

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Міністерство освіти і науки України

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Міністерство освіти і науки України

Кваліфікаційна наукова
праця на правах рукопису

ЛЕВЕНЧУК ЛЮДМИЛА БОРИСІВНА

УДК 004.896

ДИСЕРТАЦІЯ
СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ АНАЛІЗУ
ОПЕРАЦІЙНИХ РИЗИКІВ

124 – Системний аналіз
12 – Інформаційні технології

Подається на здобуття наукового ступеня доктора філософії
Дисертація містить результати власних досліджень. Використання ідей,
результатів і текстів інших авторів мають посилання на відповідне джерело
_____ Л.Б. Левенчук

Науковий керівник
Бідюк Петро Іванович, доктор технічних наук, професор.

Київ – 2023

АНОТАЦІЯ

Левенчук Л.Б. Система підтримки прийняття рішень для аналізу операційних ризиків. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії з галузі знань 12 Інформаційні технології за спеціальністю 124 Системний аналіз. – Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, 2023.

Операційний ризик стає все більш актуальним у різних сферах, включаючи фінанси, медицину, транспорт і виробництво. Він представляє собою ризик збитків, спричинений невдалими внутрішніми процесами, людьми, системами або зовнішніми подіями, і може мати серйозний фінансовий та репутаційний вплив на організацію. Система підтримки прийняття рішень (СППР) для аналізу операційного ризику є ключовим інструментом для використання в різних організаціях. Вона дає можливість виявити потенційні ризики, оцінити їхню ймовірність та вплив на протікання процесів, а також визначити оптимальні стратегії для їхнього пом'якшення.

Процеси, які не відносяться до стаціонарних та мають нелінійний характер, є загально поширеними і мають різноманітні застосування у демографії, екології, економіці, фінансах та інших сферах. Ці процеси характеризуються великою кількістю невизначеностей, таких як невизначеність в даних, невизначеність в структурі моделі та параметрична невизначеність. Невизначеність у всіх цих аспектах може впливати на точність та надійність системи підтримки прийняття рішень і призводити до неточних результатів аналізу даних та прогнозування.

Для ефективного управління нестаціонарними та нелінійними процесами необхідно використовувати адаптивні підходи, які дозволяють системі підтримки прийняття рішень адаптувати свою структуру та

параметри до змінних умов. Загальна класифікація нелінійних та нестационарних процесів допомагає розуміти природу цих процесів у різних сферах та визначати належні методи аналізу та прогнозування. Усвідомлення наявності різних видів невизначеностей і їхнього впливу на результати аналізу даних є важливим етапом у розробці надійних та точних системи підтримки прийняття рішень.

Метою дисертаційного дослідження є створення та програмна реалізація у формі інтелектуальної СППР системної методології побудови математичних моделей для формального опису і прогнозування розвитку нелінійних нестационарних фінансово-економічних процесів за допомогою ймовірісно-статистичного підходу, регресійного аналізу і методів інтелектуального аналізу даних (ІАД) з використанням інтелектуальної СППР.

В дисертації отримані такі нові наукові результати:

1. Розроблені **нові** моделі у формі байєсівської мережі (БМ) і байєсівської регресії, які відрізняються можливістю врахування ймовірісно-статистичних невизначеностей даних і забезпечують підвищення якості оцінювання можливих операційних втрат.

2. **Вперше** застосовано метод структурно-параметричної адаптації моделей у формі БМ до нових даних (повторне оцінювання структури і параметрів моделі з використанням заданої множини критеріїв) з метою підвищення адекватності ймовірісних моделей.

3. **Удосконалено** алгоритми обчислювальних процедур, що використовуються для оцінювання структури і параметрів моделей (удосконалення методу Монте-Карло для марковських ланцюгів) операційних ризиків.

Практична цінність дисертаційної роботи полягає у тому, що **розроблено оригінальну** СППР на основі принципів системного аналізу для розв'язання задачі моделювання, оцінювання і прогнозування можливих втрат внаслідок реалізації операційних ризиків на основі

статистичних даних та експертних оцінок, розроблених математичних моделей і множини статистичних критеріїв якості.

Всі результати роботи доведено до практичного інженерного рівня і впроваджено у навчальний процес інституту прикладного системного аналізу НТУУ «КПІ імені Ігоря Сікорського» та у фінансових компаніях з метою застосування методології оцінювання ринкових ризиків на практиці. Запропонована методологія дає змогу обґрунтовано підійти до розв'язання задачі прогнозування розвитку сучасних фінансових процесів, оцінювання можливих фінансових ризиків та їх врахування у виробленні тактичних і стратегічних управлінських рішень. Всі теоретичні і практичні результати дисертаційної роботи у повній мірі опубліковано у фахових вітчизняних та закордонних наукових виданнях, що входять до відповідного встановленого переліку, а також виконано їх належну апробацію на міжнародних наукових конференціях і семінарах.

У дисертаційній роботі розглянуто питання актуальності дослідження операційного ризику та розробки системи підтримки прийняття рішень для його аналізу. Встановлено, що правильна оцінка прогнозів є надзвичайно важливим завданням, оскільки невірні рішення при оцінюванні прогнозів можуть призвести до серйозних матеріальних втрат. Проведено дослідження методів боротьби з невизначеністю та опрацьоване питання розробки ефективних стратегій обробки даних для моделювання, прогнозування і оцінювання ризиків в нестаціонарних та нелінійних процесах.

Зроблено огляд існуючих математичних моделей для формального опису операційних фінансових ризиків. Розроблено ймовірнісно-статистичну модель операційного ризику у формі динамічної байєсівської мережі (БМ) і подано приклад її застосування до оцінювання операційного ризику у страхуванні. Розроблено комбіновану модель: лінійна + нелінійна регресія для операційного ризику, яка будується на основі статистичних даних.

Для побудови моделей у формі БМ та регресії використано теоретичні положення, що стосуються аналізу статистичних/експериментальних даних з метою оцінювання структури і параметрів такої моделі. Формування (обчислення) ймовірнісного висновку виконувалось за допомогою теореми Байєса.

Розроблено структуру та складові елементи інтелектуальної СППР для аналізу операційних ризиків. Створено метод моделювання операційних ризиків на основі комбінованої моделі (ймовірнісні фільтри + регресія + динамічна байєсівська мережа).

Запропоновано вдосконалену системну багатокрокову методологію побудови моделей фінансових процесів і фінансових ризиків довільного походження. Наданий детальний огляд теоретичних основ і методів адаптації байєсівських мереж до даних в контексті аналізу ймовірнісних розподілів для дискретних випадкових величин і динамічних систем. Наведено приклад застосування методології, який демонструє її ефективність в застосуванні до аналізу ризиків актуарних процесів. Застосування запропонованої методології для моделювання фінансових процесів з використанням узагальнених лінійних моделей та байєсівського оцінювання параметрів гарантує високу якість оцінювання ризиків з мінімальними похибками.

Сформульовано метод моделювання ризиків на основі системного підходу до побудови моделей операційних ризиків та його застосування при побудові системи підтримки прийняття рішень. Створено модель операційного ризику на основі узагальненої лінійної регресійної (байєсівської) моделі. Побудовано (розширення функцій) інтелектуальну СППР для аналізу операційних ризиків.

За матеріалами дисертації опубліковано 14 робіт, з яких 6 – це статті у журналах і збірниках наукових праць, що входять до переліку фахових видань затверджених МОН України за спеціальністю дисертації або у періодичних виданнях іноземних держав (1 з них включена до міжнародної

наукометричної бази Scopus, 1 до Web of Science), та 8 – публікації у матеріалах конференцій (у тому числі, міжнародних).

Ключові слова: операційні ризики, система підтримки прийняття рішень, адаптивна байєсівська мережа, математична модель, прогнозування, теорема Байєса, нелінійні нестационарні процеси.

SUMMARY

Levenchuk L. Decision support system for operational risk analysis. – Qualification scientific work, the manuscript.

PhD thesis in the field of knowledge 12 Information Technology in specialty 124 System Analysis – National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, 2023.

Operational risk is becoming increasingly relevant in various fields, including finance, medicine, transportation, and manufacturing. It is the risk of loss caused by failed internal processes, people, systems, or external events, and can have a serious financial and reputational impact on an organization. A decision support system (DSS) for operational risk analysis is a key tool for an organization. It allows to identify potential risks, assess their likelihood and impact, and determine the best strategies to mitigate them.

Non-stationary and nonlinear processes are widespread and have various applications in demography, ecology, economics, finance, and other fields. These processes are characterized by a large number of uncertainties, such as uncertainty in the data, uncertainty in the model structure, and parametric uncertainty. Uncertainty in all these aspects can affect the accuracy and reliability of a decision support system and lead to inaccurate data analysis and forecasting results.

For effective management of non-stationary and nonlinear processes, it is necessary to use adaptive approaches that allow the decision support system to adapt its structure and parameters to changing conditions. A general classification of nonlinear and nonstationary processes helps to understand the nature of these processes in different areas and to determine the appropriate methods of analysis and forecasting. Awareness of the presence of different types of uncertainties and their impact on the results of data analysis is an important step in developing a reliable and accurate decision support system.

The aim of the thesis is to create and programmatically implement in the form of an intelligent DSS a systematic methodology for building mathematical models for the formal description and forecasting of the development of nonlinear non-stationary financial and economic processes using a probabilistic-statistical approach, regression analysis and methods of data mining using intelligent DSS.

The following new scientific results were obtained in the dissertation:

1. **New** models have been developed in the form of a Bayesian network (BN) and Bayesian regression, which are distinguished by the ability to take into account probabilistic and statistical uncertainties of data and provide an improvement in the quality of estimating possible operational losses.

2. **For the first time**, the method of structural and parametric adaptation of models in the form of a BN to new data (re-estimation of the structure and parameters of the model using a given set of criteria) is applied to improve the adequacy of probabilistic models.

3. The algorithm of computational procedures used to evaluate the structure and parameters of models (improvement of the Monte Carlo method for Markov chains) of operational risks **has been improved**.

The practical significance of the thesis is that an **original** DSS based on the principles of system analysis **has been developed** to solve the problem of modeling, assessing and forecasting possible losses due to the realization of operational risks based on statistical data and expert opinions, developed mathematical models and a set of statistical quality criteria.

The results of the thesis fulfillment have been brought to a practical engineering level and implemented in the educational process of the MMSA Department of Institute for Applied System Analysis (IASA) at the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", and in financial companies to apply the methodology for assessing market risks in real life. The proposed methodology provides the reasonable approach to the task of forecasting the development of modern financial processes, assessing

possible financial risks and taking them into account in the development of tactical and strategic management decisions.

All theoretical and practical results of the thesis were published in professional domestic and foreign scientific journals included in the relevant list, as well as their proper approbation at international scientific conferences and seminars.

The thesis addresses the issue of the relevance of the study of operational risk and the development of a decision support system for its analysis. It is established that the correct assessment of forecasts is an extremely important task, since incorrect decisions in assessing forecasts can lead to serious material losses. The methods of dealing with uncertainty are studied and the issue of developing effective data processing strategies for modeling, forecasting and risk assessment in non-stationary and nonlinear processes is addressed.

The existing mathematical models for the formal description of operational financial risks are reviewed. A probabilistic-statistical model of operational risk in the form of a dynamic Bayesian network (BN) is developed and an example of its application to the assessment of operational risk in insurance is presented. A combined model is developed: linear + nonlinear regression for operational risk, which is based on statistical data.

Theoretical provisions related to the analysis of statistical/experimental data are used to build models in the form of BN and regression in order to assess the structure and parameters of such a model. The formation (calculation) of a probabilistic inference was performed using Bayes theorem.

The structure and constituent elements of an intelligent DSS for operational risk analysis were developed. The method for modeling operational risks based on a combined model (probabilistic filters + regression + dynamic Bayesian network) was created.

The improved systemic multi-step methodology for building models of financial processes and financial risks of arbitrary origin is proposed. A detailed review of the theoretical foundations and methods of adaptation of Bayesian

networks to data in the context of analyzing probability distributions for discrete random variables and dynamic systems is provided. The example of application of the methodology is given, which demonstrates its effectiveness in the analysis of risks of actuarial processes. The application of the proposed methodology for modeling financial processes using generalized linear models and Bayesian parameter estimation guarantees high quality risk assessment with minimal errors.

The method of risk modeling based on a systematic approach to building operational risk models and its application in building a decision support system is formulated. The operational risk model based on a generalized linear regression (Bayesian) model is created. An intelligent DSS for operational risk analysis has been built (its functions extended).

Keywords: operational risks, decision support system, adaptive Bayesian network, mathematical model, forecasting, Bayes theorem, nonlinear nonstationary processes.

List of principal publications of the applicant:

1. Л.Б. Левенчук, П.І. Бідюк Байєсівський аналіз даних у моделюванні і прогнозуванні нелінійних нестационарних процесів, KPI Science News, 2020, No. 3, P. 14-23.

Здобувачкою запропоновано підхід для моделювання задачі оцінювання прогнозів фінансового процесу на фондовому ринку для опису еволюції цін на акції для відомої компанії. Побудована комплексна модель з використанням байєсівської мережі надає можливість удосконалити ймовірнісне оцінювання прогнозів при здійсненні торговельних операцій на фондовому ринку.

2. Л.Б. Левенчук, В.Г. Гуськова, П.І. Бідюк Ймовірнісне моделювання операційних ризиків, KPI Science News, 2021, No. 3, P. 26-37.

Здобувачкою описано методи дослідження операційних ризиків (ОР) і їх модифікації. Запропоновано застосування методу моделювання для

побудови моделі ОР страхового шахрайства. Функціонування моделі продемонстровано на прикладі.

3. N.V. Kuznietsova, O.M. Trofymchuk, P.I. Bidyuk, O.M. Terentiev, L.B. Levenchuk, Bayesian modeling of risks of various origin, KPI Science News, 2021, No. 4, P. 7-18.

Здобувачкою запропоновано застосування підходу до моделювання операційного ризику, пов'язаного з некоректною обробкою інформації. Для побудови та застосування моделі до оцінки ризику було проаналізовано проблему ризику, обрано відповідні змінні та оцінено попередні умовні ймовірності. Функціонування побудованих моделей було продемонстровано на наочних прикладах.

4. Kuznietsova N. V., Huskova V. H., Bidyuk P. I., Matsuki Y., Levenchuk L. B. Modelling risk factors interaction and risk estimation with copulas, Radio Electronics, Computer Science, Control, 2022, N. 2, P. 44-52.

В роботі запропоновано концепцію побудови розподілів за допомогою копули, для створення підґрунтя для розв'язання задач прогнозування можливих збитків та прийняття відповідних рішень щодо управління ризиками. Отримані результати будуть корисними для подальших теоретичних досліджень, а також для практичних застосувань у сфері ризик-менеджменту.

5. Kasyanov, P., Levenchuk, L.B. Formalization and Development of Autonomous Artificial Intelligence Systems. In: Zgurovsky, M., Pankratova, N. (eds) System Analysis and Artificial Intelligence. Studies in Computational Intelligence, 2023, vol 1107. Springer, Cham, P. 153-164.

У роботі запропонована концепція формалізації задачі розробки автономних систем штучного інтелекту та запропонована математично обґрунтована методологія для побудови наближених ϵ -оптимальних стратегій із заданою точністю.

6. Levenchuk L. The Bayesian approach to analysis of financial operational risk, ScienceRise, 2022, No. 2, P. 11–20.

Здобувачкою проаналізовані основні методи побудови байєсівських мереж для моделювання операційного ризику в умовах невизначеності а також запропонована методологія у вигляді послідовності дій, необхідних для створення моделі у вигляді мережі.

ЗМІСТ

АНОТАЦІЯ.....	2
SUMMARY	7
ЗМІСТ	13
ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ.....	16
ВСТУП.....	17
РОЗДІЛ 1. АКТУАЛЬНІСТЬ ДОСЛІДЖЕННЯ І МЕТОДИ ОЦІНЮВАННЯ ОПЕРАЦІЙНИХ РИЗИКІВ	22
1.1 Актуальність дослідження операційного ризику та розробки системи підтримки прийняття рішень для його аналізу	22
1.2 Огляд існуючих моделей для аналізу та оцінювання операційного та фінансового ризику	30
1.2.1 Методологія байєсівського програмування	31
1.2.2 Value at risk (VaR).....	33
1.2.3 Метод Монте-Карло	35
1.2.4 Метод внутрішніх вимірів (Internal Measurements Approach) ...	40
1.2.5 Scenario-based Advanced Measurement Approach (Sb-AMA)	41
1.2.6 Байєсівська модель з автоматичним оцінюванням невизначеності.....	42
1.3 Програмні інструменти побудови системи підтримки прийняття рішень. Огляд програмних засобів.....	42
1.3.1 Мова програмування Python	47
1.3.2 Мова програмування R.....	50
1.4 Висновки до розділу.....	55
РОЗДІЛ 2. МОДЕЛІ ОПЕРАЦІЙНИХ РИЗИКІВ	57

	14
2.1 Невизначеності у моделюванні фінансових процесів	58
2.2 Байєсівське програмування.....	65
2.3 Узагальнені лінійні моделі.....	73
2.4 Регресійні моделі на основі статистичних даних	78
2.5 Моделювання нелінійних нестационарних процесів	82
2.5.1 Деякі моделі нелінійних нестационарних процесів.....	84
2.6 Комбіновані моделі	85
2.7 Метод Монте-Карло для Марківських Ланцюгів	87
2.8 Висновки до розділу	94
РОЗДІЛ 3: СТРУКТУРНО-ПАРАМЕТРИЧНА АДАПТАЦІЯ	
БАЙЄСІВСЬКИХ МОДЕЛЕЙ.....	96
3.1 Побудова адаптивних байєсівських моделей	96
3.2 Адаптація байєсівської мережі до даних	99
3.3 Навчання мережі для випадку повних вибірок даних	103
3.3.1 Навчання ймовірнісної мережі	104
3.3.2 Оцінки Байєса-Діріхле (БДе) для ймовірнісної мережі.....	106
3.3.3 Оцінки Байєса-Діріхле (БДе) для динамічної ймовірнісної мережі	109
3.4 Навчання на практиці	112
3.4.1 Навчання для випадку наявності неповних вибірок даних	113
3.5 Байєсівська фільтрація нелінійних даних	118
3.5.1 Модель системи даних	118
3.5.2 Методи гранулярної (particle) фільтрації	119
3.5.3 Основний алгоритм фільтрації	122
3.5.4 Фільтр вибірки за значимістю з відсівом	125

	15
3.5.6 Регуляризований гранулярний фільтр	129
3.6 Висновки до розділу	133
РОЗДІЛ 4. Створення системи підтримки прийняття рішень для моделювання і використання обчислювальних експериментів.....	135
4.1 Архітектура і функції системи підтримки прийняття рішень.....	135
4.2 Побудова функцій прогнозування в СППР	142
4.3 Оцінювання адекватності моделей і якості прогнозів.....	146
4.3.1 Адаптивне оцінювання прогнозів.....	148
4.4 Приклади оцінювання операційних ризиків	150
4.5 Висновки до розділу	158
ВИСНОВКИ	160
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	162

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

БМ – байєсівська мережа

ГФ – гранулярний фільтр

МКМЛ – метод Монте-Карло для марковських ланцюгів

ММ – математична модель

ММП – метод максимальної правдоподібності

МНК – метод найменших квадратів

ННП – нелінійні нестационарні процеси

ОР – операційний ризик

ПММ – приховані марковські моделі

САПП – середня абсолютна похибка у процентах

СКП – середньоквадратична похибка

СППР – система підтримки прийняття рішень

УЛМ – узагальнені лінійні моделі

CDF – cumulative distribution function

DBN – Dynamic Bayesian Networks

IMA – Internal Measurements Approach

LDA – Loss Distribution Approach

MCMC – Markov Chain Monte Carlo

PRA – Probabilistic Risk Assessment

VaR – Value at risk

IMA – Internal Measurements Approach

ВСТУП

Актуальність роботи. Проблема правильної оцінки прогнозів втрат є першочерговим завданням багатьох досліджень у всьому світі. Невірні рішення з оцінки прогнозів призводять до значних матеріальних втрат. Розробка методології, результатом якої є побудована адекватна математична модель (ММ) для визначення ризику фінансових та інших видів процесів і використання цих моделей для оцінювання необхідних прогнозів, є однією з ключових проблем розвитку методів аналізу даних.

Дослідження операційного ризику (ОР) стає все більш актуальним у різних сферах, включаючи фінанси, медицину, транспорт і виробництво. Він представляє собою ризик збитків, спричинений невдалими внутрішніми процесами, людьми, системами або зовнішніми подіями, і може мати серйозний фінансовий та репутаційний вплив на організацію. Аналіз операційного ризику є актуальною задачею у багатьох галузях. ОР – це ризик збитків, спричинених неадекватними або невдалими внутрішніми процесами, людьми чи системами, або зовнішніми подіями.

Розробка системи підтримки прийняття рішень (СППР) для аналізу ОР допоможе організаціям більш ефективно виявляти, оцінювати та управляти цими ризиками. Вона дозволяє виявити потенційні ризики, оцінити їхню ймовірність та вплив, а також визначити оптимальні стратегії для їхнього пом'якшення. Використовуючи моделі та алгоритми, засновані на фактичних даних, СППР можуть допомогти фінансовим організаціям виявити потенційні ризики, оцінити їхню ймовірність та вплив, а також визначити найкращі стратегії для їхнього пом'якшення.

Мета і задачі наукового дослідження. Метою дисертаційної роботи є створення та програмна реалізація у формі інтелектуальної СППР системної методології побудови математичних моделей для формального опису і прогнозування розвитку нелінійних нестационарних фінансово-економічних процесів за допомогою ймовірнісно-статистичного підходу,

регресійного аналізу і методів інтелектуального аналізу даних (ІАД) з використанням інтелектуальної СППР.

Для досягнення поставленої мети необхідно розв'язати такі задачі:

- Проаналізувати наявні методи моделювання, оцінювання і прогнозування операційних ризиків за допомогою СППР.
- Сформулювати перелік особливостей протікання фінансових процесів.
- Запропонувати методику моделювання і прогнозування операційних ризиків на основі статистичних даних і з використанням ймовірнісних методів.
- Провести побудову математичних моделей операційних ризиків, оцінювання і прогнозування можливих операційних втрат на основі статистичних даних та експертних оцінок.
- Створити СППР для виконання обчислювальних експериментів, комбінування моделей, тестування функцій системи.
- Запропонувати практичне застосування розробленої методики.

Об'єкт дослідження: операційні ризики та нелінійні нестационарні процеси в економіці та фінансах.

Предмет дослідження:

- Методи інтелектуального аналізу даних (байєсівські мережі (БМ), байєсівська регресія, фільтри байєсівського типу, комбіновані моделі).
- Статистичний і регресійний аналіз: тобто методи, які використовуються для аналізу даних: теорія оцінювання; перевірка гіпотез; статистичні критерії якості оцінок параметрів і моделей.
- Теорія оцінювання: методи оцінювання структури і параметрів математичних моделей, оцінювання змінних на фоні шумів, оцінювання прогнозів, ризиків можливих втрат тощо.

Методи дослідження: методи теорії ймовірностей, математичної статистики, теорії прийняття рішень, теорія байєсівських ймовірнісно-статистичних рішень, регресійний аналіз та дерева рішень.

Наукова новизна роботи. Наукова новизна дисертаційної роботи полягає у наступному:

1. Розроблені **нові** моделі у формі байєсівської мережі (БМ) і байєсівської регресії, які відрізняються можливістю врахування ймовірно-статистичних невизначеностей даних і забезпечують підвищення якості оцінювання можливих операційних втрат.

2. **Вперше** застосовано метод структурно-параметричної адаптації моделей у формі БМ до нових даних (повторне оцінювання структури і параметрів моделі з використанням заданої множини критеріїв) з метою підвищення адекватності ймовірнісних моделей.

3. **Удосконалено** алгоритми обчислювальних процедур, що використовуються для оцінювання структури і параметрів моделей (удосконалення методу Монте-Карло для марковських ланцюгів) операційних ризиків.

Обґрунтованість і достовірність наукових результатів забезпечується **розробкою оригінальної СППР** на основі принципів системного аналізу для розв'язання задачі моделювання, оцінювання і прогнозування можливих втрат внаслідок реалізації операційних ризиків на основі статистичних даних та експертних оцінок, розроблених математичних моделей і множини статистичних критеріїв якості.

Практичне значення отриманих результатів. Результати роботи доведено до практичного інженерного рівня і впроваджено у навчальний процес інституту прикладного системного аналізу НТУУ «КПІ імені Ігоря Сікорського» та у фінансових компаніях з метою застосування методології оцінювання ринкових ризиків на практиці. Запропонована методологія дає змогу обґрунтовано підійти до розв'язання задачі прогнозування розвитку сучасних фінансових процесів, оцінювання можливих фінансових ризиків та їх врахування у виробленні тактичних і стратегічних управлінських рішень. Всі теоретичні і практичні результати дисертаційної роботи у повній мірі опубліковано у фахових вітчизняних та закордонних наукових

виданнях, що входять до відповідного встановленого переліку, а також виконано їх належну апробацію на міжнародних наукових конференціях і семінарах.

Особистий внесок користувача. Всі основні результати дисертаційної роботи отримані автором особисто. 5 фахових публікацій написано у співавторстві, одна публікація одноосібна. Здобувачці належать наступні результати. В роботі [1] здобувачкою запропоновано підхід для моделювання задачі оцінювання прогнозів фінансового процесу на фондовому ринку для опису еволюції цін на акції для відомої компанії. Побудована комплексна модель з використанням байєсівської мережі надало можливість удосконалити ймовірнісне оцінювання прогнозів при здійсненні торговельних операцій на фондовому ринку. В роботі [2] здобувачкою описано методи дослідження операційних ризиків (ОР) і їх модифікації. Запропоновано застосування методу моделювання для побудови моделі ОР страхового шахрайства. В роботі [3] Здобувачкою запропоновано застосування підходу до моделювання операційного ризику, пов'язаного з некоректною обробкою інформації. Для побудови та застосування моделі до оцінки ризику було проаналізовано проблему ризику, обрано відповідні змінні та оцінено попередні умовні ймовірності. Функціонування побудованих моделей було продемонстровано на наочних прикладах. В роботі [4] запропоновано концепцію побудови розподілів за допомогою копули, для створення підґрунтя для розв'язання задач прогнозування можливих збитків та прийняття відповідних рішень щодо управління ризиками. У роботі [5] запропонована концепція формалізації задачі розробки автономних систем штучного інтелекту та запропонована математично обґрунтована методологія для побудови наближених ϵ -оптимальних стратегій із заданою точністю. У публікації [6] Здобувачкою проаналізовані основні методи побудови байєсівських мереж для моделювання операційного ризику в умовах невизначеності а також

запропонована методологія у вигляді послідовності дій, необхідних для створення моделі у вигляді мережі.

Апробація результатів дисертації. Результати та основні положення роботи подавалися та обговорювалися на: 15-й Міжнародній науково-технічній конференції (НТК) «Проблеми інформатизації» у Парижі, 2020 [7,8]; 24-му Міжнародному молодіжному форумі «Радіoeлектроніка і молодь у ХХІ столітті», м.Харків, 7-9 квіт. 2020 р. [9]; 10-й Міжнародній НТК «Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління», м. Баку, м. Харків, м. Жиліна, 9-10 квіт. 2020 р. [10]; 2020 IEEE 2nd International Conference on System Analysis & Intelligent Computing (SAIC), 5-9 Oct. 2020 [11]; 25-му Міжнародному молодіжному форумі «Радіoeлектроніка і молодь у ХХІ столітті», м.Харків, 20-22 квіт. 2021 р. [12]; 8-й Міжнародній НТК «Сучасні методи, інформаційне, програмне та технічне забезпечення систем керування організаційно-технічними та технологічними комплексами», Київ, 25-26 листопада 2021 р. [13]; 6-й Міжнародній НТК «Прикладні системи та технології в інформаційному суспільстві», 30 вересня 2022 [14].

Публікації. За матеріалами дисертації опубліковано 14 робіт, 6 з яких – статті у журналах і збірниках наукових праць, що входять до переліку фахових видань затверджених МОН України за спеціальністю дисертації, 2 статті – у виданнях, проіндексованих у базах даних Scopus і Web of Science, або у періодичних виданнях іноземних держав, та 8 публікацій у матеріалах конференцій (у тому числі, міжнародних).

Структура та обсяг дисертації. Дисертація складається із анотації, вступу, чотирьох розділів, висновків, списку використаних джерел. Робота містить 168 сторінок, у тому числі: 140 сторінок основного тексту, 13 рисунків, 10 таблиць, список використаних джерел із 75 найменувань на 7 сторінках.

РОЗДІЛ 1. АКТУАЛЬНІСТЬ ДОСЛІДЖЕННЯ І МЕТОДИ ОЦІНЮВАННЯ ОПЕРАЦІЙНИХ РИЗИКІВ

1.1 Актуальність дослідження операційного ризику та розробки системи підтримки прийняття рішень для його аналізу

Проблема правильної оцінки прогнозів є першочерговим завданням багатьох досліджень у всьому світі. Невірні рішення стосовно оцінювання прогнозів призводять до значних матеріальних втрат. Розробка методології, результатом якої є побудована адекватна математична модель для визначення ризику фінансових та інших видів процесів і використання цих моделей для оцінки необхідних прогнозів, є однією з ключових проблем розвитку методів аналізу даних [15].

Дослідження операційного ризику є актуальним для багатьох галузей, зокрема для фінансової, медичної, транспортної та виробничої сфер. Операційний ризик – це ризик збитків, спричинених неадекватними або невдалими внутрішніми процесами, людьми чи системами, або зовнішніми подіями. Ці ризики можуть мати значний фінансовий та репутаційний вплив на організацію.

Розробка системи підтримки прийняття рішень для аналізу операційних ризиків може допомогти організаціям більш ефективно виявляти, оцінювати та управляти цими ризиками. Використовуючи моделі та алгоритми, засновані на даних, системи підтримки прийняття рішень можуть допомогти організаціям виявити потенційні ризики, оцінити їхню ймовірність та вплив, а також визначити найкращі стратегії для їхнього пом'якшення. Більше того, оскільки бізнес стає складнішим, а обсяг даних, які він генерує, зростає, стає все більш важливим використовувати передові методи аналітики, щоб отримати уявлення про операційні ризики.

Важливим етапом у вирішенні цього завдання є застосування системного підходу до оцінки прогнозів, що передбачає побудову

спеціалізованої системи підтримки прийняття рішень. Машинне навчання, штучний інтелект та байєсівські підходи можуть допомогти бізнесу зрозуміти сенс даних, виявити закономірності та аномалії, а також приймати більш обґрунтовані рішення щодо управління ризиками. Застосування ймовірісно-статистичних методів необхідно розглядати як важливу трудомістку частину процесу прийняття рішень, яка дозволяє приймати обґрунтовані тактичні та стратегічні рішення. Найпоширеніші з них включають:

- Теорему Байєса: процедура обчислення умовних ймовірностей, що використовується для оновлення переконань на основі нових даних і широко застосовується в байєсівській статистиці та прийнятті рішень.

- Дерева рішень: графічне представлення процесу прийняття рішень, яке відображає різні доступні варіанти, ймовірності, пов'язані з кожним результатом, та очікувану корисність і цінність кожного рішення.

- Марковські процеси: математичні моделі для прийняття рішень в умовах невизначеності, де результати кожного рішення залежать від попередніх рішень і поточного стану системи.

- Моделювання за методом Монте-Карло: обчислювальна техніка, яка використовує випадкову вибірку для моделювання можливих результатів рішення. Моделювання за методом Монте-Карло часто використовується для оцінки ймовірності різних результатів і визначення найкращого способу дій.

- Лінійну та нелінійну регресію: статистичний метод для моделювання зв'язку між залежною змінною та однією або декількома незалежними змінними. Регресійний аналіз також може бути використаний для виявлення закономірностей у даних і прогнозування майбутніх результатів.

Ці методи ґрунтуються на глибокому аналізі наявної інформації, а також на знаннях, досвіді та інтуїції фахівців (експертні оцінки). Ймовірнісний аналіз процесів, подій і даних різних типів передбачає два підходи [16–18]:

- частотний, який ґрунтується на ймовірнісному аналізі за класичним підходом – дані накопичуються у процесі виконання експериментів (або збору статистики) і обробляються так званими частотними методами теорії ймовірностей;

- байєсівський, в основу якого покладається той чи інший варіант теореми Байєса; цей підхід не виключає використання класичних методів, а вхідна інформація для аналізу може бути подана у вигляді статистичних (експериментальних) даних, експертних оцінок, окремих фактів і т. ін.

Коректне використання ймовірнісно-статистичних методів надає значні переваги стосовно прийняття коректних рішень практично в усіх напрямках людської діяльності: в конкуренції за підвищення якості та продаж нової продукції, дає можливість отримати високоякісні оцінки прогнозів, обґрунтувати фінансові, макроекономічні, а також рішення стосовно раціонального ведення домашнього господарства і розв'язати багато інших задач. Послідовність реалізації етапів розв'язування задачі моделювання і подальше застосування побудованої моделі при використанні ймовірнісно-статистичних методів залишається такою ж як і при використанні інших методів, наприклад, регресійного аналізу [19 – 21].

Весь процес моделювання і подальше застосування побудованої моделі складається з таких етапів:

- вивчення теоретичних основ можливих методів моделювання, придатних для розв'язування задач конкретного типу;
- вибір методу (методів) моделювання;
- збір даних і побудова моделей-кандидатів, що належать до вибраного класу моделей; вибір кращої моделі з множини побудованих кандидатів за статистичними критеріями адекватності;

- обчислення оцінок прогнозів за допомогою побудованої моделі;
- використання прогнозів для автоматичного (або іншого типу) керування або для підтримки прийняття управлінських, особистих та інших типів рішень.

Система підтримки прийняття рішень виконує такі функції:

- попередня обробка даних;
- виявлення та врахування можливих невизначеностей даних;
- застосування оптимізаційних процедур для оцінки структури та параметрів моделі;
- моніторинг обчислювальних процесів у рамках системи підтримки прийняття рішень для досягнення найкращих результатів щодо адекватності моделі та оцінки прогнозів.

У випадку, коли досліджувані процеси є нестаціонарними та нелінійними, системи підтримки прийняття рішень передбачають адаптацію структури моделі і параметрів для покращення результатів роботи моделі. Спрощена класифікація нелінійних нестаціонарних процесів у демографії, екології, економіці та фінансах представлена на Рис. 1.1.

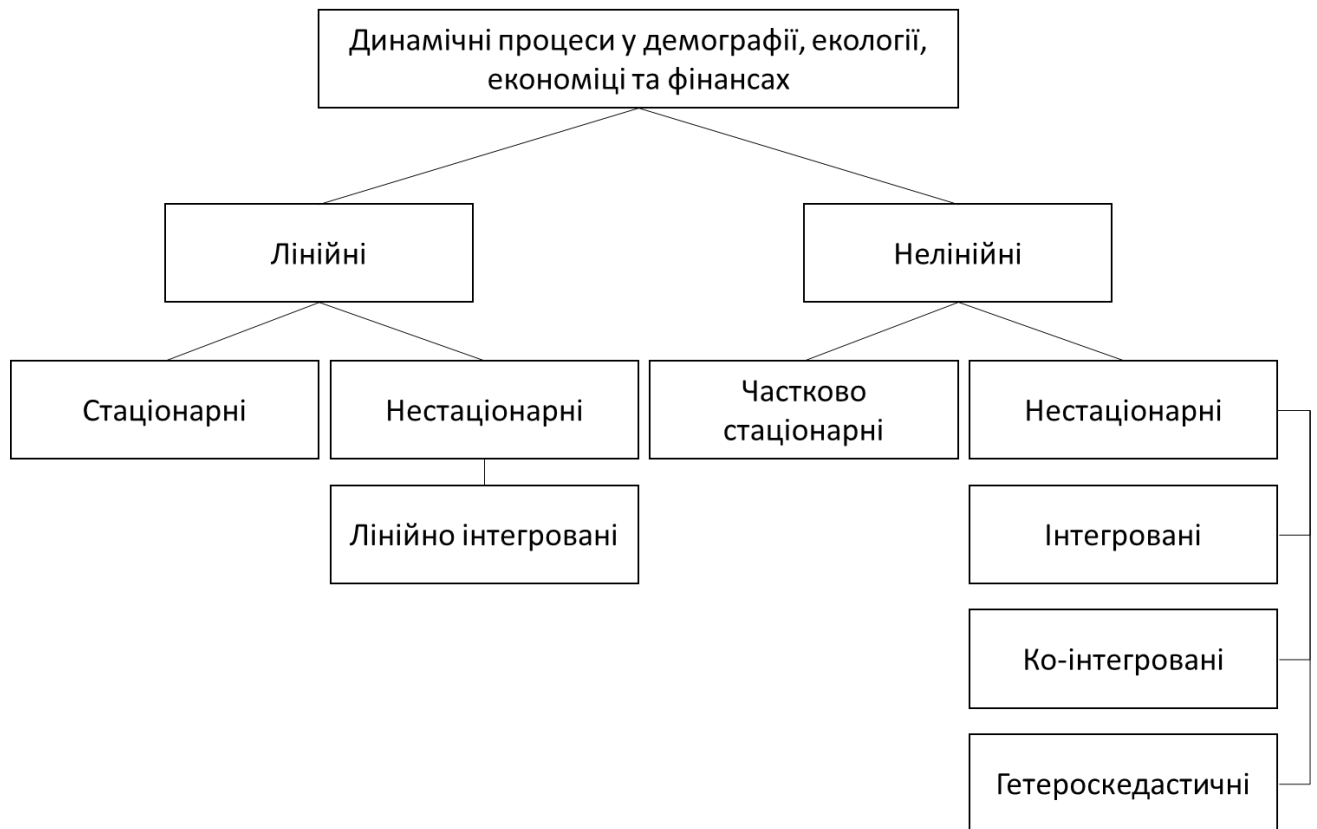


Рисунок 1.1 Спрощена класифікація динамічних процесів

Нелінійні нестационарні процеси, що виникають у різних сферах діяльності людини, пов'язані з великою кількістю невизначеностей, нечіткістю, неповнотою та неточністю даних.

