

Класифікація статистичних рядів як елементів багатомірного масиву даних

Розглянуто статистичні ряди як одномірні масиви та як елементи трьохмірного масиву даних, на основі чого доповнено існуючу класифікацію рядів поняттям «монографічний ряд», елементами якого є значення множини досліджуваних ознак одиниці сукупності в статистиці.

Statistical rows as one-dimensional arrays and as elements of a three-dimensional array are considered. Also it is added existing classification of statistical rows by concept "a monographic row". Values of a monographic row is a set of indications one elements of a population in a statics.

Ключові слова: *статистичні ряди, масиви, багатомірний аналіз, монографічний ряд.*

Вступ. Сучасне економічне життя нерозривно пов'язане з процесами накопичення та аналізу інформації. За оцінками деяких експертів, світові об'єми інформації подвоюються кожні два роки [1]. Водночас традиційні засоби аналізу, що оперують даними таблиць реляційних баз даних, уже не є достатньо ефективними, так як все частіше постає потреба побудови залежностей між параметрами, число яких може варіювати у широких межах. Ще у 1993 році засновник реляційної моделі бази даних Е.Ф. Кодд розглянув її недоліки та вказав на неможливість «об'єднувати, проглядати та аналізувати дані з точки зору множини вимірів, способом, котрий є найприйнятнішим для аналітиків» [2]. Під виміром тут розуміється послідовність значень одного з параметрів, по якому аналізується досліджувана ознака (наприклад, для параметру «час» це може бути послідовність календарних днів). Множинність вимірів, у свою чергу, потребує актуалізації даних у вигляді багатомірної моделі.

Як це буде розкрито нижче, проведення статистичного дослідження передбачає формування саме багатомірного масиву даних як множини значень досліджуваних ознак сукупності одиниць в динаміці.

Циганець В.В., к.е.н., викладач кафедри обліку і аудиту, Прикарпатський національний університет ім. В.Стефаника

Постановка завдання. Багатомірна модель представлення даних вимагає перегляду теорії статистичних рядів у новому цілісному дефрагментованому світі багатомірного аналізу.

Результати. Загальноприйнятою є класифікація статистичних рядів на ряди розподілу та ряди динаміки [3;4;5;6].

Ряд розподілу – це результат ранжування та групування первинного ряду даних. Так званий первинний ряд – це множина значень ознаки сукупності одиниць у статистиці. Ряд динаміки – це множина значень ознаки одиниці сукупності в динаміці. Але окрім ряду розподілу, що описує ознаку досліджуваної сукупності одиниць та динамічного ряду, що описує ознаку одиниці досліджуваної сукупності, теоретично існує ряд, що описує не певну ознаку, але множину ознак одиниці сукупності, тому існуюча класифікація статистичних рядів, на наш погляд, є недосконалою та потребує доопрацювання.

Статистичний ряд являє собою множину елементів x_{kmt} , кожний з яких – це певне значення (у кількісному або якісному виразі) k -ознаки m -одиниці за t -період часу.

Первинний ряд даних ряду розподілу можна представити у вигляді множини M :

$$M = \{x_1, x_2, x_3, \dots, x_m\}, \quad k = 1, \quad m > 1, \quad t = 1$$

де M – одномірний масив значень k -ознаки ($k = 1$) m -одиниць ($m > 1$) у статистиці ($t = 1$).

Ряд динаміки являє собою множину T :

$$T = \{x_1, x_2, x_3, \dots, x_t\}, \quad k = 1, \quad m = 1, \quad t > 1$$

де T – одномірний масив значень k -ознаки ($k = 1$) m -одиниці ($m = 1$) за t -періодів часу ($t > 1$).

Сукупність первинних рядів (t -рядів M_t), рівно як і сукупність динамічних рядів (m -рядів T_m) утворить матрицю MT (TM):

$$MT = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1t} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2t} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3t} \\ \dots & \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & x_{m3} & \dots & x_{mt} \end{pmatrix}$$

$$TM = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1m} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2m} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3m} \\ \dots & \dots & \dots & \dots & \dots \\ x_{t1} & x_{t2} & x_{t3} & \dots & x_{tm} \end{pmatrix}$$

$$MT = \{M_1, M_2, M_3, \dots, M_t\}$$

$$TM = \{T_1, T_2, T_3, \dots, T_m\}$$

$$MT = TM$$

$$k = 1, \quad m > 1, \quad t > 1$$

де MT (TM) – двохмірний масив значень k -ознаки ($k = 1$) m -одиниць ($m > 1$) за t -періодів часу ($t > 1$).

Сукупність k -матриць MT_k , рівно як і матриця mt -рядів K_{mt} утворить масив MTK (KMT):

$$MTK = \{MT_1, MT_2, MT_3, \dots, MT_k\}$$

$$KMT = \begin{pmatrix} K_{11} & K_{12} & K_{13} & \dots & K_{1t} \\ K_{21} & K_{22} & K_{23} & \dots & K_{2t} \\ K_{31} & K_{32} & K_{33} & \dots & K_{3t} \\ \dots & \dots & \dots & \dots & \dots \\ K_{m1} & K_{m2} & K_{m3} & \dots & K_{mt} \end{pmatrix}$$

$$MTK = KTM$$

$$k > 1, \quad m > 1, \quad t > 1$$

де MTK (KMT) – трьохмірний масив значень k -ознак ($k > 1$) m -одиниць ($m > 1$) за t -періодів часу ($t > 1$).

