

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДОНЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАСИЛЯ СТУСА

А. В. Сидорова, Д. В. Біленко, Н. В. Буркіна

Навчально-методичний посібник
«БІЗНЕС-АНАЛІТИКА»

Вінниця
ДонНУ
2019

УДК 005.511 (083.92) (075.8)

С 347

*Рекомендовано до друку Вченою радою
Донецького національного університету імені Василя Стуса
(протокол № 11 від 29.05.2019 р.)*

Авторський колектив: *А. В. Сидорова, д-р екон. наук, професор (розділ 1).
Д. В. Біленко, канд. екон. наук, доцент (розділ 2).
Н. В. Буркіна, канд. пед. наук, доцент (розділ 3).*

Рецензенти: *О. С. Гринькевич, д-р екон. наук, доцент кафедри
статистики ЛНУ імені Івана Франка.
Л. Є. Момотюк, д-р екон. наук, проф., проректор
Національної академії статистики, обліку та аудиту.*

Сидорова А. В., Біленко Д. В., Буркіна Н. В.

С 347 Бізнес-аналітика: навчально-методичний посібник. Вінниця: ДонНУ імені Василя Стуса. 2019. 104 с.

У навчально-методичному посібнику розглянуто теоретико-методологічні засади та прикладні аспекти бізнес-аналітики. Зміст включає статистичні методи економічної аналітики, методи бізнес-аналітики й оцінку ризиків у діяльності суб'єктів господарювання, основи програмування статистичних методів аналізу й моделей в R-середовищі.

Для студентів економічних та управлінських спеціальностей закладів вищої освіти, викладачів, аспірантів, наукових працівників, роботодавців, усіх, хто цікавиться питаннями бізнес-аналітики у різних сферах професійної діяльності.

УДК 005.511 (083.92) (075.8)

© Сидорова А. В., 2019

© Біленко Д. В., 2019

© Буркіна Н. В., 2019

© ДонНУ імені Василя Стуса 2019

ЗМІСТ

ВСТУП	4
Розділ 1. Економічна аналітика	5
1.1. Методологічні засади бізнес-аналітики.....	5
1.2. Багатовимірні угруповання й розрахунок інтегрального показника	7
1.3. Кластерний аналіз	12
1.4. Аналітика часових рядів.....	18
1.5. Аналітика макроекономічних процесів	26
1.5.1. Валовий внутрішній продукт (ВВП) як об'єкт статистичного дослідження та його аналіз.....	26
1.5.2. Напрями використання ВВП у макроекономічному аналізі.....	28
1.6. Прикладні індексні функціональні моделі	33
1.7. Кореляційно-регресійний аналіз. Множинна регресія.....	39
Список використаних джерел до розділу 1	48
Розділ 2. Методологія бізнес-аналітики в ризик-менеджменті	49
2.1. Методи бізнес-аналітики в оцінці ризиків	49
2.2. Інструменти аналізу ризиків	57
2.3. Методи бізнес-аналітики в управлінні ризиками	69
Список використаних джерел до розділу 2	76
Розділ 3. Бізнес-аналітика в середовищі R	77
3.1. Мова програмування R та середовище R Studio.....	77
3.2. Описова статистика засобами R.....	80
3.3. Кореляційно-регресійний аналіз засобами R.....	91
Список використаних джерел до розділу 3	101
Додаток А	102
Додаток Б	103

ВСТУП

Професія аналітика вже має широкий попит серед компаній великого й середнього бізнесу, соціальних установ та урядових структур. Усе більша кількість вакансій на українському ринку праці пропонується саме для аналітиків, тому їхня підготовка дуже перспективна. Крім того, підготовка фахівців з вищою освітою в сучасних умовах усе більше орієнтується не на фахівця в галузі конкретної науки, а на оволодіння компетентностями професії й посади, яку буде обіймати випускник ЗВО.

Аналітика – це наука аналізу, яка застосовується у різних сферах життя – в політиці, бізнесі, фінансах, на ринках товарів і послуг, у страховій справі, у соціальній сфері тощо, на різних рівнях управління: національному, регіональному, на рівні підприємств усіх форм власності. Особлива увага приділяється бізнесу, тобто суб'єктам господарювання, метою діяльності яких є отримання прибутку, й аналізу окремих бізнес-процесів.

Бізнес-аналітика передбачає застосування під час аналізу явищ та процесів широкого кола аналітичних показників, статистичних і математичних методів, комп'ютерних технологій і програмного забезпечення для прийняття управлінських рішень. Враховуючи складні умови невизначеності та ризиків, у яких працюють підприємства, студент повинен оволодіти методологією бізнес-аналітики і методами аналізу та управління ризиками, а також основами програмування. Саме на вирішення цих завдань направлено зміст навчально-методичного посібника.

Навчально-методичний посібник містить три розділи: «Економічна аналітика», «Методи бізнес-аналітики в оцінці ризиків», «Програмування в середовищі R», тобто охоплює програмні питання курсу «Бізнес-аналітика».

За структурою він включає теоретичний матеріал з кожної теми, контрольні питання, задачі для розв'язання в аудиторії та для самостійного виконання, завдання для виконання лабораторних робіт, список літературних джерел, перелік електронних ресурсів.

РОЗДІЛ 1 ЕКОНОМІЧНА АНАЛІТИКА

1.1. МЕТОДОЛОГІЧНІ ЗАСАДИ БІЗНЕС-АНАЛІТИКИ

Бізнес-аналітика – це «наука аналізу» (поглибленого аналізу). Вона передбачає застосування під час аналізу явищ та процесів великих баз даних (*big data*), широкого кола аналітичних показників, статистичних і математичних методів, комп'ютерних технологій і програмного забезпечення для прийняття управлінських рішень.

Аналітика як наука поглибленого аналізу використовується в усіх сферах життя, вона не тільки констатує, яка ситуація склалася, а й обов'язково доводить взаємозв'язок між явищами та процесами, чому склалася така ситуація, дає підстави для обробки великих масивів даних, дає можливість виконувати моделювання та прогнозування досліджуваних явищ і на цій основі розробляти управлінські рішення.

Економічна аналітика – це наука аналізу економічних процесів, які відбуваються в усіх видах діяльності й життя суспільства, на різних рівнях управління: національному, регіональному, на рівні підприємств усіх форм власності (державна, приватна, колективна). Вона використовує широке коло методів для виявлення тенденцій, закономірностей і трансформацій, доведення взаємозв'язку між явищами та процесами, виявляє причини цього, дає підстави для обробки великих масивів даних і прийняття ефективних управлінських рішень.

Бізнес-аналітика є наукою аналізу у бізнес-сфері, тобто вона виконує аналіз завдань, які притаманні бізнесу – підприємствам, організаціям, установам, діяльність яких пов'язана з отриманням прибутку.

Бізнес розглядається як сукупність статистичних одиниць, які здійснюють свою діяльність в умовах ринку, з метою отримання прибутку. У вітчизняній економічній науці та практиці синонімом поняття «бізнес» (англ. *business*) є підприємництво. Згідно з Господарським кодексом України, підприємництво – це самостійна, ініціативна, систематична, на власний ризик господарська діяльність, яку провадять суб'єкти господарювання (підприємці) з метою досягнення економічних і соціальних результатів та отримання прибутку.

Поняття «бізнес» вважають ширшим, ніж підприємництво, оскільки до бізнесу відносять здійснення будь-яких комерційних угод у різних сферах діяльності з метою отримання прибутку.

Типами суб'єктів підприємництва в Україні є підприємство (юридична особа) та фізична особа-підприємець (СПД або ФОП). Розрізняють суб'єктів

малого, середнього й великого бізнесу. Критеріями визначення типу підприємств вважають чисельність персоналу й річний дохід (млн євро).

Юридичні особи:

малого бізнесу – це підприємства чисельністю до 50 осіб включно з річним доходом до 10 млн євро включно. У складі малих виокремлюють мікропідприємства – чисельністю до 10 осіб включно;

середнього бізнесу – від 51 до 250 осіб (включно); річний дохід – більше 10 і до 50 млн євро включно;

великого бізнесу – 251 і більше осіб; річний дохід – більше 50 млн євро.

Фізичних осіб визначають і як юридичних осіб-суб'єктів малого та середнього бізнесу, але великим підприємством фізична особа бути не може.

Бізнес-процес – це сукупність логічних, послідовних і взаємопов'язаних заходів або завдань, спрямованих на створення продукту (послуги) для споживачів. Бізнес-процес визначають також як будь-яку діяльність, що має вхідний продукт, додає вартість до нього та забезпечує вихідний продукт для внутрішнього та зовнішнього споживача (Інтернет-портал для управлінців).

Для цього використовується широке коло статистичних даних у вигляді показників описової статистики для оцінки стану явищ та процесів, розробка методології розрахунків статистичних показників, виконання аналізу закономірностей і тенденцій, що склалися, моделювання і прогнозування показників за допомогою інформаційних і телекомунікаційних технологій.

Аналітика вимагає розуміння змісту, знання методів розрахунку, а також взаємозв'язків і співвідношень між показниками прикладної статистики, використання інформаційних технологій для обробки даних і прогнозування зміни стану явищ та процесів. Отже, **бізнес-аналітика** – це наука аналізу, яка вивчає кількісну та якісну сторону бізнес-процесів у різних видах діяльності, застосовуючи репродуктивну, розрахункову й аналітичну функції показників, статистичні та математичні методи аналізу, моделювання й прогнозування, інформаційні технології та програмне забезпечення для обґрунтування ефективних управлінських рішень.

Об'єктом бізнес-аналітики є явища та процеси, пов'язані з підприємницькою діяльністю та її бізнес-процесами.

Предмет бізнес-аналітики – кількісна оцінка та аналіз бізнес-процесів для розробки ефективних управлінських рішень на основі застосування статистичних і математичних методів, інформаційних та телекомунікаційних технологій.

Бізнесу доводиться вирішувати завдання підвищення темпів росту продуктивності праці, оцінки впливу чинників і одночасно зниження витрат підприємства, управління ризиками, залучення і утримання клієнтів, а також оптимізації

операційної діяльності. Ефективне рішення більшості цих завдань дає бізнес-аналітика.

Для вирішення завдань використовуються новітні технології:

- прийняття рішень у режимі реального часу (наявність інформації про клієнта під час видачі кредиту);
- система кредитного скорингу;
- аналітичні платформи для оперативного управління ризиками тощо.

Отже, бізнес-аналіз – це своєрідний набір технік, інструментів і завдань, необхідних для визначення потреб бізнесу, знаходження і розробки шляхів задоволення цих потреб.

Завдання для вивчення теми

I. Дайте коротку відповідь на питання:

1. Бізнес-аналітика – це...
2. Назвіть складові частини бізнес-аналітики.
3. Чим відрізняється економічна аналітика від бізнес-аналітики?
4. Бізнес – це...
5. Визначте поняття «бізнес-процес».
6. Що є об'єктом бізнес-аналітики?
7. Предметом бізнес-аналітики є...
8. Назвіть типи суб'єктів господарювання.
9. Які суб'єкти господарювання належать до малого й середнього бізнесу?
10. Чи може фізична особа-підприємець належати до великого бізнесу?

II. Завдання

Побудуйте логіко-структурну схему взаємозв'язків бізнес-аналітики з дисциплінами, які вже вами вивчено.

1.2. БАГАТОВИМІРНІ УГРУПОВАННЯ Й РОЗРАХУНОК ІНТЕГРАЛЬНОГО ПОКАЗНИКА

Статистична сукупність об'єднує об'єкти або явища, які відрізняються один від одного, що дає змогу поділяти її на групи. Для цього використовуються **методи угруповання**. При типологічному угрупованні формування груп проводиться шляхом послідовного розбиття сукупності спочатку за однією ознакою, потім отримані частини – за іншою і далі, тобто виконується комбінаційне угруповання (зазвичай за 2–3 ознаками) з дотриманням принципу ієрархії груп.

Комбінаційне угруповання дає змогу достатньо просто інтерпретувати результати, що широко використовується у практиці статистичного аналізу (метод розглядається в курсі «Статистика»).

Якщо групи формуються на основі близькості об'єктів за трьома і більше ознаками, угруповання називається **багатовимірним**. Цей метод застосовують для розрахунку інтегральних показників на основі багатовимірної середньої, для виконання кластерного аналізу. Класифікація об'єктів проводиться не послідовно за окремими ознаками, а одночасно за кількома ознаками. Ці ознаки утворюють «простір ознак». Як і комбінаційне угруповання, метод багатовимірного угруповання дає змогу систематизувати дані, тобто виділяти однорідні групи або класифікувати явища, об'єкти, процеси, виконувати рейтингування, на основі якого приймають управлінські рішення.

Метод багатовимірної середньої застосовується тоді, коли треба оцінити складне явище або процес і неможливо використати якийсь один показник. Якщо ознаки інформаційного простору $\{X\}$ виражені різними одиницями виміру, значення яких суттєво відрізняються, то їх згортка (агрегування) вимагає виконання **попередньої стандартизації** (або нормування). Вихідні ознаки $X_i = \{x_1, x_2, \dots, x_m\}$ замінюються стандартизованими (нормованими) значеннями $Z_i = \{z_1, z_2, \dots, z_m\}$, на основі яких розраховується інтегральний показник.

Інтегральна оцінка $\bar{\rho}_j$ визначається як середня арифметична стандартизованих значень ознак – для j -ої одиниці сукупності, і розраховується за формулами:

- середньої арифметичної простої $\bar{\rho}_j = \frac{1}{m} \sum_{i=1}^m Z_{ij}$;
- середньої арифметичної зваженої $\bar{\rho}_i^{zm} = \sum \bar{\rho}_{ik} \cdot W_k$,

де w_i – вага i -ої ознаки.

Розрахунок інтегральної оцінки ($\bar{\rho}_j$) передбачає такі етапи:

- формування ознакового простору;
- вибір способу стандартизації ознак;
- визначення вагових коефіцієнтів;
- розрахунок інтегральних показників.

При формуванні інформаційного простору необхідно забезпечити односпрямованість впливу показників на досліджуваний процес (явище). Для цього показники (або ознаки) треба поділити на стимулятори і дестимулятори. Зв'язок між оцінкою та показником-стимулятором (X_s) має бути прямим, а між оцінкою і показником-дестимулятором (X_{ds}) – зворотним, тобто під час агрегування

дестимулятори перетворюють на стимулятори, наприклад:
 $X_s = 1 - X_{ds}$ або $X_s = \frac{1}{X_{ds}}$.

Вибір вагових коефіцієнтів (W_i) ґрунтується на змістовному аналізі значущості показників (ознак) і визначається експертним шляхом або методом парних порівнянь, аналізу ієрархій та ін. Способи стандартизації базуються на порівнянні фактичних значень показника X_{ij} з певною величиною. Це може бути максимальне (X_{max}), мінімальне (X_{min}), середнє (x) або еталонне (нормативне) значення (x_0) показника.

Особливості використання конкретної методики нормування залежать від характеру вихідних показників. Зокрема, якщо представлені тільки **стимулятори (або тільки дестимулятори)**, то найзручніший спосіб – це варіант 4 табл. 1.1. Якщо існують стандарти, нормативи, порогові значення або інші еталонні значення, то стандартизацію проводять за варіантом 3 (або 4), але замість X_j^{min} , X_j^{max} беруть $X_j^{етал.}$, тобто $Z_{ij} = \frac{X_{ij}}{X_j^{етал.}}$. Інтегральна оцінка змінюється в інтервалі від 0 до 1.

Для **неоднорідних сукупностей** використовуються методи, розраховані на основі відхилень ($X_{ij} - a$), стандартизовані за розмахом варіації або за середнім квадратичним відхиленням (варіанти 1 і 2). При високих значеннях i -ої ознаки z_{ij} прагне до 1, а при низьких – до 0. Така властивість притаманна й інтегральній середній, тобто чим вищий рівень розвитку, тим ближче до 1 буде її значення. Для цього варіанта нормування можна використовувати теоретичний варіаційний розмах, тобто відхилення максимально можливого (теоретичного) від мінімально можливого (теоретичного) значення відповідних показників.

Таблиця 1.1 – Формули для нормування стимуляторів і дестимуляторів

Варіанти нормування	Методика нормування	
	для стимуляторів	для дестимуляторів
1	$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{\sigma_j}$	$Z_{ij} = \frac{\bar{X}_j - X_{ij}}{\sigma_j}$
2	$Z_{ij} = \frac{X_{ij} - X_j^{min}}{X_j^{max} - X_j^{min}}$	$Z_{ij} = \frac{X_j^{max} - X_{ij}}{X_j^{max} - X_j^{min}}$
3	$Z_{ij} = \frac{X_{ij}}{X_j^{max}}$	$Z_{ij} = \frac{X_{ij}^{min}}{X_{ij}}$
4	$Z_{ij} = \frac{X_{ij}}{X_j^{max}}$	$Z_{ij} = 1 - \frac{X_{ij}}{X_j^{max}}$
5	$Z_{ij} = \frac{X_{ij}}{X_j^{max}}$	$Z_{ij} = \frac{X_j^{min}}{X_{ij}}$

Такий підхід дає можливість проводити порівняльний аналіз у просторі й часі, оскільки знаменник формули нормування залишається постійним. Якщо аналіз динаміки не передбачено, або не проводиться зіставлення у просторі, то можна використовувати фактичний розмах варіації, тобто різницю між найбільшим і найменшим значенням i -ої ознаки з досліджуваної сукупності об'єктів.

На основі інтегральних коефіцієнтів виконується рейтингування об'єктів та їх угруповання, наприклад, з виділенням таких груп за рівнем розвитку: низький рівень; нижче середнього; середній; високий; або низький рівень; нижче середнього; середній; вище середнього; високий. У свою чергу, групування дає змогу визначити закономірності, що склалися, а також виявити додаткові зміни, які неможливо було побачити під час простого угруповання. Отримані аналітичні результати є підґрунтям для розробки управлінських рішень.

Завдання для вивчення теми

Дайте коротку відповідь на питання:

1. Багатовимірне угруповання – це...
2. Чим відрізняється багатовимірне угруповання від звичайного?
3. Що таке «простір ознак» для розрахунку багатовимірної середньої?
4. На які дві групи за напрямом впливу поділяють показники?
5. Якщо показники за одиницями виміру або за величиною значно відрізняються, то необхідно...
6. Які способи нормування (стандартизації) показників існують?
7. Якщо інтегральний показник розраховується для одного об'єкта, то за який період необхідно сформулювати систему показників?
8. Якщо інтегральний показник розраховується для сукупності об'єктів, то яким може бути період дослідження?
9. Які способи візуалізації можна використати в цих випадках?
10. Як розрахувати ефект впливу факторів на інтегральний показник?

Задачі для самостійного виконання у програмі Statistica

Задача 1. За наведеними в таблиці 1.2 даними оцініть інвестиційну привабливість країни методом багатовимірної середньої.

Таблиця 1.2 – Показники оцінки інвестиційної привабливості України за 5 років

Показники	1 рік	2 рік	3 рік	4 рік	5 рік
Індекс фізичного обсягу ВВП, у % до попереднього року	107,3	107,9	102,3	85,2	104,2
Валове накопичення (інвестиції), у % до ВВП	24,7	28,2	27,9	17,0	19,3
Сальдо експорту та імпорту товарів та послуг, млн дол. США	-3 068	-8 152	-14 350	-2 022	-3 850
Питома вага інвестицій в основний капітал у загальному обсязі капітальних інвестицій, %	84,1	84,6	85,7	78,7	79,7
Прямі іноземні інвестиції, які надійшли впродовж року, у % до ВВП	4,4	5,6	3,4	3,8	5,7
Дефіцит державного бюджету, у % до ВВП	0,7	1,4	1,3	3,9	5,9
Державний борг до ВВП, %	12,1	9,9	13,8	24,9	29,6

Для кожного року визначте інтегральний показник. Нормування вихідних показників виконайте на основі зіставлення з максимальними значеннями, а також з еталонними: для рівнів валового накопичення (20 %), прямих іноземних інвестицій (6 %), рівня дефіциту держбюджету (2,6 %) і рівня державного боргу (60 %). Візуалізацію зробіть за допомогою лінійного графіка.

Задача 2. За даними регіонів України про розвиток споживчого кредитування (Додаток А) розрахуйте інтегральний показник рівня розвитку споживчого кредитування методом багатовимірної середньої; оберіть спосіб стандартизації показників; визначте рейтинги регіонів; виділіть 5 груп регіонів за рівнем розвитку: низький рівень; нижче середнього; середній; вище середнього; високий. Виконайте угруповання регіонів, побудуйте групову таблицю. Зробіть висновки.

Задача 3. На сайті Держслужби статистики виберіть показники (6–7), які, на ваш погляд, характеризують рівень життя населення регіонів України за останній рік. Розрахуйте інтегральний показник рівня життя для кожного регіону та визначте ранги регіонів. Побудуйте групову таблицю за інтегральним показником рівня життя населення. Обґрунтуйте кількість груп і величину інтервалів. Зробіть висновки.

Задача 4. На сайті Держслужби статистики України виберіть за останні 5 років показники: грошових доходів, витрат населення, середньої зарплати, рівня захворюваності, коефіцієнта злочинності, індекса споживчих цін і

розрахуйте інтегральний показник рівня життя населення методом багатовимірної середньої. Побудуйте графік динаміки цього коефіцієнта.

Задача 5. За даними Додатка Б про розвиток споживчого кредитування методом багатовимірної середньої розрахуйте інтегральний показник рівня розвитку споживчого кредитування. Зробіть угруповання регіонів з 5 груп.

Задача 6. Виконайте багатовимірне угруповання та розрахуйте інтегральний показник методом багатовимірної середньої за темою своєї магістерської роботи.

1.3. КЛАСТЕРНИЙ АНАЛІЗ

Кластерний аналіз – це розбиття заданої вибірки об'єктів (ситуацій) на групи, які називаються кластерами, так, щоб кожен кластер складався зі схожих об'єктів, а об'єкти різних кластерів істотно відрізнялися. Основна мета **кластерного аналізу** – знаходження груп схожих об'єктів у вибірці, що дає змогу об'єднати їх у кластери для прийняття управлінських рішень і розробки спільних стратегій розвитку. **Кластерний аналіз** – це багатовимірна статистична процедура, яка виконує збір даних, що містять інформацію про вибірку об'єктів, і потім упорядковує об'єкти в порівняно однорідні групи.

Найбільш успішно кластерний аналіз як техніка дослідження ринку впроваджується на регіональному та місцевому рівнях. Аналіз заснований на ідентифікації кластерів за допомогою комп'ютерного аналізу великої кількості змінних.

Кластерний підхід у регіональному розвитку європейських країн став важливим механізмом у визначенні темпів промислового та економічного розвитку. Такий підхід визначає стратегію регіонального розвитку територій незалежно від їх розмірів у нових умовах конкуренції та глобалізації. Кластери є важливою передумовою підвищення конкурентоспроможності, продуктивності, зростання підприємств переважно малого та середнього бізнесу. Кластери є універсальним механізмом для забезпечення кращого досвіду до здобуття навичок, побудови інфраструктури знань, маркетингових досліджень, розподілу ризиків внаслідок об'єднання організацій у мережі, розробки спільних стратегій діяльності суб'єктів ринку, побудови корпоративної культури.

Кластерний аналіз можна виконувати за допомогою програми Statistica 10.0. Спочатку виконується деревоподібна кластеризація за регіонами методом Joining (tree clustering). Як метрику зазвичай використовують евклідову відстань (Euclidean distances), а при поєднанні елементів у кластери обирають метод Варда (Ward's method). Цей метод відрізняється від інших використанням дисперсійного аналізу для оцінки відстаней між кластерами. Метод мінімізує суму квадратів відхилень для будь-яких двох кластерів, які можуть бути сформованими на кожному кроці. Цей метод є досить ефективним, однак він зазвичай формує кластери малого розміру. На регіональному рівні це обмеження не є критичним у зв'язку з невеликою кількістю регіонів.

Аналіз складається з кількох етапів:

1 етап. Для аналізу необхідно сформуванати систему показників. Так, для підприємств середнього й малого бізнесу (СМП) регіонів України можна використати показники, представлені в табл. 1.3. Якщо показники значно відрізняються за величиною або за одиницями виміру, то виконується нормування (стандартизація) показників під час проведення кластерного аналізу (методи нормування розглянуто у темі 2).

Таблиця 1.3 – Абсолютні та відносні показники діяльності СМП за 2018 р.

Позначення факторів	Умовне позначення	Назва показників
x_1	Усього ЗП	Загальна чисельність зайнятого населення на всіх підприємствах
x_2	ЗП_СП	Чисельність зайнятих на середніх підприємствах
x_3	Частка СП_Зайн	Частка зайнятих на середніх підприємствах у % до загальної чисельності зайнятих у регіоні
x_4	ЗП_МП	Чисельність зайнятих на малих підприємствах
x_5	Частка МП_Зайн	Частка зайнятих на малих підприємствах у % до загальної чисельності зайнятих у регіоні
x_6	ЗП_МКП	Чисельність зайнятих на мікропідприємствах
x_7	Частка МКП_Зайн	Частка зайнятих на мікропідприємствах у % до чисельності зайнятих у регіоні
x_8	ЧН	Чисельність населення регіону
x_9	ФП_СП	Прибуток середніх підприємств регіону
x_{10}	ФЗ_СП	Збиток середніх підприємств регіону
x_{11}	ФП_МП	Прибуток малих підприємств регіону
x_{12}	ФЗ_МП	Збиток малих підприємств регіону

За показниками Державної служби статистики виокремлено сукупності регіонів зі схожою економічною ситуацією. У цьому аналізі використано 12 критеріїв СМП за регіонами України за 2018 р. Сукупність малих та середніх підприємств розглянуто в розрізі 25 областей, для цієї сукупності регіонів виділено 4 кластери. При формуванні сукупності акцентовано на чисельності зайнятого населення регіону на підприємствах. Деревоподібна кластеризація за регіонами методом Joining (tree clustering) із евклідовою метрикою (Euclidean distances), за правилом Варда (Ward's method) реалізована у вигляді дендрограми на рис. 1.1.

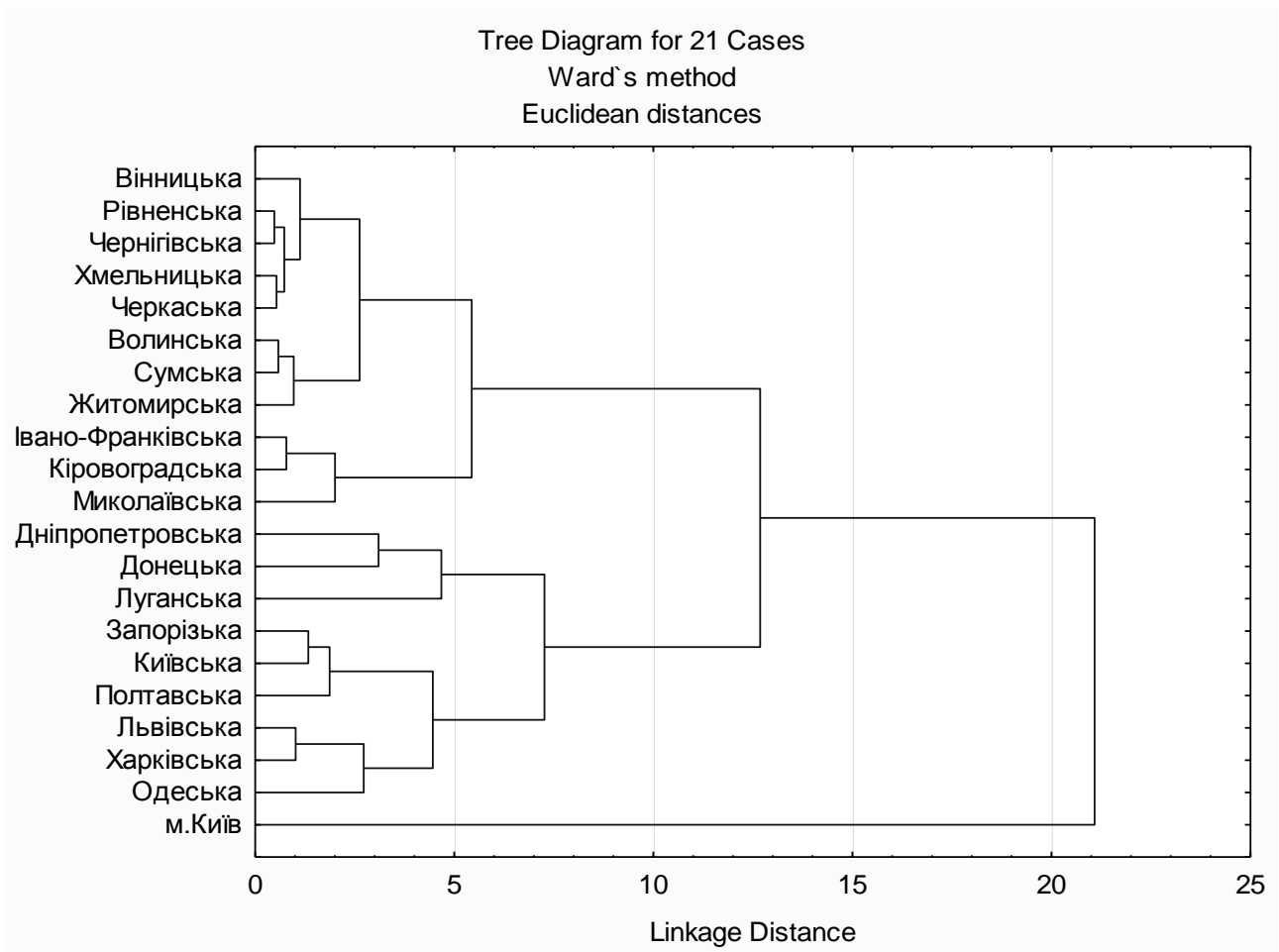


Рисунок 1.1 – Дендрограма кластеризації регіонів за показниками СМП

Кластерний аналіз за методом К-середніх (k-means) із 4 кластерами дав можливість отримати графік, наведений на рис. 1.2, який демонструє відсутність однорідності за обраними показниками. Отже, можна зробити висновок, що запропонована сукупність факторів може бути використана для сегментації регіонів з погляду СМП.

