізації потребують поглибленого аналізу можливих ризиків та математичного обґрунтування фінансової політики діяльності СК. Для отримання обґрунтованої та достовірної оцінки основних характеристик фінансової стабільності підприємств потребує використання економіко-математичних методів для аналізу та менеджменту фінансових ризиків.

1.1 Аналіз протікання фінансових процесів у страхуванні

Діяльність СК орієнтована на захист майнових інтересів як фізичних, так і юридичних осіб із грошових фондів, що формуються із страхових внесків, які сплачуються страхувальниками. При цьому фінансовий крах страховика може зачепити інтереси великого кола осіб. Під страхуванням розуміється певний вид господарської діяльності на основі солідарності і відплатності, який має метою покриття можливих потреб та нестатків, які є наслідком випадкової події. Погляди

на страхування та страхову діяльність (СД) є неоднозначними, які розглядаються в подальшому з позицій формування грошових фондів для здійснення страхових виплат.

Аналізуючи зміст основних завдань фінансового менеджменту, можна зробити висновок, що основним об'єктом фінансового управління є фінансові ресурси, а внутрішнім фактором фінансової стійкості СК є правильний вибір тактики та стратегії управління ними. Аналіз фінансової діяльності СК з позицій економічної природи страхування можна розглядати як дослідження потоків грошей, яке, з одного боку, є засобом захисту інтересів майна, а з другого – таким видом діяльності, який приносить дохід.

Фінансовий аналіз СК являє собою аналіз їх фінансового стану, а також оцінку їх надійності та фінансової стійкості, які є найважливішими характеристиками фінансово-економічної діяльності підприємства. Будь-яке підприємство платіжоспроможне та має переваги в отриманні інвестицій, залученні клієнтів і партнерів, якщо воно фінансово стійке. Результати аналізу фінансового стану СК представляють інтерес при ухваленні рішень як для внутрішніх користувачів-страховиків, так і для зовнішніх компаній, які взаємодіють з цими СК.

Сукупність показників фінансового стану й методів їх оцінювання являє собою методика фінансового аналізу підприємства у вузькому розумінні, які розрізняють за способами групування відповідних показників фінансового стану підприємства, що охоплюють показники, за допомогою яких можна оцінити платіжоспроможність, зобов'язання, ліквідність і власний капітал. Для оцінювання фінансового стану страховиків необхідні знання базових понять і показників, які вживаються у фінансовому менеджменті, а також знання специфіки СД. Тому фінансовий аналіз СК повинен орієнтуватися на загальні та специфічні показники оцінки фінансового стану, що характерні для СД. [1, 2].

Власний та залучений капітал являються основними частинами фінансових ресурсів СК, причому залучена частина у вигляді страхових премій є набагато більшою, ніж власна. Залучені засоби тимчасово знаходяться в розпорядженні СК і застосовуються на страхові виплати або стають прибутковою базою [2, 3].

На рисунку 1.1 показана схема кругообігу основних фінансових процесів у страховій діяльності.

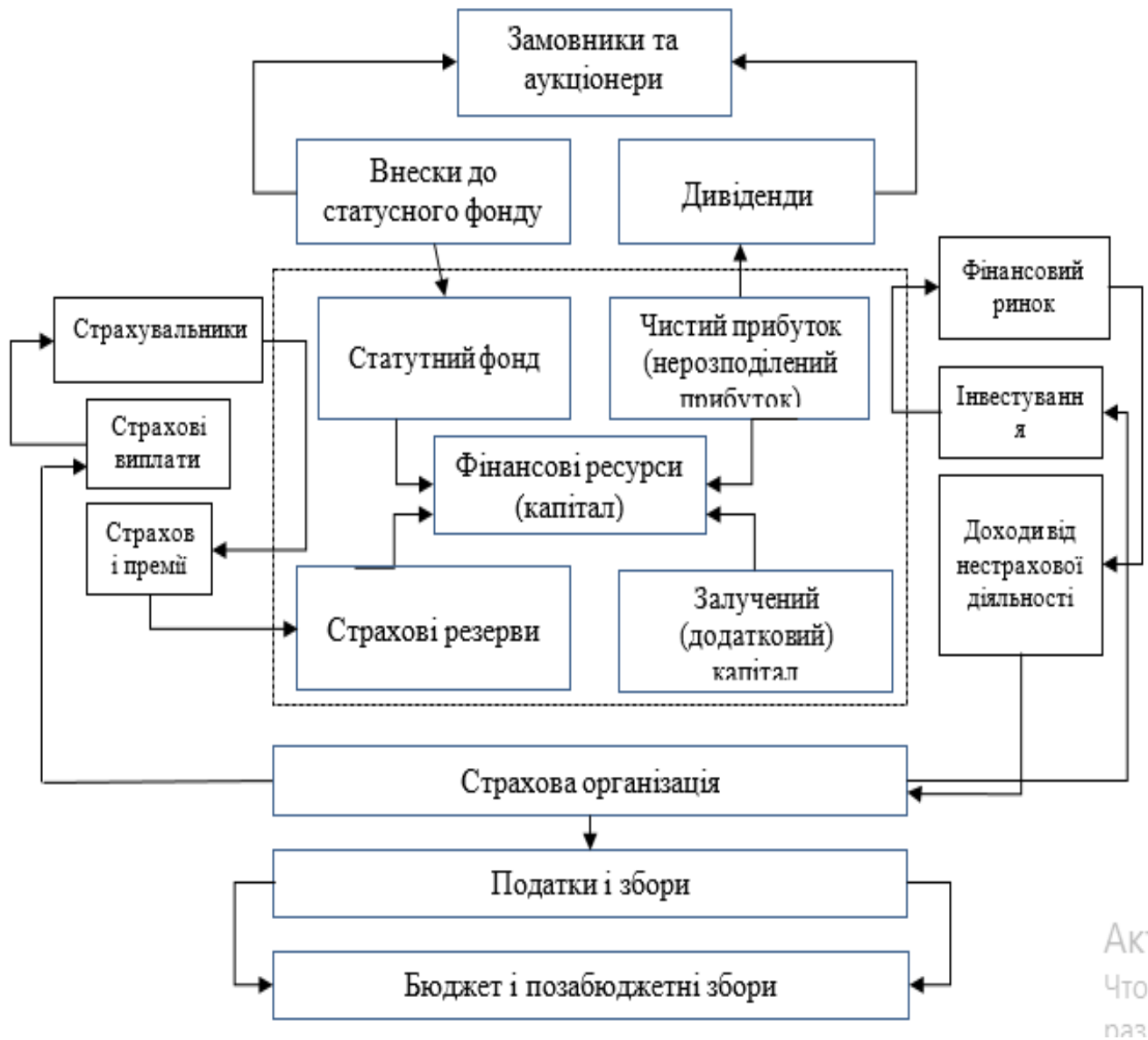


Рисунок 1.1 – Кругообіг основних фінансових процесів у СД

Страховики, акумулюючи певні грошові кошти, мають можливість виконувати страховий захист страхувальників, а також залучати інвестиційну діяльність. Тому СК перш за все є фінансовими установами, являються активними учасниками перерозподілу внутрішнього валового продукту, та перетворюються згодом на великі фінансові корпорації, які мають власну галузеву специфікацію.

1.2 Аналіз методів і моделей оцінювання фінансових ризиків у страхуванні

Для оцінювання фінансових ризиків у страхуванні можна використовувати різні спеціальні методи, які розвиваються та постійно вдосконалюються в часі. На практиці існують два основних підходи до аналізу ризиків:

- 1) з точки зору чинників ризику;
- 2) з точки зору наслідків дії ризиків.

Згідно до зазначених підходів математичні моделі для оцінювання ризиків також можна розподілити на два основних класи:

- 1) моделі, що засновані на аналізі наслідків – це високорівневі моделі;
- 2) моделі, що засновані на аналізі чинників ризику – це низькорівневі моделі.

В якості ознаки при класифікації моделей для оцінювання ризиків використана множина завдань, для вирішення яких модель може бути застосована. Розглянемо більш детально ці моделі.

1.2.1 Моделі, що засновані на аналізі наслідків

Ці моделі засновані на апараті математичної статистики, використання якого ґрунтується на зборі та дослідженні даних про втрати, які обумовлені фінансовим ризиком в попередніх періодах та екстраполяцією втрат на наступні періоди. Основними з цих моделей є наступні [4-6]:

1. ВІА (Basic Indicator Approach) – метод базових показників. Модель ВІА орієнтована на оцінювання максимально можливих втрат на заданому рівні значущості (99 %) і базується на тому, що операційні втрати, які пов'язані з ризиками, є залишковими від загальних втрат. Квантиль розподілу випадкової величини x , за допомогою якої можна характеризувати величину втрат, описується виразом:

$$\hat{Q}_{99\%} = \bar{G}_I \alpha, \quad (1.1)$$

де $\hat{Q}_{99\%}$ – оцінка 99%-го квантиля випадкової величини, яка характеризує

втрати;

\bar{G}_t – середній валовий дохід;

$\alpha = 15\%$ – визначений коефіцієнт для фінансових організацій і банків.

Видно, що вказана оцінка обсягу можливих фінансових втрат (1.1) залежить тільки від обсягу бізнесу і не залежить від особливостей страхування, що є недоліком моделі ВІА.

2. LDA (Loss Distribution Approach) – метод розподілу збитків. В моделях LDA припускається, що випадкова величину x , що характеризує розмір втрат, які мали місце протягом часу t , можна описати виразом:

$$x = \sum_{i=1}^{n(t)} L_i, \quad (1.2)$$

де $n(t)$ – випадкова величина, що характеризує кількість випадків втрат певного типу за час t ;

L_i – множина випадкових величин, які характеризують обсяги можливих втрат.

При побудові таких моделей необхідно провести аналіз втрат за обраний період для кожної пари «лінія бізнесу»/«тип втрат» і розрахувати вибіркоче середнє значення частоти виникнення ризикових подій та вибіркоче середнє значення величини втрат при їх виникненні [7].

3. ІМА (Internal Measurements Approach) – метод внутрішніх вимірів. Модель ІМА дозволяє обчислити розмір максимально можливих втрат для певного виду ризиків, не використовуючи розподіл випадкової величини x , яка характеризує розмір втрат. При побудові таких моделей лежить припущення: якщо розділити всі втрати на очікувані і непередбачені, то має місце функціональна залежність між величинами очікуваних і непередбачених втрат, якою може бути лінійна залежність:

$$\hat{Q}_{99\%} = E_I P_E L_{GE} = \gamma E_L \alpha, \quad (1.3)$$

де $\hat{Q}_{99\%}$ – оцінка 99%-го квантиля розподілу можливих втрат при цьому типі події;

P_E – ймовірність настання в даний період негативної події цього типу;

L_{GE} – середня величина втрат при настанні негативної події;

E_i – масштабний коефіцієнт;

γ – коефіцієнт оцінювання вимог до капіталу через збитки E_L .

Загальна величина можливих втрат оцінюється як сума всіх оцінок, які розраховані для ліній бізнесу і різних типів ризикових подій.

4. Статистичний метод оцінювання ризику. Цей метод ґрунтується на вивченні статистики втрат як негативних наслідків реалізації рішень, які виникали в таких же видах підприємницької діяльності. В якості показника, що розраховується на основі статистичного підходу, може бути частота втрат, які пов'язані з тим чи іншим видом діяльності:

$$f = \frac{n'}{n_{заг}}, \quad (1.4)$$

де n' – число випадків настання втрат;

$n_{заг}$ – загальне число випадків, які бралися у статистичній вибірці.

Зазначимо, що у статистичного методу є багато модифікацій і найбільшу популярність має метод статистичного моделювання (метод Монте-Карло), який дозволяє аналізувати і оцінювати різні сценарії розвитку, а також враховувати різні фактори впливу в межах певного підходу. Його недоліком є високий рівень використання ймовірних характеристик.

Одним із статистичних методів дослідження є метод середніх величин. Використання середніх величин дозволяє виразити типові розміри ознак, що змінюються, та визначити їх коливання біля середнього рівня. Для оцінювання ризику необхідне розмежування окремих ризикових груп на менші підгрупи, щоб створити аналітичну базу для визначення ризику за відповідними ознаками ризику.

До групи методів статистичного аналізу відноситься метод процентів який представляє деяку сукупність додатних та від'ємних відхилень, у термінах оцінювання ризиків, від середнього ризикового типу аналітичної бази, яка є в наявності. При цьому дані відхилення можуть виражатися в процентах або в проміле від середнього типу ризику.

1.2.2 Моделі, що засновані на аналізі факторів ризику

Ці моделі дозволяють виконати поглиблений аналіз процесів у СК і використовують інформацію про внутрішні причинно-наслідкові зв'язки. В цих математичних моделях використовують різні методи, зокрема такі: Sb-АМА (Scenario-based Advanced Measurement Approach) – сценарний аналіз; метод функціональних кореляцій; регресійні моделі; байєсівські мережі; методи нечіткої логіки та інші. Розглянемо ці методи.

Метод Sb-АМА базується на чинниках ризиків, які є основою генерації сценаріїв за правилом “що буде, якщо” [8]. На відміну від методів, що засновані на аналізі збитків, модель Sb-АМА заснована на оцінюванні можливих втрат, які матимуть місце в майбутньому. На основі даних або експертних оцінок для кожного сценарію визначають частоту і розмір втрат. Далі перевіряються дані та виправляються некоректні оцінки сценаріїв, що дозволяє їх згрупувати за чинниками ризику. На останньому етапі виконується імітаційне моделювання методом Монте-Карло, який дозволяє оцінити загальний розподіл втрат. Таким чином, виконання у такий спосіб постійного контролю за фінансовим станом дозволяє наперед прослідкувати негативні тенденції та зменшити ризик банкрутства на підприємстві.

1.2.3 Метод функціональних кореляцій

Згідно даного методу створюється структурна модель організації [9, 10]. Перехід до математичної моделі здійснюється шляхом побудови орієнтованого графа, в якому вершини відповідають співробітникам і підрозділам, а ребра відповідають потокам інформації. У графі кожній вершині ставиться у відповідність випадковий процес, який відображає її функціональність.

Згідно цього методу встановлюються стохастичні залежності між функціональностями пов'язаних вершин. Аналіз ризиків виконується не з точки зору окремих чинників ризику або наслідків, а відповідно до співробітників, систем,

чи бізнес-процесів, яким відповідають вершини графа.

