

УДК 311.172

С. Герасименко, д-р екон. наук, проф.,  
ВНЗ "Національна академія правління",  
О. Чуприна, канд. екон. наук, доц.,  
Харківський національний університет ім. В. Каразіна,  
О. Чуприна, канд. екон. наук, доц.,  
Національна юридична академія України ім. Ярослава Мудрого

## БАГАТОМІРНИЙ СТАТИСТИЧНИЙ АНАЛІЗ ТЕНДЕНЦІЙ ЕКОНОМІЧНОГО РОЗВИТКУ ТА РІВНЯ ЖИТТЯ НАСЕЛЕННЯ

*Подана характеристика об'єктивності підходів, які найчастіше використовуються для порівняльної оцінки економічного розвитку та рівня життя населення країн світу. Обґрунтована доцільність використання методів багатомірного статистичного аналізу для здійснення таких оцінок. Зокрема, за допомогою кластерного аналізу проведено порівняння економічного розвитку та рівня життя населення країн Європи та оптимізацію набору потрібних для цього показників.*

*Ключові слова: порівняльна оцінка, інтегральний показник, рейтинг, багатомірна класифікація, кластерний аналіз, стандартизація чинників.*

*Предложена характеристика объективности подходов, которые чаще всего используются для сравнительной оценки экономического развития и уровня жизни населения стран мира. Обоснована целесообразность использования для таких оценок методов многомерного статистического анализа. В частности, с помощью кластерного анализа выполнено сравнение экономического развития и уровня жизни населения стран Европы и оптимизацию набора необходимых для этого показателей.*

*Ключевые слова: сравнительная оценка, интегральный показатель, многомерная классификация, кластерный анализ, стандартизация факторов.*

*The description of the objectivity of the approaches which are the most often used for comparative estimation of the economical development and the living standards of the population of the countries of the world is suggested. The reasonability of usage for such estimation the methods of multidimensional statistical analysis is grounded. In particular, using the cluster analysis the comparison of social-economic development of the countries of Europe and the optimization of the set of indexes necessary for this are carried out.*

*Keywords: comparative valuation, integral index, rating, multivariate classification, cluster analysis, standardization of factors.*

В умовах глобалізації, коли конкурентна боротьба за ресурси та ринки збуту посилюється, прийняття рішень щодо інвестування вимагає вичерпної, оперативної та надійної інформації з метою зменшення ризиків підприємницької діяльності на всіх рівнях. Багатогранність ризиків та відсутність, в багатьох випадках, можливості їх кількісного виміру обумовили потребу у розподілі країн або територій за рейтингами, при визначенні яких враховуються ті чи інші характеристики. Вважається, що такі рейтинги та інтегральні оцінки дозволяють скласти враження про певну країну, її природні, економічні, соціальні, політичні, демографічні, науково-технічні особливості.

Одним з найвідоміших інтегральних показників є індекс людського розвитку (ІЛР), або індекс розвитку людського потенціалу (ІРЛП), який дозволяє оцінити розвиток країни за можливістю населення реалізувати три головні цілі людського розвитку: прожити довге життя та мати добре здоров'я; отримати знання та реалізувати власні здібності; мати гідний рівень життя та забезпечувати свої фізіологічні, соціальні та духовні потреби. Відповідно до звіту 2010 року Україна посіла 69 місце за рівнем ІРЛП та увійшла до другої групи країн за рівнем людського розвитку. При цьому вона поступилася прибалтійським країнам, Росії, Білорусі, Казахстану та Вірменії. В той час, як ще у 1990 році Україна посідала 45 місце серед 173 країн світу та входила до складу 55 держав, що мали найвищий рівень людського розвитку [7].