2 етап. Надати характеристику кожному показнику за 4-ма отриманими кластерами. **Позицію кожного регіону за кожним фактором визначити як високу, середню або низьку.** Результати аналізу наведено у табл. 1.4.

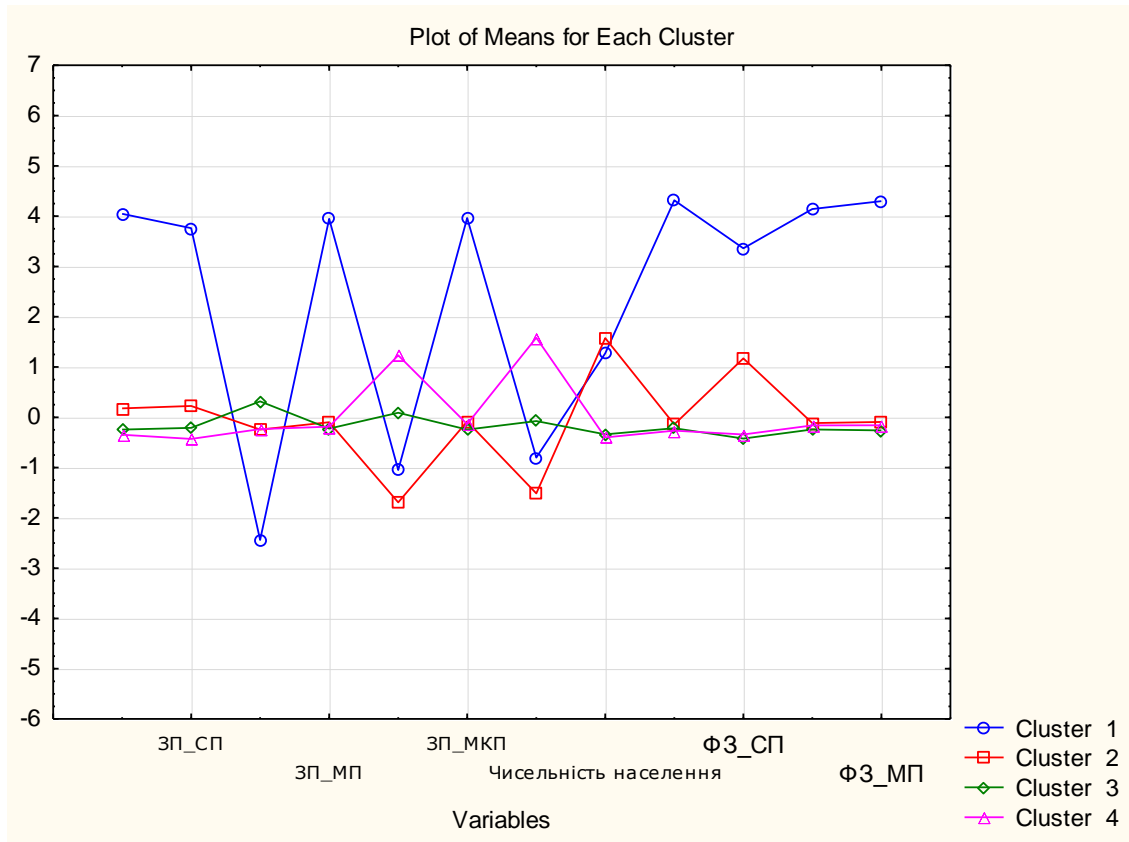


Рисунок 1.2 – Графік розподілу регіонів на 4 кластери за показниками СМП

Таблиця 1.4 – Діагностика показників за кластерами за регіонами України

Показник	Позначення показника	Кластер 1	Кластер 2	Кластер 3	Кластер 4
x_1	Усього ЗП	↑	≈	↓	↓
x_2	ЗП_СП	↑	≈	↓	↓
x_3	Частка СП_Зайн	↓	≈	↑	≈
x_4	ЗП_МП	↑	↓	↓	↓
x_5	Частка МП_Зайн	↓	↓	≈	↑
x_6	ЗП_МКП	↑	↓	↓	↓
x_7	Частка МКП_Зайн	↓	↓	≈	↑
x_8	Чисельність населення	↑	↑	↓	↓
x_9	ФП_СП	↑	↓	↓	↓
x_{10}	ФЗ_СП	↑	≈	↓	↓
x_{11}	ФП_МП	↑	↓	↓	↓
x_{12}	ФЗ_МП	↑	↓	↓	↓

↑ – високе значення показника, ↓ – низьке значення, ≈ – середнє.

Розподіл регіонів за кластерами дає змогу зробити загальну характеристику за кожним кластером (табл. 1.5). Акцент зроблено саме на характерних значеннях, притаманних кожному кластерові.

Таблиця 1.5 – Характеристика кластерів за регіонами й показниками

Кластер 1	Кластер 2	Кластер 3	Кластер 4
1) Високий рівень зайнятого населення на всіх підприємствах	1) Низький рівень зайнятості на малих та мікропідприємствах	1) Низький рівень зайнятого населення на великих підприємствах	1) Низький рівень зайнятого населення на середніх підприємствах
2) Низька частка зайнятих на середніх та малих підприємствах у % до загальної чисельності зайнятих	2) Низький прибуток на середніх та малих підприємствах	2) Найвища частка зайнятих на середніх підприємствах у % до загальної чисельності зайнятих	2) Найвища частка зайнятих на малих та мікропідприємствах у % до загальної чисельності зайнятих
3) Висока чисельність населення	3) Підвищені збитки на середніх підприємствах		
4) Високий рівень і прибутків, і збитків на середніх та малих підприємствах			
м. Київ	Дніпропетровська	Вінницька, Волинська	Івано-Франківська
	Донецька	Житомирська	Кіровоградська
	Луганська	Запорізька, Київська	Миколаївська
		Львівська, Полтавська	Одеська
		Рівненська, Сумська	
		Харківська, Хмельницька	
		Черкаська, Чернігівська	

Отже, розподіл регіонів у кластерах нерівномірний. Вінницький регіон потрапив до найбільш чисельного 3-го кластеру, який характеризується високою часткою зайнятих на середніх підприємствах у % до загальної чисельності зайнятих при відносно низькому рівні зайнятості населення.

Завдання для вивчення теми

Дайте коротку відповідь на питання:

1. Кластерний аналіз – це...
2. З якою метою використовують кластерний аналіз?
3. У яких сферах використовується кластерний аналіз?
4. Як формується система показників для кластерного аналізу?
5. Що означає «дендрограма кластеризації»?
6. Які висновки можна зробити на основі графіка розподілу регіонів на кластери за обраними показниками?
7. Яким показником оцінюють відстань між кластерами?
8. У яких випадках виконується нормування (стандартизація) показників під час проведення кластерного аналізу?
9. Як визначають необхідну або достатню кількість кластерів для аналізу?
10. Які висновки можна зробити за кожним кластером?

Задача 1. Виконайте кластерний аналіз регіонів України за рівнем розвитку споживчого кредитування (дані у Додатку А). Скористайтесь програмою Statistica 10.0.

Для кластеризації регіонів оберіть такі показники:

- X1 – середня заробітна плата 1 штатного працівника, грн;
- X2 – чисельність зайнятих у віці 15–70 років, тис. осіб;
- X3 – процентна ставка за користування споживчим кредитом, %;
- X4 – обсяг залучених депозитів домашніх господарств, млн грн;
- X5 – процентна ставка за депозитами, %;
- X6 – імпорт товарів, млн дол. США;
- X7 – індекс споживчих цін на предмети домашнього вжитку, побутову техніку і поточне утримання житла, %;
- X8 – заборгованість з виплати заробітної плати, млн грн.

Задача 2. На сайті Держслужби статистики України виберіть показники (6–7), які, на ваш погляд, характеризують рівень життя населення регіонів України за останній рік.

Для виконання кластерного аналізу обґрунтуйте, чи треба виконувати стандартизацію цих показників. За допомогою дендрограми доведіть, скільки кластерів треба виділити для аналізу та розробки управлінських рішень щодо підвищення рівня життя в регіонах. Сформулюйте першочергові заходи щодо підвищення рівня життя.

Задача 3. За даними Додатка Б виконайте кластерний аналіз регіонів за рівнем розвитку сфери послуг.

Задача 4. Виконайте кластерний аналіз за темою своєї магістерської роботи.

1.4. АНАЛІТИКА ЧАСОВИХ РЯДІВ

Динаміка – це зміна явищ або процесів у часі. Часовий або динамічний ряд – це послідовність значень статистичного показника, яка характеризує його зміну у часі. Окремі значення показника називають рівнями динамічного ряду. Вони можуть бути виражені абсолютними, відносними і середніми величинами.

Аналіз рядів динаміки передбачає розрахунок абсолютних та відносних аналітичних показників з постійною базою порівняння (базисні) та зі змінною базою порівняння (ланцюгові) (табл. 1.6). Також обчислюються їх середні значення.

Таблиця 1.6 – Аналітичні показники ряду динаміки

Показники	Спосіб розрахунку	
	базисний	ланцюговий
Абсолютні прирости	$\Delta = y_i - y_{\delta}$	$\Delta' = y_i - y_{i-1}$
Темпи росту	$T = \frac{y_i}{y_{\delta}} 100$	$T' = \frac{y_i}{y_{i-1}} 100$
Темпи приросту (зниження)	$T_{np} = \frac{\Delta}{y_{\delta}} 100$ або $T_{np} = T - 100$	$T'_{np} = \frac{\Delta'}{y_{i-1}} 100$ або $T'_{np} = T - 100$
Абсолютні значення 1 % приросту	$A = \frac{\Delta}{T_{np} (\%)} \text{ або } A = \frac{y_{\delta}}{100}$	$A'_i = \frac{\Delta'}{T'_{np} (\%)} \text{ або } A'_i = \frac{y_{i-1}}{100}$

Умовні позначки: y_i – рівень поточного періоду; y_{i-1} – рівень попереднього періоду; y_{δ} – базисний рівень (частіше за все – перший у ряду динаміки: y_1).

Між базисними та ланцюговими показниками динаміки існує взаємозв'язок: сума ланцюгових абсолютних приростів дорівнює базисному за той же відрізок часу:

$\sum_{i=1}^m \Delta'_i = \Delta_i$ або різниця сусідніх базисних дає відповідний ланцюговий:

$\Delta_i - \Delta_{i-1} = \Delta'_i$; добуток послідовних ланцюгових темпів росту дорівнює

базисному за той же відрізок часу: $T_1' \cdot T_2' \cdot \dots \cdot T_m' = T_m$ або ділення двох сусідніх базисних показників дає відповідний ланцюговий темп росту: $T_i : T_{i-1} = T_i'$.

Узагальнюючу характеристику абсолютної швидкості й інтенсивності зміни рівнів ряду динаміки дають середні значення аналітичних показників.

Середній абсолютний приріст: $\bar{\Delta} = \frac{\sum_{i=1}^m \Delta_i'}{m}$ або $\bar{\Delta} = \frac{y_n - y_1}{n-1}$.

Середній темп росту визначається за формулою **середньої геометричної**:
 $\bar{T} = \sqrt[m]{T_1' \cdot T_2' \cdot \dots \cdot T_m'}$ або $\sqrt[n-1]{y_n : y_1}$, де $m = n - 1$.

Середній темп приросту визначається на основі середнього темпу зростання: $\bar{T}_{np} = (\bar{T} - 1) 100$ або $\bar{T}_{np} = \bar{T} \% - 100$.

Середнє абсолютне значення 1 % приросту (\bar{A}): $\bar{A} = \frac{\sum A_i}{n-1}$.

З аналітичною метою оцінки зростання показників у двох рядах за однакові періоди часу розраховують **коефіцієнт випередження (відставання)** як співвідношення базисних (або середньорічних) темпів росту. У чисельнику береться більший за величиною темп:

$$K_{\text{випередж.}} = \frac{T^{(1)}}{T^{(2)}} \quad \text{або} \quad \frac{\bar{T}^{(1)}}{\bar{T}^{(2)}},$$

де $T^{(1)}$, $\bar{T}^{(1)}$ – відповідно базисний або середньорічний темпи росту одного з рядів динаміки;

$T^{(2)}$, $\bar{T}^{(2)}$ – відповідно базисний або середньорічний темпи росту другого ряду динаміки.

Для характеристики швидкості зміни рівнів одного ряду динаміки за різні проміжки часу визначають **коефіцієнт прискорення (уповільнення)**. Він розраховується на основі середньорічних темпів зростання: $K_{\text{прискор./уповільн.}} = \frac{\bar{T}_2}{\bar{T}_1}$,

де \bar{T}_2 і \bar{T}_1 – середні темпи зростання відповідно за наступний і попередній періоди одного ряду динаміки.

Порівняння інтенсивності динаміки двох взаємопов'язаних показників, коли один розглядається як причина (фактор) – x , а другий – як результат (y), здійснюють на основі співвідношення темпів приросту, яке називають **емпіричним**

коефіцієнтом еластичності. Він показує, на скільки відсотків змінюється y зі зміною x на 1 %:

$$K_{ел.} = \frac{\Delta_y}{y_0} : \frac{\Delta_x}{x_0} = \frac{T_{np.(y)}}{T_{np.(x)}}, \text{ де } \Delta_y \text{ та } \Delta_x - \text{ абсолютна зміна результативної та}$$

факторної ознак; y_0 і x_0 – базисні значення результативної й факторної ознак;

$T_{np.(y)}$ і $T_{np.(x)}$ – темпи приросту результативної і факторної ознак.

Прогнозування явищ та процесів на основі трендових моделей

Головною метою побудови трендових моделей для явища, що вивчається, є виявлення тенденції розвитку (зростання або зниження), і на цій основі розробка прогнозу про розвиток досліджуваних процесів на майбутній проміжок часу. Прогнозування на основі часового ряду належить до одновимірних методів прогнозування, що базуються на екстраполяції. У цьому разі зміни досліджуваного показника пов'язують не з чинниками, а з плином часу, що проявляється в утворенні одновимірних часових рядів.

Прогнозування за допомогою трендів – це один із найпростіших методів статистичного прогнозування. Використання методу екстраполяції на основі кривих зростання для прогнозування базується на двох припущеннях:

- часовий ряд економічного показника дійсно має тренд, тобто переважну тенденцію (зростання або зниження);
- загальні умови, що визначали розвиток показника минулого, залишаються без істотних змін протягом періоду випередження.

Під час екстраполяційного прогнозування динаміки на основі часових рядів виконуються такі етапи:

- 1) здійснення попереднього аналізу даних;
- 2) формування набору моделей і вибір кращої з них;
- 3) кількісне оцінювання параметрів моделей;
- 4) визначення адекватності, достовірності й точності моделей;
- 5) отримання точкового та інтервального прогнозів;
- 6) верифікація прогнозу.

Екстраполяцію розуміють як поширення певних закономірностей, зв'язків і висновків про них, що притаманні досліджуваному періоду, за його межі. В широкому сенсі її розглядають як спосіб отримання уявлень про майбутнє на основі інформації про минулий і поточний стан явища або процесу. Під час побудови прогнозних моделей до їх структури іноді закладають елементи майбутнього передбачуваного стану об'єкта чи явища, але загалом ці моделі відображають закономірності, що були в минулому і сьогодні, тому достовірно

прогнозувати можливо лише такі об'єкти і явища, які значною мірою детермінуються минулим і сьогоденням.

Виділяють дві основні форми детермінації: внутрішня і зовнішня. Внутрішня детермінація, або самодетермінація, більш стійка, оскільки її простіше ідентифікувати за допомогою економіко-математичних моделей. Зовнішня детермінація визначається великою кількістю факторів, саме тому врахувати їх усі практично неможливо. У той час, коли деякі методи моделювання, наприклад, адаптивні, відображають загальний сукупний вплив на явище (експорт) зовнішніх факторів, тобто відображають зовнішню детермінацію, методи, що базуються на використанні трендових моделей, представлених одновимірними часовими рядами, відбивають внутрішню детермінацію об'єктів і явищ.

Прогноз на основі трендових моделей обов'язково повинен мати двосторонні межі, тобто визначення інтервалу тих значень, у яких із достатньою імовірністю можна очікувати появу прогнозованих значень. Інтервальний прогноз на базі трендових моделей здійснюється шляхом розрахунку **довірчого інтервалу**, в якому з певною ймовірністю можна чекати появи фактичного значення прогнозованого економічного показника. **Довірчий інтервал** залежить від стандартної помилки оцінки прогнозованого показника, від часу випередження прогнозу, від кількості рівнів у часовому ряді і від рівня значущості прогнозу.

Стандартна (середня квадратична) помилка оцінки визначається за формулою:

$$S_{\hat{y}} = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n - k}},$$

де y_t – фактичне значення рівня часового ряду для часу t ;

\hat{y}_t – розрахункова оцінка показника за моделлю;

n – кількість рівнів у вихідному ряді;

k – число параметрів моделі.

У разі прямолінійного тренду для розрахунку довірчого інтервалу використовується аналогічна формула для парної регресії, отже, довірчий інтервал прогнозу матиме вигляд:

$$U_y = \hat{y}_{n+L} \quad t_\alpha \pm S_{\hat{y}} \sqrt{1 + \frac{1}{n} + \frac{3(n + 2L - 1)^2}{n(n^2 - 1)}},$$

де L – період упередження;

\hat{y}_{n+L} – точковий прогноз за моделлю на $n + L$ момент часу;

n – кількість спостережень у часовому ряді;

S_y – стандартна помилка оцінки прогнозованого показника, розраховується за формулою для числа параметрів моделі, що дорівнює 2;

t_α – табличне значення критерію Стюдента для рівня значущості α і для числа ступенів свободи, що дорівнює $n - 2$.

Найбільш оптимальним вважається період упередження, що дорівнює половині періоду, на основі якого виконувався прогноз, оскільки зі збільшенням часу, для якого складається прогноз, ситуація може змінитись і втрачається точність зроблених припущень.

Під час вирівнювання часових рядів з використанням кривих зростання необхідно **правильно визначити довжину ряду**, на основі якого буде прогнозування. Якщо брати невеликий період, наприклад, 3 роки, то це не дозволить виявити тенденцію його розвитку, а дуже тривалий часовий ряд може охоплювати періоди з різними трендами, і тому опис за допомогою лише однієї кривої зростання не буде змістовним. Періоди, в які відбувалися значні зміни (наприклад, поява нового великого покупця рослинницької продукції, різке зменшення врожаю порівняно з іншими роками, припинення торгівлі з певними країнами або регіонами через політичні негаразди) можуть спотворювати обраний ряд динаміки, а отже, – і вплинути на кінцевий результат побудови моделей та прогнозів на їх основі.

Заключний етап прогнозування – верифікація прогнозу – є вкрай важливим. Верифікація будь-яких трендових моделей зводиться до зіставлення розрахункових результатів за моделлю з відповідними реальними даними – закономірностями динаміки процесів. Верифікація прогновної моделі являє собою сукупність таких критеріїв, способів і процедур, що дають змогу на основі багатостороннього аналізу оцінювати якість отриманого прогнозу, а саме: мінімальний МНК, мінімальну помилку апроксимації, максимальний F-критерій і максимальний коефіцієнт детермінації. Однак частіше на етапі верифікації здійснюється оцінка методу прогнозування, ніж оцінка якості самого результату. Це пояснюється тим, що немає однозначного ефективного підходу до оцінки якості прогнозу до його реалізації.

Про точність прогнозу свідчить величина помилки прогнозу – різниця між фактичним значенням досліджуваного показника і його прогнозним значенням, якщо період упередження вже завершився і відоме фактичне значення прогнозованого показника. Отримані в такий спосіб помилки прогнозу вказують на точність методики прогнозування, що застосовувалась.

Наприклад, виконувалось моделювання і прогнозування експорту рослинницької продукції України до країн ЄС, що є основою для розробки політики регулювання зовнішньої торгівлі країни. Для прогнозування було апробовано декілька трендових моделей. Аналіз використаних моделей прогнозування за формальними критеріями апроксимації показав, що кращою виявилася модель Бокса–Дженкінса. За цією моделлю спочатку розробляється точковий прогноз, визначається довірчий інтервал прогнозу, а потім – межі прогнозу. Із графіка на рис. 1.3 бачимо, що в наступні 5 років обсяг експорту рослинницької продукції Вінницької області не буде знижуватися, а навпаки, приростати щороку приблизно на 0,5 млрд дол., що загалом підтверджує наявні тенденції і закономірності в зовнішній торгівлі України та гіпотезу про те, що у найближчий час не варто очікувати нових кардинальних змін у структурі експорту та втрати аграрним сектором і рослинницькою галуззю позицій найважливіших статей доходу держави.

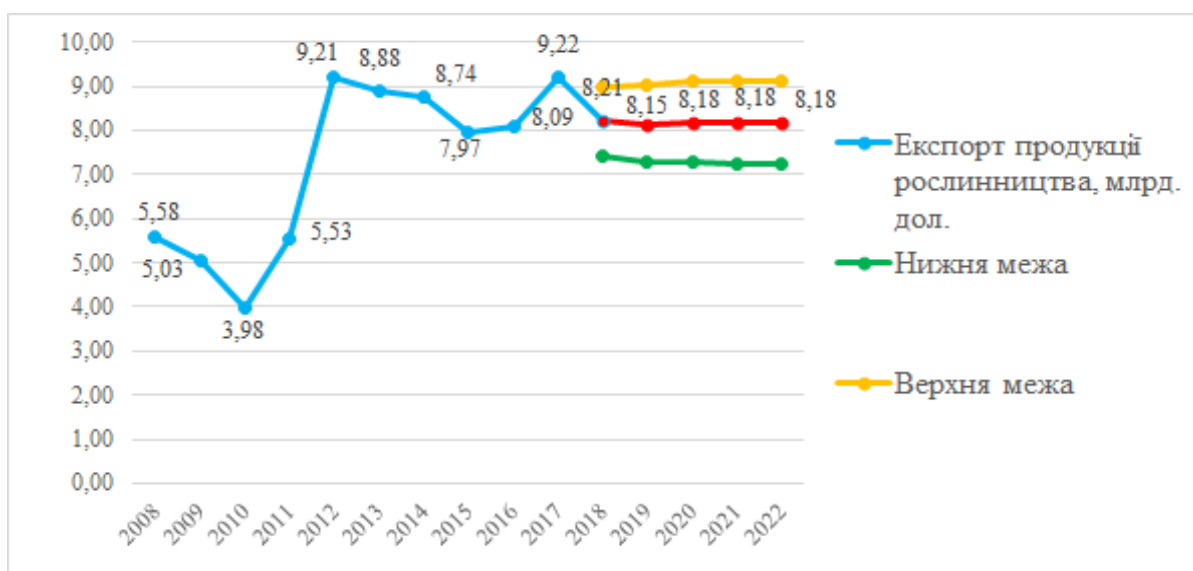


Рисунок 1.3 – Динаміка і прогноз експорту продукції рослинництва Вінницької області на 2018–2022 рр., млрд дол.

Розраховане значення для 2018–2020 рр. може відхилитися на 0,8 млрд дол., для 2021–2022 рр. – на 0,9 млрд дол.

Завдання для вивчення теми

Дайте коротку відповідь на питання:

1. Тенденція розвитку – це...
2. На основі яких показників можна визначити загальну тенденцію?

3. Якщо приблизно рівними в ряді динаміки є ланцюгові абсолютні прирости, то аналітичне вирівнювання доцільно вести за функцією...
4. У степеневій функції параметрами рівняння є...
5. У лінійному рівнянні тренду параметр «*b*» характеризує...

Задача 1. Маємо дані про виробництво харчових продуктів і перевезення вантажів автомобільним транспортом області за 2000–2019 рр. (млн т):

Роки	2000	2005	2010	2015	2018
Обсяг перевезень, млн т	118,2	125,3	170,5	131,7	149,8
Обсяг виробництва, млрд грн	105,9	119,4	150,6	114,5	130,3

Розрахуйте для кожного ряду динаміки:

- 1) середньорічні абсолютні прирости, середні темпи росту й приросту:
 - а) за 2000–2005 рр.;
 - б) за 2005–2010 рр.;
 - в) за 2010–2015 рр.;
 - г) за 2015–2018 рр.;
- 2) коефіцієнти випередження (відставання) середньорічних темпів зростання виробництва порівняно з обсягами перевезень за ті ж самі періоди;
- 3) коефіцієнти прискорення чи уповільнення за 2010–2018 рр. порівняно з 2000–2010 рр.;
- 4) коефіцієнти еластичності перевезень від обсягів виробництва за 2015–2018 рр. порівняно з 2005–2010 рр. (використайте середні темпи приросту).

Задача 2. Динаміка інвестицій в основний капітал підприємств області характеризується такими даними (у млн грн):

2010 р.	2014 р.	2015 р.	2018 р.
96	221	130	216

Визначте за 2010–2014 рр. та за 2015–2018 рр.:

- 1) середні абсолютні прирости;
- 2) середні темпи росту і приросту;
- 3) коефіцієнт прискорення (уповільнення) у 2015–2018 рр. порівняно з 2010–2014 рр.

Зробіть висновки.

Задача 3. Чисельність громадян України, які виїжджали протягом року за кордон, характеризується такими даними (у млн осіб):

2010 р.	2014 р.	2015 р.	2016 р.	2017 р.	2018 р.
16,5	15,5	15,3	17,2	19,8	21,0

Визначте за 2010–2014 та за 2014–2018 рр.:

- 1) середні абсолютні прирости;
- 2) середні темпи зростання;
- 3) середні темпи приросту;
- 4) коефіцієнт прискорення (уповільнення).

Зробіть висновки.

Задача 4. Маємо дані про споживання промисловістю області окремих видів енергетичних матеріалів:

Роки	2010	2015	2016	2017	2018
Нафта, млн т	28,0	28,5	27,4	30,2	31,7
Газ природний, млрд м ³	10,8	10,6	6,3	7,3	8,1

Для порівняльного аналізу обсягів споживання окремих видів енергетичних матеріалів за період 2010–2015 рр. і за 2016–2018 рр. розрахуйте:

- 1) середні темпи зростання споживання;
- 2) споживання якого з енергетичних матеріалів здійснювалось випереджаючими темпами?

Зробіть висновки.

Задача 5. Середній річний темп приросту виробництва тканин склав за 2010–2014 рр. 4,6 %; за 2014–2018 рр. темп зниження (–2,2 %). Визначте базисний і середньорічний темпи зростання виробництва за 2010–2018 рр.

1.5. АНАЛІТИКА МАКРОЕКОНОМІЧНИХ ПРОЦЕСІВ

1.5.1. Валовий внутрішній продукт (ВВП) як об'єкт статистичного дослідження та його аналіз

Валовий внутрішній продукт (ВВП, *Gross Domestic Product*, GDP) – макроекономічний показник, який характеризує ринкову вартість усіх кінцевих товарів і послуг, вироблених за рік усіма галузями економіки на території держави, для споживання, експорту та накопичення, незалежно від національної приналежності робочої сили та використаних факторів виробництва. Він визначається в межах території, на якій його вироблено.

ВВП слугує для оцінки економічного розвитку країни й для виявлення основних пропорцій макроекономіки. Це головний оціночний показник, центральний у системі національних рахунків. Під час його розрахунку використовується уніфікована методика, заснована на уточненій СНР 2010 р., яку використовують усі країни-учасниці програми міжнародних зіставлень ООН.

ВВП – головний показник системи національних рахунків (СНР), який слугує (поряд із валовим національним доходом – ВНД) для оцінки рівня розвитку країни. Для оцінки рівня життя населення використовується відносний показник ВВП на душу населення (в національній валюті). Під час міжнародних порівнянь ВВП країни перераховується в долари США або в паритет купівельної спроможності валют. Для цього ВВП у поточних національних цінах ділиться відповідно на курс долара до гривні або на паритет купівельної спроможності національної валюти. Перерахування здійснюється:

- за середнім валютним курсом: $ВВП_{\$} = \frac{ВВП_{\text{нац вал.}}}{\text{курс}}$,
- і в паритеті купівельної спроможності $ВВП_{\$} = \frac{ВВП_{\text{нац вал.}}}{\text{пкс}}$.

ВВП розраховується як номінальний і реальний:

- Номінальний ВВП визначається у поточних (фактичних) цінах поточного періоду ($\sum q_1 p_1$).

- Реальний ВВП, або фізичний обсяг ВВП ($\sum q_1 p_0$), виражається в цінах попереднього (або будь-якого іншого базисного) року. Це дає змогу визначити ВВП з усуненням інфляції. Тому в реальному ВВП враховується його зростання лише за рахунок реального зростання (фізичного обсягу) виробництва, виключивши зростання цін.

