МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний університет «Запорізька політехніка»

Конспект лекцій
з дисципліни «Статистика»
для здобувачів вищої освіти за освітнім ступенем «бакалавр» денної та заочної форм навчання спеціальностей 072 «Фінанси, банківська справа та страхування», 071 «Облік та оподаткування»
(Частина 2)

2021

Конспект лекцій з дисципліни «Статистика» для здобувачів вищої освіти за освітнім ступенем «бакалавр» денної та заочної форм навчання спеціальностей 072 «Фінанси, банківська справа та страхування», 071 «Облік та оподаткування» (Частина 2) / Укл. Н.Г. Фатюха – Запоріжжя: НУ«ЗП», 2021.- 59с.

Укладач: к.е.н., доцент Н.Г. Фатюха

Рецензент: к.е.н, доцент, зав. кафедри Економічна теорія та підприємництво Губарь О. В.

Відповідальний за випуск: к.е.н., доцент С.В. Шарова

Затверджено
на засіданні кафедри
“Фінанси, банківська справа
та страхування ”
Протокол № 10
від 16.04. 2021 р.

Рекомендовано
до видання НМК ФЕУ
Протокол № 30
від 19.04. 2021 р.

ЗМІСТ

Частина 2
Тема 5. Аналіз рядів розподілу 58
5.1. Частотні характеристики рядів розподілу 58
5.2. Характеристики центру розподілу 58
5.3. Характеристики варіації 60
Тема 6. Аналіз концентрації, диференціації та подібності розподілів 63
6.1. Види та взаємозв'язок дисперсій 63
6.2. Характеристики форми розподілу 65
6.3. Показники диференціації 67
6.4. Показники концентрації 68
Тема 7. Статистичні методи вимірювання взаємозв'язків 69
7.1. Види взаємозв'язків 70
7.2. Рівняння регресії 71
7.3. Визначення щільності зв'язку 72
7.4. Рангова кореляція 75
Тема 8. Аналіз інтенсивності динаміки 76
8.1. Ряди динаміки, їх суть і види 76
8.2. Показники аналізу рядів динаміки 78
8.3. Середні показники динаміки 80
Тема 9. Аналіз тенденцій розвитку та коливань 83
9.1. Характеристика основної тенденції розвитку 83
9.2. Вивчення сезонних коливань 87
Тема 10. Індексний метод 88
10.1. Суть і види індексів 89
10.2. Агрегатні індекси, взаємозв'язок індексів 91
10.3. Середньозважені індекси 94
10.4. Індекси середніх величин 94
Тема 11. Вибірковий метод 96
11.1. Суть вибіркового спостереження 96
11.2. Обчислення помилок вибірки 98
11.3. Різновиди вибірок 100
11.4. Визначення обсягу вибірки 102

Тема 12. Подання статистичних даних: таблиці, графіки, карти 103
12.1. Статистичні таблиці 103
12.2. Графічне зображення абсолютних та відносних величин 107
Список використаних джерел 112

ТЕМА 5 АНАЛІЗ РЯДІВ РОЗПОДІЛУ

План лекції

5.1. Частотні характеристики рядів розподілу
5.2. Характеристики центру розподілу
5.3. Характеристики варіації

5.1. Частотні характеристики рядів розподілу

Характерні риси та особливості структури статистичної сукупності відображаються в рядах розподілу (атрибутивних та варіаційних). Частотними характеристиками будь-якого ряду є: частота (f); частка (d).

Додатковою характеристикою варіаційних рядів є кумулятивна (нагромаджена) частота (частка), яка є результатом послідовного об'єднання груп і підсумовування відповідних їм частот (часток).

Якщо інтервали варіаційного ряду нерівні, то використовують **щільність частоти (частки)** на одиницю інтервалу, яка розраховується за формулами:

$$q = \frac{f}{h} \quad або \quad q = \frac{d}{h}, \tag{5.1}$$

де h - ширина інтервалу.

5.2. Характеристики центру розподілу

До характеристик центру розподілу відносять середню, моду та медіану.

Середня величина характеризує типовий рівень ознаки в сукупності.

Мода (Мо) - найпоширеніше значення ознаки, тобто варіанта, яка в ряду розподілу має найбільшу частоту (частку).

Моду використовують у комерційній діяльності, коли вивчають ринок попиту, у соціальних дослідженнях - встановлюється рейтинг популярності осіб чи товарів.

У **дискретному ряду** Мо визначається візуально за максимальною частотою або часткою.

В **інтервальному ряду** за найбільшою частотою визначається **модальний інтервал**. Конкретне значення моди в інтервалі обчислюється за формулою:

$$M_0 = x_0 + h\frac{fm_0 - fm_{0-1}}{(fm_0 - fm_{0-1}) + (fm_0 - fm_{0+1})}, \qquad (5.2)$$

де x_0- нижня межа модального інтервалу;
h – ширина модального інтервалу;
fm_0- частота (частка) модального інтервалу,
f_{m0-1} - частота (частка)передмодального інтервалу;.
f_{m0+1} – частота (частка) післямодального інтервалу.

Медіана (Ме) – варіант, що є серединою впорядкованого варіаційного ряду, тобто ділить його на дві рівні частини: одна частина має значення варіаційної ознаки менше, ніж середня, а друга – більше.

Медіана вказує на значення варіаційної ознаки, якого досягла половина одиниць сукупності.

Медіану в дискретному варіаційному ряді визначають за сумою всіх частот, яку треба поділити на дві і до отриманого результату додати 0,5.

$$N_{Me} = \frac{\sum f + 1}{2}. \qquad (5.3)$$

Якщо непарне число варіант записати в порядку зростання чи зменшення, то центральна з них і буде медіаною. Коли число варіант парне- медіана розраховується як середня арифметична двох централь ($\overline{X} = \sum X / n$)

У дискретному ряду медіаною буде значення ознаки, для якої кумулятивна частота Sf перевищує половину обсягу сукупності $\sum f / 2$ або кумулятивна частка Sd перевищує $\sum d / 2$ (0,5 або 50%).

Медіана не залежить ні від амплітуди коливань ряду, ні від розподілення частот у межах двох рівних частин ряду. Її обчислюють

для вирішення окремих завдань, пов’язаних із визначенням оптимуму, який співпадає з варіантою, що припадає на середину ряду.

В інтервальному ряду визначається медіанний інтервал і значення медіани в інтервалі за формулою:

$$Me = x_{ме} + i\frac{\sum \frac{1}{2} f - Sf_{Me-1}}{f_{Me}}, \qquad (5.4)$$

де і – розмір медіанного інтервалу;

Sf_{me-1} - кумулятивна частота передмедіанного інтервалу;

f_{me} - частота медіанного інтервалу.

5.3. Характеристики варіації

Коливність окремих значень варіантів характеризують показники варіації.

Варіація, тобто коливання будь-якої ознаки є властивістю статистичної сукупності. Чим менша варіація, тим однорідніша сукупність, отже, тим більш надійні і типові середні величини. Для виміру і оцінки варіації використовують абсолютні та відносні характеристики.

Характеристики варіації обчислюються для глибокого та всебічного аналізу сутності досліджуваних явищ.

Наприклад, для оцінювання ритмічності роботи підприємств, фінансового ризику, диференціації доходів населення.

Абсолютні характеристики варіації.

Варіаційний розмах - різниця між максимальним і мінімальним значеннями ознаки:

$$R = X_{\max} - X_{\min}. \qquad (5.5)$$

недоліком цього показника є те, що він фіксує лише крайні відхилення і зовсім не враховує відхилень решти варіантів від їхньої середньої.

Тому цей показник відносно нечасто використовують у практичній діяльності.

Якщо крайні значення ознаки не типові для сукупності, то використовують квартальні або децільні розмахи.

Середнє лінійне відхилення:

$$\bar{l} = \frac{\sum |x - \bar{x}| f}{\sum f}$$ - зважене, якщо дані згруповані; (5.6)

$$\bar{l} = \frac{\sum |x - \bar{x}|}{n}$$ - просте, якщо дані не згруповані. (5.7)

де x- рівень ознаки;

n-число варіантів.

Середнє лінійне відхилення не завжди характеризує розсів варіантів.

Ступінь варіації об’єктивніше відображує показник середнього квадрата відхилення (дисперсія).

Дисперсія (середній квадрат відхилення):

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f}$$ - зважена; (5.8)

$$\sigma^2 = \frac{\sum (x - \bar{x})}{n}$$ - проста. (5.9)

Середнє квадратичне (стандартне) відхилення:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}}$$ - зважене; (5.10)

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$ - просте. (5.11)

Тобто, середнє квадратичне відхилення розраховується

шляхом добування квадратного кореня з дисперсії.

Відносні характеристики.

При порівнянні варіації різних ознак або однієї ознаки в різних сукупностях використовуються коефіцієнти варіації V:

лінійний $$V_{\bar{l}} = \frac{\bar{l}}{\bar{x}} \cdot 100; \qquad (5.12)$$

характеризує частку середнього значення абсолютних відхилень від середньої величини:

квадратичний $$V_{\sigma} = \frac{\sigma}{\bar{x}} \cdot 100; \qquad (5.13)$$

коливання індивідуальних значень відносно середнього в цілому по сукупності, і відповідно типовість середньої величини.

осциляції $$V_R = \frac{R}{\bar{x}} \cdot 100. \qquad (5.14)$$

відображає відносну коливність **крайніх значень ознаки навколо середньої.**

Коефіцієнт варіації є певною мірою критерієм типовості середньої. Чим менша варіація, тим менше значення цих характеристик.

Для порівняння варіацій найчастіше використовують квадратичний коефіцієнт варіації. Вважають, що сукупність є однорідною, а середня - типовою, коли цей коефіцієнт не перевищує 33% (0,33).

Універсальним показником варіації є коефіцієнт варіації, цінність якого полягає в тому, що ним можна користуватись для характеристики і порівняння варіації різних сукупностей і різних явищ.

Питання для самоконтролю

1. Що таке мода та медіана?
2. Як визначають моду в дискретному та інтервальному рядах

розподілу?

3. Назвіть особливості визначення медіани в дискретному та інтервальному рядах розподілу?

4. В яких випадках застосовуються структурні середні?

5. Що розуміють під варіацією ознаки?

6. Які показники використовують для вимірювання варіації?

7. Як визначити амплітуду коливання ознаки?

8. Як розраховується середнє лінійне відхилення?

9. Що таке дисперсія і яке місце вона посідає у системі показників варіації?

10. Як визначити коефіцієнт варіації і яке його практичне значення?

Література [4 с.114-118, 125-134, 5 с.101-110].

ТЕМА 6. АНАЛІЗ КОНЦЕНТРАЦІЇ, ДИФЕРЕНЦІАЦІЇ ТА ПОДІБНОСТІ РОЗПОДІЛІВ

План лекції

6.1. Види та взаємозв'язок дисперсій
6.2. Характеристики форми розподілу
6.3. Показники диференціації
6.4. Показники концентрації

6.1. Види та взаємозв'язок дисперсій

Дисперсія посідає особливе місце у статистичному аналізі соціально-економічних явищ.

Для ознак метричної шкали дисперсія - середній квадрат відхилень індивідуальних значень ознаки від середньої:

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} \qquad (6.1)$$

як і будь-яка середня дисперсія має певні математичні властивості:

1. Якщо всі значення варіант *x* зменшити на сталу величину *C*, то дисперсія не зміниться.