Одномірний масив K буде являти собою статистичний ряд, що описує не лише певну ознаку, але множину ознак одиниці сукупності. Назвемо такий ряд *монографіч*

ним:

$$K = \{x_1, x_2, x_3, \dots, x_k\}, \quad k > 1, \quad m = 1, \quad t = 1$$

де K - одномірний масив значень k -ознаки ($k = 1$) m -одиниці ($m = 1$) у статистиці ($t = 1$).

Отже, трьохмірний масив даних можна представити як матрицю kt -рядів (розподілу), матрицю mk -рядів (динаміки) або ж як матрицю mt -рядів (монографічних):

Таблиця 1

Класифікація статистичних рядів як елементів трьохмірного масиву даних

Трьохмірний масив даних		
Ряди розподілу M_{kt}	Ряди динаміки T_{mk}	Монографічні ряди K_{mt}
$MKT = \begin{pmatrix} M_{11} & M_{12} & M_{13} & \dots & M_{1t} \\ M_{21} & M_{22} & M_{23} & \dots & M_{2t} \\ \dots & \dots & \dots & \dots & \dots \\ M_{31} & M_{32} & M_{33} & \dots & M_{3t} \\ \dots & \dots & \dots & \dots & \dots \\ M_{k1} & M_{k2} & M_{k3} & \dots & M_{kt} \end{pmatrix}$	$TMK = \begin{pmatrix} T_{11} & T_{12} & T_{13} & \dots & T_{1k} \\ T_{21} & T_{22} & T_{23} & \dots & T_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ T_{31} & T_{32} & T_{33} & \dots & T_{3k} \\ \dots & \dots & \dots & \dots & \dots \\ T_{m1} & T_{m2} & T_{m3} & \dots & T_{mk} \end{pmatrix}$	$KMT = \begin{pmatrix} K_{11} & K_{12} & K_{13} & \dots & K_{1t} \\ K_{21} & K_{22} & K_{23} & \dots & K_{2t} \\ \dots & \dots & \dots & \dots & \dots \\ K_{31} & K_{32} & K_{33} & \dots & K_{3t} \\ \dots & \dots & \dots & \dots & \dots \\ K_{m1} & K_{m2} & K_{m3} & \dots & K_{mt} \end{pmatrix}$

Висновки. Наведемо найпростіший приклад трьохмірного масиву даних та монографічного ряду.

Якщо журнал обліку успішності академгрупи розглядати як трьохмірний масив KMT , елементами якого є x_{kmt} значення оцінок з k -дисциплін m -студентів за t -днів, тоді кожна сторінка журналу утворить матрицю оцінок x_{mt} з певної k -дисципліни. Також сторінку журналу можна розглядати як одномірний масив t -рядів (розподілу) $MT = \{M_1, M_2, M_3, \dots, M_t\}$ або як одномірний масив m -рядів (динаміки) $TM = \{T_1, T_2, T_3, \dots, T_m\}$.

Якщо ж певний студент ($m = 1$) захоче проаналізувати свою успішність з k -дисциплін за t -днів, тоді він напевне скористається сукупністю монографічних рядів $KT = \{K_1, K_2, K_3, \dots, K_t\}$ або ж сукупністю динамічних рядів $TK = \{T_1, T_2, T_3, \dots, T_k\}$.

Література

1. Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. Методы и модели анализа данных: OLAP и Data Mining. — СПб.: БХВ-Петербург, 2004. - 336 с: ил. - ISBN 5-94157-522-X
2. Л.В. Щавелёв. Оперативная аналитическая обработка данных: концепции и технологии [Электронный ресурс] / Л.В. Щавелёв // Ивановский государственный энергетический университет. - Режим доступа до кн.: http://citforum.ru/seminars/cis99/sch_03.shtml
3. Статистика: теоретичні засади і прикладні аспекти. Навчальний посібник. Р.В. Фещур, А.Ф. Барвінський, В.П. Кічор та інші; За наук. ред. Р.В. Фещура. – 2-е вид. оновлене і доповнене. – Львів: «Інтелект-Захід», 2003. – 576 с. - ISBN 996-7597-27-X
4. Ефимова М.Р., Петрова Е.В., Румянцев В.Н. Общая теория статистики: Учебник. – М.: ИНФРА-М, 1996 – 416 с. – ISBN 5-86225-276-2
5. Никитина Н.Ш. Математическая статистика для экономистов: Учеб. пособие. – 2-е изд. перераб. и доп. – М.: ИНФРА-М; Новосибирск: Изд-во НГТУ, 2001. – 170 с. – (Серия «Высшее образование»). – ISBN 5-16-000793-8 (ИНФРА-М). – ISBN 5-7782-0349-7 (НГТУ)
6. Лялин В.С. Статистика: теория и практика в Excel: учеб. пособие / В.С. Лялин, И.Г. Зверева, Н.Г. Никифорова. — М.: Финансы и статистика; ИНФРА-М, 2010. - 448 с. - ISBN 978-5-279-03381-2 (Финансы и статистика). - ISBN 978-5-16-003908-4 (ИНФРА-М)
7. Орлов А.И. Прикладная статистика. Учебник. / А.И.Орлов.- М.: Издательство «Экзамен», 2004. - 656 с. – ISBN: 5-472-01122-1
8. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы: Учебник. - М.: Финансы и статистика, 2003. – 352 с. – ISBN 5-279-01945-3