

**МІНІСТЕРСТВО ФІНАНСІВ УКРАЇНИ
ДЕРЖАВНИЙ ПОДАТКОВИЙ УНІВЕРСИТЕТ**

Факультет податкової справи, обліку та аудиту
Кафедра облікових технологій та бізнес-аналітики

Затверджено

Вченою радою Факультету,

протокол від 12.10 2022 № 3

Голова Вченої ради Факультету

В. Краєвський

**Конспект лекцій
з навчальної дисципліни
«Прикладні програми аналізу статистичних даних та моделювання»**
для підготовки здобувачів вищої освіти першого (бакалаврського) рівня
галузь знань 07 «Управління та адміністрування»
спеціальність 071 «Облік і оподаткування»
освітня програма «Облік, податковий аудит та бізнес-аналітика»

Ірпінь 2022

Конспект лекцій з курсу «Прикладні програми аналізу статистичних даних та моделювання в» складений на основі робочої програми навчальної дисципліни «Прикладні програми аналізу статистичних даних та моделювання», затвердженої Науково-методичною радою ДПУ (18.08.2022, протокол № 7)

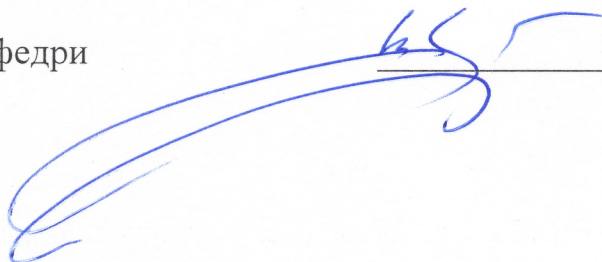
Укладач:



Я.О. Остапенко, доцент, к.е.н.,
доцент кафедри облікових технологій
та бізнес-аналітики

Розглянуто і схвалено кафедрою облікових технологій та бізнес-аналітики, протокол від 03.08. № 1
2022

В.о завідувача кафедри



Т.М. Паянок, доцент, к.е.н.

Зміст

	Стор.
ВСТУП	4
Тема 1. Основні поняття пакетів прикладних програм.	4
Тема 2. Основна характеристика прикладних програм статистичного аналізу.	7
Тема 3. Використання прикладних програм для розрахунку та аналізу описової статистики.	9
Тема 4. Використання прикладних програм для розрахунку та аналізу статистичних показників взаємозв'язку.	10
Тема 5. Використання прикладних програм для розрахунку та аналізу статистичних показників дисперсійного аналізу.	13
Тема 6. Використання прикладних програм для розрахунку та аналізу статистичних показників динаміки.	16
Тема 7. Використання прикладних програм для виявлення та аналізу тенденції розвитку та прогнозування.	19
Тема 8. Використання прикладних програм для виявлення та аналізу сезонності явища.	21
Ключові слова та терміни	22
РЕКОМЕНДОВАНА ЛІТЕРАТУРА	26

ВСТУП

Одним з обов'язкових етапів будь-якого наукового дослідження є статистичний аналіз даних. З появою і вдосконаленням сучасних програм обробки даних статистична обробка піднялася на новий рівень. Поява персонального комп'ютера третього покоління послужило базою для створення складних пакетів програм статистичного аналізу число яких продовжує зростати.

Мета навчальної дисципліни "Прикладні програми аналізу статистичних даних та моделювання": оволодіння інструментарієм оброблення статистичної інформації з використанням комп'ютерних технологій, засвоєння теоретичних і практичних знань з основ створення та функціонування інформаційних систем і технологій в статистиці.

Завдання навчальної дисципліни:

- сформувати знання про прикладні програми аналізу статистичних даних та моделювання;
- навчити вибрати пакет прикладних програм статистичного аналізу оптимальний для вирішення поставленої задачі;
- сприяти оволодінню основами роботи з прикладними статистичними пакетами програм обробки і представлення даних для ЕОМ;
- сприяти організації самостійної роботи.

Об'єкт вивчення навчальної дисципліни є пакети прикладних програм, за допомогою яких досліджують людське суспільство, явища і процеси суспільного життя.

Предметом вивчення навчальної дисципліни можливість прикладних програм для аналізу кількісної характеристики обсягу, структури, динаміки і взаємозв'язків соціально-економічних явищ в конкретних умовах місця і часу.

Матеріал предмету слід пов'язувати з такими дисциплінами як: "Інформаційні системи", "Економічна теорія", "Системний аналіз", "Фінанси", "Економічний ризик", "Страховання", "Вища математика", "Теорія ймовірностей і математична статистика".

Зміст дисципліни розкривається в темах:

Тема 1. Основні поняття пакетів прикладних програм.

Тема 2. Основна характеристика прикладних програм статистичного аналізу.

Тема 3. Використання прикладних програм для розрахунку та аналізу описової статистики.

Тема 4. Використання прикладних програм для розрахунку та аналізу статистичних показників взаємозв'язку.

Тема 5. Використання прикладних програм для розрахунку та аналізу статистичних показників дисперсійного аналізу.

Тема 6. Використання прикладних програм для розрахунку та аналізу статистичних показників динаміки.

Тема 7. Використання прикладних програм для виявлення та аналізу тенденції розвитку та прогнозування.

Тема 8. Використання прикладних програм для виявлення та аналізу сезонності явища.

Конспект лекцій з курсу «Прикладні програми аналізу статистичних даних та моделювання» має надати допомогу студентам у процесі самостійної роботи над навчальним матеріалом і передбачає послідовне викладення навчального матеріалу, навчальною метою якого є ознайомлення студентів з основним змістом, принципами і закономірностями вивчення навчальної дисципліни «Прикладні програми аналізу статистичних даних та моделювання», головними ідеями та напрямками розвитку даної науки.

Тема 1. Основні поняття пакетів прикладних програм.

Мета заняття: *засвоїти основні поняття пакетів прикладних програм, розглянути засади використання прикладних програм в статистиці. Роль прикладних програм статистичних досліджень в аналізі та прогнозуванні.*

Ключові слова: *пакет прикладних програм, програмне забезпечення*

План лекційного заняття:

1. Поняття пакету прикладних програм.
2. Програмне забезпечення статистичної інформаційної системи.
3. Застосування прикладних програм статистичного аналізу в навчальному процесі.

Одним з обов'язкових етапів будь-якого наукового дослідження є статистичний аналіз даних. Тривалий час аналіз економічних даних був долею окремих фахівців, так як це вимагало серйозної попередньої підготовки, зокрема, математичної. З появою і вдосконаленням сучасних програм обробки даних статистична обробка піднялася на новий рівень. Тепер дослідник-економіст може і не мати математичної підготовки. Досить оперувати статистичними поняттями і, найголовніше, правильно вибрати метод аналізу. Все можна реалізувати завдяки комп'ютеру і новітнім програмам.

В даний час до статистичного програмного забезпечення прийнято відносити пакети програм статистичного аналізу загального і спеціального призначення, пакети програм табулювання, пакети програм редагування, пакети програм управління даними, пакети програм для вибіркового обстеження, а також багато графічних пакетів, пакети економетричного моделювання і прогнозування, імітаційного моделювання і т.д.

Програмне забезпечення статистичної інформаційної системи (СІС) являє собою сукупність програмних засобів обчислювальної техніки, які реалізують математичні методи, моделі та алгоритми розв'язання статистичних задач, а також забезпечують відповідне сервісне обслуговування типових процедур введення-виведення, обчислення, передавання та всіх інших видів оброблення статистичної інформації. Нова інформаційна технологія у статистичній інформаційній системі припускає використання програмного забезпечення, побудованого на основі принципів єдності операційного середовища, відкритості, адаптованості, модульності, універсальності тощо.

Програмне забезпечення ЕОМ можна поділити на два класи: системне програмне забезпечення та прикладні системи (пакети прикладних програм).

У свою чергу серед системного програмного забезпечення вирізняють: операційні системи та сервісні програми; операційні оболонки; мережеві операційні системи; мови та системи програмування.

Операційні системи доповнюють апаратні засоби комп'ютера, даючи змогу прикладним програмам звертатися до зовнішніх пристроїв, а людині — користувачеві комп'ютера — управляти роботою машини. Ядро операційної системи доповнюється набором сервісних програм, які слугують для різних цілей. Так, з їхньою допомогою виконуються попереднє форматування диска, встановлення параметрів зовнішніх пристроїв, тестування оперативної пам'яті та інших пристроїв.

Операційні оболонки призначені для створення інтерфейсу, який спрощує користувачеві процес спілкування з операційною системою. Вони звільняють користувача від необхідності запам'ятовувати команди операційної системи та вводити їх з клавіатури.

Мережеві операційні системи забезпечують доступ до ресурсів, що знаходяться за межами локального комп'ютера.

Мови та системи програмування — це категорія програмних засобів, за допомогою яких створюються всі інші програми.

Прикладні системи, або пакети прикладних програм, становлять категорію програмних засобів, призначених для розв'язання повсякденних задач користувачів.

Серед таких систем вирізняють *спеціалізовані прикладні програми* — програмні засоби, що використовуються користувачем — непрограмістом для розв'язання реальних прикладних задач у статистиці.