Перерахунок ВВП у співставні ціни попереднього року здійснюється діленням ВВП у поточних цінах на індекс-дефлятор ВВП цього періоду. Індекс-

дефлятор – це своєрідний індекс цін, основний показник, який характеризує зміну цін усіх виробників товарів і послуг в усіх видах економічної діяльності,

$$I_p = \frac{\sum q_1 p_1}{\sum q_1 p_0}$$

тобто враховує індекси цін продукції промисловості, сільського господарства, будівництва, освіти, охорони здоров'я, культури тощо.

Дефлятор може відхилитися від індексу споживчих цін та індексу цін виробників, оскільки він точніше враховує реальну структуру виробництва. Як індекс споживчих цін дефлятор дає змогу розрахувати рівень інфляції (U):

$$U = I_{def} - 1 \text{ або, якщо дефлятор у відсотках, то } U = I_{def} - 100 \%$$

Безперечно, індекс споживчих цін та дефлятор ВВП дають різну характеристику зміни цін. Це пояснюється тим, що між ними є суттєві відмінності:

1) дефлятор ВВП відображає **зміни цін на всі вироблені** товари та надані послуги, а індекс споживчих цін – тільки на реалізовані товари й послуги, що входять до складу споживчого набору товарів (послуг)-представників, тобто тих товарів і послуг, які включено до цього набору;

2) дефлятор ВВП **не відображає зміни цін на імпортні товари**, оскільки імпорт не входить до складу ВВП, але до споживчого набору належать і імпортні товари, тому в індексі споживчих цін відображається зміна цін і на імпортні товари;

3) у дефляторі ВВП **вагою є кількість** вироблених товарів і послуг поточного (звітного) року (індекс Пааше), а в індексі споживчих цін вагою є структура споживчого набору товарів і послуг базисного періоду (модифікований індекс Ласпейреса). Структура споживчого набору переглядається кожні 5 років. Згідно з методикою розрахунку інфляції, у споживчий набір включали 335 найменувань, за новою (2017 р.) методикою – 328, тобто тільки ті товари, що складають не менше 0,1 % вартості споживчого набору.

Перерахунок ВВП у співставні ціни ($\sum q_1 p_0$) якогось базисного року здійснюється діленням ВВП в поточних цінах останнього року на індекс-дефлятор ВВП цього останнього року до базисного, з яким здійснюється порівняння:

$$\text{ВВП}_{\text{реал}} = \frac{\text{ВВП}_{\text{ном}}}{I_{\text{деф}}} = \frac{\sum q_1 p_1}{I_{\text{деф}}}$$

Реальна динаміка ВВП обчислюється за допомогою **індекса фізичного обсягу**. Для його розрахунку треба ВВП поточного року в співставних цінах ($\sum q_1 p_0$) поділити на обсяг ВВП попереднього року ($\sum q_0 p_0$) або базисного року залежно від мети.

1.5.2. Напрями використання ВВП у макроекономічному аналізі

ВВП є базою для розрахунку багатьох важливих показників, що дає змогу використовувати його в макроекономічному аналізі. По відношенню до ВВП обчислюються відносні показники, які відображають найважливіші макроекономічні пропорції. До них належать:

- маса грошей, необхідних для обігу в країні: $M = \frac{ВВП}{n} * 100$.

де n – число оборотів грошової маси.

Ця формула використовується для визначення зміни грошової маси за рахунок ВВП у такий спосіб: $M_{ВВП} = (ВВП_1 - ВВП_0) \cdot \frac{1}{n_0}$;

Вплив другого чинника прискорення (зниження) оборотності грошової маси: $Mn = \left(\frac{1}{n_1} - \frac{1}{n_0}\right) ВВП_1$.

- рівень бюджетного дефіциту (профіциту) країни: $У = \frac{Д-В}{ВВП} * 100$;

де $В, Д$, – державні видатки й доходи.

- рівень державного зовнішнього боргу: $У = \frac{ЗБ}{ВВП} * 100$,

де $ЗБ$ – величина зовнішнього боргу країни (валового, державного чи державного й гарантованого державою зовнішнього боргу);

- експортна квота: $ЕК = \frac{Е}{ВВП} * 100$; ($Е$ – експорт);

- імпортна квота (імпортна залежність): $ІК = \frac{І}{ВВП} * 100$; ($І$ – імпорт);

- інвестиційна відкритість, яка характеризує рівень іноземних інвестицій із нашої країни до ВВП (Π з країни): $ІВ = \frac{\Pi \text{ з країни}}{ВВП} * 100$;

- інвестиційна залежність від іноземного капіталу: $ІЗ = \frac{\Pi \text{ в країну}}{ВВП} \cdot 100$;

- частка податків у ВВП;

• **узагальнюючий показник ефективності економіки** характеризує, скільки грн ВВП припадає на одиницю ресурсів:

$$E = \frac{ВВП}{ОК + С + \Phi} ;$$

де $ОК, С$ – середньорічна вартість основного та оборотного капіталу;

Φ – вартість робочої сили (фонд оплати праці).

Напрями визначення інтенсивності розвитку процесів в окремих видах економічної діяльності на основі ВВП:

- у страхуванні розраховується рівень страхових премій і страхових виплат у % до ВВП;
- у туризмі – рівень доходів від туристичної діяльності до ВВП;
- у дослідженні питань екології – рівень екологічних інвестицій до ВВП;
- продуктивність праці визначається діленням величини ВВП (у порівнянних цінах попереднього року або якогось базисного року) на чисельність зайнятого населення. Перерахунок ВВП у порівнянні ціни здійснюється, щоб уникнути впливу цін на зростання продуктивності. Отже, продуктивність праці

в економіці:
$$W = \frac{ВВП}{T},$$

де W – продуктивність праці;

T – чисельність зайнятих в економіці.

Зміна ВВП під впливом чисельності зайнятих визначається:

$$ВВП_T = ВВП_0 (I_T - 1); \quad \text{або} \quad (T_1 - T_0) \cdot W_0;$$

за рахунок зростання (чи зниження) продуктивності праці:

$$ВВП_W = ВВП_0 \cdot I_T \cdot (I_W - 1); \text{ або } (W_1 - W_0) \cdot T_1,$$

де I_T, I_W – індекс чисельності зайнятих і індекс продуктивності праці.

Отже, макроекономічні аналітичні розрахунки в економіці, у видах економічної діяльності, окремих галузях і сферах базуються на використанні ВВП як головного показника системи національних рахунків.

Приклад

1) перерахувати номінальний ВВП України в ціни попереднього року й у ціни 2010 р.;

2) розрахувати реальну динаміку ВВП за допомогою індексу фізичного обсягу (у цінах попереднього року);

3) розрахувати продуктивність праці, а також темпи її динаміки в економіці України.

Роки	Номінальний ВВП, млрд грн	Індекс-дефлятор, %	ВВП у цінах попереднього року, млрд грн	Індекси фізичного обсягу ВВП (% до попереднього року)	Чисельність зайнятих в економіці, млн осіб	ВВП у % до 1990 р.
А	1	2	3 = 1 : 2	4	5	6
2010	1 083	113,80	951	104,1	18 436,5	65,8
2011	1 317	115,70	1 138	105,1	18 516,2	69,2
2012	1 409	108,04	1 304	99,05	18 736,9	68,5
2013	1 454	103,10	1 410	100,1	18 901,8	68,6
2014	1 567	114,78	1 365	93,8	17 188,1	64,3
2015	1 979	138,40	1 430	91,3	15 742,0	58,7
2016	2 383	117,14	2 034	102,8	15 626,1	60,35
2017	2 983	121,97	2 446	102,6	15 495,9	62,04
2018	3 558,7	115,49	3081	103,3	15 718,6	64,1

Перерахунок номінального ВВП у реальний (у ціни попереднього року):

- 1) реальний ВВП 2010 р. у цінах 2009 р.: $1\,083 : 1,138 = 951$ млрд грн;
- 2) реальний ВВП 2011 р. у цінах 2010 р.: $1\,316 : 1,157 = 1\,138$ млрд грн;

далі аналогічно.

Індекси фізичного обсягу, тобто реальна динаміка ВВП визначається відношенням реального обсягу ВВП (тобто у цінах попереднього року) до номінального ВВП попереднього року.

Наприклад, для 2013 р. $Iq = 1\,410 : 1\,409 = 1,001$ або 100,1 %, тобто за 2013 р. реальний ВВП зріс тільки на 0,1 %.

У 2014 р. $Iq = 1365 : 1454 = 0,938$ або 93,8 %, тобто падіння реального ВВП на 6,2 % (93,8 % – 100). У 2015 р. падіння склало ще 8,7 % (0,913 – 100).

Отже, за 2 роки відбулося падіння ВВП: $0,938 \cdot 0,913 = 0,856$, або 85,6 %, тобто на 14,4 %.

У 2016 р. – незначне зростання на 2,8 %, у 2017 р. – зростання на 2,6 %.

Щоб визначити, як змінився ВВП у 2018 р. порівняно з 2010 р., треба за цей період розрахувати базисний індекс фізичного обсягу ВВП. Враховуючи взаємозв'язок, згідно з яким базисний індекс дорівнює добутку ланцюгових індексів, визначимо, на скільки відсотків зріс ВВП 2018 р. порівняно з 2010 р.

Добуток ланцюгових індексів за 2010–2018 рр:

$1,041 \cdot 1,051 \cdot 0,9905 \cdot 1,001 \cdot 0,938 \cdot 0,913 \cdot 1,028 \cdot 1,026 \cdot 1,033 = 1,012$
або 101,2 %.

Отже, за 2010–2018 рр. реальний ВВП зріс лише на 1,2 %.

Щоб перерахувати ВВП 2010 р. у % до 1990 р., треба додатково визначити базисний індекс фізичного обсягу ВВП за 1990–2010 рр. За даними Державної статистичної служби України, він дорівнює 0,658 або 65,8 %, тобто реальний

ВВП у 2010 р. порівняно з 1990 р. скоротився на 34,2 %. Подальший розрахунок індекса:

$$0,658 \cdot 1,041 \cdot 1,051 \cdot 0,9905 \cdot 1,001 \cdot 0,938 \cdot 0,913 \cdot 1,028 \cdot 1,026 \cdot 1,033 = 0,641 \text{ або } 64,1 \%$$

Отже, у 2018 р. ВВП України не досяг рівня 1990 р., його падіння за цей період склало 35,9 %.

Завдання для вивчення теми

Дайте коротку відповідь на питання:

1. Сформулюйте визначення ВВП.
2. Чим відрізняється ВВП від ВНД?
3. Для яких показників ВВП слугує базою розрахунків?
4. Чим відрізняється індекс-дефлятор від індексу споживчих цін?
5. Як розраховується рівень інфляції?
6. Як розраховують узагальнюючий показник ефективності економіки?
7. Як перераховується номінальний ВВП у ціни попереднього року?
8. Який показник відбиває реальну динаміку ВВП?
9. У чому особливість розрахунку продуктивності праці для економіки країни?
10. Як виконується факторний аналіз впливу чисельності зайнятих і продуктивності праці на зміну виробництва ВВП?

Задача 1. Розрахуйте для кожного року індекс-дефлятор ВВП України і оцініть реальну динаміку ВВП за допомогою індекса фізичного обсягу. Виконайте прогноз індекса фізичного обсягу ВВП на 2019–2020 рр.

Роки	Номінальний ВВП, млрд грн	Індекс-дефлятор, %	ВВП у цінах попереднього року, млрд грн	Індекси фізичного обсягу ВВП (до попереднього року), %
2010	1 083		951	–
2011	1 317		1 138	
2012	1 409		1 304	
2013	1 454		1 410	
2014	1 567		1 365	
2015	1 979		1 430	
2016	2 383		2 034	
2017	2 983		2 446	
2018	3 559		3 081	

Задача 2. За наведеними даними виконайте дії:

- 1) розрахуйте для кожного року число оборотів грошової маси в Україні;
- 2) визначте зміну грошової маси під впливом ВВП і за рахунок прискорення чи уповільнення оборотності у 2018 р. порівняно з 2015 р. і в 2014 р. порівняно з 2010 р.;
- 3) визначте за вказані періоди коефіцієнт прискорення чи уповільнення (у %);
- 4) виконайте прогноз грошової маси на 2019–2020 рр. на основі трендової моделі.

Зробіть висновки.

Роки	Номінальний ВВП, млрд грн	Грошова маса (агрегат М2), млрд грн	Число оборотів грошової маси
2010	1 083	597	
2011	1 317	682	
2012	1 409	771	
2013	1 454	906	
2014	1 567	955	
2015	1 979	994	
2016	2 383	1 102	
2017	2 983	1 103	
2018	3 559	1 208	

Задача 3. За наведеними даними виконайте дії:

- 1) перерахуйте номінальний ВВП України в ціни 2014 р.;
- 2) визначте реальну динаміку ВВП за допомогою індекса фізичного обсягу;
- 3) перерахуйте ВВП 2014–2018 рр. у ціни 1990 р.

Зробіть висновки.

Роки	Номінальний ВВП, млрд грн	Індекс- дефлятор, %	ВВП у цінах 2014 р., млрд грн	Індекси фізичного обсягу ВВП (до попереднього року), %	ВВП у % до 1990 р.
2014	1 567	114,8			—
2015	1 979	138,4			58,7
2016	2 383	117,1			60,4
2017	2 983	122,0			62,0
2018	3 559	115,5			64,1

Задача 4. За наведеними даними виконайте дії:

Роки	ВВП у цінах попереднього року, млрд грн	Середня реальна заробітна плата, грн	Чисельність працюючих, млн осіб	Продуктивність праці, тис. грн/ос.	Темпи росту продуктивності праці, %
2016	2034	5 546	15,63		
2017	2446	6 193	15,50		
2018	3081	8 334	15,72		

- 1) розрахуйте темпи росту реальної заробітної плати в Україні;
- 2) розрахуйте для кожного року продуктивність праці;
- 3) визначте темпи росту продуктивності праці;
- 4) визначте зміну ВВП за рахунок чисельності працюючих і продуктивності праці;
- 5) розрахуйте коефіцієнти випередження темпів росту продуктивності праці над темпами росту заробітної плати;
- 6) розрахуйте коефіцієнти еластичності.

1.6. ПРИКЛАДНІ ІНДЕКСНІ ФУНКЦІОНАЛЬНІ МОДЕЛІ

Статистичне моделювання показників можна здійснювати за допомогою детермінованої (функціональної) залежності результативного показника та факторів, що його формують. Якщо залежність виражена математичною функцією у вигляді суми або добутку факторів, методологію такого аналізу розглядають у рамках теорії статистичних індексів, а сам аналіз називають індексним факторним аналізом. Індексний метод є одним із найбільш потужних, інформативних і поширених інструментів економічного аналізу на всіх його рівнях. Використовуючи індексний метод, можна охарактеризувати зміну різних економічних показників у часі та просторі: обсягів виробленої продукції, виробничих потужностей, цін, вартості і собівартості продукції, продуктивності праці тощо.

В основі побудови сучасних прикладних індексних функціональних моделей принципі моделі Дюпона. Модель Дюпона (англ. *The DuPont System of Analysis*, формула Дюпона) – це метод фінансового аналізу для оцінки впливу ключових факторів, які визначають рентабельність власного капіталу. Цей метод було розроблено й використано компанією «DuPont» у 1920 р. Індексний аналіз залежностей між показниками за допомогою детермінованих моделей є найпоширенішим видом аналізу для суб'єктів господарювання.

Результативним показником моделі DuPont була рентабельність власного капіталу, а факторами – показники виробничої й фінансової діяльності, а саме: операційна рентабельність, ефективність використання активів (оборотність активів) і коефіцієнт капіталізації (співвідношення активів і власного капіталу).

$$ROE = \frac{Пч}{BK} = \frac{Пч}{BP} \cdot \frac{BP}{A} \cdot \frac{A}{BK}.$$

Модель дає змогу визначити, які фактори більшою або меншою мірою впливають на результат. Використовуючи методологію побудови цієї моделі, можна адаптувати її для інших процесів, що досліджуються.

Прикладні моделі будуються розкладанням результативного показника на декілька субфакторів. Якщо результативний показник визначається підсумовуванням факторних ознак, то модель називається адитивною: $y = a + b + c + d$; якщо добутком факторів – мультиплікативною: $y = a \cdot b \cdot c \cdot d$; а також змішаного

типу: $y = \frac{a}{b + c + d}$.

Під час побудови моделей необхідно дотримуватись таких умов:

1) усі фактори, які включені в модель, мають бути односпрямованими, пов'язаними з результатом прямою (або зворотною) залежністю. Інакше неможлива інтерпретація отриманих результатів;

2) усі фактори моделі повинні мати реальний економічний зміст;

3) добуток двох факторів, які стоять поруч, і послідовне включення у розрахунок нового фактору має давати новий показник, що має реальний економічний зміст;

4) для забезпечення математичної точності моделі знаменник першого фактора має дорівнювати знаменнику результативної ознаки. Чисельник останнього фактора має співпадати з чисельником результативної ознаки. Використовується і зворотна побудова моделі: чисельник першого фактора є чисельником результативного, а знаменник останнього є знаменником результативної ознаки.

Вирішуються такі моделі різними методами: послідовних різниць, ланцюгових підстановок тощо. Але ці методи мають певні недоліки, перш за все, результати розрахунку залежать від порядку розташування в моделі факторів, що знижує об'єктивність оцінок. Метод пов'язаних ланцюгових індексів точніший, придатний для застосування і для мультиплікативних, і для адитивних моделей. Сутність інтегрального методу полягає у розкладанні абсолютного відхилення (приросту) результативного показника на окремі показники-фактори з урахуванням взаємозв'язку між ними.

Оцінити за допомогою детермінованих моделей можна різні процеси, а ефективність управління багато в чому залежить від того, наскільки раціонально здійснюються процеси збору, обробки, передачі і використання інформації, наскільки ідентифікована система показників, що характеризує ефективність діяльності суб'єктів господарювання, і головне, які методи використовувалися для оцінки ефективності з метою ухвалення управлінських рішень.

На першому етапі на основі змістового, логічного аналізу причинно-наслідкових зв'язків словесно визначають первинний набір ознак-факторів, на другому – аналізують структуру зв'язків між ними, на третьому – здійснюють послідовне уточнення первинного набору факторів, вибір серед них статистично значущих.

Наприклад, необхідно виявити конкретні фактори, які є джерелом зростання чи падіння прибутку підприємства, щоб їх вплив надалі можна було усунути або послабити. Ґрунтуючись на положеннях індексної теорії, побудуємо індексну функціональну модель, у якій результативним показником є прибуток підприємства ($P_{\text{чист}}$), а факторами – прибутковість активів ($ВД/А$), валова рентабельність продажів ($P_{\text{вал}}/ВД$), частка чистого прибутку у валовому ($P_{\text{чист}}/P_{\text{вал}}$). Усі фактори, включені в модель, однонаправлені, тобто вони пов'язані з результатом прямою залежністю. Модель чистого прибутку має вигляд:

$$P_{\text{чист}} = A \times \frac{ВД}{А} \times \frac{P_{\text{вал}}}{ВД} \times \frac{P_{\text{чист}}}{P_{\text{вал}}}.$$

Включення цих факторів в модель пояснюється так:

- активи характеризують основний і оборотний капітал, тобто матеріальні ресурси підприємства;
- дохідність активів характеризує ефективність використання грошових доходів підприємства;
- валова рентабельність продажів продукції показує, скільки валового прибутку припадає на кожну гривню виручки від реалізації (валового доходу);
- частка чистого прибутку у валовому дає змогу оцінити, скільки податків сплачує підприємство, і яка частка залишається у підприємства.

Розрахунки виконано методом послідовних ланцюгових індексів зв'язку.

Виявлення впливу кожного конкретного фактора на величину чистого прибутку у процесі дослідження розраховано у такий спосіб:

- 1) зміна прибутку, яка відбувається під впливом збільшення активів:

$$\Delta P A = P_0 \times (I A - 1);$$

- 2) зміна обсягу чистого прибутку під впливом дохідності активів:

$$\Delta\Pi_{ВД/А} = \Pi_0 \times (I_{ВД} - I_A);$$

3) підвищення рентабельності продажів сприяє збільшенню чистого прибутку, і навпаки, під впливом зниження рентабельності продажів чистий прибуток буде скорочуватися:

$$\Delta\Pi_{Пв/ВД} = \Pi_0 \times (I_{П} - I_{ВД});$$

4) зміна чистого прибутку під впливом частки чистого прибутку у валовому:

$$\Delta\Pi_{Пч/Пв} = \Pi_0 \times (I_{Пч} - I_{Пв}).$$

Для аналізу чистого прибутку необхідні дані за два періоди (табл. 1.7).

Таблиця 1.7 – Динаміка фінансових показників діяльності підприємства

№ з/п	Показники	2017	2018	Індекси
1.	Активи, тис. грн	5 313	4 990	0,939
2.	Валовий дохід, тис. грн	4 687	4 129	0,881
3.	Прибуток валовий, тис. грн	2 012	1 703	0,846
4.	Прибуток чистий, тис. грн	1 650	1 396	0,846
5.	Рентабельність активів, %	31,1	28,0	0,900

Розрахунки показали, що чистий прибуток у 2018 р. порівняно з 2017 р. знизився на: $1\,396 - 1\,650 = -254$ тис. грн.

Вплив кожного фактора оцінено на основі ланцюгових індексів зв'язку. Зміна активів вплинула негативно:

$$\Delta\Pi = 1\,650 (0,939 - 1) = -100,6 \text{ тис. грн};$$

вплив зміни дохідності активів виявився також негативним:

$$\Delta\Pi = 1\,650 (0,881 - 0,939) = -95,8 \text{ тис. грн.}$$

Падіння валової рентабельності продажів призвело до зниження чистого прибутку:

$$\Delta\Pi = 1\,650 (0,846 - 0,881) = -57,5 \text{ тис. грн.}$$

Вплив зміни частки чистого прибутку у валовому:

$$\Delta\Pi = 1\,650 (0,8461 - 0,8462) = -0,1 \text{ тис. грн.}$$

Загальний вплив факторів на зниження чистого прибутку:

$$-100,6 - 95,8 - 57,5 - 0,1 = -254 \text{ тис. грн.}$$

Результати розрахунків наведено у табл. 1.8.

Таблиця 1.8 – Результати факторного аналізу чистого прибутку

№ з/п	Фактори моделі	2017	2018	Вплив факторів, +, – тис. грн
1.	Активи, тис. грн	5 313	4 990	-100,6
2.	Дохідність активів, грн/грн	0,882	0,828	-95,8
3.	Валова рентабельність продажів, %	42,9	41,2	-57,5
4.	Частка чистого прибутку у валовому	82,0	82,0	-0,1
5.	Разом: чистий прибуток, тис. грн	1 650	1 396	-254,0

Отже, зниження активів вплинуло на зменшення прибутку на 100,6 тис. грн. Інші фактори також призвели до негативного результату.

Отже, індексні функціональні моделі дають змогу поглибити аналіз результативності діяльності підприємств; їх побудову можна здійснювати на основі відкритої інформації органів державної статистики, обираючи певні періоди для порівняння (кризи, падіння чи зростання); аналіз доцільно виконувати окремо для прибуткових і збиткових підприємств, що дасть можливість відстежувати конкретні збиткові підприємства і впродовж якого терміну вони мають збиток; розробляти конкретні управлінські рішення щодо прибуткових і збиткових підприємств.

Завдання для вивчення теми

Дайте коротку відповідь на питання:

1. З якою метою використовуються індексні функціональні моделі?
2. Чим характеризуються індексні функціональні моделі адитивного та мультиплікативного типу?
3. Які умови висуваються під час побудови індексних функціональних моделей?
4. Як визначається вплив факторів моделі на результативний показник?

Задача 1. За наведеною формулою виконайте факторний аналіз ефективності використання банківських ресурсів:

$$E_{\text{банк.рес.}} = \frac{ЧП}{ВК + ЗК + ПК.}$$

Показники	Базисний рік	Звітний рік
Чистий прибуток (ЧП), млн грн	5 308	4 880
Власний капітал (ВК), млн грн	42 623	74 957
Позикові кошти (ПК), млн грн	160 300	230 824
Запозичені кошти (ЗК), млн грн	176 220	290 550
Ефективність використання банківських ресурсів, %	1,40	0,82

Задача 2. Розрахуйте для кожного року чисту рентабельність банку; оцініть вплив факторів на рентабельність активів. Використайте наведену модель: $ROA = \frac{ЧП}{А} = \frac{ЗК}{А} \cdot \frac{ВП}{ЗК} \cdot \frac{ЧП}{ВП}$, де ЧП – чистий прибуток банку; А – вартість активів; ЗК – запозичені кошти; ВП – валовий прибуток (у млн грн).

Показники	Умовне позначення	Базисний період	Звітний період	Індекси, %
Запозичені кошти	ЗК	235	504	
Активи	А	3 930	6 295	
Валовий прибуток	ВП	67	70	
Чистий прибуток	ЧП	53	49	

Задача 3. За наведеними даними та моделлю $P_{СМП} = T \cdot \frac{T_{СМП}}{T} \cdot \frac{П_{СМП}}{T_{СМП}} = T \cdot d_T \cdot W$

розрахуйте:

- 1) індекси вихідних показників;
- 2) другий і третій фактори моделі;
- 3) абсолютну зміну прибутку за рахунок кожного фактора і загалом.

Показники моделі прибутку підприємств середнього й малого бізнесу

№ з/п	Показники моделі	Умовне позначення	Фактори		Індекси вихідних показників
			2017	2018	
1	Загальна чисельність зайнятих у регіоні, осіб	T	254	286	
2	Чисельність зайнятих у середніх та малих підприємствах, осіб	T _{СМП}	246	248	
3	Прибуток, отриманий середніми та малими підприємствами, млн грн	П _{СМП}	1 425	1 948	

Задача 4. За наведеними даними виконайте факторний аналіз ефективності витрат банків на персонал: $E = \frac{ЧП}{\Phi OT} = \frac{В}{\Phi OP} \cdot \frac{Р}{В} \cdot \frac{ВК}{Р} \cdot \frac{ЧП}{ВК}$.

Показники	2017	2018	Індекси
Витрати банків, млн грн (В)	12 360	21 577	
Фонд оплати праці, млн грн (ФОП)	6 720	11 404	
Ресурси банку, млн грн (Р)	297 312	588 420	
Власний капітал, млн грн (ВК)	42 623	74 957	
Чистий прибуток, млн грн (ЧП)	5 308	4 880	
Ефективність витрат на персонал, грн/грн (Е)	78,99	42,79	

1.7. КОРЕЛЯЦІЙНО-РЕГРЕСІЙНИЙ АНАЛІЗ. МНОЖИННА РЕГРЕСІЯ

Дослідження реальних явищ та процесів, які відбуваються в суспільстві, показує, що практично кожне суспільне явище знаходиться у тісному зв'язку з іншими явищами, якими б випадковими вони не здавались на перший погляд. Вивчити взаємозв'язки між показниками, закономірності, що склалися, визначити залежність, вплив факторів на результати діяльності дає змогу кореляційно-регресійний аналіз (КРА). Наприклад, ефективність використання основних фондів (фондовіддача) залежить від таких факторів: частка активної частини ОФ, величина інвестицій в основний капітал, коефіцієнт зносу та інші. Розглянемо послідовність виконання аналізу й моделювання фондовіддачі на основі даних табл. 1.9.

Етапи виконання КРА:

1. Перевірка факторів на відповідність вимогам.
2. Вибір виду рівняння регресії й оцінка параметрів у натурально-уречевленому й стандартизованому вигляді.
3. Оцінка показників щільності зв'язку.
4. Перевірка статистичної достовірності та точності рівняння регресії, параметрів рівняння, показників щільності зв'язку.
5. Змістовний аналіз результатів статистичного моделювання.

На основі теоретичного аналізу обґрунтовуємо, що між факторами є причинно-наслідкова залежність, тобто фондовіддача як залежна величина (y) знаходиться у стохастичній залежності від факторних ознак (x_i) (табл. 1.9). Щільність зв'язку підтверджуємо розрахунком парних коефіцієнтів (r_{yx_i}) кореляції (табл. 1.10). Зв'язок фондовіддачі з кожним фактором (x_i) достатньо сильний, що дає змогу включити в модель усі фактори. Але між факторами x_1 та x_3 існує дуже сильний зв'язок (0,923), що свідчить про наявність мультиколінеарності. Отже один з коефіцієнтів треба виключити з подальшого аналізу.

Таблиця 1.9 – Показники залежності фондівіддачі від економічних факторів по підприємствах однієї галузі

№ з/п	Фондовіддача, грн (y)	Частка активної частини основних фондів, % (x ₁)	Обсяги інвестицій в основний капітал, млн грн (x ₂)	Коефіцієнт зносу основних фондів (x ₃)	Теоретичні значення $\bar{y}_{x_i} = -4,589 + 0,712x_1 + 1,939x_2$
A	1	2	3	4	5
1	25,8	29,3	4,6	48,4	25,19
2	32,1	42,1	5,3	64,1	35,66
3	21,8	23,9	6,4	42,8	24,84
4	29,6	32,8	8,5	63,2	35,25
5	37,5	35,0	8,6	56,5	37,01
6	48,3	40,0	9,2	44,5	41,73
7	27,1	28,7	4,8	67,2	25,15
8	30,6	33,1	5,6	62,5	29,84
9	30,0	28,8	6,0	60,4	27,55
10	19,2	26,6	4,8	70,3	23,66
11	29,3	25,0	8,2	66,9	29,11
12	29,6	26,1	8,9	58,5	31,25
13	23,7	24,5	5,4	61,7	23,33
14	26,5	24,9	6,1	66,3	24,97
15	20,7	19,8	4,0	58,3	17,26

Середні для підприємств: $\bar{x}_2 = 6,427$ млн грн, $\bar{x}_1 = 29,37\%$, $\bar{y} = 28,79$ грн.

