




Аналіз структури сукупності та її змін



Структура – взаємне розміщення груп, зв'язки між якими визначають специфіку сукупності, що вивчається.

Графічно доцільно зображати структуру у вигляді кругових або стовпчастих діаграм.

Метою структурного аналізу є опис закономірностей формування сукупності одиниць, що спостерігаються, тобто проведення типологізації елементів сукупності та виділення пріоритетного, домінантного та мозозначимого типів одиниць.

Система показників для вимірювання структури та відмінностей структури включає:

- *індивідуальні характеристики* – частка, накопичена частка, абсолютна та відносна зміна часток груп, коефіцієнти координації, коефіцієнти навантаження;
- *узагальнюючі характеристики* - характеристики властивостей однієї структури; узагальнюючі індекси порівняння двох структур (у часі та просторі); узагальнюючі індекси порівняння декількох (більше двох) структур (у часі та просторі).

Характеристики властивостей однієї структури – система узагальнюючих показників розподілу, в тому числі екстремальних значень показника у сукупності в цілому та за групами, середній рівень показника, що варіюється, (з урахуванням асиметричності розподілу), характер варіації та форму розподілу.

Для узагальненого аналізу зміни структури сукупності в часі або зміни просторових відмінностей за двома порівняльними рівнями використовують різні узагальнюючі показники структурних зсувів.

Узагальнюючі показники абсолютного значення структурних зсувів

Індекс відмінностей:

$$I_{\text{відм}} = \frac{1}{2} \sum |d_1 - d_0|$$

d_1, d_0 - показники питомої ваги, виражені у простому кратному вираженні

У міжнародній статистичній практиці найбільш часто розраховують:

Лінійний коефіцієнт абсолютних структурних зсувів (Казинца):

$$S_d = \frac{\sum |d_1 - d_0|}{n}$$

d_1, d_0 - показники питомої ваги (у %) окремих елементів сукупності у періоді, що розглядається, та у попередньому періоді;
 n - кількість елементів сукупності, що виділяються.

У міжнародній статистичній практиці найбільш часто розраховують:

Квадратичний коефіцієнт абсолютних структурних зсувів (Казинца):

$$S_{\sigma} = \sqrt{\frac{\sum (d_1 - d_0)^2}{n}}$$

d_1, d_0 - показники питомої ваги (у %) окремих елементів сукупності у періоді, що розглядається, та у попередньому періоді;
 n - кількість елементів сукупності, що виділяються.

Лінійний та квадратичні коефіцієнти абсолютних структурних зсувів показують, на скільки відсоткових пунктів у середньому відхиляються один від одного питомі ваги, що порівнюються.

За відсутності зсувів у структурі сукупності ці показники дорівнюють 0.

Верхньої межі зміни коефіцієнти не мають: чим більші зміни структури, тим вищим є значення коефіцієнтів.

Використання квадратичного коефіцієнта є переважним, оскільки він більш чуйно на сильні коливання структури.

Узагальнюючі показники відносного значення структурних зсувів:

- індекс Салаї;
- індекс Гатєва;
- індекс Рябцева;
- індекс кореляції Спірмена.

Узагальнюючі показники відносного значення структурних зсувів

Інтегральний коефіцієнт структурних зсувів К. Гатєва :

$$J_G = \sqrt{\frac{\sum (w_1 - w_0)^2}{\sum (w_1^2 + w_0^2)}}, \quad 0 < J_G < 1$$

w_1, w_0 - частка окремих елементів сукупності у періоді, що розглядається, та у попередньому періоді

Узагальнюючі показники відносного значення структурних зсувів

Інтегральний коефіцієнт структурних зсувів Салаї:

$$J_S = \sqrt{\frac{\sum \left(\frac{(w_1 - w_0)}{(w_1 + w_0)} \right)^2}{n}}, \quad 0 < J_S < 1$$

n - кількість груп

Узагальнюючі показники відносного значення структурних зсувів

Інтегральний коефіцієнт структурних відмінностей

Рябцева:

$$J_R = \sqrt{\frac{\sum (w_1 - w_0)^2}{\sum (w_1 + w_0)^2}}, \quad 0 < J_R < 1$$

n - кількість груп

Узагальнюючі показники відносного значення структурних зсувів

Коефіцієнт рангової кореляції Спірмена:

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

у випадку незв'язних рангів

Узагальнюючі показники відносного значення структурних зсувів

Коефіцієнт рангової кореляції Спірмена:

$$\rho = 1 - \frac{6 \sum d^2 - A - B}{\sqrt{(n^3 - n - 12A)(n^3 - n - 12B)}},$$

$$A = \frac{1}{12} \sum_j (A_j^3 - A_j), B = \frac{1}{12} \sum_k (B_k^3 - B_k)$$

у випадку зв'язних рангів

j - номер зв'язок для першої ознаки

A_j - кількість однакових рангів для першої ознаки у відповідній зв'язці

k - номер зв'язок для другої ознаки

B_k - кількість однакових рангів для другої ознаки у відповідній зв'язці

Показники одночасного співставлення декількох структур (у часі та просторі)

Коефіцієнт нерівномірності:

$$k_{\text{нерівн}} = \frac{K \cdot L}{K - L} \sum (w_i - p)^2, \quad 0 < k_{\text{нерівн}} < 1$$

p – частка груп при абсолютно рівномірному розподілі ($p = \frac{1}{K}$);

K – кількість груп у структурах, що порівнюються;

L – кількість домінантних груп у структурах.

Показники одночасного співставлення декількох структур (у часі та просторі)

Коефіцієнт нерівномірності:

$k_{\text{нерівн}} \rightarrow 0$ – рівномірний розподіл (різноманітність у структурі);

$k_{\text{нерівн}} \rightarrow 1$ – посилення одноманітності у структурі (нерівномірний розподіл).

Для аналізу структури сукупності важливо оцінити форму розподілу:

- коефіцієнт асиметрії за Пірсоном;
- коефіцієнт асиметрії за Ліндбергом;
- коефіцієнт асиметрії на основі моменту розподілу 3-го порядку;
- коефіцієнт ексцесу.

Для оцінки рівня диференціації та концентрації аналізується ряд розподілу:

- коефіцієнт Лоренца;
- коефіцієнт Джині;
- коефіцієнт Херфіндаля-Хіршмана.

Для оцінки міри невизначеності атрибутивної ознаки використовується коефіцієнт ентропії:

абсолютна ентропія:

$$H(x) = - \sum_{i=1}^k p(x_i) \log_2 p(x_i)$$

x_i - частка одиниць i -тої категорії

k - кількість категорії

$$0 \leq H(x) \leq H_{\max}$$

$$H_{\max} = - \sum_{i=1}^k \frac{1}{k} \log_2 \frac{1}{k}$$

Для оцінки міри невизначеності атрибутивної ознаки використовується коефіцієнт ентропії:

відносна ентропія:

$$H^* = \frac{H(x)}{H_{\max}}$$

k - кількість категорії

$$H_{\max} = - \sum_{i=1}^k \frac{1}{k} \log_2 \frac{1}{k}$$