У загальному випадку невизначеність – це фактор негативного впливу на процес попередньої обробки даних, побудови математичної моделі, оцінювання прогнозу та генерування рішення на його основі, який призводить до погіршення якості проміжних та остаточних результатів аналізу даних та експертних оцінок.

Поява невизначеності зумовлена недостатністю або спотворенням даних та/або експертних оцінок на будь-якому етапі згаданого процесу. Це може бути коротка вибірка даних, якої недостатньо для побудови адекватної моделі; спотворення вимірів шумовими складовими (збурення станів і похибки вимірів); некоректність встановлення типу розподілу даних і, як наслідок, неправильний вибір методу оцінювання параметрів

моделі; помилки дослідника при виборі методу обчислення оцінок прогнозів або формулювання альтернативних рішень на їх основі тощо.

У байєсівському аналізі невизначеність зазвичай визначається шляхом обчислення розподілу ймовірностей для невідомих параметрів або змінних у моделі. Цей розподіл ймовірностей називається апостеріорним розподілом і виводиться з апіорного (попереднього) розподілу та функції правдоподібності [16].

Апіорний розподіл представляє попередні знання або переконання щодо невідомих параметрів, а функція правдоподібності представляє ймовірність спостереження даних при заданих значеннях невідомих параметрів. Апостеріорний розподіл поєднує в собі попередній розподіл і функцію правдоподібності для отримання розподілу ймовірності невідомих параметрів, який відображає як попередні переконання, так і спостережувані дані.

Апостеріорний розподіл забезпечує міру невизначеності в моделі, оскільки він представляє діапазон значень, яких можуть набувати невідомі параметри, враховуючи спостережувані дані та попередні припущення. Апостеріорний розподіл можна використовувати для прогнозування нових даних або для оцінювання значень невідомих параметрів з невизначеністю [21].

У деяких випадках апостеріорний розподіл можна обчислити аналітично, за допомогою математичних формул. Однак у багатьох випадках для апроксимації апостеріорного розподілу може знадобитися використання чисельних методів, таких як метод марковських ланцюгів або метод Монте-Карло.

Для прогнозування таких процесів необхідно коректно опрацьовувати невизначеності, тому актуальною є задача розробки і застосування нових методів, які дають можливість здійснювати належну обробку вхідних даних з метою моделювання, прогнозування досліджуваних процесів, а також оцінювання відповідних ризиків.

У загальному випадку для боротьби з невизначеностями використовують нечітку логіку, ймовірісно-статистичне моделювання, оптимальну і цифрову фільтрацію даних, методи заповнення пропусків і обробки екстремальних значень, альтернативні методи оцінювання параметрів та інші. При побудові математичних моделей на основі даних зустрічаються три основні типи невизначеностей: статистична, структурна і параметрична [19 – 21].

- Статистична невизначеність зумовлена самими даними або вимірами, тобто наявністю пропусків вимірів, короткою вибіркою, наявністю екстремальних значень, впливом зовнішніх випадкових збурень на досліджуваній процес та похибками вимірів. У деяких випадках виникає необхідність оцінювати невимірювані компоненти вектора стану досліджуваної системи (змінні), які мають значення для побудови адекватної математичної моделі. Для того щоб досягти прийнятних за точністю результатів оцінювання значень таких змінних необхідно застосовувати альтернативні методи оцінювання, комплексування та комбінування вимірів.

- Структурна невизначеність або невизначеність моделі виникає при оцінюванні структури моделі на основі даних. Так, наприклад, оцінити порядок авторегресії можна тільки наближено, оскільки значення автокореляційної функції, а також часткової автокореляційної функції – це випадкові величини. Це ж стосується оцінювання лагу (часу запізнення) по входу, типу нелінійності і типу нестационарності (які оцінюються за відповідними статистиками). На кожному етапі аналізу даних ми маємо справу з оцінками параметрів і змінних, які є випадковими величинами, що вносить похибки у проміжні та остаточні результати обчислень. Тому виникає необхідність ідентифікації таких невизначеностей та вибору (створення) методів мінімізації їх впливу на результати аналізу даних.

- Параметрична невизначеність є наслідком наявності двох попередніх типів невизначеності. Вона зумовлена наближеними оцінками

параметрів, що характеризують структуру моделі (порядок, значення запізнення по входу, оцінки параметрів розподілу), наявністю випадкових збурень і похибок вимірів.

Крім того, більш рідко, але також виділяють такі типи невизначеностей [22,23]:

- Невизначеність процесу – невизначеність щодо фізичних, біологічних або соціальних процесів, які лежать в основі даних, а також щодо того, наскільки добре ці процеси зрозумілі та змодельовані.

- Епістемічна невизначеність – виникає через неповні або непевні знання і відображає обмеження людського розуміння і знань про певну систему.

- Алеаторна невизначеність – виникає через притаманну системі мінливість або випадковість, і яку неможливо зменшити навіть за умови досконалого знання та розуміння.

Досить часто неможливо встановити коректно тип розподілу даних внаслідок наявності коротких або спотворених вибірок – все це призводить до зміщення (зсуву) оцінок параметрів моделі від точних значень та підвищення дисперсії цих оцінок (відхилення від ефективності). Тому до вибору методу оцінювання параметрів на основі наявних даних завжди необхідно підходити обґрунтовано, враховуючи існуючі обмеження кожного методу.

При обробці статистичних даних, частіше за все необхідно обробляти такі невизначеності:

- зовнішні випадкові впливи (шум або збурення);
- похибки вимірів.

Похибки вимірів наявні в обох випадках: коли дані збирають люди або вимірювання доставляють датчики. Крім того, окреме джерело невизначеності може з'явитися через відсутність спостережень і можливу наявність екстремальних значень. Усі ці фактори негативного впливу мають бути правильно визначені та належним чином враховані в

обчислювальних процедурах, які використовуються в рамках спеціалізованих систем підтримки прийняття рішень [24,25].

1.2 Огляд існуючих моделей для аналізу та оцінювання операційного та фінансового ризику

В першу чергу, коли мова йде про оцінювання операційних фінансових ризиків, мається на увазі оцінювання ймовірності настання факту втрати, а також розмір можливих збитків, якщо цей факт все ж таки трапився. Оскільки вони можуть статися в будь-який момент часу, слід мати на увазі випадковий характер можливих втрат.

Щодо методів визначення ймовірності настання події, існують об'єктивні та суб'єктивні методи. Об'єктивний метод ґрунтується на обчисленні частоти, з якою відбувається подія, і може бути застосований у випадках, коли є достатньо статистичних даних. Суб'єктивний метод, з іншого боку, ґрунтується на суб'єктивних критеріях, які можуть бути побудовані завдяки різним припущенням. Цей метод може бути застосований у випадках, коли немає достатньо статистичних даних або коли експерти мають досвід і знання, що дозволяють їм зробити суб'єктивні оцінки.

Розглядаючи залежність між визначеними розмірами втрат і ймовірністю їх появи, ця залежність має місце в кривій ймовірностей виникнення визначеного рівня втрат (ризик).

Сьогодні існують математичні моделі, які дозволяють ефективно оцінювати фінансові та операційні ризики. Такі моделі постійно розвиваються і вдосконалюються, і вони є однією з важливих складових для стимулювання розвитку фінансової сфери. Методи оцінювання ризику, можна класифікувати з точки зору чинників ризику або наслідків дії ризикових подій. Математичні моделі для оцінювання ризиків також можна класифікувати на два основні класи в залежності від того, чи

базуються вони на аналізі наслідків, чи чинників ризику. Класифікація моделей ґрунтується на множині завдань, для яких вони можуть використовуватись.

1.2.1 Методологія байєсівського програмування

Байєсівський підхід до оцінювання ризику включає у себе використання умовних ймовірностей для визначення ймовірності настання ризику та для прийняття рішень про те, які заходи стосовно зменшення ризику слід прийняти.

Байєсівське програмування – це особливий підхід до моделювання та аналізу, який використовує байєсівську теорію ймовірностей для побудови моделей і прогнозування. Сьогодні байєсівська методологія моделювання та формування ймовірнісного висновку (остаточного результату) на основі моделі отримала назву байєсівського програмування.

Одним із прикладів підходу байєсівського програмування до оцінювання ризиків є система ймовірнісної оцінки ризиків (PRA - Probabilistic Risk Assessment). PRA широко використовується в ядерній промисловості для оцінювання та управління ризиками, пов'язаними з атомними електростанціями. Система PRA використовує байєсівські мережі для моделювання взаємозв'язків між різними компонентами і підсистемами електростанції, а також для оцінки ймовірності різних типів аварій та їхніх потенційних наслідків.

Іншим прикладом підходу байєсівського програмування є використання методів на основі марковських ланцюгів Монте-Карло (MCMC – Markov Chain Monte Carlo) для оцінювання параметрів складних моделей. Методи MCMC передбачають використання вибірки з апостеріорного розподілу моделі, що дозволяє врахувати попередні знання та невизначеність у моделі. Цей підхід використовується у фінансах для

оцінювання параметрів складних фінансових моделей, таких як моделі ціноутворення опціонів і моделі ціноутворення активів.

Методологія байєсівського програмування означає, що для аналізу даних використовується множина методів, які забезпечують розв'язання таких задач [25, 26]:

- побудову ймовірнісно-статистичних моделей різних типів (оцінювання структури і параметрів) з використанням статистичних даних і експертних оцінок;
- обчислення остаточних результатів на основі створеної моделі згідно з постановкою задачі: оцінок прогнозів, керуючих впливів, оцінок змінних і параметрів на виході фільтрів, розпізнавання образів, знаходження рішень стосовно управління досліджуваними процесами і об'єктами і т. ін.;
- аналіз коректності отриманих результатів за відповідними множинами статистичних критеріїв якості;
- практичне застосування отриманих результатів та їх корегування (повторне оцінювання) залежно від остаточного результату.

До методології байєсівського програмування відносять методи, подані в розділі 2.1. Основне рівняння оцінювання має такий вигляд:

$$P(S(k)|O(0) \dots O(k)) = P(O(k)|S(k)) \times \sum_{S(k-1)} [(P(S(k)|S(k-1)))(P(S(k-1)|O(0) \dots O(k-1))], \quad (1.1)$$

де $S(0), \dots, S(k)$ – часовий ряд змінних стану; $O(0), \dots, O(k)$ – часовий ряд спостережень; $P(S(k)|S(k-1))$ – модель системи або модель переходів; $P(O(k)|S(k))$ – модель спостережень, яка показує яким буде спостереження у момент k , якщо система знаходиться у стані $S(k)$ [1].

За цією моделлю можна знайти $P(S(k+l)|O(0) \dots O(k))$ – тобто, яким буде розподіл ймовірностей станів у момент $k+l$, маючи

спостереження $O(0), \dots, O(k)$. Якщо $l = 0$, то реалізується процедура фільтрації; якщо $l > 0$, то реалізується процедура прогнозування; а при $l < 0$ відбувається згладжування – відновлення минулого стану на основі спостережень, зроблених до або після моменту згладжування.

Загалом, байєсівське програмування є гнучким і потужним підходом до моделювання та аналізу, який можна застосовувати до розв’язання широкого кола проблем, включаючи оцінку ризиків та прийняття рішень. Однак він вимагає глибокого розуміння байєсівської теорії ймовірностей і статистичного моделювання, а також досвіду в програмуванні та аналізі даних.

1.2.2 Value at risk (VaR)

Один з підходів до оцінювання операційних фінансових ризиків передбачає використання методів розподілу збитків LDA (Loss Distribution Approach). В їх основі, лежить припущення про те, що випадкову величину, яка характеризує розмір втрат, що сталися протягом часу, можна визначити так:

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

де $n(t)$ – випадкова величина, яка характеризує кількість випадків втрат конкретного типу за певний період часу t ;

L_i – множина випадкових величин, що характеризують величини можливих окремих втрат.

Модель Value at risk (VaR) оцінює потенційну втрату вартості портфеля або активів організації через несприятливі події протягом певного періоду часу. VaR зазвичай розраховується за допомогою статистичної моделі, яка включає історичні дані, припущення щодо

ринкових умов та інші вхідні дані. Перевага цього підходу полягає в тому, що він надає кількісну оцінку потенційного ризику, але недоліком є те, що він ґрунтується на деяких припущеннях і може бути неточним у всіх можливих ситуаціях [27, 28].

Використання цього підходу передбачає виконання таких етапів:

- Визначення часового проміжку для дослідження. Першим кроком є визначення часового періоду, за який буде розраховуватися VaR. Зазвичай використовується кратність – день, тиждень, місяць або рік.

- Визначення рівня довіри. Довірчий інтервал – це ймовірність того, що збиток перевищує допустимий рівень. Наприклад, довірча ймовірність 95% означає, що існує 5% ймовірність того, що збиток перевищить обчислене значення VaR. Вибір рівня довіри залежатиме від відношення до ризику організації або окремої особи (експерта).

- Збір історичних даних. Для розрахунку VaR необхідні історичні дані про вартість портфеля або інвестиції, що аналізуються. Ці дані повинні бути репрезентативними для часового горизонту та ринкових умов, що розглядаються.

- Розрахунок прибутковості: найпростіший спосіб – різниця початкової вартості від кінцевої та ділення на початкову. Наприклад, якщо портфель мав вартість \$10,000 на початку місяця і \$11,000 в кінці місяця, дохідність буде $(11,000 - 10,000) / 10,000 = 0,1$ або 10%.

- Визначення розподілу ймовірностей: зазвичай використовуються статистичні методи, такі як: нормальний розподіл, лог-нормальний розподіл або інший розподіл, який відповідає даним.

- Власне розрахунок VaR виконується за виразом:

$$VaR = A * Pr * \sqrt{T}, \quad (1.3)$$

де A – вартість портфеля;

Pr – процентиль розподілу;

T – часовий проміжок (горизонт).

Процентиль розподілу є оберненою величиною до кумулятивної функції розподілу ймовірностей (CDF – cumulative distribution function) на обраному довірчому рівні. Наприклад, якщо рівень довіри становить 95%, процентиль розподілу буде значенням розподілу ймовірностей, де CDF дорівнює 0,05.

Припустимо, що портфель має вартість \$100 000, а дохідність за один місяць розподілена за нормальним законом із середнім значенням 1% і стандартним відхиленням 2%. Необхідно розрахувати VaR з довірчою ймовірністю 95%.

Спочатку визначаємо процентиль розподілу, використовуючи стандартну таблицю нормального розподілу або калькулятор. Для довірчої ймовірності 95% процентиль розподілу дорівнює 1,645. Далі обчислюємо квадратний корінь з часового горизонту, який дорівнює одному місяцю або 1/12 року. Квадратний корінь з 1/12 дорівнює приблизно 0,29.

Підставляємо ці значення у формулу для обчислення VaR і отримуємо:

$$VaR = \$100\,000 \times 1,645 \times 0,29 \times 2\% = \$954,1.$$

Це означає, що існує 5% ймовірність того, що портфель втратить більше \$954,1 протягом наступного місяця.

1.2.3 Метод Монте-Карло

Моделювання методом Монте-Карло – це обчислювальний метод, який використовує випадкову вибірку для імітації низки можливих результатів у моделі.

В задачах з використанням методів розподілу збитків припускається, що величини в множині є незалежними та однаково розподіленими для конкретного типу збитків. Для створення таких моделей проводиться аналіз втрат за вибраний період по кожній парі “втрати” - “подія”. На

основі даних про частоту втрат і величину збитків, що були спостережені у минулому, для кожної пари розраховується вибіркове середнє значення частоти виникнення ризикових подій і вибіркове середнє значення величини втрат при виникненні ризикової події. Після визначення функцій розподілу випадкових величин будується функція розподілу випадкової величини зазвичай саме за допомогою методу Монте-Карло. Функція розподілу випадкової величини, що характеризує загальний обсяг втрат за вибраним видом ризиків, дає можливість знайти точкову оцінку математичного сподівання втрат і розрахувати квантиль заданого рівня, тобто значення OpVaR (Operational Value at Risk) [27,28].

Метод Монте-Карло часто використовують у фінансах та інженерії для моделювання складних систем з багатьма змінними. Основні етапи моделювання включають в себе такі:

- Визначення проблеми і побудова моделі. Першим кроком є чітке визначення проблеми і створення моделі для її представлення. Модель повинна включати всі змінні, які впливають на результат, і взаємозв'язки між ними. Модель може бути виражена у вигляді набору рівнянь або блок-схеми. Наприклад проблема полягає в тому, щоб оцінити діапазон можливих прибутків від інвестиції в акції протягом наступного року. Модель може бути виражена наступним чином:

$$I = A + D * V, \quad (1.4)$$

де I – дохідність;

A – середня дохідність;

D – стандартне відхилення;

V – випадкове число, що використовуються для імітації різних можливих результатів роботи системи, що моделюється.

- Визначення розподілу ймовірностей вхідних змінних. Для цього проводиться аналіз історичних даних або експертних оцінок для

визначення ймовірного діапазону значень для кожної вхідної змінної та ймовірності появи кожного значення. Найпоширеніші розподіли, що використовуються в моделюванні методом Монте-Карло, включають нормальний, лог-нормальний, та рівномірний розподіли. Наприклад, вважається, що середня дохідність та середньоквадратичне відхилення розподілені за нормальним законом. На основі історичних даних середня дохідність оцінюється як 10% зі стандартним відхиленням 15%.

- Створення випадкових вибірок. При моделюванні методом Монте-Карло випадкові вибірки генеруються з розподілу ймовірностей кожної вхідної змінної. Кількість зразків залежить від складності моделі та бажаного рівня точності. Для кожної змінної випадкові вибірки генеруються за допомогою генератора випадкових чисел за допомогою програмних продуктів. Наприклад, припустимо, що ми генеруємо 1000 вибірок середньої дохідності та 1000 вибірок стандартного відхилення.

- Обчислення результатів моделі: після того, як випадкові вибірки згенеровані, вони використовуються як вхідні дані для моделі. Після чого модель виконується для кожного набору вхідних даних, щоб розрахувати вихідні дані. Цей процес повторюється для всіх вибірок. В нашому прикладі, для кожного набору випадкових вибірок розраховується дохідність за допомогою рівняння моделі. Наприклад, якщо випадкова вибірка середньої прибутковості становить 12%, а випадкова вибірка стандартного відхилення - 10%, то прибутковість розраховується таким чином:

$$I = 12\% + 10\% * V. \quad (1.5)$$

- Аналіз результатів. Останнім кроком є аналіз результатів імітаційного моделювання (симуляції). Результатом моделювання є множина даних, які представляють діапазон можливих результатів. Ці дані можна використовувати для обчислення статистичних показників, таких як

середнє значення, стандартне відхилення та процентилі. Середня дохідність розраховується як середнє значення всіх точок даних, а стандартне відхилення – як стандартне відхилення всіх точок даних. VaR можна розрахувати як 5-й процентиль точок даних, що представляє максимальний очікуваний збиток з довірчою ймовірністю 95%.

Метод Монте-Карло також називають статистичним методом оцінювання ризику, тобто таким, що полягає у вивченні статистики втрат або ж негативних наслідків реалізації рішень, що мали місце в аналогічних видах окремої діяльності, яка піддається аналізу. Основним показником, який розраховується на підставі статистичного методу, є частота втрат, пов'язаних із певним видом діяльності:

$$f = \frac{n'}{n_{\text{заг}}}, \quad (1.6)$$

де n' – кількість випадків настання втрат у статистичній вибірці;

$n_{\text{заг}}$ – загальна кількість випадків, що розглядались у статистичній вибірці.

При прийнятті рішення на основі цього методу показник частоти витрат переноситься на прогнозні дані і розглядається вже як ймовірність настання втрат певного рівня.

Приклад моделювання за методом Монте-Карло

Припустимо, необхідно оцінити вартість європейського опціону “колл” на акції, який дає власнику право купити акції за фіксованою ціною (“ціна виконання”) на визначену дату закінчення терміну дії. Вартість опціону залежить від кількох змінних, таких як ціна акції, ціна виконання, волатильність ціни акції та безризикова відсоткова ставка.

Для оцінювання вартості опціону ми можемо використовувати метод Монте-Карло, моделюючи можливі майбутні ціни на акції та

розраховуючи виплати за опціоном для кожної симуляції. Основні кроки моделювання за методом Монте-Карло в цьому випадку подано нижче.

Генерування множини випадкових чисел, які підпорядковуються нормальному розподілу з середнім значенням 0 і стандартним відхиленням 1. Ці випадкові числа представлятимуть випадкові шоки для ціни акцій.

Для кожної симуляції розраховуються майбутні ціни акцій, використовуючи поточну ціну акцій, безризикову відсоткову ставку та волатильність. Майбутня ціна акції S_t в момент часу t задається формулою:

$$S_t = S_0 \times e^{\left(\left(r - \frac{\sigma^2}{2}\right)t + \sigma M_t \sqrt{t}\right)}, \quad (1.7)$$

де S_0 – поточна ціна акції; r - безризикова відсоткова ставка; σ - волатильність ціни акції; t - час до закінчення опціону; M_t - випадкове число з проміжку 0 – 1.

Для кожної симуляції обчислюється виграш від опціону, взявши максимальне значення 0 та різницю між майбутньою ціною акції та ціною виконання опціону. Виграш P_t у момент часу t задається формулою:

$$P_t = \max(0, S_t - K), \quad (1.8)$$

де K - ціна виконання.

Кроки 2 і 3 повторюються для великої кількості симуляцій (наприклад, 10 000), щоб отримати розподіл можливих виплат.

Оцінюється вартість опціону, взявши середнє значення змодельованих виплат, дисконтованих до теперішньої вартості за безризиковою відсотковою ставкою. Розрахункова вартість опціону V визначається за формулою:

$$V = e^{(-r*T)} * E^{P_t}, \quad (1.9)$$

де T – час до закінчення терміну дії опціону, а E^{P_t} – очікувані виплати при закінченні терміну дії на основі змодельованих виплат.

Таке моделювання за методом Монте-Карло дозволяє оцінити вартість опціону, враховуючи невизначеність у ціні акцій та інших змінних.

1.2.4 Метод внутрішніх вимірів (Internal Measurements Approach)

Також слід зазначити, що варіанти методів розподілу збитків передбачають також формулювання, що базуються на розширеній множині змінних, до якої входять і чинники ризику. Одним із таких варіантів є метод ІМА, який дозволяє оцінити максимально можливі втрати для певного виду ризиків, не вимагаючи побудови розподілу випадкової величини. Натомість, він полягає в розділенні всіх збитків на очікувані та непередбачені втрати. При цьому припускається, що існує функціональна залежність між величинами очікуваних та непередбачених збитків. Очікувані збитки є тими, що входять у суму близької до математичного сподівання суми збитків за період, тоді як непередбачені збитки є тими, що перевищують середнє значення та належать до залишків статистичного розподілу. Простим випадком є лінійна залежність, тобто:

$$\hat{Q}_{99\%} = E_I P_E L_{GE} = \gamma E_L \alpha, \quad (1.10)$$

де $\hat{Q}_{99\%}$ – оцінка 99%-го квантиля розподілу можливих втрат (величини капіталу, необхідного для покриття ризику) за цим типом події; P_E – ймовірність настання негативної події цього типу в даний період; L_{GE} – середня величина збитку конкретного типу за умови настання негативної події; E_I – коефіцієнт масштабу; γ – коефіцієнт для оцінювання

вимог до капіталу через оцінку очікуваних збитків E_L . Сукупна оцінка максимально можливих втрат визначається сумою усіх оцінок, розрахованих для різних типів ризикових подій і джерела виникнення самої події [29].

1.2.5 Scenario-based Advanced Measurement Approach (Sb-AMA)

Це один з методів побудови моделі на основі аналізу факторів ризику, що передбачає ґрунтовний аналіз процесів суб'єкта, що піддається аналізу (підприємство, діяльність), з урахуванням внутрішніх причинно-наслідкових зв'язків та їх визначення. Метод Sb-AMA застосовується в операційних ризиках шляхом визначення сценаріїв, що можуть призвести до збитків. Наприклад, можна визначити сценарії, пов'язані зі зловживанням співробітниками, некомпетентністю персоналу, технічними збоями в програмному забезпеченні тощо. Для кожного сценарію на основі історичних даних або експертних оцінок визначають частоту і розмір збитків. Після перевірки даних і виправлення некоректних оцінок сценарії групують за чинниками ризику і для кожної групи сценаріїв оцінити параметри статистичних розподілів частоти і величини збитків. Завершальним етапом є імітаційне моделювання за методом Монте-Карло, що дає можливість оцінити загальний розподіл збитків. Далі, для кожного з таких сценаріїв розробляється план дій для управління ризиком та використовуються результати дослідження для передбачення можливих втрат.

Оцінювання можливих втрат може бути проведено за допомогою різних методів, таких як історична аналітика, аналіз витрат, рейтинг-шкали та інші. Застосування методу Sb-AMA для менеджменту операційних ризиків дозволяє компаніям виявляти можливі ризики та вчасно приймати заходи для їх управління, зменшуючи при цьому ризики збитків [30].

1.2.6 Байєсівська модель з автоматичним оцінюванням невизначеності

Метод SBAMARD (Sparse Bayesian Additive Model with Automatic Relevance Determination), або як його ще називають SBLARD (Sparse Bayesian Learning with Automatic Relevance Determination) – це статистичний метод машинного навчання, який використовується для регресійного аналізу даних. В основі методу лежить підхід з використанням Байєсівської статистики, який дозволяє оцінювати невизначеності ймовірностей параметрів моделі на основі апріорної інформації та даних. Використання подібних методів дозволяє враховувати неоднорідність даних та їх залежності від багатьох змінних, шляхом розкладу моделі на адитивні складові. Кожна складова може бути апріорно задана як залежність від певних змінних, що дозволяє підібрати оптимальну модель для даних з високою точністю та ефективністю.

Метод SBLARD також використовує автоматичне визначення релевантних змінних, тобто знаходить змінні, які дійсно впливають на результат, і виключає змінні, які не мають великого впливу. Це дозволяє зменшити кількість змінних та складність моделі, що знижує ризик перенавчання та покращує її універсальність. Загалом, цей застосовується в різних областях, де необхідно робити прогнози на основі даних, зокрема в медицині, біології, фінансах та інших галузях [31].

1.3 Програмні інструменти побудови системи підтримки прийняття рішень. Огляд програмних засобів

У цьому розділі розглядаються основні програмні інструменти та мови розробки, що використовуються в процесі створення та програмування систем підтримки прийняття рішень а також оцінювання операційних ризиків.

Найбільшими гравцями на існуючому ринку є всього декілька постачальників програмних засобів, що тісно конкурують один з одним в області побудови моделей динамічних процесів. Нижче наведено перелік цих програмних продуктів.

1. FICO Analytic Modeler Scorecard. Скорингові (оціночні) карти (Score card) використовуються як потужна та зручна технологія прогностичного моделювання з широким спектром бізнес-додатків. Вони допомагають отримати уявлення про свої дані та прогностні відносини всередині них, а також вирішити проблеми моделювання, які, найімовірніше, виникнуть у практиці експертного оцінювання. FICO Analytic Modeler Scorecard створена для бізнес-аналітиків, які тільки починають займатися предиктивним моделюванням. Це інтуїтивно зрозуміле програмне забезпечення, розміщене в FICO Analytic Cloud і доступне для локального розгортання як частина пакета FICO Decision Management Suite, що може автоматично створювати та досліджувати ефективні прогностичні моделі [32]. На відміну від інших рішень для побудови карт показників, FICO Analytic Modeler Scorecard надає інтерактивні візуалізації, які дають представлення про найбільш важливі прогностичні закономірності та фактори, що визначають бізнес-результати (Рис. 1.2).

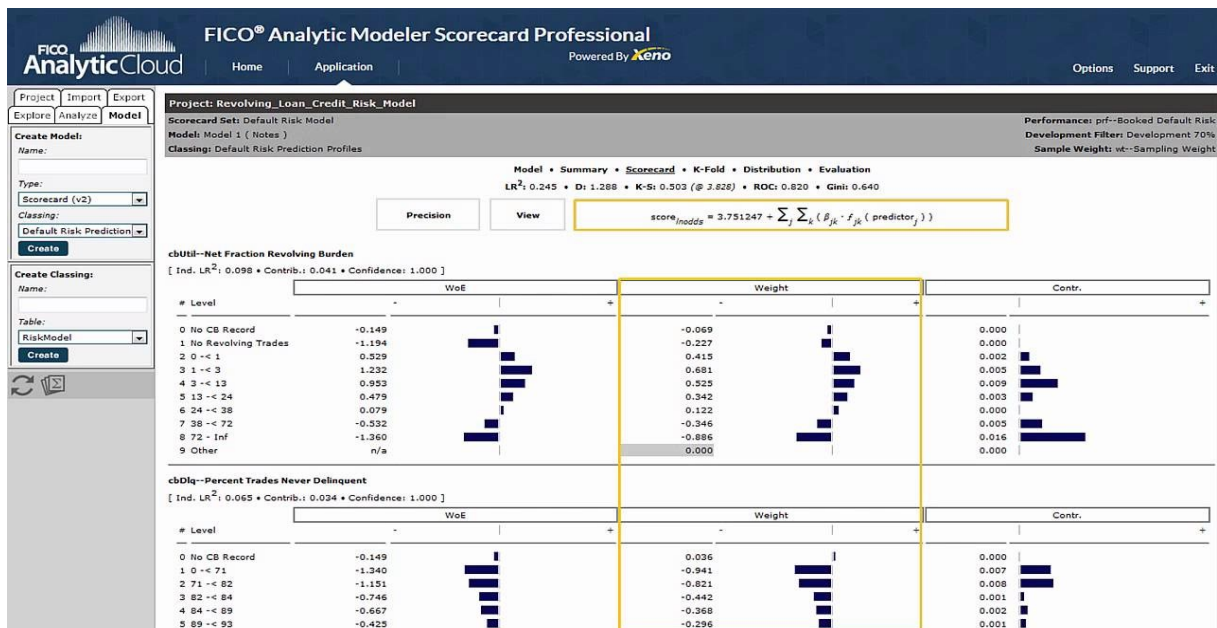


Рисунок 1.2 Приклад інтерфейсу програмного продукту FICO Analytic Modeler Scorecard

Кожен набір даних пов'язаний із розробленими на його основі моделями, для того щоб моделі можна було згрупувати відповідно до того, які дані використовувалися для створення моделі. Поля в наборах даних визначаються за типом основної та додаткової інформації, а також може ідентифікуватися як цільова або виняткова.

Побудова моделі вимагає від користувача вибору цільової змінної, встановлення необхідних параметрів і запуску розрахунку. Сторінка результатів вказує на те, наскільки сильним є прогноз у нетехнічних термінах, надаючи опис разом із змінною кривою важливості, розподілу та ранжирування. Кожен із цих елементів надає певні додаткові деталі – наприклад, користувачі можуть бачити вагомість елементів і інтервали для змінної, досліджувати поділ на різних рівнях тощо.

2. Minitab Statistical Software. Програмне забезпечення Minitab® Statistical Software забезпечує ефективний спосіб введення в систему статистичних даних, маніпулювання цими даними, визначення тенденцій і моделей, екстраполяції відповідей на поточні проблеми, візуалізацію, статистичний аналіз, аналітику прогнозів і покращень, що дозволяє

приймати рішення на основі даних. Це одне з найбільш широко використовуваних програмних забезпечень для бізнесу будь-якого розміру – малого, середнього та великого.

Система дозволяє використовувати додаткові інтерактивні інструменти для візуалізації даних на графіках, змінювати тип діаграм - з гістограми на корелограму, бульбашкову діаграму тощо, а також дозволяє робити імпорт з Microsoft Excel і Access та експортувати до Microsoft PowerPoint і Word, Minitab Engage і Minitab Workspace [33]. Приклад інтерфейсу програмного продукту Minitab Statistical Software подано на Рис. 1.3.

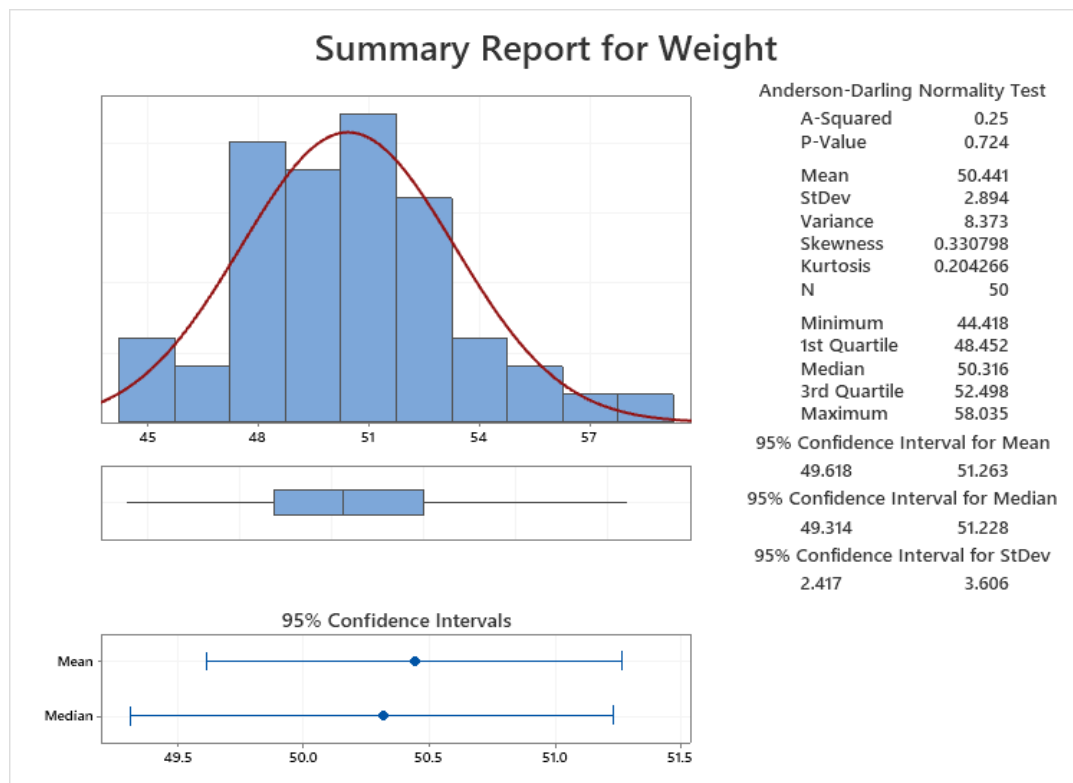


Рисунок 1.3 Приклад інтерфейсу програмного продукту Minitab Statistical Software

3. Credit Scoring for SAS Enterprise. Під час проведення оцінювання капіталу важливою складовою оцінки вимог щодо його формування є моделювання та оцінювання кредитів. Різні фінансові установи, зокрема банки, створюють окремі відділи, що займаються питаннями та

проблемами подібного моделювання. Credit Scoring for SAS дозволяє створювати, перевіряти та розгортати кращі моделі кредитного ризику, використовуючи передову прогностну аналітику та власний досвід. За допомогою цього інструменту, спеціалісти мають змогу точно оцінити та контролювати ризик у своїх існуючих споживчих портфелях, розуміючи специфічні характеристики ризику, які призводять до прострочення, дефолту та безнадійної заборгованості. Оптимізувавши процеси затвердження кредитів і удосконалюючи стратегії придбання, утримання та збору, фінансові установи, що використовують подібні програмні продукти, мають змогу примножувати свій прибуток та різко скорочувати свої витрати [34].

