

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БУДІВНИЦТВА І АРХІТЕКТУРИ

СТАТИСТИКА

Конспект лекцій

для студентів які навчаються за напрямом підготовки 6.030510
«Товарознавство і торговельне підприємництво»

Київ 2010

УДК 519
ББК 22.172
С78

Укладачі: Н.Д. Федоренко, канд. техн. наук, професор
О.В. Доля, канд. фіз.-мат. наук, доцент
С.В. Білощицька, канд. техн. наук, доцент
А.О. Білощицький, канд. техн. наук, доцент
С.А. Теренчук, канд. фіз.-мат. наук, доцент

Рецензент В.М. Михайленко, доктор техн. наук, професор

Відповідальний за випуск В.В. Демченко, канд. техн. наук,
доцент

*Затверджено на засіданні кафедри прикладної математики,
протокол №13 від 23 червня 2010 року.*

Видається в авторській редакції.

Статистика: конспект лекцій / уклад.: Н.Д. Федоренко, О.В. Доля,
С78 С.В. Білощицька та ін. – К.: КНУБА, 2010. – 56 с.

Розглянуто теоретичні і методичні основи побудови статистичних показників, а також розкриті найважливіші статистичні методи дослідження соціально-економічних явищ і процесів.

Призначено для студентів, які навчаються за напрямом підготовки 6.030510 «Товарознавство і торговельне підприємництво».

УДК 519
ББК 22.172

© Н.Д. Федоренко, О.В. Доля, С.В. Білощицька,
А.О. Білощицький, С.А. Теренчук, 2010
© КНУБА, 2010

ВСТУП

Навчальна дисципліна «Статистика» відіграє провідну роль у формуванні висококваліфікованого спеціаліста, озброюючи його знаннями методів і способів збирання, опрацьовування та аналізу економічної інформації.

У сучасних умовах, коли головним завданням економічної науки є вивчення основних закономірностей ринкової системи в їх розвитку і взаємодії, особливої актуальності набувають уміння грамотно опрацьовувати та аналізувати статистичні дані. У зв'язку з цим підвищуються вимоги до спеціалістів щодо досконалого вивчення дисципліни «Статистика», головна мета якої – викладення методологічних основ кількісного дослідження закономірностей економічних явищ і процесів.

Запропонований конспект лекцій містить шість тем, які розподілені на окремі питання. В кінці кожної теми наведено перелік запитань для самоконтролю знань. Текст супроводжується рисунками, таблицями, графіками та прикладами.

Лекція 1. ВСТУП. СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ. ТАБЛИЧНИЙ І ГРАФІЧНИЙ МЕТОДИ ПОДАННЯ СТАТИСТИЧНИХ ДАНИХ

План викладу і засвоєння матеріалу:

1. Предмет і метод статистики.
2. Методи статистичного спостереження.
3. Табличний метод подання статистичних даних.
4. Графічний метод подання статистичних даних.
5. Контрольні запитання для самоперевірки.

Ключові поняття та терміни

Статистика, статистична сукупність, варіація ознаки, статистичне спостереження, помилки репрезентативності, статистична таблиця, статистичний графік, діаграма.

1. ПРЕДМЕТ І МЕТОД СТАТИСТИКИ

Мета курсу «Статистика» – розвинути науково-дослідницькі компоненти статистичного мислення, тобто допомогти зрозуміти основні правила, методи, а також способи кількісного аналізу різного виду інформації.

Вивчення дисципліни «Статистика» забезпечує:

- глибоке засвоєння методів статистичного аналізу та методики розрахунку статистичних характеристик під час застосування певного статистичного методу;
- набуття практичних навичок застосування статистичних методів у процесі самостійного вивчення певного економічного явища на основі фактичних даних;
- розвиток умінь на підставі узагальнення й аналізу результатів дослідження роботи висновки й подавати їх у формі статистичних таблиць і графіків.

Теоретичну основу будь-якої науки, у тому числі і статистики, становлять поняття і категорії, в сукупності яких виражаються основні принципи цієї науки. У статистиці до найважливіших категорій і понять належать: **закономірність, сукупність, варіація, ознака.**

Важливою категорією статистики є **статистична закономірність**, яка розглядається як кількісна закономірність зміни у просторі та часі масових явищ і процесів, елементів, що складаються з одиниць сукупності.

Статистична сукупність – це множина однорідних з погляду якої-небудь ознаки явищ, існування яких обмежене у просторі і часі. Кількісні зміни значень ознаки під час переходу від однієї одиниці сукупності до іншої називаються **варіацією**. Вона виникає під впливом випадкових насамперед зовнішніх причин.

У більшості статистичних досліджень широко використовуються **показники варіації**, які демонструють, як групуються значення ознаки навколо середньої величини. Показники варіації є одночасно і характеристиками однорідності сукупностей.

Статистичні сукупності мають певні властивості, які характеризуються деякими **ознаками**. За формою зовнішнього вираження вони поділяються на: номінальні (описові, якісні) та кількісні.

Статистична інформація ґрунтується на єдиній системі статистичних показників, які обчислюються за єдиними рекомендаціями, розробленими і схваленими міжнародними організаціями як міжнародні статистичні стандарти.

Незалежно від характеру дослідження всі статистичні дослідження проводяться в чотири етапи:

- проектування дослідження.
- збір та обробка статистичної інформації.
- статистичне оцінювання.
- перевірка статистичних гіпотез.

2. МЕТОДИ СТАТИСТИЧНОГО СПОСТЕРЕЖЕННЯ

Статистичне спостереження це планомірне, науково організоване спостереження за масовими явищами, яке полягає в реєстрації відібраних ознак у кожної одиниці сукупності.

Під час підготовки спостереження, як правило, ставляться такі запитання:

1. Якого типу повинна бути інформація?
2. У який спосіб зібрати і звести інформацію?
3. У який спосіб опрацювати зібрану інформацію і зробити відповідні висновки?

Статистичні спостереження умовно можна поділити на:

- первинні, що спираються на реєстрацію даних у безпосередній спосіб на такому об'єкті, наприклад, продукція на підприємстві тощо;
- вторинні, які спираються на використання раніше зібраних даних, наприклад, банківські звіти, статистичні щорічники тощо.

Головні вимоги до статистичних даних: правдоподібність і регулярність; комплектність; актуальність; однорідність, порівнюваність, тотожність одиниць виміру даних, доступність.

Отримання статистичних даних передбачає підготовку і планування спостереження, реєстрацію даних, складання бази даних. Кожен етап вимагає опрацювання певного методу. На початку роботи потрібно визначити:

- мету статистичного дослідження;
- об'єкт спостереження;
- властивості об'єкта спостереження;
- одиниці спостереження.

У вітчизняній статистиці використовуються три організаційні форми статистичного спостереження:

- статистична звітність (підприємств, організацій, установ тощо);
- спеціально організовані статистичні спостереження суцільного і несуцільного характеру;
- реєстри.

Статистична інформація може бути отримана різними способами, найважливішими з яких є **безпосереднє спостереження, документальний облік фактів і опитування.**

У статистиці застосовуються такі види опитування: **усний, самореєстрація, кореспондентський, анкетний і явочний.**

Статистичні спостереження можна поділити на групи за такими ознаками: час реєстрації фактів та охоплення одиниць сукупності.

Щодо охоплення **одиниць сукупності**, то статистичне спостереження буває **суцільне і несуцільне.**

Суцільне спостереження полягає в отриманні інформації про всі одиниці досліджуваної сукупності. Однак такий вид має недоліки: високу вартість отримання й оброблення всього обсягу інформації, недостатню оперативність інформації.

Несуцільне спостереження передбачає, що обстеженню підлягає лише частина одиниць сукупності, що вивчається.

Одним з видів несуцільного спостереження є **вибіркове спостереження**. Воно ґрунтується на принципі випадкового відбору одиниць сукупності, що вивчається. Чисельність вибіркової сукупності залежить від природи (характеру) досліджуваного явища.

Під час підготовки спостереження, крім мети, слід встановити **об'єкт спостереження**. Під **об'єктом спостереження** розуміється деяка статистична сукупність, в якій відбуваються досліджувані масові явища і процеси.

Одиницею спостереження називають елемент об'єкта, який є носієм ознак, що підлягають реєстрації. Наприклад, у бюджетних обстеженнях – домашнє господарство.

Програма статистичного спостереження – це перелік ознак (або питань), що підлягають реєстрації в процесі спостереження. Для визначення складу реєстрованих ознак розробляють програму спостереження.

Статистичний формуляр – це документ єдиного зразка, що містить програму і результати спостереження. Його обов'язковими елементами є титульна й адресна частини, які повинні містити найменування статистичного спостереження та органу провідного спостереження, адресу звітної одиниці, її підлеглість.

Формуляри мають такі назви: звіт, картка, переписний лист, експериментний бланк, анкета тощо.

До формуляра розробляється **інструкція**, що визначає порядок проведення спостереження і заповнення форми звітності, переписного листа, анкети. Формуляр та інструкція після їх заповнення разом становлять **інструментарій статистичного спостереження**.

Помилкою спостереження називається розбіжність між розрахунковим і реальним значеннями величин, що вивчаються.

Залежно від причин виникнення розрізняють помилки реєстрації і помилки репрезентативності.

Помилки репрезентативності характерні тільки для несуцільного спостереження. Вони виникають тому, що відібрана й обстежена сукупність недостатньо точно відтворює генеральну сукупність.

3. ТАБЛИЧНИЙ МЕТОД ПОДАННЯ СТАТИСТИЧНИХ ДАНИХ

Для компактного і наочного викладу результатів статистичного зведення та їх аналізу застосовуються таблиці, а основним засобом узагальнення й аналізу статистичної інформації є табличний метод.

Статистична таблиця – не лише форма раціонального і компактного викладу статистичних даних, але й могутній засіб їх аналізу, тому що дає змогу здійснювати наочні зіставлення даних, констатувати факти, встановлювати взаємозалежності між показниками, виявляти закономірності розвитку явищ.

Статистичну таблицю із заголовком таблиці та назвами рядків і граф, але без заповнених клітин, називають **макетом статистичної таблиці** (рис. 1.1).

За своєю сутністю, статистична таблиця розглядається як «статистичне речення», де **підметом** є об'єкт дослідження, а **присудком** – система показників, ознак що характеризують об'єкт.

Таблиця (вказати номер)

Заголовок (назва таблиці)

Підмет присудок	Місце проживання		Стать		Міське населення		Сільське населення	
	міське	сільське	чоловіча	жіноча	чоловіки	жінки	чоловіки	жінки
А	1	2	3	4	5	6	7	8
Перелік одиниць (груп) сукупності								
1								
2								
3								
Підсумковий рядок								



Рис. 1.1. Макет статистичної таблиці

4. ГРАФІЧНИЙ МЕТОД ПОДАННЯ СТАТИСТИЧНИХ ДАНИХ

Статистичний графік – це особливий спосіб наочного подання і викладу за допомогою геометричних знаків та інших графічних засобів статистичної інформації з метою її узагальнення та аналізу (рис. 1.2).

Аналізуючи статистичні дані, графіки можна використовувати з метою:

- вивчення структури явищ і її зміни (структурні зрушення);
- порівняння розмірів різних явищ та їх окремих частин;
- відображення розподілу одиниць сукупності за значеннями ознаки, що вивчається;
- відображення розвитку явищ у часі, їх тенденцій, сезонних явищ;
- виявлення зв'язку між явищами, а також напрямку і ступеня його щільності.

