

М.В. Грисенко, О.А. Чугаєв

**КІЛЬКІСНІ МЕТОДИ АНАЛІЗУ
МІЖНАРОДНИХ ЕКОНОМІЧНИХ
ВІДНОСИН**

Фрагмент

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ
ТАРАСА ШЕВЧЕНКА**

ІНСТИТУТ МІЖНАРОДНИХ ВІДНОСИН

М.В. Грисенко, О.А. Чугаєв

**КІЛЬКІСНІ МЕТОДИ АНАЛІЗУ
МІЖНАРОДНИХ ЕКОНОМІЧНИХ
ВІДНОСИН**

2012

УДК 004.9:303:311:33

Р е ц е н з е н т и:

Майборода Р.Е., доктор фізико-математичних наук, професор кафедри теорії ймовірностей, статистики та актуарної математики, механіко-математичного факультету Київського національного університету імені Тараса Шевченка

Дзюба П.В., кандидат економічних наук, доцент кафедри міжнародних фінансів Інституту міжнародних відносин Київського національного університету імені Тараса Шевченка

*Рекомендовано до друку
вченою радою Інституту міжнародних відносин
Київського національного університету імені Тараса Шевченка
(протокол № 5 від 28 грудня 2011 року)*

*Рекомендовано до друку науково-методичною радою
Київського національного університету імені Тараса Шевченка
(14 листопада 2012 року)*

Грисенко М.В., Чугаєв, О.А.

Кількісні методи аналізу міжнародних економічних відносин: Навчальний посібник. - К.: Інститут міжнародних відносин Київського національного університету імені Тараса Шевченка, 2012. - 235 с.

У книзі розглядаються проблеми аналізу міжнародних економічних відносин кількісними методами. Видання має практичну орієнтацію і побудоване на міждисциплінарній основі. Поєднуються елементи таких дисциплін як міжнародні економічні відносини, економіко-математичний

і статистичний аналіз, інформатика та іноземна мова. Структура включає джерела міжнародної економічної статистики, основи використання спеціалізованого програмного забезпечення, теоретичні аспекти окремих кількісних методів та приклади їх використання за допомогою програмного забезпечення.

Книга може бути корисною для студентів, аспірантів, викладачів, науковців у сфері міжнародних економічних відносин та економіки в цілому, аналітиків та практиків, зацікавлених у кількісних методах як універсальному інструменті аналізу.

© Грисенко М.В., Чугаєв О.А., 2012

ЗМІСТ

ПЕРЕДМОВА.....	7
----------------	---

ГЛАВА 1. Реальні моделі у сфері міжнародних економічних відносин

1.1. Фактори експорту.....	9
1.2. Фактори імпорту.....	10
1.3. Фактори цін експорту та імпорту.....	11
1.4. Поточний рахунок і ВВП.....	12
1.5. Поточний рахунок, активи та заощадження.....	12
1.6. Фінансовий рахунок.....	13
1.7. Валютний курс і процентні ставки.....	14
1.8. Міграція робочої сили.....	15

ГЛАВА 2. Дослідження міжнародних економічних відносин методами кореляційного аналізу

2.1. Основи кореляційного аналізу.....	16
2.2. Побудова кореляційної матриці для зовнішньоекономічних показників.....	19
2.3. Подолання проблем у застосуванні кореляційного аналізу світової економіки.....	20
2.4. Непараметричні методи кореляційного аналізу зв'язку якісних змінних політичного клімату з припливом інвестицій.....	26
2.5. Кореляційний аналіз, як метод первинного дослідження.....	27

ГЛАВА 3. Дослідження факторів валютних криз методами частотного аналізу

3.1. Метод частотного аналізу на прикладі аналізу впливу політичної стабільності та валютних резервів.....	29
3.2. Таблиці частот для аналізу взаємодії факторів: процентної ставки і валютних резервів.....	30
3.5. Частотний аналіз впливу зовнішнього боргу у Microsoft Excel.....	37
3.6. Алгоритм дослідження взаємодії факторів валютних криз методами частотного аналізу.....	38
3.7. Переваги та недоліки методу частотного аналізу.....	46

ГЛАВА 4. Сигнальний метод аналізу факторів валютних та фіскальних криз

4.1. Основи сигнального методу.....	48
4.2. Дослідження валютних криз сигнальним методом у Microsoft Excel.....	49
4.3. Застосування сигнального методу для аналізу факторів фіскальних криз.....	50

ГЛАВА 5. Дослідження іноземних інвестицій методами аналізу середніх

5.1. Методи порівняння середніх.....	55
5.2. Критерії Стьюдента.....	55
5.4. Аналіз чинників припливу інвестицій у Microsoft Excel.....	64
5.5. Непараметричні аналоги t -критерію Стьюдента.....	67

ГЛАВА 6. Дисперсійний аналіз міжнародної торгівлі та інвестицій

6.1. Теоретичні основи методу дисперсійного аналізу.....	71
--	----

6.3. Метод багатofакторного одномірного дисперсійного аналізу впливу торговельних обмежень на імпорт.....	73
6.4. Умови методу дисперсійного аналізу.....	74
6.6. Багатомірний дисперсійний аналіз (MANOVA).....	84
6.8. Дисперсійний аналіз впливу рівня економічного розвитку на приплив інвестицій у Microsoft Excel.....	93
6.9. Непараметричні методи дисперсійного аналізу.....	94

ГЛАВА 7. Дослідження впливу міжнародних економічних відносин на економічне зростання методами лінійного регресійного аналізу

7.1. Основи методу лінійного регресійного аналізу	99
7.2. Регресійний аналіз регіональних особливостей формування процентної ставки за зовнішніми кредитами з використанням бінарних змінних.....	105
7.3. Аналіз взаємодії факторів приросту імпорту.....	106
7.5. Модель лінійної регресії впливу платіжного балансу на економічне зростання у Microsoft Excel.....	118
7.6. Прогнозування динаміки валютного курсу в умовах невизначеності значень факторів у лінійній регресії.....	121

ГЛАВА 8. Дослідження міжнародної торгівлі та валютних криз методами нелінійного регресійного аналізу

8.1. Методи нелінійної регресії.....	124
8.3. Дослідження факторів валютних криз за допомогою логіт-регресії	128
8.4. Функції втрат	136
8.5. Модель нелінійної регресії факторів високотехнологічного експорту у Microsoft Excel.....	137
8.6. Алгоритм дослідження наслідків утворення зони вільної торгівлі	138

ГЛАВА 9. Класифікація країн за рівнем відкритості та зовнішньої збалансованості економіки методами кластерного аналізу

9.1. Метод кластерного аналізу.....	141
9.2. Види кластерного аналізу.....	142

ДОДАТКИ

Додаток А. Джерела міжнародної економічної статистики

A.1. Формат даних.....	154
A.2. Статистика Світового банку. Світові індикатори розвитку і Фінанси глобального розвитку.....	154
A.3. Міжнародна фінансова статистика Міжнародного валютного фонду.....	162
A.4. Статистика Світової організації торгівлі.....	162
A.5. Статистичні органи України.....	164
A.6. Основні джерела комплексної статистики	165
A.7. Комплексна статистика по групі країн	166
A.8. Комплексна статистика: конкурентоспроможність та інтернаціоналізація.....	168
A.9. Статистика міжнародної торгівлі.....	169
A.10. Фінансова статистика.....	171
A.11. Статистика технологій і знання.....	173
A.12. Статистика інфраструктури.....	173

А.13. Статистика населення.....	174
А.14. Статистика соціальної сфери.....	174
А.15. Статистика політичної сфери та державного управління.....	175
А.16. Екологічна статистика.....	175

Додаток Б. Основи роботи у програмному забезпеченні Microsoft Office Excel

Б.1. Формули.....	176
Б.2. Функції.....	177
Б.3. Перерахунок формул.....	178
Б.4. Посилання на комірки у формулах.....	178
Б.5. Функції масиву.....	179
Б.6. Переміщення і копіювання формул.....	180
Б.7. виправлення помилок.....	180
Б.8. Впливаючі та залежні комірки.....	181
Б.9. Основні логічні функції.....	181
Б.10. Основні математичні функції.....	182
Б.11. Функції суми та добутку.....	183
Б.12. Табличні функції.....	185
Б.13. Надбудови.....	187
Б.14. Діаграми.....	189

Додаток Г. Формування бази вхідних даних

Г.1. Підготовка вхідних даних із зовнішніх джерел.....	204
Г.2. Підготовка додаткових розрахованих показників.....	205
Г.3. Характеристики показників.....	207
Г.4. Формування узагальнюючої таблиці.....	208
Г.5. Проблеми з відсутніми даними та їх діагностика.....	209
Г.6. Вирішення проблем з відсутніми даними.....	210
Г.7. Викиди.....	212

Додаток Д. Описова статистика і розподіл даних

Д.2. Описова статистика у Microsoft Excel.....	216
Д.3. Розподіл даних.....	220
Д.6. Розподіл даних у Microsoft Excel.....	227
Д.7. Генерація випадкових чисел.....	229

СПИСОК ЛІТЕРАТУРИ.....	232
-------------------------------	------------

ПЕРЕДМОВА

Міжнародна економіка є однією з важливих сфер застосування кількісних методів та моделей. На сучасному етапі важко собі уявити дослідження та прогнозування світогосподарських процесів без використання економетричного моделювання, регресійного аналізу, трендових та згладжувальних моделей та інших методів. Саме тому, необхідність застосування кількісних методів та аналіз числової інформації у зовнішньоекономічних дослідженнях та практиці розробки економічних рішень у сфері міжнародних економічних відносин обумовлені наступними чинниками:

- збільшення складності та взаємозв'язків процесів, що відбуваються у сфері міжнародних економічних відносин;
- збільшення невизначеності при прийнятті управлінських рішень у сфері міжнародних економічних відносин;
- посилення дії непередбачуваних чинників;
- стрімка зміна умов діяльності на зовнішніх ринках;
- виникнення ситуацій, які раніше ніколи не існували (такі, як світові фінансові кризи).

Як наслідок вищезазначених чинників, є суттєве підвищення вимог до ефективності та обґрунтованості використання кількісних методів в сфері міжнародних економічних відносин до глибини та адекватності економічного аналізу. Кількісні методи використовуються в економіко-математичних моделях зовнішньоекономічних процесів і явищ. Ці моделі можна поділити на класи:

1. Моделі процесів на макро-, мезо- рівнях (у макроекономічній сфері: інфляція, спад, зростання, які можуть суттєво вплинути на характер зовнішньоекономічних зв'язків).

2. Моделі умов зовнішньоекономічної діяльності (моделі процесів на окремих сегментах національного ринку та у сфері міждержавних стосунків, які можуть суттєво впливати на явища у зовнішньоекономічній сфері: динаміка валютних ринків щодо основних видів конвертованої валюти, процеси у позаекономічній сфері держави (характер реформ у країнах з перехідною економікою).

3. Прогнозування та аналіз обсягів та структури експорту та імпорту окремих груп товарів та послуг (в залежності від конкретних чинників).

4. Формування окремих аспектів зовнішньоекономічної політики держави (що дають відповідь на те, які дії слід здійснювати організаціям державного управління, щоб досягти певної мети).

5. Моделі фінансової діяльності, що дають опис процесам, які відбуваються на окремих сегментах валютного ринку, ринку капіталів, фондового ринку, ринку ризиків та страхових послуг тощо.

6. Прогнозування дій параметрів у конфліктних та непередбачуваних ситуаціях (допомагають дослідити, як діятимуть суб'єкти господарювання за певних обставин, які характеризуються відмінністю їх інтересів).

Кількісні методи є потужним інструментом дослідження будь-якої сфери. Водночас кожний предмет дослідження має свою специфіку щодо використання цих методів. Тому в даному навчальному посібнику автори використовували міждисциплінарний підхід, який поєднує чотири складові:

1. Міжнародні економічні відносини як предмет дослідження. Використання універсальних кількісних методів проілюстровано на прикладах аналізу різних форм міжнародних відносин. Окрема частина книги присвячена джерелам статистичної інформації щодо показників міжнародних економічних відносин та суміжних сфер.

2. Практичні аспекти використання кількісних методів: можливості цих методів, етапи аналізу, приклади застосування, застереження, інтерпретація результатів, способи вирішення проблем, поєднання різних методів.

3. Програмне забезпечення, яке дозволяє автоматизацію розрахунків та представляти результати у зручній формі. Існує достатньо велика кількість програмного забезпечення, яке значно полегшує використання кількісних методів, особливо в умовах необхідності дослідження великого масиву даних. Зокрема, використані приклади для Microsoft Excel як найбільш розповсюдженого і знайомого для читача програмного продукту.

4. Іноземна мова. Ключові терміни подані як українською мовою, так і англійською (іноді російською). Знання іноземної мови в цій сфері є корисним зважаючи на те, що:

- міжнародні економічні статистичні бази даних переважно представлені англійською мовою;
- програмне забезпечення переважно доступне англійською чи російською мовою;
- читач може виявити бажання поглибити знання за допомогою книг українською, англійською чи російською мовами;
- значна кількість досліджень у сфері міжнародної економіки, в яких використовуються кількісні методи, написані іноземною мовою, а ознайомлення з ними потребує знання спеціальної термінології.

Навчальний посібник орієнтує читача також на активне формування своїх навичок: ознайомлення з базами статистичних даних та самостійне використання методів кількісного аналізу. Автори мали на меті показати можливості різних кількісних методів, можливі проблеми з їх використанням на практиці та способи вирішення цих проблем. Саме тому значну частину книги становлять приклади. Для прикладів використані умовні або реальні дані (якщо не вказано джерело, використовуються умовні дані).

Робота з навчальним посібником передбачає, що читач вже має базові знання з таких дисциплін як: „Світова економіка”, „Міжнародні економічні відносини”, „Міжнародні фінанси”, „Математика для економістів”, „Теорія ймовірності і математична статистика”, „Теорія статистики”, „Економетрика”, „Сучасні інформаційні системи і технології”, „Економіко-математичне моделювання світогосподарських процесів”, „Іноземна мова (англійська)”.

Навчальний посібник написаний спеціально для навчального курсу „Кількісні методи аналізу міжнародних економічних відносин”, що не виключає можливості використання окремих складових книги при викладанні вказаних вище курсів.

Дисципліна „Кількісні методи аналізу міжнародних економічних відносин” ставить перед собою такі завдання:

- ознайомити студентів з основними кількісними методами, які використовують в сфері міжнародних економічних відносин;
- розширити систему знань щодо статистичних показників в сфері міжнародних економічних відносин;
- сформуванати навички пошуку статистичної інформації для аналізу міжнародних економічних відносин;
- розширити теоретичні знання в сфері методики кількісного аналізу;
- сформуванати навички підготовки та здійснення кількісного аналізу дослідження в міжнародних економічних відносин з використанням спеціального програмного забезпечення;
- надати понятійно-термінологічний апарат у сфері соціально-економічних статистичних показників та методів кількісного аналізу українською та іноземною мовою (англійська, російська).

Під час підготовки навчального посібника використовуватися матеріали підручників, навчальних посібників; наукові публікації; статистична інформація офіційних веб-сайтів міжнародних організацій, державних органів та недержавних структур; інформаційні сайти, присвячені спеціалізованому програмному забезпеченню.

На сучасному етапі розвитку суспільства, завдяки сучасним комп'ютерним технологіям, можливості математичного моделювання практично безмежні, але скористатися ними повною мірою вдається лише тим фахівцям, які вільно володіють кількісними математичними методами, що передбачає насамперед самостійну роботу із навчальною літературою.

Навчальний посібник рекомендовано для студентів Інституту міжнародних відносин Київського національного університету імені Тараса Шевченка, які навчаються за програмою магістрів зі спеціальності міжнародні економічні відносини. Може бути корисний як для аспірантів, викладачів, так і для економістів-практиків, зокрема для економістів-міжнародників. Навчальний посібник може бути корисним фінансистам, актуаріям, менеджерам, бізнесменам, співробітникам страхових компаній, соціологам і політологам, керівникам різного рівня при плануванні та аналізі повсякденної діяльності. Видання можна використовувати не тільки як практичний посібник, але і як довідник.

ГЛАВА 1. Реальні моделі у сфері міжнародних економічних відносин

Світогосподарські процеси є складноорганізованою динамічною підсистемою функціонування світової економіки, що віддзеркалює зовнішньоекономічну діяльність держав і структурних підрозділів їхніх господарств, які базуються на міжнародному поділі праці.

Вони характеризуються динамізмом у часі та просторі, мають галузеву і територіальну структуру. Ці сторони зовнішньої економічної діяльності легко піддаються математичній формалізації, яка зберігає ідентичність реальних співвідношень та масштабність складних явищ і процесів, які досліджуються в просторовому та часовому аспектах. На цій основі за допомогою кількісних методів можлива побудова економіко-математичних моделей, які досліджують і встановлюють конкретні економічні закономірності й взаємозалежності, які спостерігаються у світогосподарських процесах.

В даній главі розглядаються моделі, у яких змінними виступають показники міжнародних економічних відносин: функції експорту та імпорту; фактори цін експорту та імпорту; поточний рахунок і ВВП; фінансовий рахунок валютний курс і процентні ставки; міграція робочої сили.

Для використання розглянутих моделей, які розраховувалися за даними, що не включають останні періоди або потрібні країни, у сучасних умовах для України варто зробити перерахунок коефіцієнтів, здійснити додаткову перевірку на адекватність моделі, модифікувати склад змінних (якщо це потрібно).

1.1. Фактори експорту

Функцію експорту в неявному вигляді можна подати так:

$$X = F \left(ER \cdot \frac{P_F}{P_D}; YWR; Y_p; HD \right), \quad (1.1)$$

де X – реальний експорт;

ER – валютний курс;

P_D – внутрішні ціни;

P_F – експортні ціни або ціни закордоном;

YWR – реальний світовий попит;

Y_p – виробничий потенціал експортоорієнтованого сектору;

HD – реальний внутрішній попит.

Також функцію експорту можна подати у вигляді:

$$X = F (GDP; GNP; DEF; DEFNP), \quad (1.2)$$

де X – реальний (в постійних цінах експорт);

GDP – реальний ВВП;

GNP – реальний світовий ВВП;

DEF – індекс експортних цін (дефлятор експорту);

$DEFNP$ – індекс цін світового ВВП.

В роботі [32] розрахована модель для України:

$$X_{96} = -19369.08 + 2.2629 \cdot GNP_{96} + 2027.91 \cdot (EO(DEF/DEFGDP)), \quad (1.3)$$

де X_{96} – експорт у цінах 1996 року;

GNP_{96} – світовий ВВП у цінах 1996 року;

EO – валютний курс;

DEF – дефлятор експорту, 1996 рік = 1;

$DEFGDP$ – дефлятор ВВП, 1996 рік = 1.
В роботі [19] розрахована модель для України експорту газу з Росії:

$$V_t = 57.14 - 0.07P_t + 0.64 \cdot GDP_t, \quad (1.4)$$

де V_t - експорт газу з Росії до ЄС, млрд. куб. м;

P_t – експортна ціна газу, дол./тис. куб. м;

GDP_t – сумарний ВВП країн ЄС, % до 1994 року.

Оцінка робилася за вибіркою 1995 – 2009 рр.

В підручнику [16] розглядають функцію експорту туристичних послуг:

$$T_e = a_0 + a_1 y_r^f + a_2 (P_d / E_{d/f} \cdot P_f), \quad (1.5)$$

де T_e – експорт туристичних послуг;

y_r^f – реальний дохід у країнах, звідки прибувають туристи (зважений відповідно до часток країн у сумарній кількості туристів, що прибувають);

P_d – індекс внутрішніх цін (індекс цін послуг готелів і ресторанів або дефлятор ВВП);

$E_{d/f}$ – валютний курс

P_f – індекс цін у країнах, звідки прибувають туристи.

В моделі мають бути передбачені достатні лаги, наприклад, рішення про поїздку часто формуються заздалегідь, виходячи із заробленого раніше доходу.

1.2. Фактори імпорту

Функцію імпорту в неявному вигляді можна подати так:

$$M = F(ER \cdot P_f / P_d; Y_d), \quad (1.6)$$

де M – реальний імпорт;

ER – валютний курс;

P_d – внутрішні ціни;

P_f – експортні ціни або ціни закордоном;

Y_d – реальний попит у країні (або реальний ВВП).

Функцію імпорту також можна подати у вигляді:

$$M = F(GDP; DEFM / DEFGDP), \quad (1.7)$$

де M – реальний імпорт;

DGP – реальний ВВП;

$DEFM$ – індекс імпортних цін;

$DEFGDP$ – індекс цін (дефлятор) ВВП.

Можлива побудова функцій експорту в розрізі окремих галузей (наприклад машинобудування, решти товарів, послуг).

В роботі [32] для України розрахована модель:

$$M_{96} = 45925.27 - 0.11805 \cdot DGP_{96} - 1160.87 \cdot (EO(DEFX / DEFGDP)), \quad (1.8)$$

де M_{96} – імпорт у цінах 1996 року;

DGP_{96} – ВВП у цінах 1996 року;

EO – валютний курс;
 $DEFM$ – дефлятор імпорту, 1996 рік = 1;
 $DEFGDP$ – дефлятор ВВП, 1996 рік = 1.
 А імпорт з Німеччини до України описує модель:

$$\begin{aligned} \text{LOG}(1/MG_{90}) = & 4.082 + 1.014 \cdot \text{LOG}(1/BIP_{90}) - 0.18 \cdot \text{LOG}(1/(PMG/PBIT)) + \\ & + 0.009 \cdot \text{TREND} + 0.22F_{92}, \end{aligned} \quad (1.9)$$

де MG_{90} – загальний імпорт з Німеччини у цінах 1990 р.;

BIP_{90} – ВВП України у цінах 1990 р.;

PMG – дефлятор загального імпорту з Німеччини;

$PBIT$ – дефлятор ВВП України;

$TREND$ – фактор часу;

F_{92} – екзогенна фіктивна змінна.

В підручнику [16] розглядають функцію імпорту туристичних послуг:

$$im_r^d = a_0 + a_1 y_r^f + a_2 (E_{d/f} (P_{im}/P_d)) + a_3 D_e, \quad (1.10)$$

де im_r^d – попит на реальний імпорт;

y_r^f – реальний ВВП;

$E_{d/f}$ – номінальний валютний курс (од.нац.вал./дол.), середній за період;

P_{im} – дефлятор ВВП, або індекс споживчих цін у країнах-контрагентах;

P_d – індекс внутрішніх цін;

D_e – надлишковий попит: реальний ВВП мінус трендовий реальний ВВП.

В роботі [9] імпорт Росії по окремій групі товарів розраховано за моделлю:

$$imRDA = -1035.79156 + 1.44 \cdot preVT - 10.09 \cdot imprDA + 19.22 \cdot sezimRDA, \quad (1.11)$$

де $imRDA$ – імпорт харчових продуктів і тютюну в цінах 2003 року;

$preVT$ – споживання домогосподарств у цінах 2003 року;

$imprDA$ – індекс імпортних цін на харчові продукти;

$sezimRDA$ – сезонний фактор.

1.3. Фактори цін експорту та імпорту

В роботі [9] імпортні та експортні ціни Росії по окремих групах товарів описує модель:

$$imprDN = 1.35 + 0.68 \cdot dpce(3) - 0.03 \cdot rateusdm(3), \quad (1.12)$$

де $imprDN$ – індекс імпортних цін на меблі;

$dpce(3)$ – дефлятор споживання домогосподарств з лагом 3 квартали;

$rateusdm(3)$ – номінальний курс рубля до долара з лагом 3 квартали.

$$\exp rDJ = -33.47 + 0.12 \cdot asteelpw + 0.08 \cdot alpw(2), \quad (1.13)$$

де $\exp rDJ$ – індекс експортних цін на продукцію металургії;

$asteelpw$ – ціна на арматурну сталь ФО ЄС, дол./т;

$alpw(2)$ – ціна на алюміній ЛБМ, дол./т.

1.4. Поточний рахунок і ВВП

Розглянемо модель ВВП:

$$BIP = c \cdot (BIP - T) + G + I + X - m \cdot BIP, \quad (1.14)$$

де BIP – ВВП;

$BIP = c$ – гранична схильність до споживання;

T – чисті податки;

G – державні витрати;

I – приватні інвестиції;

X – експорту;

M – частка імпорту у ВВП.

Після модифікації, дістанемо:

$$BIP = -c \cdot T / (1 - c + m) + G / (1 - c + m) + (I + X) / (1 - c + m), \quad (1.15)$$

Тоді в роботі [32] показано, що валовий національний використовуваний дохід і ВВП пов'язані співвідношенням:

$$BII = BIP + Yf + Trf, \quad (1.16)$$

де BII – валовий національний використовуваний дохід; BIP – ВВП;

Yf – чисте надходження доходів;

Trf – чисте надходження зовнішніх трансфертів.

Залежність зростання реального ВВП України від торговельного балансу та тягаря зовнішнього боргу:

$$\begin{aligned} GDP = & -0.01328INF_ENERGY - 15.7808 \cdot DISBALANSBANKS_{t-1} + \\ & + 0.3178EXP_IMP + 0.0290 \cdot REAL_VITRATI - 0.1497 \cdot REFINRATE_{t-1} - \\ & - 0.2023 \cdot DEBTSERV_{t-3} + 92.6171, \end{aligned} \quad (1.17)$$

де GDP – зростання реального ВВП, %;

INF_ENERGY – індекс цін виробників у паливній промисловості (до 1996 року) або індекс цін виробників у сфері виробництва і розподілу електроенергії, газу і води (з 1997 року), % грудень до грудня попереднього року;

$DISBALANSBANKS_{t-1}$ – коефіцієнт дисбалансів у національній банківській системі (співвідношення загальних обсягів депозитів і кредитів за всіма видами) – див. [20];

EXP_IMP – коефіцієнт покриття товарного імпорту експортом товарів %;

$REAL_VITRATI$ – зростання реальних кінцевих споживчих витрат населення, % (проте коефіцієнт при цій змінній статистично незначущий);

$REFINRATE_{t-1}$ – ставка рефінансування комерційних банків НБУ, середньозважена за всіма інструментами, %;

$DEBTSERV_{t-3}$ – коефіцієнт обслуговування довгострокового зовнішнього боргу (витрати на обслуговування зовнішнього боргу / експорт товарів та послуг), %;

Результати отримані в роботі [4] Оцінка робилася за вибіркою 1991 – 2008 рр.

1.5. Поточний рахунок, активи та заощадження

В роботі [70] поточний рахунок платіжного балансу та чисті зовнішні активи:

$$\frac{CA_t}{Y_t} = \frac{TB_t}{Y_t} + r_t \cdot \frac{NFA_{t-1}}{Y_t}, \quad (1.18)$$

де CA – поточний рахунок платіжного балансу;

TB – торговельний баланс;

r – номінальна процентна ставка;

NFA – чисті закордонні активи;

Y – номінальний ВВП.

В роботі [32] описана модель чистих внутрішніх заощадження і поточний рахунок платіжного балансу:

$$(SPR - IANPR) + (T - G) = X - M + Yf + Trf = CA, \quad (1.19)$$

де SPR – приватні заощадження;

$IANPR$ – приватні інвестиції;

T – чисті податки;

G – державні витрати;

X – експорт;

M – імпорт;

Yf – чисте надходження доходів;

Trf – чисте надходження зовнішніх трансфертів;

CA – поточний рахунок.

Модель впливу платіжного рахунку на організовані заощадження населення отримано в роботі [6]:

$$Sorg = -0.37 + 0.308 \cdot \pi + 0.167 \cdot Sunorg + 2.417 \cdot curr_acc + 0.359 \cdot Sorg_{t-1}, \quad (1.20)$$

де $Sorg$ – схильність до організованих заощаджень населення, тобто відношення таких заощаджень до доходів населення (вклади у банки та небанківські депозитні установи, придбання цінних паперів, вкладення у страхові поліси);

$Sunorg$ – схильність до неорганізованих заощаджень населення (готівку тобто грошовий агрегат $M0$);

π – інфляція;

$curr_acc$ – сальдо рахунку операцій з капіталом і фінансових операцій, % від ВВП.

Всі часові ряди сезонно згладжено. В моделі охоплено період 1999:2 кв. – 2008:1 кв.

1.6. Фінансовий рахунок

Грошова маса і валютні резерви пов'язані співвідношенням:

$$dDCg + dCp + dR = dM, \quad (1.21)$$

де $dDCg$ – зміна кредиту державному сектору;

dCp – зміна кредиту приватному сектору;

dR – зміна валютних резервів

dM – зміна обсягу грошей.

В роботі [32] описана модель залежність зростання реального ВВП України від зовнішніх позик:

$$GDP = 0.08 \cdot INVEST_{t-1} + 0.09 \cdot LOAN + 0.68 \cdot AR(4), \quad (1.22)$$

де GDP – зростання реального ВВП порівняно з попереднім кварталом, (%);

$INVEST_{t-1}$ – зростання інвестицій в основний капітал порівняно з попереднім кварталом з

лагом у 1 період, (%);

$LOAN$ – залучення іноземного довгострокового капіталу, % від ВВП;

$AR(4)$ – компонент авто регресії 4-го порядку для урахування фактору сезонності.

Оцінка в розрахунках робилася за вибіркою 1997:1 кв. – 2006:3 кв.

Залежність реального ефективного курсу гривні від чистого припливу портфельних інвестицій:

$$REER = 0.21 \cdot REER_{t-1} + 0.40 \cdot PLS + 0.94 \cdot MA(4), \quad (1.23)$$

де $REER$ – зміна реального ефективного курсу гривні за квартал, %;

$REER_{t-1}$ – те ж у попередньому періоді;

PLS – сальдо портфельних інвестицій в Україну за квартал, % ВВП;

$MA(4)$ – компонент плинної (ковзної) середньої 4-го порядку.

Оцінка робилася за вибіркою 1996:2 кв. – 2006:3 кв.

Залежність зростання грошової маси в Україні від припливу іноземного капіталу:

$$M2M = 0.04 \cdot GDP + 94.79 \cdot FDI + 335.9 \cdot LOANpr, \quad (1.24)$$

де $M2M$ – реальний приріст грошової маси за квартал (у цінах 1995 р.);

GDP – реальний ВВП за квартал (у цінах 1995 р.), млн. грн.;

FDI – приріст ПІІ в Україну за квартал, % ВВП;

$LOANpr$ – нетто-залучення іноземного позикового капіталу корпоративним сектором. економіки (включаючи банки), % ВВП.

Оцінка робилася за вибіркою 1996:2 – 2006:3.

Інша специфікації цієї моделі:

$$M2M = -17.44 \cdot INT + 98.48 \cdot FDI + 334.3 \cdot LOANpr + 1449, \quad (1.25)$$

де INT – процентна ставка на кінець кварталу по гривневих кредитах, % річних. Оцінка робилася за вибіркою 1996:2 кв. – 2006:3 кв.

Залежність фондового індексу ПФТС від впливу капіталу за кордон:

$$PFTS = 0.002 \cdot GDP_{t-1} - 1.35 \cdot OUTFLOW - 1.05 \cdot DEPOSIT_2 - 26.55, \quad (1.26)$$

де $PFTS$ – зміна індексу ПФТС за квартал, %;

GDP_{t-1} – реальний ВВП у попередньому кварталі (у цінах 1995 р.), млн. грн.;

$OUTFLOW$ – приріст фінансових активів за прямими, портфельними та іншими інвестиціями за квартал, % ВВП;

$DEPOSIT_2$ – приріст строкових депозитів у банках України за квартал, % ВВП.

Ці результати отримані в роботі [5]. Оцінка робилася за вибіркою 1997:3 кв. – 2006:3 кв.

1.7. Валютний курс і процентні ставки

В роботі [58] визначення рівноважного валютного курсу відбувається згідно монетарної моделі:

$$s_t = Ms_t^* - ms_t + \varphi y_t - \varphi^* y_t^* - \lambda r_t + \lambda^* r_t^*, \quad (1.27)$$

де s – номінальний валютний курс;

ms – пропозиція грошей;

y – ВВП;

r – номінальна процентна ставка;

φ, λ – регресійні коефіцієнти;

* – означає іноземні показники;

Всі змінні крім процентної ставки – натуральні логарифми.

Визначення процентної ставки за режиму таргетування грошової маси та валютного курсу:

$$rs_t = rs_{t-1} + a(m_t - m_t^*) + b(er_t - er_t^*), \quad (1.28)$$

де rs – короткострокова номінальна процентна ставка;

m – логарифм грошових балансів;

er – номінальний валютний курс;

a, b – коефіцієнти для цілей таргетування грошової маси та валютного курсу; * - означає цільовий рівень, а не фактичний.

Визначення процентної ставки за режиму таргетування інфляції:

$$rs_t = rr_t^* + \pi_t^e + m(\pi_t - \pi_t^*) - v(y_t - y_t^*), \quad (1.29)$$

де rs – короткострокова номінальна процентна ставка;

rr – реальна процентна ставка;

π^e - очікувана інфляція;

π – інфляція;

y – розрив ВВП (output gap);

m, v – коефіцієнти для цілей таргетування інфляції та розриву ВВП;

* - означає цільовий рівень, а не фактичний.

Результати (1.28) і (1.29) отримані в роботі [70].

1.8. Міграція робочої сили

Залежність трудової еміграції від показників ринку праці:

$$EM = 113887.7 + 40.45 \cdot WAGE - 0.004 \cdot POPUL + 600.4 \cdot UNEMPL, \quad (1.30)$$

де EM – річна кількість емігрантів;

$WAGE$ – номінальна заробітна плата;

$POPUL$ – економічно активне населення;

$UNEMPL$ – навантаження незайнятих трудовою діяльністю громадян на одне вільне робоче місце.

Оцінка робилася за вибіркою 1996-2006 рр.

Результат (1.30) отримано в роботі [17].

ГЛАВА 2. Дослідження міжнародних економічних відносин методами кореляційного аналізу

2.1. Основи кореляційного аналізу

В економічних дослідженнях однією з основних задач є аналіз залежностей між змінними. Але довільна залежність в певній мірі є абстракцією, оскільки в оточуючому світі, частиною якої є світова економіка, значення конкретної величини не визначається незмінною формулою її залежності від набору інших величин. Завжди є декілька величин, які визначають головні тенденції зміни величини, яка розглядається, і в економічній теорії і практиці обмежуються тим чи іншим колом таких величин (пояснювальних змінних). Але завжди існує взаємодія великої кількості інших, менш важливих або важко ідентифікованих факторів, які приводять до відхилень значень пояснювальної (залежної змінної) від конкретної формули її зв'язку з пояснювальними змінними, наскільки точною ця формула не була. Знаходження, оцінка і аналіз таких зв'язків, ідентифікація пояснювальних змінних, побудова формул залежності і оцінка їх параметрів є свого роду мистецтво, що враховує в кожній конкретній галузі знань (зокрема в міжнародних економічних відносинах) її внутрішні закони та потреби.

Нехай потрібно оцінити зв'язок між змінними X та Y (наприклад, зв'язок показників безробіття і інфляції в даній країні за певний проміжок часу). Зокрема може стояти питання, чи пов'язані між собою ці показники, і при позитивній відповіді на це питання, природно постає задача знаходження формули цього зв'язку. Основою для відповіді на це запитання є статистичні дані про динаміку цих показників (річні, квартальні, місячні і т.п.). Ці дані утворюють деяку випадкову вибірку з генеральної сукупності, тобто з сукупності всіх можливих показників інфляції та безробіття в даних умовах. Питання про наявність зв'язку між економічними змінними постає як питання про наявність конкретної формули (специфікації) такої залежності, стійкої до числа спостережень. Нехай, наприклад є припущення про те що дві змінні пов'язані між собою лінійною залежністю, тобто:

$$Y = a_0 + a_1 X \quad (2.1)$$

Розглянемо спочатку питання про лінійну залежність двох змінних:

- 1) Чи пов'язані між собою лінійно змінні X і Y ?
- 2) Яка формула зв'язку змінних X і Y ?

Для відповіді на поставлені запитання існують спеціальні статистичні методи, і відповідно показники, значення яких спеціальним чином (з певною ймовірністю) свідчать про наявність або відсутність лінійного зв'язку між змінними. В першому випадку це коефіцієнт кореляції величин X та Y , в другому – коефіцієнти лінійної регресії a_0 та a_1 , їхні стандартні похибки і t -статистики, за значеннями яких перевіряється гіпотеза про відсутність зв'язку величин X та Y .

Пояснимо логіку появи такого показника, як коефіцієнт кореляції. Припустимо, що між змінними X та Y існує лінійна залежність. Наявність такої залежності можна інтерпретувати таким чином. Якщо змінна X приймає значення більші, ніж її середнє значення, і зв'язок додатний (на мові формул це означає, що коефіцієнт $a_1 > 0$), то значення змінної Y також повинно бути більше її середнього значення і співвідношення відхилень X та Y від їх середніх значень повинно бути сталим. Якщо змінна X приймає значення менші, ніж її середнє значення, то значення змінної Y також повинно бути менше її середнього значення з тим самим коефіцієнтом пропорційності цих відхилень. Якщо зв'язок між змінними X та Y від'ємний, то додатне відхилення X від її середнього значення повинно узгоджуватись з від'ємним відхиленням Y від її середнього значення, і навпаки. Якщо лінійної залежності між змінними X та Y немає, то додатні відхилення змінної X від її середнього значення повинно узгоджуватись як з додатними так і від'ємними відхиленнями Y від її середнього значення. Те саме можна сказати і про від'ємні відхилення X від її середнього значення. За міру ступеня лінійного зв'язку між двома змінними використовують коефіцієнт їх кореляції. Він обчислюється за формулою:

$$r = \frac{\frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})(y_k - \bar{y})}{\sqrt{\frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2} \cdot \sqrt{\frac{1}{n} \sum_{m=1}^n (y_m - \bar{y})^2}} \quad (2.2)$$

В цій формулі x_i, y_j - значення n змінних X та Y відповідно, а \bar{x}, \bar{y} - їх середні арифметичні.

Парний коефіцієнт кореляції Пірсона (2.2) є безрозмірною величиною, тобто його величина не залежить від одиниць вимірювання обох змінних. Він приймає значення $|r| \leq 1$, тобто змінюється від -1 у випадку строгої лінійної від'ємної залежності до +1 у випадку строгої лінійної додатної залежності. Якщо він дорівнює нулю (або дуже близький до нуля), то це означає, що між змінним відсутній лінійний зв'язок (але залишається ймовірність, що вони пов'язані нелінійним зв'язком). Чим ближче кореляція до -1 чи 1, тим вона є сильнішою. Від'ємне значення коефіцієнта кореляції означає, що зі збільшенням значень однієї змінної значення іншої змінної, зазвичай, зменшуються. Додатне значення коефіцієнта кореляції означає, що зі збільшенням значень однієї змінної значення іншої змінної, зазвичай, збільшуються.

Прийнято вважати, що якщо $|r| \leq 0.25$, то кореляція слабка, якщо $0.25 < |r| \leq 0.75$, то кореляція середня, якщо $|r| > 0.75$ - кореляція сильна. Якщо $r=0$, то говорять що змінні некорельовані.

Візуально силу кореляції між двома змінними можливо представити за допомогою діаграми розсіювання, де по осях відкладені значення двох змінних. Якщо форма сукупності крапок-спостережень наближається до прямої лінії, кореляція є сильною. Якщо – до круглої хмари – кореляція є слабкою.

Нижче на діаграмах розсіювання показані приклади: сильної негативної кореляції (Рис.2.1), сильної позитивної кореляції (Рис.2.2), майже відсутньої кореляції (Рис.2.3):

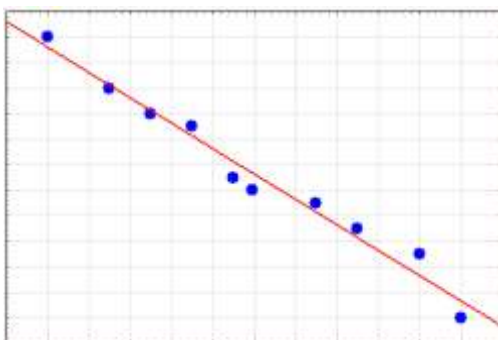


Рис.2.1.

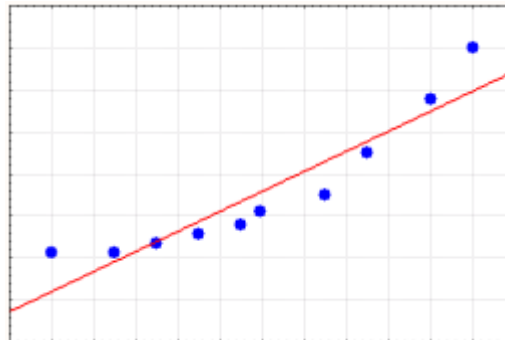


Рис.2.2.

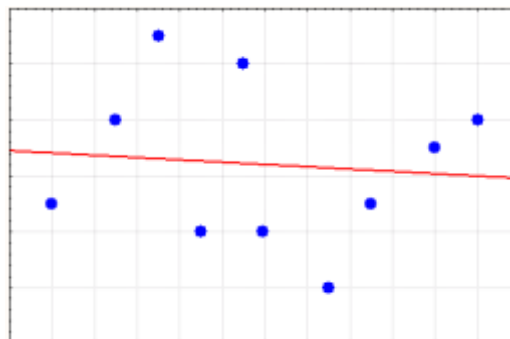


Рис.2.3.

Далі, при аналізі коефіцієнта кореляції виникає наступне питання. Якщо він дорівнює нулю для генеральної сукупності, це зовсім не означає, що він також дорівнює нулю для вибірки. Навпаки він буде обов'язково відхилитися від справжнього значення, але чим більше відхилення, тим менш ймовірно воно при даному обсягу вибірки. Для вибіркового коефіцієнта кореляції r

будується t -статистика, яка має розподіл Стюдента з $n - 2$ ступенями вільності, і розраховується за формулою:

$$t = r \cdot \frac{\sqrt{n-2}}{\sqrt{1-r^2}} \quad (2.3)$$

Існує тест для визначення значущості коефіцієнта кореляції (наскільки є впевненість, що він відрізняється від нуля). Перевіряється нульова гіпотеза H_0 про те, що істинне значення коефіцієнта кореляції дорівнює нулю: $r = 0$. Альтернативною гіпотезою H_1 є гіпотеза про те, що $r \neq 0$. Порівнюючи обчислене за вибіркою значення t -статистики за формулою (2.3) з критичними точками, які визначаються за таблицями розподілу Стюдента ми можемо прийняти або відхилити нульову гіпотезу. Для двосторонньої критичної області заданого рівня значущості α критична точка t_{kp} знаходиться з таблиці як $t_{kp} = t_{\alpha/2, m}$ для кількості ступенів вільності $m = n - 2$. Якщо $|t| \leq t_{kp}$, то гіпотеза H_0 приймається, якщо $|t| > t_{kp}$ – гіпотеза H_0 відхиляється.

Рівень значущості $\alpha = 1 - p$ – це імовірність здійснити помилку 1-го роду, тобто відхилити нульову гіпотезу, коли вона є насправді вірною. В даному випадку – ймовірність того, що ми вважатимемо кореляцію ненульовою, коли вона насправді дорівнює нулю.

Число α задають наперед, і найчастіше його обирають рівним 0,1; 0,05; 0,01; 0,001. Значення $\alpha = 0,05$ означає, що ймовірність допустити помилку першого роду є малою: ми ризикуємо її допустити у 5 випадках зі 100.

Як правило, допустимим рівнем значущості є 0.05 і менше (менше 0.10 – менш суворий критерій, менше 0.01 – більш суворий). Тоді кореляція вважається статистично значущою. Значущість коефіцієнту кореляції тим більша, чим сильніше його розраховане значення відрізняється від нуля і чим більша вибірка спостережень, за якою він розрахований.

Якщо мова йде про взаємозв'язок однієї змінної Y з декількома m змінними $X_i, i = 1, 2, \dots, m$, то використовують коефіцієнт множинної кореляції R , як міру тісноти взаємозв'язку однієї змінної Y (показника) з сукупністю інших змінних (факторів) $X_i, i = 1, 2, \dots, m$. Його обчислюють за формулою:

$$R = \frac{\frac{1}{n} \sum_{k=1}^n (y_k - \bar{y})(\hat{y}_k - \bar{\hat{y}})}{\sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \bar{y})^2} \cdot \sqrt{\frac{1}{n} \sum_{m=1}^n (\hat{y}_m - \bar{\hat{y}})^2}} \quad (2.4)$$

Квадрат коефіцієнта множинної кореляції R^2 називають коефіцієнтом детермінації. Він показує, скільки відсотками варіація залежної змінної Y визначається варіацією незалежних змінних $X_i, i = 1, 2, \dots, m$. Коефіцієнт детермінації приймає значення $R^2 \in (0, 1)$. Чим ближче його значення до одиниці, тим істотніший зв'язок між змінними в моделі.

Взаємозв'язок між двома змінними, обчисленими при фіксованих значеннях всіх інших змінних, характеризується коефіцієнтом частинної кореляції. Тоді записують кореляційну матрицю:

$$r = \begin{pmatrix} r_{yy} & r_{yx_1} & \dots & r_{yx_m} \\ r_{x_1y} & r_{x_1x_1} & \dots & r_{x_1x_m} \\ \dots & \dots & \dots & \dots \\ r_{x_my} & r_{x_mx_1} & \dots & r_{x_mx_m} \end{pmatrix}, \quad (2.5)$$

де r_{y_k} - парні коефіцієнти кореляції між залежною і незалежними змінними, а r_{x_k, x_i} - парні коефіцієнти кореляції між незалежними змінними.

2.2. Побудова кореляційної матриці для зовнішньоекономічних показників

В Microsoft Excel є можливість скористуватися опцією *Корреляция* у надбудові *Пакет аналіза* для побудови кореляційної матриці для декількох змінних. Увага, для зручності ми включили у вхідній діапазон комірки з назвами змінних, але при цьому включили опцію *Метки в первой строке*. Також ми включили опцію *Группирование по столбцам*, так як змінні у нас розташовані по стовпчиках, а не рядках. Нажаль, ми не можемо автоматично порахувати рівень значущості коефіцієнтів кореляції. На Рис.2.6.-2.8 вказані вхідні дані, діалогове вікно аналізу та результат.

	A	B	C	D
1	Експорт	Імпорт	ПІІ	Зовнішній борг
2	12	12	3	10
3	22	34	4	35
4	33	23	5	30
5	21	22	2	40
6	34	15	4	50
7	56	22	2	15
8	43	13	1	20

Рис.2.6.

Або можливо скористатися функцією:

=КОРРЕЛ(діапазон першої змінної;діапазон другої змінної)

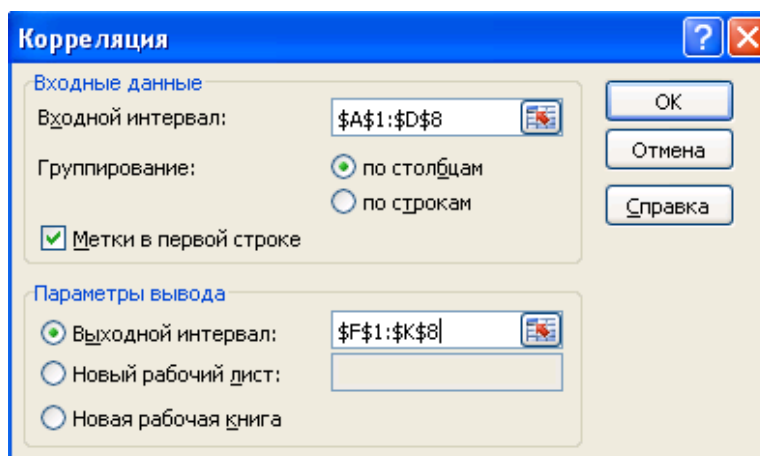


Рис.2.7.

	F	G	H	I	J
	Експорт	Імпорт	ПІІ	Зовнішній борг	
Експорт	1				
Імпорт	-0.05517	1			
ПІІ	-0.32556	0.385287	1		
Зовнішній	-0.15202	0.290725	0.410601	1	

Рис.2.8.

Функція повертає звичайний коефіцієнт кореляції Пірсона між двома змінними, значення яких прописані у відповідних двох діапазонах. Діапазони мають містити однакову кількість комірок (бути одного розміру). Приклад: =КОРРЕЛ(A2:A25;B2:B25)

2.3. Подолання проблем у застосуванні кореляційного аналізу світової економіки

Кореляційний аналіз може бути пов'язаний з деякими труднощами.

1. Нелінійний зв'язок. При певних видах нелінійних зв'язків (наприклад квадратична залежність), навіть якщо вони сильні, коефіцієнт кореляції може бути низьким або навіть дорівнювати 0. При інших видах нелінійної залежності (наприклад, експоненціальна) коефіцієнт кореляції буде високим, хоча не наблизатиметься до 1.

Вирішенням цієї проблеми є трансформація змінних: наприклад, розрахунок кореляції між логарифмом однієї змінної та іншою змінною; між однією змінною та кубом іншої змінної; між однією змінною та квадратом відхилення від середньої іншої змінної тощо. Тобто спочатку ми визначаємо, яка функція найкраще відображає залежність, а потім робимо перетворення з однією чи обома змінними. Визначити таку функцію може допомогти діаграма розсіювання. Другий варіант – використання непараметричних аналогів звичайного коефіцієнту кореляції Пірсона (наприклад, коефіцієнту кореляції Спірмена). Третій варіант перетворити одну зі змінних на категоріальну (розбити всі значення на групи: наприклад, великі, середні і малі) і визначити як відрізняються середні значення другої змінної в цих групах за допомогою методів аналізу середніх чи дисперсійного аналізу.

2. Економічні дані, як і будь-які інші, можуть мати грубі помилки (аномальні значення). Викиди (впливові спостереження). Вони можуть штучно завищити чи занижити значення коефіцієнту кореляції, або навіть призвести до того, що кореляція буде не того знаку. Наприклад, випадок, коли, якщо ми включаємо до аналізу поряд звичайними спостереженнями викиди, кореляція буде позитивною, а вилучення 1-2 викидів з аналізу призводить до того, що кореляція стає негативною.

На Рис. 2.9-2.10 наведемо графічний приклад: дві діаграми розсіювання для тієї самої пари змінних з викидом та без нього. Як бачимо кореляція з викидом позитивна (0.51), а без нього слабо негативна (-0.24).

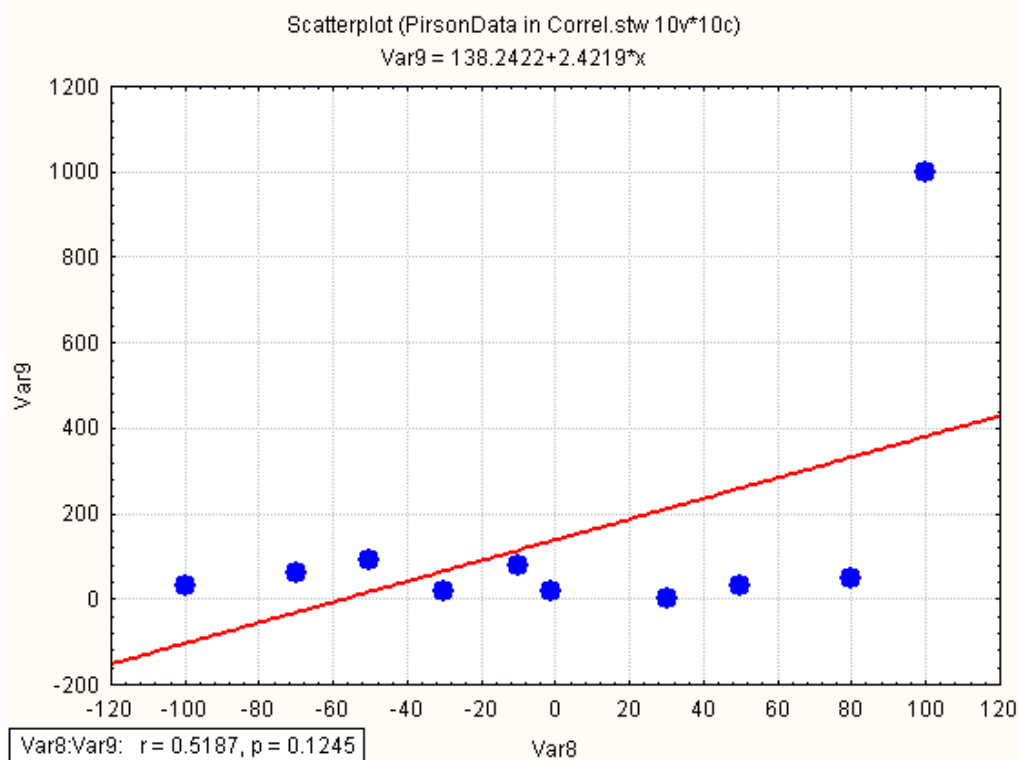


Рис.2.9.

Для подолання проблеми викидів можна запропонувати такі методи:

- обрати велику вибірку, де викиди вже не створюють великого впливу;
- провести кореляційний аналіз з викидами і без них, порівнюючи результати (наприклад, враховувати середній коефіцієнт кореляції як кращу оцінку справжньої кореляції для генеральної сукупності);

- вилучити з аналізу викиди, якщо вони є результатом помилки чи унікальної події, яка у майбутньому навряд чи повторюватиметься.

3. Кореляції у неоднорідних групах. Наприклад, якщо велике значення коефіцієнта кореляції викликано даними з зовсім неоднорідних груп. Припустимо, у нас є дві групи країн: високорозвинуті та слаборозвинуті. При цьому між двома показниками (зростання ВВП та інфляція) у високорозвинутих країнах існує слабопозитивний зв'язок, а у слаборозвинутих країнах – сильний негативний. Якщо розрахувати коефіцієнт кореляції для всієї сукупності країн, він буде ближче

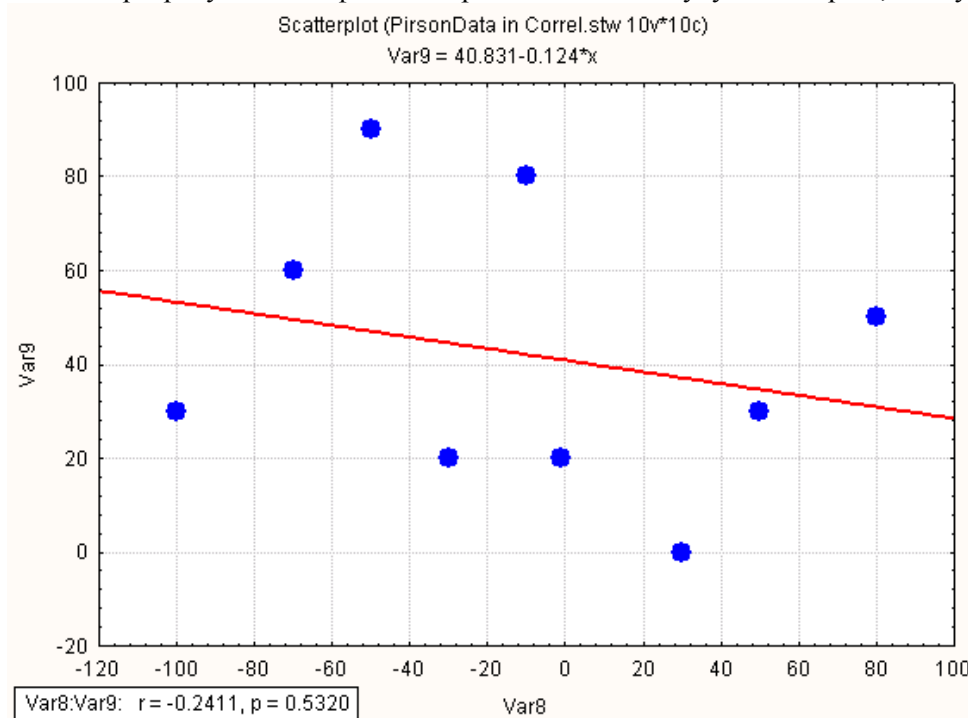


Рис.2.10.

до нейтрального. Але цей результат не буде з практичної точки зору корисним. Для слаборозвинутих країн буде недооціненою небезпека інфляції, а для розвинутих країн – дефляції.

Графічно проілюструвати цей приклад можливо на Рис.2.11 (вхідні дані) і 2.12 (графік).

10 GDPgr	11 Inflation	12 Development
-1	0	High
1	1	High
1.5	2	High
3	2.5	High
3	4	High
9	3	Low
10	4	Low
6	9	Low
5	20	Low
2	25	Low

Рис.2.11.

Для всіх країн кореляція дуже слабка (0.0588). Але якщо виділити дві групи країн (кола – високо розвинуті, трикутники – слаборозвинуті), ми побачимо, що залежність є сильною, але неоднаковою по групах країн.

Вирішенням проблеми неоднорідних груп може бути використання діаграм розсіювання для ідентифікації різних груп спостережень за іншою змінною, або розрахунок коефіцієнтів кореляції для різних груп спостережень без побудови діаграми розсіювання. Використання на додаток кластерного аналізу (див. главу 12) для визначення різних груп спостережень також може бути корисним.

4. Випадкові кореляції при масовому аналізі. Якщо будується кореляційна матриця (таблиця, де рядкам і стовпчикам відповідають змінні, на перетині кожного рядка та стовпчика знахо-

диться коефіцієнт кореляції між відповідною парою змінних) для багатьох змінних, ймовірно певна частина високих кореляцій буде високими лише випадково, коли насправді (в генеральній сукупності) такої кореляції між відповідними парами змінних не буде. Наприклад, коефіцієнт кореляції значущий на рівні 0.05 може бути випадковим із ймовірністю 5%. Тому, якщо у нас, наприклад, у кореляційній матриці 60 коефіцієнтів кореляцій із рівнем значущості 0.05, ймовірно, 3 з них відрізняються від нуля лише випадково внаслідок збігу обставин.

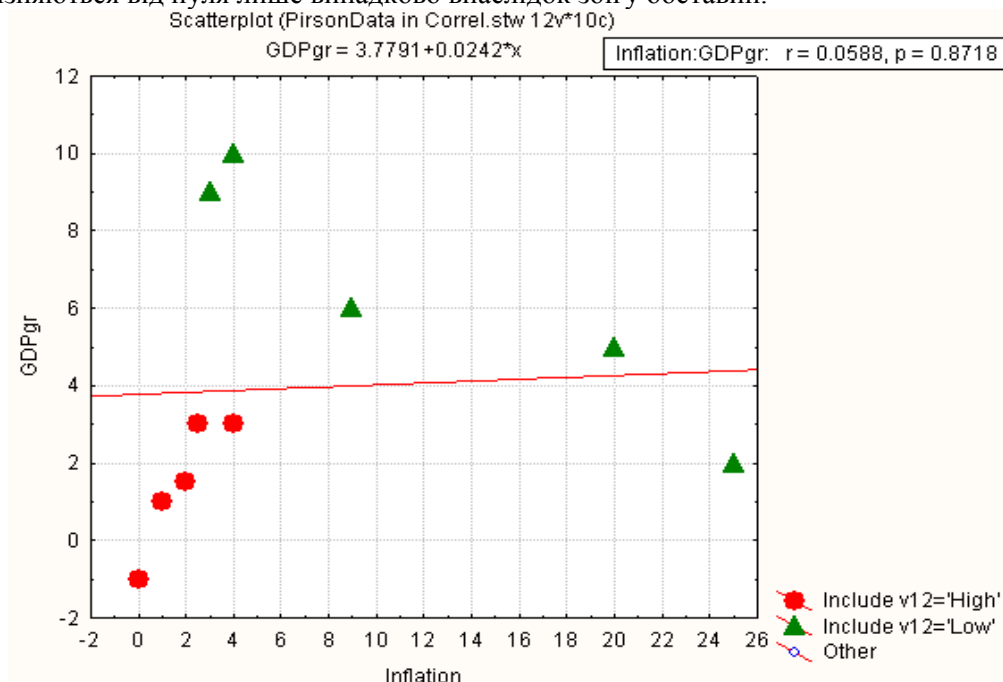


Рис.2.12.