Нижче, на Рис. 1.4, наведено оглядову схему архітектури програмного продукту, яка використовується для моделювання діяльності фінансових установ.

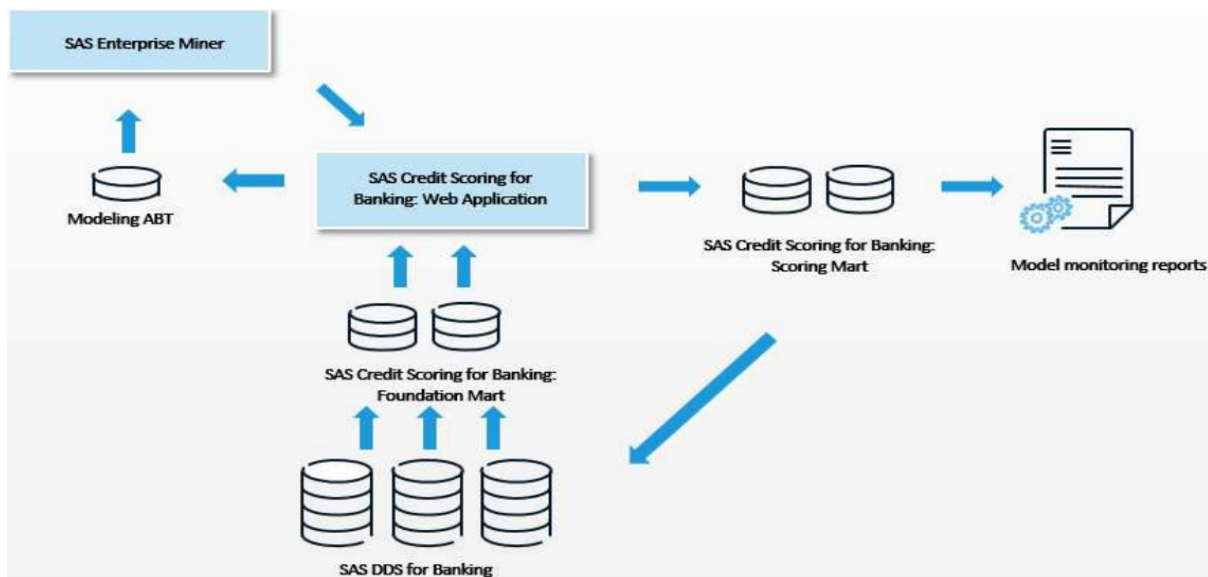


Рисунок 1.4 – Загальна схема роботи Credit Scoring for SAS Enterprise

4. IBM SPSS Modeler. The IBM SPSS Modeler це провідне рішення для візуалізації даних та машинного навчання. Цей інструмент допомагає підприємствам пришвидшити час до отримання цінності та бажаного результату, пришвидшивши виконання операційних завдань для

спеціалістів із обробки даних. Провідні організації в усьому світі покладаються на IBM для виявлення даних, прогнозової аналітики, керування моделями та розгортання, а також машинного навчання для монетизації активів даних. IBM SPSS Modeler дає змогу організаціям використовувати дистрибутиви даних і сучасні програми за допомогою повних готових алгоритмів і моделей, які підходять для гібридних багатохмарних середовищ із надійним керуванням і безпекою [35].

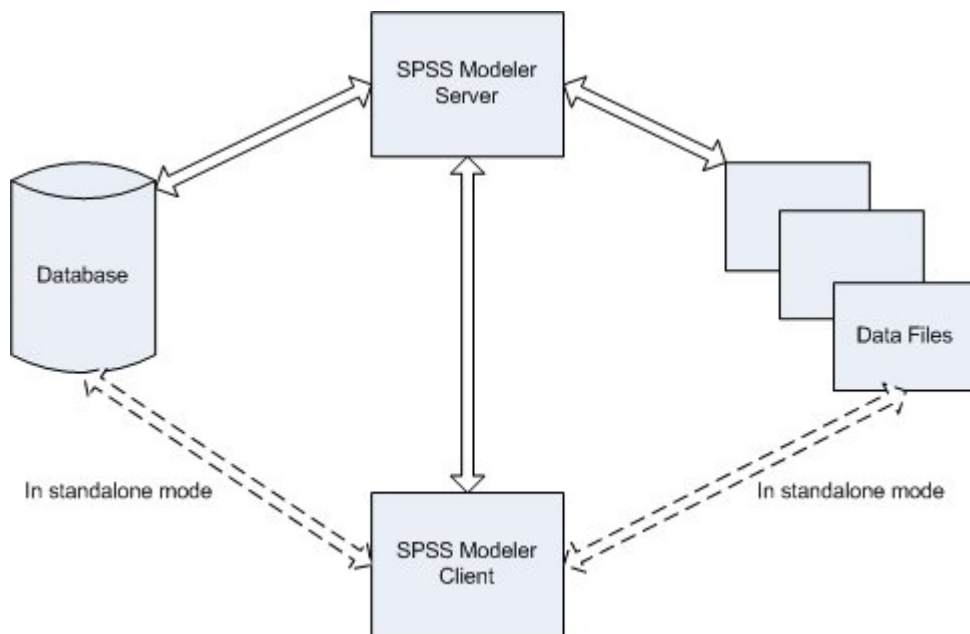


Рисунок 1.5 Загальна архітектура інструменту IBM SPSS Modeler

Якщо жоден з існуючих програмних засобів не підходить для вирішення конкретних задач, є можливість розробити власну програму, використовуючи одну з двох найпоширеніших мов програмування - R або Python. Щоб не створювати програму з нуля, існує широкий спектр бібліотек та шаблонів, що дозволяють значно спростувати розробку.

1.3.1 Мова програмування Python

Python – це мова програмування загального призначення, яка зарекомендувала себе в багатьох галузях, включаючи фінанси та ризик-

менеджмент. Синтаксис ядра Python мінімальний. У той час, коли стандартна бібліотека включає великий об'єм корисних функцій. Python підтримує технологічне, об'єктно-орієнтоване, функціональне та логічно орієнтоване програмування. Загалом, Python є потужним інструментом для оцінки операційних ризиків. Він дозволяє легко обробляти та аналізувати великі обсяги даних, створювати візуалізації та графіки для кращого розуміння результатів, а також програмувати власні скрипти та програми для розрахунку показників ризику. Для оцінки операційних ризиків використання Python має кілька переваг.

Перш за все, в Python є множина пакетів та бібліотек для обробки та аналізу даних, які можуть бути корисні для оцінки ризиків. Наприклад, NumPy та Pandas можуть допомогти зчитувати, обробляти та аналізувати великі обсяги даних, а Matplotlib та Seaborn дозволяють створювати візуалізації та графіки для кращого розуміння результатів.

Бібліотека NumPy додає підтримку великих багатовимірних масивів і матриць, разом з великою бібліотекою високорівневих математичних функцій для операцій з цими масивами. Математичні алгоритми, реалізовані на Python, часто працюють набагато повільніше тих же алгоритмів, реалізованих на компільованих мовах (наприклад, Фортран, Сі, Java). Бібліотека NumPy надає реалізації обчислювальних алгоритмів (у вигляді функцій і операторів), оптимізовані для роботи з багатовимірними масивами [36].

Pandas є потужним інструментом для аналізу даних. Пакет дає можливість будувати зведені таблиці, виконувати групування даних, надає зручний доступ до табличних даних, а за наявності пакета matplotlib дає можливість рисувати графіки на отриманих наборах даних. Основні можливості бібліотеки:

- об'єкт DataFrame для маніпулювання індексованими масивами двовимірних даних;

- інструменти для обміну даними між наявними структурами в пам'яті і файлами різних форматів;
- засоби інтегрування даних і способи обробки пропусків;
- переформатування наборів даних, в тому числі створення зведених таблиць;
- зріз даних за значеннями індексу, розширені можливості індексування, вибірка з великих наборів даних;
- вставка і видалення стовпців даних;
- можливості групування дозволяють виконувати триетапну операцію типу «поділ, модифікація, об'єднання»;
- злиття і об'єднання масивів даних;
- ієрархічне індексування дозволяє працювати з даними високої розмірності в структурах меншої розмірності;
- робота з часовими рядами: формування часових періодів і зміна інтервалів.

Matplotlib – це бібліотека Python для побудови високоякісних двовимірних графіків. Matplotlib є гнучким, легко конфігурованим пакетом, який разом з NumPy, SciPy і IPython надає можливості, подібні MATLAB. В даний час пакет працює з кількома графічними бібліотеками, включаючи wxWindows і PyGTK [36]. Пакет підтримує багато видів графіків та діаграм:

- графіки (line plot);
- діаграми розкиду (scatter plot);
- стовпчасті діаграми (bar chart)
- гістограми (histogram);
- кругові діаграми (pie chart);
- лист діаграми (stem plot);
- контурні графіки (contour plot);
- поля градієнтів (quiver);
- спектральні діаграми (spectrogram).

Користувач може вказати осі координат, додати написи і пояснення, використовувати логарифмічну шкалу або полярні координати.

Нескладні тривимірні графіки можна будувати за допомогою набору інструментів (toolkit) `mplot3d`. Є й інші набори інструментів: для картографії, для роботи з Excel, утиліти для GTK і інші.

Друга перевага полягає в тому, що Python є відкритим джерелом, тобто він безкоштовний та доступний для всіх. Це передбачає можливість використовувати Python для оцінювання ризиків, не витрачаючи кошти на дороге програмне забезпечення.

Третя перевага полягає в тому, що Python дає можливість програмістам створювати власні скрипти та програми для розрахунку показників ризику. Наприклад, існує можливість створити скрипт, який буде обчислювати показник VaR (Value at Risk).

1.3.2 Мова програмування R

R – це мова програмування та середовище аналізу даних, яке широко використовується для статистичного аналізу, моделювання та візуалізації даних. У контексті оцінки операційних ризиків, R також є потужним інструментом з кількома перевагами. Перш за все, R має широкий вибір пакетів та бібліотек, що спеціалізуються на обробці та аналізі даних, включаючи ризик-менеджмент. Наприклад, пакети, такі як "operationalRisk" надають функціональність для моделювання та оцінки операційних ризиків, включаючи методи оцінки капіталу ризику та сценарний аналіз, а "riskclustr" дозволяє проводити кластерний аналіз ризиків, що допомагає групувати схожі типи ризиків та оцінювати їх вплив. Використовуючи різні методи та моделі для оцінки ризиків, можливо створювати графіки та діаграми для візуалізації результатів.

Мова R використовує всі базові принципи об'єктно орієнтованого програмування, тобто додаток складається з об'єктів, кожен з яких має свій

стан, містить певні атрибути та виконує певну дію. Подібно до Python, R має відкритий код, що дозволяє використовувати мову програмування безкоштовно та вносити власні зміни у вихідний код. Це сприяє гнучкості та можливості вдосконалення функціональності для конкретних потреб у оцінці ризиків. Саме тому мова R широко використовується серед статистиків та аналітиків даних для розробки статистичного програмного забезпечення і аналізу даних.

Ще одна перевага – широка спільнота користувачів та документація. R має активну спільноту аналітиків даних та ризик-менеджерів, які активно обговорюють та обмінюються знаннями. Існує велика кількість документації, пакетів, прикладів та навчальних матеріалів, що сприяє легкому освоєнню та використанню R для оцінки операційних ризиків. На додаток до вищезгаданих, широко використовуваних загальних методів, слід відзначити також наступні:

- VaR: пакет використовується для обчислення Value at Risk (VaR), який є метрикою ризику, що вимірює максимальну можливу втрату у заданий інтервал часу з певною ймовірністю.

- QRM: надає функціональність для квантильного управління ризиками, включаючи обчислення квантилей, моделювання ризикових факторів та оцінку портфелів ризиків.

- RiskAnalytics: Цей пакет містить набір інструментів для аналізу ризиків, включаючи методи для обробки даних, моделювання ризиків, оцінки та візуалізації результатів.

Крім того, існують додаткові пакети для мови R, які можна використовувати для оцінки операційних ризиків, наприклад такі:

1. CARET (Classification And REgression Training): набір інструментів для розв'язання задач класифікації та регресії. Він дозволяє використовувати різні алгоритми машинного навчання для моделювання та прогнозування ризикових подій. Цей набір розроблений для формування комплексних моделей прогнозування. В ньому наявні алгоритми, придатні

для різних завдань, а також можливість автоматичного підбору оптимальних параметрів для побудованої моделі, на основі даних про результати проведених експериментів. CARET є одним з найпопулярніших пакетів мови R для машинного навчання. Його основна мета полягає у спрощенні процесу моделювання, налаштування та оцінки моделей класифікації та регресії. Основні завдання, що вирішуються за допомогою цього пакету, наступні:

- Підготовка даних: CARET має вбудовані функції для обробки та попередньої обробки даних перед моделюванням. Існує можливість виконати масштабування, нормалізацію, видалення пропущених значень та інші операції з даними.

- Вибір моделей: надаються засоби для автоматичного вибору моделей, які найкраще підходять для вашої задачі. Він дозволяє використовувати різні алгоритми класифікації та регресії і автоматично визначати найкращі параметри моделей.

- Налаштування параметрів: CARET пропонує функції для автоматичного налаштування параметрів моделей за допомогою методів, таких як перехресна перевірка та генетичний алгоритм. Це допомагає покращити ефективність моделі та уникнути перенавчання.

- Оцінка та порівняння моделей: кожен побудовану модель можна оцінити та порівняти на основі запропонованих метрик ефективності, таких як точність, чутливість, специфічність, середньоквадратична помилка та інші.

- Валідація моделі: CARET дозволяє виконувати валідацію моделі на основі розбиття даних на тренувальний та тестовий набори, а також застосовувати перехресну перевірку для оцінки стабільності та узагальненості моделі.

CARET підтримує багато алгоритмів машинного навчання, таких як логістична регресія, дерева рішень, градієнтний бустінг, нейронні мережі

та інші. Він також інтегрується з іншими пакетами мови R, такими як `ggplot2` для візуалізації результатів.

2. MICE (Multiple Imputation by Chained Equations) є пакетом мови R, призначеним для розв'язання проблеми відновлення пропущених даних шляхом множинного заповнення. Основна мета MICE полягає в тому, щоб створити багато можливих заповнених наборів даних, які відображають невизначені значення пропущених даних. Використовуючи ці набори, можна оцінити невизначеність, пов'язану з пропущеними даними, і провести аналіз на основі повного набору даних.

Основні функції MICE включають:

- Заповнення пропущених даних: MICE використовує метод ланцюгових рівнянь для заповнення пропущених значень. Він використовує ітеративний підхід, де кожна змінна замінюється прогнозованим значенням на основі інших доступних змінних.

- Моделювання: для заповнення пропущених значень пропонується використання різних моделей: лінійна регресія, дерева рішень та інші моделі в залежності від характеру даних та типу змінних.

- Повторення: MICE виконує кілька ітерацій процесу заповнення пропущених значень, щоб забезпечити більш точний результат. Кожна ітерація включає у себе моделювання та заповнення пропущених значень змінних.

- Оцінка результатів: Після завершення процесу система надає набір заповнених даних, що відображає різні можливі значення пропущених даних.

3. Nnet є пакетом мови R, який надає інструменти для побудови та навчання штучних нейронних мереж. Він дозволяє виконувати прогнозування, класифікацію та апроксимацію за допомогою нейронних мереж з різними архітектурами. Nnet є корисним інструментом для роботи з нейронними мережами в мові R та дозволяє виконувати різноманітні

завдання прогнозування та класифікації на основі складних наборів даних [37].

Основні функції пакету Nnet включають:

- Побудова нейронної мережі: пакет дозволяє визначити архітектуру нейронної мережі, включаючи кількість шарів, кількість нейронів у кожному шарі та тип активаційної функції. Ви можете створювати як прості одношарові мережі, так і складніші багатшарові мережі з рекурентними зв'язками.
- Навчання мережі на основі набору даних, з доступним вибором параметрів навчання, такі як швидкість навчання і кількість ітерацій, для покращення прогнозової точності мережі.
- Прогнозування та класифікація: Після навчання нейронної мережі за допомогою Nnet, її можна використовувати для прогнозування та класифікації нових даних. Мережа може дати прогнозні значення або прийняти рішення щодо класифікації на основі вхідних даних.
- Можливість налаштування різних параметрів нейронної мережі, таких як регуляризація, критерії зупинки навчання та інші. Це дозволяє контролювати процес навчання та підлаштовувати мережу для досягнення оптимальних результатів.

Більшість пакетів для мови R будуються саме на мові програмування R, адже R є інтерпретованою мовою програмування, спеціально розробленою для аналізу даних і статистичних обчислень. Також, в мові R є можливість писати пакети, які містять код, написаний на інших мовах програмування, наприклад на мові програмування C. Це дає можливість використовувати швидкість і ефективність низькорівневих мов програмування для обробки великих обсягів даних або виконання обчислювально важких завдань.

У мові R існує спеціальний механізм, який дозволяє інтегрувати сторонній код у пакети R. Це здійснюється за допомогою таких інструментів, як Rcpp (R-C++ інтеграція) і RcppArmadillo (для

використання бібліотеки Armadillo у мові C++). Ці пакети дозволяють зручно писати функції на C++ та викликати їх з коду R.

Коли пакет містить код на C або C++, він може забезпечити покращену швидкодію та оптимізацію виконання, особливо для обчислювально вимогливих завдань. Це особливо важливо при роботі з великими наборами даних або при реалізації складних алгоритмів.

Пакети для R зазвичай розповсюджуються через Comprehensive R Archive Network (CRAN) або інші репозиторії пакетів, такі як Bioconductor для біоінформатики або GitHub для розробок спільноти. Вони можуть бути встановлені та використані користувачами R для виконання специфічних завдань аналізу даних, візуалізації, машинного навчання, статистичного моделювання та багато іншого.

Розробка пакетів для R зазвичай відбувається відкритою спільнотою, що означає, що будь-який користувач R може взяти участь у розробці, внести свій внесок, виявити помилки або навіть створити свій власний пакет для вільного поширення. Цей принцип сприяє активному розвитку екосистеми R та постійному збагаченню її функціональності.

1.4 Висновки до розділу

У даному розділі розглянуто питання актуальності дослідження операційного ризику та розробки системи підтримки прийняття рішень для його аналізу. Також були розглянуті різні моделі для аналізу і методи оцінювання операційного та фінансового ризику. Так, наприклад, була розглянута методологія байєсівського програмування як підхід до моделювання, аналізу та оцінювання ризиків. Іншим прикладом є використання методів на основі марковських ланцюгів Монте-Карло (MCMC) для оцінювання параметрів складних моделей, що використовується у фінансовому аналізі та моделюванні ціноутворення фінансових інструментів.

Розглянуто модель Value at Risk (VaR) як один з потужних інструментів для кількісної оцінки потенційних фінансових втрат у портфелях або активах організації внаслідок негативних ринкових подій. Крім цього, в розділі були розглянуті інші методи оцінки операційних ризиків - Internal Measurements Approach, Scenario-based Advanced Measurement Approach та Байєсівська модель з автоматичним оцінюванням невизначеності (SBAMARD або SBLARD)

Також у розділі було розглянуто основні програмні інструменти та мови розробки, які використовуються для створення систем підтримки прийняття рішень та проведення оцінки операційних ризиків. Створено детальний опис передумов для використання таких інструментів як FICO Analytic Modeler Scorecard, Minitab® Statistical Software, Credit Scoring for SAS Enterprise та IBM SPSS Modeler. Формалізовані підходи до застосування програмних мов R та Python для вирішення задач розробки систем моделювання операційних ризиків.

РОЗДІЛ 2. МОДЕЛІ ОПЕРАЦІЙНИХ РИЗИКІВ

Фінансове моделювання та аналіз ризиків відносяться до популярних дослідницьких тем через практичну необхідність використання їх математичних моделей, оцінки можливих втрат у багатьох сферах людської діяльності, прогнозування та прийняття високоякісних управлінських рішень у фінансовій та багатьох інших сферах, де обертається капітал, зобов'язання, акції, облігації та інші види діяльності. Загалом фінансові процеси можуть демонструвати досить складні форми еволюції в часі, що вимагає застосування сучасних методів, технологій і процедур моделювання, прогнозування та підтримки прийняття рішень. Їх можна класифікувати як нелінійні нестационарні процеси (ННП), які вимагають застосування сучасних підходів до їх аналізу і дуже часто складного аналітичного опису.

Можливі нестационарності процесу виникають через наявність детермінованих і стохастичних тенденцій, зміну дисперсії в часі; а нелінійності можна розділити на два великі класи: нелінійності по відношенню до змінних і нелінійності по відношенню до параметрів. З нелінійностями відносно змінних легше впоратися (побудувати модель), оскільки моделі (параметри), що їх містять, можна оцінити легше, ніж моделі, нелінійні відносно параметрів. Наприклад, модель, що описує процес з нелінійним трендом (поліноміальний тренд другого або вищого порядку), може бути коректно оцінена за допомогою нелінійного методу найменших квадратів (МНК) або методу максимальної правдоподібності (ММП). З іншого боку, моделі, нелінійні за параметрами, можуть вимагати для своєї оцінки застосування складних оптимізаційних процедур, включаючи байєсівські методи, такі як метод Монте-Карло для ланцюгів Маркова або метод байєсівської оптимізації. Однак більшість фінансових процесів, з якими сьогодні доводиться мати справу дослідникам, є нелійними та нестационарними, що вимагає особливої уваги до їх

структури та оцінювання параметрів. Зазвичай для аналізу цих процесів будуються спеціалізовані інтелектуальні системи підтримки прийняття рішень (СППР), що базуються на системних принципах, які допомагають суттєво підвищити адекватність моделей, якість прогнозів, відповідних рішень та покращити загальне розуміння практичних ситуацій, пов'язаних з проблемами моделювання та оцінювання прогнозів.

2.1 Невизначеності у моделюванні фінансових процесів

У різних сферах діяльності суб'єкта, так чи інакше виникають певні невизначеності, що неодмінно впливають на подальший процес моделювання та аналізу побудованої системи. Коли, в досліджуваній системі чи процесі відсутній повний або достатній рівень знань, впевненості або передбачуваності, ми неодмінно стикаємося з невизначеностями. Це означає, що неможливо точно визначити або передбачити результат або поведінку системи, оскільки існують непередбачувані фактори, що впливають на ситуацію.

Невизначеність – це фактор негативного впливу на процес попередньої обробки даних, побудови математичної моделі, оцінювання прогнозу та генерування рішення на його основі, який призводить до погіршення якості проміжних та остаточних результатів моделювання і прогнозування [1, 32, 38, 39].

Управління невизначеністю включає в себе розробку методів та стратегій для врахування та управління невизначеністю, зменшення ризиків та прийняття рішень в умовах невизначеності.

Невизначеність у моделюванні відноситься до незрозумілих або непередбачуваних аспектів системи, які можуть впливати на результати створення моделі. Це важливий аспект в багатьох галузях, включаючи науку, технологію, економіку, фінанси, соціальні науки та інженерію.

Деякі основні причини виникнення невизначеності в моделюванні включають наступне:

- Неповнота і недостовірність даних. Моделі часто ґрунтуються на даних, які можуть бути неповними, неточними або недостовірними. Невірні дані можуть призвести до неправильних висновків та прогнозів.
- Відсутність повного розуміння системи. У деяких випадках система може бути настільки складною, що повне розуміння всіх її аспектів стає неможливим. Це може включати незрозумілі взаємодії між різними компонентами системи або непередбачувані наслідки певних подій.
- Випадковість та стохастичність. Багато процесів мають стохастичну природу, де результати не можуть бути передбаченими з абсолютною впевненістю. Це може включати фінансові ринки, де ціни акцій та валют можуть коливатися випадковим чином.
- Вплив зовнішніх факторів. Моделі можуть не враховувати або недостатньо враховувати вплив зовнішніх факторів, таких як політичні події, зміни законодавства, екологічні фактори тощо. Ці фактори можуть мати суттєвий вплив на систему та призвести до невизначеності в моделюванні.
- Суб'єктивність та припущення. Моделі можуть ґрунтуватися на суб'єктивних оцінках, припущеннях та спрощеннях, які можуть впливати на точність результатів. Різні моделі можуть використовувати різні припущення, що призводить до різних рівнів невизначеності.

Урахування невизначеності є важливим аспектом моделювання. Деякі підходи до управління невизначеністю включають використання статистичних методів, введення додаткових змінних для моделювання невизначених факторів, проведення чутливості аналізу та використання сценарного аналізу.

Невизначеність є важливою складовою моделювання фінансових процесів через багатофакторну природу ринків та пов'язану зі змінними економічними умовами. Основні джерела невизначеності включають такі [26]:

- непередбачуваність ринкових умов: фінансові ринки піддаються впливу багатьох факторів, таких як економічні зміни, політичні події, природні катаклізми та інші зовнішні фактори. Ці фактори можуть бути складними для передбачення та моделювання, оскільки вони часто є непередбачуваними та невизначеними;
- недостатня якість та недостовірність даних: моделі фінансових процесів побудовані на основі історичних даних, прогнозів та різних параметрів. Однак, ці дані можуть бути неповними, неточними або недостовірними; недостовірність або неправильне трактування даних може призвести до невірної моделювання та невірних прогнозів;
- вплив людського фактора: поведінка та рішення учасників фінансових ринків можуть бути непередбачуваними; рішення, прийняті інвесторами, трейдерами та іншими учасниками ринку, можуть впливати на ціни активів та фінансові ризики; людські фактори, такі як емоції, страх, жадібність та недостатнє раціональне мислення, можуть призвести до непередбачуваних коливань на ринках;
- системний ризик: фінансові ринки та інституції є чутливими до системних ризиків; на них можуть впливати фінансові кризи, коливання валютних курсів та процентних ставок, ризики ліквідності та кредитні ризики; такі системні ризики можуть мати негативний вплив на моделі та прогнози фінансових процесів;
- статистична невизначеність: моделі фінансового моделювання часто базуються на статистичних методах та припущеннях; однак, статистичні моделі можуть мати обмежену прогностичну точність, особливо при моделюванні екстремальних подій або рідкісних явищ.

Категорія невизначеності характеризується деякими змінними параметрами, які описують різні види невизначеностей: глобальну невизначеність, ситуативну, політичну, соціальну тощо. Для розв'язання задач прийняття рішень в умовах наявності невизначеностей необхідно встановити рівень аналізу і типи невизначеностей, що розглядаються.

Невірно вважати, що невизначеність стосується тільки відсутності повної інформації про досліджуваний об'єкт, це недостатність знання станів об'єкту. Крім цієї невизначеності розглядають невизначеність цілей і невизначеність критеріальної бази вибору рішень. Необхідно зазначити, що часто невизначеність ототожнюють лише з відсутністю повної інформації про той чи інший об'єкт. Насправді недостатні знання станів об'єкту – це не єдина невизначеність. Поряд з цим іноді можна розглядати, наприклад, невизначеність цілей та невизначеність критеріїв вибору рішень.

Складність прийняття рішень визначає кількість альтернативних варіантів вибору та різноманітність критеріїв оцінювання цих варіантів. У багатьох реальних задачах складність прийняття рішення визначається насамперед кількістю альтернативних варіантів та кількістю і різноманітністю критеріїв оцінювання цих варіантів.

Існує кілька типів невизначеності, які можуть бути наявними у моделюванні. Основні типи невизначеності включають такі:

- невизначеність цілей – це невизначеність вибору цілей в багатокритеріальних задачах.
- ситуаційна невизначеність – це невизначеність впливу неконтрольованих факторів, що позначаються на процесах практичної діяльності. Зокрема, невизначеність може бути зумовлена: – невизначеністю природи за відсутності достатніх знань про оточення та зовнішні фактори; – ненадійністю очікувань, наприклад, як невизначеність розвитку певних подій у майбутньому;

- стратегічна невизначеність – це невизначеність цілей і дій активного або пасивного партнера чи противника (так звана невизначеність конфліктів).
- інформаційна невизначеність – це нечіткість та розпливчастість процесів і явищ та інформації про досліджувану систему, відсутність відомостей про достовірність інформації.

Крім цього, можна виділити додаткові типи невизначеності:

- статистична невизначеність: Це пов'язано зі стохастичними або випадковими варіаціями в даних або процесах. Статистична невизначеність може виникати через випадковість або непередбачуваність результатів, які не можуть бути повністю передбаченими або вимірними.
- епістемічна невизначеність: вона пов'язана з обмеженою або неповною інформацією, яка доступна для моделювання. Епістемічна невизначеність виникає, коли не маємо достатньої кількості даних, або коли є обмеження щодо якості, точності або достовірності даних.
- параметрична невизначеність: це пов'язано з незнанням або непевністю стосовно значень параметрів у моделі. Параметрична невизначеність виникає, коли значення параметрів моделі є неясними або потребують оцінки. Це апріорна невизначеність параметрів моделі системи, складність оцінювання і аналізу якості параметрів моделі. Сюди можна віднести, також, невизначеності, зумовлені наявністю збурюючих впливів та похибок (шуму) вимірів
- структурна невизначеність: це пов'язано з незнанням або непевністю стосовно структури або взаємозв'язків у моделі. Структурна невизначеність виникає, коли не маємо повного розуміння системи, що моделюється, або коли є непередбачувані або складні взаємодії між компонентами системи.
- невизначеність вхідних даних: це пов'язано з непевністю або варіаціями вхідних даних, які використовуються для моделювання.

Вхідні дані можуть містити помилки, шум, артефакти або бути залежними від зовнішніх факторів, що може призвести до невизначеності в результаті моделювання.

- методична невизначеність – невизначеність (неоднозначність), притаманна методу оброблення даних чи методу розв’язання задачі;
- комбінаторна невизначеність – неможливість знання всіх можливих варіантів. Комбінаторна невизначеність пов’язана із усіма іншими типами невизначеностей і найчастіше впливає з них. Необхідно зазначити, що в реальних практичних задачах прийняття рішень і системного аналізу часто наявними є різноманітні види невизначеностей, які разом утворюють деякий комплекс невизначеностей, і називають системною невизначеністю [1, 40].

Деякі типи невизначеностей, причини виникнення та методи їх подолання подані у Таблиці 2.1

Таблиця 2.1

Типи невизначеностей у моделюванні і методи їх подолання

	Тип невизначеності	Причини виникнення	Методи подолання
	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; – байєсівські мережі, узагальнені лінійні моделі;

Продовження таблиці 2.1

	Тип невизначеності	Причини виникнення	Методи подолання
	Статистичні невизначеності у процесі побудови моделей	<ul style="list-style-type: none"> – похибки вимірів; – стохастичні зовнішні збурення; – мультиколінеарність; – екстремальні значення; – пропуски вимірів; 	<ul style="list-style-type: none"> – цифрові та оптимальні фільтри; байєсівські фільтри; – уточнення типів спільних розподілів; – метод головних компонент; – теорія екстремальних значень; – методи заповнення пропусків;
	Параметричні невизначеності	<ul style="list-style-type: none"> – некоректний вибір методу оцінювання; – короткі вибірки; – похибки вимірів; – випадкові зовнішні збурення; 	<ul style="list-style-type: none"> – забезпечення вибору альтернативних методів оцінювання параметрів; – метод Монте-Карло для марковських ланцюгів; – розмноження вибірок;
	Ймовірнісні невизначеності	<ul style="list-style-type: none"> – складні механізми виникнення причинно-наслідкових зв'язків; – відсутність детермінованості; 	<ul style="list-style-type: none"> – статичні і динамічні мережі Байєса; – марковські моделі; – ймовірнісні фільтри; – умовні багатовимірні розподіли;
	Невизначеність амплітудного типу	<ul style="list-style-type: none"> – наявність змінних, що не вимірюються; 	<ul style="list-style-type: none"> – методи обробки нечіткої інформації; – байєсівські мережі [1,3].

Ці типи невизначеності можуть взаємодіяти між собою та впливати на результати моделювання. Для кращого розуміння і керування невизначеністю важливо враховувати ці типи та використовувати відповідні методи та підходи для їх врахування.

2.2 Байєсівське програмування

Байєсівське програмування (англ. Bayesian programming) – це підхід до моделювання та аналізу, який базується на ідеях байєсівської статистики та ймовірнісних моделей. В основі байєсівського програмування лежить використання байєсівських мереж (або графічних моделей) для представлення залежностей між змінними і використання байєсівських методів для оцінки ймовірностей та формування висновків.

Основні принципи байєсівського програмування включають:

- використання апріорних знань: байєсівське програмування дозволяє враховувати апріорні знання або переконання про систему або процес, який моделюється. Апріорні ймовірності можуть бути використані для визначення початкових уявлень про стан системи перед отриманням даних.
- оновлення за допомогою правила Байєса: байєсівське програмування використовує правило Байєса для оновлення ймовірностей після отримання нової інформації або даних. Це дозволяє адаптувати ймовірності на основі нових спостережень та оновлених апріорних знань.
- висновки та прогнозування: байєсівське програмування дозволяє здійснювати побудову висновків, тобто виведення нової інформації або оцінювання невідомих змінних на основі наявних даних і апріорних знань. Також можна здійснювати прогнози на майбутнє на основі моделей та методів формування висновків.

Байєсівське програмування знайшло застосування у багатьох галузях, таких як штучний інтелект, машинне навчання, розпізнавання образів, обробка природної мови, фінансова аналітика та інші. Його перевагою є здатність враховувати невизначеність, оновлювати моделі на основі нових даних та використовувати апріорні знання для кращого аналізу та прогнозування.

Байєсівський підхід до побудови моделей та формування ймовірнісного висновку передбачає виконання етапів, поданих на рис 2.1. У байєсівському аналізі даних передбачається, що інформація надходить з двох джерел: апріорна інформація від дослідника стосовно досліджуваної задачі і нові статистичні (експериментальні) дані, отримані в результаті виконання експериментів [1,2].

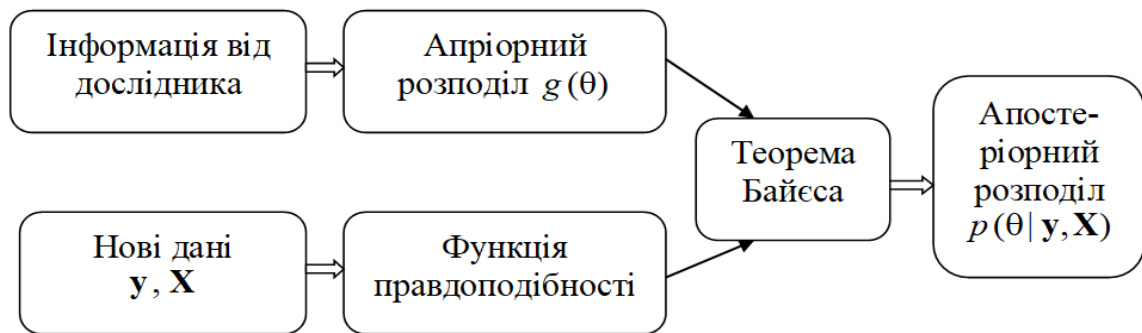


Рисунок 2.1 Інформаційні потоки у системі моделювання на основі теореми Байєса

Апріорна інформація від дослідника відображає результати попередніх досліджень; теоретичні основи досліджуваних явищ, процесів та об'єктів; додаткові неформалізовані дані, отримані з різних джерел. Загалом – це додаткова інформація стосовно досліджуваних процесів (об'єктів, яка не повинна ґрунтуватись на експериментальних даних, що стосуються конкретної задачі [41].

В центрі байєсівської методології аналізу даних знаходиться теорема Байєса (ТБ). Запишемо її з урахуванням позначень, що широко використовуються у економетричному аналізі даних:

$$p(\theta|y, X) = \frac{g(\theta)p(y|\theta, X)}{p(y)} \propto g(\theta)p(y|\theta, X), \quad (2.1)$$

де \mathbf{y} – вектор вимірів основної (залежної) змінної досліджуваного процесу; \mathbf{X} – матриця вимірів незалежних змінних (регресорів), які визначають поведінку основної змінної; $\boldsymbol{\theta}$ – вектор випадкових параметрів, які (разом з \mathbf{X}) визначають функцію щільності розподілу ймовірностей змінної y (у класичній регресійній моделі $\boldsymbol{\theta}$ – вектор параметрів регресійної моделі); $p(\cdot)$ – функція щільності розподілу ймовірностей (ЩРЙ); $p(\boldsymbol{\theta})$ – апіорна щільність розподілу ймовірностей для значень випадкового параметра $\boldsymbol{\theta}$, яка ґрунтується на апіорних знаннях дослідника до використання даних (\mathbf{y}, \mathbf{X}) ; $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X})$ – умовна щільність даних \mathbf{y} при конкретних значеннях $\boldsymbol{\theta}$ і \mathbf{X} , іншими словами, це функція правдоподібності для даних \mathbf{y} ; $p(\mathbf{y})$ – *маргінальна правдоподібність* для \mathbf{y} (оскільки у процесі обчислень $p(\mathbf{y})$ вилучено вплив незалежних змінних і параметрів); символ \propto означає «наближення».