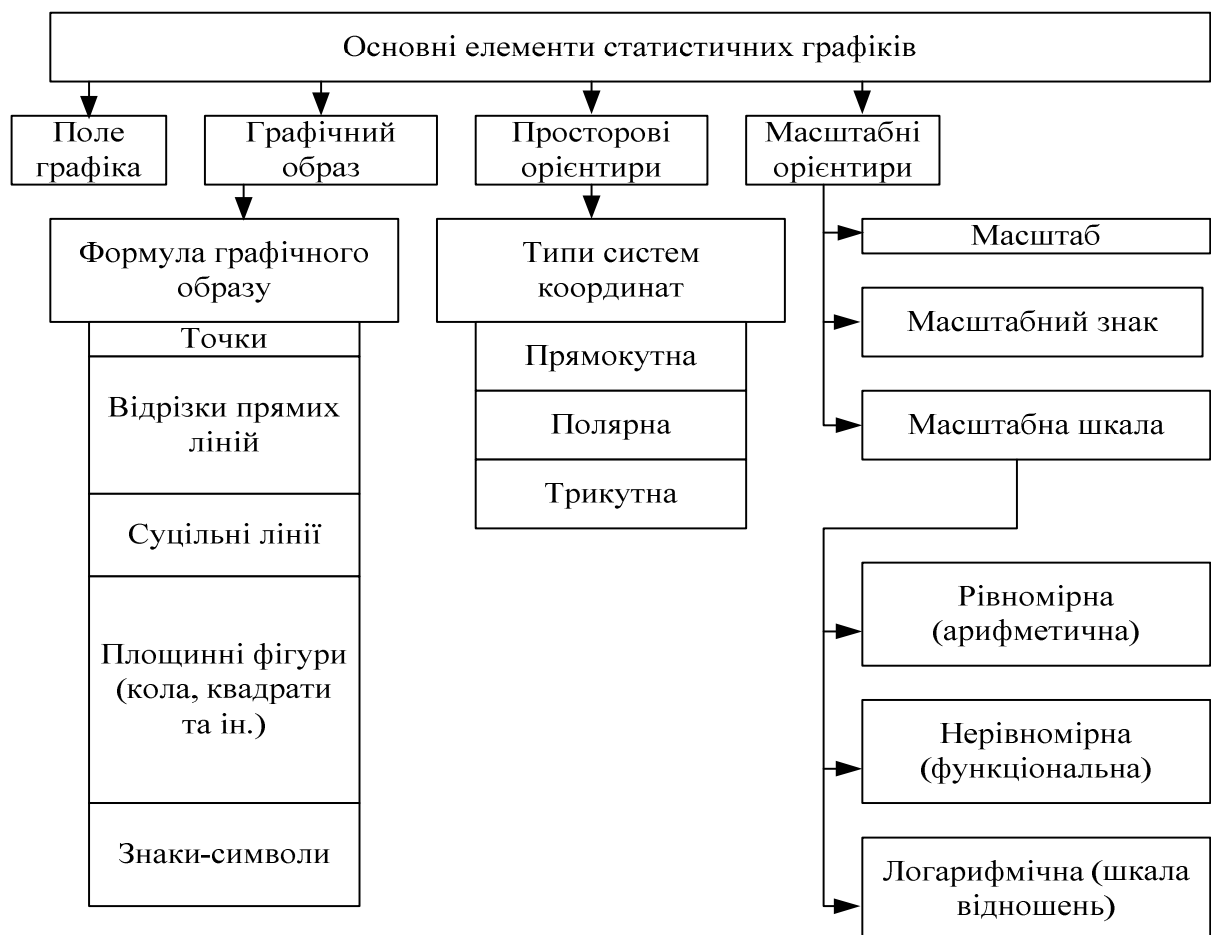


Рис. 1.2. Основні елементи статистичних графіків

5. КОНТРОЛЬНІ ЗАПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Дайте визначення предмета статистики.
2. Що є теоретичною основою статистичної науки?
3. Що таке статистична закономірність? Наведіть приклади:
4. Дайте визначення статистичної сукупності.
5. Що є одиницею статистичної сукупності?
6. Що таке статистична ознака?
7. Чим зумовлена варіація ознак, наведіть приклади варіюючих ознак?
8. Які етапи статистичного дослідження?
9. Розкрийте сутність і завдання статистичного спостереження.
10. Які є види статистичного спостереження?
11. У чому полягає план статистичного спостереження?
12. Перелічіть програмно-методологічні питання плану статистичного спостереження.
13. Вкажіть організаційні питання статистичного спостереження.
14. Які є організаційні форми статистичного спостереження?
15. Як трактується поняття «статистична звітність»?
16. Вкажіть спеціально організовані статистичні спостереження та їх види.
17. Що таке «реєстри»?
18. Розкрийте сутність помилок статистичного спостереження та методів перевірки достовірності зібраних даних.
19. Що таке статистична таблиця і які її функції?
20. Які таблиці не є статистичними?
21. Що є макетом статистичної таблиці? Вкажіть його елементи?
22. Що називається підметом і присудком в статистичній таблиці?
23. Що таке статистичний графік і яке його значення в поданні та аналізі статистичних даних?
24. Які завдання статистичного дослідження вирішуються за допомогою статистичних графіків?
25. Які графіки не є статистичними?

Лекція 2. СТАТИСТИЧНІ ПОКАЗНИКИ. ВИБІРКОВЕ СПОСТЕРЕЖЕННЯ

План викладу і засвоєння матеріалу

1. Середні величини. Зміст і техніка обчислення.
2. Мода і медіана.
3. Варіація ознаки, її сутність.
4. Вибіркове спостереження.
5. Інтервальна оцінка параметра досліджуваної ознаки.
6. Контрольні запитання для самоперевірки.

Ключові поняття та терміни

Середні арифметичне, геометричне, хронологічне, мода і медіана, варіація ознаки, коефіцієнт варіації, генеральна і вибіркова сукупності, інтервальна оцінка.

1. СЕРЕДНІ ВЕЛИЧИНИ. ЗМІСТ І ТЕХНІКА ОБЧИСЛЕННЯ

Одним із найпростіших статистичних методів, який дає можливість кількісно охарактеризувати якість товарів або послуг, є **метод середніх величин**. **Середня величина** є значенням, яке дає узагальнюючу характеристику певному явищу, процесу. До найбільш уживаних належать такі: середнє арифметичне просте і зважене; середнє геометричне; середнє хронологічне.

Середнє арифметичне просте для певної сукупності визначається за формулою:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n},$$

де x_1, x_2, \dots, x_n – значення ознаки, яка досліджується; n – кількість спостережень.

Значення, яких може набувати певна ознака, називаються її **варіантами**.

Якщо певні значення ознаки зустрічаються не один раз, то їх можна згрупувати і розраховувати середнє значення ознаки як **середнє арифметичне зважене**, надаючи кожній варіанті ознаки вагу відповідно до частоти її появи:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_n f_n}{f_1 + f_2 + f_3 + \dots + f_n} = \frac{\sum x f}{\sum f},$$

де f_1, f_2, \dots, f_n — кількість разів, з якою зустрічається в сукупності відповідне значення ознаки (частота).

Подані у певному порядку варіанти досліджуваної ознаки та відповідні їм частоти називаються **рядом розподілу ознаки**.

Якщо узагальнююча величина формується як добуток індивідуальних значень, то середнє значення ознаки визначається за формулою **середнього геометричного**:

$$\bar{x} = \sqrt[n]{x_1 x_2 x_3 \dots x_n}.$$

У тому разі, коли значення ознаки наводяться через рівні проміжки часу, користуються середнім хронологічним:

$$\bar{x} = \frac{\frac{1}{2}x_1 + x_2 + x_3 + \dots + \frac{1}{2}x_n}{n-1},$$

де n — кількість моментів спостережень.

2. МОДА І МЕДІАНА

Крім власне середніх величин, охарактеризувати певну ознаку можна за допомогою ще двох величин — **моди** і **медіани**.

Мода (M_o) — це найпоширеніше значення ознаки, тобто варіанта, яка в ряду розподілу має найбільшу частоту.

В інтервальному ряду за найбільшою частотою визначається **модальний інтервал**. Значення моди обчислюється за формулою

$$M_o = x_0 + h \cdot \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})},$$

де x_0 — нижня межа модального інтервалу; h — ширина модального інтервалу; f_{M_o} — частота модального інтервалу; f_{M_o-1} — частота попереднього інтервалу; f_{M_o+1} — частота інтервалу, наступного за модальним.

Медіана (Me) — це варіанта, яка припадає на середину впорядкованого ряду розподілу і ділить його на дві рівні за обсягом частини. В інтервальному ряду **медіанний інтервал** визначається з умови:

$$S_j \geq \frac{1}{2} \sum_{j=1}^n f_j,$$

де S_j — кумулятивна частота.

Значення медіани визначається за формулою:

$$Me = x_0 + h \cdot \frac{\frac{1}{2} \sum_{j=1}^n f_j - S_{f_{Me-1}}}{f_{Me}},$$

де x_0 – нижня межа медіанного інтервалу; h – ширина медіанного інтервалу; f_{Me} – частота медіанного інтервалу; $S_{f_{Me-1}}$ – кумулятивна частота попереднього до медіанного інтервалу.

Приклад 1. На підприємстві збирають дані про кількість дефектів у партіях виготовленої продукції. За результатами попереднього місяця отримано дані, наведені у табл. 2.1.

Таблиця 2.1

Кількість дефектів у партіях виготовленої продукції

№п/п	Кількість дефектів у партії	Кількість партій продукції
1	1	17
2	2	8
3	3	6
4	4	5
5	5	4
Разом	–	40

Керівництво підприємства хоче знати середню кількість дефектів в окремій партії продукції.

Розв’язання. Скористаємося формулою середнього арифметичного зваженого:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_n f_n}{f_1 + f_2 + f_3 + \dots + x_n} = \frac{\sum x f}{\sum f} = \frac{1 \cdot 17 + 2 \cdot 8 + 3 \cdot 6 + 4 \cdot 5 + 5 \cdot 4}{17 + 8 + 6 + 5 + 4} = \frac{91}{40} = 2,275.$$

Отже, середня кількість дефектів на партію продукції становить 2,275 одиниці.

3. ВАРІАЦІЯ ОЗНАКИ, ЇЇ СУТНІСТЬ

Крім середньої величини, іншою важливою характеристикою будь-якого явища або процесу є його варіація.

Варіація – це набуття ознакою різних значень внаслідок дії численних, часто випадкових факторів. Варіація існує в усіх явищах і процесах.

Найпростішим показником, що характеризує варіацію ознак, є **розмах варіації R** :

$$R = X_{\max} - X_{\min}.$$

Середнє з абсолютних відхилень називається **середнім лінійним відхиленням d** , яке, аналогічно середнім величинам, буває простим і зваженим. Просте середнє лінійне відхилення використовується в тому разі, коли розрахунок здійснюється за незгрупованими даними:

$$d = \frac{\sum |x_i - \bar{x}|}{n}.$$

Середнє лінійне відхилення зважене використовується тоді, коли дані, за якими здійснюється розрахунок, згруповані:

$$d = \frac{\sum |x_i - \bar{x}| f}{\sum f},$$

де f – частота повторюваності ознаки.

Середнє з квадратів відхилень називається **дисперсією σ^2** і розраховується за такою формулою:

$$\text{просте: } \sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_i - \bar{x})^2}{n} = \frac{\sum (x_i - \bar{x})^2}{n};$$

$$\text{зважене: } \sigma^2 = \frac{(x_1 - \bar{x})^2 f_1 + (x_2 - \bar{x})^2 f_2 + \dots + (x_i - \bar{x})^2 f_i}{f_1 + f_2 + \dots f_i} = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f}.$$

Для того щоб охарактеризувати варіацію в тих самих одиницях виміру, що й досліджувана ознака, потрібно з дисперсії добути квадратний корінь. Отриманий показник називається **середнім квадратичним відхиленням** і розраховується так:

$$\sigma = \sqrt{\sigma^2}.$$

Середнє квадратичне відхилення (зважене):

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 f_1 + (x_2 - \bar{x})^2 f_2 + \dots + (x_i - \bar{x})^2 f_i}{f_1 + f_2 + \dots f_i}} = \sqrt{\frac{\sum (x_i - \bar{x})^2 f_i}{\sum f}}.$$

Середнє квадратичне відхилення ще має назву **стандартного відхилення** і показує, на скільки одиниць у середньому індивідуальні значення ознаки відхиляються від середнього значення. Далі обидва терміни будемо вживати як синоніми.

Лінійний коефіцієнт варіації визначається за формулою:

$$V = \frac{d}{x} \cdot 100, \text{ квадратичний} - \text{за формулою: } V = \frac{\sigma}{x} \cdot 100.$$

Коефіцієнт варіації показує, на скільки відсотків у середньому індивідуальні значення ознаки відрізняються від середнього значення ознаки.

Для альтернативної ознаки, коли p – частка елементів з певною ознакою у сукупності, дисперсія визначається за формулою:

$$\sigma^2 = p(1-p).$$

Приклад 2. Виробник батарейок формату АА досліджує час їх роботи. Проведено 10 вимірів тривалості роботи, результати яких наведено у табл. 2.2. Керівництво хоче знати варіацію часу роботи батарейок.

Таблиця 2.2

Результати вимірів тривалості роботи батарейок

№ п/п	1	2	3	4	5	6	7	8	9	10
Тривалість роботи, год	25,5	26,8	24,2	25,0	27,3	26,1	23,2	28,4	27,8	25,7

Розв’язання. За цими даними розмах варіації становить:

$$R = X_{\max} - X_{\min} = 28,4 - 23,2 = 5,2 (\text{год}).$$

Для знаходження середнього квадратичного відхилення спочатку потрібно визначити середній час роботи:

$$\bar{x} = \frac{\sum x}{n} = \frac{25,5 + 26,8 + 24,2 + 25,0 + 27,3 + 26,1 + 23,2 + 28,4 + 27,8 + 25,7}{10} = \frac{260}{10} = 26.$$

Далі визначаємо середнє квадратичне відхилення (табл. 2.3).

Таблиця 2.3

Розрахунки для визначення варіації тривалості роботи батарейок

№ п/п	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	25,5	-0,50	0,25
2	26,8	0,80	0,64
3	24,2	-1,80	3,24
4	25,0	-1,00	1,00
5	27,3	1,30	1,69
6	26,1	0,10	0,01
7	23,2	-2,80	7,84
8	28,4	2,40	5,76
9	27,8	1,80	3,24
10	25,7	-0,30	0,09
Разом	—	—	23,76

Маємо:

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}} = \sqrt{\frac{23,76}{10}} = \sqrt{2,376} = 1,54.$$

Отже в середньому час роботи батарейки відрізняється від середнього часу роботи у 26 годин на 1,54 год.

4. ВИБІРКОВЕ СПОСТЕРЕЖЕННЯ

Вибірковим називається вид несущільного спостереження, за характеристиками відібраної частини одиниць якого судять про всю сукупність.

Сукупність, з якої відбираються елементи для обстеження, називають **генеральною**, а сукупність, яку безпосередньо обстежують, **вибірковою**. Оскільки вибірка сукупності не точно відтворює структуру генеральної, то її характеристики не збігаються з характеристиками генеральної сукупності. Розбіжності між ними називають **помилками репрезентативності**. За причинами виникнення ці помилки поділяються на **систематичні** (тенденційні) та **випадкові**. **Систематичні помилки** виникають тоді, коли при формуванні вибіркової сукупності порушується принцип випадковості вибору (упереджений вибір елементів, недосконала основа вибірки тощо). **Випадкові помилки** неминуче виникають при дотриманні принципу випадковості вибору елементів для дослідження.

У практиці вибірових спостережень використовують два типи вибірових оцінок – **точкові** та **інтервальні**. **Точкова оцінка** – це значення параметра за даними вибірки: вибіркове середнє \bar{x} або вибіркова частка p . Достовірнішими є інтервальні оцінки параметрів.

5. ІНТЕРВАЛЬНА ОЦІНКА ПАРАМЕТРА ДОСЛІДЖУВАНОЇ ОЗНАКИ

Інтервальна оцінка – це інтервал значень параметра, розрахований за даними вибірки для певної ймовірності, тобто довірчий інтервал. На основі точкового параметра вибіркової сукупності можна визначити інтервальний параметр генеральної сукупності.

Для генерального середнього довірчий інтервал для певного рівня ймовірності визначається за формулою:

$$\bar{x} - t \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + t \frac{\sigma}{\sqrt{n}};$$

для генеральної частки:

$$p - t \sqrt{\frac{p(1-p)}{n}} \leq W \leq p + t \sqrt{\frac{p(1-p)}{n}},$$

де \bar{x} – вибіркове середнє певної ознаки; p – вибіркова частка певної ознаки; μ – генеральне середнє; W – генеральна частка; t – коефіцієнт кратності помилки або коефіцієнт довіри; $\sqrt{\frac{p(1-p)}{n}}$ – середня помилка вибірки для частки; $\frac{\sigma}{\sqrt{n}}$ – середня помилка вибірки для середнього.

Дисперсія вибірових середніх у n разів менша від дисперсії ознаки у генеральній сукупності, тому між вибірковим і середньоквадратичним відхиленням генеральної сукупності існує співвідношення:

$$\sigma_x = \frac{\sigma}{\sqrt{n}},$$

де σ_x – вибіркове середнє квадратичне відхилення; σ – середнє квадратичне відхилення для генеральної сукупності; n – обсяг вибірки.

