

РОЗДІЛ 1
Статистичні методи в географії: основні поняття і терміни.
Статистичні показники

ТЕМА 1

Статистичні методи в географії: основні поняття і терміни

Лекція 1. Статистичні методи в географії: основні поняття і терміни. Об'єкт та мета курсу. Основні поняття і терміни курсу «Статистичні методи і обробка геоінформації». Методи, методика і методологія. Статистика і геостатистика. Модель і моделювання. Географічні, математичні і статистичні моделі. Географічне, математичне і статистичне моделювання. Математичні, статистичні та математико-статистичні методи. Географічна і статистична інформація, математичні і статистичні дані.

Класифікації моделей: натурні, аналогові, математичні. Історія розвитку і використання моделей в природознавстві і суспільних науках. Три рівні математизації географії: рівень параметризації географічних явищ та об'єктів, рівень емпіричних моделей, рівень теоретичних моделей. Місце методів ідеалізації, формалізації, математичних методів, методів моделювання в методології географічного дослідження. Протиріччя системного і синергетичного підходів. Перспективи розвитку моделювання в географії.

Географічна інформація та її види.

Статистична інформація. Способи отримання статистичної інформації: документальний (викопіювання даних); опитувальний спосіб (анкетний, кореспондентський, самореєстрації, експедиційний); спосіб безпосереднього спостереження.

Статистичне дослідження і статистичне спостереження. Види статистичного дослідження: в залежності від характеру проведення спостереження за часом, в залежності від повноти обліку фактів спостереження, за способами отримання статистичної інформації. Відмінні риси (специфіка) статистичного дослідження.

Етапи статистичного дослідження. I етап: визначення мети, завдань, розробка плану та програми дослідження. II етап: збір статистичного матеріалу. III етап: зведення й обробка матеріалу. IV етап: аналіз, інтерпретація та порівняння даних, літературне оформлення та висновки за результатами статистичного дослідження і пропозиції щодо впровадження в практику.

ТЕМА 2

Статистичні показники.

Абсолютні, відносні і середні статистичні показники

Лекція 2-3. Статистичні показники. Абсолютні і відносні статистичні показники. Поняття «статистичний показник». Якісний, кількісний, часовий і просторовий зміст статистичних показників. Функції статистичних показників. Класифікація статистичних показників за різними критеріями: за способом одержання (первинні і вторинні), за часовою ознакою (моментні та інтервальні), за суттю досліджуваних явищ (об'ємні та якісні), за ступенем агрегування досліджуваних явищ (індивідуальні, групові, загальні, зведені), за об'єктом дослідження (територіальні, галузеві та сукупні), за взаємозалежністю (прямі та обернені), за можливістю узагальнення / сумування (адитивні та неадитивні), за формою (абсолютні, відносні та середні).

Абсолютні показники, їх види та форми. Одиниці вимірювання: натуральні, умовно-натуральні, вартісні, трудові, одиниці часу. Прості і складні одиниці вимірювання.

Відносні показники, їх види та форми. Одиниці вимірювання: коефіцієнти, відсотки, проміле, продециміле, іменовані числа.

Відносні показники динаміки (темпи зростання). Темпи росту і темпи приросту, розраховані ланцюговим і базисним способами. Відносні показники структури і координат. Відносні показники порівняння. Просторове порівняння і порівняння зі стандартом. Відносні показники планового завдання і виконання плану. Відносні показники інтенсивності. Відносні показники диференціації.

Індекси та їх використання в географічних дослідженнях.

Лекція 4. Статистичні показники. Середні статистичні показники. Середні показники: визначення поняття, види та форми. Ступеневі і структурні середні показники. Прості і зважені середні показники. Умови застосування середніх показників. Переваги і недоліки середніх показників.

Середня арифметична: проста і зважена; для дискретного та інтервального рядів (із закритими та відкритими і інтервалами). Властивості середньої арифметичної. Інші види ступеневих середніх показників: середня гармонічна, середня геометрична, середня квадратична.

Мода: визначення поняття, методика обчислення для дискретних та інтервальних рядів. Переваги і недоліки моди.

Медіана: визначення поняття, методика обчислення для дискретних (з парним та непарним числом варіант, з різною частотою варіант) та інтервальних рядів. Переваги і недоліки медіани.

ТЕМА 3

Статистична база даних. Матриці даних

Лекція 5. Статистична база даних. Матриці даних. Статистична інформація, статистичні дані, статистичні показники, статистична сукупність: визначення і сутність понять.