Спеціалізовані прикладні програми мають відповідати таким вимогам:

- ✓ модульність розробки програмного забезпечення;
- ✓ розвинута система допомоги за вибору способу оброблення даних;
- ✓ використання простої проблемно-орієнтованої мови для формулювання завдання користувача;
- ✓ автоматична організація процесу оброблення даних і зв'язків модулями пакета;
- ✓ наявність засобів ведення банку даних користувача і результатів проведеного аналізу;

- ✓ можливість діалогового режиму роботи користувача з пакетом;
- ✓ сумісність з іншим програмним забезпеченням.

Характеристики функціональних можливостей спеціалізованих прикладних програм розглядаються з погляду спектра задач, розв'язуваних за допомогою математико-статистичного інструментарію.

Так, спеціалізовані статистичні пакети реалізують основні задачі первинного статистичного оброблення даних:

- аналіз змішаної природи багатомірної ознаки та уніфікацію запису вихідних даних;
- аналіз різко відмінних спостережень;
- відновлення пропущених («стертих») спостережень;
- перевірку статистичної незалежності спостережень;
- визначення основних числових характеристик і частотне оброблення вихідних даних (побудова гістограм, полігонів частот, обчислення вибіркового середнього, дисперсій і т.д.);
- розрахунок критеріїв в однорідності (середніх, дисперсій, законів розподілу, непараметричні чи кількох вибірок);
- розрахунок критеріїв згоди (хі-квадрат, Колмогорова та ін.);
- статистичне оцінювання параметрів;

обчислення найбільш розповсюджених модельних законів розподілу ймовірностей (біноміального, геометричного, Пуассона, нормального, лог-нормального, хі-квадрат, Стюдента, Фишера, бета-, гамма-експонентного, Релея, Вейбулла, Максвелла, рівномірного і деяких інших).

До категорії математико-статистичного інструментарію, що їх реалізують статистичні пакети, належить і статистичне дослідження залежностей, а саме: кореляційно-регресійний аналіз, дисперсійний і коваріаційний аналіз, планування регресійних експериментів і вибіркового обстеження, аналіз часових рядів, аналіз залежностей марковського типу та ін.

Крім того, статистичні пакети здійснюють класифікацію та зниження розмірності, а саме: дискримінантний аналіз, статистичний аналіз сумішей розподілів, кластер-аналіз, зниження розмірності (методи головних компонентів, факторного аналізу тощо).

До функцій статистичних пакетів також можуть належати задачі планування експерименту і вибіркового обстеження.

Питання для самоконтролю:

1. Що вивчає дисципліна “Прикладні програми аналізу статистичних даних та моделювання”? Охарактеризуйте предмет дисципліни..
2. Що являє собою пакет прикладних програм?
3. Які принципи використання прикладних програм для статистичного аналізу та моделювання?
4. Які існують види програмного забезпечення?
5. Які особливості застосування прикладних програм для статистичного аналізу та моделювання?
6. Програмне забезпечення ЕОМ.
7. Класифікація програмного забезпечення ЕОМ.
8. Законодавчі та нормативні документи, які регламентують проведення статистичного дослідження.

Рекомендовані літературні джерела :

основна [3-5,7,10]

допоміжна [4,7]

інформаційні ресурси Інтернет [1,3]

міжнародні видання [1,6]

Тема 2. Основна характеристика прикладних програм статистичного аналізу.

Мета заняття: розглянути особливості та характеристик прикладних програм статистичного аналізу та моделювання. Класифікація прикладних програм для статистичного аналізу та моделювання. Особливості використання прикладних програм в статистиці.

Ключові слова: прикладні програми статистичного аналізу та моделювання, класифікація прикладних програм

План лекційного заняття:

1. Пакети прикладних програм для статистичного аналізу та моделювання та їх особливості.
2. Класифікація прикладних пакетів для статистичного аналізу та моделювання.
3. Огляд основних прикладних програм статистичного аналізу та моделювання.

Поява персонального комп'ютера третього покоління послужило базою для створення складних пакетів програм статистичного аналізу число яких продовжує зростати. Конструкція пакетів статистичного аналізу залежить в значній мірі від типу використовуваного персонального комп'ютера, його конфігурації, оперативної системи, а також від рівня підготовки користувача пакета в області програмування для персонального комп'ютера.

Основною вимогою до організації вирішення завдань економіки і управління, наукових, соціологічних та інших завдань, є мінімізація ресурсів, споживаних для досягнення поставленої мети. Залежно від характеру завдання на обсяг споживаних ресурсів впливають тимчасові і вартісні обмеження. Дотримання цих обмежень може в значній мірі сприяти використанню персонального комп'ютера на всіх або окремих етапах виконання завдання.

На сьогоднішній день, за різними оцінками, на ринку представлено близько тисячі комп'ютерних програм для статистичної обробки даних (далі - статистичні пакети). Різноманітність статистичних пакетів обумовлено багатоплановістю задач обробки даних із застосуванням різних типів статистичних процедур аналізу для пошуку відповідей на питання з різних областей людської діяльності.

При виборі пакета рекомендується враховувати такі параметри:

- відповідність характеру вирішуваних завдань;
- обсяг оброблюваних даних;
- вимоги, що ставляться до кваліфікації користувача (рівень знань в області статистики);
- наявне комп'ютерне обладнання.

Статистичні пакети за ознакою функціональності можуть бути розділені на три основні групи.

1. Універсальні пакети, або пакети загального призначення (наприклад, SPSS, STATA, STATISTICA, S-PLUS, Stadia, STATGRAPHICS, SYSTAT, Minitab).
2. Професійні пакети (наприклад, SAS, BMDP).
3. Спеціалізовані пакети (наприклад, BioStat, MESOSAUR, DATASCOPE).

Універсальні пакети не орієнтовані на специфічну предметну область і можуть застосовуватися для аналізу даних з різних областей діяльності. Як правило, вони пропонують широкий діапазон статистичних методів і мають відносно простий інтерфейс. З такими пакетами рекомендується працювати користувачам-початківцям, які володіють лише базовими знаннями в області статистики, а також досвідченим користувачам на початкових етапах роботи з даними, коли ще чітко не визначені статистичні методи, які будуть застосовуватися для вирішення того чи іншого питання. Багатофункціональність універсального пакета дозволяє провести пробний аналіз різних типів даних з використанням широкого діапазону статистичних методів. Більшість

існуючих універсальних пакетів мають багато спільного за складом вбудованих статистичних процедур.

Для того щоб статистичний пакет вважався універсальним, він повинен задовольняти певні вимоги:

- містити досить широкий набір стандартних статистичних методів;
- бути досить простим для швидкого освоєння та використання непрофесійним користувачем;
- працювати з досить великими базами даних і відповідати високим вимогам до введення, перетворення і організації зберігання даних;
- здійснювати обмін даними з широко поширеними пакетами і базами даних;
- мати великий набір засобів графічного представлення даних і результатів їх аналізу;
- мати докладний документаційний супровід і довідкову систему, що дозволяє з легкістю знаходити відповіді на питання, пов'язані з роботою програми і можливостями застосування засобів аналізу даних.

Професійні пакети відрізняються від універсальних тим, що дозволяють працювати з надвеликими обсягами даних, застосовувати вузькоспеціалізовані методи аналізу, створювати власну систему обробки даних. Як правило, подібні пакети складні в освоєнні для непрофесіоналів. У той же час підготовленим користувачам робота з професійним пакетом надасть більше можливостей для глибокого і детального аналізу даних, побудови складних моделей і адаптації системи до власних потреб. Професійні пакети більш дорогі, ніж універсальні. Наприклад, вартість покупки SAS Analytics Pro на один рік для індивідуального користування становить біля п'яти тисяч євро. Ці фактори роблять сучасні професійні статистичні пакети занадто важкими для масового застосування в різних областях діяльності.

Спеціалізовані пакети дозволяють проводити аналіз з використанням обмеженого числа спеціалізованих статистичних методів або застосовні до використання для вирішення питань, що відносяться до окремо взятої предметної області. Як правило, з подібними статистичними пакетами працюють фахівці, добре знайомі з методами аналізу даних в тій області, на яку орієнтований пакет. Так, статистичний пакет BioStat створений для аналізу даних в області біології. Статистичний пакет MESOSAUR спеціалізується на аналізі одновимірних і багатовимірних часових рядів і побудові регресійних моделей. Ще один статистичний пакет DATASCOPE спеціалізується на проведенні аналізу багатовимірних даних.

Користуватися відповідними спеціалізованими пакетами доцільно тоді, коли потрібно систематично вирішувати завдання з конкретної області знань або застосовувати обмежене коло складних статистичних процедур для аналізу даних з декількох областей людської діяльності.

Перелік питань для самоконтролю

1. Сутність пакету прикладних програм для статистичного аналізу та моделювання.
2. Види прикладних програм статистичного аналізу та моделювання.
3. Універсальні прикладні програми та їх характеристика.
4. Спеціалізовані прикладні програми та їх характеристика.
5. Професійні прикладні програми та їх характеристика.
6. Загальна характеристика програмного забезпечення статистичної інформаційної системи.
7. Спеціалізовані прикладні програми статистичного аналізу та моделювання.
8. Професійні прикладні програми статистичного аналізу та моделювання.
9. Аналіз сучасного стану на ринку прикладних програм.