Щільність $p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{X})$ – апостеріорний розподіл $\boldsymbol{\theta}$, який ґрунтується на оновленні апіорного розподілу завдяки даним (\mathbf{y}, \mathbf{X}) . Тобто ТБ показує як можна об'єднати інформацію, що надходить з двох джерел (апіорний розподіл і дані) з метою уточнення апіорного розподілу. Фактично, ТБ не тільки дає можливість об'єднувати інформацію з двох (або більше) джерел, але її суть у тому, що будь-яке інше комбінування даних буде порушувати логічну (і математичну) сутність правил оперування з розподілами ймовірностей.

У виразі (2.1) перший варіант теореми записано як рівність, оскільки у ньому наявний знаменник $p(\mathbf{y})$ (безумовна щільність для \mathbf{y}), який відіграє роль *нормуючої константи* і забезпечує те, що апостеріорна умовна щільність для $\boldsymbol{\theta}$ є *належною* і інтегрується до 1 по області визначення параметрів. Другий варіант ТБ у виразі (2.1) подано з точністю до константи пропорційності, тобто без нормування, що часто робиться для спрощення представлення виразів. При розв'язуванні практичних задач спочатку обчислюється чисельник, а потім, за необхідності, нормуюча константа. У багатьох прикладах існує необхідність у

обчисленні відношень результатів, а тому нормуюча константа скорочується [42, 43].

Кількість потоків інформації, що подаються на ТБ з метою обчислення апостеріорного розподілу параметра θ , може бути більшою одиниці. Тобто на вході може бути вектор змінних. Відносний вплив двох і більше джерел інформації на остаточний результат залежить від точності їх представлення. Наприклад, чим меншою буде дисперсія апіорного розподілу, тим більшу роль (з точки зору точності остаточного результату) він відіграватиме у формуванні апостеріорного розподілу.

Байєсівське програмування використовує байєсівські методи для моделювання, аналізу та формування висновків у різних задачах. Основні (найбільш відомі) методи байєсівського програмування включають:

- Байєсівські мережі: байєсівські мережі (або графічні моделі) використовуються для візуалізації та моделювання залежностей між змінними. Вони складаються з вузлів, які представляють змінні, та сполучень між вузлами, які відображають ймовірнісні залежності. Байєсівські мережі дозволяють здійснювати побудову висновків про невідомі змінні на основі наявних даних та апіорних знань.
- Марківські ланцюги Монте-Карло: метод Марківських ланцюгів Монте-Карло (МКМЛ, англ. Markov Chain Monte Carlo, MCMC) використовується для оцінки розподілу параметрів моделі на основі вибірки з цього розподілу. Вони дозволяють отримувати наближені значення параметрів та оцінки ймовірностей шляхом стохастичного аналізу у просторі параметрів.
- Апроксимація висновків: апроксимаційні методи, такі як варіаційний метод Байєса або методи згорткового навчання, використовуються для наближення складних розподілів або апостеріорних розподілів з використанням спрощених моделей або апроксимаційних функцій.
- Байєсівська оптимізація: байєсівська оптимізація використовує байєсівські методи для пошуку оптимального рішення у просторі

параметрів. Вона комбінує формування висновків та оптимізацію для знаходження найкращого розв'язку з урахуванням невизначеності та апіорних знань.

- Байєсівське навчання: Байєсівське навчання використовує байєсівські методи для оновлення (адаптації) моделі на основі нових даних. Воно дозволяє адаптувати модель до середовища, що змінюється, та здійснювати прогнози з урахуванням невизначеності.

Ці методи можуть використовуватися окремо або комбінуватися в залежності від конкретної задачі та контексту моделювання. Байєсівське програмування дозволяє враховувати невизначеність, оновлювати моделі та робити розумні висновки на основі наявних даних та апіорних знань [44, 45].

До відомої методології байєсівського програмування відносять такі методи:

- Рекурсивне байєсівське оцінювання: фільтрація, прогнозування, згладжування змінних;
- Приховані марковські моделі (ПММ) – це модифікація байєсівського фільтра, у якому припускається, що дані дискретні; моделі переходів станів і спостережень задаються матрицями ймовірностей або таблицями умовних ймовірностей. Якщо спостережувані змінні неперервні, то такі моделі називають напівнеперервними ПММ.
- Оптимальні рекурсивні фільтри Калмана (КФ); змінні неперервні або дискретні; моделі переходів станів і спостережень задаються з використанням гаусівських процесів – зовнішніх випадкових збурень і похибок (шумів) вимірів. У випадку нелінійних моделей використовують розклад у ряд Тейлора, що дає можливість лінійно локальних моделей. Для одночасного оцінювання станів досліджуваних процесів та їх параметрів використовують розширений фільтр Калмана (РКФ).

- Гранулярні (particle) фільтри (ГФ); розподіл ймовірностей станів описують такою моделлю:

$$P(S(k-1)|O(0), \dots, O(k-1)), \quad (2.2)$$

де $S(\cdot)$ – матриця станів; $O(\cdot)$ – матриця спостережень, апроксимується множиною гранул (particles), вагові коефіцієнти яких пропорційні ймовірностям їх появи. Для оновлення ймовірностей станів використовується рекурсивна процедура.

- Статичні байєсівські мережі (static Bayesian networks, BN) – це ймовірнісно-статистичні моделі для опису ймовірнісної і статистичної інформації в умовах наявності невизначеності. На змінні мережі практично не накладаються обмеження і немає спеціальної семантики для їх опису – тобто існує певна свобода вибору змінних для побудови мережі. Змінними мережі можуть дискретні і неперервні змінні, а також експертні оцінки, приведені до необхідної числової форми.

Байєсівська мережа (БМ) подається у формі спрямованого ациклічного графа, вершинами якого є змінні досліджуваного процесу, а дуги вказують на існуючі умовні залежності між змінними. Формально мережу можна описати трійкою:

$$BN = V, G, T, \quad (2.3)$$

де V – змінні (дані) для побудови мережі (база даних); G – спрямований ациклічний граф; T – таблиця умовних ймовірностей для вершин графа (змінних моделі).

Параметрами такої моделі є умовні ймовірності у таблицях умовних ймовірностей. Для батьківських (незалежних) змінних це таблиці безумовних ймовірностей. Неперервні змінні подаються відповідними

розподілами. Якщо БМ містить дискретні і неперервні змінні, її називають гібридною. Неперервні змінні у таких випадках, як правило, дискретизують, що дає можливість суттєво спростити обчислювальні операції [46].

Послідовність побудови моделі у формі БМ можна представити у вигляді таких кроків:

- поглиблений аналіз досліджуваного процесу (об'єкта) з метою встановлення особливостей його функціонування та виявлення батьківських і дочірніх змінних;
- виявлення існуючих моделей процесу та аналіз можливості їх подальшого використання в інтелектуальних системах підтримки і прийняття рішень;
- встановлення існуючих зв'язків між змінними процесу за допомогою спеціальних тестів та експертного оцінювання;
- скорочення розмірності задачі моделювання;
- масштабування і дискретизація змінних;
- визначення семантичних обмежень для моделі;
- оцінювання структур моделей-кандидатів з використанням оптимізаційних процедур, тобто пошук альтернативних моделей у формі БМ;
- аналіз якості і вибір кращої з моделей-кандидатів;
- застосування обраної моделі для розв'язання поставленої задачі;
- формування ймовірнісних висновків стосовно вибраних змінних за побудованою моделлю (моделями), аналіз якості отриманого результату.

Результатом побудови і використання БМ є ймовірнісний висновок у формі $P(X^i | \text{Known})$, де *Known* – підмножина інших змінних мережі, ймовірності станів яких відомі на момент обчислення ймовірнісного висновку. Загалом ймовірнісний висновок в БМ полягає у поширенні ймовірностей і параметрів гаусівських законів розподілу по всій мережі

залежно від отриманих *свідчень* (додаткової інформації про стани мережі). В основі процесу формування ймовірнісного висновку покладено досить складні математичні алгоритми.

Динамічні байєсівські мережі (Dynamic Bayesian Networks, DBN) створюються для того, щоб враховувати динаміку процесів (їх зміни у часі), а також можливі стохастичні впливи на їх протікання. Фактично, динамічна байєсівська мережа (ДБМ) – це розширення звичайних (статичних) мереж. Спочатку будується звичайна БМ для наявних змінних, структура якої передбачається інваріантною стосовно часу, тобто залишається сталою. Така структура повторюється для кожного наступного моменту часу з надходженням нових спостережень. Таким чином досягається відтворення динаміки (змін у часі) досліджуваних процесів.

Частину графа, що відповідає конкретному моменту часу t_k або просто k , називають часовим перерізом. Якщо приймається гіпотеза стосовно того, що стан поточного часового перерізу залежить тільки від попереднього, то таке припущення називають *марковським припущенням першого порядку*. Якщо структура всіх часових перерізів однакова, то таку ДБМ називають *стаціонарною*. У такому випадку модель, що відповідає одному часовому перерізу, називають *локальною* та *інваріантною стосовно часу* або *гомогенною* [47, 48].

Моделі *марковської локалізації* (Markov localization (ML) models) – це моделі типу байєсівських фільтрів, які додатково включають керуючі змінні $\mathbf{u}(0), \dots, \mathbf{u}(k-1)$. Іноді їх називають ще прихованими марковськими моделями по входу-виходу.

У такій моделі ймовірності станів уточнюються за допомогою керуючих змінних таким чином:

$$P(S(k)|\mathbf{u}(k-1), S(k-1)). \quad (2.4)$$

Таку модель називають ще моделлю відтворення дії (action model). Форми таких моделей можуть бути різними. Загалом це матричні моделі; а якщо вони подібні до тих, що використовуються у гранулярних фільтрах, то їх називають моделями Монте-Карло з марковською локалізацією (Monte Carlo Markov Localization, MCML).

Побудована модель призначена для того щоб давати відповідь на запитання стосовно ймовірності поточного стану досліджуваного об'єкта

$$P(S(k)|u(0), \dots, u(k-1), O(0) \dots O(k)) \quad (2.5)$$

на основі попередніх керуючих дій та спостережень за динамікою об'єкта. Термін «локалізація» пов'язаний із застосуванням у робототехніці, тобто розглядається задача локалізації (місцезнаходження) робота у навколишньому середовищі.

Основне рівняння моделі подібне до основного рівняння фільтрації і має вигляд:

$$\begin{aligned} P(S(k)|u(0), \dots, u(k-1), O(0) \dots O(k)) &= \\ &= P(u(k-1)P(O(k)|S(k)) \times \\ &\times \sum_{S(k-1)} [P(S(k)|u(k-1), S(k-1))(P(S(k-1)|u(0), \dots, u(k-2), O(0) \dots O(k-1))] \end{aligned} \quad (2.6)$$

2.3 Узагальнені лінійні моделі

Термін «узагальнена» лінійна модель зазвичай відноситься до звичайних лінійних регресійних моделей для неперервної змінної відгуку з неперервними та/або категоріальними предикторами. Узагальнені лінійні моделі (УЛМ) є статистичними моделями, які розширюють класичні

лінійні моделі, дозволяючи моделювати залежність між вихідними змінними з нелінійними залежностями та розподілами помилок. Вони є потужним інструментом для моделювання нелінійних залежностей та аналізу даних, які не відповідають передумовам класичних лінійних моделей.

Узагальнені лінійні моделі (УЛМ) створюють певною мірою універсальний підхід до моделювання різноманітних процесів, включаючи лінійні та нелінійні, стаціонарні та нестаціонарні. Це можливо завдяки тому, що до цієї групи моделей входять моделі різної структури: чисто лінійна байєсівська множинна регресія, пуассонівська регресія, дисперсійний та коваріаційний аналіз, нелінійні структури, що включають нелінійні поліноміальні елементи, нелінійні структури типу логіт-моделей та пробіт-моделей, які здатні врахувати нелінійні залежності в досліджуваних процесах. Поєднання лінійних і нелінійних структурних елементів в одній моделі здатне описати складні нелінійні нестаціонарні залежності, притаманні багатьом фінансовим процесам. Наприклад, нелінійні логіт-моделі та пробіт-моделі можуть поєднуватися з парною або множинною лінійною регресією, байєсівськими мережами (статичними або динамічними), байєсівськими (або статистичними) процедурами фільтрації даних, які можуть бути лійними або нелійними [49].

Узагальнені лінійні моделі побудовані на основі трьох компонентів: лінійної комбінації пояснювальних змінних, функції перетворення та розподілу похибок.

Форма має вигляд $y_i \sim N(x_i^T \beta, \sigma^2)$.

Основні компоненти узагальнених лінійних моделей такі:

- Лінійна комбінація пояснювальних змінних: використовується лінійна комбінація пояснювальних змінних для моделювання залежності між вихідною змінною і пояснювальними змінними. Ця комбінація може включати як лінійні, так і нелінійні компоненти.

- Функція перетворення: використовується функція перетворення для зв'язку між математичним сподіванням вихідної змінної та лінійною комбінацією пояснювальних змінних. Це дозволяє моделювати нелінійні залежності між змінними.
- Розподіл похибок: вибирається розподіл похибок моделі, який описує варіацію між спостереженнями та їхніми прогнозованими значеннями. Це дозволяє враховувати нелінійність та розподіл похибок моделі.

Узагальнені лінійні моделі включають логістичну регресію (для бінарних вихідних змінних), пуассонівську регресію (для лічильних даних), гаусівську регресію (для неперервних нормально розподілених даних) та інші. Кожен з цих типів моделей використовує відповідний розподіл похибок та функцію перетворення для моделювання конкретної ситуації.

Методологія моделювання з використанням УЛМ включає такі кроки:

- системний аналіз досліджуваного процесу на основі наявної базової теорії та експертних знань, аналіз поведінки входів і виходів, представлених інформацією часових рядів, вивчення попередніх існуючих структур моделей та іншої доступної інформації;
- попередня обробка наявних даних, яка базується на застосуванні відповідних процедур нормування та фільтрації, відповідного аналізу можливих екстремальних значень, заповнення пропущених значень тощо. Зазвичай включає в себе нормалізація та можливе коригування амплітуд вимірювань; корекція даних (врахування пропущених значень та обробку екстремальних значень); обчислення, за необхідності, різниць першого та вищих порядків, необхідних для аналізу відповідних елементів досліджуваного фінансового часового ряду.

- аналіз статистичних даних на стаціонарність та нелінійність за допомогою відповідних статистичних тестів;
- генерація та оцінювання структур моделей-кандидатів з урахуванням можливої структури УЛМ: ідентифікація стохастичних і системних компонентів та функції зв'язку; обчислення описових характеристик; визначення статистичної значущості незалежних змінних за допомогою статистики Вальда; оцінка статистичних характеристик інших структурних елементів можливої математичної моделі (наприклад, залишків);
- розробка/вибір та застосування методу (або методів) оцінювання параметрів моделі; це може бути лінійний або нелінійний метод найменших квадратів (МНК), метод максимальної правдоподібності (ММП) або байєсівський підхід на основі методу Монте-Карло;
- вибір кращої моделі з множини оцінених кандидатів за допомогою відповідних статистичних критеріїв.

Враховуючи можливі розподіли узагальнених лінійних моделей, існує декілька основних типів УЛМ, поданих нижче в Таблиці 2.2.

Таблиця 2.2 Типи узагальнених лінійних моделей

Тип моделі	Функція посилення	Розподіл залежної змінної
GLM	$g(\mu) = \mu$	нормальний
Логіко-лінійна модель	$g(\mu) = \ln(\mu)$	Пуассона
Логістична модель	$g(\mu) = \ln\left(\frac{\mu}{1-\mu}\right)$	біноміальний
Пробіт-аналіз	$g(\mu) = \Phi^{-1}\mu$	біноміальний
Аналіз "виживання"	$g(\mu) = \mu^{-1}$	Гамма-розподіл, експоненціальний

При розгляді узагальнених лінійних моделей, структура моделі оцінюється з точки зору її компонентів і включає наступні елементи:

- стохастичний компонент – незалежна змінна, яка характеризується розподілом, що належить до експоненціального сімейства;
- систематичний компонент, який включає незалежні змінні, що створюють так званий «лінійний предиктор»

$$\eta = X * \beta; \quad (2.7)$$

- особливості оцінювання функції зв'язку з урахуванням її класифікації.

Крім зазначеного вище, поняття структури моделі включає такі елементи: порядок моделі (максимальний порядок рівнянь, що утворюють модель); розмірність моделі або кількість рівнянь у моделі; вибір типу функції зв'язку, дисперсійної функції та елементів лінійного предиктора [50–53].

Після оцінювання структури моделі наступним завданням є правильна оцінка параметрів моделі з використанням наявних (модифікованих або новостворених) методів та відповідних статистичних/експериментальних даних. Визначена структура моделі надає необхідну інформацію про кількість параметрів, що підлягають оцінюванню, а також про метод оцінювання, який слід використовувати в кожному конкретному випадку. Практика моделювання свідчить, що в деяких випадках доцільно використовувати так званий принцип економії, який полягає в наступній логіці: кількість параметрів моделі, що оцінюються, не повинна перевищувати їх необхідної кількості. Під "необхідністю" тут можна розуміти необхідність збереження в побудованій моделі основних статистичних характеристик досліджуваного процесу (середнє, дисперсія та коваріація).

Для оцінювання параметрів УЛМ широко використовуються методи МНК, зважений МНК (ЗМНК), метод максимальної правдоподібності (ММП), метод моментів та методи Монте-Карло (ітеративні та неітеративні). Оскільки МНК в деяких випадках призводить до зміщених оцінок параметрів через свою чутливість до пропусків даних та пропущених значень, альтернативною процедурою для оцінювання зазвичай є ММП.

Коректність умов, необхідних для оцінювання параметрів, можна проаналізувати після обчислення оцінок параметрів. Після оцінювання параметрів отримуємо оцінки залишкового випадкового процесу у вигляді, що дає можливість проаналізувати його статистичні характеристики, які показують коректність значень параметрів.

В якості критеріїв під час діагностики моделей та вибору найкращої з побудованих моделей-кандидатів використовуються статистика Дарбіна-Уотсона; статистика Стюдента; коефіцієнт детермінації; сума квадратів помилок; інформаційний критерій Акаїке та критерій Байєса-Шварца; статистику Фішера і т.д.

Коректне використання наведеної вище методології побудови моделі дає можливість побудувати адекватну математичну модель в класі узагальнених лінійних моделей, якщо зібрані статистичні дані відповідають вимогам адекватного представлення системи і містять необхідну інформацію.

2.4 Регресійні моделі на основі статистичних даних

Лінійна регресійна модель встановлює зв'язок між залежною змінною y , k та змінними, $x_1 \dots x_n$, у вигляді:

$$y = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon, \quad (2.8)$$

де ε – похибка регресії, а x_1 неявно дорівнює 1, щоб врахувати перехоплення.

Будь-якому економісту неважко придумати багато прикладів, коли певна змінна залежить від інших. Наприклад, заробітна плата людини залежить від її освіти, досвіду та інших характеристик. Рівень ВВП в країні залежить від кількості та якості робочої сили, основного капіталу та багатьох інших характеристик. Виробничі витрати фірми залежать від кількості виробленої продукції, а також цін на ресурси тощо. Основна увага зосереджена на фактори, які впливають на залежну змінну (y). При великій кількості параметрів, визначати змінну стає складним завданням. Тому для вирішення подібних задач використовується матричну алгебру. Для того, щоб представити основні поняття та принципи побудови лінійної регресійної моделі, варто розглянути простий випадок, коли є лише одна незалежна змінна (x).

Нехай y_i та x_i позначають спостережувані дані про залежну та незалежні змінні, відповідно, для випадку i , де $i = 1, \dots, N$. Тоді, отримаємо лінійну регресійну модель:

$$y_i = \beta x_i + \varepsilon_i, \quad (2.9)$$

де ε_i – похибка. Існує багато обґрунтувань для включення члена похибки. Вона може відображати похибку вимірювання або той факт, що лінійна залежність між x та y є лише наближенням до істинної залежності.

Припущення щодо x_i та ε_i визначає форму функції правдоподібності. Стандарти припущення такі:

1. ε_i – нормально розподілений з математичним сподіванням 0 та дисперсією σ^2 , і ε_i та ε_j незалежні одна від одної для $i \neq j$. Такі похибки називають незалежними та однаково розподіленими за $N(0, \sigma^2)$.

2. x_i або фіксовані (не випадкові змінні), або, якщо ж вони випадкові, то незалежні від ε_i з розподілом ймовірностей щільністю $p(x_i|\lambda)$ де λ – вектор параметрів, що не включають β та σ^2 .

Припущення, що незалежні змінні не є випадковими, є стандартним у фізичних науках, де поширені експериментальні методи. Тобто в рамках експерименту дослідник обирає певні значення для x і вони не є випадковими. У більшості економічних застосувань таке припущення не є обґрунтованим. Однак, припущення, що розподіл x є незалежним від похибки і має розподіл, який не залежить від параметрів, що нас цікавлять, часто є обґрунтованим.

Функція правдоподібності визначається як об'єднана функція щільності ймовірності для всіх даних, залежних від невідомих параметрів. За допомогою скороченого запису ми можемо скласти всі наші спостереження за залежною змінною у вектор довжиною N :

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (2.10)$$

або, еквівалентно (і більш компактно), $y = (y_1, y_2, \dots, y_N)'$. Аналогічно, для незалежної змінної визначимо $x = (x_1, x_2, \dots, x_N)'$. Функція правдоподібності набуває вигляду

$$p(y, x|\beta, \sigma^2, \lambda) \quad (2.11)$$

Або

$$p(y, x|\beta, \sigma^2, \lambda) = p(y|x, \beta, \sigma^2)p(x|\lambda). \quad (2.12)$$

Оскільки розподіл x нас не цікавить, ми можемо працювати з функцією правдоподібності, залежною від x , $p(y|x, \beta, \sigma^2)$. Для простоти позначень не будемо явно включати x в нашу множину умов для регресійної моделі. Слід пам'ятати, що регресійна модель (незалежно від того, обробляється вона байєсівським чи частотними методами) неявно передбачає роботу з умовним розподілом y , заданим x , а не спільним розподілом цих двох випадкових векторів.

Припущення про похибки можуть бути використані для визначення точної форми функції правдоподібності. Зокрема, використовуючи деякі основні правила теорії ймовірностей, знайдемо:

- $p(y_i|\beta, \sigma^2)$ – нормальний розподіл;
- $E(y_i|\beta, \sigma^2) = \beta x_i$;
- $var(y_i|\beta, \sigma^2) = \sigma^2$.

Використовуючи означення нормального розподілу отримаємо:

$$p(y_i|\beta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(y_i - \beta x_i)^2}{2\sigma^2} \right] \quad (2.13)$$

або, враховуючи незалежність ε_i та ε_j , з якої випливає незалежність y_i та y_j для $i \neq j$, отримуємо:

$$p(y|\beta, \sigma^2) = \frac{1}{(2\pi)^{\frac{N}{2}}} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^N (y_i - \beta x_i)^2 \right] \quad (2.14)$$

Для спрощення запису також використовується наступна форма:

$$\sum_{i=1}^N (y_i - \beta x_i)^2 = v s^2 + (\beta - \hat{\beta})^2 \sum_{i=1}^N x_i^2, \quad (2.15)$$

$$\text{де } v = N - 1, \hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2} \text{ та } s^2 = \frac{\sum (y_i - \hat{\beta} x_i)^2}{v}.$$

Якщо ж розглядати похибку як точність вимірів, тобто $h = \frac{1}{\sigma^2}$ отримаємо:

$$p(y|\beta, h) = \frac{1}{(2\pi)^{\frac{N}{2}}} \left\{ h^{\frac{1}{2}} \exp \left[-\frac{h}{2} (\beta - \hat{\beta})^2 \sum_{i=1}^N x_i^2 \right] \right\} \left\{ h^{\frac{y}{2}} \exp \left[-\frac{h v}{2 s^{-2}} \right] \right\} \quad (2.16)$$

Нескладно помітити, що $\left\{ h^{\frac{1}{2}} \exp \left[-\frac{h}{2} (\beta - \hat{\beta})^2 \sum_{i=1}^N x_i^2 \right] \right\}$ – це ядро нормального розподілу для β , а $\left\{ h^{\frac{y}{2}} \exp \left[-\frac{h v}{2 s^{-2}} \right] \right\}$ – гама розподіл для h .

Попередні дані повинні відображати будь-яку інформацію, яку дослідник має до того, як побачить дані, які він хоче включити. Отже, попередні дані можуть мати будь-яку форму. Однак, зазвичай обирають певні класи передумов, які легко інтерпретувати та/або спрощують обчислення. Спряжений попередній розподіл – це такий розподіл, який у поєднанні з ймовірністю дає апостеріорний, який потрапляє в той самий клас розподілів. Природний спряжений попередній розподіл має додаткову властивість, що він має таку саму функціональну форму, що і функція правдоподібності. Ці властивості означають, що її попередньо можна інтерпретувати так само, як і інформацію про функцію правдоподібності. Іншими словами, попередню інформацію можна інтерпретувати як таку, що виникає з деякого фіктивного набору даних з того самого процесу, що й фактичні дані [53].

2.5 Моделювання нелінійних нестационарних процесів

Ймовірнісний аналіз процесів, подій і даних різних типів передбачає два підходи: частотний, який базується на класичному підході та байєсівський, в основу якого покладається теорема Байєса. У

байєсівському аналізі даних передбачається, що інформація надходить з двох джерел: апіорна – від доступних джерел стосовно досліджуваної задачі і нові статистичні дані – в результаті виконання експериментів [3].

Оптимальні рекурсивні фільтри Калмана використовуються для моделювання процесів, які функціонують під впливом випадкових зовнішніх збурень та за наявності похибок вимірів. Гранулярні (particle) фільтри застосовуються у випадках, коли розподіл ймовірностей станів апроксимується множиною гранул (particles), вагові коефіцієнти яких пропорційні ймовірностям їх появи. Ці методи дають можливість будувати моделі за наявності множини невизначеностей (фактори негативного впливу на процес, що погіршують якість проміжних та остаточних результатів). Крім того, для боротьби з невизначеностями додатково можна скористатись байєсівськими мережами, нечіткою логікою, цифровими та оптимальними фільтрами, методами заповнення пропусків вимірів і т. ін. Моделі марковської локалізації – це моделі типу байєсівських фільтрів, які додатково включають керуючі змінні і також дають можливість моделювати і прогнозувати нелінійні процеси.

Формалізована постановка задачі оцінювання моделі волатильності наступна:

Для нестационарного процесу

$$\{y(k)\} \sim N(\bar{y}, \sigma_y^2), \quad (2.17)$$

де $E[y] \neq \text{Const}$, $\sigma_y^2 \neq \text{Const}$, $k \in [0, T]$.

Оцінити структуру моделі:

$$y(k) = f_1(\psi, \varepsilon(k), h(k), u(k)), \quad (2.18)$$

$$h(k) = f_2(\theta, h(k-i), \sigma_y^2), \quad (2.19)$$

$$\{u(k)\} \sim N(0,1), \quad (2.20)$$

і невідомі вектори параметрів θ, ψ . За вимірами $y(k), k \in [0, T]$.

Гетероскедастичні процеси, які поширені у фінансовому аналізі, є нестационарними та нелінійними (за означенням) і у загальному випадку містять тренди, а тому ставиться задача моделювання і оцінювання нелінійних нестационарних процесів. Тут $y(k)$ – основна (залежна) змінна досліджуваного процесу; $h(k)$ – умовна дисперсія процесу; θ, ψ – вектори параметрів моделі; $u(k)$ – нормовані випадкові збурення.

2.5.1 Деякі моделі нелінійних нестационарних процесів

Для опису детермінованих трендів широко застосовуються комбінації: АР+поліном; АРКС+поліном; крім поліномів застосовуються інші детерміновані функції: експоненти, комбінації гармонічних функцій.

Популярні білінійні моделі, якщо нелінійність не перевищує 2-й порядок. В Таблиці 2.3 Наведено деякі моделі нелінійних нестационарних процесів.

Таблиця 2.3 Деякі моделі нелінійних нестационарних процесів

1	АР + Поліном часу	$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + b_1 k + \dots + b_m k^m + \varepsilon(k),$ $k = 0, 1, 2, \dots$ – дискретний час; $t = kT_s$; T_s – час дискретизації
2	Узагальнена білінійна модель	$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=1}^q b_j v(k-i) +$ $+ \sum_{i=1}^m \sum_{j=1}^s c_{ij} y(k-i) v(k-j) + \varepsilon(k)$
3	Логістична регресія	$\varphi(x(k, z)) = \frac{1}{1 + \exp(-x(k, z))},$ $x(k) = \alpha_0 + \alpha_1 z_1(k) + \dots + \alpha_m z_m(k) + \varepsilon(k)$
4	Нелінійна розширена економетрична авторегресія	$y_1(k) = a_0 + a_1 y_1(k-1) + b_{12} \exp(y_2(k)) + a_2 x_1 x_2 + \varepsilon_1(k)$ $y_2(k) = c_0 + c_1 y_2(k-1) + b_{21} \exp(y_1(k)) + c_2 x_1 x_2 + \varepsilon_2(k)$
5	Узагальнена авторегресія з умовною гетероскедастичністю	$h(k) = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon^2(k-1) + \sum_{i=1}^p \beta_i h(k-i)$
6	Експоненційна узагальнена авторегресія з умовною гетероскедастичністю	$\log[h(k)] = \alpha_0 + \sum_{i=1}^p \alpha_i \frac{ \varepsilon(k-i) }{\sqrt{h(k-i)}} + \sum_{i=1}^p \beta_i \frac{ \varepsilon(k-i) }{\sqrt{h(k-i)}} +$ $+ \sum_{i=1}^q \gamma_i \log[h(k-i)] + v(k)$
7	Радіальна базисна функція	$f_\theta(x(k)) = \sum_{i=1}^M \lambda_i \exp\left(-\frac{(x(k) - \mu_i)^2}{2\sigma_i^2}\right) + \varepsilon(k),$ $\theta = [\mu_i, \sigma_i, \lambda_i]^T; M = 2, 3, \dots$

8	Представлення простору станів	$x(k) = F[a(k), x(k-1)] + B[b(k), u(k-d)] + w(k)$
9	Нейронні мережі	Обрана (побудована) структура мережі
10	Нечіткі набори та нейро-нечіткі моделі	Поєднання нечітких змінних та нейромережевої моделі
11	Динамічні Байєсівські мережі	Ймовірнісна структура байєсівської мережі, побудована на основі даних та/або експертних оцінок
12	Багатовимірні розподіли	Наприклад, застосування копули для опису багатовимірного розподілу
13	Імунні системи	Імунні алгоритми та їх поєднання з іншими моделями

2.6 Комбіновані моделі

Комбіновані моделі операційного ризику мають на меті надати цілісне і всебічне уявлення про схильність організації до операційного ризику шляхом інтеграції різних джерел ризику, врахування залежності і кореляції та використання як кількісних, так і якісних підходів. Такий підхід дозволяє організаціям приймати більш обґрунтовані рішення щодо управління ризиками, розподілу ресурсів та стратегічного планування.

Ідея комбінованих моделей полягає в тому, щоб надати комплексне уявлення про схильність організації до операційного ризику, розглядаючи кілька факторів ризику разом. Операційні ризики можуть виникати з різних джерел, включаючи людський фактор (наприклад, помилки співробітників, шахрайство), перебої в функціонуванні процесів і систем (наприклад, перебої в роботі ІТ), правові та регуляторні питання (наприклад, порушення нормативних вимог), а також зовнішні події (наприклад, стихійні лиха). Комбіновані моделі мають на меті врахувати ці різноманітні джерела ризику. Операційні ризики можуть також корелювати або залежати один від одного. Наприклад, технологічний збій може спричинити помилки у фінансовій звітності, що призведе до проблем з дотриманням нормативних вимог. Комбіновані моделі намагаються врахувати ці залежності та кореляції, щоб забезпечити більш точну оцінку загального профілю ризику.

Організації часто мають справу з операційним ризиком, використовуючи різні моделі, метрики та джерела даних. Комбіновані моделі мають на меті інтегрувати ці окремі моделі та джерела даних у цілісну систему, яка забезпечує єдиний погляд на операційний ризик. Це може допомогти організаціям уникнути розрізненого управління ризиками і забезпечити більш точну загальну оцінку. Об'єднання операційних ризиків передбачає агрегування ризиків різних бізнес-ліній, департаментів або підрозділів організації. Це передбачає обчислення кількісних оцінок та агрегування ризиків з різних джерел (факторів) для отримання консолідованої картини загальної схильності до ризику.

Сценарний аналіз є поширеним методом у комбінованих моделях. Він передбачає створення гіпотетичних сценаріїв, які імітують потенційні події операційного ризику. Ці сценарії враховують множину факторів і допомагають оцінити, як різні ризики (фактори) можуть взаємодіяти або посилювати один одного. Аналізуючи ці сценарії, організації можуть виявити вразливі місця та сфери, де комбінований вплив ризиків може бути особливо сильним. Стрес-тестування тісно пов'язане з аналізом сценаріїв. Воно передбачає піддавання операцій організації екстремальним або несприятливим умовам, щоб оцінити її стійкість. У комбінованих моделях стрес-тестування може передбачати одночасний розгляд декількох стресових факторів та їхнього потенційного спільного впливу на операційний ризик [4].

Комбіновані моделі можуть включати як кількісні, так і якісні підходи. Кількісні методи передбачають використання даних і статистичних методів, тоді як якісні методи включають експертні судження та якісні оцінки. Поєднання обох підходів дає більш повне розуміння операційного ризику. У деяких галузях регуляторні органи можуть вимагати, щоб організації мали комплексне уявлення про операційний ризик. Комбіновані моделі можуть допомогти задовольнити ці вимоги, пропонуючи більш систематичний і цілісний підхід до

управління ризиками та звітності. Це передбачає розгляд того, як різні типи операційних ризиків можуть вплинути на досягнення бізнес-цілей.

Технічно, комбіновані моделі операційних ризиків реалізуються різними способами. Одним із таких способів є використання різних типів фільтрів, що дає можливість отримати дві лінійні моделі, виходи яких подаються на вхід нелінійної логістичної регресії у якості незалежних змінних (Рис. 2.2). Нелінійна регресія може прогнозувати, наприклад, напрямок руху тренду [1].

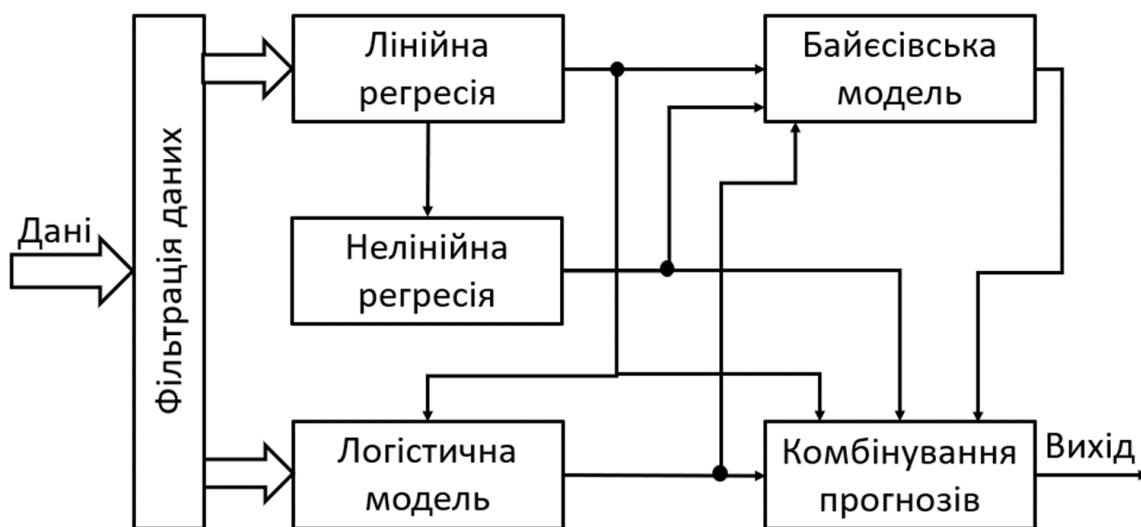


Рисунок 2.2 Комбінована модель для прогнозування нелінійних процесів

2.7 Метод Монте-Карло для Марківських Ланцюгів

За останні десятиріччя, кількість джерел формування різних даних зросла в рази, а вартість їх зберігання, разом із розвитком інформаційних технологій, навпаки, надто дешева. Це призвело до різкої зміни кількості даних у світі, що вимагає послідууючого аналізу. У зв'язку з цим, широкої популярності набувають дослідження методів Монте-Карло для роботи з даними.

