

**Мукачівський державний університет
Факультет економіки, управління та інженерії
Кафедра фінансів**



ОПОРНИЙ КОНСПЕКТ ЛЕКЦІЙ

з дисципліни

«СТАТИСТИКА»

**для студентів
напрямів підготовки**

6.030508 «Фінанси і кредит»

6.030509 «Облік і аудит»

6.030507 «Маркетинг»

денної та заочної форм навчання

**Мукачево
2016**

Опорний конспект лекцій з дисципліни «статистика» для студентів напрямів підготовки 6.030508 «Фінанси і кредит», 6.030509 «Облік і аудит», 6.030507 «Маркетинг» денної та заочної форм навчання / Укладач М.І. Стегней. – Мукачево: МДУ, 2016.– 36 с. (1,1 авт.арк).

Рекомендовано до друку Науково-методичною радою Мукачівського державного університету, протокол №7 від «23» лютого 2016 р.

Обговорено і схвалено на засіданні кафедри фінансів, протокол № 8 від «7 » грудня 2015 р.

Укладач: М.І. Стегней, доктор економічних наук, професор,
Мукачівський державний університет

Відповідальний

за випуск: Т.Д. Черничко, зав. кафедрою фінансів
доктор економічних наук, професор,
Мукачівський державний університет

Рецензент: В.В. Гоблик, доктор економічних наук,
Мукачівський державний університет

В опорному конспекті лекцій наведено короткий зміст лекційних тем з нормативної навчальної дисципліни циклу природничо-наукової та загальноекономічної підготовки «Статистика».

ЗМІСТ

	Стр.
Передмова	4
Предмет та методи статистичної науки	5
Статистичне спостереження	6
Зведення і групування статистичних даних	9
Статистичні таблиці	12
Графічний метод у статистиці	13
Узагальнюючі статистичні показники	15
Статистичне вивчення варіації	21
Вибіркові спостереження	23
Статистичні методи вивчення взаємозв'язків	27
Ряди динаміки	28
Статистичні індекси	31
Список рекомендованої літератури	35

ПЕРЕДМОВА

Статистика виявляє кількісні зміни масових суспільних явищ. Основним прийомом в статистичному дослідженні є формування і вивчення узагальнюючих статистичних показників. Такі показники, з одного боку, виступають предметом дослідження, а з іншого – служать спеціальним прийомом дослідження. Статистичні показники відображають суспільно-економічні явища, які знаходяться в складних взаємозв'язках. Взаємозв'язок явищ вимагає сумісного застосування різних статистичних методів і показників. Абсолютні, відносні, середні показники, показники структури і динаміки, взаємозв'язані системи показників, що їх доповнюють, в дослідженні повинні застосовуватися комплексно.

Тому основним завданням розроблених методичних рекомендацій є

- формування у студентів навичок розрахунків та аналізу статистичних показників,
- вимірювання об'ємів і рівня, структурних зрушень, динаміки і зв'язку суспільних явищ, чинників, що впливають на загальну динаміку та величини їх впливу.

Подальше вдосконалення методології розрахунку показників, розширення застосування статистичних методів при соціально-економічній оцінці організаційно-технічних і господарських рішень та в цілому в економічній практиці є важливою передумовою підвищення рівня управління суспільним виробництвом.

Мета вивчення курсу: формування знань щодо методів збирання, обробки та аналізу інформації про соціально-економічні явища і процеси.

Завдання вивчення курсу: вивчення принципів організації статистичних

спостережень, методик розрахунків показників статистичного аналізу соціально-економічних явищ і процесів.

Предмет курсу: розміри й кількісні співвідношення масових явищ і процесів у економіці.

Предмет і методи статистики

Статистика - суспільна наука, яка обіймає такі сфери:

- статистичну інформацію, отриману шляхом спостережень;
- статистичну практику (діяльність статистичних установ, які збирають і аналізують цю інформацію);
- статистичну науку.

Предметом статистики є розміри і кількісні співвідношення масових суспільних явищ, закономірності їх формування і розвитку. Важливо наступне: статистика вивчає кількісну сторону суспільних явищ і не одинокі, а масові явища. Кількісну сторону явищ статистика відображає за допомогою показників. Вивчення кількісної сторони суспільних явищ нерозривно пов'язано з якісним їх змістом, так як кількісна розмірність не існує без якісної визначеності в конкретних умовах простору і часу;

Статистика має наступні складові частини:

- загальну теорію статистики;
- економічну статистику;
- галузеві статистики (промисловості, будівництва, сільського господарства, тощо);
- соціальну статистику.

Основними в статистиці є поняття статистичної закономірності і статистичної сукупності. Статистичним еквівалентом притаманних елементам сукупності властивостей є ознака. Ознака, що приймає в межах сукупності різні значення, називається варіюючою, а коливання значень ознаки - варіацією. Ознаки є кількісні і атрибутивні (якісні). Кількісні ознаки вимірюють використовуючи еталони і одиниці виміру. Для атрибутивних ознак вимірювання - це реєстрація наявності (відсутності) властивості, що вивчається. набір властивостей явища і відповідних їм чисел називають шкалою вимірювання: матричною, номінальною чи ранговою.

Статистична методологія - це сукупність засобів дослідження, які ґрунтуються на філософських (логіка, діалектика) і наукових принципах (аналіз, синтез, порівняння). Статистичне дослідження має такі етапи:

1. Збір первинних даних (статистичне спостереження).
2. Систематизація і групування цих даних.
3. Обчислення узагальнюючих показників.
4. Аналіз варіації, динаміки, взаємозв'язків.

Найуживаніше поняття статистики - показник, це узагальнююча характеристика явища, в якій поєднується його якісна і кількісна визначеність; якісна визначеність відображається назвою показника.

Показник складається:

- з моделі кількісної сторони явища;
- числового значення.

Модель встановлює: "що, де, коли, яким чином вимірюється". Найважливішою характеристикою показника є його вірогідність, яка має два аспекти: адекватність і точність.

Статистичне спостереження

Статистичне спостереження означає збір статистичної інформації на початковій стадії статистичного дослідження.

Статистична інформація - це сукупність статистичних даних, які відображають хід соціально-економічних процесів і використовуються в управлінні економіки суспільством. Статистичне спостереження - це спланований систематичний науково-організований збір масових даних про суспільно-економічні явища і процеси. Основне завдання одержання точних і безсумнівних фактів.

Уявлення про форми і види спостереження дає рис. 1. Перед проведенням статистичного спостереження складають його план, який має дві частини: програмно-методологічну і організаційну.