Як і в методі LDA вводиться розподіл величин збитків (цей розподіл реалізується випадковою величиною L_i для вершини i), які пов'язані з певними випадками реалізації ризиків. Ці випадки для вершини i описуються випадковим процесом $n_i(t)$, який може мати два стани: $n_i(t) = 0$ – нормальне функціонування, а $n_i(t) = 1$ – збій. Загальні втрати $X_i(t + \Delta t)$, які у зв'язку зі збоями на i -й вершині графа понесені на момент $(t + \Delta t)$, можна описати так:

$$X_i(t + \Delta t) = X_i(t) + n_i(t + \Delta t)L_{i,t+\Delta t}, \quad (1.5)$$

де $L_{i,t+\Delta t}$ – реалізація випадкової величини L_i .

Для визначення зв'язку між функціональними елементами введемо поняття “підтримки”, що дозволить її отримати кожному функціональному елементу від пов'язаних з ним елементів.

Якщо обсяг “підтримки” буде меншим певної величини, тоді відповідний функціональний елемент переходить у стан відмови: для поставленої йому у відповідність i -ї вершини графа приймається $n_i(t) = 0$. Щоб відобразити цю структуру, вводяться в модель випадкові процеси $h_i(t)$:

$$h_i(t) = \mathcal{G}_i - \sum_j w_{ij} n_j(t) + \varepsilon_i(t), \quad (1.6)$$

де \mathcal{G}_i – середній обсяг «підтримки», яку i -й процес може отримати у системі, що повністю функціональна – коли всі $n_i(t) = 1$;

w_{ij} – величина підтримки, яку i -й процес може отримати від j -го процесу;

$\varepsilon_i(t)$ – гауссів білий шум.

Зазначимо, що цей метод орієнтований на виявлення збоїв та аналіз функціональності процесів. Крім того, фінансові ризики пов'язані з відсутністю коштів на відшкодування страхового випадку і мають іншу природу. Проведений аналіз показав, що цю модель недоцільно застосовувати для оцінювання фінансових ризиків.

1.2.4 Регресійні моделі

Ці моделі засновані на виявленні причинно-наслідкових зв'язків між індикаторами, що спостерігаються, і рівнем ризику. Слід розрізняти дві основні групи показників, які можна використовувати як індикатори, що спостерігаються: 1) змінні оточення – числові показники, які характеризують бізнес-процеси підприємства; 2) фактори ризику – числові показники, що характеризують випадки реалізації ризиків, що спостерігаються. Така регресійна модель має вигляд:

$$x = Af + b + \varepsilon, \quad (1.7)$$

де x – величина збитків, пов'язаних з фінансовим ризиком;

f – вектор значень змінних, що спостерігаються;

ε – випадкова величина, яка показує рівень похибки моделі;

A і b – параметри, що оцінюються, які характеризують залежність між зміною x та факторними змінними f .

Для застосування цього методу з метою забезпечення високої точності оцінок необхідно мати необхідний обсяг даних.

1.2.5 Методи експертного оцінювання

При дослідженні фінансових систем як складних систем можуть виникати задачі, які знаходяться за межами формальних математичних задач. Основна ідея методу експертного оцінювання полягає у використанні знань людей для знаходження рішення слабо формалізованих завдань.

Проведення експертного оцінювання складається з наступних кроків:

- формулювання мети оцінювання;
- постановка задачі;
- створення групи для управління процесом оцінювання;
- створення форми отримання результатів оцінювання;
- підбір групи експертів та визначення їх компетентності.

Підбір експертів, які увійдуть в групу, здійснюється таким чином, щоб фахівці, по-перше, були знайомі зі специфікою роботи підприємства (об'єкта рішень); по-друге, щоб вони не були зацікавлені в результатах оцінювання. Тому, як правило, включають до групи експертів 2–3-х фахівців підприємства та 2–3-х зовнішніх експертів. Шляхом аналізу професійної, наукової та іншої діяльності визначається компетентність експертів, яка об'єктивно впливає на ступінь їхньої кваліфікації у певній області знань. Крім того, можна використати суб'єктивний метод оцінювання компетентності, який полягає у взаємному оцінюванні кожним експертом своєї кваліфікації і кваліфікації своїх колег і за певною шкалою. Встановлюється компетентність експертної групи після оброблення результатів опитування, що дозволяє визначити можливу похибку оцінювання.

Розглянемо методи обробки результатів опитування.

1. *Метод надання переваг.* При його використанні експерти нумерують показники ризику (об'єкти) в порядку їх характерності, згідно якого найменш характерний елемент отримує 1-й. При обробці результатів розраховують коефіцієнт характерності (відносної важливості) j -го елемента за виразом:

$$K_j = \frac{\sum_{i=1}^n K_{ij}}{\sum_{j=1}^s \sum_{i=1}^n ij}, \quad (1.8)$$

де K_{ij} – місце, на яке поставлена характерність j -го показника (елемента) у i -го експерта;

n – число експертів;

s – число елементів, які розглядаються.

Показником ризику буде той показник, для якого цей коефіцієнт найбільший, який, на думку експертів, більшою мірою характеризує об'єкт.

2. *Метод рангів.* Згідно цього методу експерти в наперед обумовленому діапазоні: 0–1, 0–10, 0–100 оцінюють важливість кожного елемента за шкалою відносної значущості. На відміну від попереднього, цей метод дозволяє оцінити не тільки пріоритетність j -го елемента, а й ступінь його вагомості.

Для оцінювання фінансового ризику експертні методи застосовують у тому

випадку, якщо відсутні на підприємстві необхідні інформативні дані, щоб здійснити розрахунки економіко-статистичними методами. Ці методи ґрунтуються на опитуванні кваліфікованих фахівців (фінансових, страхових, інвестиційних менеджерів у відповідних спеціалізованих організаціях тощо), щоб виконати математичну обробку результатів опитування.

На практиці широко застосовуються групові методи експертизи: наради, консилиуми, бізнес-тренінги, закриті обговорення тощо. Слід зазначити, що методи колективного обговорення дають більше ідей, ніж їх можна одержати від експертів у процесі індивідуальної експертизи. Незалежно від форми проведення експертизи всі вони базуються на виставленні бальних оцінок окремим факторам ризику та визначенні їх частки. Недоліком експертного методу є суб'єктивність оцінок, які отримуються.

3. Аналітичний метод. В його основі лежить правило ринкової економіки про те, згідно якого більший ризик пов'язаний з більшим доходом. Тому застосування будь-якого аналітичного методу зводиться до оцінювання приросту ризику проекту і приросту доходу проекту, тобто оцінки корисності.

При застосуванні аналітичного підходу використовуються багато методів, серед яких можна виділити такі:

- метод аналізу абсолютних і відносних показників;
- метод аналізу чутливості.

Серед аналітичних методів розглянемо метод аналізу чутливості моделі, який використовується в такій послідовності:

1-крок: вибір головного показника, відносно якого буде здійснюватися оцінка чутливості;

2-крок: вибір факторів ризику (вхідних параметрів), які впливають на головний показник, змінюючи його від бажаного значення;

3-крок: розрахунок конкретних значень головного показника на різних етапах створення проекту, змінюючи різні значення факторів ризику.

Таким чином формуються потоки надходжень і витрат, які дають можливість розрахувати показники ефективності в кожний момент часу. Потім будуються

діаграми, які показують залежності між результуючим показником і змінними вхідними параметрами. Порівнюючи між собою ці діаграми, можна вказати фактори ризику, зміни в яких найбільше впливають на значення головного параметра. Менеджер, враховуючи критичні для проекту значення факторів ризику, визначає слабкі місця і розробляє програму дій для усунення цих слабких місць. Наприклад, якщо ціна на продукцію буде критичним фактором, то можна змінити програму маркетингу або зменшити вартість проекту.

Недоліком аналітичного методу є те, що він не являється всеохоплюючим, так як не може врахувати всі можливі обставини реалізації. Також він не враховує ймовірності створення альтернативних проектів.

3. *Метод аналогів.* Для аналізу ризиків можна використати інформацію про вплив факторів ризику на проекти, що були виконані раніше. Для цього необхідно створити інформаційну базу, на підставі якої можна зробити узагальнення і прийняти рішення щодо виконання проекту. Недоліком методу аналогій є його описовий характер. Крім того, вплив факторів ризику на подібні проекти з часом може змінитися, як може змінитися і сама множина факторів ризику.

3. *Байєсівські мережі (БМ).* БМ дають можливість відобразити причинно-наслідкові зв'язки між різними чинниками ризику, що впливають, та змінами середовища. Байєсівські мережі, на відміну від регресійних моделей, дозволяють враховувати не лише безпосередні залежності ризику від факторів ризику, а й безпосередньо самі залежності між факторами ризику. Крім того, клас моделей БМ дозволяє формування висновку також на основі неповних даних.

З математичної точки зору БМ – це орієнтований граф, вершини якого відповідають чинникам ризику і змінам середовища, а ребрам являють собою виявлені або передбачувані взаємозв'язки. Слід зазначити, що БМ описується множиною умовних розподілів випадкових величин, які характеризують чинники ризику і змінні середовища.

Нехай існує n випадкових величин X_1, \dots, X_n . Спільна ймовірність розподілу значень цих величин може бути виражена як добуток n умовних імовірностей:

$$P(x_1, \dots, x_n) = P(x_1) \prod_{j=2}^n P(x_j | (x_1, \dots, x_{j-1})). \quad (1.9)$$

До переваг БМ можна віднести можливість одночасного використання як експертного оцінювання (наприклад, вибір типів апріорних розподілів для змінних та оцінювання структури мережі шляхом визначення залежностей між відповідними змінними), так і математичних методів для отримання висновку по БМ. Це дозволяє зв'язувати вибірки статистичних даних з експертними знаннями.

Слід зазначити, що формування ймовірнісного висновку на основі БМ може здійснюватися через поширення інформації у будь-якому напрямку шляхом розрахунку умовної ймовірності отримання значень для частини випадкових величин за умови відомих значень інших величин.

З математичної точки зору це завдання можна сформулювати як обчислення $P(y, x)$, де X – множина значень, що спостережуються, а Y – множина змінних, які потрібно оцінити. Таким чином, задача зводиться до пошуку умовних ймовірностей за виразом:

$$P(y | x) = \frac{\sum_s P(y, x, s)}{\sum_{y,s} P(y, x, s)}, \quad (1.10)$$

де S – множина всіх змінних, крім змінних X і Y .

Виконання таких обчислень є трудомісткими, а задача формування висновку є NP – повною. Для забезпечення високої точності та спрощення обчислень було розроблено та апробовано багато алгоритмів побудови висновку у БМ.

Залежно від повноти інформації, якою володіє суб'єкт підприємницької діяльності, методи оцінювання фінансових ризиків із врахуванням таких умов умовно можна поділити на такі групи (рисунок 1.2):

- *визначеності*, коли інформація про ризик достатньо повна, наприклад, представлена у вигляді бухгалтерського балансу, звітів про збитки і прибутки, та ін.;
- *часткової невизначеності*, коли інформація про ризик існує у вигляді частот появи ризикових подій;
- *повної невизначеності*, коли інформація про ризик повністю відсутня, але для часткового подолання невизначеності є можливість залучення фахівців та

експертів.

Таким чином, зменшення обсягу вихідної інформації про середовище діяльності робить складним розв'язок задачі оцінки ступеню ризику, використовуючи скорочення можливостей та зменшення достовірності результатів, що отримуються. Розглянемо зміст кожної із груп методів оцінювання ризику наведених на рисунку 1.2.

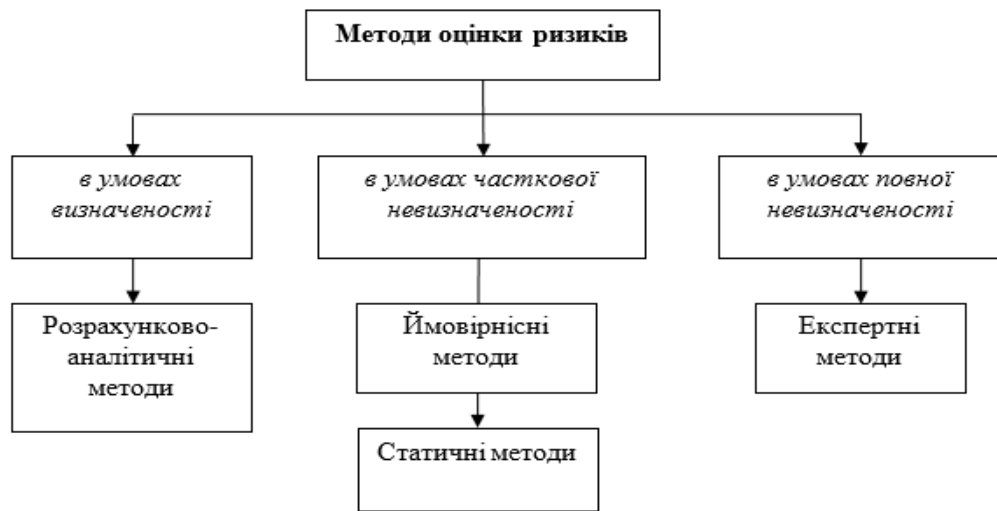


Рисунок 1.2 – Класифікація методів оцінювання ризиків

В умовах визначеності як правило використовуються розрахунково-аналітичні методи, які застосовують для розрахунку показників ризику за даними звітності. У цих випадках показники оцінки ризику можуть виражатися у вигляді абсолютних, відносних або середніх величин.

Абсолютні величини характеризують результати різних ризикових подій:

- у вигляді грошового (вартісного) чи фізичного (матеріального) вираження, якщо відповідні збитки піддаються таким одиницям виміру;
- у вигляді балансових відношень, які відображають результати фінансово-господарської діяльності підприємства: спроможність фінансових джерел формувати матеріально-оборотні засоби або ліквідність балансу організації, та ін.

Зрозуміло, що показники в такому разі представлені у вигляді *абсолютних величин*.

У відносному представленні ризик можна характеризувати як величину можливих втрат, яка віднесена до певної бази, за яку можна вважати або очікуваний дохід (прибуток) від діяльності, або затрати ресурсів на даний вид підприємницької діяльності. Відносні показники по відношенню до абсолютних є похідними (вторинними).

Показники *середніх величин* ризику являються узагальненими показниками, в яких можуть мати місце фактори ризику, діючі причини, закономірності підприємницької діяльності тощо. Але при цьому згладжується різниця результатів діяльності окремих підприємців/організацій, що може відобразитися або як загальне, або як властиве тільки всій масі підприємців/організаціям у певній сфері діяльності за відповідною ознакою.

