

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

**Н. В. Кузнєцова, П. І. Бідюк**

**ТЕОРІЯ І ПРАКТИКА  
АНАЛІЗУ ФІНАНСОВИХ РИЗИКІВ:  
СИСТЕМНИЙ ПІДХІД**

Монографія

Київ  
Видавництво Ліра-К  
2020

**УДК 519.766.4: 519.816: 303.732.4**

**К89**

*Рекомендовано Вченою Радою КПІ ім. Ігоря Сікорського  
(протокол № 4 від 10.03.2020)*

*Рекомендовано Вченою Радою Інституту прикладного системного аналізу  
(протокол № 2 від 24.02.2020)*

**Рецензенти:**

**Яценко В. О.** – доктор технічних наук, професор, Інститут космічних досліджень НАН України та Національного космічного агентства України;

**Положаєнко С. А.** – доктор технічних наук, професор, Одеський національний політехнічний університет.

**К89      Кузнєцова Н.В., Бідюк П.І.**

Теорія і практика аналізу фінансових ризиків: системний підхід: монографія / Кузнєцова Н.В., Бідюк П.І. Київ: Видавництво Ліра-К, 2020. 400 с.

**ISBN 978-617-7910-35-9**

Монографія присвячена розв'язанню актуальних задач математичного моделювання і оцінювання фінансових ризиків, притаманних фінансово-економічній галузі в цілому, підприємствам, що діють в умовах ризиків, пов'язаних з невизначеністю функціонування фінансової системи та зовнішніх впливів, дій конкурентів. У монографії висвітлено і узагальнено аналітичний та практичний досвід авторів з менеджменту ризиків, показано застосування системного підходу до аналізу і мінімізації ризиків. Досліджено сучасні практики та інструментарій розробки моделей оцінювання ризиків на основі методів інтелектуального аналізу даних. На базі регресійного моделювання виконано прогнозування фінансово-економічних процесів та побудову моделей оцінювання їх волатильності. Аналіз ймовірності виникнення ризиків висвітлено на основі скорингових та інтегрованих моделей. Запропоновано динамічний підхід до оцінювання фінансових ризиків на основі моделей теорії виживання та теорії часових рядів. Запропоновано адаптивний підхід до менеджменту ризиків на основі структурно-параметричної адаптації моделей оцінювання ризиків, пов'язаної з появою нових фактів та свідчень. Наведено приклади розроблених авторами інформаційних технологій і інформаційних систем підтримки прийняття рішень.

Рекомендована для аналітиків і фахівців з ризик-менеджменту, а також студентів і аспірантів, які цікавляться задачами математичного моделювання і прогнозування динаміки фінансово-економічних процесів.

**УДК 519.766.4: 519.816: 303.732.4**

**ISBN 978-617-7910-35-9**

© Кузнєцова Н.В., Бідюк П.І., 2020

© Видавництво Ліра-К, 2020

# ЗМІСТ

<b>ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ І СКОРОЧЕНЬ .....</b>	<b>14</b>
<b>ВСТУП.....</b>	<b>15</b>
<b>РОЗДІЛ 1. СУТНІСТЬ ПРОБЛЕМИ АНАЛІЗУ І МІНІМІЗАЦІЇ РИЗИКІВ.....</b>	<b>21</b>
1.1. Системний підхід до аналізу ризиків .....	22
1.1.1. Категорія «ризик»: основні властивості.....	22
1.1.2. Ризики фінансових систем: поняття і означення.....	24
1.1.3. Загальна характеристика процесу аналізу та опрацювання ризику .....	25
1.1.4. Міжнародні та національні стандарти ризик-менеджменту для фінансових систем .....	29
1.1.5. Формалізація рейтингу та рівня ризику .....	33
1.2. Причини виникнення, класифікація і види ризиків .....	36
1.3. Роль невизначеностей у формуванні ризиків різної природи .....	41
1.4. Вплив інформаційних ризиків на виникнення та результати оцінювання фінансових ризиків .....	43
1.5. Особливості задач оцінювання і прогнозування фінансових ризиків .....	46
1.6. Методологічні засади і актуальні напрями розвитку досліджень фінансового ризик-менеджменту .....	47
<b>РОЗДІЛ 2. МЕТОДИ ТА МОДЕЛІ ДОСЛІДЖЕННЯ І ОЦІНЮВАННЯ РИЗИКІВ ФІНАНСОВИХ СИСТЕМ .....</b>	<b>51</b>
2.1. Моделі оцінювання втрат від реалізації ризиків .....	51
2.2. Методи і технології оцінювання ризиків в сучасній економіці ...	61
2.3. Інтелектуальний аналіз даних як основа для дослідження фінансових процесів і систем .....	67
2.3.1. Факторний аналіз фінансових процесів .....	68
2.3.2. Нечітка логіка – інструмент для формалізації невизначеностей у моделюванні.....	69
2.3.3. Нейромережеві моделі .....	70
2.3.4. Мережі Байеса – ймовірнісні моделі причинно-наслідкових зв'язків у задачах аналізу ризиків.....	72

2.3.5. Древа рішень – математичний апарат для автоматичного аналізу фінансових даних.....	74
2.3.6. Методи оптимальної фільтрації.....	76
2.3.7. Регресійні моделі у задачах прогнозування.....	77
2.3.8. Узагальнені лінійні моделі для формалізованого опису розподілу фінансових змінних.....	79

### **РОЗДІЛ 3. МАТЕМАТИЧНІ МОДЕЛІ ОЦІНЮВАННЯ ВОЛАТИЛЬНОСТІ ФІНАНСОВИХ ПРОЦЕСІВ ..... 88**

3.1. Огляд моделей оцінювання волатильності фінансових процесів...	88
3.1.1. Модель авторегресії з умовною гетероскедастичністю (ARUG).....	91
3.1.2. Модель узагальненої авторегресії з умовною гетероскедастичністю (UARUG) .....	96
3.1.3. Експоненціальна узагальнена модель авторегресії з умовною гетероскедастичністю (EUARUG).....	99
3.1.4. Частково інтегрована узагальнена авторегресія з умовною гетероскедастичністю (CIUARUG).....	100
3.1.5. Частково інтегрована експоненціальна узагальнена авторегресія з умовною гетероскедастичністю (CIEUARUG).....	101
3.1.6. Частково інтегрована авторегресія з ковзним середнім (CIARKC).....	101
3.1.7. Моделі стохастичної волатильності (MSB).....	102
3.2. Приклад застосування моделей для прогнозування значень волатильності фінансового часового ряду.....	107

### **РОЗДІЛ 4. НЕВИЗНАЧЕНОСТІ В ЗАДАЧАХ МОДЕЛЮВАННЯ ФІНАНСОВИХ РИЗИКІВ ..... 114**

4.1. Ідентифікація та класифікація невизначеностей.....	115
4.2. Ілюстрація розкриття невизначеностей.....	118
4.2.1. Неповнота даних як реалізація невизначеності .....	118
4.2.2. Існуючі методи відновлення даних .....	122
4.3. Байєсівський підхід до подолання невизначеностей .....	129
4.4. Оптимальний фільтр Калмана.....	139
4.5. Комбінований метод обробки неповних даних при моделюванні фінансових ризиків.....	145

<b>РОЗДІЛ 5. СКОРИНГОВІ ТА ІНТЕГРОВАНІ МОДЕЛІ ДЛЯ</b>	
<b>ОЦІНЮВАННЯ ФІНАНСОВИХ РИЗИКІВ .....</b>	<b>151</b>
5.1. Експертний і скоринговий підходи до оцінювання	
фінансових ризиків .....	151
5.2. Побудова інтегрованих моделей для аналізу ризиків .....	157
5.3. Скорингові карти ризику .....	169
5.3.1. Скорингова карта поведінки .....	171
5.3.2. Методологічні аспекти розробки скорингових карт .....	172
5.3.3. Нейро-нечіткий метод доповнення даних відхиленнями	
заявками при моделюванні кредитних ризиків .....	176
5.3.4. Приклад побудови скорингової карти	
для аплікаційного скорингу .....	180
5.3.5. Скорингова карта як узагальнена міра ризику .....	183
<b>РОЗДІЛ 6. СИСТЕМНА МЕТОДОЛОГІЯ МЕНЕДЖМЕНТУ</b>	
<b>ФІНАНСОВИХ РИЗИКІВ .....</b>	<b>186</b>
6.1. Складові системної методології менеджменту ризиків .....	186
6.1.1. Принципи менеджменту ризиків .....	187
6.1.2. Урахування інформаційного ризику у процесі	
оцінювання фінансового ризику .....	188
6.1.3. Критерії якості у задачах аналізу та оцінювання	
фінансових ризиків .....	190
6.2. Елементи системної методології аналізу, оцінювання	
і прогнозування фінансових ризиків .....	200
6.3. Опрацювання невизначеностей за системною методологією	
аналізу ризиків .....	204
6.4. Статичне оцінювання фінансових ризиків.....	206
6.5. Динамічне оцінювання ризиків.....	208
6.6. Ймовірно-статистичний метод оцінювання ризику	
фінансових втрат .....	211
6.6.1. Приклад оцінювання ринкового ризику на основі	
ймовірно-статистичного методу .....	219
<b>РОЗДІЛ 7. ДИНАМІЧНЕ ОЦІНЮВАННЯ РИЗИКІВ</b>	
<b>НА ОСНОВІ ТЕОРІЇ ВИЖИВАННЯ.....</b>	<b>222</b>
7.1. Теоретичні і прикладні аспекти теорії виживання.....	222
7.1.1. Припущення і означення теорії виживання .....	223
7.1.2. Формалізація задачі динамічного прогнозування часу	
успішності обслуговування банківського кредиту .....	224

7.2. Динамічні моделі прогнозування ризиків .....	227
7.2.1. Модель пропорційних ризиків Кокса .....	227
7.2.2. Узагальнена лінійна модель для прогнозування часу настання ризику.....	231
7.2.3. Розвиток моделі пропорційних ризиків .....	232
7.2.4. Непараметрична модель оцінювання фінансових ризиків..	234
7.3. Оцінки для порівняння динамічних моделей .....	235
7.4. Алгоритми прогнозування часу настання ризику .....	237
7.5. Метод динамічного оцінювання ризиків.....	238
7.6. Приклади динамічного моделювання фінансових ризиків .....	241
7.6.1. Динамічне моделювання кредитних ризиків .....	241
7.6.2. Динамічне моделювання ризиків телекомунікаційної компанії .....	251