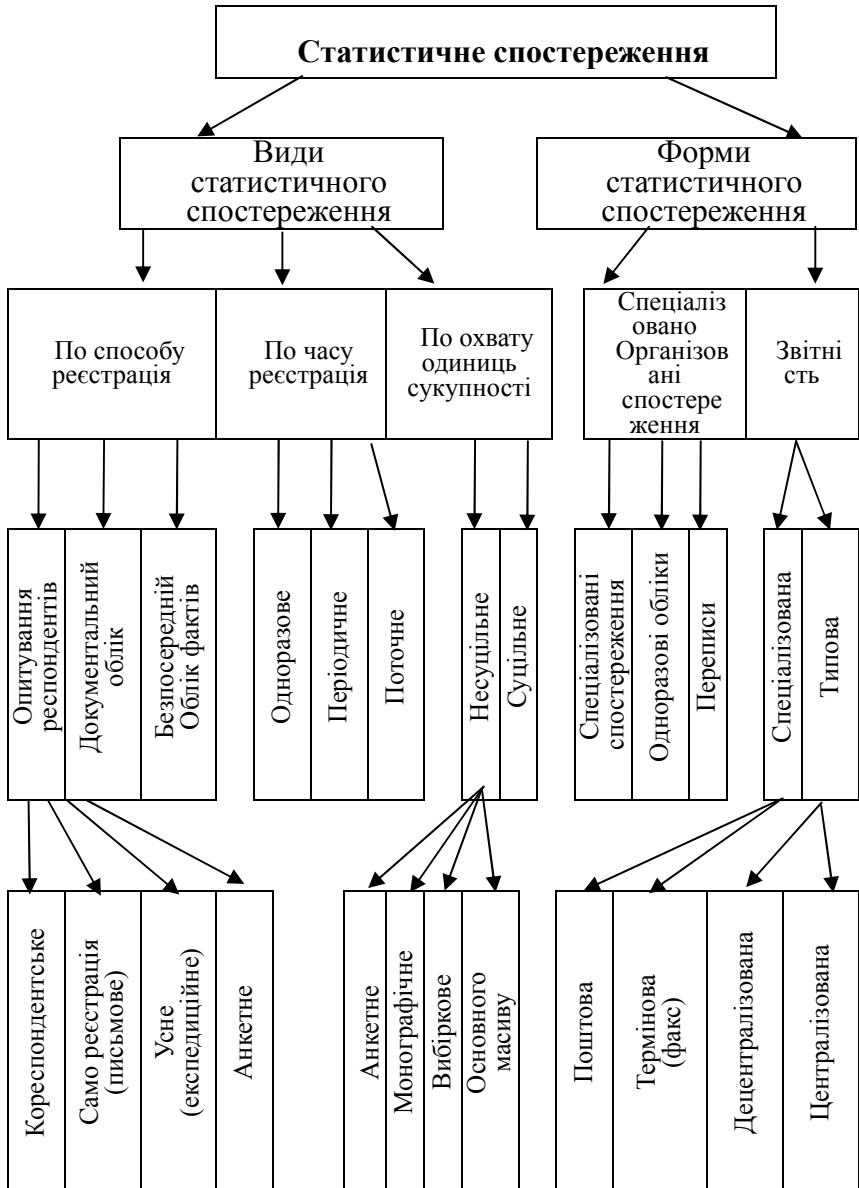


Рис. 1. Форми і види статистичних спостережень.

Програмно-методологічна частина плану визначає мету спостереження: встановлення об'єкта, одиниці, елемента спостереження: складання програми спостереження.

Мета спостереження визначається потребами статистичних даних і формується чітко і конкретно урахуванням завдань спостереження. Об'єкт спостережень - це сукупність явищ, які вивчаються. Важливо чітко знайти його межі, риси, істотні ознаки. Одиниця спостереження джерелом інформації (підприємство, установа). Елемент спостереження є носієм ознаки, яка підлягає реєстрації. Програма спостереження дає перелік ознак, які підлягає реєстрації. Запитання програми містяться у статистичних формулярах з адресною частиною, яка відображає місце збору даних.

Організаційна частина плану статистичного спостереження визначає: місце, час, органи спостереження, графік підготовки і інструктаж кадрів, матеріально-технічну базу спостереження.

Місце спостереження - пункт, де реєструю одиниць сукупності у формулярах. Час спостереження поділяють на об'єктивний і суб'єктивний. Об'єктивним називають час, до якого відносяться дані спостереження, - це певний момент або період. Момент часу, на який проводиться реєстрація ознаки, називається критичним.

Період, протягом якого реєструються ознаки спостереження, називається суб'єктивним часом.

Важливою вимогою до спостережень є їх точність і вірогідність. Точність - це міра відповідності даних спостереження дійсній їх величині. Вірогідність - міра об'єктивного відображення статистичним спостереженням суті явища. Розбіжності між даними спостереження і дійсним "значенням показників називають помилками спостереження. Розрізняють помилки реєстрації і репрезентативності. Помилки реєстрації - ті, які виникли внаслідок неправильного запису або встановлення факту. Вони діють випадково або систематично. Випадкові помилки не є небезпечними, оскільки вони мають різні знаки, тому їх вплив гаситься при масових спостереженнях.

Систематичні помилки виникають внаслідок нечіткого формулювання програми спостереження і можуть привести до значного спотворення загальних підсумків. Помилки репрезентативності виникають лише при несу цільному спостереженні і тоді, коли вибірка неправильно відтворює структуру досліджуваної сукупності. Помилки спостереження виявляються внаслідок зовнішнього контролю, логічного контролю і арифметичного контролю підсумків.

Зведення і групування статистичних даних

Статистичне зведення і групування є другим етапом статистичного дослідження. Статистичним зведенням називають: систематизацію, упорядкування, наукову обробку статистичних даних. Зведення буває просте і складне. Просте - просто підсумок первинного статистичного матеріалу. Складне передбачає групування, підрахунок групових і загальних підсумків і викладання зведення у вигляді статистичних таблиць і графіків. Групування - це розподіл сукупності на групи за істотними для них ознаками. При використанні статистичних групувань вирішують наступні задачі:

- визначають типові риси і відмінності груп;
- визначають істотні ознаки групування;
- визначають інтервали групування.

Групування використовують для вирішення таких проблем:

- визначення структури і структурних зрушень;
- виявлення соціально-економічних типів, явищ;
- дослідження взаємозв'язку ознак.

Відповідно до цих завдань групування поділяють на структурні, типологічні, аналітичні.

Структурні - характеризують розподіл якісно-однорідної сукупності на групи за певною ознакою. За допомогою типологічних групувань виділяють найхарактерніші групи, типи

явищ, які складають неоднорідну статистичну сукупність. Аналітичні – використовують при вивченні взаємозв'язків, найчастіше між двома ознаками. Одна з ознак відображає величину, інша – наслідок. Відповідно до цього розрізняють: факторну і результативну ознаку. Аналітичні групування завжди проводять за факторною ознакою і в кожній групі визначається середнє результативної ознаки.

Першочерговим питанням теорії групувань є вибір групувальних ознак. Групування проводиться або за атрибутивними, або за кількісними ознаками. Для атрибутивної ознаки число груп відповідає числу її різновидів. При групуванні за варіаційною (кількісною) ознакою постає питання щодо кількості груп і ширини інтервалів групування; останні можуть бути рівні і нерівні. Нерівні можуть прогресивно зменшуватись (збільшуватись). Ширину рівних інтервалів визначають за формулою:

$$h = \frac{(x_{\max} - x_{\min})}{M} \quad (1)$$

де x_{\max} , x_{\min} – найбільше і найменше значення ознаки

в сукупності; M – кількість груп. Орієнтовно число груп визначають за формулою:

$$M = 1 + 3,332 \lg n \quad (2)$$

де n – обсяг сукупності. Якщо групування проводять за одною ознакою, то воно є простим, за кількома – складними. Останнє може бути комбінаційним, якщо в його основі скомбіновано дві і більше ознак, або багатомірним, якщо воно проводиться за кількома ознаками одночасно.

Схематично всі види групувань і зведень вказані на рис. 2.

Важливим поняттям цієї теми є: ознаки (факторні, результативні, кількісні, атрибутивні) інтервали (відкриті, закриті, рівні, нерівні).