Зазвичай виключають із подальшого аналізу фактор, що слабкіше пов'язаний з результатом. Таким фактором є x_3 (коефіцієнт зносу основних фондів), його виключаємо з подальшого аналізу. В моделі фондівіддачі залишається 2 фактори. Коефіцієнти парної кореляції 0,335 та 0,433 свідчать, що між факторами слабкий зв'язок, тобто відібрано незалежні фактори, а між результатом (y) і факторами зв'язок сильний.

Таблиця 1.10 – Парні коефіцієнти кореляції між факторами

	Y	X ₁	X ₂	X ₃
Y	1			
X ₁	0,773	1		
X ₂	0,683	0,335	1	
X ₃	0,723	0,923	0,433	1

Високі значення парних коефіцієнтів кореляції ($r_{yx_1} = 0,78; r_{yx_2} = 0,68$) свідчать, що в рівняння регресії можна включити фактори x_1 та x_2 , оскільки вони суттєво впливають на зміну фондоддачі.

Перевірка показників на однорідність показала, що фактори (x_i і y):

- статистично однорідні, оскільки коефіцієнти варіації не перевищують критичні значення ($V_{x_i} \leq 33\%$ і $V_y \leq 33\%$);
- відповідають закону нормального розподілу;
- відсутня мультиколінеарність між факторними ознаками.

Для перевірки відповідності емпіричного розподілу факторів нормальному закону використовують χ^2 -критерій або надають висновок за допомогою t -критерія. Зокрема, використання t -критерія засновано на виконанні таких вимог: y і x_i відповідають нормальному закону розподілу, якщо мінімальне та максимальне значення факторів не виходить за межі: $y_{\min}; y_{\max} \in [\bar{y} \pm 3\sigma_y]$; $x_{i\min}; x_{i\max} \in [\bar{x}_i \pm 3\sigma_{x_i}]$. Дані для розрахунку в табл. 1.11.

$$\text{Для результату (y): } \left. \begin{array}{l} y_{\min} = 19,2 \\ y_{\max} = 48,3 \end{array} \right\} \in [28,79 \pm 3 \cdot 6,9426] = [7,96 \div 49,61].$$

Таблиця 1.11 – Розрахункові показники за даними табл. 1.9

Показники	Середнє значення	Середнє квадратичне відхилення	Коефіцієнт варіації, %
Y	28,79	6,943	24,12
X ₁	29,37	5,977	20,35
X ₂	6,43	1,709	26,59

$$\text{Для (x}_1\text{): } \left. \begin{array}{l} x_{1\min} = 19,8 \\ x_{1\max} = 42,1 \end{array} \right\} \in [29,37 \pm 3 \cdot 5,977] = [11,4423 \div 47,3043];$$

$$\text{для (x}_2\text{): } \left. \begin{array}{l} x_{2\min} = 4,0 \\ x_{2\max} = 9,2 \end{array} \right\} \in [6,43 \pm 3 \cdot 1,7086] = [1,301 \div 11,5525].$$

Отже, з імовірністю 0,997 можна стверджувати, що гіпотезу про відповідність нормальному закону за всіма факторами підтверджено. Відсутність мультиколінеарності підтверджують коефіцієнти кореляції між факторами ($r_{x_i x_j}$) 0,335 та 0,443 $|r_{x_i x_j}| \leq 0,8$, тобто менше 0,8.

Вибір рівняння множинної регресії здійснюють на основі теоретичного (змістовного) аналізу або за критеріями апроксимації: мінімальний критерій МНК – $\min \sum (y - \hat{y}_{x_i})^2$; максимальний критерій Фішера–Снедекора –

max F-критерій, найбільший коефіцієнт детермінації, мінімальна відносна помилка апроксимації (табл. 1.12).

Зазвичай обирають лінійну чи степеневу функцію, яка дає ширшу інтерпретацію для аналізу. Зазначимо, якщо вплив факторів (x_i) **змінює напрям зв'язку**, то треба використовувати квадратичну функцію; якщо зміна (y) **має обмеження зростання або зниження**, то гіперболічну; якщо інтенсивність впливу факторів із їх зростанням уповільнюється, – логарифмічну функцію (табл. 1.12).

$$\text{Степенева регресійна залежність: } \bar{y}_{x_i} = e^{0,278} x_1^{0,696} x_2^{0,392}.$$

При визначенні параметрів цього рівняння за МНК спочатку всі фактори (y та x_i) логарифмуються, тобто $\ln(y)$, $\ln(x_1)$, $\ln(x_2)$, а потім визначаються параметри a_1 і a_2 , тобто: $\ln \bar{y}_{x_i} = 0,278 + 0,696x_1 + 0,392x_2$.

Для визначення параметрів квадратичної моделі фактори спочатку зводять у квадрат (x_1^2, x_2^2), а для гіперболічної перетворюють в обернені значення: $\frac{1}{x_1}, \frac{1}{x_2}$ (табл. 1.12). Перші три моделі задовольняють критеріям апроксимації, найгірша з моделей за всіма параметрами – гіперболічна. Враховуючи, що критерії для квадратичної залежності не дуже відрізняються від лінійної та степеневі моделі, зупинимось на змістовній інтерпретації цих моделей.

Параметри a_i лінійної моделі показують, на скільки одиниць свого виміру зміниться y , якщо x_i збільшиться на одиницю свого виміру за умови, що й інші фактори, включені в модель, впливають на y , але не варіюють, тобто зафіксовані на рівні свого середнього значення. Отже, $a_1 = 0,712$ показує, що збільшення частки активної частини ОФ (x_1) на 1 % супроводжується в середньому на 15 підприємствах зростанням фондівіддачі на 0,712 грн за умови, що одночасно на зміну фондівіддачі позитивно впливає й обсяг інвестицій.

Таблиця 1.12 – Формальні критерії апроксимації регресійних моделей

№ з/п	Форма рівняння регресії	$R_{yx_i}^2$	F_p	$\varepsilon_{\text{відн}}, \%$
1	Лінійна модель $\bar{y}_{x_i} = -4,589 + 0,712x_1 + 1,939x_2$	0,800	21,94	8,8
2	Степенева модель $\bar{y}_{x_i} = e^{0,278} x_1^{0,696} x_2^{0,392}$	0,792	21,8	2,5
3	Квадратична модель $\bar{y}_{x_i} = 12,07 + 0,011x_1^2 + 0,150x_2^2$	0,815	25,52	8,8
4	Гіперболічна модель $\bar{y}_{x_i} = 60,845 - 588,52 \frac{1}{x_1} - 67,354 \frac{1}{x_2}$	0,703	13,2	9,5

При збільшенні частки активної частини ОФ на 1 %, тобто $\bar{x}_{1,розр} = 30,37\%$ забезпечить значення фондівдачі в середньому на 15 підприємствах на рівні 29,5 грн:

$$\bar{y}_{x_i} = -4,589 + 0,712 \cdot 30,37 + 1,939 \cdot 6,427 = 29,5 \text{ грн.}$$

Аналогічно можна розрахувати зміну фондівдачі під час зміни частки активної частини ОФ за кожним підприємством. Зокрема, за першим підприємством, при $x_1 = 30,3\%$ ($29,3 + 1 = 30,3\%$) і $x_2 = 6,427$, рівень фондівдачі складе 29,44 грн, що на 3,64 грн більше від фактичного значення.

Параметр $a_2 = 1,939$ показує, що в середньому за всіма підприємствами кожний млн грн приросту інвестицій супроводжується зростанням фондівдачі на 1,939 грн за умови, що одночасно впливає і частка активної частини ОФ, котра зафіксована на рівні свого середнього значення.

Значення параметрів рівняння в натуральній формі не співставні між собою, оскільки показники виражені у різних одиницях виміру факторних ознак, тому **додатково розраховується стандартизоване рівняння регресії**, в якому параметри співставні, оскільки виражені у стандартних відхиленнях (σ_{x_i}, σ_y представлені в табл. 1.11).

У загальному вигляді це рівняння: $t_{y_{x_i}}^- = \alpha_1 t_{x_1} + \alpha_2 t_{x_2}$;

у стандартизованому вигляді $t_{y_{x_i}}^- = 0,613 t_{x_1} + 0,477 t_{x_2}$;

де $\alpha_1 = a_1 \frac{\sigma_{x_1}}{\sigma_y} = 0,712 \frac{5,977}{6,9426} = 0,613$; $\alpha_2 = a_2 \frac{\sigma_{x_2}}{\sigma_y} = 1,939 \frac{1,7086}{6,9426} = 0,477$.

Співвідношення параметрів α_1, α_2 показує, що x_1 сильніше впливає на зміну y , тобто зміна частки активної частини ОФ у 1,285 раза ($0,613 / 0,477 = 1,285$) сильніше впливає на рівень фондівдачі, ніж обсяги інвестицій.

Параметри степеневі моделі (2) є частковими коефіцієнтами еластичності. Наприклад, $a_1 = 0,696$ показує, що кожен відсоток приросту частки активної частини ОФ супроводжується зростанням фондівдачі на 0,696 % за умов, що одночасно впливає і обсяг інвестицій, котрий зафіксований на рівні свого середнього значення. Аналогічний висновок можемо зробити за параметром $a_2 = 0,392$.

Оцінка показників щільності зв'язку:

- парні коефіцієнти кореляції r_{yx_i} і детермінації: $d_{yx_i} = r_{yx_i}^2$;
- множинні коефіцієнти кореляції та детермінації;
- часткові коефіцієнти еластичності.

Парні коефіцієнти (табл. 1.10) оцінюють напрям і силу зв'язку y з кожною факторною ознакою окремо, тобто ізольовано від впливу інших факторів, що

включені в модель. Наприклад, $r_{yx_1} = 0,77$ показує, що залежність фондівдачі від частки активної частини ОФ сильна (за шкалою Чеддока) й пряма, тобто на 59,3 % ($d_{yx_1} = 0,593$) зміна фондівдачі за 15 підприємствами обумовлена впливом частки активної частини ОФ. За обсягом інвестицій сила зв'язку помірна ($r_{yx_2} = 0,68$, $d_{yx_2} = 0,462$), а саме: 46,2 % зміни фондівдачі обумовлюється зміною величини інвестицій.

Множинні коефіцієнти кореляції та детермінації оцінюють щільність зв'язку всіх факторів із результатом (y). Для лінійної залежності коефіцієнт кореляції $R_{yx_i}^- = 0,894$ (зв'язок щільний), а коефіцієнт детермінації $R_{yx_i}^2 = 0,8$ свідчать, що варіація фондівдачі за об'єктами, що аналізуються, на 80 % обумовлена зміною частки активної частини ОФ і обсягами інвестицій. Решта 20 % – це вплив факторів, котрі не включались у модель.

Статистичний аналіз результатів моделювання. Перевірка статистичної достовірності лінійного рівняння регресії виконується за допомогою F-критерія (Фішера–Снедекора). Розрахункове значення складає:

$$F_p = \frac{\sum(y-\bar{y})^2 - \sum(y-\bar{y}_{x_i})^2}{\sum(y-\bar{y}_{x_i})^2} \cdot \frac{n-m}{m-1} = \frac{723-144,9}{144,9} \cdot \frac{12}{2} = 23,94; \text{ табличне} - 3,89;$$

$$V_1 = m - 1 = 2, V_2 = 15 - 3 = 12; \quad F_\alpha \left\{ \begin{array}{l} \alpha = 0,05 \\ V_1 = 2 \\ V_2 = 12 \end{array} \right\} = 3,89 \quad F_p > F_\alpha.$$

З імовірністю 0,95 можна стверджувати, що лінійна регресійна модель статистично достовірна. Статистична точність моделі оцінюється відносною помилкою апроксимації: $\varepsilon_{\text{відн}} = \left(\frac{\sum |y - \bar{y}_{x_i}|}{y} \right) : n \cdot 100$.

Якщо помилка апроксимації менше 10 %, то рівняння відповідає критерію статистичної точності моделі. У нашому випадку – $\varepsilon_{\text{відн}} = 8,8 \%$.

Перевірка достовірності параметрів рівняння здійснюється на основі t-критерія (Стьюдента): $t_{a_i} = \frac{|a_i|}{m_{a_i}}$, де $m_{a_i} = \frac{\sigma_\varepsilon}{\sigma_{x_i} \sqrt{n}} \sqrt{\frac{1}{1 - R_{x_i, x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_m}^2}}$,

m_{a_i} – стандартна помилка параметра a_i ;

$$\sigma_\varepsilon = \sqrt{\frac{\sum (y - \bar{y}_{x_i})^2}{n - m}} - \text{стандартна помилка рівняння регресії};$$

$$\sigma_\varepsilon = \sqrt{\frac{144,9}{15 - 3}} = 3,47;$$

$$\sigma_{x_i} = \sqrt{\frac{\sum (x_i - \bar{x}_i)^2}{n-1}}, \quad \sigma_{x_1} = 6,4, \quad \sigma_{x_2} = 1,83.$$

Параметри a_1, a_2 з імовірністю 0,95 статистично достовірні, тобто за ними характер і щільність зв'язку мають об'єктивний характер.

Перевірка статистичної достовірності показників щільності зв'язку

Статистична достовірність показників щільності зв'язку перевіряється за допомогою t-критерія. Для множинного коефіцієнта кореляції:

$$t_{R_{y_{x_i}}} = \frac{|R_{y_{x_i}}| \sqrt{n-m}}{\sqrt{1-R_{y_{x_i}}^2}} = \frac{0,8\sqrt{12}}{\sqrt{1-0,64}} = 4,61; \quad t_{\alpha} \left\{ \begin{array}{l} \alpha = 0,05 \\ V = 12 \end{array} \right\} = 2,18; \quad t_{R_{y_{x_i}}} > t_{\alpha}.$$

Отже, усі статистичні характеристики відповідають критеріям статистичної точності та достовірності. Це свідчить про об'єктивність і адекватність оцінок залежностей за допомогою багатofакторного рівняння регресії.

Під час дослідження явищ або процесів у динаміці часто стикаються з автокореляцією, тобто залежністю показника від попередніх даних. Перевірка на автокореляцію здійснюється за допомогою критерія Дарбіна–Уотсона. Для усунення цього явища в модель вводять додатково фактор часу (1, 2 ... n).

Крім аналітичних функцій, багатofакторні регресійні моделі використовують для прогнозування явищ або процесів, що вивчаються. За допомогою рівняння регресії можна розрахувати, наприклад, теоретичні значення обсягу виробництва продукції сільського господарства. Негативну різницю між фактичними й теоретичними значеннями можна вважати втратами або невикористаними резервами збільшення обсягів за рахунок впливу випадкових, циклічних, сезонних чи інших причин.

За допомогою кореляційно-регресійного аналізу, разом з використанням трендових моделей, можна розрахувати реалістичний прогноз розвитку соціально-економічного явища, яке буде враховувати вплив досліджуваних факторів на результат. Для виконання прогнозу фондovіддачі спочатку необхідно за допомогою трендової моделі (чи іншим способом) розрахувати прогнозні значення відібраних факторів та підставити їх у регресійну модель. Далі розрахувати точковий прогноз, довірчі інтервали та виконати прогнозування у довірчих інтервалах.

Дайте коротку відповідь на питання:

1. Яким показником оцінюється щільність зв'язку?
2. Який показник щільності зв'язку змінюється в межах від -1 до $+1$?
3. Яка задача вирішується за допомогою коефіцієнта детермінації?

4. Якщо лінійний коефіцієнт кореляції становить 0,84, чому дорівнює коефіцієнт детермінації?
5. Як інтерпретується коефіцієнт еластичності?
6. За яким показником перевіряється однорідність розподілу ознаки?

Задача 1. Виконайте кореляційно-регресійний аналіз (КРА) чистого доходу металургійного підприємства (y). Побудуйте багатофакторну регресійну модель, обравши з кожної підсистеми найбільш впливовий показник. Перевірте модель і фактори на достовірність і точність. Виконайте прогнозування чистого доходу в довірчих інтервалах на наступні 3 роки (дані в табл. 1.13).

➤ **Ресурсний потенціал підприємства:**

- x_1 – середньорічна вартість необоротних активів (НОА);
- x_2 – чисельність працівників.

➤ **Стан валютного ринку, що відображається на величині експортної виручки металургійних підприємств:**

- x_3 – середньорічний курс долара за даними НБУ.

➤ **Потреба машинобудування в металопродукції:**

- x_4 – обсяг реалізованої продукції машинобудівних підприємств.

➤ **Цінова кон'юнктура світового ринку сировинної продукції:**

- x_5 – індекс сировинних цін за даними МВФ (Commodity Price Index);
- x_6 – індекс цін на метали за даними МВФ (Commodity Metals Price);
- x_7 – ціна імпортованих контрактів на залізну руду.

Таблиця 1.13 – Вихідні дані для кореляційно-регресійного аналізу чистого доходу підприємства

Рік	ЧД (y), (млн грн)	НОА x_1 , (млн грн)	Чисельність працюючих, x_2 (осіб)	Курс \$, x_3 , (грн)	Реаліз. продук. x_4 (млн грн)	Індекс цін сировини x_5 (%)	Індекс цін x_6 (%)	Ціна руди x_7 (дол./т)
1	2	3	4	5	6	7	8	9
2009	4 722	1 678	18 995	5,12	59 668	100,0	100,0	28,11
2010	5 754	1 898	18 954	5,05	68 731	120,6	156,2	33,45
2011	10 215	2 156	18 465	5,05	98 340	134,8	183,3	36,63
2012	7 902	2 852	17 729	5,27	121 780	172,2	169,0	61,57
2013	9 380	3 197	16 399	7,79	85 833	120,5	136,5	79,99
2014	10 752	3 326	15 700	7,94	97 057	152,0	202,3	146,7
2015	15 543	3 412	15 235	7,97	130 848	191,9	229,7	167,8

Закінчення таблиці

1	2	3	4	5	6	7	8	9
2016	14 478	3 302	13 941	7,99	140 539	185,6	191,0	128,5
2017	11 864	3 547	12 239	7,99	113 927	182,9	182,9	135,3
2018	14 208	4 341	11 432	11,89	101 925	171,45	164,1	96,8

Задача 2. Виконайте КРА впливу факторів цінової кон'юнктури світового ринку сировинної продукції (x_5, x_6, x_7) на зміну чистого доходу підприємства (за даними таблиці 1.13). Перевірте модель і фактори на достовірність і точність. Виконайте прогнозування чистого доходу в довірчих інтервалах на наступні 3 роки (дані в табл. 1.13).

Зробіть висновки.

Задача 3. Виконайте КРА впливу ресурсного потенціалу та стану валютного ринку (x_1, x_2, x_3) на зміну чистого доходу підприємства (за даними таблиці 1.13). Перевірте модель і фактори на достовірність і точність. Виконайте прогнозування чистого доходу в довірчих інтервалах на наступні 3 роки (дані в табл.1.13).

Зробіть висновки.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ДО РОЗДІЛУ 1

1. Семенова К. Д., Тарасова К. І. Бізнес-статистика: підручник. К.: ФОП Гуляева В. М. 2018. 210 с.
2. Сидорова А. В., Глущенко А. М. Статистичне забезпечення управління змінами на підприємствах великого бізнесу: монографія. Вінниця: ДонНУ імені Василя Стуса, 2017. 169 с.
3. Єріна А. М. Статистичне моделювання та прогнозування: навчальний посібник. К.: КНЕУ, 2001. 170 с.
4. Бізнес-статистика: навчальний посібник / Матковський С. О., Гринькевич О. С., Вдовин М. Л., Вільчинська О. М., Марець О. Р., Сорочак О. З. К.: Алерта, 2016. 280 с.
5. Єріна А. М., Пальян З. О. Статистика: підручник. К.: 2010. 384 с.
6. Методические рекомендации для выполнения лабораторных работ по дисциплине «Статистическое моделирование и прогнозирование» / Л. Л. Шамилева, А. Ю. Дейна. Винница: ДонНУ им. Васыля Стуса, 2017. 104 с.
7. Практикум зі статистики: навчальний посібник / А. В. Сидорова, Г. В. Анісімова, Л. О. Масіч, Л. Л. Шамілева та ін. Донецьк: Каштан, 2014. 284 с.
8. Сидорова А. В., Кіосак Я. В. Міжнародна статистика: підручник. Донецьк: Каштан, 2013. 384 с.
9. Воронцов К. В. Алгоритмы кластеризации и многомерного шкалирования. Курс лекций. МГУ, 2007.
10. Чубукова И. А. Курс лекций «Data Mining», Интернет-университет информационных технологий. URL: www.intuit.ru/department/database/datamining

РОЗДІЛ 2

МЕТОДОЛОГІЯ БІЗНЕС-АНАЛІТИКИ В РИЗИК-МЕНЕДЖМЕНТІ

2.1. МЕТОДИ БІЗНЕС-АНАЛІТИКИ В ОЦІНЦІ РИЗИКІВ

Дослідження питань, що пов'язані з ризиком, почалося лише за часів Ренесансу, коли люди звільнилися від багатьох заборон і поставили під сумнів багатовікові застигли вірування. Слово «ризик» походить від староіталійського *risicare*, що означає «наважуватися». У XVI ст. завдяки блискучим досягненням Кардано, а саме його трактату *Liber de Ludo Aleae* «Книга про випадкові ігри», де було створено теоретичні основи комбінаторики, виникла сама ідея і можливість управління ризиком.

Перша серйозна робота, присвячена детальному аналізу прийняття рішень в умовах невизначеності, опублікована в 1921 р. Френком Найтом – «Ризик, невизначеність і прибуток», де вперше було розрізнено ризик та невизначеність. Ризик Найт розумів як вимірну невизначеність.

Протягом наступної чверті століття серйозний прогрес у розумінні ризику й невизначеності був досягнутий у рамках теорії стратегічних ігор, яка показала, що істинним джерелом невизначеності є наміри інших.

Приблизно десять років тому головним критерієм прийняття рішень була якість, на зміну якій прийшла ціна, що призвело до глобального аутсорсингу, та сукупна вартість володіння. В сучасних умовах на зміну зазначеним критеріям прийшов ризик-менеджмент, головною метою якого є передбачення та запобігання, ефективне управління результатами й очікуваннями зацікавлених сторін в умовах глобалізації та аутсорсингу.

Важливість використання методології аналізу та управління ризиками підтверджується тим, що в кожній версії єдиного всесвітньо визнаного стандарту бізнес-аналізу – Керівництво BABOK (*The Essential Standart for Business Analysis*) – зазначена методологія є незмінною та включена у спектр рекомендованих інструментів аналізу (*Techniques*).

Стандартами в галузі ризик-менеджменту передбачається уніфікація термінології, що використовується в цій сфері, складових частин процесу управління ризиками та підходів до побудови організаційної структури ризик-менеджменту. Відповідно до стандарту ISO 31000: 2009 ризик – це вплив невизначеності на цілі. Вплив розглядається як відхилення від очікуваного і з позитивними, і з негативними наслідками. Цілі можуть мати різні аспекти (наприклад, фінансові аспекти, що стосуються професійної безпеки і здоров'я; екологічні завдання тощо) і можуть ставитися до різних рівнів (зокрема до стратегічного рівня, організаційного, рівня проекту, продукції та процесу). Невизначеність – це стан,

при якому відсутня інформація щодо розуміння або знання події, її наслідків або ймовірності.

Ризик-менеджмент передбачає виконання послідовності кроків при оцінці ризиків:

- ідентифікація ризику – процес знаходження, розпізнавання й опису ризику;
- аналіз ризику – процес розуміння природи ризику та визначення рівня ризику;
- визначення ступеня ризику – процес порівняння результатів аналізу ризику з критеріями ризику для визначення того, чи можна прийняти величину ризику;
- обробка ризику – процес модифікації ризику.

Ідентифікація ризику – це процес визначення елементів ризику, складання їх переліку та опис кожного з елементів ризику. Метою ідентифікації ризику є складання переліку джерел ризику і подій, які можуть вплинути на досягнення кожної зі встановлених цілей організації або зробити виконання цих цілей неможливим. Після ідентифікації ризику організація повинна здійснювати ідентифікацію суттєвих особливостей проекту, персоналу, процесів, систем та засобів управління.

Методи ідентифікації ризику можуть включати в себе:

- методи оцінки ризику на основі документальних свідчень, прикладами яких є аналіз контрольних листів, аналіз експериментальних даних, а також даних і подій, що сталися в минулому;
- підхід, згідно з яким група експертів здійснює встановлений процес ідентифікації ризику за допомогою структурованих підказок та запитань;
- індуктивні методи.

Для підвищення точності і повноти ідентифікації ризику можуть бути використані різні допоміжні методи, наприклад, метод мозкового штурму та метод Дельфі.

Контрольні листи є переліком небезпек, ризиків або відмов засобів управління, які зазвичай розробляють на основі отриманого раніше досвіду, результатів.

Використання контрольних листів при ідентифікації ризиків передбачає виконання такої процедури:

- визначення галузі застосування;
- складання контрольного листа так, щоб він охоплював усю сферу застосування. Контрольні листи мають бути ретельно складені для досягнення поставленої мети. Наприклад, складений раніше контрольний лист не може бути використаний при ідентифікації нових небезпек або ризику;

- особа або група осіб повинні застосовувати контрольний лист послідовно до кожного елементу процесу або системи для визначення того, чи представлений цей елемент у контрольному листі.

Переваги методу контрольних листів полягають у тому, що контрольні листи можуть використовувати особи, які не є експертами. Якщо контрольні листи добре розроблені, то вони об'єднують різноманітні види експертних оцінок у просту для використання форму оцінки.

Серед недоліків виділяють те, що робота з контрольними листами часто стримує свободу думок при ідентифікації небезпек, а сам метод заснований на спостереженнях, тому існує стійка тенденція не бачити або не помічати проблеми.

У *структурованому інтерв'ю* опитуваному задають питання з заздалегідь підготовленого переліку, заохочуючи в такий спосіб всебічний аналіз ситуації та більш повну ідентифікацію небезпек і ризиків. Частково структуроване інтерв'ю аналогічне структурованому, однак воно забезпечує більшу свободу при обговоренні проблеми, що досліджується.

Структуровані і частково структуровані інтерв'ю корисні в ситуаціях, коли важко зібрати людей для обговорення або коли вільне обговорення у групі неможливе.

Для виконання інтерв'ю необхідно скласти перелік питань. Питання мають бути викладені зрозумілою для опитуваного мовою та охоплювати тільки одну проблему. Відповіді на питання не мають бути обмежені за часом. Необхідно стежити за тим, щоб постановка питання не підказувала опитуваному певну відповідь.

Переваги структурованого інтерв'ю полягають у можливості обміну інформацією «один на один», що дає змогу розглянути проблему з усіх боків, та залучення до обговорення проблеми більшої кількості причетних сторін, ніж мозковий штурм. Водночас при структурованому інтерв'ю важко застосувати способи стимулювання і фантазії людини, які є особливістю мозкового штурму.

Отже, існують і спеціально розроблені методи ідентифікації ризиків, і загальновідомі методи експертної оцінки. Однак незалежно від методів, що будуть використовуватися, важливо враховувати людські й організаційні чинники та відповідні викликані відхилення.

Аналіз ризику включає в себе аналіз і дослідження інформації про ризик. Аналіз ризику забезпечує вхідні дані процесу загальної оцінки ризику, допомагає у прийнятті рішень щодо необхідності обробки ризику, а також допомагає вибрати відповідні стратегії і методи обробки ризику.

Аналіз ризику включає аналіз імовірності і наслідків ідентифікованих небезпечних подій з урахуванням наявності та ефективності застосовуваних методів

управління. Дані про ймовірність подій і їх наслідки використовують для визначення рівня ризику.

Методи, які використовуються при аналізі ризику, можуть бути якісними, кількісними або змішаними.

При якісній оцінці ризику визначають наслідки, ймовірність і рівень ризику за шкалою «високий», «середній» і «низький», оцінка наслідків і ймовірності може бути об'єднана, а порівняльну оцінку рівня ризику в цьому разі проводять відповідно до якісних критеріїв.

У змішаних методах використовують числову шкалу оцінки наслідків, ймовірності та їх поєднання для визначення рівня ризику за відповідною формулою. Шкали можуть бути лінійними, логарифмічними або можуть бути побудовані за іншими принципами. Формули, що використовуються, можуть бути різними.

При *кількісному аналізі* оцінюють практичну значимість і вартість наслідків, їх ймовірності, і отримують значення рівня ризику в певних одиницях, установлених при розробці сфери застосування менеджменту ризику.

Для оцінки ймовірності зазвичай застосовують такі три загальні підходи, які можуть бути використані і самостійно, і спільно:

1. Використання відповідних хронологічних даних для ідентифікації події або ситуації, що сталися в минулому, і допускання можливості екстраполяції ймовірності їх появи в майбутньому. Якщо відповідно до наявних даних частота появи події дуже низька, то всі оцінки ймовірності матимуть високу невизначеність.