В умовах *часткової невизначеності* ризик розглядається як ймовірнісна характеристика, до якої доцільно використовувати ймовірнісні та статистичні методи оцінювання ризику, використання яких дозволяє розрахувати ймовірнісні та статистичні показники оцінки ризику. Ймовірнісні показники можна розглядати як міру можливості появи/настання ризику та його наслідків. Як правило, на основі відповідної інформації здійснюється розрахунок частоти ризикової події, яка виникає. Відповідні наслідки ризикової події можуть бути представлені у вигляді *точкових* або *інтервальних* оцінок.

В умовах *повної невизначеності* використовуються експертні методи, які являються суб'єктивними і допомагають вжити відповідних заходів для зниження ступеня невизначеності та прийняття обґрунтованих рішень.

1.3 Аналіз типів даних і невизначеностей в актуарному моделюванні

У прикладній статистиці об'єктом дослідження являються статистичні дані, які являють собою сукупність дослідів (об'єктів) та змінних (ознак), що їх характеризують та отримуються при виконанні досліджень чи експериментів. Об'єктом дослідження можуть бути, наприклад, економічні показники країни, тобто доля ВВП на одиницю населення, видатки та доходи бюджету, страхові виплати і

премії в актуаріїв, індекси якості життя та ін. Ознаками, які є змінними, являються величини, які отримуються в результаті дослідів та можуть приймати різні значення.

При моделюванні та оцінюванні фінансових ризиків можуть зустрічатися різні невизначеності, типи яких, причини їх виникнення та методи їх подолання приведені у таблиці 1.1.

Таблиця 1.1 – Типи невизначеностей при моделюванні та оцінюванні фінансових ризиків

№	Тип невизначеності	Причини виникнення	Методи подолання
1	Структурна невизначеність моделі	– неможливість встановлення всіх причинно-наслідкових зв'язків між змінними; – наближені значення елементів структури моделі.	– експертні методи; – застосування статистичних тестів; – застосування теорії перевірки гіпотез;
2	Статистичні невизначеності у процесі побудови моделей	– похибки вимірів; – стохастичні зовнішні збурення; – мультиколінеарність; – екстремальні значення; – пропуски вимірів.	– цифрові та оптимальні фільтри; – уточнення типів розподілів; – метод головний компонент; – теорія екстремальних значень; – методи заповнення пропусків.
3	Параметричні невизначеності	– некоректний вимір методу оцінювання; – короткі вибірки.	– забезпечення вибору альтернативних методів оцінювання параметрів; – розмноження вибірок.
4	Ймовірнісні невизначеності	– складні механізми виникнення причинно-наслідкових зв'язків; – відсутність детермінованості.	– статистичні і динамічні мережі Байеса; – марківські моделі; – ймовірнісні фільтри; – умовні розподіли.
5	Невизначеність амплітудного типу	– наявність змінних, що не вимірюються.	– методи обробки нечіткої інформації.

Слід враховувати характер даних, які включаються до аналізу (типи змінних, форми їх розподілів та залежності між ними), щоб зробити правильний вибір статистичного тесту. Змінні процесу (ознаки можуть) можуть вимірюватися в різних шкалах (номінальна, інтервальна, порядкова, шкала відношень та ін.).

Розрізняють три основних типи даних: дискретні, неперервні та номінальні (категоріальні). Дискретні змінні представляють собою кінцеву множину упорядкованих значень, які класифікуються, в свою чергу, по ступіні прояву на ранговій шкалі (наприклад, курс валют, індекси в економіці та ін.). Неперервні дані можуть приймати будь-які числові значення, які упорядковуються природним способом на числовій осі. Категоріальні дані являються неупорядкованими та, як правило, використовуються для якісної класифікації (адреса проживання, стать та ін.). В категоріальних даних можна виділити бінарні значення та категоричні значення (наприклад, є/немає, так/ні та ін.).

2 ПОБУДОВА МАТЕМАТИЧНИХ МОДЕЛЕЙ ДЛЯ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ АКТУАРНИХ ПРОЦЕСІВ

Середовище актуарних процесів являє собою сукупність процесів, що мають місце в діяльності страхових компаній, яка спрямована на фінансову компенсацію наслідків випадкових подій, що несуть за собою матеріальні втрати. Страхова діяльність орієнтована на перерозподіл коштів та їх акумуляцію безпосередньо для страхової діяльності, з однієї сторони, та інвестування цих коштів в різні галузі діяльності, з іншої сторони, що сприяє їх розвитку. Тому для цієї сфери, згідно характеру та різноманітності діяльності СК, необхідна розробка і вдосконалення математичних моделей для прогнозування за умов невизначеності, а також оцінювання ризиків реалізації актуарних процесів. Виходячи з цього побудуємо математичні моделі для аналізу та оцінювання ризиків СК.

2.1 Аналіз методів і моделей побудови математичних моделей актуарних процесів

На практиці застосовують найбільш часто структурний або функціональний методи побудови математичних моделей. Згідно структурного метода виконується моделювання внутрішнього механізму взаємодії змінних, яке відображає їх фактичні взаємозв'язки. За критерій оптимальності структурної моделі обирається максимально наближений характер поведінки основних змінних моделі до реального процесу.

Функціональний метод застосовується для формального опису процесу, коли інформація про його структуру і взаємодії між змінними відсутня або тільки частково береться до уваги. У цьому випадку процес називають «чорною скринькою». Для побудови математичних моделей будемо використовувати структурний метод, який базується на системному підході [11].

Структурна схема процедури побудови математичних моделей актуарних

процесів, яка основана на використанні структурного методу та системного підходу до аналізу даних [12], наведена на рисунку 2.1.

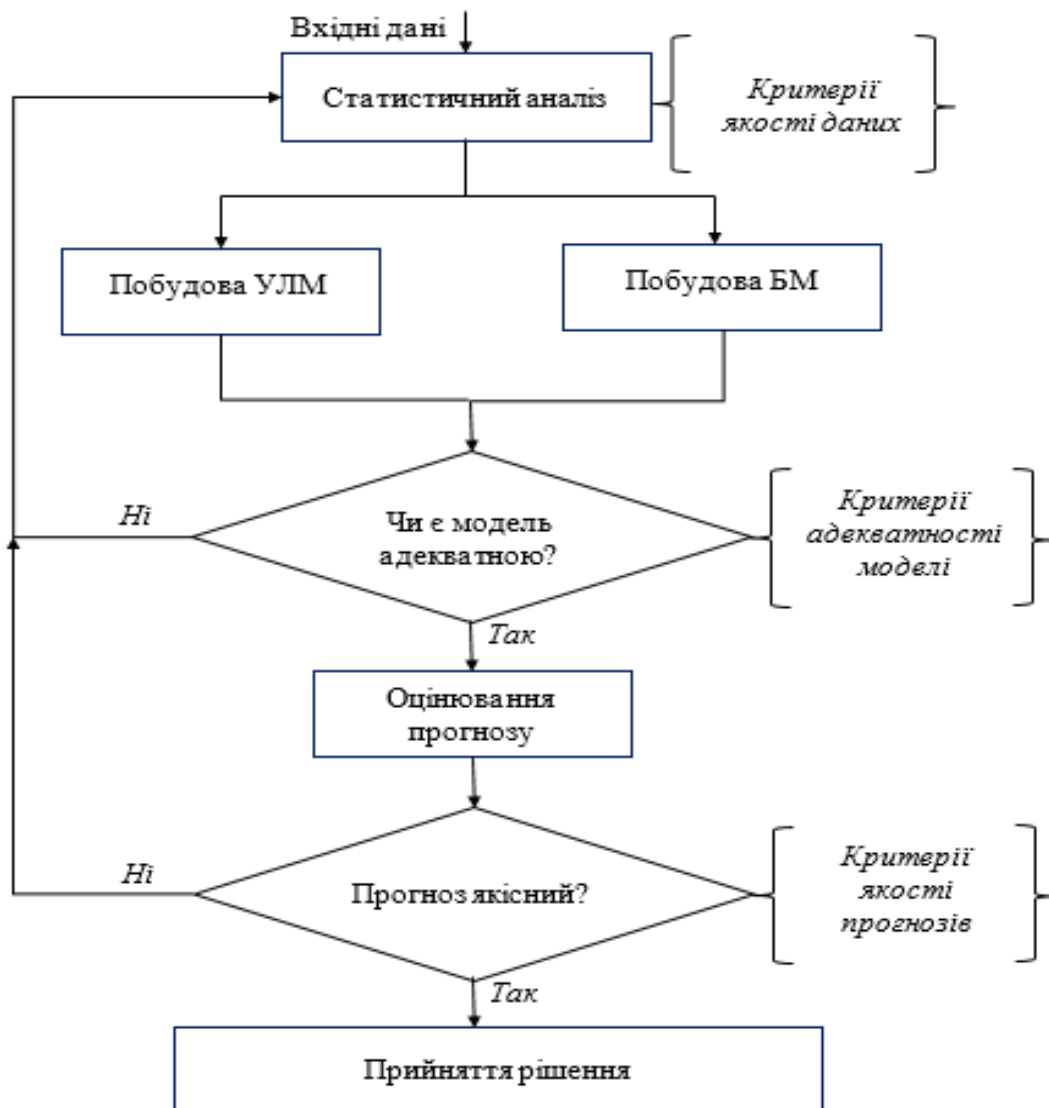


Рисунок 2.1 – Структура схема процедури побудови математичних моделей актуарних процесів

2.2 Визначення структури математичної моделі актуарних процесів

Задача оцінки структури моделі, яка є адекватною досліджуваному процесу, являється не простою і вирішується, як правило, ітераційно. Спочатку, на підставі дослідження закономірностей протікання процесу, візуального аналізу даних та аналізу кореляційних функцій, структуру моделі оцінюють наближено. Вибирають

кілька найбільш ймовірних структур-кандидатів, використовуючи відповідні статистичні характеристики моделей, а потім вибирають кращу з них, обчислюючи оцінки параметрів моделей-кандидатів. Якщо жодна з моделей-кандидатів не являється адекватною, то слід дослідити на інформативність експериментальні дані, виявити типи випадкових зовнішніх збурень, оцінити похибки вимірів, а також умови, в яких функціонує досліджуваний процес.

При оцінці неадекватності моделі найпоширенішою помилкою може бути недостатній об'єм статистичних даних або вибірок. У такому випадку необхідно зробити повторний або додатковий збір експериментальних даних, визначити впливові фактори тощо.

З метою забезпечення необхідної точності оцінок прогнозів актуарних процесів пропонується методика побудови математичної моделі актуарного процесу, яка передбачає наступні етапи [12, 13]:

1) проведення системного аналізу процесу, для якого формується модель, з використанням експертних оцінок протікання процесу, візуальне дослідження графіків вхідних і вихідних змінних, які подані часовими рядами, використання вже існуючих моделей та іншої інформації;

2) підготовка даних для побудови моделі і попередня обробка експериментальних даних;

3) за допомогою описових статистик і візуального аналізу проведення аналізу часових рядів на стаціонарність, можливу наявність екстремальних значень та виродженість;

4) знаходження та обробка екстремальних значень;

5) враховуючи базову структуру УЛМ, оцінювання структури моделей-кандидатів, що включає:

– ідентифікацію систематичної і стохастичної складових моделі та функцій зв'язку;

– забезпечення виконання припущень щодо законів розподілу залежних змінних;

– використовуючи критерії відповідності, обчислення основних статистик та

оцінювання відповідності моделі або класу моделей, що оцінюються, статистичним даним;

– оцінювання характеристики інших елементів структури математичної моделі (лагів, залишків тощо);

б) вибір методу оцінювання невідомих параметрів математичних моделей структур, які обрано. Це може бути метод найменших квадратів (МНК) або його модифікації, метод максимальної правдоподібності (ММП), метод Монте-Карло, байєсівський підхід та ін.

7) за допомогою множини статистичних критеріїв адекватності моделей вибір та обґрунтування кращої з моделей-кандидатів, які були оцінені.

Застосування цієї методики допоможе побудувати математичну модель актуарних процесів та оцінити її ефективність та практичну значимість.

2.2.1 Системний аналіз актуарного процесу

У загальному випадку математична модель системи містить формальний опис множини можливих її станів та закон переходу з одного стану в інший. Щоб побудувати відповідну математичну модель, необхідно на етапі системного аналізу здійснити ідентифікацію об'єкта дослідження як системи, яка передбачає виконання якісного аналізу актуарного процесу.

Аналіз актуарного процесу являється надзвичайно важливим етапом, якісне виконання якого потребує, перш за все, практичного досвіду дослідження реальних процесів тієї чи іншої природи, і спрямовується на розв'язання таких задач [14]:

а) визначення розмірності процесу та його моделі, тобто, визначення кількості її входів і виходів;

б) встановлення зв'язків між змінними математичної моделі та аналіз можливостей їх математичного опису;

в) встановлення типу та кількості зовнішніх збурень (детерміновані чи стохастичні; види їх розподілів як випадкових процесів тощо);

г) встановлення можливості розподілу актуарного процесу на підпроцеси, які

повинні бути простішими як з точки зору математичного опису, так і з точки зору їх функціонування;

д) за умови, що процес має ієрархічну структуру, слід відокремити ці рівні, при цьому визначивши функції кожного з рівнів та встановити типи зв'язків, які існують між ними;

е) аналіз та використання знань щодо відомих законів та закономірностей протікання процесу, особливостей його функціонування, виявлення моделей процесу, що існують, та досвіду теоретичного та експериментального його дослідження;

ж) якщо існують розроблені моделі досліджуваного процесу, необхідно встановити їх переваги та недоліки, а також визначити можливість їх подальшого використання або модифікації, що суттєво скорочує час та інші витрати на побудову і використання моделі.

Якщо не виконати цей етап, то це може призвести до ускладнень або до неможливості побудови моделі, яка мала б високу ступінь адекватності процесу, або призвести до отримання хибної моделі, яка не відображає досліджуваний процес. Крім того, використання такої моделі може призвести до неправильного прогнозу, і як наслідок, до великих втрат.

2.2.2 Попередня обробка експериментальних даних

Експериментальні дані з практичної точки зору являються випадковими величинами або наближеними до них. Тому в ході експериментальних досліджень можливе отримання інформації, яка спотворена певними перешкодами, в результаті вихідні дані можуть суттєво відрізнитись за одних і тих же умов. Метою попередньої обробки даних є відсіювання грубих похибок та оцінювання достовірності експериментальних даних, а також їх підготовка для оцінювання структури і параметрів моделі.