Після цього можна перейти до рядів розподілу, які є основою будь-якого групування. Ряд розподілу складається з двох елементів: варіантів і частот. Варіантами значення групувальної ознаки; частоти – числа, що показують як часто повторюються окремі варіанти. Замість частот може бути частка, виражена коефіцієнтом чи відсотком. Накопичену частоту (частку) називають кумулятивною. Залежно від статистичної природи групувальної ознаки (атрибутивна чи кількісна) ряди розподілу поділяють на атрибутивні і варіаційні. Останні поділяють на дискретні і інтервальні.

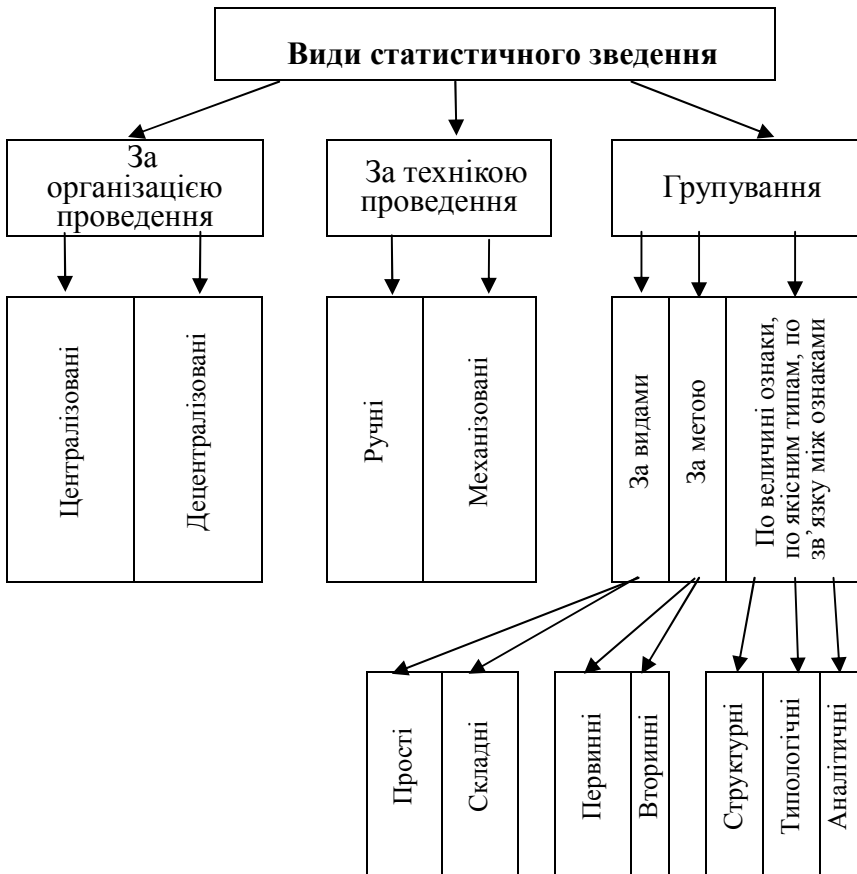


Рис. 2. Види зведень і групувань

Статистичні таблиці

Результати зведення статистичних даних зображаються у вигляді статистичних таблиць. Не всі таблиці є статистичними, наприклад, таблиці логарифмів, вигравів в лотерею. Статистичними є лише ті таблиці, які містять наслідки статистичного аналізу соціально-економічних процесів.

У статистичній таблиці розрізняють підмет і присудок. Підмет таблиці - це ті об'єкти, які характеризують числовими показниками. Самі ж показники, які характеризують ці об'єкти її - присудок. Обов'язковим атрибутом статистичної таблиці є її загальний і внутрішній заголовки: загальний - чітко і коротко характеризує її зміст, внутрішні розміщаються збоку і зверху. У бічних заголовках розкривається зміст підмета, у верхніх - зміст присудка. Найголовнішими правилами складання статистичних таблиць є такі:

- таблиця має включати лише необхідні дані, і по можливості, має бути невеликою;
- якщо число показників присудка велике, то їх нумерують арабськими цифрами; графи з переліком об'єктів позначають великими літерами;
- при відсутності відомостей про розмір явища у відповідній клітинці пишуть: "немає відомостей " або ставлять крапки (...); відсутність явища позначають (—); число 0.0 ставлять якщо величина показника не більше 0.05; x-клітинка не заповнюється;
- кількісні показники в межах графи наводяться з однаковою точністю;
- якщо показники мають різні одиниці виміру, то їм виділяють спеціальну графу;
- таблиці мають бути замкненими (за винятком аналітичних таблиць, в яких підсумки не обов'язкові). За побудовою підмета таблиці поділяють на: прості, групові і комбінаційні.

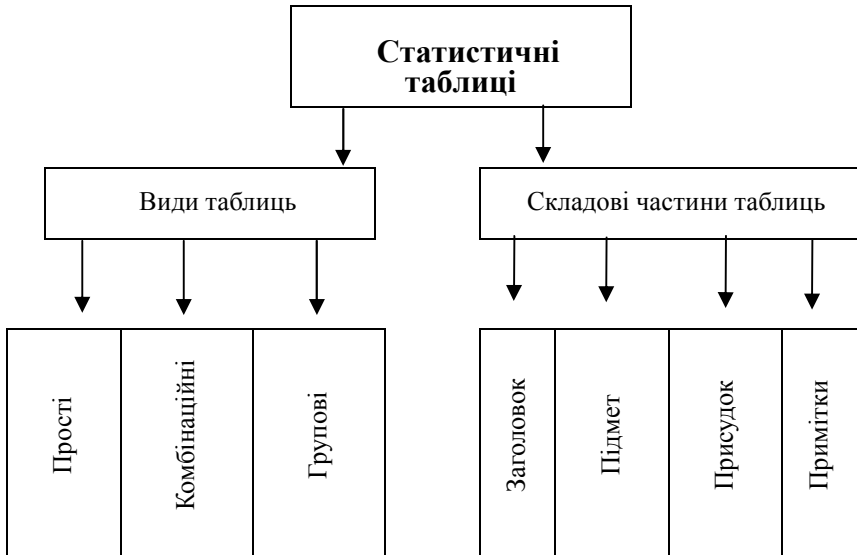


Рис. 3. Види і складові частини таблиць.

Схема на рис. 3 дає відомості про види і складові частини таблиць

Графічний метод у статистиці

Статистичний графік - це спосіб наочного подання і викладення статистичних даних за допомогою геометричних знаків та інших графічних засобів з метою узагальнення їх аналізу. Графічній мові притаманні важливі риси як лаконічність, змістовна однозначність тлумачення символічного запису, відносна простота, кодування. Разом з тим, графічна мова має такі особливості: двомірність запису, неперервність виразу, відокремленість викладу.

Поле графіка - простір, у якому розміщуються геометричні або інші графічні знаки і який обмежений аркушем чи картою.