Рекомендовані літературні джерела :

основна [2,5,8]

допоміжна [2,5,7]

інформаційні ресурси Інтернет [1,3,6]

міжнародні видання [1,4,6]

Тема 3. Використання прикладних програм для розрахунку та аналізу описової статистики.

Мета заняття: розглянути показники описової статистики в рядах розподілу. Особливості використання пакетів прикладних програм для розрахунку та аналізу показників описової статистики.

Ключові слова: показники описової статистики, ряди розподілу

План лекційного заняття:

1. Поняття статистичного ряду розподілу. Види рядів розподілу.
2. Характеристики центру розподілу в статистичних рядах розподілу.
3. Розрахунок та аналіз показників описової статистики за допомогою пакетів прикладних програм.

Статистичним рядом розподілу називається ряд чисел, що характеризує розподіл одиниць досліджуваної сукупності на групи за якоюсь ознакою, різновидності якої розташовані у певному порядку.

Ряди розподілу складаються з двох елементів: *варіантів і частот*.

Варіанта – окреме значення групувальної ознаки, а **частоти** – кількість елементів у групі з відповідним значенням ознаки.

Частоти, які відповідають певній ознаці, можуть подаватись як в абсолютних значеннях, так і у відносних, виражених коефіцієнтом, або відсотком (часткою). Накопичену частоту(частку) називають **кумулятивною**.

Кумулятивні ряди, які будуються за накопиченими частотами (частками). На підставі таких рядів визначають структурні середні: моду та медіану, вивчають процес концентрації досліджуваного явища.

За характером розподілу варіаційні ряди бувають *симетричними і асиметричними*.

Ряд розподілу, в якому частоти спочатку наростають, а потім так само спадають, називають **симетричним**. Якщо ж розміщення частот в обидві сторони від середньої неоднакове, такий ряд називають **асиметричним**, або **скошеним**.

Ряди розподілу допомагають досліджувати структуру явищ. Вони мають самостійне значення при вивченні варіації групувальної ознаки.

Аналіз закономірностей розподілу ґрунтується на характеристиках: *центру розподілу, варіації, форми розподілу(асиметрії, концентрації)*.

До характеристик **центру розподілу** відносять *середню, моду, медіану*.

Середня величина характеризує типовий рівень ознаки в сукупності. За даними ряду розподілу середня обчислюється як арифметична зважена, де вагами є частоти f_j або частки d_j :

$$\bar{x} = \frac{\sum_1^m x_j f_j}{\sum_1^m f_j}, \bar{x} = \sum_1^m x_j d_j,$$

де j – номер групи; m – кількість груп.

Моду в статистиці називається ознака, що зустрічається в досліджуваній сукупності найчастіше.

Для *дискретного ряду розподілу* модою буде ознака, яка має найбільшу частоту (f).

В *інтервальному ряду розподілу* мода обчислюється за інтерполяційною формулою:

$$M_o = x_0 + i \frac{f_2 - f_1}{(f_2 - f_1) + (f_2 - f_3)}; \quad (5.1)$$

де M_o – мода (конкретне значення);

x_0 – нижня межа модального інтервалу;

i – ширина модального інтервалу;

f_2 – частота модального інтервалу;

f_1 – частота інтервалу, що стоїть перед модальним;

f_3 – частота інтервалу, що стоїть після модального.

Медіаною або серединною варіантою називається ознака, яка знаходиться в середині ранжированого ряду значень ознаки.

У дискретному ряді медіаною є значення ознаки, кумулятивна частота якої перевищує половину обсягу сукупності, тобто $s_{f_j} \geq 0,5 \sum_1^m f_j$ (для кумулятивної частки $s_{d_j} \geq 0,5$).

Медіана в інтервальному ряду розподілу обчислюється за формулою

$$M_e = x_0 + i \frac{\frac{\sum f}{2} - S_{Me-1}}{f_{Me}}, \quad (5.2)$$

де x_0 – нижня межа медіанного інтервалу;

i – ширина медіанного інтервалу;

S_{Me-1} – сума частот, що стоять перед медіанною частотою (кумулятивна частота);

f_{Me} – частота медіанного інтервалу;

$\frac{\sum f}{2}$ – півсума частот.

Програма PSPP дає змогу розрахувати практично всі застосовувані сьогодні описові статистики. Створення дискретного ряду розподілу та визначення частот у програмі PSPP здійснюють за алгоритмом «Аналіз → описова статистика → частоти». У віконечку «Частоти» необхідно зазначити ознаку для групування (визначення частот) та за потреби можна позначити міткою необхідні розрахунки описової статистики та побудови.

Результати розрахунків система відобразить окремим файлом, про що повідомить мерехтіння значка PSPP на панелі.

Перелік питань для самоконтролю

1. Складові статистичних рядів розподілу.
2. Суть та види середньої величини.
3. Особливості розрахунку моди та медіани в рядах розподілу.
4. Використання прикладних програм для розрахунку та аналізу показників описової статистики.
5. Поняття атрибутивного ряду розподілу.
6. Поняття варіаційного ряду розподілу.
7. Мода і медіана в статистиці.
8. Особливості використання пакетів прикладних програм для розрахунку показників описової статистики.

Рекомендовані літературні джерела :

основна [1-5,7-10]

допоміжна [1-,7]

інформаційні ресурси Інтернет [1-3]

міжнародні видання [1-6]

Тема 4. Використання прикладних програм для розрахунку та аналізу статистичних показників взаємозв'язку.

Мета заняття: вивчити взаємозв'язки між явищами та процесами суспільного життя. Способи визначення взаємозв'язків. Особливості використання пакетів прикладних програм при визначенні взаємозв'язків

Ключові слова: взаємозв'язки між явищами, види взаємозв'язків

План лекційного заняття

1. Поняття взаємозв'язків між явищами і процесами, що відбуваються в суспільному житті.
2. Види взаємозв'язків.
3. Статистичні показники визначення та аналізу взаємозв'язків між явищами.
4. Використання прикладних програм для розрахунку та аналізу статистичних показників взаємозв'язку.

Завдання статистики у вивченні взаємозв'язків між явищами полягає в тому, щоб: виявити наявність зв'язку; визначити характер зв'язку і його напрям; кількісно виміряти зв'язок.

За статистичною природою всі зв'язки можна поділити на два види: *функціональні та стохастичні* (нефункціональні). Перший вид ще називають повним зв'язком.

Функціональний - це такий зв'язок, при якому кожному конкретному значенню факторної ознаки x відповідає певне значення результативної ознаки y . Наприклад, залежність обсягу статутного капіталу від кількості акцій та ціни однієї акції.

Функціональний зв'язок можна алгебраїчно виразити формулою, яка встановлює повну відповідність між факторною та результативною ознаками, тобто причиною і наслідком. (площа круга залежить від радіуса).

Стохастичний - це зв'язок, при якому кожному значенню факторної ознаки x відповідають декілька значень результативної ознаки y . (виразити функцією неможливо).

Різновидом стохастичного зв'язку є кореляційний зв'язок.

Кореляція (від англ. *Співвідношення, відповідність*) - взаємозв'язок між ознаками, що полягає в зміні середнього значення однієї з них залежно від зміни іншої, тобто при кореляційному зв'язку кожному значенню ознаки x відповідає середнє значення ознаки \bar{y} . Наприклад, залежність між прибутковістю та ліквідністю активів компанії.

Розрізняють дві форми кореляційних зв'язків: *прямолінійний та криволінійний (обернений)*.

При **прямолінійному** зв'язку - рівній зміні факторної ознаки відповідає рівна зміна результативної. Його можна точно або наближено зобразити рівнянням будь-якої прямої лінії. Наприклад, зв'язок між фондоозброєністю та продуктивністю праці, між собівартістю продукції і рівнем рентабельності.

При **криволінійному (оберненому)** зв'язку - рівній зміні факторної ознаки відповідає нерівна зміна результативної, тобто значення результативної ознаки змінюється в протилежному напрямку відносно зміни значення факторної.

Криволінійні зв'язки можуть бути виражені рівнянням будь-якої кривої лінії: параболи, гіперболи, степеневі функції.

Характеристикою кореляційного зв'язку є *лінія регресії*, яка розглядається у двох моделях: *аналітичного групування та регресійного аналізу*.

Метод аналітичних групувань полягає у тому, що спочатку обирають факторну ознаку і результативну, потім проводять групування за факторною ознакою та обчислення середніх у кожній групі за результативною ознакою. Зіставленням характеру зміни факторної та результативної ознак можна дійти висновку про наявність зв'язку, його напрям та тісноту.

Кореляційний метод застосовується для вимірювання тісноти (щільності) зв'язків між ознаками за допомогою спеціальних співвідношень, що базуються на правилі додавання дисперсій. Ці співвідношення можна обчислити для кількісних ознак. Числові характеристики кореляційного зв'язку: кореляційне відношення; індекс кореляції; лінійний коефіцієнт кореляції.

Кореляційне відношення показує питому вагу міжгрупової дисперсії у загальній дисперсії, тобто визначає, наскільки тісний зв'язок факторної ознаки, за якою проводилося групування, та результативної ознаки.