Важливим індикатором, що дозволяє оцінити сприятливість інвестиційного клімату, є конкурентоспроможність країни. Дослідженнями конкурентоспроможності займаються:

– Міжнародний інститут розвитку менеджменту (Institute of management development, IMD) в Лозанні (Швейцарія). Результати враховують стан економіки, ефективність уряду, стан ділового середовища та стан інфраструктури: за даними 2010 року Україна за рівнем конкурентоспроможності посіла 57 місце з 59 провідних

економік світу, що були обстежені. Гірші результати мали лише Хорватія та Венесуела [8].

– спеціалісти Всесвітнього економічного форуму в Давосі (Швейцарія), які здійснюють оцінку конкурентоспроможності країн за 12 складовими: у 2010-2011 році наша країна посіла 89 місце з 139, причому її рейтинг порівняно з 2009-2010 роком знизився на 7 пунктів, що засвідчує низьку конкурентоспроможність вітчизняної економіки [9].

Аналітики британського дослідного центру New Economic Foundation проводять дослідження, розраховуючи "Всесвітній індекс щастя", що враховує три складові: задоволеність громадян своїм рівнем життя, тривалість життя, екологічне благополуччя. Рейтинг, складений у 2009 році, містить дані за 143 країнами. Найщасливішими в світі були визнані жителі Коста-Ріки, а наша країна в цьому рейтингу обіймає лише 95 місце. Цілком очевидною за таких умов є необхідність розроблення та здійснення ефективної соціально-економічної політики [10].

Таким чином, використання міжнародних інтегральних оцінок та рейтингів дозволяє зробити певні висновки стосовно стану та перспектив соціально-економічного розвитку окремої країни. В той же час, існують недоліки побудови та розрахунків інтегральних показників, що досить часто створюють обмеження у їх використанні з метою порівнянь. Зокрема, якщо країни займають близькі позиції у рейтингу, то зовсім необов'язково, що соціально-економічна ситуація в цих країнах схожа. Так, відповідно до звіту 2010 року Албанія за ІЛР посіла 64 місце, а Російська Федерація – 65. При цьому рівні складових ІЛР цих країн зовсім несхожі. Зокрема, очікувана тривалість життя в Албанії на 10 років більша, ніж в Росії, в той час як ВНД на душу населення – вдвічі менший [7].

Проблеми побудови міжнародних інтегральних оцінок та рейтингів досліджувало багато вчених, зокрема Данилишин Б., Грішнова О., Лібанова Е., Мірошніченко О., Ревенко А., Шишкін В. та інших [1-6]. Більшість з них в своїх роботах приділяли увагу методиці побудови

та аналізу окремих інтегральних показників. Дослідження, побудовані на одночасному використанні багатьох інтегральних індикаторів, що дозволяє отримати більш ґрунтовні та системні висновки, практично відсутні. В зв'язку з вказаним є потреба у розгляді підходів до застосування методів багатомірних класифікацій, інформації для використання яких зараз достатньо в міжнародній статистиці, з метою здійснення об'єктивної порівняльної оцінки економічного розвитку окремих країн, рівня життя населення та тенденцій їх змін.

Постановка завдання. Інвестори велику увагу приділяють не стільки оцінці рівня розвитку на певний момент часу, скільки тенденціям, притаманним економіці країни. Як було зауважено, використання для цього інтегральних рейтингів і оцінок не завжди дає надійні результати в зв'язку з тим, що переважна більшість рейтингів визначається на базі експертних оцінок, які – крім того, що самі по собі мають суб'єктивний характер – ще і в різні періоди можуть надаватися різними експертами. Крім того, набори показників, що застосовуються для таких рейтингових оцінок можна також називати авторським, адже пропонуються і використовуються вони виключно авторами певних рейтингових методик, до того ж часто без суттєвого обґрунтування. Більше того, без доведення або перевірки об'єктивності результатів, отриманих за допомогою обраних експертами методик та наборів показників. Найчастіше автори методик не є статистиками, а тому з множини наявних методів багатовимірних оцінок обирають найпростіший – розрахунок багатомірної середньої. При цьому, в більшості випадків, зважування показників не здійснюється, а коли здійснюється – визначення ваг відбувається експертним шляхом, щодо обґрунтованості якого важко знайти якісь аргументи. Ще більші складності виникають при використанні багатомірної середньої для порівняльної оцінки соціально-економічного розвитку, коли має бути використана велика кількість досить різномірних показників. Здійснити більш об'єктивну порівняльну оцінку дозволяє застосування методів багатомірної класифікації, зокрема кластерного аналізу. Кластерний аналіз також доцільно використовувати для оптимізації набору показників з метою підвищення ефективності подібних оцінок.