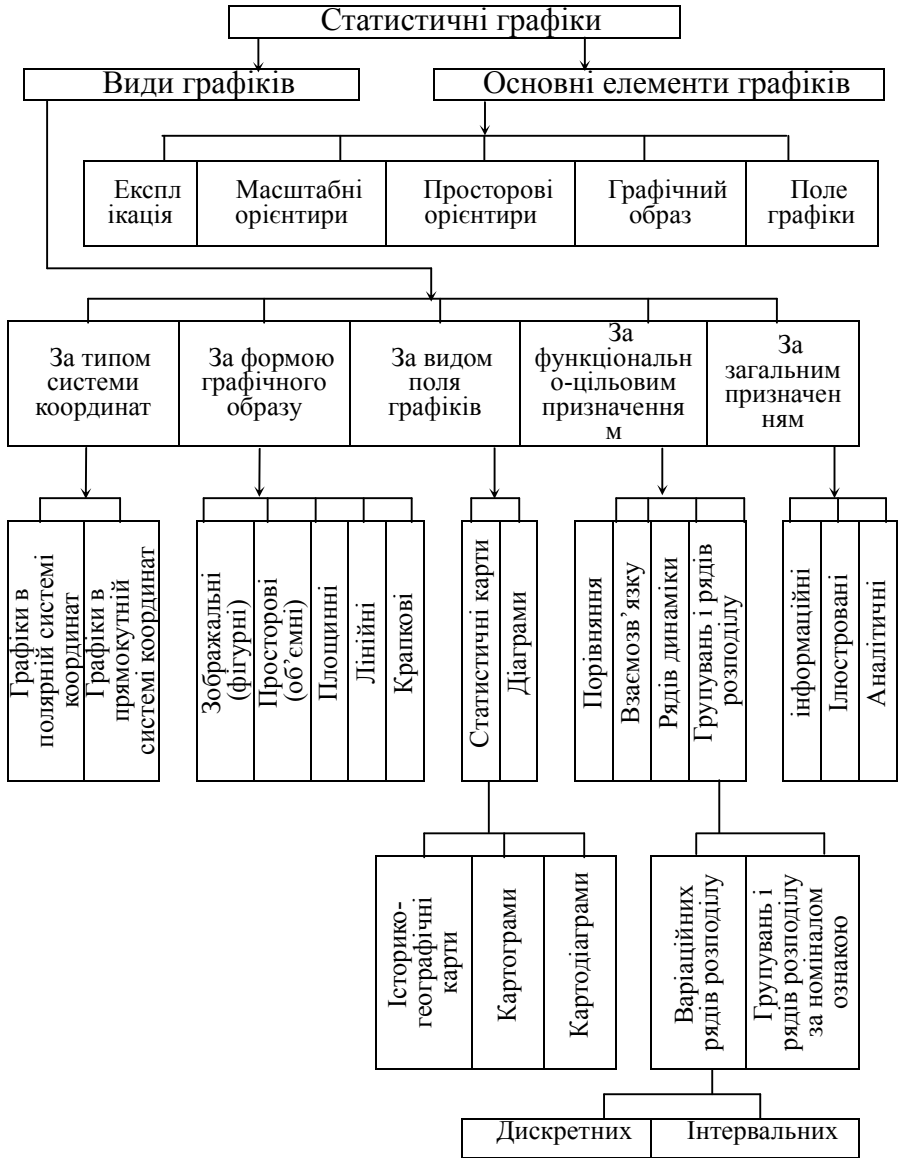


Рис. 4. Основні елементи і види графіків

Графічний образ - сукупність геометричних або графічних знаків, що зображують статистичні дані. Це є мова графіка, його основа. Вибір графічного образу залежить від того, який аспект має висвітлити графік. Наприклад, стовпчики точніше відображають зміну обсягів видобутку, а відрізки прямих - тенденцію динаміки.

Просторові орієнтири використовують для визначення порядку розміщення графічних знаків у полі графіка. Цей порядок визначається характером і особливостями статистичних даних, а також завданнями аналізу інтерпретації їх і задається системою координат: прямокутною чи полярною.

Масштабні орієнтири - масштаб, масштабна шкала і масштабний знак.

Масштаб - це умовна міра переводу числового значення статистичної величини у графічну, і навпаки.

Масштабна шкала - лінія, поділена відповідно до прийнятого масштабу.

Масштабні знаки - це знаки-еталони, за допомогою яких зображують статистичні величини у вигляді квадратів, кругів, силуетів.

Експлікація графіка - це словесні пояснення його змісту і основних елементів, яка включає в себе загальний заголовок, підписи вздовж масштабних шкал, пояснювальні підписи, які розкривають зміст елементів графіка.

Класифікація графіків по видам показана на рис. 4.

Узагальнюючі статистичні показники

Статистичне дослідження, незалежно від його масштабу і мети, завершується розрахунком і аналізом різних по виду і формі статистичних показників.

Показник є узагальнюючою характеристикою суспільних явищ, в якій поєднується їх якісна і кількісна визначеність.

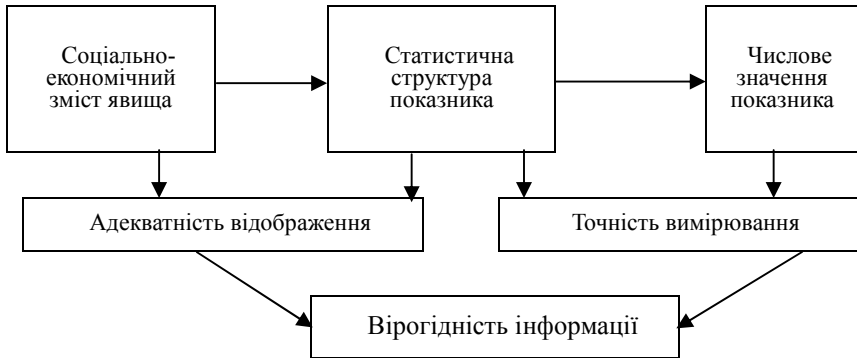


Рис. 5. Статистична модель показника

Статистичний показник складається з моделі кількісної сторони явища (яке вивчається) і числового значення. Модель встановлює: "що, де, коли, яким чином вимірюється" в конкретних умовах простору і часу. Важливою характеристикою показника є його вірогідність, що має два аспекти - адекватність і точність. На рис. 5 дана модель показника, що поєднує якісну і кількісну визначеність економічних категорій і точність їх вимірювання.

Статистичні показники можуть мати форму абсолютних або відносних величин; перші відображають розміри явищ іменованими числами (тн, л, м). Іменовані числа - це вимірники ознак. Є три групи вимірників: натуральні (умовно-натуральні), трудові, вартісні. Перерахунок натуральних величин в умовно-натуральні здійснюється за формулою:

$$y = c + \sum_{i=1}^n k_i \cdot x_i \quad (3)$$

де y - обсяг в умовно-натуральних одиницях; c - обсяг різновиду, прийнятого за еталон; x_i - обсяг i -того не еталонного різновиду; k_i - коефіцієнт перерахунку.

Трудові вимірники (людино-годин, людино-днів) використовують для обчислень витрат праці.

Вартісні - дають змогу узагальнити і зіставити різноманітні явища.

Відносні величини - відображають кількісні співвідношення соціально-економічних явищ і є часткою від ділення двох однойменних або різнойменних величин. Співвідношення різнойменних абсолютних величин дає відносну величину інтенсивності. Це іменована величина, в якій поєднуються одиниці виміру чисельника і знаменника (густота населення, виробництво продукції на душу населення).

Порівняння однойменних статистичних величин використовують для:

- характеристики структури сукупності;
- оцінки інтенсивності розвитку явищ;
- оцінки виконання плану;
- оцінки дотримання норм;
- порівняння.

Відповідно до цих завдань виділяють відносні величини: координації, динаміки, планового завдання і виконання плану; дотримання норм, порівняння. Нехай маємо сукупність, яка складається з кількох підмножин $d_1+d_2+d_3=d$. Тоді, відносні величини структури це відношення типу $d_i:d$, $i=1, 2, 3$; відносні величини координації - це відношення типу $d_i+d_{i=1}$.