Саме тому побудова кореляційної матриці – це лише попередній крок при аналізі зв'язку між змінними, а не остаточний доказ наявності зв'язку. Аналіз, який проводиться без теоретичного обґрунтування називається розвідковим (*exploratory*). Якщо виявлений високий коефіцієнт кореляції між змінними, але йому немає теоретичного пояснення, варто ставитися обережно до такого результату. Доцільно перевірити наявність кореляції на іншій вибірці чи на підвибірках цієї ж вибірки.

5. Відсутні дані. Звичайним способом вирішення проблеми з відсутніми даними при кореляційному аналізі є попарне (*pairwise*) видалення спостережень з відсутніми даними (оскільки повне (*casewise*) видалення спостережень, по яких відсутнє значення принаймні по одній змінній, може призвести до суттєвого зменшення вибірки). Але може виникнути небезпека систематичного місцезнаходження відсутніх даних (спостереження з відсутніми даними залежать від однієї чи декількох змінних). У результаті середнє значення певної змінної А, яке використовується для розрахунку її кореляції зі змінною В, може відрізнитися від середнього значення А, яке використовується для розрахунку її кореляції зі змінною С. В результаті коефіцієнти кореляції між різними змінними можуть бути непорівнюваними (враховуючи також те, що вони обраховані за різною кількістю спостережень). Ще один спосіб (заміна відсутніх даних на середні значення змінних) призводить до заниження кореляції та дисперсії. Підсумовуючи, скажемо, що оптимальним все ж є попарне видалення відсутніх даних. Але можливо спробувати всі три способи і порівняти результати.

6. Структурні зміни. З часом зв'язки між показниками можуть змінюватися. Наприклад, існують періоди, коли курс долара позитивно корелює з цінами на нафту, і періоди, коли негативно. Це пояснюється, наприклад, яким чином витрачають доларові надходження нафтовидобувні країни свої доларові надходження від нафти: вкладають у доларові активи чи купують імпорتنі товари, розраховуючись в інших валютах, зокрема євро. Тому розрахований коефіцієнт кореляції в один період часу може не мати практичної корисності в інший період часу.

Варіантом вирішення цієї проблеми є використання плинних коефіцієнтів кореляції (за аналогією з плинною середньою), наприклад, за 1990-2000, 1991-2001, 1992-2002 рр. і т.д. Або поділ всього досліджуваного періоду на підперіоди, наприклад, 1971-2010 рр. на 1971-1980, 1981-1990,

1991-2000, 2001-2010 рр. Це дозволить прослідкувати як коефіцієнт кореляції змінюється з часом і оцінити наскільки він є надійним.

Наприклад, візьмемо дані по поточному рахунку у % ВВП та приростом реального ВВП (за даними World Economic Outlook) по Україні за 1995-2010 рр. Розрахуємо плинну кореляцію і кореляції за три підперіоди як на Рис. 2.13 (вхідні дані та результат) і 2.14 (формули).

Як бачимо кореляція була достатньо стійкою в усі роки крім останніх, коли вона навіть змінила знак.

	A	B	C	D	E
1	Рік	Приріст ВВП	ПР/ВВП	Плинна кореляція	Кореляції за
2	1995	-12.151	-1.152	(останні 5 років)	3 підперіоди
3	1996	-10.044	-1.184		
4	1997	-2.988	-1.335		
5	1998	-1.949	-1.296		
6	1999	-0.224	1.658	0.507436458	
7	2000	5.85	1.481	0.71467758	0.71038821
8	2001	9.046	1.402	0.730392507	
9	2002	5.253	3.173	0.58181892	
10	2003	9.595	2.891	0.209615127	
11	2004	12.019	6.909	0.652640041	
12	2005	2.944	2.531	0.535655522	0.53565552
13	2006	7.534	-1.617	0.441221415	
14	2007	7.518	-5.272	0.373767736	
15	2008	1.945	-12.763	0.647517052	
16	2009	-14.462	-1.732	-0.112867883	
17	2010	4.187	-2.884	-0.182088832	-0.18208883

Рис.2.13.

D	E
Плинна кореляція (останні 5 років)	Кореляції за 3 підперіоди
=КОРРЕЛ(B2:B6;C2:C6)	
=КОРРЕЛ(B3:B7;C3:C7)	=КОРРЕЛ(B2:B7;C2:C7)
=КОРРЕЛ(B4:B8;C4:C8)	
=КОРРЕЛ(B5:B9;C5:C9)	
=КОРРЕЛ(B6:B10;C6:C10)	
=КОРРЕЛ(B7:B11;C7:C11)	
=КОРРЕЛ(B8:B12;C8:C12)	=КОРРЕЛ(B8:B12;C8:C12)
=КОРРЕЛ(B9:B13;C9:C13)	
=КОРРЕЛ(B10:B14;C10:C14)	
=КОРРЕЛ(B11:B15;C11:C15)	
=КОРРЕЛ(B12:B16;C12:C16)	
=КОРРЕЛ(B13:B17;C13:C17)	=КОРРЕЛ(B13:B17;C13:C17)

Рис.2.14.

7. Складність визначення причини та наслідку. Сам по собі коефіцієнт кореляції не відповідає на питання, коливання якої з двох змінних спричиняють коливання іншої, наприклад інфляція зміни валютного курсу чи навпаки.

Вирішенням цієї проблеми є використання лагів. Наприклад, припустимо, існує кореляція між інфляцією в поточному періоді та зміною валютного курсу у наступному періоді (чи через період, через два періоди тощо).


Це ймовірно означає, що інфляція спричиняє зміни валютного курсу. Проте без теоретичних міркувань неможливо визначити що є причиною, а що наслідком, у рамках одного й того ж періоду. Спеціальним методом визначення причинних зв'язків є тест Грейнджера (зокрема, доступний у програмі E-views).

8. Спільна причина. Несправжня кореляція може бути результатом дії спільної причини. Наприклад, можуть корелювати зміни валютних резервів двох країн, які розташовані в різних кінцях світу, не мають суттєвих прямих чи опосередкованих торговельних чи фінансових зв'язків між собою, які б могли пояснити достатній вплив однієї країни на іншу. А насправді ці коливання викликані коливаннями глобальної схильності інвесторів до ризику.

В цьому випадку важливим є теоретичне обґрунтування зв'язків і виявлення ймовірної спільної причини. Далі можливо поділити всі спостереження на дві групи: в умовах великої схильності інвесторів до ризику та низької. В межах кожної з груп, можливо, кореляція між змінами валютних резервів обох країн буде вже майже відсутньою. Інший спосіб вирішення проблеми використання часткової кореляції (з поправкою на вплив інших змінних).

9. Неадитивність. Коефіцієнти кореляції не є адитивними. Але вже коефіцієнти детермінації є адитивними. Хоча в цьому випадку може діяти спотворюючий ефект мультиколінеарності. Але є випадки коли може бути доцільним сумувати чи знаходити середнє коефіцієнтів кореляції (або детермінації), але одержане таким чином значення може вже й не інтерпретуватися як власне коефіцієнт кореляції. Наприклад, ми хочемо дізнатися, як корелює зростання ВВП різних країн, але при цьому надати більшу вагу останньому періоду. Ми можемо розрахувати кореляцію, наприклад, за 1990-2010 та за 2000-2010 рр. і знайти середнє. Хоча оптимальним рішенням, якщо дозволяє програмне забезпечення, розраховувати кореляцію за весь період 1990-2010 рр. вже з урахуванням штучної змінної, яка збільшується з часом і використовуватиметься для зважування спостережень. Наприклад, 1990-му року надаватиметься вага 10, 1991-му – 11, ..., 2010-му – 30.

Наведемо приклад, візьмемо дані по поточному рахунку у % ВВП та приростом реального ВВП (за даними World Economic Outlook) по Україні за 1995-2010 рр. Створимо штучну змінну спеціально для зважування. Її значення обираються довільно, але так, щоб вони були більшими в пізніші періоди (див. Рис. 2.15). Побудуємо кореляційну матрицю спочатку без зважування, а по-

тім зі зважуванням (останній варіант потребує натиснення кнопки  та заповнення опцій у вікні на Рис. 2.16).

	1	2	3
	GDPgr	CurAc_GDP	Weight
1995	-12.15	-1.15	10
1996	-10.04	-1.18	11
1997	-2.99	-1.34	12
1998	-1.95	-1.30	13
1999	-0.22	1.66	14
2000	5.85	1.48	15
2001	9.05	1.40	16
2002	5.25	3.17	17
2003	9.60	2.89	18
2004	12.02	6.91	19
2005	2.94	2.53	20
2006	7.53	-1.62	21
2007	7.52	-5.27	22
2008	1.95	-12.76	23
2009	-14.46	-1.73	24
2010	4.19	-2.88	25

Рис.2.15.

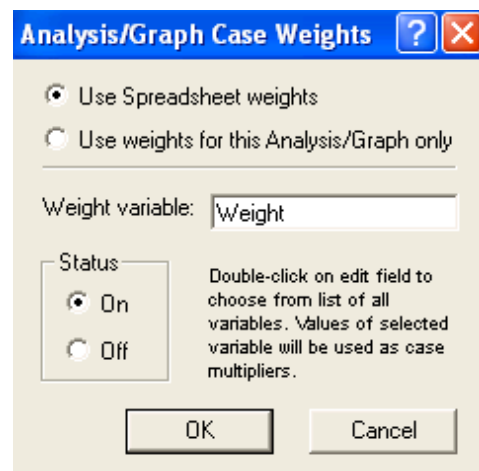


Рис.2.16.

Одержимо результати як на Рис.2.17-2.18. В даному прикладі вони лише трохи відрізняються (0.26 без зважування і 0.24 зі зважуванням). Проте рівень значущості зваженої кореляції рахується за іншою кількістю спостережень (280, а не 16), тому кореляція може бути позначена червоним кольором помилково, як значуща.

Correlations (DataWeighted in Correl.stw)				
Marked correlations are significant at p < .05000				
N=16 (Casewise deletion of missing data)				
Variable	GDPgr	CurAc_GDP		
GDPgr	1.00	0.26		
CurAc_GDP	0.26	1.00		

Рис.2.17.

Correlations (DataWeighted in Correl.stw)				
Marked correlations are significant at p < .05000				
N=280 (Casewise deletion of missing data)				
Variable	GDPgr	CurAc_GDP		
GDPgr	1.00	0.24		
CurAc_GDP	0.24	1.00		

Рис.2.18.

2.4. Непараметричні методи кореляційного аналізу зв'язку якісних змінних політичного клімату з припливом інвестицій

Коефіцієнт кореляції Пірсона є параметричним критерієм, оскільки припускає, що змінні кількісні і вимірюють в інтервальній шкалі, шкалі відношень, або в абсолютній шкалі і мають нормальний розподіл. Непараметричні аналоги коефіцієнту кореляції Пірсона дозволяють аналізувати зв'язки також між змінними, що вимірюються у метричній шкалі, якщо їх розподіл невідомий чи не є нормальним, або у порядковій шкалі. В цьому випадку використовують так звані рангові коефіцієнти кореляції. Якщо змінні кількісні, але їх розподіл не відповідає нормальному розподілу, також рекомендують використовувати рангові кореляції. Використовуються такі аналоги:

1. Кореляція Спірмена (ρ читається «ро»). Є прямим аналогом кореляції Пірсона, але розраховується не за абсолютними значеннями показників, а за їх рангами (порядковими номерами у переліках, відсортованих за відповідними змінними). Він обчислюється за формулою:

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}, \quad (2.6)$$

де $d = R_x - R_y$ - різниці між рангами змінних, які характеризують об'єкти вибірки, n - обсяг вибірки.

2. Коефіцієнт кореляції рангів Кендала (τ читається «тау») ґрунтується не на різницях рангів, як попередній коефіцієнт, а на відхиленнях кількості рангів, більшої від порівнюваної s_1 , і меншої від порівнювальної s_2 , за результативною ознакою Y :

$$\tau = \frac{2 \sum_{k=1}^n (s_1 - s_2)}{n(n-1)} \quad (2.7)$$

Коефіцієнти Спірмена і Кендала можуть набувати значень від -1 до $+1$, тобто характеризують як тісноту, так і напрям зв'язку. Якщо ρ і τ дорівнюють $+1$, то йдеться про повний прямий зв'язок між ознаками, якщо ρ і τ дорівнюють -1 , то про повний зворотній зв'язок між ними.

3. Гамма статистика краща за статистики ρ -Спірмена або τ -Кендала, якщо в даних є багато значень, які збігаються. Гамма-статистика – це ймовірність, точніше це різниця між ймовірністю того, що ранговий порядок двох змінних збігається і ймовірністю того, що він не збігається, поділена на вираз: 1 мінус ймовірність їх збігів.

Для даних, які вимірюються у номінальній шкалі, непараметричними аналогами можливо вважати χ^2 -квадрат, ϕ -коефіцієнт, точний критерій Фішера, коефіцієнт конкордації Кендалла.

Наведемо приклад. Припустимо у нас є дані про політичну стабільність та прямі іноземні інвестиції (див. 2.19). Перша змінна вимірюється у порядковій шкалі (бали).

2.5. Кореляційний аналіз як метод первинного дослідження

Для отримання перших уявлень про взаємозв'язки між змінними можливо провести кореляційний аналіз для залежної змінної та решти показників. Показники з високими значеннями коефіцієнту кореляції (звичайного, Пірсона) із залежною змінною варто далі використовувати для детальнішого аналізу. Але так ми врахуємо лише лінійні зв'язки. Для того, щоб одержати уявлення про наявність нелінійного зв'язку варто також розрахувати коефіцієнти кореляції Спірмена, а також коефіцієнти кореляції Пірсона між залежною змінною та квадратом відхилення від середнього значення кожного показника.

Наведемо приклад. Ми маємо змінні Y та X , між якими майже відсутній лінійний зв'язок. Якби ми користувалися лише звичайною кореляцією Пірсона для попереднього відбору змінних для подальшого аналізу, то ми би на цьому і зупинилися. Але з діаграми розсіювання у прикладі на Рис. 2.22 ми бачимо, що є явний нелінійний зв'язок:

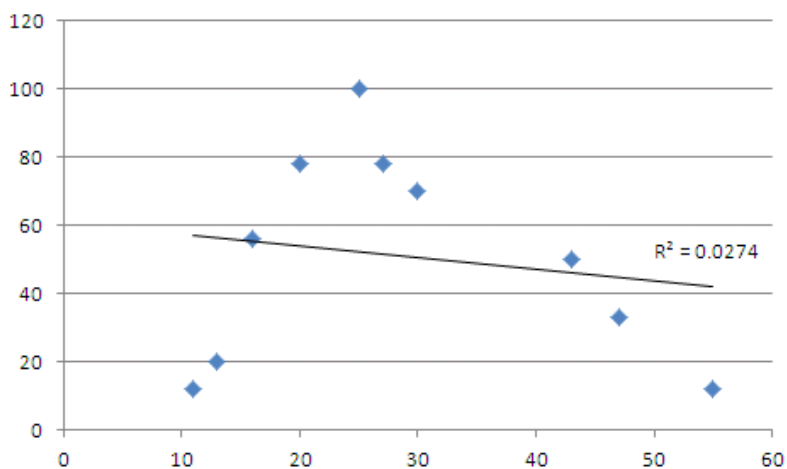


Рис.2.22.

Тому додатково до коефіцієнту кореляції Пірсона між Y та X , можливо розрахувати кореляцію між Y квадратом відхилення X від свого середнього значення як на Рис. 2.23 (вхідні дані і результат) та Рис.2.24 (формули).

	A	B	C	D	E	
1	Y	X	(X-avX)2			
2		12	11	313.29	коррел1	-0.16554
3		20	13	246.49	коррел2	-0.83594
4		56	16	161.29		
5		78	20	75.69		
6		100	25	13.69		
7		78	27	2.89		
8		70	30	1.69		
9		50	43	204.49		
10		33	47	334.89		
11		12	55	691.69		

Рис.2.23.

C	D	E
(X-avX)2		
=СТЕПЕНЬ((B2-CPЗНАЧ(\$B\$2:\$B\$11));2)	коррел1	=КОРРЕЛ(B2:B11;A2:A11)
=СТЕПЕНЬ((B3-CPЗНАЧ(\$B\$2:\$B\$11));2)	коррел2	=КОРРЕЛ(C2:C11;A2:A11)
=СТЕПЕНЬ((B4-CPЗНАЧ(\$B\$2:\$B\$11));2)		
=СТЕПЕНЬ((B5-CPЗНАЧ(\$B\$2:\$B\$11));2)		
=СТЕПЕНЬ((B6-CPЗНАЧ(\$B\$2:\$B\$11));2)		
=СТЕПЕНЬ((B7-CPЗНАЧ(\$B\$2:\$B\$11));2)		
=СТЕПЕНЬ((B8-CPЗНАЧ(\$B\$2:\$B\$11));2)		
=СТЕПЕНЬ((B9-CPЗНАЧ(\$B\$2:\$B\$11));2)		
=СТЕПЕНЬ((B10-CPЗНАЧ(\$B\$2:\$B\$11));2)		
=СТЕПЕНЬ((B11-CPЗНАЧ(\$B\$2:\$B\$11));2)		

Рис.2.24.

Якщо принаймні один з двох коефіцієнтів кореляції є достатньо високим, то ми навіть без діаграми розсіювання можемо знати, що зв'язок між Y та X існує і його варто далі аналізувати детальніше.

Далі варто видалити перед проведенням подальшого основного аналізу сильнокорелюючі між собою незалежні змінні (використовуємо коефіцієнти кореляції Пірсона, а якщо у подальшому аналізуватимемо непараметричними методами, то Спірмена). Ефективним засобом є спеціальний метод аналізу основних компонент, але іноді його використання ускладнене важкістю інтерпретації некорелюючих штучних незалежних змінних, які він пропонує.

У такому випадку звичайний кореляційний аналіз може мати свої переваги. В рамках нього всі незалежні змінні поділяються на умовно сильнокорелюючі між собою змінні. З кожної групи змінних залишаємо одну (як фактор вона представлятиме не тільки себе, а й решту змінних у групі). Наприклад, якщо ми обираємо зростання індексу споживчих цін, а решта змінних у групі – це зростання індексу цін виробників, зростання грошової маси, необхідно розуміти, що у подальшому аналізі ми аналізуватимемо вплив не просто зростання індексу цін виробників, а в цілому фактору, який ми можемо назвати інфляція та зростання грошової маси.

З групи варто обирати змінну, виходячи з таких міркувань:

- більша кількість спостережень із наявними даними по змінній;
- більший коефіцієнт кореляції змінної із залежною змінною;
- більший коефіцієнт кореляції квадрату відхилення змінної від середньої із залежною змінною;
- більша кількість сильнокорелюючих з нею змінних.

ГЛАВА 3. Дослідження валютних криз методами частотного аналізу

3.1. Метод частотного аналізу на прикладі аналізу впливу політичної стабільності та валютних резервів

Таблиці частот (їх ще називають одно вхідні таблиці) є засобом аналізу зв'язку між категоріальними змінними (які вимірюються у неметричній шкалі). Але таблиці частот можуть з успіхом використовуватись для дослідження кількісних змінних, хоча при цьому можуть виникнути труднощі з інтерпретацією результатів. Вони показують як часто проявляється певна ознака/подія у різних групах спостережень. Цей вид аналізу можна ще назвати аналізом часток або аналізом пропорцій. Даний метод дослідження часто використовують як одну з процедур розвідувального аналізу, щоб розглянути, яким чином різні групи спостережень у вибірці, або який розподіл має значення показника на проміжку між мінімального до максимального значення.

Наприклад, ми хочемо дізнатися яким чином урядові кризи впливають на валютні кризи. Урядова криза (бінарна змінна) – буде незалежною змінною, вона ж групуюча змінна у таблиці частот. Валютна криза (також бінарна змінна) – залежна змінна, частоти якої ми розраховуємо емпіричним шляхом за наявними спостереженнями (бажано брати з часовим лагом, наприклад валютна криза протягом наступних двох років). Кожне спостереження – країно-рік. Припустимо, ми систематизували спостереження у формі таблиці – Табл.3.1 (цей процес називається крос-табуляція або групування даних):

Таблиця 3.1.

	Валютна криза є (1)	Валютної кризи немає (0)	Сума
Урядова криза є (1)	30	120	150
Урядової кризи немає (0)	50	600	650
Сума	80	720	800

Це найпростіший варіант крос-табуляції, коли створюється таблиця 2x2 (по два значення кожної з двох змінних). Кожна комірка цієї таблиці представляє комбінацію значень двох змінних. У цих комірках містяться дані про те, як часто зустрічається у досліджуваній вибірці відповідна комбінація значень змінних. Частоти по краях таблиці – це маргінальні частоти, вони показують як розподіляються спостереження при групуванні лише за однією змінною, без урахування іншої.

У процентному вигляді таблиця частот (де показані вже не абсолютні частоти, а відносні частки) виглядає як у табл.3.2. (її також можна виразити не у відсотках, а в частках одиниці):

Таблиця 3.2.

	Валютна криза є (1)	Валютної кризи немає (0)	Сума
Урядова криза є (1)	3.75	15	18.75
Урядової кризи немає (0)	6.25	75	81.25
Сума	10	90	100

Ще один варіант подання вказаний у табл.3.3

Таблиця 3.3.

	Урядова криза є (1)	Урядової кризи немає (0)	Середня
Частка спостережень, коли валютна криза є (1), %	20 =30/150	7.7 =50/650	10 =80/800
Кількість спостережень	150	650	

Частку спостережень, коли валютна криза є (за існуючими спостереженнями) ми використовуємо як оцінку ймовірності виникнення валютної кризи у майбутньому, якщо країна опиняється в аналогічних умовах. З цього ми бачимо, що країна з урядовою кризою має більше шансів опинитися у стані валютної кризи, ніж країна без урядової кризи. Але наскільки надійний цей ви-

сновок? Для перевірки цього існує χ^2 тест (хі-квадрат тест), який показує чи різниця у частотах статистично значуща.

Цей тест базується на перевірці наскільки відрізняються фактичні частоти по комірках від очікуваних, якщо припустити, що зв'язку між досліджуваними змінними немає. Очікувані частоти для нашого прикладу розраховуються як у табл.3.4. (фактичні частоти множаться на частки маргінальних частот у всіх спостереженнях):

Таблиця 3.4.

	Валютна криза є (1)	Валютної кризи немає (0)	Сума
Урядова криза (1)	$15 = 800 \cdot (150/800) \cdot (80/800)$	$135 = 800 \cdot (150/800) \cdot (720/800)$	150
Відсутність урядової кризи (0)	$65 = 800 \cdot (650/800) \cdot (80/800)$	$535 = 800 \cdot (650/800) \cdot (720/800)$	650
Сума	80	720	800

Отже, фактичні частоти: 30, 120, 50, 600 відрізняються від очікуваних: 15, 135, 65, 535.

Якщо цей тест показує рівень значущості p менше 0.05, ми умовно вважаємо, що різниця значуща. Результат буде тим більш надійним і значущим, чим більшою є розрахована нами різниця у частотах на основі вибірки і чим більшою є кількість спостережень у порівнюваних комірках таблиці частот (групах спостережень при групуванні за групувальною змінною). На практиці достатньо надійними можна вважати результати, якщо у кожній порівнюваній групі спостережень (при групуванні за незалежною змінною) більше 50-80 спостережень. Якщо їх 30-50 результати будуть відносно надійними. При меншій кількості треба бути обережним при інтерпретації результатів, оскільки часто χ^2 тест не може достатньо гарантувати надійність результатів.

Незалежна змінна не обов'язково має бути категоріальною. Але в такому випадку її потрібно перетворити на категоріальну. Наприклад, ми хочемо перевірити наскільки впливає на валютні кризи величина валютних резервів у % ВВП. Це континуальний показник, який вимірюється у метричній шкалі. Поділимо всі значення на два діапазони значень, наприклад, більше 15% та менше 15%. Таким чином ми створюємо категоріальну змінну, яка приймає значення 1 (великі валютні резерви) або 0 (малі валютні резерви). Далі аналізуємо подібно до того, як ми аналізували вплив урядових криз.

При виборі значення, яке ділитиме всі значення на два діапазони, варто користуватися такими міркуваннями:

- щоб кількість спостережень у двох групах не відрізнялася сильно;
- щоб це значення добре психологічно сприймалося: наприклад, 15% сприйматиметься краще, ніж скажімо 14.3241%, навіть якщо останнє поділить спостереження рівно навпіл;
- щоб це значення ділило спостереження на такі групи, які дозволяють максимізувати різницю у частках досліджуваної ознаки у формі залежної змінної (це дозволяє сигнальний метод, про який йдеться далі).

Можливо групувати і за більшою кількістю груп. Наприклад, поділимо валютні резерви на три діапазони значень (до 10%; 10-20%, більше 20%). Таблиця частот матиме вигляд (в одній з форм презентації) як у табл.3.5.

Таблиця 3.5.

	Малі валютні резерви	Середні валютні резерви	Великі валютні резерви
Частка спостережень, коли валютна криза є (1), %	22	14	10
Кількість спостережень	370	298	314

3.2. Таблиці частот для аналізу взаємодії факторів: процентної ставки і валютних резервів

Більш складний варіант – побудова таблиці частот для декількох змінних, одночасно. Наприклад, див. табл.3.6. Або в іншій формі представлення (для наочності ми її показуємо з розрахунком часток) – див. табл.3.7.

Таблиця 3.6.

	Малі валютні резерви (0)			Великі валютні резерви (1)			Сума
	Валютна криза є (1)	Валютної кризи немає (0)	Сума	Валютна криза є (1)	Валютної кризи немає (0)	Сума	
Низька процентна ставка (0)	30	170	200	15	285	300	500
Висока процентна ставка (1)	70	280	350	25	125	150	500
Сума	100	450	550	40	410	450	1000

Таблиця 3.7.

Частка спостережень, коли валютна криза є у % (у дужках кількість спостережень)			
	Малі валютні резерви (0)	Великі валютні резерви (1)	Середня
Низька процентна ставка (0)	30/200= 15 (200)	15/300= 5 (300)	(30+15)/500=9 (500)
Висока процентна ставка (1)	70/350= 20 (350)	25/150= 16.7 (150)	(70+25)/500=19 (500)
Середня	(30+70)/550=18.2 (550)	(15+25)/450=8.9 (450)	(100+40)/1000=14 (1000)

За допомогою подібної таблиці можливо побачити ефект взаємодії двох факторів. Так, найменша імовірність валютної кризи за низької (номінальної) процентної ставки та високих валютних резервів. Найбільша – за високої процентної ставки та малих валютних резервів.

При цьому, для країн з низькою процентною ставкою валютні резерви більшою мірою впливають на імовірність валютної кризи (15-5=10), ніж для країн з високою процентною ставкою (20-16.7=3.3). Тобто нарощування валютних резервів часто не є ефективним рятувальним заходом, якщо існує проблема високих номінальних процентних ставок. Так само, в цьому прикладі за умовними даними процентна ставка грає невелику роль, якщо валютних резервів недостатньо. Детальніше про ефект взаємодії див. у Додатку Б.

Таблицю можна ускладнити також, якщо додати групу спостережень з відсутніми даними, якщо такі є. Наприклад, див. табл.3.8.

Таблиця 3.8.

Частка спостережень, коли валютна криза є у % (у дужках кількість спостережень)				
	Малі валютні резерви (0)	Великі валютні резерви (1)	Середня	Відсутні дані по валютних резервах
Низька процентна ставка (0)	15 (200)	5 (300)	9.4 (500)	15 (20)
Висока процентна ставка (1)	20 (350)	16.7 (150)	18.3 (500)	24 (150)
Середня	18.2 (550)	8.9 (450)	140 (1000)	
Відсутні дані по процентній ставці	21 (120)	8 (130)		

З цього прикладу ми бачимо, що незважаючи на досить велику частку спостережень з невідомою процентною ставкою, ймовірності валютної кризи є близькими до середніх таких ймовірностей по спостереженнях з відомою процентною ставкою (21 до 18.2, а 8 до 8.9). Щодо відсутніх даних за валютними резервами тут є два випадки:

1. Суттєва відмінність у групі низьких процентних ставок (15 і 9.4), але кількість відсутніх даних невелика (20). Тест χ^2 у частотах не покаже значущої відмінності між цими двома частотами (15 і 9.4), але навіть, не використовуючи χ^2 тест, можливо сказати що кількість відсутніх даних відносно відомих даних тут мізерна (20/500) і така все ж не занадто велика різниця у частотах не викликає занепокоєння щодо дії спотворюючого ефекту відсутніх даних.

2. Відмінність у групі високих процентних ставок (24 і 18.3) є досить великою при тому, що кількість відсутніх даних є великою (150/520). Це викликає занепокоєння через те, що розміщення

відсутніх даних може бути залежним від частоти валютних криз. У результаті висновок, про те, що великі валютні резерви лише трохи зменшують ймовірність валютної кризи при великих процентних ставках, може бути ненадійним. Насправді, можуть бути варіанти:

- при великих процентних ставках великі валютні резерви лише *трохи зменшують* ймовірність валютної кризи (як нам і здається);
- при великих процентних ставках великі валютні резерви *суттєво зменшують* ймовірність валютної кризи (вплив недооцінений);
- при великих процентних ставках великі валютні резерви *не впливають* на ймовірність валютної кризи (вплив переоцінений);
- при великих процентних ставках великі валютні резерви *збільшують* ймовірність валютної кризи (характер впливу визначено невірно, хоча це малоймовірний випадок).

3.4. Частотний аналіз впливу зовнішнього боргу у Microsoft Excel

Для початку потрібно самостійно побудувати таблицю частот, після чого використати функцію:

=ХИ2ТЕСТ(діапазон фактичних частот; діапазон очікуваних частот)

Ця функція повертає значущість різниці між очікуваними та фактичними частотами.

Наведемо приклад за умовними даними. Припустимо, ми хочемо дізнатися, як різні рівні зовнішнього боргу у % ВВП впливають на ймовірність валютної кризи. На одному аркуші побудуємо таблиці, які міститимуть дані про фактичні частоти, очікувані частоти та частки спостережень з валютною кризою. На Рис. 3.14 зображені вхідні дані та результати, на Рис.3.15 – формули.

Як бачимо різниця є значущою (0.03816 менше 0.05), а отже ми маємо докази на користь того, що величина зовнішнього боргу впливає на ймовірність валютної кризи. Якщо зовнішній борг невеликий (менше 50% ВВП) ймовірність валютної кризи становить лише 10.7%, за середньою боргу (50-80%) – 21%, за великого (більше 80%) – 40%.

	A	B	C	D	E	F
1	Фактичні частоти		зовнішній борг > 80%	50 - 80%	< 50%	
2		валютна криза є	10	8	3	21
3		немає	15	30	25	70
4			25	38	28	91
5						
6	Очікувані частоти					
7			зовнішній борг > 80%	50 - 80%	< 50%	
8		валютна криза є	5.77	8.77	6.46	0.23
9		немає	19.23	29.23	21.54	0.77
10			0.27	0.42	0.31	1.00
11						
12	Частки та кількість спостережень					
13			зовнішній борг > 80%	50 - 80%	< 50%	
14		валютна криза є	0.400	0.211	0.107	0.231
15		кількість	25	38	28	91
16						
17						
18	Значущість χ^2 критерію	0.038163205				

Рис.3.14.

	A	B	C	D	E	F
1	Фактичні частоти		зовнішній борг > 80%	50 - 80%	< 50%	
2		валютна криза є	10	8	3	=СУММ(C2:E2)
3		немає	15	30	25	=СУММ(C3:E3)
4			=СУММ(C2:C3)	=СУММ(D2:D3)	=СУММ(E2:E3)	=СУММ(F2:F3)
5						
6	Очікувані частоти		зовнішній борг > 80%	50 - 80%	< 50%	
7						
8		валютна криза є	=F4*F8*C10	=F4*F8*D10	=F4*F8*E10	=F2/F4
9		немає	=F4*F9*C10	=F4*F9*D10	=F4*F9*E10	=F3/F4
10			=C4/F4	=D4/F4	=E4/F4	=F4/F4
11						
12	Частки та					
13	кількість спостережень		зовнішній борг > 80%	50 - 80%	< 50%	
14		валютна криза є	=C2/C4	=D2/D4	=E2/E4	=F2/F4
15		кількість	=C4	=D4	=E4	=F4
16						
17						
18	Значущість χ^2 критерію	=ХИ2ТЕСТ(C2:E3;C8:E9)				

Рис.3.15.

3.5. Алгоритм дослідження взаємодії факторів валютних криз методами частотного аналізу

Далі буде описаний методологічний алгоритм для дослідження факторів валютних криз застосований у книзі [47], а також можливості для використання наукових результатів у ній для аналізу поточної ситуації. Ці результати можуть бути використані для пояснення впливу факторів в умовах їх взаємодії, для визначення заходів з метою попередження валютних криз. Використання результатів з прогностичною метою можливе, але має суттєві обмеження.

Для визначення заходів запобігання валютним кризам доцільно з'ясувати не тільки, які фактори впливають на ризик виникнення валютних криз за інших рівних умов, а й умови, від яких залежить дія цих факторів. З цією метою на основі широкої вибірки країн проведена оцінка важливості факторів, характеру їх впливу на виникнення валютної кризи з урахуванням взаємодії цих факторів.

Як залежна змінна використовується логічна змінна наявності валютної кризи у відповідному році у відповідній країні. Якщо принаймні в один із місяців відповідного року в країні є валютна криза, то залежній змінній присвоюється значення 1. Якщо в жодному місяці цього року немає валютної кризи, то присвоюється значення 0. Наявність валютної кризи у відповідному місяці визначається за індексом I_{ch2} . Перевищення індексом I_{ch2} нуля свідчить про наявність валютної кризи. Цей індекс розраховується за формулою:

$$I_{ch2} = E + R + I - 3 + 0,5 \times D, \quad (3.1)$$

де E – компонент зменшення валютного курсу:

$$E = \frac{e_t - \bar{e}}{0,06 \cdot \bar{e} + \sigma_e}; \quad (3.2)$$

R – компонент валютних інтервенцій:

$$R = \frac{\bar{r} - r_t + L_t - \bar{L}}{0,12 \cdot \bar{r} + \sigma_r}; \quad (3.3)$$

I – компонент процентної ставки:

$$I = \frac{m_t - \bar{m}}{0,03 \cdot (100 + \bar{m}) + \sigma_m}, \quad (3.4)$$

або якщо немає даних про m , то:

$$I = \frac{c_t - \bar{c}}{0,017 \cdot (100 + \bar{c}) + \sigma_c}; \quad (3.5)$$

D – компонент прискорення зменшення валютного курсу:

$$D = \frac{d_t - \bar{d}}{0,037 \cdot \bar{d} + \sigma_d}, \quad (3.6)$$

$e_t, r_t, L_t, m_t, c_t, d_t$ – курс національної валюти до СДР, золотовалютні резерви (у СДР, включаючи золоті резерви за ринковими цінами), зовнішні зобов'язання монетарної влади (центрального банку), процентна ставка грошового ринку, облікова ставка центрального банку і темпи зниження валютного курсу в t -й місяць (джерело даних для розрахунку International Financial Statistics Online Service (<http://ifs.spdi.net/imf>) або можливе використання іншого джерела, наприклад статистики центрального банку);

$\bar{e}, \bar{r}, \bar{L}, \bar{m}, \bar{c}, \bar{d}$ – їх середні значення протягом попередніх 12 місяців;

$\sigma_e, \sigma_r, \sigma_m, \sigma_c, \sigma_d$ – їх стандартні відхилення (крім для зобов'язань монетарної влади) протягом попередніх 12 місяців.

Темпи зниження валютного курсу розраховуються як величина $\frac{e_t}{e_{t-1}} - 1$.

У разі відсутності даних про R або I , один з цих компонентів відповідно прирівнюється до одиниці. Якщо відсутні дані про зобов'язання монетарної влади, їх зміна прирівнюється до нуля. Золотовалютні резерви при можливості обраховуються за ринковою ціною золота. Абсолютні значення порогових рівнів (0,06; 0,12; 0,03; 0,017; 0,037) розраховані для кожного з показників як середнє плинних коефіцієнтів варіації відповідного показника для всіх країн протягом періоду 1989–2003 рр.

Якщо розрахувати за даними по Україні з початку 2008 року по липень 2009 року (з урахуванням процентної ставки на міжбанківському ринку, а не облікової ставки, та середньомісячного валютного курсу), то індекс I_{ch2} давав сигнал про наявність валютної кризи у листопаді 2008 – січні 2009 року. Вже в жовтні 2008 року компонент індексу по валютному курсу E (якщо рахувати по курсу на кінець місяця, а не середньомісячному) практично дорівнював 1. Кожний з компонентів E і R перевищували 1 до квітня 2009.

Усього досліджена дія 154 факторів. Використані річні дані. При оцінці впливу на залежну змінну враховується лаг у 1 рік. Для аналізу використано спостереження за період 1990–2003 рр. для змінної наявності валютної кризи (і за період 1989–2002 рр. для незалежних змінних).

Ураховано спостереження по 109 країнах які умовно поділені на п'яти регіонів: Центральна та Східна Європа і СНД; Близький Схід і Північна Африка; Латинська Америка; Східна Азія; Африка, південніше Сахари. Переважна більшість країн – це країни середнього, а також низького рівня розвитку. З аналізу виключені малі країни (на ВВП та населенням), а також ті, які не мають власної національної валюти.

Виключено з розгляду спостереження, щодо яких відсутні дані за залежною змінною з лагом у 1 рік, а також щодо яких дані за залежною змінною поточного року відсутні або вже дорівнюють 1. Таким чином, результати аналізу можна використовувати для прогнозу наявності чи валютної кризи у наступному році лише, якщо у поточному році її немає. Тобто дослідження не мало на меті прогнозування тривалості або занадто швидкого повторення валютної кризи. Таким чином, залишається 989 спостережень, з них у 143 значення залежної змінної в наступному році дорівнює 1. Кожне спостереження представляє собою країно-рік.

Побудована таблиця, яка містить важливі характеристики змінних:

- назва змінної (наприклад, курс національної валюти до долара);
- позначка, чи є показник незалежною змінною, що використовується в аналізі, або лише використаний для розрахунку незалежних змінних;
- одиниці виміру, використані при розрахунках (наприклад, од. національної валюти / дол.);

- коротке позначення показника (наприклад, $e1Exr\$a$);
- джерело (з кодом показника у джерелі) або формула розрахунку;
- примітки – додаткові дані про характер показника (наприклад, на кінець періоду чи в середньому за рік), спосіб обрахунку (наприклад, чи використовувалися плинні середні або значення інших періодів для заміни відсутніх даних).

Наступним кроком стало зменшення мультиколінеарності незалежних змінних: з аналізу вилучена частина незалежних змінних. Проведення стандартного аналізу основних компонент в умовах великої кількості змінних створювало проблему ускладненої інтерпретації нових штучно побудованих некорелюючих між собою змінних. Тому для зменшення мультиколінеарності було обрано інший метод: виділено декілька груп сильнокорелюючих між собою змінних (парний коефіцієнт кореляції Пірсона або Спірмана більше 65%). З кожної групи сильнокорелюючих між собою факторів залишається лише одна (сурогатна змінна). Така сурогатна змінна у кожній групі у подальшому розглядається як фактор і значною мірою відображає всі інші змінні у групі. Бажано, щоб сурогатна змінна мала:

- значну кількість спостережень із наявними даними по цій змінній;
- більший коефіцієнт кореляції із залежною змінною наступного року;
- більший коефіцієнт кореляції квадрату відхилення змінної від середньої із залежною змінною наступного року (для врахування можливих нелінійних зв'язків);
- більшу кількість сильно корелюючих з нею змінних, які потім виключаються з аналізу.

Альтернативним способом замість використання сурогатної змінної могла би бути побудова штучної змінної у формі лінійної комбінації з усіх сильнокорелюючих між собою змінних. Цей варіант був би більш точним, але ускладнює інтерпретацію результатів.

У результаті як фактори розглядаються 154 незалежні змінні. Побудована нова таблиця, де по кожному фактору подана така інформація:

- Назва, сурогатна змінна та сильнокорелюючі змінні. Назва фактору може відрізнитися від точної назви сурогатної змінної. Наприклад, фактор з назвою «зростання грошей та інфляція» базується на сурогатній змінній «зростання грошей і квазігрошей» (фактично агрегату M2), яка сильно корелює з декількома показниками грошової маси, обсягів кредитування, процентних ставок, інфляції. Тобто, наприклад, в більшості випадків інфляція пов'язана саме з пропозицією грошей.

- Межі діапазонів значень у рамках зазвичай трьох груп спостережень (діапазону найменшої значення (група спостережень I), середні значення (II група), найбільші значення (III група). Наприклад, для фактору «Зростання грошей і інфляція» (у відсотках): -21.4 (7) 12 мінімальне; 12.01 (17) 22 середнє; 22.1 (34) 142 значне. В дужках – медіани по кожній групі значень. Бажано при цьому вказувати межі діапазонів значень сурогатної змінної (яка розглядається як фактор), а й відповідні їм межі діапазонів значень сильнокорелюючих з нею інших змінних (наприклад, інфляції за індексом споживчих цін).

Наступним кроком проводиться аналіз частот (часток, пропорцій). За допомогою кросс-табуляції отримуються дані за кожною групою значень незалежних змінних: кількість спостережень і частка спостережень, коли залежна змінна наступного року має значення 1.

Наприклад у табл. 3.9, ми бачимо результати кросс-табуляції по двох факторах: по стовпчиках – частка інвестицій у ВВП (в основний капітал: 5.9 (15,6) 18 низька; 18.01 (21) 23 середня; 23.01 (27,4) 65.1 значна), по рядках – серед по процентних ставках (різниця між кредитною і депозитною ставкою: -42 (3) 5 мінімальний; 5.01 (7) 10 середній; 10.01 (13) 164 великий).

Таблиця 3.9.

Rows:	m2Spread	Columns:	r3Invfix		
	1	2	3	6	All
1	69	46	115	22	252
	0.2319	0.0435	0.0522	0	0.0952
2	61	87	104	8	260
	0.2131	0.1839	0.1442	0	0.1692

3	92	68	50	9	219
	0.1739	0.1618	0.18	0.5556	0.1872
6	46	70	36	106	258
	0.2609	0.1429	0.1111	0.0755	0.1318
All	268	271	305	145	989
	0.2127	0.1439	0.1115	0.0897	0.1446

Спостереження згруповані по кожній змінній поділені на 3 діапазони/групи (I, II і III або 1, 2 і 3 арабськими цифрами), 6 – позначає групу спостережень з відсутніми даними по даній групі. На перетині комірок показані – кількість спостережень (зверху) і частка спостережень, що передують виникненню валютної кризи у наступному році (знизу). Крос-табуляцію можна представити й в іншій формі (див. табл. 3.10): зверху частоти спостережень, що передують спокійним періодам, а знизу частоти спостережень, що передують валютним кризам.

Спочатку розглядається, який вплив здійснює групуюча змінна по стовпчиках у рядку з усіма спостереженнями (далі фактор впливу), у цьому прикладі – це частка інвестицій.

Якщо значущі відрізняються частки в I та III групах, або лише в одній з них, вплив відповідного фактора постійно прямий або обернений і цілком можливо, що лінійний. Якщо значущі відрізняється частка у II групі, то, імовірно, вплив фактора нелінійний та існують найбільш або найменш оптимальні значення незалежної змінної всередині діапазону значень у II групі.

В цьому прикладі і частка 0.21 у групі I і частка 0.11 у групі III значущі відрізняються (навіть з рівнем значущості менше 0.01). Тобто чим більшою є частка інвестицій у ВВП, тим більш захищеною є країна від валютних криз.

Для визначення значущості різниці між частками у групах використовується χ^2 тест. Статистика χ^2 тесту розраховується за формулою:

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e}, \quad (3.7)$$

Таблиця 3.10.

Rows:	m2Spread	Columns:	r3Invfix		
	1	2	3	6	All
1	53	44	109	22	228
	16	2	6	0	24
2	48	71	89	8	216
	13	16	15	0	44
3	76	57	41	4	178
	16	11	9	5	41
6	34	60	32	98	224
	12	10	4	8	34

All	211	232	271	132	846
	57	39	34	13	143

де f_0 – частоти, що спостерігаються у в комірках таблиці 2×2 (з групуванням спостережень за фактором впливу по горизонталі та за залежною змінною по вертикалі); кожна з них дорівнює частці спостережень із залежною змінною = 1 або 0 у відповідній групі за фактором впливу, помноженій на кількість спостережень у відповідній групі за фактором впливу;

f_e – очікувані частоти, якщо припустити, що вплив фактора відсутній; кожна з них дорівнює частці спостережень із залежною змінною = 1 або 0 у всіх групах за фактором впливу, помноженій на кількість спостережень у відповідній групі за фактором впливу.

Різниця вважається значущою, якщо тест показує рівень значущості менше 0.05. Додаткові застереження робляться у випадку невеликої кількості спостережень у порівнюваних групах (менше 20).

Потім порівнюються частки у групах з наявними даними з урахуванням дії решти незалежних змінних, тобто після групування по рядках кожного разу за іншою змінною (далі фактором умов). У нашому прикладі фактором умови виступає серед по процентних ставках. Це дозволяє побачити, як залежить частка спостережень, що передують валютній кризі (яку ми розглядаємо як оцінку імовірності виникнення валютної кризи), від розглянутого фактора впливу залежно від того, яке значення приймає фактор умови.

Значущість відмінності впливу у групах за факторами умов від впливу за інших рівних умов перевіряється за допомогою модифікованого χ^2 тесту для різниці часток у n групах за фактором впливу одночасно. Перевіряється, чи відрізняється вплив у кожній групі окремо за кожним фактором умов від впливу без урахування факторів умов. Статистика цього тесту розраховується за формулою (3.7), де:

f_0 – частоти, що спостерігаються, в комірках таблиці $2 \times n$ (з групуванням спостережень за фактором впливу по горизонталі та за залежною змінною по вертикалі, при цьому в цій таблиці відображені лише спостереження у відповідній групі за фактором умов); але при цьому для їх розрахунку частки спостережень із залежною змінною = 1 у групах за фактором впливу замінюються на ці ж частки спостережень із залежною змінною = 1 у групах за фактором впливу, що помножені на частку спостережень із залежною змінною = 1 в усіх спостереженнях і поділені на частку спостережень із залежною змінною = 1 у відповідній групі за фактором умов; частка спостережень із залежною змінною = 0 у кожній групі за фактором впливу обчислюється як одиниця мінус частка спостережень із залежною змінною = 1 у відповідній групі за фактором впливу;

f_e – очікувані частоти, якщо припустити, що вплив фактора не відрізняється; кожна з них дорівнює частці спостережень із залежною змінною = 1 або 0 у відповідних групах за фактором впливу (але по всіх групах за фактором умов), помноженій на кількість спостережень у відповідній групі за фактором впливу.

При цьому не враховуємо спостереження, щодо яких немає даних по незалежній змінній. Так само відмінність впливу вважається значущою, якщо тест показує рівень значущості менше 0.05.

З формулами в Microsoft Office Excel це виглядатиме як на Рис.3.16.

В нашому прикладі, в групі I за фактором умов (тобто коли серед по процентних ставках мінімальний) вплив фактору впливу (частки інвестицій) значуще відрізняється від його впливу за інших рівних умов – цей вплив є сильнішим. У цій групі ймовірність валютної кризи ще більше зменшується із зростанням частки інвестицій. У даному випадку, країни виграють більше, якщо поєднують високу частку інвестицій з низьким процентним середом (різницею між депозитними і кредитними ставками).

В узагальнюючій табл. 3.11. вказані ще декілька прикладів груп спостережень за факторами умов, у яких вплив частки інвестицій відрізняється.

Ми бачимо, що історично при відпливі позичкового іноземного капіталу з банків імовірність валютної кризи була більшою при середній частці інвестицій. З одного боку цей результат є

достатньо статистично значущим. З іншого боку складно знайти теоретичне підґрунтя, яке пояснювало б таку закономірність.

За низької частки депозитів до запитання у всіх депозитах (а отже при високій довірі до банківської системи), як і при значних золотовалютних резервах висока частка інвестицій вже не є оптимальним варіантом.

Виявлені закономірності у впливі факторів на імовірність виникнення валютних криз з урахуванням умов, характерних для України в поточний період, дають можливість визначити фактори ризику для валютної стабільності наступного року та рекомендації щодо забезпечення валютної стабільності в Україні. Проте, слід зазначити, що з міркувань урахування структурних змін доцільно при цьому використовувати лише найбільш значущі закономірності, які при цьому не викликають проблем з інтерпретацією з теоретичної точки зору.

Оскільки статистичні дані за 2011 рік були ще не відомі на момент проведення аналізу, можливе використання оцінок або прогнозу на основі даних за місяці з наявними даними або даних за попередній рік, якщо показник мало змінюється у часі. Поточні дані по Україні розраховані за такими джерелами: Національний банк України, Державний комітет статистики України, Міністерство фінансів України, World Development Indicators, World Economic Outlook Database, Economic Freedom of the World: 2010 Annual Report.

Наприклад, у Табл. 3.12 для подорожчання національної валюти відносно СДР (в частках одиниці: 0,55 (0,88) 0,93 значне здешевлення; 0,9301 (0,97) 1 невелике здешевлення; 1,0001 (1,03) 1,44 подорожчання) ми залишимо лише найбільш значущі відхилення від загальної закономірності.

Поряд із загальною закономірністю про небажаність значного здешевлення національної валюти, є характерні для України умови в поточному періоді: відносно значні валютні резерви та відносно швидке зростання кінцевого споживання, де оптимальне значення динаміки курсу валюти інше, у той час як значні надходження від туризму і зростання процентних ставок по кредитах не характерні для України. Для оцінки ймовірності валютної кризи з урахуванням конкретних умов, можемо розрахувати середні значення для груп спостережень.

Ми можемо припустити, що найменша ймовірність виникнення валютної кризи буде при невеликому здешевленні курсу національної валюти (на 0-7%). З жовтня 2010 по жовтень 2011 гривня відносно СДР знецінилася на 3%. Тобто курсова динаміка в поточний момент є близькою до оптимальної. Хоча слід зазначити, що це може значно залежати від курсу євро.

Наприклад, якщо євро подешевшає на 8-16% і більше порівняно з жовтневим рівнем 1.415 дол./євро (частка євро у СДР 37%, а отже 3% різницю з межею діапазону груп II і III та 6% різницю з медіаною діапазону групи III ділимо на 0.37), курс гривні вийде за межі оптимального діапазону. Наприклад, курс євро до долара 1.22, за нашими оцінками, практично вдвічі (або на 7 п.п.) збільшує ймовірність виникнення валютної кризи порівняно з курсом 1.415. Так само, наприклад,

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Row	s:m2Spread	C	columns:r3	Invfix														
2																			
3		1	2	3	6 All														
4																			
5	1	69	46	116	22	262	230	69	46	116			0.00479206	23.6293	2.95494	8.86481	14.6763	6.6194	12.8226
6		0.2319	0.0435	0.0522	0	0.0952	0.10431	0.34245	0.06424	0.07709			0 ИСТИНА	46.3707	43.0451	106.135	54.3237	39.3806	102.178
7																			
8	2	61	87	104	8	260	252	61	87	104			0.69776164	11.4697	14.1169	13.2323	12.9747	12.5193	11.696
9		0.2131	0.1839	0.1442	0	0.1692	0.17467	0.18803	0.16226	0.12723			ЛОЖЬ	49.5303	72.8831	90.7677	48.0263	74.4807	92.404
10																			
11	3	92	68	50	9	219	210	92	68	50			0.22033766	14.3767	9.86886	8.06748	19.6684	9.7852	5.675
12		0.1739	0.1618	0.18	0.5556	0.1672	0.17141	0.16627	0.14639	0.16175			ЛОЖЬ	77.6233	58.1131	41.9125	72.4316	58.2148	44.425
13																			
14	6	46	70	36	106	258	152	46	70	36	106								
15		0.2609	0.1429	0.1111	0.0755	0.1318	0.17109	0.23493	0.12867	0.10004	0.06798								
16																			
17	All	268	271	305	145	969	844	268	271	305	145								
18		0.2127	0.1439	0.1115	0.0897	0.1446	0.15403	0.2127	0.1439	0.1115	0.0897								
	G	H	I	J	K	L	M	N	O	P	Q	R	S						
1																			
2																			
3																			
4																			
5	=F5-E5	=B5	=C5	=D5	=E5	=D6	=E6	=E5	=E6	=E5	=E6	=E5	=X/2TECT(N5:P6;O5:S6)	=H5*H6	=I5*I6	=J5*J6	=K5*K6	=L5*L6	=M5*M6
6	=(F8*F9-E9*E8)/(F8-E8)	=B8*\$G\$18:\$G8	=C8*\$G\$18:\$G8	=D8*\$G\$18:\$G8	=E8*\$G\$18:\$G8	=D9*\$G\$18:\$G9	=E9*\$G\$18:\$G9	=E8	=E9	=E8	=E9	=E8	=X/2TECT(N8:P9;O8:S9)	=H8*H9	=I8*I9	=J8*J9	=K8*K9	=L8*L9	=M8*M9
7	=F8-E8	=B8	=C8	=D8	=E8	=D9	=E9	=E8	=E9	=E8	=E9	=E8		=H8*N8	=I8*O8	=J8*P8	=K8*Q8	=L8*R8	=M8*S8
8	=(F9*F8-E9*E8)/(F8-E8)	=B9*\$G\$18:\$G9	=C9*\$G\$18:\$G9	=D9*\$G\$18:\$G9	=E9*\$G\$18:\$G9	=D10*\$G\$18:\$G10	=E10*\$G\$18:\$G10	=E9	=E10	=E9	=E10	=E9		=H9*N9	=I9*O9	=J9*P9	=K9*Q9	=L9*R9	=M9*S9
9	=F9-E9	=B9	=C9	=D9	=E9	=D10	=E10	=E9	=E10	=E9	=E10	=E9		=H9*N9	=I9*O9	=J9*P9	=K9*Q9	=L9*R9	=M9*S9
10																			
11	=F11-E11	=B11	=C11	=D11	=E11	=D12	=E12	=E11	=E12	=E11	=E12	=E11	=X/2TECT(N11:P12;Q11:S12)	=H11*H12	=I11*I12	=J11*J12	=K11*K12	=L11*L12	=M11*M12
12	=(F12*F11-E12*E11)/(F11-E11)	=B12*\$G\$18:\$G12	=C12*\$G\$18:\$G12	=D12*\$G\$18:\$G12	=E12*\$G\$18:\$G12	=D13*\$G\$18:\$G13	=E13*\$G\$18:\$G13	=E12	=E13	=E12	=E13	=E12		=H11*N11	=I11*O11	=J11*P11	=K11*Q11	=L11*R11	=M11*S11
13																			
14	=F14-E14	=B14	=C14	=D14	=E14	=D15	=E15	=E14	=E15	=E14	=E15								
15	=(F15*F14-E15*E14)/(F14-E14)	=B15*\$G\$18:\$G15	=C15*\$G\$18:\$G15	=D15*\$G\$18:\$G15	=E15*\$G\$18:\$G15	=D16*\$G\$18:\$G16	=E16*\$G\$18:\$G16	=E15	=E16	=E15	=E16								
16																			
17	=F17-E17	=B17	=C17	=D17	=E17	=D18	=E18	=E17	=E18	=E17	=E18								
18	=(F18*F17-E18*E17)/(F17-E17)	=B18*\$G\$18:\$G18	=C18*\$G\$18:\$G18	=D18*\$G\$18:\$G18	=E18*\$G\$18:\$G18	=D19*\$G\$18:\$G19	=E19*\$G\$18:\$G19	=E18	=E19	=E18	=E19								

Рис.3.16.

Таблиця 3.11.

Вплив частки інвестицій у ВВП

Групи	Імовірність валютної кризи, %		
	I	II	III
Vci	21*+	14+	11*+
c6Othlbank/GDP III** (від'ємні інші інвестиції в країну в банки відносно ВВП)	8	20	8+
m2Spread I**(мінімальний серед по процентних ставках)	23*	4	5+
b2DDep/BDep I** (депозити до запитання / всі депозити)	6	12+	12+
a1FRestr/GDP III (значні золотовалютні резерви)	8	10+	10+

Примітки:

1. По стовпчиках – дані по групах спостережень за фактором впливу, вказаним у назві таблиці.
2. По рядках – дані по всіх спостереженнях або групах спостережень за факторами умов, де вплив фактору впливу значуще відрізняється від впливу по всіх спостереженнях.
3. Жирним шрифтом позначені частки, що значуще відрізняються від інших груп у рядку (при рівні значущості менше 0.05), якщо кількість спостережень не менше 20 (додаткові критерії використовуються, якщо кількість спостережень становить 15-19).
4. Зіркою (*) позначені частки, що особливо значуще відрізняються (при рівні значущості менше 0,01) від інших груп у рядку, якщо кількість спостережень у групах не менше 30.
5. Двома зірками (**) позначені рядки, де вплив фактору впливу особливо значуще відрізняється (при рівні значущості менше 0.01) від впливу за усіма спостереженнями, якщо кількість спостережень у кожній з груп у рядку не менше 30.
6. Знаком (^) позначено випадки, коли можлива наявність спотворюючого ефекту відсутніх даних.
7. Знаком (+) позначені частки в групах спостережень, якщо кількість цих спостережень не менше 80.
8. Знаком (-) позначені частки в групах спостережень, якщо кількість цих спостережень від 10 до 40 (частки не показані при кількості спостережень менше 10).

Таблиця 3.12

Вплив подорожчання національної валюти

Групи	Імовірність валютної кризи, %		
	I	II	III
Vci	18+	13+	13+
m2LRatech III** (зростання процентних ставок по кредитах)	15	14+	28*+
t5Tour/GDP III** (значні надходження від туризму відносно ВВП)	7	3+	12+
r2FConsgr III** (швидке реальне зростання кінцевого споживання)	12	8+	21+
a1FRestr/GDP III** (значне трендове значення валютних резервів у майбутньому році: тобто величина цього співвідношення плюс його зміна порівняно з попереднім роком)	13	4+	13+
Середні з урахуванням умов України	14.3	8.3	15.7

знецінення курсу гривні відносно долара на 9% (при незмінному курсі євро до долара) на 70% (або на 6 п.п.) збільшує ймовірність переходу до стану валютної кризи наступного року. Далі за аналогічною методикою зробимо й інші оцінки за іншими факторами:

- динаміка курсу долара протягом попередніх 12 місяць була в цілому у сприятливому діапазоні, але подорожчання долара до 1.30 може на 55% (або 5 п.п.) збільшити ймовірність валютної кризи;
- динаміка реального ефективного курсу гривні була в цілому у сприятливому діапазоні, але в 4 рази (або на 28 п.п.) може збільшити ймовірність кризи реальна ревальвація гривні на 7% (наприклад, у разі прискорення інфляції або знецінення євро до долара до рівня 1.17);
- величина валютних резервів відносно ВВП знаходиться в сприятливому діапазоні, про-

те вихід з нього можливий у разі зниження до рівня 29 – 32 млрд. дол. із зростанням ймовірності валютної кризи на 60% (або на 5.5. п.п.);

- величина валютних резервів відносно короткострокового зовнішнього боргу знаходиться в відносно несприятливому діапазоні, зниження ймовірності валютної кризи на 5п.п. або на 30% можливе при збільшенні цього співвідношення в 2.5 рази, що в короткостроковій перспективі не є реалістичним;
- величина валютних резервів відносно грошового агрегату М2 (у липні 2011 року 215%) знаходиться у сприятливому діапазоні;
- річні темпи зростання агрегату М2 знаходяться у сприятливому діапазоні (в липні 2011 року 18.7%), проте кожне додаткове таке зростання на 1 п.п. може збільшити ймовірність валютної кризи на 0.6 п.п.;
- динаміка ставки рефінансування знаходиться у сприятливому діапазоні (в липні 2011 року 9.3%). Зниження або зростання ставки рефінансування протягом року на 4-5% супроводжується збільшенням ймовірності валютної кризи майже на 10 п.п. або вдвічі;
- частка депозитів до запитання у всіх депозитах знаходиться у відносно несприятливому діапазоні, її зниження в 1.5 рази може супроводжуватися зниженням ймовірності виникнення валютної кризи майже на 40% або на 7 п.п.