Його позначають малою грецькою буквою η — («тета») і обчислюють за формулою:

$$\eta = \sqrt{\frac{\delta^2}{\sigma^2}}, \quad (4.1)$$

де δ^2 - міжгрупова дисперсія; σ^2 - загальна дисперсія.

Кореляційне відношення змінюється від 0 до 1.

Чим ближче η до 1, тим зв'язок між ознаками тісніший.

Індекс кореляції визначають зіставленням внутрішньогрупової дисперсії та загальної, позначають буквою R і обчислюють за формулою:

$$R = \sqrt{1 - \frac{\sigma^2}{\sigma^2}}, \quad (4.2)$$

де σ^2 - внутрішньо групова дисперсія; σ^2 - загальна дисперсія.

Чим R ближче до 1, тим тісніший зв'язок між ознаками.

Лінійний коефіцієнт кореляції використовують для вимірювання тісноти прямолінійних зв'язків.

Мірою тісноти зв'язку як лінійного, так і нелінійного є **коефіцієнт детермінації R^2** — співвідношення факторної і загальної дисперсії:

$$R^2 = \frac{\sigma_y^2}{\sigma^2}, \quad (4.3)$$

де $\sigma_y^2 = \frac{\sum (Y - \bar{y})^2}{n}$ (4.4)

Корінь квадратний з коефіцієнта детермінації є **індекс кореляції R**

У кореляційно-регресійному аналізі істотність зв'язку перевіряється так само, як і в аналітичному групуванні за допомогою **R^2** чи **F -критерія**:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{k_2}{k_1} \quad (4.5)$$

В основі кореляційно-регресійного аналізу лежить припущення, що залежність між факторною і результативною ознаками може бути виражена функцією $Y=f(x)$, яка називається **рівнянням регресії**.

Рівняння регресії — аналітичне рівняння, за допомогою якого можна виразити зв'язок між ознаками. Тобто це економіко-схематична модель залежності результативної ознаки від факторної. Графіком рівняння регресії є **лінія регресії**, яка описує кореляційний зв'язок. При побудові графіка значення факторної ознаки відкладаються на горизонтальній осі (OX), а результативної — на вертикальній (OY).

Для визначення парного коефіцієнта кореляції в PSPP на вкладці «Аналіз» відзначаємо курсором «Двовірна кореляція»:

Заповнюємо віконечко для обрахування коефіцієнта кореляції.

Окремим файлом PSPP виведе результати обчислення, про що повідомить мерехтіння значка PSPP.

Регресійний аналіз полягає у наближенні досліджуваного ряду розподілу результативного показника до ряду, який приблизно описує відповідність між результативною та факторними ознаками при цьому завдяки наближенню за можливості виключається дія випадкових факторів.

Для визначення показників регресійної статистики в PSPP необхідно виконати такі кроки:

«Аналіз» → «Регресія» → «Лінійна».

Заповнюємо віконечко.

Результати будуть надані окремим файлом результатів. Заключним етапом усіх статистичних розрахунків є аналіз отриманих результатів.

Перелік питань для самоконтролю

1. Поняття функціонального зв'язку між ознаками.
2. Стохастичний зв'язок між ознаками.
3. Поняття кореляційного зв'язку.
4. Графічний метод виявлення взаємозв'язку.
5. Особливості використання пакетів прикладних програм для розрахунку показників для виявлення взаємозв'язку між ознаками.
6. Визначення взаємозв'язку між ознаками в статистиці.
7. Види взаємозв'язків.
8. Статистичні показники для визначення наявності, тісноти та напрямку взаємозв'язків.
9. Використання прикладних програм для розрахунку та аналізу показників описової статистики.

Рекомендовані літературні джерела :

основна [2-7,9]

допоміжна [3-5]

інформаційні ресурси Інтернет [2-4]

міжнародні видання [3,6]

Тема 5. Використання прикладних програм для розрахунку та аналізу статистичних показників дисперсійного аналізу.

Мета заняття: розглянути основи дисперсійного аналізу в статистиці. Визначення взаємозв'язку. Особливості використання прикладних програм в дисперсійному аналізі.

Ключові слова: дисперсійний аналіз, дисперсія, види дисперсій.

План лекційного заняття

1. Сутність дисперсійного аналізу в статистиці.
2. Види та значення дисперсій.
3. Визначення взаємозв'язку в дисперсійному аналізі.
4. Використання прикладних програм для дисперсійного аналізу

Дисперсійний аналіз є сукупністю статистичних методів, призначених для перевірки гіпотез про зв'язок між певною ознакою та досліджуваними факторами, які не мають кількісного опису, а також для встановлення ступеня впливу факторів та їх взаємодії. У спеціальній літературі дисперсійний аналіз часто називають ANOVA (від англomовної назви Analysis of Variations). Вперше цей метод було розроблено Р. Фішером в 1925 р. Факторами називають контрольовані чинники, що впливають на кінцевий результат. Рівнем фактора, або способом обробки, називають значення, що характеризують конкретний прояв цього фактора. Ці значення зазвичай подають у номінальній або порядковій шкалі вимірювань. Значення вимірюваної ознаки називають відгуком. Часто вихідні значення факторів вимірюють у кількісних або порядкових шкалах. Тоді постає проблема групування вихідних даних у ряди спостережень, що відповідають приблизно однаковим значенням фактора. Якщо кількість груп взяти надмірно великою, то кількість спостережень у них може виявитися недостатньою для отримання надійних результатів.

Якщо її взяти надмірно малою, це може призвести до втрати суттєвих особливостей впливу досліджуваного фактора на систему. Загальну методологію групування описано в розділі 1. Вибір конкретного способу групування даних залежить від їх обсягу і характеру варіювання значень фактора. Кількість і розміри інтервалів при однофакторному аналізі найчастіше визначають за принципом рівних інтервалів або за принципом рівних частот. При багатфакторному аналізі застосовують три типи групування: – групи з рівною кількістю спостережень; – групи з різною кількістю спостережень; – групи, кількості спостережень у яких відповідають певній пропорції.

Основною метою дисперсійного аналізу є виявлення впливу окремих факторів чи умов, які визначають варіацію ознаки. В основі дисперсійного аналізу лежить закон розкладання загальної дисперсії на складові, згідно з яким загальна дисперсія результативної ознаки складається з міжгрупової дисперсії та середньої з групових дисперсій (залишкової): $\sigma^2 = \overline{\sigma^2} + \delta^2$, де

Загальна дисперсія – характеризує варіацію значень ознаки y_s навколо загальної середньої \bar{y} :

$$\sigma^2 = \frac{\sum_1^n (y_j - \bar{y})^2}{n}$$

Середня з групових дисперсій – узагальнює внутрішню групову варіацію:

$$\overline{\sigma^2} = \frac{\sum_1^m \sigma_j^2 \cdot f_j}{\sum_1^m f_j}$$

Між групова дисперсія – характеризує варіацію групових середніх \bar{y}_j навколо загальної середньої:

$$\delta^2 = \frac{\sum_1^m (\bar{y}_j - \bar{y})^2 \cdot f_j}{\sum_1^m f_j}$$

За допомогою однофакторного дисперсійного аналізу досліджується вплив одного фактора, тобто перевіряється значущість відмінностей між собою значень середніх величин декількох вибірових сукупностей. В однофакторному дисперсійному аналізі набір даних складається з k незалежних одновимірних вибірок із k генеральних сукупностей (або з однієї генеральної сукупності), елементи яких вимірюються в однакових одиницях.

Багатфакторний дисперсійний аналіз дає можливість оцінити ефект кожного фактора, тобто вплив кожного фактора на результативний показник, а також ефект взаємодії факторів (показати як фактори пов'язані між собою). При чому чим більше факторів, тим більше видів взаємодії між ними. Допустимо, при трьохфакторному дисперсійному аналізі крім чистого впливу кожного фактора на результативний показник визначається взаємодія між кожними двома факторами (першим і другим, другим і третім, першим і третім), а також між трьома факторами разом. Таким чином, багатфакторний дисперсійний аналіз застосовується тоді, коли в одному дисперсійному аналізі необхідно одночасно досліджувати не лише вплив факторів, а також можливу коваріацію змінних. При цьому змінні не є незалежними одна від іншої, а навпаки, корелюють між собою. Для коректного застосування багатфакторного дисперсійного аналізу повинні виконуватися такі ж вимоги, як і для однофакторного.

Для перевірки гіпотези щодо суттєвості відмінностей між середніми величинами у вибірках, що сформовані за факторною ознакою, використовується F-критерій Фішера. В однофакторному дисперсійному аналізі статистичні гіпотези формулюються так: – нульова гіпотеза (H_0): усі середні рівні між собою; – альтернативна (дослідницька) гіпотеза (H_1): не всі середні рівні між собою, принаймні існує хоча б одна пара вибірових сукупностей, середні величини в яких відрізняються між собою.

Для коректного використання однофакторного дисперсійного аналізу повинні виконуватися такі вимоги: 1. Кількісні дані повинні бути неперервними, дискретні дані є менш бажаними. 2. Вибіркові сукупності повинні бути незалежними. 3. Незалежними повинні бути спостереження в кожній із вибірок. 4. Вибіркові сукупності повинні мати нормальний розподіл. 5. Дисперсії генеральних сукупностей, із яких здійснювалися вибірки, повинні бути однаковими. Це означає, що всі вибірки здійснюються або з однієї генеральної сукупності, або середні в генеральних сукупностях, з яких здійснюються вибірки, рівні між собою, що дозволяє використовувати для перевірки гіпотез стандартні статистичні таблиці.