Добуток середньої помилки вибірки на коефіцієнт довіри має назву **граничної помилки вибірки** і позначається:

$$\text{для середньої} - \Delta = t \frac{\sigma}{\sqrt{n}}, \text{ для частки} - \Delta = t \sqrt{\frac{p(1-p)}{n}}.$$

6. КОНТРОЛЬНІ ЗАПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Що таке статистичний показник?
2. У чому полягає суть і значення середніх величин?
3. Назвіть основні види середніх величин.
4. Що таке мода і медіана?
5. В яких випадках застосовують структурні середні?
6. Що розуміють під варіацією ознаки?
7. Як визначити амплітуду коливання ознаки?
8. Як визначається коефіцієнт варіації, і яке його практичне значення?
9. Які властивості притаманні дисперсії?
10. Якими показниками характеризують форми розподілу?
11. Назвіть види і способи відбору у вибіркиму сукупність.
12. Як визначається помилка вибірки?

13. Як поширюються дані вибіркового спостереження на генеральну сукупність?
14. На чому ґрунтується обчислення необхідної чисельності вибірки?
15. Як визначаються межі довірчих інтервалів для середньої величини і частки?

Лекція 3. СТАТИСТИЧНА ПЕРЕВІРКА ГІПОТЕЗ

План викладу і засвоєння матеріалу

1. Статистична гіпотеза. Основні поняття і принципи перевірки.
2. Перевірка гіпотез про рівність дисперсій.
3. Перевірка гіпотез про рівність середніх
4. Перевірка гіпотез про частку сукупності.
5. Контрольні запитання для самоперевірки.

Ключові поняття та терміни

Статистична гіпотеза, рівень значущості, помилки I та II роду, критичні точки та області.

1. СТАТИСТИЧНА ГІПОТЕЗА. ОСНОВНІ ПОНЯТТЯ І ПРИНЦИПИ ПЕРЕВІРКИ

Статистична гіпотеза – це певне припущення щодо властивостей генеральної сукупності, яке перевіряють за даними вибіркового спостереження. Найчастіше це гіпотеза про характеристики ознаки якості та про параметри відомого розподілу.

Гіпотеза, яку потрібно перевірити, формулюється як відсутність розбіжностей між параметром генеральної сукупності C і заданою величиною a . Ця висунута гіпотеза має назву **нульової** і записується так:

$$H_0 : C = a .$$

Кожній нульовій гіпотезі H_0 висувається **конкуруюча (альтернативна) гіпотеза** $H_1(Ha)$, яка вступає у протиріччя з H_0 . Залежно від виду і вагомості відхилень вона може бути сформульована так:

$$H_1 : C > a, \quad H_1 : C < a, \quad H_1 : C \neq a.$$

У результаті перевірки гіпотези можна припуститися помилок двох видів.

Помилка першого виду полягає в тому, що буде відкинута правильна гіпотеза H_0 . Ймовірність помилки першого роду називають **рівнем значущості** (істинності) і позначають α .

Помилка другого виду полягає в тому, що буде прийнято неправильну гіпотезу, її ймовірність позначається як β . Тобто прийнято гіпотезу H_0 , коли насправді правильною є альтернативна гіпотеза H_1 .

Правило, за яким статистична гіпотеза H_0 відкидається або не відкидається, називають **статистичним критерієм**. Таким критерієм є певна випадкова величина, яка застосовується для перевірки гіпотези, і закон розподілу якої відомий.

Емпіричним (спостережуваним) значенням $K_{\text{спост}}$ називають значення критерію, яке обчислене за вибіркою.

Областю прийняття гіпотези (допустимих значень) називають сукупність значень критерію, при яких нульову гіпотезу H_0 приймають.

Критичною областю називають сукупність значень критерію, при яких нульову гіпотезу відкидають (рис.3.1).

Загальний принцип перевірки статистичних гіпотез такий: якщо спостережуване значення критерію належить критичній області, то нульову гіпотезу H_0 відкидають і приймають альтернативну гіпотезу H_1 .

Якщо значення критерію, отримане на основі вибірових даних, належить області прийняття гіпотези, то приймають нульову гіпотезу H_0 .

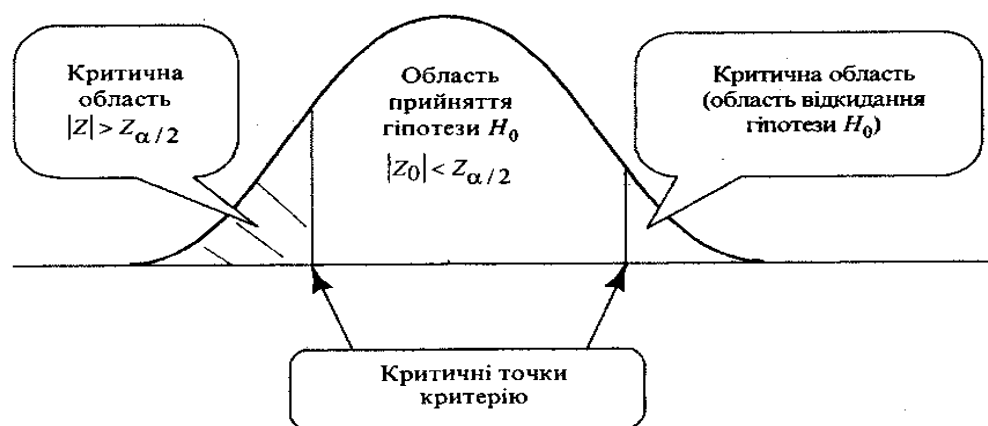


Рис. 3.1. Критерії перевірки гіпотез

Критичні точки $K_{\text{кр}}$ – це ті значення критерію, які відокремлюють критичну область від області прийняття гіпотези.

Критичні області бувають такі:

- правосторонні, тоді $K > K_{кр}$;
- лівосторонні, у цьому разі $K < K_{кр}$;
- двосторонні, при цьому $K < K_1$ або $K > K_2$;
- симетричні відносно 0, тоді $K < -K_{кр}$ або $K > K_{кр}$.

2. ПЕРЕВІРКА ГІПОТЕЗ ПРО РІВНІСТЬ ДИСПЕРСІЙ

На практиці часто виникають ситуації, коли потрібно перевірити точність устаткування, пристроїв перевірки, стабільність технологічних процесів. З цією метою на основі вибіркової дисперсії потрібно перевірити гіпотезу про рівність генеральної дисперсії досліджуваної сукупності певному очікуваному значенню.

Припустимо, що з певної генеральної сукупності взята вибірка, де n – обсяг вибірки і s^2 – незміщена вибіркова дисперсія для цього обсягу. Якщо потрібно перевірити гіпотезу про рівність генеральної дисперсії гіпотетичній генеральній дисперсії, тобто

$$H_0 : \sigma^2 = \sigma_0^2, \quad H_1 : \sigma^2 \neq \sigma_0^2,$$

де σ^2 – невідома генеральна дисперсія; σ_0^2 – гіпотетичне значення генеральної дисперсії, то необхідно обчислити спостережуване значення критерію хі-квадрат χ^2 Пірсона:

$$\chi_{\text{спост}}^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

і з таблиці критичних точок χ^2 – розподілу для рівня значущості α і числа степенів вільності $k = n-1$ знайти критичну точку $\chi_{кр}^2(\alpha, k)$. Дисперсія генеральної сукупності σ^2 оцінюється на основі вибіркової дисперсії s^2 .

У тому разі, якщо $\chi_{\text{спост}}^2 > \chi_{кр}^2$, нульова гіпотеза H_0 відкидається на користь альтернативної. Якщо $\chi_{\text{спост}}^2 < \chi_{кр}^2$ – приймають нульову гіпотезу.

Приклад 1. На основі хронометражу часу, потрібного на встановлення певного елемента у виріб різними працівниками, визначено, що дисперсія дорівнює $\sigma_0^2 = 2$ (хв.). Результати хронометражу часу роботи нового працівника за 20 спостереженнями наведені у таблиці:

Час, x	56	58	60	62	64
Кількість, f	1	4	10	3	2

Чи можемо ми вважати, що за рівня значущості $\alpha = 0,05$ новий робітник працює ритмічно?

Розв'язання. Потрібно перевірити гіпотезу, що дисперсія часу, який витрачає на встановлення елемента новий робітник, істотно не відрізняється від дисперсії часу інших робітників (гіпотетичної дисперсії).

Визначаємо дисперсію витрат часу нового робітника. Середній час за формулою середнього арифметичного зваженого дорівнює:

$$\bar{x} = \frac{56 \cdot 1 + 58 \cdot 4 + 60 \cdot 10 + 62 \cdot 3 + 64 \cdot 2}{1 + 4 + 10 + 3 + 2} = \frac{1202}{20} = 60,1.$$

№п/п	x	f	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
1	56	1	16,81	16,81
2	58	4	4,41	17,64
3	60	10	0,01	0,1
4	62	3	3,61	10,83
5	64	2	15,21	30,42
Середнє	60,1	-	-	3,98

Дисперсія становить: $s^2 = 3,98$.

Знаходимо розрахункове значення критерію χ^2 :

$$\chi_{\text{спост}}^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(20-1) \cdot 3,98}{2} = 37,81.$$

Оскільки альтернативна гіпотеза має вид рівності, то потрібно встановити ліву і праву критичні точки. В нашому випадку для 19 степенів вільності ліва критична точка $\chi_{\text{лів.кр}}^2 = 8,91$ і права критична точка $\chi_{\text{прав.кр}}^2 = 32,9$. Спостережуване значення перевищує праву критичну точку, тому є підстави відкинути гіпотезу про рівність вибіркової дисперсії гіпотетичній і зробити висновок, що новий робітник працює неритмічно.

Коли виникає потреба порівняти два методи вимірювання, точність роботи двох видів обладнання, однорідність двох сукупностей тощо, доцільно скористатися перевіркою гіпотези про рівність двох дисперсій генеральних сукупностей.

Нехай з двох генеральних сукупностей отримано вибірки обсягом n_1 і n_2 , для кожної з яких визначено скориговану (незміщену) вибірку дисперсію s_x^2 та s_y^2 відповідно. Потрібно порівняти дисперсії цих сукупностей, тобто перевірити гіпотезу:

$$H_0 : \sigma_1^2 = \sigma_2^2, \quad H_1 : \sigma_1^2 \neq \sigma_2^2,$$

за рівня значущості α .

Для цього потрібно знайти спостережуване значення критерію як відношення більшої дисперсії до меншої:

$$F_{\text{спост}} = \frac{S_6^2}{S_M^2}.$$

Потім за таблицею критичних точок розподілу Фішера – Снедекора для рівня значущості α і числа степенів вільності $k_1 = n-1$, $k_2 = n-1$ (де значення k_1 беремо для більшої дисперсії) потрібно знайти критичну точку $F_{\text{кр}}(\alpha, k_1, k_2)$ і порівняти спостережуване і критичне значення критерію:

якщо $F_{\text{спост}} > F_{\text{кр}}$ – відкидаємо гіпотезу H_0 ;

якщо $F_{\text{спост}} < F_{\text{кр}}$ – приймаємо гіпотезу H_0 .

Приклад 2. За допомогою двох методів виміряно певну величину. За першим методом отримано такі результати – 9,6; 10,0; 9,8; 10,2; 10,6, за другим – 10,4; 9,7; 10,0; 10,3.

Чи можемо ми вважати, що обидва методи забезпечують однакову точність вимірів, якщо приймемо рівень значущості $\alpha = 0,1$?

Розв’язання. Для перевірки точності методів порівняємо їхні дисперсії. Якщо дисперсії однакові, то можемо вважати, що методи забезпечують однакову точність вимірювання. Перевіримо нульову гіпотезу про рівність дисперсій:

$$H_0 : \sigma_1^2 = \sigma_2^2;$$

за альтернативної гіпотези:

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

за рівня значущості $\alpha = 0,1$.

Спочатку знайдемо вибірові дисперсії.

Дисперсії для першого методу			Дисперсія для другого методу		
№ п/п	Спостереження	$(x - \bar{x})^2$	№п/п	Спостереження	$(x - \bar{x})^2$
1	9,6	0,1936	1	10,4	0,09
2	10	0,0016	2	9,7	0,16
3	9,8	0,0576	3	10	0,01
4	10,2	0,0256	4	10,3	0,04
5	10,6	0,3136	Середнє	10,1	0,075
Середнє	10,04	0,1184			

Знаходимо спостережуване значення критерію як відношення більшої дисперсії до меншої:

$$F_{\text{спост}} = \frac{S_6^2}{S_M^2} = \frac{0,1184}{0,075} = 1,58.$$

За умовою конкуруюча гіпотеза має вид нерівності, тому критична область є двосторонньою, і для визначення критичного значення критерію потрібно брати рівень значущості вдвічі менший, ніж заданий.

З таблиці критичних точок розподілу Фішера – Снедекора для рівня значущості $\alpha/2 = 0,1/2 = 0,05$ і числа степенів вільності $k_1=5-1=4$ і $k^2=4-1=3$ знаходимо критичну точку $F(0,05; 4; 3) = 9,12$.

Оскільки $F_{\text{спост}} < F$, то немає підстав відкидати гіпотезу про рівність дисперсій. Отже, можемо вважати, що обидва методи забезпечують однакову точність вимірювання.

3. ПЕРЕВІРКА ГІПОТЕЗ ПРО РІВНІСТЬ СЕРЕДНІХ

3.1. Порівняння середнього генеральної сукупності з його очікуваним значенням

Виготовлена продукція повинна відповідати вимогам стандартів, виробничим специфікаціям. Кількісною характеристикою міри такої відповідності є результати перевірки гіпотези про рівність середнього значення досліджуваної сукупності певному очікуваному або стандартно встановленому значенню.

У тому разі, якщо дисперсія генеральної сукупності відома, для перевірки гіпотези

$$H_0 : \mu_1 = \mu_0, \quad H_1 : \mu_1 \neq \mu_0$$

про рівність середньої величини гіпотетичному генеральному середньому, де μ_1 – невідоме середнє із сукупності з відомою дисперсією; μ_0 гіпотетичне значення середнього цієї сукупності, потрібно обчислити спостережуване значення критерію:

$$t_0 = \frac{(\bar{x} - \mu_0)\sqrt{n}}{\sigma}$$

і з таблиці функції Лапласа знайти критичну точку $u_{\text{кр}}$ з рівняння:

$$\Phi(u_{\text{кр}}) = \frac{1 - \alpha}{2}.$$

Якщо $|t_0| > u_{\text{кр}}$ – відкидаємо гіпотезу H_0 , і навпаки, коли $|t_0| < u_{\text{кр}}$ – приймаємо H_0 за визначеного рівня значущості α .