2. Якщо всі значення варіант *x* змінити в *C* разів, то дисперсія зміниться в C^2 раз: $\sigma^2 = \sigma^2 C^2$.

3. Якщо частоти замінити частками, дисперсія не зміниться.

Якщо сукупність розбито на групи за певною ознакою *x*, то для будь-якої іншої ознаки *y* можна обчислити дисперсію як у цілому по сукупності, так і в кожній групі та дисперсію між групами (міжгрупову).

Для визначення впливу постійного фактора на розмір варіації потрібно розбити всю сукупність на групи та знайти, як змінюється результат під дією чинника, покладеного в основу групування.

Для цього попередньо необхідно обчислити для кожної групи середню величину ознаки, групові (часткові) дисперсії, середню з групових та між групову дисперсію.

Групова дисперсія дорівнює середньому квадрату відхилень окремих значень ознаки всередині групи від середньої арифметичної відповідної групи.

$$\sigma_i^{\,2} = \frac{\sum (x - \bar{x}_i)^2 f}{\sum f} \qquad (6.2)$$

$$\sigma_i^{\,2} = \frac{\sum (x - \bar{x}_i)^2}{\sum n} \qquad (6.3)$$

розраховується окремо для кожної групи.

або спрощеним способом: $\sigma^2{}_i = \overline{x^2} - \bar{x}^2$ (6.4)

Ця дисперсія відображує варіацією ознаки лише за рахунок умов і причин, що діють всередині групи.

Середня дисперсія з групових- середня арифметична зважена з групових дисперсій:

$$\bar{\sigma}_i^{\,2} = \frac{\sum \sigma^2{}_i f}{\sum f}. \qquad (6.5)$$

Міжгрупова дисперсія дорівнює середньому квадрату відхилень групових середніх $\bar{x}_i$ від загальної середньої $\bar{x}$.

$$\delta^2 = \frac{\sum (\bar{x}_i - \bar{x})^2 f_i}{\sum f_i}, \qquad (6.6)$$

де $\bar{x}_i$ - середня і-ї групи,

$\bar{x}$ - загальна середня,

fi - частота і-ї групи.

Характеризує варіацію ознаки *y* за рахунок впливу всіх причин (факторів).

Загальна дисперсія складається з двох частин: сумі середньої з групових дисперсій та між групової дисперсії:

$$\sigma^2 = \bar{\sigma^2}_i + \delta^2 \qquad (6.7)$$

- **це правило складання дисперсій.**

Загальну дисперсію можна визначити безпосередньо за формулою:

$$\sigma^2 = \bar{x^2} - \bar{x}^2 \qquad (6.8)$$

6.2. Характеристики форми розподілу

Аналіз закономірностей розподілу передбачає оцінку ступеня однорідності сукупності, асиметрії та ексцесу розподілу.

Однорідними вважаються такі сукупності, елементи яких мають спільні риси і належать до одного типу, класу. Критерієм однорідності сукупності вважається квадратичний коефіцієнт варіації (V_σ < 33%).

Асиметрія як відносна статистична характеристика дорівнює різниці між середнім значенням і медіаною або модою, поділеними на середнє квадратичне відхилення.

Стандартизовані відхилення:

$$A_1 = \frac{\bar{x} - M_0}{\sigma} \; або \; A_2 = \frac{\bar{x} - M_e}{\sigma}. \qquad (6.9)$$

характеризують напрям і міру скошеності розподілу. Очевидно, що в симетричному розподілі А=0, при правосторонній асиметрії А більше 0, при лівосторонній $A\langle 0$

Чим більша асиметрія, тим більше відхилення $\bar{x}$ – Мо.

У симетричних та помірно асиметричних розподілах вимірюється ексцес розподілу.

Асиметрія та ексцес - дві пов'язані з варіацією властивості форми розподілу. Комплексна їх оцінка виконується на основі центральних моментів розподілу.

Момент 2-го порядку є дисперсією, яка характеризує варіацію:

$$M_2 = \frac{\sum (x - \bar{x})^2 f}{\sum f}. \qquad (6.10)$$

Момент 3-го порядку характеризує асиметрію:

$$M_3 = \frac{\sum (x - \bar{x})^3 f}{\sum f}. \qquad (6.11)$$

тоді коефіцієнт асиметрії:

$$A = \frac{M_3}{\sigma^3}. \qquad (6.12)$$

При правосторонній асиметрії А > 0, при лівосторонній - А < 0. Вважається, що при А < 0,25 асиметрія низька, якщо А не перевищує 0,5 - середня, при А > 0,5 - висока.

Момент 4-го порядку характеризує ексцес:

$$M_4 = \frac{\sum (x - \bar{x})^4 f}{\sum f}. \qquad (6.13)$$

тоді коефіцієнт ексцесу:

$$E = \frac{M_4}{\sigma^4}. \qquad (6.14)$$

У симетричному, близькому до нормального, розподілі E = 3. При гостровершинному розподілі E > 3, при плосковершинному E < 3. Якщо E > 3 - позитивний розподіл, якщо E < 3 - негативний розподіл.

6.3. Показники диференціації

Якщо виникає необхідність вивчити структуру варіаційного ряду більш детально, розраховують значення ознаки, аналогічні медіані. Такі значення ознаки, які ділять всі одиниці розподілу на рівні чисельності, називаються квантилями або дицилями.

Вони визначаються на основі кумулятивних частот (часток) за аналогією з медіаною. В аналізі закономірностей розподілу використовуються також інші порядкові характеристики: квартилі та децилі.

Квартилі Q - варіанти, які поділяють обсяги сукупності на чотири рівні частини.

Децилі D - на десять рівних частин. Вони визначаються на основі кумулятивних частот (часток) за аналогією з медіаною:

$$Q = x_0 + h\frac{0{,}25\sum f - S_{Q-1}}{f_Q}, \qquad (6.15)$$

$$D = x_0 + h\frac{0{,}25\sum f - S_{D-1}}{f_D} \qquad (6.16)$$

Квартилі: 0,25; 0,5; 0,75. Децилі: 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; 0,9. Другий квартиль та п'ятий дециль збігаються з медіаною.

6.4. Показники концентрації

Оцінка нерівності розподілу значень ознаки між окремими складовими сукупності ґрунтується на порівнянні часток двох розподілів:

- за кількістю елементів сукупності d;
- за обсягом значень ознаки D.

Відхилення часток свідчить про певну нерівномірність розподілу, яка вимірюється коефіцієнтами:

- локалізації:

$$L = \frac{D}{d} \tag{6.17}$$

- концентрації:

$$K = \frac{1}{2}\sum\left|d - D\right| \quad \text{або} \quad K = \frac{1}{2}\sum\left|xd_i - d_i\right| \tag{6.18}$$

Коефіцієнт локалізації розраховується для кожної складової сукупності, L > 1 свідчить про концентрацію значень ознаки в окремій складовій сукупності.

Коефіцієнт концентрації є узагальнюючою характеристикою відхилення розподілу від рівномірного. Значення його коливаються в межах від 0 до 1. Чим помітніша концентрація, тим більше значення К відхиляється від 0.

Коефіцієнти концентрації та локалізації є ефективним засобом вимірювання диференціації сукупності за даними інтервальних рядів з нерівними інтервалами та за даними атрибутивних рядів.

Для оцінки інтенсивності структурних зрушень у часі використовують абсолютні міри варіації - **коефіцієнти структурних зрушень:**

$$\text{лінійний: } l_d = \frac{\sum\left|d_1 - d_2\right|}{n}, \tag{6.19}$$

$$\text{квадратичний: } \sigma_d = \sqrt{\frac{\sum(d_1 - d_2)}{n}}. \tag{6.20}$$

де d_1, d_2 - частки розподілу за два періоди;

n - число складових сукупності.

Коефіцієнт подібності (схожості) структур двох сукупностей:

$$P = 1 - \frac{1}{2}\sum_{1}^{m}\left|d_j - d_r\right|, \qquad (6.21)$$

де d_j, d_r - частка розподілу елементів сукупності.

Якщо структури однакові, P=1.

Чим більше відхилення структур, тим менше значення P.

Питання для самоконтролю

1. Що таке дисперсія?
2. Які математичні властивості притаманні дисперсії?
3. У чому полягає суть способу моментів для визначення дисперсії?
4. В яких випадках застосовується середня гармонічна і як вона обчислюється?
5. Які розрізняють види дисперсій?
6. Чим відрізняються внутрішньогрупова та міжгрупова дисперсії від загальної дисперсії?
7. Як і в якій послідовності визначаються різні види дисперсій?
8. У чому полягає суть правила додавання дисперсій?
9. Які розподіли називають симетричними, асиметричними?
10. Якими показниками характеризують форми розподілу?
11. Які моменти розподілу використовуються у статистичному дослідженні?
12. Що таке ексцес?
13. Який показник є мірою концентрації і як він обчислюється?

Література [4 с.134-144, 5 с.106-141].

ТЕМА 7. СТАТИСТИЧНІ МЕТОДИ ВИМІРЮВАННЯ ВЗАЄМОЗВ'ЯЗКІВ

План лекції

7.1. Види взаємозв'язків

7.2. Рівняння регресії

7.3. Визначення щільності зв'язку
7.4. Рангова кореляція

7.1. Види взаємозв'язків

Усі соціально-економічні явища взаємопов'язані та взаємообумовлені. Зв'язок між ними має причинно-наслідковий характер. Ознаки, що характеризують причини та умови зв'язку, називаються факторними.

Ознаки, що характеризують наслідки зв'язку, називаються результативними.

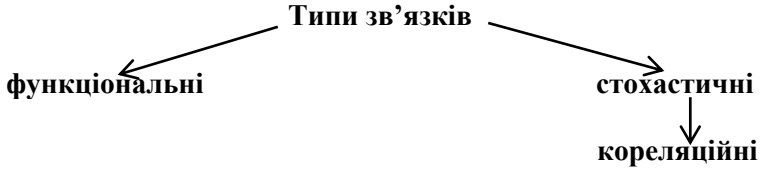

Рисунок 7.1 – Типи зв’язків

При функціональному зв'язку кожному значенню факторної ознаки *x* відповідає одне або кілька чітко визначених значень результативної ознаки *y*.

На відміну від функціональних, стохастичні зв'язки неоднозначні. При стохастичному зв'язку кожному значенню ознаки *x* відповідає певна множина значень *y,* які утворюють так званий **умовний розподіл**.

Якщо умовні розподіли замінюються одним параметром - середньою, то такий зв'язок називають кореляційним. Кореляційний зв'язок є різновидом стохастичного і виявляється у зміні середніх, умовних розподілів.

Наявність стохастичного зв'язку можна виявити на основі комбінаційного розподілу елементів сукупності.

За напрямком розрізняють прямі і обернені зв'язки.

Прямий зв'язок - зв'язок, коли із зростанням факторної ознаки *x* зростає результативна ознака *y*.

Обернений зв'язок - зв'язок, якщо зі зростанням факторної ознаки результативна зменшується або навпаки, із зменшенням

факторної ознаки результативна зростає.

7.2. Рівняння регресії

Залежно від форми зв'язку кореляційні залежності бувають прямолінійними і криволінійними.