F-критерій Фішера для однофакторного дисперсійного аналізу визначається як співвідношення незміщених оцінок міжгрупової та середньої з внутрішньогрупових дисперсій.

Тобто, F-критерій показує, у скільки разів середні величини кожної з вибірок відрізняються між собою, тобто наскільки більше вони відрізняються між собою, ніж можна було цього очікувати, якби їх розбіжності були суто випадковими. Очевидно, що його значення можуть бути або додатними, або дорівнювати нулю. Дорівнювати нулю F-критерій може в тому випадку, коли міжгрупова варіація ознаки відсутня. Для того, щоб перевірити істотність розходження дисперсій необхідно отримане значення F-критерію порівняти з табличним.

Якщо фактичне значення F-критерію перевищує його табличне значення, то слід відхилити нульову гіпотезу і прийняти альтернативну, а це означає, що групові вибіркові середні суттєво відрізняються між собою, причому ці відмінності не можна пояснити лише випадковістю, тобто результат є статистично значущим.

В двохфакторному дисперсійному аналізі використовуються три різних критерії: 1) для перевірки гіпотези про відсутність ефекту фактора А; 2) для перевірки гіпотези про відсутність ефекту фактора В; 3) для перевірки гіпотези про відсутність ефекту взаємодії факторів А і В. Після визначення F-критеріїв Фішера перевірка гіпотез здійснюється за таким самим алгоритмом, як і в однофакторному дисперсійному аналізі. Двохфакторний дисперсійний аналіз може мати два різновиди: 1) без повторень; 2) з повтореннями.

У першому випадку кожному рівню факторів відповідає тільки одна вибірка даних, у другому – певним рівням факторів може відповідати більше однієї вибірки даних. Унаслідок складності розрахунків, особливо при великій кількості рівнів кожного фактора, для двохфакторного аналізу слід застосовувати або Excel, або спеціалізоване програмне забезпечення. Для здійснення розрахунків в Excel необхідно активізувати «Пакет аналізу» і далі здійснити наступні кроки: 1. Обрати опцію «Двохфакторний дисперсійний аналіз без повторень (або з повтореннями)». 2. У діалоговому вікні з'являється адреси ознак, їх необхідно вказати (із попередньо складеного в Excel масиву вихідних даних). 3. Вказати кількість рядків для вибірки (враховуючи кількість ступенів свободи). 4. Встановити рівень надійності коефіцієнтів регресії (за замовчуванням – 95 %). 5. Командою ОК вивести результати на новий робочий аркуш.

Дисперсійний аналіз зручно проводити в *Microsoft Excel*.

Для розрахунку внутрішньогрупової дисперсії необхідно застосувати функції: **КВАДРОТКЛ** (число_1; число_2...) - визначає суму квадратів відхилення заданого числа від середнього:

❖ число_1; 2... - число, для якого обраховується квадрат відхилення;

ОКРУГЛ - для заокруглення показників.

ДИСПР (число1;число 2...) - визначає дисперсію для даної сукупності:

❖ число 1;2... - число сукупності.

СУММПРОИЗВ (масив1; масив 2...) - визначає суму добутків діапазонів та масивів:

❖ масив1; 2... - масиви, які потрібно спочатку перемножити, а потім скласти.

Перелік питань для самоконтролю

1. Значення дисперсійного аналізу для статистичного дослідження.
2. Характеристика взаємозв'язків при використанні дисперсійного аналізу.
3. Особливості використання пакетів прикладних програм для дисперсійного аналізу.
4. Поняття дисперсійного аналізу.
5. Види дисперсій в статистиці.

6. Визначення взаємозв'язку в дисперсійному аналізі.
7. Особливості використання прикладних програм в дисперсійному аналізі.

Рекомендовані літературні джерела :

основна [4-9]
 допоміжна [2-5]
 інформаційні ресурси Інтернет [2,4-6]
 міжнародні видання [2,3]

Тема 6. Використання прикладних програм для розрахунку та аналізу статистичних показників динаміки.

Мета заняття: вивчити показників рядів динаміки. Абсолютні, відносні та середні показники рядів динаміки. Взаємозв'язок показників. Особливості використання прикладних програм для аналізу рядів динаміки.

Ключові слова: ряди динаміки, показники рядів динаміки

План лекційного заняття

1. Поняття рядів динаміки. Види рядів динаміки.
2. Показники рядів динаміки та їх характеристика.
3. Взаємозв'язок показників динаміки.
4. Використання прикладних програм для аналізу рядів динаміки.

Рядом динаміки називається ряд чисел, що характеризують зміну величини суспільного явища в часі.

Розрізняють ряди динаміки абсолютних, відносних і середніх величин.

Ряди динаміки абсолютних величин бувають моментними і періодичними.

Кожний ряд динаміки можна охарактеризувати за допомогою таких показників: абсолютний приріст, коефіцієнт зростання, темп зростання, темп приросту, абсолютне значення одного процента приросту, середній рівень ряду динаміки, середній абсолютний приріст, середній темп зростання та приросту.

Абсолютний приріст обчислюється як різниця між поточним (звітним) і базисним рівнями.

Базисний

$$A = Y_n - Y_0$$

Ланцюговий

$$A = Y_n - Y_{n-1}$$

де A – абсолютний приріст;

Y_n – поточний (звітний) рівень;

Y_0 – базисний рівень;

Y_{n-1} – попередній рівень.

Коефіцієнт зростання обчислюється як відношення рівня досліджуваного періоду (Y_n) до рівня, з яким порівнюють.

Базисний

$$K_p = \frac{Y_n}{Y_0}$$

Ланцюговий

$$K_p = \frac{Y_n}{Y_{n-1}}$$

Темп зростання обчислюється як процентне відношення рівня досліджуваного періоду (Y_n) до рівня, з яким порівнюють.

Базисний

$$T_p = \frac{Y_n}{Y_0} 100$$

Ланцюговий

$$T_p = \frac{Y_n}{Y_{n-1}} 100$$

Не важко помітити, що $T_p = K_p * 100$

Темп приросту обчислюється шляхом ділення абсолютного приросту, помноженого на 100, на величину рівня, з яким порівнюють.

Обчислюється також базисним і ланцюговим способом.

Базисний

$$T_{np} = \frac{Y_n - Y_0}{Y_0} 100$$

Ланцюговий

$$T_{np} = \frac{Y_n - Y_{n-1}}{Y_{n-1}} 100$$

Темп приросту можна обчислити шляхом віднімання "100" від значення темпів росту:

$$T_{np} = T_p - 100$$

Абсолютне значення одного процента приросту обчислюється шляхом ділення абсолютного приросту на темп приросту.

$$A\% = \frac{Y_n - Y_{n-1}}{T_{np}}$$

Для узагальнюючої характеристики динаміки досліджуваного явища за ряд періодів визначають різного роду середні показники. Серед них середні рівні ряду та середні показники змін рівнів ряду.

Середні рівні використовують, насамперед, для узагальнення коливних рядів, для забезпечення порівнянності чисельника і знаменника, побудови динамічних рядів похідних показників.

Метод обчислення середнього рівня динамічного ряду залежить від виду ряду динаміки.

В *інтервальному* ряді абсолютних величин з рівними періодами часу використовується середня арифметична проста:

$$\bar{y} = \frac{\sum x}{n},$$

де n – число рівнів ряду

x – рівні ряду.

У *моментному* ряді при умові рівномірної зміни показника між датами (рівновіддаленими рівнями), середня величина між двома датами розраховується як півсума значень на початок і кінець періоду (середня арифметична проста):

$$\bar{y} = \frac{Y_0 + Y_n}{2};$$

для різних проміжків часу – середню арифметичну зважену.

Якщо моментний ряд динаміки має однакові проміжки (інтервали) часу між датами, розрахунок середнього рівня виконується за формулою середньої хронологічної:

$$\bar{y} = \frac{(\frac{y_1}{2} + y_2 + y_3 + \dots + y_{n-1} + \frac{y_n}{2})}{n-1};$$

У моментних та інтервальних рядах динаміки з нерівними періодами (проміжками) часу для обчислення середнього рівня ряду використовують середню арифметичну зважену:

$$\bar{y} = \frac{\sum yt}{\sum t}$$

де yt – рівні ряду;

t – проміжки часу між суміжними датами або періоди часу.

Середній абсолютний приріст обчислюється як середня арифметична проста із ланцюгових абсолютних приростів.

$$\bar{A} = \frac{\sum A}{n},$$

де A – величини ланцюгових приростів;

n – число ланцюгових приростів.

При відсутності ланцюгових приростів середній абсолютний приріст можна обчислити за формулою:

$$\bar{A} = \frac{Y_n - Y_0}{n-1},$$

де n – число календарних дат.

Середні темпи (коефіцієнти) зростання обчислюються за формулами середньої геометричної:

$$\bar{K}_p = \sqrt[n]{K_1 K_2 \dots K_n} = \sqrt[n]{PK},$$

де \bar{K} – середній коефіцієнт зростання,

K – ланцюгові коефіцієнти зростання,

n – число ланцюгових коефіцієнтів,

P – знак добутку, який показує, що величини, які стоять після нього, потрібно перемножити і знайти їх добуток.