2. Використання для оцінки ймовірності методів прогнозування (зокрема аналізу дерева помилок і аналізу дерева подій). Якщо хронологічні дані недоступні або недостовірні, то для оцінки ймовірності головної події числові дані отримують методом прогнозування.

3. Використання експертних оцінок у систематизованому та структурованому процесі оцінки ймовірності. Для отримання експертних оцінок треба використовувати всю доступну інформацію. Існують формалізовані методи отримання експертних оцінок, які допомагають формулювати відповідні питання, наприклад, методи Дельфі, попарного порівняння, ранжування за категоріями оцінки та абсолютних оцінок.

Порівняльна оцінка ризику включає в себе зіставлення рівня ризику з критеріями ризику, що встановлені при визначенні сфери застосування менеджменту ризику, для визначення типу ризику і його значущості.

Порівняльна оцінка ризику використовує інформацію про ризик, отриманий при аналізі ризику. Результати порівняльної оцінки ризику використовують для прийняття рішень про майбутні дії.

Найпростіша структура для визначення критеріїв ризику – це встановлення одного рівня, що розділяє небезпеки і ризик, що вимагають обробки, від тих, які подібних дій не вимагають. Застосування такої структури призводить до простих і зрозумілих результатів, однак не відображає невизначеність, притаманну оцінці ризику і встановленому прикордонному рівню ризику.

Відповідно до загальних підходів ризик поділяється на три групи:

- вища група, в якій рівень ризику є неприпустимим, безвідносно до переваг прийняття ризику і доходів, отримуваних від діяльності організації, обробка ризику є необхідною незалежно від витрат;
- середня група («сіра» область), для якої витрати і переваги прийняття ризику треба враховувати, а можливості співвідносити з наслідками;
- нижча група, в якій рівень ризику незначний або настільки малий, що необхідність обробки ризику відсутня.

Одним з найпопулярніших методів ранжування і порівняння ризиків є *матриця наслідків і ймовірностей*. Матрицю зазвичай застосовують як засіб попередньої оцінки, якщо було виявлено кілька видів ризиків, наприклад, для визначення того, який ризик вимагає подальшого або більш докладного аналізу, який ризик необхідно обробляти насамперед, а який треба розглядати на більш високому рівні менеджменту. Цю матрицю також застосовують для відбору видів ризику, які не потребують подальшого розгляду, а також для визначення прийнятності чи неприйнятності ризику відповідно до матриці.

Перевагами використання матриці наслідків і ймовірностей є відносна простота використання та забезпечення швидкого ранжування ризику за рівнями значущості. Водночас вона не позбавлена певних недоліків, серед яких головним є те, що застосування матриці досить суб'єктивне і значною мірою залежить від фахівця, що виконує оцінку. Крім того, матриця має бути розроблена для конкретних обставин, у таких умовах досить важко скласти універсальну матрицю, яку підприємство може застосувати в будь-яких обставинах.

Отже, основною метою аналізу та порівняльної оцінки ризику є зосередження ресурсів на найважливіших видах небезпечних подій та ризиків. Важливо не пропустити події з високою частотою появи і істотним сукупним ризиком. Аналіз має бути заснований на критеріях, встановлених у галузі застосування менеджменту ризику.

На конкретному прикладі будівництва заводу реалізуємо методи бізнес-аналітики в оцінці ризиків.

Наприклад, інвестиційна компанія досліджує можливість будівництва нового заводу з виробництва ліків. Вартість будівництва складає 3,4 млн грн. Згідно з прогнозами аналітиків компанія в перший, другий і третій роки проекту про-

дасть 8 020 тис., 9 670 тис. та 11 320 тис. упаковок ліків за ціною 6 грн, 6,05 грн і 6,1 грн з урахуванням ПДВ за упаковку відповідно.

Економісти проекту визначили, що ставка податку на прибуток становить 30 %, ставка дисконтування – 10 %, собівартість становить 55 % від ціни препарату, адміністративні витрати, витрати на збут та інші операційні витрати складають 10 % від виручки.

Необхідно підготувати фінансовий аналіз ефективності будівництва заводу та провести ідентифікацію ризиків.

Щоб прийняти рішення щодо ефективності проекту, необхідно розрахувати показник *NPV* – чистої приведеної цінності, який виражає в «сьогоднішній» оцінці ефект від проекту.

Формули, для розрахунку показника *NPV* в ПП MS Excel наведені на рис. 2.1.

	A	B	C	D	E	F
1		0 год	1 год	2 год	3 год	%
2	Ціна упаковки		6	6,05	6,1	
3	Обсяг продажу		802000	967000	1132000	
4	Виручка		=C2*C3	=D2*D3	=E2*E3	
5	Собівартість		=C2*0,55*C3	=D2*0,55*D3	=E2*0,55*E3	0,55
6	Валовий прибуток		=C4-C5	=D4-D5	=E4-E5	
7	Операційні витрати		=C2*C3*0,1	=D2*D3*0,1	=E2*E3*0,1	0,1
8	Чистий дохід до опадат		=C6-C7	=D6-D7	=E6-E7	
9	Подати		=C8*0,3	=D8*0,3	=E8*0,32	0,3
10	Інвестиції	3400000				
11	Чистий дохід		=C8-C9	=D8-D9	=E8-E9	
12	NPV		=ЧПС(0,1;-B10;C11;D11;E11)			

Рисунок 2.1 – Розрахунок показника *NPV* для проекту будівництва заводу в ПП MS Excel

Розрахунок ефективності проекту за наведеними формулами (рис. 2.1) свідчить про те, що проект окупається та має позитивний показник *NPV*.

З метою ідентифікації ризиків було розроблено контрольні листи та проведено структуроване інтерв'ю.

На основі документальних свідчень та подій, що сталися в минулому, відомо, що існування на фармацевтичному ринку жорсткої конкуренції може призвести до зниження ціни нижче прогнозованої. Структуроване інтерв'ю, що використовувалися для отримання зазначеної інформації, містило в собі такі питання:

1. Чи вважаєте Ви, що ціна на продукт, який виводять на ринок, може не відповідати прогнозованому значенню в кожний рік продажів (6 грн, 6,05 грн і 6,1 грн з урахуванням ПДВ за упаковку в 1, 2 та 3 роки відповідно)?

- Так, ціна може бути іншою.

- Ні, ціна залишиться на рівні прогнозованої (6 грн, 6,05 грн і 6,1 грн з урахуванням ПДВ за упаковку в 1, 2 та 3 роки відповідно).

2. Якщо ціна для продукту, який виводять на ринок, може бути іншою, то який діапазон коливання ціни є ймовірним для кожного року продажів?

- $\pm 5\%$.
- $\pm 10\%$.
- $\pm 15\%$.
- $\pm 20\%$.
- Інша відповідь.

3. Якщо ціна для продукту, який виводять на ринок, може бути іншою, з певним діапазоном коливання, то яка ймовірність коливання від прогнозованої ціни для кожного строку продажів?

- Ймовірність коливання однакова в певних діапазонах та не залежить від прогнозованої ціни.
- Ймовірність прогнозованої ціни – вища, а коливання в певних діапазонах однакове. Ціна за зазначені діапазони не вийде.
- Зі впевненістю 95 % можна стверджувати, що ймовірність прогнозованої ціни – вища, а коливання в певних діапазонах однакове, але ціна може вийти за зазначені діапазони.
- Зі впевненістю 95 % можна стверджувати, що ймовірність прогнозованої ціни – вища, а коливання в певних діапазонах однакове, відхилення вгору є більш ймовірним. Ціна може вийти за зазначені діапазони.
- Інша відповідь.

4. Який діапазон коливання ціни вище прогнозованої може зробити продукт неконкурентоспроможним на наявному ринку:

- [0 %; 5 %].
- [0 %; 10 %].
- [0 %; 15 %].
- [0 %; 20 %].

За допомогою контрольних листів експерти також встановили, що обсяг продажів та майбутня собівартість може змінюватися. Ймовірність зміни цих показників є високою на думку експертів.

Також можлива зміна собівартості може бути обумовлена тим фактом, що 10 % від собівартості продукту має бути оплачено валютою, оскільки ця сировина закуповується за кордоном. Вважається, що курс буде дорівнювати 20 грн за 1 дол., але є загроза того, що курс може досягти 30 грн. Експерти оцінили, що ймовірність цього 25 %.

Додаткові ризики, що пов'язані з будівництвом заводу, полягають у тому, що люди, які мешкають у зоні будівництва, можуть бути проти того, щоб з'являлось загрознає для їх життя та довкілля виробництво. Ймовірність цього складає 20 %. Попередні переговори свідчать про те, що ймовірність зупинки будівництва з причини блокування проекту з боку громад може бути знижено до 2 % за умови будівництва дитячого майданчика за кошти фармакологічної компанії, ціна якого складатиме 1 % від інвестицій. Якщо будівництво не буде розпочато вчасно, то в першому році продукцію не буде реалізовано при тому, що інвестиції в будівництво вже будуть застосовані.

В останній рік реалізації проекту передбачається заміна обладнання, ймовірність аварії при заміні складає 1 %. Якщо аварія трапиться, то штрафні санкції складуть 1 млн грн.

Отже, проведена ідентифікація ризиків дає змогу встановити, що з високою ймовірністю ефективність проекту може бути значно нижчою, ніж це було доведено за допомогою розрахунку показника *NPV* без врахування впливу факторів ризиків та невизначеності.

Завдання для вивчення теми

I. Дайте коротку відповідь на питання:

1. Чим поняття «ризик» відрізняється від поняття «невизначеність»?
2. Як називається єдиний всесвітньо визнаний стандарт бізнес-аналізу?
3. Яку послідовність кроків передбачає ризик-менеджмент при оцінці ризиків?
4. Яка мета ідентифікації ризиків?
5. У чому полягають переваги та недоліки методу контрольних листів?
6. У яких ситуаціях корисні структуровані і частково структуровані інтерв'ю?
7. З якою метою використовують дані про ймовірність подій і їх наслідки?
8. Назвіть три загальні підходи, які застосовують для оцінки ймовірності.
9. Яка структура визначення критеріїв ризику є найпростішою?
10. Як називається група ризиків, в якій рівень ризику є неприпустимим?

II. Завдання

Побудуйте структуру контрольних листів для ідентифікації ризиків збільшення майбутньої собівартості інвестиційного проекту, в тому числі з урахуванням валютних ризиків.

2.2. ІНСТРУМЕНТИ АНАЛІЗУ РИЗИКІВ

Одною з важливих конкурентних переваг будь-якого економічного суб'єкта на ринку є швидка реакція на всі зміни, відповідно до яких змінюються та ускладнюються відповідні фактори ризику, відкриваються їх нові, раніше невідомі сторони та характеристики. Ризики стають багатofакторними, міждисциплінарними, мають низку складних внутрішніх залежностей. Звідси виникає необхідність використання нових комп'ютерних технологій та складних інструментів моделювання.

До інструментів аналізу ризиків, що реалізуються у спеціальних програмних продуктах, належать мультикритеріальний аналіз рішень та метод Монте-Карло.

Метою **мультикритеріального аналізу** є ранжування критеріїв для об'єктивної та прозорої оцінки різних варіантів рішень з урахуванням факторів ризику та невизначеності з метою розташування за перевагами доступних варіантів рішень. Аналіз включає в себе розробку матриці варіантів та критеріїв, які треба ранжувати і об'єднати для здійснення загальної оцінки кожного варіанта рішення.

Вхідними даними є набір варіантів рішень для проведення аналізу. Зазвичай до них належать:

$V = \{v_1, v_2, \dots, v_n\}$ – безліч варіантів (аналогів), які підлягають аналізу;

$C = \{c_1, c_2, \dots, c_m\}$ – безліч кількісних та якісних критеріїв, за якими оцінюються варіанти.

Завдання полягає в тому, щоб упорядкувати елементи безлічі V за критеріями з множини C .

Зазвичай зазначений вище процес включає в себе виконання групою компетентних фахівців, які представляють причетні сторони, такої послідовності дій:

- встановлення мети (одна або декілька);
- визначення якісних ознак (критеріїв, показників оцінки або якісних характеристик виконання роботи), відповідних кожній меті;
- структурування якісних ознак за ієрархічним принципом;
- розробка варіантів рішень, які необхідно оцінити відповідно до обраних критеріїв;
- визначення важливості критеріїв і призначення для кожного з них вагового коефіцієнта;
- оцінка альтернативних варіантів рішень з урахуванням критеріїв, які можуть бути представлені у вигляді матриці бальних оцінок;
- об'єднання множинних бальних оцінок для кожної якісної ознаки в об'єднану бальну оцінку, що враховує безліч якісних ознак;
- оцінка отриманих результатів.

Існують різні методи, відповідно до яких кожному критерію може бути призначений ваговий коефіцієнт, і різні способи об'єднання оцінок за критеріями для кожного варіанта рішення в єдину бальну оцінку. Серед таких методів найбільш популярними є:

- розгляд критеріїв як нечітких множин, які задані на універсальних множинах варіантів за допомогою функції приналежності;
- визначення функцій приналежності нечітких множин на основі експертної інформації про парні порівняння варіантів за допомогою 9-бальної шкали Сааті;
- ранжування варіантів на основі перетину нечітких множин – критеріїв, які відповідають схемі Беллмана–Заде.

Відповідно до першого методу критерій c_i можна представити у вигляді нечіткої множини \tilde{c}_i , яку задано на універсальній множині V у такий спосіб:

$$\tilde{c}_i = \left\{ \frac{\mu^l(v_1)}{v_1}, \frac{\mu^l(v_2)}{v_2}, \dots, \frac{\mu^l(v_n)}{v_n} \right\},$$

де $\mu^l(v_i)$ – ступінь приналежності елемента v_i до нечіткої множини \tilde{c}_i .

Згідно з методом Сааті вибір пріоритетних рішень здійснюється за допомогою парних порівнянь. Для фіксації результату порівняння пари альтернатив використовується шкала такого типу:

- 1 – рівноцінність;
- 3 – помірна перевага;
- 5 – сильна перевага;
- 7 – дуже сильна перевага;
- 9 – вища перевага.

Щоб визначити ступінь приналежності, формуються матриці парних порівнянь A з елементами a_{ij} варіантів v_i за кожним критерієм c_j . Загальна кількість таких матриць має збігатися з кількістю критеріїв і дорівнює m .

Після визначення всіх елементів a_{ij} матриці A ступеня приналежності, що необхідні для формування нечіткої множини, використовується формула:

$$\mu^l(v_i) = \frac{1}{a_{i1}^l + a_{i2}^l + \dots + a_{in}^l},$$

де a_{ij}^l – елемент матриці, що оцінюється експертом за 9-бальною шкалою Сааті.

Базуючись на принципі Беллмана–Заде, найкращою системою будемо вважати ту, яка одночасно найкраща за критеріями c_1, c_2, \dots, c_m . Тому нечітка

множина, що необхідно для рейтингового аналізу, визначається у вигляді перетину:

$$D = \left\{ \frac{\min[\mu^l(v_1)]}{v_1}, \frac{\min[\mu^l(v_2)]}{v_2}, \dots, \frac{\min[\mu^l(v_n)]}{v_n} \right\}.$$

Згідно з отриманою безліччю D найкращим треба вважати той варіант, для якого ступінь приналежності (чисельник) є найбільшим.

Перевагою методу мультикритеріального аналізу рішень можна вважати те, що метод одночасно забезпечує просту структуру ефективного прийняття рішень та подання припущень і висновків, і дає змогу вирішувати складні проблеми пошуку оптимального рішення. Метод дає можливість досягти компромісу в ситуації, коли причетні сторони мають різні цілі і, отже, критерії. Серед недоліків є те, що метод схильний до впливу упередженого та неповного вибору критеріїв для прийняття рішення. Крім того, більшість багатокритеріальних проблем не мають остаточного або однозначного рішення.

Сьогодні існує ціла низка академічних та комерційних комп'ютерних систем, що реалізують різні моделі і додаткові засоби для вирішення всіх основних завдань мультикритеріального аналізу рішень (багатокритеріальних задач вибору, ранжування, сортування та ін.).

До засобів структурування завдань, включно з засобами візуальної підтримки процесу формування ідей, концепцій, критеріїв і їх взаємозв'язків з використанням 2D та 3D-графіки, є прикладні програмні продукти Decision Explorer (www.banxia.com) та Mind Manager 4.0 (www.mind-map.com).

До програмних продуктів більш високого класу складності належать:

- Expert Choice (www.expertchoice.com);
- Criterium DecisionPlus (www.infoharvest.com);
- Decision Deck (www.decision-deck.org).

III Expert Choice є класичним варіантом системи, що реалізує стохастичну модель узагальненого методу аналітичної ієрархії. В III Expert Choice підтримуються методи структурування (дерево критеріїв) та аналіз чутливості результатів до зміни вагових коефіцієнтів. Розроблено версії веб-систем підтримки групового методу аналізу рішень (Team Expert Choice).

Criterium DecisionPlus – це зручний інструмент і для структурування багатокритеріальної задачі з використанням дерева критеріїв, і для реалізації подальшого процесу визначення випадкових значень критеріїв, що нормовані з використанням локальних функцій корисності, які потім агрегуються в інтегральне значення корисності альтернативи за допомогою обчислення очікуваної корисності. Також III Criterium DecisionPlus дає змогу використовувати стохас-

тичні моделі узагальненого методу аналітичної ієрархії, проводити аналіз невизначеності та чутливості до результатів ранжування альтернатив, до зміни вагових коефіцієнтів критеріїв. Для цього програмного продукту розроблено версії веб-системи підтримки групового методу аналізу рішень з обмеженими можливостями.

В рамках проекту **Decision Deck** розробляються стандарти і платформа для створення відкритого програмного забезпечення (Open Source software), що реалізує різні етапи розробки і використання методів мультикритеріального аналізу рішень.

Наявні програмні продукти, що реалізують різні моделі та методи мультикритеріального аналізу рішень, є ефективним інструментом для багатокритеріального аналізу альтернатив при вирішенні широкого кола науково-прикладних задач, та сприяють впровадженню сучасних методів бізнес-аналітики при обґрунтуванні та вибору рішень з урахуванням компромісів і узгодження різних поглядів зацікавлених сторін.

Більшість систем є занадто складними для дослідження впливу невизначеності з використанням аналітичних методів. Однак такі системи можна досліджувати, якщо розглядати вхідні дані у вигляді випадкових змінних, повторюючи велику кількість обчислень (ітерацій), для отримання результату з необхідною точністю. Такий підхід був розроблений у 1946 р. Станіславом Уламом та названий Монте-Карло.

Метод Монте-Карло – загальна назва групи числових методів, заснованих на отриманні великої кількості реалізацій стохастичного (випадкового) процесу, який формується так, щоб його ймовірнісні характеристики збігалися з аналогічними величинами задачі, яку потрібно розв'язати.

Метод Монте-Карло є способом оцінки впливу невизначеності оцінки параметрів системи в широкому діапазоні ситуацій. Метод зазвичай використовують для оцінки діапазону зміни результатів та відносної частоти значень у цьому діапазоні для кількісних величин, як-от вартість, тривалість, продуктивність, попит та ін. Моделювання методом Монте-Карло може бути використане для двох різних цілей:

- трансформування невизначеності для звичайних аналітичних моделей;
- розрахунку ймовірностей, якщо аналітичні методи не можуть бути використані.

Для використання методу Монте-Карло необхідно спочатку визначити модель, яка найбільш точно описує поведінку системи, що досліджується.

Далі проводиться багаторазове застосування моделі із використанням генератора випадкових чисел для отримання вихідних даних моделі. При цьому

використовуються випадкові числа, що мають відповідати таким характеристикам:

1. Якщо $\tilde{r}_i (i = 1, 2, \dots)$ є випадкові числа, то їх кумулятивний розподіл (позначимо F) задовольняє рівняння:

$$F(r_i) = P(\tilde{r}_i < r_i),$$

$$F(r_i) = \begin{cases} r_i & \text{для } 0 \leq r_i \leq 1, \\ 0 & \text{для } r_i < 0, \\ 1 & \text{для } r_i > 1. \end{cases}$$

2. Змінні $r_1, r_2, \dots, r_i, \dots, r_n$ – незалежні, якщо їх спільний кумулятивний розподіл (позначимо G) можна представити у вигляді добутку індивідуальних функцій розподілу:

$$G(r_1, r_2, \dots, r_n) = F_1(r_1) F_2(r_2) \dots F_n(r_n).$$

Для практичних цілей використовуються псевдовипадкові числа, які відповідають вказаним вище характеристикам, і генеруються за допомогою такого правила: числа можна вважати випадковими, рівномірно розподіленими в інтервалі $(0; 1)$ і незалежними у припущенні, що довжина циклу буде досить велика. За довжину циклу приймається число псевдовипадкових чисел, отриманих до того, як числа, що генеруються, почнуть повторюватися. Для генерації псевдовипадкових чисел застосовується така формула:

$$x_i = ax_{i-1} \pmod{m} \quad (i = 1, 2, \dots),$$

$$x_0 = b.$$

Цей метод називається *мультиплікативно-конгруентним, або методом степеневих відрахувань*. Рівняння використовує попереднє число x_{i-1} , перемножуючи його на константу a і ділячи результат на m , залишок береться як нове число x_i ; b слугує початковим значенням для x_i . Відповідне псевдовипадкове число r_i , яке має знаходитися в інтервалі $(0; 1)$, виходить від ділення x_i на m .

Перевага псевдовипадкових чисел над чисто випадковими числами полягає в тому, що їх можна генерувати за допомогою комп'ютерів, що виключає необхідність зберігання в пам'яті машини великих таблиць випадкових чисел і витрат часу на їх читання.

Після того, як псевдовипадкові числа були визначені, записується модель у формі рівняння, що виражає співвідношення між вхідними та вихідними параметрами. Значення, відібрані як вхідні дані, отримують, зважаючи на відповідні розподіли ймовірностей, що характеризують невизначеності даних. За допомо-

гою спеціальних прикладних програм багаторазово використовують модель (часто до 10 000 разів) з різними вхідними даними, і отримують вихідні дані, які можуть бути оброблені за допомогою статистичних методів для отримання оцінок середнього, стандартного відхилення, довірчих інтервалів.

Перевагами методу Монте-Карло є те, що він може бути адаптований до будь-якого розподілу вхідних даних, включно з емпіричними розподілами, що побудовані на основі спостережень за відповідними системами. Моделі відносно прості для роботи і можуть бути за необхідності розширені, а сам метод дає змогу врахувати будь-які дії і взаємозв'язки, включно з тонкими умовними залежностями.

Недоліки методу полягають у тому, що невизначеність даних можна описати відомим розподілом, а точність рішень залежить від кількості ітерацій, які можуть бути виконані (цей недолік стає менш значущим зі збільшенням швидкодії комп'ютера та відповідних програмних засобів).

Сьогодні існують сучасні програмні засоби, що задовольняють високі вимоги. Реалізується метод Монте-Карло, виділяють MS Excel, Statistica, GPSS, @RISK Palisade, Oracle Crystal Ball.

В MS Excel легко може бути здійснено моделювання методом Монте-Карло, оскільки MS Excel надає широкі математичні можливості, зокрема й можливості використання функції розподілу випадкової величини. В MS Excel для побудови вибіркової функції розподілу використовуються спеціальна функція ЧАСТОТА і процедура «Гістограма» з пакета аналізу.

Модуль «Моделювання структурними рівняннями (SEPATH)» системи Statistica містить потужні засоби імітаційного моделювання методом Монте-Карло: є можливість породжувати (і зберігати) набори даних для зумовлених моделей, заснованих на нормальному розподілі або на скошених розподілах. За допомогою методу Монте-Карло можна обчислювати розподіл різних інформаційних статистик, оцінки параметрів і т. д. Для візуалізації результатів прогонів методу Монте-Карло (наприклад, розподілів параметрів) є різні графічні засоби.

Система GPSS широко використовується для вирішення практичних завдань. Динамічним елементом моделі є транзакт – абстрактний об'єкт, який переміщується між статичними елементами, відтворюючи різні події реального об'єкта, що моделюється. Під час роботи моделі накопичується статистика, що автоматично виводиться по завершенні процесу моделювання. Модель являє собою програму, а значить, не має графічної інтерпретації, що ускладнює процес розробки моделі методом Монте-Карло.

Програмний комплекс @RISK Palisade виконує аналіз ризиків з використанням моделювання за методом Монте-Карло, щоб продемонструвати якомога більше результатів у моделі на базі електронної таблиці, показуючи ймовірність

кожного з них. Сутність програмного комплексу полягає в тому, що невизначені значення в електронній таблиці замінюються функціями розподілу ймовірності (будь-якими з наявних 65 функцій). @RISK передбачає широкий набір графіків для інтерпретації та представлення отриманих результатів іншим користувачам.

Oracle Crystal Ball – це один з найпоширеніших табличних додатків, який надає можливості аналізу критичних чинників, що впливають на рівень ризику. Алгоритм його роботи аналогічний попередньому програмному комплексу.

Отже, наявність різнопланових за своєю структурою та можливостями інструментів аналізу ризиків (прикладних програмних продуктів та веб-платформ), у яких реалізовані методи мультикритеріального аналізу рішень та Монте-Карло, сприяють зростанню культури і практичних навичок опрацювання рішень в умовах ризиків та невизначеності на основі методології бізнес-аналітики.

На прикладі будівництва заводу, що був розглянутий в попередньому пункті, **застосуємо інструменти аналізу ризиків методом Монте-Карло.**

Відповідно до послідовності кроків, що вимагає метод Монте-Карло, необхідно спочатку визначити модель, у якій він буде реалізовуватися. За модель візьмемо розрахунок, що наведено на рис. 2.1.

Далі необхідно визначити функції розподілу для кожного фактора ризику та невизначеності.

Для моделювання ціни продажу (у 1-й, 2-й та 3-й роки окремо) використовується рівномірний розподіл. Ціна продажу в 1-й рік має мінімальне значення 5,9 грн, максимальне – 6,1 грн. Аналогічно ціна продажу для 2-го року має рівномірний розподіл із параметрами 5,95 грн і 6,15 грн, у 3-й рік – 6,0 грн та 6,2 грн.

На відміну від ціни, яка хоч і коливається, але перебуває під контролем менеджерів компанії, обсяг продажів залежить від неконтрольованих чинників. Тому використовується нормальний розподіл. Обсяг продажів у 1-й рік має нормальний розподіл із середнім значенням (математичним очікуванням) 8 020 тис. упаковок і стандартним відхиленням 250 тис. упаковок. Для 2-го і 3-го року є такі пари значень: 9 670 тис./300 тис. та 11 320 тис./250 тис. упаковок відповідно.

Собівартість (як відсоток від продажів) має рівномірний розподіл із мінімальним значенням 35 % і максимальним значенням 55 %.

Для оцінки валютного ризику та ризику зупинки будівництва з причини блокування проекту з боку громад доцільно використовувати дискретний розподіл. Завдання аварій потребує функції Пуассона.

Для багаторазового застосування моделі з використанням генератора випадкових чисел можна використовувати або MS Excel, або Oracle Crystal Ball.

Розглянемо можливості кожного з запропонованих програмних інструментів з аналізу ризиків.

Для використання методу Монте-Карло в *MS Excel* необхідно організувати попередню розрахункову таблицю (рис. 2.1) так (рис. 2.2):

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
1	№	1 рік	2 рік	3 рік	1 рік	2 рік	3 рік	1 рік	2 рік	3 рік	1 рік	2 рік	3 рік	1 рік	2 рік	3 рік	1 рік	2 рік	3 рік	1 рік	2 рік	3
2	ексер.	Ціна упаковки			Обсяг продажів			Виручка			Собівартість			Валовий прибуток			Операційні витрати			Собівартість		
3	1																					
4	2																					
5	3																					
6	4																					
7	5																					

Рисунок 2.2 – Вихідна таблиця для розрахунку показника NPV методом Монте-Карло для проекту будівництва заводу в ПП MS Excel

Діапазон A3:A103 (рис. 2.2) необхідно заповнити послідовністю чисел від 1 до 100 (кількість експериментів). За допомогою функції «Генерація випадкових чисел» (Дані → Аналіз даних) згенерувати випадкові числа для кожного чинника невизначеності (рис. 2.3).