Попередня обробка експериментальних (статистичних) даних, як правило, включає наступні операції:

- а) візуальний аналіз даних i , за необхідності, їх корегування та нормування;
- б) під корегуванням розуміється заповненні пропусків та зменшення екстремальних значень (викидів), які виходять за заданий діапазон значень змінних;
- в) формування різниць різних порядків, які необхідні для аналізу складових часового ряду, який описує процес.

Використання різниць дозволяє будувати моделі, коли має місце швидкість та прискорення основної змінної. Інколи віднімають із значень ряду його середнє з тим, щоб отримати можливість працювати не з повними значеннями вимірів змінних, а з їх відхиленнями. Виникає необхідність це робити при побудові моделей у просторі станів.

Для оцінювання середнього можна використати метод оновлюваного середнього, оснований на використанні рекурентної формули обчислення величини середнього арифметичного. Якщо випадкова величина X надходить у вигляді дискретних значень i для $(N - 1)$ -го виміру обчислюється середнє значення, то поява нового значення змінює попереднє середнє значення на величину $\frac{1}{N}(X_N - X_{N-1})$, тобто виконується згладжування. Таким чином, при згладжуванні згідно цього метода експериментальне значення у кожній точці на числовій осі замінюється на величину середнього, яке розраховане на цей момент часу. Послідовність отриманих в такий спосіб середніх значень являє собою статистичний ряд щодо змінної x , який може бути використаний при подальшій обробці.

Застосування того чи іншого методу попередньої обробки даних для моделювання визначається окремо для кожного випадку.

2.2.3 Визначення наявності нелінійностей

Для визначення характеру та наявності нелінійностей у досліджуваному процесі можуть бути використаними візуальний аналіз даних та формальні тести. Візуальний аналіз дозволяє виявити існування ділянок з лінійним або нелінійним трендом, а також можливу наявність гетероскедастичності та значних викидів, які

можуть вплинути суттєво на якість отриманої моделі [14].

Існують багато формальних тестів на визначення наявності нелінійності. Розглянемо простий тест, який можна застосувати у випадку, коли до уваги беруться для одного і того ж процесу кілька вибірок спостережень:

$$\hat{F} = \frac{\frac{1}{m-2} \sum_{i=1}^m n_i (\bar{y}_i - \hat{y}_i)^2}{\frac{1}{n-m} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2},$$

де \bar{y}_i – середнє значення для i -ї вибірки або групи даних;

\hat{y}_i – середнє для лінійної апроксимації даних;

m – число груп даних;

n_i – число вимірів в i -й групі;

n – загальне число вимірів.

Якщо отримана статистика \hat{F} з $\nu_1 = m - 2$ та $\nu_2 = n - m$ ступенями свободи досягає або більше рівня значущості, то гіпотезу про лінійність слід відхилити. Недоліком цього підходу є те, щоб його застосувати, необхідно мати декілька (не менше трьох) груп даних для процесу, які можна отримати при виконанні повторних експериментів.

За допомогою вибірових нелінійних кореляційних функцій (НКФ) також можна встановити наявність нелінійності. Кореляційні функції розраховуються за вибірками експериментальних (статистичних) даних. Наприклад, якщо дискретна НКФ

$$r_{yx^2}(s) = r_{y(k) x^2(k-s)} = \frac{1}{N} \frac{\sum_{k=s+1}^N \{ [y(k) - \bar{y}] [x(k-s) - \bar{x}]^2 \}}{\sigma_y \sigma_x^2}, \quad s = 0, 1, 2, 3, \dots,$$

містить значення, які в статистичному смислі суттєво відрізняються від нуля, то можна стверджувати, що процес містить квадратичну нелінійність стосовно регресора x .

Також можна визначити наявність нелінійного детермінованого тренду у

процесі шляхом оцінювання рівняння:

$$y(k) = a_0 + c_1 k + c_2 k^2 + \dots + c_m k^m,$$

яке стосовно дискретного часу представляє собою поліном порядку m .

Якщо хоча б один із коефіцієнтів c_i , $i=1, \dots, m$ являється статистично значущим, то гіпотеза про відсутність тренду відхиляється. Якщо тренд відносно швидко змінюється в напрямку свого руху і для нього важко знайти відповідний функціональний опис, то можна застосувати моделі випадкових трендів, що засновані на комбінаціях випадкових величин [11, 14].

Слід зазначити, що метод групового врахування аргументів (МГВА) автоматично оцінює структуру математичної моделі. Цей метод широко застосовується для формалізованого опису широкого класу процесів, а також для моделювання процесів різної природи з нелінійностями та нестационарностями. Подальшим розвитком цього методу є нечіткий МГВА, який використовує нечітке представлення параметрів оцінюваної моделі [14].

2.2.4 Обробка екстремальних значень

Часто зустрічаються недостатні обсяги інформації при дослідженні процесів фізичного, економічного, природничого походження, що стає причиною труднощів при розв'язанні задач при побудові моделей для екстремальних даних. Наприклад, економісту потрібно знайти максимальне значення з деякої вибірки з виродженими даними. Існує кілька можливих способів, використовуючи які експерт із знаннями досліджуваного процесу може надати необхідну інформацію, яка має відношення до екстремальної поведінки і залежить від наявних даних. Разом з тим така інформація часто супроводжується наближеними вимірами, які можуть відрізнитися від дійсних значень та робити майбутні прогнози хибними, що на їх основі будуються. Саме теорія екстремальних значень (ТЕЗ) є Одним із сучасних методів аналізу таких екстремальних значень [15].

ТЕЗ дозволяє виконати цілеспрямований аналіз та оцінювання ймовірності

появи випадкових величин, які пов'язані з екстремальними та рідкісними подіями. Екстремальні значення – це нові випадкові величини, які не є фіксованими та залежать від характеру протікання процесу, об'ємів вибірки та типу вихідного розподілу. Наприклад, в галузі страхування майна рідкісною, але ймовірною подією є настання такого страхового випадку, який супроводжується значними виплатами страхових премій.

Математичну модель екстремальних даних можна представити наступним чином [15]:

$$M_n = \max\{X_1, \dots, X_n\}, \quad (2.1)$$

де X_1, \dots, X_n – послідовність незалежних випадкових величин, що мають функцію розподілу F .

У виразі (2.1) величина M_n є максимум досліджуваного процесу на інтервалі часу n , яка має розподіл [11]

$$\begin{aligned} \Pr\{M_n \leq z\} &= \Pr\{X_1 \leq z, \dots, X_n \leq z\} = \\ \Pr\{X_1 \leq z\} \times \dots \times \Pr\{X_n \leq z\} &= \{F(z)\}^n. \end{aligned} \quad (2.2)$$

Так як функція F невідома, то розглядається наближена оцінка для F^n . Якщо послідовність констант $\{a_n > 0\}$ та $\{b_n > 0\}$ таких, що

$$\Pr\left\{\frac{M_n - b_n}{a_n} \leq z\right\} \rightarrow F(a_n x + b_n)^n \rightarrow G(z),$$

при $n \rightarrow \infty$, то G є невиродженою функцією розподілу, яка належить до одного з розподілів екстремальних значень, наприклад, до узагальненого розподілу екстремальних значень (Generalized extreme value – GEV) [16]:

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right]^{-1/\xi}\right\},$$

де μ – параметр розподілу,

ξ – параметр форми розподілу,

σ – параметр масштабованості.

Згідно з теоремою про типи розподілів екстремальних значень можна виділити

три типи таких розподілів, а саме:

1) розподіл Гумбела:

$$G(z) = \exp \left\{ - \exp \left(- \left(\frac{z-b}{a} \right) \right) \right\}, \quad -\infty < z < \infty;$$

2) розподіл Фреше:

$$G(z) = \begin{cases} 0, & z \leq b; \\ \exp \left(- \left(\frac{z-b}{a} \right)^{-\alpha} \right), & z > b; \end{cases}$$

3) розподіл Вейбулла:

$$G(z) = \begin{cases} \exp \left(- \left(- \left(\frac{z-b}{a} \right) \right)^\alpha \right), & z < b. \\ 1, & z \geq b \end{cases}$$

Для всіх трьох випадків $a > 0$, b – дійсне число. Для другої та третьої функцій $\alpha > 0$. Ці класи розподілів називають розподілами екстремальних значень.

В окремий клас можна виділити узагальнений розподіл Парето (*Generalized Pareto Distribution – GPD*), який отримується при наступній умові: X – це розподіл, що умовно перевищує деякий поріг u :

$$F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)}, \quad (2.3)$$

де $u \rightarrow w_F = \sup\{x : F(x) < 1\}$, що найчастіше зводиться до пошуку границі:

$$F_u(y) \approx G(y, \sigma_u, \xi),$$

де G – узагальнений розподіл Парето, який еквівалентний виразу [16]:

$$G(y, \sigma, \xi) = 1 - \left(1 + \xi \frac{y}{\sigma} \right)_+^{-1/\xi},$$

1) якщо $\xi > 0$, то маємо довгий хвіст $x^{-1/\xi}$, що еквівалентно розподілу Парето;

2) якщо $\xi = 0$ та спрямовуючи $\xi \rightarrow 0$, отримаємо

$G(y, \sigma, 0) = 1 - \exp\left(-\frac{y}{\sigma}\right)$, тобто експоненціальний розподіл з середнім σ ;

3) якщо $\xi < 0$, то кінцева верхня точка знаходиться на рівні $-\frac{\sigma}{\xi}$.

Слід зазначити, що однією із переваг практичного застосування *GEV*-розподілів виступає інваріантність кожного з розподілів, що належать до певного класу.

2.2.5 Формування структури моделі

Необхідно на даному етапі вибрати структури моделей кандидатів, поняття яких слід розглядати, виходячи із складових УЛМ, і включає таке:

а) визначення залежної змінної (стохастичної складової), яка розподілена з середнім μ . відповідно до законів розподілу із сімейства експоненціальних законів розподілу. Її називають розподілом вихідної змінної;

б) встановлення p -незалежних змінних (систематичної складової), які об'єднуються в один «лінійний предиктор» [17]:

$$\eta = X \cdot \beta.$$

в) визначення характеру *функції зв'язку*, виходячи з того, що вона відображає взаємозалежність між припущенням систематичності та випадковості.

До структури УЛМ належать показники, за допомогою яких можна визначити:

- 1) порядок моделі або порядок рівнянь, що її утворюють;
- 2) розмірність або кількість рівнянь моделі;
- 3) залишки в моделі, які використовуються для оцінки адекватності моделі прогнозування, вибору функції дисперсії, функції зв'язку та елементів лінійного предиктора;
- 4) наявність нелінійностей та їх типи;
- 5) наявність зовнішніх збурень та їх типів (детерміновані або випадкові; адитивні та мультиплікативні).

В загальному випадку кореляційна функція, або коефіцієнт кореляції дають можливість встановити наявність зв'язку між змінними, а кореляційна матриця дозволяє встановити наявність зв'язку між залежною змінною та незалежними змінними в моделі. Тому для визначення необхідності включення незалежних змінних (регресорів) в праву частину рівняння, слід обчислити коефіцієнт кореляції між залежною та незалежними змінними моделі.

Загальний вплив випадкових факторів, що не вимірюються, можна описати випадковою змінною $\varepsilon(k)$. Так як вона не вимірюється, то оцінку її значення, або похибки моделі, або залишок можна отримати після оцінювання коефіцієнтів моделі:

$$\hat{\varepsilon}(k) = e(k) = y(k) - \hat{y}(k),$$

де $\hat{y}(k)$ – оцінка змінної $y(k)$, що отримана по моделі;

$y(k)$ – фактичний вимір.

Для аналізу змінних моделі, крім кореляційної матриці, можна використовувати наступні описові статистики: коефіцієнт ексцесу, коефіцієнт асиметрії, та інші, що дозволяють виконати перевірки гіпотези про нормальний розподіл.

Цей етап закінчується формуванням структур декількох моделей-кандидатів, які можуть мати різні закони розподілу: гамма, нормальний, розподіл Пуассона, та відповідні види функцій зв'язку (тотожна, логарифмічна тощо). Так як встановити адекватну структуру одноразово, як правило, неможливо, то кандидатів може бути багато. В цілому побудова моделі високого ступеня адекватності являє собою ітераційний процес, який вимагає великих зусиль.

Далі необхідно оцінити параметри моделей-кандидатів.

2.2.6 Оцінювання параметрів моделей

На шостому етапі оцінюють коефіцієнти моделі, на основі принципу економії або збереження, який означає, що кількість коефіцієнтів, що оцінюються, не повинна бути більшою необхідного числа [11, 14].

Як правило, оцінювання параметрів моделі виконується з використанням МНК. Якщо дані нормально розподілені, то це еквівалентно оцінюванню за методом максимальної правдоподібності. Але при моделюванні необхідно до уваги брати те, що поведінку процесу слід апроксимувати з використанням рівнянь, а не описувати його до найменших дрібниць. В ході оцінювання зазвичай використовують відхилення змінних від середнього, не абсолютні їх значення, тобто

$$y(k) = Y(k) - \mu_y,$$

де $Y(k)$ – значення виміру,

μ_y – середнє значення ряду.

Для оцінювання параметрів моделей можна використати рекурсивні процедури; тоді поточне середнє може бути обчислене за рекурсивною формулою:

$$\mu_y(k) = \mu_y(k-1) + \frac{1}{k}[y(k) - \mu_y(k-1)].$$

Методи оцінювання параметрів моделі, що найбільш поширені, наступні: метод найменших квадратів (МНК); метод максимальної правдоподібності (ММП); зважений метод найменших квадратів (ЗМНК); метод моментів; методи Монте-Карло (ітеративні та неітеративні). Згідно МНК, оцінки параметрів моделей обчислюють за допомогою виразу:

$$\hat{\theta} = [X^T X]^{-1} X^T y,$$

де $\theta [p]$ – вектор оцінок параметрів вимірності p ;

$X[N \times p]$ – матриця вимірів;

$y[N]$ – вектор вимірів залежної змінної.

