

С. В. Чугасвська, к. е. н., доцент,
О. П. Русак, к. е. н., доцент,
кафедра аналізу і статистики ім. І. В. Поповича

ТЕОРЕТИЧНІ ЗАСАДИ МЕТОДИЧНОЇ ОБРОБКИ СТАТИСТИЧНОЇ ІНФОРМАЦІЇ

Постановка проблеми Активізація євроінтеграційного механізму розвитку економіки країни вимагає застосування сучасних методологічних підходів щодо обробки інформації, які б сприяли ефективному прийняттю управлінських рішень. Враховуючи експоненційне зростання обсягів економічної інформації, різноманітність критеріїв оцінювання її якості та принципи моніторингу соціальних процесів актуалізується питання щодо розробки і удосконалення методичного інструментарію статистичної обробки даних на макро- і мікрорівнях.

Аналіз останніх досліджень та публікацій Сучасні методичні підходи щодо статистичної обробки інформації значною мірою відображені у вітчизняних і зарубіжних наукових працях А. Бююля, М. Портера, Г. І. Купалової, П. Т. Саблука, О. В. Ульяновченка та інших. Однак нові виклики ринкового середовища вимагають системного вивчення та постійного удосконалення методичного інструментарію, як допоміжного елементу перспективного бізнесу.

Результати дослідження. В сучасних умовах господарювання обробка масивів статистичних даних здебільшого пов'язана із громіздкими та досить складними розрахунками. Зокрема, при обробці результатів проведеного анкетування виникає необхідність визначення відсоткової величини по кожному запитанню, проведення сортування схожих і близьких між собою відповідей респондентів та виділення найбільш вагомих запитань (факторів), які респонденти вважають найвагомішими. В таких випадках застосовується функція «факторний аналіз» програми SPSS – Statistica [1, с. 368–375].

В науковій літературі від факторами, здебільшого, розглядають сукупність незалежних величин, які зумовлюють вплив на даний процес. У випадку обробки результатів анкетування, під факторами слід розуміти кожне окреме запитання в анкеті, оскільки усі вони є незалежними між собою і повністю розкривають зміст поставлених в роботі завдань. Сутність факторного аналізу вимагає вибрати із більшої кількості факторів меншу так, щоб виділені фактори найтісніше корелювали між собою, тобто мали найтісніший зв'язок.

На практиці проведення факторного аналізу ґрунтується на здійсненні класифікації об'єктів дослідження за допомогою численних обчислювальних процедур. В результаті цього утворюються “факторні групи” або типові групи схожих між собою об'єктів. На відміну від інших методів, цей вид аналізу дає можливість класифікувати об'єкти за однією ознакою, якою виступає в нашому випадку середній бал фактора в анкетуванні. Громіздкі та досить складні розрахунки значно спрощуються з використання спеціальних статистичних програм. Результати обробки матриці вихідних даних програма SPSS – Statistica зображує у вигляді сукупності таблиць: розкладу сукупної дисперсії в одиничному та відсотковому вираженні, прямої та оберненої кореляційних матриць. Обернена кореляційна матриця компонентів дає підстави виділити по кожному фактору номер групи шляхом вибору найвищого коефіцієнта кореляції Пірсона із застосуванням методу обертаного з нормалізацією Кайзера.

Першочерговим етапом проведення факторного аналізу є побудова кореляційної матриці, де по кожному фактору визначаються кореляційні коефіцієнти Пірсона між ними. Далі власні значення сортуються в порядку зменшення, для чого відбирається стільки типових груп, скільки є власних значень змінних, більших за одиницю. Останній етап ґрунтується на побудові факторного навантаження по кожній типовій групі, який відображається коефіцієнтами кореляції між даними змінними, покладеними в основу групування та вибраними факторами.

Процес обробки результатів проведеного анкетування є важливим, оскільки дає змогу здійснити розподіл визначених факторів на певну кількість груп. Кожна виділена група має одну головну компоненту і декілька інших факторів, скупчених біля неї.

Для економіко-статистичного аналізу результатів дослідження не менш актуальним є проведення кластеризації, що забезпечує розвиток особливої форми мислення в багатьох галузях політики та економіки. Значний вклад в теорію кластерів здійснив М. Портер, який спираючись на теорію конкурентоспроможності товарів та послуг, дає таке визначення: “Кластери – це сконцентровані за географічними ознаками групи взаємопов'язаних компаній, спеціалізованих постачальників та постачальників послуг; фірм у відповідних галузях, а також пов'язаних з їхньою діяльністю організацій у певних галузях, що конкурують, але разом з тим ведуть спільну роботу” [2, с. 206].