Якщо відсутні ланцюгові коефіцієнти, то можна скористатися іншою формулою:

$$\bar{K} = \sqrt[n-1]{\frac{Y_n}{Y_0}},$$

де n – число календарних дат, Y_n і Y_0 – рівні відповідно звітного і базисного періодів. Одержані коефіцієнти перетворюють в темпи зростання за формулою $T_{np} = K_p * 100$ і в темпи приросту за формулою $T_{np} = K_p * 100 - 100$.

Якщо швидкість розвитку явища в межах розглядуваного періоду неоднакова, то зіставленням однойменних характеристик за різні інтервали часу визначають **прискорення** чи **уповільнення динаміки**. Різниця абсолютних ланцюгових приростів $\Delta_t - \Delta_{t-1}$ характеризує **абсолютне прискорення (+)** чи **уповільнення (-)**. Зіставленням темпів зростання (приросту) дістають **відносне прискорення (уповільнення)**.

Для розрахунку показників в *Microsoft Excel* використовують функції:

СРЗНАЧ (число 1; число 2...) яка обраховує середнє арифметичне заданих чисел:

❖ число 1; 2... - числа, для яких обраховується середнє значення

СТЕПЕНЬ (число; степінь) - підносить значення показника до степеня:

❖ число - показник, який необхідно піднести до степеня;

❖ степінь - показник степеня, до якого підноситься показник

розрахуємо середні показники динаміки.

Перелік питань для самоконтролю

1. Абсолютний приріст та його значення в статистиці.
2. Темп зростання та його значення в статистиці.
3. Темп приросту та його значення в статистиці.
4. Абсолютне значення одного відсотка приросту та його значення в статистиці.
5. Особливості використання прикладних програм для аналізу рядів динаміки.
6. Складові ряди динаміки. Види рядів динаміки.
7. Характеристики рядів динаміки, їх розрахунок та значення.
8. Середня характеристики рядів динаміки.
9. Особливості використання прикладних програм для аналізу рядів динаміки.

Рекомендовані літературні джерела :

основна [4,8,10]

допоміжна [2,4,7]

інформаційні ресурси Інтернет [1,3,6]

міжнародні видання [1,4,6]

Тема 7. Використання прикладних програм для виявлення та аналізу тенденції розвитку та прогнозування.

Мета заняття: вивчення показників тенденції розвитку явищ і процесів суспільного життя. Прогнозування показників на основі виявленої тенденції. Особливості використання прикладних програм для виявлення тенденції розвитку.

Ключові слова: тенденція розвитку, показники тенденції розвитку, прогнозування

План лекційного заняття

1. Поняття тенденції розвитку.
2. Методи та способи виявлення тенденції розвитку.
3. Використання прикладних програм для виявлення тенденції розвитку.

Для встановлення загальних закономірностей розвитку суспільних явищ за даними динамічних рядів їх обробляють за допомогою методів, які можна розділити на механічні та аналітичні.

Механічне вирівнювання рядів динаміки здійснюють за допомогою таких прийомів: укрупнення періодів і обчислення за ними середніх показників з наступним їх аналізом; переведення абсолютних показників динамічних рядів у відносні, за рахунок чого досягається порівнянність багатомірних динамічних рядів.

Найпростішим способом виявлення загальної тенденції розвитку явища є **укрупнення інтервалів часу динамічного ряду**. Суть цього прийому полягає в тому, що первинний ряд динаміки перетворюється і змінюється іншим, показники якого відносяться до більш триваліших періодів часу. Новостворений ряд може складатися із абсолютних величин за укрупнені періоди часу (ці величини утворюють шляхом додавання рівнів первинного ряду абсолютних величин), або із середніх величин по інтервалах.

Для визначення плинної середньої формують укрупнені інтервали, які складаються з однакової кількості рівнів. Кожний наступний інтервал одержують, поступово зсуваючись від початкового рівня динамічного ряду на один рівень. У сформованих укрупнених інтервалах визначають суми значень рівнів, на основі яких обчислюють плинні середні.

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

При аналітичному вирівнюванні ряду фактичні значення Y_t замінюються обчисленими на основі певної функції $Y = f(t)$, яку називають **трендовим рівнянням**.

Іноді виникає необхідність у знаходженні відсутніх проміжних рівнів ряду. Ця процедура має назву **інтерполяції** і проводиться з огляду загальної тенденції розвитку за період, що досліджується.

Інтерполяція – це знаходження невідомого рівня динамічного ряду.

При прогнозуванні економічних показників використовують інший статистичний засіб — **екстраполяцію**. При цьому обчислюють значення рівнів за межами наявних фактичних даних. При екстраполяції виходять із припущення, що виявлена тенденція буде зберігатися і надалі. Для проведення цієї операції треба лише у рівняння тренду підставити необхідне значення згідно з продовженням вихідного ряду та розрахувати Y .

Іноді виникає потреба порівняти між собою рівнів динамічних рядів кількох споріднених або взаємопов'язаних явищ. Для цього переводять абсолютні показники рядів динаміки у відносні, прийнявши рівні будь – якого періоду за одиницю або сто. Таке перетворення динамічних рядів називають **зведенням до однієї основи**.

У процесі аналізу рядів динаміки важливо виявити загальну тенденцію розвитку суспільно – економічного явища, тобто встановити, в якому напрямі воно змінюється: зростає чи знижується.

Тенденція – це певний напрям розвитку, тривала еволюція, яка набуває вигляду більш – менш плавної траєкторії.

Серед методів статистичного описування тенденцій особливо широко застосовують **трендові криві** – певні математичні функції, за допомогою яких описується основна тенденція. Вибір функції залежить від характеру динаміки. Так, у разі порівняно стабільних абсолютних приростів беруть лінійний тренд $Y_t = a + bt$ (10.1), у разі стабільних темпів приросту – експоненту $Y_t = ab^t$ (10.2). У лінійному рівнянні параметр b характеризує середній абсолютний приріст, в експоненті – середній темп зростання. Параметр a в обох функціях – це теоретичне значення рівня при $t=0$. У лінійному рівнянні параметр b характеризує швидкість динаміки: середню абсолютну (середній абсолютний приріст) в лінійній функції і середню відносну (середній темп зростання) в експоненті. Параметр a в обох функціях – це теоретичне значення рівня при $t=0$. Коли характеристики швидкості розвитку зростають (чи зменшуються), використовують інші функції (парабола 2-го степеня, модифікована експонента тощо).

Визначають параметри трендових кривих, розв'язуючи системи нормальних рівнянь. Для лінійної функції вона має такий вигляд:

$$\begin{aligned} an + b \sum t &= \sum y, \\ a \sum t + b \sum t^2 &= \sum yt. \end{aligned} \quad (7.3)$$

Параметри a і b рівняння прямої можна обчислити простіше, скориставшись методом відліку від умовного нуля. За умовний нуль беруть значення t , що міститься посередині ряду. Тоді значення t , розміщені вище середини, будуть від'ємними, а нижче додатними. При непарному числі членів ряду (наприклад, $n=5$) змінній t надаються значення з інтервалом одиниця: **-2, -1, 0, 1, 2**; при парному (наприклад, $n=6$) – з інтервалом два: **-5, -3, -1, 1, 3, 5**. В обох випадках $\sum t = 0$, а система рівнянь набуває вигляду:

$$\begin{aligned} na &= \sum y \\ b \sum t^2 &= \sum yt \end{aligned} \quad (7.4)$$

$$\text{Тоді, } a = \frac{\sum y_i}{n}; \quad (7.5)$$

$$b = \frac{\sum_{i=1}^n y_i t_i}{\sum_{i=1}^n t_i^2}. \quad (7.6)$$

Багато суспільних явищ, які є предметом статистики, мають сезонний характер. Найбільше сезонність поширена в сільському господарстві.

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

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

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

Для вивчення загальної тенденції сезонності за деякий період часу потрібно користуватись узагальнюючим показником, яким може бути середньорічний коефіцієнт сезонності:

$$\bar{I}_s = \frac{\bar{d}}{\bar{y}_{заг}}, \quad (7.7)$$

де \bar{d} – середнє лінійне відхилення квартальних рівнів ряду динаміки від середнього рівня.

$$\bar{d} = \frac{\sum |\bar{y}_i - \bar{y}_{заг}|}{n} \quad (7.8)$$

Чим ближче значення I_s до нуля, тим менший рівень сезонності.

Використовуючи середньорічний коефіцієнт сезонності, можна визначити **коефіцієнт стабільності**:

$$I_{st} = 1 - I_s \quad (7.9)$$

Використовуючи Майстер діаграм Microsoft Excel шляхом підбору, обирають лінійну тренду, для цього обов'язково у віконечку "Формат лінії тренду", активуємо: "поместить на діаграму величину достоверности аппроксимации R^2 " та "показывать уравнение на диаграмме". Для знаходження прогнозного показника на наступний рік за часовим трендом необхідно у формулу отриманого рівняння замість "x" підставити відповідний період часу: для прогнозу.