Слід зазначити, що обчислення елементів матриці вимірів для кожної конкретної моделі має свої особливості. В деяких випадках використання МНК призводить до отримання зміщених оцінок, так як він має чутливість оцінок до різких викидів, які зустрічаються у вихідних даних. До МНК альтернативним є метод максимальної правдоподібності (ММП), який являється один із основних, що застосовується для оцінювання параметрів моделі. Якщо похибка має нормальний

розподіл, то логарифмічна функція правдоподібності l , виходячи з n -спостережень, може бути представлена у вигляді:

$$-2l = n \log(2\pi\sigma^2) + \sum_{i=1}^n \frac{(y_i - \mu_i)^2}{\sigma^2}.$$

Для фіксованого σ^2 максимізація параметра l для відхилень μ є еквівалентною мінімізації суми квадратів $\sum (y - \mu)^2$. Тоді для лінійної моделі:

$$\eta_i = \mu_i = \sum_{j=1}^p x_{ij} \beta_j.$$

В останні роки для оцінювання невідомих параметрів моделей набув популярності метод Монте-Карло, що відносяться до класу УЛМ. Основна ідея цього методу шляхом введення неспостережуваних (згенерованих) величин у певні моменти часу полягає у рівномірному розбитті інтервалу спостереження даних $[t_i; t_{i+1}]$ на множину менших відрізків [11].

Алгоритм оцінювання параметрів моделей функціонує в такий спосіб:

- 1) генерування значення X_u^{i+1} з розподілу $P(X_u | X_0, \theta^i)$, де X_0 – експериментальні значення процесу;
- 2) генерування значення параметрів θ^{i+1} з розподілу $P(\theta | X_0, X_u^i)$.

Згенерований в такий спосіб статистичний ряд має граничний стаціонарний розподіл $P(X_u, \theta | X_0)$, який перетворюється в $P(\theta | X_0)$ за допомогою відомих арифметичних операцій. Недоліком методів Монте-Карло є велика ресурсоемність. Зазначимо, що методи цієї групи широко застосовуються завдячуючи їх універсальності, здатності урахувувати неспостережувані змінні, хорошій масштабованості, можливості застосування паралельних обчислень для прискорення процесу оцінювання, а також незначним похибкам оцінювання.

Відзначимо, що можна перевірити виконання вказаних умов тільки після оцінювання параметрів моделі, а до цього можна тільки постулювати їх виконання. Таким чином, після оцінювання параметрів моделі оцінка значень випадкового процесу буде визначатися похибками моделі: $\hat{\varepsilon}(k) = e(k) = y(k) - \hat{y}(k)$, що дозволяє виконати аналіз характеристик випадкового процесу $\{\varepsilon(k)\}$.

2.2.6 Діагностика моделей та вибір кращої з оцінених кандидатів

На останньому етапі виконується перевірка оцінених кандидатів на адекватність процесу тобто аналізується якість моделі. Діагностика моделей складається з наступних кроків:

1. Дослідження графіка похибок моделі $e(k) = y(k) - \hat{y}(k)$ у візуальний спосіб, де $\hat{y}(k)$ – оцінка змінної, яка отримана за допомогою побудованої моделі. На графіку не повинно бути довгих інтервалів суттєвої неадекватності значних викидів, на яких похибка приймає великі значення та довгих інтервалів [14].

2. Похибки моделі не повинні бути корельовані між собою. Для аналізу кореляційних зв'язків між значеннями похибок слід обчислити АКФ та ЧАКФ для ряду $\{e(k)\}$, і використовуючи Q – статистики визначити ступінь корельованості. Так, вважається Q – статистика несуттєвою до рівня 10%.

3. Перевірка статистичної значущості параметрів моделі, яка виконується з використанням статистики Стюдента або t – статистики, що використовується в статистичному сенсі для визначення значимості оцінки кожного коефіцієнта та обчислюється за виразом:

$$t = \frac{\hat{a} - a^0}{SE_{\hat{a}}},$$

де \hat{a} – оцінка коефіцієнта моделі;

a^0 – нуль-гіпотеза або початкова гіпотеза щодо цієї оцінки;

$SE_{\hat{a}}$ – стандартна похибка оцінки.

Щодо значимості оцінки за нуль-гіпотезу можна висувати наступну: коефіцієнт значущий, тобто $H_0 : a^0 \neq 0$, або не значущий, тобто $H_0 : a^0 = 0$. Згідно статистичної теорії перевірки гіпотез доцільно висувати нуль-гіпотезу, яка є протилежною бажаному результату. У нашому випадку бажаним результатом є значущість коефіцієнтів моделі та її адекватність. Таким чином, необхідно висувати нульову гіпотезу про незначущість коефіцієнтів, що дозволяє підійти коректно до

визначення значущості оцінок коефіцієнтів та спростити розрахунки.

4. Сума квадратів похибок повинна бути мінімальною для вибраної моделі:

$$\sum_{k=1}^N e^2(k) = \sum_{k=1}^N [\hat{y}(k) - y(k)]^2 \rightarrow \min_{\hat{\theta}},$$

у порівнянні з іншими моделями, де N – об'єм вибірки даних ($N \geq 100$).

5. Для оцінювання адекватності моделі можна використовувати інформаційний критерій Акайке:

$$AIC = N \ln \left(\sum_{k=1}^N e^2(k) \right) + 2n,$$

або критерій Байєса-Шварца:

$$BSC = N \ln \left(\sum_{k=1}^N e^2(k) \right) + n \ln(N),$$

де $n = p + q + 1$ – кількість параметрів моделі, що оцінюються з використанням статистичних даних (p – кількість параметрів моделі;

N – об'єм вибірки; q – число параметрів ковзного середнього).

6. Для визначення адекватності моделі в цілому можна застосувати F – статистику Фішера, що пропорційна відношенню:

з якої слідує, що якщо $R^2 \rightarrow 1$, то $F \rightarrow \infty$.

Порядок застосування F – статистики такий же, як і t – статистики. В даному випадку нуль-гіпотезою є припущення, що модель неадекватна в цілому.

Побудову адекватної математичної моделі процесу забезпечує коректне застосування методики Бокса-Дженкінса у випадку, коли експериментальні дані відповідають вимогам інформативності та представництва. Вимога інформативності означає, що вибірка в собі містить об'єм інформації, який достатній для оцінювання коефіцієнтів моделі. Вимога представництва означає, що вибірка охоплює досить довгий проміжок часу, щоб можна було відобразити поведінку необхідного режиму функціонування процесу, для якого будується модель.

Для формування статистичного висновку часто використовують байєсівський коефіцієнт, що представляє собою відношення апостеріорних ймовірностей до апріорних. Перевага однієї моделі над іншою визначається згідно значення байєсівського коефіцієнта.

Якщо значення байєсівського коефіцієнта суттєво більше одиниці, то приймається відповідне рішення щодо прийняття або відхилення моделі [14, 18].

2.3 Обробка екстремальних значень

Для аналізу та прогнозування актуарних процесів згідно четвертого етапа при визначенні структури математичних моделей передбачається обробка масивів екстремальних значень. Розглянемо обробку екстремальних значень. При обробці статистичного ряду з n – незалежних змінних X_1, \dots, X_n , які однаково розподілені, будемо виконувати наступну послідовність дій:

1) групування вибірок даних з n спостережень, які можуть містити від 50 до 100 значень;

2) для кожного i -го блоку визначення максимуму Z_i ;

3) наближення кожного i -го блоку максимумів до GEV -розподілу.

Як довжина блоку можна взяти величину першого року, але часто використовують для зручності дані максимуму i -го року Z_i .

Для кожного з річних максимумів, виконавши апроксимацію GEV -розподілом, розраховується функція квантилю [19, 20]:

$$z_p = \begin{cases} \mu - (\sigma / \xi)(1 - (-\log(1 - p))^{-\xi}), & \xi \neq 0; \\ \mu - \sigma \log(-\log(1 - p)), & \xi = 0; \end{cases}$$

Якщо припустимо, що $y_p = -\log(1 - p)$, тоді квантиль-функція прийме вигляд:

$$z_p = \begin{cases} \mu - (\sigma / \xi)(1 - (y_p)^{-\xi}), & \xi \neq 0; \\ \mu - \sigma \log(y_p), & \xi = 0. \end{cases}$$

При $\xi = 0$ графік буде мати лінійний характер, Якщо z_p зобразити в залежності від $\log(y_p)$. При $\xi < 0$ будемо мати випуклу криву з асимптотичною границею $(\mu - \sigma) / \xi$ при $p \rightarrow 0$, а при $\xi > 0$ будемо мати увігнутий графік без

кінцевої границі. Такий графік називають графіком повернення рівня (return level plot), який можна вважати інструментом або способом представлення згладженої моделі [20].

4. Оцінювання параметрів моделі та розв'язання задачі пошуку оптимальної довжини блоку. Остання може зводитися до пошуку співвідношення між величинами дисперсії та відхилення.

За умов, коли довжина блоків невелика, то наближення розподілів погане до границь, що може призвести до відхилень в екстраполяції та оцінюванні. коли довжина блоків велика, то породжуються значення зі значними оцінками дисперсії.

Для оцінювання параметрів моделей можна використати метод максимальної правдоподібності (ММП). Але при цьому при застосуванні ММП до *GEV*-розподілів не буде задоволена умова регулярності оцінювання, так як кінцева точка розподілів залежатиме від значення параметрів. Це говорить про те, що за методом максимальної правдоподібності стандартні асимптотичні результати аналізу недоцільно застосовувати до *GEV*-розподілів.

Логарифмічна функція правдоподібності для *GEV*-розподілів при $\xi \neq 0$ описується у вигляді:

$$l(\mu, \sigma, \xi) = -m \log \sigma - (1 + 1/\xi) \sum_{i=1}^m \log \left(1 + \xi \frac{z_i - \mu}{\sigma} \right) - \sum_{i=1}^m \left(1 + \xi \frac{z_i - \mu}{\sigma} \right)^{-1/\xi}$$

за умови, що $\left(1 + \xi \frac{z_i - \mu}{\sigma} \right) > 0$ при $i = 1, \dots, m$.

Як тільки ця умова не виконується, то функція правдоподібності стане рівною нулю і логарифмічна функція правдоподібності прийме значення нескінченності.

При розподілі Гумбела $\xi = 0$ і логарифмічна функція правдоподібності прийме вигляд:

$$l(\mu, \sigma) = -m \log \sigma - \sum_{i=1}^m \left(\frac{z_i - \mu}{\sigma} \right) - \sum_{i=1}^m \left(- \frac{z_i - \mu}{\sigma} \right). \quad (2.4)$$

Використовуючи методи чисельної оптимізації та максимізації виразу (2.4), можна отримати оцінку максимальної правдоподібності.

5. Графічна перевірка наближення GEV -моделей. Можна скористатись способами графічного аналізу даних, щоб обґрунтувати екстраполяції GEV -моделей.

Графік щільності розподілу, основу якого становить порівняння апроксимуючої та емпіричної функцій щільності розподілу. На графіку щільності розподілів абсциса точки являється емпіричною функцією розподілу, в яку підставляють дані з вибірки замість аргументу, а ординатою є це теоретична функція розподілу, в яку замість аргументу аналогічно підставляють дані із статистичної вибірки. В упорядкованому i -му блоці максимумів Z_i оцінюється функція емпіричного розподілу, яка має вигляд:

$$\tilde{G}_i(Z_i) = i/(m+1)$$

В тій же самій точці апроксимуюча функція щільності розподілу виглядає так:

$$\hat{G}(Z_i) = \exp \left\{ - \left(1 + \hat{\xi}_5 \left(\frac{z_{(i)} - \hat{\mu}}{\hat{\sigma}} \right) \right)^{-1/\hat{\xi}} \right\}.$$

Щоб отримати найкраще наближення моделі, необхідне виконання рівності $\tilde{G}(Z_i) = \hat{G}(Z_i)$. Використання цього графіка на практиці дозволяє запобігти ефекту «виродженості». Коли множина точок $\{\tilde{G}(Z_i), \hat{G}(Z_i)\}$, $i = 1, \dots, m$ знаходиться близько до першої діагоналі і обидві функції являються обмеженими біля одиниці, то значення абсциси z збільшуються.

Графік квантилів (Q-Q plot). Припущення про нормальність розподілу та наявність симетрії у розподілі являється недоліком класичної методології VaR при оцінюванні фінансових ризиків. На практиці фінансові ряди мають вироджений хвіст, а більшість економічних процесів асиметричні. Тому графік квантилів дозволяє оцінити певну ступінь довіри для параметричних моделей. Графік квантилів визначається через множину точок [19]:

$$\left\{ X_{k,n}, F^{-1} \left(\frac{n-k+1}{n} \right), k = 1, \dots, n \right\}.$$

Графік має лінійну форму у випадку, коли параметрична модель дає

прийнятне згладжування. Якраз графік дозволяє виконати порівняння оцінених моделей та вибрати найкращу модель, а також оцінити обрану модель по якості апроксимації хвоста емпіричного розподілу. У випадку, коли емпіричні дані мають вироджений хвіст, а ряд апроксимується нормальним розподілом, то графік квантилів характеризується кривою на дні лівого кінця розподілу, або на вершині правого кінця.

Крім розглянутих видів графічного аналізу можна використовувати графік рівня процесу (return level plot), а також середню функцію ексцесу (mean excess function) [19, 20].

6. Визначення порогу екстремального значення. Щоб забезпечити більш ефективний результат наближення екстремальних даних до одного з *GEV*-розподілів, можна застосувати так звані порогові моделі.

З умов, що множина статистичних даних перевищує деякий поріг u , а X_1, \dots, X_n являється послідовністю незалежних однаково розподілених змінних з функцією розподілу F . умовну ймовірність можна визначити так:

$$F_u(y) = P(X \leq u + y \mid X > u),$$

або

$$F_u(y) = \frac{F(u + y) - F(u)}{1 - F(u)}.$$

Використання цього виразу дозволяє визначити для великих значень порогу u ступінь наближення значень ймовірності.

Задача вибору оптимального порога може розглядатися як задача визначення розміру блока. Ці задачі спрямовані на визначення балансу між дисперсією та відхиленням. Високий рівень забезпечує велику дисперсію, а низький рівень може призвести до порушень асимптотичної апроксимації.

Вибору порогу ґрунтується на основі середнього GPD розподілу. Якщо Y у *GPD*-розподілі є випадковою змінною з параметрами σ і ξ , при $\xi < 1$, то математичне сподівання $E(Y) = \sigma / (1 - \xi)$. В інших випадках середнє GPD розподілу є нескінченністю.

Для визначення порогу екстремального значення можна використати метод умовно прийняттого вибору, який ґрунтується на наступному правилі: поріг встановлюється у тому регіоні, де хвіст від усієї вибірки становить 5-10%. При цьому він не повинен бути більшим ніж 10-15% [13].