Через ефект усереднення розрахувати конкретну ймовірність виникнення валютної кризи складно. Але ця методика дозволяє визначити ряд основних факторів ризику, такі як динаміка курсу основних світових валют, субоптимальне покриття валютними резервами зовнішнього боргу, субоптимальна структура банківських депозитів.

3.7. Переваги та недоліки методу частотного аналізу

Вказана методика аналізу має такі переваги:

- охоплено значну кількість потенційних факторів;
- враховано ефект взаємодії факторів;
- дає можливість з'ясувати нелінійні зв'язки;
- здійснюється спроба обрати різні способи виміру схожих факторів (наприклад, абсолютне значення показника, його зміна, приріст, трендові значення, відхилення від теоретичних або середніх значень, співвідношення з іншими показниками, поточні та кумулятивні показники, застосування даних з різних джерел).

В той же час існують елементи недосконалості:

- частина закономірностей не мають теоретичного пояснення;
- частина виявлених зв'язків є результатом дії випадковості, а не справжньої закономірності;
- межі діапазонів значень при проведенні крос-табуляції можуть бути обрані не оптимально;
- кореляція між незалежними змінними не сильна, але існує;
- змінні з більшою кількістю відсутніх даних мають менше шансів дати значущі результати;
- в окремих випадках виявлені закономірності могли би бути іншими, якщо б ми мали інформацію про відсутні дані;
- тест про різницю між частками може дати зміщений результат за умови малої кількості спостережень і близької до нуля частки.
- іноді важко відрізнити що є причиною, а що наслідком: валютна криза чи певна величина незалежної змінної;
- можуть бути випадки, коли заходи держави, виражені в певних значеннях показника, насправді мають антикризовий вплив, але внаслідок їх великої політичної чи економічної вартості застосовуються лише за умови безпосередньої загрози кризи, і тому проведений аналіз характеризує ці заходи як ті, що сприяють виникненню валютної кризи;
- фактори можуть мати інший лаг впливу (наприклад, негативний вплив середньої, а не великої кількості валютних криз у інших країнах, який спричиняє ефект поширення, може пояснюватися тим, що у разі великої кількості валютних криз по країнах у регіоні вони поширюються на слабкіші країни дуже швидко і ці країни випадають з аналізу в цьому ж році, як ті, що вже мають валютну кризу);

- розглядаючи річні дані, можливо не помітити короткострокові великі флуктуації значень показників;
- ефект занадто низької чи високої бази порівняння для відносних показників;
- інші фактори можуть компенсувати дію фактору, що розглядається (наприклад, у Східній Європі, за результатами аналізу, більше обслуговування зовнішнього боргу здійснює позитивний вплив. Але на початковому етапі трансформації зовнішній борг був значно меншим (хоча і швидко зростає) і відповідно менше мав обслуговуватися, водночас значення інших показників економіки в цілому були в дуже несприятливому діапазоні. Із завершенням основного періоду трансформації зовнішній борг вже був більшим, а відповідно і його обслуговування, але значення інших показників перейшли до сприятливого діапазону);
- показники можуть бути недосконало виміряні (неточність офіційної статистики, використання інших показників як наближення замість необхідного нам показника, неточність бальних оцінок, відсутністю даних по частині показників, якщо вони застосовуються для підрахунку сумарного показника);
- для показників глобальної дії (світові процентні ставки, зростання курсу долара тощо) кількість спостережень істотно завищують, також їх видимий вплив може бути спотворений ефектом поширення/зараження кризою, недоскональними оцінками кореляції між факторами глобальної дії;
- з часом у світі відбуваються структурні зміни, тому частина виявлених закономірностей може зникнути або виникнуть нові закономірності;
- використання закономірностей для прогнозу зіштовхується з проблемою усереднення ймовірності валютної кризи, оскільки країна за одними показниками може знаходитися у сприятливому діапазоні, за іншими у несприятливому.

ГЛАВА 4. Сигнальний метод аналізу факторів валютних та фіскальних криз

4.1. Основи сигнального методу

Методологія сигнального методу (*Signalling approach*) дає можливість обрати змінні, які найкраще передбачають настання у майбутньому певної події. Наприклад, це може бути криза (валютна, банківська тощо). Наявність кризи розглядається як бінарна змінна, яка приймає значення 0 (кризи немає або спокійний період) або 1 (є криза).

Для кожної досліджуваної незалежної змінної (показника) встановлюється певний пороговий рівень значення. При перевищенні показником (незалежною змінною) цього порогового значення (*threshold* або *cut-off*) вважається, що подається сигнал (*signal*), що ймовірно буде криза (наприклад, протягом наступних 24 місяців). Сигнал вважається добрим, якщо криза дійсно відбувається, і поганим, якщо криза не відбувається. Порогове значення обирається таким шляхом, щоб мінімізувати відношення поганих (шуму – *noise*) до добрих сигналів, тобто прогнозу здатність сигналів. Можемо це проілюструвати за допомогою таблиці матриці сигналів (див. табл. 4.1):

Таблиця 4.1

Сигнал	Криза відбувається	Криза не відбувається
Є	A	B - помилка I роду (type I error)
Немає	C - помилка II роду (type II error)	D

Тут A – кількість випадків, коли показник давав сигнал і криза потім відбувалася; B – кількість випадків, коли показник давав сигнал і криза потім не відбувалася; C – кількість випадків, коли показник не давав сигнал і криза потім відбувалася; D – кількість випадків, коли показник не давав сигнал і криза потім не відбувалася.

Функція шум/сигнал *Noise-to-signal ratio* (*NSR*), яку ми маємо мінімізувати має вигляд:

$$NSR = \frac{B/(B+D)}{A/(A+C)} \quad (4.1)$$

Якщо вдається підібрати таке порогове значення показника, що співвідношення шуму до сигналів може бути менше 1, то показник можна використовувати для прогнозування. Чим менше це співвідношення, тим краще показник пояснює ймовірність настання кризи. В літературі іноді також можливо знайти обернений показник ($SNR = 1/NSR$), який максимізується.

Альтернативною функцією, яка може мінімізуватися, є функція загальної кількості помилок *Total misclassified errors* (*TME*), яка має вигляд:

$$TME = \frac{C}{A+C} + \frac{B}{B+D} \quad (4.2)$$

Сила сигналу *Signalling power* (*SP*) показника розраховується так:

$$SP = 1 - TME \quad (4.3)$$

Вона може використовуватися як зважувальний коефіцієнт при розрахунку ймовірності виникнення кризової події за формулою:

$$CP = \sum w_i I_i \quad (4.4)$$

де w_i – вага i -того показника (незалежної змінної); I_i – розрахована ймовірність кризи, коли значення i -того показника знаходиться у небезпечній зоні, якщо воно зараз знаходиться у небезпечній зоні; або розрахована ймовірність кризи, коли значення i -того показника знаходиться у безпечній зоні, якщо воно зараз знаходиться у безпечній зоні.

Показники повинні не корелювати сильно між собою, але навіть за сильної кореляції сигнальний підхід може використовуватися за умови, що сильнокорелюючі між собою показники мають одержати меншу вагу.

Але слід враховувати, що розрахована за допомогою *CP* ймовірність має недоліки внаслідок ефекту усереднення через те, що, наприклад, половина показників є проблемними, а інша половина безпечними. При цьому не враховується ефект взаємодії факторів.

Альтернативний спосіб розрахунку ймовірності є спочатку побудова певного композитного показника, за формулою:

$$CI = \sum w_i d_i, \quad (4.5)$$

де w_i – вага i -того показника (незалежної змінної); d_i – логічна змінна по i -тому показнику (приймає значення 1, якщо значення i -того показника зараз знаходиться у небезпечній зоні; або 0, якщо значення i -того показника зараз знаходиться у безпечній зоні).

Після цього будується логіт-регресія, де залежною змінною є логічна змінна наявності кризи, а незалежною – значення CI (про побудову логіт-регресії йдеться у главі, присвяченій нелінійній регресії).

4.2. Дослідження валютних криз сигнальним методом у Microsoft Excel

Наведемо приклад у Microsoft Office Excel. Вважатимемо, що в країні відбувається валютна криза, якщо протягом року її валюта знецінилася більше, ніж на 25%. Як показник, який має давати сигнал оберемо поточний рахунок відносно ВВП у % для спрощення лише у 2007 році. Сигнал має давати інформацію, що в країні відбудеться валютна криза протягом наступних трьох років (2008-2010 роки).

Вводимо дані в стовпчики ліворуч, а праворуч прописуємо необхідні формули для одержання результату. Знизу на малюнку вказані фрагменти таблиці в двох видах: вхідні дані та результати обробки (Рис.4.1) та текст формул (Рис.4.2).

	A	B	C	D	E	F	G
1	Країна	ПР/ВВП2007	Криза2008_10	Сигнал	Порогове значення	Шум/Сигнал	
2	Angola	17.50	0	0	-6	0.529411765	
3	Albania	-10.75	0	1		Криза	Немає кризи
4	Argentina	2.82	0	0	Сигнал	6	36
5	Armenia	-6.40	0	1	Немає сигналу	3	66
6	Australia	-6.77	0	1			
7	Azerbaijan	27.29	0	0	Ймовірність кризи, якщо сигнал є		
8	Burkina Faso	-15.81	0	1	0.167		
9	Bangladesh	1.25	0	0			
10	Bulgaria	-27.16	0	1	Ймовірність кризи, якщо сигналу немає		
11	Bahrain	15.73	0	0	0.045		
12	Belarus	-6.70	1	1			
13	Bolivia	12.13	0	0	Загальна кількість помилок		
14	Brazil	0.11	0	0	0.686		
15	Brunei Darussalan	39.42	0	0			
16	Botswana	14.51	0	0	Сила сигналу		
17	Canada	0.84	0	0	0.314		
18	Switzerland	9.30	0	0			

Рис.4.1.

Якщо, наприклад у комірку E2 (порогове значення) поставити значення 5, то в комірці F2 (функція шум/сигнал) значення буде 0.98. Ітераційним шляхом ми маємо змінювати значення у комірці E2 так, щоб мінімізувати функцію шум/сигнал або функцію загальної кількості помилок. Нажаль, автоматично знайти рішення складно, оскільки може існувати декілька локальних мінімумів або максимумів функції.

В нашому прикладі після перебору варіантів ми одержали значення порогового рівня для поточного рахунку -6% ВВП.

	D	E	F	G
1	Сигнал	Порогове значення	Шум/Сигнал	
2	=ЕСЛИ(B2<=\$E\$2;1;0)	-6	=(G4/(G4+G5))/(F4/(F4+F5))	
3	=ЕСЛИ(B3<=\$E\$2;1;0)		Криза	Немає кризи
4	=ЕСЛИ(B4<=\$E\$2;1;0)	Сигнал	=СЧЕТЕСЛИМН(C2:C112;"=1";D2:D112;"=1")	=СЧЕТЕСЛИМН(C2:C112;"=0";D2:D112;"=1")
5	=ЕСЛИ(B5<=\$E\$2;1;0)	Немає сигналу	=СЧЕТЕСЛИМН(C2:C112;"=1";D2:D112;"=0")	=СЧЕТЕСЛИМН(C2:C112;"=0";D2:D112;"=0")
6	=ЕСЛИ(B6<=\$E\$2;1;0)			
7	=ЕСЛИ(B7<=\$E\$2;1;0)	Ймовірність кризи, якщо сигн		
8	=ЕСЛИ(B8<=\$E\$2;1;0)	=F4/G4		
9	=ЕСЛИ(B9<=\$E\$2;1;0)			
10	=ЕСЛИ(B10<=\$E\$2;1;0)	Ймовірність кризи, якщо сигн		
11	=ЕСЛИ(B11<=\$E\$2;1;0)	=F5/G5		
12	=ЕСЛИ(B12<=\$E\$2;1;0)			
13	=ЕСЛИ(B13<=\$E\$2;1;0)	Загальна кількість помилок		
14	=ЕСЛИ(B14<=\$E\$2;1;0)	=F5/(F4+F5)+G4/(G4+G5)		
15	=ЕСЛИ(B15<=\$E\$2;1;0)			
16	=ЕСЛИ(B16<=\$E\$2;1;0)	Сила сигналу		
17	=ЕСЛИ(B17<=\$E\$2;1;0)	=1-E14		
18	=ЕСЛИ(B18<=\$E\$2;1;0)			

Рис.4.2.

Йому відповідає значення функції шум/сигнал 0.53. Воно менше 1, тобто розглянутий показник може використовуватися для прогнозування валютної кризи. Ми бачимо, що якщо сальдо поточного рахунку становить більше ніж -6%ВВП, ймовірність валютної кризи протягом наступних трьох років невелика – лише 4.5%. Якщо ж поточний рахунок становить менше -6%ВВП (великий дефіцит), то ймовірність валютної кризи в 3-4 рази вища і становить 16.7%.

Увага: формули таблиці прописані так, що нижче значення показника розглядається як більш небезпечне, ніж вище, оскільки небезпечним є дефіцит поточного рахунку. Для показників, у яких небезпечними є вищі значення (наприклад відношення зовнішнього боргу до ВВП), формули у рядку D мають бути іншими – типу: =ЕСЛИ(B2>=\$E\$2;1;0)

Тепер на прикладі за умовними даними розглянемо, як можливо урахувати дані по декільком показникам одночасно (див. Рис.4.3-4.4). Припустимо, у нас є три показники, по яких ми зробили подібні розрахунки щодо ймовірності валютної кризи, якщо сигнал є або немає, порогового рівня та напряму небезпечного діапазону (наприклад для першого показника значення більше 80 означає більшу ймовірність кризи), сили сигналу. В рядку 9 ми проставляємо поточні значення показників у країні для того, щоб спрогнозувати ймовірність кризи. В решті решт у клітинці A14 ми одержуємо ймовірність кризи 0.0515. Значення невелике переважно через те, що показник 1 (він має найбільшу силу сигналу) не подає сигнал.

	A	B	C	D
1				
2		Показник 1	Показник 2	Показник 3
3				
4				
5	Ймовірність кризи, якщо сигнал є	0.5	0.1	0.2
6	Ймовірність кризи, якщо сигналу немає	0.03	0.07	0.06
7	Напряв небезпечного діапазону	>	<	<
8	Пороговий рівень	80	-3	1
9	Поточне значення показника	60	-5	0
10	Наявність сигналу	0	1	1
11	Сила сигналу	0.55	0.05	0.15
12		0.733	0.067	0.200
13		0.022	0.007	0.040
14	Ймовірність кризи			
15		0.068666667		

Рис.4.3.

4.3. Застосування сигнального методу для аналізу факторів фіскальних криз

Далі буде описаний методологічний алгоритм для дослідження факторів фіскальних криз застосований у [56], а також можливості для використання наукових результатів у ній для аналізу поточної ситуації. Ці результати можуть бути використані для пояснення впливу факторів, для визначення заходів з метою попередження фіскальних криз, прогнозу.

	A	B	C	D
1				
2		Показник 1	Показник 2	Показник 3
3				
4				
5	Ймовірність кризи, якщо сигнал є	0.5	0.1	0.2
6	Ймовірність кризи, якщо сигналу немає	0.03	0.07	0.06
7	Напрямок небезпечного діапазону		>	<
8	Пороговий рівень	80	-3	1
9	Поточне значення показника	80	-5	0
10	Нааявність сигналу	=ЕСЛИ(B9>B8;1;0)	=ЕСЛИ(C9<C8;1;0)	=ЕСЛИ(D9<D8;1;0)
11	Сила сигналу	0.55	0.05	0.15
12		=B11/СУММ(\$B\$11:\$D\$11)	=C11/СУММ(\$C\$11:\$D\$11)	=D11/СУММ(\$B\$11:\$D\$11)
13		=ЕСЛИ(B10=1,B5*B12,B6*B12)	=ЕСЛИ(C10=1,C5*C12,C6*C12)	=ЕСЛИ(D10=1,D5*D12,D6*D12)
14	Ймовірність кризи			
15	=СУММ(B13:D13)			

Рис.4.4.

В якості залежної змінної використовується логічна змінна, яка базується на чотирьох змінних: дефолт за боргом або його реструктуризація, неявний дефолт, надзвичайне фінансування МВФ та значна обмеженість ринкового фінансування. Остання змінна враховується, щоб виявити епізоди, коли формально ситуація не розглядалася як боргова криза, але ставки на ринках урядових облігацій створювали суттєве напруження. Детальніше визначення компонентів фінансової кризи вказані у табл. 4.2. Фінансова криза в країні відбувається, якщо виконується принаймні одна з чотирьох подій.

Таблиця 4.2.

Компоненти фінансової кризи

Подія	Критерій	Визначення критерію	Джерело
Дефолт за боргом або його реструктуризація	Неможливість обслуговування боргу або вимушена конвертація боргу	Визначення S&P	Standards and Poor's and IMF's Private Market Financing for Developing Countries
Значне фінансування	Значна програма, підтримувана МВФ	Доступ до більше 100% квоти	IMF's Finance Department database
Неявний дефолт	Висока інфляція	Річна інфляція більше 35% для розвинутих країн або більше 500% для новітніх ринків	Standards and Poor's and IMF's Private Market Financing for Developing Countries
Значна обмеженість ринкового фінансування	Тиск на доходність суверенного боргу	Спред за суверенним боргом більше 1000 базових пунктів або більше ніж на 2 стандартних відхилення середнього рівня для країни	IFS, Bloomberg and Datastream, річні та місячні дані

Для розвинутих країн довгострокові спреди по внутрішнім облігаціям (*Long-term domestic bond spreads*) і спреди за 5-річними свопами кредитного дефолту (*credit default swap (CDS)*). Для новітніх ринків – індекс облігацій новітніх ринків (*EMBI spreads*) та довгострокові спреди по внутрішнім облігаціям. Спреди розраховуються як різниця у доходності облігацій порівняно з доходністю аналогічних облігацій США.

Фінансові кризи в одній країні вважаються різними, якщо між ними проходить принаймні 2 роки. При цьому, лише перший рік кризи береться до уваги. Аналізуються дані по 29 розвинутих країнам (*advanced economies*) і 52 новітнім ринкам (*emerging markets*) за 1970-2009 роки. Розглянуто 47 змінних, розподілені по 4 групах: фінансовий вплив, фінансова стабільність, вразливість і демографічні тенденції. У табл.4.3 вказані визначення показників, які є найбільш корисними у прогнозі.

Автори надають перевагу мінімізації функції загальної кількості помилок - *Total misclassified errors (TME)*. У табл. 4.4. вказані результати розрахунків по кожному показнику (по групі країн новітні ринки) серед числа тих, які є найбільш корисними у прогнозі (тут ми наводимо лише результати для новітніх ринків).

З метою узагальнення прогнозу по всіх показниках одночасно, замість безпосереднього розрахунку ймовірності, будується індекс фінансового стресу (*fiscal stress index*), який обчислюють за формулою:

$$FSI = \sum w_i d_i, \quad (4.6)$$

Таблиця 4.3.

Показник українською мовою	Показник англійською мовою	Джерело: примітки
Первинний баланс з поправкою на циклічність	Cyclically adjusted primary balance	Оцінки персоналу: доходи загального уряду – витрати загального уряду (мінус процентні платежі) з поправкою на тренд випуску
Фіскальний імпульс	Fiscal Impulse	Оцінки персоналу, від'ємне значення зміни структурного балансу
Частка податкових надходжень	Public tax revenue share	WEO
Розрив ВВП (%потенційного ВВП)	Output gap (% potential GDP, in abs. value)	Оцінки персоналу
Державний борг (% багатства фінансового сектору)	Public debt (% private sector wealth)	Eurostat (фінансове багатство приватного сектору)
Валовий зовнішній борг (державний і приватний)	Gross external debt (public and private)	WEO (новітні ринки), JEDH (розвинуті країни)
Чистий державний борг	Net public sector debt	WEO
Державний борг на душу населення (дол.)	Public debt stock per capita (US dollars)	WBWDI (населення)
Поточний рахунок	Current account	WEO
Короткостроковий державний борг (% від загального державного боргу)	Short-term public debt (% total public debt)	VEE (новітні ринки), BIS (розвинуті країни)
Короткостроковий зовнішній борг (% від валових резервів)	Short-term external debt (% gross reserves)	BIS
Зовнішній державний борг (%загального державного боргу)	External public debt (% total public debt)	BIS
Волатильність частки державних витрат	Volatility of public expenditure share	Оцінки персоналу: стандартне відхилення витрат загального уряду у% до ВВП поділене на середнє за останні 5 років
Кредитний рейтинг	Credit Ratings	S&P: AAA=10, AA+=9, AA=8, AA-=7, A+=6, A=5, A-=4, BBB+=3, BBB=2, BBB-=1, BB+=0, BB=-1, BB=-2, B+=-3, B=-4, B=-5, CCC=-6, CC=-7, C=-8, SD=-9.2
Зміна рейтингу	Rating Actions	Оцінки персоналу: сума балів "+1" у кредитному рейтингу
Витрати на державні пенсії (%ВВП)	Public pension spending (% GDP)	OECD
Витрати на охорону здоров'я (%ВВП)	Healthcare spending (% GDP)	OECD
Державні витрати на охорону здоров'я (% всіх витрат)	Public health spending (% total spending)	OECD
Ефективний вік виходу на пенсію - чоловіки	Effective Retirement Age- Men	OECD
Ефективний вік виходу на пенсію - жінки	Effective Retirement Age - Women	OECD

Примітка: часто показники беруться у відносному вимірі, наприклад у % ВВП.

де w_i – вага i -того показника (незалежної змінної);

d_i – логічна змінна по i -тому показнику (приймає значення 1, якщо значення i -того показника зараз знаходиться у небезпечній зоні; або 0, якщо значення i -того показника зараз знаходиться у безпечній зоні).

Використавши прогнози дані на 2010 рік і методику індексу фіскального стресу, автори зробили рейтинг країн. Середній FSI становив для розвинутих країн 0.62, для новітніх ринків 0.35.

Серед розвинутих країн найвищий FSI був у Португалії, Ісландії (більше 0.8), Японії, Іспанії, Італії, Нової Зеландії, Німеччини, Греції (більше 0.75), Бельгії, Австрії, Ірландії, Італії, Великобританії (більше 0.7), Сінгапурі, Франції (більше 0.62).

Серед новітніх ринків найвищий FSI був у Латвії, Лівані, Литві, Естонії (більше 0.5), Грузії, Вірменії, Хорватії, Угорщині, Румунії, Сербії (більше 0.4), Ямайці, Македонії, Україні, Росії (більше 0.35).

Результати розрахунків по показниках

Показник	Напрямок безпечного діапазону значень	Пороговий рівень	Ймовірність фіскальної кризи для безпечного діапазону	Ймовірність фіскальної кризи для небезпечного діапазону	Сила сигналу (1-ТМЕ)	Ваговий коефіцієнт
Первинний баланс з поправкою на циклічність	>	-2.9	0.152	0.078	0.11	2.3
Фіскальний імпульс	<	2.0	0.136	0.072	0.14	3.0
Частка податкових надходжень	>	11.4	0.173	0.068	0.16	3.4
Розрив ВВП (% потенційного ВВП)	<	1.8	0.141	0.055	0.24	5.0
Державний борг (% багатства фінансового сектору)	<	6.4	0.123	0.000	0.15	3.2
Валовий зовнішній борг (державний і приватний)	<	59.7	0.133	0.075	0.11	2.4
Чистий державний борг	<	62.4	0.145	0.079	0.13	2.8
Державний борг на душу населення (дол.)	<	1715.4	0.141	0.063	0.19	4.0
Поточний рахунок	>	-3.7	0.133	0.050	0.25	5.2
Короткостроковий державний борг (% від загального державного боргу)	<	35.1	0.215	0.084	0.14	3.0
Короткостроковий зовнішній борг (% від валових резервів)	<	79.0	0.150	0.056	0.25	5.2
Зовнішній державний борг (% від загального державного боргу)	<	16.1	0.140	0.038	0.26	5.5
Волатильність частки державних витрат	<	8.5	0.130	0.073	0.13	2.9
Кредитний рейтинг	>	-1.0	0.119	0.067	0.14	3.1
Зміна рейтингу	>	-1.0	0.225	0.083	0.10	2.1
Витрати на державні пенсії (%ВВП)	<	5.2	0.088	0.033	0.23	5.0
Витрати на охорону здоров'я (%ВВП)	<	5.9	0.233	0.099	0.21	4.4
Державні витрати на охорону здоров'я (% всіх витрат)	<	70.6	0.233	0.080	0.25	5.4
Ефективний вік виходу на пенсію - чоловіки	>	66.1	0.173	0.065	0.22	4.7
Ефективний вік виходу на пенсію - жінки	>	65.7	0.148	0.041	0.16	3.5
Решта показників	-	-	-	-	-	28.3

Примітка: ймовірності фіскальної кризи для безпечного і небезпечного діапазону кожного показника розраховані нами, виходячи з даних у дослідженні про кількість кризових і спокійних періодів та часток помилок I та II типу.

Інтуїтивно можемо сказати, що в цілому FSI тяжіє до адекватного відображення поточної ситуації, але не завжди (наприклад, хоча Греція займає одне з перших місць, вона знаходиться на одному рівні з Німеччиною).

Спробуємо для оцінки FSI використати поточні значення показників України (див. табл. 4.5). Проте будемо враховувати не всі показники, а лише ті, які є найбільш корисними у прогнозі (дають меншу кількість помилок) і по яких можемо знайти попередні/оцінені дані за поточний період (2011 рік). Для цього відповідно потрібно збільшити ваговий коефіцієнт кожного показника за рахунок усунення вагових коефіцієнтів тих показників, які дають більшу кількість помилок у прогнозі.

FSI для України становить у поточний момент 0.55. Він є більшим ніж у розрахований нами аналогічним шляхом за обмеженою кількістю показників FSI 2010 році 0.42. Розрахована ймовірність виникнення фіскальної кризи як зважена середня ймовірностей становить 0.11. Але слід враховувати, що розрахована таким чином ймовірність має недоліки через ефект усереднення через те, що половина показників є проблемними, а інша половина безпечними. При цьому не враховується ефект взаємодії факторів.

Дані по Україні

Показник	Попереднє / оцінене значення у 2011 р.	На основі джерела	Чи у небезпечному діапазоні	Скоригований ваговий коефіцієнт
Частка податкових надходжень	20	Міністерство фінансів України, WEO	0	8.7
Валовий зовнішній борг (державний і приватний)	76	Національний банк України, WEO	1	6.2
Чистий державний борг	38	WEO	0	7.2
Державний борг на душу населення (дол.)	1420	WEO	0	10.3
Поточний рахунок	-3.9	WEO	1	13.4
Короткостроковий зовнішній борг (% від валових резервів)	155	Національний банк України	1	13.4
Зовнішній державний борг (% від загального державного боргу)	56 (64 включаючи гарантований державою)	Міністерство фінансів України	1	14.1
Волатильність частки державних витрат	4	WEO	0	7.5
Кредитний рейтинг	-3	http://chartsbin.com	1	8.0
Витрати на охорону здоров'я (%ВВП)	3.3	Міністерство фінансів України	0	11.3

Кращу оцінку ймовірності могла би дати побудова логіт-регресії, де залежною змінною є логічна змінна наявності фіскальної кризи, а незалежною – значення FSI. Проте для її побудови необхідно мати повні дані по всіх країнах про FSI та наявність фіскальної кризи з лагом, що доступне, зазвичай лише авторам дослідження.

Наприклад, на Рис. 4.5 вказана залежність між ймовірністю виникнення фіскальної кризи та FSI знайдена у іншому дослідженні [57]:

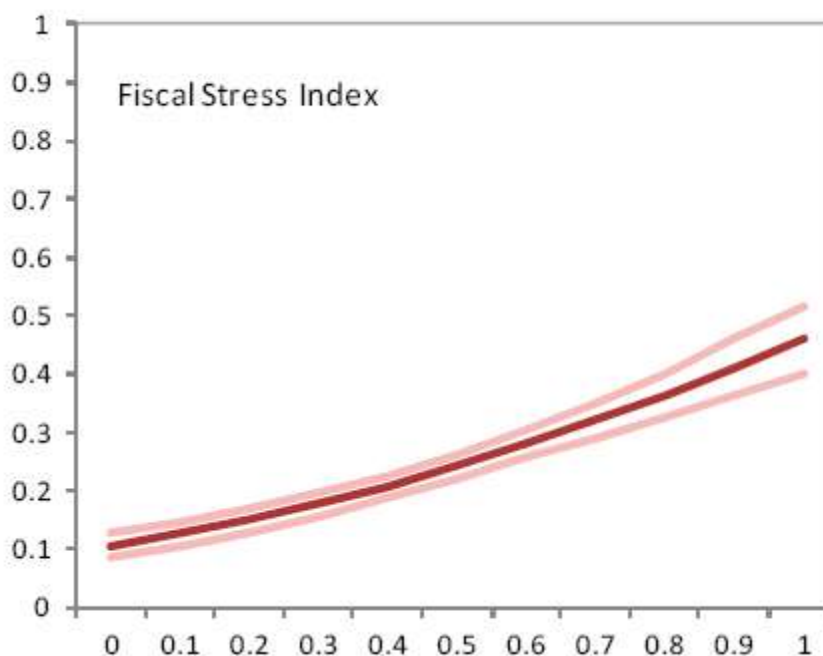


Рис.4.5.

Але для наших розрахунків вона не може бути використана, оскільки розраховувалася за частково іншими даними та іншим набором показників.

ГЛАВА 5. Дослідження іноземних інвестицій методами аналізу середніх

5.1. Методи порівняння середніх

В практичних економічних дослідженнях часто зустрічаються випадки, коли середній результат деякої ознаки однієї серії експериментів відрізняється від середнього результату іншої серії. Оскільки середнє – це результати вимірювань, то, як правило, вони завжди відрізняються, питання тільки в тому, чи можна пояснити виявлену розбіжність середніх неминучими випадковими помилками експерименту або воно викликане певними причинами.

Аналіз середніх нагадує аналіз частот, але замість категоріальної залежної змінної використовується залежна змінна, яка вимірюється у метричній шкалі. Можливе також використання аналізу середніх і для залежної змінної у порядковій шкалі. Незалежна змінна має бути категоріальною або штучно перетворена у категоріальну, подібно до того, як це ми робили в рамках частотного аналізу.

Замість χ^2 тесту як у частотному аналізі, у випадку аналізу середніх зазвичай використовується t -тест (t -критерій) Стьюдента, який показує наскільки значуще відрізняються середні залежної змінної у різних групах спостережень при групуванні за незалежною змінною. Наприклад, середній приріст імпорту в країнах з фіксованим валютним курсом від середнього приросту імпорту в країнах з плаваючим валютним курсом (див. табл.5.1).

Таблиця 5.1.

Група	Фіксований валютний курс (I група)	Плаваючий валютний курс (II група)
Середній приріст імпорту, %	11	7

Якщо різниця є значущою (зазвичай йдеться про рівень значущості p менше 0.05), це означає, що режим валютного курсу впливає на приріст імпорту (хоча можливий і зворотній вплив приросту імпорту на вибір режиму валютного курсу). Більш точно рівень значущості (p) різниці між середніми – це імовірність того, що дві вибірки (два діапазони, масиви значень) взяті з генеральних сукупностей, які мають однакові середні.

t -критерій буде показувати тим більш значущу різницю, чим:

- більша кількість спостережень у досліджуваній вибірці;
- більша різниця у середніх;
- менша дисперсія залежної змінної в кожній групі.

При проведенні аналізу середніх за допомогою звичайного t -критерію Стьюдента мають виконуватися умови:

1. Має бути достатньо спостережень у кожній групі спостережень (бажано більше 30, щоб результати були більш надійними).

2. Залежна змінна має нормальний розподіл у кожній групі. Якщо ця умова не виконується, кращим варіантом є використання непараметричних аналогів t -критерію Стьюдента.

3. Дисперсія залежної змінної має бути приблизно однаковою в різних групах. Різниця у дисперсіях перевіряється за допомогою F -критерію або тесту Левена. Якщо ця умова не виконується, використовується спеціальний варіант t -критерію Стьюдента для груп з різною дисперсією.

4. Вибірки (спостереження) мають бути незалежними. Якщо ця умова не виконується, використовується спеціальний варіант t -критерію Стьюдента для залежних вибірок.

5.2. Критерій Стьюдента

Розглянемо один з варіантів перевірки гіпотези про рівність середніх, у припущенні, що вибірки незалежні, ознаки мають нормальний розподіл порівнюються між собою тільки дві сукупності. Якщо одне з цих припущень не виконується, то застосовують непараметричні критерії.

Нехай є дві сукупності даних (генеральні сукупності) X та Y , які характеризуються середніми \bar{x}_0 та \bar{y}_0 і відомими дисперсіями σ_x^2 та σ_y^2 . З цих сукупностей беруть дві незалежні вибі-

рки обсягами n_1 і n_2 , за якими знайдені середні арифметичні \bar{x} та \bar{y} і виправлені вибіркві дисперсії S_x^2 та S_y^2 . Необхідно перевірити гіпотезу H_0 про рівність генеральних середніх $H_0: \bar{x}_0 = \bar{y}_0$. Іншими словами, перевіряють гіпотезу про те, що вибірки належать одній і тій же генеральній сукупності.

Відомо, що достатньо великих n_1 і n_2 , середні арифметичні \bar{x} та \bar{y} мають наближено нормальний розподіл. Тобто, $\bar{x} \approx N(\bar{x}_0; \sigma_x/n_1)$ і $\bar{y} \approx N(\bar{y}_0; \sigma_y/n_2)$.

У випадку справедливості гіпотези H_0 статистика:

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}} \quad (5.1)$$

має нормальний розподіл $N(0;1)$. Якщо приймають конкуруючу гіпотезу $H_1: \bar{x}_0 \neq \bar{y}_0$, то критична область критерію будується таким чином:

- якщо виконується нерівність $|t| > t_{kp}$, то гіпотезу H_0 відхиляють;
- якщо виконується нерівність $|t| \leq t_{kp}$, то гіпотеза H_0 приймається.

При цьому t_{kp} визначається з рівняння $\Phi(t_{kp}) = \Phi(t_{1-p}) = 1 - p$, де p – рівень значущості критерію, а $\Phi(x)$ - функція Лапласа.

Якщо σ_x^2 та σ_y^2 невідомі, але рівні, тобто $\sigma_x^2 = \sigma_y^2 = \sigma$, то використовують статистику:

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{n_1 S_x^2 + n_2 S_y^2}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}, \quad (5.2)$$

де S_x^2 та S_y^2 - вибіркві дисперсії (оцінки σ_x^2 та σ_y^2).

Відомо, що t -критерій (5.2) має розподіл Стьюдента з $k = n_1 + n_2 - 2$ ступенями вільності. При цьому:

- якщо виконується нерівність $|t| > t_{k,p}$, то гіпотезу H_0 відхиляють;
- якщо виконується нерівність $|t| \leq t_{k,p}$, то гіпотезу H_0 приймають.

Критерії (5.1) і (5.2) застосовують для перевірки гіпотези про рівність середніх, якщо обсяги обох вибірок малі і дисперсії рівні, в протилежному випадку застосовують іншу формулу. При великих обсягах, умова рівності дисперсій втрачає свою актуальність. Обсяги вибірок вважають малими, якщо $n_1 < 30$ і $n_2 < 30$. Якщо обсяги вибірок рівні, тобто $n_1 = n_2$, формула для обчислення t -критерію значно спрощується і набуває вигляду:

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{S_x^2 + S_y^2}} \cdot \sqrt{\frac{n-1}{n}}. \quad (5.3)$$

Розглянемо процедуру порівняння дисперсій у двох сукупностях X та Y , які мають нормальний розподіл с дисперсіями σ_x^2 та σ_y^2 . Перевіряють гіпотезу про рівність дисперсій, тобто $H_0: \sigma_x^2 = \sigma_y^2 = \sigma$. Для перевірки цієї гіпотези з сукупностей беруть дві незалежні вибірки обсягами n_1 і n_2 . За вибірками розраховують виправлені вибіркві дисперсії S_x^2 та S_y^2 . Розглядають F -критерій, тобто величину:

$$F = \frac{S_x^2}{S_y^2}, \quad (5.4)$$

(в чисельнику ставлять більшу дисперсію). Відомо, що величина F має розподіл Фішера з $k_1 = n_1 - 1, k_2 = n_2 - 1$ ступенями вільності. При цьому:

- якщо виконується нерівність $F > F_{p,k_1,k_2}$, то гіпотезу H_0 відхиляють;
- якщо виконується протилежна нерівність $F \leq F_{p,k_1,k_2}$, то гіпотезу H_0 приймають.

Отже, t -критерій застосовують, коли розміри вибірки невеликі, змінні мають нормальний розподіл, а дисперсії спостережень в групах не дуже відрізняються. Припущення про нормальний розподіл можна перевірити, досліджуючи розподіл (наприклад, візуально за допомогою гістограми) або застосовуючи інші критерії. Рівність дисперсій в двох групах можна перевірити за допомогою F -критерію або використовувати більш стійкий критерій Левена. Рівень значущості p критерію Стьюдента дорівнює ймовірності помилково відхилити гіпотезу про рівність середніх двох вибірок, якщо в дійсності ця гіпотеза вірна.

5.4. Аналіз чинників припливу інвестицій у Microsoft Excel

Для аналізу середніх у надбудові *Пакет аналіза* оберіть трьох один з можливих варіантів t -тесту. *Двухвыборочный t-тест с одинаковыми дисперсиями* та *Двухвыборочный t-тест с различными дисперсиями*. *Парный двухвыборочный t-тест для средних*. Розміри вибірок мають бути обов'язково однаковими лише у випадку парного тесту, в інших варіантах тесту вибірки можуть бути й різного розміру.

Використаємо для ілюстрації цих тестів ті самі приклади. Спочатку обираємо *Двухвыборочный t-тест с одинаковыми дисперсиями* для даних, вказаних на Рис. 5.20. Заповнюємо діалогове вікно аналізу (див. Рис.5.21).

Результати будуть виводитися у новий робочий аркуш (ми одразу даємо йому назву, наприклад, *ttestsamevar*). Як видно з Рис.5.22, середні ПІ по двох групах становлять 2.3 і 3.5 (% від ВВП). З результатів ми бачимо рівень значущості 0.093769 (для двостороннього t -тесту), який більший за 0.05, але менше 0.1. Отже різниця між середнім відносно незначуща.

Для того, щоб перевірити значущість різниці дисперсії у двох групах, потрібно скористатися опцією *Двухвыборочный F-тест для дисперсии* у *Пакет аналіза* для тих самих даних (див. Рис.5.23).

Як бачимо результатів на Рис.5.24, різниця у дисперсії є значущою (0.007529 менше 0.05). Але в Microsoft Excel використовується односторонній тест. Тому можливо порівнювати 0.007529 не з критичним рівнем 0.05, а з критичним рівнем 0.025.

Н	I
FDIlowincome	FDIhighincome
3	5
5	9
2	4
3	2
2	3
5	11
1	6
1	3
4	4
4	5
3	2
4	2
0	1
1	1
2	2
1	4
2	3
2	2
1	0
0	1

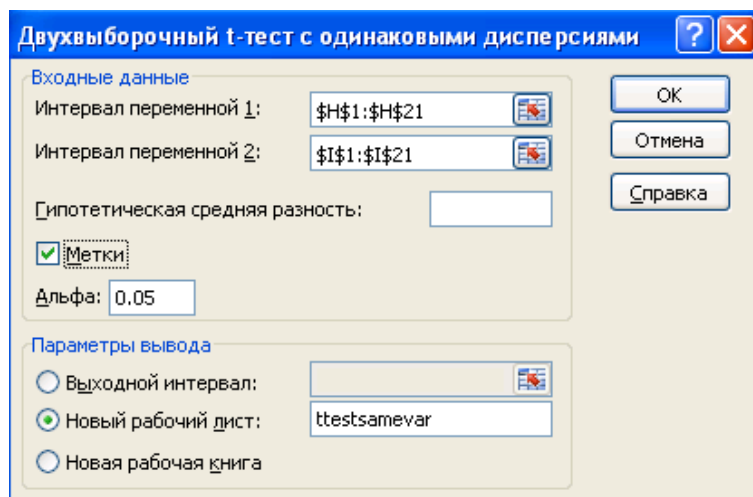


Рис.5.20.

	A	B	C
1	Двухвыборочный t-тест с одинаковыми дисперсиями		
2			
3		<i>FDlowincome</i>	<i>FDhighincome</i>
4	Среднее	2.3	3.5
5	Дисперсия	2.326315789	7.421052632
6	Наблюдения	20	20
7	Объединенная дисперсия	4.873684211	
8	Гипотетическая разность средних	0	
9	df	38	
10	t-статистика	-1.718907685	
11	P(T<=t) одностороннее	0.046884553	
12	t критическое одностороннее	1.68595446	
13	P(T<=t) двухстороннее	0.093769106	
14	t критическое двухстороннее	2.024394164	

Рис.5.22.

Рис.5.23.

Тепер для даного прикладу використаємо *Двохвибірковий t-тест с різними дисперсіями* (див. Рис.5.25). З результатів на Рис. 5.26, ми бачимо, що рівень значущості для *t*-тесту з різною дисперсією в групах становить 0.095989.

	A	B	C
1	Двухвыборочный F-тест для дисперсии		
2			
3		<i>FDlowincome</i>	<i>FDhighincome</i>
4	Среднее	2.3	3.5
5	Дисперсия	2.326315789	7.421052632
6	Наблюдения	20	20
7	df	19	19
8	F	0.313475177	
9	P(F<=f) одностороннее	0.007529238	
10	F критическое одностороннее	0.461201089	

Рис.5.24.

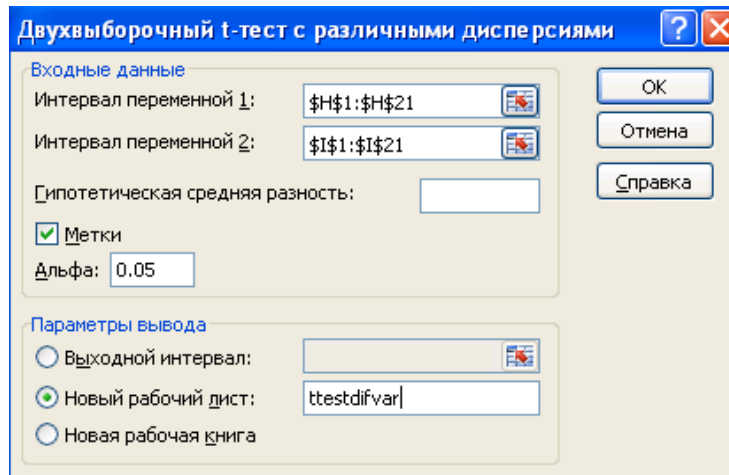


Рис.5.25.

	A	B	C
1	Двухвыборочный t-тест с различными дисперсиями		
2			
3		<i>FDlowincome</i>	<i>FDhighincome</i>
4	Среднее	2.3	3.5
5	Дисперсия	2.326315789	7.421052632
6	Наблюдения	20	20
7	Гипотетическая разность средних	0	
8	df	30	
9	t-статистика	-1.718907685	
10	P(T<=t) одностороннее	0.047967861	
11	t критическое одностороннее	1.697260887	
12	P(T<=t) двухстороннее	0.095935722	
13	t критическое двухстороннее	2.042272456	

Рис.5.26.

Тепер використаємо *Парний двухвыборочный t-тест для средних* для даних на Рис. 5.27. Діалогове вікно вказане на Рис. 5.28. З таблиці результатів (Рис. 5.29) видно, що різниця між середніми ПІ по двох вибірках значуща ($p=0.001576$, тобто менше 0.05).

Аналіз середніх також можливо проводити без надбудови *Пакет анализа* за допомогою функцій:

=ТТЕСТ(діапазон1;діапазон2;хвости;тип)

Повертає рівень значущості різниці між середніми. Аргументами діапазон 1 і діапазон 2 є два набори значень, середні по яких порівнюються. Хвости: якщо 1 – функція повертає результати одностороннього t-тесту, якщо 2 – двостороннього. Тип може приймати значення:

- 1 – парний тест для залежних вибірок;
- 2 – тест для незалежними вибірок з однаковими дисперсіями;
- 3 – тест для незалежними вибірок з різними дисперсіями.

Приклад: =ТТЕСТ(H2:H23;I2:I21;2;2).

F	G
FDIbeforeLib	FDIafterLib
6	7
5	9
4	6
2	3
4	3
11	15
5	8
5	6
7	5
8	8
2	7
0	3
1	3
1	2
2	5
4	5
3	3
2	3
2	4
1	0

Рис.5.27.

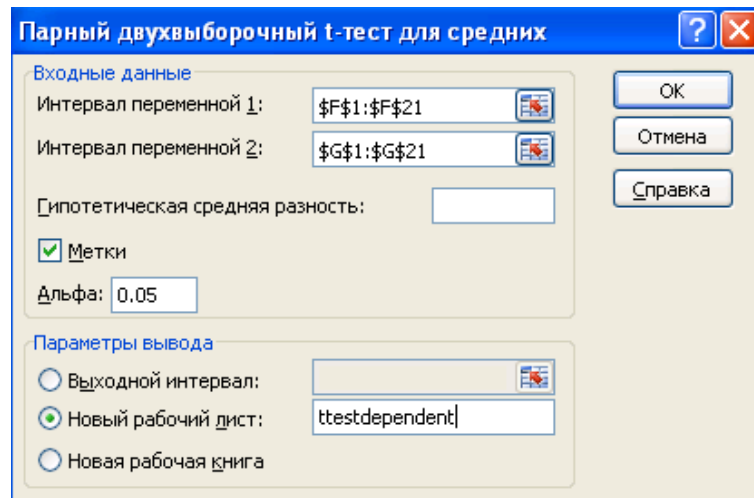


Рис.5.28.

	A	B	C
1	Парный двухвыборочный t-тест для средних		
2			
3		FDIbeforeLib	FDIafterLib
4	Среднее	3.75	5.25
5	Дисперсия	7.56578947	10.61842105
6	Наблюдения	20	20
7	Корреляция Пирсона	0.82942622	
8	Гипотетическая разность средних	0	
9	df	19	
10	t-статистика	-3.68394199	
11	P(T<=t) одностороннее	0.0007882	
12	t критическое одностороннее	1.72913281	
13	P(T<=t) двухстороннее	0.0015764	
14	t критическое двухстороннее	2.09302405	

Рис.5.29.

=ФТЕСТ(диапазон1;диапазон2)

Повертає рівень значущості різниці між дисперсіями. Аргументами діапазон 1 і діапазон 2 є два набори значень, дисперсії по яких порівнюються.

5.5. Непараметричні аналоги *t*-критерію Стьюдента

Непараметричні аналоги *t*-тесту використовуються, якщо розподіл залежної змінної відрізняється від нормального розподілу або невідомий. Для вказаних тут непараметричних методів залежні змінні мають вимірюватися у порядковій чи метричній шкалі. Якщо, вони вимірюються у номінальній шкалі, то використовуються методи частотного аналізу.

Аналогом *t*-тесту для різниці середніх у незалежних вибірках за групами є тести:

- Вальда-Волфовіца (*Wald-Wolfowitz runs test*);
- Колмогорова-Смірнова (*Kolmogorov-Smirnoff two-sample test*);
- Манна-Уїтні (*Mann-Whitney U test*).

Ці методи можуть давати інформацію про відмінність не тільки середніх, а й параметрів розподілу в цілому у двох вибірках.

ГЛАВА 6. Дисперсійний аналіз міжнародної торгівлі та інвестицій

6.1. Теоретичні основи методу дисперсійного аналізу

Дисперсійний аналіз можна визначити як параметричний метод, призначений для оцінки впливу різних факторів на результат експерименту, а також для наступного планування експериментів. Тому в дисперсійному аналізі можна досліджувати залежності кількісної ознаки від одного чи багатьох ознак-факторів.

Дисперсійний аналіз (або аналіз варіації або *ANOVA – analysis of variance*), як і аналіз середніх за допомогою *t*-тесту, має мету – визначити наскільки значуще відрізняються середні залежної змінної у різних групах спостережень при групуванні за незалежною змінною (групуючою змінною або фактором). Дисперсійний аналіз має більше можливостей. Він дозволяє:

- виділяти як дві групи значень за фактором (незалежною змінною), так і більше;
- досліджувати взаємодію факторів;
- досліджувати вплив факторів на декілька залежних змінних одночасно.

В основі дисперсійного аналізу лежить поділ загальної дисперсії залежної змінної на дві складові: внутрішню групову дисперсію та міжгрупову дисперсію. Міжгрупова дисперсія, яка є відображенням різниці у середніх по групах, є наслідком наявності ймовірного впливу фактору. Внутрішньогрупова дисперсія – є наслідком дії інших факторів, які ми не враховуємо у дослідженні. Чим більшою є міжгрупова дисперсія порівняно з внутрішньою груповою, тим більш значущою буде різниця між середніми залежної змінної у групах, а отже більшим буде вплив фактору. Ця значущість перевіряється за допомогою *F*-тесту.

Найпростішим варіантом дисперсійного аналізу є однофакторний одномірний дисперсійний аналіз (*one-way univariate ANOVA*). Якщо за групуючою змінною всі спостереження поділяються лише на дві групи, то дисперсійний аналіз дасть той самий результат, що й *t*-тест для аналізу середніх. Припустимо, ми аналізуємо вплив митних тарифів на відношення імпорту до ВВП і поділяємо всі спостереження на три групи, залежно від середньозваженої (по всіх товарах) ставки митного тарифу (див. Табл.6.1).

Таблиця 6.1.

Група	Низька ставка митного тарифу (I група)	Середня ставка митного тарифу (II група)	Висока ставка митного тарифу (III група)
Середня залежної змінної – імпорт/ВВП, %	50	32	18

Ми маємо альтернативи (не враховуючи ще й можливості дослідження за допомогою непараметричних методів):

1. Досліджувати різницю між всіма середніми експорту/ВВП одночасно за допомогою дисперсійного аналізу.

2. Досліджувати попарно різницю між середніми з кожної групи (середня в I групі і середня в II групі; середня в II групі і середня в III групі; середня в I групі і середня в III групі) за допомогою дисперсійного аналізу або *t*-тесту.

3. Досліджувати по черзі різницю між кожною з середніх і середньою по решті груп (середня в I групі (50) і середня в II і III групах (розраховується додатково як середньозважена між 32 і 18); середня в II групі і середня в I і III групах; середня в III групі і середня в I і II групах) за допомогою дисперсійного аналізу або *t*-тесту.

На практиці найбільш ефективним є використання на початку першого способу. Якщо він показує значущу різницю між середніми, тоді за допомогою Апостеріорних критеріїв (*Post Hoc Methods*) у рамках дисперсійного аналізу визначається, саме за рахунок яких з груп виникає різниця. Використовують саме апостеріорні критерії, а не серію звичайних *t*-тестів для різниці між середніми, оскільки в останньому випадку при великій кількості порівнянь більшою мірою можливий варіант, коли невелика частка значущих різниць між середніми є насправді випадковістю. Але *t*-тест може давати кращі результати, лише якщо для різниці між середніми є теоретичне обґрунтування і ми перевіряємо наскільки практика відповідає теорії.

6.3. Метод багатofакторного одномірного дисперсійного аналізу впливу торговельних обмежень на імпорт

Складнішим варіантом дисперсійного аналізу є багатofакторний одномірний дисперсійний аналіз (*n-way univariate ANOVA*). Припустимо, ми аналізуємо вплив як митних тарифів, так і нетарифних бар'єрів на відношення імпорту до ВВП (див. табл.6.2).

Таблиця 6.2.

Середня залежної змінної по групах – імпорт/ВВП, %			
Група	Низька ставка митного тарифу (I група)	Середня ставка митного тарифу (II група)	Висока ставка митного тарифу (III група)
Низькі нетарифні бар'єри (I група)	60	37	20
Високі нетарифні бар'єри (II група)	38	25	17

Тут ми можемо досліджувати три види ефектів:

- вплив першого фактору (митні тарифи) на залежну змінну (в даному випадку високі митні тарифи зменшують імпорт відносно ВВП);
- вплив другого фактору (нетарифні бар'єри) на залежну змінну (високі нетарифні бар'єри зменшують імпорт відносно ВВП);
- взаємодію факторів, тобто яким чином ті чи інші значення одного фактору впливають на вплив іншого фактору на залежну змінну (високі нетарифні бар'єри більшою мірою зменшують імпорт/ВВП при нижчих ставках митного тарифу, а вплив митних тарифів менш сильний при високих нетарифних обмеженнях).

Перші два ефекти – це головні ефекти, а останній – це ефект взаємодії. Якщо ми проводимо групування за трьома факторами, будуть три головні ефекти і три парні ефекти взаємодії (кожний фактор з кожним фактором). Якщо по чотирьох – то чотири головні ефекти та шість парних ефектів взаємодії. Взаємодію вищих порядків складніше проілюструвати і пояснити.

Досліджуючи ефект взаємодії, потрібно дивитися, як перший фактор змінює вплив другого фактору на залежну змінну, а також як другий фактор змінює вплив першого фактору на залежну змінну.

Ефект взаємодії може бути різним:

- відсутній, коли один фактор не змінює вплив іншого фактору на залежну змінну;
- коли певні значення одного фактору зменшують чи збільшують силу впливу іншого фактору на залежну змінну, але характер впливу не міняється;
- коли певні значення одного фактору змінюють характер впливу іншого фактору на залежну змінну (позитивний стає негативним або навпаки).

З одного боку в одній таблиці є прагнення відобразити вплив якомога більшої кількості факторів. Але при кожному додатковому групуванні за додатковим фактором у комірках середні розраховуються за все меншою кількістю спостережень, внаслідок чого результати стають менш значущими і надійними. Тому чим більшу кількість факторів ми бажаємо одночасно досліджувати з урахуванням їх взаємодії, тим більшою має бути кількість спостережень.

6.4. Умови методу дисперсійного аналізу

Умови для використання одномірного дисперсійного аналізу такі:

1. Кількість спостережень у кожній групі (чи комірці у таблиці при групуванні за декількома факторами) спостережень має бути достатньою. Бажано більше 30 спостережень, хоча можливо проводити аналіз і при меншій кількості з менш надійними результатами. Бажано, щоб у кожній групі (комірці) кількість спостережень була приблизно однаковою.

2. Спостереження мають бути незалежними, тобто відсутність серійної кореляції. Якщо беруть значення за різний період часу, навіть для різних об'єктів, серійна кореляція може бути ймовірною. Інший приклад, якщо певна частина спостережень корелює одне з одним внаслідок схожих особливостей (наприклад, якщо дані одержуються за допомогою опитування, а певна одна з груп респондентів одержує заплутані інструкції, як потрібно відповідати). Вирішенням проблеми

може бути введення додаткового групування за іншим впливовим фактором або використання більш суворого критерію значущості (наприклад менше 0.01, а не 0.05).

3. В усіх групах (комірках) дисперсія має бути однаковою (гомоскедастичність). Порушення цієї вимоги – означає, що середні і дисперсія по групах корелюють між собою. Тоді F -критерій може давати неправильну оцінку значущості різниці між середніми. Наприклад, якщо дисперсія є більшою у групі з більшою середньою залежною змінною, це може означати, що така середня, ймовірно, здається більшою випадково, оскільки більша дисперсія означає менш надійну оцінку середньої для генеральної сукупності за середньою, розрахованою на основі вибірки. Більша дисперсія, зокрема, може бути наслідком наявності викидів.

Порушення цієї вимоги не є критичним, якщо групи (чи комірки) мають приблизно однакову кількість спостережень: найбільша від найменшої відрізняється не більше ніж у 1.5 рази. Існують тести для перевірки різниці у дисперсії. Якщо гетероскедастичність (різна дисперсія) є проблемою, потрібно:

- трансформувати залежну змінну;
- використовувати більш суворий критерій значущості ніж 0.05;
- перевірити на наявність викидів за залежною змінною, провести повторно дисперсійний аналіз без викидів;
- скористатися замість дисперсійного аналізу його непараметричним аналогом.

4. Залежна змінна повинна підпорядковуватися нормальному розподілу всередині кожної групи (комірки). Порушення цієї передумови не є критичним, якщо вибірка є великою. Якщо ексцес більше нуля, рівень значущості p , який дає F -критерій завищений, тобто дисперсійний аналіз вказує на відсутність закономірності, коли вона є. Якщо менше нуля, то рівень значущості буде заниженим. Асиметрія розподілу мало впливає на надійність результатів дисперсійного аналізу.

5. Додатково розглядають лінійність зв'язку між факторами і залежною змінною та відсутність мультиколінеарності.

6.6. Багатомірний дисперсійний аналіз (MANOVA)

У багатомірному дисперсійному аналізі (MANOVA – *multivariate analysis of variance*) досліджується вплив фактора або факторів не по черзі, а одночасно на декілька залежних змінних. Наприклад, вплив обмежень руху капіталу одночасно на прямі та портфельні інвестиції в країну. У MANOVA замість звичайного F -тесту використовується такі критерії:

- Багатомірний F -критерій (лямбда-критерій Уїлкса – Wilks).
- Roy's gcr - Greatest Characteristic Root;
- Hotelling's trace;
- Pillai's criterion – має переваги, якщо вибірка невелика, кількість спостережень у комірках є достатньо різною або є гетероскедастичність.

Якщо вони показують значущість впливу факторів на одночасно декілька залежних змінних, далі розглядається яким чином фактори впливають на кожну залежну змінну окремо за допомогою звичайного одномірного F -критерію з можливими поправками (Bonferroni inequality або stepdown analysis).

Проведення MANOVA має перевагу перед проведенням серії одномірних ANOVA для кожної залежної змінної у тому, що забезпечує від прийняття випадкової різниці у середніх за закономірність.

Наведемо ілюстративний приклад. Припустимо, ми аналізуємо 40 залежних змінних. Якщо серія одномірних ANOVA показує, що фактори впливають на 20 із рівнем значущості 0.05, цілком імовірно, що насправді одна із залежних змінних не залежить від досліджуваних факторів.

Також MANOVA може виявити, що існує певна лінійна комбінація залежних змінних, середні якої суттєво відрізняються по групах, навіть якщо середні індивідуальних залежних змінних мало відрізняються по групах. Тобто MANOVA може виявити комбіновану відмінність між групами. Серія одномірних ANOVA ігнорує наявність кореляції між залежними змінними.

Цілі використання MANOVA можуть бути такими:

1. Дослідження різнорідних залежних змінних (наприклад, валютний курс, приріст ВВП, міграція, інфляція тощо), коли потрібно знизити ризик сприйняття випадковостей за закономірності. У разі виявлення впливу на всі змінні одночасно, досліджується вплив на кожну залежну змінну окремо.

2. Дослідження тієї самої залежної змінної у різні періоди часу, але для тих самих об'єктів, тобто в умовах повторних спостережень. Різні залежні змінні в такому випадку є фактично однією змінною, вимірюваною в різні періоди часу. Наприклад, якщо перша залежна змінна – це зростання імпорту у 2009 р., друга – зростання імпорту у 2010 р., третя – зростання імпорту у 2011 р. по тих самих країнах. Або перша змінна – зростання експорту перед кризою, друга – під час кризи, третя – після кризи. Фактором може бути, наприклад, дефіцит державного бюджету.