## **РОЗДІЛ 8. АДАПТИВНИЙ ПІДХІД ДО ПОБУДОВИ МОДЕЛЕЙ РИЗИКІВ ФІНАНСОВИХ СИСТЕМ .....**

258

8.1. Принцип адаптивного менеджменту ризиків фінансових систем .....	258
8.2. Адаптивний підхід до менеджменту фінансових ризиків .....	259
8.3. Структурно-параметрична адаптація ймовірно-статистичних моделей .....	269
8.3.1. Метод структурно-параметричної адаптації у менеджменті ризиків.....	269
8.3.2. Використання структурно-параметричної адаптації у процесі моделювання кредитних ризиків .....	279

## **РОЗДІЛ 9. ОГЛЯД СУЧАСНИХ ІНСТРУМЕНТІВ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ ДЛЯ ФІНАНСОВИХ СИСТЕМ .....**

282

9.1. Інструментальні засоби аналізу даних і оцінювання ризиків у фінансовій сфері.....	282
9.1.1. Програмні засоби ІАД компанії Microsoft .....	285
9.1.2. Аналітична платформа Tableau.....	286
9.1.3. Qlik – рішення для бізнес-аналітики.....	287
9.1.4. Інтелектуальні рішення для бізнесу компанії IBM .....	288
9.1.5. SAP – аналітика та консолідація даних.....	289
9.1.6. Аналітичні можливості рішень TIBCO .....	290

9.1.7. AgenaRisk – аналітичне рішення з використанням мереж Байєса .....	290
9.1.8. RapidMiner – інструмент машинного навчання.....	291
9.1.9. STATISTICA – засоби для ризик-менеджменту.....	291
9.1.10. Angoss Predictive Analytics – платформа для прогнозової аналітики .....	292
9.1.11. Analytic Solver – інструмент бізнес-аналітики .....	293
9.1.12. Аналітична платформа Alteryx.....	293
9.1.13. OLAP-технології.....	295
9.1.14. Безкоштовні аналітичні засоби .....	296
9.2. Продукти компанії SAS Institute для аналізу даних .....	297
9.2.1. SAS Enterprise Miner – базове аналітичне рішення компанії SAS .....	297
9.2.2. SAS Credit Scoring for Banking .....	299
9.2.3. Хмарна технологія SAS® Viya.....	302
9.3. SPSS (PASW) Statistics – інструментарій для аналізу даних.....	303

## **РОЗДІЛ 10. ПОБУДОВА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ОЦІНЮВАННЯ, МОНІТОРИНГУ ТА МЕНЕДЖМЕНТУ ФІНАНСОВИХ РИЗИКІВ .....**

10.1. Формування вимог до моделюючого інструментарію.....	307
10.2. Розширена інформаційна технологія менеджменту ризиків фінансових систем.....	309
10.2.1. Клієнт-серверна архітектура .....	309
10.2.2. Реалізація інформаційної технології у вигляді мікросервісів .....	311
10.3. Базова архітектура інформаційної системи підтримки прийняття рішень (ІСППР).....	313
10.3.1. Основні рівні ІСППР .....	314
10.3.2. Структура ІСППР для менеджменту фінансових ризиків .....	315
10.4. Приклад інформаційної технології на основі мікросервісів для оцінювання ризиків втрати клієнтів (з інтеграцією в ERP-систему) .....	318
10.5. Інформаційна технологія динамічного оцінювання ризиків відтоку клієнтів.....	327
10.6. Технологія аналізу фінансових даних підприємства .....	330

<b>РОЗДІЛ 11. ПРАКТИЧНІ ЗАДАЧІ ОЦІНЮВАННЯ І ПРОГНОЗУВАННЯ ФІНАНСОВИХ РИЗИКІВ ....</b>	<b>337</b>
11.1. Оцінювання ринкового ризику.....	337
11.2. Аналіз електронних закупівель на платформі ProZorro.....	343
11.3. Прогнозування банкрутства банку .....	354
11.4. Прогнозування часу відтоку абонентів телекомунікаційної компанії.....	361
11.5. Оцінювання актуарних ризиків страхової компанії.....	367
<b>ЗАКЛЮЧЕННЯ:</b>	
<b>РЕЗУЛЬТАТИ І ПЕРСПЕКТИВИ ДОСЛІДЖЕНЬ.....</b>	<b>380</b>
<b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....</b>	<b>383</b>



# CONTENTS

<b>LIST OF ABBREVIATIONS</b> .....	14
<b>INTRODUCTION</b> .....	15
<b>CHAPTER 1. THE ESSENCE OF THE PROBLEM OF RISKS ANALYSIS AND MINIMIZATION</b> .....	21
1.1. The systemic approach to risk analysis .....	22
1.1.1. Category "risk": basic properties.....	22
1.1.2. Risks in financial systems: concepts and definitions.....	24
1.1.3. General characteristics of the risk analysis and risk treatment process .....	25
1.1.4. International and national risk management standards for financial systems .....	29
1.1.5. Formalization of risk rating and level .....	33
1.2. Causes, classification and types of risks.....	36
1.3. The role of uncertainties in the formation of risks of different nature .....	41
1.4. Information risks impact on the occurrence and results of financial risk assessment .....	43
1.5. Features of financial risks assessment and forecasting problems.....	46
1.6. Methodological bases and actual directions of financial risk management research.....	47
<b>CHAPTER 2. METHODS AND MODELS OF FINANCIAL SYSTEMS RISKS ASSESSMENT</b> .....	51
2.1. Models for estimating risks losses.....	51
2.2. Risk assessment methods and techniques in the modern economy ...	61
2.3. Data science as a basis for financial processes and systems research.....	67
2.3.1. Factor analysis of financial processes .....	68
2.3.2. Fuzzy logic as a tool for formalizing uncertainties in modelling .....	69
2.3.3. Neural network models .....	70
2.3.4. Bayesian networks as probabilistic models of causation in risk analysis problems.....	72
2.3.5. Decision trees as a mathematical apparatus for financial data automatic analysis.....	74

2.3.6. Optimal filtering methods .....	76
2.3.7. Regression models in forecasting problems .....	77
2.3.8. Generalized linear models for a formalized description of the financial variables distribution .....	79

**CHAPTER 3. MATHEMATICAL MODELS FOR FINANCIAL PROCESSES VOLATILITY EVALUATION .....** 88

3.1. Models review for estimating the financial processes volatility .....	88
3.1.1. Autoregressive conditional heteroskedasticity model .....	91
3.1.2. Generalized autoregressive conditional heteroskedasticity model .....	96
3.1.3. Exponential generalized autoregressive conditional heteroskedasticity model .....	99
3.1.4. Fractionally integrated generalized autoregressive conditional heteroskedasticity .....	100
3.1.5. Fractionally integrated exponential generalized autoregressive conditional heteroskedasticity model .....	101
3.1.6. Autoregressive fractionally integrated moving average model.....	101
3.1.7. Stochastic volatility models.....	102
3.2. Models implementation for financial time series volatility prediction .....	107

**CHAPTER 4. UNCERTAINTIES IN FINANCIAL RISK SIMULATION TASKS .....** 114

4.1. Uncertainties identification and classification.....	115
4.2. Uncertainty disclosure illustration.....	118
4.2.1. Incomplete data as a realization of uncertainty .....	118
4.2.2. Existing data recovery methods.....	122
4.3. Bayesian approach to overcoming uncertainties .....	129
4.4. Optimal Kalman filter .....	139
4.5. Combined method of incomplete data processing in financial risk modelling .....	145

**CHAPTER 5. SCORING AND INTEGRATED MODELS FOR FINANCIAL RISK ASSESSMENT .....** 151

5.1. Expert and scoring approaches to financial risk assessment .....	151
5.2. Construction of integrated models for risk analysis .....	157

5.3. Scoring risk cards .....	169
5.3.1. Scoring behavioural card.....	171
5.3.2. Methodological aspects of scoring card development.....	172
5.3.3. Neuro-fuzzy method of supplementing data with rejected applications when modelling credit risks.....	176
5.3.4. An example of building a scoring card for application scoring.....	180
5.3.5. Scoring card as a generalized measure of risk.....	183

**CHAPTER 6. THE SYSTEMIC METHODOLOGY  
OF FINANCIAL RISK MANAGEMENT ..... 186**

6.1. Risk management systemic methodology components .....	186
6.1.1. Risk management principles. ....	187
6.1.2. Information risk accounting in financial risk assessment.....	188
6.1.3. Quality criteria in the tasks of financial risks analysis and assessment.....	190
6.2. Elements of system methodology of financial risks analysis, assessment and forecasting.....	200
6.3. Uncertainties elaboration based on the systemic methodology of risk analysis .....	204
6.4. Financial risks static assessment .....	206
6.5. Dynamic risk assessment.....	208
6.6. Probabilistic-statistical method of estimating the risk of financial losses .....	211
6.6.1. An example of market risk assessment based on the probabilistic-statistical method. ....	219

**CHAPTER 7. DYNAMIC RISK ASSESSMENT BASED  
ON SURVIVAL THEORY ..... 222**

7.1. Survival theory: theoretical and applied aspects .....	222
7.1.1. Survival theory: assumptions and definitions.....	223
7.1.2. Dynamic forecasting of bank loan service time: task formalization .....	224
7.2. Dynamic models for risk forecasting .....	227
7.2.1. Cox proportional hazards model .....	227
7.2.2. Generalized linear model for the moment of risk onset forecasting.....	231
7.2.3. Proportional hazards model development .....	232
7.2.4. Non-parametric model of financial risk assessment.....	234

7.3. Criteria for dynamic models comparison .....	235
7.4. Algorithms for predicting the time of risk onset .....	237
7.5. Dynamic risk assessment method.....	238
7.6. Examples of dynamic modelling the financial risks.....	241
7.6.1. Credit risks dynamic modelling.....	241
7.6.2. Dynamic modelling of a telecommunications company risks .....	251

<b>CHAPTER 8. ADAPTIVE APPROACH TO FINANCIAL SYSTEMS RISK MODELS CONSTRUCTION .....</b>	<b>258</b>
8.1. The principle of adaptive risk management in financial systems ....	258
8.2. Adaptive approach to financial risk management .....	259
8.3. Structural-parametric adaptation of probabilistic and statistical models .....	269
8.3.1. Method of structural-parametric adaptation in risk management.....	269
8.3.2. Structural-parametric adaptation in credit risks modelling ....	279