Ланцюг Маркова – це послідовність випадкових величин, де розподіл ймовірностей кожної змінної залежить тільки від попередньої змінної в послідовності. Методи Монте-Карло для ланцюгів Маркова – це клас алгоритмів, що використовуються у вибірках з розподілами ймовірностей, заснованими на побудові ланцюга Маркова, який має бажаний розподіл в якості рівномірного розподілу. Ланцюг Маркова побудовано таким чином, що його стаціонарний розподіл є цільовим розподілом, з якого ми хочемо зробити вибірку [50]. Стан цього ланцюга після великої кількості кроків використовується як вибірка бажаного розподілу. Якість вибірки покращується зі збільшенням кількості кроків.

У багатьох наукових застосуваннях ми маємо доступ до деяких даних, які ми хочемо використати для того, щоб робити висновки про навколишній світ. Найчастіше ми хочемо інтерпретувати ці дані у світлі базової моделі, яка може робити передбачення щодо даних, які очікуємо побачити як функцію деяких параметрів цієї моделі. Об'єднуючи ці частини разом, можемо оцінити ймовірність того, що ми дійсно побачимо дані, які ми зібрали, припускаючи певний вибір параметрів з нашої моделі. Іншими словами, якщо припустити, що побудована модель M правильна і параметри коректно описують дані, ми оцінюємо ймовірність появи конкретних параметрів на основі даних спостережень. Припускаючи, що різні значення дадуть різні ймовірності, вони мають вказати нам, яка множина параметрів найкраще описує дані, які ми спостерігаємо [48].

Методи Монте-Карло для Марківських Ланцюгів в основному використовуються для обчислення числових наближень багатовимірних інтегралів, наприклад, у байєсівській статистиці, обчислювальній фізиці, обчислювальній біології та обчислювальній лінгвістиці. Вони також використовуються для генерації вибірок, які поступово охоплюють область рідкісних відмов у вибірках рідкісних подій [49].

У байєсівському висновку нас цікавить зворотна величина. Вона описує ймовірність того, що базові параметри дійсно є шуканим заданим

набором, враховуючи наші дані і припускаючи певну модель. Використовуючи факторизацію ймовірності, можливо пов'язати цю нову ймовірність з ймовірністю, описаною вище, так:

$$P(\boldsymbol{\theta}_M|\mathbf{D}, M)P(\mathbf{D}|M) = P(\boldsymbol{\theta}_M, \mathbf{D}|M) = P(\mathbf{D}|\boldsymbol{\theta}_M, M)P(\boldsymbol{\theta}_M|M) \quad (2.21)$$

де $P(\boldsymbol{\theta}_M, \mathbf{D}|M)$ представляє спільну ймовірність того, що ми маємо базовий набір параметрів, які описують дані, і спостерігаємо конкретний набір даних \mathbf{D} , який ми вже ми вже зібрали. В більш звичній формі ми отримуємо відому форму теорему Байєса:

$$P(\boldsymbol{\theta}_M|\mathbf{D}, M) = \frac{P(\mathbf{D}|\boldsymbol{\theta}_M, M)P(\boldsymbol{\theta}_M|M)}{P(\mathbf{D}|M)}. \quad (2.22)$$

В даному випадку, $P(\boldsymbol{\theta}_M|M)$ вважається апіорною ймовірністю, що описує вірогідність появи набору $\boldsymbol{\theta}_M$ параметрів в побудованій моделі M . Враховуючи незалежність даних, апіорні переконання про вигляд $\boldsymbol{\theta}_M$ ґрунтуються на попередніх вимірах, фізичних міркуваннях та інших наперед відомих факторах.

Знаменник – це гранична ймовірність для нашої моделі M , маргіналізована (тобто інтегрована) за всіма можливими значеннями параметра $\boldsymbol{\theta}_M$. У широкому сенсі, це спроба кількісно оцінити, наскільки добре наша модель M пояснює дані після усереднення за всіма можливими значеннями $\boldsymbol{\theta}_M$:

$$P(\mathbf{D}|M) = \int P(\mathbf{D}|\boldsymbol{\theta}_M, M)P(\boldsymbol{\theta}_M|M)d\boldsymbol{\theta}_M. \quad (2.23)$$

Іншими словами, якщо спостереження, передбачені нашою моделлю, виглядають схожими на початкові дані, то M є хорошою (адекватною)

моделлю. Моделям, в яких це найчастіше так, також надається перевага над моделями, які дають відмінну збіжність час від часу, але не збігаються більшу частину часу.

Окрім байєсівського застосування розглядають, в якості основного метода Монте-Карло для Марківських ланцюгів, розглядається підклас – випадкових блукань. Методи випадкових блукань використовуються для вибірки зі складних розподілів ймовірностей. Вони передбачають ітеративне пропонування нових станів на основі випадкових збурень і широко використовуються в різних галузях, таких як статистика, фізика і машинне навчання для вирішення таких завдань, як формулювання байєсівського висновку і оцінювання параметрів.

Концепція методів випадкових блукань передбачає такі основні ключові складові:

- Алгоритм Метрополіса-Гастінгса – є фундаментальною технікою методів Монте-Карло для Марківських ланцюгів, яка включає в себе методи випадкових блукань. Він передбачає пропозицію нових станів на основі поточного стану ланцюга і прийняття або відхилення цих пропозицій на основі критерію прийняття. Ймовірність прийняття визначається шляхом порівняння ймовірностей запропонованого і поточного станів відповідно до цільового розподілу.
- Пропозиції випадкового блукання: у методі випадкового блукання запропоновано, що новий стан генерується шляхом додавання випадкового збурення до поточного стану. Збурення може бути згенероване з симетричного розподілу пропозицій, наприклад, з розподілу Гауса. Розмір кроку випадкового блукання можна регулювати для того щоб контролювати дослідження простору станів.
- Для того щоб підтримувати детальну умову балансу, необхідну для збіжності ланцюга Маркова до правильного стаціонарного розподілу,

методи випадкових блукань часто використовують симетричні розподіли пропозицій. Ці розподіли генерують пропозиції, які з однаковою ймовірністю рухаються в будь-якому напрямку.

- Адаптивні методи: однією з проблем методу випадкових блукань є вибір відповідного розміру кроку для формування розподілу пропозицій. Адаптивні методи динамічно підлаштовують розмір кроку під час процесу вибірки, щоб оптимізувати дослідження простору станів і підвищити ефективність вибірки.
- Ефективність методу випадкового блукання Монте-Карло для Марківських Ланцюгів залежить від того, наскільки ефективно він досліджує простір станів і вибірки з цільового розподілу. Правильне налаштування розподілу пропозицій та розміру кроку має вирішальне значення для досягнення хорошого змішування, що означає здатність ланцюга досліджувати всі релевантні області розподілу.
- Збіжність: як і всі методи МСМС, метод випадкових блукань потрібно запускати на достатній кількості ітерацій, щоб забезпечити збіжність до цільового розподілу. Діагностика збіжності, така як графіки трасування та статистика Гельмана-Рубіна, може бути використана для оцінювання того, чи досяг ланцюжок (що формується) необхідного стабільного розподілу.

Найвідомішими прикладами методів випадкових блукань [20] є алгоритм Random Walk Metropolis і семплер (вибірка) Гіббса.

Вибірка Гіббса – це спосіб формування вибірки з розподілу ймовірностей з двома або більше вимірами або багатовимірною розподілу. Це метод Монте-Карло для Марківських ланцюгів, тобто є різновидом алгоритму залежної вибірки. Через це, якщо нам потрібні незалежні вибірки, деякі вибірки з самого початку зазвичай відкидаються, оскільки вони можуть неточно представляти бажаний розподіл.

Суть вибірки Гіббса полягає у тому, що за наявності багатовимірного розподілу простіше зробити вибірку з умовного розподілу, ніж зі спільного. Наприклад, замість того, щоб формувати вибірку безпосередньо зі спільного розподілу $P(x, y)$, вибірка Гіббса пропонує вибірку з двох умовних розподілів $P(x|y)$ і $P(y|x)$. В такому випадку, для визначення $P(x, y)$, алгоритм починається з випадкової вибірки $(x^{(0)}, y^{(0)})$. Після цього, вибірка $x^{(1)}$ з умовного розподілу $P(x|y^{(0)})$ та $y^{(1)}$ з умовного розподілу $P(y|x^{(1)})$. Аналогічно отримується вибірка $x^{(k)}$ з умовного розподілу $P(x|y^{(k-1)})$ та $y^{(k)}$ з умовного розподілу $P(y|x^{(k)})$, тобто:

$$x_k^{(i)} \sim p\left(x_k \mid x_1^{(i)}, \dots, x_{k-1}^{(i)}, x_{k+1}^{(i-1)}, \dots, x_d^{(i-1)}\right). \quad (2.24)$$

В основі ж алгоритму Ніколаса Метрополіса лежить припущення, що розподіл

$$P(x^{(t-1)}|y^{(t)}) = P(y^{(t)}|x^{(t-1)}), \quad (2.25)$$

тобто він симетричний, та

$$P_t = \min\left(\frac{f(y_t)}{f(x_{t-1})}, 1\right). \quad (2.26)$$

Особлива умова – розподіл повинен бути здатним генерувати всі значення, що належать до супорта (підтримки) цільового розподілу. Якщо розподіл пропозиції ніколи не генерує значення x_t , якому цільовий розподіл присвоює позитивну ймовірність, то, звичайно, стаціонарний розподіл ланцюга не може дорівнювати цільовому розподілу. З технічної точки зору, якщо значення x_t належить до супорта цільового розподілу, але ніколи не вилучається з розподілу пропозиції, то $P(x_t|x_{t-1}) = 0$, що

викликає проблеми з діленням на нуль при доведенні умови детальної рівноваги, робить її недійсною.

Алгоритм Random Walk Metropolis широко використовується в байєсівському висновку, оцінюванні параметрів і статистичному моделюванні. Він забезпечує гнучкий фреймворк для формування вибірки зі складних розподілів ймовірностей, де пряма вибірка неможлива, і має застосування в різних галузях, таких як фізика, біологія, економіка та машинне навчання. Етапи функціонування алгоритму схожі з іншими алгоритмами цього класу. Початковий стан в просторі параметрів може бути згенерований випадково або обраний на основі попередніх знань. На кожній ітерації алгоритму пропонується новий стан, додаючи випадкове збурення до поточного стану. Збурення (випадкове значення) береться із симетричного розподілу пропозицій, найчастіше це гаусівський розподіл із заданою дисперсією. При цьому обчислюється ймовірність прийняття запропонованого стану. Ця ймовірність є мінімальним з двох значень: відношення цільового розподілу ймовірностей при запропонованому стані і поточному стані та 1. Якщо ймовірність прийняття більша або дорівнює 1, то запропонований стан приймається. В іншому випадку він приймається з ймовірністю, що дорівнює розрахованій ймовірності прийняття. Якщо запропонований стан прийнято, поточний стан встановлюється у запропонований. Якщо запропонований стан відхилено, залишається поточний стан без змін. Цей новий стан стає відправною точкою для наступної ітерації. Ці кроки повторюються відповідно до попередньо визначеної кількості ітерацій, щоб згенерувати необхідну послідовність станів.

Алгоритм використовує випадкове блукання для дослідження простору параметрів. Розмір кроку випадкового блукання можна налаштувати для керування дослідженням. Більший розмір кроку може призвести до швидшого дослідження, але може призвести до нижчого рівня прийнятності, тоді як менший розмір кроку може призвести до

вищого рівня прийнятності, але повільнішого дослідження. Розмір кроку випадкового блукання також може бути адаптований під час роботи алгоритму. Адаптивні методи можуть регулювати розмір кроку для оптимізації ефективності та дослідження простору параметрів.

Як і всі методи Монте-Карло для Марківських ланцюгів, алгоритм RWM підтримує умову детального балансу, гарантуючи, що згенерований ланцюг Маркова збігається до бажаного цільового розподілу. Ефективність алгоритму залежить від вибору розподілу та розміру кроку. Правильне налаштування цих параметрів має важливе значення для досягнення гарного змішування та збіжності до цільового розподілу.

2.8 Висновки до розділу

У другому розділі розглянуто причини виникнення фінансових операційних ризиків у фінансових організаціях, а також різноманітні моделі для їх формалізації, особливості застосування і побудови кожної з них. Показано, що актуальним завданням для організацій є розробка та практичне використання систем управління ризиками на основі сучасних математичних моделей, зокрема імовірнісних моделей байєсівського типу, які надають можливість ідентифікувати та враховувати невизначеності, пов'язані з випадковим характером подій у фінансовому аналізі.

Сформульовані різні джерела невизначеності у фінансовому моделюванні та прийнятті рішень. Введена класифікація типів невизначеності, що можуть виникати в різних аспектах фінансового моделювання та відповідно впливати на процеси прийняття рішень.

Крім того, в розділі детально розглянуті різні моделі операційних ризиків та принципи їх побудови. Байєсівське програмування є потужним інструментом для моделювання, аналізу та прийняття рішень в умовах невизначеності, і воно має широкий спектр застосувань у різних галузях. Використання апіорних знань, правила Байєса та інформаційних потоків

дозволяє ефективно оновлювати моделі та здійснювати аналіз на основі нових даних. Узагальнені лінійні моделі є розширеним класом лінійних моделей, які дозволяють моделювати залежність між вихідними змінними та пояснювальними змінними з урахуванням нелінійних залежностей та розподілів помилок.

Для моделювання нелінійних процесів, подій і даних існують два основних підходи - частотний та байєсівський. Частотний підхід ґрунтується на класичній статистиці, тоді як байєсівський підхід базується на теоремі Байєса. Для моделювання процесів, що піддаються впливу випадкових збурень та мають похибки вимірювань, розглянуто використання оптимальних рекурсивних фільтрів Калмана. У випадках, коли розподіл ймовірностей станів може бути наближено множиною гранул, використовують гранулярні фільтри. Комбіновані моделі сприяють створенню цілісного та всебічного уявлення про схильність організації до операційного ризику. Вони спрямовані на створення більш повного та систематичного підходу до управління операційним ризиком в організаціях. Вони дозволяють враховувати різноманітні джерела ризику, залежності між ними та проводити аналіз в екстремальних умовах для забезпечення більшої стійкості та безпеки діяльності організації.

Також детально розглянуто метод Монте-Карло для Марківських ланцюгів. Алгоритм Метрополіса-Гастінгса використовується для пропозиції нових станів на основі поточного стану ланцюга і прийняття або відхилення цих пропозицій на основі критерію прийняття. Для покращення ефективності вибірки в методі випадкового блукання, можуть використовуватися адаптивні методи, які динамічно налаштовують розмір кроку. Це допомагає оптимізувати дослідження простору станів та підвищувати ефективність вибірки. Для досягнення збіжності до цільового розподілу, метод випадкового блукання Монте-Карло для Марківських ланцюгів повинен бути запущений на достатній кількості ітерацій.

РОЗДІЛ 3: СТРУКТУРНО-ПАРАМЕТРИЧНА АДАПТАЦІЯ БАЙЄСІВСЬКИХ МОДЕЛЕЙ

3.1 Побудова адаптивних байєсівських моделей

Оскільки фінансові процеси відносяться до класу нелінійних нестационарних, то структура і параметри статистичних даних, які використовуються для формалізованого опису та оцінювання ризиків можливих втрат, як правило, змінюються у часі. Це підтверджується відповідним статистичним тестуванням наявних даних. Так, наприклад, наявність нелінійного тренду процесу (тренд другого або вищого порядку) або наявність гетероскедастичності переводять досліджуваний процес до класу нелінійних та нестационарних. Тому у процесі використання побудованих математичних моделей для формального опису та оцінювання ризиків можливих втрат виникає завдання адаптації створених моделей до нових даних. Це стосується всіх типів моделей, що використовуються у роботі: узагальнених лінійних моделей, байєсівських мереж, звичайної регресії та інших можливих типів моделей, що використовуються для математичного опису досліджуваних фінансових процесів [41].

Задача адаптації відноситься до структури і параметрів моделі, що вже використовується, тобто, як правило, це структурно-параметрична адаптація, яка може виконуватись за допомогою різних обчислювальних процедур та критеріїв якості для моделей різних типів. У загальному випадку це одна із задач системного підходу до побудови систем підтримки прийняття рішень (СППР) для моделювання, прогнозування розвитку фінансових процесів та оцінювання ризиків можливих втрат. Серед інших задач системного підходу можна вказати на такі: створення ієрархічної багаторівневої СППР; підготовка даних до моделювання із врахуванням можливих невизначеностей даних структурного,

параметричного і статистичного характеру; оцінювання та оптимізація структури і параметрів моделей, що будуються у системі; використання окремих множин критеріїв якості на кожному етапі виконання аналізу даних (критерії якості даних, критерії адекватності моделей, якості оцінок прогнозів та якості альтернативних рішень, що приймаються/підтримуються); функціональна повнота системи підтримки прийняття рішень та інші. Використання окремих множин критеріїв якості на кожному етапі виконання обчислень стосовно аналізу даних дає можливість підвищити якість проміжних та остаточних обчислень, що програмується у системі. В цілому, це гарантує високу якість остаточних результатів, які можна отримати за допомогою створеної системи.

Незважаючи на велику кількість систем підтримки прийняття рішень, які створені і пропонуються на ринку комерційними підприємствами з розвитку інформаційних технологій для аналізу ризиків, як правило, далеко не всі, необхідні для високоякісного менеджменту ризиків, задачі розв'язуються завдяки використанню цих систем. Одна з причин виникнення такої ситуації полягає у відсутності фахівців необхідної кваліфікації, що особливо стосується аналізу складних сучасних фінансових процесів, які потребують створення інтелектуальних СППР та нових адекватних моделей за їх допомогою, ефективних моделей і методів оцінювання можливих фінансових втрат і генерування альтернативних управлінських рішень (альтернатив). Оскільки аналіз фінансових ризиків – це задачі високої складності і актуальності, то такі задачі розв'язуються численними фахівцями з теорії оцінювання, математичного моделювання, прогнозування та інформаційних технологій практично у всьому світі.

Далі розглянемо можливий варіант розв'язання задачі адаптації формалізованого опису процесів у формі узагальнених лінійних моделей (УЛМ) і байєсівської мережі до нових даних на прикладі динамічної мережі, яка є прийнятною для аналізу динамічних фінансових даних.

Структурна адаптація УЛМ (лінійних та нелінійних) передбачає аналіз та врахування таких елементів моделей: кількість рівнянь, що використовуються для створення моделі; наявність нелінійності у процесі (що підтверджується відповідними тестами); кількість незалежних змінних (регресорів); випадкове збурення та його тип (тип розподілу ймовірностей для цього випадкового процесу); порядок авто регресійної частини моделі (за умови її наявності); порядок ковзного середнього (за умови його наявності); порядок інтегрованості процесу (за умови наявності тренду); можливі обмеження на значення змінних та параметрів. Всі вказані елементи структури моделей досить просто виявляються за допомогою статистичного тестування даних з використанням відомих обчислювальних процедур.

Після корегування структури моделі необхідно виконувати повторне оцінювання параметрів з коректним використанням відповідних методів параметричної ідентифікації (оцінювання параметрів). До цих методів можна віднести такі: звичайний метод найменших квадратів (МНК); нелінійний МНК (НМНК); метод максимальної правдоподібності (ММП); метод допоміжної змінної (МДП); метод Монте-Карло для марківських ланцюгів (МКМЛ), який здобув популярність завдяки відносній універсальності його застосування до оцінювання параметрів лінійних та нелінійних процесів. Підбір/модифікація методу для оцінювання параметрів моделі ґрунтується на аналізі типу розподілу даних та виявлення наявності не лінійності та встановлення її типу.

Структурна адаптація байєсівських мереж (БМ) передбачає аналіз та корегування таких елементів цих ймовірнісно-статистичних моделей: додавання/скорочення кількості змінних (вузлів); додавання/видалення кількості дуг, що з'єднують вузли мережі; корегування типів локальних розподілів ймовірностей, що відносяться до окремих вузлів [42, 43]. Адаптація структури мережі відбувається з використанням необхідних критеріїв якості мережі (моделі), які надають інформацію про те, що

відбувається з моделлю при виконанні тієї чи іншої дії, згаданої вище. Наприклад, при додаванні/видаленні змінних або дуг може відбуватись збільшення/зменшення загальної ймовірності мережі (як спільного багатовимірного розподілу використаних змінних), що підтверджується відповідним вибраним критерієм якості моделі, який використовується у методі адаптації.

Параметрична адаптація байєсівської мережі передбачає корегування значень умовних ймовірностей у відповідних таблицях (таблиці умовних ймовірностей або ТУЙ), значення яких можуть змінюватись в результаті надходження нових даних та/або попереднього корегування структури моделі. Значення у таблицях безумовних ймовірностей, що визначаються батьківськими вузлами, також можуть корегуватись з появою нових батьківських вузлів та зміни границь їх допустимих значень. Необхідно зазначити, що загалом структурно-параметрична адаптація байєсівських мереж – це досить складна задача, яка потребує використання спеціальних обчислювальних процедур і контролю коректності їх виконання. Всі ці процедури спрямовуються на досягнення необхідної адекватності моделей (це стосується також УЛМ та інших типів моделей), що будуються для формального опису вибраних фінансових процесів. Висока ступінь адекватності моделей забезпечує високу якість оцінювання необхідних прогнозів та ризиків можливих фінансових втрат.

3.2 Адаптація байєсівської мережі до даних

У подальшому аналізі будемо розглядати дискретні змінні, які є зручнішими з точки зору простоти представлення розподілів та менших обчислювальних витрат. Такий підхід дає можливість спростити, також, процедури обчислення ймовірнісного висновку – отримання остаточного результату за мережею. Розглянемо розподіли дискретних випадкових величин, у яких кожна змінна X_i може приймати значення із скінченної

множини спостережень $Val(X_i)$. Позначимо розмірність $Val(X_i)$ через $\|X_i\|$. Надалі будемо використовувати заголовні літери, наприклад, X, Y, Z для позначення імен змінних і малі літери x, y, z для позначення конкретних значень, що приймаються цими змінними. Множини змінних будемо позначати напівжирними заголовними літерами $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$, з відповідними значеннями, позначеними напівжирними малими літерами $\mathbf{x}, \mathbf{y}, \mathbf{z}$. Змінні мережі і спільні розподіли будемо позначати так: $\mathbf{X} = \{X_1, \dots, X_n\}$.

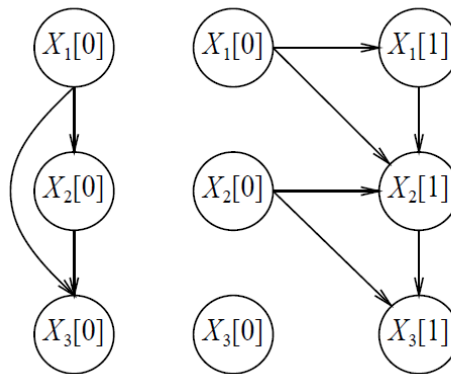


Рисунок 3.1. Простий приклад ймовірнісної мережі

Ймовірнісна (байєсівська) мережа (ЙМ або БМ) – це орієнтований ациклічний граф, який відображає спільний розподіл ймовірностей для змінних \mathbf{X} . Формально ймовірнісна мережа для \mathbf{X} може бути представлена парою $B = (G, \theta)$. Перший компонент цієї пари, G , представляє орієнтований ациклічний граф, вершини якого відповідають випадковим змінним X_1, \dots, X_n . Цей компонент кодує таку множину припущень стосовно умовної незалежності: кожна змінна X_i не залежить від його нащадків, враховуючи його батьків $\mathbf{Pa}(X_i)$ в графі G . Другий компонент, θ , являє собою множину параметрів змінних і визначає кількісно мережу. У простому випадку він містить параметри

$$\theta_{i,j_i,k_i} = Pr(X_i = k_i | \mathbf{Pa}(X_i) = j_i) \quad (3.1)$$

для кожного можливого значення k_i від X_i і кожного можливого набору значень j_i від $\mathbf{Pa}(X_i)$. Кожен умовний розподіл ймовірностей можна представити у вигляді таблиці, яка називається таблицею умовних ймовірностей (ТУЙ). Формалізовані представлення, які вимагають менше параметрів, наприклад таких як дерева, також можуть бути використані. Ми використовуємо їх надалі в розділі аналізу результатів експерименту, але для простоти позначень будемо дотримуватися згаданим позначенням ТУЙ.

Враховуючи G і θ , ймовірнісна байєсівська мережа (ЙМ або БМ) B визначає унікальний спільний розподіл ймовірності для змінних \mathbf{X} , що визначається за виразом [50]:

$$P_B(x_1, \dots, x_n) = \prod_{i=1}^n P_B(x_i | \mathbf{pa}(X_i)) \quad (3.2)$$

ЙМ описує розподіл ймовірностей за допомогою фіксованої множини змінних. Динамічна ймовірнісна мережа (ДЙМ) розширює це подання (мережі) для формального опису (моделювання) досліджуваних процесів у часі. Для простоти будемо вважати, що зміни мають місце між дискретними моментами часу, які проіндексовані невід'ємними цілими числами. Тут ми припускаємо, що $\mathbf{X} = \{X_1, \dots, X_n\}$ являє собою множину атрибутів, що змінюють процес. При цьому $X_i[t]$ є випадковою величиною, що визначає значення атрибута X_i у момент часу t і $\mathbf{X}[t]$ є множиною випадкових величин $X_i[t]$.

Для отримання моделі можливих траєкторій розвитку досліджуваного процесу, необхідно знайти розподіл ймовірностей для випадкових величин $\mathbf{X}[0] \cup \mathbf{X}[1] \cup \mathbf{X}[2] \cup \dots$. Зазвичай, такий спільний

розподіл може бути досить складним. У поданому дослідженні припускаємо, що процес є *марківським* стосовно \mathbf{X} , тобто

$$P(\mathbf{X}[t+1]|\mathbf{X}[0], \dots, \mathbf{X}[t]) = P(\mathbf{X}[t+1]|\mathbf{X}[t]). \quad (3.3)$$

Припускаємо також, що процес, який моделюється, є стаціонарним стосовно ймовірностей переходів, тобто у даному випадку стаціонарність означає, що ймовірність переходу $P(\mathbf{X}[t+1]|\mathbf{X}[t])$ не залежить від часу t .

Виходячи з поданих припущень, динамічна ймовірнісна мережа представляє собою спільний розподіл для всіх можливих траєкторій еволюції процесу і складається з двох частин:

- апріорна структура мережі (моделі) B_0 , що визначає розподіл за відомими початковими станами $\mathbf{X}[0]$;
- мережа переходів B_{\rightarrow} для значень змінних $\mathbf{X}[0] \cup \mathbf{X}[1]$, з метою визначення ймовірності переходу $P(\mathbf{X}[t+1]|\mathbf{X}[t])$ для всіх t .

Рисунок 3.1 ілюструє простий приклад мережі. В мережах переходів (але не в апріорних мережах), змінні в $\mathbf{X}[0]$ не мають батьків. Ймовірність переходу формально можна представити таким виразом:

$$P_{B_{\rightarrow}}(\mathbf{x}[1]|\mathbf{x}[0]) = \prod_{i=1}^n P_{B_{\rightarrow}}(x_i|pa(\mathbf{X}_i[1])). \quad (3.4)$$

Зазначимо, що ДЙМ, яка визначається парою (B_0, B_{\rightarrow}) , відповідає напів-нескінченній мережі для змінних $\mathbf{X}[0], \dots, \mathbf{X}[\infty]$. На практиці робимо висновок (знаходимо остаточний результат) тільки стосовно скінченного інтервалу $0, \dots, T$. Для того щоб мати можливість отримати цей результат, можемо умовно «розгорнути» структуру ДЙМ в ЙМ над змінними $\mathbf{X}[0], \dots, \mathbf{X}[T]$. Для множини даних 0 , батьками $X_i[0]$ будуть ті, які вказані в

попередній мережі B_0 ; в секції $t + 1$, батьками $X_i[t + 1]$ є ті ж вузли, в секціях t і $t + 1$ відповідно до батьків $X_i[1]$ в B_{\rightarrow} . Тобто, умовні розподіли можна скопіювати для цих змінних аналогічним чином. Рисунок 3.2 показує результат розгортання мережі на Рис.3.1 протягом трьох часових інтервалів.

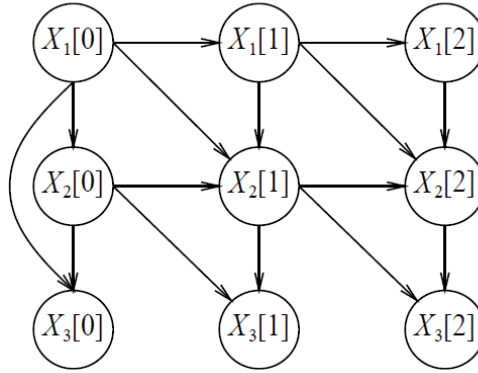


Рисунок 3.2. Розгортання мережі протягом трьох часових інтервалів

Враховуючи, що модель ДЙМ це спільний розподіл по $X[0], \dots, X[T]$, процес розгортання можна формально представити у такому вигляді [44, 45]:

$$P_B(x[0], \dots, x[T]) = P_{B_0}(x[0]) \prod_{t=0}^{T-1} P_{B_{\rightarrow}}(x[t+1]|x[t]), \quad (3.5)$$

де $P_{B_{\rightarrow}}(x[t+1]|x[t])$ можна отримати очевидним чином з моделі переходу.

3.3 Навчання мережі для випадку повних вибірок даних

У цьому підрозділі розглянемо задачу навчання ДЙМ для випадку наявності *повних вибірок даних*. Результатом цього розділу є те, що навчання ДЙМ для випадку повних вибірок даних використовує ті ж

методи, що і навчання ЙМ для випадку повних вибірок даних. Але фактично навчання ДЙМ не є повністю таким самим, що і застосування методів навчання статичної ЙМ у формі розгорнутої мережі у зв'язку з обмеженням на можливість створення повторної структури моделі і обчисленням відповідних параметрів цієї структури.

3.3.1 Навчання ймовірнісної мережі

Задача навчання ймовірнісної мережі формулюється таким чином. Використовуючи навчальну вибірку з D зразків (значень) X , необхідно побудувати мережу $B = (G, \theta)$, яка найкращим чином відповідає D . Поняття "найкраща відповідність" визначається за допомогою функції підрахунку відповідних функцій оцінок. В літературі запропоновано кілька різних функцій оцінювання. Найбільш часто використовуються байєсівський інформаційний критерій (БІК) оцінки і Bayesian-Dirichlet (БДе) критерій [46]. Обидві оцінки поєднують ймовірність даних із структурою мережі,

$$L(B: D) = \log \Pr(D|B). \quad (3.6)$$

БІК і оцінки БДе є похідними від ймовірності отримання належної структури мережі. Нехай випадкова величина G змінюється над можливими варіантами структури мережі, які можна отримати за допомогою фактичних даних. Використовуючи правило Байєса, можна отримати такий розподіл для G :

$$\Pr(G|D) \propto \Pr(D|G) \Pr(G), \quad (3.7)$$

де ймовірність даних для цієї структури мережі може бути обчислена за відповідними параметрами розподілу змінних мережі:

$$Pr(D|G) = \int Pr(D|G, \theta) Pr(\theta|G) d\theta. \quad (3.8)$$

Очевидно, що оцінювання апіорних параметрів і обчислення цього інтегралу може бути у загальному випадку складною задачею. Одним з підходів до запобігання повного обчислення інтеграла (3.8) є аналіз асимптотичної поведінки цього виразу. Враховуючи можливу наявність великої кількості значень даних, можна показати, що наступна ймовірність є нечутливою до вибору апіорної ймовірності (за умови, що апіорна ймовірність є ненульовою в будь-якому випадку). У роботі [40] отримано асимптотичну оцінку, зручну для аналізу апіорної ймовірності:

$$\log Pr(D|G) = \log Pr(D|G, \hat{\theta}_G) - \frac{\log N}{2} \#G + O(1), \quad (3.9)$$

де $\hat{\theta}_G$ – параметр налаштування для G , який максимізує ймовірність даних; N є кількістю навчальних зразків (значень); $\#G$ – розмірність G (яка в разі повної інформації є тільки кількістю параметрів); $O(1)$ – постійний член, який не залежить від N і G . Оцінка БІК використовує рівняння (3.9) для ранжирування кандидатів на структуру мережі, що усуває необхідність пошуку апіорних параметрів.

Альтернативний підхід до аналізу даних і параметрів полягає в оцінюванні виразу (3.8) в замкнутій формі. Наближено, апіорні ймовірності параметрів для цієї структури G передбачається розкласти на множники з використанням незалежних апіорних ймовірностей для кожного умовного розподілу $P(X_i | pa(X_i))$. У випадку наявності повних вибірок даних це означає, що апостеріорні фактори діють так само незалежно. Таким чином, кожен умовний розподіл може бути створено і оновлено/модифіковано окремо, що спрощує аналіз даних. Оцінювання цих розподілів може бути обчислене в замкнутій формі, якщо додатково припустити, що апіорні ймовірності кожного умовного розподілу

належать відомому спряженому сімейству, наприклад, широко використовуваному розподілу Діріхле. Деякі подробиці такого підходу подано нижче.

Оскільки існує можливість дослідити велику кількість можливих структур мережі, а також уникнути необхідності призначати попередньо оцінений розподіл параметрів для кожної можливої структури, то необхідно забезпечити множину припущень, які дозволять визначати апіорні ймовірності для всіх структур за допомогою єдиної мережі на основі розподілу Діріхле, який дає впевненість стосовно коректного встановлення значень апіорних ймовірностей [40]. Цей підхід має ту бажану властивість, що оцінки деяких двох мереж будуть еквівалентними, тобто однаковими стосовно можливого формального опису з використанням однакових незалежних припущень [47].

3.3.2 Оцінки Байєса-Діріхле (БДе) для ймовірнісної мережі

Опишемо можливе оцінювання за критерієм БДе для динамічної ймовірнісної мережі. У цьому підрозділі припускається, що відома навчальна вибірка D , яка складається з N_{seq} повних (необхідних) послідовностей спостережень. Така l послідовність має довжину N_ℓ і визначає значення для змінних $\mathbf{x}^\ell[0], \dots, \mathbf{x}^\ell[N_\ell]$. Така множина даних дає N_{seq} початкових зразків (вимірів), на яких можна навчити структуру B_0 , а також $N = \sum \ell N_\ell$ випадків можливих переходів (структурних змін), на яких можна навчити модифіковану структуру моделі B_\rightarrow .

Введемо позначення:

$$\theta_{i,j_i',k_i'}^{(0)} = Pr(X_i[0] = k_i' \mathbf{Pa}(X_i[0]) = j_i') \quad (3.10)$$

і так само для моментів часу $t = 1, \dots, T$:

$$\theta_{i,j_i',k_i'}^{\rightarrow} = Pr(X_t[t] = k_i' \mathbf{Pa}(X_i[t]) = j_i). \quad (3.11)$$

Також необхідно ввести позначення для достатньої статистики для кожної наявної множини даних:

$$N_{i,j_i',k_i'}^{(0)} = \sum \ell I(X_i[0] = k_i' \mathbf{Pa}(X_i[0]) = j_i', \mathbf{x}^\ell) \quad (3.12)$$

$$N_{i,j_i',k_i'}^{\rightarrow} = \sum \ell \sum t I(X_i[t] = k_i, \mathbf{Pa}(X_i[t]) = j_i, \mathbf{x}^\ell), \quad (3.13)$$

де $I(\cdot; \mathbf{x}^\ell)$ – індикаторна функція, яка приймає значення 1, якщо подія відбувається (має місце) в послідовності \mathbf{x}^ℓ , і дорівнює 0 – у протилежному випадку.

Використовуючи вираз (3.7) і перегрупування термів, знайдемо, що функція правдоподібності розкладається, згідно зі структурою ДЙМ, так само, як і у випадку статистичної ймовірнісної мережі [51]:

$$\begin{aligned} Pr(D|G, \theta_G) = \\ = \prod_i i \prod_{j_i'} j_i' \prod_{k_i'} k_i' (\theta_{i,j_i',k_i'}^{(0)})^{N_{i,j_i',k_i'}^{(0)}} \\ \times \prod_i i \prod_{j_i} j_i \prod_{k_i} k_i (\theta_{i,j_i,k_i}^{\rightarrow})^{N_{i,j_i,k_i}^{\rightarrow}} \end{aligned} \quad (3.14)$$

а, отже, логарифмічну функцію правдоподібності можна задати таким чином [10]:

$$L(B:D) = \sum_{i,j_i',k_i'}^{(0)} \sum_{i,j_i,k_i}^{\rightarrow} \sum_{j_i} \sum_{k_i} k_i N_{i,j_i,k_i}^{\rightarrow} \log \sum_i \sum_{j_i'} \sum_{k_i'} k_i' N_{i,j_i',k_i'}^{(0)} \log \quad (3.15)$$

Таке розкладання полегшує обчислення критерію БІК і оцінок БДе різними способами. По-перше, ймовірність виражається у вигляді суми термів, де кожен терм залежить тільки від умовної ймовірності змінної, враховуючи конкретні значення її батьківських змінних. Таким чином, якщо необхідно знайти для розподілу параметри максимальної правдоподібності, то їх можна максимізувати в межах кожної множини окремо. По-друге, розкладання означає, що можна навчити структуру B_0 незалежно від B_{\rightarrow} . Нарешті, можна навчити структуру B_{\rightarrow} , таким же чином, як і виконати навчання звичайної мережі на основі наявної множини вимірів.