3. Дослідження однорідних залежних змінних, коли цікавить саме одночасний вплив факторів на усі змінні, а не на кожен окрему. Наприклад, у випадку дослідження валютних криз, такими однорідними змінними можуть бути знецінення національної валюти, зменшення чистих валютних резервів, зростання процентної ставки. Ці всі змінні є основними індикаторами наявності валютної кризи, і нас може цікавити одночасний вплив факторів на всі ці залежні змінні в сукупності. Особливо корисним *MANOVA* може бути, коли значення кожної залежної змінної неточні, а використання комбінації залежних змінних нівелює ці неточності. Зокрема, коли індивідуальні залежні змінні є результатом суб'єктивних оцінок, наприклад, як такі дві змінні: ступінь задоволення іноземних туристів перебуванням у країні та їх бажання приїхати ще раз у країну протягом наступних декількох років.

Особливості передумов для *MANOVA*. Більшість з них співпадає з передумовами для одномірного *ANOVA*, але є й відмінності.

1. *MANOVA* потребує більші вибірки ніж одномірний *ANOVA*.

2. Кореляція для змінних у кожній групі (комірці) має бути однаковою.

3. Залежні змінні повинні підпорядковуватися багатомірному нормальному розподілу, тобто будь-яка лінійна комбінація залежних змінних має підпорядковуватися нормальному розподілу. Оскільки нормальність багатомірного розподілу важко перевірити, на практиці, тестують на нормальність розподілу кожен залежну змінну окремо.

6.8. Дисперсійний аналіз впливу рівня економічного розвитку на приплив інвестицій у Microsoft Excel

Припустимо, у нас є дані про ПІІ в країні з різним доходом на душу населення (низьким, високим, але також виділяємо окремо і з середнім) – див. Рис.6.50. Не обов'язково, щоб групи мали однакову кількість спостережень.

У надбудові *Пакет аналізу* оберемо *Однофакторный дисперсионный анализ*. Обираємо *Группирование по столбцам* (оскільки у наших вхідних даних – один стовпчик – одна група) та інші опції, як на Рис.6.51.

Н	І	Ж
FDIlowincome	FDIhighincome	FDImidincome
3	5	6
5	9	11
2	4	3
3	2	6
2	3	1
5	11	10
1	6	8
1	3	4
4	4	5
4	5	3
3	2	6
4	2	8
0	1	6
1	1	4
2	2	3
1	4	4
2	3	1
2	2	4
1	0	5
0	1	3

Рис.9.50.

У робочому аркуші з результатами (див. Рис.6.52) ми бачимо, що найбільші ППП – у країнах с середнім доходом на душу населення (5.05% ВВП), далі у країнах з високим доходом (3.5 % ВВП), далі з низьким доходом (2.3 % ВВП). Рівень значущості F-критерію становить 0.00230427, тобто менше 0.05. Це означає, що різниці є значущими.

У надбудові *Пакет анализа* можливо використати *Двухвыборочный F-тест для дисперсии* для попарного порівняння дисперсії (в цій програмі ми не можемо порівняти дисперсію у всіх групах одночасно). Він показує, що вона значуще відрізняється у групах 1 і 2 та групах 1 і 3. Але це порушення гомоскедастичності не є критичним, враховуючи однакову кількість спостережень у групах.

У надбудові *Пакет анализа* можливо використати опцію *Гистограмма* для візуальної перевірки значень у кожній групі на відповідність нормальному розподілу.

Оскільки апостеріорні критерії у Microsoft Excel не доступні, далі можливо попарно порівняти середні у кожній групі, за допомогою t-тесту, як ми робили це раніше, або за допомогою дисперсійного аналізу кожний раз для двох груп. Виявиться, що значуще відрізняються середні у 1 і 3 групах (країнах з низьким доходом і середнім доходом) з рівнем значущості 0.000278. Значущість відмінності середніх у групах 1 і 2 та 2 і 3 є меншою (0.094 та 0.076).

	A	B	C	D	E	F	G
1	Однофакторный дисперсионный анализ						
2							
3	ИТОГИ						
4	<i>Группы</i>	<i>Счет</i>	<i>Сумма</i>	<i>Среднее</i>	<i>Дисперсия</i>		
5	FDlowincome	20	46	2.3	2.326315789		
6	FDhighincome	20	70	3.5	7.421052632		
7	FDmidincome	20	101	5.05	7.102631579		
8							
9							
10	Дисперсионный анализ						
11	<i>Источник вариации</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-Значение</i>	<i>F критическое</i>
12	Между группами	76.033333	2	38.01667	6.768545994	0.00230427	3.158842719
13	Внутри групп	320.15	57	5.616667			
14							
15	Итого	396.18333	59				

Рис.6.52.

Microsoft Excel також дозволяє проводити дисперсійний аналіз за допомогою опцій *Двухфакторный дисперсионный анализ с повторениями* та *Двухфакторный дисперсионный анализ без повторений* у надбудові *Пакет анализа*.

6.9. Непараметричні методи дисперсійного аналізу

Непараметричними аналогами однофакторного дисперсійного аналізу є:

- Тест Краскела-Уолліса (*Kruskal-Wallis test*);
- Медіанний тест (*Median test*).

Для змінних, які вимірюються у номінальній шкалі, використовуються описані раніше методи аналізу частот.

ГЛАВА 7. Дослідження впливу міжнародних економічних відносин на економічне зростання методами лінійного регресійного аналізу

7.1. Основи методу лінійного регресійного аналізу

Стан міжнародних економічних відносин характеризується великою кількістю економічних показників, які в свою чергу, залежать від різноманітних факторів. Як було показано в главі 2 кореляційний аналіз досліджує тісноту взаємозв'язків між змінними. Регресійний аналіз дозволяє виявити характер цих взаємозв'язків. В регресійному аналізі моделюється взаємозв'язок однієї залежної економічної змінної (показника, результуючої ознаки або відгуку) від однієї або декількох незалежних економічних змінних (факторів). Вибір або призначення залежної і незалежних змінних є довільним і здійснюється дослідником в залежності від поставленої задачі.

В загальному випадку регресійну модель можна записати у вигляді:

$$Y = f(x_1, x_2, \dots, x_n) \quad (7.1)$$

де Y – залежна змінна (відгук); $x_i, i = 1, 2, \dots, n$ – незалежні змінні (фактори).

За допомогою регресійного аналізу можна вирішувати важливі задачі для економічної проблеми, яка досліджується. Серед них можна виділити такі задачі:

1. Зменшення розмірності простору змінних, які аналізуються (факторного простору), за рахунок заміни частини факторів однієї змінної – відгуком. Більш повно ця задача розв'язується факторним аналізом.

2. Кількісна зміна ефекту кожного фактору, тобто множинна регресія дозволяє досліднику задати питання і отримати відповідь про те, «що є кращим фактором для...». При цьому стає більш зрозумілим дія окремих факторів на відгук, і дослідник краще розуміє природу економічного явища, яке вивчається.

3. Виявлення прогностичних значень відгуку при певних значеннях факторів, тобто регресійний аналіз, створює базу для обчислення з метою отримання відповідей на питання типу «що буде, якщо...?».

4. В регресійному аналізі в більш явній формі виступають причинно-наслідковий механізм. Прогноз при цьому краще піддається змістовній інтерпретації.

Регресійний аналіз складається з наступних основних етапів:

- з множини незалежних змінних вибір факторів, які суттєво впливають на залежну змінну – відгук;
- визначення форми рівняння регресії;
- оцінювання параметрів регресійної моделі.

В загальному випадку класичний регресійний аналіз є параметричним методом, тобто припускає, що змінні в моделі кількісні, закон розподілу яких відповідає нормальному закону.

В дослідження у сфері міжнародних економічних відносин частіше використовують лінійні регресійні моделі. Існує декілька причин широкого застосування лінійних моделей:

- припущення про лінійність є простим, тому легко сприймається і інтерпретується;
- метод оцінки параметрів найменш трудомісткий і мена проводити розрахунки «в ручну» за допомогою калькуляторів;
- багато математичних методів дослідження пристосовані до лінійних моделей.

Лінійна модель множинної регресії має вигляд:

$$Y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n + \varepsilon, \quad (7.2)$$

де Y – залежна змінна (відгук);

x_i – i -та незалежна змінна (i -й фактор) $i = 1, 2, \dots, n$;

b_i – i -й коефіцієнт регресії (коефіцієнт при i -му факторі),

b_0 – вільний член – константа (*intercept*);

ε – похибка спостереження (залишок), яка є різницею між спостережувальним значенням залежної змінної та розрахованим за регресією (теоретичним або передбаченим) значенням залежної змінної. В ідеальному випадку залишки є випадковими величинами, які мають нормальний розподіл з математичним сподіванням, яке дорівнює нулю.

Спочатку за відомими значеннями залежної та незалежних змінних розраховуються найкращі коефіцієнти регресії (методом найменших квадратів або іншими методами), за яких функція регресії стає максимально можливо достовірною. Далі регресія з відомими коефіцієнтами регресії, підставляючи нові значення незалежних факторів, використовується для розрахунку невідомих значень залежної змінної.

Є дві основні мети регресійного аналізу:

- прогноз.
- пояснення зв'язку між змінними.

Вибір найбільш підходящої регресійної моделі здійснюється через критерії – коефіцієнти детермінації, множинної кореляції, статистики Дарбіна-Уотсона, розподілу залишків та інше. Розраховані значення залежної змінної можуть не співпадати з фактичними її значеннями, але регресія тим більше буде точною, чим меншими будуть залишки.

Введемо позначення y_i - вихідні дані залежної змінної Y (відгука), \hat{y}_i - передбачувальні значення, тобто значення розраховані за регресійною залежністю, $i = 1, 2, \dots, n$; n - обсяг вибірки, або кількість спостережень.

Коефіцієнт детермінації обчислюється за формулою:

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (7.3)$$

де $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ - середнє значення залежної змінної Y .

Коефіцієнт детермінації характеризує відношення розсіювання значень залишків навколо лінії регресії по відношенню до загального розсіювання значень, тим, очевидно, краще прогноз.

Коефіцієнт детермінації R^2 , який приймає значення від 0 до 1, може свідчити про величину залишків. Чим ближче він до 1 (за умов, що виконуються умови для регресійного аналізу), тим тісніший лінійний кореляційний зв'язок між залежною змінною та факторами. Коефіцієнт детермінації показує, яку частину дисперсії залежної змінної можливо пояснити за допомогою дисперсії включених до регресії факторів. Значущість коефіцієнту детермінації перевіряється за допомогою F -тесту.

Коефіцієнт детермінації R^2 оцінює частку варіації даних навколо середнього значення, яка може бути пояснена за допомогою вибраного рівняння регресії. Якщо, наприклад, $R^2 = 0.4$, то 40% можуть бути пояснені рівнянням регресії, 60% залишкової мінливості залишаються не поясненими. В ідеалі бажано мати пояснення не всієї, а хоча б більшої частини вихідної мінливості. Значення R^2 є індикатором ступені підгонки до даних. Значення $R^2 \approx 1$, показує, що модель пояснює майже всю мінливість відповідних змінних.

Коефіцієнт множинної кореляції $R = \sqrt{R^2}$ характеризує тісноту зв'язку між змінними, а також є оцінкою якості передбачення. Якщо коефіцієнт множинної кореляції, як і коефіцієнт парної кореляції, характеризують ступінь лінійного взаємозв'язку. Коефіцієнт детермінації R^2 оцінює адекватність регресійної моделі будь-якого виду. Тому коефіцієнт детермінації використовують також як оцінку ступеня нелінійної залежності між змінними.

Для порівняння якості регресійних рівнянь з різною кількістю факторів або розрахованих за різною кількістю спостережень варто використовувати скоригований коефіцієнт детермінації (*adjusted coefficient of determination*). Тоді можливо уникнути ситуації, коли для збільшення коефіцієнту детермінації, дослідник включатиме до регресії якомога більшої кількості факторів, незважаючи на те, що вони можуть здійснювати мінімальний вплив на залежну змінну. Також регресія, розрахована за меншою кількістю спостережень, може мати більший звичайний коефіцієнт детермінації ніж регресія, розрахована за більшою кількістю спостережень.

Знак коефіцієнтів $b_i, i = 1, 2, \dots, n$ вказує на характер впливу i -го фактора на залежну змінну. Якщо $b_i > 0$ додатне число, фактор впливає позитивно (більше значення фактору пов'язане з більшим значенням залежної змінної), якщо від'ємне – негативно (більше значення фактору пов'язане з меншим значенням залежної змінної), якщо дорівнює нулю – фактор не впливає на залежну змінну. Наскільки значно кожен з коефіцієнтів b_i відрізняється від нуля показує t -тест, який проводиться для кожного такого коефіцієнту. Тобто якщо F -тест вказує на значущість зв'язку між залежною змінною та сукупністю всіх факторів, включених до регресії, то t -тест вказує на значущість зв'язку між залежною змінною та окремим фактором у рамках побудованої регресії.

Аналогами звичайних b -коефіцієнтів є β -коефіцієнти (бета-коефіцієнти), якщо в регресію замість абсолютних значень змінних підставляти стандартизовані значення змінних. Тому β -коефіцієнти не залежать від того, в яких одиницях вимірюються змінні. β -коефіцієнти для різних показників у різних одиницях виміру є порівнюваними. Той фактор, чий β -коефіцієнт більший, здійснює сильніший вплив.

Якісна регресійна модель повинна давати достатньо точні розраховані значення залежної змінної на основі даних по факторах за межами сукупності, на основі якої будувалася регресія. Тому варто перевірити, як працює регресійна модель на іншій вибірці. Часто це можливо зробити лише з часом, коли будуть одержані дані за новий період.

Тому можливо первинну вибірку поділити на дві чи більше частин і розрахувати дві або більше регресії на основі цих підвбірок. Далі перевіряються наскільки співпадають чи відрізняються коефіцієнти регресій при факторах (величина, знак, рівень значущості) та коефіцієнти детермінації регресій (величина, рівень значущості).

Статистика Дарбіна-Уотсона, оцінюючи наявність серійної кореляції між залишками сусідніх спостережень, характеризує ступінь залежності між спостереженнями. Регресійні моделі будуються в припущенні, що вибірка випадкова. Тобто спостереження незалежні. В протилежному випадку коефіцієнти побудованого рівняння регресії можуть бути не стійкими до зміни вихідних даних, що в результаті впливає на достовірність результатів аналізу. Чим менша статистика Дарбіна-Уотсона і модуль серійної кореляції, тим менш залежні спостереження у вибірці, а значить і більш адекватна побудована регресійна модель.

Якщо побудована адекватна регресійна модель, то залишки між вихідними і прогнозними значеннями, обчислені за рівнянням регресії, повинні бути випадковими величинами з математичним сподівання, рівним нулю (білий шум). Тому за ступенем відхилення розподілу залишків від нормального закону розподілу також можна судити про якість побудованої моделі.

Обираючи змінні для аналізу, важливо включити до регресійного рівняння всі важливі фактори (але уникаючи при цьому проблеми мультиколінеарності) і утримуватися від включення факторів, які мало впливають на залежну змінну. Для попереднього визначення важливості фактору можливо використати кореляційний аналіз та графічний аналіз за допомогою діаграм розсіювання (або точкових діаграм – *scatterplots*). Бажано, щоб включалися незалежні змінні, зв'язок яких із залежною змінною має існуюче або потенційне теоретичне пояснення. Якщо регресійний аналіз проводиться з метою прогнозу, варто включати лише ті змінні, за якими можливо буде знайти у майбутньому дані.

Визначення, які саме змінні варто включати в регресійну модель, а які ні, можливо автоматизувати. Для цього використовуються покрокові процедури із включенням або виключенням факторів. Але цілковито покладатися на них не варто. Може виявитися так, що обрана за допомогою цих процедур форма регресії виявиться неоптимальною з практичної точки зору.

Покроковий метод включення змінних передбачає, що спочатку будується однофакторна регресія з тим фактором, який має найбільший кореляційний зв'язок із залежною змінною. Далі будується двофакторна регресія: додається з числа інших факторів той фактор, який пояснює найбільшу частину дисперсії, яка не пояснюється дією першого фактору (за допомогою часткового коефіцієнту кореляції – *partial correlation coefficient*). Далі будується трифакторна регресія: додається з числа решти факторів той фактор, який пояснює найбільшу частину дисперсії, яка не пояснюється дією першого та другого фактору. Аналогічно далі додаються четвертий фактор і т.д. Процес зупиняється, коли додаткова користь від включення нової змінної стає мінімальною (коли частковий F -критерій свідчить про незначущість внеску нового фактору). Як варіант, при вклю-

ченні нових факторів можуть перевірятися вже включені в регресію фактори на предмет необхідності їх залишення чи виключення.

Покроковий метод виключення змінних передбачає побудову регресії з усіма можливими факторами. Потім фактори з найменшим внеском виключаються.

Альтернативною покроковим процедурам є побудова всіх можливих регресій. Наприклад, для 10 факторів можливо побудувати 1024 варіантів регресій з повним або частковим включенням факторів.

Більша кількість факторів потребує більшої кількості спостережень. Бажано, щоб при побудові однофакторної регресії було більше 20-30 спостережень. Умовна рекомендація для багатфакторної регресії передбачає наявність більше 10-20 (в крайньому випадку мінімум 5) спостережень на кожний фактор, включений у модель. Наприклад, для п'ятифакторної моделі бажано мати більше 50-100 спостережень.

Якщо кількість спостережень є дуже великою, наприклад, 1000, то статистичні тести (F, t) будуть показувати, що зв'язок між факторами та залежною змінною буде значущий, навіть якщо він є дуже слабким. Тому при великих вибірках потрібно забезпечити не лише статистичну, а й практичну значущість результатів. Для визначення практичної значущості варто звернути увагу не тільки на значущість, а й на саму величину коефіцієнту детермінації, парних коефіцієнтів кореляції та β -коефіцієнти.

Тому, навіть якщо F -тест показує значущість коефіцієнту детермінації, а сам коефіцієнт детермінації невеликий (наприклад, 0.10), то регресія є недостатньо адекватною для прогнозування чи визначення причин змін залежної змінної. Тоді варто змінити специфікацію моделі – трансформувати її: спробувати додати інші фактори або використати нелінійну регресію.

Навіть якщо t -тест вказує на значущість b -коефіцієнта для відповідного фактора, варто пересвідчитися чи не є замалим парний коефіцієнт кореляції або β -коефіцієнт для цього фактора. Також якщо парний коефіцієнт кореляції та b -коефіцієнт мають різні знаки, то регресія швидше за все не є адекватною: вона неправильно визначає характер впливу цього фактору. Тоді варто перерахувати регресію без цього фактору або інших факторів, які з ним корелюють.

Як й інші види аналізу лінійний регресійний аналіз виходить з певних припущень. Для впевненості, що регресія адекватно відображає зв'язки між змінними і може використовуватися для прогнозу або пояснення впливу, фактично потрібно перевірити на відповідність припущенням не тільки кожен змінну окремо, а й всю лінійну комбінацію факторів (якщо йдеться про багатфакторну регресію).

Тестування на відповідність припущенням відбувається, як перед проведенням регресійного аналізу, так і в процесі його проведення з можливим коригуванням специфікації рівняння регресії або трансформацією змінних.

Умови для регресійного аналізу такі:

1. Лінійний характер зв'язку.
2. Гомоскедастичність (незалежність дисперсії від значень змінних).
3. Незалежність залишків (відсутність серійної кореляції).
4. Нормальний розподіл залишків.
5. Відсутність мультиколінеарності.

Відповідність припущенням часто перевіряється, зокрема, за допомогою діаграм розсіювання, де на вертикальній осі відкладаються залишки, а на горизонтальній – розраховані значення залежної змінної y , кожної із залежних змінних $x_i, i = 1, 2, \dots, n$, або часу - t . Іноді для кращої порівнюваності використовують не абсолютні значення залишків, а їх стандартизовані значення або трансформовані відповідно до t -розподілу Стьюдента (*studentized residual*). Наведемо приклади відповідних діаграм розсіювання на Рис. 7.1.

Лінійність варто перевіряти за допомогою діаграм розсіювання, де на горизонтальній осі відкладені значення одного з факторів (*partial regression plots*), а не залежної змінної. І таких діаграм розсіювання варто побудувати стільки, скільки факторів є у регресії. Це пов'язане із тим, що нелінійність зв'язку із залежною змінною може бути у частини факторів, а не у всіх.

Лінійність зв'язку також можливо перевірити за допомогою звичайної діаграми розсіювання (а не залишків), де по вертикальній осі відкладаються значення залежної змінної (y), а по горизонтальній – одного з факторів (x_i).

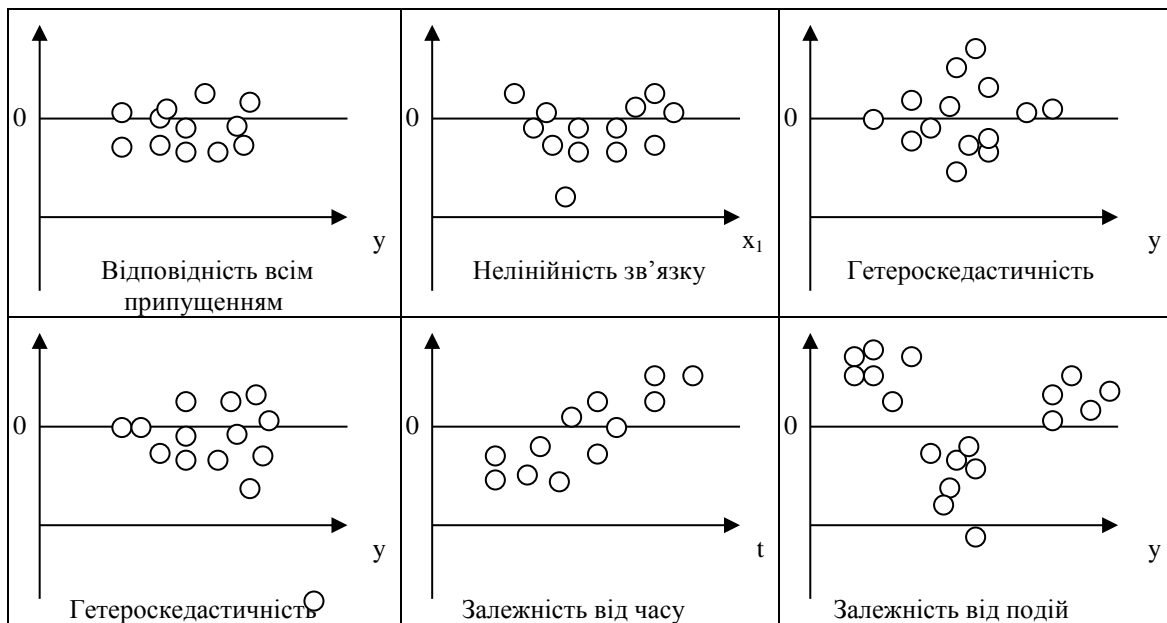


Рис.7.1.

Якщо виявлений нелінійний зв'язок залежної змінної із певними факторами, ці фактори варто трансформувати. Наприклад, замість фактору x_i використати $\ln(x_i)$, $1/x_i$, $\sqrt{x_i}$, або x_i^2 . За необхідності те саме можливо зробити й із залежною змінною. Можливо також спробувати заміну абсолютних значень змінних на їх прирости, наприклад: y_t на $(y_t - y_{t-1})/y_t$, а x_{it} на $(x_{it} - x_{it-1})/x_{it}$. Або на різницю логарифмів: $\ln(y_t)$ на $\ln(y_t) - \ln(y_{t-1})$, а $\ln(x_{it})$ на $\ln(x_{it}) - \ln(x_{it-1})$.

Гетероскедастичність передбачає залежність залишків від значення залежної змінної, що пов'язане з залежністю дисперсії залежної змінної від її значень. Наприклад, чим більші значення залежної змінної, тим більша її дисперсія, і відповідно залишки при побудові регресії. На Рис. 10.1 представлена трикутників та ромбовидна форма гетероскедастичності. Для тестування гетероскедастичності також використовується тест Левена.

Для вирішення проблеми гетероскедастичності можливо замість звичайного методу найменших квадратів використати метод зважених найменших квадратів (*weighted least squares*) або спробувати трансформувати залежну змінну. Наприклад, $\ln(y_i)$, $1/y_i$, $\sqrt{y_i}$.

Часто якщо існує порушення однієї умови, порушуються й інші. І вирішення проблем з однією умовою часто вирішує проблеми з іншими. Наприклад, вирішення проблем з нелінійністю зв'язку, може призвести до того, що гетероскедастичність зникає. Або вирішення проблеми з відхиленням від нормального розподілу змінних може вирішити проблему гетероскедастичності.

Незалежність залишків передбачає, що розраховані значення залежної змінної не залежать від інших розрахованих значень, а відповідно залишки не залежать один від одного. Порушення цієї умови передбачає, що є закономірність у тому, що певні залишки є додатними, а решта від'ємними. На Рис. 10.1 представлені:

- залежність залишків від часу – час часто використовується як змінна, яка упорядковує спостереження;
- залежність від подій – наприклад, коли не враховано фактор сезонності у моделі (зимові та літні місяці), якщо він є насправді важливим (наприклад, якщо розглядається зовнішня торгівля окремими видами одягу), або коли спостереження взяті щодо зовсім різних груп об'єктів (наприклад, країни з низькою часткою неформального сектору економіки та країни з високою його часткою).

Вирішенням проблеми залежності залишків можуть бути:

- заміна абсолютних значень на прирости типу: y_t на $(y_t - y_{t-1})/y_t$, а x_{it} на $(x_{it} - x_{it-1})/x_{it}$;

- урахування факторів умов (наприклад, сезону);
- спеціально сконструйовані регресійні моделі, адаптовані до залежності залишків.

Залишки повинні мати нормальний розподіл. За недостатньої кількості спостережень відхилення від нормального розподілу змінних можуть призводити до відхилень від нормального розподілу залишків. Нормальність розподілу перевіряється за допомогою графіку нормального розподілу (*normal probability plot*).

Мультиколінеарність передбачає кореляцію факторів між собою. Включення сильнокорельюючих факторів у регресійну модель є недоцільним. Уявімо екстремальний випадок: якщо включити в модель такі фактори: ціну нафти за тону і ціну нафти за барель. Кореляція між цими факторами буде дорівнювати 1. Але включити обидва фактори недоцільно. Ціна нафти за барель не несе ніякої додаткової варіації (а відповідно й інформації) порівняно з ціною нафти за тону. Аналогічно приріст ВВП і приріст ВНД будуть сильно корелювати, і включення цих обох факторів у модель недоцільне: другий фактор дає лише мінімум інформації порівняно з першим. У цьому плані приріст ВВП і приріст ВНД практично одним фактором і нам достатньо включити в модель лише одну змінну, яка його представляє.

Мультиколінеарність є суттєвою проблемою, оскільки регресійні коефіцієнти при сильнокорельюючих факторах можуть бути розраховані неправильно, і навіть мати неправильний знак (замість додатного від'ємний або навпаки, тобто спотворюватимуть висновки про характер впливу фактору, визначений за коефіцієнтом кореляції його із залежною змінною).

Найпростішим способом визначення мультиколінеарності є вивчення кореляційної матриці факторів. Не можна включати до моделі одночасно фактори, кореляція між якими більше 0.8-0.9. Але відсутність високих парних кореляцій не виключає мультиколінеарність. Наприклад, один із факторів може сильно корелювати з лінійною комбінацією декількох інших факторів, хоча з кожним окремим фактором кореляція буде помірною.

Іншими методами є толерантність (*tolerance value*) і фактор росту дисперсії (*variance inflation factor (VIF)*). Ці два показники є оберненими величинами одна одної. Вони розраховуються для кожного фактору по черзі окремо. Для цього будується регресія, де відповідний фактор розглядається як залежна змінна від решти факторів. Толерантність є частиною дисперсії такого залежного фактору, яка не пояснюється дисперсією решти факторів. Толерантність нижче 0.1 та VIF більше 10 означають наявність сильної кореляції між відповідним залежним фактором і рештою факторів.

В модель варто включати ті фактори, які сильно корелюють із залежною змінною і слабо корелюють між собою. Як виняток, регресію з мультиколінеарністю можливо використовувати, але лише з метою прогнозу, а не оцінки величини і характеру впливу факторів на залежну змінну. Також вирішенням проблеми мультиколінеарності є використання гребеневої регресії (*ridge regression*) або регресії на баз основних компонент (*regression on principal components*).

Як бачимо лінійний регресійний аналіз фактично має ті самі передумови і обмеження, що й кореляційний аналіз. Так як ці два методи взаємопов'язані, часто їх розглядають як один: регресійно-кореляційний аналіз. Наприклад, так само як і кореляційний аналіз регресійний аналіз не дає відповідь, що є причиною, а що наслідком. Чи зміни у залежній змінній є наслідком змін у факторах, чи навпаки. Відповідь на це можливо одержати, якщо включити в модель часові лаги. Наприклад, припустимо, як незалежна змінна розглядається приплив інвестицій у поточному році (t), а як залежна – приріст ВВП у наступному році ($t + 1$). Тоді, якщо параметри якості (коефіцієнт детермінації, F -тест, t -тест щодо значущості регресійного коефіцієнту для приросту ВВП) регресії покажуть значущий зв'язок, ми можемо сказати, що саме приплив інвестицій впливає на приріст ВВП. З іншого боку, навіть лаги не гарантують встановлення на пряму зв'язку, внаслідок того, що

- зменшення припливу інвестицій може також бути спричинено очікуваннями зменшення приросту ВВП;
- дії спільної причини (наприклад, економічний спад за кордоном може призвести як до зменшення припливу інвестицій так і зменшення приросту ВВП через інший канал (наприклад зменшення експорту).

Також коли розраховується значення залежної змінної, варто враховувати довірчі інтервали для прогнозного значення. Вони можуть бути достатньо широкими та не давати можливість зробити точний прогноз.

Як правило, регресійна модель дозволяє достатньо точно розрахувати значення залежної

змінної на основі типових для первинної вибірки (на основі якої розраховувалися коефіцієнти детермінації) значень факторів. Наприклад, припустимо, що регресійна модель залежності приросту імпорту від приросту світового ВВП та зміни валютного курсу розраховувалася за вибіркою, де приріст світового ВВП варіювався від 1 до 6%, а зростання валютного курсу від -10 до 10%. Якщо нам потрібно розрахувати приріст імпорту на основі зовсім інших нових значень факторів (скажімо, приріст світового ВВП -3%, а зростання валютного курсу -40%), розрахований приріст імпорту швидше за все буде неточним.

При використанні регресійної моделі потрібно пересвідчитися, що умови не змінилися, порівняно з тими, які були характерні для спостережень, на основі яких розраховувалася регресія. Наприклад, навряд чи можливо буде використати регресію, розраховану лише на умови адміністративно-командної економіки, в умовах ринкової економіки, або розраховану лише на умови стабільного економічного зростання – в умовах рецесії, розраховану лише для високорозвинутих країн – для країн з низьким рівнем розвитку.

Впливові спостереження (викиди зокрема) можуть створювати суттєві проблеми в регресійному аналізі. Існує декілька варіантів, як впливові спостереження можуть впливати на результати регресійного аналізу (див. Рис.7.2).

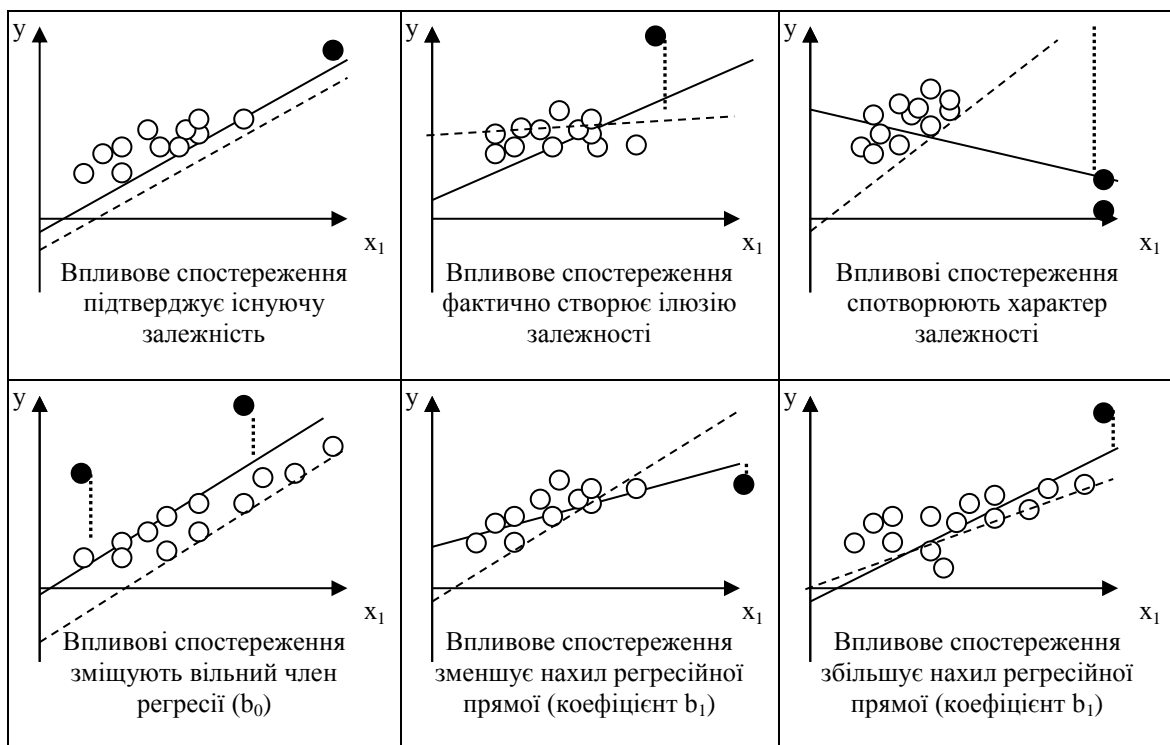


Рис.7.2.

У випадку регресійного аналізу викиди легше за все визначити за величиною залишків (графічно величина залишку може бути представлена як розмір перпендикуляру від спостереження до регресійної прямої). Хоча цей метод не дає повної картини про впливові спостереження. Впливові спостереження також можливо визначити графічно на вказаних вище діаграмах розсіювання.

Ті спостереження, для яких залишки є великими (більше 2-3 стандартних відхилень) можливо вважати викидами. За наявності викидів варто побудувати дві регресії: з викидами та без них і так само порівняти коефіцієнти при факторах та коефіцієнт детермінації.

7.2. Регресійний аналіз регіональних особливостей формування процентної ставки за зовнішніми кредитами з використанням бінарних змінних

У лінійній та більшості нелінійних регресійних моделях залежна змінна є кількісною змінною. Якщо залежна змінна є якісною, використовуються спеціальні види нелінійних регресійних моделей (логіт-регресія або пробіт-регресія).

Зазвичай для лінійної регресії використовуються лише кількісні фактори (які вимірюються у метричній, рідше – у порядковій шкалі). Для включення якісних факторів використовуються бінарні змінні (псевдозмінні, фіктивні змінні – *dummy variables*). Бінарна змінна приймає значення 1 або 0 (або рідше 1 і -1). Наприклад, 1 – є зона вільної торгівлі, 0 – нема зони вільної торгівлі.

Якщо категоріальна змінна, яка вимірюється у номінальній шкалі, може приймати декілька значень k , то для її включення в регресію потрібно $k - 1$ бінарних змінних. Наприклад, створимо для категоріальної змінної *Region*, яка може приймати 4 значення 3 бінарні змінні, які можливо включити в регресійну модель (див. Табл.7.1).

Таблиця 7.1.

Значення категоріальної змінної Region	Значення бінарної змінної Європа (D_1)	Значення бінарної змінної Азія, Австралія і Океанія (D_2)	Значення бінарної змінної Америка (D_3)
Європа	1	0	0
Азія, Австралія і Океанія	0	1	0
Америка	0	0	1
Африка	0	0	0

Наприклад, якщо йдеться про Америку, то всі бінарні змінні, крім Америка матимуть значення 0. У випадку Африки всі бінарні змінні прийматимуть значення 0. Наведемо приклад інтерпретації результатів. Припустимо, розрахована за умовними даними регресія має вигляд:

$$i = 5 + 1.5\pi - 3D_1 - 2D_2 + 1.5D_3, \quad (7.4)$$

де i – процентна ставка (по зовнішніх кредитах) у процентах;

π – інфляція у процентах.

Це означає, що:

- в Африці за нульової інфляції (коли всі змінні приймають значення 0) процентна ставка дорівнюватиме 5%;
- збільшення інфляції на кожний 1% (а точніше 1 п.п.) підвищує процентну ставку на 1.5% (1.5 п.п.);
- приналежність до регіону Європа (порівняно з тим, якщо країна знаходилася в Африці) зменшує процентну ставку на 3%;
- приналежність до регіону Азія, Австралія і Океанія (порівняно з Африкою) зменшує процентну ставку на 2%;
- приналежність до регіону Америка (порівняно з Африкою) збільшує процентну ставку на 1.5%.

Якщо, наприклад, потрібно порівняти наскільки збільшує процентну ставку приналежність до Америки порівняно з Європою, то дістанемо $4.5 = 1.5 - (-3)$.

Звичайно, як і у випадку звичайних змінних, бінарні змінні слід включати до регресії, якщо t -тест показує значущість коефіцієнту при відповідній бінарній змінній.

Альтернативний варіант урахування якісних змінних – побудувати декілька регресійних моделей. У нашому прикладі тоді потрібно побудувати 4 регресії (за спостереженнями для кожного з 4 регіонів). У такому випадку може статися й так, що b -коефіцієнт при змінній π буде різний у різних 4 моделях (тобто буде виявлено ефект взаємодії між факторами регіон та інфляція), що також може представляти інтерес. Більш доцільно будувати окремі регресії після побудови регресії з бінарними змінними, пересвідчившись, що відповідна бінарна змінна значуще впливає на залежну змінну.

7.3. Аналіз взаємодії факторів приросту імпорту

Для урахування ефекту взаємодії (*interaction effect* або *moderator effect*) між факторами модель може бути трансформована таким чином (наприклад, у випадку двофакторної регресії):

$$Y = b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2, \quad (7.5)$$

Так само потрібно перевіряти за допомогою t-тесту значущість коефіцієнту b_3 . Якщо він значущий, то ефект взаємодії існує, тобто вплив одного з факторів залежить від того, яке значення приймає інший фактор. Інший варіант визначення значущості ефекту взаємодії полягає у побудові спочатку регресії без x_1x_2 , а потім з x_1x_2 . Якщо коефіцієнти детермінації цих двох регресій значуще відрізняються, то ефект взаємодії значущий.

Наведемо приклад за умовними даними. Припустимо, розрахована регресія має вигляд:

$$IMP = 2 + 4GDP + 2E + 0.5GDP \cdot E \quad (7.6)$$

де IMP – приріст імпорту, %;

GDP – приріст ВВП, %;

E – зростання (реального ефективного) валютного курсу, %.

Сумарний вплив приросту ВВП на приріст імпорту залежить від зростання валютного курсу: $4 + 0.5 \cdot E$. Наприклад, якщо валютний курс зростає на 5%, приріст ВВП на кожний додатковий 1 п.п. призводитиме до зростання імпорту на $6.5\% = 4 + 0.5 \cdot 5$. Якщо валютний курс залишається незмінним, приріст ВВП на кожний додатковий 1 п.п. призводитиме до зростання імпорту на $4\% = 4 + 0.5 \cdot 0$. Якщо валютний курс зменшується на 6%, приріст ВВП на кожний додатковий 1 п.п. призводитиме до зростання імпорту лише на $1\% = 4 - 0.5 \cdot 6$.

Сумарний вплив зростання валютного курсу на приріст імпорту залежить від приросту ВВП: $2 + 0.5 GDP$.

Трифакторна регресія з урахуванням взаємодії може мати вигляд:

$$Y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_1x_2 + b_5x_2x_3 + b_6x_1x_3 + b_7x_1x_2x_3. \quad (7.7)$$

Розгляд ефектів взаємодії може ускладнюватися обмеженою кількістю спостережень. Кожний ефект взаємодії – це ніби додатковий фактор, а чим більше факторів у регресії, тим більше має бути спостережень для забезпечення надійності результатів.

Інша проблема – мультиколінеарність. Наприклад, якщо x_1 та x_1x_2 сильно корелюють, тоді x_1 та x_1x_2 одночасно не можуть перебувати у рівнянні регресії. У цьому випадку варто спробувати спочатку включити в регресійну модель лише x_1 поряд з іншими факторами. А потім – лише x_1x_2 поряд з іншими факторами. Далі визначаємо, яка з моделей є більш адекватною.

7.5. Модель лінійної регресії впливу платіжного балансу на економічне зростання у Microsoft Excel

З минулого прикладу спробуємо побудувати залежність GNI_gr_next від $PrivGap_ch$ та $TradeBal$, за даними (фрагмент) на Рис.7.32. Попередньо ми видалили спостереження (рядки), з пропущеними даними. В надбудові *Анализ даних* оберіть *Регрессия*. Вкажіть опції як на Рис. 7.33. *Константа* – *ноль* обирається рідко: лише якщо в теорії передбачається, що вільний член регресії (константа) дорівнює нулю або якщо він статистично незначущий згідно *t*-тесту.

	A	B	C
1	GNI_gr_next	PrivCap_ch	TradeBal
2	4.432849018	0.04295526	-4.83165389
3	4.06961789	-0.08486369	-5.90668014
4	1.083916811	2.999932997	-6.42207319
5	1.66467976	-3.761545283	-3.6647674
6	4.479724594	-0.709414795	-3.46309028
7	7.564802222	0.179191146	-2.69073512
8	2.576589461	5.205072487	-2.35087248
9	6.889065205	-1.891069086	-0.74116576
10	7.825935282	-4.144650705	-1.80668233
11	3.653991103	0.492307626	-2.87401591
12	3.134376285	-1.253825362	-3.96994023
13	3.715991891	3.879372183	0.09951204
14	6.106408588	-7.988104575	6.670216
15	8.937283309	-3.651521724	17.0467739
16	3.788332458	-4.23689431	20.0271025
17	5.228103443	4.014575918	12.6761036

Рис.7.32.

Рис.7.33.

У новому аркуші побачимо результати (див. Рис.7.34-7.41). З першої частини (Рис.7.34) видно, що коефіцієнт детермінації становить лише 0.1998, а скоригований 0.1635. Використано 46 спостережень. З другої частини (Рис.7.35) видно рівень значущості коефіцієнту детермінації (0.0074 менше 0.05), тобто він є значущим.

<i>Регрессионная статистика</i>	
Множественный R	0.447028557
R-квадрат	0.199834531
Нормированный R-квадрат	0.163463373
Стандартная ошибка	4.41628501
Наблюдения	47

Рис.7.34.

З третьої частини (Рис.7.36) можливо побудувати рівняння регресії:

$$GNI_gr_next = 4.094 + 0.55PrivGap_ch + 0.230TradeBal \quad (7.11)$$

Дисперсионный анализ					
	df	SS	MS	F	Значимость F
Регрессия	2	214.3175	107.1587	5.4943132	0.007412347
Остаток	44	858.1572	19.50357		
Итого	46	1072.475			

Рис.7.35.

	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%
Y-пересечение	4.094344659	0.669788508	6.112891767	2.303E-07	2.744474617	5.4442147
PrivCap_ch	0.557205704	0.228911487	2.434153526	0.0190538	0.095864915	1.018546493
TradeBal	0.229950936	0.09537579	2.410999006	0.020152	0.037733661	0.422168212

Рис.7.36.

При цьому всі коефіцієнти регресії є значущими (рівні значущості: 0.000, 0.019, 0.020). Вказані 95%-довірчі межі для кожного коефіцієнту. Наприклад, коефіцієнт при *PrivGap_ch* знаходиться в межах від 0.096 до 1.019.

З четвертої частини (Рис.7.37 – фрагмент), можливо побудувати гістограму за допомогою опції *Гистограмма* надбудови *Пакет анализа*, прописавши перед цим інтервал карманів, як ми робили це раніше. З гістограми (Рис.7.38) видно, що розподіл залишків неідеально, але близький до нормального. В п'ятій частині (Рис. 7.39 – фрагмент) вказаний процентний ранг кожного значення залежної змінної.

Наблюдение	Предсказанное GNI_gr_next	Остатки	Стандартные остатки
1	3.007236241	1.425612777	0.330063168
2	2.6888115	1.38080639	0.319689426
3	4.289162695	-3.205245884	-0.742090438
4	1.155673476	0.509006284	0.117847026
5	2.902713836	1.577010758	0.365115391
6	3.575453928	3.989348294	0.923628741
7	6.454055412	-3.877465951	-0.897725326
8	2.870198415	4.01886679	0.930462973
9	1.36947335	6.456461931	1.494824058
10	3.707778627	-0.053787524	-0.012453087

Рис.7.37.



Рис.7.38.

ВЫВОД ВЕРОЯТНОСТИ	
Персентиль	GNI_gr_next
1.063829787	-13.7417618
3.191489362	-8.01906589
5.319148936	-4.73133101
7.446808511	-4.64593308
9.574468085	-3.06393297
11.70212766	0.546342169
13.82978723	0.69190861
15.95744681	1.083916811
18.08510638	1.443758314
20.21276596	1.66467976

Рис.7.39.

З графіків залишків та підбору (Рис.7.40-7.41) можливо дізнатися про лінійний або нелінійний характер зв'язку по кожному фактору і непрямо про наявність чи відсутність гетероскедастичності.

В Microsoft Excel Можливо також самостійно будувати регресію за допомогою функцій **ЛИНЕЙН**, **ФРАСП**, **СТЬЮДРАСПОБР**, **ТЕНДЕНЦИЯ**, **ПРЕДСКАЗ**.

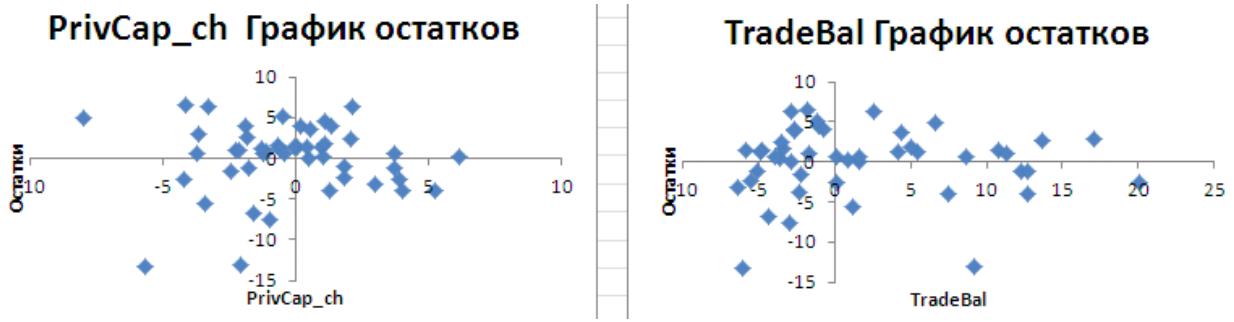


Рис.7.40.

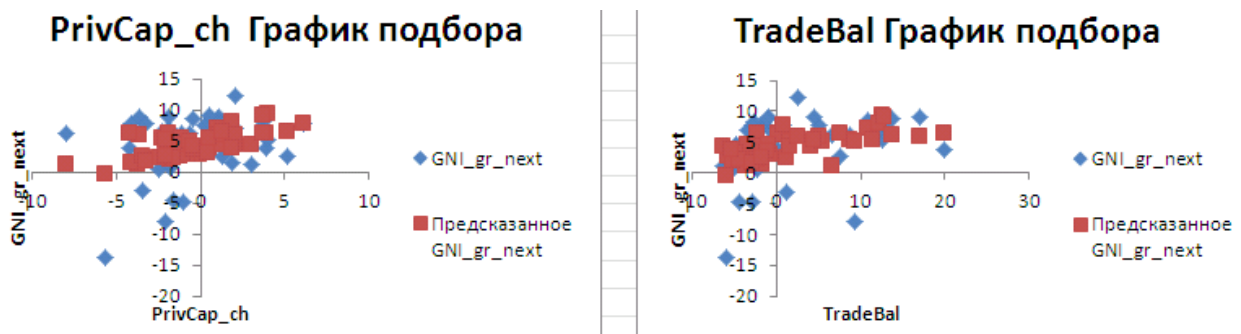


Рис.7.41.

7.6. Прогнозування динаміки валютного курсу в умовах невизначеності значень факторів у лінійній регресії

Генерація випадкових чисел може бути корисною для прогнозування в умовах ризику. Наприклад, можливо до моделей включати на вході багаторазово різні значення екзогенних незалежних змінних (припускаємо, що вони мають відомий для нас розподіл) і дивитися весь спектр значень, які моделі дають на виході (визначаючи, наприклад, середньоочікуване значення, коефіцієнт варіації, 95%-довірчий інтервал для прогнозних значень тощо).

Наведемо приклад. Припустимо, що приріст індексу валютного курсу у % визначається за формулою:

$$EXR = 0.1 + 0.4GDP - 0.5PR - 30CC \quad (7.12)$$

де PR – інфляція у %;

GDP – приріст ВВП у %;

CC – бінарна змінна Валютна криза (1 – є; 0 – немає).

Припустимо, ми прогнозуємо приріст ВВП наступного року на рівні +3% (середня або математичне сподівання), приріст ВВП має нормальний розподіл, 95%-довірчий інтервал для прогнозного значення становить (-3%;9%). Тобто стандартне відхилення будемо брати як 3%, оскільки розподіл нормальний (ділимо різницю між верхньою межею 95%-довірчого інтервалу та середньою приблизно навпіл).

Інфляція прогнозується на рівні 10%. Інфляція має нормальний розподіл. 95%-довірчий інтервал прогнозного значення становить (0%;20%). Тобто стандартне відхилення становить 5%.

В Microsoft Excel за допомогою опції *Генерація случайных чисел* подібно тому, як ми робили раніше, створюємо випадкові значення для двох змінних: приріст ВВП та інфляція в стовпчиках А та В.

У нас немає даних буде чи не буде валютна криза, але ми припускаємо що ймовірність валютної кризи становить 0.1, а її відсутності 0.9. Запишемо ймовірності в окремому місці в таблиці і створимо випадкові значення змінної «валютна криза», але обравши дискретний розподіл (див. Рис.7.42 і 7.43).

	F	G
Валютна криза		
	1	0.1
	0	0.9

Рис.7.42.

Генерация случайных чисел ? X

Число переменных: OK

Число случайных чисел: Отмена

Распределение: Справка

Параметры

Входной интервал значений и вероятностей:
 📄

Случайное рассеивание:

Параметры вывода

Выходной интервал: 📄

Новый рабочий лист:

Новая рабочая книга

Рис.7.43.

Зверніть увагу на те, що ми вказали в полі *Входной интервал значений и вероятностей*. Також замість виводу результатів у новий робочий лист ми вивели його одразу у потрібний нам інтервал.

Третю змінну (валютний курс) розраховуємо за допомогою формул. У результаті одержуємо таблицю на Рис. 7.44 (стовпчик D у двох видах з результатами і формулами).

	A	B	C	D	D
1	Приріст ВВП	Інфляція	Валютна криза	Приріст валютного курсу	Приріст валютного курсу
2	5.51	2.99	0	0.81	=0.1+0.4*A2-0.5*B2-30*C2
3	2.13	-5.08	0	3.49	=0.1+0.4*A3-0.5*B3-30*C3
4	6.62	12.06	0	-3.28	=0.1+0.4*A4-0.5*B4-30*C4
5	5.78	11.95	0	-3.56	=0.1+0.4*A5-0.5*B5-30*C5
6	3.24	10.96	0	-4.08	=0.1+0.4*A6-0.5*B6-30*C6
7	6.28	5.63	0	-0.20	=0.1+0.4*A7-0.5*B7-30*C7
8	7.74	8.98	0	-1.29	=0.1+0.4*A8-0.5*B8-30*C8
9	4.09	7.44	0	-1.98	=0.1+0.4*A9-0.5*B9-30*C9
10	5.21	3.11	0	0.63	=0.1+0.4*A10-0.5*B10-30*C10
11	-0.84	10.22	1	-35.34	=0.1+0.4*A11-0.5*B11-30*C11
12	1.52	5.01	0	-1.80	=0.1+0.4*A12-0.5*B12-30*C12
13	3.16	4.25	0	-0.76	=0.1+0.4*A13-0.5*B13-30*C13
14	2.93	16.51	0	-6.98	=0.1+0.4*A14-0.5*B14-30*C14
15	7.07	21.35	0	-7.75	=0.1+0.4*A15-0.5*B15-30*C15
16	5.28	11.71	0	-3.65	=0.1+0.4*A16-0.5*B16-30*C16
17	7.49	-0.59	1	-26.61	=0.1+0.4*A17-0.5*B17-30*C17
18	1.51	8.01	0	-3.30	=0.1+0.4*A18-0.5*B18-30*C18
19	6.57	4.03	0	0.72	=0.1+0.4*A19-0.5*B19-30*C19
20	5.74	9.63	0	-2.42	=0.1+0.4*A20-0.5*B20-30*C20
21	0.78	3.85	0	-1.52	=0.1+0.4*A21-0.5*B21-30*C21

Рис.7.44.

Отже, середнє приросту валютного курсу є майже -5, тобто курс зменшиться на 5% (див. Рис 7.45). Але медіана відрізняється і показує, що курс лише на 2%. Оскільки медіана і середня відрізняються, ми вже бачимо, що розподіл прогнозного приросту валютного курсу відрізняється від нормального. Максимальне зниження курсу становить 35%, максимальне зростання 3.5%. Із ймовірністю 70% динаміка курсу буде в межах від -7.7% до +0.7%.

<i>Столбец1</i>	
Среднее	-4.94434
Стандартная ошибка	2.098026
Медиана	-2.19959
Мода	#Н/Д
Стандартное отклонение	9.382656
Дисперсия выборки	88.03423
Эксцесс	6.562796
Асимметричность	-2.60446
Интервал	38.83578
Минимум	-35.3428
Максимум	3.492969
Сумма	-98.8868
Счет	20
Наибольший(3)	0.715363
Наименьший(3)	-7.74566
Уровень надежности(95.0%)	4.391218

Рис.7.45.

Ми навели простий приклад. На практиці моделі часто мають складнішу структуру. Також кількість випадкових значень, які потрібно створювати, є набагато більшою – декілька сотень або навіть тисяч, залежно від складності моделей. Це робиться для того, щоб охопити максимально більшу кількість можливих комбінацій значень екзогенних змінних.

ГЛАВА 8

Дослідження міжнародної торгівлі та валютних криз методами нелінійного регресійного аналізу

8.1. Методи нелінійної регресії

Серед усіх багатовимірних регресійних моделей найбільш вивчена лінійна, хоча лише деякі з економічних процесів в сфері міжнародних економічних відносин можна моделювати за допомогою такої моделі. Її вибір залежить від типу процесу і тривалості спостереження за ним. Багато процесів в сфері міжнародних економічних відносин при нетривалому спостереженні за ними можна певним наближенням моделювати за допомогою лінійної багатофакторної моделі.

Значна частина нелінійних регресійних моделей може бути перетворена на лінійні за допомогою трансформації змінних. Після цього вони можуть бути побудовані як звичайні лінійні регресійні моделі. Різниця буде лише у інтерпретації моделі.

Розглянемо деякі випадки однофакторної нелінійної моделі, які шляхом перетворень зводяться до лінійної моделі:

$$y = b_0 + b_1 x. \quad (8.1)$$

У Табл.8.1 вказано, яким чином (за допомогою яких замінь) можуть перетворюватися змінні та коефіцієнти регресії для ряду розповсюджених нелінійних однофакторних моделей, щоб звести до лінійної моделі (8.1).

Таблиця 8.1.

Функція	y	x	b ₀	b ₁
$y = b_0 + b_1/x$	y	1/x	b ₀	b ₁
$y = 1/(b_0 + b_1 x)$	1/y	x	b ₀	b ₁
$y = x/(b_0 + b_1 x)$	x/y	x	b ₀	b ₁
$y = b_0 + b_1^x$	ln(y)	x	ln(b ₀)	ln(b ₁)
$y = 1/(b_0 + b_1 e^x)$	1/y	e ^x	b ₀	b ₁
$y = b_0 x^{b_1}$	ln(y)	ln(x)	ln(b ₀)	b ₁
$y = b_0 + b_1 \ln(x+1)$	y	ln(x+1)	b ₀	b ₁
$y = b_0 x(b_1 + b_2 x)$	1/y	1/x	b ₁ /b ₀	b ₂ /b ₀
$y = b_0 e^{b_1/x}$	ln(y)	1/x	ln(b ₀)	b ₁
$y = b_0 + b_1 x^n$	y	x ⁿ	b ₀	b ₁

Наприклад, двофакторну модель

$$y = b_0 x_1^{b_1} x_2^{b_2} \quad (8.2)$$

можна перетворити шляхом логарифмування до моделі

$$\ln y = \ln b_0 + b_1 \ln x_1 + b_2 \ln x_2, \quad (8.3)$$

яка після заміни $Y = \ln y$, $B_0 = \ln b_0$, $X_1 = \ln x_1$, $X_2 = \ln x_2$ буде записана у вигляді:

$$Y = B_0 + b_1 X_1 + b_2 X_2 \quad (8.4)$$

5. Кусоково-лінійна регресія (*Piecewise linear regression*). Використовується для розрахунку регресії типу:

$$Y = (b_{01} + b_{11}x_1 + b_{21}x_2 + \dots) \cdot (Y < Y^*) + (b_{02} + b_{12}x_1 + b_{22}x_2 + \dots) \cdot (Y > Y^*), \quad (8.6)$$

де $Y < Y^*$ та $Y > Y^*$ – логічні вирази, які повертають 1, якщо відповідний вираз відповідає дійсності, або 0, якщо не відповідає. Тобто за значень залежної змінної меншої за вказаний рівень буде використана форма регресії $b_{01} + b_{11}x_1 + b_{21}x_2 + \dots$, а у разі більшого – $b_{02} + b_{12}x_1 + b_{22}x_2 + \dots$. Точка розриву Y^* (*Breakpoint*) може бути вказана серед початкових значень коефіцієнтів.

8.3. Дослідження факторів валютних криз за допомогою логіт-регресії

Моделі бінарних відгуків (логіт або пробіт) використовуються, якщо залежна змінна є бінарною, наприклад (1-валютна криза є, 0 – валютної кризи немає). Але в регресії можливо використати залежну змінну, яка є неперервною і може приймати значення від 0 до 1.

У логіт-регресії це досягається за допомогою такого рівняння:

$$p = y = \frac{\exp(b_0 + b_1x_1 + \dots + b_nx_n)}{1 + \exp(b_0 + b_1x_1 + \dots + b_nx_n)}, \quad (8.9)$$

де y – це неперервна залежна змінна на відрізку $[0;1]$ – аналог ймовірності p того, що бінарна залежна змінна прийме значення 1.

Для того щоб одержати звичайну неперервну залежну змінну на відрізку $(-\infty, +\infty)$ використовується логіт-перетворення:

$$p' = \ln\left(\frac{1}{1-p}\right) = b_0 + b_1x_1 + \dots + b_nx_n. \quad (8.10)$$

8.4. Функції втрат

В окремих випадках програма дозволяє користувачу обирати Функцію втрат (*Loss function*), яка мінімізується при визначенні коефіцієнтів регресії. Відхилення спостерігаємих значень від розрахованих і є фактично втратами точності передбачення за допомогою моделі.