<b>CHAPTER 9. MODERN DATA MINING TECHNIQUES FOR FINANCIAL SYSTEMS.....</b>	<b>282</b>
9.1. Tools for data analysis and risk assessment in the financial sector.....	282
9.1.1. Microsoft data analysis tools .....	285
9.1.2. Tableau analytical platform.....	286
9.1.3. Qlik - business intelligence solution.....	287
9.1.4. IBM intelligent solutions for business.....	288
9.1.5. SAP – data analytics and consolidation.....	289
9.1.6. Analytical capabilities of TIBCO solutions .....	290
9.1.7. AgenaRisk – analytical solution with Bayesian networks.....	290
9.1.8. RapidMiner – machine learning technique.....	291
9.1.9. STATISTICA – tools for risk management .....	291
9.1.10. Angoss Predictive Analytics – platform for forecasting analytics.....	292
9.1.11. Analytic Solver – business intelligence tool .....	293
9.1.12. Analytical platform Alteryx .....	293
9.1.13. OLAP-technologies .....	295
9.1.14. Free analytical tools.....	296
9.2. SAS Institute products for data analysis.....	297
9.2.1. SAS Enterprise Miner – SAS basic analytical solution.....	297
9.2.2. SAS Credit Scoring for Banking .....	299

9.2.3. SAS® Viya cloud technology .....	302
9.3. SPSS (PASW) Statistics – data analysis tools.....	303
<b>CHAPTER 10. INFORMATION TECHNOLOGIES FOR EVALUATION, MONITORING AND FINANCIAL RISK MANAGEMENT .....</b>	<b>307</b>
10.1. Requirements for modeling tools formation.....	307
10.2. Advanced information technology for financial systems risk management .....	309
10.2.1. Client-server architecture .....	309
10.2.2. Information technology implementation in the form of microservices .....	311
10.3. Basic architecture of information decision support system (IDSS) .....	313
10.3.1. The main levels of IDSS .....	314
10.3.2. IDSS structure for financial risk management .....	315
10.4. Example of the information technology based on microservices for the customer churn risk assessing (with integration into ERP-system) .....	318
10.5. Information technology for customer churn dynamic assessment .....	327
10.6. Technology of the enterprise financial data analysis.....	330
<b>CHAPTER 11. PRACTICAL TASKS OF ASSESSING AND FORECASTING FINANCIAL RISKS .....</b>	<b>337</b>
11.1. Market risk assessment.....	337
11.2. Analysis of e-procurement on the ProZorro platform .....	343
11.3. The bank bankruptcy forecasting .....	354
11.4. Forecasting the churn time of telecommunications company subscriber.....	361
11.5. Actuarial risks of the insurance company assessment.....	367
<b>CONCLUSIONS: RESEARCH RESULTS AND PROSPECTS .....</b>	<b>380</b>
<b>LIST OF REFERENCES .....</b>	<b>383</b>

$$\mathbf{K}(k)\mathbf{v}(k) = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \\ K_{31} & K_{32} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

і оцінка вектора стану визначається за рівняннями:

$$\begin{aligned} \hat{x}_1(k) &= \hat{x}_1(k, k-1) + K_{11}v_1 + K_{12}v_2, \\ \hat{x}_2(k) &= \hat{x}_2(k, k-1) + K_{21}v_1 + K_{22}v_2, \\ \hat{x}_3(k) &= \hat{x}_3(k, k-1) + K_{31}v_1 + K_{32}v_2. \end{aligned}$$

Невимірювана третя компонента вектора стану може бути оцінена, якщо елементи матриць  $\mathbf{P}'(k)$ ,  $\mathbf{P}(k)$  і  $\mathbf{K}(k)$  мають ненульові значення.

#### 4.5. Комбінований метод обробки неповних даних при моделюванні фінансових ризиків

Метод реалізується у вигляді послідовного виконання кроків, що реалізують перевірку повноти, інформативності та якості даних. Здійснюється аналіз систематичності втрачених або пропущених даних та їх відновлення за допомогою побудови множини регресійних моделей та обрання кращої з них.

**Крок 1.** Оцінювання неповноти даних в цілому по вибірці для кожної характеристики.

Якщо *критерій оцінювання кількості пропусків*, описаний у п. 6.1.3,  $I_{j(missing)} > p_{cut-off} \%$ , то змінна-характеристика виключається з моделювання і пропущені значення не має сенсу відновлювати.

**Крок 2.** Аналіз змінних та систематичності пропущених значень.

2.1. Для категоріальної змінної виділення пропущених значень в окрему категорію – заповнення пропусків значенням:

$$V_{категор} := "Missing".$$

2.2. Для всіх числових змінних, які містять пропущені значення здійснюємо аналіз їх появи (S-systematic):

$$S_{jnum} = \begin{cases} 1 - \text{для систематичних пропусків, де } I_{j(missing)} \geq p_{cut-off} \%, \\ 0 - \text{для несистематичних пропусків, де } I_{j(missing)} < p_{cut-off} \%. \end{cases}$$

**Крок 3.** Аналіз причин та наслідків появи пропусків.

Будується МБ для встановлення причинно-наслідкових зв'язків між змінними та аналізу наслідків появи пропуску. Цільова (прогнозована) змінна – наслідки появи пропуску (реалізація інформаційного ризику).

3.1. Для  $S_{j_{\text{num}}}$  аналізують причини та наслідки появи пропусків.

$$C_j = \begin{cases} 1 - \text{випадкові}; \\ 2 - \text{критичні}; \\ 3 - \text{катастрофічні}. \end{cases}$$

3.2. Якщо для  $j$ -ї змінної  $S_{j_{\text{num}}} = 0$ ,  $C_j = 1$ , то всі  $i$ -ті пропущені значення замінюють як:

$$v_{ji} = \begin{cases} 0, \\ \text{мода}. \end{cases}, \text{ де } V_{j_{\text{num}}} = \begin{pmatrix} v_{j1} \\ v_{j2} \\ \blacktriangle \\ v_{j4} \end{pmatrix} \quad i\text{-те значення є пропуском.}$$

3.3. Інакше застосовують регресійне рівняння для прогнозування пропущених значень.

**Крок 4.** Побудова та вибір кращої моделі для відновлення пропущених значень.

#### 4.1. Регресійне моделювання

Для відновлення втрачених даних пропонується використати прогнозу модель регресійного типу. Для цього можлива побудова авторегресійної моделі (АР), авторегресії з ковзним середнім, або множини гетероскедастичних моделей (АРУГ, УАРУГ, ЕУАРУГ, тощо).

Будується модель регресійного типу, функція прогнозування на 1, 2 кроки, на  $s$ -кроків вперед. При побудові моделі регресії першого порядку AR(1) [75]:

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k), \quad E[\varepsilon(k)] = 0, \quad (4.13)$$

прогнозування на один крок вперед:

$$y(k+1) = a_0 + a_1 y(k) + \varepsilon(k+1),$$

Якщо коефіцієнти  $a_0, a_1$  відомі, то прогноз втраченого значення знаходять як умовне математичне сподівання:

$$\begin{aligned} \hat{y}(k+1, k) &= E_k[y(k+1)] = E_k[y(k+1) | y(k), y(k-1), \dots, \varepsilon(k), \varepsilon(k-1), \dots)] = \\ &= a_0 + a_1 E_k[y(k)] = a_0 + a_1 y(k), \end{aligned}$$

Використовуючи знову той самий підхід, можна знайти прогноз на два кроки вперед (4.13):

$$y(k+2) = a_0 + a_1 y(k+1) + \varepsilon(k+2),$$

$$\hat{y}(k+2, k) = E_k[y(k+2)] = a_0 + a_1 E_k[y(k+1)] = a_0 + a_1 E_k[a_0 + a_1 y(k)] = a_0 + a_0 a_1 + a_1^2 y(k).$$

За індукцією прогноз на три кроки вперед:

$$\hat{y}(k+3, k) = E_k[y(k+3)] = a_0 + a_0 a_1 + a_0 a_1^2 + a_1^3 y(k).$$

На  $s$ -кроків вперед прогноз можна обчислити за функцією [297]:

$$\hat{y}(k+s, k) = E_s[y(k+s)] = a_0 \left( \sum_{i=0}^{s-1} a_1^i \right) + a_1^s y(k) = a_0 \sum_{i=0}^{s-1} a_1^i + a_1^s y(k). \quad (4.14)$$

Рівняння (4.14) називають функцією прогнозування для довільного числа кроків вперед. Послідовність прогнозів є збіжним процесом, якщо виконується умова:  $|a_1| < 1$ , тобто

$$\lim_{s \rightarrow \infty} E_k[y(k+s)] = \frac{a_0}{1-a_1}, \quad |a_1| < 1, \quad (4.15)$$

де  $a_1$  – знаменник геометричної прогресії в правій частині (4.14). Вираз (4.15) показує, що для будь-якого стаціонарного АР процесу або процесу АРКС послідовність умовних прогнозних оцінок асимптотично збігається  $s \rightarrow \infty$ , до безумовного середнього значення [76].

Виконується розширення функції прогнозування в процесі АР ( $p$ ) у вигляді рекурсії:

$$\hat{y}(k+s, k) = a_0 + \sum_{i=1}^p a_i \hat{y}(k+s-i),$$

де  $\hat{y}(k+s-i) = E_k[y(k+s-i)]$ .

За аналогічним підходом формується множина функцій прогнозування регресійного типу, які застосовують для відновлення пропущених значень, а також оцінюється точність прогнозування.

4.2. Застосування альтернативних методів відновлення пропусків (наприклад, заповнення середнім значенням, нулем або ЕМ-алгоритмом). Оцінка якості прогнозування за критеріями оцінки точності прогнозів (MAPE, MSE, RMSE тощо).

4.3. Обрання кращої моделі за сукупністю критеріїв якості прогнозів та використання її для відновлення втрачених або пропущених значень.

**Крок 5.** Застосування відновлених даних для моделювання.

Відновлені дані включаються у вибірку даних і застосовуються для подальшої розробки моделей оцінювання фінансових ризиків.



Комбінований метод заповнення пропусків надає вищі оцінки точності заповнення пропущених даних та їх подальшого моделювання, оскільки виконується глибинний аналіз причин появи пропусків та використовуються ідеологічно різні методи для відновлення втрачених даних (регресійні, статистичні, ймовірнісні і ймовірнісно-статистичні).

Проілюструємо це прикладом. Якщо відновлення втрачених даних здійснюється за допомогою регресійного моделювання, то це однозначно зсуває вибірку в бік регресійного рівняння, що використовувалось для заповнення пропусків (наприклад, рівняння  $AR(2)$ ). Якщо частка пропущених значень становила 10%, то в подальшому, при застосуванні цієї самої вибірки для моделювання ризиків різної природи, приблизно на 10% збільшується ймовірність обрання в якості кращої моделі регресійного рівняння типу  $AR(2)$ . Для навчальної вибірки ця модель дає кращі результати, оскільки частина даних згенерована за законом регресійного рівняння другого порядку, проте на тестовій вибірці розподіл даних може мати деякі відмінності.