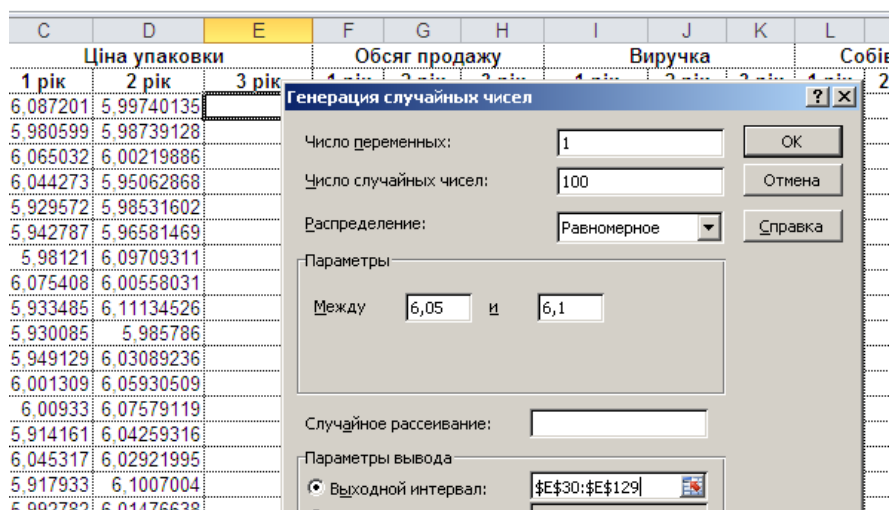


Рисунок 2.3 – Генерація випадкових чисел показника «ціна» за рівномірним розподілом для розрахунку показника NPV методом Монте-Карло для проекту будівництва заводу в ПП MS Excel

Дискретний розподіл для оцінки валютного ризику використовується в такий спосіб (рис. 2.4):

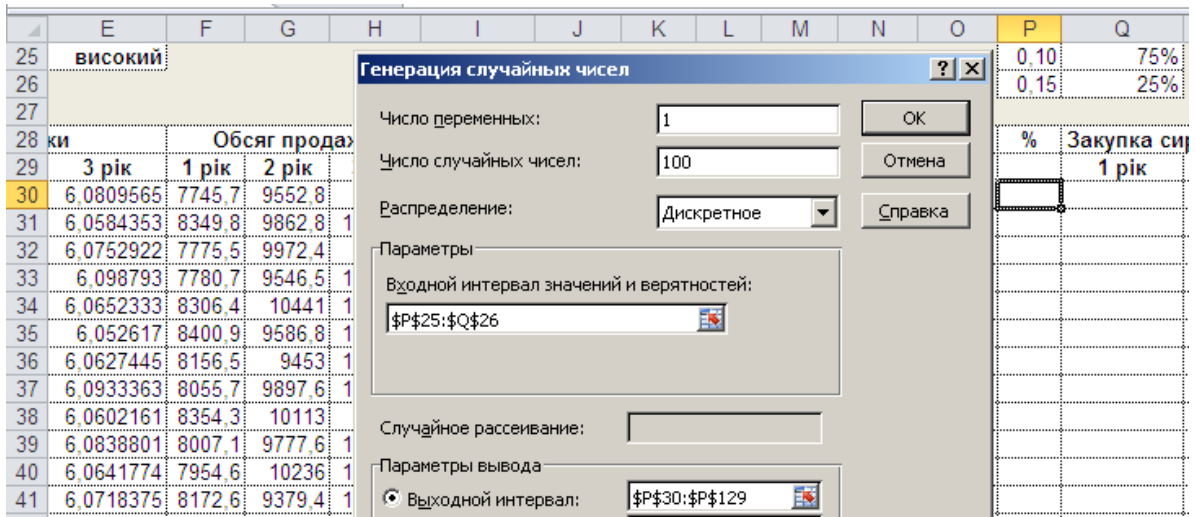


Рисунок 2.4 – Генерація випадкових чисел показника «закупка сировини за кордоном» за дискретним розподілом для розрахунку показника NPV методом Монте-Карло для проекту будівництва заводу в ПП MS Excel

В аналогічний спосіб задайте всі інші функції розподілу в ПП MS Excel.

Останнім кроком є розрахунок кількості показників NPV, що прийняли значення менше 0. Відношення загальної кількості експериментів (100) до кількості експериментів, де $NPV < 0$, є ймовірністю отримання негативного NPV.

За результатами розрахунку необхідно розрахувати середнє значення, стандартне відхилення та коефіцієнт варіації для чинників невизначеності (ціна упаковки, обсяг продажу, собівартість) а також для NPV. Для отриманих показників NPV розрахуйте ймовірність того, що проект може бути збитковим, та побудуйте такий графік (рис. 2.5):

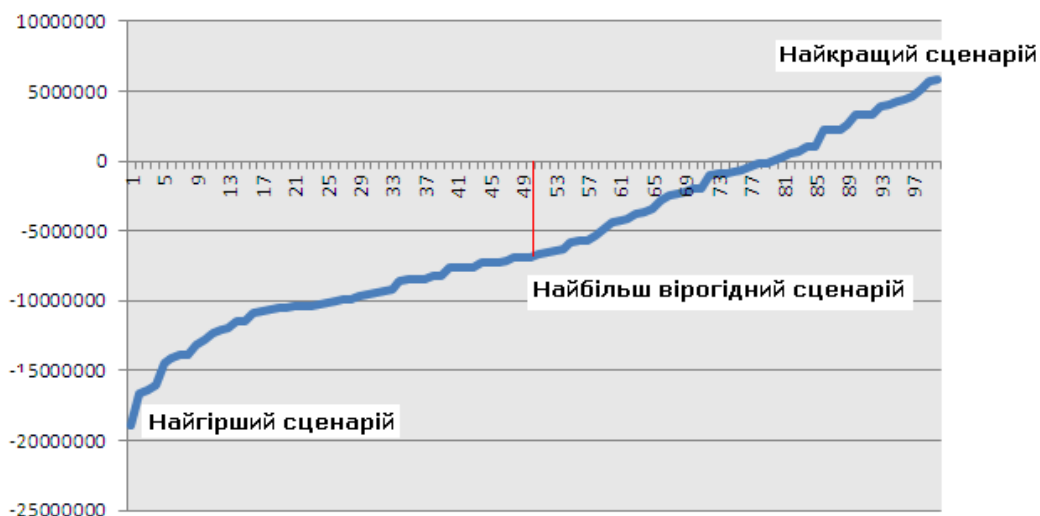


Рисунок 2.5 – Розподіл показника NPV методом Монте-Карло для проекту будівництва заводу в ПП MS Excel

Висновком із цього графіка (рис. 2.5) є те, що найімовірніше, *NPV* буде дорівнювати приблизно – 7 000 000 грн, що є неприпустимим для інвестора. Без врахування факторів ризиків *NPV* прогнозувався на рівні 1 158 123 грн. Відповідно до розрахунків методом Монте-Карло, в найкращому випадку *NPV* може досягти 5 000 000 грн, а позитивна ефективність проекту можлива лише у 23 % випадків (100–77).

Для використання методу Монте-Карло в *Oracle Crystal Ball* необхідно з електронної адреси офіційного сайту завантажити демоверсію програмного продукту *Oracle Crystal Ball* (<https://www.oracle.com>).

Перш ніж запустити програму, необхідно закрити всі вікна Excel та запустити *Crystal Ball*.

Першим кроком є задання параметрів розподілу незалежних змінних. У першій розрахунковій таблиці (рис. 2.1) зробіть розрахунки за допомогою вбудованих функцій із галереї зображень розподілу. Для цього поставте в комірку C2 курсор і на вкладці *Crystal Ball* клацніть «Определение допущения». У вікні виберіть «Равномерное» і натисніть *Ok* (рис. 2.6).

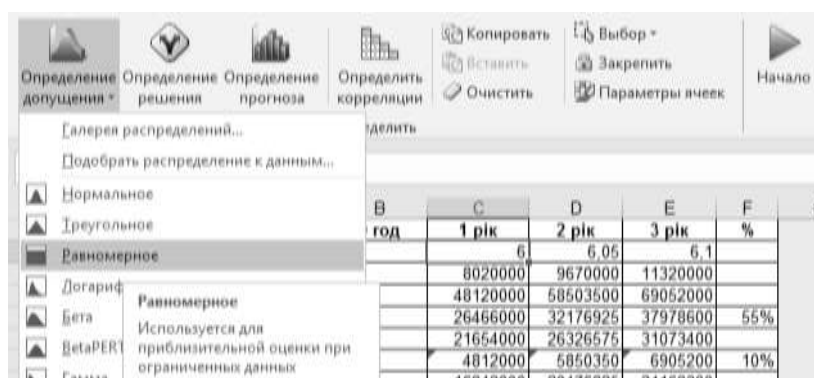


Рисунок 2.6 – Генерація випадкових чисел показника «ціна» за рівномірним розподілом для розрахунку показника *NPV* методом Монте-Карло для проекту будівництва заводу в *Oracle Crystal Ball*

Для моделювання ціни продажу використовується рівномірний розподіл із такими ж параметрами, як у попередньому прикладі (рис. 2.7).

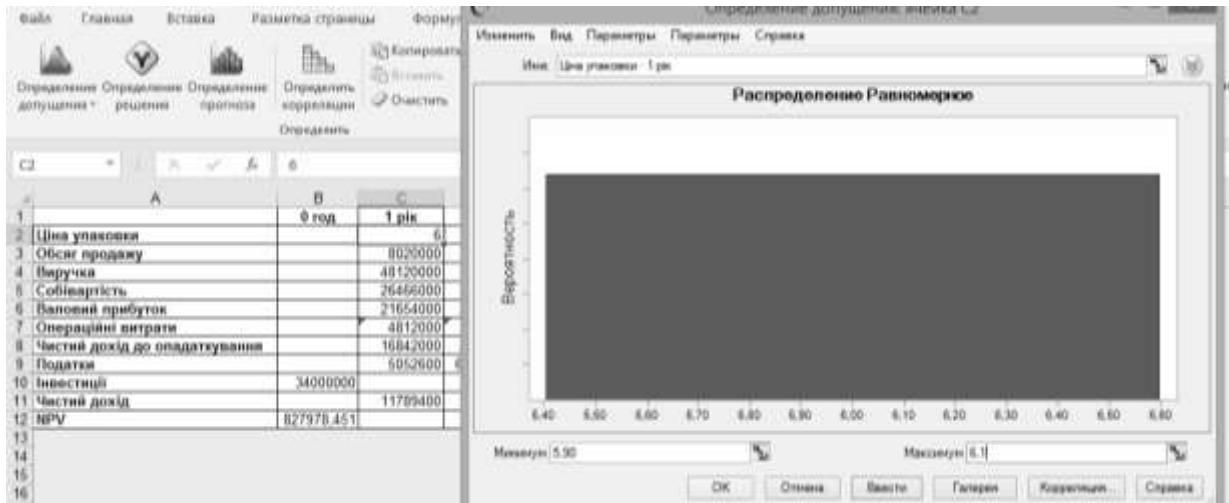


Рисунок 2.7 – Параметри рівномірного розподілу показника «ціна» в Oracle Crystal Ball

В аналогічний спосіб задайте функції розподілу для показника «обсяг продажів» та «собівартість».

Після задання параметрів комірки фарбуються в зелений колір.

Уже на цьому етапі можна переглянути розподіл показника *NPV*. Для цього необхідно обрати його як залежну змінну. Щоб вибрати залежну змінну, необхідно вставити курсор у комірку *B12*, що містить формулу розрахунку *NPV*, та клацніть «Определение прогноза» (рис. 2.8).

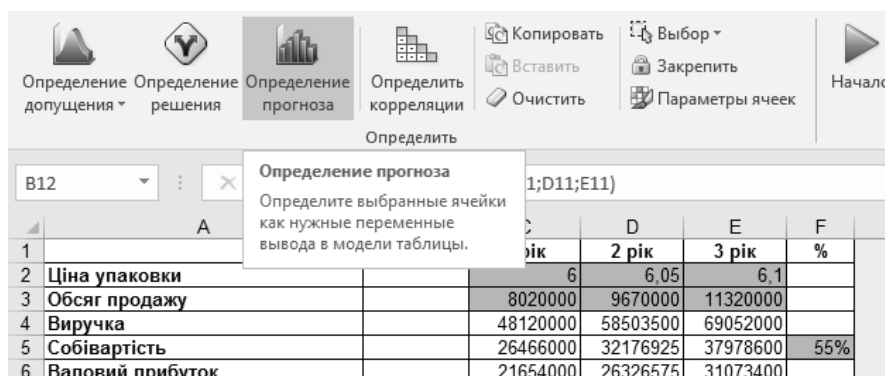


Рисунок 2.8 – Визначення залежної змінної в Oracle Crystal Ball

Щоб запустити процес моделювання, оберіть кількість ітерацій та клацніть на кнопку «Начало». Після заданої кількості ітерацій (5 000) програма виведе результати у графічному вигляді. Також можна створити звіт про моделювання (в окремому файлі Excel), де буде виведений графік розподілу показника *NPV* та відповідна статистика (рис. 2.9).

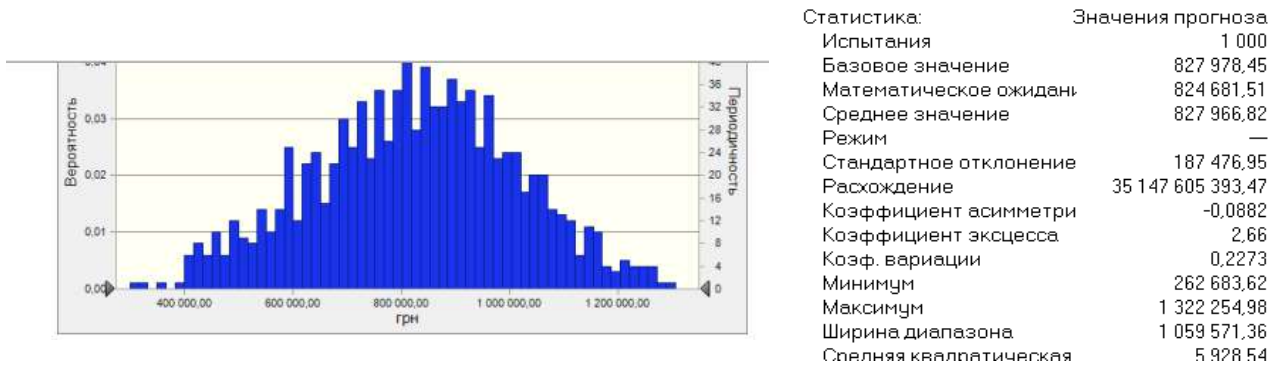


Рисунок 2.9 – Розподіл показника NPV методом Монте-Карло для проекту будівництва заводу в ПП Oracle Crystal Ball

Отже, аналіз розподілу показника NPV методом Монте-Карло з використанням відповідного інструментарію довів необхідність розробки заходів з управління ризиками.

Завдання для вивчення теми

I. Дайте коротку відповідь на питання:

1. З якою метою проводиться мультикритеріальний аналіз рішень?
2. Скільки якісних ознак може визначити фахівець при використанні методу мультикритеріального аналізу рішень?
3. Чи дає змогу метод мультикритеріального аналізу рішень досягти компромісу в ситуації, коли причетні сторони мають різні цілі?
4. Назвіть відомі програмні продукти, що реалізують моделі вирішення основних завдань мультикритеріального аналізу рішень.
5. Які програмні продукти дають змогу використовувати стохастичні моделі узагальненого методу аналітичної ієрархії?
6. Для яких цілей може застосовуватися моделювання методом Монте-Карло?
7. Як має записуватися модель, щоб її можна було використовувати для моделювання методом Монте-Карло?
8. У чому полягають переваги методу Монте-Карло?
9. Як у ПП MS Excel легко може бути здійснено моделювання методом Монте-Карло?
10. У чому полягають основні недоліки використання методу Монте-Карло в системі GPSS?

II. Завдання

Для проекту будівництва заводу в ПП Oracle Crystal Ball задайте функції розподілу для ризику зупинки будівництва з причини блокування проекту з боку громад та аварії. За результатами моделювання створіть звіт (в окремому файлі Excel), де буде виведений графік розподілу показника *NPV* та відповідна статистика.

Порівняйте результати, що були отримані в ПП MS Excel, з кількістю ітерацій 100 та в ПП Oracle Crystal Ball з кількістю ітерацій 5 000.

2.3. МЕТОДИ БІЗНЕС-АНАЛІТИКИ В УПРАВЛІННІ РИЗИКАМИ

Управління ризиками підприємства розуміється як система принципів та методів розробки і реалізації ризикових господарських рішень, що забезпечують всебічну оцінку різних видів ризиків і нейтралізацію їх можливих негативних наслідків. Чим складніша організаційна структура підприємства, тим вищі вимоги до процесу управління ризиками та актуальніше завдання підвищення його якості за рахунок ефективного використання сучасних методів бізнес-аналітики.

Управління ризиком можна здійснювати шляхом його уникнення або шляхом запобігання.

Уникнення ризику – це метод, який полягає в розробці заходів, які повністю виключають конкретний вид ризику, наприклад:

- відмова від здійснення операцій, рівень ризику за якими надмірно високий. Необхідно зазначити, що використання цього заходу має обмежений характер, оскільки більшість операцій підприємства пов'язані зі здійсненням основної виробничо-комерційної діяльності, що забезпечує регулярне надходження доходів та формування прибутку;

- відмова від використання в високих обсягах позикового капіталу. Зниження частки позикових фінансових коштів у господарському обороті дає змогу уникнути втрати фінансової стійкості підприємства. Водночас таке уникнення ризику тягне за собою зниження можливості отримання додаткової суми прибутку на вкладений капітал;

- відмова від надмірного використання оборотних активів у низьколіквідних формах. Підвищення рівня ліквідності активів дає змогу уникнути ризику неплатоспроможності підприємства в майбутньому періоді. Однак це позбавляє підприємство додаткових доходів від розширення обсягів продажу продукції в кредит і породжує нові ризики, пов'язані з порушенням ритмічності операцій-

ного процесу через зниження розміру страхових запасів сировини, матеріалів, готової продукції;

- відмова від використання тимчасово вільних грошових активів у короткострокові фінансові вкладення. Цей захід дає змогу уникнути депозитного і процентного ризику, проте породжує інфляційний ризик, а також ризик упущеної вигоди.

Форми уникнення ризику позбавляють підприємство додаткових джерел формування прибутку, та відповідно, негативно впливають на темпи його економічного розвитку та ефективність використання власного капіталу.

Тому в системі внутрішніх механізмів підприємства нейтралізації ризиків їх уникнення має здійснюватися дуже виважено у разі, якщо відмова від одного ризику не тягне виникнення іншого ризику вищого чи однозначного рівня або якщо втрати з цього виду ризику перевищують можливості їх відшкодування за рахунок власних фінансових коштів підприємства. Якщо розмір доходу від операції, яка генерує певні види ризику, є несуттєвим, тобто займає невідчутну питому вагу у формованому позитивному грошовому потоці підприємства, або операції, не характерні для діяльності підприємства, мають інноваційний характер і за ними відсутня інформаційна база, що необхідна для визначення рівня ризиків і прийняття відповідних управлінських рішень, то ризиків також доцільно уникати.

При управлінні ризиками шляхом запобігання існує можливість застосовувати великий набір методів, до яких належать:

- запобігання збиткам;
- розподіл ризику;
- прийняття ризику;
- передача ризику.

Методи запобігання збиткам є методи, які дають можливість завчасно запобігти виникненню потенційних збитків. До них відносяться: попередження ризиків, мінімізація втрат та пошук інформації.

Попередження ризиків (збитків) означає можливість уберегтися від випадковостей за допомогою конкретного набору превентивних дій. До найефективніших методів попередження збитків належать розробка та впровадження програми превентивних заходів, прогнозування зовнішньої економічної обстановки, метод стратегічного планування, постійний контроль і аналіз інформації, навчання і тренінг персоналу, консалтинг.

Метод мінімізації втрат означає зменшення значної частини збитків інвестора, коли ризику уникнути неможливо або він виникає незаплановано. Методами мінімізації втрат є диверсифікація та лімітування.

Диверсифікація є процесом розподілу капіталу між різними об'єктами вкладення, які безпосередньо не пов'язані між собою, що є найбільш обґрунтованим і менш витратним способом зниження ступеня ризику. Зазвичай диверсифікація використовується для нейтралізації негативних наслідків специфічних видів ризиків.

Лімітування концентрації ризику – це встановлення ліміту. Цей метод використовується за тими видами ризиків, які виходять за межі їх допустимого рівня, тобто за операціями, які здійснюються в зоні критичного чи катастрофічного ризику. Лімітування реалізується шляхом встановлення на підприємстві відповідних внутрішніх нормативів у процесі розробки фінансової політики. Ця система нормативів може включати:

- граничний розмір (питома вага) позикових коштів, що використовуються в господарській діяльності. Цей ліміт встановлюється окремо для операційної та інвестиційної діяльності підприємства, а в низці випадків і для окремих операцій (фінансування реального інвестиційного проекту; фінансування формування оборотних активів тощо);

- мінімальний розмір (питома вага) активів у високоліквідній формі. Цей ліміт забезпечує формування «ліквідної подушки», що характеризує розмір резервування високоліквідних активів з метою майбутнього погашення невідкладних фінансових зобов'язань підприємства. «Ліквідною подушкою» насамперед є короткострокові фінансові вкладення підприємства, а також його короткострокова дебіторська заборгованість;

- максимальний розмір товарного (комерційного) або споживчого кредиту, наданого одному покупцеві. Розмір кредитного ліміту встановлюється при формуванні кредитної політики підприємства;

- максимальний розмір депозитного вкладу, що розміщується в одному банку. Лімітування депозитного ризику здійснюється під час використання цього фінансового інструменту інвестування капіталу підприємства;

- максимальний розмір вкладення коштів у цінні папери одного емітента. Ця форма лімітування направлена на зниження концентрації несистематичного (специфічного) ризику при формуванні портфеля цінних паперів;

- максимальний період відволікання коштів у дебіторську заборгованість. За рахунок цього нормативу забезпечується лімітування ризику неплатоспроможності, інфляційного ризику, а також кредитного ризику.

Пошук інформації – це метод, спрямований на зниження ризику шляхом знаходження і використання необхідної інформації для прийняття ризикового рішення. До методів пошуку інформації належать: метод моніторингу нормативно-правового, соціально-економічного, технічного, психологічного середовища.

Методи розподілу (дисипації) ризику – це методи, за яких ризик можливої шкоди або втрат ділиться між учасниками так, що можливі втрати кожного невеликі.

Методи прийняття ризику (поглинання ризику, утримання ризику) являють собою методи добровільного або вимушеного, свідомого чи незапланованого залишення на своїй відповідальності ризиків. До них належать локалізація та самострахування.

Методи локалізації ризику включають у себе створення венчурних підприємств (венчурів) та створення спеціальних структурних підрозділів.

Самострахування означає компенсацію збитків самостійно за рахунок використання поточного прибутку, нерозподіленого прибутку минулих років, коштів від розрахунків за дивідендами, створення системи спеціальних резервних (ризикових) фондів тощо.

Методи передачі ризику – це методи трансферту ризику іншим економічним суб'єктам. До них належать хеджування, страхування та передача контролю ризику.

Хеджування – це система укладання термінових контрактів (ф'ючерсних, форвардних контрактів і опціонів) і угод, що враховує ймовірні в майбутньому зміни цін, курсів і має на меті уникнути несприятливих наслідків цих змін.

Страхування – це метод передачі ризику страховій компанії, яка відшкодує різноманітні збитки, викликані несприятливими подіями, шляхом виплати страхового відшкодування і страхових сум. За ці послуги за договором страхування вона отримує від інвестора страхову премію.

При передачі контролю ризику інвестор може делегувати повноваження контролю ризику іншій особі (групі осіб) шляхом передачі: реальної власності, напрямів діяльності, пов'язаної з ризиком, або відповідальності за ризик.

Отже, для ефективного управління всім різноманіттям ризиків у діяльності підприємства необхідно застосовувати цілий комплекс методів бізнес-аналітики, які сприяють більш ефективній адаптації системи управління ризиком підприємства до зміни умов функціонування навколишнього середовища.

На прикладі будівництва заводу, що розглядається в кожному пункті цього розділу, розробимо заходи з управління ризиками.

Для визначення найбільш загрозливих видів ризиків необхідно заповнити таблицю.

Таблиця 2.1 – Таблиця наслідків та ймовірності реалізації ризиків

№	Ризик	Фактор ризику	Ймовірність реалізації ризику	Збиток від реалізації ризиків
1	Валютний ризик	Збільшення частки собівартості продукту з причини збільшення курсу на покупку долара США (10 % від собівартості оплачується у валюті)	висока	високий
2	...			

Ймовірність реалізації ризику може бути надано експертно (висока, середня, низька), або ж з огляду на кількісну оцінку: низький рівень імовірності – до 30 %, середній – від 30 % до 70 %, високий – вище 70 %.

Щоб розрахувати збиток від реалізації ризику, можна надати найгірше значення показнику, що його характеризує. Якщо негативне відхилення NPV коливається в межах від 0 % до 25 % від початкового значення, то можна зробити висновок, що збиток від реалізації ризиків – низький; якщо від 25 % до 80 %, то середній; якщо перевищує 80 %, то високий.

Наприклад, для валютного ризику розрахунок збитку буде мати такий вигляд (рис. 2.10):

C6		fx =C2*\$F\$6*C3*1,5				
	A	B	C	D	E	F
1		0 рік	1 рік	2 рік	3 рік	%
2	Ціна упаковки		6	6,05	6,1	
3	Обсяг продажу		8020000	9670000	11320000	
4	Виручка		48120000	58503500	69052000	
5	Собівартість		21654000	26326575	31073400	45%
6	Закупка сировини за кордоном		7218000	8775525	10357800	10%
7	Валовий прибуток		19248000	23401400	27620800	
8	Операційні витрати		4812000	5850350	6905200	10%
9	Чистий дохід до опадаткування		14436000	17551050	20715600	
10	Податки		4330800	5265315	6214680	30%
11	Інвестиції	34000000				
12	Чистий дохід		10105200	12285735	14500920	
13	NPV	-3422908				
14						
15	NPV	1158122,6				
16	Δ	-396%				
17	Збитки	високий				

Рисунок 2.10 – Розрахунок збитку від реалізації валютного ризику для проекту будівництва заводу

Результат розрахунку ввести в матрицю (рис.2. 11):

Збитки	Високі			Валютний ризик
	Середні			
	Низькі			
		Низька	Середня	Висока
		Ймовірність		

Рисунок 2.11 – Матриця ризиків для проекту будівництва заводу

На прикладі валютного ризику розглянемо схему хеджування валютних ризиків за допомогою опціону. Інвестиційна компанія купує валютний опціон, який надає право (але не обов'язок) купити певну кількість валюти за фіксованим курсом у певний день, тобто у компанії існує можливість купити долари США за курсом 20 грн за 1 доллар. Опціонна премія складає 3,5 %, ціна цього опціону наведена на рис. 2.12:

B14		fx =ЧПС(0,1;-B12;C13;D13;E13)				
	A	B	C	D	E	F
1		0 рік	1 рік	2 рік	3 рік	%
2	Ціна упаковки		6	6.05	6.1	
3	Обсяг продажу		8020000	9670000	11320000	
4	Виручка		48120000	58503500	69052000	
5	Собівартість		21654000	26326575	31073400	45%
6	Закупка сировини за кордоном		4812000	5850350	6905200	10%
7	Ціна опціону		4980420	5850350	6905200	3.5%
8	Валовий прибуток		21654000	26326575	31073400	
9	Операційні витрати		4812000	5850350	6905200	10%
10	Чистий дохід до опадаткування		16842000	20476225	24168200	
11	Податки		5052600	6142867.5	7250460	30%
12	Інвестиції	34000000				
13	Чистий дохід		11789400	14333357.5	16917740	
14	NPV	1.158.122,57р.				

Рисунок 2.12 – Розрахунок хеджування валютних ризиків за допомогою опціону для проекту будівництва заводу

Як показали розрахунки (рис. 2.12), в цьому разі ймовірність негативного NPV зменшується на 3 % при тому, що інші види ризиків залишилися на тому ж рівні.

Результати розрахунку можна подати у вигляді таблиці.

Таблиця 2.2 – Таблиця заходів з управління ризиками

№	Ризик	Фактор ризику	Заходи з управління ризиком	Ціна заходу	Зменшення збитку від реалізації ризику
1	Валютний ризик	Збільшення частки собівартості продукту з причини збільшення курсу на покупку долара США (10 % від собівартості продукції)	Покупка валютного опціону	614 864 тис. грн	3 %

Отже, хеджування, як один із заходів управління валютним ризиком, дає змогу зменшити збитки від можливої реалізації зазначеного ризику на 3 %.

Завдання для вивчення теми

I. Дайте коротку відповідь на питання:

1. Як розуміється управління ризиками підприємства?
2. Назвіть два основні шляхи управління ризиком.
3. Як називається метод, який полягає в розробці заходів, що повністю виключають конкретний вид ризику?
4. Які дії дають змогу уникнути ризику неплатоспроможності підприємства в майбутньому періоді?
5. Які є недоліки та переваги в відмові від використання тимчасово вільних грошових активів у короткострокові фінансові вкладення?
6. Назвіть методи, які дають можливість завчасно запобігти виникненню потенційних збитків.
7. Який спосіб розподілу капіталу є найбільш обґрунтованим і менш витратним способом зниження ступеня ризику?
8. Який метод використовується за тими видами ризиків, які виходять за межі їх допустимого рівня?
9. Які існують методи локалізації ризику?
10. Як називається метод передачі ризику страховій компанії?

II. Завдання

Для проекту будівництва заводу за попередніми даними, що були наведені в п. 2.1, заповніть таблицю наслідків та ймовірності реалізації (табл. 2.1) для кожного виду ризику (зміна ціни реалізації, обсягів продажу, собівартості, соціальні ризики, що пов'язані з блокуванням проекту з боку громад та аварії під час заміни обладнання), матрицю ризиків, проведіть відповідні розрахунки та заповніть таблицю заходів з управління ризиками (табл. 2.2).