При прямолінійній залежності рівним змінам середніх значень факторної ознаки відповідають рівні зміни середніх значень результативної ознаки.

При криволінійній залежності рівним змінам середніх значень факторної ознаки відповідають нерівні зміни середніх значень результативної ознаки.

Залежно від характеру зв'язку в економічних дослідженнях використовують лінійні і нелінійні рівняння.

Лінійне рівняння має таку формулу:

$$Y = a + bx, \tag{7.1}$$

де Y - значення результативної ознаки;

a - вільний член рівняння, реального економічного змісту не має;

b - коефіцієнт регресії, який показує на скільки одиниць в середньому зміниться результативна ознака при зміні факторної ознаки на одиницю;

x - значення факторної ознаки.

Параметри *a* і *b* такого рівняння визначаються методом найменших квадратів шляхом розв'язання системи нормальних рівнянь:

$$\begin{cases} \sum y = na + b\sum x, \\ \sum xy = a\sum x + b\sum x^2. \end{cases} \tag{7.2}$$

Розв'язавши систему, отримуємо:

$$a = \frac{\sum x^2 \sum y - \sum x \sum xy}{n\sum x^2 - \sum x \sum x}, \tag{7.3}$$

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \sum x \sum x}. \qquad (7.4)$$

Якщо b має позитивний знак, то зв'язок прямий, якщо від'ємний - зв'язок обернений.

Для економічної інтерпретації лінійних і нелінійних зв'язків між двома явищами використовується **коефіцієнт еластичності**, який показує, на скільки процентів у середньому змінюється результативна ознака зі зміною факторної ознаки на 1%.

Коефіцієнт еластичності розраховується за формулою:

$$E = b\frac{\bar{x}}{\bar{y}}. \qquad (7.5)$$

7.3. Визначення щільності зв'язку

Для вимірювання щільності зв'язку і визначення його напрямку при лінійній залежності використовують лінійний коефіцієнт кореляції, який визначається за формулою:

$$r = \frac{\overline{xy} - \bar{x} \times \bar{y}}{\sigma_x \times \sigma_y} \text{ або } r = b\frac{\sigma_x}{\sigma_y}. \qquad (7.6)$$

Значення r коливається в межах від -1 до +1. Якщо лінійний коефіцієнт кореляції із знаком «+» - зв'язок між ознаками прямий, а якщо із знаком «-» - зв'язок зворотній.

Мірою щільності кореляційного зв'язку є коефіцієнт детермінації, кий розраховується як відношення факторної дисперсії до загальної за формулою:

$$R^2 = \frac{\sigma_Y^2}{\sigma_y^2}, \qquad (7.7)$$

Корінь квадратний з коефіцієнта детермінації називають **індексом кореляції**, тобто:

$$R = \sqrt{R^2}. \qquad (7.8)$$

Індекс кореляції показує тісну залежність витрат від віку устаткування.

Мірою щільності зв'язку є також **кореляційне відношення**

$$\eta^2 = \frac{\delta^2}{\sigma^2}, \qquad (7.9)$$

де δ^2 - міжгрупова дисперсія;

σ^2 - загальна дисперсія.

Перевірка істотності кореляційного зв'язку ґрунтується на порівнянні фактичних значень R^2 і η^2 з критичними, які могли б виникнути за відсутності зв'язку.

Якщо фактичне значення R^2 чи η^2 перевищує критичне, то зв'язок між ознаками не випадковий.

Критичні значення R^2 і η^2 для рівня істотності a = 0,05 і відповідного числа ступенів свободи k_1 і k_2 наведено в таблиці.

Таблиця 7.1 - Критичні значення коефіцієнта детермінації R^2 і кореляційного відношення η^2 для рівня істотності a = 0,05

k_1 / k_2	1	2	3	4	5
5	0,569	699	764	806	835
6	500	632	704	751	785
7	444	575	651	702	739
8	399	527	604	657	697
9	362	488	563	618	659
10	332	451	527	582	624
12	283	394	466	521	564
14	247	348	417	471	514
16	219	312	378	429	477
18	197	283	345	39 4	435
20	179	259	318	364	404

Продовження таблиці 7.1.

k1 / k2	1	2	3	4	5

24	151	221	273	316	353
28	130	193	240	279	314
32	115	171	214	250	282
36	102	153	192	226	256
40	093	139	176	207	234
50	075	113	143	170	194
60	063	095	121	144	165
80	047	072	093	110	127
100	038	058	075	090	103
120	032	049	063	075	087
200	019	030	038	046	053

$$k_1 = m - 1, \; k_2 = n - m, \tag{7.10}$$

де n - обсяг сукупності;

m - число груп або параметрів функції.

Істотність зв'язку R^2 або η^2 перевіряють також за допомогою таблиці критерію F для 5%-ного рівня значимості.

Таблиця 7.2- Таблиця критерію F для 5%-го рівня значимості (ймовірність «нульової гіпотези» 0,05)

k_2 - ступені свободи для внутрішньо-групової дисперсії	k_1 - ступені свободи для міжгрупової дисперсії								
	1	2	3	4	5	6	7	8	9
2	18,51	19	19,16	19,25	19,3	19,33	19,26	13,97	19,38
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6
6	5,99	5,14	4,76	4,53	4,31	4,28	4,21	4,15	4,1
8	5,32	4,46	4,07	3,84	3,69	3,58	3,5	3,44	3,39
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18

Продовження таблиці 7.2

k_2 - ступені свободи для внутрішньо-групової дисперсії	k_1 - ступені свободи для міжгрупової дисперсії								
	1	2	3	4	5	6	7	8	9
10	4,96	4,1	3,71	3,48	3,33	3,32	3,14	3,0/	3,02
12	4,75	3,88	3,49	3,26	3,11	*a*	2,92	2,85	2,8
14	4,6	3,74	3,34	3,11	2,96	2,85	2,77	2,7	2,65
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46
20	4,35	3,49	3,1	2,87	2,71	2,6	2,52	2,45	2,4
24	4,26	3,4	3,01	2,78	2,62	2,51	2,42	2,3Ь	2,3
25	4,24	3,38	2,99	2,76	2,6	2,49	2,41	2,34	2,28
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,3	2,25
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12
50	4,03	3,18	2,79	2,56	2,4	2,29	2,2	2,13	2,07
60	4	3,15	2,76	2,52	2,37	2,25	2,1/	2,1	2,04
80	3,96	3,11	2,72	2,48	2,33	2,21	2,12	2,05	1,99
100	3,94	3,09	2,7	2,46	2,3	2,19	2,1	2,03	1,97
200	3,89	3,04	2,65	2,41	2,26	2,14	2,05	1,98	1,92

F(0,95) = 5,99 значно менше від фактичного. Це також підтверджує істотність кореляційного зв'язку.

7.4. Рангова кореляція

Для виявлення та вимірювання зв'язків використовують методи рангової кореляції. Рангами називають числа натурального ряду, які згідно зі значеннями ознаки надаються елементам сукупності і певним чином упорядкують її. Ранжування проводиться за кожною ознакою окремо: перший ранг надається найменшому значенню ознаки, останній - найбільшому або навпаки. Найбільш поширений **коефіцієнт рангової кореляції Спірмена** за формулою:

$$P = 1 - \frac{6\sum d^2}{n\left(n^2 - 1\right)}, \qquad (7.11)$$

де d - відхилення рангів факторної ознаки R_x і результативної R_y $(d_i = R_{xi} - R_{yi})$;

n - кількість рангів.

Коефіцієнт рангової кореляції змінюється в межах від -1 до +1, тобто одночасно оцінює щільність зв'язку та його напрям.

Питання для самоконтролю

1. Чим відрізняються стохастичні зв'язки від функціональних?
2. В чому суть кореляційного зв'язку?
3. Що показує коефіцієнт еластичності?
4. За даними аналітичного групування 20 комерційних банків загальна дисперсія прибутковості активів - 16, міжгрупова - 10, кількість груп - 4. Визначте кореляційне відношення, поясніть його зміст та перевірте істотність зв'язку з імовірністю 0,954.
5. Що характеризує коефіцієнт регресії?

Література [4 с.247-270, 5 с.141-186].

ТЕМА 8. АНАЛІЗ ІНТЕНСИВНОСТІ ДИНАМІКИ

План лекції

8.1. Ряди динаміки, їх суть і види
8.2. Показники аналізу рядів динаміки
8.3. Середні показники динаміки

8.1. Ряди динаміки, їх суть і види

Динаміка - це поступальний розвиток, зміна явищ у часі.

Ряд динаміки - ряд статистичних показників, які розміщені у хронологічній послідовності.

Складовими елементами ряду динаміки є:

- ряд хронологічних дат (моментів) або інтервалів часу;
- ряд числових значень статистичного показника, які називаються

рівнями ряду.

Залежно від виду статистичних показників - рівнів, ряди динаміки поділяють на: ряди абсолютних величин; ряди відносних величин; ряди середніх величин.

За ознакою часу динамічні ряди поділяються на два види:інтервальні ряди; моментні ряди.

Інтервальний (періодичний) ряд - такий ряд динаміки, який характеризує явище за будь-які інтервали часу (за місяць, квартал, рік).

Характерними особливостями інтервального динамічного ряду є: залежність величини рівня ряду від величини проміжку часу; показники періодичного (інтервального) ряду динаміки, як правило, можна додавати і ці показники мають реальний конкретний зміст; наприклад, в результаті додавання можна одержати новий динамічний ряд динаміки, кожний показник якого характеризує величину явища за збільшені періоди часу.

Моментний ряд - такий ряд динаміки, який характеризує явище за станом на будь-які моменти часу (на початок або на кінець періоду або на певну дату).

Характерною особливістю моментного ряду динаміки є те, що кожний наступний рівень ряду частково або повністю містить в собі попередній і тому підсумування (додавання) послідовних рівнів ряду не дає реальних показників.

За допомогою моментних рядів динаміки вивчають товарні запаси, стан кадрів на підприємстві, кількість обладнання, оборотні кошти, основні фонди та ін..

Підсумовування рівнів моментного ряду не має змісту.

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 доход 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 ознаками багатомірного динамічного ряду буває **функціональним**

(адитивним чи мультиплікативним) або **кореляційним.**

8.2. Показники аналізу рядів динаміки

У процесі аналізу рядів динаміки використовують такі показники: абсолютний приріст; темп зростання; темп приросту; абсолютне значення одного процента приросту; абсолютне та відносне прискорення (уповільнення); коефіцієнт прискорення (уповільнення); коефіцієнт випередження.

Абсолютний приріст показує, на скільки одиниць збільшився або зменшився наступний рівень ряду у порівнянні з попереднім або базисним. Обчислюється як різниця рівнів ряду ланцюговим і базисним методом:

$$\text{ланцюговий: } \Delta_y = y_i - y_{i-1}, \tag{8.1}$$

$$\text{базисний: } \Delta_y = y_i - y_0, \tag{8.2}$$

де Δ - абсолютний приріст;
y_i - наступний рівень ряду;
y_{i-1} - попередиш рівень;
y_0 - базисний рівень.