Обмеженням для використання даного методу (як і для більшості методів відновлення) є необхідність достатньо потужної вибірки даних, у якій є можливість вручну видалити пропущені значення, а потім їх відновити для того, щоб перевірити коректність розробленої моделі. Тоді буде можливість встановити і точність прогнозу відновлення, бо будуть наявні початкові значення вхідних змінних. За такого підходу можна будувати множини моделей-кандидаток на тестовому (видаленому вручну та відновленому) наборі, і обирати кращу модель.

Оскільки невідомо, який метод дасть кращий кінцевий результат, бажано використовувати альтернативні процедури і обирати кращу модель для конкретного випадку, використовуючи відомі статистичні критерії якості. Доцільно будувати і застосовувати комбіновані критерії якості, які містять складові для оцінювання адекватності моделі і якості оцінок прогнозів. Це полегшує створення автоматизованих процедур моделювання і прогнозування у системах підтримки прийняття рішень.

Було виконано множину обчислювальних експериментів із застосуванням даного методу, де на четвертому кроці будувалась множина регресійних моделей і застосовувався ЕМ-алгоритм. Отримані результати відновлення штучно видалених даних з високою точністю. Ілюстрація прикладу моделювання на фактичних даних подана нижче.

#### ***Ілюстрація функціонування методу заповнення пропусків.***

Обрано вибірку з 500 значень (ціни акцій). Спочатку модель прогнозування була побудована для всієї вибірки даних. Потім 25 точок

даних були штучно видалені (усунені значення з 301 по 325); потім кілька альтернативних методів були використані для оцінювання цих пропущених значень (результати наведені у таблиці 4.2).

Таблиця 4.2

**Результати відновлення даних та прогнозування за комбінованим методом**

Умови побудови моделей	Найкраща модель	Якість моделі		Якість прогнозу	
		$R^2$	$DW$	$RMSE$	$MAPE$
Повна вхідна вибірка	AR(1) + t3	0,993	2,12	25,92	3,19
Втрачені дані заповнені нулем (реальні пропуски)	AR(1) + t3	0,31	0,10	118,14	6,51
Заміна пропусків середнім значенням вибірки	AR(1) + t3	0,990	1,97	27,97	3,45
Заміна пропусків прогнозом	AR(1) + t3	0,991	2,18	26,78	3,27
Заміна пропусків за EM-алгоритмом	AR(1) + t3	0,992	2,12	25,97	3,21

Для такої вибірки даних найкращою моделлю за критерієм адекватності виявилась авторегресійна модель першого порядку AR(1) з додатковим членом, що описував тренд третього порядку (t3); а кращим методом відновлення пропущених даних виявився EM-алгоритм. Середня абсолютна похибка у відсотках (MAPE) для одного кроку прогнозування становить близько 3,21%, що цілком можна порівняти з 3,19%, обчисленими для повної вибірки даних [147].

Адекватність моделі перевірялась за коефіцієнтом детермінації, статистикою Дарбіна-Уотсона, і виявилась майже однаковою в обох випадках. Найгірший результат був отриманий при заміні відсутніх значень нулями. В цьому випадку MAPE = 6,51%, що в два рази більше порівняно з повною вибіркою, а коефіцієнт детермінації зменшився більш ніж в три рази (з 0,993 до 0,31) [147].

Враховуючи значні обсяги даних, які потрібно обробляти під час аналізу фінансових ризиків, слід намагатися автоматизувати процес аналізу даних, спрямовувати його на поліпшення якості результатів, пов'язаних із заповненням відсутніх значень та обробкою інших видів невизначеностей даних та обчислювальних процедур. Автоматична (автоматизована) процедура аналізу даних передбачає застосування різних методів, що забезпечують не лише відновлення пропущених даних, а й попереднє згладжування, боротьбу з шумами вимірів та збуреннями даних і при цьому забезпечують високоякісні альтернативні

результати обчислень, такі як короткострокове прогнозування та комплексування даних і оцінок прогнозів, генерування відповідних рекомендацій стосовно керування досліджуваними процесами.

\* \* \* \* \*

Поява нових задач з великими обсягами вхідних даних, які не можна розв'язати з використанням існуючих методів, викликає потребу у розробці нових інтегрованих методів та підходів до моделювання та прогнозування. У монографії запропоновано системний підхід до моделювання нелінійних нестационарних фінансових процесів, який ґрунтується на принципах системного аналізу та використанні звичайних лінійних моделей, узагальнених лінійних та ймовірнісних моделей, альтернативних методів оцінювання параметрів із застосуванням модифікованих процедур Монте-Карло для марковських ланцюгів, умовних розподілах випадкових величин, байєсівській регресії та статистичних і динамічних мережах Байєса. Системний підхід до аналізу даних передбачає створення і використання спеціалізованої системи підтримки прийняття рішень, яку буде розглянуто нижче.

таких пропусків можуть бути: (1) випадковий пропуск інформації (наприклад, не є обов'язковими для заповнення клієнтом ІПН через релігійні переконання); (2) клієнт не хоче надати інформацію; (3) клієнт робить помилки в наданні інформації, і є можливість їх встановити і відновити (умисне спотворення інформації). Останні дві причини істотні при аналізі і можуть бути підставою для віднесення цих клієнтів до ненадійних і відмові їм у наданні певних послуг або товарів [6]. Такі випадки повинні включатися до окремої категорії на наступному кроці.

**Крок 4.** Глибинний аналіз причин появи відсутніх значень і визначення найгіршого випадку (причини 2 і 3 на кроці 3) з заміною відсутнього значення найменшим або найбільшим значенням, яке є найгіршим варіантом для змінної.

**Крок 5.** Використання спеціальних інструментів і методів для відновлення відсутніх значень. Для різних інтелектуальних платформ існує досить багато методів заповнення відсутніх даних, їх детально описано у четвертому розділі. Алгоритми, такі як Zet, Zetbraid і т.д., встановлюють пропущене значення, знаходячи найбільш подібні характеристики інших показників у наборі даних, або замінюють середнім значенням для цієї характеристики [7,29].

**Крок 6.** Використання відновлених значень для моделювання.

Для застосування запропонованої методики при розв'язуванні практичних задач у четвертому розділі було розроблено комбінований метод, який дозволяє врахувати описані вище невизначеності та відновити втрачені дані.

## 6.4. Статичне оцінювання фінансових ризиків

Наявні зараз методи і підходи дозволяють оцінювати ризики на певний момент часу і передбачають таку послідовність етапів (рис. 6.5). На першому етапі здійснюється оцінювання фінансового ризику, виявляються притаманні даним фінансовій системі ризики і розробляються моделі їх кількісної оцінки. Оцінювання фінансових ризиків дозволяє визначити можливі втрати при коливаннях параметрів ринку і розмір капіталу, який необхідно резервувати для покриття цих втрат. Розроблено множину економічних та статистичних підходів до оцінювання фінансових ризиків, зокрема різноманітні варіації VaR (Value-at-Risk), методи на основі IRB-підходу, Shortfall, LDA, методи з використанням байєсівського програмування та нечіткої логіки, розглянуті у другому розділі.

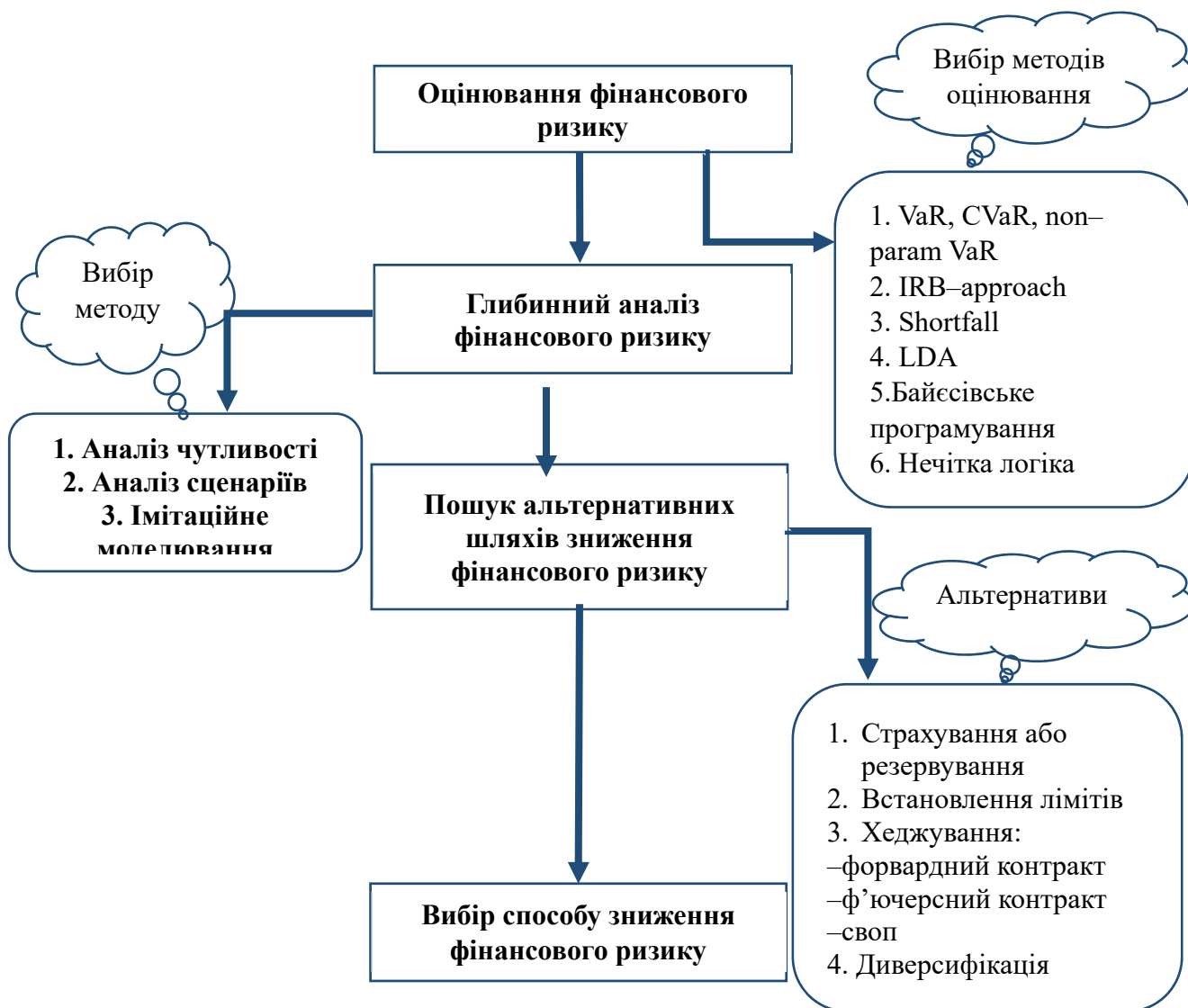


Рис. 6.5. Статичне оцінювання фінансових ризиків

Методи VaR оцінюють ризик як очікуваний максимальний збиток протягом встановленого періоду часу із встановленим рівнем ймовірності; Shortfall є більш консервативним, ніж VaR-метод оцінювання ризику, тобто оцінювання ризику того, що фактичний прибуток інвестиції буде меншим, ніж очікуваний дохід; за методом LDA виконується оцінювання розподілу втрат для розрахунку суми капіталу банку для операційного ризику [74].