За звичайного методу найменших квадратів мінімізується сума квадратів залишків (відхилень спостерігаємих значень залежної змінної від значень, розрахованих за допомогою регресії). Але можливо, наприклад, замість квадратів залишків використати абсолютні величини залишків. Цей метод зменшить вплив великих залишків на форму регресії.

Метод зважених найменших квадратів передбачає, що залишки можуть бути зважені за допомогою певної іншої змінної, наприклад, з метою здійснення поправки на гетероскедастичність або серійну кореляцію.

У нелінійному оцінюванні часто використовується метод найбільшої правдоподібності. Він дає ті самі оцінки параметрів регресії, що й метод найменших квадратів, якщо виконуються передумови для регресійного аналізу.

До функції втрат може включатися штрафна функція, яка дорівнює 0, якщо оцінювані параметри знаходяться в межах допустимих значень, і дорівнює дуже великій величині, якщо параметри виходять за межі допустимих значень. Наприклад, якщо нам важливо, щоб вільний член b_0 був невід'ємним функція втрат може мати вигляд $((b_0 < 0))$ – логічний вираз, який приймає значення 1, якщо оцінений вільний член стає від'ємним):

$$L = (obs - pred)^2 + (b_0 < 0) \cdot 1000000. \quad (8.12)$$

Перед початком пошуку мінімуму користувач вказує початкові значення коефіцієнтів регресії і величину кроків, відштовхуючись від яких, програма ітеративним шляхом намагається

знайти мінімум функції втрат. Критерій сходження визначає момент, коли ітерації можливо припинити.

На жаль, часто програма знаходить локальний мінімум і зупиняється у пошуках, а глобальний (абсолютний) мінімум залишається незнайденим. Відповідно, розраховані коефіцієнти регресії виявляються неоптимальними. Візуально локальні мінімуми є невеликими впадинами на графіку функції втрат. Наприклад, на Рис.8.27 показано, як при певному первинному значенні коефіцієнту регресії, програма зупинить пошук оптимального значення b_i , коли знайде другий локальний мінімум:

Тому варто спробувати різні початкові значення.

Розповсюдженим алгоритмом мінімізації функцій втрат є квазі-ньютонівський метод (*quasi-Newton*). Він передбачає розрахунок першої та другої похідної функції втрат для пошуку її мінімуму. Окрім нього доступні й інші алгоритми:

- Симплекс метод (*Simplex*) – є менш чутливим до локальних мінімумів;
- Симплекс і квазі-ньютонівський метод (*Simplex and quasi-Newton*);

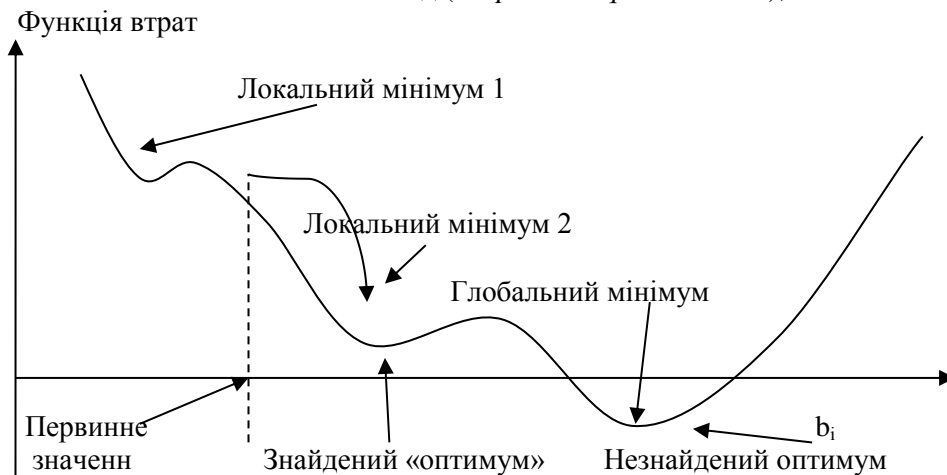


Рис.8.27.

- Метод Хука-Дживіса (*Hooke-Jeeves pattern moves*);
- Метод Хука-Дживіса і квазі-ньютонівський метод (*Hooke-Jeeves and quasi-Newton*);
- Метод Розенброка (*Rosenbrock pattern search*) або метод обертання координат;
- Метод Розенброка і квазі-ньютонівський метод (*Rosenbrock and quasi-Newton*).

8.5. Модель нелінійної регресії факторів високотехнологічного експорту у Microsoft Excel

В більшості випадків, як фактори, потрібно використовувати вже трансформовані змінні (наприклад квадрат незалежної змінної незалежної змінної як фактор), лінеаризуючи модель, і далі використовувати опцію *Регресія* у надбудові *Пакет Аналіза*.

Але є також вбудовані функції: ЛГРФПРИБЛ та РОСТ. Наведемо приклад використання такої функції масиву:

=ЛГРФПРИБЛ(відомі значення у;відомі значення х;константа;статистика)

Аргументи функції такі. Відомі значення у – діапазон значень залежної змінної. Відомі значення х – діапазон значень незалежних змінних. Константа – якщо ИСТИНА, то b розраховується, якщо ЛОЖЬ, то вважається, що $b=1$. Статистика – якщо ИСТИНА, розраховується додаткова статистика.

Функція вводиться як функція масиву за допомогою команди CTRL+SHIFT+ВВОД.

Рівняння функції має вигляд: $y = b m_1^{x_1} m_2^{x_2}$.

Припустимо, ми ставимо задачу розрахувати залежність частки високотехнологічного експорту у експорті товарів (*HTExp* – High-technology exports (% of manufactured exports)) від змінних: *Lit* – рівень писемності населення - Literacy rate, adult total (% of people ages 15 and above); *Internet* – кількість інтернет-користувачів на 100 людей – Internet users (per 100 people).

Дані за 2009 р. взяті з World Development Indicators. На Рис.8.28 показана таблиця вхідних даних (вибірка):

Вводимо функцію =ЛГРФПРИБЛ(В2:В15;С2:Д15;ИСТИНА;ИСТИНА). Одержуємо результати (див. Рис.8.29). Отже залежність має вигляд:

$$HTExp = 5.574(0.963127^{Lit}) \cdot (1.06635^{Internet}). \quad (8.13)$$

	A	B	C	D
1		HTExp	Lit	Internet
2	Panama	0.02	93.61	27.73
3	Portugal	4.16	94.91	48.61
4	Romania	10.07	97.65	36.25
5	Russian Federation	9.33	99.56	42.09
6	Saudi Arabia	0.26	86.13	36.55
7	Singapore	49.06	94.71	73.35
8	El Salvador	5.03	84.10	14.43
9	Slovenia	6.51	99.68	63.66
10	Trinidad and Tobago	0.26	98.74	36.29
11	Turkey	1.87	90.82	36.76
12	Uruguay	5.37	98.27	55.46
13	Samoa	0.33	98.78	4.93
14	Yemen, Rep.	0.38	62.39	1.80
15	Zambia	1.56	70.88	6.42

Рис.8.28.

F	G	H
1.06635	0.963127	5.574194
0.0296609	0.056953	4.607095
0.3360428	1.80654	#Н/Д
2.783666	11	#Н/Д
18.169464	35.89944	#Н/Д
#Н/Д	#Н/Д	#Н/Д

Рис.8.29.

Ми тепер можемо розрахувати, підставляючи відомі значення факторів залежну змінну (див. Рис.8.30).

J	K	L
HTExp	Lit	Internet
23.05164169	99	80

J	K	L
HTExp	Lit	Internet
=H2*СТЕПЕНЬ(G2;K3)*СТЕПЕНЬ(F2;L3)	99	80

Рис.8.30.

Проте слід із застереженням використовувати ці результати. Модель потребує трансформації, враховуючи, що коефіцієнт 0.963127 менший за 1, отже писемність негативно впливає на частку високотехнологічного експорту. У той же час кореляція між писемністю та часткою високотехнологічного експорту слобопозитивна 0.209 (якщо скористатися функцією КОРРЕЛ). Також негативний вплив є дивним з теоретичної точки зору. Тому ймовірно, в моделі існує спотворюючий ефект мультиколінеарності.

Коефіцієнт детермінації становить 0.336. *F*-статистика 2.78. Кількість ступенів свободи 11. 4.61, 0.057, 0.297 – стандартні похибки для кожного з коефіцієнтів регресії. 1.81 – стандартна похибка для залежної змінної. Але слід враховувати, що додаткова статистика розраховується для лінеаризованої моделі:

$$\ln(HTExp) = \ln(5.574) + Lit \cdot \ln(0.963127) + Internet \cdot \ln(1.06635). \quad (8.14)$$

Для того, щоб дізнатися рівень значущості *F*-критерію, скористаємося функцією:

$$=ФРАСП(F; v1; v2),$$

де *F*-значення *F*-статистики, *v1* – кількість факторів, *v2* – ступені свободи.

В нашому прикладі =ФРАСП(F5; 2; G5) повертає 0.1051, тобто коефіцієнт детермінації незначущий (більше 0.05).

8.6. Алгоритм дослідження наслідків утворення зони вільної торгівлі

Далі буде описаний методологічний алгоритм для дослідження потенційних наслідків утворення зони вільної торгівлі (ЗВТ) між Україною та Туреччиною, використаний автором у роботі [42]. Ці результати можуть бути використані для оцінки доцільності утворення зони вільної торгівлі та наслідків для взаємного торговельного балансу.

За основу береться гравітаційна модель у такому загальному модифікованому вигляді (дані станом на 2010 рік, якщо не вказане інше):

$$Trade = b_0 GDP^{b_1} Dist^{b_2} e^{b_3 TR} \quad (8.15)$$

де: $TRAdе$ – експорт Exp (або імпорт Imp) Туреччини до (з) відповідної країни-партнера;

GDP – ВВП країни-партнера;

$Dist$ – відстань від столиці країни-партнера до Анкари;

TR – одна зі змінних, що характеризують регулювання міжнародної торгівлі (див. далі).

Якщо лінеаризувати рівняння, воно матиме вигляд (всі змінні, крім TR , у формі натурального логарифму):

$$TRAdе = b_0 + b_1 GDP + b_2 Dist + b_3 TR, \quad (8.16)$$

Змінна $Dist$ в окремих моделях не використовується. Це обумовлене тим, що:

- імпорт Туреччини мало залежить від відстані до країни походження імпорту;
- відстань в окремих випадках корелює зі змінною регулювання торгівлі, тоді ми залишаємо лише змінну регулювання торгівлі, хоча це зменшує надійність оцінки впливу саме зони вільної торгівлі (вплив її важче відрізнити від впливу відстані);

Ми використовуємо по черзі декілька варіантів визначення TR :

- FTA – бінарна змінна яка приймає значення 1 (є ЗВТ або митний союз) або 0 (немає ЗВТ або митного союзу);
- TF – зважена середня ставка митного тарифу, що застосовується для всіх продуктів (Tariff rate, applied, weighted mean, all products (%)) за даними World Development Indicators у країнах-контрагентах Туреччини (для країн із ЗВТ вважаємо, що $TF = 0$);
- TN – аналогічний показник імпортного тарифу, але Туреччини перед утворенням відповідних ЗВТ (для країн без ЗВТ вважаємо, що $TN = 0$);
- $FTAT$ – час перебування у зоні вільної торгівлі у роках.

Перша модель для експорту Туреччини до країн-партнерів ($TR = FTA$) виражає вплив зони вільної торгівлі без урахування факторів величини тарифів та часу:

$$Exp = 17.500 + 0.825GDP - 1.072Dist + 0.445FTA. \quad (8.17)$$

Враховані параметри якості моделі. Скоригований коефіцієнт детермінації 0.50. Значущість F -статистики 0.00000. Згідно t -статистики всі коефіцієнти значущі. Суттєві мультиколінеарність і гетероскедастичність не спостерігаються. Розподіл залишків близький до нормального. Серійна кореляція 0.14. Впливових спостережень (викидів серед залишків) не виявлено.

Якщо підставити дані по Україні (логарифми ВВП та відстані) і $FTA = 0$ експорт дорівнюватиме 12.179 (95%-довірчий інтервал: 9.18;15.18), якщо $FTA = 1$, то 13.580 (10.59;16.57). Враховуючи, що це логарифми, відповідні абсолютні значення становлять близько 195000 і 790000 (нас цікавить саме співвідношення цих двох розрахованих величин експорту, навіть якщо вони суттєво відрізняються від фактичних значень). Тобто за умов ЗВТ, як очікується, експорт Туреччини до України (він же імпорт України з Туреччини) має бути більшим на 305%, але цей результат не є точним, враховуючи широкі довірчі інтервали. Це є першою оцінкою потенціалу зростання експорту Туреччини до України за умов утворення ЗВТ.

Аналогічно будуємо дві моделі для імпорту (за умов включення до аналізу впливових спостережень (США і Китай, залишки (різниця між розрахованим і фактичним значенням залежної змінної – імпорт, по яких є достатньо великими) і виключення цих впливових спостережень):

$$Imp = 5.490 + 1.404GDP + 1.478FTA, \quad (8.18)$$

$$Imp = 5.440 + 1.34GDP + 1.823FTA . \quad (8.19)$$

Ці дві моделі дещо відрізняються за коефіцієнтами. Аналогічно, за умов ЗВТ, як очікується, імпорт Туреччини з України має бути більшим на 339% та 519% відповідно. Обидві моделі імпорту показують що потенціал зростання імпорту Туреччини з України більший за потенціал зростання експорту Туреччину в Україну.

Альтернативний варіант для оцінки впливу ЗВТ міг би бути таким. Пересвідчившись, що вплив ЗВТ значущий, використавши бінарну змінну, можливо побудувати по дві окремі пари моделей (для експорту та імпорту) для країн, що входять у ЗВТ, та країн, що не входять до ЗВТ. Цей варіант може дати можливість урахувати взаємодію між приналежністю до ЗВТ та іншими змінними (коефіцієнти при ВВП та відстані ймовірно будуть відрізнятися).

Проте цифри потенціалу зростання виглядають завеликими, але цьому є пояснення: стартові ставки митного тарифу при утворенні Туреччиною раніше ЗВТ та сучасні ставки митного тарифу країн, які не входять з Туреччиною в ЗВТ, як правило більші за поточні ставки у випадку України. А ефект від утворення ЗВТ тим більший, чим більші ставки мита, які ліквідуються при взаємній вільній торгівлі.

Тому наступним кроком модифікуємо моделі для експорту так, що $TR = TN$ (з та без впливових спостережень):

$$Exp = 17.883 + 0.858GDP - 0.992Dist - 0.0867FTA , \quad (8.20)$$

$$Exp = 17.934 + 0.863GDP - 1.108Dist - 0.0664FTA . \quad (8.21)$$

За умов ЗВТ, як очікується, експорт Туреччини до України має бути більшим на 27 або 20%. Ми вже бачимо, що потенціал зростання експорту з урахуванням низького рівня тарифних ставок в Україні не такий високий.

Модифікуємо моделі для імпорту так, що $TR = TN$:

$$Imp = 5.812 + 1.377GDP + 0.220FTA , \quad (8.22)$$

$$Imp = 5.949 + 1.324GDP + 0.241FTA . \quad (8.23)$$

За умов ЗВТ, як очікується, імпорт Туреччини з України має бути більшим на 70 або 79%, але цей результат не є точним, враховуючи широкі довірчі інтервали. Ми бачимо, що потенціал зростання імпорту Туреччини також є меншим ніж за першого методу, але так само більший за потенціал зростання експорту Туреччини.

Трансформуємо тепер моделі так, що $TR = FTA$. Для експорту:

$$Exp = 8.487 + 0.874GDP + 0.048FTA \quad (8.24)$$

Якщо підставити дані по Україні і $FTA = 0$ експорт дорівнюватиме 12.621, якщо $FTA = 1$ (рік після утворення ЗВТ), то 12.669, і далі додаючи по 1 року: 12.717, 12.765 ... Тобто за рахунок ЗВТ, як очікується, щорічно експорт Туреччини до України має додатково зростати на 5%.

Моделі для імпорту Туреччини з країн-партнерів:

$$Imp = 5.793 + 1.42GDP + 0.0657FTA , \quad (8.25)$$

$$Imp = 5.923 + 1.402GDP + 0.0621FTA . \quad (8.26)$$

Аналогічно, щорічно імпорт Туреччини з України має додатково зростати майже на 7 або 6.5%.

Розраховані моделі експорту та імпорту Туреччини до (з) країн-контрагентів пояснюють приблизно половину коливань експорту та імпорту Туреччини. Вони відповідають або принаймні майже відповідають критеріям якості регресійних моделей. Всі варіанти моделей показують, що виходячи зі значень незалежних змінних, характерних для України, утворення ЗВТ з Туреччиною на характерних для Туреччини умовах призведе швидше до покращання сальдо торговельного балансу України.

Статичні моделі в цілому показують потенціал для зростання зовнішньої торгівлі (враховуючи стартові тарифи: експорт Туреччини в Україну на 20-27%, а, імпорт Туреччини з України на 70-79%), а динамічні – що цей потенціал можливо досягнути лише після проходження тривалого періоду (щорічне зростання за рахунок утворення зони вільної торгівлі: експорту Туреччини в Україну на 5%, а, імпорту Туреччини з України на 6.5-7%). Ці цифри є лише середньо очікуваними, фактичні результати можуть суттєво від них відхилятися. В окремих моделях позитивний ефект від власне угод про ЗВТ може бути наслідком не стільки угоди, скільки географічного розташування країн-партнерів.

Недоліком використаного підходу, є те, що він не враховує галузеву структуру взаємної торгівлі, тому для остаточного прийняття рішення варто використовувати цей підхід як додатковий до більш детальних методів.

ГЛАВА 9

Класифікація країн за рівнем відкритості та зовнішньої збалансованості економіки методами кластерного аналізу

9.1. Метод кластерного аналізу

Кластерний аналіз – це метод класифікаційного аналізу. Під час застосування цього методу відбувається розбиття множини досліджуваних об'єктів та ознак на однорідні в деякому розумінні групи або кластери. Це багатомірний статистичний метод, тому припускається, що вихідні дані можуть бути достатньо великого обсягу, тобто суттєво більшим може бути як кількість дослідження (спостережень), так і ознак, які характеризують ці об'єкти.

Кластерний аналіз призначений для класифікації спостережень, наприклад, країн, фірм, товарів, міст, регіонів, періодів часу або комбінованих просторово-часових спостережень (наприклад, країно-місяців). При цьому він дає можливість класифікувати спостереження одночасно за декількома ознаками (змінними). Змінні можуть бути будь-якого характеру. Всі спостереження поділяються на декілька кластерів (груп спостережень), так що спостереження всередині одного кластеру схожі між собою і відрізняються від спостережень з інших кластерів. Кількість кластерів може бути обрана дослідником априорі або в процесі кластерного аналізу.

Перевага кластерного аналізу в тому, що він дає можливість зробити розбиття об'єктів не за однією ознакою, а за декількома ознаками. Крім того, при застосуванні кластерного аналізу, на відміну від більшості математико-статистичних методів, не накладають ніяких обмежень на вид об'єктів, що дозволяє досліджувати множину вихідних даних практично довільної природи.

Кластери – це групи однорідності і задача кластерного аналізу полягає в тому, щоб на основі ознак об'єктів розбити їх множину на m (m - ціле) кластерів так, щоб кожен об'єкт належав тільки одній групі розбиття. При цьому об'єкти, які належать одному кластеру, повинні бути однорідними (подібними), а об'єкти, які належать різним кластерам - різнорідні.

Якщо об'єкти кластеризації подати як точки в n -вимірному просторі ознак (n - кількість ознак, які характеризують об'єкти), то схожість між об'єктами визначається через поняття відстані між точками, оскільки інтуїтивно зрозуміло, що чим менша відстань, тим вони більш схожі.

Наприклад, ми хочемо класифікувати 6 країн за двома ознаками: середня ставка митного тарифу та частка імпорту у ВВП. Класифікацію можливо зробити і візуально за допомогою діаграми розсіювання (див. Рис. 9.1).

З малюнку ми бачимо, що ближче всього розташовані спостереження E і F. А чим ближче об'єкти у цьому двомірному просторі, тим більш схожими вони є. Ми можемо розглядати їх як один кластер, а всі решта спостережень як окремі кластери. В сумі у нас будуть 5 кластерів. Якщо нам потрібні 4 кластери, спостереження будуть об'єднані таким чином: E+F, B+C, D, A. Якщо нам потрібні 3 кластери, спостереження будуть об'єднані таким чином: E+F, B+C+D, A. Якщо нам потрібні 2 кластери, спостереження будуть об'єднані так: E+F, B+C+D+A. Звичайно можливо й всі спостереження розглядати як один кластер.

Кластерний аналіз дозволяє здійснювати те саме, враховуючи точне місце розташування спостережень відносно один одного у двомірному просторі, де виміри відповідають двом змінним: середня ставка митного тарифу та частка імпорту у ВВП. Але графічний метод практично стає неможливо використовувати, якщо спостережень дуже багато, а особливо, якщо кількість вимірів (змінних), за якими відбувається класифікація, є великою.

Використовуються різні функції відстаней між спостереженнями у багатомірному просторі.

1. Найбільш розповсюдженою є *евклідова відстань* – евклідова метрика (*Euclidean distances*). Наприклад, у двомірному просторі (на площині) з осями Ox та Oy евклідова відстань між певними точками $A(x_1, y_1)$ і $B(x_2, y_2)$ розраховується за формулою:

$$d = AB = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (9.1)$$

Вона зазвичай вимірюється за абсолютними значеннями змінної, а не стандартизованими, але тоді на відстані можуть впливати одиниці виміру. Дана метрика, як і більшість інших, чутлива до одиниць вимірювання осей. Наприклад, якщо сантиметри перевести в міліметри, то зміниться і

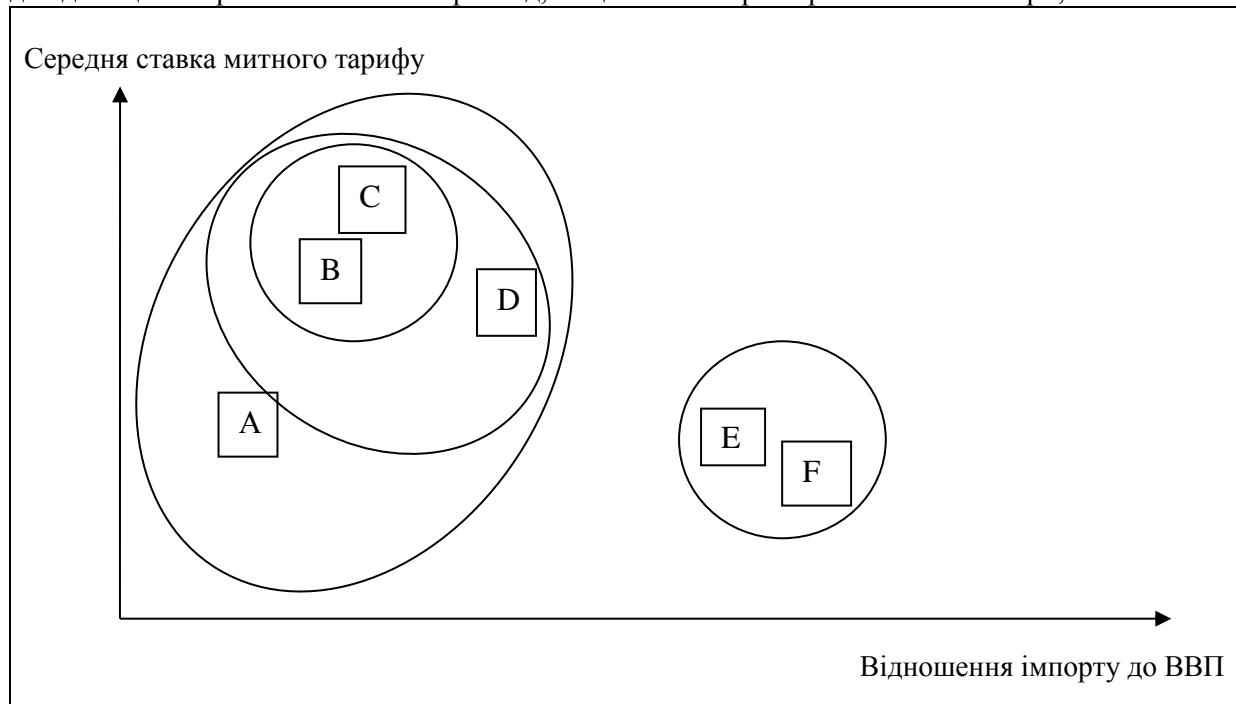


Рис.9.1.

відстань.

2. *Квадрат евклідової відстані* використовують, якщо необхідно надати більшої ваги більш віддаленим один від одного об'єктам.

3. *Манхеттенівська відстань* міських кварталів зменшує вплив окремих великих різниць між однойменними координатами точок, так як при обчисленні відстані ці різниці не підносяться до квадрату (на відміну від евклідової метрики).

4. *Відстань Чебишева* застосовують, коли бажають визначити два об'єкти, якщо вони відрізняються за якоюсь однією координатою.

5. *Процент незгоди (Percent disagreement)* використовується, коли змінні є категоріальними, і враховує частку змінних, за якими значення спостережень не дорівнюють.

6. *1 мінус коефіцієнт кореляції Пірсона (1-Pearson r)*. За цією опцією до одного кластеру відноситимуться не ті спостереження, по яких значення змінних близькі, а ті, які мають більш тісний лінійний зв'язок.

Наприклад, є три країни, які характеризуються такими значеннями змінних (відносно ВВП): експорт, прямі інвестиції, зовнішній борг і валютні резерви – A(40, 5, 60, 20), B(45, 2, 66, 15), C(80, 9, 100, 41). Ми бачимо, що країни A і B мають невелику евклідову відстань (значення змінних у країні B мало відрізняються від значень змінних у країні A), але країни A і C мають більш тісний лінійний зв'язок (значення змінних у країні C практично вдвічі більші за значення змінних у країні A).

9.2. Види кластерного аналізу

Алгоритмів кластерного аналізу достатньо багато. Їх можна поділити на *ієрархічні* (дерево-видні) та *неієрархічні*.

Ієрархічні процедури – найбільш поширені алгоритми кластерного аналізу за їх реалізації на комп'ютері. Розрізняють агломеративні (від слова *agglomerate* – збирати) та ітеративні дивизивні (від слова *division* – розділяти) процедури.

Принцип роботи ієрархічних агломеративних процедур полягає в послідовному об'єднанні груп елементів спочатку самих близьких, потім все більш віддалених один від одного. Принцип роботи ієрархічних дивизивних процедур, навпаки, полягає в послідовному розділенні груп еле-

ментів спочатку самих відділених, потім все більш близьких один від одного. Більшість цих алгоритмів виходить з матриці відстаней. До недоліків цих алгоритмів слід віднести громіздкість їх реалізації. На кожному кроці алгоритми потребують обчислення матриці відстаней, а значить, великої кількості машинної пам'яті та часу. В зв'язку з цих реалізація таких алгоритмів за великої кількості спостережень, недоцільна, а в деяких випадках – неможлива.

Загальний принцип роботи агломеративного алгоритму наступний. На першому кроці кожне спостереження розглядають як окремий кластер. В подальшому на кожному кроці роботи алгоритму відбувається об'єднання двох самих близьких кластерів, і з урахуванням прийнятої відстані за формулою перераховується матриця відстаней, розмірність якої, очевидно зменшується на одиницю. Робота алгоритму закінчується, коли всі спостереження об'єднані в один клас. Більшість програм, які реалізують алгоритм ієрархічної класифікації, передбачає графічне подання у вигляді дендрограми.

Нижче на Рис. 9.2 представлена дендрограма, яка вказує на ступінь схожості (яка є оберненою величиною відстані) між окремими спостереженнями чи кластерами, вказаними у прикладі на Рис. 9.1.

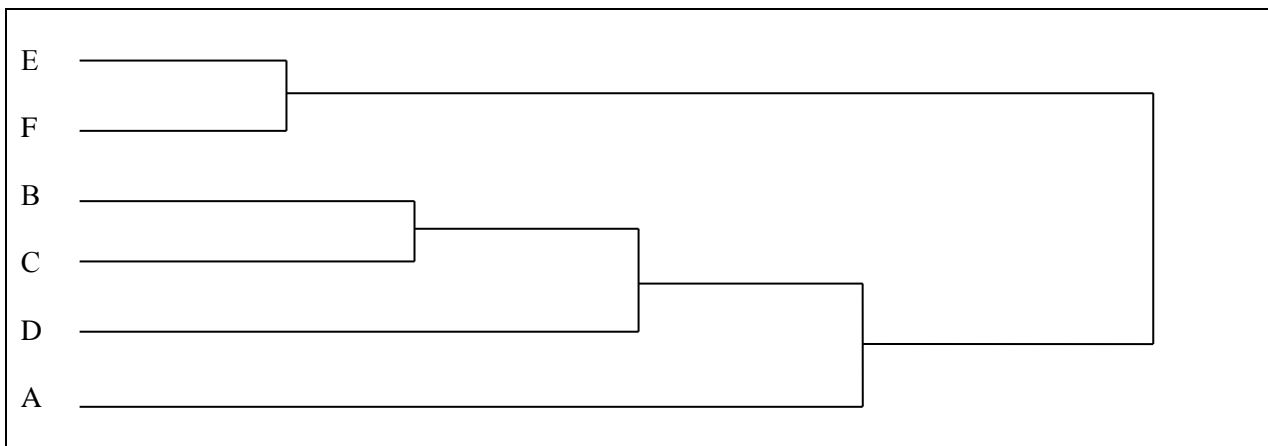


Рис.9.2.

Як розрахувати відстань між окремими спостереженнями, ми вказали раніше. Але додатково потрібно знати, як розраховувати відстань між кластерами, при вирішенні, які саме кластери слід об'єднувати на кожному кроці. Існує декілька методів, зокрема:

- *Правило окремого зв'язку (single linkage)*. На першому кроці об'єднують два кластери відстань між якими найменша. Відстань між кластерами визначають як відстань між їх найближчими елементами. На наступному кроці приєднують найближчий кластер і т.д. Цей метод ще називають *методом «найближчого сусіда»*, так як відстань між двома порівнюваними кластерами визначається як відстань між найближчими об'єктами в різноманітних кластерах. Цей метод має недолік: утворюються занадто продовгуваті кластери, а не колоподібні.
- *Правило повних зв'язків*. Суть методу в тому, що два об'єкти, які належать до однієї групи (кластеру), мають коефіцієнт схожості, який більший за деяке порогове значення S . В термінах евклідової відстані це означає, що відстань між двома точками (об'єктами) кластера не повинна перевищувати деяке порогове значення d . Таким чином, d визначає максимально допустимий діаметр підмножини, що утворює кластер. Цей метод ще називають *методом «найбільш віддалених сусідів»*, так як при достатньо великому пороговому значенні d відстань між кластерами визначається найбільшою відстанню між довільними об'єктами в різних кластерах.
- *Правило незваженого парного середнього (unweighted pair-group average)*. Відстань між двома порівнюваними кластерами визначається як середня відстань між всіма парами об'єктів з обох цих кластерів. Метод ефективний, коли об'єкт в дійсності формують різні групи, але він працює однаково добре і у випадках кластерах ланцюгового типу.

- *Правило зваженого попарного середнього (weighted pair-group average)*. Метод ідентичний попередньому, за виключенням того, що при обчисленні відстань між двома порівнюваними кластерами визначається так само як і у попередньому методі, але з урахуванням зважування на основі величини кластерів.

Припустимо, що є гіпотези відносно числа m (m - ціле) кластерів. Тоді можна створити рівно m кластерів так, щоб вони були настільки різні, наскільки це можливо. Саме для розв'язування задач цього типу призначений метод k -середніх (k -means). Гіпотеза може базуватись на теоретичних міркуваннях або на здогадах. Виконуючи послідовно розбиття на різне число кластерів, можна порівнювати якість отриманих розв'язків.

Неієрархічний метод k -середніх передбачає, що дослідник апіорі визначає кількість кластерів, на які хоче поділити спостереження. Наприклад, на 4 кластери (як варіант, рішення про кількість кластерів може бути прийняте після ієрархічного аналізу, коли вже є певне уявлення про відстані між спостереженнями). Програма спочатку випадково розподіляє спостереження по цих кластерах, а далі ітеративним шляхом перерозподіляє спостереження між кластерами. Нарешті всі спостереження розподіляються по чотирьох кластерах так, що у кожному кластері спостереження є максимально схожими (близькими) один до одного і максимально відмінними (далекими) від спостережень з решти кластерів.

Після цього можливо визначити середні кожної змінної у одному кластері, другому, третьому тощо. По деяких змінних середні можуть суттєво відрізнятися по кластерах, по інших – майже не відрізнятися. Якщо середні змінних відрізняються достатньою мірою, це означає, що програма успішно кластеризувала спостереження на неоднорідні кластери (групи). Значущість відмінностей у кластерах визначається за допомогою F -критерію. Якщо середні змінних по кластерах мало відрізняються (а F -критерій при цьому швидше за все сигналізуватиме про незначущість кластеризації), це означає, що принципових відмінностей між спостереженнями немає.

Кластерний аналіз може бути попереднім етапом для проведення подальшого аналізу. Наприклад, якщо ми хочемо побудувати не одну регресійну модель для всіх спостережень, а декілька – по одній для кожного кластера. Наприклад, ми хочемо розрахувати регресійну модель залежності внутрішніх цін на бензин від світових цін на нафту, ступеню монополізації ринку нафтопродуктів та частки імпорту у споживанні бензину. Перед цим можливо кластеризувати всі спостереження на основі декількох інших змінних: наприклад, ВВП на душу населення і приріст ВВП (кластеризацію варто проводити на основі змінних, які, як передбачається, впливають на вплив факторів на залежну змінну). Припустимо, ми поділили всі спостереження на 4 кластери:

- кластер 1 – країни з високим або середнім ВВП на душу населення в роки з економічним зростанням;
- кластер 2 – країни з високим або середнім ВВП на душу населення в роки з рецесією або стагнацією;
- кластер 3 – країни з низьким або середнім ВВП на душу населення в роки з швидким економічним зростанням
- кластер 4 – країни з низьким ВВП на душу населення в роки з рецесією або повільним економічним зростанням.

Далі для кожного кластеру будується окрема регресійна модель, або приналежність до відповідного кластеру вводиться в регресійну модель як бінарна псевдозмінна.

Звичайно, можливо було б безпосередньо ввести в модель як фактори ВВП на душу населення та приріст ВВП. Але тоді потрібно було б вводити багато добутків факторів для урахування ефектів взаємодії, що б вимагало більшої кількості спостережень для надійності результатів в умовах, коли часто кількість доступних спостережень обмежена. Особливо це було би проблемою, якщо замість двох змінних (ВВП на душу населення та приріст ВВП) таких змінних було би багато.

ДОДАТКИ

Додаток А.

Джерела міжнародної економічної статистики

В цьому додатку представлений широкий спектр джерел міжнародної статистики, умовно згрупованих за сферами. Представлені джерела, які надають переважно порівнювану інформацію по різних країнах. В додатку описані переважно безкоштовні бази даних. Але є і джерела, доступ до яких обмежений. На початку розглянуті детальніше декілька основних джерел.

Звичайно, в посібнику охоплені не всі джерела. Корисною також може бути інформація статистичних органів окремих країн, центральних банків, інтеграційних об'єднань, недержавних установ, які не охоплені в цій книзі.

Під час перегляду статистичної інформації важливо розуміти точні визначення показників. Наприклад, якщо йдеться про ВВП, потрібно розуміти за який період, у якій валюті, який спосіб перерахунку, в яких цінах тощо. Інший приклад: якщо йдеться про короткостроковий борг, потрібно розуміти чи включає він частину довгострокового боргу, до погашення якої залишається менше року.

Більшість джерел статистичної інформації містять інформацію про показники та методику їх розрахунку, наприклад у розділі Metadata. Іноді опції щодо способу виміру показника обираються автором у процесі формування запиту. Іншим варіантом є деталізація інформації в окремих частинах електронної таблиці (аркуші, рядки або стовпчики) або у додатках до текстового файлу.

А.1. Формат даних

Доступ до даних може бути організований різними способами:

- доступні для скачування файли з електронними таблицями у форматі xls,xlsx, csv тощо.
- доступні для скачування текстові файли формату pdf;
- онлайн таблиці;
- онлайн таблиці, які формуються на основі багатокрокового запиту, де користувач вказує параметри даних, які він бажає одержати, результати в такому випадку іноді можливо скачати у формі електронних таблиць;
- карти або діаграми та графіки – деякі сайти мають інструменти для візуалізації статистичних даних.

Якщо електронні таблиці зберігаються у іншому форматі ніж xls або xlsx – наприклад csv, вони можуть не відкриватися належним чином у Microsoft Excel. Універсальним вирішенням цієї проблеми є такі дії. По-перше, змініть розширення файлу на розширення txt у файловому менеджері. Далі у Microsoft Office Excel в меню *Файл* оберіть *Открить* і знайдіть у діалоговому вікні потрібний файл (перед цим у діалоговому вікні в полі *Тип файла* оберіть *Все файлы*, щоб побачити файли всіх форматів) і відкрийте його. У новому діалоговому вікні (на Рис. А.1.) в полі *Укажите формат данных* оберіть опцію *с разделителями* (в більшості випадків) і натисніть кнопку *Далее*.

Оберіть у полі *символом-разделителем является* опцію *запятая*. Внизу вікна текст та цифри мають бути тепер правильно розподілені по стовпчиках, як зображено на Рис.А.2. Натисніть *Готово* і таблиця буде відображатися правильним чином. За бажанням тепер можливо файл зберегти у форматі xls або xlsx.

Якщо після вказаних дій таблиця не відображається правильним чином, спробуйте інші опції (*фиксированной ширины* або інші опції в полі *символом разделителем является*).

А.2. Статистика Світового банку. Світові індикатори розвитку і Фінанси глобального розвитку

Світовий банк на своєму сайті дає можливість скачати велику кількість міжнародних економічних статистичних даних. Для цього скористайтесь сторінкою <http://data.worldbank.org/data-catalog>. Деякі дані можливо одержати через сторінку <http://econ.worldbank.org>

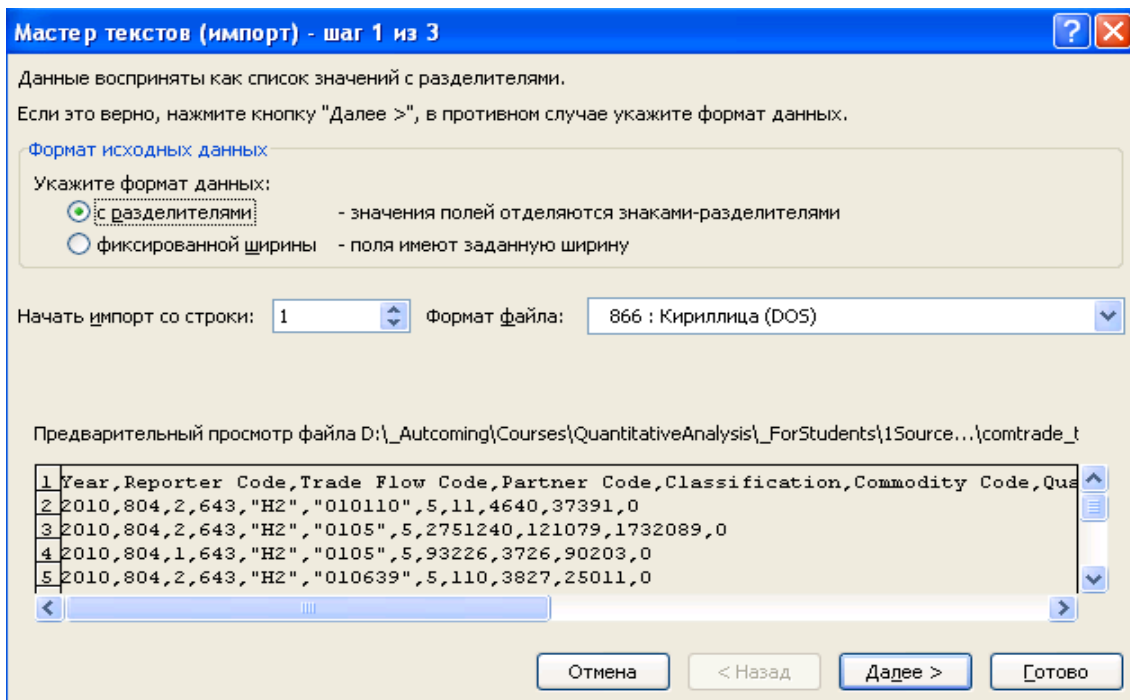


Рис.А.1.

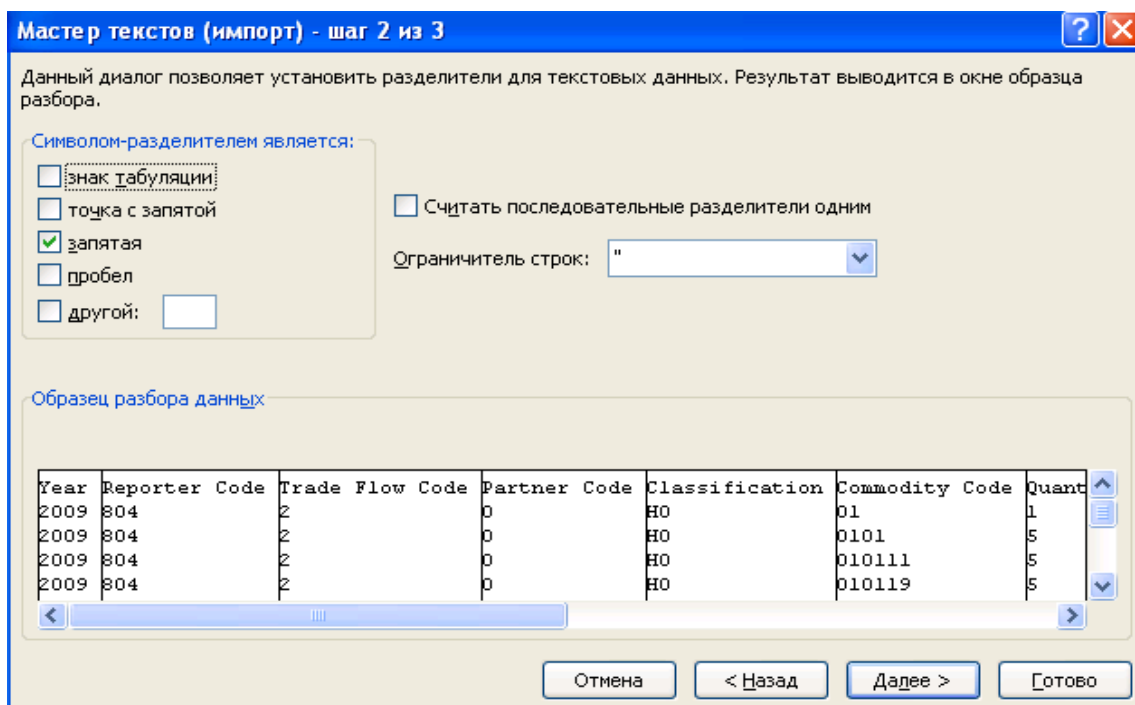


Рис.А.2.

Окрему категорію статистичних баз даних становлять дослідницькі бази даних та аналітичні інструменти – Research Datasets and Analytical Tools (<http://go.worldbank.org/D353HLMTO0>), які створюються під потреби конкретного дослідження.

Вони лише створюються один раз і, зазвичай, не оновлюються. Тим не менш вони також можуть бути корисними, так як можуть містити інформацію, яку важко або неможливо знайти в інших джерелах. Приклад: Міграційна база даних з гендерним розподілом – Migration Database with Gender Breakdown, 1990 – 2000 (<http://go.worldbank.org/FCU1VDGOA0>).

На сторінці (<http://data.worldbank.org>) доступні й інструменти візуалізації даних. Для цього оберіть пункт *Indicators*. Далі виберіть конкретний показник. З'явиться таблиця з даними по цьому показнику. Натисніть кнопку *Map* або *Graph*. Наприклад, для експорту товарів та послуг у % ВВП, одержимо карту на Рис А.3 та графік на Рис.А.4. На карті можливо змінювати масштаб.

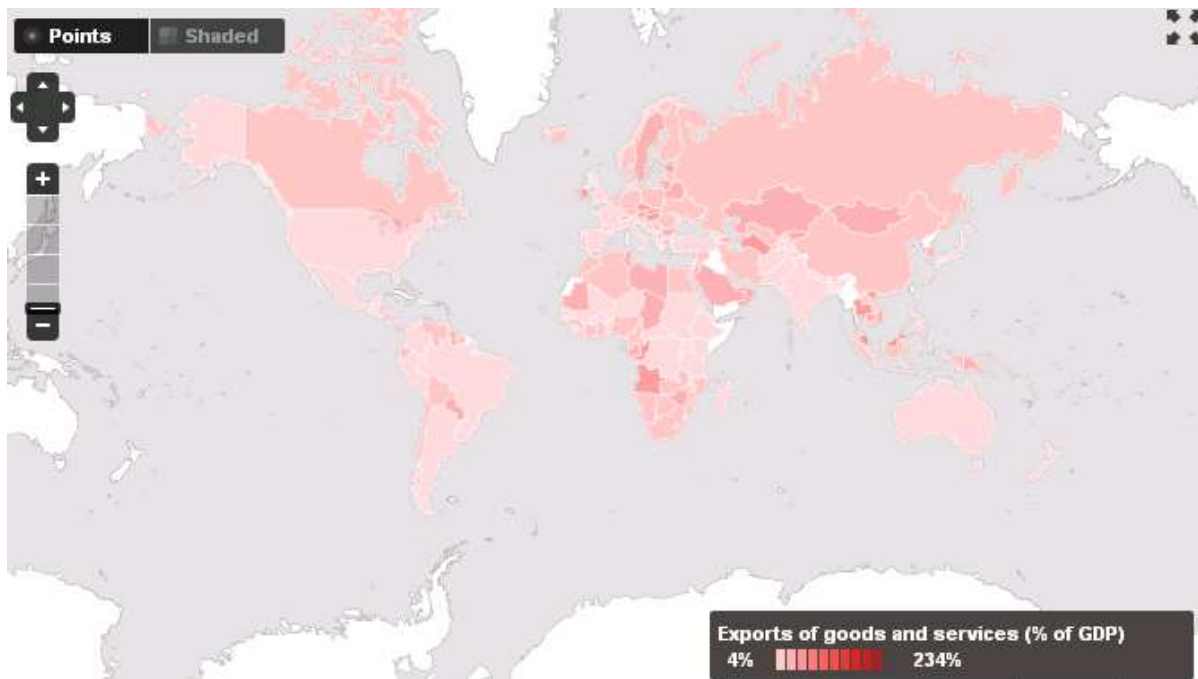


Рис.А.3



Рис.А.4.

Ряд статистичних баз даних Світового банку будуть розглянуті в інших підрозділах цієї глави, а в цьому підрозділі ми сконцентруємося на основній базі даних, яка є фактично об'єднанням двох баз даних Світові індикатори розвитку і Фінанси глобального розвитку – World Development Indicators and Global Development Finance (<http://data.worldbank.org/data-catalog/world-development-indicators>). Файл доступний для скачування в форматі *xlsx* або *csv*, які є зручними для проведення подальших розрахунків з цими даними.

Як і більшість баз даних Світового банку ця база даних містить річні дані, починаючи з 1960 р. У ній присутні більше 1150 показників (*Series*) для майже всіх країн світу. Деякі з них є узагальненими показниками з інших баз даних Світового банку чи Міжнародного валютного фонду.

Пріоритет надається одиницям, а не тисячам, мільйонам чи мільярдам. Хоча це незручно візуально (числа типу 2354168908 або 2.35E09), з такими одиницями виміру зручніше проводити подальші розрахунки, не турбуючись про те, що, наприклад, перед тим як поділити мільярди на мільйони, мільярди потрібно перетворити на мільйони, помноживши на 1000. Файл бази даних містить також пояснення щодо країн, показників та примітки.

Показники в цій базі даних можливо класифікувати за такими групами (див. Табл. А.1.):

Таблиця А.1.

Economic Policy & Debt – економічна політика і борг	Balance of payments – платіжний баланс	Current account – поточний рахунок	Balances – баланси
			Goods, services & income – товари, послуги та доходи
			Transfers – Поточні трансферти
		Capital & financial account – рахунок операцій з капіталом і фінансовий рахунок	
		Reserves & other items – резервні активи та інші статті	
	External debt – зовнішній борг	Amortization – амортизація – виплата основної частини боргу	
		Arrears, reschedulings, etc. – прострочені платежі та відстрочений борг	
		Commitments – зобов’язання кредиторів надати позику	
		Currency composition – структура боргу за валютою	
		Debt outstanding – величина боргу	
		Debt ratios & other items – коефіцієнти боргу	
		Debt service – обслуговування боргу – виплата основної частини боргу та відсотків	
		Disbursements – одержання позик	
		Interest – процентні платежі	
		Net flows – чисті потоки по позиках (одержання мінус виплати основної частини боргу)	
		Net transfers – чисті трансферти по позиках (одержання мінус виплати основної частини боргу і процентів)	
		Terms – строки	
		Undisbursed debt – невикористаний борг (виділені, але ще невикористані суми позики)	
	National accounts – національні рахунки – ВВП та його складові	Atlas GNI & GNI per capita – валовий національний дохід (в цілому і на душу населення)	
		Growth rates – темпи зростання	
		Local currency at constant prices – у національній валюті в постійних цінах	Aggregate indicators – узагальнені показники
			Expenditure on GDP – ВВП за витратами
			Other items – інші статті
		Local currency at current prices – те саме у національній валюті в поточних цінах	
		US\$ at constant 2000 prices – те саме у постійних цінах у доларах у цінах 2000 року	
		US\$ at current prices – те саме у доларах у поточних цінах	
		Shares of GDP & other – частки ВВП	
Official development assistance – офіційна допомога у розвитку			
Purchasing power parity – паритет купівельної спроможності			
Financial Sector – фінансовий сектор	Assets – активи		
	Capital markets – ринки капіталу		
	Exchange rates & prices – валютні курси та ціни		
	Interest rates – процентні ставки		
	Monetary holdings (liabilities) – монетарні активи та зобов’язання		

Private Sector & Trade – приватний сектор і торгівля	Business environment – регуляторне середовище для бізнесу	
	Exports – експорт	
	Imports – імпорт	
	Private infrastructure investment – приватні інвестиції в інфраструктуру	
	Tariffs – тарифи	
	Total merchandise trade – торгівля товарами	
	Trade facilitation – стимулювання торгівлі	
	Trade indexes – торговельні індекси	
	Travel & tourism – подорожі та туризм	
Public Sector – державний сектор	Defense & arms trade – оборона і торгівля зброєю	
	Government finance – Державні фінанси	Deficit & financing – дефіцит бюджету та його фінансування
		Expense – видатки
		Revenue – доходи
Policy & institutions – політика та інституції		
Labor & Social Protection – праця і соціальний захист	Economic activity – економічна активність	
	Labor force structure – структура робочої сили	
	Migration – міграція	
	Unemployment – безробіття	
Poverty – бідність	Conflict & fragility – конфлікти та слабкість	
	Income distribution – розподіл доходу	
	Poverty rates – коефіцієнти бідності	
Infrastructure – інфраструктура	Communications – комунікації	
	Technology – технологія	
	Transportation – транспорт	
Education – освіта	Efficiency – ефективність	
	Inputs – показники на вході	
	Outcomes – результати	
	Participation – участь	
Environment – навколишнє середовище	Adjusted savings – заощадження з поправкою на зміни в навколишньому середовищі та витрати на освіту	
	Agricultural production – сільськогосподарське виробництво	
	Biodiversity & protected areas – біорізноманіття і заповідники	
Environment – навколишнє середовище	Density & urbanization – щільність розселення і урбанізація	
	Emissions – викиди газів	
	Energy production & use – виробництво і споживання енергії	
	Freshwater – прісна вода	
	Land use – використання землі	
	Water pollution – забруднення води	

Health – охорона здоров'я	Disease prevention – попередження хвороб	
	Health services – медичні послуги	
	Mortality – смертність	
	Nutrition – харчування	
	Population – населення	Dynamics – динаміка
		Structure – структура
	Reproductive health – репродуктивне здоров'я	
Risk factors – фактори ризику		
Social Development – соціальний розвиток		

На Рис. А.5. покажемо приклад фрагменту цієї бази даних.

SeriesCode	Series Name	Country Code	Country Name	2007	2008	2009
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	ADO	Andorra			
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	AFG	Afghanistan			0.137977
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	AGO	Angola			
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	ALB	Albania	13.55953	32.63314	9.557746
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	ARB	Arab World	1.329666	1.298855	
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	ARE	United Arab Emi	1.248819	0.941098	
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	ARG	Argentina	3.856929	3.235335	3.909664
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	ARM	Armenia	31.62122	28.63984	46.54003
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	ATG	Antigua and Bar	0.547547		
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	AUS	Australia	27.34319	26.72057	27.42659
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	AUT	Austria	3.462063	3.485933	3.156185
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	AZE	Azerbaijan	1.765431	0.288224	0.24995
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	BDI	Burundi	3.124748	9.250727	4.832546
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	BEL	Belgium	3.709195	3.367288	2.755986
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	BEN	Benin			
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	BFA	Burkina Faso			0.64209
TX.VAL.MMTL.ZS.UN	Ores and metals exports (% of merchandise exports)	BGD	Bangladesh	0.427076		

Рис.А.5.

Наведемо деякі корисні пояснення щодо термінології та скорочень:

Net – чистий (надходження мінус витрати, або активи мінус зобов'язання)

% of GDP – у % від ВВП

BoP – платіжний баланс

Trade in services – торгівля послугами (експорт + імпорт)

% of service imports – у % від імпорту послуг

Payments – виплати нерезидентам

Receipts – надходження від нерезидентів

Workers' remittances – грошові перекази працюючих

Current transfers – поточні трансферти

Outflows – відплив

Inflows – надходження

Foreign direct investment – прямі іноземні інвестиції

Profit remittances – перекази прибутків

Portfolio investment – портфельні інвестиції

LCFAR – валютні резерви інших держав

Changes in net reserves – зміна валютних резервів (мінус означає збільшення валютних резервів країни)

Errors and omissions – помилки і упушення

Total reserves – валютні резерви

Stocks – станом на певну дату, наприклад, прямі інвестицій, накопичені до вказаного моменту за всі

попередні роки

Flow – за період, наприклад прямі інвестиції, які надійшли протягом року

External debt stocks, total – загальний зовнішній борг (короткостроковий + державний + гарантований державою + приватний негарантований + використання кредитів МВФ)

Public and publicly guaranteed (PPG) – державний або гарантований державою

Private nonguaranteed (PNG) – приватний негарантований борг

variable rate – зі змінною процентною ставкою

External debt stocks, short-term – короткостроковий зовнішній борг, який має строк менше року з самого початку. Не включає частину довгострокового боргу, яка має погашатися протягом року, тому дані щодо короткострокового боргу є фактично заниженими

Multilateral – перед міжнародними організаціями (борг)

Concessional – на пільгових умовах

IBRD – Міжнародний банк реконструкції та розвитку

IDA – Міжнародна асоціація розвитку

Use of IMF credit – використання кредитів МВФ

Present value – теперішня вартість

AMT – амортизація боргу

IMF repurchases – повернення кредитів МВФ

Principal arrears – вчасно не виплачена сума основної частини боргу

Principal forgiven – списана основна частина боргу

Principal rescheduled – реструктурована частина основної частини боргу, строк виплати якої відстро-
чено

Debt forgiveness or reduction – списання боргу плюс викуп боргу позичальником

Interest arrears – акумульовані невикладені процентні платежі за боргом

Debt stock rescheduled – реструктурована частина боргу

Interest forgiven – процентні платежі списані

Interest rescheduled (capitalized) – капіталізація – означає, що несплачені процентні платежі стають
частиною основної частини боргу

Cross-currency valuation – зміни, обумовлені зміною курсу валюти деномінації боргу до долара

SDR – спеціальні права запозичення (СПЗ або СДР)

GNI – валовий національний дохід

IMF repurchases and charges – повернення кредитів МВФ і процентів

IFC – Міжнародна фінансова корпорація

EBRD – Європейський банк реконструкції та розвитку

Average grace period – середній пільговий період

Average grant element – середня грантова складова позик у %

Average maturity – середній строк нових позик (пільговий період + період виплат)

Atlas method – метод Атлас передбачає перерахунок у долари не за поточним курсом, а за середнім
курсом останнім часом

per capita – на душу населення

growth (annual %) – приріст у % на рік

General government – загальний уряд (включає не тільки центральний уряд)

Fixed capital – основні засоби

Gross capital formation – інвестиції в основні засоби та збільшення запасів

Manufacturing – обробна промисловість

Industry – промисловість у цілому

LCU – в одиницях національної валюти

Constant – у постійних цінах

Gross domestic savings – валові внутрішні збереження

Net taxes on products – чисті податки на продукцію (за мінусом субсидій)

Changes in inventories – зміна запасів

Exports as a capacity to import – експорт як здатність імпортувати = експорт в поточних цінах, зменшений на зростання імпорتنих цін

Terms of trade adjustment – поправка на умови торгівлі = здатність імпортувати мінус експорт у постійних цінах

Current – у поточних цінах

Trade – сума експорту та імпорту (зовнішньоторговельний оборот)

2000 US\$ – у доларах у цінах 2000 року

External balance on goods and services – зовнішньоторговельний баланс (експорт мінус імпорт)

Net bilateral aid flows from DAC donors – чиста двостороння допомога від донорів офіційної допомоги у розвитку (за мінусом повернення основної частини позик)

Net official flows from UN agencies – чисті офіційні надходження від окремих структур ООН

PPP – у перерахунку за коефіцієнтом паритету купівельної спроможності (ПКС)

PPP conversion factor – коефіцієнт ПКС (співвідношенням цін у даній країні та США)

PPP conversion factor (GDP) to market exchange rate ratio – відношення коефіцієнту ПКС до валютного курсу

Foreign assets – закордонні активи

Claims on governments and other public entities – чисті вимоги до державного сектору (за мінусом депозитів державного сектору)

Annual growth as % of M2 – річне зростання у % від грошового агрегату M2

S&P Global Equity Indices – ціни на акції в доларах

Market capitalization of listed companies – ринкова капіталізація (кількість акції всіх компаній (крім інституціональних інвесторів) на біржах помножена на їх ціну)

Stocks traded – обсяги торгівлі акціями

Stocks traded, turnover ratio – відношення обсягів торгівлі акціями до ринкової капіталізації

Consumer price index (2005 = 100) – індекс споживчих цін (базовий рік 2005)

Period average – середній за період

Real interest rate – реальна процентна ставка по кредитах (номінальна процентна ставка мінус інфляція за дефлятором ВВП)

Risk premium on lending (prime rate minus treasury bill rate, %) – премія за ризик (ставка по кредитах першокласним позичальниками мінус ставка по короткостроковим облігаціям уряду)

Money – гроші – агрегат M1

Money and quasi money – гроші та квазігроші – агрегат M2

Liquid liabilities – ліквідні зобов'язання – агрегат M3

Age dependency ratio (% of working-age population) – коефіцієнт вікової залежності (людей до 15 і з 64 років до працюючого населення)

High-technology exports – високотехнологічний експорт (аерокосмічний, фармацевтичний, наукових приладів, електротехніки, комп'ютерів)

Burden of customs procedure, WEF (1=extremely inefficient to 7=extremely efficient) – легкість митних процедур (7 балів – найлегша)

Freight – вантажні перевезення

International migrant stock – кількість іммігрантів у країні (які народилися за кордоном, для СРСР – в іншій республіці)

Income share held by lowest 10% – частка доходу або споживання, яке належить найбіднішим 10% населення

GINI index – індекс джині (0 – повна рівність у доходах населення, 100 – повна нерівність)

Average time to clear exports through customs (days) – середня кількість днів для розмитнення експорту

Cost to export (US\$ per container) – вартість експортних митних процедур для 1 контейнера (за винятком мита та податків на торгівлю)

Merchandise exports – експорт товарів

Bound rate, simple mean, all products – максимальні ставки митного тарифу, погоджені внаслідок переговорів – середня незважена по всіх товарах

Share of tariff lines with international peaks – частка продуктових ліній, за якими ставка митного тарифу більше 15 %

Tariff rate, applied, simple mean – середня незважена ставок митного тарифу, які фактично застосовуються

Most favored nation – для країн, які користуються режимом найбільшого сприяння

Weighted mean – зважена середня

Documents to export (number) – кількість документів, необхідних для відправки товарів на експорт

Logistics performance index: Overall (1=low to 5=high) – індекс ефективності логістики: загальний (1 – низький, 5 – високий)

Net barter terms of trade index (2000 = 100) – умови торгівлі (відношення експортних цін до імпортних – відносно базового 2000 р.)