Щоб пояснити відносні величини динаміки k_d , планового завдання $k_{пз}$, виконання плану, $K_{вп}$ позначимо фактичний рівень поточного періоду через Y_1 , базового - Y_0 , передбаченого планом $Y_{пл}$, тоді:

$$K_d = \frac{Y_2}{Y_1}; K_{пз} = \frac{Y_{пл}}{Y_0}; K_{вп} = \frac{Y_1}{Y_{пл}} \quad (4)$$

$$\text{Ясно, що} \quad K_d = K_{пз} * K_{вп} \quad (5)$$

Відносна величина дотримання норм - це співвідношення виду

$$K_n = U_f: U_n,$$

де U_f - фактичний рівень, U_n - норма.



Рис. 6. Види відносних величин і їх класифікація

Відносні величини порівняння обчислюють як співвідношення однойменних показників, що характеризують різні об'єкти, території і мають однакову часову визначеність. Інтерпретація цих величин залежить від бази порівняння, вибір якої довільний.

Класифікація різних типів відносних величин дана на рис. 6.

Найбільш поширеною формою статистичних показників є середня величина, яка являє собою узагальнену кількісну характеристику ознаки в статистичній сукупності в конкретних умовах простору і часу. Найважливіша властивість середньої

величини полягає в тому, що вона відображає те загальне, що властиве усім одиницям сукупності. Типовість середньої безпосереднім чином пов'язана з однорідністю статистичної сукупності. Середня величина лише тоді відображає типовий рівень ознаки, якщо вона обчислена по якісно однорідній сукупності.

Види середніх і формули для їх обчислення показані в табл. 1.

В табл. 1 позначено через: x_i – індивідуальні значення варіюючої ознаки; f_i - відповідні їм частоти; Π - добуток; n - обсяг варіант; x_0 - нижня межа модального (медіанного) інтервалу; f_m - частота модального (медіанного) інтервалу; f_{m-1} , f_{m+1} - частоти попереднього і наступного інтервалів відносно модального; S_{m-1} - кумулятивна частота інтервалу, що передує медіанному; h - ширина модального (медіанного) інтервалу.

Найважливішою середньою є середня арифметична, або просто середня. Зауважимо, що M_e ділить варіаційний ряд на дві однакові за обсягом частини.

Якщо статистичний ряд можна поділити на 4 рівні за обсягом частини, то варіанти, які відсікають 25%, 50% і 75% ранжированого ряду називаються квантилями: Q_1 -першим, Q_2 -другим, Q_3 - третім. Від Q_1 виключно до Q_3 включно розташовано 50% центральних елементів статистичного ряду.

При виборі форми середньої найчастіше використовують принцип вихідного співвідношення середньої (ВСС), або логічну формулу середньої. Тобто, перш ніж обчислювати середню, потрібно спочатку чітко уявити співвідношенням яких величин вона є.

$$\text{ВСС} = \frac{\text{Сумарне значення чи обсяг ознаки, яка осереднюється}}{\text{Число одиниць чи обсяг сукупності}}$$

Таблиця 1. Види середніх величин

Аналітичні середні			
Назва середньої	Формули		№ формули
	прості	зваженої	
Середня арифметична	$\bar{X} = \frac{\sum x_i}{n}$	$\bar{x} = \frac{n}{\sum \frac{1}{x_i}}$	6-7
Середня гармонійна	$\bar{X}_r = \frac{n}{\sum \frac{1}{x_i}}$	$\bar{X}_r = \frac{\sum w_i}{\sum w_i/x_i}$	8-9
Середня квадратична	$\bar{X}_k = \sqrt{\frac{\sum x_i^2}{n}}$	$\bar{X}_k = \sqrt{\frac{\sum x_i^2 f_i}{\sum f_i}}$	10-11
Середня геометрична	$\bar{X}_{re} = \sqrt[n]{\prod_{i=1}^n x_i}$		12
Структурні середні			
Мода	$M_0 = x_0 + h \frac{\Delta_1}{\Delta_1 + \Delta_2}$ $\Delta_1 = f_m - f_{m-1}$ $\Delta_2 = f_m - f_{m+1}$		13
Медіана	$M_e = x_0 + h \frac{0,5 \sum f_i - S_{m-1}}{f_m}$		14

Статистичне вивчення варіації

Коливання або мінливість величини ознаки у одиниць сукупності називають варіацією. Ступінь мінливості ознак характеризується за допомогою показників варіації (табл. 3).

Таблиця 3.

Основні показники варіації

Назва показника	Формула	№ формули
Розмах варіації	$R = x_{\max} - x_{\min}$	18
Дисперсія: проста	$\sigma^2 = \sum (x_i - \bar{x})^2 / n$	19
– зважена	$\sigma^2 = \sum (x_i - \bar{x})^2 \cdot f_i / \sum f_i$	20
Середнє квадратичне відхилення: просте	$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$	21
– зважене	$\sigma = \sqrt{\sum (x_i - \bar{x})^2 \cdot f_i / \sum f_i}$	22
Середнє лінійне відхилення: просте	$l = \sum x_i - \bar{x} / n$	23
– зважене	$l = \sum x_i - \bar{x} \cdot f_i / \sum f_i$	24
Коефіцієнти варіації: квадратичний	$v_{\sigma} = \frac{\sigma}{\bar{x}} \cdot 100\%$	25
– лінійний	$v_i = \frac{l}{\bar{x}} \cdot 100\%$	26

Назва показника	Формула	№ формули
– осциляції	$v_R = \frac{R}{\bar{x}} \cdot 100\%$	27
– кварталний	$v_R = \frac{0,5R_Q}{M_e} \cdot 100\%$ $R_Q = Q_3 - Q_1$	28
Дисперсія альтернативної ознаки	$\sigma^2 = p \cdot q$	29

В табл. 3 позначено через: x_{max} , x_{min} - максимальні і мінімальне значення варіюючої ознаки; x_i , f_i - варіанти і їх частоти; p – доля одиниць сукупності, що мають дану ознаку; q – доля одиниць, які цієї ознаки не мають; Rq – між квартална широта.

Нехай маємо сукупність розбиту на k груп з числом спостережень f_j в j -тій групі. Обчислимо позначенням y_i загальну дисперсію

Обчислимо позначенням y_i загальну дисперсію

$$\sigma_y^2 = \frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N} = \frac{\sum_{i=1}^N y_i^2}{N} - \bar{y}^2 \quad (30)$$

середню з групових дисперсій

$$\bar{\sigma}^2 = \frac{\sum_{j=1}^N \sigma_j^2 \cdot f_j}{\sum_{j=1}^N f_j} \quad (31)$$

міжгрупову дисперсію

$$\delta^2 = \sum_{j=1}^N (\bar{y}_j - \bar{y})^2 \cdot f_j / \sum_{j=1}^N f_j \quad (32)$$

де $\bar{y} = \sum_{i=1}^N y_i / N$ y_i - середня по j-тій групі.

Тоді правило складання дисперсій має такий вигляд

$$\sigma_y^2 = \bar{\sigma}^2 + \delta^2 \quad (33)$$

В статистичному аналізі широко використовується кореляційне відношення, яке визначається за формулою:

$$\eta^2 = \frac{\delta^2}{\sigma_y^2} \quad (34)$$

Цей коефіцієнт показує долю загальної варіації ознаки, яка обумовлена варіацією груповальної ознаки.