Методи байєсівського програмування та нечіткої логіки є більш універсальними і можуть застосовуватись до різних видів фінансового ризику, із встановленням правил та критеріїв, і формуванням висновку щодо рівня ризику (нечітка логіка) [78,92]. Байєсівське програмування включає сукупність таких методів, як мережі Байєса для прогнозування ймовірності ризику, байєсівська регресія для оцінювання рівня ризику і

можливих втрат, гранулярна фільтрація та байєсівський класифікатор для обробки вхідних даних і моделювання багатовимірних розподілів.

На наступному етапі здійснюють глибинний аналіз фінансових ризиків [5], встановлення ключових параметрів, які впливають на бізнес (аналіз чутливості), аналіз різноманітних сценаріїв на основі альтернативних наборів вхідних даних; імітаційне моделювання для оцінки фінансових показників.

На третьому етапі відбувається пошук шляхів зниження ризику: за рахунок страхування або резервування (без зменшення ймовірності появи ризиків, а лише з орієнтацією на відшкодування матеріальних збитків від ризиків); встановленням лімітів операцій; хеджуванням (мінімізацією цінового ризику з метою фіксації певного рівня цін) та диверсифікацією (зменшення сукупної схильності до ризику за рахунок розподілу коштів між різними активами, ціна або прибутковість яких слабо корельовані між собою). Диверсифікація полягає у зниженні значення максимально можливих втрат за одну подію, однак одночасно зростає кількість видів ризику, які необхідно контролювати. Для хеджування використовують біржові (товарні ф'ючерси та опціони) та позабіржові інструменти (форвардні контракти, свопи тощо).

На четвертому етапі приймається управлінське рішення щодо фінансового ризику [5]. На основі всіх проаналізованих альтернатив, методів зниження збитків обирається та альтернатива (або комбінація альтернатив), яка дозволяє отримати найменші фінансові втрати.

## 6.5. Динамічне оцінювання ризиків

В системній методології аналізу ризиків для оцінювання і моделювання ризиків було запропоновано враховувати час, що надасть можливість для їх динамічного оцінювання.

*Принцип динамічного оцінювання ризиків* передбачає, що для оцінювання рівня втрат слід визначити очікувані моменти настання допустимого, критичного і катастрофічного ризиків, ймовірність настання таких ризиків і розмір можливих втрат. Момент часу, коли змінюється рівень ризику від допустимого до катастрофічного, є найважливішим для оцінювання рівня втрат. Ймовірність появи ризиків теж змінюється в часі, і момент, в який ймовірність прояву ризику різко зростає, може бути визначений. На рис. 6.6 подана характеристика основних зон ризику через ймовірність ризику та обсяг втрат.

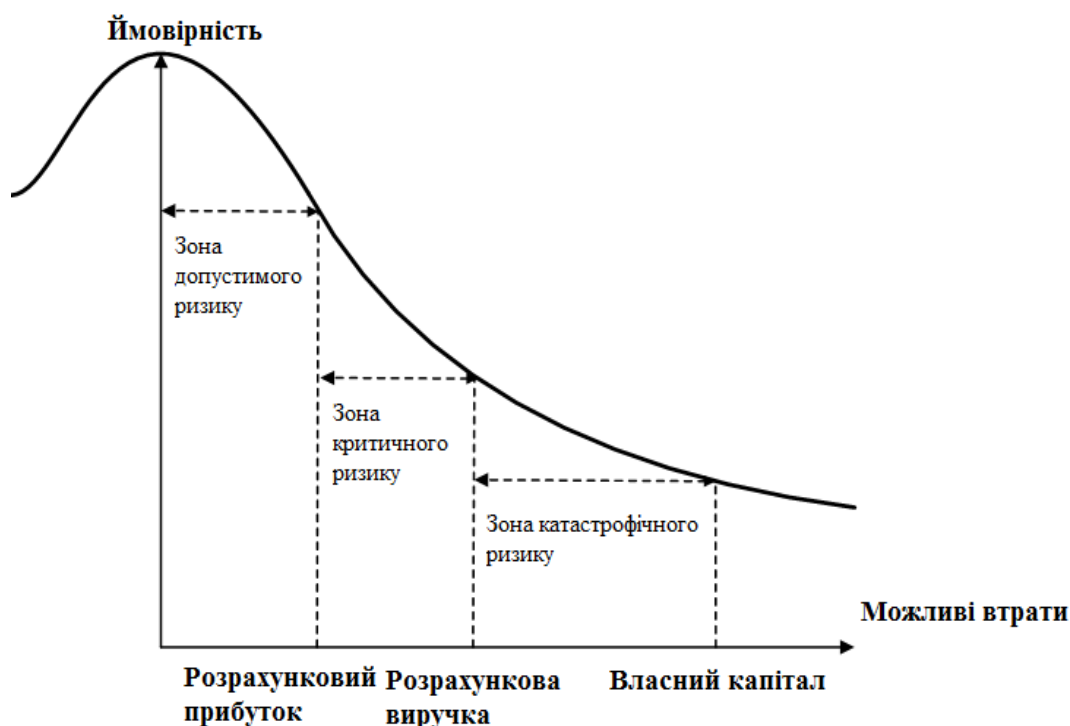


Рис. 6.6. Характеристика основних зон ризику

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

Динамічне оцінювання відрізняється від статичного можливістю оцінювання ризиків у явному вигляді в динаміці, тобто прогнозування функції втрат та ймовірності ризику (переходу на вищий ступінь: критичний, катастрофічний) – як функцій часу.

Стратегія компанії суттєво залежить від її фінансових можливостей та толерантності до ризику, тобто того, який рівень ризику і, відповідно, можливих втрат, вона може собі дозволити прийняти [5]. Формально це множина таких кривих – графіків функції виживання:

$S(t_1|x)$  – умовна функція функціонування фінансової системи на рівні допустимого ризику, в якому вона успішно працює до моменту часу  $t_1$ . В цей момент розмір можливих втрат оцінюється через недоотримання розрахункового прибутку.

$S(t_2 | x)$  – умовна функція функціонування фінансової системи на рівні критичного ризику (в часовому інтервалі від  $t_1$  до  $t_2$ ). Можливі втрати вираховуються з доходу компанії.

$S(t_3 | x)$  – умовна функція функціонування ФС на рівні катастрофічного ризику (в момент часу від  $t_2$  до  $t_3$ ), коли система несе втрати з власного капіталу.

$\lambda(t | x)$  – умовний рівень небезпеки [5] (перехід у зону катастрофічного рівня), що призводить до великого відтоку клієнтів і великих фінансових втрат.

Ключові моменти, які потребували розробки та доопрацювання для статичного та динамічного оцінювання ризику, подані у таблиці 6.2.

Таблиця 6.2

### Статичне та динамічне оцінювання ризиків

Статичне оцінювання $EL = \sum_{i=1}^N P(R_i) \cdot CE_i \cdot LGD_i$	
Проблеми та обмеження	Методи та способи подолання
Неповнота вхідних даних	Адаптація для коротких вибірок, комбінований метод обробки неповних даних, урахування інформаційних ризиків
Неструктурованість вхідних даних	Критерії для формування структури моделі $R^2, \chi^2, IV, WOE$
Недостатня ефективність існуючих методів ІАД	Оцінювання ризиків інтегрованими та комбінованими моделями (нейро-нечіткі методи, дерева рішень, регресійні та байєсові моделі)
Визначення міри ступеня ризику	Скорингові карти
Критерії якості	$GINI, CA, BS$ та власні критерії (квадратичний критерій та критерій ефективності)
Динамічне оцінювання $DE = \langle PR, Losses, t, S(t   x), \lambda(t   x) \rangle$	
Момент настання ризику	Визначення часу з використанням параметричних, напівпараметричних та непараметричних моделей
Визначення ймовірності фінансового ризику	Використання функції ризику (hazard function)
Оцінювання втрат на конкретний момент часу	1) Функції виживання; 2) Ймовірісно-статистичний метод оцінювання втрат



Визначення рівня небезпеки і ключових моментів часу, які характеризують допустимий, критичний та катастрофічний рівень ризику, є задачею системного аналізу, яку необхідно вирішувати в рамках кожного виду ризику незалежно від типу ризику та галузі, в якій він спостерігається [5]. Пропонується підхід, що базується на визначенні втрат компанії як допустимих  $\lambda(t_1 | x) = c_1$ , критичних  $\lambda(t_2 | x) = c_2$  та катастрофічних  $\lambda(t_3 | x) = c_3$ , де  $c_1, c_2, c_3$  – певні константи, які визначаються компанією залежно від її фінансових оборотів, потужностей тощо (наприклад, обсягу власного капіталу).

Для визначення моментів часу  $t_1, t_2, t_3$  запропоновано метод динамічного оцінювання ризиків на основі динамічних моделей виживання та алгоритми визначення моментів часу на основі допустимих втрат та ймовірності, які детально описані у розділі 7.

## **6.6. Ймовірісно-статистичний метод оцінювання ризику фінансових втрат**

Для математичного моделювання ризиків фінансових втрат застосовують різні методи залежно від типу ризику, даних щодо розвитку фінансових процесів та методів їх обробки, наявності необхідних інструментальних засобів для моделювання, оцінювання і прогнозування фінансових процесів та досвіду використання тих чи інших методів і підходів. Широко застосовуються ймовірісно-статистичні методи і моделі, які ґрунтуються на комбінуванні різних підходів [101, 177, 184], що дає можливість використовувати переваги ідеологічно різних методів та отримувати оцінки можливих втрат у формі точкових оцінок та ймовірностей настання ризикових ситуацій.

Статистичні дані потребують попередньої обробки, заповнення пропусків вимірів, нормування, фільтрації даних з метою зменшення впливу похибок вимірів та випадкових зовнішніх збурень, бутстреп-аналізу. Математичні моделі потребують адаптації у процесі надходження нових вимірів і застосування специфічних оптимізаційних процедур для оцінювання їх параметрів [5].

Пропонується побудова комбінованої моделі на основі оптимального фільтра, регресійної моделі і байєсівської мережі для оцінювання ризику можливих фінансових втрат. Структура ймовірісно-статистичного методу оцінювання ризику фінансових втрат подана на рис. 6.7.