International tourism, expenditures – імпорт туристичних послуг

International tourism, receipts – експорт туристичних послуг

Arms imports – імпорт зброї

Cash surplus/deficit – профіцит (плюс) або дефіцит (мінус) державного бюджету

Highest marginal tax rate, individual rate – найвища гранична ставка податку на дохід фізичних осіб

Taxes on international trade – податки на міжнародну торгівлю (експортні та імпортні мита, прибутки експортних та імпортних монополій, валютні прибутки та податки)

А.3. Міжнародна фінансова статистика Міжнародного валютного фонду

Міжнародний валютний фонд – International Monetary Fund (<http://www.imf.org/external/data.htm>) надає можливість знайти широкий спектр показників по країнах, переважно в фінансовій сфері. На вказаній вище сторінці є посилання на декілька баз даних. Ми розглянемо більшість з них у подальших підрозділах цієї глави, а зараз зосередимо увагу на базі даних Міжнародна фінансова статистика – International Financial Statistics (IFS) (<http://elibrary-data.imf.org/FindDataReports.aspx?d=33061&e=169393>). У базі даних показники представлені у розрізі окремих країн. Присутні річні, квартальні та місячні дані. Для окремих показників є загальносвітові таблиці.

Доступ до неї обмежений – платний. Альтернативами платному онлайн доступу є такі:

- користування обмеженими можливостями в рамках вільного доступу або 5-денним пробним періодом;
- для обмеженої кількості показників з IFS також існує вільний доступ через портал UNData (<http://data.un.org/Default.aspx>);
- декілька показників з IFS присутні в World Development Indicators and Global Development Finance;
- користування друкованими версіями IFS, хоча це менш зручно для аналізу;
- відвідування установ, які мають доступ до цієї бази даних.

Наведемо також деякі корисні пояснення щодо термінології та скорочень (перелік показників може варіюватися залежно від країни) у Табл.А.2.

А.4. Статистика Світової організації торгівлі

Світова організація торгівлі – World Trade Organization (http://www.wto.org/english/res_e/statis_e/statis_e.htm) має декілька джерел статистичної інформації, зокрема:

1. Статистична база даних – Statistics database (<http://stat.wto.org/Home/WSDBHome.aspx?Language=E>) включає складові, вказані у Табл. А.3.

2. Міжнародна торговельна статистика – International Trade Statistics 2010 (http://www.wto.org/english/res_e/statis_e/its_e.htm) – є оглядовим виданням, для якого доступні ряд графіків, таблиць, карт по торгівлі товарами та послугами.

3. Детальні дані по тарифах – Comprehensive tariff data (<http://tariffdata.wto.org/ReportersAndProducts.aspx>) дає можливість побачити статистику по тарифах для кожної країни-члена СОТ з деталізацією по товарах.

Окрім власне статистичних даних наявні інформаційні системи іншого типу. Наприклад, Інформаційна система регіональних торговельних угод – Regional Trade Agreements Information System (<http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx>) або інформація по торговельних спорах – Dispute settlement (http://www.wto.org/english/tratop_e/dispu_e/dispu_e.htm).

Таблиця А.2.

Exchange Rates – офіційний або ринковий валютний курс, валютні курси на кінець періоду та середній за період, до долара та до СДР (СПЗ), окрім звичайних курсів вказується ефективний валютний курс, номінальний та реальний, для країн зони євро окремо подані курси колишніх національних валют та євро
Fund Position – позиція країни в МВФ, включаючи квоту, СДР, резервну позицію, непогашені кредити МВФ
International Liquidity – дані про валютні резерви в розрізі складових, подані також інші закордонні активи та зобов'язання центрального банку та інших банків
Monetary Authorities (за старою класифікацією) або Central Bank (за новою класифікацією) – структура активів та зобов'язань центрального банку або іншого органу, який виконує його функції Foreign Assets – закордонні активи Claims on General Government – внутрішні вимоги до загального уряду Claims on Nonfin.Pub.Enterprises – внутрішні вимоги до нефінансових державних підприємств Claims on Private Sector – внутрішні вимоги до приватного сектору Claims on Banks – внутрішні вимоги до банків Reserve Money – грошова база of which: Currency Outside Banks – з них готівка поза банками Time, Savings, & Fgn.Currency Dep. – строкові, ощадні та валютні депозити Foreign Liabilities – зобов'язання перед нерезидентами General Government Deposits – депозити загального уряду Capital Accounts – капітал Other Items (Net) – інші статті (на нетто основі)
Deposit Money Banks або Depository Corporations – активи та зобов'язання комерційних банків і подібних установ, які приймають гроші на депозити до запитання; в цьому розділі за деякими винятками схожа структура, як і у розділі Central Bank, наприклад, Reserves – банківські резерви (готівка і депозити в центральному банку) Demand Deposits – депозити до запитання
Other Banking Institutions або Other Financial Corporations – активи та зобов'язання інших банків та подібних фінансових установ, аналогічно по складових
Banking Survey або Financial Corporations – консолідовані дані по активах і зобов'язаннях всієї банківської системи (включаючи центральний банк, депозитні корпорації, та інші фінансові корпорації), аналогічно по складових, зокрема, представлені дані й по грошових агрегатах
Interest Rates – процентні ставки Refinancing Rate – ставка рефінансування комерційних банків центральним банком Money Market Rate – ставка грошового (міжбанківського) ринку в національній та іноземній валюті представлені також депозитні та кредитні ставки
Prices, Production, Labor – індекси цін, промислове виробництво, показники ринку праці
Industrial Employment – зайнятість у промисловості
International Transactions – дані по експорту (ціни FOB) та імпорту (ціни FOB – f.o.b. та CIF – c.i.f.) за митною статистикою – можуть відрізнятися від даних по зовнішній торгівлі за платіжним балансом
Balance of Payments – структура платіжного балансу по статтях, наприклад, Trade Balance – торговельний баланс (тут лише по товарах) Balance on Goods & Services – торговельний баланс (по товарах та послугах)
Credit – надходження за відповідною статтею (доходів від нерезидентів, поточних чи капітальних трансфертів) Debit – виплати за відповідною статтею Dir. Invest. in Rep. Econ. – прямі інвестиції в країну, яка розглядається Assets – закордонні активи по відповідній статті (плюс означає зменшення активів (приплив капіталу), мінус – збільшення (відплив капіталу)) Liab. – зобов'язання по відповідній статті (плюс означає – збільшення зобов'язань (приплив капіталу), мінус – зменшення (відплив капіталу)) Reserves and Related Items – Резерви та подібні активи (валютні резерви, використання кредитів МВФ, надзвичайне фінансування)
International Investment Position – міжнародна інвестиційна позиція: закордонні активи резидентів та їх зобов'язання перед нерезидентами в розрізі видів інвестицій

Government Finance – державні фінанси (баланс, доходи, витрати, зовнішнє та внутрішнє фінансування дефіциту бюджету, державний борг)
National Accounts – національні рахунки – складові ВВП
Population – населення
Для країн єврозони є особливості, наприклад, Banking Survey або Financial Corporations подані в двох варіантах залежно від концепції резидента (лише цієї країни або всієї єврозони): Nat'l Residency – закордонними активами вважаються вимоги до нерезидентів цієї країни, зовнішніми зобов'язаннями вважаються зобов'язання перед нерезидентами цієї країни EA-Wide Residency – закордонними активами вважаються вимоги до нерезидентів єврозони, зовнішніми зобов'язаннями вважаються зобов'язання перед нерезидентами єврозони

Таблиця А.3.

Trade Profiles – Торговельні профілі – інформація по основних показниках торгівлі, її регулювання, структури торгівлі товарами та послугами, інтелектуальній власності
Tariff Profiles – Тарифні профілі: - основні дані по митних тарифах у країні та розподіл ставок імпорتنних тарифів за їх величиною (Tariffs and imports: Summary and duty ranges); - розподіл ставок імпорتنних тарифів та імпорту по групах товарів (Tariffs and imports by product groups); - розподіл експорту та ставок імпорتنних тарифів, з якими стикається експорт країни по країнах-контрагентах (Exports to major trading partners and duties faced). Основні позначення: Ag – для сільського господарства Non-Ag – для секторів крім сільського господарства Final bound – зв'язаний (рівень, який країна зобов'язалася не перевищувати) MFN applied – який застосовується в рамках режиму найбільшого сприяння (Most-Favoured-Nation) AVG – середній Binding in % – частка зв'язаних тарифів Max – максимальний Duty-free in % – частка без мита
Services Profiles – Профілі послуг – загальна інформація по послугах та деталізація лише для таких видів послуг: - транспортних; - телекомунікаційних; - фінансів та страхування. Деталізація охоплює приплив інвестицій, показники ринку, виробництво і зайнятість, зовнішню торгівлю, інвестиції закордон
Попередні модулі дають можливість по кожній країні дізнатися ситуацію станом на певний рік. Існує й інший модуль Time Series – Часові ряди, який дає можливість побачити ряд показників торгівлі та її структури за декілька років.

А.5. Статистичні органи України

Основним органом в цій сфері є Державна служба статистики України (<http://www.ukrstat.gov.ua/>). Для перегляду варто користуватися браузером Internet Explorer, інакше кнопки меню можуть не спрацьовувати.

В розділі *Статистична інформація* доступні дані по таких розділах:

- Макроекономічні показники;
- Національні рахунки;
- Промисловість;
- Сільське господарство;
- Рибне господарство;
- Інвестиції та будівельна діяльність;
- Житловий фонд;
- Наука та інновації;

- Транспорт і зв'язок;
- Структурні зміни в економіці;
- Зовнішньоекономічна діяльність (включаючи географічну та товарну структуру експорту та імпорту товарів або послуг, географічну структуру прямих інвестицій);
- Споживчий ринок;
- Ціни і тарифи;
- Фінанси;
- Витрати та ресурси домогосподарств;
- Доходи населення;
- Ринок праці;
- Демографічна ситуація;
- Охорона здоров'я;
- Соціальне забезпечення;
- Освіта;
- Туризм (включаючи міжнародний);
- Культура;
- Навколишнє середовище;
- Правопорушення;
- Соціально-економічне становище регіонів України;
- Основні соціально-економічні показники країн СНД (щомісячна інформація).

В розділі *Новини* можливо знайти найбільш останні статистичні дані у вигляді таблиць або текстових файлів. Також детальні статистичні дані представлені в окремих статистичних виданнях.

Іншим основним джерелом статистичної інформації є Національний банк України (<http://www.bank.gov.ua>). У розділі *Статистика* доступні:

- Фінансовий сектор (грошово-кредитна статистика і фінансові ринки);
- Зовнішній сектор (валютний курс, платіжний баланс, міжнародна інвестиційна позиція, зовнішній борг);
- Реальний сектор;
- Фінансові рахунки (компоненти балансових звітів депозитних корпорацій (Національного банку України та інших депозитних корпорацій));
- Статистичний бюлетень (основні соціально-економічні показники, а також детальніше показники грошово-кредитної сфери);
- Додаток до статистичного бюлетеня (по кредитах, депозитах, процентних ставках у регіональному розрізі);
- Статистичні випуски;
- Індикатори фінансової стійкості.

Дані про державний бюджет та державний борг вказані на сайті Міністерства фінансів України (<http://www.minfin.gov.ua/control/uk/index>) у розділах *Усе про бюджет* та *Державний борг та гарантований державою борг*.

А.6. Основні джерела комплексної статистики

Система доступу до даних баз даних ООН UNData (<http://data.un.org/Default.aspx>) дає можливість доступу до принаймні частини показників з більше 30 баз даних. Оберіть опцію *Databases*, якщо бажаєте продивитися доступні бази даних. Оберіть країну в полі *Country data services*, якщо бажаєте перейти на сайт статистичного органу відповідної країни.

Euromonitor International Passport GMID (<http://www.euromonitor.com/passport-gmid>) є джерелом комплексної інформації по країнах, компаніях та ринках. Доступ є платним. Ця база даних є більш зорієнтованою на аналітику для бізнесу в сфері маркетингу, а не тільки макроекономічний аналіз. Окрім звичайних макроекономічних даних має інформацію по таких показниках як:

- витрати споживачів у розрізі груп товарів або послуг;
- індекси цін по окремих групах товарів;
- ціни на конкретні товари;
- структура ВВП більш деталізовано по галузях;

- детальна статистика по споживанню, виробництву, запасам енергії по видах палива чи джерел електроенергії;
- детальні екологічні показники (наприклад, обсяги використання вторинної сировини з алюмінію, групи тварин під загрозою зникнення тощо);
- експорт та імпорт у розрізі країн-контрагентів і груп товарів;
- детальні показники охорони здоров'я;
- види домогосподарств (наприклад, з власним житлом, кредитним, орендованим, з кухнею тощо);
- популяція домашніх тварин;
- кількість людей, які мають ті чи інші товари (наприклад комп'ютер, піаніно, кондиціонер, мікрохвильова піч тощо)
- показники розподілу доходу, такі як доходи людей у розрізі вікових груп, кількість населення з відповідним доходом (наприклад від 10 до 20 тис доларів);
- виробництво конкретних видів товарів;
- показники галузі розваг (наприклад, відвідування кінотеатрів, оборот прокату DVD дисків, кількість музеїв за видами тощо);
- тренди стосовно споживачів (наприклад, індивідуалізація; сегментація споживачів за різноманітними параметрами: студенти, молоді професіонали, батьки, геймери, туристи тощо);
- демографічні показники;
- показники зайнятості;
- показники освіти;
- показники транспорту (наприклад, ціни на паливо, новозареєстровані двоколісні транспортні засоби, кількість човнів за видами, автобани тощо).

Світові економічні перспективи МВФ – IMF World Economic Outlook (<http://www.imf.org/external/ns/cs.aspx?id=28>) надає історичні, поточні та прогнозні дані по ряду показників у розрізі країн та груп країн. Показники включають ВВП за різними способами вимірювання, інфляцію, доходи та витрати державного бюджету, поточний рахунок платіжного балансу та деякі інші показники залежно від країни.

Конференція ООН з торгівлі та розвитку ЮНКТАД – UNCTAD (http://unctadstat.unctad.org/ReportFolders/reportFolders.aspx?sCS_referer=&sCS_ChosenLang=en) надає інформацію по різних сферах, таких як міжнародна торгівля товарами та послугами, зовнішні фінансові ресурси, ВВП, населення, інформаційна економіка, креативна економіка, морський транспорт тощо.

Глобальний економічний монітор – World Bank Global Economic Monitor (<http://data.worldbank.org/data-catalog/global-economic-monitor>) надає денні, місячні або кварталні дані по обмеженому колу показників.

Світова книга фактів ЦРУ – CIA The World Factbook (<https://www.cia.gov/library/publications/the-world-factbook/index.html>) дає можливість одержати статистичну та фактологічну інформацію по кожній країні в найрізноманітніших сферах (історія, люди, економіка, географія, комунікації, транспорт, збройні сили, транснаціональні проблеми). Містить як фактичні дані так і оціночні. Дає також можливість ранжування країн за конкретним показником.

Nation Master (<http://www.nationmaster.com>) містить інформацію по широкому колу статистичних показників та фактологічної інформації. Має орієнтацію на ілюстративність даних. Наприклад, дає можливість одержати дані по країні, порівнювати країни за різними показниками, продивлятися рейтинги країн за обраним показником, продивлятися ілюстративні карти.

Gap Minder (<http://www.gapminder.org/>) містить інформацію по близько 500 показниках країн у різних сферах з різних джерел. Надає широкі можливості для візуалізації, наприклад побудувати діаграму розсіювання, яка показує залежність імпорту (% ВВП) від ВВП на душу населення (див. Рис. А.6). За допомогою кнопки *Play* можливо прослідкувати як ця діаграма змінювалася у часі.

А.7. Комплексна статистика по групі країн

Євростат – Eurostat (<http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/themes>) є спеціалізованим статистичним органом Європейського Союзу. Дає можливість знайти велику кількість показників, але переважно лише по країнах ЄС.

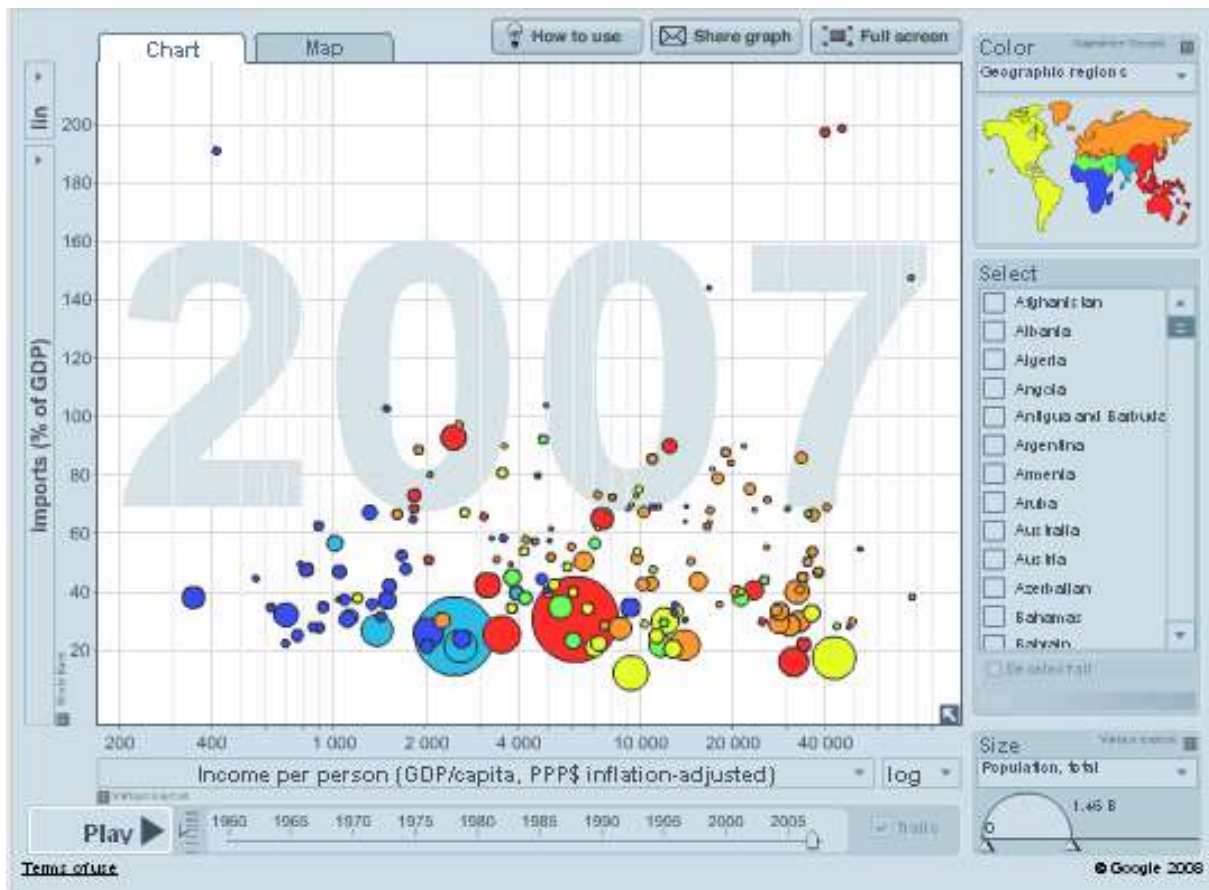


Рис.А.6.

Основна форма доступу – запит (*Query*). Для здійснення запиту оберіть опцію *Browse / Search database*, далі у ієрархічному дереві показників оберіть потрібний. Наприклад, *Economy and finance*, далі *Balance of payments – International transactions*, далі *Balance of payments statistics and International investment positions (bop_q)*, далі *Balance of payments by country*. У вікні, що з'явиться оберіть вкладку *Select Data* (див. Рис.А.7).

View Table Select Data Metadata Information Download Preview Boo

Balance of payments by country

Last update: 10-11-2011

Interactive extraction size limit: 500000
 Current extraction size: 700 Selection overview Update
 Dimension selection: 1/2

CURRENCY FLOW GEO PARTNER POST TIME

Filtering
 Filtering type: Code range Pattern
 Search in: Codes Labels Both
 Search Show all

Code	Label
<input checked="" type="checkbox"/> MIO_EUR	Millions of euro (from 1.1.1999)/Millions of ECU (up to 31.12.1998)
<input type="checkbox"/> MIO_NAC	Millions of national currency (including 'euro fixed' series for euro area countries)

CURRENCY: Millions of euro (from 1.1.1999)/Millions of ECU (up to 31.12.1998)
 POST: Current account
 FLOW: Net
 PARTNER: All countries of the world

GEO	TIME		
	2009Q2	2009Q3	2009Q4
Belgium	1.676	-3.861	-54
Bulgaria	-1.190	250	-902
Czech Republic	-1.827	-1.360	-570
Denmark	2.167	2.558	2.243
Germany (including...)	26.201	32.190	50.485
Estonia	120	233	209
Ireland	-1.458	-686	-14
Greece	-6.986	-3.543	-8.190
Spain	-10.439	-9.500	-14.546
France	-13.493	-1.276	-10.864
Italy	-6.281	-4.519	-3.580
Cyprus	: (e)	: (e)	: (e)
Latvia	648	412	486

Рис.А.7.

У вкладці *Currency* оберіть валюту (одиницю виміру), наприклад *Millions of euro*. У вкладці *Flow* оберіть вид потоку: по кредиту *Credit* (надходження з закордону), по дебету *Debit* (виплати за кордон) або чисті *Net* (надходження мінус виплати). Припустимо ми обираємо *Net*. У вкладці *Geo* оберіть країну або групу країн, наприклад, позначимо всі можливі країни.

У вкладці *Partner* оберіть країну або групу країн-партнерів, наприклад *World* (тобто платіжний баланс обраних країн по відношенню до всього світу, а не двосторонній). У вкладці *Post* оберіть конкретний вид показника, наприклад поточний рахунок (*Current account*). У вкладці *Time* оберіть періоди часу, наприклад, *2011Q2*, *2011Q1*, *2010* (тобто перші два квартали 2011 р. і в цілому 2010 р.). Натисніть кнопку *Update*. Таблиця, яка знаходиться праворуч буде трансформована відповідно до умов Вашого запиту (на Рис.А.8 зображено фрагмент).

CURRENCY: Millions of euro (from 1.1.1999)/Millions of ECU (up to 31.12.1998)
POST: Current account
FLOW: Net
PARTNER: All countries of the world

	TIME ▶	2010	2011Q1	2011Q2
GEO ▼				
Belgium		5.173	-1.563	6.700
Bulgaria		-476	147	97
Czech Republic		-4.664	876	-2.047
Denmark		13.003	2.798	4.177
Germany (includin...		141.104	35.335	27.201
Estonia		513	-53	34
Ireland		761	-1.031	-488
Greece		-22.971	-7.216	-5.744
Spain		-48.404	-17.463	-9.910
France		-33.657	-9.428	-14.252
Italy		-54.074	-23.396	-13.807
Cyprus		: (e)	: (e)	:
Latvia		535	47	45

Рис.А.8.

Натисніть кнопку *Download* для того, щоб закачати файл з цією таблицею. Оберіть формат файлу, наприклад *xls*, додаткові опції файлу за потреби.

Інші інтеграційні об'єднання також часто надають статистичну інформацію по своїх країнах-членах. Також скажемо про ще два джерела даних по обмеженому колу країн:

- Організація економічного співробітництва та розвитку – ОЕСД (<http://www.oecd-ilibrary.org/statistics>) є джерелом широкого спектру статистичних даних по країнах-членах цієї організації.
- Основні глобальні індикатори – IMF Principal Global Indicators (<http://www.principalglobalindicators.org/default.aspx>) надає інформацію по країнах Великої двадцятки, тобто по країнах, які є системно важливими для світової економіки, для того, щоб відслідковувати їх економічний та фінансовий розвиток.

А.8. Комплексна статистика: конкурентоспроможність та інтернаціоналізація

Глобальний звіт з конкурентоспроможності Світового економічного форуму – World Economic Forum Global Competitiveness Report (<http://www.weforum.org/reports>) оцінює рейтинги конкурентоспроможності країн за таким складовими: інституції, інфраструктура, макроекономічне середовище, охорона здоров'я і початкова освіта, вища освіта і підготовка, ефективність ринку товарів, ефективність ринку праці, розвиток фінансового ринку, технологічна готовність, розмір ринку, розвиток бізнесу, інновації.

Щорічник світової конкурентоспроможності Міжнародного інституту розвитку менеджменту – IMD (International Institute for Management Development) World Competitiveness Yearbook (<http://www.imd.org/research/publications/wcy/index.cfm>) також оцінює рейтинги конкурентоспроможності країн, але за іншою методикою – за таким складовими: економічна ефективність, ефективність уряду, ефективність бізнесу, інфраструктура.

Індекс європейської конкурентоспроможності – European Competitiveness Index (<http://www.cforic.org/pages/european-competitiveness.php>) містить інформацію про конкурентоспроможність країн більшості країн Західної та Центральної Європи. Вказані узагальнюючий індекс конкурентоспроможності та дані по його складовим. Особливістю є те, що це джерело містить інформацію як на національному рівні так і на рівні регіонів у межах країн.

Приватний сектор – World Bank Private Sector (<http://go.worldbank.org/YDSTBSIU00>) дає можливість порівняти показники по кожній країні з аналогічними показниками групи країн з аналогічним рівнем доходу на душу населення та показниками всього регіону.

Індекс економічної свободи Мережі економічної свободи – Economic Freedom Network: Economic Freedom Index (<http://www.freetheworld.com/index.html>) містить оцінку економічної свободи для країн за такими складовими: величина уряду – витрати, податки і підприємства; правова структура і захист прав власності; доступ до надійних грошей; свобода міжнародної торгівлі; регулювання кредитного ринку, ринку праці та бізнесу. Кожна із складових поділяється більш детально. Більшість показників є не точними кількісними показниками, а бальними оцінками якісних показників.

Фундація «Спадщина» – Heritage Foundation (<http://www.heritage.org/index/about>) також розраховує індекс економічної свободи, але за іншою методикою.

Здійснення бізнесу – World Bank Doing Business (<http://www.doingbusiness.org/Data>) фактично є ще одним індексом економічної свободи, який враховує складність бюрократичних процедур та інших аспектів у сфері заснування бізнесу, реєстрації власності, одержання кредиту, захисту інвесторів, сплати податків, транскордонної торгівлі, забезпечення виконання контрактів, вирішення проблем неплатоспроможності.

Огляд підприємств – World Bank Enterprise Surveys (<http://data.worldbank.org/data-catalog/enterprise-surveys>) містить дані про компанії країн у таких сферах: бюрократичні складнощі, податковий тягар, корупційний тягар, втрати від кримінальних злочинів, конкуренція з боку неформального сектору, відключення або складність підключення електроенергії або водопостачання, способи фінансування, напрями капіталовкладень, способи комунікації, завантаженість виробничих потужностей, важливість експорту та імпорту, структура персоналу, характер власності, розмір тощо.

Міжнародний гід щодо країнового ризику – PRS Group International Country Risk Guide (<http://www.prsgroup.com/ICRG.aspx>) містить інформацію щодо політичних, фінансових та економічних ризиків по країнах.

Рейтинг країнового ризику Євромані – Euromoney Country Risk Ratings (<http://www.euromoneycountryrisk.com/>) враховує такі фактори: політичний ризик, економічна ефективність, структурні оцінки, доступ до банківського фінансування та ринків капіталу, індикатори боргу, кредитні рейтинги.

В сфері інтернаціоналізації компанія The A.T. Kearney випускає такі джерела інформації:

- Індекс глобальних міст – The A.T. Kearney Global Cities Index (<http://www.atkearney.com/index.php/Publications/global-cities-index.html>) показує ступінь інтеграції провідних міст з рештою світу за такими складовими: активність бізнесу, людський капітал, інформаційний обмін, культурний досвід, політичне залучення.
- Індекс глобалізації – The A.T. Kearney Globalization Index (<http://www.atkearney.com/index.php/Publications/globalization-index.html>) показує ступінь глобалізації країн за такими складовими: політичне залучення (міжнародна допомога, договори, організації, миротворча діяльність), персональні контакти (телефонні дзвінки, подорожі, грошові перекази), технологічна сполученість (поширеність інтернету), економічна інтеграція (міжнародна торгівля та прямі іноземні інвестиції).

А.9. Статистика міжнародної торгівлі

Статистична база даних торгівлі товарами ООН – UN Comtrade – United Nations Commodity Trade Statistics Database (<http://comtrade.un.org/db/default.aspx>) – дає можливість одержати інформацію за допомогою запиту за такими параметрами: класифікація товарів, вид товару, країна, країна-партнер, тип потоку (експорт, імпорт, реекспорт, реімпорт). Для запиту оберіть *Data Query* і далі *Basic Selection*. Після здійснення покрокового запиту створюється таблиця в форматі csv, яку можливо закачати. Оскільки в таблиці замість назв країн, товарів містяться коди, для їх розшиф-

ровки скористайтеся *Metadata&Reference*.

Статистика напрямів торгівлі – IMF Directions Of Trade Statistics (<http://elibrary-data.imf.org/FindDataReports.aspx?d=33061&e=170921>) містить дані про обсягу двосторонньої торгівлі між країнами: експорту та імпорту. Показана географічна структура зовнішньої торгівлі для кожної країни.

Існують й спеціалізовані бази даних по цінах:

- Ціни на сировину – IMF Primary Commodity Prices (<http://www.imf.org/external/np/res/commod/index.aspx>) містить тижневі, місячні, квартальні або річні дані по цінах сировинних товарів.
- Дані цін на товари – World Bank Commodity Price Data (<http://data.worldbank.org/data-catalog/commodity-price-data>) мають схоже призначення.

Дані та інструменти торгівлі – World Bank Trade Data and Tools (<http://go.worldbank.org/B1795PHK00>) є набором декількох складових по конкретних аспектах, пов'язаних з міжнародною торгівлею. Наприклад, База даних тимчасових бар'єрів для торгівлі – World Bank Temporary Trade Barriers Database (<http://data.worldbank.org/data-catalog/temporary-trade-barriers-database>) містить інформацію про випадки антидемпінгових процедур, встановлення компенсаційного мита, застосування захисних заходів та торговельних суперечок у рамках СОТ.

Індикатори світової торгівлі – World Bank World Trade Indicators (<http://go.worldbank.org/7F01C2NTP0>) містить інформацію про широкий спектр показників, які пов'язані з міжнародною торгівлею, об'єднані у такі групи: торговельна політика, зовнішнє середовище, інституціональне середовище, сприяння торгівлі, результати торгівлі.

Рішення світової інтегрованої торгівлі – World Bank World Integrated Trade Solution (<http://data.worldbank.org/data-catalog/wits>) – є програмним продуктом, який дозволяє доступ до баз даних UN Comtrade, СОТ та ЮНКТАД. Окрім доступу до баз даних WITS є аналітичним інструментом для оцінки наслідків зміни ставок митних тарифів. Містить можливості для візуалізації, наприклад, експорту України по країнах-імпортерах у динаміці (див. Рис.А.9).

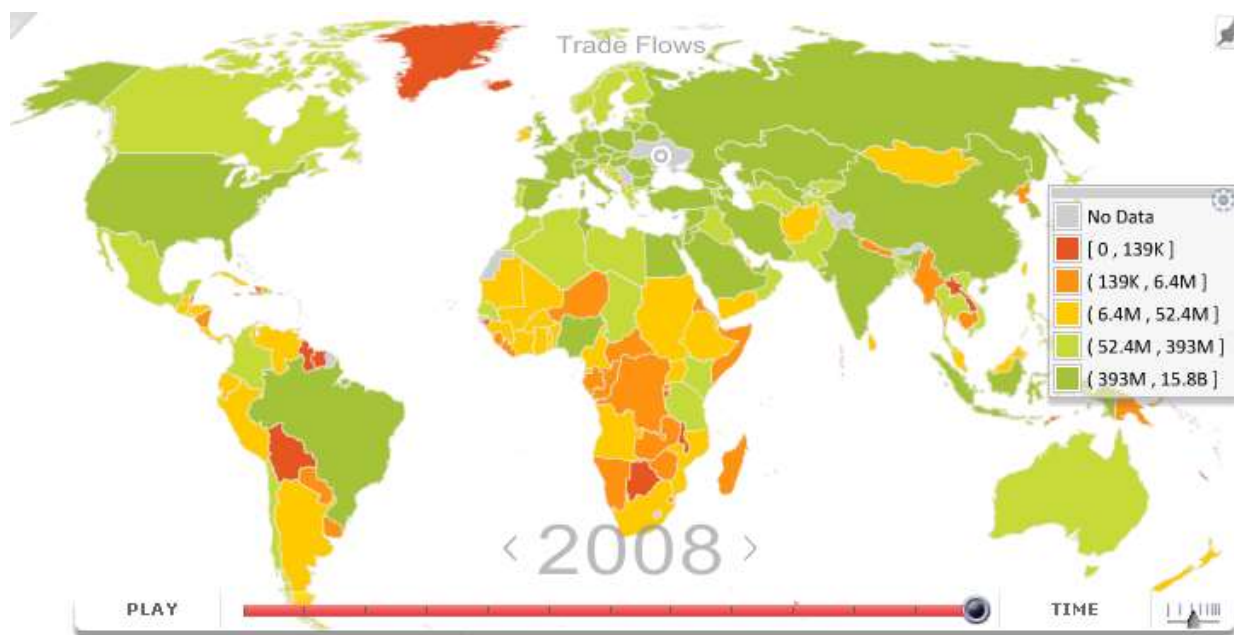


Рис.А.9.

Міжнародний торговельний центр – International Trade Center (<http://www.intracen.org/>) дає можливість одержати для кожної країни та галузі дані по міжнародній торгівлі (обсяги, темпи зростання, рівень диверсифікації, спеціалізації), прямих інвестиціях, митних тарифах тощо. Інформація центру може бути корисною при прийнятті рішень про вихід на зовнішні ринки за допомогою міжнародної торгівлі або прямих інвестицій.

В сфері привабливості розвитку торгівлі компанія The A.T. Kearney випускає такі джерела:

- Індекс глобального розвитку роздрібної торгівлі – The A.T. Kearney Global Retail Development Index (<http://www.atkearney.com/index.php/Publications/global-retail>)

development-index.html) містить рейтинги країн за їх комерційною привабливістю для розвитку роздрібною торгівлі.

- Індекс глобального розміщення послуг – The A.T. Kearney Global Services Location Index (<http://www.atkearney.com/index.php/Publications/global-services-location-index-gsli.html>) містить рейтинги країн за їх привабливістю для аутсорсингу послуг, таких як інформаційно-технологічні послуги, послуги контакт-центрів тощо.

Світова організація туризму – World Tourism Organization (<http://www.unwto.org/statistics/home.htm>) містить інформацію про розвиток міжнародного туризму.

А.10. Фінансова статистика

Міжнародний валютний фонд оприлюднює декілька джерел у сфері міжнародних та внутрішніх фінансів:

- Статистика платіжного балансу – IMF Balance Of Payments Statistics (<http://elibrary-data.imf.org/FindDataReports.aspx?d=33061&e=170784>) містить детальну інформацію про структуру платіжних балансів кожної країни.
- Погоджений огляд прямих інвестицій – IMF Coordinated Direct Investment Survey (<http://cdis.imf.org/>) надає інформацію по країнах у розрізі географічної структури прямих інвестицій та структури за інструментами.
- Погоджений огляд портфельних інвестицій – IMF Coordinated Portfolio Investment Survey (<http://www.imf.org/external/np/sta/pi/datarsl.htm>) надає таку саму інформацію по портфельних інвестиціях.
- Валютні курси – IMF Exchange Rates (http://www.imf.org/external/np/fin/data/param_rms_mth.aspx) містить інформацію по курсах валют відносно долара та СДР на денній основі. Нажаль, не містить інформацію по гривні.
- Валютна структура офіційних валютних резервів – IMF Currency Composition of Official Foreign Exchange Reserves (<http://www.imf.org/external/np/sta/cofer/eng/index.htm>) подається для світу в цілому, для розвинутих країн, для новітніх ринків та країн, що розвиваються, але в розрізі окремих країн не вказується.
- Міжнародні резерви та валютна ліквідність – IMF International Reserves and Foreign Currency Liquidity (<http://www.imf.org/external/np/sta/ir/IRProcessWeb/colist.aspx>) містить дані по складових валютних резервів та інших валютних активах центрального банку та центрального уряду, чисті зміни валютних активів у короткостроковій перспективі (детерміновані та залежні від умов).
- Огляд доступу до фінансування – IMF Financial Access Survey (<http://fas.imf.org/>) містить інформацію про показники доступності фінансових послуг для клієнтів (наприклад, кількість людей, які мають кредити або депозити на 1000 дорослих, кількість філій банків або банкоматів на 1000 км²).
- Індикатори фінансової міцності – IMF Financial Soundness Indicators (<http://fsi.imf.org/>) надають інформацію про такі показники фінансової системи, як відношення регуляторного капіталу до активів з урахуванням ризику, частка непрацюючих кредитів, прибутковість активів, частка кредитів на нерухомість у всіх кредитах тощо.
- Статистика урядових фінансів – IMF Government Finance Statistics (<http://elibrary-data.imf.org/FindDataReports.aspx?d=33061&e=170809>) містить детальні статистичні дані по доходах, витратах, активах і зобов'язаннях загального уряду та його складових.
- Фінансові дані – IMF Financial Data (<http://www.imf.org/external/data.htm>) містить інформацію про операції самого МВФ, вартість СДР та процентні ставки.
- Моніторинг механізмів фонду – IMF Monitoring of Fund Arrangements (<http://www.imf.org/external/np/pdr/mona/index.aspx>) містить дані про програми фінансування країн за підтримки МВФ: тип механізму, строки дії, цілі, обсяги, наслідки для економічної ситуації тощо.

Світовий банк також пропонує декілька джерел статистичної інформації в цій сфері:

- Банківське регулювання та нагляд – World Bank Bank Regulation and Supervision (<http://go.worldbank.org/SNUSW978P0>) надає інформацію по переважно якісних параметрах банківського регулювання та нагляду за такими складовими: вхід на ринок, власність, капітал, діяльність, зовнішній аудит, внутрішній менеджмент / організаційні вимо-

ги, ліквідність і диверсифікація, захист вкладників, забезпечення, бухгалтерській облік і розкриття інформації, дисципліна / проблемні інституції / вихід з ринку, нагляд.

- Фінансовий розвиток та структура – World Bank Financial Development and Structure (<http://go.worldbank.org/X23UD9QUX0>) містить інформацію про різноманітні відносні показники фінансової системи.
- База даних приватизації – World Bank Privatization Database (<http://data.worldbank.org/data-catalog/privatization-database>) створена для відображення приватизаційних операцій у країнах, що розвиваються, в розрізі проданих об'єктів вартістю більше 1 млн. дол.
- Квартальні дані боргу державного сектору – World Bank Quarterly Public Sector Debt (<http://data.worldbank.org/data-catalog/quarterly-public-sector-debt>) є спільною з МВФ базою даних щодо боргу загального уряду, центрального уряду, нефінансових державних корпорацій, фінансових державних корпорацій у розрізі за інструментами, строками та валютами.
- Квартальні дані зовнішнього боргу – World Bank Quarterly External Debt Statistics (<http://data.worldbank.org/data-catalog/quarterly-external-debt-statistics-ssds>) та (<http://data.worldbank.org/data-catalog/quarterly-external-debt-statistics-gdds>) містить інформацію про зовнішній борг у розрізі секторів економіки, інструментів, валют; короткостроковий зовнішній борг, обслуговування боргу, прострочені виплати.
- Ціни на грошові перекази – World Bank Remittance Prices (<http://data.worldbank.org/data-catalog/remittance-prices>) містить дані про ціни на грошові перекази невеликих сум грошей у розрізі країни платника та країни одержувача.
- Потоки допомоги – World Bank Aid Flows (<http://siteresources.worldbank.org/CFPEXT/Resources/299947-1266002444164/index.html>).
- Портфель проектів – World Bank Projects Portfolio (<http://data.worldbank.org/data-catalog/projects-portfolio>) інформує про проекти Світового банку за такими характеристиками: назва, країна, сектор, обсяги фінансування тощо.
- Фінанси Світового банку – World Bank Finances (<http://data.worldbank.org/data-catalog/world-bank-finances>) містить дані про позики Міжнародного банку реконструкції та розвитку та кредити Міжнародної асоціації розвитку, фінансові звіти, внески глобальних фондів тощо.

Банк міжнародних розрахунків – Bank for International Settlements охоплює декілька складових статистики міжнародних фінансів:

- Платіжні системи – BIS Payment Systems (http://www.bis.org/statistics/payment_stats.htm) містить дані по платіжних системах, торговельних платформах, клірингових домах, та системах розрахунків по цінних паперах.
- Статистика цін на нерухомість – BIS Property Price Statistics (<http://www.bis.org/statistics/pp.htm>) вказує ціни за видами нерухомості (від місячних до річних даних).
- Статистика цінних паперів – BIS Securities Statistics (<http://www.bis.org/statistics/secstats.htm>) містить інформацію по міжнародних та внутрішніх боргових цінних паперах, оголошених міжнародних випусках акцій, міжнародних синдигованих кредитах.
- Банківська статистика – BIS Banking Statistics (<http://www.bis.org/statistics/bankstats.htm>) надає інформацію про зовнішні фінансові вимоги (активи) та зобов'язання банків, у тому числі в розрізі валют, країн та країн-контрагентів.
- Статистика деривативів – BIS Derivatives Statistics (<http://www.bis.org/statistics/derstats.htm>) містить дані по ринку деривативів країн Великої десяти та Швейцарії.
- Ефективні валютні курси – BIS Effective Exchange Rates (<http://www.bis.org/statistics/eer/index.htm>) – вказуються номінальні та реальні ефективні курси.
- Статистика валютного ринку – BIS Foreign Exchange Statistics (<http://www.bis.org/publ/rpfx10t.htm>) подається у рамках Трирічного огляду центральних банків діяльності валютного ринку та ринку деривативів.

Спільний хаб зовнішнього боргу – Joint External Debt Hub (<http://www.jedh.org/>) або (<http://data.worldbank.org/data-catalog/joint-external-data-hub>) є спільною базою даних Світового банку, Банку міжнародних розрахунків, Міжнародного валютного фонду та Організації економічного розвитку та співробітництва. Статистика представлена у подвійному вимірі: за даними країни-боржника та за даними кредиторів або ринків.

Світова федерація бірж – Worlds Federation Of Exchanges (<http://www.world-exchanges.org/statistics>) надає інформацію про ринкову капіталізацію акцій, ринкову капіталізацію облігацій, оборот торгівлі акціями, оборот торгівлі облігаціями, капітал залучений за допомогою випуску акцій, кількість компаній у лістингу, індекси цін на акції тощо.

Звіт з фінансового розвитку Світового економічного форуму – World Economic Forum Financial Development Report (<http://www.weforum.org/reports>) оцінює рейтинги розвитку фінансової сфери країн за такими складовими: інституційне середовище, бізнес середовище, фінансова стабільність, банківські фінансові послуги, небанківські фінансові послуги, фінансові ринки, доступність фінансування.

Суверенний рейтинг Фітч – Fitch Sovereign Rating (http://www.fitchratings.com/index_fitchratings.cfm) публікує рейтинги боргу урядів у національній або іноземній валюті.

Інший приклад кредитної рейтингової інформації – Суверенний рейтинг Стендард енд Пуерс – Standard And Poors Sovereign Rating List (<http://www.standardandpoors.com/ratings/en/us>)

Індекс впевненості для прямих іноземних інвестицій – The A.T. Kearney Foreign Direct Investment Confidence Index (<http://www.atkearney.com/index.php/Publications/foreign-direct-investment-confidence-index.html>) містить дані, які характеризують умови, що впливають рішення про здійснення прямих іноземних інвестицій.

Огляд підприємництва – World Bank Entrepreneurship Survey (<http://go.worldbank.org/T8G73Z9ZM0>) містить інформацію про кількість нових компаній, які реєструються: в абсолютному вимірі та відносно населення працездатного віку. Для зручності офшорні фінансові центри виділяються червоним кольором.

A.11. Статистика технологій і знання

Індекс економіки знань – World Bank Knowledge Economy Index (<http://data.worldbank.org/data-catalog/KEI>) складається з таких компонентів: економічні стимули та інституційний режим, освіта, інновації, інформаційні та телекомунікаційні технології.

Світова організація інтелектуальної власності – World Intellectual Property Organization (<http://www.wipo.int/ipstats/en/>) надає інформацію щодо показників, пов'язаних з патентами, корисними моделями (малими винаходами), торговельними марками, промисловими зразками, сортами рослин, мікроорганізмами.

Світовий індекс конкурентоспроможності знань – World Knowledge Competitiveness Index (<http://www.cforic.org/pages/wkci.php>) містить інформацію по цьому індексу для регіонів всередині країн. Основними складовими індексу є людський капітал, капітал у формі знань, випуск регіональної економіки, фінансовий капітал, сталість знань.

A.12. Статистика інфраструктури

Міжнародна організація цивільної авіації – International Civil Aviation Organization (<http://www.icaoata.com/Trial/WhatIsICAO.aspx>) містить дані по комерційних авіаперевізниках (перевезення, фінансові дані, персонал тощо), аеропортах (перевезення, фінансові дані), провайдерах авіанавігаційних послуг, зареєстрованих цивільних літаках.

Міжнародна дорожня федерація – International Road Federation (<http://www.irfnet.org/statistics.php>) оприлюднює дані по довжині, якості та щільності мережі доріг; обсягах перевезень за видами автотранспорту; забезпеченості автотранспортом; дорожньо-транспортних пригодах; виробництву, реєстрації, експорту та імпорту автотранспорту; витратах на дороги; цінах та споживанню автомобільного пального.

Міжнародний транспортний форум – International Transport Forum (<http://www.internationaltransportforum.org/statistics/index.html>) надає інформацію про обсяги перевезень, інвестиції в інфраструктуру, транспортні податки, викиди вуглекислого газу транспортом тощо.

Індекс ефективності логістики – World Bank Logistic Performance Index (<http://data.worldbank.org/data-catalog/logistics-performance-index>) включає такі складові як ефективність розмитнення, якість торговельної та транспортної інфраструктури, легкість організації відвантаження товарів за прийнятними цінами, якість логістичних послуг, здатність відслідковувати переміщення вантажів, вчасність доставки вантажу.

Енергетична статистика – UNSD Energy Statistics (<http://unstats.un.org/unsd/energy/default.htm>) містить дані про виробництво, міжнародну торгівлю, зміни запасів, перетворення, кі-

нцеве споживання (промисловістю, транспортом, домогосподарствами) енергоносіїв за їх видами; виробництво, торгівлю та споживання електроенергії, потужність електростанцій.

Міжнародний телекомунікаційний союз – International Telecommunication Union (<http://www.itu.int/ITU-D/ict/>) оприлюднює дані по фіксованих телефонних лініях, мобільному зв'язку, телебаченню, інтернету та комп'ютерах, персоналу, тарифах, якості, доходах та інвестиціях у сфері телекомунікацій.

Інформаційні та телекомунікаційні технології – World Bank Information and Communication Technology (<http://go.worldbank.org/7S1VU9WNZ0>) містить інформацію про рівень конкуренції, витрат і доходів у цій галузі, доступ, використання, якість, доступність інформаційно-телекомунікаційних послуг, зовнішню торгівлю товарами та послугами в цій сфері тощо.

Приватна участь у інфраструктурі – World Bank Private Participation in Infrastructure (<http://data.worldbank.org/data-catalog/private-participation-in-infrastructure>) містить дані про кількість інфраструктурних проектів та обсяги інвестицій у розрізі секторів (енергетика, транспорт, телекомунікації, водопостачання та водовідведення).

A.13. Статистика населення

Міжнародна організація праці – International Labour Organization (<http://www.ilo.org/global/statistics-and-databases/lang--en/index.htm>) оприлюднює інформацію про показники зайнятості, безробіття, робочих годинах, зарплатах, вартості робочої сили, індексах споживчих цін, травмах на виробництві, страйках, доходах і витратах домогосподарств, трудовій міграції тощо.

Міграція і грошові перекази – World Bank Migration And Remittances (<http://data.worldbank.org/data-catalog/migration-and-remittances>) містить інформацію про кількість мігрантів, їх склад, основні напрями міграції, доходи і перекази працюючих за кордоном.

Прогнози населення – World Bank Population Projections (<http://data.worldbank.org/data-catalog/population-projection-tables>) містить довгострокові прогнозні дані щодо кількості населення по країнах, його вікової та статевої структури, показників народжуваності, міграції та смертності.

Глобальна база даних двосторонньої міграції – World Bank Global Bilateral Migration Database (<http://data.worldbank.org/data-catalog/global-bilateral-migration-database>) містить дані по кількості мігрантів, які емігрували в період 1960-2000 рр., у розрізі країни народження та країни імміграції. Для країн колишнього СРСР враховано міграцію між республіками.

A.14. Статистика соціальної сфери

Міжнародні показники людського розвитку – International Human Development Indicators (<http://hdr.undp.org/en/data/about/>) – містить інформацію про широкий спектр показників соціального спрямування: наскільки зручним для людини є умови в тій чи іншій країні. Основний акцент робиться на доходах населення, освіті, розвитку охорони здоров'я, екологічних аспектах.

Інститут статистики ЮНЕСКО – UNESCO Institute of Statistics (<http://www.uis.unesco.org/Pages/default.aspx>) надає інформацію в сфері освіти, науки і технології, культури та комунікацій.

Статистика освіти – World Bank Education Statistics (<http://data.worldbank.org/data-catalog/ed-stats>) охоплює показники доступності та результативності освіти, забезпеченості вчителями, витрат на освіту.

World Bank Health Nutrition and Population (<http://data.worldbank.org/data-catalog/health-nutrition-and-population-statistics>) охоплює показники динаміки кількості населення, харчування, репродуктивного здоров'я, фінансування охорони здоров'я, ресурсів медицини, рівня імунізації, поширеності хвороб тощо.

Гендерна статистика – World Bank Gender Statistics (<http://data.worldbank.org/data-catalog/gender-statistics>) містить дані про вікові групи населення, показники зайнятості, тривалості життя, доступності освіти, доступ до ресурсів, зарплати, участь в управлінні тощо залежно від статевої приналежності.

Динаміка глобального розподілу доходів – World Bank Global Income Distribution Dynamics (<http://go.worldbank.org/15GL9L8SJ0>) містить дані про розподіл доходів або споживання серед різних верств населення та дозволяє порівняти відповідні верстви населення за іншими показниками, такими як освіта, вік, місце проживання.

База даних нерівності доходу – UNU-WIDER Income Inequality Database (http://www.wider.unu.edu/research/Database/en_GB/database/) містить дані про коефіцієнт Джині та частки доходу за верствами населення.

A.15. Статистика політичної сфери та державного управління

Дієві індикатори управління – World Bank Actionable Governance Indicators (<http://data.worldbank.org/data-catalog/actionable-governance-indicators>) містять інформацію про управління державним сектором, політичну підзвітність, взаємозалежність та взаємні обмеження інституцій влади, громадянське суспільство і засоби масової інформації, взаємодію з приватним сектором, децентралізацію і місцеву участь.

Політичні інституції – World Bank Political Institutions (<http://go.worldbank.org/2EAGGLRZ40>) є дослідницькою базою даних, яка містить інформацію про особливості політичної системи країни, строк знаходження президента / прем'єр-міністра та правлячої партії при владі, рівень підтримки на виборах голови держави / уряду, характер правлячої партії, розподіл місць у парламенті за партіями, особливості виборчого законодавства, стабільність, систему взаємозалежності та взаємних обмежень інституцій влади, рівень федералізації.

Загальносвітові індикатори управління – World Bank Worldwide Governance Indicators (<http://data.worldbank.org/data-catalog/worldwide-governance-indicators>) охоплюють такі сфери: голосування та підзвітність, політична стабільність та відсутність насилля, ефективність уряду, якість регулювання, верховенство права, контроль над корупцією.

Freedom House (<http://www.freedomhouse.org/template.cfm?page=15>) у рамках звіту Свобода у світі Freedom in the World проводить рейтинг країн за ступенем додержання політичних прав та громадянських свобод.

Мережа Інституту дослідження миру в Осло – PRIO Network (<http://www.prio.no/Data/>) містить дані про дати та характеристики збройних конфліктів та обсяги торгівлі зброєю.

Transparency International (http://www.transparency.org/policy_research/surveys_indices) здійснює оцінки ступеню поширення корупції, зокрема й у розрізі окремих сфер: політичні партії, законодавчий орган, правоохоронні органи, бізнес, засоби масової інформації, державні службовці, система судочинства, недержавні організації, релігійні організації, збройні сили, система освіти.

A.16. Екологічна статистика

Індекс екологічної ефективності – Environmental Performance Index (<http://sedac.ciesin.columbia.edu/es/epi/index.html>) враховує такі складові: втрати здоров'я від несприятливого навколишнього середовища, вплив забруднення на людину та навколишнє середовище, біорізноманіття, ліси, рибне господарство, сільське господарство, кліматичні зміни.

Скориговані чисті заощадження – World Bank Adjusted Net Savings (<http://data.worldbank.org/data-catalog/environmental-accounting>) надають оцінку справжнього рівня заощаджень в економіці з урахуванням поправок на інвестиції в людський капітал, вичерпання природних ресурсів та втрати від забруднення навколишнього середовища.

Дані про навколишнє середовище – World Bank Environmental Data (<http://data.worldbank.org/data-catalog/environmental-data-by-topic>) містить інформацію про сільське господарство, ліси та біорізноманіття, енергію, викиди та забруднення, водопостачання і водовідведення, навколишнє середовище і здоров'я.

Навколишнє середовище – короткий огляд – World Bank Environment At Glance (<http://data.worldbank.org/data-catalog/country-environmental-factsheets>) містить дані про складові багатства країни, склад природного капіталу, зобов'язання щодо міжнародних договорів у сфері захисту навколишнього середовища, забруднення, доступ до водопостачання і водовідведення, транспорт і енергетику, навколишнє середовище і здоров'я, природні ресурси та їх використання.


Багатство націй – World Bank Wealth of Nations (<http://data.worldbank.org/data-catalog/wealth-of-nations>) містить оцінки багатства країни та її складових: вироблений капітал, невідчутний капітал та природний капітал. Останній поділяється на активи надр, ресурси деревини, інші ресурси лісу, землі під захистом, орні землі, пасовища.

Додаток Б.

Основи роботи у програмному забезпеченні Microsoft Office Excel

В цій главі будуть розглянуті лише ті аспекти роботи у Microsoft Office Excel, які безпосередньо стосуються кількісного аналізу.

Увага: в даній книзі використовується західний стиль позначення розділового знаку для дрібних розрядів чисел, тобто використовується крапка, а не кома, так як більшість джерел міжнародної економічної статистики є закордонними і використовують саме цей стиль. Цей параметр

можна обрати в Microsoft Windows. Для цього потрібно після натиснення кнопки *Пуск*  потрібно обрати *Панель управління*, далі обрати *Язык и региональные стандарты*, обрати *Настройка*, в полі *Разделитель целой и дробной части* обрати крапку (.) або кому (,).

Наприклад, якщо в Microsoft Windows як розділовий знак дрібної частини використовується крапка, то Microsoft Excel сприйматиме запис у комірці або формулі типу 3.4, як число, а запис типу 3,4 як текст. Через це можуть стати неможливими розрахунки з числами типу 3,4. Тому потрібно пересвідчитись, що Microsoft Excel сприймає відповідні записи як число для того, щоб проводити з ними розрахунки. Наприклад, відмітити всі комірки з числами типу 3,4 і замінити коми на крапки.

Якщо в Microsoft Windows як розділовий знак дрібної частини використовується кома, то виникає протилежна проблема: Microsoft Excel сприйматиме запис у комірці або формулі типу 3,4, як число, а запис типу 3.4 як текст.

Б.1.Формули

Формули дозволяють проводити розрахунки та інші операції з даними. Кожна формула починається зі знаку =. Формула повертає результат розрахунку в комірку, в якій вона прописана.

Аргументами формул можуть, наприклад, бути числа, текст, посилання на інші комірки. Посилання на інші комірки можливо вводити по різному:

- прописуючи їх назву у формулі;
- відмічаючи потрібні комірки в процесі написання формули.

Приклади формул показані у Табл.Б.1

Таблиця Б.1.

=3+7*2	Додає 3 до добутку 7 та 2
= ABS(-3)	Обраховує модуль числа -3, тобто повертає 3
= A1+B3	Додає значення комірок A1 та B3
= ЕСЛИ(A3>5)	Перевіряє чи значення у комірці A3 є більшим за 5

Елементи формул. Приклад:

= ABS(-3)*A1^5

де ABS(-3) – функція, яка повертає значення 3;

A1 – посилання на комірку A1;

-3 та 5 – константи, тобто числа або текст, які вводяться в формулу безпосередньо;

* та ^ - оператори, тобто символи, які позначають тип розрахунків (* – множення, а ^ – введення значення перед цим оператором у ступінь, вказаний після оператора)

Оператори. Приклади операторів вказані у Табл.Б.2. Розрахунки у формулах здійснюються згідно правил арифметики (наприклад, спочатку множення, а потім додавання). Зокрема операції здійснюються в такому порядку : пробіл , - % ^ * та / + та - & = < > < = > = < >

Якщо є оператори з одноковим пріоритетом, то першою здійснюється операція, оператор якої є лівішим. За допомогою дужок можливо коригувати порядок операцій. Наприклад, у наступній формулі спочатку буде виконуватися додавання, а не множення = 7*(3+5) і результатом формули буде 56.

Таблиця Б.2.

Оператор	Тип операції	Приклад
Арифметичні (результатом є число)		
+	додавання	1+2
-	віднімання	3-2
*	добуток	5*2
/	ділення	1/2
%	процент	9%
^	введення у ступінь	2^2
Порівняння (результатом є логічне значення ИСТИНА або ЛОЖЬ)		
=	дорівнює	A1=A2
>	більше	A1>A2
<	менше	A1<A2
>=	більше або дорівнює	A1>= A2
<=	менше або дорівнює	A1<=A2
<>	не дорівнює	A1<> A2
Текстові		
&	конкатенація – об'єднання тексту	"Експорт"&"України" дає результат "Експорт України"
Посилання		
:	діапазон	A1:A10
,	об'єднання множин	A1:A10,A20:A22
(пробіл)	перетин множин (спільні комірки двох діапазонів)	A1:B20 B5:K10

Б.2. Функції

Функції – це стандартні формули, які є вбудовані у програмне забезпечення. Вони можуть бути використані для розрахунку числових значень, одержання поточної дати, пошуку, перетворення тексту, порівняння. Елементи формул. Приклад:

=СТЕПЕНЬ(3;G6)

Як і формули функції починаються зі знаку =. Далі йде назва функції. В дужках вказані аргументи, розділені знаком ;. Часто важливим є порядок, у якому прописуються аргументи. Аргументами можуть бути:

- числа;
- текст;
- логічні значення (ИСТИНА або ЛОЖЬ);
- масиви (множини даних у діапазоні комірок);
- помилки.