Дослідження присвячене розгляду підходів щодо підвищення надійності порівняльних оцінок соціально-економічного розвитку країн шляхом використання для цього кластерного аналізу. Він був застосований до порівняльної оцінки розвитку країн Європи в 2005-2009 рр. При цьому використовувалася лише офіційна інформація Євростату, що обумовлювало об'єктивність отриманих оцінок, тому що показники всіх країн були обчислені за однаковою методологією. Був також врахований часовий лаг дії показників, що характеризували вплив чинників на досягнуті результати в соціальній та економічній сферах. Кластерний аналіз був застосований до абсолютних рівнів показників і до темпових характеристик змін їх рівнів у часі.

Для порівняльної характеристики соціально-економічного розвитку була використана інформація Євростату та Держкомстату України за попередньо відібраними 15 показниками. При цьому було застосовано тільки один показник, для вимірювання якого використовувалася грошова одиниця – валовий внутрішній продукт (ВВП) в розрахунок на душу населення. Це обумовлено тим, що цінність грошей є примарною і, на відміну від речей, що задовольняють певні наші потреби, і обумовлюється суб'єктивними чинниками. Тому було вирішено використовувати для оцінювання показники, рівень яких, з одного боку, реально відображає результати господарської діяльності, а з іншого обумовлює умови життя і добробут населення країни.

Оцінювання було здійснене за 3-ма наборами показників:

1) загальна порівняльна оцінка соціально-економічного розвитку – 13 показників: рівень безробіття, рівень інфляції, частка інвестицій у ВВП, сальдо міграції, частка фінансування НДДКР у ВВП, частка заробітної плати у ВВП, частка витрат домогосподарств у ВВП, ВВП на душу населення, рівень зайнятості, рівень безробіття, рівень безробіття молоді, середня очікувана тривалість життя жінок, середня очікувана тривалість життя чоловіків, природний приріст населення;

2) порівняльна оцінка економічного розвитку країн – 8 показників: рівень безробіття, рівень інфляції, частка інвестицій у ВВП, сальдо міграції, частка фінансування НДДКР у ВВП, частка заробітної плати у ВВП, частка витрат домогосподарств у ВВП, ВВП на душу населення;

3) порівняльна оцінка соціального стану – 6 показників: рівень безробіття молоді, середня очікувана тривалість життя жінок, середня очікувана тривалість життя чоловіків, природний приріст населення, рівень народжуваності, рівень смертності.

Побудова дендрограм здійснювалася за стандартизованими показниками на базі Евклідової відстані за методом Уорда, що дозволило отримати кластери приблизно однакових розмірів. Кластеризація за 13 соціально-економічними показниками дала такі результати (рис.1). Сусідами України в I кластері виявилися країни, що утворилися з республік колишнього СРСР – Естонія, Латвія, Литва. Слід звернути увагу, що до того ж за величиною міжкластерної Евклідової відстані, кластери I та II (до якого увійшли країни Східної Європи – Болгарія, Угорщина, Польща, Словаччина, Румунія) суттєво відрізняються від інших.

Співставлення змісту кластерів, утворених за вказаними вище 3-ма наборами показників, виявило результати, які – на перший погляд – важко пояснити: рух країн з кластеру до кластеру в разі використання всього набору показників в деяких випадках відрізнявся від руху, що спостерігався при використанні окремих складових набору – економічних та соціальних. Для окремих країн ці відмінності наведено в табл.1.