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ДО РОЗДІЛУ 2

1. A Guide to the Business Analyses Body of Knowledge (BABOC) Version 3. URL: <http://www.iiba.org/babok-guide/core-standard.aspx>
2. ISO 31000 «Менеджмент рисков. Принципы и руководящие указания». URL: <https://intercert.com.ua/articles/regulatory-documents/311-iso-31000-risk-management>
3. IEC 31010 «Менеджмент рисков. Методы оценки рисков». URL: <https://intercert.com.ua/articles/regulatory-documents/311-iso-31000-risk-management>
4. Інтегрована система управління ризиками банку: монографія / Н. П. Шульга, В. І. Міщенко, Л. Л. Анісімова та ін.; за заг. ред. Н. П. Шульги. К.: Київ. нац. торг.-екон. ун-т, 2018. 440 с.
5. Дэвенпорт Т. Харрис Дж. Аналитика как конкурентное преимущество. Новая наука побеждает. СПб: BestBusinessBooks, 2010. 264 с.
6. Карлберг К. Бизнес-анализ с использованием Excel. М: Диалектика Вильямс, 2017. 576 с.
7. Charnes J. Financial Modeling with Crystal Ball and Excel. New Jersey: John Wiley & Sons, Inc., 2007. 290 с.

РОЗДІЛ 3 БІЗНЕС-АНАЛІТИКА В СЕРЕДОВИЩІ R

3.1. МОВА ПРОГРАМУВАННЯ R ТА СЕРЕДОВИЩЕ R STUDIO

Необхідною умовою сучасного статистичного аналізу даних як основного процесу бізнес-аналітики є ефективне використання комп'ютерних програм, від функціональної повноти й алгоритмічної продуманості яких залежить підсумкова інтерпретація результатів дослідження та надійність висновків. Однією з найзручніших для досягнення цієї мети є вільно розповсюджуване статистичне середовище R, яке є найбільш повним, надійним, динамічно розвивається та об'єднує мову програмування високого рівня і потужні бібліотеки програмних модулів для обчислювальної та графічної обробки даних.

Сьогодні статистичне середовище R (<http://www.r-project.org>) є безумовним лідером серед некомерційних систем статистичного аналізу і поступово стає незамінним при проведенні аналітичних розрахунків у більшості західних університетських центрів і багатьох провідних фірмах.

R – це універсальна мова програмування, розроблена для застосування в галузях сучасної бізнес-аналітики з аналізом даних, класичної статистики і якісної модерної графіки. Завдяки своїй великій бібліотеці пакетів, що безперервно розширюється, мова підтримує широкий спектр статистичних та чисельних методів.

R широко використовується як аналітичне програмне забезпечення для аналізу даних і фактично стала стандартом для різних статистичних програм. У 2010 р. R увійшла до списку переможців конкурсу журналу Infoworld у номінації на краще відкрите програмне забезпечення для розробки статистичних моделей. Ще однією особливістю R є графічні можливості, які полягають у здатності створення якісної графіки. Для зручності роботи з R розроблена низка графічних інтерфейсів, зокрема R Studio. R і додаткові пакети поширюються через CRAN (акронім *Comprehensive R Archive Network*).

Переваги R

R – це потужна скриптова мова з притаманними їй швидкою дією, гнучкістю та кросплатформеністю.

R – лідер напрямку. Чимало нових розробок у сфері статистики та аналітики спочатку з'являються як пакети для платформи R («R-пакети») і тільки потім переходять на комерційні платформи.

Інтеграція з різними форматами даних, зокрема з засобами публікації документів. R органічно використовує вхідні дані в різних форматах, а також інтегрується з системами публікації документів, що дає змогу вбудовувати статистичні результати і графіки з середовища R у звітні документи та презентації.

R – безкоштовний програмний засіб із відкритим кодом.

У *R* наявні численні пакети для застосування у різних галузях знань для моделювання, статистичного аналізу та обробки даних, які розширюють базовий функціонал.

Інсталяція *R*:

- перейдіть за посиланням <https://cloud.r-project.org/>;
- оберіть свою операційну систему;
- завантажте відповідний пакунок;
- інсталюйте його.

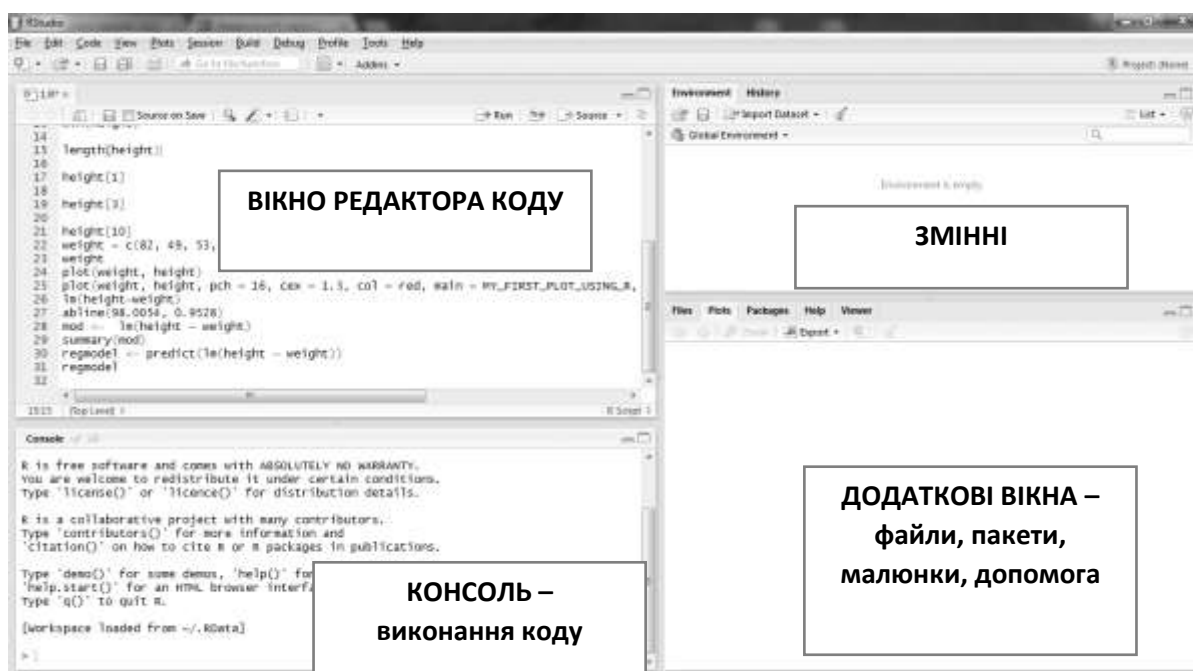
Інсталяція *R Studio*:

- перейдіть за посиланням <https://www.rstudio.com/products/rstudio/download/>;
- оберіть свою операційну систему;
- завантажте відповідний пакунок;
- інсталюйте його.

Створення нового файлу в *RStudio*:

- в меню оберіть пункт меню File;
- далі New File;
- оберіть пункт R Script.

Інтерфейс *R Studio*



*Рисунок 3.1 – Інтерфейс *R Studio*
Структура вікон *R Studio* (рис. 3.1):*

1. Вікно редактора коду (Source) – містить безпосередньо команди для виконання.
2. Консоль R Studio (Console) надає цілу низку опцій, які здійснюють роботу з R просто і продуктивно.
3. Вікно змінних (Environment) – містить усі введені в проєкті змінні, а також їх розмірності та значення.
4. Додаткові вікна (Files, Plots, Packages, Help, Viewer) – файли (всі файли проєкту), пакети (які додатково підключаються до проєкту), малюнки (побудовані графіки та діаграми проєкту), допомога (довідкова система).

Виконання коду

1. R Studio підтримує виконання коду безпосередньо з вікна Редактора (виконувані команди надсилаються в Консоль, де з'являється також результат їх виконання). Для виконання поточного рядка коду можна скористатися поєднанням клавіш Ctrl + Enter або кнопкою Run Line (s), розташованою у верхній частині вікна Редактора. Після виконання коду курсор автоматично перейде на наступний рядок. Такий механізм забезпечує можливість покрокового виконання команд.
2. Є дві можливості для одночасного виконання кількох рядків коду:
 - Виділити необхідні рядки і натиснути клавіші Ctrl + Enter (або скористатися кнопкою Run Line (s), розташованою у верхній частині вікна Редактора);
 - Натиснути Ctrl + Shift + Enter для виконання всіх рядків, наявних у скрипті.
3. Також код можна набирати в консолі (нижня ліва панель в R Studio).

Таблиця 3.1 – Основні дії в R

Дія в R	Назва дії	Приклад
1	2	3
<i>Арифметичні дії</i>		
+	додавання, сума	$5+3=8$
-	віднімання, різниця	$5-3=2$
*	множення	$5*3=15$
/	ділення, відношення	$6/2=3$
^ або **	піднесення до степеня	$5^2 = 25$ або $5**2 = 25$
abs(n)	модуль	$ -5 =5$
<i>Логічні дії</i>		
<	менше	$3<5$
<=	менше або дорівнює	$3\leq 5$
>	більше	$5>3$
>=	більше або дорівнює	$5\geq 3$
==	перевірка на рівність	якщо $x==y$, то ...

Закінчення таблиці

1	2	3
!=	не дорівнює	5≠3
!x	не x (для випадкових подій або умов)	якщо не виконується умова, то ...
x y	x або y (для випадкових подій або умов)	якщо виконується умова 1 або умова 2, то ...
x & y	x та y (для випадкових подій або умов)	якщо виконується умова 1 та умова 2, то ...
TRUE або T	істина, правда (завжди правильно)	
FALSE або F	хибність, помилковість, неправда	

3.2. ОПИСОВА СТАТИСТИКА ЗАСОБАМИ R

Основні функції описової статистики та їх реалізація засобами R наведені у таблиці 3.2.

Таблиця 3.2 – Функції описової статистики

Функція в R	Значення функції
mean(x)	Середнє значення (середнє арифметичне)
sum(x)	Сума значень
var(x)	Вибіркова дисперсія
sd(x)	Вибіркове середньоквадратичне відхилення
length(a)	Обсяг вибірки
median(x)	Медіана
sort(x)	Значення варіанта за зростанням
sort(unique(x))	Унікальні значення варіанта за зростанням
x.t<-table(x)	Дискретний закон розподілу
sort(unique(x))[which.max(x.t)]	Мода
summary(a)	Підсумкова статистика
min(x), max(x)	Мінімальний, максимальний елементи вибірки
range(x)	Розмах вибірки
cor(x,y)	Лінійний коефіцієнт кореляції

Типи даних

Типи даних в порядку збільшення пріоритету:



Логічні (logical)
 Цілочисельні (integer)
 Дійсні числа (numeric)
 Комплексні числа (complex)
 Текстові (character)
 Списки (list)

Способи створення векторів

Для створення сукупності даних у вигляді вектору існує декілька способів, які наведені у таблиці 3.3.

Таблиця 3.3 – Способи створення векторів

Команда	Інтерпретація	Реалізація
<code>c(1,3,7)</code>	Конкатенація (поєднання окремих значень у вектор)	1 3 7
<code>1:15</code>	Прямий діапазон цілих додатних чисел (від першого до останнього)	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
<code>15:1</code>	Обернений діапазон цілих чисел (від останнього до першого)	15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
<code>-10:-1</code>	Прямий діапазон цілих від'ємних чисел	-10 -9 -8 -7 -6 -5 -4 -3 -2 -1
<code>-3.8:7</code>	Прямий діапазон дробових чисел	-3.8 -2.8 -1.8 -0.8 0.2 1.2 2.2 3.2 4.2 5.2 6.2
<code>seq(from=2, to=20, by=2)</code> або <code>seq(2, 20, 2)</code>	Прямий діапазон чисел із заданим кроком	2 4 6 8 10 12 14 16 18 20
<code>c(1:5, 8, 11:15)</code>	Конкатенація векторів	1 2 3 4 5 8 11 12 13 14 15

Завдання для вивчення теми

I. Завдання

1. Який вектор формується командою: `c(1:5, seq(10, 20, 2))`?

У мові R після виконання команди на наступному рядку з'являється номер команди у квадратних дужках та її інтерпретація.

Команда `class` дає змогу перевірити тип даних, до якого належать дані, які задані в дужках як параметри функції `class`.

2. Задайте наступні команди і перевірте тип даних для кожного вектора, заданого у таблиці 3.4.

Таблиця 3.4 – Приклади векторів різних типів

Логічні	Цілочисельні	Текстові
c(TRUE, TRUE, FALSE) [1] TRUE TRUE FALSE class(c(TRUE, TRUE, FALSE)) [1] "logical"	c(1, 2, TRUE) [1] 1 2 1 class(c(1, 2, TRUE)) [1] "numeric"	c(2, "boo") [1] "2" "boo" class(c(2, "boo")) [1] "character"

У мові R кожна команда в консолі починається зі знака >.

Для перегляду значень, які містить змінна, необхідно ввести ім'я змінної для відображення значень та натиснути Enter.

II. Завдання

Який клас буде мати вектор c(1, 2, "tom", 2+8i, TRUE)?

Функція *summary*

Основні параметри функції *summary* наведені у таблиці 3.5.

Таблиця 3.5 – Параметри функції *summary*

Компонента функції <i>summary</i>	Значення компоненти
Min	Найменше значення
1st Qu.	Перший квантиль
Median	Медіана (другий квантиль)
Mean	Середнє значення
3rd Qu.	Третій квантиль
Max	Найбільше значення

Приклад. У масиві *income* задано прибутки фірми за 10 місяців у тис. грн, а в масиві *outcome* – збитки.

> *income* = c(176, 154, 138, 196, 132, 176, 181, 169, 150, 175);

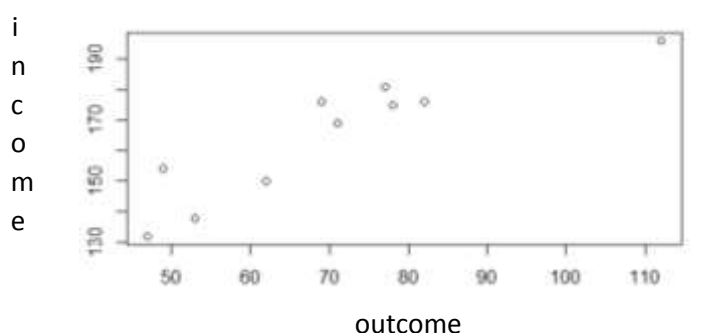
> *outcome* = c(82, 49, 53, 112, 47, 69, 77, 71, 62, 78).

У таблиці 3.6 наведено основні розглянуті команди, застосовані до даних прикладу.

Таблиця 3.6 – Приклад застосування команд описової статистики

Команда	Реалізація
summary(income)	Min. 1st Qu. Median Mean 3rd Qu. Max. 132.0 151.0 172.0 164.7 176.0 196.0
range(income)	132 196 (мінімальне та максимальне значення)
mean(income)	164.7 (середнє значення)
sd(income)	20.38 (середньоквадратичне відхилення)
length(income)	10 (обсяг вибірки)
max(income)	196 (максимальне значення)
min(income)	132(мінімальне значення)
income [1]	176 (перший елемент вибірки)
income [3]	138 (третій елемент вибірки)

Команда `plot(outcome, income)` дає змогу зобразити кореляційне поле (точковий графік залежності), що продемонстровано на рис. 3.2.

Рисунок 3.2 – Кореляційне поле із використанням команди `plot`

Для зберігання даних у мові R, як і в будь-якій іншій мові програмування, використовуються змінні. Команда присвоєння значення змінній у R виглядає так: **Змінна** `<- значення`.

Завдання для вивчення теми

1. У векторі `my_vector` відберіть лише ті спостереження, які відхиляються від середнього менше, ніж на одне стандартне відхилення. Збережіть ці спостереження в нову змінну `my_vector_2`. При цьому вихідний вектор `my_vector` залиште без змін.

Відповідь. `my_vector_2 <- my_vector[my_vector > mean(my_vector) - sd(my_vector) & my_vector < mean(my_vector) + sd(my_vector)]`.

2. Візьміть на сайті Державної статистики України будь-яку динамічну величину (що задана за роками) або просторову величину (що задана за регіо-

нами) і розрахуйте для неї всі основні характеристики описової статистики. Зробіть висновки.

Матриці

Матриця – двомірний набір елементів одного типу (аналог таблиці). Команда для створення матриці у мові R виглядає так:

matrix(data=Nmin:Nmax, nrow=m, ncol=n),

де Nmin – мінімальне значення в матриці, Nmax – максимальне, m і n – кількість рядочків та стовпчиків матриці відповідно. При цьому матриця заповнюється числами від Nmin до Nmax по стовпцях. Щоб задати цілий стовбчик даних, необхідно в квадратних дужках поруч із ім'ям матриці номер рядочка пропустити, і відповідно, щоб задати рядочок, треба пропустити номер стовпця.

Приклад. Команда **matrix(data=1:12, nrow=3, ncol=4)** задає таку матрицю:

```

      [,1] [,2] [,3] [,4]
[1,]  1   4   7  10
[2,]  2   5   8  11
[3,]  3   6   9  12

```

Щоб задати матрицю за рядочками, до попередньої команди треба додати параметр byrow=TRUE.

Приклад. Команда **matrix(data=1:12, nrow=3, ncol=4, byrow=TRUE)** задає таку матрицю:

```

      [,1] [,2] [,3] [,4]
[1,]  1   2   3   4
[2,]  5   6   7   8
[3,]  9  10  11  12

```

Data.frames

Data.frame – двомірний набір даних (також аналог таблиці). Відрізняється від матриць тим, що стовпці в data.frame можуть містити дані різного типу. Проте тип даних в кожному стовпці може бути лише один.

Приклад. Сформуємо датафрейм, який містить дані за трьома країнами (змінна country) про кількість тисяч підприємств малого бізнесу (змінна SmallBusiness) та частку працівників малого підприємництва у загальній кількості працюючих, задану у відсотках (змінна PartMBemp_emp) за 2017 р.

data.frame(country("Ukraine", "France", "USA"), SmallBusiness=c(323,1980, 19200), PartMBemp_emp=c(28.5, 54, 54)).

Результат команди наведено нижче:

	country	SmallBusiness	PartMBemp_emp
1	Ukraine	323	28.5
2	France	1980	54
3	USA	19200	54

Формули

Формули – це спеціальна форма вираження відношень між змінними у рівнянні. Формули використовуються при побудові моделей для визначення функціональної залежності між параметрами.

Лінійна комбінація (задається за допомогою знака +):

$$y = a * x_1 + b * x_2 + c.$$

Команда для формування лінійної комбінації:

formula("y ~ x1 + x2").

Результат виконання команди: $y \sim x_1 + x_2$.

Лінійна комбінація без вільного члена (задається за допомогою символів +0): $y = a * x_1 + b * x_2$.

Команда для формування лінійної комбінації без вільного члена:

formula("y ~ x1 + x2 + 0").

Результат виконання команди: $y \sim x_1 + x_2 + 0$.

Формули можуть містити математичні функції:

$$\log(y) = a * \log(x_1) + b * x_2 + c.$$

Команда для формування лінійної комбінації з логарифмом:

formula("log(y) ~ log(x1) + x2").

Результат виконання команди: $\log(y) \sim \log(x_1) + x_2$.

Завдання для вивчення теми

Як задати формулу для експоненційної регресії?

Векторизація

В R дії виконуються поелементно відразу над усім набором даних (вектор, матриця, data.frame). Векторні обчислення дуже ефективні, тому завжди, коли потрібно виконати дії над елементами вектора (матриці і т. д.) краще використовувати векторизацію, а не цикли.

Наприклад, для заданого масиву даних, який містить числа від 1 до 10, однією командою додано число 2 до кожного елементу масиву.

```
a <- 1:10
a + 2
[1] 3 4 5 6 7 8 9 10 11 12
```

Так само додавання та віднімання, а також лінійна комбінація двох векторів однакової довжини відбувається поелементно:

```
a <- 1:10
b <- 11:20
a + b
[1] 12 14 16 18 20 22 24 26 28 30
m <- matrix(1:12, nrow=4)
```

Помножимо всі елементи матриці на 2, або піднесемо до другого степеня:

m * 2				m ^ 2			
[,1]	[,2]	[,3]		[,1]	[,2]	[,3]	
[1,]	2	10	18	[1,]	1	25	81
[2,]	4	12	20	[2,]	4	36	100
[3,]	6	14	22	[3,]	9	49	121
[4,]	8	16	24	[4,]	16	64	144

Особливості векторизації

Якщо два вектори мають різну кількість елементів, то вектор меншої довжини буде повторюватися стільки разів, щоб відповідати довжині більшого вектора.

Якщо довжина більшого вектора *кратна* довжині меншого вектора, така операція буде проведена неявно, без повідомлення користувача:

```
a <- 1:10
a
[1] 1 2 3 4 5 6 7 8 9 10
b <- 1:5
b
[1] 1 2 3 4 5
a + b
[1] 2 4 6 8 10 7 9 11 13 15
```

У разі, якщо довжини комбінованих векторів *різні*, то буде зроблено циклічне суміщення елементів меншого вектора щодо елементів більшого вектора і буде згенеровано попередження:

```
a <- 1:10
b <- 1:3
a + b
```

Увага: довжина більшого об'єкта не є добутком довжини меншого об'єкта.

```
[1] 2 4 6 5 7 9 8 10 12 11
```

Імена елементів векторів, матриць, data.frames

Звичайний та іменованний вектор:

```
a <- 1:5
a
[1] 1 2 3 4 5
names(a) <- c("one", "two", "three", "four", "five")
a
one two three four five
1 2 3 4 5
```

Інший спосіб створення іменованого вектора:

```
a <- c("one"=1, "two"=2, "three"=3, "four"=4, "five"=5)
a
one two three four five
1 2 3 4 5
```

Аналогічно векторам матриці та data.frames властивості **rownames** і **colnames**, які дають змогу змінювати імена стовпців і рядочків:

```
df <- data.frame(1:2, 3:4, 5:6, 7:8)
rownames(df) <- c("case_1", "case_2")
colnames(df) <- c("var_1", "var_2", "var_3", "var_4")
df
var_1 var_2 var_3 var_4
case_1 1 3 5 7
case_2 2 4 6 8
```

Видалення імен здійснюється присвоєнням спеціального типу NULL:

```
colnames(df) <- NULL
df
```

```

NA NA NA NA
case_1 1 3 5 7
case_2 2 4 6 8
Або для векторів
a <- 1:5
names(a) <- c("one", "two", "three", "four", "five")
a
one two three four five
1 2 3 4 5
names(a) <- NULL
a
[1] 1 2 3 4 5

```

При цьому до елементів вже не можна буде звертатися за іменем, а тільки за індексом.

Індексація векторів, матриць, data.frames

Індексація – ефективний і потужний інструмент для роботи з даними.

Індекси можуть бути: числовими, логічними і текстовими.

Для індексування використовується три типи виразів:

[– вибирає елементи вектора / списку / масиву і т. д.

\$ – вибирає один елемент з data.frame / списку за його ім'ям.

[[– вибирає елементи з вектора / списку / масиву і т. д, але відкидає імена, якщо вони є.

Числові індекси

Вибір елементів вектора за їх індексом:

```

a <- 11:20
a
[1] 11 12 13 14 15 16 17 18 19 20
a[1:5]
[1] 11 12 13 14 15
a[c(2,4:7,9)]
[1] 12 14 15 16 17 19

```

З цього прикладу видно, що для індексування та вибору елементів насправді використовується вектор індексів:


```

a <- 11:20
a
[1] 11 12 13 14 15 16 17 18 19 20
b <- 1:5
b
[1] 1 2 3 4 5
a[b]
[1] 11 12 13 14 15

```

Щоб вибрати кожен другий елемент за індексом, треба згенерувати вектор, що складається з парних чисел, і використовувати його для індексації вихідного вектора:

```

a <- 11:20
a
[1] 11 12 13 14 15 16 17 18 19 20
b <- seq(1, 10, 2)
b
[1] 1 3 5 7 9
a[b]
[1] 11 13 15 17 19

```

Для видалення елементів за значенням індексу перед ними додають знак мінус:

```

a <- 11:20
a
[1] 11 12 13 14 15 16 17 18 19 20
a[-1]
[1] 12 13 14 15 16 17 18 19 20
a[-c(2:4,7)]
[1] 11 15 16 18 19 20

```

Особливості індексування дають змогу змінювати положення елементів і дублювати їх. Це дуже ефективний прийом, про який часто забувають:

```

a <- 11:20
a[5:1]
[1] 15 14 13 12 11
a[c(5,1,1,1,2,3)]
[1] 15 11 11 11 12 13

```

З погляду індексування матриці і `data.frames` майже нічим не відрізняються. Створимо тестовий набір даних:

```
df <- data.frame(var1=c(11,21,31), var2=c(12,22,32), var3=c(13,23,33),
  var4=c(14,24,34), row.names=c("case1", "case2", "case3"))
```

```
df
  var1 var2 var3 var4
case1  11  12  13  14
case2  21  22  23  24
case3  31  32  33  34
```

Оберемо елемент рядочка 1 і стовпчика 2:

```
df[1,2]
[1] 12
```

Оберемо всі значення рядка 1. Результатом буде новий `data.frame`:

```
df[1,]
  var1 var2 var3 var4
case1  11  12  13  14
```

Оберемо всі значення стовпця 1. Результатом буде вектор.

`Data.frame` – це сукупність векторів-стовпців, і при виборі одного стовпця відбувається автоматичне перетворення результату до вектора:

```
df[,1]
[1] 11 21 31
```

Щоб уникнути цього, необхідно додати опцію `drop`. Тепер результатом буде `data.frame`:

```
df[,1, drop=FALSE]
  var1
case1  11
case2  21
case3  31
```

Для вибору блоків даних індексами рядків і стовпців (знак `#` на початку рядочка позначає коментар, тобто команда не виконується, а лише пояснює наступні дії):

```
# вибір рядочків 1 і 2 та стовпців 1 і 2
df[1:2, 1:2]
```

```

var1 var2
case1 11 12
case2 21 22
# вибір рядочків 1, 2 і 4 та стовпців 1 і 3
df[c(1,3), c(1:2,4)]
var1 var2 var4
case1 11 12 14
case3 31 32 34

```

Від'ємні індекси використовуються для видалення відповідних стовпців і рядків (зверніть увагу, що видаляється не один елемент, а стовпці і рядки):

```

df[-2, -3]
var1 var2 var4
case1 11 12 14
case3 31 32 34

```

Або аналогічно попередньому прикладу, що приводить до того ж результату:

```

df[c(1,3), c(1:2,4)]
var1 var2 var4
case1 11 12 14
case3 31 32 34

```

3.3. КОРЕЛЯЦІЙНО-РЕГРЕСІЙНИЙ АНАЛІЗ ЗАСОБАМИ R

Лінійна регресія засобами R

Лінійна регресія – це найпростіший варіант регресії. Розглянемо, для прикладу, лінійну регресію з одним предиктором (аргументом, фактором) і одним прогнозним значенням. Таку регресію легко намалювати на графіку ХОУ. Для цього по осі абсцис Х ми відзначаємо значення предиктора, а по осі ординат Y значення прогнозованої величини. Тоді проста лінійна регресія – це пряма, проведена в такий спосіб, щоб мінімізувати розбіжність між істинними значеннями прогнозованої величини і точками на лінії, відповідними значеннями предикторів.

Для побудови моделі лінійної регресії у R використовується функція **lm(formula=результат~аргумент)**. Продемонструємо на прикладі побудови парної лінійної моделі залежності прибутків від витрат, які були розглянуті у попередніх прикладах.

```
> # Обчислення параметрів рівняння регресії
> income<-c(156,175,136,186,178)
> outcome<-c(68,78,76,86,79)
> lm.model<-lm(formula = outcome ~ income)
> lm.model
```

Call:

```
lm(formula = outcome ~ income)'
```

Coefficients:

```
(Intercept)    income
  43.3785      0.2047
```

Внаслідок цього були отримані коефіцієнти регресії: Intercept – вільний член регресії, income – значення основного коефіцієнта регресії. Отримане рівняння регресії має такий вигляд: $outcome = 43.3785 + 0.2047 \text{ income}$.