Аргументи формул також можна класифікувати як:

- константи (постійне значення, яке не розраховується, наприклад, текст або число);
- посилання на комірки;
- формули;
- функції.

Приклад вставки однієї функції у іншу:

=СТЕПЕНЬ(2+ABS(-3);G6)

У деяких функціях потрібно вказувати лише аргументи певного типу. Деякі функції не потребують аргументів.

При введенні тексту функції у комірку можуть з'являтися підказки (по початкових літерах – повна назва функції, після розкриття дужки – характер аргументів, наприклад *СТЕПЕНЬ* (чис-

ло;степень)). Ще один спосіб вставити функцію – натиснути кнопку *Вставити функцію* у меню формули, обрати з переліку потрібну функцію і у вікні, що з’явиться після натискання кнопки *ОК*, вказати аргументи:

- набрати з клавіатури;
- вставити попередньо скопійовані значення;
- або натиснути мишею на комірку (або виділити діапазон комірок, натиснувши лівою кнопкою миші і протягнувши її по відповідних комірках), посилання на яку (які) потрібно використати як аргумент.

Після введення функції (як і формули) в комірці буде відображатися результат розрахунку або значення помилки. Сам текст функції (або формули) буде відображатися у рядку формул, коли відповідна комірка є обраною. Наприклад, як на Рис.Б.1. В цьому ж рядку формул можна редагувати формулу (функцію).

	A	B	C	D	E	F
1	27	3				
2						

Рис.Б.1.

Б.3. Перерахунок формул

При зміні значень комірок, на які є посилання в формулі, (впливаючих комірок) Microsoft Office Excel автоматично буде перераховувати результат у комірці, в якій прописана формула (залежна комірка). Якщо посилання є циклічним (залежна комірка сама впливає на свої впливаючі комірки), автоматичний перерахунок неможливий.

Вибір способу перерахунку можна змінити. Для цього у вкладці *Файл* оберіть *Параметри*, а далі – *Формули*. Оберіть необхідний спосіб у розділі *Параметри вычислений*:

- *Автоматически*;
- *Автоматически, кроме таблиц данных*;
- *Вручную* (в останньому випадку автоматично поряд буде включений режим *Пересчитывать книгу перед сохранением*, але його можливо відключити).

Зміна способу перерахунку впливатиме на всі відкриті книги. Для перерахунку при ручному способі перерахунку оберіть у вкладці *Формули* в групі *Вычисление* кнопку *Пересчет* (для перерахунку всіх відкритих книг, включаючи таблиці даних та оновлення всіх відкритих аркушів діаграм) або *Произвести вычисления* (для перерахунку активного аркушу, діаграм та аркушів діаграм, пов’язаних з цим аркушем).

Б.4. Посилання на комірки у формулах

Для того, щоб вказати посилання на комірку, потрібно набрати назву посилання з клавіатури або під час введення формули натиснути на відповідну комірку, на яку здійснюється посилання (в останньому випадку, якщо комірок багато, потрібно виділити діапазон, натиснувши лівою кнопкою миші на крайню комірку діапазону і провести мишею до протилежної крайньої комірки діапазону). Увага: в посиланнях використовуються латинські літери, тому якщо замість латинської А буде стояти А кирилицею, формула повертатиме помилку замість результату розрахунку.

Приклади посилань на комірки вказані у Табл.Б.3. При копіюванні, заповненні по рядкам (або стовпчикам) або перенесенні формули в іншу комірку посилання можуть залишатися незмінними (якщо використані абсолютні посилання) або зміщуватися (якщо використані відносні посилання).

Відносні посилання. Припустимо в комірці А1 є формула = 2*А2. А2 = приклад відносного посилання. При копіюванні формули в комірку В3, формула в В3 буде виглядати як =2*В4. Тобто у відносному посиланні важливо, де розташована комірка, на яку є посилання, відносно комірки з формулою.

Таблиця Б.3.

B3	Комірка на перетині стовпчика В і рядку 3
A2:A5	Діапазон комірок у стовпчику А, рядки 2-5
B3:K3	Діапазон комірок у рядку 3, стовпчики В-К
A3:C7	Прямокутний діапазон комірок у стовпчиках А-С, рядках 3-7
Експорт!A2:A5	Діапазон A2:A5 в аркуші під назвою <i>Експорт</i> у тій самій книзі
'Експорт%'!A2:A5	Діапазон A2:A5 в аркуші під назвою <i>Експорт%</i> у тій самій книзі – одинарні лапки використовуються, якщо назва аркушу містить не тільки літери
[Україна.xlsx]'Експорт%'!A2:A5	Діапазон A2:A5 в аркуші під назвою <i>Експорт%</i> у книзі <i>Україна.xlsx</i>
Лист5:Лист7!A1:A8	Діапазон A1:A8 в аркушах <i>Лист5</i> , <i>Лист6</i> та <i>Лист7</i>

Абсолютні посилання. Припустимо в комірці А1 є формула =2*\$A\$2. Абсолютні посилання позначаються знаком \$ перед літерою стовпчика і числом рядка. При копіюванні формули в комірку В3, формула в В3 буде виглядати як =2*\$A\$2. Тобто зміщення посилання не відбудеться.

Змішані посилання. Такі посилання містять абсолютний стовпчик і відносний рядок або відносний стовпчик і абсолютний рядок. Припустимо в комірці А1 є формула =2*\$A2. При копіюванні формули в комірку В3, формула в В3 буде виглядати як =2*\$A4. Приклади змішаних посилань \$A2 або B\$6.

Б.5. Функції масиву

Формула масиву повертає одне або декілька значень, які з'являються у одній чи декількох комірках, відповідно. Аргументами формули масиву є декілька наборів значень (аргументи масиву). Потрібно, щоб аргументи масиву у формулі включали однаково кількість стовпчиків та рядків. Формули масиву вводяться не кнопкою *Enter*, а одночасним натисканням трьох клавіш *CTRL+SHIFT+Enter*. Приклад розрахунку одного значення показаний у Табл.Б.4

Таблиця Б.4.

	А	В	С
1		Експорт (фізичний обсяг), тони	Ціна, долари
2	Нафта	100	400
3	Чавун	300	500
4	Загальний експорт (вартісний обсяг)	=СУММ(В2:В3*С2:С3)	

Після натискання *CTRL+SHIFT+Enter* формула у рядку формул матиме вигляд {=СУММ(В2:В3*С2:С3)}, тобто буде у фігурних дужках.

Приклад розрахунку декількох значень. Припустимо у рядку А експорт за п'ять років. У рядку В ми хочемо дізнатися розраховані значення експорту в ці роки згідно лінійного тренду. Відмітимо діапазон В2:В5 (треба перед вводом формули вказати діапазон комірок, який міститиме результати). В рядку формул пропишемо формулу =ТЕНДЕНЦИЯ(А1:А5). Натиснемо *CTRL+SHIFT+Enter*. У рядку В з'являться результати розрахунку, вказані у Табл.Б.5.

Таблиця Б.5.



	А	В
1	35	39
2	50	49
3	60	59
4	80	69
5	70	79


Після вводу формули масиву, значення у комірках, де містяться результати не можна змінити. Також не можна вставляти чи видаляти рядки чи стовпчики посеред діапазону з результатами.

З одного боку це не зручно, з іншого боку це забезпечує від внесення випадкових помилкових змін.

Для видалення формули масиву, потрібно виділити комірку з результатом розрахунку, виділити формулу у рядку формул, натиснути клавішу *DELETE*, а потім ще натиснути *CTRL+SHIFT+Enter*.

Б.6. Переміщення і копіювання формул

Для вирізання або копіювання формул (або чисел, тексту) можливо використати стандартні кнопки *Вирізати*  та *Скопіювати*  у вкладці *Главная*. Для вставки після виділення діапазо-

ну, куди потрібно скопіювати або перемістити, використовується кнопка  *Вставити*. Можна обрати спосіб вставки, натиснувши на трикутничок під цією кнопкою і обравши *Спеціальная вставка*. З'являються кнопки над якими з'являються підказки. Наприклад, можливо вставити:



- формули (відобразяться значення, але нові комірки міститимуть формули);



- лише значення (відобразяться значення і нові комірки не міститимуть формули – це зручно, коли нас цікавлять лише значення і їх потім не потрібно перераховувати, якщо нас цікавлять лише значення, а формула містить відносні посилання і при переміщенні формули результат може змінитися);



- значення і формат;



- лише формат;



- транспонувати (порівняно з первинним діапазоном комірок у нового діапазону замість стовпчиків будуть рядки, а замість рядків стовпчики, наприклад, як з Табл. Б.6 у Табл.Б.7).

Таблиця Б.6

Первинний діапазон	
6	7
2	1

Таблиця Б.7.

Діапазон після копіювання з транспонуванням	
6	2
7	1

Копіювати формули до сусідніх комірок можна також за допомогою маркера заповнення (невеликого чорного квадрату праворуч знизу виділеного діапазону або комірки). Після виділення комірки чи діапазону з потрібною нам формулою натисніть на маркер заповнення і перемістити в потрібному напрямі. Формули копіюються до тих комірок, які виділяються при переміщенні цього маркера заповнення.

Після копіювання або переміщення варто впевнитися, що формули використовують потрібні аргументи та комірки містять потрібні значення.

Б.7. виправлення помилок

Замість значень формули можуть повертати помилки. Приклади помилок:

#ДЕЛ/0! – якщо відбувається ділення на нуль або ділення на посилання на комірку з відсутнім значенням.

#ЗНАЧ! – якщо формула містить посилання на комірки з різними типами даних; у формулі введене посилання як аргумент, а значення у комірці, на яку вказане посилання, іншого типу, ніж дозволяється. Наприклад, в результаті формули =LOG(H13), якщо H13 містить текст замість числа.

#ИМЯ? – якщо назва формули вказана неправильно; текст як аргумент введений не в лапках; у посиланні на діапазон комірок відсутній знак : ; посилання на інший аркуш без одинарних лапок (') тощо.

#Н/Д – якщо в формулі масиву аргумент з іншою кількістю строк або рядків ніж у діапазоні, якій містить формулу масиву; не вказані обов'язкові аргументи для функції тощо.

#ССЫЛКА! – якщо посилання на комірку не дійсне (наприклад посилання є, а комірка була видалена, тому що був видалений відповідний аркуш, стовпчик тощо; або після копіювання відносно посилання не можуть бути адекватно відображені).

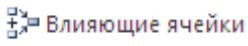
#ЧИСЛО! – якщо в функції використано не той тип даних у аргументі (наприклад, замість числа текст); результатом функції є число занадто велике або мале для того, щоб його можливо було б представити в програмі.

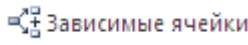
– позначка комірки, якщо ширини стовпчика недостатня для відображення всіх цифр числа, для її виправлення достатньо розширити стовпець до необхідної ширини.

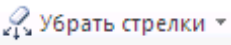
#ПУСТО! – якщо вказаний перетин двох множин, які насправді не перетинаються, наприклад =СУММ(B2:B3 D2:D3)


Б.8. Впливаючі та залежні комірки

Приклад. Якщо в комірці А1 є формула =А2*2, то комірка А1 буде залежною відносно А2, а А2 буде впливаючою відносно А1. Зв'язки між комірками також можливо відобразити графічно стрілками (стрілка йде від впливаючої до залежної комірки).

Оберіть комірку, для якої потрібно знайти залежності. На вкладці *Формулы* в групі *Зависимости формул* натисніть кнопку *Влияющие ячейки* . З'являться стрілки до цієї комірки з комірок, на які є посилання. Якщо натиснути на кнопку *Влияющие ячейки* ще раз, будуть показані впливаючі комірки другого порядку, які впливають на впливаючі комірки для первинно обраної нами комірки. Кожний раз натискаючи на *Влияющие ячейки*, з'являтимуться залежності все більш глибокого рівня.

Аналогічно натискаючи на кнопку *Зависимые ячейки*  можливо дізнатися, в яких комірках є посилання на виділену комірку. Так само можна натискати на неї декілька разів, щоб дізнатися залежності більшого рівня.

За допомогою кнопки *Убрать стрелки*  можна повернутися до режиму відображення, коли стрілки між комірками відсутні.

Сині стрілки вказують на комірки без помилок, червоні на комірки, які спричиняють помилки. Якщо залежна або впливаюча комірка знаходиться на іншому аркуші або книзі, відображатиметься чорна стрілка із значком 

Б.9. Основні логічні функції

Логічні функції можуть бути, зокрема, корисними для перевірки на відсутність помилок або при створенні нових змінних, особливо неметричних. Розглянемо декілька корисних функцій.

=ЕСЛИ(логічний вираз; значення якщо ИСТИНА; значення якщо ЛОЖЬ)

Повертає значення (наприклад, число або текст), вказане як другий аргумент, якщо логічний вираз у першому аргументі вірний, інакше повертає значення, вказане як третій аргумент). Приклад – у комірці D2 пропишемо формулу (а потім її скопіюємо в нижчі комірки) =ЕСЛИ(C2="ЗВТ";1;0)

Ця функція корисна, якщо ми хочемо створити псевдозмінну (для подальшої побудови регресійної моделі), яка приймає значення 1, якщо відповідному спостереженню відповідає режим зони вільної торгівлі, і 0, якщо зони вільної торгівлі немає. В таблиці з вхідними даними наявність зони вільної торгівлі позначена текстом "ЗВТ" як у Табл.Б.8.

Таблиця Б.8

	А	В	С	Д
1	Країни	Зовнішньоторговельний оборот	Регіональна торговельна угода	Наявність ЗВТ
2	Україна-Росія	25000	ЗВТ	1
3	Україна-Бразилія	1000	-	0

Інші приклади використання формули вказані у Табл.Б.9

Таблиця Б.9

=ЕСЛИ(A1=100;A2+A3;"")	Якщо A1=100, розраховується сума A2 і A3, інакше комірка буде пустою
=ЕСЛИ(A1>10;"Високий тариф"; "Низький тариф")	Якщо значення A1 більше 10, у комірці з формулою буде текст "Високий тариф", інакше "Низький тариф")
=ЕСЛИ(A1=B1; "ОК"; "Помилка")	Якщо A1=B1 повертається "ОК", інакше "Помилка"
=ЕСЛИ(A1>10;"Високий тариф"; ЕСЛИ(A1>5;"Середній тариф"; "Низький тариф")	Приклад вкладеної функції. Якщо значення A1 більше 10, у комірці з формулою буде текст "Високий тариф", якщо більше 5 і до 10 включно – "Середній тариф", до 5 включно "Низький тариф")

= **И**(логічний вираз; логічний вираз; ...)

Повертає значення **ИСТИНА**, якщо кожний з аргументів дорівнює **ИСТИНА**. Якщо хоча б один з аргументів дорівнює **ЛОЖЬ**, повертає **ЛОЖЬ**. Приклади:

Таблиця Б.10

=И(1<A2;A2<100)	
=ЕСЛИ(И(0<A1;A1<100); A1;"Помилка")	Повертатиме значення A1, лише, якщо 0<A1<100, інакше повідомлятиме про помилку

= **ИЛИ**(логічний вираз; логічний вираз; ...)

Повертає значення **ИСТИНА**, якщо хоча б один з аргументів дорівнює **ИСТИНА**. Якщо всі аргументи дорівнюють **ЛОЖЬ**, повертає **ЛОЖЬ**.

= **НЕ**(логічний вираз)

Повертає значення **ИСТИНА**, якщо аргумент дорівнює **ЛОЖЬ**. Якщо аргумент дорівнює **ЛОЖЬ**, повертає **ИСТИНА**.

Б.10. Основні математичні функції

= **ABS**(число)

Повертає модуль (абсолютну величину) числа (тут і далі замість числа може бути посилання на комірку, в якій міститься число або формула, яка повертає число). Наприклад =ABS(-10) повертає 10.

= **ЗНАК**(число)

Повертає 1, якщо аргумент додатне число; 0, якщо він дорівнює нулю; -1, якщо аргумент від'ємне число. Наприклад =ЗНАК(-10) повертає -1.

= **ОКРУГЛ**(число для округлення;кількість десятинних розрядів для округлення)

Повертає округлене число. Приклади:

Таблиця Б.11.

=ОКРУГЛ(4.2;0)	Повертає 4
=ОКРУГЛ(A5;0)	Наприклад, якщо в комірці A5 значення дорівнює 5,8, то повертає 6
=ОКРУГЛ(-4,2;0)	Повертає -4
=ОКРУГЛ(1.337;2)	Повертає 1,34
=ОКРУГЛ(53721;-3)	Повертає 54000

Для округлення вгору або вниз використовуються схожі функції: ОКРУГЛВВЕРХ та ОКРУГЛВНИЗ. Приклад, =ОКРУГЛВНИЗ(-371;-1) повертатиме -370. Для округлення до заданого кратного використовується функція:

= **ОКРУГЛТ**(число для округлення;кратне до якого потрібно округлити)

Наприклад, =ОКРУГЛТ(23; 5) повертатиме 25. Увага: обидва аргументи мають бути одного знаку (обидва додатні або від'ємні). Другий аргумент може бути нецілим числом, наприклад 4.2 або 0.007.

=**СТЕПЕНЬ**(число-основа;число-показник ступеню)

Повертає результат зведення числа у ступінь (аналог оператора ^). Наприклад, =СТЕПЕНЬ(10;2) повертає 100.

=**EXP**(число)

Повертає результат зведення числа e (2,728182845904) у ступінь, задану аргументом. Наприклад, =EXP(2.5) повертає 12.182494.

=**LN**(число)

Повертає натуральний логарифм додатного числа (логарифм з основою 2,728182845904). Наприклад, =LN(5) повертає 1.6094379.

=**LOG10**(число)

Повертає десятковий логарифм додатного числа (логарифм з основою 10). Наприклад, =LOG10(1000) повертає 3.


=**LOG**(число;основа)

Повертає логарифм додатного числа (логарифм з вказаною основою). Наприклад, =LOG(8; 2) повертає 3.

Б.11. Функції суми та добутку

=**СУММ**(числа)

Повертає суму чисел. Наприклад, =СУММ(A1:A10;C5:C9)

Для функції суми також можна використовувати кнопку автосуми  у вкладці *Главная*, якщо діапазон чисел знаходиться у суміжних комірках. Перед натисканням цієї кнопки потрібно виділити комірку під стовпчиком чисел або праворуч від рядка чисел, що додаються. В цій комірці буде підрахована сума.

=**ПРОИЗВЕД**(числа)

Повертає добуток чисел.

=**СУММЕСЛИ**(діапазон;критерій;діапазон додавання)

Повертає суму чисел, які відповідають певному критерію.

Діапазон – посилання на комірки, які оцінюються за критерієм. Критерій для оцінки може бути числом, виразом, посиланням на комірку, текстом, функцією.

Критерій у вигляді тексту або з логічними чи математичними знаками мають бути у лапках. В аргументі критерій можна використовувати знак ? (означає будь-який знак) та * (будь-яка послідовність знаків).

Діапазон сумування – необов’язковий аргумент, вказує на комірки, які додаються, якщо вони відрізняються від комірок, вказаних у аргументі діапазон. Якщо діапазон сумування не вказаний, додаються значення у комірках, вказаних у аргументі діапазон.

Приклади:

Таблиця Б.12.

=СУММЕСЛИ(A1:A5;"Експорт";C1:C5)	Повертає суму тих чисел у діапазоні C1:C5, якщо у відповідних комірках в одному й тому рядку з цими числами у діапазоні A1:A5 вказаний текст <i>Експорт</i> . Наприклад, це може бути корисним, якщо в рядках йдуть різні показники по різних країнах і нам потрібно підсумувати один показник по декількох країнах
=СУММЕСЛИ(A1:A100;">0")	Повертає суму тих чисел у діапазоні A1:A100, які перевищують 0. Наприклад якщо у вказаному діапазоні – баланс поточного рахунку по різних країнах, а нас цікавить сума лише профіцитів балансу поточного рахунку по декількох країнах (інакше сума дефіцитів і профіцитів буде наближатися до нуля).
=СУММЕСЛИ(A1:A10;0;B1:B10)	Наприклад, якщо у A1:A10 вказано ставку імпортного мита по конкретних товарах, а в B1:B10 імпорт цих товарів, за допомогою цієї функції ми можемо підрахувати величину імпорту товарів, з яких не стягується мито.
=СУММЕСЛИ(A1:A10;">"&E20;B1:B10)	Повертає суму тих чисел у діапазоні B1:B10, якщо у відповідних комірках в одному й тому рядку з цими числами у діапазоні A1:A10 числа перевищують число вказане у комірці E20.
=СУММЕСЛИ(A1:A5;"*порт";C1:C5)	Повертає суму тих чисел у діапазоні C1:C5, якщо у відповідних комірках в одному й тому рядку з цими числами у діапазоні A1:A5 вказаний текст, який закінчується на <i>порт</i> . Наприклад, якщо ми хочемо підрахувати суму по країнах значень обох показників експорту та імпорту (обидва слова закінчуються літерами <i>порт</i>).
=СУММЕСЛИ(A1:A5;"";C1:C5)	Наприклад, якщо у A1:A5 вказані товари, які імпортуються, а в C1:C5 величини імпорту цих товарів, функція повертає суму імпорту тих товарів, назви яких не вказані.

=СУММЕСЛИМН(діапазон сумування; діапазон умови; критерій; діапазон умови; критерій;...)

Повертає суму чисел, які відповідають одразу декільком критеріям (приклад див. у Табл. Б.13).

Таблиця Б.13.

	А	В	С	Д
1	Країна-експортер	Країна-імпортер	Товар	Обсяг
2	Україна	Польща	чавун	100
3	Україна	ФРН	чавун	200
4	Росія	Польща	нафта	1500
5	Україна	ФРН	зерно	100
6	США	Мексика	автомобілі	2000

Функція =СУММЕСЛИМН(D2:D7;A2:A7;"Україна";C2:C7;"чавун") повертатиме суму експорту з України чавуна.

=СУММКВ(числа)

Повертає суму квадратів чисел. Наприклад, =СУММКВ(A1:A2). Може бути корисною, якщо самостійно розраховувати деякі статистичні показники. Або, наприклад, якщо ми з метою попереднього аналізу хочемо дізнатися наскільки тісно пов’язані два показники (наприклад зростання прямих іноземних інвестицій у двох країнах), а в комірках A1 та A2 вказані кореляція між зростанням прямих іноземних інвестицій у двох країнах: в A1 за один період (наприклад 1991-2010 рр. або 1991-2000), а в A2 за інший період наприклад за 2001-2010 рр.). Сума квадратів дозволяє уникнути ситуації, коли від’ємна і додатна кореляція при сумуванні взаємно компенсують одна одну і ми б тоді не побачили наявності суттєвого зв’язку, який міг би змінюватися з часом.

=СУММКВРАЗН(масив1;масив2)

Повертає суму квадратів різниці чисел. Як і в інших подібних функціях, масиви мають бути однакового розміру за кількістю комірок. Фактично використовується формула: $\Sigma(x-y)^2$. Приклади:

Таблиця Б.14.

=СУММКВРАЗН(A1:A8;B1:B8)	Повертає суму: $(A1-B1)^2+(A2-B2)^2+\dots+(A8-B8)^2$
=СУММКВРАЗН({2;3;9};{6;5;11})	Повертає суму квадрату різниць двох масивів констант

=СУММПРОИЗВ(масив1;масив2;...)

Повертає суму добутків пар чисел у відповідних масивах. Приклад:

Таблиця Б.15.

	A	B	C	D	E
1	Місяць	Експорт, тон	Імпорт, тон	Експортна ціна за тону	Імпортна ціна за тону
2	Січень	100	30	100	80
3	Лютий	80	40	90	70
4	Березень	60	20	60	80
5	Квітень	70	30	70	75

=СУММПРОИЗВ(B2:C5;D2:E5) Повертає вартісний зовнішньоторговельний оборот за 4 місяці: $100*100+30*80+80*90+40*70+60*60+20*80+70*70+30*75$, аналогом цієї функції є функція =СУММ(B2:C5*D2:E5)

=СУММПРАЗНКВ (масив1;масив2)

Повертає суму різниць квадратів чисел. Фактично використовується формула: $\Sigma(x^2-y^2)$. Приклад: =СУММПРАЗНКВ(A1:A2;B1:B2) повертає суму: $(A1^2-B1^2)+(A2^2-B2^2)$

=СУММСУММКВ (масив1;масив2)

Повертає суму сум квадратів чисел. Фактично використовується формула: $\Sigma(x^2+y^2)$. Наприклад: =СУММСУММКВ(A1:A2;B1:B2) Повертає суму: $(A1^2+B1^2)+(A2^2+B2^2)$.

Б.12. Табличні функції

=СЧЁТ(діапазон)

Повертає кількість комірок у діапазоні, які містять числа. Приклад:

Таблиця Б.16.

Індекс експортних цін (2000=100)	A	B	C	D
1	Країна/рік	1993	1994	1995
2	Україна	50	60	75
3	Білорусь	н.д.	70	65

=СЧЁТ(B2:D3) повертає кількість спостережень (країно-років) з наявними числовими даними (5).

Якщо потрібно підрахувати також й комірки, що містять й інші дані (логічні значення, текст, значення помилок), використовується функція =СЧЁТЗ(діапазон).

=СЧЁТЕСЛИ(діапазон; критерій)

Повертає кількість комірок, які відповідають певному критерію. Приклади:

Таблиця Б.17.

=СЧЁТЕСЛИ(B1:B100;"ЗВТ")	Повертає кількість комірок у діапазоні B1:B100, в яких присутній текст ЗВТ (зона вільної торгівлі)
=СЧЁТЕСЛИ(B1:B100;"<"&E3)	Повертає кількість комірок у діапазоні B1:B100, в яких значення не дорівнює значенню у комірці E3
=СЧЁТЕСЛИ(B1:B100;">=10")-СЧЁТЕСЛИ(B1:B100;">30")	Повертає кількість комірок у діапазоні B1:B100, в яких значення становить від 10 до 30
=СЧЁТЕСЛИ(A1:A5;"*порт")	Повертає кількість комірок у діапазоні A1:A5, у яких вказаний текст, який закінчується на <i>порт</i> .
=СЧЕТЕСЛИ(A1:A5;"*")	Повертає кількість комірок у діапазоні A1:A5, у яких вказаний будь-який текст
=СЧЕТЕСЛИ(A1:A5;"???pc")	Повертає кількість комірок у діапазоні A1:A5, у яких вказане будь-яке слово з 5 букв, яке закінчується на <i>pc</i> (наприклад <i>GDPpc</i>)
=СЧЕТЕСЛИ(A1:A5;"так")/ЧСТРОК(A1:A5)	Повертає середню кількість комірок, коли зустрічається слово <i>так</i> (наприклад, якщо в стовпчику A вказується наявність валютної кризи (<i>так</i> або <i>ні</i>) у певній країні у певний рік)

=СЧЁТЕСЛИМН(діапазон умови; критерій; діапазон умови; критерій;...)

Повертає кількість комірок, які відповідають одразу декільком критеріям. Приклади:

Таблиця Б.18.

=СЧЁТЕСЛИМН(B1:B5;"=так";C1:C5;">0")	Повертає кількість рядків, у яких у діапазоні B1:B5 комірки мають значення <i>так</i> , а в діапазоні C1:C5 комірки мають число більше 0
=СЧЁТЕСЛИМН(B1:B5;"=так";C1:C5;"<"&E4)	Повертає кількість рядків, у яких у діапазоні B1:B5 комірки мають значення <i>так</i> , а в діапазоні C1:C5 комірки мають число менше, ніж вказане у комірці E4
=СЧЁТЕСЛИМН(B1:B5;"=так";C1:C5;">31.12.2000")	Повертає кількість рядків, у яких у діапазоні B1:B5 комірки мають значення <i>так</i> , а в діапазоні C1:C5 комірки мають дату пізніше 31 грудня 2000 року

=ЧАСТОТА(масив даних; масив інтервалів)

Повертає масив чисел, які вказують на частоту появи даних у певних діапазонах значень. Є формулою масиву (а отже вводиться за допомогою CTRL+SHIFT+ENTER). Приклад:

Таблиця Б.19.

	А	В	С
1	Товари	Ставка імпорного тарифу	Інтервали
2	Товар 1	0	5
3	Товар 2	5	10
4	Товар 3	0	20
5	Товар 4	25	
6	Товар 5	12	
7	Товар 6	2	
8		3	
9		1	
10		1	
11		1	

Дані в діапазоні B8:B11 з'явилися після вставки туди формули масиву **=ЧАСТОТА(B2:B7;C2:C4)**

Функція ЧАСТОТА у поєднанні з функціями СУММ та ЕСЛИ дозволяє проводити сумування унікальних значень. Приклад:

Таблиця Б.20.

	А
1	3
2	8
3	7
4	8
5	1
6	3
7	1

=СУММ(ЕСЛИ(ЧАСТОТА(A1:A7;A1:A7)>0;A1:A7)) В сумі дає 3+8+7+1=15.

Функція ЧАСТОТА разом з іншими функціями дозволяє проводити розрахунок кількості унікальних значень. Наприклад, див. Табл.Б.21-Б.22:

Таблиця Б.21.

	А	В
1	Україна	3
2	Білорусь	8
3	Болгарія	
4	Польща	8
5	Україна	1
6	Польща	3
7	Україна	1

Таблиця Б.22.

=СУММ(ЕСЛИ(ЧАСТОТА(A1:B7;A2:B7)>0;1))	Повертає кількість унікальних числових значень у діапазоні A1:B7 без урахування пустих комірок та тексту (3)
=СУММ(ЕСЛИ(ЧАСТОТА(ПОИСКПОЗ(A1:A7;A1:A7;0);ПОИСКПОЗ(A1:A7;A1:A7;0))>0;1))	Повертає кількість унікальних текстових або числових значень у діапазоні A1:A7, діапазон не має містити пусті комірки (4)
=СУММ(ЕСЛИ(ЧАСТОТА(ЕСЛИ(ДЛСТР(A1:A7)>0;ПОИСКПОЗ(A1:A7;A1:A7;0);""));ЕСЛИ(ДЛСТР(A1:A7)>0;ПОИСКПОЗ(A1:A7;A1:A7;0);""))>0;1))	Повертає кількість унікальних числових або текстових значень у діапазоні A1:B7 без урахування пустих комірок та тексту (4)

=СЦЕПИТЬ(масив даних; масив інтервалів)

Об'єднує декілька стовпчиків з даними. Наприклад, у D2 вставимо формулу =СЦЕПИТЬ(A2;" ";B2;" ";C2) і скопіюємо її у D3.

Таблиця Б.23.

	А	В	С	Д
1	Показник	Одиниця виміру	Ціни	Повна характеристика показника
2	Експорт	долар	за місяць	Експорт,долар-за місяць
3	М2	національна валюта	на кінець місяця	М2,національна валюта-на кінець місяця

Б.13. Надбудови

Надбудовами в Microsoft Excel є компоненти, які надають додаткові можливості. Стандартними надбудовами є "Пакет аналіза" і "Поиск решения", які включені вже до Microsoft Excel (існують й зовнішні надбудови, наприклад PhStat). Але у вже інстальованому Microsoft Excel надбудови потрібно додатково інсталиувати або активувати.

Для активації стандартних надбудов потрібно відкрити вкладку *Файл*, натиснути *Параметри*, обрати *Надстройки*. У полі *Управление* обрати *Надстройки Excel* і натиснути *Перейти*. Встановіть відмітки проти *Пакет аналіза* та *Поиск решения* та натисніть *ОК*. Після цього у вкладці *Данные* в групі *Анализ* можна побачити кнопки *Анализ данных* та *Поиск решения*. При натисканні

кнопки *Анализ данных* з'являється вікно у якому можемо вибрати конкретний інструмент аналізу (див. Рис.Б.2).

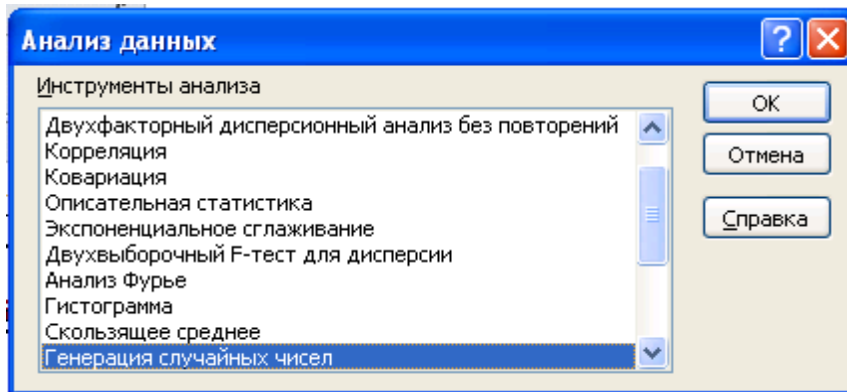


Рис.Б.2.

Наприклад, інструмент *Выборка* створює з вихідного діапазону (який розглядається як генеральна сукупність) вибірку. Це може бути корисним, наприклад, коли побудований за всіма даними графік буде занадто перевантажений графічними елементами. Цей інструмент дозволяє вибрати кожне n -не значення з діапазону, який вказаний у полі *Входной интервал* (якщо обрати *Периодический* й вказати у полі *Период* число n , наприклад 5) або створити вибірку певного розміру за випадковим принципом (якщо обрати *Случайный* і вказати кількість елементів вибірки у *Число выборок*). Якщо у першому рядку вказаного діапазону вказані назви стовпчиків (роки, назви показників тощо) потрібно поставити позначку у полі *Метки*. Наприклад, нижче на малюнках вказані вікно інструменту (Рис. Б.3), вхідний інтервал (Рис. Б.4) та результати (Рис. Б.5) у новому аркуші, якщо ми хочемо одержати вибірку з 4 спостережень.

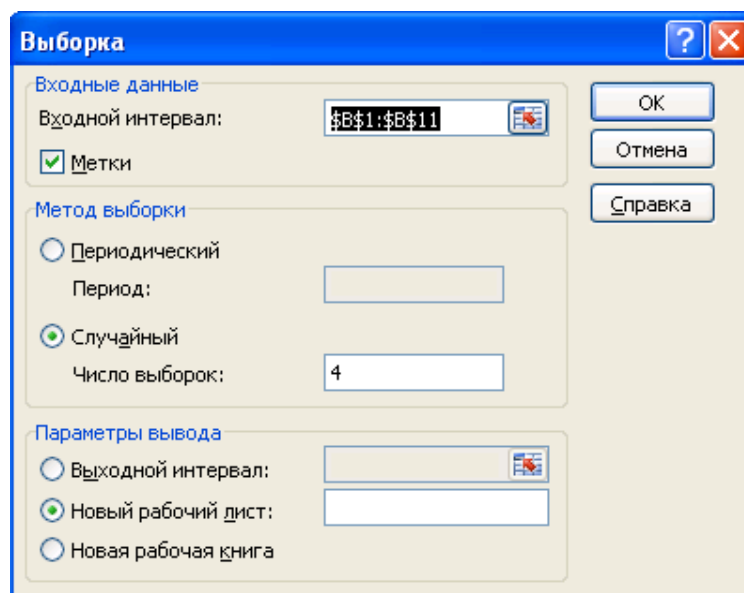


Рис.Б.3.

Інструмент *Скользящее среднее* використовується для розрахунку плинної середньої. Наприклад, зробимо це для даних експорту товарів та послуг України (за даними World Development Indicators). У полі *Интервал* вкажемо кількість періодів, за які розраховується плинна середня (наприклад, 3). Додатково можемо розрахувати стандартні похибки та показати графік. На Рис. Б.6 вказані вхідні дані, на Рис Б.7 – результат у таблиці, на Рис.Б.8 – діалогове вікно аналізу, на Рис.Б.9 – графік.

Схоже призначення має інструмент *Экспоненциальное сглаживание*.

	A	B
1	Країна	Експорт
2	Країна1	100
3	Країна2	200
4	Країна3	5000
5	Країна4	500
6	Країна5	1000
7	Країна6	40
8	Країна7	10
9	Країна8	1200
10	Країна9	2000
11	Країна10	300

Рис.Б.4.

	A
1	1200
2	500
3	5000
4	300

Рис.Б.5.

	A	B
1	Роки	Експорт
2	2000	19.5
3	2001	21.1
4	2002	23.4
5	2003	29.0
6	2004	41.3
7	2005	44.4
8	2006	50.2
9	2007	64.0
10	2008	85.6
11	2009	54.3
12	2010	69.3

Рис.Б.6.

	A	B	C
1	Роки	Експорт	
2	2000	19.5	#Н/Д
3	2001	21.1	#Н/Д
4	2002	23.4	21.3
5	2003	29.0	24.5
6	2004	41.3	31.2
7	2005	44.4	38.2
8	2006	50.2	45.3
9	2007	64.0	52.9
10	2008	85.6	66.6
11	2009	54.3	68.0
12	2010	69.3	69.7

Рис.Б.7.

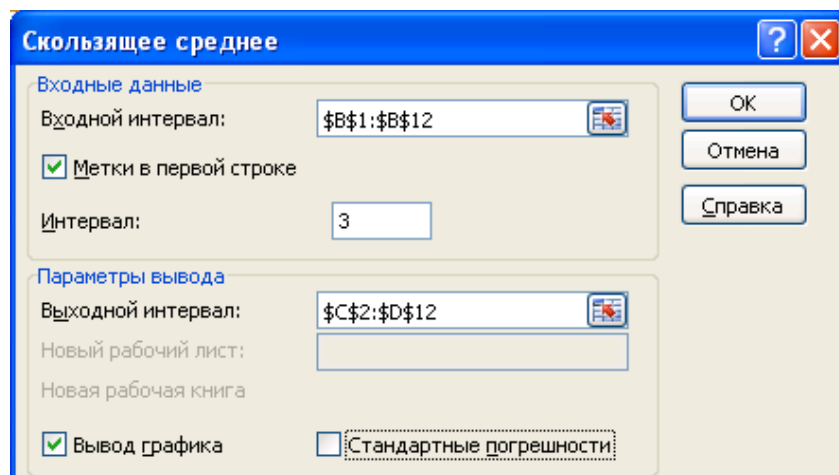


Рис.Б.8.

Б.14. Діаграми

В програмі доступні такі види діаграм:

- Гістограми (стовпчасті діаграми);
- Лінійні діаграми;
- Графіки (для визначення динаміки змінних з часом або по впорядкованих категоріях);
- Діаграми з областями (мають схоже призначення);
- Точкові (крапкові) діаграми або діаграми розсіювання (для визначення залежностей між змінними);

Скользящее среднее



Рис.Б.9

- Біржові діаграми;
- Поверхневі діаграми (тривимірні);
- Кругові діаграми (для відображення часток цілого);
- Кільцеві діаграми (мають схоже призначення, але дають можливість відобразити декілька рядів даних);
- Бульбашкові діаграми (тривимірні);
- Пелюсткові діаграми (багатовимірні).

Для створення діаграм потрібно:

- ввести вхідні дані (залежно від типу діаграми вони мають бути впорядковані певним чином);
- виділити дані;
- на вкладці *Вставити* обрати *Діаграмми*.
- обрати тип діаграми.

Далі за потреби елементи діаграми редагуються. До діаграм можливо додавати лінії тренду та рівняння регресійної залежності.

Додаток Г. Формування бази вхідних даних

Г.1. Підготовка вхідних даних із зовнішніх джерел

Варіантів формування бази даних може бути декілька. Ми опишемо один з шляхів, з яким часто можна зустрітися, аналізуючи дані по багатьох об'єктах (наприклад, країнах) та періодах часу. На кожному кроці бажано проводити процедури перевірки, щоб не допустити помилки. Помилку набагато легше виправити одразу після її виникнення, ніж пізніше шукати, де саме вона виникла.

Наприклад, всі показники або окрема група показників можуть бути представлені в одній книзі Microsoft Excel. Кожному показнику виділяється один аркуш. У другому стовпчику зверху до низу йдуть країни в алфавітному порядку (як правило англійською мовою). У другому рядку зліва направо перераховуються роки (або квартали чи місяці). Ми радимо робити це, наприклад, у другому стовпчику та другому рядку (а не першому) для того, щоб у таблиці було вільне місце для таких коментарів, як назва показника, одиниці виміру, спосіб розрахунку тощо.

На Рис Г.1 зображений фрагмент аркушу, де розміщені дані по одному показнику: коротко-строковий зовнішній борг (за даними World Development Indicators). У даному прикладі, для зручності відображення великих значень, числа показані у науковому стилі. Якби великі числа відображалися би не у науковому стилі, а стовпчики були б занадто вузькими, ми би бачили ##### замість значень (див. підрозділ 2.7).

	A	B	C	D	E	F	G	H	I	J	K	L
1	External debt stocks, short-term (DOD, current US\$)											
2				2000	2001	2002	2003	2004	2005	2006	2007	2008
3		Afghanistan								18152000	20997000	16926000
4		Angola		1.32E+09	1.45E+09	1.21E+09	1.07E+09	1.2E+09	2.32E+09	2.13E+09	2.27E+09	2.42E+09
5		Albania		36666000	30619000	29046000	1.49E+08	1227000	2.83E+08	5.91E+08	8.27E+08	7.79E+08
6		Argentina		2.83E+10	2E+10	1.48E+10	2.23E+10	2.65E+10	3.48E+10	3.36E+10	3.81E+10	3.75E+10
7		Armenia		44419000	41972000	2.2E+08	4.04E+08	4.1E+08	2.98E+08	3.07E+08	4.59E+08	4.65E+08
8		Azerbaijan		1.56E+08	1.03E+08	82380000	1.03E+08	1.38E+08	1.86E+08	5.2E+08	1.04E+09	1.17E+09
9		Burundi		65000000	88395000	96346000	47562000	22560000	34059000	37617000	13785000	19303000
10		Benin		65170000	78650000	73610000	33606000	28606000	44045000	44068000	5184000	37606000
11		Burkina Faso		84312000	63472000	12918000	14139000	24096000	22222000	89063000	1.55E+08	1.1E+08
12		Bangladesh		3.34E+08	3.61E+08	5.72E+08	6.17E+08	7.12E+08	6.88E+08	1.18E+09	1.38E+09	1.99E+09
13		Bulgaria		1.45E+09	1.22E+09	1.84E+09	2.66E+09	3.26E+09	4.44E+09	8.04E+09	1.4E+10	1.85E+10
14		Bosnia and Herzegovina		49331000	59730000	3.49E+08	1.13E+08	3.6E+08	8.37E+08	1.17E+09	1.69E+09	9.12E+08
15		Belarus		1.24E+09	1.31E+09	1.63E+09	1.97E+09	2.94E+09	3.5E+09	3.65E+09	6.88E+09	6.96E+09
16		Belize		50000000	50844000	45055000	80000000	319000	5996000	6615000	6358000	6558000
17		Bolivia		4.02E+08	3.8E+08	3.7E+08	3.32E+08	2.72E+08	1.82E+08	2.2E+08	1.77E+08	1.66E+08
18		Brazil		3.1E+10	2.83E+10	2.34E+10	2.46E+10	2.53E+10	2.4E+10	2.03E+10	3.92E+10	3.67E+10

Рис.Г.1.

Потрібно одразу визначитися із колом об'єктів і періодів, які будуть досліджуватися. Наприклад, часто не включають країни, які невеликі за розміром, по яких відсутня переважна частина даних, або країни які різко відрізняються за своїми характеристиками, наприклад, від України, якщо нас цікавить застосування результатів подальшого аналізу для України. Щодо періодів, то для більшості показників доступні лише річні дані, але мета аналізу може вимагати використання коротших періодів (місяці, квартали, дні). Бажано включати до всього досліджуваного періоду, з одного боку, більше років. З іншого боку дані за ранні роки можуть бути недоступними (для України, наприклад, раніше 1995 року важко знайти дані по багатьох показниках, або їх надійність викликає суттєвий сумнів), так само як і за останній період (хоча іноді можливо використовувати більш менш надійні попередні оцінки чи прогнозні дані, наприклад, у жовтні місяці прогнозні дані на весь поточний рік). Також система зв'язків між явищами з часом суттєво змінюється. Тому часто варто не включати до аналізу занадто давні періоди минулого або принаймні надавати їм меншу вагу у розрахунках.

Потрібно пересвідчитися, що у різних аркушах (або книгах) по різних показниках значення по тій самій країні у той самий період розташовані у комірках з тією самою адресою в межах аркуша. Це забезпечується, наприклад, тим що, якщо в усіх аркушах дані по Україні знаходяться у рядку з одним й тим самим номером, а дані за 2010 рік – у стовпчику з одним й тим самим номе-

ром. Нажаль, особливо використовуючи дані з різних джерел, часто це досягається вручну. Перевірка на однаковий склад і порядок країн можливо зробити, скопіювавши стовпчик з назвами країн з однієї таблиці (який вважаємо за еталон) у решту таблиць, наприклад у перший стовпчик кожної таблиці. У останньому стовпчику прописуються формули типу =ЕСЛИ(A2=B2;1;0). Якщо ми побачимо в результаті у цьому стовпчику в певній комірці 0 замість 1, це означатиме, що послідовність країн не співпадає з еталонною і це потрібно виправити.

Наприклад, на Рис. Г.2 зображений фрагмент аркушу де ми розмістили дані по іншому показнику: гранти у формі анулювання боргу (за даними World Development Indicators), але при цьому у другому стовпчику скопіювали стовпчик з еталонним переліком країн, а в комірці А3 прописали функцію =ЕСЛИ(B3=C3;1;0) і скопіювали нижче у стовпчику. Як бачимо у першому стовпчику всюди результатом є 1.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Debt forgiveness grants (current US\$)											
2				2000	2001	2002	2003	2004	2005	2006	2007	2008
3	1	Afghanistan	Afghanistan							0	31300000	14900000
4	1	Angola	Angola	1320000	60000	0	0	0	0	0	0	30020000
5	1	Albania	Albania	2040000	0	0	0	0	0	0	0	0
6	1	Argentina	Argentina	0	0	0	9210000	0	0	0	3410000	0
7	1	Armenia	Armenia	0	0	0	0	0	0	0	0	0
8	1	Azerbaijan	Azerbaijan	0	0	0	0	0	0	0	0	0
9	1	Burundi	Burundi	5630000	4490000	3980000	2940000	7080000	14190000	3980000	24440000	44150000
10	1	Benin	Benin	32720000	21360000	24240000	58100000	88760000	22500000	1.01E+09	10840000	9820000
11	1	Burkina Faso	Burkina Faso	1.9E+08	31160000	39090000	48580000	55530000	40510000	1.22E+09	16710000	21250000
12	1	Bangladesh	Bangladesh	1.79E+08	1.56E+08	1.74E+08	93660000	2.71E+08	40950000	2.39E+08	1.29E+08	7.55E+08
13	1	Bulgaria	Bulgaria	1560000	0	0	0	0	0	0	0	0
14	1	Bosnia and He	Bosnia and He	1.25E+08	4250000	3590000	24020000	3590000	3790000	4610000	5990000	6830000
15	1	Belarus	Belarus	0	0	0	0	0	0	0	0	0
16	1	Belize	Belize	0	0	2190000	2380000	2670000	1750000	1770000	1290000	0
17	1	Bolivia	Bolivia	24090000	1.52E+08	4.06E+08	1.17E+08	5.27E+08	37590000	1.82E+09	1.18E+09	2990000
18	1	Brazil	Brazil	0	0	0	0	0	0	0	0	0

Рис.Г.2.

На цьому етапі також варто пересвідчитися, що числа прописані належним чином. Наприклад, якщо в системі Windows розділовим значком дробної частини є крапка, а число має вигляд як 3,45 (тобто з комою), то воно сприйматиметься як текст. Так само як і число типу 3 500, оскільки воно містить пробіл. Це можливо виправити командою *Заменить*, але потрібно впевнитися, що заміни зроблено правильно.

Г.2. Підготовка додаткових розрахованих показників

Припустимо, ми сформували базу даних запозичених із зовнішніх джерел показників. Наступним кроком є вставка нових аркушів, де формуються таблички по нових, розрахованих, показниках. На перетині стовпчиків-років і рядків-країн прописуються формули. Наприклад, якщо ми хочемо розрахувати показник Валютні резерви (частка від ВВП) для комірки С3, вставляємо туди формулу типу: =FReserveUSD!C3/'GDPUSD!C3, де FReserveUSD та GDP – назви відповідних аркушів, де містяться дані про валютні резерви (в дол.) та ВВП (в дол.). Далі ця формула копіюється у решту потрібних комірок аркушу, де міститимуться дані про Валютні резерви (частка від ВВП).

Наприклад, на Рис Г.3 зображений фрагмент аркушу, де ми розмістили дані про розрахований нами подібним чином показник: відношення грантів у формі анулювання боргу до короткострокового зовнішнього боргу. Залишається лише видалити значення з помилками (комірки з даними по Афганістану в 2000-2005 рр. стануть порожніми).

Перевірка на помилки може бути проведена, зокрема, перевіряючи, чи не вийшли за межі допустимих значень показники. Наприклад, не може бути більше 1 чи 100% частка сільськогосподарської продукції у експорті. Також можливо перевірити результат розрахунку вручну у певній комірці. Якщо помилка викликана неправильною формулою, потрібно скоригувати формулу. Якщо – неправильним чи незвичним значенням показника, на основі якого проводиться розрахунок формули, бажано за можливості виправити це значення, якщо йому не можна довіряти, інакше його потрібно сприймати як відсутні дані.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Debt forgiveness grants (current US\$)/External debt stocks, short-term (DOD, current US\$)											
2				2000	2001	2002	2003	2004	2005	2006	2007	2008
3			Afghanistan	#ДЕЛ/0!	#ДЕЛ/0!	#ДЕЛ/0!	#ДЕЛ/0!	#ДЕЛ/0!	#ДЕЛ/0!	0	1.490689	0.880302
4			Angola	0.000998	4.13E-05	0	0	0	0	0	0	0.012409
5			Albania	0.055637	0	0	0	0	0	0	0	0
6			Argentina	0	0	0	0.000413	0	0	0	8.96E-05	0
7			Armenia	0	0	0	0	0	0	0	0	0
8			Azerbaijan	0	0	0	0	0	0	0	0	0
9			Burundi	0.086615	0.050795	0.041309	0.061814	0.31383	0.41663	0.105803	1.772942	2.287209
10			Benin	0.502072	0.271583	0.329303	1.728858	3.102846	0.510841	22.87465	2.091049	0.261129
11			Burkina Fas	2.252111	0.490925	3.02601	3.435887	2.304532	1.822968	13.66157	0.107795	0.193106
12			Bangladesh	0.536766	0.433328	0.304216	0.151799	0.380112	0.059483	0.202784	0.093869	0.380351
13			Bulgaria	0.001076	0	0	0	0	0	0	0	0
14			Bosnia and I	2.531877	0.071154	0.010296	0.212566	0.009972	0.004528	0.003937	0.003551	0.007489
15			Belarus	0	0	0	0	0	0	0	0	0
16			Belize	0	0	0.048607	0.02975	8.369906	0.291861	0.267574	0.202894	0
17			Bolivia	0.05986	0.399289	1.098081	0.35272	1.935499	0.206823	8.268	6.641243	0.018012
18			Brazil	0	0	0	0	0	0	0	0	0

Рис.Г.3.

Можливо, що в формулах потрібно використати додаткові коефіцієнти. Наприклад, якщо ми хочемо дані у відсотках, а не у частках одиниці, формула у попередньому прикладі виглядати-ме як ='FReserveUSD'!C3*100/'GDPUSD'!C3. Вимірювання у відсотках сприймаються візуально легше. Але вимірювання у частках одиниці зручніше, якщо розрахований показник, буде далі ви-користовуватися для розрахунку нових показників.

Якщо, наприклад, ВВП вимірюється у дол., а валютні резерви у млн. дол., потрібна корекція на одиниці виміру: ='FReserveUSD'!C3*1000000/'GDPUSD'!C3. Зокрема, у World Development Indicators часто показники вимірюються безпосередньо у доларах чи одиницях національної ва-люти, у той час як в інших джерелах зазвичай вимірюють у млн. чи тис. доларів одиниць націона-льної валюти.

Якщо показники вимірюються у різних валютах (наприклад, ВВП у національній валюті), використовуватиметься формула типу: ='FReserveUSD'! C3*'ExR'!C3/'GDPLCU'!C3, де ExR назва аркушу де містяться дані про валютний курс у одиницях національної валюти (*local currency unit*) за долар. При використанні валютних курсів потрібно пересвідчитися, що використовується пот-рібний вид котировки, при цьому для всіх країн. Не варто, щоб наприклад, у аркуші з даними про валютний курс у рядку Великобританія курс був у доларах за фунт стерлінгів, а у рядку Україна курс був у гривнях за долар – варто робити всі курси у одиницях національної валюти за долар. Для показників за період варто використовувати середній курс за період, для показників на кі-нець періоду – курс на кінець періоду. Але іноді важко робити вибір, якщо порівнюється один показник за період, а інший на кінець періоду.

Розраховані показники можуть бути і достатньо складними. Наприклад для оцінки впливу валютних криз на економічне зростання як залежна змінна в [47] розраховувався коефіцієнт прис-корення економічного зростання (*КПЕЗ₀*) за формулою:

$$КПЕЗ_0 = \frac{(1 + y_{t+1})^{1/3} \times (1 + y_{t+2})^{1/3} \times (1 + y_{t+3})^{1/3}}{\sqrt{\left((1 + y_{t-1})^{1/2} \times (1 + y_{t-2})^{1/3} \times (1 + y_{t-3})^{1/6} \right) \times \left(1 + y_{t-1} + \frac{2}{3}(y_{t-1} - y_{t-2}) + \frac{1}{3}(y_{t-2} - y_{t-3}) \right)}}$$

де y_t – приріст ВВП року t (з лагом 0 років відносно фактора).

За допомогою цієї формули порівнюється економічне зростання в наступні три роки (переве-дене на річну основу) – у чисельнику, з геометричним середнім економічного зростання попередніх трьох років і трендовими його значеннями – у знаменнику. Останнім спостереженням надається більша вага, ніж більш давнім. Якщо $КПЕЗ_0 > 1$, за інших рівних умов відповідна подія (наприклад, криза) у рік t , імовірно, сприяє економічному зростанню, якщо $КПЕЗ_0 < 1$, то – протидіє.

Також на цьому етапі потрібно зробити порожнімі комірки, по яких дані відсутні (оскільки іноді вони позначаються текстом, наприклад *н.д.*, -, *NA* тощо) або прописана помилка в результаті розрахунків. Це важливо, оскільки у подальшому аналіз може не дати результати або дати не ті результати, якщо будуть комірки з текстом. Але також варто пересвідчитися, що формули не сприймають порожнімі комірки як ті, що містять значення 0.

Далі ми можемо діяти, наприклад, таким чином. Розрахувавши всі потрібні нам показники у книзі, ми можемо створити нову книгу, куди скопіювати всі дані з попередньої книги, але вже у формі значень, а не формул. Подібну операцію варто зробити, оскільки інакше при подальших трансформаціях бази даних може статися ситуація, коли після переміщення формули можуть посылатися не на потрібні комірки або посилання стають недійсними. В результаті виникають помилки у результатах.

Г.3. Характеристики показників

Паралельно варто зробити систематизований опис показників і надати їм скорочену назву, яку буде зручно використовувати для назв аркушів чи у формулах для опису ходу дослідження. Кожний дослідник може сам обрати зручні для нього позначення (наприклад, GDP_n – ВВП у національній валюті, $FRes/GDP$ – валютні резерви відносно ВВП, GDP_{pc} – ВВП на душу населення, Ex_{grg} – зростання експорту, Ex_{rgrg} – зростання реального експорту, $PrivFDebt/FDebt$ – частка приватного боргу у зовнішньому боргу тощо).

Опис може включати:

- повну і скорочену назви;
- джерело (можливо з кодом рядка у джерелі), якщо показник запозичений;
- одиниці виміру;
- формулу розрахунку за потреби;
- примітки – додаткові дані про характер показника (наприклад, на кінець періоду чи в середньому за рік), чи використовуються лагові значення, спосіб заміни відсутніх даних (якщо замінювалися відсутні дані на розрахункові), максимальні та мінімальні значення (наприклад 10 – найвищий рівень економічної свободи, 0 – найнижчий), варіанти значень (наприклад, 1 – відсутні збройні конфлікти, 2 – обмежений збройний конфлікт, 3 – війна), особливості для окремої групи країн, за який період відсутні дані тощо.

Показники можливо класифікувати залежно від шкали виміру (перші дві використовуються для якісних змінних (такі змінні ще можливо назвати категоріальними), остання – для кількісних):

- номінальна шкала визначає приналежність об'єкту до певного класу, які не мають відносин підпорядкованості між собою (наприклад, показник регіон (Європа, Азія, Африка тощо)) і не можуть бути трансформовані у кількісні;
- порядкова шкала – визначає порядок (ранжує) об'єкти, але не визначає відстані між ними (наприклад, наявність чи відсутність регіональної торговельної угоди або, детальніше, відсутність регіональної торговельної угоди, угода с обмеженим предметом дії, зона вільної торгівлі, митний союз); якщо використовувати бальні оцінки, то такі якісні показники можливо умовно трансформувати в кількісні;
- метрична шкала – визначає величину інтервалу між значеннями показника (наприклад, експорт).

Можливо спробувати один й той самий загальний показник включити у базу даних у різних формах: наприклад, валютні резерви:

- в абсолютному вимірі;
- відносно інших змінних (ВВП, зовнішній борг, короткостроковий зовнішній борг, імпорт та доходи сплачені, грошовий агрегат M_2);
- приріст у відсотках чи частках одиниці;
- приріст відносно інших змінних (наприклад відносно ВВП: валютні резерви/ВВП у поточному періоді мінус валютні резерви/ВВП у попередньому періоді);
- як частка відносно сукупного показника (наприклад частка країни у валютних резервах світу, або частка валютних резервів у всіх активах у міжнародній інвестиційній позиції країни);
- логарифм;
- відносно середнього рівня у світі;
- трендові значення наступного періоду (наприклад, $=2 \cdot$ валютні резерви поточного року – валютні резерви попереднього року = валютні резерви поточного року + зміна валютних резервів порівняно з попереднім роком);
- відносно теоретичного рівня на основі регресії (приклад для іншого показника: співвідношення валютного курсу і коефіцієнту паритету купівельної спроможності – наскільки

його фактичне значення відрізняється від теоретичного, розрахованого на основі регресії, де незалежною змінною є відношення ВВП на душу населення країни до середньосвітового рівня ВВП на душу населення);

- стандартизоване значення (у стандартних відхиленнях відносно середнього значення: за всі періоди у певній країні, по всіх країнах за певний період або по всіх країнах за всі періоди);
- категоризоване значення одного з вказаних вище варіантів (наприклад, 1 – низькі валютні резерви, 2 – середні, 3 – високі), коли змінна у метричній шкалі перетворюється на змінну у порядковій шкалі.

Певну проблему можуть створити показники, які мало змінюються з часом. З ними може виникнути спотворюючий ефект завищення кількості спостережень. Припустимо, ми маємо 10 років і 3 країни, показник приймає такі значення: 0 – мир, 1 – війна. Перші дві країни завжди у стані миру. А третя – завжди у стані війни. В результаті у нас є 30 спостережень, але вони несуть інформацію як і 3 спостереження взяті за будь-який рік. Якщо ми аналізуємо взаємозв'язок такого проблемного показника зі звичайним, то спотворюючий ефект буде менший. Але якщо ми досліджуємо взаємозв'язок двох чи більше таких проблемних показників, то спотворюючий ефект буде суттєвим. Наприклад, критерії якості регресії можуть вказувати на її адекватність, розраховуючи на те, що є 30 спостережень, хоча насправді вони мали б бути розраховані виходячи з 3 спостережень. Аналогічну проблему можуть створювати глобальні показники, наприклад світові ціни на нафту будуть однакові для всіх країн протягом одного періоду.