Сума ланцюгових абсолютних приростів дорівнює кінцевому базисному:

$$\sum (y_i - y_{i-1}) = y_n - y_0. \tag{8.3}$$

Темп зростання показує, у скільки разів наступний рівень більший або менший від попереднього або базисного рівня. Обчислюється співвідношенням рівнів ланцюговим та базисним методом.

$$\text{Ланцюговий: } T_{зр} = \frac{y_i}{y_{i-1}}. \tag{8.4}$$

$$\text{Базисний: } T_{зр} = \frac{y_i}{y_0}. \tag{8.5}$$

Темпи зростання виражають або у коефіцієнтах, або у відсотках.

Добуток ланцюгових темпів зростання дорівнює кінцевому базисному:

$$\frac{y_1}{y_0}\cdot\frac{y_2}{y_1}\cdot\frac{y_2}{y_2}\times\ldots\times\frac{y_n}{y_{n-1}}=T_1\cdot T_2\cdot T_3\times\ldots\times T_n=\prod_{i-1}^{n}T=\frac{y_n}{y_0} \qquad (8.6)$$

Ящо $T_{зр} > 1$ - збільшення. Якщо $T_{зр} < 1$ - зменшення.

Темп приросту показує, на скільки процентів наступний рівень $y_{(}$ більший або менший від попереднього y_{i-1} або базисного y_0. Вимірюється тільки у відсотках. Обчислюється як співвідношення абсолютного приросту до рівня порівняння, або як різниця між темпом зростання і 100%:

$$\text{Ланцюговий:}\quad T_{пр}=\frac{\Delta_y}{y_{i-1}}\times 100\%. \qquad (8.7)$$

$$\text{Базисний:}\quad T_{пр}=\frac{\Delta_y}{y_0}\times 100\%. \qquad (8.8)$$

$$\text{або}\quad T_{пр}=T_{зр}-100\%. \qquad (8.9)$$

Абсолютне значення одного процента приросту показує, яка абсолютна величина відповідає кожному проценту приросту. Обчислюється як відношення абсолютного приросту до темпу приросту *ланцюговим методом*:

$$A\%=\frac{\Delta_y}{T_{пр}}=\frac{y_i-y_{i-1}}{\left(\frac{y_i-y_{i-1}}{y_{i-1}}\right)\cdot 100}=\frac{y_{i-1}}{100}. \qquad (8.10)$$

Узагальнюючою характеристикою ознаки динамічного ряду є середня величина

Якщо порівняти ланцюгові абсолютні прирости шляхом різниці між наступним і попереднім, то отримаємо **абсолютне прискорення чи уповільнення** швидкості розвитку явища:

$$\Delta' = \Delta_i - \Delta_{i-1}. \qquad (8.11)$$

Знак “+” вказує на прискорення абсолютної швидкості динаміки, а знак “-” - навпаки - на уповільнення.

Порівняння абсолютних приростів шляхом співвідношення показує **відносне прискорення чи уповільнення** швидкості динаміки:

$$k_{\Delta} = \frac{\Delta_i}{\Delta_{i-1}} \qquad (8.12)$$

Порівняння ланцюгових темпів зростання дає **коефіцієнт прискорення (уповільнення)** відносної швидкості розвитку:

$$k_t = \frac{T_{зр.i}}{T_{зр.i-1}} \qquad (8.13)$$

Порівняння інтенсивності розвитку ознак у різних рядах розраховують за допомогою коефіцієнта випередження: $ТЗ'' \div ТЗ'$. За допомогою останнього порівнюють відносну швидкість динамічних рядів однакового змісту по різних об’єктах (регіони, країни тощо) або різного змісту по одному об’єкту.

Щодо темпів приросту, то співвідношення їх використовують лише для взаємопов’язаних показників х і у. Таке співвідношення називають **коефіцієнтом еластичності К=ТПу/ТПх,** він показує, на скільки відсотків змінюється **у** зі зміною **х** на один відсоток.

Ряди динаміки можна подавати не тільки в таблицях, а й у вигляді графічних зображень. У цьому разі на осі абсид відкладають шкалу часу, а на осі ординат – шкалу рівнів ряду.

8.3. Середні показники динаміки

З часом змінюються як рівні рядів динаміки, так і обчислені на їх основі показники аналізу. Тому для узагальненої характеристики розвитку явищ використовують середні показники: середні рівні; середні абсолютні прирости; середні темпи зростання.

Розглянемо обчислення середніх рівнів рядів динаміки.

В **інтервальному ряді** абсолютних величин використовується

середня арифметична проста:

$$\bar{y} = \frac{\sum y_i}{n}, \qquad (8.14)$$

де n - число рівнів ряду;

уі- рівні ряду.

Для моментних рядів динаміки середній рівень обчислюється залежно від характеру статистичної інформації.

Якщо є **два рівня моментного ряду** (на початок і на кінець періоду) ,то середній рівень розраховують за середньою арифметичною простою:

$$\bar{y} = \frac{y_0 + y_n}{2}. \qquad (8.15)$$

Якщо **рівнів моментного ряду динаміки більше ніж 2** (n > 2), то середній рівень обчислюється за формулою середньої хронологічної:

$$\bar{y} = \frac{\frac{y_1 + y_n}{2} + y_2 + y_3 + \ldots + y_{n-1}}{n-1}. \qquad (8.16)$$

Для обчислення середнього рівня моментного ряду іноді використовують формулу середньої арифметичної зваженої, **якщо інтервали часу нерівні:**

$$\bar{y} = \frac{\sum y_i t_i}{\sum t_i}, \qquad (8.17)$$

де y_i - рівень ряду;

t_i - інтервал часу між датами.

Середній абсолютний приріст обчислюється як середня арифметична проста із ланцюгових абсолютних приростів або як відношення кінцевого базисного абсолютного приросту до кількості

відрізків часу:

$$\overline{\Delta}_n = \frac{\sum \Delta_y}{n} = \frac{y_n - y_0}{n}. \qquad (8.18)$$

Середній темп зростання обчислюється як середня геометрична з ланцюгових темпів зростання:

$$T_{зр} = \sqrt[n]{T_1 \times T_2 \times .. T_n} = \sqrt[n]{\prod_{i=1}^{n} T_i}, \qquad (8.19)$$

де n - кількість ланцюгових темпів зростання.

$$\overline{T}_{зр} \sqrt[n]{\frac{y_n}{y_0}}, \qquad (8.20)$$

де n - число відрізків часу.

При обчисленні середніх показників рядів динаміки необхідно вказувати період часу, за який обчислюється середня величина та часову одиницю вимірювання (рік, квартал, місяць).

Питання для самоконтролю

1. Які завдання виконуються статистикою за допомогою рядів динаміки?
2. Що називають рядом динаміки?
3. З яких елементів складається ряд динаміки?
4. Які види рядів динаміки існують?
5. Наведіть приклади моментного та інтервального динамічних рядів, зазначте їх елементи та особливості.
6. Як визначаються базисні та ланцюгові характеристики динаміки? Який між ними існує взаємозв’язок?
7. Які середні поеазники використовуються в рядах динаміки?
8. Які види середніх використовуються для визначення середнього рівня в моментних та інтервальних рядах?

Література [4 с.172-189, 5 с.187-204].

ТЕМА 9. АНАЛІЗ ТЕНДЕНЦІЙ РОЗВИТКУ ТА КОЛИВАНЬ

План лекції

9.1. Характеристика основної тенденції розвитку
9.2. Вивчення сезонних коливань

9.1. Характеристика основної тенденції розвитку

Тенденція - основний напрямок розвитку (зростання, стабільність або зниження рівнів явища).

У статистичній практиці виявлення основної тенденції розвитку явищ використовують такі методи: укрупнення інтервалів часу; обчислення плинної середньої; аналітичне вирівнювання.

Суть методу укрупнення інтервалів часу, до яких відносяться рівні динамічного ряду, полягає в тому, що рівні об'єднуються в групи по інтервалах, розраховується середній рівень за інтервал - 3 роки, 5 років і т.д.

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.

Другий метод виявлення основної тенденції розвитку є згладжування ряду динаміки за допомогою плинної середньої. Суть його також полягає в укрупленні інтервалів, але воно проводиться шляхом послідовних зміщень на одну дату при збереженні постійного інтервалу часу.

Вибір інтервалу згладжування переважно здійснюється мимовільно, однак при цьому слід враховувати кількість рівнів у динамічному ряді, який аналізується, тому що при використанні прийому плинної середньої згладжений ряд скорочується порівняно з первинним рядом на (m— 1). Разом з тим, чим довший інтервал згладжування, тим сильніше досягається усереднення і, відповідно,

виявлена тенденція розвитку буде більш плавною. Переважно інтервал згладжування може складатися з трьох, п'яти або семи рівнів.

Вивчення основної тенденції розвитку методом плинної середньої є прийомом попереднього аналізу. Розглянуті прийоми згладжування динамічних рядів можна розглядати як важливий допоміжний засіб, який полегшує застосування інших методів виявлення тенденції. Для того, щоб мати кількісну модель, яка виражає загальну тенденцію зміни рівнів динамічного ряду у часі, використовується **вирівнювання ряду динаміки**.

Найбільш ефективним методом виявлення основної тенденції є **аналітичне вирівнювання.** Його використовують у рядах з чітко визначеною тенденцією. При цьому рівні ряду динаміки розглядаються як функція часу:

$$Y=f(t), \qquad (9.1)$$

де Y - теоретичні рівні ряду,

t - змінна часу (0, 1, 2, 3,,.., n).

Таку функцію називають трендовим рівнянням. Коли абсолютні прирости відносно постійні використовують лінійне рівняння $Y = a + bt$, а при відносно постійних темпах приросту - показникову функцію $Y = ab^t$.

Розглянемо лінійне рівняння (тренд). Параметри тренду: a - початковий рівень; b – характеризує швидкість динаміки (щорічний абсолютний приріст, або середня відносна).

Коли характеристики швидкості розвитку зростають (чи зменшуються), використовують інші функції (парабола 2-го степеня, модифікована експонента тощо.)

Для знаходження параметрів a і b необхідно розв'язати систему нормальних рівнянь за методом найменших квадратів.

Суть його полягає в знаходженні такої прямої або кривої, ординати точок якої були б найближчі до значень фактичного динамічного ряду. При такому вирівнюванні досліджуване явище збільшуватиметься або зменшуватиметься в кожному періоді на однакову величину в арифметичній прогресії.

Згідно з умовою мінімізації суми квадратів відхилень фактичних рівнів ряду y_t від теоретичних Y_t параметри визначаються

розв'язуванням системи нормальних рівнянь, Для лінійної функції вона записується так:

$$\begin{cases} \sum y = na + b\sum t \\ \sum yt = a\sum t + b\sum t^2 \end{cases}, \qquad (9.2)$$

де у - фактичні рівні ряду динаміки;
п - число рівнів ряду.

Дану систему рівнянь можна легко спростити, якщо відлік часу брати з середини ряду таким чином, щоб сума часу дорівнювала нулю ($\sum t$ =0).

При непарному числі рівнів серединна точка (початок відліку часу) приймається за 0, тоді попередні періоди позначаються відповідно через -1, -2, -3 і т. д., а наступні за серединним періодом відповідно + 1, +2, +3 і т. д.

При парному числі рівнів два середніх моменти часу позначаються через -1 і +1, всі решта - через два інтервали, тобто попередні періоди до середини через -3, -5, -7 і т. д., а наступні - відповідно +3, +5, +7 і т. д.