Уточнимо ці аргументи. Використовуючи звичайну оцінку максимальної правдоподібності для поліноміальних розподілів, можна отримати такий вираз для кількості параметрів $\hat{\theta}_G$ [48]:

$$\hat{\theta}_{i,j_i',k_i'}^{(0)} = \frac{N_{i,j_i',k_i'}^{(0)}}{\sum k_i' N_{i,j_i',k_i'}^{(0)}}. \quad (3.16)$$

За аналогією можна записати вираз для випадку перехідної структури.

У випадку побудови таблиці умовних ймовірностей кількість параметрів мережі можна оцінити за виразом:

$$\#G = \#G_0 + \#G_{\rightarrow}, \quad (3.17)$$

де

$$\#G_0 = \sum i \sum p \in \mathbf{Pa}(X_i[0]) \|X_p\| \times (\|X_i[0]\| - 1). \quad (3.18)$$

За аналогією обчислюється кількість параметрів для перехідної структури.

Нарешті, підставляючи знайдені оцінки у вираз (3.9), знаходимо, що критерій БІК буде задаватися у даному випадку так [52]:

$$BIC(G; D) = BIC_0 + BIC_{\rightarrow} \quad (3.19)$$

де

$$BIC_0 = \sum_i \sum_{j'_i} \sum_{k'_i} k'_i N_{i,j'_i,k'_i}^{(0)} \widehat{\log \theta_{i,j'_i,k'_i}^{(0)}} - \frac{\log N_{seq}}{2} G_0 \quad (3.20)$$

$$BIC_{\rightarrow} = \sum_i \sum_{j_i} \sum_{k_i} k_i N_{i,j_i,k_i}^{\rightarrow} \widehat{\log \theta_{i,j_i,k_i}^{\rightarrow}} - \frac{\log N}{2} G_{\rightarrow} \quad (3.21)$$

Звернемо увагу, що погіршення значення критерію для множин у вихідний часовий інтервал є функцією від кількості послідовностей N_{seq} , оскільки погіршення оцінки для перехідної моделі є функцією загальної кількості вимірів N .

3.3.3 Оцінки Байєса-Діріхле (БДе) для динамічної ймовірнісної мережі

Нагадаємо, що для обчислення оцінки БДе, необхідно оцінити такий інтеграл:

$$Pr(D|G) = \int Pr(D|G, \theta_G) Pr(\theta_G|G) d\theta_G. \quad (3.22)$$

Перший член всередині інтеграла розкладається, як в рівнянні (3.14). Якщо припустити, що апостеріорна ймовірність по кожній умовній ймовірності не залежить від інших, то такий терм також має розклад у вигляді [52]:

$$Pr(\theta_G|G) = \prod_i i \prod_{j'_i} j'_i Pr(\theta_{i,j'_i,k'_i}^{(0)}) \times \prod_i i \prod_{j_i} j_i Pr(\theta_{i,j_i,k_i}^{\rightarrow}). \quad (3.23)$$

Підставляючи цей вираз у попереднє рівняння, бачимо, що цілісний вираз складається з інтеграла по добутку незалежних термів. Тому вираз $Pr(D|G)$ можна записати у вигляді добутку двох інтегралів таким чином [52]:

$$\prod_i i \prod_{j'_i} j'_i \int \prod_{k'_i} k'_i \theta_{i,j'_i,k'_i}^{(0) N_{i,j'_i,k'_i}^{(0)}} \times Pr(\theta_{i,j'_i,k'_i}^{(0)}) \times d\theta_{i,j'_i,k'_i}^{(0)}, \quad (3.24)$$

і аналогічно можна поступити для випадку перехідної структури.

Для того щоб отримати рішення в замкнутій формі, приймаємо апріорні ймовірності по Діріхле. Для поліноміального розподілу Діріхле змінна X задається множиною гіперпараметрів $N'_x: x \in Val(X)$ таким чином:

$$Pr(\theta_X) = \text{Dirichlet}(\{N'_x: x \in Val(X)\}) \propto \prod_x \theta_x^{N'_x-1}. \quad (3.25)$$

Інтуїтивно, гіперпараметри можна розглядати як “псевдо-відліки” (значення), оскільки вони відіграють аналогічну роль з фактичними відліками і впливають з наявних даних. При використанні апріорних ймовірностей по Діріхле, ймовірність спостереження послідовності значень X з кількістю $N(x)$ є такою [53]:

$$\begin{aligned} \int \prod_x \theta_x^{N(x)} Pr(\theta_X|G) d\theta_X &= \\ &= \frac{\Gamma(\sum_x N'_x)}{\Gamma(\sum_x (N'_x + N(x)))} \times \prod_x \frac{\Gamma(N'_x + N(x))}{\Gamma(N'_x)}, \end{aligned} \quad (3.26)$$

де $\Gamma(x) \int_0^\infty t^{x-1} e^{-t} dt$ – функція Гамма, що задовольняє властивостям $\Gamma(1) = 1$ і $\Gamma(x+1) = x\Gamma(x)$.

Повертаючись до оцінювання структури за критерієм БДе, припустимо, що для кожної структури G , вибрано гіперпараметри $N'_{i,j_i,k_i} \rightarrow$. Тепер можна переписати $Pr(D|G)$ у вигляді добутку двох термів таким чином [54]:

$$\prod_i \prod_{j'_i} \frac{\Gamma(\sum_{k'_i} N'_{i,j'_i,k'_i}(0))}{\Gamma(\sum_{k'_i} N'_{i,j'_i,k'_i}(0) + N'_{i,j_i,k_i} \rightarrow)} \times \prod_{k'_i} \frac{\Gamma(N'_{i,j'_i,k'_i}(0) + N_{i,j'_i,k'_i}(0))}{\Gamma(N'_{i,j'_i,k'_i}(0))} \quad (3.27)$$

Можна поступити аналогічно для випадку перехідної структури.

Цей підхід, як і раніше, вимагає підстановки гіперпараметрів Діріхле для кожного кандидата у модифіковану структуру ДЙМ. Оскільки фактична кількість можливих структур ДЙМ є досить великою, то практично отримати оцінки апіорних ймовірностей може виявитись важко визначити. Тому можна призначити всі апіорні ймовірності для ДЙМ $B' = (B'_{(0)}, B'_{\rightarrow})$ у вигляді двох еквівалентних вибірок $\hat{N}^{(0)}$ і \hat{N}^{\rightarrow} . Враховуючи ці компоненти, можна призначити вагові коефіцієнти для розподілу Діріхле, $G = (G_{(0)}, G_{\rightarrow})$, таким чином:

$$N'_{i,j'_i,k'_i}(0) = \hat{N}^{(0)} \times P_{B'_0}(X_i[0]) = k'_i \mathbf{Pa}(X_i[0]) = j'_i \quad (3.28)$$

$$N'_{i,j_i,k_i} \rightarrow = \hat{N}^{\rightarrow} \times P_{B'_{\rightarrow}}(X_i[1]) = k_i \mathbf{Pa}(X_i[1]) = j_i. \quad (3.29)$$

Вибір батьківських вузлів тут ґрунтується на структурі G , і може відрізнятися від структури B' . Інтуїтивно, можна розглянути міру довіри в B' як еквівалентну тому, що раніше використовувалась в послідовності $\hat{N}^{(0)}$ з \hat{N}^{\rightarrow} переходами.

Досить просто можна перевірити, що при виборі апіорних ймовірностей вказаним чином зберігається основна властивість моделі, яка вимагається у даному випадку: дві структури ДЙМ, які мають ідентичні припущення стосовно незалежності змінних, отримують однакові кількісні оцінки адекватності. Таким чином, можна стверджувати, що розглянуте визначення критерію є природним продовженням процесу оцінювання уточненої структури за допомогою критерію БДе у випадку аналізу статичних і динамічних ймовірнісних мереж.

3.4 Навчання на практиці

В обох видах оцінок структури мережі, які розглянуто вище, є дві важливі властивості. По-перше, оцінку ДЙМ можна записати у вигляді суми термів, (у випадку аналізу БДе розглядаємо логарифм $Pr(D|G)$), де кожен член визначає вклад певного вибору (множини) батьків для конкретної змінної. Таким чином, локальна зміна однієї множини (наприклад, додавання або видалення дуги) впливає тільки на один з цих термів. По-друге, член, який оцінює $X_i[t]$ із врахуванням його батьків, є функцією відповідних відгуків (або $N^{(0)}(\cdot)$ або $N^{\rightarrow}(\cdot)$) для цієї множини. Таким чином, за рахунок кешування цих відгуків можна ефективно оцінювати вплив багатьох множин даних.

Ці дві властивості можуть бути використані у процедурах пошуку екстремуму використаного функціоналу, що здатні поступово покращити потенційну структуру, застосовуючи інформативно найкращі дуги додавання, видалення або повернення видалених елементів. У випадку аналізу динамічних ймовірнісних мереж, на відміну від статичних ймовірнісних мереж, потрібно враховувати додаткове обмеження, яке структура мережі повинна повторити протягом відносно тривалого часу

(можливе часткове або повне повторення структури). Це дає можливість зменшити кількість варіантів структур у кожній точці простору структур конкретного типу. Крім того, можна знайти кращі інформативно початкові структури B_0 незалежно від пошуку можливого кращого варіанту для модифікованої структури B_{\rightarrow} .

3.4.1 Навчання для випадку наявності неповних вибірок даних

Розглянемо навчання динамічної ймовірнісної мережі у випадку наявності неповних вибірок даних. Неповні вибірки даних мають вирішальне значення, оскільки в більшості практичних задач немає повних спостережень процесу, що моделюється. Це означає, що якщо досліджуваний процес – це стаціонарний марківський процес, то часткові спостереження, які є у наявності, можуть не бути марківськими. Припустимо, що відстежується рух автомобіля по шосе, і що спостерігається швидкість автомобіля, відносна відстань до інших автомобілів та інші параметри. Враховуючи вже наявні спостереження, можна стверджувати, що цей приклад не марківський. З іншого боку, марківська модель може бути коректна за умови, якщо включити інформацію про відомі цілі водія, наприклад, їхати по лівій смузі, виїхати на сусідню автомагістраль і т. ін. Вивчаючи приховані змінні, можна перехопити необхідні відомості про стан процесу. Це дозволяє моделі «пам'ятати» додаткову інформацію про минулі стани і краще передбачати майбутні стани досліджуваної системи.

Основними труднощами з навчанням від часткових спостережень є те, що більше немає властивостей розкладу оцінки (3.14). Це означає, що оптимальний вибір параметрів в одній частині мережі залежить від вибору параметрів в іншій частині мережі. Ця проблема стає зрозумілою краще, якщо відомо, що, як тільки у дослідженні є випадок часткових спостережень, то більше не можна говорити про оптимальне оцінювання

на навчальних даних. Для більшості подій, що представляють інтерес для навчання моделі, оцінки є достатньо невизначеними, оскільки невідомі точні значення змінних у наявних вибірках.

Для розв'язання цієї задачі досить часто використовується *ЕМ*-алгоритм [50, 52]. На *Е*-кроці *ЕМ*-алгоритму оцінюються і використовуються наближені параметри для повних вибірок даних, які отримують шляхом обчислення їх очікуваної кількості. На *М*-кроці знову оцінюються значення параметрів максимальної правдоподібності за припущення, що очікувана кількість була до істинних оцінок спостережень. Основне припущення, що лежить в основі поведінки *ЕМ*-алгоритму, є таким, що кожен цикл *ЕМ*-алгоритму гарантовано підвищить ймовірність появи даних, поданих (використаних) у моделі, поки не буде досягнє локального максимуму.

ЕМ-алгоритм традиційно розглядається як метод налаштування параметрів фіксованої структури моделі. Однак, основне припущення можна узагальнити для застосування до довільних структурних, а також параметричних модифікацій моделей (тобто до їх структурної і параметричної адаптації). Так, структурний *ЕМ*-алгоритм (СЕМ) Фрідмана [44] містить той же *Е*-крок, що і класичний *ЕМ*-алгоритм, заповнює дані шляхом обчислення очікуваної кількості вимірів на основі поточної структури і параметрів. Крім того, для того щоб повторно оцінити параметри, на *М*-кроці структурний *ЕМ*-алгоритм може використовувати очікувану кількість параметрів відповідно до наявної поточної структури для того щоб оцінити будь-яку іншу структуру-кандидата, при цьому, по суті, виконуючи структурний пошук повних вибірок даних у внутрішньому циклі виконання обчислень. Фрідман показує, що в результаті модифікована мережа повинна мати вищу оцінку, ніж в оригіналі. Це вірно, навіть у тому випадку, коли очікувана кількість параметрів, що використовується для оцінювання нової структури, обчислюється з використанням інформації про попередню структуру.

Результати Фрідмана можна розширити на випадок ДЙМ таким чином. Нехай $E[BIC((B'_0, B'_\rightarrow): D^+): D, (B_0, B_\rightarrow)]$ є очікуваною оцінкою БІК для ДЙМ (B'_0, B'_\rightarrow) на основі всіх можливих доповнень D^+ у вигляді нових даних, тобто можна використати також значення можливих прихованих вимірів. Очікування береться по відношенню до $Pr(D^+|D, (B_0, B_\rightarrow))$, тобто ймовірність відноситься до цього поповнення на основі попередньої структури ДЙМ. Використовуючи теорему [44], можна довести, що (теорема 1):

$$\begin{aligned} & BIC((B'_0, B'_\rightarrow): D) - BIC((B_0, B_\rightarrow): D) > \\ & > E[BIC((B'_0, B'_\rightarrow): D^+): D, (B_0, B_\rightarrow)] - \\ & - E[BIC((B_0, B_\rightarrow): D^+): D, (B_0, B_\rightarrow)]. \end{aligned} \quad (3.30)$$

Тобто, якщо вибирається модифікована (нова) структура ДЙМ, яка має більш високі очікувані оцінки, ніж очікувані оцінки старої ДЙМ, то істинна оцінка нової (модифікованої) ДЙМ також буде вищою, ніж істинна оцінка старої ДЙМ. Крім того, різниця в очікуваних оцінках є тільки нижньою межею поліпшення в термінах фактичних оцінок, які оптимізуються; тобто залишається простір для подальшого покращення оцінок нової структури мережі завдяки наявним новим та прихованим вимірам.

Вирішальною властивістю очікуваних оцінок, отриманих за допомогою БІК є те, що вони теж розкладаються в суму місцевих термів таким чином, що в силу лінійності оператора очікування, можемо "спрямувати" оператор E через знаки обчислення суми неявних членів таким чином, щоб отримати рівняння, яке включає в себе терми у такому вигляді:

$$E[N_{i,j_i,k_i}^\rightarrow] = \sum \ell \sum t E[I(X_i[t] = k_i, \mathbf{Pa}(X_i[t]) = j_i; \mathbf{x}^\ell)] =$$

$$= \sum \ell \sum Pr(X_i[t] = k_i, \mathbf{Pa}(X_i[t]) = j_i | \mathbf{x}^\ell); \quad (3.31)$$

і можна за аналогією зробити це для членів $E[N_{i,j_i',k_i'}^{(0)}]$. Вони називаються очікуваною достатньою статистикою для модифікації структури моделі. Тому ключовою вимогою для застосування структурного ЕМ-алгоритму (СЕМ), у випадку існування неповних даних, є можливість обчислити ймовірність всіх допустимих сімейств ймовірностей, $Pr(X_i[t] = k_i, \mathbf{Pa}(X_i[t]) = j_i | \mathbf{x}^\ell)$, для всіх мереж, які можна оцінити, тобто саму поточну мережну структуру і «сусідні» мережі у допустимому полі пошуку структур ДЙМ.

Для ефективного обчислення загальної ймовірності для конкретних сімейств ймовірностей (розподілів), можна перетворити ДЙМ в об'єднані дерева і використати динамічний двопрхідний алгоритм [55] аналізу структур, схожий на алгоритм прямого-зворотнього ходу, який використовується при оцінюванні прихованих марківських моделей (ПММ) [54]. Для того щоб ефективно обчислити повну ймовірність множини вузлів, які ще не містяться у жодному вузлі об'єднаних дерев, необхідно використовувати складніші методи [55]. Простіший підхід, який використовується сьогодні, повинен об'єднати разом всі змінні, що містяться в одному інтервалі, в одному вузлі (тобто для того щоб перетворити ДЙМ у ланцюг Маркова), обчислити очікувану кількість переходів між кожною парою послідовних станів в цьому ланцюгу, а потім ізолювати ці кількості станів для того щоб отримати кількості переходів для кожного сімейства змінних.

Також можна виконувати кешування (збереження) очікуваних кількостей структур, обчислених на основі поточної моделі поповнення, і використати надалі цей кеш для того щоб уникнути повторного обчислення очікуваних кількостей. Методика, яка використовується сьогодні, для обчислення кількості і значень очікуваних достатніх

статистик, залишається чисельно складною. Таким чином, процедура застосування алгоритму СЕМ з використанням згенерованих або фактичних даних полягає у вигляді, поданому нижче.

Вибрати початкову структури $(B_0^0, B_{\rightarrow}^0)$ (можливо випадково).

Проаналізувати значення вимірів для $n = 0, 1, \dots$ на збіжність.

Відслідкувати поліпшення параметрів $(B_0^n, B_{\rightarrow}^n)$ за допомогою ЕМ.

Реалізувати пошук по можливих структурах ДЙМ (з використанням очікуваних оцінок, які обчислені за допомогою структур $(B_0^n, B_{\rightarrow}^n)$). Нехай кращою оцінкою для ДЙМ, при цьому, буде варіант $(B_0^{n+1}, B_{\rightarrow}^{n+1})$: якщо якість

$$(B_0^{n+1}, B_{\rightarrow}^{n+1}) = (B_0^n, B_{\rightarrow}^n), \quad (3.32)$$

то необхідно повернутись до структури $(B_0^{n+1}, B_{\rightarrow}^{n+1})$.

Таблиця 3.1. Значення вибраних критеріїв для одного набору даних

Розмір даних		Критерій БІК			Критерій BDe		
N _{seq}	N	0	1	2	0	1	2
300	11387	-1.4527	-1.5198	-1.9842	-1.5588	-1.5138	-1.359
600	28043	-1.1638	-1.0443	-1.0369	-1.2762	-1.0235	-1.0458
1200	61864	-0.9883	-0.9977	-0.9885	-0.9672	-0.9587	-0.9869
1800	102906	-0.9623	-0.9746	-0.9889	-0.9637	-0.9618	-0.9913

Поданий вище результат отримано для випадку застосування алгоритму СЕМ з метою отримання оцінок за байєсівським інформаційним критерієм (БІК). Процедура навчання на основі алгоритму СЕМ також розширена на використання оцінок БДе, але у цьому випадку потрібно пам'ятати про нелінійність цих оцінок. Однак, це не відіграє критичної ролі при їх практичному використанні, тобто їх можна використовувати,

пам'ятаючи, що виконується дослідження нелінійних нестационарних процесів, які мають свої особливості моделювання.

3.5 Байєсівська фільтрація нелінійних даних

3.5.1 Модель системи даних

Ще однією прикладною задачею, яка відноситься до підсистеми адаптації моделі, є фільтрація даних, тобто покращення їх якості перед моделюванням, прогнозуванням та оцінюванням ризику. Існують різні підходи та алгоритми фільтрації статистичних/експериментальних даних і серед них байєсівські методи, які широко використовуються на практиці.

Прикладна нелінійна (зокрема гранулярна) байєсівська фільтрація базується на нелінійній моделі в просторі станів у дискретному часі, яка пов'язує прихований стан x_k зі спостереженням z_k [57]:

$$x_k = f(x_{k-1}, v_{k-1}) \quad v_{k-1} \sim p_{v_{k-1}}, x_1 \sim p_{x_1} \quad (3.33)$$

$$z_k = h(x_k) + n_k, \quad n_k \sim p_{n_k} \quad (3.34)$$

Тут k означає номер дискретного моменту часу; v_k – процес випадкового шуму з відомою щільністю розподілу p_{v_k} ; n_k – адитивний шум вимірів з відомою щільністю p_{n_k} . Перший вимір позначається як z_1 , тому перший невідомий стан це x_1 , він має щільність розподілу p_{x_1} . Модель може залежати також від відомого керуючого впливу u_k , що в даному випадку опущено для спрощення викладок. Запис $s_{1:k}$ означає послідовність s_1, s_2, \dots, s_k (тут s є одним із сигналів x, v, z, n).

У статистичній літературі модель часто записується в термінах умовної щільності розподілу

$$x_k \sim p(x_k | x_{k-1}) \quad (3.35)$$

$$z_k \sim p(z_k | z_{k-1}) \quad (3.36)$$

У деякому сенсі така модель є більш загальною стосовно досліджуваного процесу. Завданням методів гранулярної байєсівської фільтрації є апроксимація апостеріорної щільності розподілу невідомого стану системи (процесу) за умови наявних вимірів, яка позначається так: $p(x_k | z_{1:k})$.

3.5.2 Методи гранулярної (particle) фільтрації

Алгоритм генерування послідовної вибірки за значимістю

Алгоритм послідовної вибірки за значимістю (Sequential Importance Sampling (SIS) Algorithm) є методом реалізації рекурсивного байєсівського фільтра з використанням симуляцій Монте-Карло. Це базовий алгоритм гранулярної фільтрації. Основною ідеєю є представлення необхідної апостеріорної щільності розподілу як множини випадкових елементів з асоційованими вагами та обчислення оцінок на основі цих елементів та вагів. При збільшенні кількості елементів вибірки така Монте-Карло характеристика стає еквівалентною до звичайного функціонального опису апостеріорної щільності та SIS фільтр наближається до оптимальної байєсівської оцінки.

Нехай $\{x_{1:k}^i, w_k^i\}_{i=1}^{N_s}$ позначає випадкову міру, яка характеризує апостеріорну щільність $p(x_{1:k} | z_{1:k})$, де $\{x_{1:k}^i, i = 0, \dots, N_s\}$ – множина k -крокових траєкторій з опорних точок (гранул, частинок) з асоційованими нормованими вагами $\{w_k^i, i = 0, \dots, N_s\}$, $\sum_{i=1}^{N_s} w_k^i = 1$; N_s – задана кількість гранул (particles). Тоді апостеріорна щільність в момент часу k може бути апроксимована так:

$$p(x_{1:k}|z_{1:k}) \approx \sum_{i=1}^{N_s} w_k^i \delta(x_{1:k} - x_{1:k}^i), \quad (3.37)$$

де $\delta(x)$ – δ -функція Дірака. Тобто $p(x_{1:k}|z_{1:k}) \approx w_k^i$

Вагові коефіцієнти обираються за принципом *вибірки за значимістю*. Він полягає в наступному. Припустимо, що $p(x) \propto \pi(x)$ (тут \propto означає пропорційність) – це щільність розподілу випадкової величини, реалізацію якої з її справжнього розподілу складно згенерувати, але для якої $\pi(x)$ можна обчислити. Нехай також $x^i \sim q(x), i = 0, \dots, N_s$ – реалізації випадкової величини, які легко генеруються з розподілу зі щільністю $q(\cdot)$, яка називається *пропонуною щільністю* або *значимою щільністю* (*proposal density, importance density*). Тоді зважена апроксимація щільності $p(\cdot)$ виглядатиме таким чином:

$$p(x) \propto \sum_{i=1}^{N_s} w^i \delta(x - x^i), \quad (3.38)$$

де $w^i \propto \frac{\pi(x^i)}{q(x^i)}$ є нормованою вагою i -ої гранули (після обчислення відношення проводиться нормування вагів: $\sum_{i=1}^{N_s} w_k^i = 1$).

Тоді, якщо реалізації $x_{1:k}^i$ були згенеровані з розподілу з пропонуною щільністю $q(x_{1:k}|z_{1:k})$, то ваги в (3.37), відповідно до (3.38), будуть такими:

$$w^i \propto \frac{p(x_{1:k}|z_{1:k})}{q(x_{1:k}|z_{1:k})}. \quad (3.39)$$

У випадку послідовних обчислень, на кожній ітерації маючи зважену вибірку, яка апроксимує $p(x_{1:k-1}|z_{1:k-1})$, можна було б апроксимувати $p(x_{1:k}|z_{1:k})$ новою вибіркою. Але якщо пропонує щільність обрана так, що вона розкладається на множники:

$$q(x_{1:k}|z_{1:k}) = q(x_k|x_{1:k-1}, z_{1:k})q(x_{1:k-1}|z_{1:k-1}), \quad (3.40)$$

то можна отримати елементи $x_{1:k}^i \sim q(x_{1:k}|z_{1:k})$ доповнюючи кожен з існуючих елементів $x_{1:k-1}^i \sim q(x_{1:k-1}|z_{1:k-1})$ новим станом $x_k^i \sim q(x_k|x_{1:k-1}, z_{1:k})$.

У практичних застосуваннях найчастіше вимагається лише фільтрована оцінка розподілу $p(x_k|z_{1:k})$ на кожному кроці, а не умовний розподіл одразу всієї траєкторії $p(x_{1:k}|z_{1:k})$. Тому надалі розглядати саме цей випадок.

Використовуючи прямий наслідок з теореми Байєса, отримаємо:

$$p(x_k|z_{1:k}) = \frac{p(z_k|x_k)p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})} \quad (3.41)$$

за виконання певних умов, зокрема, якщо $q(x_k|x_{1:k-1}, z_{1:k}) = q(x_k|x_{k-1}, z_k)$, тобто пропонує щільність залежить лише від x_{k-1} та z_k , можна показати [41], що для оновлення вагових коефіцієнтів використовується таке співвідношення:

$$w_k^i \propto w_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{k-1}^i, z_k)}, \quad (3.42)$$

і фільтрований апостеріорний розподіл може бути апроксимований так:

$$p(x_k|z_{1:k}) \approx \sum_{i=1}^{N_s} w_k^i \delta(x_k - x_k^i). \quad (3.43)$$

Ще раз необхідно наголосити, що вагові коефіцієнти w_k^i мають бути нормовані щоб забезпечити виконання умови: $\sum_{i=1}^{N_s} w_k^i = 1$. Це гарантує коректність виконання обчислень. Вибір пропонуемого розподілу є одним із найбільш важливих етапів проектування гранулярного фільтра. Можливі варіанти вибору, їх переваги та недоліки розглянуті у [58]. Зокрема важливо, щоб дисперсія вагових коефіцієнтів w_k^i була невеликою.

Найчастіше в якості пропонуемого розподілу використовують апіорний з (3.35):

$$q(x_k|x_{k-1}^i, z_k) = p(x_k|x_{k-1}^i), \quad (3.44)$$

що є зручним вибором. В такому разі (3.42) спрощується до вигляду:

$$w_k^i \propto w_{k-1}^i p(z_k|x_k^i). \quad (3.45)$$

Необхідно зазначити, що *не в усіх задачах* такий вибір пропонуемого розподілу є однозначно оптимальним.

3.5.3 Основний алгоритм фільтрації

Підсумуємо алгоритм послідовної вибірки за значимістю. Елементи $x_{1:1}^i$ у зваженій вибірці на 1-му кроці $\left\{x_{1:1}^i, \frac{1}{N_s}\right\}_{i=1}^{N_s}$ генеруються з початкового розподілу p_{x_1} (3.33). Оскільки цей розподіл істинний, то вводити корекції не потрібно і всі ваги рівні $w_1^i = \frac{1}{N_s}$.

Процедуру генерування зваженої вибірки на k -му кроці, маючи зважену вибірку на $(k - 1)$ -у кроці, можна представити наступним псевдокодом:

Алгоритм 1: SIS Particle Filter

$$\left[\{x_k^i, w_k^i\}_{i=1}^{N_S} \right] = \text{SIS} \left[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^{N_S}, z_k \right] \quad (3.46)$$

FOR

— Згенерувати $x_k^i \sim q(x_k | x_{k-1}^i, z_k)$

— Присвоїти гранулі x_k^i вагу w_k^i відповідно до (3.42)

END FOR

Для цього та всіх подальших алгоритмів фільтрації на кожному кроці апостеріорний розподіл можна апроксимувати за формулою (3.43). Оцінкою умовного математичного сподівання стану x_k буде:

$$\widehat{x_k} = \sum_{i=1}^{N_S} w_k^i x_k^i \quad (3.47)$$

3.5.3.1 Відсів або виключення гранул

Частою проблемою SIS гранулярного фільтра є виродження вагових коефіцієнтів, коли після певної кількості ітерацій усі гранули, крім однієї, мають незначну вагу. Експериментально встановлено, що дисперсія вагових коефіцієнтів може лише збільшуватись з часом, а тому феномену виродження уникнути неможливо. Виродження спричиняє те, що багато обчислень витрачається на оновлення гранул, внесок яких у апроксимацію $p(x_k | z_{1:k})$ майже нульовий.

Одним із способів зменшення ефекту виродження є проведення *відсіву* (resampling). Основною ідеєю відсіву є виключення гранул, що мають малу вагу та концентрація на гранулах з великою вагою. На кроці відсіву

відбувається генерація нової множини $\{x_k^{i*}\}_{i=1}^{N_S}$ вибором (з повтореннями) N_S реалізацій випадкової величини x_k з наближеного дискретного представлення розподілу $p(x_k|z_{1:k})$, заданого формулою (3.43), тобто таким чином, що $P\{x_k^{i*} = x_k^j\} = w_k^j$. Отримана вибірка є насправді вибіркою незалежних однаково розподілених величин з дискретного розподілу (3.43), тому ваги тепер встановлюються рівними, $w_k^i = \frac{1}{N_S}$.

Наведемо псевдокод процедури відсіву. Такий варіант відсіву, *систематичний відсів* (systematic resampling), обрано через простоту реалізації та оцінку складності алгоритму $O(N_S)$. Для кожної частинки нової вибірки зберігається також її індекс у попередній вибірці, i^j , що знадобиться в наступних алгоритмах.

Алгоритм 2: Resampling Algorithm

$$\left[\{x_k^{j*}, w_k^j, i^j\}_{i=1}^{N_S}\right] = \text{RESAMPLE} \left[\{x_k^i, w_k^i\}_{i=1}^{N_S}\right] \quad (3.48)$$

Ініціалізувати функцію розподілу (ФР): $c_1 = 0$

FOR $i = \overline{2, N_S}$

— Побудувати ФР: $c_i = c_{i-1} + w_k^i$

END FOR

Почати з початку ФР: $i = 1$

Згенерувати початкову точку: $u_1 \sim U[0, N_S^{-1}]$

FOR $j = \overline{1, N_S}$

— Рухатись уздовж ФР: $u_j = u_1 + N_S^{-1}(j - 1)$

— WHILE $u_j > c_i$

— — $i = i + 1$

— END WHILE

— Присвоїти елемент: $x_k^{j*} = x_k^i$

— Присвоїти вагу: $w_k^j = N_S^{-1}$

— Присвоїти батька: $i^j = i$

END FOR

Процедура відсіву має деякі недоліки. По-перше, вона звужує можливості паралелізму обчислень, оскільки на кроці відсіву всі гранули мають комбінуватися. По-друге, гранули з великою вагою будуть обиратися багаторазово, що може призвести до втрати різноманітності вибірки, результуюча вибірка міститиме багато повторів. Ця проблема, відома як *збідніння вибірки* (sample impoverishment) і є особливо гострою у випадку малої випадкової складової процесу. Тоді всі гранули можуть зліпитися в одну через кілька ітерацій.

3.5.4 Фільтр вибірки за значимістю з відсівом

Фільтр на основі вибірки за значимістю з відсівом (Sampling Importance Resampling Filter (SIR)) – це метод Монте-Карло, який може бути застосованим до задач рекурсивної байєсівської фільтрації. Обмеження, накладені на його використання, є дуже слабкими. Функції $f(\cdot, \cdot)$ та $h(\cdot, \cdot)$ в (3.33) і (3.34) мають бути відомими, необхідно мати змогу генерувати реалізації з розподілу шуму $p_{v_{k-1}}$ та апіорного розподілу $p(x_k | x_{k-1})$ і визначати значення щільності розподілу $p(z_k | x_k)$ в потрібних точках, принаймні з точністю до спільної константи.

SIR алгоритм може бути легко виведений з алгоритму SIS за рахунок відповідного вибору таких елементів:

- 1) пропонує розподіл $q(x_k | x_{k-1}^i, z_k)$ – в якості цього розподілу обирається апіорна щільність $p(x_k | x_{k-1}^i)$;
- 2) кроку відсіву, який здійснюється у кожний момент часу.

Такий вибір пропонує розподіл обумовлює необхідність генерування реалізацій з $p(x_k | x_{k-1}^i)$. Реалізацію $x_k^i \sim p(x_k | x_{k-1}^i)$ можна

отримати, якщо згенерувати реалізацію шуму $v_{k-1}^i \sim p_{v_{k-1}}$ та покласти $x_k^i = f(x_{k-1}^i, v_{k-1}^i)$.

Для такого спеціального вибору пропонуємої щільності формула оновлення вагів набуває вигляду (3.45). Але, взявши до уваги, що відсів відбувається в кожний момент часу (на кожній ітерації алгоритму), отримуємо, що $w_{k-1}^i = \frac{1}{N_S} \forall i$, а тому:

$$w_k^i \propto p(z_k | x_k^i). \quad (3.49)$$

Ваги задані в (3.49) нормуються перед фазою відсіву.

Подамо псевдокод ітерації цього алгоритму:

Алгоритм 3: SIR Particle Filter

$$\left[\{x_k^i, w_k^i\}_{i=1}^{N_S} \right] = \text{SIR} \left[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^{N_S}, z_k \right] \quad (3.50)$$

FOR $i = \overline{1, N_S}$

— Згенерувати $x_k^i \sim p(x_k | x_{k-1}^i)$

— Обчислити $w_k^i = p(z_k | x_k^i)$

END FOR

Обчислити загальну вагу: $t = \sum_{i=1}^{N_S} w_k^i$

FOR $i = \overline{1, N_S}$

— Нормувати i -ту вагу: $w_k^i = t^{-1} w_k^i$

END FOR

Провести відсів, використовуючи алгоритм 2 (Resampling Algorithm):

— $\left[\{x_k^{j*}, w_k^j, i^j\}_{i=1}^{N_S} \right] = \text{RESAMPLE} \left[\{x_k^i, w_k^i\}_{i=1}^{N_S} \right]$.

3.5.5 Допоміжний фільтр вибірки за значимістю з відсівом

Допоміжний фільтр вибірки за значимістю з відсівом (Auxiliary Sampling Importance Resampling Filter (ASIR)) є варіантом стандартного фільтру для вибірки за значимістю з відсівом. Він може бути отриманий з SIR фільтра введенням пропонууючої щільності $q(x_k, i|z_{1:k})$, з якої генеруються реалізації пар $\{x_k^j, i^j\}_{j=1}^{N_s}$, де i^j позначає номер гранули в $(k-1)$ -ий момент. i^j в даному фільтрі називають допоміжною змінною, оскільки її єдина мета допомогти в задачі симуляції. Звідси назва фільтру.

Використовуючи теорему Байєса, можна показати, що

$$p(x_k, i|z_{1:k}) \propto p(z_k|x_k)p(x_k|x_{k-1}^i)w_{k-1}^i. \quad (3.51)$$

ASIR фільтр генерує реалізацію зі спільного розподілу $p(x_k, i|z_{1:k})$, а потім опускає номер i в парі (x_k, i) , щоб отримати вибірку $\{x_k^j\}_{j=1}^{N_s}$ з маргінального розподілу $p(x_k|z_{1:k})$. Пропонууюча щільність для $\{x_k^j, i^j\}_{j=1}^{N_s}$ має задовольняти пропорційність:

$$q(x_k, i|z_{1:k}) \propto p(z_k|\mu_k^i)p(x_k|x_{k-1}^i)w_{k-1}^i, \quad (3.52)$$

де μ_k^i – певна характеристика x_k за умови x_{k-1}^i . Наприклад, це може бути умовне математичне сподівання $\mu_k^i = E[x_k|x_{k-1}^i]$, мода або реалізація $\mu_k^i \sim p(x_k|x_{k-1}^i)$.