Г.4. Формування узагальнюючої таблиці

Перед зведенням всіх даних до єдиної таблиці варто передбачити лаги. Наприклад, створити декілька аркушів з даними по залежній змінній, де в одному аркуші дані будуть зміщені на 1 стовпчик праворуч (лаг 1 період), у другому – на 2 (лаг два періоди) і т.д.

Наприклад, ми будемо розглядати як залежну змінну зовнішній борг (короткостроковий). Решту змінних – як незалежні змінні. Створимо новий аркуш, де міститимуться дані по залежній змінній з лагом у 1 рік. Скопіюємо туди дані з аркушу де містяться дані по залежній змінній (користуючись опцією *значення* під час вставки), після чого змістимо на 1 стовпчик ліворуч всі значення (крім найбільш раннього періоду-стовпчика) за допомогою операцій вирізати та вставити. Одержимо таблицю як на Рис.Г.4.

	A	B	C	D	E	F	G	H	I	J	K	L
1	External debt	X		2000	2001	2002	2003	2004	2005	2006	2007	2008
2		Y		2001	2002	2003	2004	2005	2006	2007	2008	
3		Afghanistan							18152000	20997000	16926000	
4		Angola		1.45E+09	1.21E+09	1.07E+09	1.2E+09	2.32E+09	2.13E+09	2.27E+09	2.42E+09	
5		Albania		30619000	29046000	1.49E+08	1227000	2.83E+08	5.91E+08	8.27E+08	7.79E+08	
6		Argentina		2E+10	1.48E+10	2.23E+10	2.65E+10	3.48E+10	3.36E+10	3.81E+10	3.75E+10	
7		Armenia		41972000	2.2E+08	4.04E+08	4.1E+08	2.98E+08	3.07E+08	4.59E+08	4.65E+08	
8		Azerbaijan		1.03E+08	82380000	1.03E+08	1.38E+08	1.86E+08	5.2E+08	1.04E+09	1.17E+09	
9		Burundi		88395000	96346000	47562000	22560000	34059000	37617000	13785000	19303000	
10		Benin		78650000	73610000	33606000	28606000	44045000	44068000	5184000	37606000	
11		Burkina Faso		63472000	12918000	14139000	24096000	22222000	89063000	1.55E+08	1.1E+08	
12		Bangladesh		3.61E+08	5.72E+08	6.17E+08	7.12E+08	6.88E+08	1.18E+09	1.38E+09	1.99E+09	
13		Bulgaria		1.22E+09	1.84E+09	2.66E+09	3.26E+09	4.44E+09	8.04E+09	1.4E+10	1.85E+10	
14		Bosnia and Herzegovina		59730000	3.49E+08	1.13E+08	3.6E+08	8.37E+08	1.17E+09	1.69E+09	9.12E+08	
15		Belarus		1.31E+09	1.63E+09	1.97E+09	2.94E+09	3.5E+09	3.65E+09	6.88E+09	6.96E+09	
16		Belize		50844000	45055000	80000000	319000	5996000	6615000	6358000	6558000	
17		Bolivia		3.8E+08	3.7E+08	3.32E+08	2.72E+08	1.82E+08	2.2E+08	1.77E+08	1.66E+08	
18		Brazil		2.83E+10	2.34E+10	2.46E+10	2.53E+10	2.4E+10	2.03E+10	3.92E+10	3.67E+10	

Рис.Г.4.

Так само можливо створити додаткові аркуші, де залежна змінна (або змінні якщо їх багато) матиме значення зміщені на 2 роки (2 стовпчики), 3 роки і т.д.). Незалежні змінні залишаємо без зміщення.

Далі таблиця по кожному показнику перетворюється у стовпчик. Тобто у стовпчику будуть йти дані по всіх по черзі країнах за перший період, нижче по всіх країнах за другий період і т.д. або інший варіант – по першій країні за всі роки, далі по другій країні за всі роки і т.д. Наприклад, для змінної короткостроковий зовнішній борг результат буде як на Рис. Г.5 (фрагмент):

	A	B	C	D	E
1			External debt stocks, short-term (DOD, current US\$)		
2	Afghanistan	2000			
3		2001			
4		2002			
5		2003			
6		2004			
7		2005			
8		2006	18152000		
9		2007	20997000		
10		2008	16926000		
11	Angola	2000	1322922000		
12		2001	1451099000		
13		2002	1208386000		
14		2003	1074998000		
15		2004	1196102000		
16		2005	2315756000		
17		2006	2132157000		
18		2007	2271190000		
19		2008	2419205000		
20	Albania	2000	36666000		
21		2001	30619000		
22		2002	29046000		

Рис.Г.5.

Це можливо робити вручну. Але з метою автоматизації можливо прописати макрос, зробивши вручну це за лише однією таблицею під час запису макросу, а для решти таблиць викликати макрос.

Такі стовпчики з кожної таблиці копіюються в одну узагальнюючу таблицю, де кожному стовпчику відповідає змінна, а кожному рядку спостереження (наприклад, країно-рік), при цьому резервуємо перший рядок під назви змінних, а перший стовпчик – під назви спостережень. Тепер таблицю можливо аналізувати у Microsoft Excel, або скопіювати у іншу програму та аналізувати там.

Г.5. Проблеми з відсутніми даними та їх діагностика

Повернемося детальніше до проблеми відсутніх даних. Іноді частка відсутніх даних настільки велика, що вибірка з наявними даними є занадто малою. В такому випадку потрібно знайти спосіб вирішити цю проблему.

Відсутні дані може бути результатом декількох причин, зокрема:

- Відсутності статистичних даних.
- Виходом розрахованого дослідником показника за межі допустимих значень. Наприклад, якщо потрібно розрахувати логарифм показника, який у певному випадку приймає від'ємне значення.
- Неприйнятності певного показника для певної групи спостережень (об'єктів або років). Наприклад, такий показник як кумулятивний (сумарний за періоди) дефіцит поточного рахунку платіжного балансу у % від ВВП з часу виникнення такого дефіциту можливо виміряти лише для країн у яких такий дефіцит існував хоча б один раз за всю історію визначення статистичних даних. Або кількість років до наступних виборів можливо виміряти лише для країн, де вибори відбуваються.

Відсутні дані можуть становити небезпеку з точки зору практичного застосування результатів аналізу, якщо відсутність даних є результатом певної закономірності. Наприклад, припустимо, ми аналізуємо вплив рівня розвитку країни на її міжнародну інвестиційну позицію. При цьому, як правило, у менш розвинутих країнах ці статистичні дані є менш доступними. В результаті використовуючи регресійний або інший вид аналізу ми можемо одержати певну закономірність, але наскільки вона буде надійною, враховуючи те, що країни з низьким рівнем розвитку та частина країн з середнім рівнем розвитку не розглядаються? Тому безпечними для результатів аналізу можуть бути лише відсутні дані, які не є результатом важливої для нас закономірності. Через подібні проблеми регіон Африка часто не є достатньою мірою представленим у дослідженнях.

Діагностика закономірності у відсутніх даних. Припустимо, ми аналізуємо зв'язок між припливом інвестицій (у % від ВВП) та податком на прибуток і у нас відсутні дані по частині спостережень (наприклад, країн) за зміною податок на прибуток. Діагностика закономірності у відсутніх даних може бути проведена декількома шляхами:

1. Поділимо вибірку на дві групи (див. Рис. Г.6): перша з наявними даними за податком на прибуток і друга з відсутніми даними за податком на прибуток. Порівняємо середні значення припливу інвестицій цих двох груп. Додатково можемо перевірити за допомогою t-тесту (описаний в Главі 5) наскільки значуща ця різниця. Якщо різниця значна, закономірність існує: відсутність даних обумовлена певним рівнем одного з показників – припливом інвестицій. Тому при інтерпретації результатів подальшого аналізу потрібно бути дуже обережним.

	A	B	C	D	E	F	G	H	I
1				Група з відомими даними			Група з відсутніми даними		
2	Інвестиції	Податок		Інвестиції	Податок		Інвестиції	Податок	
3		4	20		4	20			
4		6	15		6	15			
5		1						1	
6		2						2	
7		2	30		2	30			
8		4	35		4	35			
9		3	25		3	25			
10		7	0		7	0			
11									
12			середнє	4.333		середнє	1.5		

Рис.Г.6.

2. Створимо бінарну змінну, яка приймає значення 1, якщо за податком на прибуток наявні дані, і 0, якщо відсутні (див. Рис.Г.7). Розрахуємо коефіцієнт кореляції (наприклад, Пірсона, але бажано Спірмена – див. Главу 2) між цією бінарною змінною та припливом інвестицій. Якщо кореляція висока або середня – закономірність у відсутніх даних існує.

Г.6. Вирішення проблем з відсутніми даними

Вирішити проблему відсутніх даних можливо декількома способами (спосіб вирішення варто описати в примітках у дослідженні):

1. Використовувати для аналізу лише спостереження з усіма наявними даними за всіма змінними, що нас цікавлять (*complete case approach* або *casewise*). Наприклад на Рис. Г.8, ми будемо використовувати лише позначені рамками 3 спостереження.

	A	B	C
1	Інвестиції	Податок	
2		4	20
3		6	15
4		1	0
5		2	0
6		2	30
7		4	35
8		3	25
9		7	0
10			
11			
12			
13		Кореляція Пірсона	0.634877

Рис.Г.7.

Експорт	Імпорт	ПІІ	Зовнішній борг
12	12	3	10
22	34	4	
33	23	5	30
21	22		40
34		4	50
56	22	2	
43	13	1	20

Рис.Г.8.

Цей спосіб можливо використовувати, якщо немає закономірності у відсутніх даних. Недоліком також є те, що в результаті кількість спостережень може скоротитися до неприйнятно низького рівня. В такому випадку варто додати спостереження, наприклад, за інші роки чи по інших країнах, хоча це не завжди можливо.

2. Видалити лише ті спостереження та змінні, за якими частка відсутніх спостережень є надто великою, залишаючи спостереження та змінні без відсутніх даних або з мінімальною кількістю відсутніх даних. У прикладі на Рис.Г.9 ми використовуємо лише перші три змінні, а останню змінну (зовнішній борг), де частка відсутніх даних найбільша, використовувати не будемо.

Часто це є оптимальним шляхом і дозволяє суттєво зменшити частку відсутніх даних. Також за цього способу видаляються спостереження з відсутніми даними по залежній змінній. Замість важливих видалених змінних бажано включити до аналізу змінні, які значно корелюють з видаленими змінними.

3. Попарне (*Pairwise*) неврахування спостережень при аналізі. Наприклад, використовується при побудові кореляційної матриці. При розрахунку парної кореляції між кожними двома змінними враховуються всі спостереження, за якими наявні дані принаймні по лише цих двох змінних, хоча по інших змінних у таких спостереженнях можуть бути відсутні дані. Наприклад, у прикладі на Рис. Г.10 для розрахунку кореляції між експортом та імпортом будуть використані всі спостереження крім 5-го, а кореляції між прямими іноземними інвестиціями та зовнішнім боргом – 1-е, 3-е, 5-е та 7-е.

Експорт	Імпорт	ПІІ	Зовнішній борг
12	12	3	10
22	34	4	
33	23	5	30
21	22		40
34		4	50
56	22	2	
43	13	1	20

Рис.Г.9.

Експорт	Імпорт	ПІІ	Зовнішній борг
12	12	3	10
22	34	4	
33	23	5	30
21	22		4
34		4	50
56	22	2	
43	13	1	20

Рис.Г.10.

4. Вставка замість відсутніх даних інших даних (*Imputation Methods*):

- Вставка даних з іншого джерела. Проте потрібно пересвідчитися, що дані є порівнюваними, оскільки іноді один й той самий показник у різних джерелах може суттєво відрізнятися, наприклад, дані про короткостроковий зовнішній борг можуть відрізнятися в декілька разів залежно від джерела та способу розрахунку .
- Заміна відсутніх даних на середні значення змінної (*Mean Substitution*). Недоліком є зменшення дисперсії змінної, а також кореляції з іншими змінними. Його варто використовувати лише, якщо вид аналізу не дозволяє взагалі мати відсутні дані або якщо дисперсія змінної невелика. Наприклад, цей спосіб зазвичай можливо використати щодо приросту населення. Для нашого попереднього прикладу, результат буде як на Рис Г.11.

	Експорт	Імпорт	ПІІ	Зовнішній борг
	12	12	3	10
	22	34	4	30
	33	23	5	30
	21	22	3.166667	40
	34	21	4	50
	56	22	2	30
	43	13	1	20
Середнє	31.57143		21	3.166667

Рис.Г.11.

- Заміна відсутніх даних на значення за попередній період. Це можливо робити, якщо є упевненість, що воно мало змінилося. Наприклад, цей спосіб можна використати щодо індексу економічної свободи, але вже не можна використати щодо його зміни.
- Заміна відсутніх даних на значення, розраховані на основі тренду (плинної середньої). Це можливо робити, якщо є упевненість, що тренд є достатньо стійким. Наприклад, у нас є часовий ряд з трьома відсутніми даними (див. Рис.Г.12). Ми бачимо що показник достатньо рівномірно змінюється з часом. Розрахуємо одразу середню зміну за період з відсутніми даними. Тепер прописуємо у порожніх комірках відповідні формули (Рис.Г.13) і одержуємо результат (Рис.Г.14).

	D	E	F	G	H	I	J	K	L	M	N	O
1												
2	2000	2001	2002	2003	2004	2005	2006	2007	2008	Середня зміна за період:		
3	23	34	45	40				50	56		2.5	

Рис.Г.12.

2000	2001	2002	2003	2004	2005	2006	2007	2008
23	34	45	40	=G3+2.5	=H3+2.5	=I3+2.5	50	56

Рис.Г.13.

2000	2001	2002	2003	2004	2005	2006	2007	2008
23	34	45	40	42.5	45	47.5	50	56

Рис.Г.14.

- Заміна відсутніх даних на розраховані за допомогою регресії, знаючи значення інших змінних, які суттєво впливають на змінну, по якій відсутні дані (*Regression Imputation*). Наприклад, у нас немає даних про частку високотехнологічних товарів у експорті по деяких країнах у вибірці, але по всіх країнах з вибірки є дані по частці витрат на дослідження і розробки у ВВП. По всіх доступних спостереженнях з відомими даними ми можемо побудувати регресію, яка визначає залежність частки високотехнологічних товарів у експорті від частки витрат на дослідження і розробки у ВВП. Далі розраховуємо за допомогою регресії значення частки високотехнологічних товарів для спостережень з відсутніми даними за цим показником. Цей метод має ряд недоліків. По-перше, він посилює існуючі залежності між змінними, і результати подальшого аналізу будуть менш корисними при застосуванні їх для інших випадків, які знаходяться за межами вибірки. По-друге, занижується дисперсія змінної, якщо тільки не включити до розрахованих її значень стохастичну складову. По-третє, регресійно-кореляційний зв'язок з іншими змінними може бути не настільки надійним. По-четверте, розраховані значення можуть опинитися за межами допустимих значень (наприклад, у якомусь випадку частка сільськогосподарської продукції у експорті може бути більше 100%).
 - Комбінований спосіб – поєднання попередніх способів для різних змінних чи спостережень.
5. Процедури на основі моделювання (*Model-Based Procedures*):
- Підхід ЕМ (*EM approach*) – ітеративна двокрокова процедура. Е-стадія передбачає найкращу оцінку відсутніх даних. М-стадія розраховує середні, стандартні відхилення та кореляції з поправкою на те, що відсутні дані були заміщені.
 - Включення відсутніх даних до аналізу за умови, що вони виділяються в окрему групу спостережень, яка порівнюється з іншими групами (наприклад, при частотному аналізі, аналізі середніх чи дисперсійному аналізі).
 - Вставка замість відсутніх інших даних (див. вище), але при цьому створюється нова бінарна змінна (або псевдозмінна), яка приймає значення 1, якщо спостереження має відсутні дані, та 0, якщо ні. Наприклад, це можливо використати у регресійному аналізі, включивши до числа незалежних змінних таку додаткову бінарну змінну.

Г.7. Викиди

Викиди (*Outliers*) – це спостереження з незвичними характеристиками, які явно відрізняються від решти спостережень (нетипові спостереження). З одного боку вони також представляють сукупність спостережень, з іншого боку можуть створювати проблеми щодо надійності результатів аналізу. Викиди можна поділи на декілька груп:

- Викиди внаслідок помилки. Наприклад, замість частки сільськогосподарських товарів у експорті 30% ми могли помилково ввести 130%. Якщо помилки не можливо виправити, такі спостереження потрібно вилучити з аналізу або вважати такими, що дані по ним відсутні.

- Викиди як результат неординарної події (наприклад, війна або розпад СРСР). Дослідник має вирішити, чи включати такі викиди до аналізу чи ні. Якщо є впевненість, що неординарна подія не буде повторюватися, то варто такі викиди виключити з аналізу.
- Викиди, які дослідник не може пояснити. Якщо він представляє важливу частину генеральної сукупності (*population*), такі викиди варто залишити або провести аналіз двічі з викидами та без них, порівнюючи результати.
- Викиди, які індивідуально по кожній змінній, наче б то не є викидами, але в сукупності по всіх змінних мають таку комбінацію значень, яка є незвичною. Наприклад, країна з достатньо високою часткою високотехнологічної продукції у експорті та достатньо низьким рівнем розвитку може виглядати незвично. В більшості випадків такі викиди включають до аналізу або проводять аналіз двічі з викидами та без них.

Варто виключати з аналізу лише дуже невелику частину спостережень (до декількох відсотків максимум) з міркувань того, що вони можуть бути нерепрезентативними викидами. Проведення аналізу двічі з викидами та без них є страховкою від неправильного рішення щодо викидів.

Ідентифікація викидів:

1. Одномірна (*Univariate Detection*). Проводиться аналіз розподілу спостережень за кожною змінною окремо. Якщо вибірка для аналізу є невеликою (приблизно до 80 спостережень), варто вважати значення змінних викидами, якщо вони відрізняються від середньої більше ніж на 2.5 стандартних відхилення. Якщо вибірка є великою, викидами слід вважати значення, що відхиляються від середньої більше ніж на 3-4 стандартних відхилення. Наприклад, у прикладі на Рис Г.15 явним викидом є спостереження, де змінна частка короткострокового зовнішнього боргу приймає значення 70 (що відповідає стандартизованому значенню +3.11).

2. Двомірна (*Bivariate Detection*). Створюється діаграма розсіювання, де по осях X та Y відкладаються дві змінні. Малюється довірчий еліпс, куди попадає переважна більшість спостережень (наприклад 90-95%). Ті спостереження, які знаходяться за межами довірчого еліпсу, вважаються викидами. Далі створюється нова діаграма розсіювання для іншої пари змінних і т.д..

Е	Ф
Частка короткострокового у зовнішньому боргу	Стандартизовані значення
3	-0.85
6	-0.67
8	-0.56
21	0.21
22	0.27
10	-0.44
15	-0.14
13	-0.26
70	3.11
18	0.04
11	-0.38
20	0.15
9	-0.50

Рис.Г.15

3. Багатомірна (*Multivariate Detection*). Уявимо собі побудову діаграми розсіювання не у двовимірному просторі, а у багатомірному просторі (за багатьма змінними). Кожне спостереження може характеризуватися відстанню від центру всіх спостережень у багатомірному просторі, яка може вимірюватися як Відстань Махаланобіса (*Mahalanobis D²*). Спостереження з найбільшою відстанню можуть вважатися викидами.

Додаток Д. Описова статистика і розподіл даних

Д.2. Описова статистика у Microsoft Excel

Описову статистику можливо розрахувати за допомогою опції *Описательная статистика* у надбудові *Пакет анализа*. Наприклад, на Рис.Д.2 показані вхідні дані за поточним рахунком у % ВВП для вибірки країн (за даними World Development Indicators), а далі діалогове вікно (Рис. Д.3) та результат аналізу (Рис.Д.4). Як додаткові опції ми вказали, зокрема, друге найменше і друге найбільше спостереження.

	A	B	C
1	Current account balance (% of GDP)		2010
2		Slovak Republic	-3.38
3		Slovenia	-0.81
4		Sweden	6.28
5		Thailand	4.63
6		Tajikistan	-6.79
7		Turkey	-6.49
8		Tanzania	-8.58
9		Uganda	-10.23
10		Ukraine	-2.09
11		Uruguay	-0.40
12		United States	-3.23
13		Venezuela, RB	3.71
14		Vietnam	-4.14
15		South Africa	-2.78
16		Zambia	3.80

Рис.Д.2.

Рис.Д.3.

Інший варіант – використати самостійно відповідні функції.

	A	B
1		Столбец1
2		
3	Среднее	-2.03410903
4	Стандартная ошибка	1.282189731
5	Медиана	-2.781737214
6	Мода	#Н/Д
7	Стандартное отклонение	4.965899474
8	Дисперсия выборки	24.66015759
9	Эксцесс	-0.830323556
10	Асимметричность	0.149811536
11	Интервал	16.50411551
12	Минимум	-10.2281257
13	Максимум	6.275989806
14	Сумма	-30.51163544
15	Счет	15
16	Наибольший(2)	4.627208613
17	Наименьший(2)	-8.579656744
18	Уровень надежности(95.0%)	2.750023466

Рис.Д.4.

= СРЗНАЧ(діапазон)

Повертає середню арифметичну вказаних чисел чи діапазону. Приклади: =СРЗНАЧ(В2:В10) або =СРЗНАЧ(35;121;76;18)

= СРЗНАЧЕСЛИ(діапазон;умова;діапазон усереднення)

Повертає середню арифметичну для комірок у діапазоні, які відповідають вказаній умові. Середня розраховується для діапазону усереднення (третій аргумент функції), а якщо він не вказаний, то для діапазону (перший аргумент функції). Приклади:

Таблиця Д.1

=СРЗНАЧЕСЛИ(В3:В20;"<100")	Розрахунок середньої тих значень з діапазону В3:В20, які менші 100
=СРЗНАЧЕСЛИ(В3:В20;">0";С3:С20)	Розрахунок середньої значень з діапазону С3:С20 з тих рядків, де в діапазоні В3:В20 значення більші 0
=СРЗНАЧЕСЛИ(В3:В20;"=індекс цін*"; С3:С20)	Розрахунок середньої значень з діапазону С3:С20 з тих рядків, де в діапазоні В3:В20 є текст, який починається з <i>індекс цін</i>
=СРЗНАЧЕСЛИ(В3:В20;" <>Ukraine";С3:С20)	Розрахунок середньої значень з діапазону С3:С20 з тих рядків, де в діапазоні В3:В20 не написано слово Ukraine

= СРЗНАЧЕСЛИМН(діапазон усереднення;діапазон умов1;умова1; діапазон умов2;умова2;...)

Повертає середню арифметичну для комірок у діапазоні, які відповідають декільком умовам. Середнє розраховується для діапазону усереднення. Діапазони умов та діапазон усереднення мають бути одного розміру. Приклад:

Таблиця Д.2

	А	В	С	Д
1	Країна експортер	Країна імпортер	Товар	Ціна
2	Україна	Польща	чавун	500
3	Україна	ФРН	чавун	480
4	Росія	Польща	нафта	400
5	Україна	ФРН	зерно	200

Функція =СРЗНАЧЕСЛИМН(Д2:Д7;А2:А7;"Україна";С2:С7; "чавун") повертатиме середню ціну чавуна, експортованого з України.

=ДОВЕРИТ(альфа;стандартне відхилення;розмір вибірки)

Повертає довірчий інтервал для середньої генеральної сукупності з нормальним розподілом. Альфа – рівень значущості для розрахунку рівня надійності. Наприклад, 0.05 – це 95% рівень надійності. Насправді, для одержання максимального (або мінімального) значення довірчого інтервалу потрібно результат розрахунку за цією функцією додати до (або відняти від) середньої за вибіркою. На Рис Д.5. вказаний приклад (ліворуч результат, а праворуч показані формули у стовпчику В):

	А	В	В
1	Середня	0.31	0.31
2	Стандартне відхилення	0.14	0.14
3	Розмір вибірки	26	26
4	Довірчий інтервал середньої	0.054	=ДОВЕРИТ(0.05;В2;В3)
5	Нижня межа довірчого інтервалу	0.256	=В1-В4
6	Верхня межа довірчого інтервалу	0.364	=В1+В4

Рис.Д.5.

=СРГЕОМ(діапазон)

Повертає середню геометричну.

=СРГАРМ(діапазон)

Повертає середню гармонійну (обернена величина середньої арифметичної обернених величин значень). Не може застосовуватися, якщо одне із значень дорівнює 0.

Розрахунок середньозваженої за допомогою двох формул. Приклад:

Таблиця Д.3

	А	В	С
1	Місяць	Експорт, тис. тон	Ціна за тону
2	Січень	100 000	400
3	Лютий	120 000	430
4	Березень	150 000	380
5	Квітень	110 000	390

Формула **=СУММПРОИЗВ(B2:B5;C2:C5)/СУММ(B2:B5)** повертає середньозважену ціну.

=МЕДИАНА(діапазон)

Повертає медіану.

=МОДА(діапазон)

Повертає моду.

=СТАНДОТКЛОН(діапазон)

Повертає стандартне відхилення для вибірки.

=ДИСП(діапазон)

Повертає дисперсію для вибірки.

=СРОТКЛ(діапазон)

Повертає середнє абсолютних відхилень значень від середньої: фактично – це $(\sum|x-x_{cp}|)/n$

=КВАДРАТОТКЛ(діапазон)

Повертає суму квадратів відхилень значень від середньої: $\sum(x-x_{cp})^2$

=МАКС(діапазон)

Повертає найбільше значення. Приклад: **=МАКС(Н2:Н16;100)** – повертає найбільше з чисел зі вказаного діапазону, але якщо найбільше число менше 100, то повертає 100.

=МИН(діапазон)

Повертає найменше значення.

=НАИМЕНЬШИЙ(діапазон;n)

Повертає вказане n-е найменше значення з діапазону. Наприклад, **=НАИМЕНЬШИЙ(B2:B100;2)** повертає друге найменше значення з діапазону B2:B100.

=НАИБОЛЬШИЙ(діапазон;n)

Повертає вказане n-е найбільше значення з діапазону.

Можливо комбінувати дві останні функції з іншими, наприклад:

=СУММ(НАИМЕНЬШИЙ(B2:B100;{1;2;3;4}))	Повертає суму чотирьох найменших значень з діапазону B2:B100
=СРЗНАЧ(НАИБОЛЬШИЙ(B2:B100;{1;2}))	Повертає середню арифметичну двох найбільших значень з діапазону B2:B100

=СКОС(діапазон)

Повертає асиметрію розподілу відносно середньої.

Для розрахунку рангів та перцентилей (квантилей) існує в надбудові *Пакет аналіза* опція *Ранг и перцентиль*. Скористаємося одним з попередніх прикладів. Нижче показані вхідні дані (Рис. Д.6), результат (Рис. Д.7) та діалогове вікно аналізу (Рис. Д.8).

A	B	C
Current account balance (% of GDP)		2010
	Slovak Republic	-3.38
	Slovenia	-0.81
	Sweden	6.28
	Thailand	4.63
	Tajikistan	-6.79
	Turkey	-6.49
	Tanzania	-8.58
	Uganda	-10.23
	Ukraine	-2.09
	Uruguay	-0.40
	United States	-3.23
	Venezuela, RB	3.71
	Vietnam	-4.14
	South Africa	-2.78
	Zambia	3.80

Рис.Д.6.

	A	B	C	D
1	Точка	Столбец1	Ранг	Процент
2	3	6.28	1	100.00%
3	4	4.63	2	92.80%
4	15	3.80	3	85.70%
5	12	3.71	4	78.50%
6	10	-0.40	5	71.40%
7	2	-0.81	6	64.20%
8	9	-2.09	7	57.10%
9	14	-2.78	8	50.00%
10	11	-3.23	9	42.80%
11	1	-3.38	10	35.70%
12	13	-4.14	11	28.50%
13	6	-6.49	12	21.40%
14	5	-6.79	13	14.20%
15	7	-8.58	14	7.10%
16	8	-10.23	15	0.00%

Рис.Д.7.

Перший стовпчик у результатах показує номер спостереження у діапазоні вхідних даних. Другий – значення змінної, яку ми досліджуємо. Третій – звичайний ранг (найбільшому значенню присвоюється ранг 1). Четвертий – процентний ранг (найбільшому значенню присвоюється 100%).

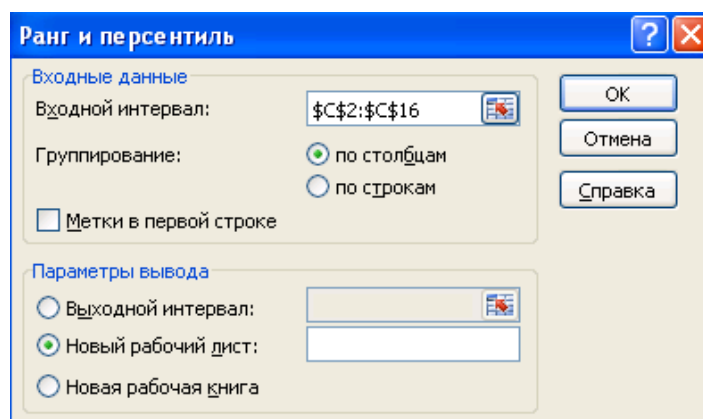


Рис.Д.8.

Якби у вхідному діапазоні було би декілька змінних – стовпчиків, то у результатах значення та ранги по решті змінних розміщувалися би далі праворуч. Але також існують функції, за допомогою яких це можливо зробити самостійно.

=РАНГ(число;діапазон;порядок)

Повертає ранг числа в переліку чисел (діапазоні). Тобто повертає порядковий номер числа, якби всі числа в діапазоні були би відсортовані у певному порядку.

Якщо порядок=0 або не вказується, то найбільшому числу у діапазоні присвоюється ранг 1. Якщо порядок=1, то ранг 1 присвоюється найменшому числу. Однаковим числам присвоюється однаковий ранг. Наприклад, =РАНГ(C2;C16:C6;1)

=ПЕРСЕНТИЛЬ(діапазон;k)

Повертає k-ту персентиль для значень з діапазону. k приймає значення від 0 (найменше значення) до 1 (найбільше значення). Наприклад, у попередньому прикладі =ПЕРСЕНТИЛЬ(C2:C16;0.857) повертатиме 3.8 – у даному випадку третє найбільше значення.

=ПРОЦЕНТРАНГ(діапазон;число;розрядність)

Зворотна до попередньої функція. Повертає частку чисел у діапазоні, що є меншими за вказане число в аргументі. Розрядність – необов'язковий аргумент, що визначає кількість значущих цифр для значення, що повертається, якщо він опущений, то повертається значення у тисячних (наприклад, 0.333, якщо третина значень у діапазоні є нижчими за вказане число). Наприклад, у попередньому прикладі функція =ПРОЦЕНТРАНГ(C2:C16;3.8;3) повертатиме 0.857.

=КВАРТИЛЬ(діапазон;частина)

Повертає кватиль діапазону. Якщо частина дорівнює 0, повертається мінімальне значення; якщо 1 – перша кватиль (25%-персентиль), якщо 2 – медіана (50%-персентиль), якщо 3 – третя кватиль (75%-персентиль), якщо 4 – максимальне значення.

Д.3. Розподіл даних

Значна частина методів аналізу підходять для аналізу даних, які підлягають нормальному розподілу. Багато показників у міжнародній економіці підлягають нормальному розподілу або близькі до нього (наприклад, сальдо поточного рахунку відносно ВВП).

Але не завжди, наприклад динаміка валютного курсу: нерідкі випадки, коли національна валюта під час кризи суттєво знецінюється, але дуже рідко можна спостерігати суттєві подорожчання валют. Також навіть якщо генеральна сукупність підлягає нормальному розподілу, невелика вибірка з неї може суттєво відхилитися від нормального розподілу. Оскільки у нас є лише інформація про вибірку, ми не можемо з упевненістю говорити про розподіл генеральної сукупності. Нормальному розподілу також не підлягають змінні, які вимірюються у неметричній шкалі (тобто у номінальній чи порядковій).

За умов відхилення від нормального розподілу результати звичайних (параметричних) методів аналізу (наприклад, кореляція Пірсона, дисперсійний аналіз) будуть давати зміщені результати. В такому випадку краще використовувати непараметричні аналоги цих методів. Для змінних, які вимірюються у неметричній шкалі, варто використовувати непараметричні методи.

Але у великих вибірках наявність нормального розподілу вже не є обов'язковою умовою для надання переваги звичайним параметричним методам. Також якщо є впевненість про наявність нормального розподілу, кращі результати дають звичайні параметричні методи, які є більш чутливими. Також апіорі, якщо є можливість, можливо використати як параметричні, так і непараметричні методи і врахувати результати обох методів.

Про відхилення від нормального розподілу може свідчити, наприклад, значна різниця між середньою арифметичною та медіаною. Про характер розподілу також свідчать такі показники, як асиметрія, ексцес, кватилі.

Існують спеціальні статистичні тести, які вказують з якою упевненістю можливо стверджувати, що розподіл змінної у вибірці відрізняється від нормального: Тест Шапіро-Уїлкса (*Shapiro-Wilks test*) або Тест Колмогорова-Смірнова (*Kolmogorov-Smirnov test*). Проте ці тести варто використовувати лише для відносно невеликих вибірок. Якщо вибірка велика, спеціальні тести будуть сигналізувати про відхилення від нормального розподілу, навіть якщо воно несуттєве. Тому уні-

версальним способом визначення нормальності розподілу є графічний: гістограма частот, а ще краще спеціальна Діаграма нормального розподілу (*Normal probability plot*).

Нормальність розподілу може бути одномірною (для лише однієї змінної) та багатомірною (додатково вимагається, що комбінації декількох змінних також мають бути нормально розподілені) – остання є важливою для методів багатомірного статистичного аналізу. Багатомірну нормальність складніше перевірити. Частковою гарантією багатомірної нормальності є одномірна нормальність розподілу, кожної змінної, що використовується у аналізі.

Якщо існує відхилення від нормального розподілу (змінної чи залишків), вона часто пов'язана з порушенням інших припущень (наприклад, гетероскедастичність), необхідних для того чи іншого виду аналізу. виправлення інших порушень, може призвести й до виправлення проблем з відхиленням від нормального розподілу.

Якщо розподіл є занадто плоский (немає явного піку), вирішенням проблеми є використання у аналізі зворотної величини (наприклад, $1/Y$ або $1/X$). Якщо розподіл асиметричний, рішенням може бути використання коренів квадратних змінних, або логарифмів чи також зворотних величин. Корінь квадратний краще використовувати за від'ємної асиметрії, а логарифм – за додатної асиметрії.

Крім нормального розподілу існують й інші, які використовуються у статистичному аналізі:

- χ^2 (Пірсона);
- t (Стьюдента);
- F (Фішера);
- лог-нормальний;
- експоненціальний;
- гамма;
- біноміальний;
- Пуассона;
- Бернуллі;
- геометричний.

Д.6. Розподіл даних у Microsoft Excel

В Microsoft Excel аналіз розподілу можливо здійснювати за допомогою опції *Гистограмма* надбудови *Пакет анализа*. Таблиця гістограми складається з меж інтервалів значень показника. Показані частоти значень у межах кожного інтервалу. Використаємо це до нашого прикладу з поточним рахунком платіжного балансу у % ВВП (див. Рис.Д.21). Ми маємо додатково прописати межі бажаних для нас інтервалів для поля *Интервал карманов* у стовпчику D та додаткові опції у діалоговому вікні на Рис.Д.22: *Интегральный процент* та *Вывод графика*.

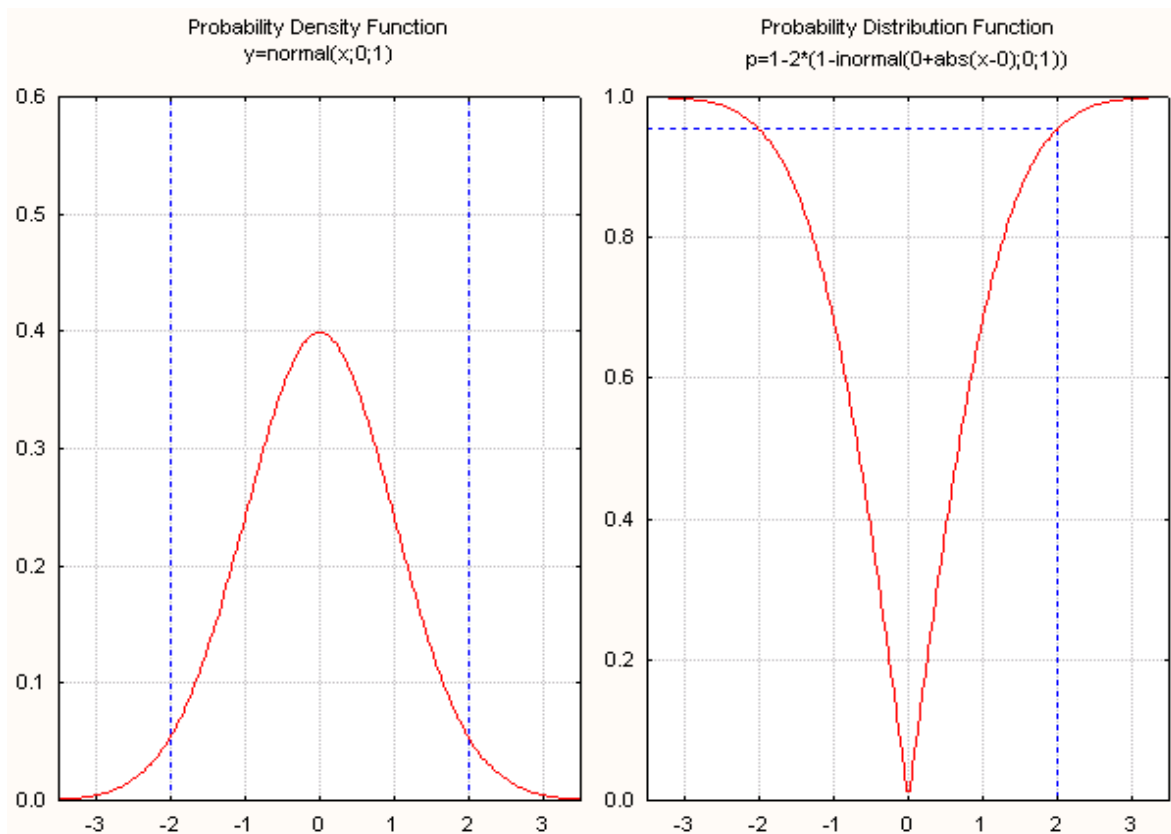


Рис.Д.20.

	A	B	C	D
1	Current account balance (% of GDP)		2010	Інтервали
2		Slovak Republic	-3.38	-10
3		Slovenia	-0.81	-7
4		Sweden	6.28	-4
5		Thailand	4.63	-1
6		Tajikistan	-6.79	2
7		Turkey	-6.49	5
8		Tanzania	-8.58	
9		Uganda	-10.23	
10		Ukraine	-2.09	
11		Uruguay	-0.40	
12		United States	-3.23	
13		Venezuela, RB	3.71	
14		Vietnam	-4.14	
15		South Africa	-2.78	
16		Zambia	3.80	

Рис.Д.21.

Результати вказані на Рис Д.23.

Також є окремі функції, пов'язані з характером розподілу даних. Наприклад,

=НОРМАЛІЗАЦІЯ(число;середня;стандартне відхилення)

Повертає нормалізоване (стандартизоване) значення для розподілу з відомими середньою та стандартним відхиленням.

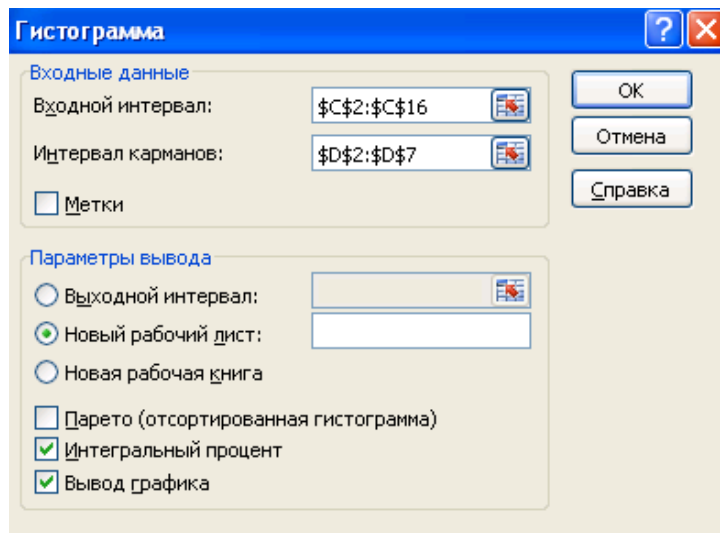


Рис.Д.22.

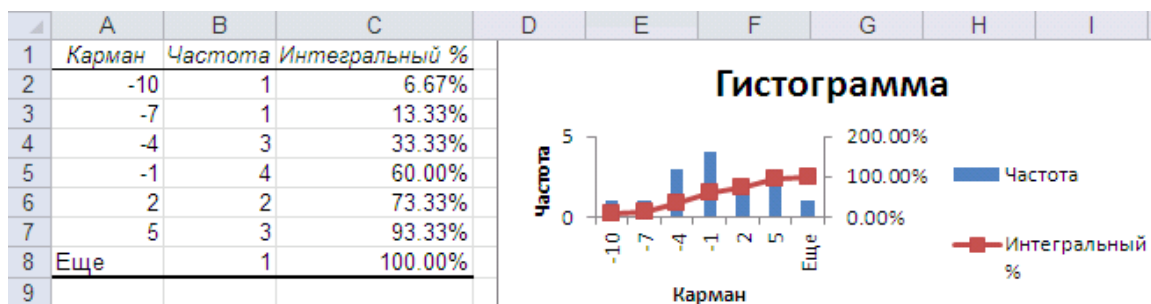


Рис.Д.23.

Д.7. Генерація випадкових чисел

В Microsoft Excel у надбудові *Пакет анализа* існує опція *Генерація случайных чисел*. Вона дозволяє заповнити діапазон випадковими числами, з урахуванням потрібного розподілу значень.

У прикладі на Рис.Д.26 результатом стане у новому аркуші стовпчик з 20 комірками, які міститимуть випадкові числа. Ці числа підпорядковуватимуться нормальному розподілу із середньою 50, та стандартним відхиленням 20. Тобто близько 95% чисел будуть знаходитися в межах від 10 до 90 (+/- 2 стандартних відхилення).

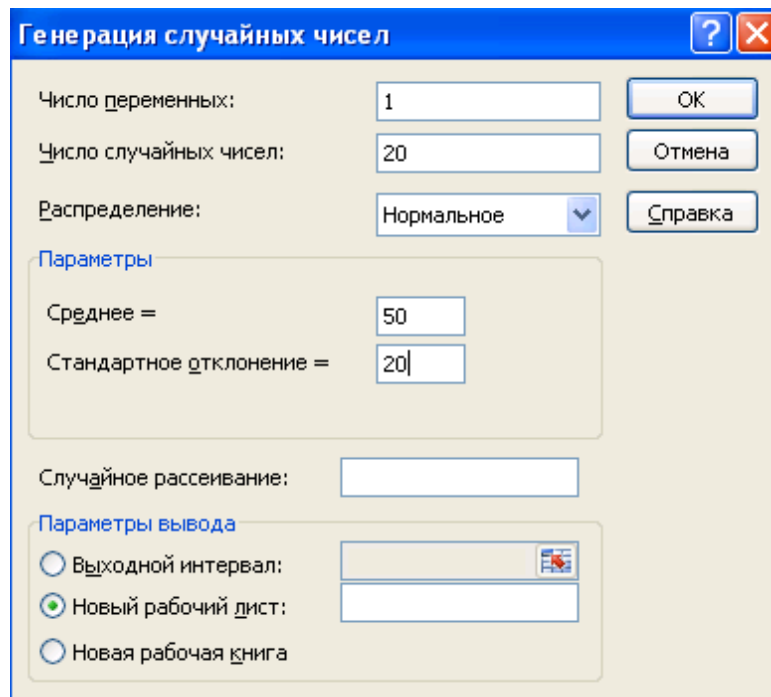


Рис.Д.26.

Також існують спеціальні функції:

=СЛЧИС()

Повертає випадкове число у діапазоні від 0 до 1. Після кожного натиснення клавіші Enter для вводу функції значення, яке вона повертає змінюється.

=СЛУЧМЕЖДУ(нижня межа; верхня межа)

Повертає випадкове ціле число у діапазоні чисел з вказаними межами. Наприклад, **=СЛУЧМЕЖДУ(0;1000)** повертає випадкове число між 0 та 1000. Після кожного натиснення клавіші Enter для вводу функції значення, яке вона повертає змінюється.

СПИСОК ЛІТЕРАТУРИ

1. Белобродский А.В., Гриценко М.А. Поиск решений с Excel 2000. – Воронеж: Воронежский государственный университет, 2001. – 76 с.
2. Боровиков В. Statistica. Искусство анализа данных на компьютере: для профессионалов. – СПб: Питер, 2003. – 688 с.
3. Боровиков В.П., Ивченко Г. И. Прогнозирование в системе STATISTICA в среде Windows. – М.: Финансы и статистика, 2000. – 384 с.
4. Бурлай Т., Білоцерківець О. Прогнозна економетрична модель для оцінки змін реального ВВП України з урахуванням потенціальних шокових впливів // Економіка України. – 2009. – №10. – С.62-68.
5. Вахненко Т. Моделювання макроекономічних факторів зовнішніх запозичень та їх впливу на розвиток економіки України // Економіка України. – 2007. – №7. – С.15-24.
6. Вдовиченко А. Визначення детермінантів заощаджень та споживання населення України на основі емпіричного дослідження // Економіка України. – 2009. – №9. – С.40-52.
7. Грисенко М.В., Рижов А.Ю. Теорія ймовірностей для економістів-міжнародників. – К.: ВПЦ «Київський університет». – 2009. – 211 с.
8. Грисенко М.В., Рижов А.Ю. Математична статистика для економістів- міжнародників. – К.: ВПЦ «Київський університет». – 2011. – 261 с.
9. Гусев М.С., Широков А.А. Прогноз внешней торговли в системе среднесрочного прогнозирования российской экономики // Проблемы прогнозирования. – 2009. – № 01. – С. 3-16.
10. Дослідження операцій в економіці: підручник / І. К. Федоренко, О. І. Черняк, Г. О. Черноус, О. О. Карагодова, О. В. Горбунов. За ред. І. К. Федоренко, О. І. Черняка. – К.: Знання, КОО, 2007. – 558 с.
11. Емельянов А.А. Имитационное моделирование экономических процессов. – М.: Финансы и статистика, 2006. – 416 с.
12. Иллюстрированный самоучитель по SPSS (<http://www.learnspss.ru/handbooks.htm>)
13. Исследование операций в экономике: Учеб. пособие для вузов/ Под. ред. Н.Ш.Кремера. – М.: ЮНИТИ. – 2003. – 407 с.
14. Карпов Ю. Имитационное моделирование систем. Введение в моделирование с AnyLogic 5. – СПб.: БХВ-Петербург, 2005. – 400 с.
15. Качинський А.Б. Безпека, загрози і ризик: наукові концепції та математичні методи. – К.: 2003. – 472 с.
16. Киреев А.П. Прикладная макроэкономика. – М.: Междунар. отношения, 2006. – 456 с.
17. Клименко О., Зубка Т. Моделювання та дослідження динаміки еміграції трудових ресурсів//Банківська справа. – 2009. – №2. – С.67-70.
18. Комп'ютерна підтримка прийняття рішень на різних рівнях державного управління: Методичні рекомендації та збірник завдань / Укл. В.П. Тронь, А.М. Панчук, С.Ю. Марков, Є.М.Нужний, В.Г. Чорна. – К. Вид-во УАДУ, 1998. – 56 с.
19. Кузовкин А.И., Яценко В.М.. Моделирование зависимости спроса на газ от цены и темпов роста ВВП на внутреннем и внешнем рынках // Проблемы прогнозирования. – 2011. – №4. – С.125-136.
20. Ліквідність банківської системи України: Науково-аналітичні матеріали. Випуск 12. – К.: НБУ, Центр наукових досліджень, 2008. – 180с.
21. Лук'яненко І.Г., Краснікова Л.І. Економетрика: Підручник. – К.:Товариство «Знання», КОО, 1998. – 494 с.
22. Ляшенко О.І. Математичне моделювання динаміки відкритої економіки: монографія. – Рівне: Волинські обереги, 2005. – С.360.
23. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс: Учебник. – М.: Дело, 2005. – 504 с.
24. Макарова Н.В., Трофимец В.Я. Статистика в Excel. – М.Финансы и статистика, 2002. – 368 с.
25. Макроекономічне моделювання та короткострокове прогнозування / За ред. І. В. Крючкової – Харків: Форт, 2000. – 336 с.
26. Маланичев А.Г., Воробьев П.В. Прогнозирование мировых цен на сталь // Проблемы прогнозирования. – 2011. – № 3. – С.108-117.
27. Михалевич М.В. Моделирование переходной экономики: модели, методы, информационные

- технологии. – К.: Наукова думка, 2005. – 671 с.
28. Моделі і методи соціально-економічного прогнозування: підручник / В. М. Геєць, Т. С. Клебанова, О. І. Черняк, А. В. Ставицький та інші. 2-е вид., виправ. – Х.: ВД «ІНЖЕК», 2008. – 396 с.
 29. Моисеев С.Р. Международные валютно-кредитные отношения. – 2-е изд., перераб. и доп. – М.: Издательство «Дело и Сервис», 2007. – 816 с.
 30. Сельцовский В.Л. Экономико-статистические методы анализа внешней торговли. – М.: Финансы и статистика, 2004. – 512 с.
 31. Системный анализ актуальных проблем экономики / Завельский М.Г. (Ред.). – М.: Эдиториал УРСС, 2002. – 128 с.
 32. Скрипниченко М.І. Секторальні та міжкраїнні моделі економічного розвитку. – К.:Фенікс, 2004. – 256 с.
 33. Сошникова Л.А. Многомерный статистический анализ. – Минск: БГЭУ, 2004. – 162 с.
 34. Справка и инструкции по Excel 2010 (<http://office.microsoft.com/ru-ru/excel-help>)
 35. Статистика: теория и практика в Excel: учеб. пособие / В.С.Лялин, И.Г. Зверева, Н.Г. Никифорова. – М.: Финансы и статистика: ИНФРА-М, 2010. – 448 с.
 36. Теорія статистики: Навчальний посібник / Вашків П.Г, Пастер П.І, Сторожук В.П., Ткач Є.І. – К.: Либідь, 2001. – 320 с.
 37. Точилін В.О., Подолець Р.З., Дячук О.А., Олександренко Ю.А. Прикладна економіко-математична модель "Times-Україна" для оптимізації енергетичних потоків та прогнозування енергетичного балансу України // Наука та інновації. – 2010 – № 2. – С.48-66.
 38. Тронь В.П. Нечітка стратегія чітких рішень. – К.: Національна академія державного управління при Президентові України, Українська академія наук з державного управління, 2007. – 748 с.
 39. Ульяновченко О. В. Дослідження операцій в економіці: Підручник для студентів вузів/Харк. нац. аграр. ун-т ім. В. В.Докучаєва. – Харків: Гриф, 2002. – 580 с.
 40. Учебник по программе STATISTICA (<http://www.hr-portal.ru/statistica/>)
 41. Халафян А.А. Statistica 6. Статистический анализ данных. 3-е изд. Учебник. – М.: ООО «Бином-Пресс», 2007. – 512 с.
 42. Хмара М.П., Чугаєв О.А. Оцінка потенційних наслідків торговельної інтеграції на прикладі формування зони вільної торгівлі між Україною та Туреччиною // Зони вільної торгівлі на початку ХХІ століття: Монографія / А.С. Філіпенко, В.С. Будкін, О.І.Шнирков; за ред. Шниркова О.І., Філіпенка А.С. – К.: Видавничо-поліграфічний центр «Київський університет», 2013.
 43. Черняк О. І. Методи прогнозування експорту-імпорту / О. І. Черняк, Л. В. Воронова, А. В. Ставицький // Економіка і прогнозування. – 2001. – № 2. – С. 96-109.
 44. Черняк О. І. Методологія розрахунку прогнозу платіжного балансу / О. І. Черняк, Н. М. Іванік, А. В. Ставицький // Вісник Національного банку України. – 1998. – № 12. – С.42-44.
 45. Черняк О. І. Прогнозування зовнішньо-торгівельних сальдо України: комбінований підхід із застосуванням VAR-моделей / О. І. Черняк, Є. С. Корнієнко // Вісник Національного банку України. – 2001. – № 11. – С. 17-20.
 46. Черняк О. І., Кучерук Л. В. Застосування байєсівських мереж в економіці // Вісник Харківського національного університету імені В. Н. Каразіна. Економічна серія. – 2009. – № 869. – С.199-209.
 47. Чугаєв О.А. Валютні кризи на межі ХХ – ХХІ століть: Монографія. – К.: „МП Леся”, 2007. – 416 с. (<http://www.sites.google.com/site/achugaiev/stati/stati-1>)
 48. Швандар К.В. Анализ и прогноз конъюнктуры мировых товарных рынков в центральных банках (зарубежный опыт) // Деньги и кредит. – 2010. – № 2. – С.64-72.
 49. Швандар, К.В. Возможности моделирования при прогнозировании конъюнктуры мировых товарных рынков / К. В. Швандар, С. В. Плотников, А. Д. Бородин // Деньги и кредит: теорет. научно-практ. журнал. – 2010. – № 4. – С.63-69.
 50. Шишлянникова Л.М. Математическое сопровождение научной работы с помощью статистического пакета SPSS for Windows 11.5.0. – М., 2005 – 107 с. (<http://www.matlab.mgppu.ru/work/0022.htm>)
 51. Экономико-математические методы и модели: Компьютерные технологии решения: Учебное пособие для экономических специальностей вузов / И. Л. Акулич, и др. – Мн.: БГЭУ, 2003. –

348 с.

52. Электронный учебник Statsoft (<http://www.statsoft.ru/home/textbook/>)
53. AnalystSoft (<http://www.analystsoft.com/ru/>)
54. A.T.Kerney (<http://atkerney.com>)
55. Bank for International Settlements (<http://www.bis.org>)
56. Belhocine N., Dobrescu G., Mazraani S., Petrova I. Assessing Fiscal Stress. IMF Working Paper, August 2010. – 34 p. (http://www.ewi-ssl.pitt.edu/econ/files/faculty/wp/Samah_Assessing%20Fiscal%20Stress%20WP%20v4.pdf)
57. Baldacci E., Petrova I., Belhocine N., Dobrescu G., Mazraani S. Assessing Fiscal Stress. IMF Working Paper WP/11/100, May 2011. – 41 p. (<http://www.imf.org/external/pubs/ft/wp/2011/wp11100.pdf>).
58. Driver R., Westaway P. Concepts of equilibrium exchange rates. – Working Paper no. 248. – Bank of England, 2004. – 64 p.
59. Euromonitor International Passport GMID (<http://www.euromonitor.com/passport-gmid>)
60. Eurostat (<http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/themes>)
61. General Algebraic Modeling System (GAMS) Home Page (<http://www.gams.com/default.htm>)
62. Grcic B. The Pollak's Macroeconomic Monetary Model. – Faculty of Economics Split. – 12 p. (http://www.systemdynamics.org/conferences/2001/papers/Grcic_1.pdf)
63. Hair J., Anderson R., Tatham R., Black W. Multivariate Data Analysis. – Fifth edition. – Upper Saddle River, New Jersey: Prentice Hall Inc., 1998.
64. Hill T., Lewicki P. (2007). STATISTICS: Methods and Applications. StatSoft, Tulsa, OK. (StatSoft, Inc. (2011). Electronic Statistics Textbook. Tulsa, OK: StatSoft (<http://www.statsoft.com/textbook/>)
65. <http://www.exponenta.ru/soft/Statist/Statist.asp>
66. IBM SPSS Statistics (<http://www-01.ibm.com/software/analytics/spss/products/statistics>)
67. International Human Development Indicators (<http://hdr.undp.org/en/data/about/>)
68. International Monetary Fund Statistics (<http://www.imf.org/external/data.htm>)
69. Kaminsky G., Lizondo S., Reinhart C. Leading Indicators of Currency Crises. Working Paper № 97/79. – Wash. D.C.: IMF, 1997. – 43 p.
70. Laxton D., Isard P., Faruqee H., Prasad E., Turtelboom B. MULTIMOD Mark III The Core Dynamic and Steady-State Models. Occasional paper, no. 164. – Washington, DC: International Monetary Fund, 1998. – 73 p.
71. Levine D., Ramsey P., Smidt R. Applied Statistics for Engineers and Scientists: Using Microsoft Excel & Minitab. – Upper Saddle River, New Jersey: Prentice Hall Inc., 2001.
72. Miller W. Statistics and Measurements Using OpenStat, 2011. – 593 p. (<http://www.statprograms4u.com/ATextBook.pdf>)
73. National Bank of Ukraine (<http://www.bank.gov.ua>)
74. NCSS (<http://www.ncss.com/information.html>)
75. State Statistics Service of Ukraine (<http://www.ukrstat.gov.ua/>)
76. STATGRAPHICS Online Statistical Analysis System. – StatPoint Technologies, Inc. 2008 – 35 p. (<http://www.statpoint.com/statgraphics%20online.pdf>)
77. Statistical Services Centre of the University of Reading (<http://www.reading.ac.uk/ssc/home.html>)
78. Statpoint Technologies, INC. (<http://www.statgraphics.com/index.htm>)
79. StatSoft Russia (<http://www.statsoft.ru>)
80. Stern R., Leidi S., Dale I., Grayer C. Instat Tutorial. – Statistical Services Centre of The University of Reading. – 40 p. (<http://www.reading.ac.uk/ssc/n/software/instat/tutorial.pdf>)
81. Systat Software, Inc. (http://www.systat.com/Contact_Us.aspx)
82. The IMF-FSB Early Warning Excercise. Design and methodological Toolkit. – IMF, September 2010. – 41 p.
83. UNData (<http://data.un.org/Default.aspx>)
84. United Nations Conference for Trade and Development Statistics (http://unctadstat.unctad.org/ReportFolders/reportFolders.aspx?sCS_referer=&sCS_ChosenLang=en)
85. Using Excel for Statistical Analysis (<http://www.scribd.com/doc/49234740/Stats-Using-Excel1>)
86. Web Pages that Perform Statistical Calculations (<http://statpages.org/>)
87. World Bank Statistics (<http://data.worldbank.org/data-catalog>)
88. World Economic Forum Global Competitiveness Report (<http://www.weforum.org/reports>)
89. World Trade Organization Statistics (http://www.wto.org/english/res_e/status_e/status_e.htm)
90. XJ Technologies (<http://www.xjtek.com/>)

Навчальне видання

**Грисенко Марина Віталіївна,
Чугаєв Олексій Анатолійович**

**КІЛЬКІСНІ МЕТОДИ АНАЛІЗУ МІЖНАРОДНИХ
ЕКОНОМІЧНИХ ВІДНОСИН**

НАВЧАЛЬНИЙ ПОСІБНИК

Підписано до друку 20.11.2012

Наклад 100 прим.

10,9 друк.арк.

Відділ оперативної поліграфії
Інституту міжнародних відносин
Київського національного університету імені Тараса Шевченка
Тел. 483-11-25