При відліку часу від середини ряду, коли $\sum t$ =0, система нормальних рівнянь спрощується і має такті вигляд:

$$\begin{cases} \sum y = na \\ \sum yt = b\sum t^2 \end{cases}. \qquad (9.3)$$

Наступним кроком є обчислення параметрів тренду: **Y = a + bt:**

$$a = \frac{\sum y}{n}; \qquad (9.4)$$

$$b = \frac{\sum yt}{\sum t^2}. \qquad (9.5)$$

Послідовно підставивши у рівняння тренду значення t (-3, -2, -1 і т.д.), дістанемо згладжений (теоретичний ряд) динаміки, абстрагований від випадкових коливань.

Якщо обчислення виконано правильно, то $\sum Y = \sum y$. Отже, значення рівнів згладженого динамічного ряду знайдено правильно.

Оцінка адекватності трендового рівняння реальному процесу.

Для оцінки адекватності трендового рівняння використовують:

- стандартну похибку $S_e = \sqrt{\dfrac{\sum_1^n (y_t - Y_t)^2}{n-m}}$; (9.6)

- коефіцієнт детермінації $R^2 = 1 - \dfrac{\sum (y_t - Y_t)^2}{\sum (y_t - \bar{y})^2}$, (9.7)

де n — довжина ряду,

m — кількість параметрів функції (для лінійної функції m = 2).

Екстраполяція тренду.

При аналізі динамічних рядів часто застосовують такі методи, як інтерполяція та екстраполяція.

Екстраполяція — розрахунок (прогноз) показників, які можуть знаходитись за межами досліджуваного ряду динаміки. Такі розрахунки здійснюють, виходячи з припущення, що виявлена тенденція у фактичному досліджуваному динамічному ряді матиме місце і надалі.

Методом екстраполяції визначається точковий прогноз показника; довірчі його межі залежать від похибки прогнозу S_p та імовірності висновку t, тобто $\mathbf{Y_{t+V} \pm ts_p}$. Похибка прогнозу S_p є функцією стандартної похибки S_e, довжини динамічного ряду (передісторії) і періоду упередження. В даному прикладі:

$$S_p = S_e \sqrt{\frac{n+1}{n} + \frac{3(n+2\upsilon-1)^2}{n(n^2-1)}} \ . \qquad (9.8)$$

Знаючи похибку прогнозу розраховують довірчий інтервал для прогнозного рівня динамічного ряду.

Інтерполяція — це знаходження невідомого (відсутнього) рівня динамічного ряду. Його можна знайти двома способами:

1. Визначити середньорічний абсолютний приріст і припустити, що збільшення обсягів має рівномірний зростаючий характер.

2. Визначити невідомий рівень динамічного ряду можна як півсуму із значень двох суміжних рівнів.

Таким чином, як інтерполяція, так і екстраполяція ґрунтуються на припущенні, що наявні величини цілком достатньо визначають темп зростання досліджуваного явища і, отже, його можна поширювати на невідомі рівні динамічного ряду.

В процесі аналізу рядів динаміки можлива ситуація, коли рівні динамічного ряду за колом охоплюваних об'єктів не співставимі. Непорівнюванність може бути наслідком зміни територіальних меж або зміни підпорядкування деяких об'єктів тощо.

За умови, що два відрізки динамічного ряду є непорівнянними, вдаються до такого прийому, як **зімкнення динамічних рядів**., тобто розраховується коефіцієнт зімкнення

$$K_з = \text{рівень дин. ряду в нових межах / рівень дин. ряду в старих межах.} \quad (9.9)$$

У розвитку соціально-економічних явищ поряд з тенденцією існують відхилення від тренду, структурні зміни, сезонні коливання.

9.2. Вивчення сезонних коливань

При вимірюванні сезонних коливань обчислюють індекси сезонності.

Індекс сезонності - процентне відношення фактичного рівня *y* за місяць (квартал) року до середньомісячного (середньоквартального):

$$I_c = \frac{y_i}{\bar{y}}. \quad (9.10)$$

У сукупності ці індекси утворюють **сезонну хвилю**.

Абсолютною мірою сезонних коливань є амплітуда коливань:

$$R = I_{\max} - I_{\min} \,. \qquad (9.11)$$

Для порівняння інтенсивності сезонних коливань використовують **середнє лінійне відхилення:**

$$\bar{l} = \frac{\sum |I_c - 100|}{12} \,. \qquad (9.12)$$

або **середнє квадратичне відхилення:**

$$\sigma = \sqrt{\frac{\sum (I_c - 100)^2}{12}} \,. \qquad (9.13)$$

Питання для самоконтролю

1. Дайте визначення тенденції розвитку. Відповідь проілюструйте прикладами.
2. У чому полягає суть методу укрупнення періодів ряду динамік?
3. Як здійснюють згладжування рядів динаміки способом плинної середньої?
4. У чому полягає суть методу аналітичного згладжування рядів динаміки?
5. Якими методами досліджуються сезонні коливання?

Література [4 с.189-205, 5 с.204-244].

ТЕМА 10. ІНДЕКСНИЙ МЕТОД

План лекції

10.1. Суть і види індексів
10.2. Агрегатні індекси, взаємозв'язок індексів
10.3. Середньозважені індекси
10.4. Індекси середніх величин

10.1. Суть і види індексів

У перекладі з латинської мови “INDEX” означає “показник”, проте в статистиці він набуває специфічного значення.

Індекс - відносна величина, яка характеризує зміну соціально-економічного явища у часі і просторі або порівняно зі стандартом. Вимірюється в коефіцієнтах або у відсотках.

За допомогою індексного методу вирішують такі завдання:

- одержують порівняльну характеристику зміни явища у часі, де індекси виступають як показники динаміки;
- характеризують виконання норми, затвердженого стандарту чи плану. Отже, індекси є засобом оперативного висвітлення виробничого процесу;
- оцінюють роль окремих факторів, що формують складне явище;
- дають порівняльну характеристику зміни явищ у просторі. У цьому разі індекси забезпечують територіальні порівняння.

Класифікація індексів.

<u>За характером порівнянь індекси</u> поділяються на такі групи: динамічні індекси; територіальні індекси; міжгрупові індекси.

Динамічний індекс характеризує зміну явища у часі. Обчислюється як співвідношення числових значень однойменних показників поточного (звітного) періоду до попереднього (базисного). Підрядкова позначка для показника поточного (звітного) періоду “1”, а базисного - “0”.

Територіальний індекс відбиває результат порівняння між об'єктами, країнами, регіонами. Показники, що порівнюються, позначаються певними літерами.

Міжгруповий індекс характеризує відхилення від стандарту. <u>За ступенем охоплення елементів сукупності</u> індекси поділяють на індивідуальні та зведені.

Індивідуальні індекси (і) характеризують зміну у часі, просторі або відхилення від стандарту окремих елементів сукупності.

Приклади економічних індивідуальних індексів:

динамічні територіальні

ціни, p

$$i_p = \frac{p_1}{p_0} \qquad i_p = \frac{p_A}{p_B}$$

фізичного обсягу (кількості), q

$$i_q = \frac{q_1}{q_0} \qquad i_q = \frac{q_A}{q_B}$$

вартості (товарообороту), pq

$$i_{pq} = \frac{p_1 q_1}{p_0 q_0} \qquad i_p q = \frac{p_A q_A}{p_B q_B}$$

собівартості, c

$$i_c = \frac{c_1}{c_0} \qquad i_c = \frac{c_A}{c_B}$$

грошових витрат, cq

$$i_{cq} = \frac{c_1 q_1}{c_0 q_0} \qquad i_{cq} = \frac{c_A q_A}{c_B q_B}$$

трудомісткості, t

$$i_{tp} = \frac{t_1}{t_0} \qquad i_t = \frac{t_A}{t_B}$$

продуктивності праці, w

$$i_w = \frac{w_1}{w_0} \qquad i_w = \frac{w_A}{w_B}$$

Зведені індекси (I) характеризують зміну певної множини елементів і поділяються на групові і загальні.

Наприклад, підприємство випускає 5 видів продукції. Індекс

собівартості по 3-х видах продукції буде груповим, а по всіх 5-ти – загальним.

Зведені індекси характеризують зміну сукупності, до якої входять різнорідні елементи, які не можна безпосередньо підсумовувати. Результат такого підсумовування не має економічного змісту.

За характером досліджуваних об'єктів індекси поділяють на такі види:

- індекси об'ємних показників (фізичного обсягу);
- індекси якісних показників (цін, собівартості, продуктивності праці та ін.).

Індекси об'ємних показників - це індекси фізичного обсягу продукції, роздрібного товарообороту, споживання окремих продуктів тощо. В них міститься характеристика зміни обсягу того чи іншого явища, який виражають у певних одиницях виміру.

Індекси якісних показників - це індекси цін, собівартості і продуктивності праці тощо. В таких індексах міститься характеристика зміни якісної ознаки, тобто такої, що відображує особливості розвитку явища.

За методом обчислення зведені індекси поділяють на:

- агрегатні;
- середньозважені;
- індекси середніх величин (індекси змінного складу, фіксованого складу, структурних зрушень).

Залежно від бази порівняння розрізняють ланцюгові і базисні.

10.2. Агрегатні індекси, взаємозв'язок індексів

Основною формою економічного індексу є агрегатна.

Агрегатний індекс - співвідношення двох агрегатів. Агрегат являє собою суму добутків показників, що індексуються, на їх співмножники. Наприклад, агрегат вартості товарів $\sum pq$. Побудуємо індекс цін товарів:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_0}, \qquad (10.1)$$

де p - показник ціна - індексується (змінюється);
q - показник фізичного обсягу - співмножник (вага);
p_1 - ціна поточного (звітного) періоду;
p_o - ціна попереднього (базисного) періоду.

Фізичні обсяги (кількість товарів) q повинні бути незмінними для чисельника і знаменника, щоб можна було виявити зміну цін товарів.

У статистичній практиці використовують дві рівноправні системи агрегатних індексів:

за базисною вагою Лайспереса		**за поточною (звітною) вагою Пааше**
	цін	
$I_p = \frac{\sum p_1 q_0}{\sum p_0 q_0}$		$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$
	фізичного обсягу	
$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$		$I_q = \frac{\sum q_1 p_1}{\sum q_0 p_1}$

Аналогічно можна побудувати індекси собівартості, продуктивності праці, трудомісткості.

У вітчизняній статистичній практиці прийнято: при побудові індексів якісної ознаки його ваги фіксують на рівні звітного періоду; у разі індексів об'ємних ознак —ознаки - сумірники фіксують на рівні базисного періоду. Це зумовлено тим, що кожен із співмножників відіграє різну роль.

Оскільки добуток ціни на кількість товару дає вартість товару, то використовується взаємозв'язок індексів $I_{pq} = I_p \cdot I_q$. При цьому необхідно звернути увагу на те, що індекси-співмножники повинні бути в різних вага. Якщо індекс ціни I_p розрахований за системою Лайспереса, то індекс фізичного обсягу I_q - за системою Пааше:

$$\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum q_1 p_1}{\sum q_0 p_1} = \frac{\sum p_1 q_1}{\sum p_0 q_0}, \qquad (10.2)$$

$I_{pq} = \dfrac{\sum p_1 q_1}{\sum p_0 q_0}$ - індекс вартості товарів,

$\sum p_1 q_1$ - сума фактичних вартостей поточного (звітного) періоду;

$\sum p_0 q_0$ - сума фактичних вартостей попереднього (базисного) періоду.