7. Оцінювання невідомих параметрів моделі. Після визначення порогу необхідно виконати оцінку невідомих параметрів узагальненого розподілу Парето, для чого можна використати метод максимальної правдоподібності.

Якщо y_1, \dots, y_k є значення k -залишків з порогу; то логарифмічна функція правдоподібності при $\xi \neq 0$ має вигляд:

$$l(\sigma, \xi) = -k \log \sigma - (1 + 1/\xi) \sum_{i=1}^k \log(1 + \xi y_i / \sigma),$$

при $(1 + \xi y_i / \sigma) > 0$, а для інших випадків $l(\sigma, \xi) = -\infty$.

При $\xi = 0$ логарифмічна функція правдоподібності описується так:

$$l(\sigma) = -k (\log \sigma - \sigma^{-1} \sum_{i=1}^k y_i).$$

В якості іншого метода оцінювання невідомих параметрів можна використати байєсівський підхід, перевагою якого є його незалежність від регулярності припущень щодо характеру початкового розподілу при застосуванні до обробки екстремальних значень, як це потребує використання методу максимальної правдоподібності. Для випадків, коли необхідні припущення щодо необхідності застосування методу максимальної правдоподібності та ймовірності зважених моментів не виконуються, байєсівський підхід надає обґрунтовану альтернативу.

3 ЕКСПЕРЕМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

3.1 Обґрунтування та вибір середовища для моделювання

Перш ніж розпочати працювати із статистичними даними, необхідно обрати відповідне програмне середовище для аналізу та моделювання актуарних процесів. В теперішній час існує багато універсальних програмних засобів обробки та аналізу статистичної інформації, серед яких можна обрати наступні: STATGRAPHICS, STATISTICA, EViews, STADIA, MINITAB, EXCEL, SAS, STATA, а також мови програмування R-package, WinBUGs.

Для правильного вибору середовища необхідно в залежності від розв'язуваних задач виконати їх порівняльний аналіз, враховуючи також простоту та зручність їх виконання, вказуючи при цьому їх недоліки і переваги. Проведемо аналіз існуючих середовищ для моделювання.

Eviews надає особливо складний та тонкий інструментарій для обробки даних, який дозволяє виконувати регресійний аналіз і будувати прогнози у *Windows*-орієнтованому середовищі. З використанням цього програмного середовища можна досить швидко виявити статистичні залежності в даних, які треба аналізувати, а потім використовуючи залежності, що отримані, робити прогноз показників, які вивчаються.

Також, програма *Eviews* має зручний інтерфейс, проста в керуванні та інтерпретації результатів. Структура програми достатньо монолітна і в ній відсутня певна кількість модулів, які, наприклад, є в пакетах STATISTICA або SPSS. Це робить програму *Eviews* більш гнучкою у використанні. Крім цього, в програмі *Eviews* є широкий спектр методів і моделей економетричного аналізу, а саме:

- 1) методи ARCH, GMM, LS, NLS, Binary, Censored, Count, Ordered, ML та TSLS;
- 2) моделі ARIMA, LRM, Logit, ECM, GRM, Probit, Tobit, VAR, VECM та Pooled model;

Не дивлячись на свою простоту, графічні можливості програмного пакету

Eviews, забезпечують основні формати представлення даних, які необхідні для роботи аналітика.

Сфера застосування програмного пакету *Eviews* охоплює майже всі аспекти сучасної теорії та практики бізнесу. Пакет дозволяє працювати з різноманітними типами змінних, але його можливості найкраще розкриваються при розв'язанні задач прогнозування показників, які представлені у вигляді того чи іншого часового ряду. Високі функціональні можливості пакету *Eviews* при обробці кількісних змінних, дозволяють стверджувати про те, що він є надійним інструментом для прогнозування продажів акцій, динаміки розвитку як ресурсної бази, так і фінансово-економічних показників.

Програмний пакет *Eviews* включає в себе повний набір методів для виявлення та боротьби з типовими для поставлених вище задач проблемами, такими як:

- 1) автокореляція (DW, LM-test);
- 2) наявність коінтеграції та нестационарність (ADF, DF, cointegration test).

Пакет *Eviews* дозволяє працювати з такими типами даних: річні, напіврічні, квартальні, місячні, тижневі (7 днів), тижневі (5 днів), щоденні а також недетерміновані спостереження. При цьому допускається перехід від одного типу даних до іншого типу з використанням різних процедур екстраполяції та інтерполяції. Що стосується можливостей управління даними, то відзначимо, що *Eviews* підтримує *RATS*, *GiveWin*, *TSP* та *Areomos TSD* файли. Так можливим є імпорт/експорт даних із *ASCIIWK1*, *WK3*, , *XLS* та *TSD*. Пакет *Eviews* дозволяє виконувати та записувати макроси, що може скоротити терміни виконання рутинної роботи.

Пакет *Eviews* дозволяє отримувати прогноз відразу після побудови моделі. Можна робити апріорні оцінки адекватності моделі, застосовуючи існуючі в пакеті «захисні» інформаційні критерії *log likelihood*, *PACF*, *ACF*, та ін. Для перевірки можливостей моделей прогнозувати додається можливість застосування механізму крос-перевірки. Таким чином, вказані особливості моделювання явищ і процесів та економетричного середовища аналізу дозволяють пакету *Eviews* ефективно вирішувати поставлені задачі, а також робити необхідні висновки.

Що стосується *R-package*, то це мова програмування, яку можна використовувати для статистичної обробки даних та працювати з графіками, а також імплементувати в рамках проекту GNU в інтегроване середовище обчислень з відкритим вихідним кодом. Широко використовується *R-package* в якості статистичного програмного забезпечення для аналізу даних, для чого він включає в себе:

- 1) середовище для ефективного зберігання даних;
- 2) набір операторів для операцій з матрицями та масивами;
- 3) великий, комплексний та когерентний інструмент для аналізу даних, які є проміжними;
- 4) графічні можливості щодо візуального аналізу статистичних даних;
- 5) мову для ефективного програмування, яка включає реалізацію різних функцій, зокрема умовні оператори, рекурсивні функції, цикли.

Розглянемо програмні пакети *BUGS* і *WinBUGS*. Пакет *WinBUGS* є спеціалізованим програмним забезпеченням, яке використовується для аналізу методів Монте-Карло для марковських ланцюгів.

Пакет *BUGS* (Bayesian inference Using Gibbs Sampling) являється аббревіатурою, яка розшифровується як Байєсівський висновок із застосуванням вибірки Гіббса.

Пакет *BUGS* є вузько спеціалізованим програмним пакетом, який застосовується для проведення повного аналізу імовірнісної моделі на базі МСМС, де всі невизначеності розглядаються як випадкові величини. Мова програмування *BUGS* є легкою для розуміння, а також дозволяє користувачеві задавати пряму специфікацію ймовірнісної моделі, виконувати калібрування моделей, оцінювати параметри та обчислювати оцінки прогнозів.

З погляду страховиків програмне забезпечення пакета *BUGS* і *WinBUGS* ще далеке від досконалості, оскільки воно не підтримує існуючий спектр аналізу ризиків, що популярні, частотні розподіли у явному вигляді, а також обробляє округлені або згруповані дані в довільний спосіб. Разом з тим, пакети *BUGS* і *WinBUGS* є дуже зручними інструментами, які дозволяють реалізувати байєсівський

аналіз для широкого спектру даних та побудову актуарних моделей, що робить їх використання більш доступними для в широкого класу актуарії.

По результатам розглянутих переваг та недоліків вище вказаних програмних середовищ для аналізу та моделювання актуарних процесів, а також враховуючи особливості задачі побудови моделей прогнозування страхових виплат, коли настає страховий випадок, та задачі оцінювання величини операційного ризику СК, було вибрано економетричний пакет *Eviews* та програмне середовище *R-package*.

3.2 Опис та структура статистичних даних

Експериментальні дослідження ефективності запропонованої методики виконувалися з використанням фактичних статистичних даних. Об'єм статистичної вибірки складав 247 вимірів, що включають наступні змінні: назва страхової компанії; статистичний рік; грошовий еквівалент страхових виплат; страхові платежі; кількість договорів, які уклала страхова компанія; кількість страхових випадків за рік.

Основною залежною змінною є страхові виплати, які при настанні страхового випадку відображають здійснення грошових переказів. Інші змінні, які включені до вибірки, являються незалежними та вважаються як фактори. Структура статистичних даних представлена в таблиці 3.1.

Також використовуються такі показники як частота страхових випадків і темпи зростання капіталу СК в якості додаткових незалежних змінних. Обидві вказані змінні не корелюють із залежною змінною страхові виплати і вимірюються у %.

Для реалізації попереднього аналізу статистичних даних застосовувались наступні програмні продукти: Microsoft Excel 2016; економетричний пакет Eviews 8.0 для попереднього оцінювання невідомих параметрів та побудови моделей; інструментальне середовище програмування R2.9.2 для роботи з графікою та статистичної обробки даних.

Таблиця 3.1 – Структура статистичних даних

№	Назва змінної	Значення
1.	Загальний розмір вибірки	247 точок;
2.	Статистичний рік	Діапазон значень, 2009-2022 рр.
3.	Назва страхової компанії	Текстова змінна;
4.	Страхові виплати	Поточна сума страхових виплат, грн;
5.	Кількість страхових договорів за статистичний рік	Ціле число, починаючи з 1;
6.	Страхові платежі	Поточна сума страхових платежів, грн;
7.	Кількість страхових випадків за рік	Ціле число, починаючи з 1;
8.	Частота страхових випадків	Величина задана у %
9.	Темпи зростання капіталу	Величина задана у %

В пакеті Eviews 8.0 використані модулі розрахунку описових статистик, побудови моделі та метод максимальної правдоподібності для оцінювання параметрів моделі. В інструментальній середовищі програмування R2.9.2 інтегровані модулі Rcmdr, evdbayes, extRemes та mcmcPack.

3.3 Попередня обробка даних

На рисунку 3.1 показано графік залежності страхових виплат від статистичного року. Видно, що мають місце різкі зміни величини «*Страхові виплати*», які можна пояснити певними коливаннями величини «*Кількість страхових випадків*» для відповідного періоду. Слід зазначити, що у страховій діяльності через відсутність відповідних грошових коштів часто мають місце не виплати страхових премій при настанні страхових випадків. Тому у графіку з наближення до нуля видно різкі коливання.

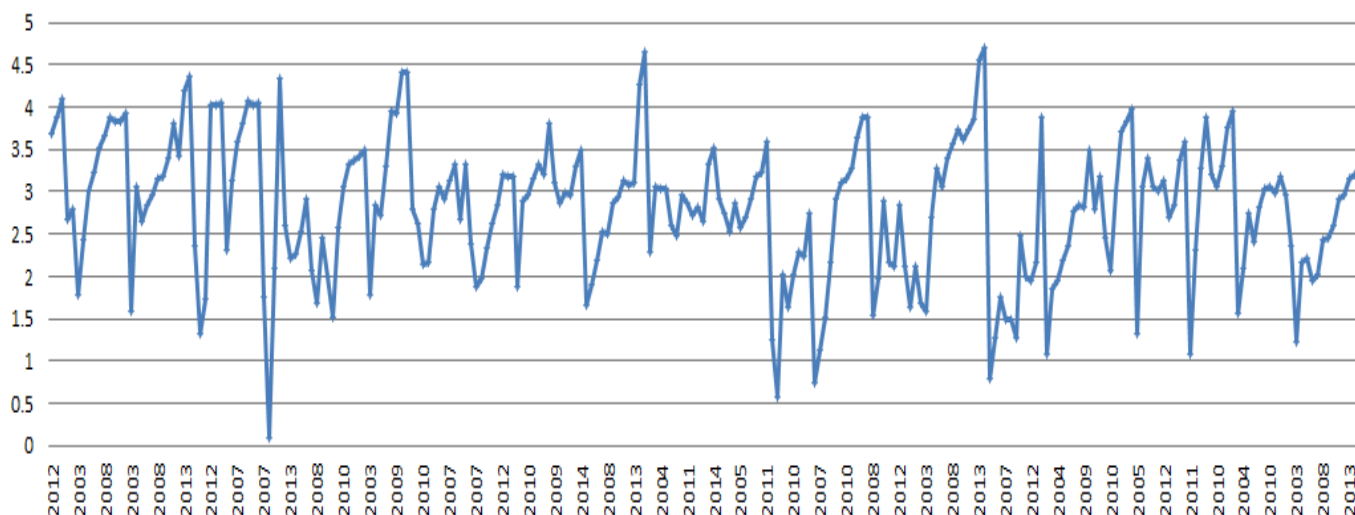


Рисунок 3.1 – Динаміка страхових виплат

На рисунку 3.2 показані значення описових статистик, в яких коефіцієнт асиметрії (Skewness) змінюється в межах від 2,839 до 8,664, що свідчить про наявність «правого хвосту» в розподілі змінних. Оскільки значення параметру ексцеса (Kurtosis) більше трьох, то це говорить про те, що розподіл є гостровершинним.

View	Proc	Object	Print	Name	Freeze	Sample	Sheet	Stats	Spec
				Q_CASES		Q_ARRANG		DAMAGES	CHARGES
Mean				248.9717		42819.83		2779.190	47154.60
Median				65.00000		2431.000		780.0000	13017.40
Maximum				3800.000		2241084.		49575.00	557884.0
Minimum				2.000000		6.000000		1.200000	469.8000
Std. Dev.				526.4061		181679.5		6121.650	81935.75
Skewness				4.241486		8.663552		4.644020	2.838897
Kurtosis				24.51794		94.69488		29.05691	12.32584
Jarque-Bera				5505.864		89621.68		7875.493	1226.855
Probability				0.000000		0.000000		0.000000	0.000000
Sum				61496.00		10576499		686459.9	11647185
Sum Sq. Dev.				68167427		8.12E+12		9.22E+09	1.65E+12
Observations				247		247		247	247

Рисунок 3.2 – Описові статистики початкових даних

Проведений попередній аналіз початкових даних вказує на сильну виродженість вибірки, що при побудові моделі проявляється у вигляді шуму

(рисунок 3.3). Саме тому, необхідний пошук методів обробки масивів екстремальних даних та прийняття рішення про необхідність попереднього логарифмування даних.

Для розв'язання задачі виродженості статистичних даних можна використовувати ТЕЗ. Відомо, що задачею теорії екстремальних значень являється напрулений аналіз та оцінювання ймовірностей появи випадкових величин, які пов'язані з екстремальними або рідкісними подіями.