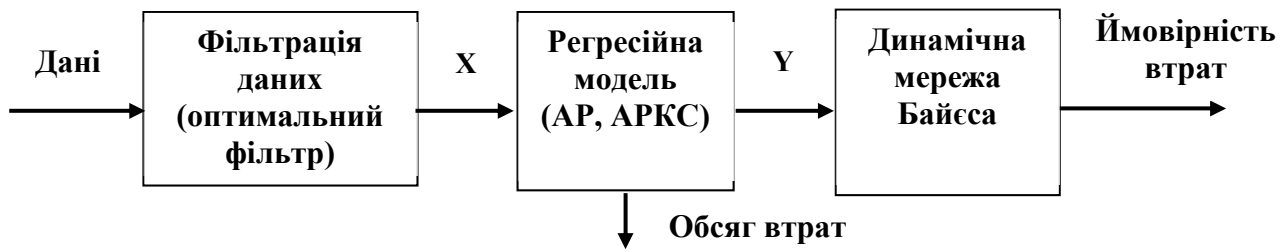


Рис. 6.7. Структурна схема ймовірнісно-статистичного методу оцінювання ризику фінансових втрат

Ймовірнісно-статистичний метод оцінювання ризику фінансових втрат реалізується у вигляді кроків, поданих нижче [185].

**Крок 1.** Вхідні дані, які характеризують можливу появу фінансових втрат, можуть бути різного типу і різних форматів. Статистичні дані характеризуються, як правило, неточністю, неповнотою, нечіткістю інформації, наявністю похибок вимірів та впливом випадкових зовнішніх збурень. Тому виникає необхідність здійснити фільтрацію даних. Для цього можна застосувати цифрові або оптимальні фільтри, наприклад, фільтр Калмана [74]. Застосування цифрових фільтрів передбачає знання смуги частот, у якій знаходяться корисні складові даних, і смуги, де знаходяться згадані шкідливі випадкові впливи, які фільтр не повинен пропускати. На основі цієї інформації проектується цифровий фільтр у вигляді лінійних рівнянь авторегресії (АР) або авторегресії з ковзним середнім (АРКС), які будуть використовуватись для попередньої обробки даних згідно схеми, поданої на рис. 6.7.

Статистичні дані, що описують фінансовий ризик за різними показниками (ліквідності, обслуговування боргів і т. ін.), можуть бути представлені у табличному вигляді, а змінна або вектор змінних стану, що характеризують можливі фінансові втрати в часі, може бути описаний рівнянням динаміки у просторі станів таким чином [185]:

$$\mathbf{x}(k+1) = \mathbf{F} \mathbf{x}(k) + \mathbf{w}(k), \quad (6.9)$$

$$\mathbf{z}(k) = \mathbf{H} \mathbf{x}(k) + \mathbf{v}(k), \quad (6.10)$$

де  $\mathbf{x}(k)$  – вектор стану досліджуваного процесу;  $\mathbf{F}$  – перехідна матриця станів;  $\mathbf{w}(k)$  – векторний процес зовнішніх випадкових збурень з коваріаційною матрицею  $E[\mathbf{w}(k)\mathbf{w}^T(j)] = \mathbf{Q}(k)\delta_{kj}$ ;  $\mathbf{z}(k)$  – вектор вимірів;  $\mathbf{H}$  – матриця коефіцієнтів вимірів;  $\mathbf{v}(k)$  – вектор похибок вимірів з

коваріаційною матрицею  $E[\mathbf{v}(k)\mathbf{v}^T(j)] = \mathbf{R}(k)\delta_{kj}$ . У класичній постановці задачі випадкові процеси  $\mathbf{w}(k)$  і  $\mathbf{v}(k)$  некорельовані між собою та вектором стану і мають нульове середнє та постійні скінченні коваріації [106]. Система матричних рівнянь (6.9), (6.10) – це модель даних у просторі станів, яка для спрощення аналізу не враховує можливих керуючих впливів. Початковим станом системи  $\mathbf{x}_0$  будемо вважати випадкові змінні з відомими статистиками:

$$E[\mathbf{x}_0] = \bar{\mathbf{x}}_0; \quad E[\mathbf{x}_0\mathbf{x}_0^T] = \mathbf{M}; \quad E[\mathbf{w}(k)\mathbf{x}_0^T] = 0, \quad \forall k.$$

Оптимальна оцінка стану  $\hat{\mathbf{x}}(k)$  повинна обчислюватися як найкраща за мінімумом середнього значення суми квадратів похибок оцінок. Тобто, оцінка повинна бути такою, щоб мінімізувати критерій:

$$E[(\hat{\mathbf{x}}(k) - \mathbf{x}(k))^T (\hat{\mathbf{x}}(k) - \mathbf{x}(k))] = \min_{\mathbf{K}}, \quad (6.11)$$

де  $\mathbf{x}(k)$  – точне значення вектора стану, яке може бути обчислене за допомогою детермінованої складової математичної моделі процесу;  $\mathbf{K}$  – оптимальний матричний коефіцієнт фільтра [106], який необхідно обчислити в результаті розв'язання оптимізаційної задачі. Основне рівняння фільтрації має такий вигляд:

$$\hat{\mathbf{x}}(k) = \mathbf{F}(k)\hat{\mathbf{x}}(k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}(k)\mathbf{F}(k)\hat{\mathbf{x}}(k-1)]. \quad (6.12)$$

На основі рівняння (6.9) можна записати функцію прогнозування на один крок як умовне математичне сподівання вектора стану [185]:

$$\hat{\mathbf{x}}(k+1|k) = E_k[\mathbf{x}(k+1)] = \mathbf{F}\hat{\mathbf{x}}(k), \quad (6.13)$$

де  $\hat{\mathbf{x}}(k+1)$  – прогноз на один крок на основі інформації на момент  $k$  включно. Функцією (6.13) можна скористатись для обчислення прогнозу на довільну кількість кроків. Так, прогноз на два кроки має вигляд [185]:

$$\hat{\mathbf{x}}(k+2) = \mathbf{F}\hat{\mathbf{x}}(k+1) = \mathbf{F} \cdot \mathbf{F}\hat{\mathbf{x}}(k) = \mathbf{F}^2\hat{\mathbf{x}}(k),$$

і на довільну кількість кроків  $s$ :

$$\hat{\mathbf{x}}(k+s) = \mathbf{F}^s\hat{\mathbf{x}}(k).$$

Дисперсія похибки прогнозу буде зростати пропорційно кількості кроків  $s$ . Так, похибка прогнозу на один і два кроки складає [76]

$$\begin{aligned} \mathbf{e}_f(1) &= \mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1) = \mathbf{F}\mathbf{x}(k) + \mathbf{w}(k) - \mathbf{F}\hat{\mathbf{x}}(k) = \mathbf{w}(k), \\ \text{і } \mathbf{e}_f(2) &= \mathbf{x}(k+2) - \hat{\mathbf{x}}(k+2) = \mathbf{F}\mathbf{x}(k+1) + \mathbf{w}(k+1) - \mathbf{F}[\mathbf{F}\hat{\mathbf{x}}(k)] = \\ &= \mathbf{F}\mathbf{x}(k+1) + \mathbf{w}(k+1) - \mathbf{F}[\mathbf{x}(k+1) - \mathbf{w}(k)] = \mathbf{w}(k+1) + \mathbf{w}(k). \end{aligned}$$

Дисперсію оцінок прогнозів на  $s$  кроків можна обчислити так:

$$\begin{aligned} \text{Var}[\mathbf{e}_f(s)] &= E\{[\mathbf{w}(k+s-1) + \mathbf{w}(k+s-2) + \dots + \mathbf{w}(k)] \times [\mathbf{w}^T(k+s-1) + \\ &+ \mathbf{w}^T(k+s-2) + \dots + \mathbf{w}^T(k)]\} = s \mathbf{Q}_w. \end{aligned}$$

Оцінювання компонент вектора стану [106], які неможливо виміряти за допомогою приладів, виконується таким чином. У випадку наявності невимірюваних компонент розмірність вектора вимірів  $\mathbf{z}(k)$  є меншою вимірності вектора стану  $\mathbf{x}(k)$ , тобто,  $\dim[\mathbf{z}] < \dim[\mathbf{x}]$ .

Оптимальний коефіцієнт фільтра обчислюється за рівнянням [185]:

$$\mathbf{K}(k) = \mathbf{P}'(k) \mathbf{H}^T [\mathbf{H} \mathbf{P}'(k) \mathbf{H}^T + \mathbf{R}]^{-1},$$

де  $\dim[\mathbf{P}'(k)] = [n \times n]$ ;  $\dim[\mathbf{H}^T] = [n \times r]$  за визначенням;  $n$  – розмірність вектора стану і  $\dim[\mathbf{H} \mathbf{P}'(k) \mathbf{H}^T + \mathbf{R}]^{-1} = [r \times r]$ . Таким чином,  $\dim[\mathbf{K}(k)] = [n \times r]$ . Вектор нев'язок  $\gamma(k)$  в рівнянні оцінювання (6.12)

$$\gamma(k) = \mathbf{z}(k) - \mathbf{H} \mathbf{F} \hat{\mathbf{x}}(k)$$

має розмірність  $[r \times 1]$ , а розмірність добутку  $\dim[\mathbf{K}(k) \gamma(k)] = [n \times 1]$ . Якщо  $\dim[\mathbf{x}] = [3 \times 1]$ , а  $\dim[\mathbf{z}] = [2 \times 1]$ , то добуток  $\mathbf{K}(k) \gamma(k)$  має вигляд:

$$\mathbf{K}(k) \gamma(k) = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \\ K_{31} & K_{32} \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}$$

і оцінка вектора стану визначається за рівняннями:

$$\begin{aligned} \hat{x}_1(k) &= \hat{x}_1(k, k-1) + K_{11} \gamma_1 + K_{12} \gamma_2, \\ \hat{x}_2(k) &= \hat{x}_2(k, k-1) + K_{21} \gamma_1 + K_{22} \gamma_2, \\ \hat{x}_3(k) &= \hat{x}_3(k, k-1) + K_{31} \gamma_1 + K_{32} \gamma_2. \end{aligned}$$

Тобто невимірювана третя компонента вектора стану може бути оцінена, якщо матриці  $\mathbf{P}'(k)$ ,  $\mathbf{P}(k)$  і  $\mathbf{K}(k)$  мають ненульові відповідні елементи [75]. Фільтр дає можливість обчислити оцінки змінних, які є оптимальними у сенсі мінімуму функціоналу (6.13) і знаходяться під значно меншим впливом випадкових збурень стану і похибок вимірів.