Вибіркові спостереження

Вибіркове спостереження — поширений вид несучільного спостереження, при якому співвідносяться дві сукупності: генеральна — з якої проводять вибір одиниць для обстеження, і вибірка, - яку безпосередньо обстежують. Статистичні характеристики вибіркової сукупності розглядають як оцінки генеральної. Отримання таких оцінок і є основним завданням вибіркового методу.

Якщо генеральна сукупність має N елементів, а для обстеження треба відібрати n, то число можливих вибірок можна обчислити за формулою:

$$C_N^n = \frac{N!}{n!(N-n)!} \quad (41)$$

Усі вибірки мають однакову імовірність і кожна з них окремо має певну похибку, властиву факту випадкового відбору. Тобто, вибірка не точно відтворює склад генеральної сукупності, тому і вибіркові оцінки не збігаються з відповідними характеристиками генеральної сукупності. Розбіжність між ними називають помилкою репрезентативності. Так, для середньої величини вона являє собою різницю:

$$\bar{x} - \bar{X}$$

де \bar{X} - генеральна, \bar{x} - вибіркова середні. Імовірність того, що випадкова величина $\bar{X} - \bar{x}$ не перевищить граничної помилки Δ (або попаде в певний інтервал) є такою,

$$P\{|\bar{x} - \bar{X}| \leq \Delta\} \geq 1 - \alpha \quad (42)$$

де гранична помилка Δ задається рівнем істотності α і залежить від величини середньої квадратичної похибки середнього арифметичного, тобто:

$$\Delta = t \cdot \mu \quad (43)$$

$$\mu = \sigma / \sqrt{n} \quad \text{де } n - \text{ коефіцієнт довіри} \quad (44)$$

Значення t , в залежності від рівня істотності α і ступенів вільності $k=n-1$, вибирають за таблицями розподілу Стьюдента (додаток 2), якщо $n \leq 30$ спостережень, або лише тільки по значенню $1 - \alpha$ за допомогою функції Лапласа (додаток 1), якщо $n > 30$.

В теорії вибіркового методу розглядається два способи відбору: повторний і безповторний. Повторний відбір здійснюється за схемою кулі, що повертається до урни. Імовірність попасти в вибірку однакова для всіх елементів генеральної сукупності. Безповторний відбір здійснюється за схемою кулі, що не повертається до урни, внаслідок чого, в процесі вибірки, обсяг генеральної сукупності зменшується, а імовірність окремих елементів попасти в вибірку збільшується. Отже, гранична помилка середньої обчислюється за різними формулами в залежності від способу відбору. Для повторного - обчислюється за формулою (43), для безповторного:

$$\Delta = t \cdot \mu \cdot \sqrt{1-D} \quad (45)$$

$$D = n/N$$

Формування вибірки здійснюється за певними правилами. Насамперед визначають основу вибірки. Одиниця основи репрезентує або окремий елемент сукупності, або певне їх угруповання. Основою вибірки визначається один з наступних видів відбору:

Простий випадковий відбір (власне випадкова вибірка) — це класичний спосіб формування вибірки за допомогою жеребка чи таблиці випадкових чисел.

Систематичний (механічний) відбір здійснюється через рівні інтервали. Крок інтервалу обчислюють за формулою N/n .

Розшарований (типовий) відбір — це вид формування вибірки з урахуванням структури генеральної сукупності.

Серійний відбір полягає в тому, що одиницею основи вибірки є серія елементів, які розглядають як одне ціле.

Крім розглянутих видів відбору в практиці статистичних досліджень застосовують комбінаційний відбір — тобто, різні комбінації розглянутих вище видів відбору, а також багатоступінчатий (багатоступеневий) відбір, при якому з генеральної сукупності спочатку відбирають великі групи, далі менші і так до того часу поки не будуть відібрані ті одиниці, які обслідуються. На відміну від багатоступеневого відбору багатофазна вибірка передбачує зберігання однієї і тієї ж одиниці відбору на всіх етапах відбору: на кожній наступній стадії відбору програма обслідування розширяється. Багатофазна вибірка ефективна при обстеженні двох і більше об'єктів, які відрізняються варіацією. Суть її полягає в тому, що для різних ознак формують вибіркові сукупності різного обсягу і поетапно: I фаза — вибірка сукупність; II фаза — підвибірка з вибіркової сукупності I фази; III фаза — підвибірка з вибіркової сукупності II фази і т.д.

Оптимальні обсяги вибірки обчислюють для повторного відбору за формулою:

$$n = t^2 \cdot \sigma^2 / \Delta^2 \quad (46)$$

і для безповторного:

$$n = \frac{t^2 \sigma^2 N}{\Delta^2 N + t^2 \sigma^2} \quad (47)$$

тобто, визначення n передбачає завчасне обчислення граничної помилки Δ .

Перейдемо до означення основних понять при статистичній перевірці гіпотез. Під статистичною гіпотезою розуміють таке припущення відносно параметрів розподілу випадкової змінної, яке можна перевірити по вибірці. Потрібно мати на увазі, що статистична перевірка гіпотез має імовірнісний характер, оскільки всі висновки базуються на вивченні властивостей розподілу випадкової змінної по вибірці, а тому завжди є ризик помилки. Але за допомогою статистичної перевірки гіпотез можна визначити імовірність прийняття помилкового рішення. Якщо вона мала, то можна вважати, що даний критерій забезпечує малий ризик помилки.

Важливо знати, що при перевірці гіпотез є можливість помилок двох видів:

- а) помилка I роду: гіпотеза що перевіряється H_0 (нуль гіпотеза) є в дійсності вірною, але результати вибірки відхиляють H_0 .
- б) помилка II роду: H_0 є помилковою, але результати перевірки приводять до її прийняття.

Зауважимо, що помилки II роду є більш небезпечними, ніж помилки I роду.

Для застосування статистичного критерію перевірки гіпотези необхідно наступне: 1) сформулювати гіпотезу H_0 ; 2) вибрати рівень істотності α , який задає імовірність помилки I роду; 3) визначити область допустимих значень і критичну область; 4) прийняти те чи інше рішення на основі порівняння фактичного і критичного значень критерію.

Рівень значущості (рівень істотності) α — це імовірність попадання критерію в критичну область при умові справедливості гіпотези. Як правило, рівень істотності приймають рівним 0,05 чи 0,01. Виходячи з величини рівня істотності можна визначити критичну область, під якою розуміють таку область значень вибіркової характеристики, попадаючи в яку вони свідчать про те, що гіпотеза H_0 повинна бути відхиленою.

Припустимо, що значення критерію попало в критичну область. Тоді, якщо гіпотеза H_0 вірна, то імовірність такої події є α і так як α є малим, то така подія є малоімовірно, і гіпотеза H_0 відхиляється. Якщо значення критерію знаходиться в допустимій області, то H_0 приймається.

Статистичні методи вивчення взаємозв'язків

Однією з основних задач статистики є вивчення вимірювання взаємних зв'язків між суспільними явищами. При вивченні зв'язків одні ознаки розглядають як фактори, що зумовлюють зміну інших ознак. Тому, перші називають факторними ознаками, другі результативними. На рис. 7 дано види зв'язків.

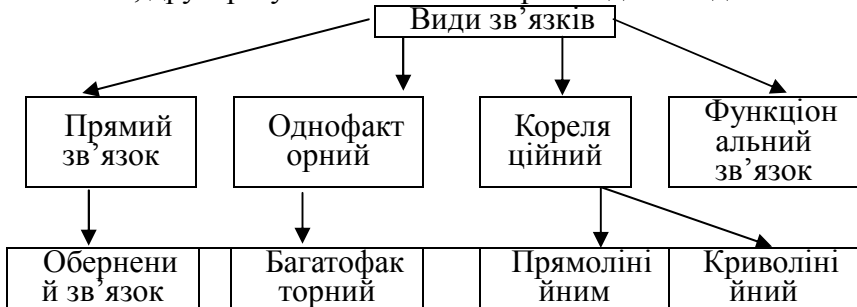


Рис. 7. Види взаємозв'язків між явищами.