У тому разі, якщо дисперсія генеральної сукупності невідома, для перевірки гіпотези про рівність невідомого генерального середнього гіпотетичному середньому потрібно використати критерій такого виду:

$$t = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s},$$

де s – незміщене вибіркове середнє квадратичне відхилення.

Цей критерій має розподіл Стюдента з $k = n - 1$ степенями вільності.

Розрахункове значення критерію порівнюється з критичними точками розподілу Стюдента за визначеного рівня значущості α та з $k = n - 1$ степенями вільності.

3.2. Порівняння середніх двох сукупностей

На виробництві часто виникають ситуації, коли потрібно порівняти між собою роботу обладнання, якість отриманої від різних постачальників сировини, якість виготовленої за різних умов продукції тощо. Для здійснення такого порівняння доцільно перевірити гіпотезу про рівність середніх двох сукупностей. Порівняння здійснюється на основі визначених з вибірки значень середніх. При цьому вибірккові середні порівнюються не з певним стандартно визначеним значенням, якого часто не існує, а одне з одним.

Коли ми вважаємо, що дисперсії є рівними, для перевірки гіпотези про рівність середніх двох сукупностей

$$H_0 : \mu_1 = \mu_2, \quad H_1 : \mu_1 \neq \mu_2$$

використовується критерій

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_y \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

який має розподіл Стюдента з двома степенями вільності $\nu = n_1 + n_2 - 2$ за рівня значущості α , де n_1, n_2 – обсяг першої та другої вибіркової сукупності відповідно; S_y – узагальнена дисперсія.

Використовувану у критерії узагальнену дисперсію визначають на основі вибіркових дисперсій першої і другої сукупностей відповідно:

$$S_y = \frac{(n-1)s_1^2 + (n-1)s_2^2}{n_1 + n_2 - 2}.$$

Нульову гіпотезу про рівність вибірових середніх H_0 відкидаємо, якщо абсолютне спостережуване значення критерію $|t|$ перевищує критичне значення $t_{\alpha/2;v}$:

$$|t| > t_{\alpha/2;v}.$$

4. ПЕРЕВІРКА ГІПОТЕЗ ПРО ЧАСТКУ СУКУПНОСТІ

Порівняння спостережуваної частки з гіпотетичною часткою сукупності

Для кількісної характеристики відповідності частоти появи певної події визначеному значенню доцільно скористатися гіпотезою про порівняння спостережуваної та стандартної частоти появи події. Припустимо, є n незалежних дослідів, ймовірність появи події p в яких є постійною, але невідомою. З цієї сукупності знайдена відносна частота кількості певних подій m у загальній кількості n одиниць. Потрібно перевірити гіпотезу, що вибіркова частота $p = m/n$ дорівнює гіпотетичній частоті p_0 , тобто:

$$H_0 : p = p_0, \quad H_1 : p \neq p_0.$$

Для перевірки цієї гіпотези за визначеного рівня значущості α потрібно визначити значення критерію

$$V_{\text{спост}} = \frac{\left(\frac{m}{n} - p_0\right)\sqrt{n}}{\sqrt{p_0 q_0}}$$

і з таблиці функції Лапласа знайти критичну точку $u_{\text{кр}}$ з рівняння

$$\Phi(u_{\text{кр}}) = \frac{1 - \alpha}{2}.$$

Якщо $|V_{\text{спост}}| > u_{\text{кр}}$ – відкидаємо гіпотезу H_0 ; якщо $|V_{\text{спост}}| < u_{\text{кр}}$ – приймаємо гіпотезу H_0 .

Приклад 3. Товарознавець магазину може прийняти партію товарів, якщо частка з певними невідповідностями не перевищує 0,03. І Серед 400 відібраних випадковим чином одиниць товару виявилось 18 бракованих. Чи може товарознавець прийняти партію?

Розв'язання. Прийmemo нульову гіпотезу $H_0 : p = p_0 = 0,03$ за альтернативної $H_1 : p > p_0 > 0,03$ і рівня значущості $\alpha = 0,05$.

Визначимо вибірку частку невідповідних товарів:

$$\frac{m}{n} = \frac{18}{400} = 0,045.$$

Розрахуємо спостережуване значення критерію:

$$V_{\text{спост}} = \frac{\left(\frac{m}{n} - p_0\right)\sqrt{n}}{\sqrt{p_0 q_0}} = \frac{(0,045 - 0,03) \cdot \sqrt{400}}{\sqrt{0,03 \cdot (1 - 0,03)}} = 1,75.$$

Знаходимо критичну точку правосторонньої області:

$$\Phi(u_{\text{кр}}) = \frac{1 - 2 \cdot 0,05}{2} = 0,45.$$

З таблиці функції Лапласа знаходимо критичну точку $u_{\text{кр}} = 1,645$. Оскільки спостережуване значення критерію $1,75 > 1,645$, то нульову гіпотезу відкидаємо і робимо висновок, що прийняти партію товарів неможливо.

5. КОНТРОЛЬНІ ЗАПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Що розуміють під статистичною гіпотезою?
2. Які є види статистичних гіпотез?
3. Як називається гіпотеза, яку необхідно перевірити?
4. Охарактеризуйте гіпотезу, протилежну нульовій.
5. Як називають випадки прийняття неправильного рішення під час перевірки статистичних гіпотез?
6. Назвіть помилки I-го роду, які можуть виникнути у процесі перевірки статистичних гіпотез.
7. Що називається потужністю критерію під час перевірки статистичних гіпотез?
8. На чому ґрунтуються помилки I та II роду?
9. Які статистичні критерії належать до непараметричних?
10. Від чого залежить вибір критичного простору?

Лекція 4. СТАТИСТИЧНІ МЕТОДИ ВИМІРЮВАННЯ ВЗАЄМОЗВ'ЯЗКІВ ЯВИЩ

План викладу і засвоєння матеріалу

1. Види взаємозв'язків між явищами.
2. Метод аналітичного групування.

3. Дисперсійний аналіз.
4. Основи кореляційно-регресійного аналізу.
5. Непараметричні методи виявлення та вимірювання зв'язків.
6. Контрольні запитання для самоперевірки.

Ключові поняття та терміни

Факторні і результативні ознаки, кореляційний зв'язок, аналітичне групування, дисперсійний аналіз, коефіцієнт детермінації, кореляційне відношення, індекс кореляції, критерії Фішера F і Стюдента t , коефіцієнт регресії, лінійний коефіцієнт кореляції, коефіцієнт Фехнера, коефіцієнти асоціації та контингенції, коефіцієнт Спірмена.

1. ВИДИ ВЗАЄМОЗВ'ЯЗКІВ МІЖ ЯВИЩАМИ

Зв'язки між явищами, окремими їх ознаками досить різноманітні, однак у будь-якому випадку одні ознаки виступають як фактори, що впливають на інші і зумовлюють їх зміну, інші – як результати дії цих факторів. Одні із них є причиною, інші наслідком.

Якщо перші прийнято називати **ознаками-факторами** або **факторними (причинними) ознаками**, то другі – **результативними (наслідковими) ознаками**.

За статистичною природою зв'язки поділяють на **функціональні** і **стохастичні**. При **функціональному** зв'язку кожному можливному значенню факторної ознаки x відповідає чітко визначене значення результативної ознаки – y , тобто функціональні зв'язки характеризуються повною відповідністю між причиною і наслідком, факторною і результативною ознаками.

На відміну від функціональних, стохастичні зв'язки неоднозначні. Стохастичні зв'язки проявляються як узгодженість варіації двох чи більше ознак. У ланці зв'язку « $x \rightarrow y$ » кожному значенню ознаки x відповідає певна множина значень ознаки y , які утворюють так званий **умовний розподіл**. Стохастичний зв'язок, відбиваючи множинність причин і наслідків, виявляється в зміні умовних розподілів.

Якщо умовні розподіли замінюються одним параметром – середньою \bar{y}_i , то такий зв'язок називають **кореляційним**. Отже, кореляційний зв'язок є різновидом стохастичного і виявляється в зміні середніх умовних розподілів, що схематично ілюструє табл. 4.1.

Види взаємозв'язків

Факторна ознака, X_i	Результативна ознака у за наявності зв'язку		
	функціонального	стохастичного	кореляційного
X_1	Y_1	$Y_1 Y_2$	$\overline{Y_1}$
X_2	Y_2	$Y_1 Y_2 Y_3$	$\overline{Y_2}$
X_3	Y_3	$Y_2 Y_3 Y_4$	$\overline{Y_3}$
X_n	Y_n	$Y_n Y_{n-1}$	$\overline{Y_4}$

За напрямком дії (спрямованістю) розрізняють зв'язок **прямий** і **обернений**. **Прямий** – це такий зв'язок, при якому зі збільшенням або зменшенням значень факторної ознаки відповідно збільшується або зменшується значення результативної ознаки, тобто факторна і результативна ознаки змінюються в одному напрямку.

Прикладом прямого зв'язку може бути зв'язок між фондоозброєністю і продуктивністю праці, між собівартістю продукції і рівнем рентабельності.

Оберненим зв'язком називають такий, при якому значення результативної ознаки змінюється в протилежному напрямку відносно зміни значення факторної ознаки. Прикладом такого зв'язку може бути продуктивність праці і собівартість продукції.

За формою аналітичного вираження в загальній класифікації виділяють зв'язки **прямолінійні** та **криволінійні**. Якщо певний зв'язок явищ можна точно або наближено зобразити рівнянням будь-якої прямої лінії, то його називають **лінійним (прямолінійним)** зв'язком, а якщо рівнянням будь-якої кривої лінії (параболи, гіперболи і т.п.) – **нелінійним (криволінійним)**.

Для відповіді на питання про наявність або відсутність кореляційного зв'язку використовують ряд специфічних методів:

- елементарні прийоми (паралельне порівняння рядів значень факторної і результативної ознак, балансовий метод, графічне зображення, метод аналітичного групування);
- дисперсійний аналіз;
- кореляційно-регресійний аналіз.

2. МЕТОД АНАЛІТИЧНОГО ГРУПУВАННЯ

Метод аналітичного групування полягає в тому, що всі елементи сукупності групують, як правило, за факторною ознакою x і в кожній групі обчислюють середні значення результативної ознаки y , тобто лінія регресії оцінюється лише в окремих точках, які відповідають певному значенню x .

Наприклад, розглянемо дані про обіг коштів і прибуток десяти банків (табл. 4.2).

Таблиця 4.2

Показники діяльності банків у звітному періоді

№ банку	Обіг коштів, млн грн	Прибуток, млн грн
1	7	3
2	11	4
3	12	4
4	16	6
5	22	8
6	27	7
7	30	11
8	38	14
9	45	18
10	46	17

Проведемо комбіноване групування банків за двома ознаками: за обігом коштів і розміром одержаного прибутку, утворивши по три групи з рівними інтервалами (табл. 4.3).

Таблиця 4.3

Розподіл банків за обігом коштів та розміром прибутку

Групи банків за обігом коштів млн грн	Кількість банків за розміром прибутку, млн грн			Разом
	3-8	8-13	13-18	
7-20	4			4
20-33	2	1		3
33-46			3	3
Разом	6	1	3	10

Побудовану таким чином таблицю називають кореляційною таблицею. Таблиця показує, що частоти розподілу банків концентруються по діагоналі, яка йде з верхнього лівого кута у правий нижній. Це свідчить про те, що існує прямий зв'язок між обігом коштів і розміром прибутку банків. Із зростанням розмірів обігу коштів збільшується прибуток банків.

На другому етапі проводиться оцінка лінії регресії – у кожній групі за факторною ознакою обчислюють середні значення результативної та факторної ознак. Групові середні обчислюємо за вихідними незгрупованими даними. У нашому прикладі:

$$\bar{y}_1 = \frac{3+4+4+6}{4} = 4,3$$

$$\bar{x}_1 = \frac{7+11+12+16}{4} = 11,5$$

$$\bar{y}_2 = \frac{8+7+11}{3} = 8,7$$

$$\bar{x}_2 = \frac{22+27+30}{3} = 26,3$$

$$\bar{y}_3 = \frac{14+18+17}{3} = 16,3$$

$$\bar{x}_3 = \frac{38+45+46}{3} = 43,0$$

Результати розрахунків оформлюємо у вигляді табл. 4.4.

Таблиця 4.4

Залежність прибутку від обігу коштів банків

№ з/п	Групи банків за обігом коштів, млн грн	Кількість банків f	Середній прибуток, млн грн \bar{y}_i	Середній обіг коштів, млн грн. \bar{x}_i
1	7-20	4	4,3	11,5
2	20-33	3	8,7	26,3
3	33-46	3	16,3	43,0
	Разом	10	9,2	25,4

Зростання групових середніх рівня прибутку від групи до групи свідчить про наявність кореляційного зв'язку між обігом коштів і розміром прибутку.

Обчислимо ефект впливу розміру обігу коштів на рівень прибутку банків за даними табл. 4.4:

$$\frac{\Delta_y}{\Delta_x} = \frac{\bar{y}_2 - \bar{y}_1}{\bar{x}_2 - \bar{x}_1} = \frac{8,7 - 4,3}{26,3 - 11,5} = 0,297; \quad \frac{\Delta_y}{\Delta_x} = \frac{\bar{y}_3 - \bar{y}_2}{\bar{x}_3 - \bar{x}_2} = \frac{16,3 - 8,7}{43,0 - 26,3} = 0,455.$$

Отже, з підвищенням обігу коштів на 1 млн. грн. прибуток збільшується: в другій групі порівняно з першою на 0,297 млн. грн., в третій групі порівняно з другою на 0,455 млн. грн.

Третій етап аналітичного групування – вимірювання тісноти зв'язку за допомогою дисперсійного аналізу.

3. ДИСПЕРСІЙНИЙ АНАЛІЗ

Основною метою дисперсійного аналізу є виявлення впливу окремих факторів чи умов, які визначають варіацію ознаки. В основі дисперсійного аналізу лежить закон розкладання загальної дисперсії на складові, згідно з яким загальна дисперсія результативної ознаки у складається з двох частин: **міжгрупової (факторної) дисперсії** та **середньої з групових (залишкової)**.