Статистична сукупність: проста статистична сукупність або статистичний ряд, генеральна (Population) і вибіркова сукупність або статистична вибірка (Sample). Одиниця статистичної сукупності та одиниця статистичного спостереження. Репрезентативна і нерепрезентативна вибірка. Методи відбору до вибіркової сукупності: випадковий відбір, механічний відбір, середній або гніздовий відбір, направлений відбір, типологічний відбір. Властивості статистичної сукупності: розподіл ознаки (частота, частка ознаки), середній рівень ознаки, різноманітність ознаки, репрезентативність ознаки, взаємозв'язок між ознаками.

Статистичні ознаки та їх класифікація. Атрибутивні (описові, якісні) та варіаційні (кількісні) ознаки. Варіація.

Статистичні ряди розподілу та їх види. Атрибутивні (якісні) та варіаційні (кількісні) ряди розподілу (дискретні, інтервальні). Ранжировані та неранжировані ряди розподілу. Статистичні шкали та їх види.

Бази даних. Матриці даних та вимоги до них.

Теорія ймовірності та закон великих чисел.

РОЗДІЛ 2

Статистичне моделювання

ТЕМА 1

Одновимірні статистичні моделі та їх особливості

Лекція 6. Одновимірний статистичний аналіз. Статистичний аналіз: поняття, методи, види. Дескриптивний (описовий) та аналітичний (індуктивний) аналіз. Методи статистичного аналізу. Одновимірний статистичний аналіз. Одновимірні статистичні моделі. Умови застосування одновимірних статистичних моделей.

Поняття про випадкову величину, дискретні і неперервні випадкові величини. Числові характеристики випадкових величин: частота, частість, функція розподілу.

Статистики одновимірного аналізу. Значення процентилей (Percentile): проценти́лі, квартилі (Quartile), міжквартильний розмах (Interquartile range). Міри центральної тенденції (Central Tendency): мода (Mode), медіана (Median), середнє (Mean). Міри варіації (Dispersion): стандартне відхилення (Std. deviation), дисперсія (Variance), розмах (Range), мінімум (Minimum), максимум (Maximum). Характеристики форми розподілу (Distribution): асиметрія (Skewness), ексцес (Kurtosis). Підсумок з п'яти значень: мінімум (Minimum), перший (або нижній) квартиль (Q25 або Q1 або lower half), медіана (Median, Q50 або Q2), третій (або верхній) квартиль (Q75 або Q3 або upper half), максимум (Maximum).

Нормальний розподіл випадкових величин (Normal distribution). Крива нормального розподілу (або крива Гаусса).

Візуалізація одновимірного аналізу. Способи подання результатів одновимірного аналізу. Ящиківі (коробчасті) діаграми (boxplot, «ящик з вусами», біржова діаграма).

Центральний момент першого, другого, третього порядків.

Лекція 7. Поточна контрольна робота

ТЕМА 2

Двовимірні статистичні моделі і особливості їх застосування

Лекція 8. Кореляційний статистичний аналіз. Двовимірні статистичні моделі. Поняття про двовимірну випадкову величину. Умови формування двовимірних випадкових величин у географічних дослідженнях. Особливості формування вибірових сукупностей. Особливості інтерпретації зв'язків компонентів двовимірних випадкових величин. Поняття про форму залежності і силу зв'язку двох випадкових величин. Функціональні, кореляційні і стохастичні залежності.

Основні поняття кореляційного аналізу: кореляція, кореляційний зв'язок, кореляційна залежність, кореляційний аналіз. Кореляційний аналіз: поняття, мета, завдання, види. Умови коректності кореляційного аналізу.

Кореляційні залежності та їх види за формою (лінійна та нелінійна), напрямом (пряма і непряма), силою (сильна і слабка), значущістю (статистично значуща і статистично незначуща). Коефіцієнт кореляції: різноманіття і правила вибору. Принципи побудови коефіцієнтів кореляції. Лінійний коефіцієнт кореляції Пірсона r (коефіцієнт парної кореляції) та коефіцієнт рангової кореляції Спірмена ρ . Формула коефіцієнту кореляції. Коефіцієнт детермінації.

Кореляція і причинно-наслідковий зв'язок. Псевдокореляція.

Візуалізація кореляційного аналізу. Діаграма розсіювання.

Особливості використання кореляційного аналізу у географічних дослідженнях.

Лекція 9. Регресійний статистичний аналіз. Кореляційний і регресійний аналіз: риси схожості і відмінності. Регресійний аналіз: поняття, мета, завдання. Види регресійного аналізу. Лінійний регресійний аналіз. Етапи регресійного аналізу: 1. Виявлення наявності взаємозв'язку між ознаками і її форми. 2. Побудова регресійної моделі: лінія регресії і рівняння регресії. 3. Прогноз.