Функціональний і кореляційний зв'язки обумовлені двома типами закономірностей: динамічними стохастичними. Динамічним зв'язкам властива жорстка механічна причинність. В суспільних явищах, проявляються стохастичні закономірності,

немає строгої залежності між факторами: кожному значенню факторної ознаки x відповідає розподіл значень результативної ознаки. Підвидом стохастичної залежності є кореляційна, коли із зміною факторної ознаки x змінюються групові середні ознаки y , тобто, замість розподілів порівнюють середні значення цих розподілів. Головною характеристикою кореляційного зв'язку є лінія регресії - це функція, яка зв'язує середні значення ознаки y із значеннями ознаки x . Залежно від форми лінії регресії розрізняють прямолінійний (лінійний) і криволінійний (нелінійний) зв'язки. Лінія регресії може мати табличне, аналітичне, графічне зображення. На табличному та аналітичному зображенні лінії регресії ґрунтуються дві основні моделі кореляційного зв'язку: аналітичного групування і регресійна.

Ряди динаміки

Ряд статистичних показників, який відображає зміну певного явища в часі називають динамічним рядом. Будь-який динамічний ряд має два основні елементи: рівні ряду (це конкретні значення показників) і час (або інтервали часу). Послідовність рівнів записують так

Y_0, Y_1, \dots, Y_n , n - число рівнів ряду.

Залежно від статистичної природи показника-рівня розрізняють динамічні ряди первинні і похідні, ряди абсолютних, середніх, відносних величин. За ознакою часу динамічні ряди поділяють на інтервальні і моментні. Рівень моментного ряду фіксує стан явища на певний момент часу t . В інтервальному ряді рівень - це агрегований результат процесу і залежить від тривалості часового інтервалу (рік, сезон).

Передумовою аналізу динамічних рядів є порівняність значень показників щодо одиниць вимірювання, методології обчислення, території, кола об'єктів, тощо.

У рядах, рівні яких варіюють, є потреба обчислення типової для ряду характеристики, якою є середній рівень.

Методи обчислення середніх рівнів динамічних рядів залежать від статистичної структури показника. В інтервальному ряді, рівні якого адитивні, використовують просту середню арифметичну:

$$\bar{y} = \sum y_i / n \quad (48)$$

де n - число рівнів динамічного ряду.

Розрахунок середнього рівня моментного ряду здійснюють за формулою хронологічної середньої:

$$\bar{y} = \frac{0,5(y_1 + y_n) + \sum_{i=1}^{n-2} y_i}{n-1} \quad (49)$$

Динамічні ряди поділяють на одно- і багатомірні. Одномірі характеризують зміну одного показника, багатомірні - двох і більше. В свою чергу багатомірні динамічні ряди поділяють на паралельні і ряди взаємозв'язаних показників. Розрахунок характеристик динамічного ряду дано на схемі (рис.8). Тенденція динамічного ряду — це напрям його розвитку його еволюція. Серед методів статистичного описуванні тенденції найпростішим є метод плинних середніх і інтервалах. Кожний наступний інтервал утворюється попереднього зрушенням на один рівень. Ряд плинних середніх коротший від первинного на $m-1$ рівнів. На практиці застосовують непарні інтервали. Плинна серед інтервалу, узагальнюючи значення $m=2p+1$ рівнів відноситься до середини інтервалу. Її обчислюють формулою:

$$\bar{y}_r = \bar{y}_{r-1} + \frac{y_{r+p} - y_{r-p-1}}{2p+1} \quad (50)$$

Метод плинних середніх має не тільки самостійне значення але може служити для попередньої обробки дуже коливних рядів динаміки.

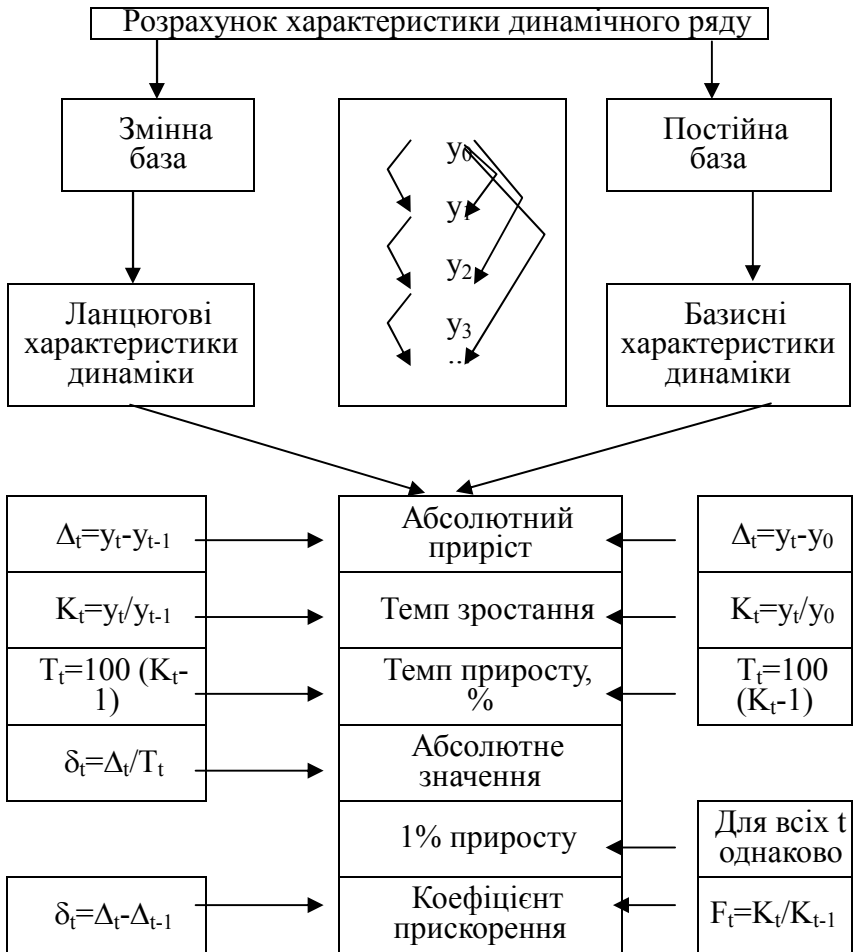


Рис. 8. Аналітичні характеристики рядів динаміки.

Статистичні індекси

Слово "індекс" означає показник. В статистиці — індекс — це відносна величина, що характеризує зміну рівня будь-якого суспільного явища в часі, просторі чи порівняно з планом, нормою, стандартом.

На практиці буквою i - позначають індивідуальні індекси, а зведені (загальні) індекси позначають великою буквою I ; перші характеризують зміни якогось одного явища, другі — характеризують співвідношення рівнів показника, до складу якого входять різномірні елементи. Такими елементами є окремі товари, види продукції, тощо. За своєю формою зведені індекси поділяють на агрегатні і середньозважені.

Для позначення показників, які індексуються, використовують певні символи, а саме:

q — кількість товарів (обсяг продукції);

p — ціна одиниці товару чи продукції;

z — собівартість одиниці продукції;

t — трудомісткість.