Якщо записати:

$$q(x_k, i|z_{1:k}) = q(x_k|i, z_{1:k})q(i|z_{1:k}), \quad (3.53)$$

та покласти

$$q(x_k|i, z_{1:k}) = p(x_k|x_{k-1}^i) \quad (3.54)$$

то з (3.52) випливає, що:

$$q(i|z_{1:k}) \propto p(z_k|\mu_k^i)w_{k-1}^i \quad (3.55)$$

Парі $\{x_k^j, i^j\}_{j=1}^{N_s}$ присвоюється вага, пропорційна відношенню (3.51) і (3.52):

$$w_k^i \propto \frac{p(z_k|x_k^i)}{p(z_k|\mu_k^i)}. \quad (3.56)$$

Наведемо псевдокод ітерації цього алгоритму:

Алгоритм 4: Auxiliary Particle Filter

$$\left[\{x_k^i, w_k^i\}_{i=1}^{N_s} \right] = \text{APF} \left[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^{N_s}, z_k \right]$$

FOR $i = \overline{1, N_s}$

— Обчислити μ_k^i .

— Обчислити $w_k^i = q(i|z_{1:k}) \propto p(z_k|\mu_k^i)w_{k-1}^i$.

END FOR

Обчислити загальну вагу: $t = \sum_{i=1}^{N_s} w_k^i$

FOR $i = \overline{1, N_s}$

— Нормувати і-ту вагу: $w_k^i = t^{-1}w_k^i$

END FOR

Провести відсів, використовуючи алгоритм 2 (Resampling Algorithm):

$$\left[\{-, -, i^j\}_{j=1}^{N_s} \right] = \text{RESAMPLE} \left[\{x_{k-1}^i, w_k^i\}_{j=1}^{N_s} \right]$$


```

FOR  $j = \overline{1, N_S}$ 
— Згенерувати  $x_k^j \sim q(x_k | i^j, z_{1:k}) = p(x_k | x_{k-1}^{ij})$ , як в SIR фільтрі
— Присвоїти вагу  $w_k^j$ , використовуючи (3.56)
END FOR

Обчислити загальну вагу:  $t = \sum_{i=1}^{N_S} w_k^i$ 
FOR  $i = \overline{1, N_S}$ 
— Нормувати  $i$ -ту вагу:  $w_k^i = t^{-1} w_k^i$ 
END FOR

```

Таким чином, ідейно, щоб згенерувати реалізацію з розподілу $p(x_k, i | z_{1:k})$ спочатку генерується номер i з ймовірністю $w_k^i \propto q(i | z_{1:k})$ (ці ймовірності називаються ймовірностями першого етапу), а потім x_k з розподілу $p(x_k | x_{k-1}^i)$. Після цього індекс i відкидається, а x_k присвоюються ваги відповідно до (3.56) – ваги другого етапу. Є експериментальні свідчення, що отримані ваги другого рівня мають меншу дисперсію, ніж ваги в оригінальному SIR алгоритмі.

3.5.6 Регуляризований гранулярний фільтр

Метод відсіву введено для зменшення ефекту виродження, яким страждають гранулярні фільтри. Але, як було зазначено при його обговоренні, відсів призводить до інших проблем, зокрема проблеми втрати різноманітності гранул. Вона виникає через те, що на кроці відсіву реалізації генеруються з дискретного, а не неперервного розподілу. Якщо не приділяти уваги цій проблемі, вона може призвести до екстремального випадку «зліплення гранул», коли всі N_S гранул знаходяться в одному й тому самому місці простору станів і погано відображають апостеріорний розподіл.

Для боротьби з цією проблемою був запропонований регуляризований гранулярний фільтр (Regularized Particle Filter (RPF)) [59]. В ньому при проведенні процедури відсіву генеруються гранули з неперервної апроксимації $p(x_k|z_{1:k})$, а не з дискретної, як в SIR фільтрі. А саме апроксимації вигляду:

$$p(x_k|z_{1:k}) \approx \sum_{i=1}^{N_S} w_k^i K_h(x_k - x_k^i), \quad (3.57)$$

де

$$K_h(x) = \frac{1}{h^{n_x}} K\left(\frac{x}{h}\right) \quad (3.58)$$

– це масштабоване *статистичне ядро* (kernel density) $K(\cdot)$, $h > 0$ – це параметр згладжування (скалярний), n_x – розмірність вектора стану x , $w_k^i, i = 1, \dots, N_S$ – нормовані ваги. Статистичне ядро $K(\cdot)$ – це симетрична функція щільності розподілу така, що:

$$\int x K(x) dx = 0, \int \|x\|^2 K(x) dx < \infty. \quad (3.59)$$

Тобто, випадковий вектор з щільністю $K(\cdot)$ має нульове математичне сподівання та скінченні моменти другого порядку.

Статистичне ядро $K(\cdot)$ та параметр згладжування $h > 0$ обираються так, щоб мінімізувати середнє інтегроване квадратичне відхилення (mean integrated square error (MISE)) між справжньою апостеріорною щільністю та відповідним регуляризованим емпіричним представленням в (3.57), яке визначається як:

$$MASE(\hat{p}) = E \left[\int (\hat{p}(x_k|z_{1:k}) - p(x_k|z_{1:k}))^2 dx_k \right], \quad (3.60)$$

де $\hat{p}(x_k|z_{1:k})$ позначає апроксимацію, визначену в (3.57). У спеціальному випадку, коли всі елементи мають однакову вагу $\frac{1}{N_S}$, оптимальним вибором ядра є ядро Єпанечнікова [60, 61]:

$$K_e(x) = \begin{cases} \frac{n_x + 2}{2c_{n_x}} (1 - \|x\|^2), & \text{якщо } \|x\| < 1 \\ 0, & \text{інакше} \end{cases} \quad (3.61)$$

де c_{n_x} – об'єм одиничної гіперсфери в просторі R^{n_x} .

Часто зручно використовувати гаусівське ядро:

$$K_n(x) = (2\pi)^{-n_x/2} e^{-\frac{1}{2}\|x\|^2}. \quad (3.62)$$

Крім того, якщо щільність, що апроксимується, є гаусівською з одиничною коваріаційною матрицею, оптимальним вибором параметра згладжування є такий:

$$h_{opt} = A(K)N_S^{1/(n_x+4)}, \quad (3.63)$$

де $A(K_e) = [8c_{n_x}^{-1}(n_x + 4)(2\sqrt{\pi})^{n_x}]^{1/(n_x+4)}$ для ядра Єпанечнікова та $A(K_n) = [4/(n_x + 2)]^{1/(n_x+4)}$ для гаусівського ядра [37].

В алгоритмі регуляризованого гранулярного фільтра відсів проводиться не на кожному кроці, а тоді, коли значення міри виродження стає нижчим деякого заданого порогу N_T . Сама міра виродження обчислюється за формулою:

$$N_{eff} = \frac{1}{\sum_{i=1}^{N_S} (w_k^i)^2}. \quad (3.64)$$

Зазначимо, що якщо встановити порогове значення N_T рівним N_S , то відсів відбуватиметься на кожній ітерації.

Наведемо псевдокод ітерації алгоритму:

Алгоритм 5: Regularized Particle Filter

$$\left[\{x_k^{i*}, w_k^i\}_{i=1}^{N_S} \right] = \text{RPF} \left[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^{N_S}, z_k \right] \quad (3.65)$$

FOR $i = \overline{1, N_S}$

— Згенерувати $x_k^i \sim q(x_k | x_{k-1}^i, z_k)$

— Присвоїти гранулі x_k^i вагу w_k^i відповідно до (3.19)

END FOR

Обчислити загальну вагу: $t = \sum_{i=1}^{N_S} w_k^i$

FOR $i = \overline{1, N_S}$

— Нормувати i -ту вагу: $w_k^i = t^{-1} w_k^i$

END FOR

Обчислити N_{eff} , використовуючи (3.64)

IF $N_{eff} < N_T$

— Обчислити емпіричну коваріаційну матрицю S_k для $\{x_k^i, w_k^i\}_{i=1}^{N_S}$

— Обчислити матрицю Σ_k таку, що $\Sigma_k \Sigma_k^T = S_k$ (наприклад, $\Sigma_k = E_k D_k^{\frac{1}{2}}$, де $S_k = E_k D_k E_k^T$ – спектральний розклад матриці S_k)

— Провести відсів, використовуючи алгоритм 2 (Resampling Algorithm):

— $\left[\{x_k^i, w_k^i, -\}_{i=1}^{N_S} \right] = \text{RESAMPLE} \{x_k^i, w_k^i\}_{i=1}^{N_S}$

— FOR $i = \overline{1, N_S}$

— — Згенерувати $\varepsilon^i \sim K$ зі статистичного ядра (використати (3.61) або (3.62))

— — $x_k^{i*} = x_k^i + h_{opt} \sum_k \varepsilon^i$ (використати (3.63))

— END FOR

END IF

Хоча результати (3.61), (3.62), (3.63) є оптимальними лише в спеціальних випадках, вони можуть використовуватись в загальному випадку для отримання субоптимального фільтра.

Введення регуляризації покращує результат порівняно зі звичайним SIR фільтром у випадках, коли наявна значна втрата різноманітності вибірки гранул, наприклад, якщо шум у процесі є невеликим. Однією з проблем такого підходу є збільшення дисперсії апостеріорного розподілу.

3.6 Висновки до розділу

В розділі запропоновано вдосконалену системну багатокрокову методологію побудови моделей фінансових процесів і фінансових ризиків довільного походження. Наданий детальний огляд теоретичних основ і методів адаптації байєсівських мереж до даних в контексті аналізу ймовірнісних розподілів для дискретних випадкових величин і динамічних систем.

Дискретний підхід для моделювання ймовірнісних розподілів змінних спрощує представлення розподілів і зменшує обчислювальні витрати. Описана байєсівська мережа (БМ) як орієнтований ациклічний граф, що відображає спільний розподіл ймовірностей для змінних. Підкреслено, що умовна незалежність графічно відображається в БМ, і це є важливим аспектом для ефективного моделювання. Динамічна байєсівська мережа складається з апіорної структури мережі (моделі) і мережі переходів для моделювання процесів у часі. Вона описує розподіл для всіх

можливих траєкторій еволюції процесу і складається з апіорної структури і мережі переходів.

В розділі розглянуто питання навчання динамічних байєсівських мереж (ДБМ) при наявності повних наборів даних. Навчання ДБМ використовує методи, подібні на навчання статичних байєсівських мереж (БМ) за наявності повних наборів даних. Оцінка відповідності моделі даним визначається за допомогою функцій оцінки, таких як байєсівський інформаційний критерій (БІК) і Bayesian-Dirichlet (БДе) критерій. Ці критерії враховують імовірність даних і структуру мережі. Один з підходів до навчання ДБМ включає апіорні ймовірності параметрів мережі, які розкладаються на множники за допомогою незалежних апіорних ймовірностей для кожного умовного розподілу. Аналіз асимптотичної поведінки оцінки функції ймовірності отримання належної структури мережі показує, що деякі оцінки можуть бути нечутливими до вибору апіорних ймовірностей при великих кількостях даних. Оцінювання апіорних параметрів може бути обчислене в замкнутій формі за умови використання спряженого сімейства, наприклад, розподілу Діріхле. Навчання ДБМ дозволяє досліджувати багато можливих структур мереж, що дає можливість знайти найліпшу структуру для задачі моделювання. Розглянуті теоретичні види оцінювання структури мережі додатково підкріплюються практичним описом процедури навчання, на чому сконцентрована робота, що описана в цьому розділі.

РОЗДІЛ 4. Створення системи підтримки прийняття рішень для моделювання і використання обчислювальних експериментів

4.1 Архітектура і функції системи підтримки прийняття рішень

Для моделювання вибраних фінансових процесів та оцінювання супутніх ризиків створена система підтримки прийняття рішень (СППР) з використанням принципів системного аналізу. Коректно побудована система дає можливість підвищити адекватність математичних моделей, що будуються, якість оцінок прогнозів та оцінок ризиків можливих фінансових втрат [62, 63]. Система будується з використанням принципів системного аналізу і має ознаки інтелектуальності, які полягають у тому, що вона фактично є учасником процесу прийняття рішень, а саме: допомагає у виборі методів обробки даних, моделювання, прогнозування, оцінювання ризиків та аналізу якості проміжних і остаточних результатів обчислень [62]. Модульна структура (архітектура) системи забезпечить реалізацію можливості її прискореної модифікації: додавання нових функцій, критеріїв, моделей, методів оцінювання ризиків і т. ін.

При побудові системи використані такі принципи системного аналізу:

- ієрархічність архітектури (структури) системи;
- принцип адаптації моделей до нових даних та нових режимів функціонування досліджуваних процесів;
- оптимізація структури і параметрів моделей у процесі їх побудови та адаптації (тобто в процедури оцінювання вводяться оптимізаційні процедури, які здатні покращити результуючі оцінки структури і параметрів моделей);
- забезпечення альтернативності застосування методів оцінювання параметрів моделей (тобто оцінювання виконується кількома методами

- з метою досягнення високої якості оцінок, а саме: незміщеності, консистентності та ефективності);
- ідентифікація (виявлення) та врахування можливих невизначеностей даних у процесі виконання обчислень;
 - забезпечення обчислень належною критеріальною базою для контролю процесів обробки даних на кожному етапі виконання обчислень;
 - функціональна повнота системи, яка забезпечує її автономність і, відповідно, високу якість обчислювальних процесів, незалежність від інших учасників процесу аналізу даних;
 - створення зручного інтерфейсу, що відповідає вимогам користувачів різного рівня підготовки і людського фактора.

Моделі і методи оцінювання можливих фінансових втрат, які широко використовуються сьогодні у інформаційно-аналітичних системах відповідного призначення, не завжди відповідають вимогам конкретних користувачів, а тому задачі підвищення адекватності моделей та якості оцінювання ризиків можливих втрат є актуальними. Розв'язати ці задачі допомагають сучасні СППР, які будуються на принципах системного аналізу і забезпечують необхідне підвищення якості результатів. Частково можливості структурно-параметричної адаптації моделей і фільтрації даних розглядаються в роботах [64–67], однак представлені результати не ґрунтуються на використанні системного підходу до аналізу даних та експертних оцінок, недостатньо враховують невизначеності даних.

Значну роль у підвищенні якості проміжних та остаточних результатів обчислень у будь-якій системі відіграє застосування методів оптимальної (байєсівської, калманівської) фільтрації даних. Подібні задачі розглядаються, зокрема, у роботі [68]. Такі процедури можуть також бути адаптивними, оскільки разом з фільтрацією даних вони можуть використовуватись для оцінювання деяких параметрів математичних моделей. Однак, необхідно зазначити, що такі застосування є досить складними з обчислювальної точки зору, оскільки часто призводять до

розбіжності відповідних обчислювальних алгоритмів та зниження якості отримуваних оцінок. Це призводить до зниження адекватності математичних моделей, зменшення якості оцінок прогнозів та можливих ризиків втрат [69].

Очевидно, що існує необхідність побудови множини нових моделей, створення нових методів моделювання та оцінювання фінансових ризиків, які дадуть можливість отримати високі результати оцінювання можливих втрат в умовах наявності невизначеностей. Інформаційна СППР, побудована на принципах системного аналізу, дасть можливість розв'язати вказані задачі на високому рівні.

Концепція створення СППР для моделювання і оцінювання ризиків можливих втрат, що ґрунтується на системному підході, подана на Рис. 4.1. Запропонований підхід передбачає виконання поглибленого аналізу досліджуваного процесу, ідентифікацію та врахування можливих невизначеностей даних, структурно-параметричну адаптацію моделі вибраного типу ризику, прогнозування можливих втрат. Більш докладну функціональну схему створеної системи ілюструє Рис. 4.2.

Аналіз даних розпочинається з вибору фінансового процесу, аналізу його поточного стану, огляду існуючих моделей його формалізованого опису, які можуть використані у подальшому виконанні дослідження ризику можливих втрат. Огляд спеціальної літератури необхідно виконати для того щоб встановити факт існування моделей для опису ризику вибраного типу. Такі моделі можуть бути використані у подальшому виконанні дослідження з метою виконання порівняльного аналізу. Серед них можуть бути моделі різних структур – нечітка логіка, нейронні мережі, нейронечіткі моделі, а також моделі у просторі станів, нелінійна регресія, нелінійні моделі у вигляді ядер ймовірнісних розподілів та ін. При цьому необхідно коректно встановити інформативність даних.

Після підготовки даних необхідно виконати їх статистичний аналіз, який включає обчислення кореляційних функцій та взаємної інформації,

застосування статистичних тестів для встановлення типу процесу: аналіз наявності нестационарності (наявності тренду та/або гетероскедастичності), встановлення наявності нелінійності, яка може стосуватись змінних або параметрів. Як правило, для такого аналізу використовуються різні методи (і алгоритми) з метою підвищення достовірності результатів тестування даних.



Рисунок 4.1. Етапи реалізації системного підходу при моделюванні і прогнозуванні

Оцінюється, також, інформативність даних, наприклад, шляхом обчислення дисперсії, яка часто використовується в якості формальної міри інформативності.

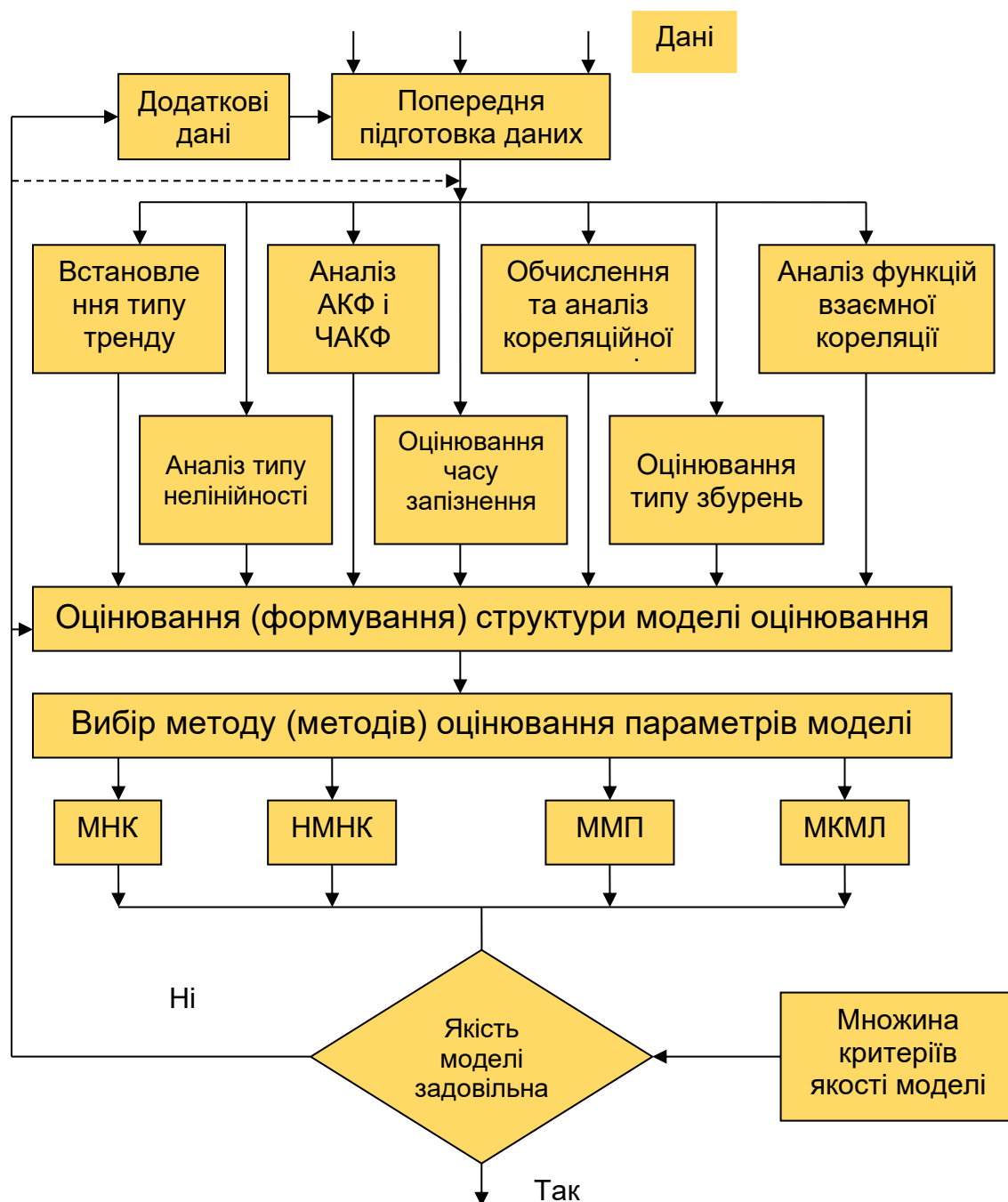


Рисунок 4.2. Схема адаптивного оцінювання моделі на основі статистичних даних

Така міра успішно використовується, зокрема, методом головних компонент, який використовується для аналізу мультиколінеарності векторів вимірів використаних змінних.

Оцінити інформативність даних, можна також шляхом оцінювання моделі наявних даних у вигляді полінома від часу:

$$y(k) = b_0 + b_1k + b_2k^2 + b_3k^3 + \dots + b_mk^m + \varepsilon(k), \quad (4.1)$$

де $y(k)$ — основна змінна; $\varepsilon(k)$ — випадковий процес, поява якого зумовлена наявністю випадкових зовнішніх збурень, похибками вимірів, неточністю структури і параметрів; $k = 0, 1, 2, \dots$ — дискретний час, який зв'язаний з неперервним часом, t , періодом дискретизації вимірів T_s : $t = kT_s$.

Така модель дає можливість оцінити інформативність даних шляхом оцінювання порядку полінома у правій частині моделі. Чим більше цей поліном містить статистично значущих оцінок коефіцієнтів $b_i, i = 0, 1, 2, \dots, m$, тим більше похідних можна обчислити за допомогою даних і тим вищою є їх інформативність. Очевидно, що якість зібраних статистичних даних є важливою передумовою для побудови їх адекватної математичної моделі [64, 66].

Підготовлені дані використовуються далі для оцінювання структури і параметрів моделей-кандидатів. Оцінювання структури моделі або її структурна ідентифікація дає можливість встановити з яких функціональних елементів (компонентів) буде складатись модель. Будемо формально описувати структуру моделі таким чином:

$$S = \{r, p, m, n, d, z, l\}, \quad (4.2)$$

де r — розмірність моделі, яка визначається кількістю рівнянь, що формують модель; p — порядок моделі, тобто максимальний порядок

рівнянь, які її утворюють; m – кількість незалежних регресорів (змінних); n – тип нелінійності, яка може бути (у нашому випадку) стосовно змінних та/або параметрів; d – лаг реакції на виході стосовно вхідного впливу; z – випадкове збурення і тип його розподілу; l – вектор можливих обмежень на параметри та/або змінні. Для одного процесу досліджують кілька можливих моделей-кандидатів, а потім вибирають кращу за множиною критеріїв адекватності. Такий підхід підвищує ймовірність побудови кращої моделі для подальшого практичного застосування з метою оцінювання можливих ризиків втрат.

Дані у формі часових рядів містять детерміновану і випадкову складові. Причина появи випадкової складової полягає у тому, що на досліджувані процеси завжди діють випадкові збурення, похибки вимірів, а також похибки обчислень, пов'язані з неточністю оцінювання структури моделі та її параметрів. Тут важливо правильно оцінити розподіл ймовірностей випадкових величин, які утворюють випадковий збурюючий процес. Експериментальні (статистичні) дані дають можливість коректно оцінити тип розподілу та його параметри, що, очевидно, сприяє підвищенню адекватності моделей процесу, який досліджується.

Після побудови моделі необхідно перевірити її на досягнення якості прогнозування, оскільки прогнозування – це задача, яка практично реалізується в СППР завжди з метою перевірки коректності функціонування побудованої моделі. Крім того, прогнозування дає можливість перевірити якість апроксимації моделлю статистичних характеристик досліджуваного процесу: дисперсії, коваріації і вибіркового середнього, яке апроксимує математичне сподівання. Для процесів з наявністю статистичних, параметричних і структурних невизначеностей доцільно застосовувати методи, які за своєю суттю близькі (або відносяться) до методів штучного інтелекту: байєсівські і нейронні мережі, метод групового урахування аргументів (МГУА), нечітка логіка, дерева рішень, комбіновані моделі, які забезпечують адекватний опис нелінійних

нестационарних процесів у фінансах і т. ін. Сьогодні все більшої популярності набирають методи ймовірнісно-статистичного аналізу даних, серед яких чільне місце займають байєсівські методи: ймовірнісні статичні і динамічні мережі, ймовірнісна фільтрація (наприклад, гранулярні фільтри), автоматичне (або автоматизоване) керування, автоматизована діагностика та підтримка прийняття рішень в бізнесі, техніці, економіці [62, 66, 70–75].

4.2 Побудова функцій прогнозування в СППР

Функції прогнозування в межах СППР будуються на основі різницевих рівнянь, які мають зручну структуру для утворення рекурсивних у часі функцій і використані для оцінювання коротко- і середньострокових прогнозів. Така прогнозуюча модель може бути подана у вигляді:

$$y(k) = f[y(k-1), \dots, y(k-p), u(k-1), \dots, u(k-q), \theta] + \varepsilon(k), \quad (4.3)$$

де $y(k)$ – залежна (основна) змінна, що прогнозується; $u(k)$ – вхідна змінна (керування); p, q – порядок авторегресії (АР) та ковзного середнього (КС); θ – вектор параметрів моделі; $\varepsilon(k)$ – випадковий процес збурення; $k = 0, 1, 2, \dots$ – дискретний час, який має наступний прямий зв'язок з неперервним часом, t , і періодом дискретизації вимірів T_s : $t = kT_s$. Функція прогнозування дає можливість спростити обчислення оцінок прогнозів завдяки зручній структурі різницевих рівнянь – побудованих моделей досліджуваних процесів. Оцінка прогнозу основної змінної на s кроків стосовно деякого довільного моменту часу k за означенням, яке дає теорія оцінювання, обчислюється за допомогою інформації про досліджуваний процес на момент k , включно таким чином:

$$\hat{y}(k+s) = E_k[y(k+s)|y(k), y(k-1), \dots, y(0), \varepsilon(k), \varepsilon(k-1), \dots, \varepsilon(0)]. \quad (4.4)$$

Як приклад, функцію прогнозування можна отримати за моделлю АР(1) таким чином:

$$\hat{y}(k+s) = E_k[y(k+s)] = a_0 \sum_{i=0}^{s-1} a_1^i + a_1^s y(k), \quad (4.5)$$

де $\hat{y}(k+s)$ – оцінка прогнозу основної (залежної) змінної $y(k)$ на s кроків вперед; E_k – позначення умовного математичного сподівання стосовно k -го (дискретного) моменту часу; a_0, a_1 – параметри моделі АР(1). Ще один приклад побудови такої функції: створення функції прогнозування процесу АРКС(2,1) на три кроки вперед:

$$\begin{aligned} \hat{y}(k+3) &= E_k[y(k+3)] = a_0 + a_1 E_k[y(k+2)] + a_2 E_k[y(k+1)] = \\ &= a_0(1 + a_1 + a_1^2 + a_2) + (a_1^3 + 2a_1 a_2)y(k) + (a_1^2 a_2 + a_2^2)y(k-1) + \\ &\quad + \beta_1(a_1^2 + a_2)\varepsilon(k), \end{aligned} \quad (4.6)$$

де a_0, a_1, a_2 і β_1 – параметри моделі; тут для обчислення прогнозу використовується оцінка випадкового процесу $\hat{\varepsilon}(k) = y(k) - \hat{y}(k)$; $y(k)$ – фактичне значення основної змінної в момент часу k (тобто її останній вимір); $\hat{y}(k)$ – оцінка залежної змінної, обчислена за моделлю: $\hat{y}(k) = \theta^T \psi(k)$, де $\psi(k)$ – вектор вимірів змінних, що містяться у правій частині рівняння. Можна отримати рекурсивну формулу для обчислення оцінок прогнозів процесу АРКС(2,1) на довільну кількість кроків s у такому вигляді:

$$\hat{y}(k+s) = E_k[y(k+s)] = a_0 + a_1 E_k[y(k+s-1)] + a_2 E_k[y(k+s-2)]. \quad (4.7)$$

Подані вище функції прогнозування та інші, створені за подібною методикою, забезпечують обчислення незміщених оцінок прогнозів. При цьому дисперсія похибок прогнозів асимптотично збігається до константи при зростанні кількості кроків прогнозування за умови, що: $E[\varepsilon(k)] = 0$ і коваріація $E[\varepsilon(k)\varepsilon(j)] = 0$, якщо $k \neq j$ [64, 68].

Адаптацію моделі у формі байєсівської мережі розглянуто у попередньому розділі; вона виконується у такій послідовності:

1. Реалізація процедури корегування структурної частини моделі:
 - процедура видалення дуг, що не відповідають множині даних;
 - додавання нових дуг.
2. Виконання процедури корегування ймовірнісної частини моделі – таблиць умовних ймовірностей.

На початковому етапі навчання байєсівської мережі ймовірнісна складова моделі створюється завдяки обчисленню таблиць умовних ймовірностей. Числові значення ймовірностей для цієї таблиці отримують шляхом частотного аналізу значень змінних у наявних спостереженнях, тому можна встановити зміни у процедурі модифікації ймовірнісної частини наявної моделі. Для виконання процедури модифікації ймовірнісної частини моделі, доцільно зберігати не повністю таблицю розподілу, а тільки значення N_{ijk} (кількість вузлів у конкретних випадках аналізу локальних структур мережі). Такий підхід надає можливість прискорено модифікувати дані стосовно значень умовних ймовірностей, а самі значення цих ймовірностей можна обчислити за допомогою виразу Діріхле:

$$p(X_i = v_{ik} | \Pi_i = \varphi_{ij}) = \frac{N_{ijk} + 1}{N_{ij} + r_i}. \quad (4.8)$$

Для модифікації структури моделі порядок обходу вузлів можна встановити за конкретним вкладом вибраного вузла в значення умовної ймовірності:

$$p(D_1|D_0, S_0) = \prod_{i=1}^n \frac{\prod_{s=1}^{R_i} \prod_{t=1}^{Q_i} \prod_{u=1}^{m_{its}} (N_{its} + u)}{\prod_{t=1}^{Q_i} \prod_{u=1}^{M_{it}} (N_{it} + r_i - 1 + u)}. \quad (4.9)$$

При цьому суть аналізу інформаційної важливості дуги мережі полягає у такому. Для перевірки дуги на можливість її видалення для кожного вузла моделі обчислюється величина $K_{delete}(S_0)$ для існуючої конфігурації наявних вузлів-предків, а також значення $K_{delete}(S_{-1}^m)$ для тих конфігурацій, які є результатом видалення однієї з M ($1 \leq m \leq M$) вхідних дуг з поточного вузла, що аналізується. Якщо при цьому виконується умова $K_{delete}(S_{-1}^m) \leq K_{delete}(S_0)$, то m -а дуга залишається в моделі, оскільки видалення цієї дуги призводить до зменшення значення локального функціоналу якості для поточного вузла. Інакше, дуга заноситься в перелік дуг, що підлягають подальшому аналізу на можливість видалення. Цей перелік надалі сортується за збільшенням значення $K_{delete}(S_{-1}^m)$. Надалі перевірка структури полягає у обчисленні локального функціоналу якості для вихідної конфігурації та конфігураціях, які створюються при видаленні однієї з дуг, що залишилися у переліку. Тактика вилучення і додавання дуг застосована в інкрементному варіанті адаптаційного алгоритму. Результатом застосування байєсівського аналізу є вибір стратегії адаптації моделі з використанням такого функціоналу:

$$P(S_1|D_1, D_0, S_0) = \arg \max_S \frac{P(S|D_0)P(D_1|S, D_0)}{P(D_1|S_0, D_0)}. \quad (4.10)$$

Тактика вилучення дуг з моделі повинна приводити до зменшення першої компоненти у чисельнику $P(S|D_0)$, оскільки це значення досягає максимуму при $S = S_0$ (у процесі формування початкової структури мережі). Для отримання позитивного результату адаптації необхідно компенсувати втрати від вилучення однієї дуги впливом від додавання нової дуги. Оскільки вихідною умовою застосованого алгоритму (K2) є

створення упорядкованої послідовності вузлів, то пошук дуги-претендента на додавання здійснюється саме в такому порядку. Оцінювання вкладу дуги виконується шляхом обчислення значення локального функціоналу якості. При цьому претендент на додавання повинен визначати таку множину вхідних дуг, що має максимальне значення локального функціоналу [69].

4.3 Оцінювання адекватності моделей і якості прогнозів

Важливим моментом процесу прогнозування є об'єктивне визначення якості отриманого прогнозу. Оскільки оцінки прогнозів – це випадкові величини, то для визначення їх якості необхідно використовувати множину відповідних статистичних критеріїв. Саме множину, а не один критерій, оскільки кожен критерій характеризує одну властивість оцінки прогнозу. Іноді якість оцінок прогнозів визначають лише за допомогою середньоквадратичної похибки (СКП). Однак, значення СКП – це лише одна із множини можливих статистик, яка залежить від масштабу даних, а тому цієї характеристики явно недостатньо для аналізу якості прогнозу.

Якість лінійних та псевдолінійних регресійних моделей оцінюють за допомогою декількох статистичних критеріїв якості, зокрема таких: коефіцієнт множинної детермінації (R^2), який характеризує інформативність моделі по відношенню до інформативності даних; статистика Дарбіна-Уотсона (DW), що визначає ступінь автокорельованості похибок моделі; інформаційний критерій Акайке (AIC) і статистика Байєса-Шварца (BSC); сума квадратів похибок моделі ($\sum e^2(k)$); F – статистика Фішера та інші.

Поглиблене оцінювання якості прогнозів за регресійними моделями досягається за рахунок використання критеріїв, які дають відносні оцінки якості (наприклад, коефіцієнт Тейла) та оцінки якості у процентах

(наприклад, середня абсолютна похибка у процентах (САПП)). Переваги їх використання полягають у тому, що вони не залежать від масштабу даних і легко інтерпретуються особою, що приймає рішення. САПП і коефіцієнт Тейла обчислюють за виразами:

$$САПП = \frac{1}{s} \sum_{i=1}^s \frac{|y(k+i) - \hat{y}(k+i, k)|}{|y(k+i)|} \times 100\% = \frac{1}{s} \sum_{i=1}^s \frac{|e(k+i)|}{|y(k+i)|} \times 100\%, \quad (4.11)$$

$$U = \frac{\sqrt{\frac{1}{s} \sum_{k=1}^s [y(k+i) - \hat{y}(k+i)]^2}}{\sqrt{\frac{1}{s} \sum_{i=1}^s y^2(k+i) + \frac{1}{s} \sum_{i=1}^s \hat{y}^2(k+i)}}, \quad (4.12)$$

де s – кількість кроків прогнозування; $y(k+i)$ – фактичні значення даних; $\hat{y}(k+i)$ – оцінки прогнозів відносно k -го моменту часу, на який наявна вся інформація про досліджуваний процес. Коефіцієнт Тейла U – це важлива характеристика якості моделі і прогнозу, за означенням $0 \leq U \leq 1$. Якщо $U \rightarrow 0$, то оцінки прогнозів наближаються до фактичних значень ряду і модель має високу ступінь адекватності. Тобто U дає можливість встановити придатність моделі (і методу оцінювання прогнозу на її основі) для оцінювання прогнозу в принципі.

Для автоматизованого вибору кращої моделі можна скористатись інтегральним критерієм якості [69]:

$$V_N(\theta, D_N) = e^{1-R^2} + \frac{SSE}{N} + \left\{ \begin{array}{l} \ln(AIC + BSC), \text{ якщо } AIC + BSC > 0 \\ e^{AIC+BSC}, \text{ якщо } AIC + BSC \leq 0 \end{array} \right\} + e^{2-DW} + \ln(CKIT) + \ln(САПП) + e^U, \quad (4.13)$$

де D_N – дані, що використовуються для оцінювання структури і параметрів моделі; СКП – середньоквадратична похибка однокрокового прогнозу на навчальній (історичній) вибірці; САПП – середня абсолютна похибка прогнозу в процентах; U – коефіцієнт Тейла (наближається до нуля, якщо модель придатна для прогнозування). Альтернативним варіантом використаного інтегрального критерію є такий:

$$V_N(\theta, D_N) = e^{|1-R^2|} + \ln\left(1 + \frac{SSE}{N}\right) + e^{|2-DW|} + \ln(1 + CKIT) + \ln(1 + CAIPT) + e^U, \quad (4.14)$$

де N – кількість вимірів часового ряду даних. Потужність цих критеріїв перевірено експериментально і встановлено, що вони дають можливість вибрати кращу модель практично із ймовірністю, близькою до одиниці.