Отже, можна зробити такі висновки.

1. При побудові загальних індексів кількісних показників, які безпосередньо не можна підсумувати, співмножниками індексованих показників є якісні показники, які виконують роль сумірників. В індексах якісних показників (цін, собівартості, продуктивності), які можна підсумовувати безпосередньо, співмножником завжди є пов'язана з ним кількісна ознака, проте вона виконує роль не сумірника, а ваги, тобто відображує значущість даного якісного показника в загальній їх сукупності.

2. У загальних індексах якісних ознак беруть кількісні ознаки, що виконують роль ваги, на рівні звітного періоду. Тільки в такий спосіб можна показати реальну зміну якісної ознаки. Водночас у загальних індексах об'ємної ознаки якісна ознака, що виконує роль сумірника, береться на рівні базисного періоду.

За допомогою індексної системи вивчається вплив факторів на результат. **Наприклад**, як вплинула зміна цін і зміна обсягів продажу на вартість товарів. Абсолютна зміна визначається як різниця між чисельником і знаменником відповідного індексу.

$\Delta_{pq} = \sum p_1 q_1 - \sum p_0 q_0$ - абсолютна зміна вартості товарів у цілому, (10.3)

в тому числі за рахунок зміни цін $\Delta_p = \sum p_1 q_0 - \sum p_0 q_0$ (10.4)

за рахунок зміни обсягів продажу $\Delta_q = \sum p_1 q_1 - \sum p_1 q_0$ (10.5)

10.3. Середньозважені індекси

Середньозважений індекс обчислюють як середню величину з індивідуальних індексів за формулою середньої арифметичної зваженої або гармонічної зваженої.

Середньозважений індекс тотожний відповідному агрегатному індексу.

Наприклад, середньозважений індекс цін за середньою арифметичною зваженою відповідає агрегатному індексу цін за системою Ласпереса:

$$I_p = \frac{\sum i_p p_0 q_0}{\sum p_0 q_0} = \frac{\sum \frac{p_1}{p_0} p_0 q_0}{\sum p_0 q_0} = \frac{\sum p_1 q_0}{\sum p_0 q_0}, \quad p_0 q_0 - \textit{вага}. \qquad (10.6)$$

Середньозважений індекс цін за середньою гармонічною зваженою відповідає агрегатному індексу цін за системою Пааше:

$$I_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}} = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{p_1 / p_0}} = \frac{\sum p_1 q_1}{\sum p_0 q_1}, \quad p_1 q_1 - \textit{вага}. \qquad (10.7)$$

Аналогічно тотожні індекси фізичного обсягу, собівартості, трудомісткості та ін.

10.4. Індекси середніх величин

Індекси середніх величин характеризують зміну середніх величин, до них відносяться: індекс змінного складу; індекс постійного (фіксованого) складу; індекс структурних зрушень.

Індекс змінного складу характеризує зміну середньої величини в цілому, тобто за рахунок зміни значень ознаки х і зміни в структурі сукупності:

$$I_{з.с.} = \frac{\bar{x}_1}{\bar{x}_0} = \frac{\sum x_1 f_1}{\sum f_1} \div \frac{\sum x_0 f_0}{\sum f_0} = \sum x_1 d_1 \div \sum x_0 d_0 \ .$$

(10.8)

Індекс постійного складу показує, як змінилась середня величина тільки за рахунок зміни у середньому значень ознаки х при постійній (фіксованій) структурі сукупності:

$$I_{п.с.} = \frac{\sum x_1 f_1}{\sum f_1} \div \frac{\sum x_0 f_1}{\sum f_1} = \sum x_1 d_1 \div \sum x_0 d_1 \ . \qquad (10.9)$$

Індекс структурних зрушень характеризує зміну середньої величини тільки за рахунок зміни структури сукупності (структурних зрушень) при постійних значеннях ознаки х:

$$I_{с.з.} = \frac{\sum x_0 f_1}{\sum f_1} \div \frac{\sum x_0 f_0}{\sum f_0} = \sum x_0 d_1 \div \sum x_0 d_0 \ . \qquad (10.10)$$

Всі три індекси ув'язуються в систему:

$$I_{з.с.} = I_{п.с.} \times I_{с.з.}$$

Питання для самоконтролю

1. Що характеризує індекс? Які завдання виконують індекси в аналізі розвитку суспільно-економічних явищ та процесів?
2. Які види індексів існують? Поясніть особливості побудови та застосування різних видів індексів.
3. Назвіть основні принципи побудови індивідуальних та зведених індексів.
4. На яких методичних принципах ґрунтується побудова агрегатних індексів? Які системи зважування індексів використовуються при побудові агрегатних індексів цін та фізичного обсягу?
5. Назвіть особливості індексного методу.

6. Побудуйте зведений індекс цін у формі індексу Е. Ласпейреса та Г. Пааше. Вкажіть, у чому різниця між цими формами індексів.

7. Поясніть суть середньозважених індексів та поясніть їх зв'язок з відповідними індексами агрегатної форми. Назвіть умови та особливості застосування середньозважених індексів.

8. Чим викликана потреба застосування індексів середніх величин? Відповідь проілюструйте прикладами.

Література [4 с.210-238, 5 с.245-267].

ТЕМА 11. ВИБІРКОВИЙ МЕТОД

План лекції

11.1. Суть вибіркового спостереження
11.2. Обчислення помилок вибірки
11.3. Різновиди вибірок
11.4. Визначення обсягу вибірки

11.1. Суть вибіркового спостереження

Вибіркове спостереження - такий вид несуцільного спостереження, при якому обстежуються не всі елементи сукупності, що вивчається, а лише їх частина, відібрана за певними правилами, що дозволяє на підставі вибіркових оцінок отримати дані для характеристики усієї сукупності.

Вибіркове спостереження має переваги порівняно з суцільним спостереженням:

- потребує менше коштів, засобів, часу;
- дозволяє більш детально вивчити кожний елемент;
- забезпечує більш точні результати, ніж при суцільному спостереженні, якщо сукупність дуже велика;
- може застосовуватись там, де не можна застосувати суцільне спостереження, якщо сукупність дуже велика, або якщо під час дослідження одиниці сукупності знищуються або псуються, наприклад, при обстеженні якості продовольчої продукції.

Вибірковий метод дозволяє через вивчення частини спеціально відібраних одиниць охарактеризувати масове явище в цілому.

Теорія і практика вибіркового спостереження показує, що воно за

правильної організації дає достовірні відомості, цілком придатні для практичного використання.

Основними етапами вибіркового спостереження є:

- обґрунтування мети вибіркового спостереження;
- складання програми спостереження і розробка відповідних даних;
- вирішення організаційних питань щодо спостереження;
- визначення частки і способу відбору одиниць у вибіркову сукупність;
- здійснення відбору;
- реєстрація ознак досліджуваних одиниць;
- узагальнення даних спостереження та визначення вибіркових характеристик;
- обчислення похибок вибірки;
- поширення кількісних характеристик вибіркового спостереження на всю сукупність.

Сукупність, з якої відбираються елементи для обстеження, називається **генеральною**.

Сукупність, яку безпосередньо обстежують, називається **вибірковою.**

Узагальнені показники генеральної сукупності називаються **генеральними**, а вибіркової **- вибірковими**.

Показники	Генеральна сукупність	Вибіркова сукупність
Обсяг сукупності	N	n
Середнє значення	$\bar{x}$	$\tilde{x}$
Частка одиниць сукупності, що мають певні значення ознаки	d	p

Рисунок 11.1- Умовні позначення для вибіркового спостереження

Розбіжності між показниками генеральної і вибіркової сукупності називаються **помилками репрезентативності** і виникають тому, що вибіркова сукупність неточно відтворює склад генеральної

сукупності. Для середньої величини помилка являє собою різницю між генеральною і вибірковою середніми, для частки - різницю між генеральною і вибірковою частками.

Помилки репрезентативності бувають **випадковими і систематичними**.

Систематичних помилок можна уникнути, якщо строго дотримуватись правил випадкового відбору.

Випадкових помилок уникнути не можна, вони виникають як наслідок того, що структура вибіркової сукупності не збігається зі структурою генеральної сукупності.

Завдання вибіркового спостереження - обчислення помилок і визначення параметрів їх зменшення.

11.2. Обчислення помилок вибірки

У теорії вибіркового методу спостереження розглядаються два методи відбору:

- **повторний відбір**, при якому відібрана у вибірку одиниця сукупності реєструється і потім повертається у генеральну сукупність і знову може бути відібрана;
- **безповторний відбір**, при якому відібрана у вибірку одиниця сукупності не повертається у генеральну сукупність і більше у відборі не приймає участі, тому безповторний відбір гарантує більш точні результати.

Ступінь вірогідності відхилення визначають на основі теорії ймовірності та закону великих чисел.

Для обчислення середньої (стандартної) помилки вибірки використовують такі формули:

Спосіб відбору	Для середньої величини	Для частки
Повторний	$\mu = \sqrt{\frac{\sigma^2}{n}}$	$\mu = \sqrt{\frac{p(1-p)}{n}}$
Безповоротний	$\mu = \sqrt{\frac{\sigma^2}{n}(1-\frac{n}{N})}$	$\mu = \sqrt{\frac{p(1-p)}{n}(1-\frac{n}{N})}$

Рисунок 11.2 – Формули для обчислення середньої (стандартної) помилки вибірки

де μ - середня (стандартна) помилка репрезентативності;
σ^2 - дисперсія у вибірці;
n/N - частка обстеженої частини вибіркової сукупності;
(1 - n/N) - необстежена частина генеральної сукупності;
p — частка одиниць, яка володіє даною ознакою:
(1 - p) - частка одиниць, яка не володіє даною ознакою.

μ - середня помилка характеризує міру відхилення вибіркової ($\tilde{x}$) середньої величини від генеральної середньої ($\bar{x}$).

Для узагальненої характеристики помилки вибірки поряд із середньою (стандартною) помилкою розраховують **граничну помилку вибірки**.

Стверджувати, що дана генеральна середня ($\bar{x}$) не вийде за межі середньої (стандартної) помилки вибірки можна лише з певним ступенем ймовірності (P).

Помилка вибірки, що обчислена із ймовірністю більш ніж 0,683, називається у статистиці **граничною** і обчислюється за формулою:

$$\Delta = t\mu, \qquad (11.1)$$

де μ - стандартна помилка;

t - коефіцієнт довіри (квантиль нормального розподілу).

Коефіцієнт довіри (t) показує як співвідносяться гранична та стандартна помилки вибірки:

$$t = \frac{\Delta}{\mu}. \qquad (11.2)$$

t залежить від ймовірності, з якою гарантується величина граничної помилки вибірки і знаходиться в спеціальних таблицях, **наприклад:**

P = 0,683, t =1, P= 0,954, t = 2, P =0,997,t = 3, P = 0,999, t = 4.

Це означає, що із ймовірністю 0,683 можна стверджувати, що гранична помилка не вийде за межі стандартної, тобто $\Delta = \pm 1\mu$, а з P= 0,954, $\Delta = \pm 2\mu$, з P = 0,997, $\Delta = \pm 3\mu$, з P = 0,999, D = $\pm 4\mu$.

Отже, гранична помилка вибірки відповідає на питання про точність вибірки з певною ймовірністю, величина якої обчислюється значенням t.