Якщо показники базисного періоду позначити підрядковим знаком "0", а поточного "1" то індивідуальні індекси фізичного обсягу товару, ціни, товарообороту і собівартості матимуть вигляд:

$$i_q = \frac{q_1}{q_0}; i_p = \frac{p_1}{p_0}; i_{pq} = \frac{p_1 q_1}{p_0 q_0}; i_z = \frac{z_1}{z_0} \quad (59)$$

Зведені індекси товарообороту і загальних витрат на виробництво продукції можна записати так:

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}; I_{zq} = \frac{\sum z_1 q_1}{\sum z_0 q_0} \quad (60)$$

Важливим положенням побудови зведених індексів є класифікація факторів-співмножників. Серед двох факторів-співмножників виділяють екстенсивний (об'ємний) інтенсивний (якісний). Так, якщо товарооборот це — ціна p , помножена на q ,

то p — інтенсивний показник, q - екстенсивний. Якщо при побудові індексу необхідно один з факторів залишити незмінним, то слід дотримувати правила: інтенсивні фактори-співмножники фіксуються на рівні базисного періоду, екстенсивні — на рівні поточного. За цієї умови індекс фізичного обсягу реалізації продукції має вигляд:

$$I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0} \quad (61)$$

а загальний індекс цін:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} \quad (62)$$

Позначимо інтенсивний показник через x , екстенсивний – w . Тоді, якщо треба охарактеризувати зміну екстенсивної показника в середньому по сукупності різнорідних елементів, то використовують середньоарифметичний зважений індекс:

$$I_q = \frac{\sum i_w x_0 w_0}{\sum x_0 w_0} \quad (63)$$

Середньозважений зведений індекс інтенсивного показника обчислюють за формулою середньогармонійного індексу:

$$I_x = \sum x_1 w_1 / \sum \frac{x_1 w_1}{i_x} \quad (64)$$

- індивідуальний індекс показника; x , w , — ваги. Індексний метод дозволяє також оцінити вплив окремих факторів. У загальному вигляді всі двофакторні зведені індекси поєднані так;

$$I_{xw} = I_x I_w \quad (65)$$

При аналізі динаміки середнього рівня інтенсивного показника використовують:

- індекс змінного складу:

$$I_{zc} = \frac{\sum x_1 w_1}{\sum w_1} / \frac{\sum x_0 w_0}{\sum w_0} \quad (66)$$

- індекс фіксованого складу:

$$I_{fc} = \frac{\sum x_1 w_1}{\sum w_1} / \frac{\sum x_0 w_1}{\sum w_1} \quad (67)$$

- індекс структурних зрушень:

$$I_{cz} = \frac{\sum x_0 w_1}{\sum w_1} / \frac{\sum x_0 w_0}{\sum w_0} \quad (68)$$

Між індексами середніх величин існує такий взаємозв'язок:

$$I_{zc} = I_{fc} I_{cz} \quad (69)$$

Аналіз і узагальнення статистичних даних заключний етап статистичного дослідження, метою якого є отримання теоретичних висновків (і практичних) про тенденції і закономірності досліджуваного явища. Як етапи статистичного аналізу виділяють: 1) формулювання цілі аналізу; 2) критична оцінка даних; 3) порівняльна оцінка і забезпечення порівняльності даних; 4) формування, узагальнюючих показників; 5) фіксація істотних особливостей і закономірностей явища; 6) висновки і практичні рекомендації.

Аналіз повинен проводитись при строгому дотриманні економічних і статистичних принципів. Економічними принципами є наступні:

- відповідність економічним законам;
- адекватне відображення суті економічної політики;
- орієнтація на кінцеві економічні результати;
- врахування специфіки об'єкта чи галузі;
- погодженість інтересів суб'єктів різних ієрархічних-1 рівнів, як підрозділів єдиного народногосподарської механізму.

До статистичних принципів відносяться;

- чітко означена мета дослідження;
- погодженість систем по горизонталі і вертикалі;
- порівняльність в часі і просторі;
- логічний зв'язок показників об'єкту;
- повнота відображення об'єкту в показниках;
- максимальний ступінь аналітичності.

Важливим є етап критичної оцінки вихідних даних точки зору їх достовірності і наукової цінності. Цей етап починають з аналізу однорідності статистичної сукупності, який виконують в такій послідовності:

- визначають ступінь однорідності усієї сукупності по одному чи двом істотним факторам;
- визначення і аналіз аномальних спостережень;
- вибір оптимального варіанту відбору однорідних сукупностей.

Важливою умовою отримання науково обґрунтованих результатів є перевірка гіпотези про відповідність розподілу емпіричних даних нормальному закону.

В наш час поряд з традиційними статистичними методами аналізу даних, широко застосовують кластерний аналіз, метод головних компонентів, факторний аналіз.

В цілому комплексна методологія статистичного аналізу даних здійснюється в такій послідовності.

1. Узагальнення вихідних даних.
2. оцінка однорідності сукупності.
3. Оцінка характеру розподілу сукупності.
4. Відбір факторних ознак для регресійного моделювання.
5. Побудова моделі зв'язків і оцінка її істотності.
6. Інтергіретапія моделі взаємозв'язків.

Методика комплексного аналізу і прогнозування на основі динамічних рядів означає виконання таких робіт:

1. Оцінка аномальних спостережень.
2. Розрахунки аналітичних показників динамічних рядів і аналіз тенденції.

3. Визначення наявності тенденції середніх і дисперсії на основі методів Форстера-Стюарта.
4. Визначення наявності тенденції автокореляції.
5. Визначення основної тенденції динамічного ряду.
6. Оцінка адекватності підбраної формули тренду.
7. Аналіз кореляційної функції тренду.
8. Спектральний аналіз кореляційної функції з метою оцінки періодичних компонент.

Ефективність проведення статистичного аналізу даних визначається, також, вмінням використовувати наявні статистичні інформаційні системи, які призначені як для вирішення регламентних задач, так і задач інформаційного сервісу і економічного аналізу. Для вирішення останніх використовують інформаційні технології, що називаються аналітичними комплексами (АК) — які реалізуються у вигляді пакетів прикладних програм, орієнтованих на застосування методів статистики.

Список рекомендованої літератури

1. *Стегней М.І., Іртищева І.О.* СТАТИСТИКА: кредитно-модульний курс. Навчальний посібник / *М.І.Стегней., І.О.Іртищева.* - К.: Кондор, 2011. – 304 с.
2. *Кулинич О.І.* Теорія статистики : підручник / *О.І.Кулинич, Р.О.Кулинич.* - 3-тє вид., переробл. і доповн. - К. : Знання, 1992. - 294 с.
3. *Тарасенко І.О.* Статистика: Навчальний посібник./ *І.О.Тарасенко* - :К.: Центр навчальної літератури, 2006.-344с.

Навчально-методичне видання

ОПОРНИЙ КОНСПЕКТ ЛЕКЦІЙ

з дисципліни

«СТАТИСТИКА»

для студентів

напрямів підготовки

6.030508 «Фінанси і кредит»

6.030509 «Облік і аудит»

6.030507 «Маркетинг»

денної та заочної форм навчання

Віддруковано у редакційно-видавничому відділі МДУ

89600 м. Мукачево

Вул. Ужгородська, 26

Тел. 2-11-09

*Свідоцтво про внесення суб`єкта видавничої справи до
державного реєстру видавців, виготовлювачів і
розповсюджувачів видавничої продукції Серія ДК № 4916 від
16.06.2015*