Взаємозв'язок факторної та залишкової варіації описується **правилом розкладання дисперсії**:

$$\sigma_y^2 = \delta_y^2 + \overline{\sigma_y^2}.$$

Загальна дисперсія характеризує варіацію результативної ознаки під впливом всіх факторів і причин, як систематично діючих, так і випадкових. Загальна дисперсія результативної ознаки обчислюється за індивідуальними значеннями ознаки у:

$$\sigma_y^2 = \frac{\sum y^2}{n} - \left(\frac{\sum y}{n} \right)^2.$$

Міжгрупова дисперсія характеризує варіацію групових середніх, тобто варіацію результативної ознаки, яка пов'язана з варіацією групувальної факторної ознаки:

$$\delta^2 = \frac{\sum (\bar{y}_i - \bar{y})^2 f}{\sum f}.$$

Середня з групових дисперсій характеризує варіацію результативної ознаки, пов'язану з варіацією всіх факторних ознак, крім тієї, яка покладена в основу групування:

$$\overline{\sigma^2} = \frac{\sum \sigma_i^2 f}{\sum f} \quad \text{або} \quad \overline{\sigma^2} = \sigma_y^2 - \delta^2.$$

Відношення міжгрупової (факторної) дисперсії до загальної розглядається як міра щільності кореляційного зв'язку і називається **коефіцієнтом детермінації**:

$$\eta^2 = \frac{\delta^2}{\sigma^2}.$$

За статистичною структурою це відношення є часткою варіації результативної ознаки у, яка пов'язана з варіацією ознаки х. Здобувши

квадратний корінь із цього відношення, одержуємо емпіричне **кореляційне відношення**:

$$\eta = \sqrt{\frac{\delta^2}{\sigma^2}}.$$

Кореляційне відношення змінюється від 0 до 1. Якщо $\eta = 0$, то міжгрупова дисперсія дорівнює нулю. Це можливо лише за умови, коли всі групові середні однакові і кореляційний зв'язок між ознаками відсутній. При $\eta = 1$ міжгрупова дисперсія дорівнює загальній, а середня з групових – нулю. В цьому випадку кожному значенню факторної ознаки відповідає єдине значення результативної ознаки, тобто зв'язок між ознаками функціональний.

Індекс кореляції визначають зіставленням внутрішньо групової і загальної дисперсії, і обчислюють за формулою:

$$R = \sqrt{1 - \frac{\sigma^2}{\sigma^2}}.$$

Для оцінки надійності кореляційних характеристик використовують критерії Фішера F або Стюдента t .

Критерій Фішера (F -критерій) визначається за формулою:

$$F_{\Phi} = \frac{\delta^2 k_2}{\sigma_i^2 k_1},$$

де σ^2 – міжгрупова дисперсія; σ^2 – середня групова (залишкова) дисперсія; k_1 і k_2 – степені вільності для великої і малої дисперсій.

Фішер знайшов розподіл відношень дисперсій і розробив відповідні математичні таблиці, в яких наводиться F -критерій теоретичний F_T при двох ймовірностях 0,95 і 0,99. Якщо $F_{\Phi} > F_T$, то з прийнятим ступенем ймовірності можна стверджувати про наявність впливу фактора, який вивчається. Коли ж $F_{\Phi} < F_T$, то різниця між дисперсіями зумовлена впливом випадкових факторів.

Розподіл у таблицях Фішера для знаходження F_T залежить від степенів вільності між групою k_1 і середньою з групових k_2 дисперсій. В аналітичному групуванні їх обчислюють за формулами:

$$k_1 = m - 1; \quad k_2 = n - m.$$

де n – кількість елементів досліджуваної сукупності; m – число груп.

Надійність кореляційного відношення за критерієм Стюдента (t – критерію):

$$t_{\eta} = \frac{\eta}{\mu_{\eta}},$$

де μ_{η} – середня похибка кореляційного відношення $\mu_{\eta} = \frac{1-\eta^2}{\sqrt{n}}$.

Якщо критерій Стюдента $t_{\eta} \geq 3$, показник кореляційного відношення вважають вірогідним (тобто зв'язок між досліджуваними явищами є доведеним). Якщо ж критерій $t_{\eta} < 3$, то висновки про вірогідність зв'язку між досліджуваними явищами сумнівні.

Комбінаційні аналітичні групування і дисперсійний аналіз використовують для вивчення зв'язку результативної ознаки з двома і більше факторними ознаками. Вони дають можливість аналізувати залежність результативної ознаки від кожного з факторів при фіксованих значеннях інших. Методи вимірювання такого зв'язку і перевірку його істотності називають **багатофакторними дисперсійними комплексами**.

4. ОСНОВИ КОРЕЛЯЦІЙНО-РЕГРЕСІЙНОГО АНАЛІЗУ

У кореляційно-регресійному аналізі оцінка лінії регресії здійснюється не в окремих точках, як в аналітичному групуванні, а в кожній точці інтервалу зміни факторної ознаки x . Тобто лінія регресії у даному випадку неперервна і зображується у вигляді певної функції $Y=f(x)$, яка називається **рівнянням регресії**, а Y – це теоретичні значення результативної ознаки.

Якщо зі зміною фактора x результат y змінюється більш-менш рівномірно, такий зв'язок описується лінійною функцією $Y=a+bx$. При нерівномірному співвідношенні варіацій взаємозв'язаних ознак (наприклад, коли прирости значень y зі зміною x прискорені чи сповільнені або напрям зв'язку змінюється), використовують нелінійні регресії, зокрема:

степеневу: $Y = ax^b$,

гіперболу: $Y = a + \frac{b}{x}$,

параболу: $Y = a + bx + cx^2$.

Параметри рівняння регресії визначаються методом найменших квадратів, основна умова якого – мінімізація суми квадратів відхилень емпіричних значень y від **теоретичних Y** :

$$\sum (y - Y)^2 = \min.$$

Математично доведено, що значення параметрів a та b , при яких мінімізується сума квадратів відхилень, визначається із системи нормальних рівнянь.

$$\begin{cases} \sum y = na + b \sum x \\ \sum xy = a \sum x + b \sum x^2 \end{cases}.$$

Для вимірювання щільності прямолінійних зв'язків використовується **лінійний коефіцієнт кореляції**:

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{(n \sum x^2 - (\sum x)^2) \cdot (n \sum y^2 - (\sum y)^2)}}.$$

Коефіцієнт кореляції можна обчислювати і за іншими формулами. Зокрема:

$$r = \frac{\overline{xy} - \bar{x} \bar{y}}{\sigma_x \sigma_y}; \quad r = \frac{\sum (x - \bar{x}) \cdot (y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}.$$

Якщо визначена форма кореляційного зв'язку і обчислений коефіцієнт регресії a , то коефіцієнт кореляції можна обчислити за

формулою:

$$r = a \frac{\sigma_x}{\sigma_y}.$$

Лінійний коефіцієнт кореляції може набувати любых значень в межах від -1 до $+1$. Якщо r близьке до 1 , то зв'язок між ознаками тісний, якщо r наближається до 0 , то зв'язок незначний. Знак лінійного коефіцієнта кореляції вказує напрямок зв'язку – знак «плюс» свідчить про прямий зв'язок, знак «мінус» – обернений зв'язок.

5. НЕПАРАМЕТРИЧНІ МЕТОДИ ВИЯВЛЕННЯ ТА ВИМІРЮВАННЯ ЗВ'ЯЗКІВ

Для виявлення та вимірювання зв'язків між якісними ознаками використовують непараметричні методи, які називають **ранговими методами кореляції**. Вони простіші в обчисленнях. Для обчислень використовують не самі значення ознак, а їх знаки, ранги, частоти.

Розглянемо чотири основних показники вимірювання зв'язків між ознаками, які є основою рангових методів: **коефіцієнт Фехнера**; **коефіцієнт асоціації**; **коефіцієнт контингенції**; **коефіцієнт Спірмена**.

Коефіцієнт Фехнера (коефіцієнт збігу знаків) визначають як відношення різниці числа знаків лінійних відхилень факторної та результативної ознак, що збігаються, та числа знаків, що не збігаються, до загального числа відхилень ознак від середніх.

Наприклад, якщо лінійні відхилення ознак x та y мають такі знаки для різних значень:

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$
x_1	y_1	+	–
x_2	y_2	+	+
x_3	y_3	+	+
x_4	y_4	–	–
x_5	y_5	+	–

Тоді коефіцієнт Фехнера дорівнює:

$$i = \frac{3-2}{5} = \frac{1}{5} = 0,2.$$

Значення коефіцієнта i змінюється від -1 до $+1$, і чим воно ближче до нуля, тим зв'язок між ознаками тісніший. Цей коефіцієнт може бути застосований як до якісних, так і до кількісних ознак.

Коефіцієнти асоціації та контингенції застосовують для вимірювання тісноти зв'язків якісних альтернативних ознак. При дослідженні щільності зв'язку між якісними альтернативними ознаками (протилежними за змістом) використовують розрахункову таблицю, яка складається із чотирьох комірок, кожна з яких відповідає відомій альтернативі того чи іншого показника.

	Так	Ні	Разом
Так	a	b	a + b
Ні	c	d	c + d
Разом	a + c	b + d	a + b + c + d

За даними такого макету таблиці можна обчислити коефіцієнт асоціації Д. Юла і коефіцієнт контингенції К. Пірсона.

Коефіцієнт асоціації (A) обчислюють за формулою:

$$A = \frac{a \cdot d - b \cdot c}{a \cdot d + b \cdot c}.$$

Коефіцієнт контингенції (K) обчислюють за формулою:

$$K = \frac{a \cdot d - b \cdot c}{\sqrt{(a+b) \cdot (b+d) \cdot (a+c) \cdot (c+d)}}.$$

Зв'язок між ознаками підтверджується, якщо:

$$A \geq 0,5 \quad \text{або} \quad K \geq 0,3.$$

Коефіцієнт Спірмена називають ще ранговим коефіцієнтом кореляції. Ранг – це порядковий номер ознаки у зростаючому ряді розподілу. Позначимо: R_x – ранг факторної ознаки; R_y – ранг результативної ознаки; n – кількість рангів (одиниць сукупності); $d = R_x - R_y$ – різниця рангів. Коефіцієнт Спірмена обчислюють за формулою:

$$R = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}.$$

Для варіантів, які повторюються, ранг визначають як середню арифметичну відповідних рангів, наприклад, ранг однакових величин, які посідають 4-те та 5-те місця, дорівнює 4,5.

Коефіцієнт рангової кореляції може набувати значення в межах $-1 < R < +1$.

6. КОНТРОЛЬНІ ЗАПИТАННЯ

1. Як виявляється причинний зв'язок?
2. У чому полягають особливості стохастичного і функціонального зв'язку?
3. Як виявляється кореляційний зв'язок?
4. У чому полягає прямий і обернений зв'язок?
5. Які є методи виявлення зв'язків між явищами?
6. В чому полягає суть методу аналітичного групування ?
7. Які етапи включає метод аналітичного групування ?
8. +Як визначити ефект впливу фактора на результат за даними аналітичного групування?
9. У чому полягає дисперсійний аналіз?
10. У чому полягає економічний зміст дисперсій: загальної, міжгрупової та середньої із групових дисперсій?
11. Що собою уявляє коефіцієнт детермінації?
12. Які функції в аналізі взаємозв'язків виконує рівняння регресії?
13. Чому лінійна регресія найбільш поширена?
14. Як визначити параметри рівняння прямої?

15. Яке значення має постійний параметр b в рівнянні прямої?
16. Що собою уявляє лінійний коефіцієнт кореляції?
17. Коли використовують непараметричні методи вимірювання зв'язків?
18. Як визначаються коефіцієнти асоціації та контингенції?

Лекція 5. АНАЛІЗ ЧАСОВИХ РЯДІВ. РЯДИ ДИНАМІКИ

План викладу і засвоєння матеріалу

7. Поняття про ряди динаміки, їх елементи, види та правила побудови.
8. Статистичні характеристики рядів динаміки.
9. Основні прийоми аналізу рядів динаміки.
10. Статистичне вивчення сезонних коливань.
11. Контрольні запитання для самоперевірки.

Ключові поняття та терміни

Динамічний ряд, моментні та періодичні ряди, абсолютний приріст, темп зростання, темп приросту, абсолютне значення 1% приросту, ланцюгові та базисні показники, коефіцієнт випередження, коефіцієнт еластичності, трендове рівняння, сезонні коливання, індекс сезонності, сезонна хвиля, коефіцієнт стабільності.

1. ПОНЯТТЯ ПРО РЯДИ ДИНАМІКИ, ЇХ ЕЛЕМЕНТИ, ВИДИ ТА ПРАВИЛА ПОБУДОВИ

Процес розвитку суспільних явищ у часі називається **динамікою**.

Динамічний ряд – це статистичні показники, розташовані в хронологічній послідовності, які характеризують розвиток того чи іншого соціально-економічного явища у часі.

Окремі числові значення розмірів явищ називають **рівнями ряду**. **Рівень ряду** відображає стан явищ, досягнутий за будь-який період або на певний момент часу. Перший показник ряду називається початковим, а останній – кінцевим.

При вивченні динаміки важливі не лише числові значення рівнів, але і послідовність їх. Як правило, часові інтервали поміж рівнями однакові (доба, декада, календарний місяць, квартал, рік). Приймаючи будь-який інтервал за одиницю, послідовність рівнів можна записати так; $y_1, y_2, y_3, \dots, y_n$, де n число рівнів (довжина динамічного ряду).

За ознакою часу динамічні ряди поділяють на **моментні** та **періодичні (інтервальні)**.

Рівень **моментного ряду** фіксує стан явища, його розмір або величину на відповідний момент часу. Прикладом таких рядів можуть бути дані про залишки продукції на складі підприємства на початок кожного кварталу звітної року (табл. 5.1):

Таблиця 5.1

	01.01	01.04	01.07	01.10	01.01
Залишки продукції, шт.	13500	14200	14960	14000	15300

Періодичні (інтервальні) ряди динаміки характеризують величину явища за відповідні періоди часу (добу, декаду, місяць, квартал, рік, п'ятирічку). Прикладами таких рядів можуть бути дані, наведені в табл. 5.2, 5.3.