Перелік питань для самоконтролю

1. Поняття екстраполяції в статистиці.
2. Поняття інтерполяції в статистиці.
3. Прогнозування за виявленою тенденцією.
4. Особливості використання прикладних програм для виявлення тенденції розвитку.
5. Тенденція розвитку та її види.
6. Методи виявлення тенденції.
7. Екстраполяція та інтерполяція в статистиці.
8. Прогнозування за тенденцією розвитку.
9. Особливості використання прикладних програм для виявлення тенденції розвитку.

Рекомендовані літературні джерела :
основна [6,8,10]

допоміжна [4-7]
інформаційні ресурси Інтернет [2-6]
міжнародні видання [5,6]

Тема 8. Використання прикладних програм для виявлення та аналізу сезонності явища.

Мета заняття: розглянути показники сезонності явищ та процесів суспільного життя. Показники виявлення сезонності. Особливості використання прикладних програм для виявлення сезонності.

Ключові слова: сезонність явища, показники виявлення сезонності

План лекційного заняття

1. Сезонні явища в статистиці.
2. Статистичні показники виявлення сезонності.
3. Використання прикладних програм для аналізу сезонності.

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

В статистиці існує ряд методів вивчення та виміру сезонних коливань:

- а) метод абсолютних різниць;
- б) метод відносних різниць;
- в) побудова індексів сезонності;
- г) побудова аналітичної моделі.

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

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

Для вивчення загальної тенденції сезонності за деякий період часу потрібно користуватись узагальнюючим показником, яким може бути середньорічний коефіцієнт сезонності:

$$\bar{I}_s = \frac{\bar{d}}{\bar{y}_{заг}} ,$$

\bar{d} – середнє лінійне відхилення кварталних рівнів ряду динаміки від середнього рівня.

$$\bar{d} = \frac{\sum |\bar{y}_i - \bar{y}_{заг}|}{n}$$

Чим ближче значення I_s до нуля, тим менший рівень сезонності.

Використовуючи середньорічний коефіцієнт сезонності, можна визначити коефіцієнт стабільності:

$$I_{st} = 1 - I_s$$

Для розрахунку індексів сезонності та побудови сезонної хвилі доцільно використовувати Microsoft Excel, зокрема, функцію **СРЗНАЧ** та Майстер діаграм програми.

Перелік питань для самоконтролю

1. Особливості сезонних явищ.
2. Способи визначення сезонності.
3. Особливості прогнозування розвитку сезонних явищ та процесів.

4. Поняття сезонності в статистиці.
5. Особливості розрахунку показників сезонності.
6. Особливості Використання прикладних програм для аналізу сезонності.

Рекомендовані літературні джерела :

основна [1,3,8]
 допоміжна [2-4]
 інформаційні ресурси Інтернет [3-5]
 міжнародні видання [1,6]

Ключові слова та терміни

Абсолютне значення одного відсотка приросту — це одна сота частина базового рівня показника або відношення абсолютного приросту до відповідного темпу приросту.

Абсолютний показник — це показник у формі абсолютної величини, яка відображає фізичні властивості, часові та вартісні характеристики соціально-економічних процесів та явищ.

Абсолютний приріст — це показник ряду динаміки, який показує, на скільки одиниць змінився поточний рівень показника порівняно з рівнем попереднього або базового періоду.

Абсолютним прискоренням приросту у статистиці називають різницю між наступним та попереднім абсолютними приростами, яка показує, на скільки дана швидкість більша (менша), ніж попередня.

Аналітичне вирівнювання — це найбільш досконалий прийом виявлення основної тенденції динаміки, що здійснюється за допомогою математичної формули, яка відображає загальну тенденцію ряду.

Варіанти — окремі значення ознаки, яких вона набуває у варіаційному ряду розподілу, тобто конкретні значення ознаки, що варіює.

Варіаційний ряд розподілу — ряд розподілу, який будується за кількісною ознакою.

Варіація — коливання значення ознаки окремих одиниць сукупності.

Варіація — коливання, різноманітність, змінюваність значення ознаки окремих одиниць сукупності явищ.

Вибіркова сукупність — це сукупність одиниць, які вибрані для обстеження.

Вибіркова частка — це питома вага одиниць, які мають певну ознаку у вибірковій сукупності.

Відносна похибка показує, на скільки відсотків вибірка оцінка може відхилитися від параметра генеральної сукупності.

Відносна стандартна похибка середньої — це коефіцієнт варіації вибірових середніх.

Відносне прискорення — це відношення абсолютного прискорення до абсолютного приросту, прийнятого за базу, або це темп приросту абсолютного приросту. Розраховується тільки у випадку, якщо абсолютний приріст, що прийнятий за базу порівняння, додатне число.

Відносні показники варіації — це коефіцієнт варіації лінійний та квадратичний, коефіцієнт осциляції.

Генеральна сукупність — це сукупність одиниць, з яких вибирають елементи для обстеження.

Гранична похибка вибірки — це максимально можлива похибка для взятої імовірності яка визначає розмір помилки залежно від того, з якою ймовірністю вона знаходиться.

Групова статистична таблиця — це таблиця, підметом якої є групування одиниць сукупності за однією (кількісною чи атрибутивною) ознакою.

Груповий добір передбачає формування вибіркової сукупності на основі добору груп одиниць з генеральної сукупності.

Децилі — варіанти, які поділяють ранжируваний ряд на десять рівних частин.

Дискретний варіаційний ряд — розподіл одиниць сукупності за дискретною ознакою.

Дисперсія — це середній квадрат відхилень індивідуальних значень ознаки від їхньої середньої величини.

Експес розподілу — це ступінь зосередженості елементів сукупності навколо центра розподілу.

Емпіричне кореляційне відношення — це корінь квадратний з коефіцієнта детермінації.

Інтервальний ряд динаміки — це числовий ряд динаміки, який характеризує зміни в часі розмірів суспільних явищ, рівні яких подано за певний період часу.

Інтерполяція — розрахунок приблизних рівнів, які знаходяться в середині ряду динаміки, але з яких-небудь причин невідомі.

Екстраполяція — визначення рівнів за межами ряду, що досліджується, тобто продовження ряду на основі виявленої закономірності рівнів за певний термін часу.

Квадратичний коефіцієнт варіації — це відсоткове відношення середнього квадратичного відхилення до середньої величини ознаки.

Квартилі — це варіанти, які поділяють ранжирувану сукупність на чотири рівновеликі частини.

Коефіцієнт детермінації — це відношення міжгрупової дисперсії до загальної.

Коефіцієнт осциляції — це відсоткове відношення розміру варіації до середньої величини ознаки.

Комбінаційна статистична таблиця — це таблиця, підметом якої є групування одиниць сукупності одночасно за двома і більше ознаками.

Кореляційна залежність є підвидом стохастичної залежності: зі зміною факторної ознаки x змінюються групові середні результативної ознаки y , тобто замість умовних розподілів порівнюються середні значення цих розподілів.

Лінійний коефіцієнт варіації — це відсоткове відношення середнього лінійного відхилення до середньої величини ознаки.

Лінія регресії — це головна характеристика кореляційного зв'язку. Лінія регресії y на x — це функція, яка зв'язує середні значення ознаки y зі значенням ознаки x . Залежно від форми лінії регресії розрізняють лінійний і нелінійний зв'язки.

Медіана — у статистиці це значення ознаки y тієї одиниці сукупності, яка знаходиться в середині упорядкованого ряду, тобто це варіанта, яка знаходиться в середині упорядкованого варіаційного ряду і поділяє його на дві рівні частини.

Метод аналітичного групування полягає в тому, що всі елементи вихідної інформації групуються за факторною ознакою x , далі в кожній групі обчислюються середні значення результативної ознаки y .

Метод плінних (ковзних) середніх — один із прийомів виявлення тенденції

Міжгрупова варіація — це результат впливу фактора, який покладено в основу групування, внутрішньогрупова — інших факторів, окрім груповального.

Основна тенденція (тренд) — це достатньо стійка зміна рівня явища в часі, більш-менш вільна від випадкових коливань. Основну тенденцію можна подати аналітично — у вигляді рівняння (моделі) тренда або графічно.

Показники варіації — це показники, які визначають міру варіації (коливання) окремих значень ознаки від середньої величини.

Правило розкладання (декомпозиції) варіації: для статистичної сукупності, яка згрупована за певною ознакою, можливо визначити три види дисперсій: загальну, внутрішньогрупову та міжгрупову. Загальна дисперсія характеризує варіацію усіх одиниць сукупності від загальної середньої, тобто варіацію ознаки навколо загальної середньої, внутрішньогрупові — варіацію ознаки у групах від групової середньої, а міжгрупова — варіацію групових середніх від загальної середньої.

Простий випадковий добір — це вибірка, за якої добір одиниць (або груп одиниць) для обстеження здійснюється з генеральної сукупності не передбачено, а випадково.

Рівень істотності — це така ймовірність, за якої ймовірність отримання значення η^2 , більшого від критичного (за умови відсутності зв'язку між ознаками), була б достатньо малою.

Рівень ряду динаміки — абсолютна (відносна, середня) величина кожного члена динамічного ряду.

Рівняння регресії показує типове в певних умовах співвідношення між розмірами ознаки-фактора і результативної ознаки, тоді як рівняння функціонального зв'язку справедливе лише для кожного окремого випадку.