Регресійна модель. Рівняння регресії і лінія регресії. Метод найменших квадратів. Показники регресійного аналізу: коефіцієнт кореляції, коефіцієнт регресії, коефіцієнт детермінації, критерій t-Ст'юдента, критерій f-Фішера, стандартизовані коефіцієнти регресії (Бета β), p-рівень статистичної значущості. Аналіз залишків: 1) випадковий характер залишків; 2) нульова середня величина залишків, незалежна від x ; 3) гомоскедастичність – дисперсія кожного відхилення ε , однакова для всіх значень x ; 4) відсутність автокореляції залишків – значення залишків розподілені незалежно один від одного; 5) залишки підкорюються нормальному розподілу

Припущення регресійного аналізу. Припущення лінійності. Припущення нормальності. Обмеження регресійного аналізу. Кількість незалежних ознак для множинного регресійного аналізу: вибір кількості змінних, принцип парсимонії, принцип здорового глузду. Вимоги до регресійних моделей. Побудова довірчого інтервалу для рівнянь регресії.

Особливості використання регресійного аналізу у географічних дослідженнях.

ТЕМА 3

Багатовимірні статистичні моделі

Лекція 10. Кластерний статистичний аналіз. Кластерний аналіз: поняття, задачі, види. Жорстка і м'яка кластеризація. Кластер та його характеристики: центр кластера, радіус кластера, розмір кластера. Стандартизація (standardization) або нормування (normalization) бази даних. Лінійне шкалювання.

Методи кластерного аналізу. Ієрархічні методи кластерного аналізу (агломераційні і дивизимні). Неієрархічні методи кластерного аналізу (ітеративні, факторні, методи згущень, методи, які використовують теорію графів). Алгоритм ієрархічного агломеративного кластерного аналізу. Алгоритм ітераційного кластерного аналізу.

Міри подібності (евклідова відстань, манхеттенська відстань, відстань Чебишева, відсоток незгоди). Сутність евклідової відстані. Методи об'єднання або зв'язки (метод Варда, метод одиничного зв'язку, метод повного зв'язку, метод середнього зв'язку, центроїдний метод). Сутність методу Варда. Дендрограма кластеризації.

Лекція 11. Факторний статистичний аналіз. Факторний аналіз: визначення поняття, мета і завдання. Види факторного аналізу. Метод головних компонент (principal component analysis, PCA) як головний метод факторного аналізу. Інші методи факторного аналізу.

Етапи факторного аналізу: 1. Підготовка даних до факторного аналізу. 2. Визначення кількості факторів (за критеріями Кайзера, кам'янистого осипу Кеттела, інтерпретації та інваріантності, кумулятивним відсотком). Власні значення (eigenvalues) факторів. 3. Обертання факторів – ортогональне і косокутне. Метод Varimax (варімакс). 4. Ідентифікація факторів. Факторні навантаження (factor loadings). 5. Обрахунок факторних ваг (factor scores). 6. Інтерпретація факторів.

РОЗДІЛ 3

Моделювання просторових змінних

ТЕМА 1

Прогнозування географічних процесів

Лекція 12. Прогнозування географічних процесів. Значення географічного прогнозування. Географічні прогнози як інформаційно-аналітична база обґрунтування заходів регіональної політики. Прогнозування як функція держави. Правові засади прогнозування соціально-економічного розвитку в Україні. Закон України «Про державне прогнозування та розроблення програм соціально-економічного розвитку». Науково-практичне значення прогнозування. «За» та «проти» прогнозування в сучасному суспільстві. Проблеми державного прогнозування регіонального розвитку.

Прогностичні ефекти: самоліквідація (саморуйнація) прогнозу, самоорганізація (самоздійснення) прогнозу, прогностична паніка, парадокс прогностичного зрушення, прогностична мода, прогностична сліпота. Ефект Едіпа: «самоздійснюваний прогноз». Ефект Пігмаліона (ефект Розенталя): «самовтілюване пророцтво». Осборн-ефект: «самоліквідація прогнозу». Ефект «чорного лебедя». Ефект (інформація) Касандри. Метод «5 ЧОМУ?» (метод тойоти).

Поняттєво-термінологічний апарат. Передбачення: наукове і ненаукове. Прогноз як наукове передбачення. Прогноз, прогнозування, прогностика. Відмінності у поняттях прогноз, гіпотеза, план, програма, проект. Період (глибина) ретроспекції, тренд, прогнозний горизонт. Довірчий інтервал прогнозу. Прогностичний фон. Варіант прогнозу. Прогнозна альтернатива. Географічний прогноз і прогнозування. Географічне прогнозування як процес наукового передбачення розвитку в часі географічних процесів, станів просторової організації регіонів. Верифікація прогнозів. Достовірність, точність прогнозів. Похибка як міра точності прогнозів. Джерела похибок. Міра якості прогнозу. Абсолютна похибка прогнозу, квадратична похибка прогнозу, коефіцієнт розбіжності.

Види та способи географічного прогнозування. Умови застосування екстраполяційного способу прогнозування.

Географічна прогностика як наука.