Так як екстремальні значення являються нефіксованими величинами, це тоді нові випадкові величини, які можуть залежити від об'ємів вибірки та типу вихідного розподілу. Наприклад, в області страхування ймовірною і рідкісною подією може бути настання страхового випадку, яка повинна супроводжуватись виплатами страхових премій.

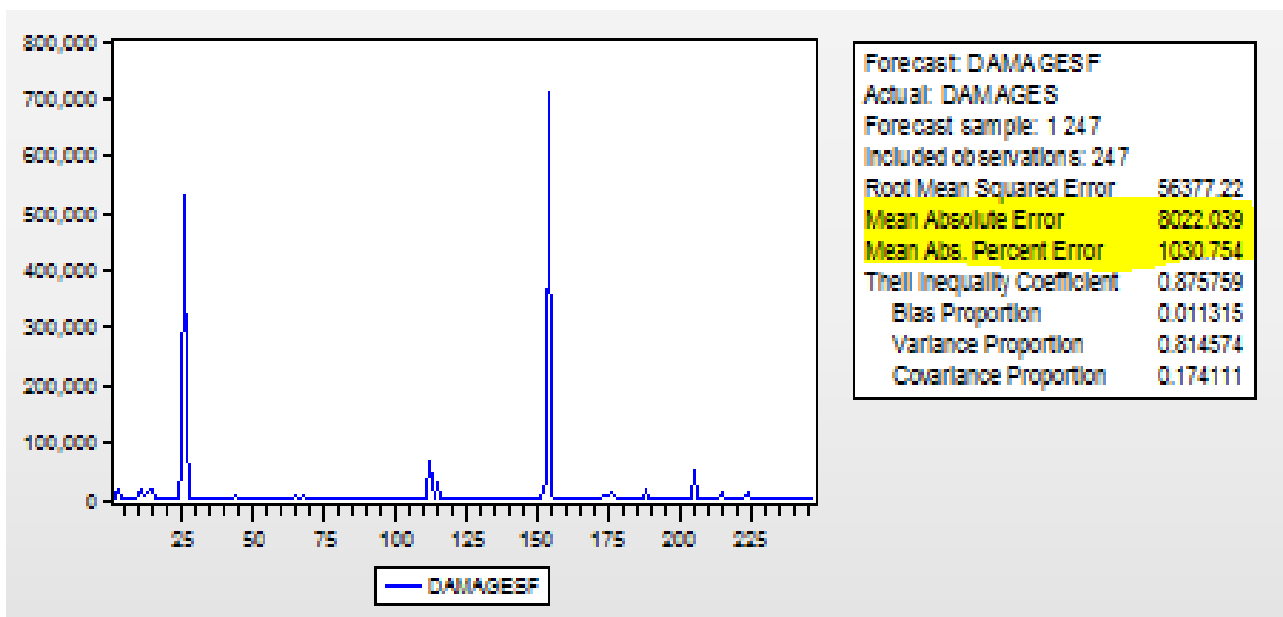


Рисунок 3.3 – Результати оцінювання моделі без попередньої обробки

Візуальний аналіз логарифмованих даних та аналіз описової статистики та, наведених на рис. 3.4, дозволяють припустити про наближення даних до *GEV*- або *GPD*-розподілів.

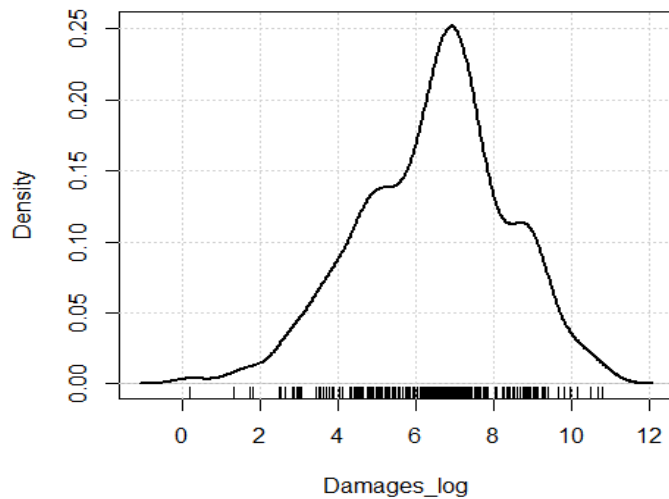


Рисунок 3.4 – Графік залежності логарифмованих страхових виплат від щільності розподілу

Щоб зменшити зміщення моделі, вибирається відносно високий поріг, що буде означати, що лише небагато досліджень використовується для оцінювання параметрів розподілу. Це гарантує збільшення оцінки дисперсії. Метою вибору величини порогу являється уникнення зміщення моделі. Для експерименту значення величини порогу прийнято рівним 6,65. Наведений на рис. 3.5 графік Mean Residual Life Plot для оціненої моделі відображає залежність порогу від середнього залишку, який використовується як важель перевірки обраного порогу. З рисунку 3.5 видно, що з'являються помітні відхилення від лінійності після значення порогу 6.

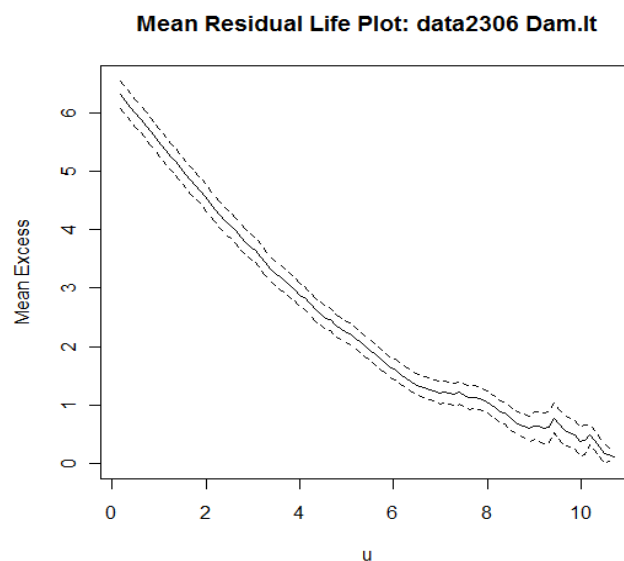


Рисунок 3.5 – Залежність значення порогу від середнього залишку моделі

В таблиці 3.2. подана порівняльна характеристика параметрів розподілу, яка показує, що оптимальним є наближення даних з використанням *GPD*-розподілу, яке має незначну похибку та характеризується максимальним наближенням емпіричної кривої до теоретичної функції розподілу, що приведено на рисунку 3.6.

Результати оцінювання побудованої моделі з використанням *GEV*-розподілів зображено на рис. 3.7. Порівнюючи графіки щільності розподілу для актуальної вибірки даних та побудованої моделі видно, що є помітні значні покращення моделі при належності до одного з *GEV*-розподілів.

Таблиця 3.2 – Порівняльна характеристика параметрів розподілів

Тип розподілу	Sigma		Xi		Log-likelihood	Exceedance rate (per year)	Number of exceedances of threshold
	Maximum likelihood estimation	Std. error	Maximum likelihood estimation	Std. error			
GEV-розподіл	1,953	0,712	-0,650	0,095	487,812	–	–
GPD-розподіл	0,777	0,346	-0,541	0,206	146,369	183,364	124

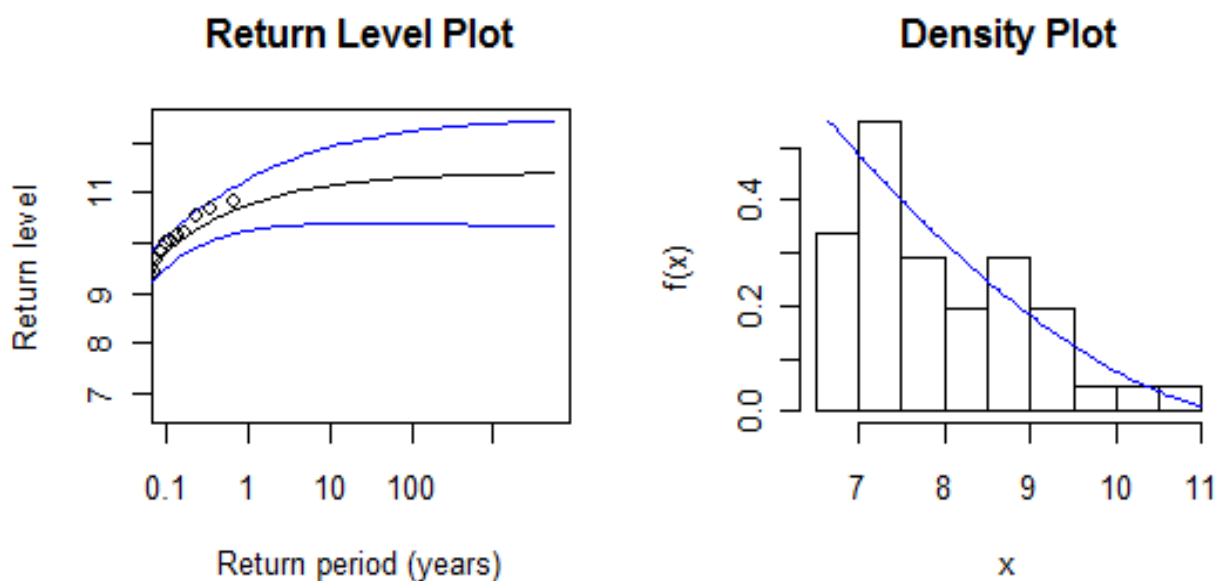


Рисунок 3.6 – Графічне представлення оціненої *GPD*-моделі

Виконавши порівняння результатів отриманих оцінок зауважимо, що байєсівський підхід має кращі результати, ніж метод максимальної правдоподібності. Це дозволяє обґрунтовано вибрати кращу модель із запропонованих GEV-розподілів, виходячи з алгоритмів вибору кращої моделі, а також зі значень апіорних параметрів.

Також вдалося успішно розв'язати задачу невиродженості даних у статистичній вибірці із застосуванням ТЕЗ та завдяки використанню запропонованої комплексної моделі обробки екстремальних статистичних даних.

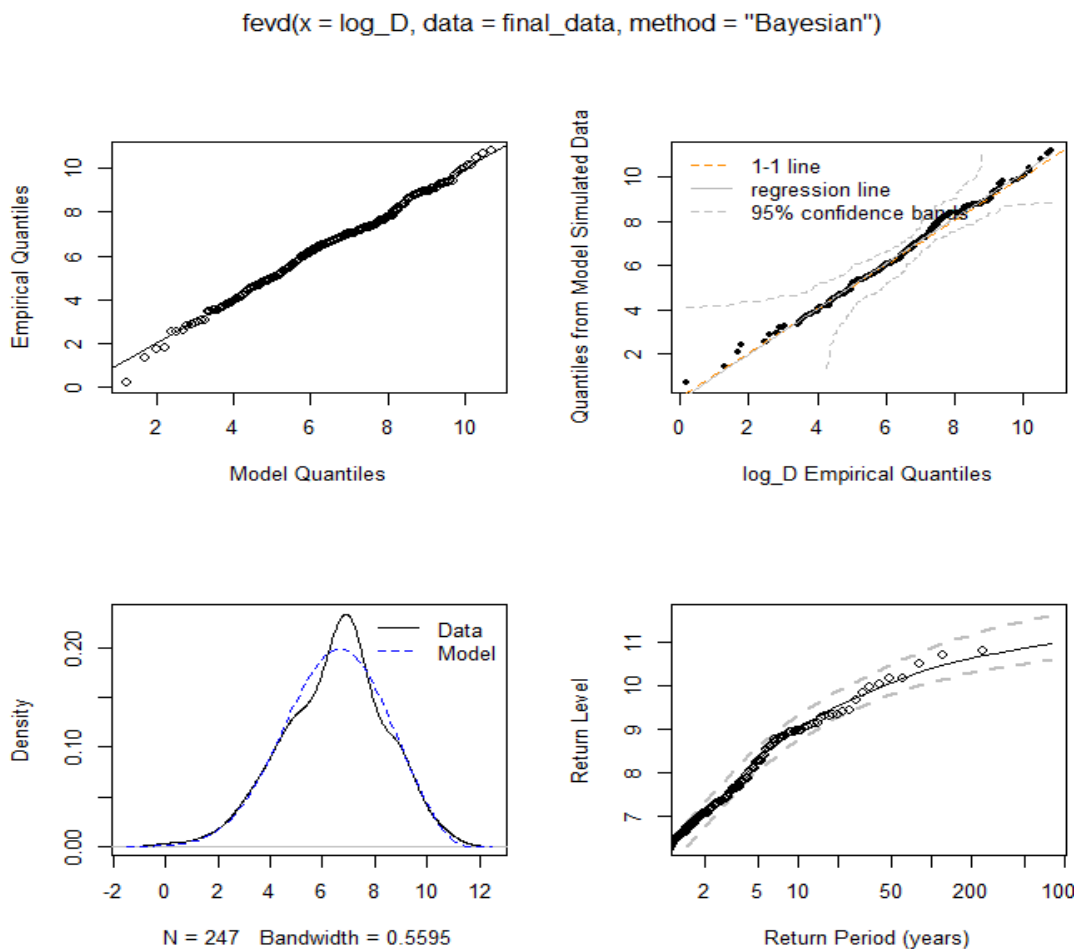


Рисунок 3.7 – Діагностика наближення моделі до одного з GEV-розподілів

Для оцінювання невідомих параметрів моделей, що побудовані і які належать до класу GEV-розподілів, можна використовувати успішно байєсівський підхід, оперуючи алгоритмами вибору кращої моделі, а також апіорними та апостеріорними розподілами параметрів.

Таким чином, використання нових комбінованих методів для оцінювання невідомих параметрів та вибору кращої моделі на основі алгоритмів зменшення порогів викидів дозволяє отримати додаткові можливості щодо дослідження особливостей методів математичного моделювання.

3.4 Прогнозування актуарних процесів за допомогою побудованих моделей

Для прогнозування величини страхових виплат СК у разі настання страхового випадку були використані сумарні річні дані за 2008-2023 рр. – страхові виплати, кількість страхових договорів, страхові платежі, темпи зростання капіталу у СК, частота появи страхових випадків.