**Крок 2.** На наступному кроці будується регресійна модель певної структури, яка здійснює прогнозування показників фінансового ризику втрат (фінансової стійкості) на наступні моменти часу [185]. Як правило, існує можливість побудови деякої множини моделей-кандидаток у

вигляді авторегресії (векторної авторегресії) або авторегресії з ковзним середнім з різними параметрами для авторегресії та ковзного середнього, які будуть здійснювати прогнозування значень фінансових змінних на 1, 2, 3, ... кроки (періоди дискретизації) вперед [76]. Так забезпечується оцінювання прогнозних значень можливих втрат. Для зручності обчислення прогнозів доцільно будувати і використовувати функції прогнозування на основі розв'язків рівнянь [30]. Так, для моделі авторегресії першого порядку функція прогнозування втрат має вигляд:

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

де  $\hat{y}(k+s)$  – оцінка прогнозу основної (цільової) змінної на  $s$  кроків;  $E_k$  – оператор умовного математичного сподівання стосовно моменту часу  $k$ ;  $a_0, a_1$  – параметри регресійної моделі. Для моделі АРКС(2,1) побудована функція прогнозування на три кроки має вигляд:

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

Табулювання функцій прогнозування дозволяє використовувати їх повторно. Похибка прогнозування  $f_k(k)$  для моделей типу АР та АРКС визначається випадковою складовою  $\varepsilon(k)$  і оцінюється так:

$$f_k(s) = \varepsilon(k+s) + a_1 \varepsilon(k+s-1) + a_1^2 \varepsilon(k+s-2) + \dots + a_1^{s-1} \varepsilon(k+1). \quad (6.14)$$

Враховуючи те, що  $E[f_k(s)] = 0$ , оцінка прогнозу, яка обчислюється за виразом (6.14), є незміщеною. Дисперсія похибки прогнозування:

$$Var[f_k(s)] = \sigma^2 [1 + a_1^2 + a_1^4 + a_1^6 + \dots + a_1^{2(s-1)}],$$

тобто дисперсія є функцією кількості кроків прогнозування  $s$ . Асимптотичне значення дисперсії похибки прогнозів для стаціонарного процесу прямує до сталої величини:

$$\lim_{s \rightarrow \infty} Var[f_k(s)] = \frac{\sigma^2}{1 - a_1^2},$$

де  $a_1^2$  – знаменник геометричної прогресії.

При наявності пояснюючих змінних лінійна регресійна модель може мати таку структуру:

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + b_1 x_1(k-d_1) + \dots + a_m x_m(k-d_m) + \varepsilon(k), \quad (6.15)$$

де  $x_i, i=1, \dots, m$  – пояснюючі змінні (предиктори);  $d_i, i=1, \dots, m$  – дискретний час запізнення для пояснюючих змінних [185]. Зазначимо, що незалежні змінні, які входять у модель з фактичним часом запізнення  $d$ , називають провідними індикаторами. Очевидно, що структура моделі (6.15) може містити складові, що описують тренд, сезонні ефекти, ковзне середнє та нелінійні ефекти. Оцінювання структури моделі здійснюється на основі кореляційного аналізу наявних даних, застосування статистичних тестів на інтегрованість, гетероскедастичність, наявність нелінійності та експертних оцінок [75]. Для подальшого аналізу використовують кілька моделей-кандидаток.

Далі аналізується якість моделі, тобто виконується перевірка оцінених кандидаток на адекватність процесу. Діагностика побудованих моделей може здійснюватися наступними способами: візуальне дослідження графіка похибок, аналіз характеристик похибок [76], перевірка, чи корелюють похибки моделі між собою (критерій Дарбіна-Уотсона), чи коефіцієнт множинної детермінації  $R^2$  прямує до 1 ( $R^2 \rightarrow 1$  коли дисперсії вимірів основної змінної та оцінок цієї ж змінної, отриманих за моделлю, збігаються), і чи є мінімальною сума квадратів похибок для вибраної моделі, тобто

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

При використанні методів рекурсивного оцінювання для лінійної моделі 2–3-го порядку оцінки параметрів мають збігатися до усталених значень після 30–40 ітерацій алгоритму оцінювання. Якщо кількість ітерацій набагато перевищує вказані значення, то процес може бути нестационарним та/або нелінійним. Методика побудови регресійних моделей та оцінювання прогнозів описана у роботах [30,186].

В результаті виконання цього кроку отримаємо прогноз величини втрат та рівняння регресійної моделі.

**Крок 3.** На третьому кроці здійснюється ймовірнісне оцінювання втрат на основі адаптивної динамічної мережі Байєса (МБ). Сьогодні МБ – це потужний інструмент ймовірнісно-статистичного моделювання, який ґрунтується на спеціальних чисельних методах, які забезпечують оцінювання альтернативних структур і параметрів моделей, формування точного чи наближеного імовірнісного висновку – остаточного результату застосування цієї моделі.

Адаптація МБ полягає у тому, що структура і параметри моделі оцінюються повторно з появою нових даних. Часовий інтервал для повторного оцінювання визначається емпірично для динаміки досліджуваного процесу і статистичних даних конкретного типу. У комбінованій моделі для оцінювання втрат пропонується застосовувати адаптивну динамічну МБ [108]. Використовуються виходи обраних на попередньому кроці моделей авторегресії (або інших структур), які подаються на вхід стаціонарного шару ймовірнісної моделі.

**Динамічні мережі Байєса** – це розширення статичних байєсівських мереж для моделювання спільних розподілів ймовірностей на множині випадкових змінних  $Z_1, Z_2, \dots$ . Змінні розбиваються на трійки  $Z_t = (U_t, X_t, Y_t)$ , що позначають множину змінних вхідного, прихованого та вихідного шарів моделі у просторі станів [108]. Надалі будемо розглядати моделі стохастичних процесів з дискретним часом, тобто індекс  $t$  буде зростати з появою кожного нового спостереження.

Оскільки динамічна мережа Байєса визначається як пара МБ  $(B_1, B_{\rightarrow})$ , де  $B_1$  – МБ, що визначає апіорну ймовірність  $P(Z_1)$ , а  $B_{\rightarrow}$  – двошарова МБ, що визначає ймовірності  $P(Z_t | Z_{t-1})$  таким чином [185]:

$$P(Z_t | Z_{t-1}) = \prod_{i=1}^N P(Z_t^i | Pa(Z_t^i)),$$

де  $Z_t^i$  –  $i$ -й вузол в момент часу  $t$ , що може бути компонентою  $X_t, Y_t$  або  $U_t$ , а  $Pa(Z_t^i)$  – батьківські вершини  $Z_t^i$  на графі. Вузли першого шару двошарової МБ не мають жодних параметрів, що з ними асоціюються, але кожен вузол другого шару має зв'язаний з ним розподіл умовної ймовірності, що визначають  $P(Z_t^i | Pa(Z_t^i))$  для всіх  $t > 1$ . Батьки вершини  $Pa(Z_t^i)$  можуть бути або ж у тому самому, або у попередньому часовому шарі. Немає суворих математичних обмежень стосовно того, що батьківські вершини знаходяться не далі, ніж в сусідньому шарі, і не можуть знаходитися, скажімо, через один шар. Дуги між шарами спрямовуються зліва направо, що позначає напрям протікання часу. Якщо існує дуга від  $Z_{t-1}^i$  до  $Z_t^i$ , то ця вершина називається сталою [185]. Дуги в межах одного шару є спрямованими, оскільки динамічна мережа є спрямованим ациклічним графом. В межах одного часового шару, як виняток, дозволяється використання неспрямованих дуг, що позначають сильні кореляції між змінними або деякі обмеження.

Для формалізації процесу адаптації моделі формулюється теорема про адаптацію динамічної мережі Байєса (ДМБ), яка полягає у тому, що

нова мережа має кращу якість за рахунок додавання нових і пропущених значень, а також модифікації структури і параметрів нової мережі [185]. Критерієм якості обрано байєсівський інформаційний критерій.

Оскільки на попередньому кроці запропонованого методу було отримано авторегресійну модель (або авторегресію з ковзним середнім), то перший (статичний) шар динамічної мережі Байєса (рис. 6.8) буде складатись зі змінних, що входять в рівняння авторегресії [185]:

$$Z(t) = a_0 + a_1 U_1(t-1) + a_2 U_2(t-2) + \dots + a_p U_p(t-p) + \beta_1 \varepsilon_1(t-1) + \beta_2 \varepsilon_2(t-2) + \dots + \beta_q \varepsilon_q(t-q),$$

де  $p$  – порядок авторегресії, отриманої на другому кроці, а  $q$  – порядок ковзного середнього. Якщо порядок ковзного середнього становить 1, то модель вироджується у звичайну авторегресію порядку  $p$ .