Для виведення або збереження коефіцієнтів регресії використовуються команди з використанням знака `$coefficient[i]`, де `i` – порядковий номер коефіцієнта регресії, починаючи з одиниці.

```
> # Виведення модельних значень:
> a0 <- lm.model$coefficient[1]
> a0
(Intercept) 43.37851
> a1 <- lm.model$coefficient[2]
> a1
income 0.2047021
```

Продемонструємо можливість побудови лінійної парної регресії у вигляді датафрейму. Для цього сформуємо датафрейм, що складається з двох компонентів – залежної та незалежної змінних, побудуємо графік регресії та нанесемо лінію регресії на нього.

```
> model<-data.frame("outcome"=outcome," income"= income)
> # Виведення графіка залежності:
> plot(model[,2], model[,1], main="Графік залежності витрат від доходів",
xlab="Доходи", ylab="Витрати")
> # Нанесення сітки:
> grid()
> # Промальовка прямої лінії регресії
> abline(a0,a1)
```

Графік залежності витрат від доходів

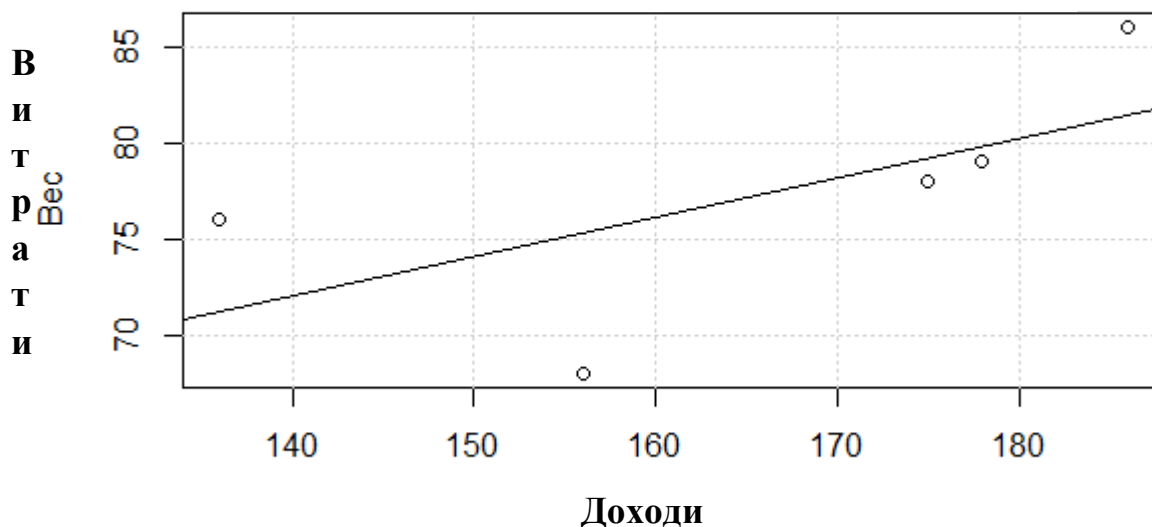


Рисунок 3.3 – Графік залежності витрат від доходів

Для перегляду розгорнутої інформації про лінійну регресію використовується функція `summary()`. Продемонструємо її дію на прикладі побудованої регресії.
`>summary(lm.model)`

Call:

`lm(formula = outcome ~ income)`

Residuals:

1	2	3	4	5
-7.3120	-1.2014	4.7820	4.5469	-0.8155

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	43.3785	23.8387	1.820	0.166
income	0.2047	0.1426	1.436	0.247

Residual standard error: 5.748 on 3 degrees of freedom

Multiple R-squared: 0.4072, Adjusted R-squared: 0.2096

F-statistic: 2.061 on 1 and 3 DF, p-value: 0.2466

Внаслідок цього отримано такі значення:

Residuals – значення залишків у кожній точці.

Coefficients – значення за коефіцієнтами містять таку інформацію:

- **Estimate** – оцінка (безпосереднє значення коефіцієнта);
- **Std. Error** – стандартна помилка коефіцієнта;

- *t value* – розрахункове значення критерія Стюдента для перевірки коефіцієнта на значущість;

- *Pr(>|t|)* – помилка критерія Стюдента при перевірці коефіцієнта на значущість.

Multiple R-squared – коефіцієнт детермінації.

Adjusted R-squared – скоригований коефіцієнт детермінації.

F-statistic – розрахункове значення критерія Фішера для перевірки моделі цілком на значущість.

p-value – помилка критерія Фішера при перевірці моделі цілком на значущість.

Аналізуючи отримані результати, робимо висновок про недоцільність використання лінійної парної регресії у такій ситуації, оскільки помилка критерія Стюдента складає 16,6 % та 24,7 % відповідно для кожного коефіцієнта, а помилка критерія Фішера – 24,7 %, що доводить незначущість отриманого рівняння загалом, а також незначущість кожного отриманого коефіцієнта.

Лінійна логарифмічна регресія

Лінійна логарифмічна регресія має такий вигляд: $y = a_0 + a_1 \ln(x)$. Тому для її реалізації також можна використовувати функцію **lm(formula = результат ~ аргумент)**, але аргументом вважати $\ln(x)$. Продемонструємо її на нашому прикладі.

```
>lm.model<-lm(formula = model[,1] ~ log(model[,2])
```

```
> summary(lm.model)
```

Call:

```
lm(formula = model[, 1] ~ log(model[, 2]))
```

Residuals:

1	2	3	4	5
-7.6308	-1.1912	4.6196	4.9202	-0.7178

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-80.81	119.80	-0.675	0.548
log(model[, 2])	30.98	23.45	1.321	0.278

Residual standard error: 5.936 on 3 degrees of freedom

Multiple R-squared: 0.3677, Adjusted R-squared: 0.157

F-statistic: 1.745 on 1 and 3 DF, p-value: 0.2782

Завдання для вивчення теми

Побудуйте аналогічно експоненційну модель, перевірте її якість та порівняйте із уже побудованими моделями. Не забувайте, що експоненційна модель

є мультиплікативною, і для її побудови необхідно перейти до адитивної засобом логарифмування і введення заміни.

Багатофакторна лінійна регресія

Для цього прикладу ми використовуємо набір даних `state.x77`, який входить до базової версії програми R та містить характеристики:

`Murder` – рівень злочинності,

`Population` – чисельність населення,

`Illiteracy` – рівень неграмотності,

`Income` – середній дохід,

`Frost` – рівень морозності (середнє число днів із температурою нижче нуля) за різними штатами США.

Оскільки функція `lm()` працює з таблицею даних (а набір даних `state.x77` зберігається у вигляді матриці), можна перетворити її за допомогою функції `as.data.frame()` із параметрами назви бази даних із сукупністю полей:

```
> states <-
as.data.frame(state.x77[c("Murder", "Population", "Illiteracy", "Income", "Frost")])
```

Для аналізу кореляційної матриці скористаємося функцією `cor()`:

```
> cor(states)
```

	Murder	Population	Illiteracy	Income	Frost
Murder	1.0000000	0.3436428	0.7029752	-0.2300776	-0.5388834
Population	0.3436428	1.0000000	0.1076224	0.2082276	-0.3321525
Illiteracy	0.7029752	0.1076224	1.0000000	-0.4370752	-0.6719470
Income	-0.2300776	0.2082276	-0.4370752	1.0000000	0.2262822
Frost	-0.5388834	-0.3321525	-0.6719470	0.2262822	1.0000000

Багатофакторна регресія будується аналогічно парній лінійній регресії із додаванням усіх факторів у формулі, при цьому слово `formula` може бути відсутнім, що продемонстровано на рис. 3.4.

Звертаємо увагу, що зірочки, які з'явилися поруч із показниками, показують якість отриманих коефіцієнтів. Чим більше зірочок, тим більш значущий коефіцієнт отриманий. Це дуже зручний показник для прийняття подальших рішень, особливо при розв'язанні задач бізнес-аналітики в умовах великих даних. Отже, у цій моделі виділено два якісних фактори – `Population` та `Illiteracy`.

```

> fit <- lm(Murder ~ Population + Illiteracy + Income + Frost,
            data=states)
> summary(fit)

Call:
lm(formula=Murder ~ Population + Illiteracy + Income + Frost,
    data=states)

Residuals:
    Min       1Q   Median       3Q      Max
-4.7960 -1.6495 -0.0811  1.4815  7.6210

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.23e+00   3.87e+00   0.32   0.751
Population   2.24e-04   9.05e-05   2.47   0.017 *
Illiteracy   4.14e+00   8.74e-01   4.74  2.2e-05 ***
Income       6.44e-05   6.84e-04   0.09   0.925
Frost        5.81e-04   1.01e-02   0.06   0.954
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 'v' 1

Residual standard error: 2.5 on 45 degrees of freedom
Multiple R-squared:  0.567,    Adjusted R-squared:  0.528
F-statistic: 14.7 on 4 and 45 DF,  p-value: 9.13e-08

```

Рисунок 3.4 – Реалізація багатofакторної регресії

Завдання для вивчення теми

Побудуйте якісну багатofакторну модель за цим прикладом засобом покрокового відбору факторів за критерієм Стьюдента. Для цього відкидаємо по одному найменш значущі фактори за критерієм Стьюдента, поки не отримаємо всі значущі фактори.

Установка додаткового пакета car

(зібраний під R 3.4.2)

Для проведення більш ґрунтовного аналізу багатofакторної моделі та визначення шляхів її покращення встановимо додатковий пакет із назвою car. Для цього в четвертому вікні R Studio оберемо блок Packages.

Packages – Install

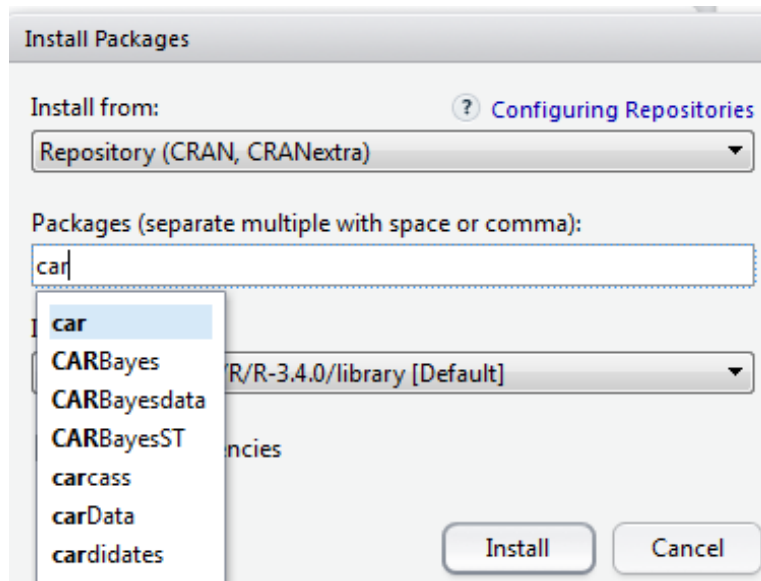


Рисунок 3.5 – Установка нового пакета

Використання пакета car

`scatterplotMatrix(states, spread=FALSE, main="Матриця діаграм розсіювання")`

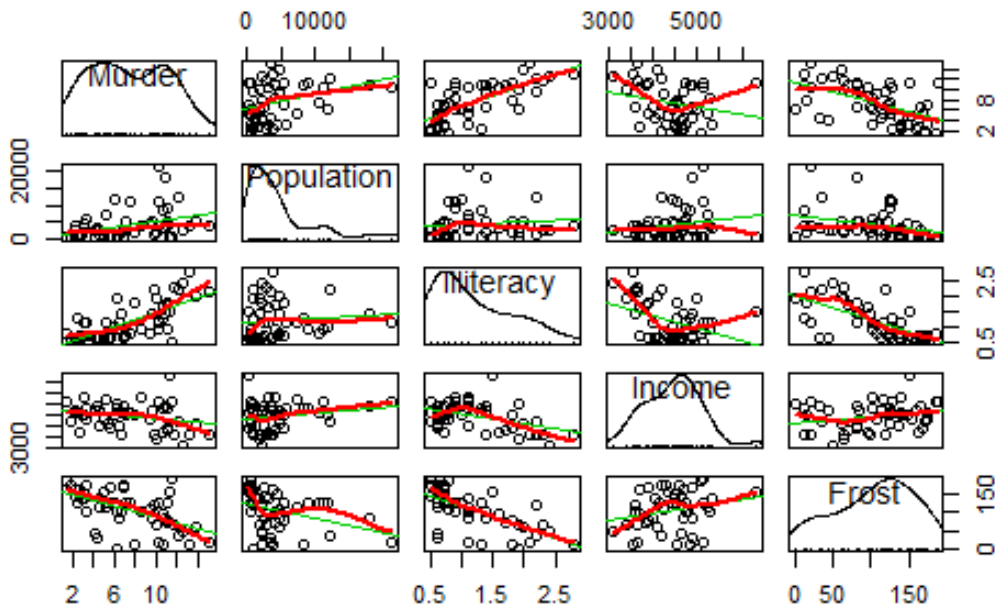


Рисунок 3.6 – Застосування пакета car

Можна бачити, що рівень злочинності у США має бімодальний розподіл, а розподіл кожної незалежної змінної тою чи іншою мірою асиметричний. Рівень злочинності зростає разом зі зростанням населення і рівнем неписьменності, а

падає – зі збільшенням доходу і морозності. Водночас більш холодні штати характеризуються нижчим рівнем неписьменності і чисельності населення, а також високими доходами:

```
> fit <- lm(Murder ~ Population + Illiteracy + Income + Frost, data=states)
```

```
> summary(fit)
```

Call:

```
lm(formula = Murder ~ Population + Illiteracy + Income + Frost,
    data = states)
```

Residuals:

```
   Min     1Q  Median     3Q      Max
-4.7960 -1.6495 -0.0811  1.4815  7.6210
```

Coefficients:

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.235e+00  3.866e+00  0.319  0.7510
Population   2.237e-04  9.052e-05  2.471  0.0173 *
Illiteracy   4.143e+00  8.744e-01  4.738 2.19e-05 ***
Income       6.442e-05  6.837e-04  0.094  0.9253
Frost        5.813e-04  1.005e-02  0.058  0.9541
```

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.535 on 45 degrees of freedom

Multiple R-squared: 0.567, Adjusted R-squared: 0.5285

F-statistic: 14.73 on 4 and 45 DF, p-value: 9.133e-08

Довірчі інтервали

Для знаходження довірчих інтервалів у R використовується функція `confint()`:

```
> fit <- lm(Murder ~ Population + Illiteracy + Income + Frost, data=states)
```

```
> confint(fit)
```

```
                2.5 %                97.5 %
(Intercept) - 6.552191e+00    9.0213182149
Population   4.136397e-05    0.0004059867
Illiteracy   2.381799e+00    5.9038743192
Income      -1.312611e-03    0.0014414600
Frost       -1.966781e-02    0.0208304170
```

Отриманий результат свідчить про те, що можна бути на 95 % впевненим у тому, що інтервал [2.38, 5.90] містить реальне значення, на яке змінюється рівень злочинності, при зміні рівня неписьменності на 1 %. Крім того, оскільки довір-

чий інтервал для морозності (Frost) містить нуль, можна зробити висновок, що зміни цього параметра не пов'язані з рівнем злочинності при постійних значеннях інших змінних.

Багатофакторна лінійна регресія зі взаємодіями

Розглянемо дані про автомобілі з таблиці `mtcars`. Вивчимо вплив ваги автомобіля і потужності двигуна на витрати палива. Можна підібрати регресійну модель, що включає обидві незалежні змінні, а також взаємодію між ними:

```
> fit <- lm(mpg ~ hp + wt + hp:wt, data=mtcars)
```

```
> summary(fit)
```

Call:

```
lm(formula = mpg ~ hp + wt + hp:wt, data = mtcars)
```

Residuals:

```
  Min    1Q  Median    3Q   Max
-3.0632 -1.6491 -0.7362  1.4211  4.5513
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	49.80842	3.60516	13.816	5.01e-14 ***
hp	-0.12010	0.02470	-4.863	4.04e-05 ***
wt	- 8.21662	1.26971	-6.471	5.20e-07 ***
hp:wt	0.02785	0.00742	3.753	0.000811 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.153 on 28 degrees of freedom

Multiple R-squared: 0.8848, Adjusted R-squared: 0.8724

F-statistic: 71.66 on 3 and 28 DF, p-value: 2.981e-13

Зі стовпця $Pr(>|t|)$ видно, що взаємодія між потужністю двигуна і вагою машини значима. Що це означає? Значуща взаємодія між двома незалежними змінними свідчить про те, що на взаємозв'язок між однією незалежною змінною і залежною впливають значення іншої незалежної змінної. У цьому разі характер залежності між витратою палива і потужністю двигуна неоднаковий для автомобілів різної ваги.

Coefficients:

модель для передбачення значень `mpg` (витрата палива):

$$\text{mpg} = 49.81 - 0.12 \times \text{hp} - 8.22 \times \text{wt} + 0.03 \times \text{hp} \times \text{wt}$$

Однак упевненість у цих результатах сильна настільки, наскільки виконуються статистичні допущення, що лежать в основі моделі.

Кластерний аналіз

```
#читаємо файл даних
x <- read.table('D:\\имя.csv', sep=';', header=TRUE)
#відкидаємо колонку з порядковим номером чи датою
x <- x[-1]
#будемо розбивати на 3 кластери, максимум 1 000 000 ітерацій
kmeans(x, 3, 10000)
hc <- hclust(dist(x))
plot(hc)
```

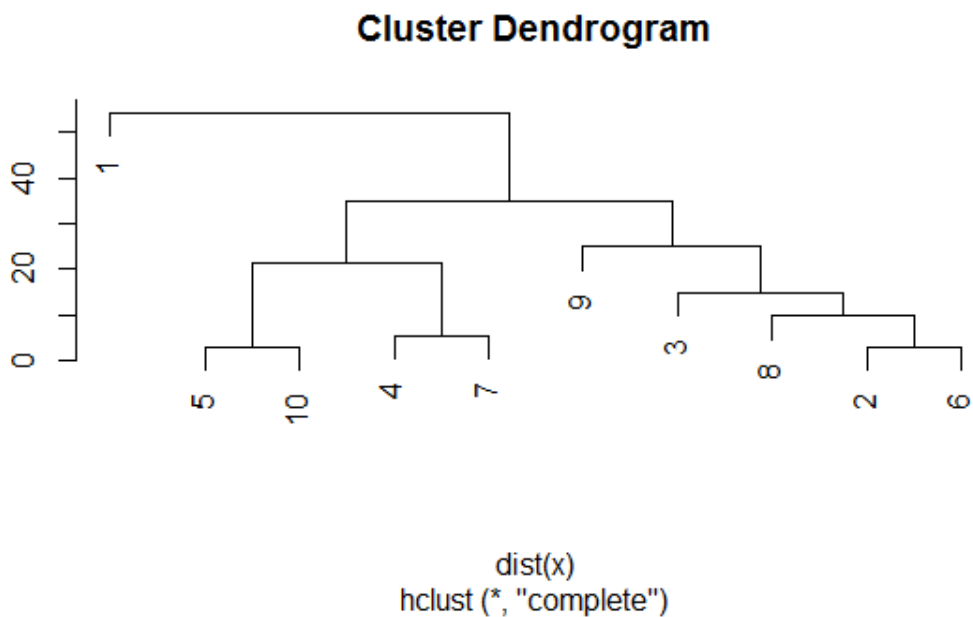


Рисунок 3.7 – Дендрограма

Важливо зазначити, що отриманий при використанні кластерного аналізу результат є одним із можливих. Цей результат необхідно порівнювати з аналогічними результатами, отриманими із застосуванням інших алгоритмів, а також з результатами інших методів аналізу даних.

Завдання для вивчення теми

Візьміть на сайті ukrstat.ua у розділі регіональної статистики декілька показників, які характеризують регіони України у різних аспектах, і проведіть регіональну диференціацію засобами кластерного аналізу.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ДО РОЗДІЛУ 3

1. Crawley J. Statistics: An Introduction using R. Wiley, 2nd edition, 2014. URL: <http://www.bio.ic.ac.uk/research/crawley/statistics/>
2. Гайдышев И. Анализ и обработка данных: специальный справочник. СПб: Питер, 2001. 752 с.
3. R: A Language and Environment for Statistical Computing. Reference Index / The R Development Core Team. URL: <http://lib.stat.cmu.edu/R/CRAN/doc/manuals/fullrefman.pdf>
4. Шипунов А. Б., Балдин Е. М., Волкова П. А. и др. Наглядная статистика. Используем R! URL: <http://ashipunov.info/shipunov/school/books/rbook.pdf>
5. Дуглас Люк: Анализ сетей (графов) в среде R. Руководство пользователя. Цветное издание (перевод с английского). URL: <http://dmkpress.com/catalog/computer/statistics/978-5-97060-428-1/>
6. Джеймс Г., Уиттон Д., Хасты Т., Тибширани Р. Введение в статистическое обучение с примерами на языке R. Цветное издание. (перевод с английского). URL: <http://dmkpress.com/catalog/computer/statistics/978-5-97060-293-5/>
7. Мастицкий С. Э., Шитиков В. К.: Статистический анализ и визуализация данных с помощью R. URL: https://github.com/ranalytics/r-tutorials/blob/master/Edition_2014/Book/Mastitsky_and_Shitikov_2014_R_tutorials.pdf
8. Шитиков В. К., Мастицкий С. Э. Классификация, регрессия и другие алгоритмы Data Mining с использованием R. URL: http://www.ievbras.ru/ecostat/Kiril/R/DM/DM_R.pdf
9. Викиучебник: Язык программирования R. URL: https://ru.wikibooks.org/wiki/Язык_программирования_R
10. Дэвид Мерц, Бред Хантинг: Статистическое программирование на R (электронный перевод, части 1, 2, 3). URL:
<http://www.ibm.com/developerworks/ru/library/1-r1/index.html>
<http://www.ibm.com/developerworks/ru/library/1-r2/index.html>
<http://www.ibm.com/developerworks/ru/library/1-r3/index.html>
11. A Handbook of Statistical Analyses Using R. URL: https://cran.r-project.org/web/packages/HSAUR/vignettes/Ch_introduction_to_R.pdf

ДОДАТОК А

РЕГІОНИ	1	2	3	4	5	6		7	8	9	10	11		12	13	14
						Обсяг залучених депозитів домашніх господарств	на душу населення, грн					Зборгова-ність з виплати заробітної плати	на 1 зайнятого, грн			
		Обсяг споживчого кредитування, млрд грн	Середня з/п 1 штатного працівника, грн	Чисельність зайнятих у віці 15-70 років, тис. осіб	Процентна ставка за споживчими кредитами, %	Обсяг залучених депозитів домашніх господарств	на душу населення, грн	Процентна ставка за депозитами, %	Імпорт товарів на душу наявного населення, дол. США	ІСД на предмети домашнього вжитку, побутову техніку, поточне утримання жителя, %	Зборгова-ність з виплати заробітної плати	на 1 зайнятого, грн	Чисельність наявного населення, тис. осіб	Чисельність зайнятих на 100 осіб наявного населення, осіб		
АР Крим (включно з м. Севастополь)		4,6	2651	914,2	28,3	16729	7202,1	13,9	680,5	98,7	63,6	69,6	2322,8	39		
Вінницька		1,8	2432	701,1	27,4	6510	3992,4	14,7	305,7	98,1	14,0	20,0	1630,6	43		
Волинська		1,7	2339	442,8	30,8	4135	3978,6	14,1	976,7	101,5	5,2	11,7	1039,3	43		
Дніпропетровська		8,6	3138	1528,5	25,9	33419	10084,2	15,2	1972,9	101,2	10,4	6,8	3314,0	46		
Донецька		10,2	3495	1985,4	26,7	37078	8447,4	14,9	949,7	100,3	172,9	87,1	4389,3	45		
Житомирська		1,7	2369	550,3	30,0	4630	3642,8	15,3	302,4	99,3	13,9	25,3	1271,0	43		
Закарпатська		1,9	2351	530,8	27,5	3451	2755,1	12,8	1603,9	100,6	10,6	20,0	1252,6	42		
Запорізька		3,2	2927	821,2	29,1	12733	7119,4	15,2	1125,7	101,2	51,8	63,1	1788,5	46		
Івано-Франківська		2,1	2540	548,5	31,2	5522	3998,6	15,2	712,4	101,3	8,9	16,2	1381,0	40		
Київська (включно з м. Київ)		51,0	4157	757,5	29,8	128153	28163,0	13,4	6792,6	101,4	82,1	108,4	4550,4	17		
Кіровоградська		1,3	2428	433,7	35,2	3503	3510,0	14,9	285,2	98,6	39,7	91,5	998,8	43		
Луганська		2,8	3090	1008,6	28,4	10808	4772,6	15,4	878,6	100,6	69,2	68,6	2264,6	45		
Львівська		5,1	2578	1099,9	25,1	18543	7298,1	13,6	1327,7	101,1	58,3	53,0	2540,8	43		
Миколаївська		2,3	2822	533,7	30,1	6427	5466,1	15,1	761,8	100,1	37,3	69,9	1175,8	45		
Одеська		6,3	2700	1060,4	26,3	20687	8630,4	12,9	1799,8	102,0	16,1	15,2	2391,7	44		
Полтавська		2,2	2850	652,7	28,4	10102	6860,4	15,9	898,3	100,1	15,4	23,6	1472,5	44		
Рівненська		1,4	2575	492,5	30,2	4133	3576,5	12,9	396,9	101,1	1,4	2,8	1155,6	43		
Сумська		1,2	2503	519,6	30,0	5078	4424,1	15,4	591,4	99,2	58,9	113,4	1147,8	45		
Тернопільська		1,2	2185	439,4	26,4	4061	3764,0	14,1	312,2	99,3	7,0	15,9	1078,9	41		
Харківська		6,0	2752	1280,6	25,0	19872	7243,8	14,4	999,2	99,8	88,7	69,3	2743,3	47		
Херсонська		1,9	2268	477,7	31,6	4760	4404,1	14,3	191,2	100,5	17,4	36,4	1080,8	44		
Хмельницька		2,1	2425	571,3	29,0	5918	4493,2	15,1	428,9	99,2	14,8	25,9	1317,1	43		
Черкаська		1,6	2508	562,7	30,7	6635	5211,7	14,7	453,1	100,3	16,1	28,6	1273,1	44		
Чернівецька		1,5	2329	387,2	30,0	3564	3932,9	13,1	193,0	99,2	0,1	0,3	906,2	43		
Чернігівська		1,2	2308	475,5	28,2	4738	4374,1	15,9	503,0	100,4	19,9	41,9	1083,2	44		

Показники споживчого кредитування за регіонами України

Показники розвитку сфери послуг за регіонами країни

ДОДАТОК Б

Регіон	Реалізація послуг на 1 особу, грн (x1)	Забезпеченість населення автомобілями на 1 000 ос., (x2)	Забезпеченість населення комп'ютерами на 100 сімей, (x3)	Забезпеченість населення житлом, м ² /ос., (x4)	Чисельність студентів на 1 000 ос. населення, (x5)	Кількість відвідувань музеїв, театрів на 10 тис. ос. (x6)	Кількість місць у готелях (x8)	Коефіцієнт захворюваності, на 10 тис. населення (x10)	Товарооборот на душу населення, грн (x11)	Експорт послуг, млн дол. США, (x12)	Споживчі кредити банків, млн грн (x13)
1	2	3	4	5	6	7	9	11	12	13	14
Вінницька	140,936	105,774	34	29,7	178,432	47	1764	95,3	729	2,5	736
Волинська	214,232	94,450	42	21,8	209,762	24	2108	99,2	956	4,4	771
Дніпропетровська	335,53	115,529	54	24,9	385,684	34	10629	144,1	1261	102,1	8032
Донецька	260,98	105,840	46	22,7	160,247	27	5424	108,6	879	66,2	5131
Житомирська	149,565	83,345	44	27,1	189,022	17	2041	118,7	786	3,7	515
Закарпатська	146,601	95,972	34	21,4	185,791	25	3007	107,4	1164	15,9	536
Запорізька	325,297	161,433	60	23,9	201,32	33	3310	116,0	1075	70,8	2300
Івано-Франківська	146,657	96,122	42	22,3	198,991	37	2315	95,2	615	18,3	790
Київська	515,128	64,985	48	31,5	176,994	24	2398	103,1	890	40,9	6255
Кіровоградська	166,879	101,499	38	25,3	172,379	33	1147	126,9	769	9,5	468
Луганська	164,95	87,520	44	22,8	166,976	32	2654	123,4	623	17	1139
Львівська	340,047	81,979	49	21,6	206,081	55	6685	112,4	1185	37,6	2025
Миколаївська	235,71	83,491	51	22,2	180,845	45	1731	103,4	772	53,3	1229
Одеська	444,073	97,645	63	23,3	206,512	52	7091	116,7	1039	505	3106
Полтавська	200,988	102,169	45	25,8	193,362	52	2402	124,7	941	10,3	1671
Рівненська	184,681	94,834	42	21,5	229,933	27	2590	111,3	849	24,1	529
Сумська	198,488	86,382	41	24,9	193,771	22	2122	116,1	828	18	687
Тернопільська	152,579	88,841	46	22,5	210,877	39	3352	105,7	640	2,5	455
Харківська	410,163	122,165	55	22,2	229,414	40	3721	119,5	1277	34,3	3864
Херсонська	210,296	113,203	47	21	196,778	53	2273	109,2	902	22,8	875
Хмельницька	184,409	75,109	43	26,5	216,782	35	1679	102,2	733	4,9	8435
Черкаська	184,319	120,980	43	28,3	187,062	50	2132	110,4	754	2,9	826
Чернівецька	758,21	286,014	44	21,9	606,154	34	1718	104,0	952	1,2	434
Чернігівська	146,627	66,429	46	28,5	176,439	60	1897	115,1	814	3,5	650
м. Київ	1366,27	162,989	88	13,9	290,461	284	13519	170,1	2037	526	29059

Навчальне видання

Сидорова Антоніна Василівна

Біленко Дар'я Вікторівна

Буркіна Наталя Валеріївна

Навчально-методичний посібник
«БІЗНЕС-АНАЛІТИКА»

Редактор
Технічний редактор

О. А. Солдатова
Т. О. Важеніна-Гопрак

Підписано до друку 16.08.2019
Формат 60 x 84/16. Папір офсетний.
Друк – цифровий. Умовн. друк. арк. 6,04
Тираж 20 прим. Зам. 79

Донецький національний університет імені Василя Стуса
21021, м. Вінниця, 600-річчя, 21
Свідоцтво про внесення суб'єкта видавничої справи
до Державного реєстру
серія ДК № 5945 від 15.01.2018