Причиною могло стати те, що при формуванні набору показників не враховувалися безпосередні та латентні зв'язки між ними. Як відомо, досить часто, при відборі показників для оцінок певних явищ і процесів не звертають уваги, що декілька показників характеризують ту ж саму складову. В такому разі, кластерний аналіз посилює вплив цієї складової на оцінку, враховуючи її стільки разів, скільки апріорно включено подібних показників до набору. І тому дія певних реальних причин для характеристики якої було задіяне декілька показників могла виявитися завищеною. Зокрема, це могло мати місце внаслідок використання 2-х рівнів безробіття – загального та молодіжного – які обумовлюються однією причиною – вмінням влади створювати місця праці. І якщо ця причина була врахована двічі, і інші – по одному разу, то і її вплив на оцінку соціально-економічного розвитку був завищений.

у 2 рази. Застосування кластерного аналізу до набору показників дозволяє сформулювати окремі їх групи, де характер дій показників на оцінку схожий між собою. В подальшому, використавши таблицю коефіцієнтів парної кореляції, є можливість замінити в наборі декілька показників з кожної групи на один, зменшивши їх загальну кількість. Тоді виключення з набору одного або декількох показників, що мають значну кореляцію з іншими, дозволяє підвищити якість аналізу, в нашому випадку – порівняльної оцінки економічного та соціального розвитку країн Європи.

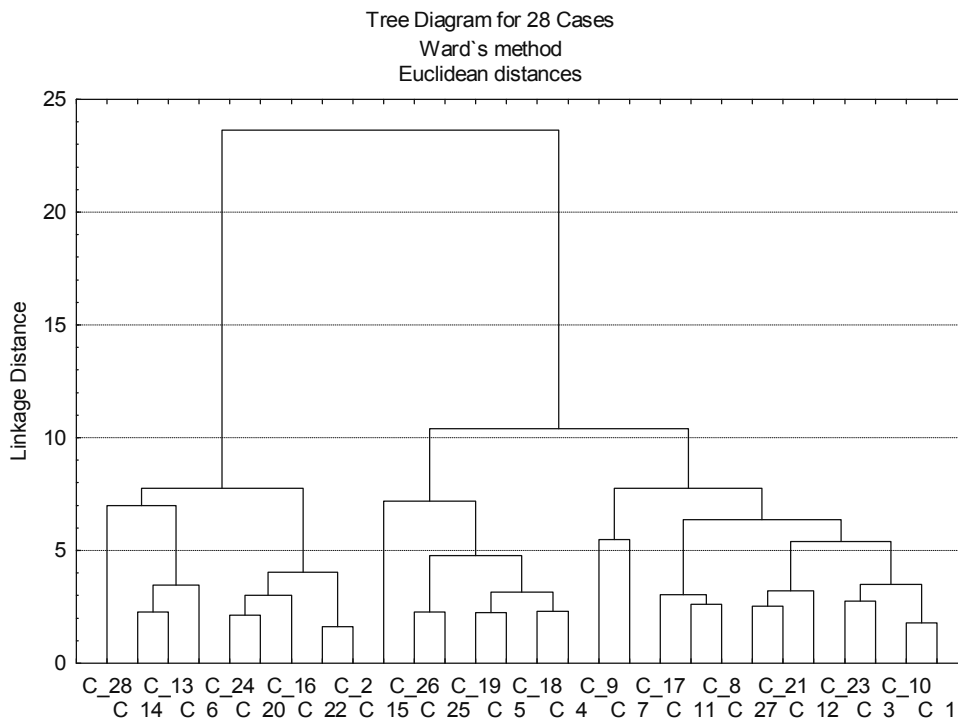


Рис.1. Дендрограма розподілу країн ЄС та України в 2009 р. за 13 соціально-економічними показниками