Показники, які були включені до моделі, являються як залежними, так і незалежними, як впливовими так і не впливовими на результативну змінну «Виплати». Ось чому для дослідження залежності або незалежності була використана кореляційна матриця, що наведена на рисунку 3.8.

Correlation					
	DAMAGES	CHARGES	FREQUENC...	INS_ARRANG	INVEST_RATE
DAMAGES	1.000000	0.688255	0.391304	0.222812	-0.357355
CHARGES	0.688255	1.000000	0.848295	0.789533	-0.621592
FREQUENCY_CASES	0.391304	0.848295	1.000000	0.907957	-0.355276
INS_ARRANG	0.222812	0.789533	0.907957	1.000000	-0.436885
INVEST_RATE	-0.357355	-0.621592	-0.355276	-0.436885	1.000000

Рисунок 3.8 – Кореляційна матриця

З використанням кореляційної матриці, наведеної на рис. 3.8 встановлено, що між змінними «Виплати» та «Страхові платежі» існує залежність, що визначається значенням коефіцієнта кореляції, який дорівнює 0,6883. Зв'язок з від'ємним значенням коефіцієнта кореляції в кореляційній матриці свідчить про наявність оберненого зв'язку, тобто чим більші страхові виплати, тим менші темпи зростання страхового капіталу. Тому, змінну «Темпи зростання страхового капіталу СК» з від'ємними показником коефіцієнта кореляції не будемо брати до уваги при прогнозуванні. Разом зтим, інші змінні не є взаємозалежними.

3.5 Результати експериментальних досліджень

Практичне застосування розглянутої методики аналізу та прогнозування актуарних процесів дозволило побудувати 6 моделей, кожна з яких відрізняється розподілами залежної змінної (страхові виплати) та видом функції зв'язку.

Для подальшого аналізу якості моделей, що побудовані, у таблиці 3.4 надані характеристики кожного з досліджень з урахуванням таких показників: середнє відхилення, стандартне відхилення, інформаційні критеріїв, відносні похибки та ін.

Для визначення того, яка із побудованих моделей є найкращим наближенням експериментальних даних до реальних процесів, були проаналізовані два класи інформаційних критеріїв:

1) критерії, що базуються на мінімізації відстані (міра відхилення моделі від реального об'єкту);

2) критерії, що базуються на байєсівському підході.

На відміну від критеріїв, що базуються на мінімізації відстані та дозволяють обирати із заданої множини моделей найбільш адекватну, критерії Байєса з певною ймовірністю вибирають «правдиву» модель, а ймовірність вибору «точної» моделі при збільшенні розміру вибірки прагне до 1.

Наступною важливою задачею аналізу актуарних процесів після оцінювання втрат, пов'язаних із страховими виплатами за полісами страхування, є оцінювання операційного ризику. Для розв'язання цієї задачі необхідно будувати модель у формі мережі Байєса.

3.6 Вибір та обґрунтування кращої моделі

Відповідно до проведених експериментальних досліджень отримані характеристики побудованих моделей наведені в таблиці 3.3.

Після побудови моделей, на основі інтеграції модулів `extRemes`, `glm` та `bestglm`, середовища програмування R package і з використанням припущень щодо початкового розподілу залежної змінної, встановлено наступне.

Таблиця 3.3 – Характеристики побудованих моделей

Критерії / моделі	Роки	Гамма-логарифмічна	Гамма-тотожна	Пуассона-квадратична	Пуассона-логарифмічна	Обернено-гаусівсько-тотожна	Гауса-логарифмічна
	2008	5956.567	1596.049	1903.869	14916.1	1587.371	24009.31
	2009	11861.295	3904.354	6281.569	16694.01	4425.157	26919.5
	2010	16877.358	22079.792	15667.493	21740.06	17926.272	34320.73
	2011	20923.09	29895.616	21654.081	25026.9	24726.812	37800.2
	2012	22771.274	22403.975	18771.973	24340.78	21963.72	33729.53
	2013	54390.238	74479.367	68189.807	52402.72	61641.887	64950.78
	2014	35443.742	76409.957	60514.376	48730.2	60285.107	65039.51
	2015	38672.291	78164.192	63422.485	51836.74	63034.589	65412.27
	2016	64922.577	93430.554	88418.321	73063.17	79040.921	76376.65
	2017	97963.514	86874.603	90318.79	82214.12	81797.671	66440.83
	2018	118690.588	101936.093	108544.809	114782.69	98590.685	72200.03
	2019	540956.727	200852.822	325641.113	346712.41	158528.148	324616.07
	2020	468048.315	196525.368	311938.275	308718.25	153129.636	315676.9
	2021	23943.557	40046.852	28821.339	28910.16	32357.339	42790.09
	2022	287910.00	163050.00	193736.00	199781.00	51820.00	266978.00
	2023	249921.00	309798.00	173053.00	176268.00	354236.00	137645.00
Середнє значення		108672.94	73471.40	86434.88	86434.88	61359.67	89305.89
Стандартне відхилення		171411.80	62670.63	104065.74	106348.93	49938.30	99396.10
Інформаційний критерій Акайке (ІКА)		23.84	325.00	Inf	Inf	324.58	25.30
Байесівський інформаційний критерій (БІК)		BICq equivalent for q in (0.0028, 0.6772)	BICq equivalent for q in (0.0293, 0.7131)	CBD	CBD	CBD	BICq equivalent for q in (0.4989, 0.7687)
Відносна похибка		3.14%	2.67%	2.42%	2.65%	2.82%	3.14%
Величина ризику втрат СК		11.27%	6.09%	8.60%	8.79%	5.81%	7.95%
Скорингова функція Фішера		4	10	5	5	6	13

Для зменшення величини ризику в побудованих моделях бажано застосувати наступні заходи:

- 1) проведення превентивних заходів, які спрямована на зменшення ризику;
- 2) застосування заходів щодо локалізації ризиків, наприклад, обмежити поліс страхування до оптимального терміну, який прорахований наперед та обґрунтований на практиці;
- 3) використовувати в критичних ситуаціях заходи, які орієнтовані на самофінансування ризиків, тобто покриття ризиків за рахунок власних коштів;
- 4) керуватися результатами оцінки критерію можливої випадковості при передачі ризику страховикам, а також критеріями однозначності, незалежності та оцінюваності у грошових одиницях.

Застосовуючи запропоновані заходи можна досягти максимізації прибутків від страхової діяльності та мінімізувати витрати на відшкодування при виникненні страхових випадків із забезпеченням вчасної виплати страхових премій.

ВИСНОВКИ

У ході виконання дослідження розв'язано актуальну задачу побудови узагальнених лінійних моделей для прогнозування та оцінювання ризиків актуарних процесів, що дозволило підвищити адекватність моделей актуарних процесів та якість оцінок прогнозів у разі настання страхових випадків або операційних втрат страхових компаній.

По результатам виконаної роботи можна зробити такі висновки:

1. Розглянуто актуальність задачі побудови адаптивних математичних моделей для прогнозування та оцінювання ризиків актуарних процесів та проведено аналіз протікання фінансових процесів у страхуванні. Виконано аналіз методів і моделей оцінювання фінансових ризиків у страхуванні, зокрема, розглянуті моделі, що засновані на аналізі наслідків; моделі, що засновані на аналізі факторів ризику, метод функціональних кореляцій, регресійні моделі, та методи експертного оцінювання, що свідчить про актуальність використання методів інтелектуального аналізу даних для розв'язання поставленої задачі.

2. Встановлено, що для розв'язання задачі аналізу та прогнозування фінансових ризиків можна використати універсальні логічні моделі, які дозволяють врахувати як безпосередні залежності від факторів ризику рівня ризику, так і залежності між факторами ризику.

3. Досліджені методики побудови для актуарних процесів довільної природи математичних моделей, які базуються на використанні узагальнених лінійних моделей та теорії екстремальних значень. Проаналізовані характеристики ефективності математичних моделей, що використовуються для аналізу актуарних процесів.

4. Побудови математичної моделі аналізу та прогнозування фінансових ризиків складається з таких етапів: формулювання мети дослідження; виділення структурних чи функціональних елементів, що відповідають поставленій меті, та визначення їх якісних характеристик; встановлення взаємозв'язків між елементами

моделі та їх опис; оцінювання статистичних характеристик побудованої моделі та визначення її адекватності; отримання результатів та їх аналіз.

5. Для моделювання та обробки фінансових статистичних даних використано економетричний пакет *Eviews* та середовище програмування *R-package*.

6. Проведені експериментальні дослідження, які показали ефективність застосування запропонованої моделі обробки екстремальних статистичних даних із застосуванням відповідних методів вибору моделі для оцінювання невідомих параметрів, обробки екстремальних даних та виконання математичного моделювання.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Вовчак О.Д. Страхування. Навч. посібник. – Львів.: «Новий світ – 2000».
2. Гаманкова О.О. Фінанси страхових організацій: Навч. посіб. – КНЕУ, 2007. – 328 с.
3. Шевчук О.В. Управління фінансовими ресурсами страховика // Економічний вісник Донбасу. – 2010. - №2. – С. 116-126.
4. Tripp et al М.Н. “Quantifying operational risk in general insurance companies”, in Giro Working Party [Presented to the Institute of Actuaries], 22 March 2004. – 137 p.
5. International Convergence of Capital Measurement and Capital Standards. A Revised Framework. Comprehensive Version. Switzerland, Basel: Basel Committee on Banking Supervision, Bank for Int. Settlements, 2006. – 158 p.
6. Sound Practices for the Management and Supervision of Operational Risk. Switzerland, Basel: Basel Committee on Banking Supervision, Bank for Int. Settlements, 2003. – 110 p.
7. Medova E. “Extreme Value Theory (Extreme values and the measurement of operational risk)”, Operational Risk, pp. 17—36, July 2000.
8. Scenario-based AMA (2003). [Online]. Available: <http://www.newyorkfed.org/newsevents/events/banking/2003/con0529d.pdf>.
9. B. Dobeli et al., “From operational risk to operational excellence”, in Advances in operational risk: Firm-wide issues for financial institutions, ed. P. Mestchian, 2nd ed. UK, London: Risk Books, Risk Water Group, 2003.
10. R. Kuhn and P. Neu, “Functional Correlation Approach to Operational Risk in Banking Organizations”, Physica A, no. 322, 2003. – pp. 650—660.
11. Трухан С. В., Бідюк П. І. Методика побудови математичних моделей актуарних процесів / С. В. Трухан, П. І. Бідюк // Східноєвропейський журнал передових технологій, №14(73), 2015.
12. Tsay R. S. Analysis of financial time series. – New York : Wiley & Sons, Inc. – 2002. – 448 p.

13. 86. Beirlant J. *Statistics of extremes: Theory and application*. – New York: John Wiley & Sons, Inc. – 2004. – 505 p.
14. Бідюк П.І., Романенко В.Д., Тимощук О.Л. *Аналіз часових рядів*. — К.: Політехніка, 2013. — 600 с.
15. Coles S. *An Introduction to Statistical Modeling of Extreme Values* / S. Coles // – London: Springer-Verlag. – 2001. – pp. 45–104.
16. Smith R. L. . *An overview of Extreme value theory.*/ R. L. Smith. – Lausanne: Bernoulli Center, 2009.
- 17 Gill J. *Generalized linear models: a unified approach*. – USA: New Delli, 2001. – 110 p.
18. Трухан С. В., Бідюк П. І. *Прогнозування актуарних процесів за допомогою узагальнених лінійних моделей* / С. В. Трухан, П. І. Бідюк // *Наукові вісті НТУУ «КПІ»*, №2, 2014.
19. R.H. Shumway and D.S. Stoffer, *Time series analysis and its applications*. New York: Springer, 2006. – 598 p.
20. Mallor F. *An introduction to statistical modeling of extreme values [Text]* / F. Mallor E. Nualart, E. Omeu // *Hub research paper*. – 2009. – No.36. – pp. 5-31.
21. Бланк І. А. *Управління фінансовими ризиками*. К.: Ніка-Центр, 2014. 600 с.
22. Клапків М.С. *Страховання фінансових ризиків: Монографія*. – Тернопіль: Економічна думка, Карт-бланш, 2002. – 570 с.
23. *Управління ризиками банків: монографія у 2 томах. Т. 2: Управління ринковими ризиками та ризиками системних характеристик* / [А.О. Єпіфанов, Т.А. Васильєва, С.М. Козьменко та ін.] /За ред. Д-ра екон. наук, проф. А.О. Єпіфанова і д-ра екон. наук, проф. Т.А. Васильєвої. — Суми: ДВНЗ "УАБС НБУ", 2012. – 299 с.
24. Бобиль В. В. *Фінансові ризики банків: теорія та практика управління в умовах кризи: монографія*. Дніпропетровськ: Дніпропетр. нац. ун-т залізн. трансп. ім. акад. В Лазаряна, 2016. – 298 с.
25. Gengjian Xue, Li Song, Jun Sun. *Foreground estimation based on linear regression model with fused sparsity on outliers*, *IEEE transactions on circuits and systems for video technology*. 2013. – pp. 1346-1357.

26. Bermingham, M.L. et al. Application of high-dimensional feature selection: evaluation for genomic prediction in man. *Sci. Rep.* 5, 10312; doi: 10.1038/srep10312 (2015).
27. Urbanowicz RJ, Meeker M, La Cava W, Olson RS, Moore JH. Relief-based feature selection: Introduction and review. *J Biomed Inform.* 2018; 85:189-203. doi:10.1016/j.jbi.2018.07.014.
28. Белевець М. С., Бідюк П. І., Кузнецова Н. В. Розробка системного підходу до менеджменту ризиків. *Системні науки та кібернетика.* 2015. №1. С. 30 – 48. URL: http://mmsa.kpi.ua/sites/default/files/ssc/issues/ssc_1_2015.pdf.
29. Tymoshchuk O. L. A combined approach to modeling nonstationary heteroscedastic processes / O. L. Tymoshchuk, V. H. Huskova, P. I. Bidyuk. // *Radio Electronics, Computer Science, Control.* – 2019. – №2. – С. 80–89.
30. Гуськова В. Г. Розробка сценарного підходу на основі моделей інтелектуального аналізу даних / В. Г. Гуськова, П. І. Бідюк // *Наукові праці Донецького національного технічного університету. Серія : Інформатика, кібернетика та обчислювальна техніка.* - 2016. –№2. - С. 158-164.
31. Бідюк П.І., Кузнецова Н.В. Основні етапи побудови і приклади застосування мереж Байєса // *Системні дослідження та інформаційні технології.* – 2007. – № 4. – С.26–39.