Таблиця 5.2

**Середня урожайність кукурудзи в одному з районів області
за 2005-2010 роки**

Урожайність	2005	2006	2007	2008	2009	2010
кукурудзи, ц/га	28,1	30,2	33,6	30,4	32,3	34,1

Таблиця 5.3

**Доля державної торгівлі в роздрібному товарообороті області
за 2005-2010 роки**

Питома вага державної торгівлі в загальному обсязі роздрібного товарообігу, %	2005	2006	2007	2008	2009	2010
	53,8	61,2	60,7	51,3	48,6	42,5

Залежно від статистичної природи показника-рівня розрізняють динамічні ряди **первинні і похідні, ряди абсолютних, середніх і відносних величин**. У табл. 5.1 наведено ряд динаміки, який складається із абсолютних величин, у табл. 5.2 – із середніх величин, у табл. 5.3 – із відносних величин. Крім цього ряди динаміки поділяють на **одно-** і **багатомірні**.

Одномірні характеризують зміну одного показника (наприклад, видобуток нафти), **багатомірні** – двох, трьох і більше показників. У свою

чергу, багатомірні динамічні ряди поділяються на два види: **паралельні та ряди взаємопов'язаних показників**.

Паралельні відображають динаміку або одного і того самого показника щодо різних об'єктів (національний дохід по країнах, прибуток по підприємствах тощо), або різних показників одного і того ж об'єкта (видобуток вугілля, нафти і газу в регіоні).

Ряди взаємопов'язаних показників характеризують динаміку декількох показників, взаємопов'язаних між собою. Зв'язок між показниками багатомірного динамічного ряду може бути **функціональним (адитивним чи мультиплікативним) або кореляційним**.

При побудові рядів динаміки потрібно дотримуватись певних вимог. Найважливішою (головною) вимогою щодо побудови динамічних рядів є забезпечення їх порівнянності.

Всі рівні ряду динаміки повинні характеризувати одне і теж явище.

Кожен рівень динамічного ряду повинен бути визначений (розрахований) за однією методологією.

Важливою умовою порівнянності рядів динаміки є вираження їх рівнів в однакових одиницях вимірювання.

2. СТАТИСТИЧНІ ХАРАКТЕРИСТИКИ РЯДІВ ДИНАМІКИ

При вивченні закономірностей соціально-економічного розвитку статистика вирішує ряд завдань:

- характеристика інтенсивності окремих змін у рівнях ряду від періоду до періоду або від дати до дати;
- визначення середніх показників динамічного ряду за той чи інший період;
- виявлення основних закономірностей динаміки досліджуваного явища на окремих етапах або за весь період, що вивчається;
- виявлення факторів, що зумовили зміни досліджуваного об'єкту у часі;
- прогноз розвитку явищ на майбутнє.

Для оцінки цих властивостей динаміки статистика використовує взаємопов'язані характеристики. Серед них **абсолютний приріст, темп зростання, темп приросту і абсолютне значення 1% приросту**.

Розрахунок характеристики динаміки ґрунтується на зіставленні рівнів ряду. Базою для порівняння може бути або попередній рівень y_{n-1} , або початковий y_1 . Показники динаміки, обчислені зіставленням із змінною базою порівняння, називаються **ланцюговими**, а з постійною базою порівняння – **базисними**.

Зміст показників динаміки розглянемо на прикладі табл. 5.4 та рис. 5.1.

Таблиця 5.4

Виробництво електроенергії за 2006-2010 роки (цифри умовні)

Виробництво електроенергії, тис. кВт- год.	2006	2007	2008	2009	2010
	17,5	20,8	22,4	19,6	23,4

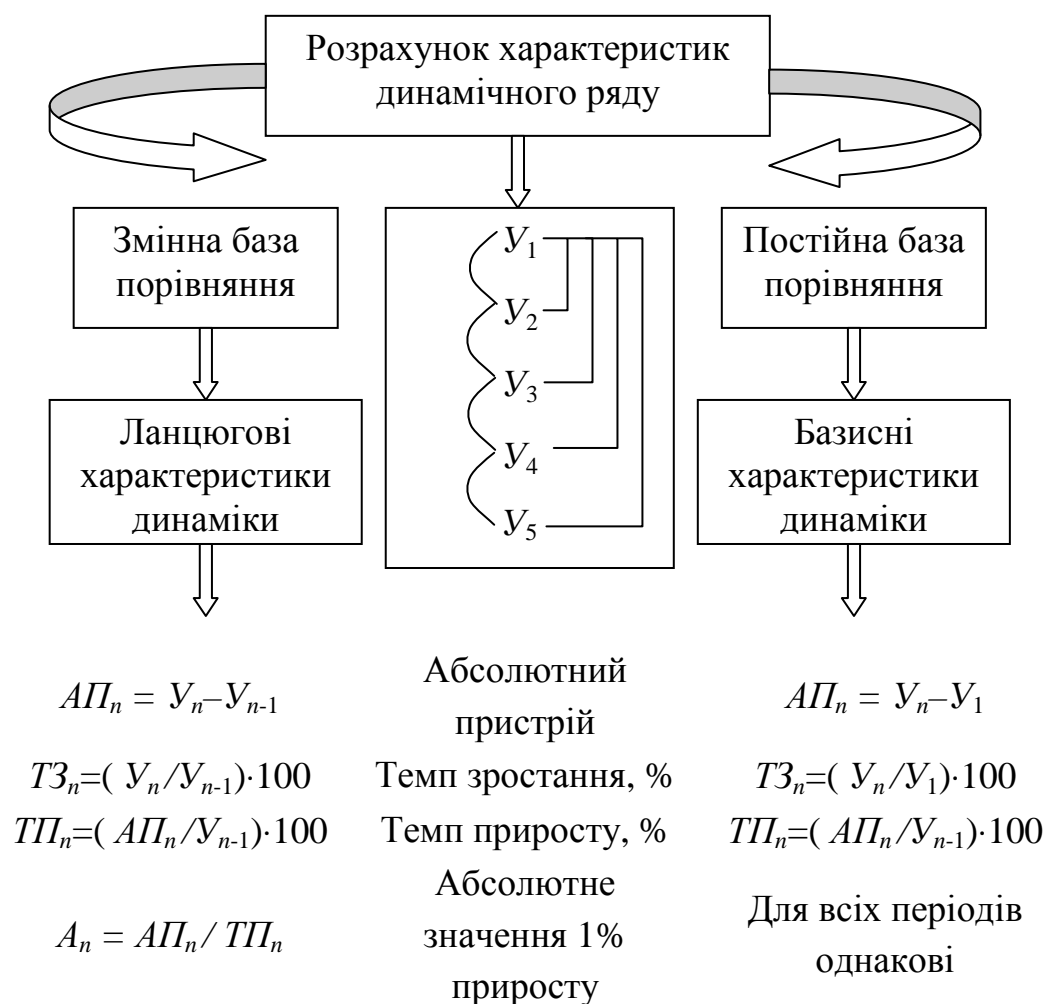


Рис. 5.1. Аналітичні характеристики рядів динаміки

У нашому прикладі початковим (базисним) рівнем є показник за 2006 рік – 17,5 тис. кВт-год. електроенергії; всі інші рівні можуть

порівнюватись або з базисним, або з попереднім, внаслідок чого будемо мати базисні показники або ланцюгові.

Абсолютний приріст ($АП_n$) відображає абсолютну швидкість змінювання рівнів ряду за певний інтервал часу. Він обчислюється як різниця рівнів ряду, знак (+,–) показує напрям динаміки.

Таким чином, абсолютний приріст може бути додатним (динаміка зростання), від’ємним (зменшення, падіння) або рівним нулю (без змін).

Абсолютний приріст вимірюється в одиницях вимірювання ознаки.

Інтенсивність зміни рівнів ряду оцінюється відносною величиною **темпом зростання**, який являє собою кратне відношення рівнів у формі коефіцієнта чи відсотка. Якщо темп зростання більший одиниці чи 100%, то це свідчить про ріст того чи іншого явища, відображеного рядом динаміки, а коли буде менший одиниці або 100%, – має місце не темп зростання, а темп зниження, зменшення, падіння.

Співвідношення абсолютного приросту і базового рівня є вимірником відносної швидкості зростання. Відносну швидкість зростання називають **темпом приросту**, який на відміну від темпу зростання завжди виражають у відсотках.

Темп приросту можна також обчислити за показниками темпів зростання:

$$ТП_n = ТЗ_n - 100\% .$$

Отже, темп приросту показує на скільки відсотків рівень звітного періоду більший (менший) від бази порівняння.

Абсолютне значення 1% приросту дає уяву про один відсоток приросту і визначається як частка від ділення абсолютного приросту на темп приросту:

$$A_n = \frac{АП_n}{ТП_n} = \frac{y_n - y_{n-1}}{100 \left(\frac{y_n - y_{n-1}}{y_{n-1}} \right)} = \frac{y_{n-1}}{100} .$$

Результати обчислень показників аналізу ряду динаміки можна подати у вигляді табл. 5.5:

Таблиця 5.5

Динаміка виробництва електроенергії за 2006-2010 роки

Роки	Виробництво електроенергії, тис. кВт-год.	Абсолютний приріст тис. кВт-год.		Темп зростання % порівняно:		Темп приросту % порівняно:		Абсолютне значення 1% приросту, тис. кВт-год
		з 2006	з попереднім роком	з 2006	з попереднім роком	з 2006	з попереднім роком	
2006	17,5	-	-	-	-	-	-	
2007	20,8	3,3	3,3	118,9	118,9	18,9	18,9	0,175
2008	22,4	4,9	1,6	128,0	107,7	28,0	7,7	0,208
2009	19,6	2,1	-2,8	112,0	87,5	12,0	-12,5	0,224
2010	23,4	5,9	3,8	133,7	119,4	33,7	19,4	0,196

Таким чином, виробництво електроенергії за період з 2006 по 2010 рр. збільшилося на 5,9 тис. кВт-год або на 33,7 %, зростання виробництва відбувалося щорічно, крім 2009 р., коли рівень виробництва електроенергії зменшився порівняно з попереднім роком на 12,5%. Найбільшого темпу зростання досягнуто у 2010 р., коли обсяг виробництва збільшився на 19,4%. Вага одного відсотку приросту з 2007 р. по 2008 р. зростала, у 2009 р. абсолютне значення 1% зниження становило 0,224 тис. кВт-год, у 2010 р. кожен відсоток приросту давав збільшення виробництва електроенергії на 0,196 тис. кВт-год. Отже, за розглянутий період характеристики абсолютної і відносної швидкості і інтенсивності зростання (зменшення) коливались, варіювали.

Якщо швидкість розвитку в межах періоду, що вивчається, неоднакова, порівнянням однойменних характеристик швидкості вимірюється **прискорення** чи **уповільнення динаміки**. На базі абсолютних приростів оцінюються абсолютне та відносне прискорення. Абсолютне – це різниця між абсолютними приростами: $r_n = \Delta\P_n - \Delta\P_{n-1}$. Прискорення характеризується додатною величиною, а уповільнення – від'ємною. Обчислимо характеристики прискорення на прикладі табл. 5.5.

2008 рік : $r_1 = 1,6 - 3,3 = -1,7$ тис. кВт-год. (уповільнення);

2009 рік : $r_2 = -2,8 - 1,6 = -4,4$ тис. кВт-год. (уповільнення);

2010 рік : $r_3 = 3,8 - (-2,8) = 6,6$ тис. кВт-год. (прискорення).

Темп зростання абсолютної швидкості обчислюється порівнянням абсолютних приростів (обчислюється тільки на підставі додатних абсолютних приростів): 2008 рік: $3,3/1,6 = 2,062$.

Порівняння темпів зростання дає коефіцієнт прискорення (уповільнення) швидкості розвитку. Для наочності та зручності тлумачення дільником є більший за значенням темп зростання.

В нашому прикладі коефіцієнти прискорення (уповільнення) відносної швидкості динаміки:

2008 рік : $118,9/107,7 = 1,104$ (уповільнення);

2009 рік : $107,7/87,5 = 1,231$ (уповільнення);

2010 рік : $119,4/87,5 = 1,365$ (прискорення).

У статистичному аналізі порівнюється також інтенсивність динаміки в різних рядах. Відношення темпів зростання TZ_2/TZ_1 називають **коефіцієнтом випередження**. За допомогою останнього порівнюють відносну швидкість динамічних рядів однакового змісту по різних об'єктах (регіони, країни тощо) або різного змісту по одному об'єкту. Наприклад, за три роки фондоозброєність праці в одній галузі зросла на 50%, в іншій – на 25%. Коефіцієнт випередження темпу зростання фондоозброєності праці в першій галузі порівняно з другою становить $1,50:1,25 = 1,20$.

Можна порівняти динаміку фондоозброєності та продуктивності праці в кожній галузі. Якщо фондоозброєність зросла на 25%, а продуктивність праці – на 37,5%, то коефіцієнт випередження зростання продуктивності праці становить $1,375:1,25 = 1,10$.

Щодо темпів приросту, то співвідношення їх використовують лише для взаємопов'язаних показників x і y . Таке співвідношення називають **коефіцієнт еластичності** $K = TP_y:TP_x$; він показує, на скільки відсотків змінюється y зі зміною x на один відсоток. Наприклад, ціна на товар А зросла на 2%, а попит зменшився на 4%. Цінова еластичність попиту на цей товар: $K = \frac{-4}{+2} = -2$, тобто із зростанням цін на 1% попит на товар зменшується на 2%.

Ряди динаміки можна подавати не тільки в таблицях, а й у вигляді графічних зображень. У цьому разі на осі абсцис відкладають шкалу часу, а на осі ординат – шкалу рівнів ряду. На підставі даних табл. 5.5 побудуємо графік (рис. 5.2).

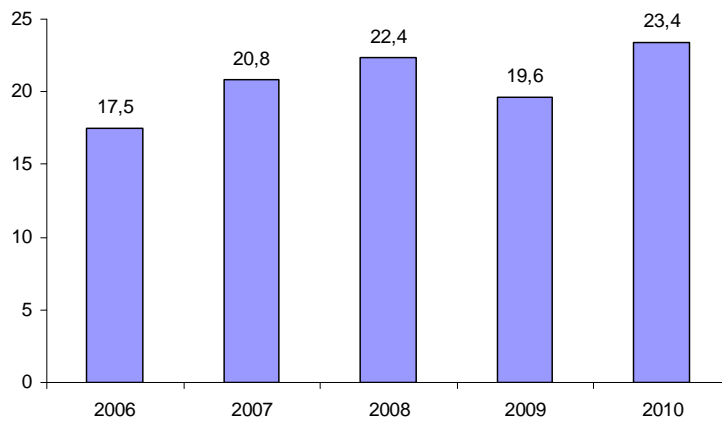


Рис. 5.2. Динаміка виробництва електроенергії за 2006-2010 роки, тис кВт-год.