Джерело: власні розрахунки автора

Таблиця 1. Економічна та соціальна складові рейтингів України, Словаччини, Чехії та Польщі за результатами кластерного аналізу у 2005-2009 рр.

|      | Складові   |       | Україна | Польща | Словаччина | Чехія |
|------|------------|-------|---------|--------|------------|-------|
| 2005 | Економічна |       | =       | =      | =          | +1    |
|      | Соціальна  |       | - 1     | + 1    | + 1        | =     |
| 2006 | Економічна | Рівні | + 1     | =      | =          | - 3   |
|      |            | Темпи | - 1     | =      | =          | =     |
|      | Соціальна  | Рівні | - 1     | + 1    | +1         | - 1   |
|      |            | Темпи | + 1     | =      | - 3        | =     |
| 2007 | Економічна | Рівні | + 1     | - 3    | - 3        | =     |
|      |            | Темпи | - 1     | - 1    | - 1        | =     |
|      | Соціальна  | Рівні | +1      | - 1    | - 1        | =     |
|      |            | Темпи | + 1     | + 1    | =          | + 1   |
| 2008 | Економічна | Рівні | =       | + 1    | +1         | =     |
|      |            | Темпи | - 3     | + 1    | =          | + 1   |
|      | Соціальна  | Рівні | + 1     | -1     | - 1        | =     |
|      |            | Темпи | - 2     | - 2    | - 1        | - 1   |
| 2009 | Економічна | Рівні | + 2     | +1     | + 1        | - 2   |
|      |            | Темпи | + 1     | =      | + 1        | - 2   |
|      | Соціальна  | Рівні | =       | =      | =          | - 1   |
|      |            | Темпи | =       | -1     | - 1        | =     |

Джерело: власні розрахунки автора

Для попередньої характеристики можливості наявності взаємної кореляції між показниками було обчислено коефіцієнти кореляції окремих показників в динаміці, що дозволило виявити 2 групи показників. Перша – показники з низьким рівнем варіації (коефіцієнти кореляції значень показників за різні роки перевищували 0,95), що дозволяє висловити припущення про подібність їх тенденцій, тобто – про можливість заміни кількох на один. Друга група – показники, для яких коефіцієнти кореляції динаміки мають низький рівень за певними фазами економічного циклу. Це обидва рівня безробіття, рівень інфляції, частка фінансування НДДКР у ВВП, сальдо міграції, рівень зайнятості щодо яких у 2008 – 2009 рр. відбулися суттєві зміни рівнів в порівнянні з попередніми роками. Зауважене може свідчити

як про подібну, так і різну скерованість їх тенденцій, що і обумовлює потребу у розрахунку парних коефіцієнтів кореляції значень показників в динаміці. Підхід до використання кластерного аналізу та парних коефіцієнтів кореляції для оптимізації набору показників проілюструємо на наборі показників, що попередньо був сформований для характеристики соціального розвитку. Кластерний аналіз визначив 2 групи подібних показників: 1) середня очікувана тривалість життя жінок, середня очікувана тривалість життя чоловіків, рівень смертності; 2) рівень безробіття молоді, природний приріст населення, рівень народжуваності. Щодо I кластеру, то найвищі значення парних коефіцієнтів кореляції з іншими показниками групи (більші, за 0,86) виявилися у показника середня очікувана тривалість життя жінок, який

– в подальшому – і використовувався в оптимізованому наборі соціальних показників. В той же час, для показників II кластеру значення парних коефіцієнтів кореляції не перевищували 0,25, що дозволило дійти висновку: їхня подібність, яку зауважив кластерний аналіз, не є настільки значною, щоб це дозволило виключити з набору який-небудь з показників.