Розмах варіації — це різниця між максимальним та мінімальним значеннями ознаки, яка варіює.

Розшарований (районований) добір — це спосіб формування з урахуванням структури генеральної сукупності; він передбачає її попередню структурування і незалежний добір елементів у кожній складовій.

Ряд розподілу — це упорядкований розподіл одиниць сукупності на групи за певною ознакою, яка варіює.

Ряди динаміки — це розміщені в часі значення явища, тобто послідовність чисел, які характеризують зміни розмірів суспільних явищ у часі.

Середнє квадратичне відхилення — це корінь квадратний з дисперсії.

Середнє лінійне відхилення — це середня арифметична з абсолютних значень відхилень варіант ознаки від їхньої середньої.

Середній або середньорічний абсолютний приріст — це показник ряду динаміки, який показує на скільки одиниць у середньому за одиницю часу (щорічно) за певний період змінювався рівень показника, що аналізується.

Середній або середньорічний темп зростання — це показник ряду динаміки, який показує у скільки разів у середньому за одиницю часу (щорічно) за певний період змінювався рівень показника, що аналізується.

Середній або середньорічний темп приросту — це показник ряду динаміки, який показує на скільки відсотків у середньому за одиницю часу (щорічно) за певний період змінювався рівень показника, що аналізується.

Середній рівень динамічного ряду — середня, обчислена на основі рівнів динамічного ряду.

Середня арифметична зважена використовується у випадках, коли окремі значення ознаки, за якою розраховується середня величина, можуть повторюватись по кілька разів, тобто в тих випадках, коли розрахунок здійснюється за згрупованими даними.

Середня арифметична проста використовується в тих випадках, коли розрахунок здійснюється за не згрупованими даними.

Середня величина — це узагальнююча характеристика сукупності однотипних явищ за ознакою, що варіює, тобто це узагальнюючий показник, який характеризує типовий рівень ознаки, що варіює, в розрахунку на одиницю однорідної сукупності.

Середня гармонічна — розраховується з відносних значень усередненої ознаки і за формою може бути простою і зваженою.

Середня інтервального ряду — розраховується за формулою середньої арифметичної простої. Середня арифметична зважена з усередненим рівнем, тобто середня модифікована, використовується для моментного ряду, коли даних недостатньо та інтервали між наданими моментами часу нерівні.

Середня хронологічна моментного ряду — розраховується як сума всіх рівнів ряду, поділеного на число членів ряду без одного, причому перший і останній члени ряду беруться у половинному розмірі за умови, що даних моментного ряду недостатньо, але інтервали між поданими моментами часу рівні.

Серійний добір передбачає вивчення не окремих одиниць сукупності, а їх серій або гнізд.

Симетричний розподіл — рівновіддалені від центра значення ознаки мають однакові частоти.

Стандартна (середня) похибка вибірки є середнім квадратичним відхиленням вибірових оцінок від значення параметра в генеральній сукупності та характеризує середню величину можливих відхилень вибіркової і генеральної середньої.

Стохастичний зв'язок виявляється зміною умовних розподілів, тобто за цього зв'язку кожному значенню ознаки x відповідає певна множина значень ознаки y , які варіюють і утворюють ряд розподілу.

Структурне групування — це групування, за якого відбувається розподіл однорідної сукупності на групи, що характеризують її структуру за певною ознакою, яка варіює (змінюється).

Темп зростання або коефіцієнт зростання — це характеристика ряду динаміки, який показує у скільки разів змінився поточний рівень показника, порівняно з рівнем попереднього або базового періоду.

Темп приросту — це показник ряду динаміки, який показує на скільки відсотків змінився поточний (порівнюваний) рівень аналізованого показника порівняно з рівнем попереднього або базового періоду.

Функціональний зв'язок між явищами характеризується повною відповідністю між причиною і наслідком, факторною і результативною ознакою, тобто за цього зв'язку кожному можливому значенню факторної ознаки x відповідає чітко визначене значення результативної ознаки y .

Частка — це частота, яка наведена відносною величиною у формі коефіцієнта чи відсотка суми.

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

Основна

1. Алан Дебни. Статистика, базовий курс в коміксах. Підручник / Алан Дебни, Грейди Клейн. — К.: Центр учбової літератури, 2016. — 240 с.

2. Заєць С.В., Томіленко В.М. Статистика - Statistics: підручник. - Ірпінь: ВЦ НУДПСУ, 2015. - 512 с.
3. Єріна А.М. Статистичне моделювання та прогнозування: підр. /А.М.Єріна, Д.Л.Єрін. – Статистика підприємств: навч. посіб./ за ред.: С.О.Матковський ,О.С. Гринькевич ., О.З.Сорочак та інш. - 2-ге вид, перероб. і доп. ; з додатками. - Київ: Алерта, 2013. - 560 с.
4. Костюк В. О. Прикладна статистика: навч. посібник / В. О. Костюк; Харк. нац. ун-т міськ. госп-ва ім. О. М. Бекетова. – Харків: ХНУМГ ім. О. М. Бекетова, 2015. – 191 с.
5. Статистика підприємств: навч. посіб./ за ред.: С.О.Матковський ,О.С. Гринькевич ., О.З.Сорочак та інш. - 2-ге вид, перероб. і доп. ; з додатками. - Київ: Алерта, 2013. - 560 с.
6. Краєвський В. М. Статистика: навчальний посібник / В. М. Краєвський, Я. О. Остапенко, Н. В. Параниця. – Ірпінь : Університет ДФС України, 2019. – 218 с.
7. Левченко Л.О., Кілянчук О.П., Повханич О.Ю. Огляд програмних продуктів фінансово-економічного аналізу діяльності енергопідприємств [Електронний ресурс]. – Режим доступу: <http://urss.knuba.edu.ua/files/zbirnyk-8/121-127.pdf> (дата звернення 01.07.2017). – Назва з екрану.
8. Лугінін О.Є. Л 83 Статистика. Підручник. 2 е видання, перероблене та доповнене – К.: Центр учбової літератури, 2007. – 608 с.
9. Майборода Р.Є., Сугакова О.В. Статистичний аналіз даних за допомогою пакету STATISTICA [Електронний ресурс]. – Режим доступу: <http://matphys.rpd.univ.kiev.ua/downloads/courses/mmatstat/StatAn.doc> (дата звернення 01.07.2017). – Назва з екрану.
10. Опря А.Т. Статистика. Підручник / А.Т. Опря. – К.: Центр учбової літератури, 2017. – 448 с2
11. Програми для математичної і статистичної обробки даних [Електронний ресурс]. – Режим доступу: <http://chem-bio.com.ua/aspirant/grant/item/> (дата звернення 01.07.2017). – Назва з екрану.
12. Функції Excel (за категоріями) [Електронний ресурс]. – Режим доступу: <https://support.office.com/uk-ua/article> (дата звернення 01.07.2017). – Назва з екрану.

Допоміжна

1. Горкавий В.К. Статистика: Навч. посібник. – К.: Алерта, 2012. – 608 с.
2. Моделі і методи соціально-економічного прогнозування : підручник / В. М. Геєць, Т. С. Клебанова, О. І. Черняк та ін. – Х. : ВД "ІНЖЕК", 2005. – 396 с
3. Яцура В. В. Соціально-економічне прогнозування : навч. посібн. / В. В. Яцура, О. С. Сенишин, М. О. Горинь. – Львів : Видавничий центр ЛНУ імені Івана Франка, 2006. – 264 с.

Інформаційні ресурси Інтернет

1. Офіційний сайт Міністерства фінансів України: <http://www.minfin.gov.ua>
2. Офіційний сайт Міністерства економічного розвитку і торгівлі України: <http://www.me.kmu.gov.ua>
3. Офіційний сайт Національного банку України: <http://www.bank.gov.ua>
4. Офіційний сайт Державної податкової служби України: <http://www.sta.gov.ua>
5. Офіційний сайт Державної служби статистики України: <http://www.ukrstat.gov.ua>
6. Офіційний сайт Державної казначейської служби України: <http://www.treasury.gov.ua>
7. Офіційний сайт Національної комісії з регулювання ринків фінансових послуг України: <http://www.dfp.gov.ua>
8. Офіційний сайт Державної комісії з цінних паперів та фондового ринку України: <http://www.ssmc.gov.ua>

Міжнародні видання

1. David Ruppert. Statistics and finance. Springer-Verlag New York LLC. 2014. – 476 p.
2. Government finance statistics manual 2014. – Washington, D.C. : International Monetary Fund, 2014. – 471 p.
3. International Financial Statistics Monthly Print Subscription. World Economic Outlook, October 2016 : Subdued Demand: Symptoms and Remedies. 2016. – 277 p.
4. Author/Editor: Maria Luz Moreno Badia Fiscal Monitor, October 2016 : Debt: Use It Wisely. 2016. – 74 p.
5. Global Financial Stability Report, October 2016 : Fostering Stability in a Low-Growth, Low-Rate Era. 2016. – 106 p.
6. Robert the Evangelist. Business Winning: A Guide for a Manager to Create a The Business of Winning: A Manager's Guide to Building a Championship Team at Work, 2005. – 248 c.