У багатьох випадках кращих результатів прогнозування можна досягти за рахунок усереднення або комбінування за допомогою вагових коефіцієнтів оцінок прогнозів, отриманих, наприклад, за допомогою оптимізаційних методів. При цьому необхідно задовольнити такі умови: похибки оцінок прогнозів, отриманих за різними методами, мають бути некорельованими, а дисперсії цих похибок близькими за своїми значеннями.

4.3.1 Адаптивне оцінювання прогнозів

Для того щоб зберегти високу якість прогнозів при дослідженні нестационарних процесів необхідно застосовувати адаптивні алгоритми обчислення оцінок прогнозів. Вхідними значеннями для алгоритму аналізу якості прогнозів є значення похибок прогнозів і вибрані статистичні характеристики їх якості. Для досягнення необхідної *структурної адаптації* прогнозуючої моделі до змін, що відбуваються у

досліджуваному процесі та до заданих вимог стосовно необхідної якості прогнозу необхідно скористатись такими можливостями СППР:

- аналіз типу розподілу ймовірностей даних і його параметрів, а також врахування отриманого результату для коректного вибору методу оцінювання параметрів моделі;

- проведення автоматизованого аналізу часткової автокореляційної функції (ЧАКФ) даних для залежної змінної з можливим подальшим корегуванням структури моделі завдяки додаванню/вилученню відповідних лагових значень змінної;

- додавання у модель можливих регресорів та аналіз їх впливу на якість прогнозу; корисними для оцінювання прогнозів є провідні індикатори, це ті регресори, що вводяться в модель з лагами більшими одиниці; вони дають можливість коректно обчислювати прогнози на таку кількість кроків вперед, що відповідає фактичному лагу змінної;

- автоматизований аналіз та вибір оптимальних вагових коефіцієнтів в алгоритмах експоненційного згладжування, побудови регресії на опорних векторах та інших методах, що використовуються у моделюванні нелінійних нестационарних процесів;

- використання автоматизованого аналізу залишків регресійних моделей, спрямованого на встановлення їх фактичної інформативності та можливе корегування структури моделі на основі отриманих результатів аналізу;

- коректне створення множини вимірів змінних стану процесу з використанням методів ієрархічного інтегрування даних, що, як правило, забезпечує підвищення їх інформативності.

Параметрична адаптація моделі до даних досягається завдяки застосуванню повторного оцінювання параметрів моделей з надходженням нових даних; це сприяє уточненню параметрів моделі та підвищенню адекватності моделі і якості прогнозу. Для оцінювання однакових структур моделей доцільно застосовувати альтернативні методи, що дає можливість

побудувати додаткові моделі-кандидати з метою виконання їх подальшого аналізу і вибрати кращі з них для практичного використання.

Створення та застосування конкретного алгоритму адаптації залежить від постановки задачі, якості та наявного об'єму статистичних даних, характеристик досліджуваного процесу, вимог до якості оцінок прогнозів і часу, який може бути наданий для виконання необхідних обчислень. Кожний алгоритм адаптивного обчислення оцінок прогнозів має свої особливості, які враховуються при проектуванні підсистеми адаптивного моделювання і прогнозування.

4.4 Приклади оцінювання операційних ризиків

Розглянемо застосування системи підтримки прийняття рішень для моделювання шахрайських виплат – операційної події, яка часто трапляється в страхових компаніях. У цьому прикладі ми моделюємо вартість шахрайських виплат, що виникають протягом тижня з портфеля комерційного страхування від вогню з річною валовою премією 1 мільйон гривень. Шахрайські виплати можуть включати випадки підпалу з боку страхувальника або штучного завищення збитків [6].

Основна гіпотеза полягає в тому, що рівень контролю при андеррайтингу нових полісів разом зі стадією економічного циклу відповідає за пояснення шахрайства. Рівень контролю при андеррайтингу визначає кількість страхувальників з більшою схильністю до подання шахрайських заяв, які приймаються на баланс страховика. Під час економічних спадів такі страхувальники можуть вдаватися до підпалів або штучного завищення страхових виплат, щоб отримати економічну вигоду. Поліси втраченого доходу, як і багато інших видів комерційного страхування від вогню, особливо вразливі до таких зловживань. Рівень андеррайтингового контролю залежить від наступних факторів:

- Навички оцінювача: старші андеррайтери краще здатні оцінити, які пропозиції можуть призвести до шахрайського відшкодування;
- Величина суб'єкта: великий розмір бізнесу чинить тиск на всю страхову систему і знижує якість андеррайтингу;
- Аутсорс: андеррайтинг менш складних ризиків може бути частково переданий на аутсорсинг співробітникам філій для зменшення витрат і підвищення ефективності, але це також збільшує ризик шахрайства через менший контроль за якістю андеррайтингу на рівні філій.

Очевидно, що відповідальність за виявлення шахрайських виплат покладається на відділ врегулювання збитків. Ймовірність виявлення залежить від рівня контролю за виплатами. Якщо відділ виявляє шахрайство, він може вжити заходів для зменшення збитків, ініціювавши розслідування для виявлення доказів, які можуть бути використані для відмови або зменшення виплати. Ці заходи є дорогими і не завжди можуть усунути шахрайські виплати страхових відшкодувань.

Рівень контролю в управлінні страховими виплатами залежить від таких факторів:

- Практика спеціаліста: більш досвідчений аварійний комісар з більшою ймовірністю виявить шахрайські заяви;
- Стороння експертна оцінка: залучення аварійних комісарів, які спеціалізуються на оцінці розміру збитків, завданих пожежею: це допомагає зменшити штучно завищені збитки, також аварійні комісари можуть визначити причину пожежі, особливо якщо можливий підпал;
- Вибірковий аудит: компанія може проводити вибіркові перевірки страхових випадків, щоб оцінити їх більш детально перед тим, як затвердити виплату.

Нарешті, вартість страхування буде залежати як від поширеності шахрайства, так і від ступеня його виявлення. Навіть якщо всі випадки шахрайства будуть виявлені, повністю уникнути витрат неможливо через неможливість спростувати всі підозрілі заяви та витрати на розслідування

шахрайських заяв. Випадкові змінні, що використовуються в моделі, наведені в Таблиці 4.1.

Таблиця 4.1. Вершини байєсівської мережі та їх можливі значення

Номер вершини	Назва вершини	Значення в мережі (назва)	Варіанти значень
1	Розмір суб'єкта	Subject volume	High \ Low
2	Довіра до аутсорсинга	Outsource reliance	Yes \ No
3	Навички оцінювача	Estimator skills	Senior \ Junior
4	Зовнішній експерт	External expert	Yes \ No
5	Аудит	Audit	Yes \ No
6	Спеціалізована практика	Specialist practice	Senior \ Junior
7	Вартість обману	Cheating value	Amount
8	Викриття обману	Cheating exposed	Yes \ No
9	Несанкціонована вимога	Unauthorized demand	Yes \ No
10	Глобальний фінансовий цикл	Global financial cycle	Up \ Down
11	Контроль вимог	Requirements control	High \ Low
12	Страховий контроль	Insurance control	High \ Low

Вершина "Вартість обману" має встановлені значення: 0, 50 000, 10 000 або 150 000 грн., але це не означає, що сума збитків може бути тільки такою. Ці значення встановлені тому, що потім необхідно застосувати лінійну інтерполяцію для розрахунку значення капіталу під ризиком (CaR) для різних довірчих інтервалів. Значення вузла "Ціна шахрайства" слід розуміти як "втрати не перевищують 50 000 грн", "втрати близькі до нуля" тощо.

Після побудови загальної структури байєсівської мережі необхідно задати розподіли ймовірностей для змінних вузлів. Всі розподіли ймовірностей є результатом експертних оцінок, оскільки статистичних даних про операційні збитки страхових компаній у відкритому доступі немає.

Умовні ймовірності для вузлів байєсівської мережі: "Навички оцінювача", "Аутсорс", "Величина суб'єкта", "Практика спеціаліста", "Вибірковий аудит", "Стороння експертна оцінка" наведені в Таблиці 4.2.

Таблиця 4.2. Умовні розподіли вузлів (змінних)

Вершини											
Навички оцінювача		Аутсорс		Величина суб'єкта		Практика спеціаліста		Вибірковий аудит		Стороння експертна оцінка	
Senior	Junior	Yes	No	High	Low	Senior	Junior	Yes	No	Yes	No
0.4	0.6	0.3	0.7	0.6	0.4	0.7	0.3	0.7	0.3	0.6	0.4

Одним з основних застосувань байєсівських мереж є моделювання різних сценаріїв розвитку подій. Створивши модель, страхова компанія може вирішити, який обсяг капіталу слід виділити для захисту компанії від усіх, окрім крайніх випадків шахрайства. На цьому етапі модель може бути скоригована для відображення фактичного відомого стану певних змінних, щоб показати реальний стан компанії. Змоделюємо варіант ситуації, коли страхова компанія не має фактичних даних і може покладатися лише на експертні оцінки.

Маргінальні значення кожного вузла показано на Рисунку 4.3. Ця модель тепер може бути використана для стрес-тестування, причинно-наслідкового моделювання та розподілу капіталу.

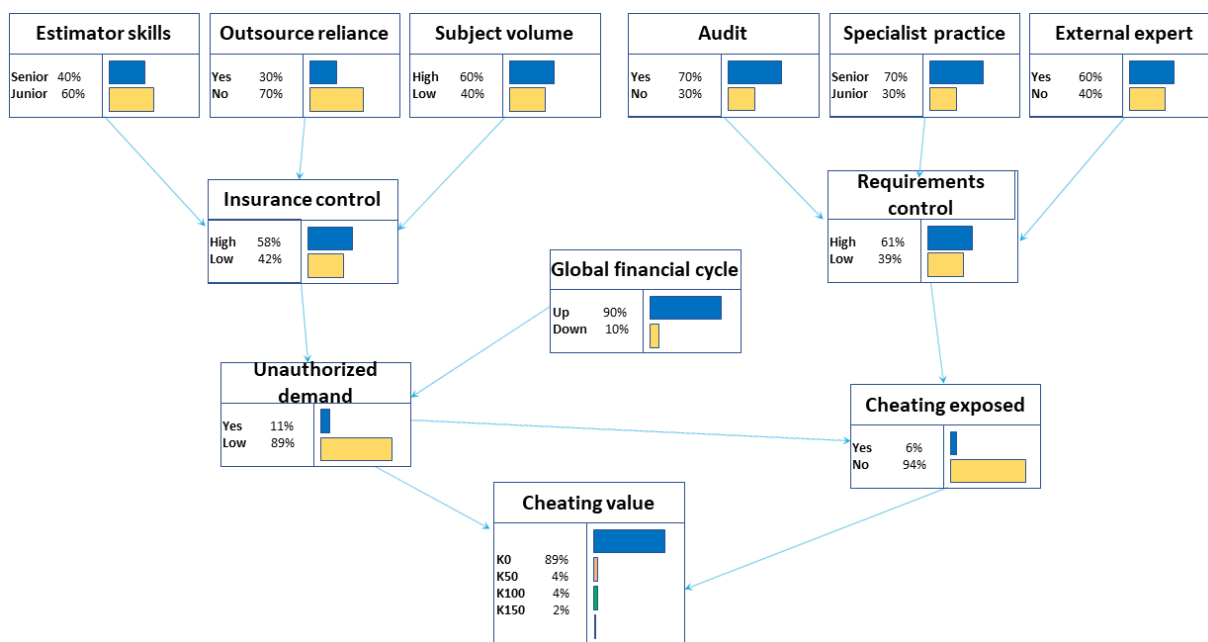


Рисунок 4.3. Маргінальні значення вузлів байєсівської мережі

На Таблиці 4.3 наведено розрахований капітал для покриття збитків, спричинених ОР CaR, за різних довірчих інтервалів: 90, 95 і 99 %. Такі рівні довіри є стандартними для розрахунку CaR у фінансових установах [6].

Таблиця 4.3. Капітал на покриття ОР із заданими рівнями довіри

Процентиль, %	90	95	99
CaR, грн	54 927	70 375	131 478

Логічно, що витрати на капітал (CaR) зростають разом із рівнем довіри. Фінансові установи, що використовують метод АМА (advanced management approach), повинні бути спроможними покрити витрати з 99,9% рівнем довіри. Але фінансові установи, зокрема й страхові компанії, розраховують капітал для 95% рівня довіри. За 95% рівня довіри CaR дорівнює 64 286 грн. При використанні методу ВІА (basic indicator approach) із річним валовим доходом в 1 млн дол. треба виділити 150 тис. грн для покриття ОР, що майже у 2,5 раза більше, ніж за АМА [2].

Побудована байєсівська мережа може бути також використана для тестування різних сценаріїв, щоб допомогти менеджерам оптимізувати профіль ризику. Наприклад, якщо страхова компанія бажає зменшити витрати, знизивши рівень “Навички оцінювача” до “Низького”, ефект цієї дії можна легко дослідити, задавши мережі значення “Навички оцінювача” = “Low”. Потім інформація поширюється по мережі, і отримуються нові умовні розподіли. Розглянемо один з можливих сценаріїв у страховій компанії. У Таблиці 4.4 показано, яких значень набувають деякі вузли.

Таблиця 4.4 – Апостеріорні значення для перевірки сценарію

Вершини	Значення
Estimator skills	Junior
Outsource reliance	Yes
Subject volume	High
Audit	No
Specialist practice	Junior
External expert	No
Global financial cycle	Down

Отримані результати зображено на Рисунку 4.4:

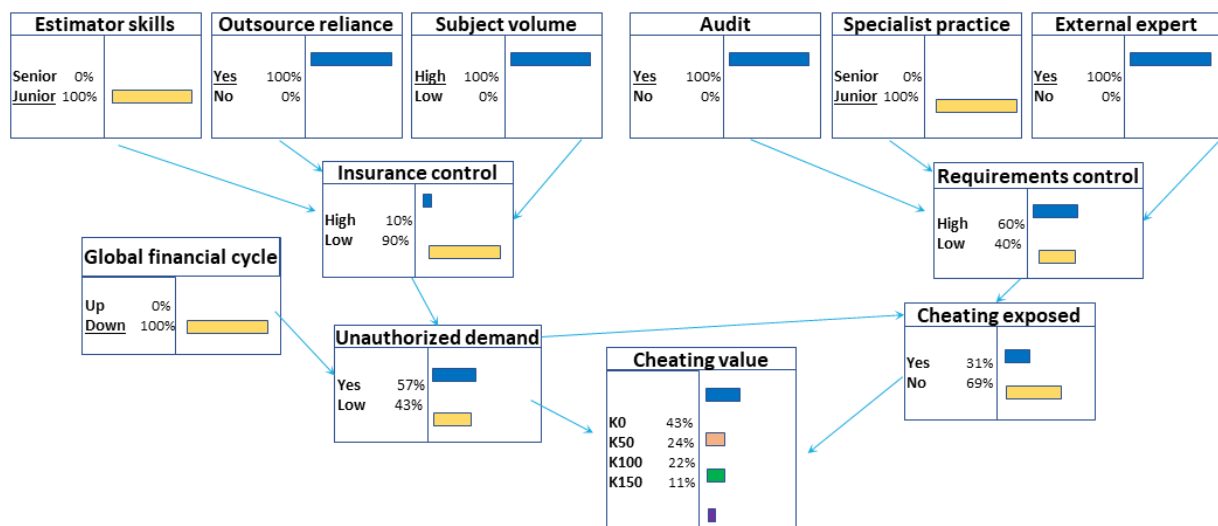


Рисунок 4.4. Умовні розподіли ймовірностей за сценарієм

Розглянемо інший приклад застосування Системи підтримки прийняття рішень. В цьому разі проведемо аналіз якості автоматизованого обслуговування клієнтів (аналіз операційного ризику): випадок дискретних даних і дискретних параметрів.

Страхова компанія запровадила автоматизований сервіс для клієнтів, який передбачає автоматичну реєстрацію клієнтів, що страхують свої автомобілі. Кількість клієнтів, які користуються послугою, сягає декількох тисяч на місяць. Розглянемо задачу оцінювання операційної похибки після обслуговування клієнтів.

Для спрощення постановки задачі припустимо, що може приймати наступні три значення: хороший результат; прийнятний результат; поганий результат, який не може бути прийнятий. Наявна статистика показує, що протягом попередніх двох років компанія забезпечувала автоматизоване обслуговування клієнтів з наступною якістю:

- протягом 60% часу ймовірність помилки обслуговування була на рівні (хороший результат);
- протягом 30% часу ймовірність помилки обслуговування була на рівні (прийнятний результат)
- протягом 10% часу ймовірність помилки обслуговування була на рівні (результат, який не може бути прийнятий).

Ці результати були використані як апіорні ймовірності для прогнозування рівня обслуговування в майбутньому. Цей розподіл наведено в Таблиці 4.5.

Таблиця 4.5. Попередні ймовірності для параметру

	Якість обслуговування		
	Хороша	Прийнятна	Неприйнятна
Ймовірність помилок в обслуговуванні (θ)	0.25	0.50	0.75
Щільність помилок $\theta, p(\theta)$	0.60	0.30	0.10

Після 10 000 випадків обслуговування клієнтів компанія вирішила провести контроль якості обслуговування. Контроль показав, що з десяти випадково вибраних випадків обслуговування два містять помилки. Питання в тому, який висновок щодо якості обслуговування слід зробити в цьому випадку. Тобто, іншими словами, яким є фактичний апостеріорний розподіл для параметра.

У цьому випадку дані мають дискретну форму і спочатку необхідно визначити (скажімо, на основі попереднього досвіду) тип розподілу для даних. На основі попереднього досвіду можна припустити, що вони мають біноміальний розподіл з параметром θ :

$$f(\theta, n, r) = \binom{n}{r} \theta^r (1 - \theta)^{n-r} = C_n^r \theta^r (1 - \theta)^{n-r}, \quad (4.15)$$

де $\binom{n}{r} = \frac{n!}{r!(n-r)!}$. Кількість успішних подій у цьому випадку дорівнює $r =$

2. Успішними є події, пов'язані з появою помилок в обслуговуванні клієнтів з 10 можливих, тобто $n = 10$ [6]. Таким чином, функція правдоподібності для даних у цьому випадку має наступний вигляд:

$$L(\theta) = \binom{n}{r} \theta^r (1 - \theta)^{n-r} = C_{10}^2 \theta^2 (1 - \theta)^{10-2}, \quad (4.16)$$

де $\theta = [0.25, 0.5, 0.75]$ – розподіл можливих подій, пов'язаних з кількістю помилок в обслуговуванні. Номінатор правила Байєса в цьому випадку наступний:

$$h(\theta | r, n) \propto L(r | \theta, n) g(\theta) \quad (4.17)$$

Тепер обчислимо необхідні значення правдоподібності, $L(\theta)$:

- якщо $\theta = 0.25$, то $L(\theta) = \frac{n!}{r!(n-r)!} \times (0.25)^2(1 - 0.25)^8 = 0.30300$;
- якщо $\theta = 0.50$, то $L(\theta) = \frac{n!}{r!(n-r)!} \times (0.50)^2(1 - 0.50)^8 = 0.04405$;
- якщо $\theta = 0.75$, то $L(\theta) = \frac{n!}{r!(n-r)!} \times (0.75)^2(1 - 0.75)^8 = 0.00039$.

Апріорні та апостеріорні щільності для наведено в Таблиці 4.6.

Таблиця 4.6. Апостеріорні ймовірності

θ	Апріорні ймовірності для θ	Ймовірність $L(\theta)$	h – апостеріорна щільність для θ
0.25	0.60	0.30300	0.955
0.50	0.30	0.04405	0.034
0.75	0.10	0.00039	0.010
Загалом:			0.999

З Табл. 4.6 видно, що найбільш ймовірним значенням θ , оціненим на основі аналізу якості обслуговування у випадку автоматизованого обслуговування, є значення: $\theta = 0.25$. Це означає, що якість обслуговування знаходиться на рівні "добре" (апостеріорна ймовірність для становить 0,955).

Зі збільшенням кількості перевірок випадків за допомогою автоматизованої системи обслуговування, якість оцінювання буде зростати.

4.5 Висновки до розділу

Розглянуто причини виникнення операційного ризику, який надзвичайно притаманний промисловості, фінансовим послугам, транспортуванню тощо. Різноманітність причин виникнення такого ризику вимагає комплексного підходу до його оцінки та управління. Представлено огляд байєсівських моделей аналізу, моделювання та оцінювання ОР.

Проведено дослідження архітектури і функцій системи підтримки прийняття рішень, а також побудову функцій прогнозування в СППР.

Створено модель операційного ризику на основі узагальненої лінійної регресійної (байєсівської) моделі. Показано, що імовірнісні байєсівські моделі працюють задовільно і дозволяють оцінити необхідний ризиковий капітал (Capital at Risk). Побудована модель ОР страхового шахрайства ілюструє можливість практичного застосування розглянутого методу моделювання.

ВИСНОВКИ

В роботі детально опрацьоване питання актуальності дослідження операційного ризику та розробки системи підтримки прийняття рішень (СППР) для його аналізу.

Сформульована класифікація джерел операційних ризиків у фінансових організаціях, а також різноманітні моделі для їх формалізації, особливості застосування і побудови кожної з них. Сформульовані різні джерела невизначеності в фінансовому моделюванні та прийнятті рішень. Введена класифікація типів невизначеності, що можуть виникати в різних аспектах фінансового моделювання та відповідно впливати на процеси прийняття рішень.

Розроблені нові моделі з урахуванням ймовірнісно-статистичних невизначеностей даних у формі байєсівської мережі (БМ) і байєсівської регресії, які відрізняються можливістю. Побудовані моделі спрямовані на покращення якості оцінювання можливих операційних втрат.

З метою підвищення адекватності ймовірнісних моделей було вперше застосовано метод структурно-параметричної адаптації моделей у формі БМ до нових даних. Виконано повторне оцінювання структури і параметрів моделі з використанням заданої множини критеріїв.

Удосконалено алгоритм обчислювальних процедур оцінювання структури і параметрів моделей операційних ризиків. Проведено удосконалення методу Монте-Карло для марковських ланцюгів.

Розроблена оригінальна СППР на основі принципів системного аналізу для розв'язання задачі моделювання, оцінювання і прогнозування можливих втрат внаслідок реалізації операційних ризиків на основі статистичних даних та експертних оцінок, розроблених математичних моделей і множини статистичних критеріїв якості.

Результати роботи впроваджено у навчальний процес інституту прикладного системного аналізу НТУУ «КПІ імені Ігоря Сікорського».

Запропоновану методологію впроваджено в фінансовій компанії ТОВ «ФІНАНСОВА КОМПАНІЯ «КОНКОРД ФАКТОРИНГ». Завдяки результатам впровадження, фахівці з фінансового аналізу компанії мають змогу обґрунтовано підійти до розв'язання задачі прогнозування розвитку сучасних фінансових процесів, оцінювання можливих фінансових ризиків та їх врахування у виробленні тактичних і стратегічних управлінських рішень.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Л.Б. Левенчук, П.І. Бідюк Байєсівський аналіз даних у моделюванні і прогнозуванні нелінійних нестационарних процесів, KPI Science News, 2020, No. 3, P. 14-23. DOI: 10.20535/kpi-sn.2020.3.209877
2. Л.Б. Левенчук, В.Г. Гуськова, П.І. Бідюк Ймовірнісне моделювання операційних ризиків, KPI Science News, 2021, No. 3, P. 26-37. DOI: 10.20535/kpism.2021.3.251681
3. N.V. Kuznietsova, O.M. Trofymchuk, P.I. Bidyuk, O.M. Terentiev, L.B. Levenchuk, Bayesian modeling of risks of various origin, KPI Science News, 2021, No. 4, P. 7-18. DOI: 10.20535/kpism.2021.4.251684
4. Kuznietsova N. V., Huskova V. H., Bidyuk P. I., Matsuki Y., Levenchuk L. B. Modelling risk factors interaction and risk estimation with copulas, Radio Electronics, Computer Science, Control, 2022, N. 2, P. 44-52. DOI: 10.20535/kpism.2021.4.251684
5. Kasyanov, P., Levenchuk, L.B. (2023). Formalization and Development of Autonomous Artificial Intelligence Systems. In: Zgurovsky, M., Pankratova, N. (eds) System Analysis and Artificial Intelligence. Studies in Computational Intelligence, vol 1107. Springer, Cham. DOI: 10.1007/978-3-031-37450-0_9
6. Levenchuk L. The Bayesian approach to analysis of financial operational risk, ScienceRise, 2022, No. 2, P. 11–20. DOI: 10.21303/2313-8416.2022.002377
7. Левенчук Л.Б., Бідюк П.І. Ймовірнісне моделювання нелінійних нестационарних процесів, 15-а Міжнародна Науково-технічна конференція «Проблеми інформатизації» у Парижі, 2020, С. 18-19.
8. P.I. Bidyuk, N.V. Kuznietsova, L.B. Levenchuk, V.G. Guskova Forecasting non-stationary processes using bayesian data analysis techniques», 15-а Міжнародна Науково-технічна конференція «Проблеми інформатизації» у Парижі, 2020, С. 26-27.

9. Левенчук Л.Б., Бідюк П.І. Байєсівський аналіз даних у моделюванні і прогнозуванні нелінійних нестационарних процесів, Радіоелектроніка і молодь у XXI столітті: матеріали XXIV Міжнародного молодіжного форуму (м.Харків, 7-9 квіт. 2020 р.). Харків: ХНУРЕ, 2020. Т. 5. С. 185-186.
10. Левенчук Л.Б., Бідюк П.І. Ймовірнісне моделювання нелінійних нестационарних процесів. Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління: тези доп. 10-ї міжнар. наук.-техн. конф. (м. Баку, м. Харків, м. Жиліна, 9-10 квіт. 2020 р.). Харків: ФОП Петров В. В., 2020. Т. 2: секції 3, 4. С. 16.
11. Petro Bidyuk, Sergii Polozhaenko, Nataliia Kuznietsova, Liudmyla Levenchuk. Probabilistic Data Analysis in Non-Stationary Processes Forecasting. 2020 IEEE 2nd International Conference on System Analysis & Intelligent Computing (SAIC), 5-9 Oct. 2020, P. 1-6.
12. Левенчук Л.Б. Моделі операційних ризиків, Радіоелектроніка і молодь у XXI столітті: матеріали XXV Міжнародного молодіжного форуму (м.Харків, 20-22 квіт. 2021 р.). Харків: ХНУРЕ, 2021. Т. 5. С. 201-202.
13. Levenchuk L.B., Bidyuk P.I., Gavrylenko V.B. Forecasts estimation: systemic approach, VIII-а Міжнародна НТК: «Сучасні методи, інформаційне, програмне та технічне забезпечення систем керування організаційно-технічними та технологічними комплексами», Київ, НУХТ, 25 – 26 листопада 2021 року, P. 157-158.
14. Бідюк П.І., Левенчук Л.Б. Методика побудови скорингових карт із використанням множини методів аналізу даних. Прикладні системи та технології в інформаційному суспільстві. VI Міжнародна наук.-практ. конференція, м. Київ, 30 вересня 2022, С.23-26.
15. Bidyuk P.I., Romanenko V.D., Tymoshchuk O.L. Time Series Analysis. – Kyiv: NTUU “Igor Sikorsky KPI”, 2012, 450 p.

- 16.P. D. Hoff, A First Course in Bayesian statistical methods. London: Springer-Verlag, 2009, pp. 31-65. ISBN 9780387922997
- 17.J. Pearl, Probabilistic reasoning in intelligent systems. San Francisco, CA: Morgan Kaufmann Publishers, 1988, pp. 77-141. ISBN 9780080514895
- 18.S. J. Press, Subjective and objective Bayesian statistics. Hoboken, New Jersey: Wiley-Interscience, 2002, pp. 264-280. ISBN: 9780471348436
- 19.M. Z. Zgurovskii and P. I. Bidyuk, Bayesian networks in decision support systems. Kyiv, Ukraine: Edelweiss, 2015, pp. 12-23.
- 20.Gibbs B.P. Advanced Kalman Filtering, Least-squares and Modeling. – New York, John Wiley & Sons, Inc., 2011, 627 p.
- 21.B. Chen, “A Bayesian sampling approach to decision fusion using hierarchical models”, IEEE Trans. Signal Process. vol. 50, no. 8, pp. 1809-1818, 2002.
- 22.S. O. Dovgiy et al., “Decision Support Systems Based on Statistical and Probabilistic Methods”, Logos, pp. 94-107, 2014.
- 23.P. I. Bidyuk et al., Mathematical statistics. Kyiv, Ukraine: Vydavnychyy dim «Personal», 2018, pp. 138-243. ISBN 9786170202345
- 24.Bidyuk P.I., Tymoshchuk O.L., Kovalenko A.Ye. Decision Support Systems: Design and Implementation. – Kyiv: NTUU “Igor Sikorsky KPI”, 2020, 600 p.
- 25.Hollsapple C.W., Winston A.B. Decision support systems. – New York: West Publishing Company, 1996. – 860 p.
- 26.J. H. Dorfman, Bayesian economics through numerical methods: a guide to econometrics and decision making with prior information. New York: Springer-Verlag, 1997, pp. 88-96. ISBN 9780387226354
- 27.Lopez, J.A., 1999b. “Methods for Evaluating Value-at-Risk Estimates,” Federal Reserve Bank of San Francisco Economic Review, forthcoming.
- 28.Lopez, J.A., 1999a. “Regulatory Evaluation of Value-at-Risk Models,” Journal of Risk, forthcoming.

29. Mori T., Harada E., 2001. “Internal Measurement Approach to Operational Risk Capital Charge”, Bank of Japan Discussion Paper.
30. Lubbe J., Snyman F.. The advanced measurement approach for banks. - *Irving Fisher Committee on Central Bank Statistics IFC Bulletin*, 2009, Vol. 33, pp. 141 – 149.
31. Rudy S., Sapsis T., 2020. Sparse Bayesian Additive Model with Automatic Relevance Determination.
32. Building Powerful, Predictive Scorecards An overview of Scorecard module for FICO® Model Builder
33. Minitab Modern Analytics and Problem-Solving Solutions [Электронный ресурс] – Режим доступа: <https://www.minitab.com/en-us/products/>
34. Paper 2751-2018 SAS® Credit Scoring for Banking – An Integrated Solution from Data Capture to Insight Ewa Nybakk, Capgemini Norway
35. IBM SPSS Modeler: Drive ROI and accelerate time to value with an intuitive, drag-and-drop data science tool [Электронный ресурс] – Режим доступа: <https://www.ibm.com/products/spss-modeler>
36. Lutz M. Learning Python. O'Reilly Media Inc, 2013. — 1643 p. — ISBN 9781449355739
37. An Introduction to Statistical Learning / with Applications in R Authors: James, G., Witten, D., Hastie, T., Tibshirani, R
38. Harvey A, Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge, UK: University of Cambridge, 1990, pp. 100—167.
39. Tsay R., Multivariate Time Series Analysis with R and Financial Applications. Hoboken: Wiley-Interscience, 2014, pp. 279—319
40. Schwarz G. Estimating the dimension of a model. – *Annals Statistics*, 1978, Vol. 6, pp. 461 – 464.
41. Binder J., Koller D., Russel S., Kanazawa K. Adaptive probabilistic networks with hidden variables. – *Machine Learning*, 1997, Vol. 29, pp. 213 – 244.
42. Buntine W. Theory refinement on Bayesian networks. – In *UAI-91*, 1991, pp. 52 – 60.

43. Cooper G., Herskovits E. A Bayesian method for the induction of probabilistic networks from data. – *Machine Learning*, 1992, Vol. 9, pp. 309 – 347.
44. Friedman N. Learning belief networks in the presence of missing values and hidden variables. – In *ICML-97*, 1997.
45. Friedman N. The Bayesian structural EM algorithm. – In *ICML-97*, 1998.
46. Heckerman D., Geiger D., Chickering D.M. Learning Bayesian networks: The combination of knowledge and statistical data. – *Machine Learning*, 1995, Vol. 20, pp. 197 – 243.
47. Kanazawa K., Koller D., Russel S. Stochastic simulation algorithms for dynamic probabilistic networks. – In *UAI-95*, 1995, pp. 346 – 351.
48. Lam W., Bacchus F. Learning Bayesian belief networks: An approach based on the MDL principle. – *Computational Intelligence*, 1994, Vol. 10, pp. 269 – 293.
49. Бідюк П.І., Трухан С.В. Розробка та використання інформаційної системи аналізу та прогнозування фінансових процесів // Системні науки та кібернетика, 2013, № 1, с. 26 - 37.
50. Demster A.P., Laird N.M., Rubin D.B. Maximum likelihood from incomplete data via the EM algorithm. – *J. Royal Stat. Society, Series B*, 1977, Vol. 39, pp. 1 – 39.
51. Kjaerulff U. A computational scheme for reasoning in dynamic probabilistic networks. – In *UAI-92*.
52. Lauritzen S.L. The EM algorithm for graphical association models with missing data. – *Computational Stat. and Data Anal.*, 1995, Vol. 19, pp. 191 – 201.
53. Malik J., Russel S. Traffic surveillance and detection technology development: New sensor technology final report. – Final Report UCB-ITS-PRR-97-6, Cal. PATH Program, 1997.
54. Smyth P., Heckerman D., Jordan M. Probabilistic independence networks for hidden Markov probability models. – *Neural Computation*, 1997, Vol. 9(2), pp. 227 – 269.

55. Xu H. Computing marginals for arbitrary subsets from marginal representations in Markov Trees. – *Artificial Intelligence*, 1995, Vol. 74, pp. 177 – 189.
56. Speagle J., 2020. “A Conceptual Introduction to Markov Chain Monte Carlo Methods”, Center for Astrophysics| Harvard & Smithsonian.
57. M.S. Arulampalam, S. Maskell, N. Gordon, T. Clapp. A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking // *IEEE Transactions on Signal Processing*, Vol. 50, No. 2, pp. 174 – 188, 2002.
58. F. Gustafsson. Particle filter theory and practice with positioning applications // *IEEE Aerospace and Electronic Systems Magazine*, Vol. 25, No. 7, pp. 53–82, 2010.
59. Fox, D., Burgard, W., Dellaert, F., & Thrun, S. (1999). Monte carlo localization: Efficient position estimation for mobile robots. *AAAI/IAAI*, 1999(343-349), 2-2.
60. Röwekämper, J., Sprunk, C., Tipaldi, G. D., Stachniss, C., Pfaff, P., & Burgard, W. (2012, October). On the position accuracy of mobile robot localization based on particle filters combined with scan matching. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on* (pp. 3158-3164). IEEE.
61. Silverman, B.W. Density estimation for statistics and data analysis. Chapman & Hall, London, 1986.
62. Holsapple C.W., Winston A.B. Decision support systems. – Saint Paul (USA): West Publishing Company, 1996. – 850 p.
63. Turban E., Aronson J.E. Decision support systems. – New Jersey: Prentice Hall, 2001. – 865 p.
64. Бідюк П.І., Романенко В.Д., Тимошук О.Л. Аналіз часових рядів. – Київ: НТУУ «КПІ», 2013. – 600 с.
65. Press S.J. Subjective and objective Bayesian statistics. – Hoboken (New Jersey): John Wiley & Sons, Inc., 2013. – 560 с.

66. Rossi P.E., Allenby G.M., McCulloch R. Bayesian statistics and marketing. – New Jersey: John Wiley & Sons, Ltd, 2005. – 348 p.
67. Diebold F.X. Forecasting. – Pennsylvania: University of Pennsylvania, 2018. – 800 p.
68. Згуровський М.З., Подладчиков В.М. Аналітичні методи калмановської фільтрації. – Київ: Наукова думка, 1995. – 285 с.
69. Ng B.M. Adaptive dynamic Bayesian networks / 2007 Joint Statistical Meetings, Salt Lake City (USA), July 29 – August 2, 2007, pp. 1 – 7.
70. Zgurovsky M.Z., Bidyuk P.I., Terentyev O.M. Method of constructing Bayesian networks based on scoring functions // Cybernetics and System Analysis, 2008, Vol. 44, No.2, pp. 219-224.
71. <http://www.mataf.net/en/tools/home>
72. Nong Y. The Handbook of Data Mining. – New Jersey: Arizona State University Publishers, 2003. – 1201 p.
73. Altman E.I., Avery R.B., Eisenbeis R.A., Sinkey J. Application of Classification Techniques in Business, Banking and Finance. – Greenwich: JAI Press, 1981. – 418 p.
74. Hosmer D.W., Lemeshow S. Applied Logistic Regression. – New York: John Wiley & Sons, Inc., 2000. – 380 p.
75. Cowell R.G., Dawid A.P., Lauritzen S.L., Spiegelhalter D.J. Probabilistic networks and expert systems. – New York: Springer, 1999. – 323 p.