Методологія проведення кластерного аналізу ґрунтується на здійсненні класифікації об'єктів дослідження за допомогою численних обчислювальних процедур. В результаті цього утворюються

“кластери” або групи дуже схожих об’єктів. На відміну від інших методів, цей вид аналізу дає можливість класифікувати об’єкти не за однією ознакою, а за декількома одночасно. Для цього використовуються відповідні показники, що характеризують певну міру близькості за всіма класифікаційними параметрами.

Купалова Г. та Бутинець Ф. пропонують проводити кластерний аналіз із застосуванням наступних методів:

- метод одиничного зв’язку (метод близького сусіда) – передбачає приєднання одиниці сукупності до кластера, якщо вона близька по відстані (знаходиться на одному рівні схожості) хоча б до одного представника даного кластера;

- метод повного зв’язку (далекого сусіда) – вимагає певного рівня подібності об’єктів (не менше граничного рівня), які передбачається включити у кластер, у порівнянні із будь-якими іншими об’єктами;

- метод середнього зв’язку – ґрунтується на використанні середньої відстані між кандидатом на включення у кластер і представниками наявного кластера;

- метод Уорда – означає, що приєднання об’єктів до кластерів здійснюється у випадку мінімального приросту внутрішньогрупової суми квадратів відхилень [3, 4].

Використовуючи метод середнього зв’язку, який є найбільш поширеним для формування регіональної кластеризації, нами було обрано функцію «кластерний аналіз» та наступний алгоритм проведення дослідження в програмі SPSS – Statistica [1, с. 387–390]:

1. Розрахунок матриці вихідних даних за досліджуваним показником, що дало змогу здійснити поділ вихідної статистичної сукупності на задану кількість кластерів у вигляді дендрограми.

2. Обчислення багатовимірних середніх $\bar{\delta}_i$ (центрів тяжіння – середніх значень даних економічних показників) виділених кластерів.

3. Розрахунок Евклідової відстані кожної одиниці сукупності до визначених центрів тяжіння кластерів та побудова матриці відстаней. Евклідова відстань передбачає, що кожен господарюючий суб’єкт можна представити у вигляді деякої точки з n -координатами у багатовимірному просторі. Тоді відстань між заданою точкою і центром тяжіння за формулою Евкліда буде визначатися таким чином:

$$r_i = \sqrt{\sum_{i=1}^n (x_{ni} - \bar{x}_i)^2}$$

З цією метою по кожному виділеному кластеру необхідно побудувати розрахункову таблицю та обчислити матрицю d –

коефіцієнтів Евклідових відстаней, які являють собою середні квадратичні відхилення досліджуваних показників та вказують на те, на скільки в середньому відхиляється кожне значення показника від його середнього кластерного значення. Теорія кластерного аналізу вимагає, щоб розраховані значення Евклідових відстаней не перевищували 30% їх середніх значень.

Висновки. Обробка масивів статистичних даних за допомогою застосування факторного та кластерного аналізу, як і інших методів дослідження стохастичних зв'язків, вимагають численних складних розрахунків, які доцільно здійснювати за допомогою сучасних інформаційних систем та з використанням спеціальних програмних продуктів. У межах поставлених завдань слід використовувати спеціальну статистичну програму SPSS – Statistica та теоретичні основи обґрунтування побудованих факторних і кластерних груп.

Список використаних джерел

1. Бююль А. SPSS: Искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей: пер. с англ./ А. Бююль, П. Цефель. – М.: Торгово-издательский дом «Dia Soft», 2005. – 603 с.

2. Портер М. Конкуренция: в 2-х томах./ М. Портер. – М. : Издательский дом «Вильямс», 2002. Т 1. – 496 с.

3. Купалова Г. І. Теорія економічного аналізу: навч. посібник / Г. І. Купалова. – К.: Знання, 2008. – С. 445-447.

4. Економічний аналіз: навч. посібник / за ред. проф. Ф. Ф. Бутинця. – Житомир: Рута, 2003. – 680 с.

5. Концепція створення кластерів в Україні від 29.08.2008 р. [Електронний ресурс]. / Міністерство економіки України. – Режим доступу: <http://www.me.gov.ua/control/uk/publish/printable>.

6. Саблук П. Т. Кластеризація як механізм підвищення конкурентоспроможності та соціальної спрямованості аграрної економіки/ П. Т. Саблук, М. Ф. Кропивко // Економіка АПК. – 2010. – № 1(183). – С. 3-12.

7. Ульянченко О. В. Підвищення конкурентоспроможності аграрного сектора економіки України на кластерних засадах / О. В. Ульянченко // Вісн. аграр. науки. – 2010. – № 10. –С. 56-59.