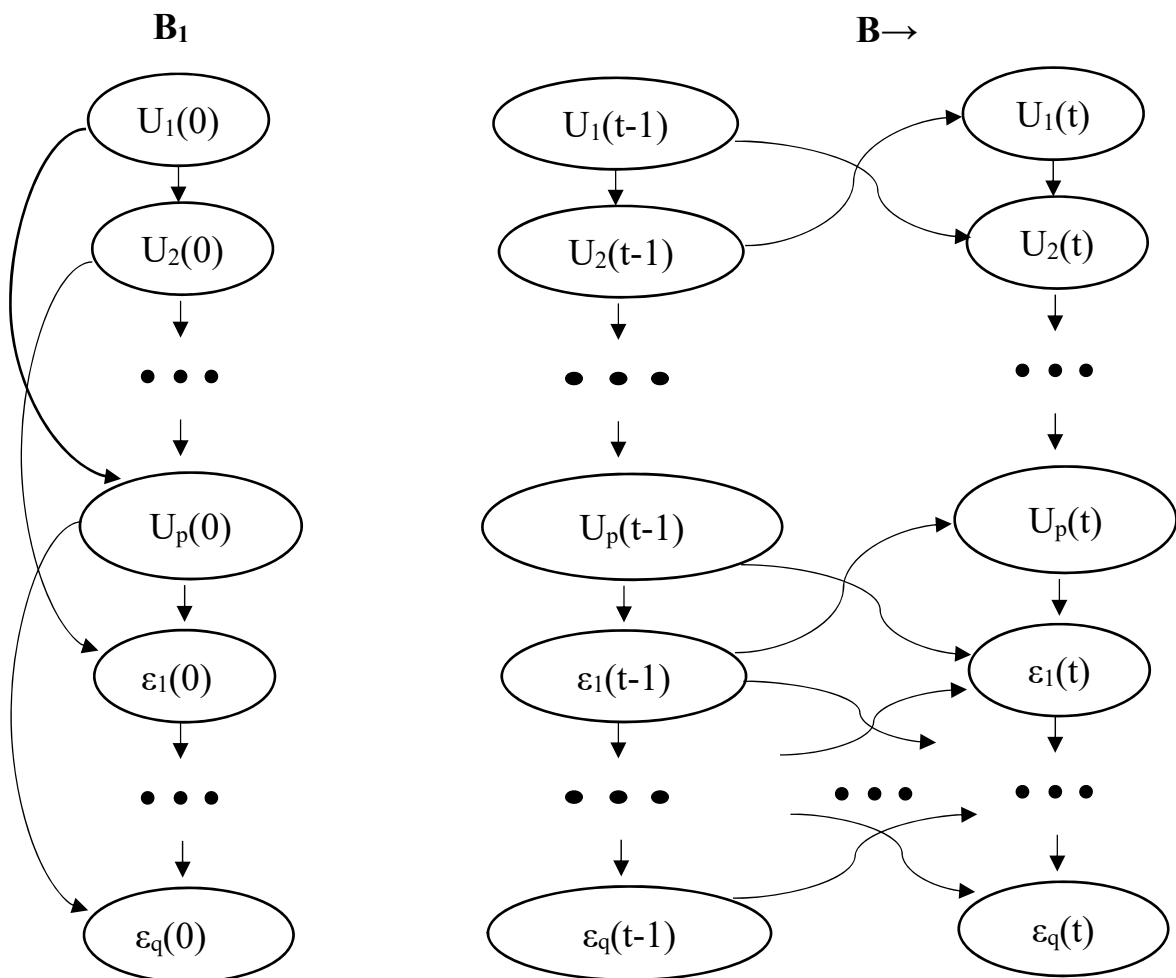


Рис. 6.8. Структура динамічної мережі Байєса



На виході отримаємо ймовірність набуття вихідними змінними значення у часовому просторі [76]. Для оцінювання фінансових ризиків це може бути ймовірність настання певного ризику або ймовірність фінансових втрат в часі (наприклад, через 30 днів). Мережа Байєса може бути використана і для короткострокового прогнозування ймовірності втрат у певному діапазоні та інтервалі значень на момент часу  $t$ . Метод може бути узагальнений на різні типи розподілів. Основною його перевагою є можливість одночасного оцінювання як ймовірності ризику (завдяки використанню МБ) [96], так і обсягів втрат (регресійними моделями). Є можливість його застосування для статичного оцінювання ризиків, використовуючи замість ДМБ дискретну або гібридну МБ [185], і прогнозуючи втрати та ймовірності настання ризику в статистиці.

### **6.6.1. Приклад оцінювання ринкового ризику на основі ймовірнісно-статистичного методу**

Розглянемо задачу оцінювання ринкового ризику при виконанні операцій з фінансовими активами. Вхідними даними є ціни акцій компаній Microsoft, Google, Cisco. Часові ряди даних, що характеризують ціни акцій при закритті торгів на фондовій біржі, взяті з історичної бази даних фінансової системи YahooFinance. Послідовність виконання операцій обробки даних, побудови моделей, прогнозування волатильності і оцінювання ризику втрат відбувається за ієрархічною схемою зверху-вниз [185]. На першому кроці здійснювалось моделювання з використанням фільтра Калмана та без нього, далі короткострокове прогнозування волатильності прибутку різними видами регресійних моделей, описаних у розділі 3.

В таблиці 6.3 подані статистичні показники точності короткострокового прогнозування волатильності на навчальній та перевіірочній вибірках для досліджених у розділі 3 моделей АРУГ, УАРУГ, ЕУАРУГ і МСВ. Значення середньої абсолютної похибки (САП) і середньої абсолютної похибки у відсотках (САПП), отримані на навчальній вибірці (НВ) для моделей УАРУГ НВ, ЕУАРУГ НВ, МСВ НВ [76], показують менші значення похибок прогнозування, ніж оцінки прогнозів моделей на перевіірочній вибірці (УАРУГ ПВ, ЕУАРУГ ПВ та МСВ ПВ), що і очікувалось. Модель АРУГ демонструє значну неточність оцінок прогнозів як на навчальній, так і перевіірочній вибірці, що пояснюється недосконалістю її структури [185], яка не відображає реальної взаємодії змінних.

## Прогнозування волатильності прибутку акцій Microsoft

Модель	САП без ФК	САПП без ФК	САПП з ФК
АРУГ НВ	0,000359	9454,4	9188,7
УАРУГ НВ	0,0000791	45,960	36,270
ЕУАРУГ НВ	0,4353	4,9400	3,7530
МСВ НВ	0,64	7,5000	4,9023
АРУГ ПВ	0,00041	5123,5	2494,4
УАРУГ ПВ	0,00013	51,993	28,396
ЕУАРУГ ПВ	0,5053	5,93	4,0710
МСВ ПВ	0,84	10,90	6,9590

На третьому кроці здійснювалось ймовірнісне оцінювання втрат. Для цього була побудована динамічна мережа Байєса (рис. 6.9), яка прогнозувала на наступні кроки ймовірність фінансових втрат [185].

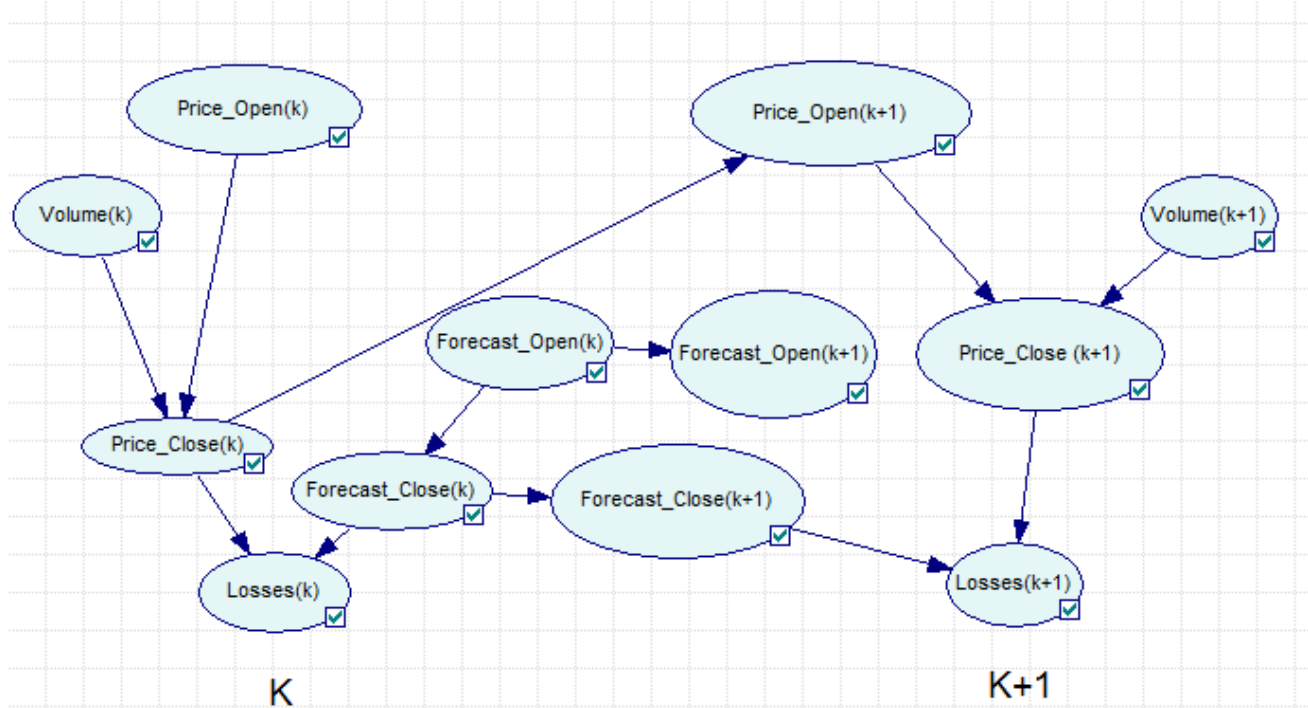


Рис. 6.9. Динамічна мережа Байєса для оцінювання ринкового ризику

Фінансові втрати за акціями в цілому за день оцінювались так:  
 $Losses(k) = (Forecast\_Close(k) - Price(k)\_Close) * Volume(k)$ .

Застосування фільтра Калмана для попередньої фільтрації даних дало можливість зменшити похибки оцінок прогнозів волатильності (без моделі АРУГ) на 7,1 – 45,4%. Подібні результати прогнозування

волатильності отримані для цін акцій компаній Google і Cisco. Оцінки прогнозів волатильності використані для обчислення величини втрат за методикою VaR. VaR дає значення можливих втрат у вибраному довірчому інтервалі, уточнення оцінки ймовірності настання втрат виконано за допомогою мережі Байєса, яка побудована на таких змінних: значення оцінок можливих втрат, обчислені на попередніх трьох інтервалах (періодах дискретизації даних); прогнози значень волатильності і фактичні втрати на попередніх десяти інтервалах.

Ймовірності можливих втрат, отримані за допомогою мережі Байєса, практично повністю співпали з фактичними ймовірностями втрат, обчисленими частотним методом на основі наявних історичних даних. Характерні значення обчислених ймовірностей знаходяться у діапазоні: 0,01 – 0,13. Застосування комплексної моделі дозволило підвищити якість оцінок прогнозів волатильності і уточнити ймовірність настання ризикових ситуацій на використаних статистичних даних.

\* \* \* \* \*

Системна методологія менеджменту ризиків узагальнює процедури статичного і динамічного оцінювання ризиків, застосування методів обробки пропущених та втрачених даних, адаптації моделей ризиків до нових даних і оцінювання їх адекватності. Методологія базується на принципах та процедурах системного аналізу, враховує основні принципи менеджменту ризиків та принципи динамічного оцінювання і адаптивного менеджменту ризиків.

Особливістю статичного оцінювання є розширення застосування скорингової карти до аналізу ризиків різної природи, використання нейро-нечіткого методу урахування відхиленіх раніше даних, критерію урахування інформаційного ризику. Динамічне оцінювання дозволяє прогнозувати рівні та ступені ризиків і ключові моменти часу (докладно описано у сьомому розділі), та оцінювати ризик фінансових втрат. Метод динамічного оцінювання передбачає комбінування різних підходів, використовує переваги ідеологічно різних методів та дозволяє оцінювати можливі втрати у формі точкових оцінок і ймовірностей настання ризикових ситуацій. Створена на засадах системної методології СППР може бути використана для прогнозування ризиків фінансового ринку, аналізу фінансової стабільності підприємств різних типів, кредитних та операційних ризиків.