3. ОСНОВНІ ПРИЙОМИ АНАЛІЗУ РЯДІВ ДИНАМІКИ

Одним із завдань статистики в процесі аналізу рядів динаміки є виявлення закономірностей зміни рівнів ряду, тобто, визначення загальної тенденції розвитку.

Найпростішим способом виявлення загальної тенденції розвитку явища є **укрупнення інтервалів часу динамічного ряду**. Суть цього прийому полягає в тому, що первинний ряд динаміки перетворюється і замінюється іншим, показники якого відносяться до більш триваліших періодів часу. Наприклад, ряд, що характеризує місячний випуск продукції, можна замінити рядом квартальних показників випуску продукції. Новостворений ряд може складатися із абсолютних величин за укрупнені періоди часу (ці величини одержують шляхом додавання рівнів первинного ряду абсолютних величин), або із середніх величин по інтервалах. При додаванні рівнів або при виведенні середніх по укрупнених інтервалах взаємозрівноважуються коливання первинного ряду, внаслідок чого тенденція розвитку вирізняється чіткіше.

Залежно від схеми формування інтервалів розрізняють ступінчасті та плинні середні. Ряди цих середніх схематично зображено на рис. 5.3 для інтервалу $m = 3$.

Очевидно, що плинна середня більш гнучка і може краще відобразити особливості тенденції розвитку явища.

Для того, щоб мати кількісну модель, яка виражає загальну тенденцію зміни рівнів динамічного ряду у часі, використовується **аналітичне вирівнювання ряду динаміки**.

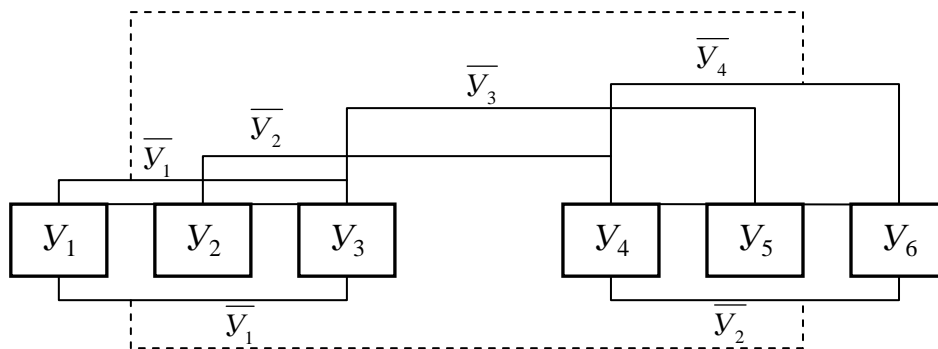


Рис.5.3. Схема утворення інтервалів згладжування рядів динаміки

При аналітичному вирівнюванні динамічного ряду фактичні значення Y_t замінюються обчисленими на основі певної функції $Y = f(t)$, яку називають **трендовим рівнянням** (t – змінна часу). Вибір типу функції ґрунтується на попередньому теоретичному аналізі суті явища, яке вивчається, і характеру його динаміки.

На практиці перевага надається функціям, параметри яких мають чіткий економічний зміст і вимірюють абсолютну чи відносну швидкість розвитку. Доцільним вважається аналіз ланцюгових характеристик інтенсивності динаміки. Якщо ланцюгові абсолютні прирости відносно стабільні, вирівнювання виконується на основі **лінійної функції**: $Y_t = a + bt$.

Якщо ж відносно стабільними є ланцюгові темпи приросту то найбільш адекватною такому характеру динаміки є **експонента** $Y_t = ab^t$. У зазначених функціях t – порядковий номер періоду (дати), a – рівень ряду при $t=0$. Параметр b характеризує швидкість динаміки: середню абсолютну в лінійній функції і середню відносну – в експоненті. Коли характеристики швидкості розвитку зростають (чи зменшуються), використовують інші функції (парабола 2-го степеня, модифікована експонента тощо).

Параметри трендових рівнянь визначають методом найменших квадратів. Згідно з умовою мінімізації суми квадратів відхилень фактичних рівнів ряду y_t від теоретичних Y_t параметри визначають розв'язуванням системи нормальних рівнянь. Для лінійно функції вона записується так:

$$\begin{cases} na + b \sum t = \sum y \\ a \sum t + b \sum t^2 = \sum yt \end{cases}$$

Система рівнянь спрощується, якщо початок відліку часу ($t = 0$) перенести в середину динамічного ряду. Тоді:

$$\begin{cases} na = \sum y \\ b \sum t^2 = \sum yt \end{cases}$$

$$\text{Отже, } a = \frac{\sum y}{n}, b = \frac{\sum yt}{\sum t^2}.$$

4. СТАТИСТИЧНЕ ВИВЧЕННЯ СЕЗОННИХ КОЛИВАНЬ

Сезонними коливаннями називають більш-менш стійкі коливання в рядах динаміки, зумовлені специфічними умовами виробництва чи споживання певного виду продукції, або іншими причинами коливань розвитку того чи іншого явища.

Сезонні коливання характеризуються спеціальним показником, який називають індексом сезонності I_s . В сукупності ці індекси утворюють **сезонну хвилю**.

Індекс сезонності – це процентне відношення фактичних рівнів рядів динаміки до середніх або вирівняних рівнів.

Аналіз сезонності розглянемо на прикладі реалізації товарів культурно-побутового призначення (табл. 5.6).

Таблиця 5.6

Розрахунок сезонної хвилі реалізації побутових холодильників торговельними підприємствами області

Квартал	Рік			Разом	У середньому \bar{y}_i	Сезонна хвиля $I_s = \frac{\bar{y}_i}{\bar{y}_{\text{заг}}} 100$
	2008	2009	2010			
I	1942	2126	2505	6573	2191,00	82,1
II	2957	2704	3704	9365	3121,67	117,0
III	2504	3291	3834	9629	3209,67	120,3
IV	2194	1745	2513	6452	2150,67	80,6
Разом	9597	9866	12556	32019	266825	×

Зобразимо сезонну хвилю на графіку (рис. 5.5).

На підставі даних табл. 5.6 і графіка (рис. 5.5) можна зробити висновок про те, що реалізація побутових холодильників суттєво зменшується в першому і четвертому кварталах і різко зростає в другому і третьому. В середньому за досліджуваний період в I кварталі холодильників продавалось на 17,9 (82,1 – 100) пунктів менше від середньо квартальної реалізації, а в IV кварталі – менше на 19,4 (80,6 – 100), тоді як в другому і третьому кварталах (літній період) обсяг реалізації

холодильників перевищує середньо квартальну реалізацію відповідно на 17,0 (117,0 – 100) пунктів і на 20,3 (120,3 – 100) пунктів.

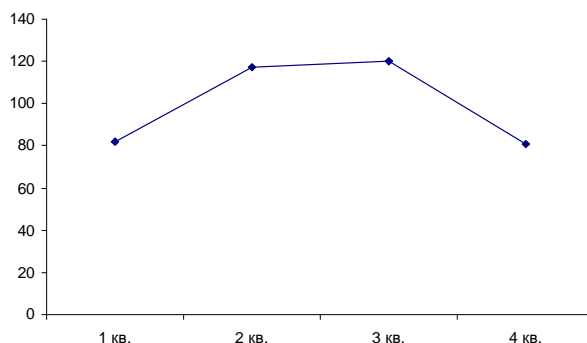


Рис. 5.5. Сезонна хвиля реалізації побутових холодильників за 2008-2010 рр.

5. КОНТРОЛЬНІ ЗАПИТАННЯ

1. Що називають рядом динаміки?
2. З яких елементів складається ряд динаміки?
3. Які Ви знаєте види рядів динаміки?
4. Які ряди динаміки називаються моментними і чому їхні рівні не можна підсумовувати?
5. У чому полягає різниця між первинними і вторинними рядами динаміки?
6. Яких умов потрібно дотримуватись при побудові рядів динаміки?
7. З яких причин виникає непорівнянність рядів динаміки?
8. У чому полягає різниця базисного і ланцюгового способів обчислення показників динаміки?
9. Як визначається абсолютний приріст, темп зростання і приросту?
10. Як обчислюється і що означає коефіцієнт випередження?
11. Як обчислюється коефіцієнт еластичності?
12. Які способи обробки та аналізу рядів динаміки Ви знаєте?
13. У чому полягає суть методу укрупнення періодів ряду динаміки?
14. В яких випадках застосовується метод приведення рядів динаміки до однієї основи?
15. Як здійснюють згладжування рядів динаміки способом плинної (ковзної) середньої?
16. У чому полягає суть методу аналітичного згладжування рядів динаміки?
17. Якими методами досліджуються сезонні коливання?

Лекція 6. ІНДЕКСНИЙ МЕТОД

План викладу і засвоєння матеріалу

1. Поняття економічних індексів, їх класифікація.
2. Індивідуальні індекси та їх взаємозв'язок.
3. Загальні індекси. Агрегатна форма загальних індексів.
4. Середньозважені індекси.
5. Індекси середніх величин.
6. Контрольні запитання для самоперевірки.

Ключові поняття та терміни

Статистичний індекс, індивідуальні і загальні індекси, ланцюгові і базисні індекси, індекс змінного складу, індекс фінансового складу, індекс структурних зрушень, система Ласпереса, система Пааше.

1. ПОНЯТТЯ ЕКОНОМІЧНИХ ІНДЕКСІВ, ЇХ КЛАСИФІКАЦІЯ

Індексом у статистиці називається відносний показник, який характеризує зміну будь-якого суспільно-економічного явища у часі, співвідношення у просторі чи порівняно з нормою, замовленням, планом, стандартом.

Насамперед індекси поділяють залежно від змісту досліджуваних явищ на індекси об'ємних (екстенсивних) показників і індекси якісних (інтенсивних) показників.

Індекси об'ємних показників характеризують зміну кількості одиниць досліджуваної сукупності, обсягу явища. До них належать індекси фізичного обсягу продукції, площі багаторічних насаджень та ін.

Індекси якісних показників відбивають зміну ознак, властивостей явища в розрахунку на одиницю сукупності. Наприклад, індекси продуктивності праці, врожайності сільськогосподарських культур тощо.

Залежно від бази порівняння індекси поділяються на:

- **планові**, що використовуються для визначення і характеристики ступеня виконання плану;
- **динамічні**, які характеризують зміни явищ у часі;
- **територіальні**, що виражають співвідношення явищ у просторі, тобто використовуються для територіальних порівнянь.

За ступенем охоплення одиниць досліджуваної сукупності (від окремих одиниць та ознак до сукупності в цілому) розрізняють індивідуальні і загальні (зведені) індекси.

За своєю формою зведені індекси поділяють на **агрегатні і середньозважені**. Вибір тієї чи іншої форми залежить від мети дослідження та наявної інформації.

Особливу групу становлять індекси середніх величин (індекси змінного та фіксованого складу, індекс структурних зрушень).

Для обчислення індексів динаміки, що характеризують зміну явищ у часі, потрібно порівняти рівні явища, що вивчаються, за два періоди. Період, з яким порівнюють, називають **базисним**, а період, який порівнюється, – **звітним** або **поточним**. Індекс обчислюють як відношення величини абсолютного показника у звітному періоді до його величини у базисному.

Показник, зміну якого характеризує індекс, називається індексованим показником, або індексованою величиною. Під час побудови індексів їх позначення супроводжується підрядковим умовним позначенням індексованої величини, тобто позначенням назви явища, зміну якого характеризує цей індекс. Отже, індексована величина визначає назву самого індексу, наприклад: індекс цін, фізичного обсягу продукції, трудомісткості, урожайності та ін.

Звітний період позначається підрядковим символом "1", базисний — символом "0" (нуль). Якщо вивчається зміна рівнів явища не за два, а за більше періодів, то кожний із періодів відповідно позначається як 0, 1, 2, 3 і тд.

Зазначимо деякі загальноприйняті умовні позначення індексованої величини:

q – кількість (обсяг) виробленої або реалізованої продукції певного виду в натуральному вираженні;

p – ціна одиниці продукції;

z або c – собівартість одиниці продукції;

t – затрати робочого часу на виробництво одиниці продукції певного виду (трудомісткість);

w – кількість продукції, виробленої за одиницю робочого часу або одним працівником (продуктивність праці);

y – урожайність окремої сільськогосподарської культури;

P або S – зібрана площа окремої сільськогосподарської культури;
 pq – вартості окремого виду продукції або товарообіг;
 zq або cq – витрати на виробництво певного виду продукції;
 $T=tq$ – затрати робочого часу на виробництво певного виду продукції;
 yP або yS – валовий збір (урожай) окремої сільськогосподарської культури тощо.

2. ІНДИВІДУАЛЬНІ ІНДЕКСИ ТА ЇХ ВЗАЄМОЗВ'ЯЗОК

Індивідуальні індекси обчислюються для характеристики зміни окремих елементів складного явища, наприклад, ціни одиниці окремого виду продукції, врожайності окремої сільськогосподарської культури.

Індивідуальні індекси позначаються малою латинською літерою "i" з підрядковим символом, що означає показник, зміна якого характеризує цей індекс. Індивідуальні індекси розраховуються шляхом зіставлення двох статистичних величин, що характеризують рівень такого явища за два порівнюваних періоди.

Якщо рівні будь-якого інтенсивного показника позначити в базисному і звітному періодах відповідно через x_0 і x_1 , а екстенсивного показника – відповідно через w_0 і w_1 то в загальному вигляді індивідуальні індекси цих показників можна записати такими формулами:

$$i_x = \frac{x_1}{x_0}; i_w = \frac{w_1}{w_0}.$$

Індекс фізичного обсягу виробництва продукції (проданого товару) в натуральному вираженні:

$$i_q = \frac{q_1}{q_0},$$

де q_0 і q_1 – кількість виробленої або реалізованої продукції певного виду в натуральному вираженні відповідно в базисному і звітному періодах.