Отже, застосовують такі формули граничної помилки вибірки:

Спосіб відбору	Для середньої величини	Для частки
Повторний	$\Delta_{\tilde{x}} = t\sqrt{\frac{\sigma^2}{n}}$	$\Delta_p = t\sqrt{\frac{p(1-p)}{n}}$
Безповторний	$\Delta_{\tilde{x}} = t\sqrt{\frac{\sigma^2}{n}(1-\frac{n}{N})}$	$\Delta_p = t\sqrt{\frac{p(1-p)}{n}(1-\frac{n}{N})}$

Рисунок 11.3- Формули граничної помилки вибірки

Гранична помилка вибірки дає можливість встановити, в яких межах знаходиться величина генеральної середньої ($\bar{x}$) або частки (d):

$$\bar{x} = \tilde{x} \pm \Delta_x \ \ або \tilde{x} - \Delta_x \le \bar{x} \le \tilde{x} + \Delta_x, \qquad (11.3)$$

$$d = p \pm \Delta_p \ \ або\, p - \Delta_p \le d \le p + \Delta_p. \qquad (11.4)$$

Довірчі межі (або довірчий інтервал) є інтервальною оцінкою параметрів вибіркової середньої або вибіркової частки, що розраховуються за даними вибірки для певної ймовірності. Причому чим менший довірчий інтервал, тим точніша оцінка.

11.3. Різновиди вибірок

У практиці вибіркових обстежень використовують різні способи формування вибіркових сукупностей: простий випадковий; механічний; типовий (розшарований, районований); серійний (гніздовий).

Простий випадковий відбір проводиться жеребкуванням або на основі таблиць випадкових чисел. Це - класичний спосіб формування вибіркової сукупності.

Він передбачає підготовку до формування вибірки. Для жереба на кожну одиницю сукупності готують фішку, а для таблиць випадкових чисел - усі одиниці сукупності нумерують. Розрізняють два

види випадкових вибірок: повторні і безповторні.

Механічний відбір - передбачає упорядкування одиниць генеральної сукупності, тобто відбір одиниць здійснюється через рівні інтервали. Крок інтервалу обчислюється діленням обсягу сукупності N на передбачений обсяг вибірки n (n/N). Початковий елемент вибірки визначається як випадкове число всередині першого інтервалу, другий елемент залежить від початкового числа й кроку інтервалу.

Помилку механічної вибірки обчислюють за формулою безповторної вибірки:

Для моментних спостережень, суть яких зводиться до фіксації стану безперевного процесу на певні моменти часу, використовують формулу помилки повторної вибірки:

Типовий (розшарований, районований) відбір - спосіб формування вибірки з урахуванням структури генеральної сукупності.

Генеральну сукупність поділяють на якісно однорідні групи за якоюсь важливою типовою ознакою. Потім із кожної групи проводять випадковий відбір елементів пропорційно питомій вазі групи у сукупності.

Помилку розшарованої вибірки обчислюють, використовуючи середню з групових дисперсій σ^2:

$$\Delta = t\sqrt{\frac{\overline{\sigma^2}}{n}\left(1-\frac{n}{N}\right)}. \qquad (11.5)$$

Як правило $\overline{\sigma}^2 \langle \sigma^2$, отже, помилка розшарованої вибірки менша, ніж механічної чи простої випадкової.

Суть **серійного відбору** полягає в тому, що відбирають не одиниці сукупності, а серії одиниць, які розглядають як одне ціле. Якщо серія потрапила у вибірку, то обстежують усі без винятку одиниці серії. Серійна вибірка застосовується у двох варіантах: 1) коли всі серії мають однакову кількість одиниць і 2) коли серії неоднакові за обсягами.

При обчисленні помилки вибірки враховується міжсерійна варіація:

$$\Delta = t\sqrt{\frac{\delta^2}{m}}(1-\frac{m}{M}), \qquad (11.6)$$

де δ^2 - міжсерійна дисперсія;
m - число серій вибіркової сукупності;
M - число серій генеральної сукупності.

11.4. Визначення обсягу вибірки

У процесі проектування вибіркових спостережень визначають мінімально достатній обсяг вибірки, при якому вибіркові оцінки репрезентували б основні властивості генеральної сукупності. Занадто великий обсяг вибірки потребує зайвих витрат, а занадто малий призведе до збільшення помилки репрезентативності.

Для визначення необхідної чисельності вибірки застосовують такі формули:

Таблиця 11.1 - Формули для визначення потрібного обсягу вибіркової сукупності

Спосіб відбору	Чисельність вибірки	
	для середньої ($\bar{x}$)	для частки (w)
Повторний	$n = \frac{t^2\sigma^2}{\Delta^2}$	$n = \frac{t^2 w(1-w)}{\Delta^2}$
Безповторний	$n = \frac{t^2\sigma^2 N}{\Delta^2 N + t^2\sigma^2}$	$n = \frac{t^2 \cdot w(1-w)N}{\Delta^2 N + t^2 w(1-w)}$

Для визначення обсягу вибірки використовують оцінки дисперсій аналогічних пробних обстежень. Якщо такі обстеження відсутні, - можна скористатися співвідношенням $\sigma^2 = 16(x_{\max} - x_{\min})$, а для частки взяти найбільше значення дисперсій σ^2 = 0,25.

Питання для самоконтролю

1. 1. Яке спостереження називають вибірковим?
2. У чому переваги вибіркового методу порівняно з іншими методами статистичних досліджень?
3. Що означає помилка репрезентативності? Які фактори визначають її величину?

4. Чому принцип випадкового відбору є визначальним при формуванні вибіркової сукупності? Які способи відбору забезпечують додержання цього принципу?

5. У чому різниця між систематичною помилкою спостереження та помилками репрезентативності? Чи можна уникнути їх появи? Зменшити?

6. Як визначається довірчий інтервал для генеральної середньої та частки?

7. Від чого залежить точність оцінювання параметрів генеральної сукупності (генеральної середньої та генеральної частки)?

8. Чим відрізняється величина стандартної помилки власне випадкової при повторному та безповторному відборі? Яка з цих помилок більша?

Література [4 с.149-168, 5 с.268-297].

ТЕМА 12. ПОДАННЯ СТАТИСТИЧНИХ ДАНИХ: ТАБЛИЦІ, ГРАФІКИ, КАРТИ

План лекції

12.1. Статистичні таблиці

12.2. Графічне зображення абсолютних та відносних величин

12.1. Статистичні таблиці

Більш ефективною формою подання статистичних даних є таблиці.

Статистична таблиця - форма найбільш раціонального, наочного та систематизованого викладення результатів зведення і групування статистичних даних.

Статистична таблиця містить **підмет** – об’єкт (окремі одиниці чи групи) і **присудок** – показники, що характеризують підмет.

За допомогою таблиць простіше провести порівняння та аналізувати зведені дані.

Складену, але не заповнену цифрами таблицю називають *макетом таблиці*.

За логічним змістом, статистична таблиця розглядається як “статистичне речення”.

Складові елементи досліджуваного об'єкта, які утворюють

статистичний підмет, розташовують у лівій частині таблиці, а показники, що формують статистичний присудок. - у правій. Однак трапляються випадки, коли статистичний підмет частково або повністю переноситься на місце статистичного присудка. В складеній і оформленій статистичній таблиці мають бути загальний, бічні та верхні заголовки.

Загальний заголовок розміщується над таблицею і виражає стисло її зміст. Розташовані зліва бічні заголовки розкривають зміст рядків підмета, а верхні - зміст граф (статистичного присудка). Одиниці вимірювання даних таблиці зазначаються або в загальному заголовку (якщо вони однакові для всіх показників таблиці), або у внутрішніх заголовках рядків і граф таблиці. До деяких таблиць подаються примітки, в яких роз'яснюється зміст окремих показників або заголовків.

Залежно від структури підмета, статистичні таблиці поділяють на: прості; групові; комбінаційні.

Проста– підметом простої таблиці є перелік одиниць статистичного спостереження, територіальний або хронологічний ряд.

Підметом простої таблиці є перелік елементів сукупності - називаються спискові (підмет — список окремих ознак або об'єктів), територіальний ряд (регіони, області) - територіальні (підмет — перелік районів, областей, країн, територій), хронологічний ряд - хронологічні (підмет — періоди часу: місяці, квартали, роки, або моменти часу, дати) (табл. 12.1).

Таблиця 12.1 - Розподіл за територією чисельності населення окремих регіонів України

Область	Площа, тис.км^2	Чисельність населення, тис. осіб	Щільність населення, осіб
Запорізька	27,2	4622,9	174
Дніпропетровська	31,9	3447,2	108
Донецька	26,5	1861,9	68
Разом	85,6	9932,0	116

Групова– підметом групової таблиці є групування за однією ознакою (табл.12.2).

Найпростішим видом групових таблиць є ряди розподілу, побудовані як за атрибутивними, так і за кількісними ознаками. У присудку таких таблиць міститься лише один показник — кількість

одиниць сукупності, що входять до кожної групи.

У складніших таблицях присудок доповнюється рядом інших показників, що характеризують підмет. Групові статистичні таблиці дають більш інформативний матеріал для аналізу досліджуваних явищ.

Таблиця 12.2- Капітальні інвестиції Запорізької області за джерелами фінансування

(у фактичних цінах)

	1 рік		2 рік	
	млн. грн.	у % до загального обсягу	млн. грн.	у % до загального обсягу
Всього	3828	100,0	4545	100,0
у тому числі за рахунок коштів державного бюджету	188	4,9	233	5,1
коштів місцевих бюджетів	174	4,5	189	4,2
власних коштів підприємств та організацій	2988	78,0	3497	76,9
коштів іноземних інвесторів	6	0,2	31	0,7
кредитів банків	382	10,0	480	10,6
інвестиційних фондів	2	0,1	8	0,2
інших джерел фінансування	88	2,3	107	2,3

Комбінаційна– підметом комбінаційної таблиці є групування за двома і більше ознаками.

Наприклад, найпростішою комбінаційною таблицею є така, що містить розподіл населення водночас за місцем проживання і статтю.

Складнішими є таблиці, частина статистичного підмета яких розташовується на місці статистичного присудка. Прикладом такої таблиці є розподіл вантажних автомобілів за тривалістю та пробігом (табл.12.3).

Таблиця 12.3 -Питома вага автомобілів за тривалістю експлуатації і пробігом

Загальний пробіг автомобілів, тис.км	Тривалість експлуатації, %				Разом
	до 3-х років	від 3-х до 8-ми включно	від 8-ми до 10-ти включно	понад 10 років	
До 100	28,4	1,0	-	-	29,4
101…250	4,4	15,1	1,9	-	21,4
251…400	-	14,8	3,4	0,8	19,0
Понад 400	-	9,3	7,0	13,9	30,2
Разом	32,8	40,2	12,3	14,7	100,0

Правила побудови таблиць

Побудова таблиці починається з макета таблиці.

1. Таблиця має містити лише ту інформацію, яка безпосередньо характеризує об'єкт дослідження. Слід уникати :зайвої, другорядної інформації.

2. Назва таблиці, заголовки рядків і граф мають бути чіткими, лаконічними, без скорочень. У назві вказується об'єкт, його часова і географічна ознаки. Якщо назви окремих граф (рядків) повторюються, мають однакові терміни або однаковий зміст, то їх доцільно об'єднати спільним заголовком.