Зауважимо, що у знаменнику розглянутого індексу може бути не лише кількість продукції, виробленої за будь-який попередній період, але і її плановий ($q_{пл}$) та нормативний (q_n) обсяги. Тоді формула індексу фізичного обсягу матиме вигляд

$$i_q = \frac{q_1}{q_{пл}}, \text{ де } i_q = \frac{q_1}{q_n}.$$

Індекс ціни одиниці продукції (товару):

$$i_p = \frac{p_1}{p_0},$$

де p_0 і p_1 – ціна одиниці продукції чи товару певного виду відповідно в базисному і звітному періодах.

Індекс собівартості одиниці продукції:

$$i_z = \frac{z_1}{z_0},$$

де z_0 і z_1 – собівартість одиниці продукції певного виду відповідно в базисному і звітному періодах.

Індекс трудомісткості (затрат робочого часу на виробництво одиниці продукції певного виду):

$$i_t = \frac{t_1}{t_0},$$

де t_0 і t_1 – трудомісткість одиниці продукції певного виду відповідно в базисному і звітному періодах.

Індекс вартості окремого виду продукції або товарообігу конкретного виду товару:

$$i_{pq} = \frac{p_1 q_1}{p_0 q_0},$$

Індекс витрат на виробництво певного виду продукції:

$$i_{zq} = \frac{z_1 q_1}{z_0 q_0}.$$

Індекс валового збору окремої сільськогосподарської культури:

$$i_{y\Pi} = \frac{y_1 \Pi_1}{y_0 \Pi_0}.$$

3. ЗАГАЛЬНІ ІНДЕКСИ. АГРЕГАТНА ФОРМА ЗАГАЛЬНИХ ІНДЕКСІВ

Загальний індекс – це співвідношення рівнів показника складного явища, до якого входять різномірні, безпосередньо несумірні елементи. Такими елементами можуть бути, наприклад, різні товари, що реалізуються, окремі види продукції, які виробляються в різних галузях народного господарства тощо. Обсяги різних видів продукції чи товарів не

підлягають порівнянню і безпосередньо їх не можна підсумовувати. Для того щоб привести різні види товарів чи продукції до порівнянного виду, необхідно обсяг кожного виду продукції чи товару в натуральному вираженні перемножити на відповідний **сумірник** (ціну, собівартість, трудомісткість). Під час множення об'ємного показника на якісний показник-сумірник кожному окремому елементу надається певна **вага**, яка відображає його значення у щойно створеному показнику. Створені таким чином показники, що є добутками, в яких хоча і поєднані різнорідні елементи (агрегати), можна підсумовувати, а отже, і порівнювати у цілому за всією сукупністю за різні періоди часу, тобто отримати загальний індекс. Такі індекси називаються **агрегатними** (від лат. aggrego — приєдную) і позначаються великою латинською літерою "I".

Агрегатним індексом називається загальний індекс, який є відношенням двох сум, кожна з яких є добутком двох різних величин. Їх і можна порівнювати.

Так, обсяг виробленої продукції або товарообігу (виручки) – це ціна, помножена на обсяг виробленої продукції або кількість проданих товарів, тобто pq :

$$I_{qp} = \frac{\sum q_1 p_1}{\sum q_0 p_0}.$$

Загальні затрати на виробництво продукції – це собівартість одиниці продукції z , помножена на кількість виробленої продукції q . Звідси формула загального індексу затрат на виробництво продукції:

$$I_{qz} = \frac{\sum q_1 z_1}{\sum q_0 z_0}.$$

Загальні затрати часу на виробництво продукції – це трудомісткість одиниці продукції t , помножена на кількість виробленої продукції q . Тому формула загального індексу затрат часу на виробництво продукції:

$$I_{qt} = \frac{\sum q_1 t_1}{\sum q_0 t_0} = \frac{\sum T_1}{\sum T_0}.$$

Залежно від того, на якому рівні – базисному чи поточному – фіксують сумірник (вагу), розрізняють дві системи побудови агрегатних індексів: **базисно-зважену**, або **систему Ласпереса** (за прізвищем німецького економіста, який у 1871 р. запропонував під час розрахунку агрегатних індексів фіксувати сумірник (вагу) на рівні базисного періоду) і **поточно-зважену систему**, або **систему Пааше** (за прізвищем німецького

економіста, який у 1874 р. запропонував для розрахунку агрегатних індексів фіксувати сумірник (вагу) на рівні поточного періоду). У загальному вигляді формули розрахунків агрегатних індексів якісних і кількісних показників за двома зазначеними системами можна записати так:

– система Ласпереса:
$$I_x = \frac{\sum x_1 \omega_0}{\sum x_0 \omega_0}; I_\omega = \frac{\sum \omega_1 x_0}{\sum \omega_0 x_0};$$

– система Пааше:
$$I_x = \frac{\sum x_1 \omega_1}{\sum x_0 \omega_1}; I_\omega = \frac{\sum \omega_1 x_1}{\sum \omega_0 x_1}.$$

де I_x і I_ω – загальні індекси відповідно якісного і кількісного показників, які характеризують зміну складного явища за рахунок зміни відповідно якісного і кількісного показників.

Значення індексів, обчислені за системами Ласпереса і Пааше, не збігаються. Це пояснюється тим, що під час їх розрахунку використовуються різні системи зваження і вони мають різний економічний зміст. Вибір системи розрахунків індексів залежить від мети дослідження та наявної інформації.

4. СЕРЕДНЬОЗВАЖЕНІ ІНДЕКСИ

Поряд з агрегатними індексами в статистиці застосовується інша їх форма – **середньозважені**. Ці індекси застосовують в тому випадку, коли наявна інформація не дає змоги обчислити загальний агрегатний індекс.

Тому агрегатна форма індексів перетворюється в інші тотожні їй форми – середньозважені індекси: арифметичні і гармонічні. Вибір тієї чи іншої форми індексу залежить від мети, з якою він визначається, та наявних вихідних даних.

Середньозважені індекси – це середні величини з індивідуальних індексів, зважені на здобутки якісного (x) і кількісного (w) показників у базисному і звітному періодах, які вибираються відповідно до системи зваження, що застосовується.

У загальному вигляді формули розрахунків середньозважених індексів кількісного і якісного показників за двома системами зваження на основі абсолютних величин мають такий вигляд:

– за системою Ласпереса:
$$I_w = \frac{\sum i_w x_0 w_0}{x_0 w_0}, I_x = \frac{\sum i_x x_0 w_0}{x_0 w_0};$$

– за системою Пааше:

$$I_w = \frac{\sum x_1 w_1}{\sum \frac{x_1 w_1}{i_w}}, \quad I_x = \frac{\sum x_1 w_1}{\sum \frac{x_1 w_1}{i_x}},$$

де i_w – індивідуальний індекс кількісного показника; i_x – індивідуальний індекс якісного показника; $x_0 w_0$ – ваги базисного періоду; $x_1 w_1$ – ваги звітнього періоду.

5. ІНДЕКСИ СЕРЕДНІХ ВЕЛИЧИН

Індекс змінного складу обчислюється як відношення середньої арифметичної зваженої звітнього періоду до середньої арифметичної зваженої базисного періоду:

$$I_{\bar{x}} = \frac{\sum x_1 f_1}{\sum f_1} : \frac{\sum x_0 f_0}{\sum f_0} = \frac{\sum x_1 \omega_1}{\sum x_0 \omega_0},$$

де x_0 і x_1 – рівні усередненого показника відповідно в базисному та звітньому періодах; f_0 і f_1 – частоти усередненого показника відповідно в базисному і звітньому періодах; w_0 і w_1 – частки усередненого показника відповідно в базисному і звітньому періодах.

Величина індексу змінного складу залежить від зміни як самого усередненого показника, так і співвідношення частот, тобто структури сукупності.

Вплив першого фактора дає змогу визначити **індекс фіксованого (постійного) складу**:

$$I_{\bar{x}} = \frac{\sum x_1 f_1}{\sum f_1} : \frac{\sum x_0 f_1}{\sum f_1} = \frac{\sum x_1 f_1}{\sum x_0 f_1} = \frac{\sum x_1 \omega_1}{\sum x_0 \omega_1}.$$

З наведених формул випливає, що індекс фіксованого складу можна виразити як агрегатний індекс якісного показника.

Вплив другого фактора дає можливість отримати **індекс структурних зрушень**:

$$I_{\omega} = \frac{\sum x_0 f_1}{\sum f_1} : \frac{\sum x_0 f_0}{\sum f_0} = \frac{\sum x_0 \omega_0}{\sum x_0 \omega_1}.$$

У цьому індексі змінюються лише ваги f_0 і f_1 , тому цей індекс характеризує вплив структурних зрушень. Між наведеними індексами є такий взаємозв'язок:

$$I_{\bar{x}} = I_x \cdot I_d.$$

Узагальнені формули індексів середніх величин конкретизуються для будь-якого якісного показника. Наприклад, система індексів середніх величин для вивчення динаміки середньої ціни виглядає так:

$$I_{\bar{p}} = \frac{\bar{p}_1}{\bar{p}_0} = \frac{\sum p_1 q_1}{\sum q_1} : \frac{\sum p_0 q_0}{\sum q_0} - \text{індекс середньої ціни змінного складу};$$

$$I_p = \frac{\sum p_1 q_1}{\sum q_1} : \frac{\sum p_0 q_1}{\sum q_1} = \frac{\sum p_1 q_1}{\sum p_0 q_1} - \text{індекс середньої ціни фіксованого складу};$$

$$I_w = \frac{\sum p_0 q_1}{\sum q_1} : \frac{\sum p_0 q_0}{\sum q_0} - \text{індекс структурних зрушень},$$

де p_0 і p_1 – ціна окремого виду товару відповідно в базисному і звітному періодах; q_0 і q_1 – відповідно їх кількості.

6. КОНТРОЛЬНІ ЗАПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Що таке економічні індекси і для чого їх застосовують?
2. Які бувають економічні індекси залежно від сутності соціально-економічних явищ? Наведіть приклади.
3. Як поділяють економічні індекси залежно від бази порівняння? Наведіть приклади таких індексів.
4. Які економічні індекси називають індивідуальними, а які — загальними (зведеними)? Наведіть приклади.
5. Які бувають економічні індекси залежно від методології їх побудови? Наведіть приклади.
6. Що таке базисний і звітний періоди і як їх позначають?
7. Що характеризують індивідуальні індекси і який між ними взаємозв'язок?
8. У чому сутність і призначення загальних індексів? Наведіть приклади.
9. Назвіть основні принципи побудови загальних індексів агрегатної форми.
10. Які є системи зваження під час розрахунку індексів агрегатної форми?
11. Як визначається агрегатний індекс фізичного обсягу, і що він характеризує? Напишіть його формулу і зазначте індексовану величину та ваги.
12. Яким чином визначається агрегатний індекс цін, і що він характеризує? Напишіть його формулу і зазначте індексовану величину та ваги.
13. Вкажіть, як визначається агрегатний індекс продуктивності праці і яка його залежність від індексу трудомісткості? Напишіть формули цих індексів і покажіть їх взаємозв'язок.
14. Які індекси називаються спряженими? Наведіть приклади.

15. За якими формулами визначаються агрегатні індекси собівартості, фізичного обсягу і затрат на виробництво? Покажіть їх взаємозв'язок.
16. Як визначається абсолютний вплив факторів-співмножників, що утворюють індекси?
17. У яких випадках застосовуються середньозважені індекси? Доведіть, що середньозважені індекси тотожні відповідним індексам агрегатної форми.
18. За якими формулами визначають середньозважений арифметичний індекс за абсолютними і відносними даними?
19. За якими формулами визначають середньозважений гармонічний індекс за абсолютними і відносними даними?
20. Що таке індекси середніх величин і в яких випадках їх використовують?
21. Як визначають і аналізують індекс змінного складу?
22. Яким чином визначається індекс фіксованого складу, і що він характеризує? Наведіть його формулу.
23. Як визначається індекс структурних зрушень, і що він характеризує? Наведіть його формулу.

Список літератури

1. Бек В.Л. Теорія статистики: навч. посіб. / В.Л. Бек. – К.: ЦУЛ, 2003. – 288 с.
2. Єріна А.М. Теорія статистики: практикум / А.М. Єріна, З.О. Пальян. – К.: Знання, 2005. – 255 с.
3. Захожай В.Б. Статистика якості: підр. для студ. вищ. навч. зал. / В.Б. Захожай, А.Ю. Чорний. – К.: МАУП, 2005. – 576 с.
4. Кулинич О.І. Теорія статистики: підручник / О.І. Кулинич, Р.О. Кулинич. – К.: Знання, 2006. – 294 с.
5. Лучінін О.Є. Статистика: підручник / О.Є. Лучінін. – К.: ЦНЛ, 2007. – 606 с.
6. Мармоза А.Т. Теорія статистики: навч. посіб. / А.Т. Мармоза. – К.: Ельга: Ніка-Центр, 2003. – 392 с.
7. Моторин Р.М. Статистика для економістів: навч. посіб. / Р.М. Моторин, Е.В. Чекотовський. – К.: Знання, 2009. – 430 с.
8. Статистика: навч. посібник для самост. вивч. дисц. / за ред. Р.М. Моторина, А.М. Єріної. – К.: КНЕУ, 2002. – 448 с.
9. Статистика : підручник / С.С. Герасименко та ін. – К.: КНЕУ, 2000. – 467 с.
10. Статистичний щорічник України за 2009р. – К.: Консультант, 2010.

Навчальне видання

ФЕДОРЕНКО Наталія Дмитрівна
ДОЛЯ Олена Вікторівна
БІЛОЩИЦЬКА Світлана Василівна
БІЛОЩИЦЬКИЙ Андрій Олександрович
ТЕРЕНЧУК Світлана Анатоліївна

СТАТИСТИКА

Конспект лекцій

для студентів які навчаються за напрямом підготовки 6.030510
«Товарознавство і торговельне підприємництво»

Комп'ютерне верстання *О.В. Кириченка*

Підписано до друку Формат 60x84 ^{1/16}.
Ум. друк. арк. 3,25. Обл.-вид. арк. 3,5.
Тираж 50 прим. Вид. № 28/І-10. Зам. №
КНУБА, Повітрофлотський проспект, 31, Київ, Україна, 03680

E-mail: red-isdat@knuba.edu.ua

Віддруковано в редакційно-видавничому відділі
Київського національного університету будівництва і архітектури

Свідоцтво про внесення до Державного реєстру суб'єктів видавничої справи
ДК № 808 від 13.02.2002 р.