3. У верхніх і бічних заголовках вказують одиниці вимірювання з використанням загальноприйнятих скорочень (т, гри.), іноді для них відводиться окрема графа.

4. Якщо одиниця вимірювання спільна для всіх даних таблиці, її зазначають над таблицею.

5. Рядки та графи доцільно нумерувати. При цьому графу з назвою підмета позначають літерою алфавіту, інші графи — номерами.

6. Інформація, що міститься в рядках (графах) таблиці узагальнюється підсумковим рядком “Разом” або “В цілому по сукупності”, який завершує статистичну таблицю; якщо підсумковий рядок розміщується першим, то деталізація його подається за допомогою словосполучення “у тому числі”.

7. Числа, по можливості, необхідно округляти у межах одного і того самого порядку чи графи обов'язково з однаковим ступенем точності.

8. Відсутність даних у таблиці позначається відповідно до причин:

а)якщо клітинка таблиці, передусім підсумкова, не може бути заповнена, ставиться знак “х”;

б)коли відомості про явище відсутні, ставиться три крапки “...” або“н. від.”;

в)відсутність самого явища позначається тире (“—”);

г)дуже малі числа записуються (0,0) або (0,00).

8. Якщо потрібна додаткова інформація, певні уточнення цифрових даних, до таблиці додається примітка.

12.2. Графічне зображення абсолютних та відносних величин

Статистичний графік - спосіб наочного зображення і узагальнення статистичних даних про соціально-економічні явища за допомогою геометричних образів.

Кожний графік, як правило, складається з таких основних елементів: шкала; масштаб; координатна сітка; заголовок; цифрові дані; пояснення умовних позначень.

Шкалою називається лінія, окремі позначки якої відповідають певним числам.

Масштаб — це довжина відрізка шкали, прийнята за числову одиницю.

Координатна сітка - прямокутна система координат, де на горизонтальній осі (вісь абсцис) відкладається час, а на вертикальній (вісь ординат) - кількісні показники за масштабом.

Експлікація графіка— це словесні пояснення, що розкривають його зміст та основні елементи — заголовок, одиниці виміру та умовні позначення показників, які представляються на графіку.

За способом побудови графіки поділяють на **діаграми та картограми**.

Діаграми — це графіки, в яких цифрові дані зображені за допомогою геометричних фігур (стовпчикові, секторні, лінійні та інші — фігурні діаграми, діаграми фігур-знаків).

Розглянемо особливості побудови та використання окремих видів графіків.

Найбільш простим видом графіків **є лінійні діаграми,** які широко використовуються для: характеристики динаміки (оцінки зміни явища в часі); характеристики варіації в рядах розподілу; оцінки виконання планових завдань; оцінки взаємозв'язку між явищами.

Лінійні діаграми будують у прямокутній системі координат. Зміна статистичних явищ у часі ілюструють динамічні графіки.

Лінійні графіки характеризують зміну явищ у часі (рис. 12.1).

Другим поширеним видом графіків є **стовпчикові діаграми (гістограми).**

Стовпчикові діаграми - статистичні дані зображуються у вигляді стовпчиків.

(грудень до грудня попереднього року; відсотків)

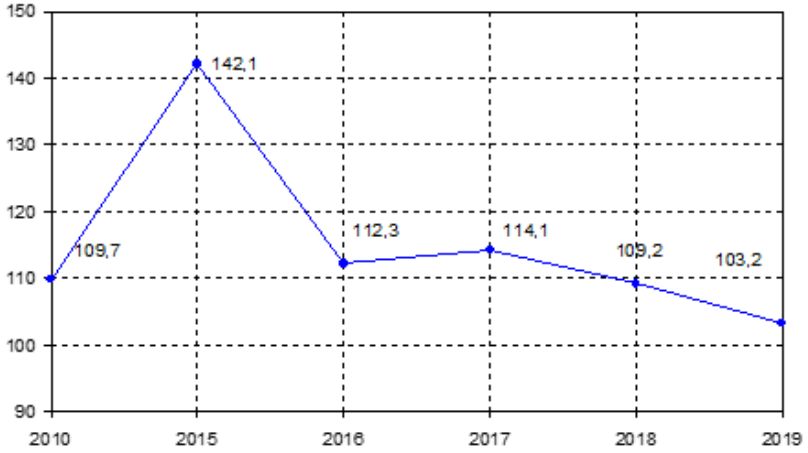

Рисунок 12.1- . Індекси споживчих цін

Стовпчикові діаграми використовуються для таких же цілей, як і лінійні діаграми (рис. 12.2).

(відсотків)

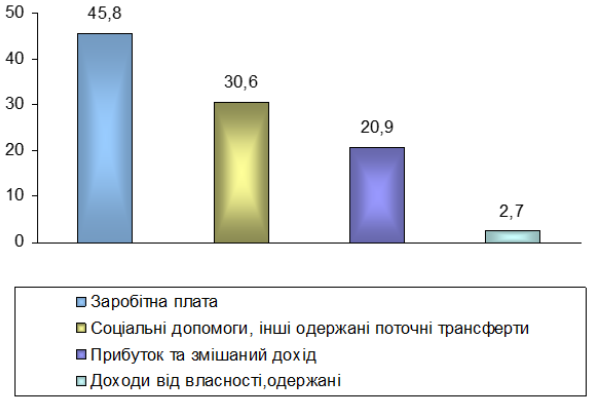

Рисунок 12.2 - Структура доходів населення у 2018 році

Для характеристики структури досліджуваних явищ значне

поширення одержали **секторні діаграми**.

Структурні діаграми - діаграми співвідношення питомої ваги, які характеризують відношення окремих частин сукупності у загальному їх обсязі (рис. 12.3).

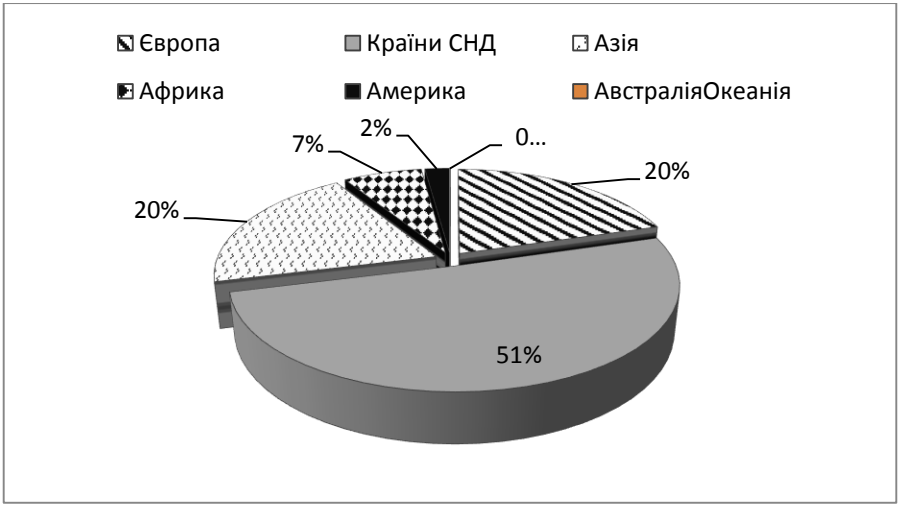

Рисунок 12.3 - Структура експорту товарів Запорізької області

Крім розглянутих видів діаграм, для цілей порівняння за регіонами, країнами, галузями іноді використовують **квадратні, кругові, фігурні діаграми (діаграми фігур-знаків).**

Для побудови **квадратної діаграми**, яка також використовується при порівняльному аналізі, необхідно знайти квадратні корені із значень порівнюваних величин статистичних показників, а потім побудувати квадрати із сторонами, пропорційними одержаним результатам.

При побудові **кругової діаграми** значення порівнюваних статистичних показників спочатку ділять на число п, тобто 3,14. Потім з одержаних величин добувають квадратні корені та будують круги, що мають радіуси, пропорційні одержаним результатам.

Діаграми фігур-знаків являють собою графічне зображення у вигляді рисунків, силуетів, фігур, які відповідають змісту статистичних даних.

Для оцінювання та порівняльного аналізу географічного

розміщення явищ та статистичних показників за територіями широко використовуються **статистичні карти**.

Статистичні карти поділяються на **картограми та картодіаграми**.

Картограма показує територіальний розподіл досліджуваного показника (явища) за окремими районами умовно (штрихуванням, зафарбовуванням, нанесенням точок). З огляду на це, картограми поділяють на два види: **фонові та точкові.**

При побудові **фонових картограм** основним питанням є групування даних за величиною ознаки (не більше 5—8 груп). Для кожної групи встановлюють певне фарбування або штрихування. При цьому додержуються такого правила: чим більша величина досліджуваної ознаки, тим інтенсивнішим повинно бути штрихування (фарбування інтенсивнішим кольором).

При побудові **точкових картограм** кожній точці відповідає певне однакове число, наприклад 100 тис.

На контур зображуваного району (території, міста, країни) наносять таку кількість точок, скільки потрібно для кількісної характеристики зображуваного явища в цьому районі.

Картодіаграма — це графічне зображення закономірностей просторового розміщення статистичних показників.

Питання для самоконтролю

1. Дайте визначення статистичної таблиці та назвіть її основні елементи.
2. Що називається підметом статистичної таблиці? Яке місце в таблиці займає підмет?
3. Що розуміють під присудком статистичної таблиці? Як у таблиці представляється присудок?
4. Назвіть особливості побудови та застосування різних видів статистичних таблиць?
5. Назвіть основні правила побудови статистичних таблиць.
6. Що таке статистичний графік? Які види графіків використовуються для наочного представлення статистичної інформації?
7. Назвіть основні елементи статистичних графіків. Дайте їх визначення.

8. Які види діаграм використовуються для представлення статистичних даних? Вкажіть особливості їх побудови.

9. Назвіть відомі вам види статистичних карт. У чому їх особливості.

Література [4 с.62-82, 5 с.45-81].

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Про державну статистику : Закон України від 17 вер. 1992 р. № 2614-XII Режим доступу : https://zakon.rada.gov.ua/laws/show/2614-12.

2. Глосарій до плану статистичного спостереження. Режим доступу:.http://www.ukrstat.gov.ua/

3. Концепція побудови національної статистики України. [Електронний ресурс]. Режим доступу:http://zakon1.rada.gov.ua/laws/show/326-93-%D0%BF.

4. Бек В. Л. Теорія статистики : курс лекцій. навчальний посібник Київ: ЦУЛ, 2003. 288 с..

5. Теорія статистики: навчальний посібник / Вашків П. Г., Пастер П.І., Сторожук., Ткач Є. І. 2-ге вид., стереотип. Київ : Либідь, 2004.- 320с.

6. Тарасенко І. О. Статистика: навчальний посібник. Київ : Центр навчальної літератури, 2005. 344 с.

7. Уманець Т. В. Загальна теорія статистики : навчальний посібник.Київ : Знання, 2006. 239 с.

8. Щурик М.В. Статистика : навчальний посібник. 2-ге вид., оновл. і доп. Львів: «Магнолія - 2006», 2011. 545 с.

9. Запорізька область за 2019 рік : статистичний щорічник / За редакцією М. Шейко. Головне управління статистики у Запорізькій області : Запоріжжя, 2020. 364 с.