Кластерний аналіз за оптимізованими вказаним способом кожним з 3-х наборів показників дав результати, які практично не відрізняються від тих, що були отримані при використанні повних наборів. Але не це стало головним висновком з проведеного дослідження та відповідних розрахунків. Як було зауважено в постановці завдання, передбачалося, що оптимізація підвищить об'єктивність оцінок. Особливо важливим це є в разі суттєвих відмінностей результатів кластеризації за повним та оптимізованим наборами показників. Тому була здійснена спроба перевірити, якою мірою виключення з набору пов'язаних взаємно показників вплинуло на оцінку подібності. Мірою такої оцінки, на наш погляд, мала б слугувати відстань між кластерами: якщо вона зменшилась, то це дає право вважати, що оптимізація дозволяє більш об'єктивно відображувати ступінь впливу окремих соціально-економічних причин на оцінку, для визначення якої у вихідному наборі використовувалось кілька показників.

Як відомо, Евклідова відстань такою мірою бути не може, тому що залежить від масштабів сукупності. В той же час, виходячи з природи обчислення – формули та змісту – Евклідової відстані, вона досить просто перетворюється (за аналогією з середнім квадратичним відхиленням) в характеристику середнього відхилення між елементами сукупності по всіх параметрах одночасно:

$$c_{jk} = \left[ \sum_{i=1}^m (z_j - z_k)^2 \right]^{1/2} \Rightarrow \frac{\left[ \sum_{i=1}^m \sum_{j=1}^k (z_j - z_k)^2 \right]^{1/2}}{m(k-1)}$$

УДК 311.15

О. Гладун, д-р. екон. наук, ст. наук. співроб., В. Хвалинська, Інститут демографії та соціальних досліджень імені М.В. Птухи НАН України

## ЗАСТОСУВАННЯ У КЛАСТЕРНОМУ ТА ДИСКРИМІНАНТНОМУ АНАЛІЗІ НЕКІЛЬКІСНИХ ЗМІННИХ

*Розглянутий підхід до використання у кластерному та дискримінантному аналізі не кількісних змінних, вимірних за ранговою та номінальною шкалою, шляхом їх перетворення у структурні змінні.*

**Ключові слова:** кластерний аналіз, дискримінантний аналіз, не кількісних змінні, структурні змінні.

*Рассмотрен подход к использованию в кластерном и дискриминантном анализе неколичественных переменных, измеренных по ранговой и номинальной шкале, на основе их преобразования в структурные переменные.*

**Ключевые слова:** кластерный анализ, дискриминантный анализ, неколичественные переменные, структурные переменные.

*The article reviews approach to use in cluster analysis and discriminate analysis non-quantitative (rang and nominal) variables based on their transformation to structure variables.*

**Keywords:** cluster analysis, discriminate analysis, non-quantitative variables, structure variables.

Однією з важливих задач при проведенні аналітичних досліджень є виділення типологічних груп. За умови невеликої кількості групувальних ознак (до чотирьох) це може бути вирішено на основі методу комбінаційних групувань. При цьому дослідник має визначити кількість груп та їх межі за кожною ознакою, що не завжди є тривіальною задачею. До того ж, загальна кількість груп дорівнює добутку кількості груп за кожною ознакою, що може призвести, по-перше, до великої кількості типологічних груп, що значно ускладнить аналітичні розвідки та інтерпретацію результатів, і, по-друге, до наявності "порожніх" груп, кількість об'єктів у яких дорівнює нулю або є занадто малою для здійснення статистичного аналізу. Слід також взяти до уваги, що при проведенні розвідувально-

Результати розрахунку такого середнього відхилення (в разі заміни 13 показників для оцінки соціально-економічного розвитку на 8 він зменшився з 1,261 до 1,099) дозволяють зробити висновок про більшу подібність, в нашому випадку, країн (тобто – кластерів, до яких їх було віднесено), ніж це виглядало в разі використання не оптимізованого набору показників.

Висновки. Кількісний вимір подібності елементів сукупності за розглянутим підходом дає можливість отримувати обґрунтовані висновки щодо ступеню їх подібності та відмінності між собою. В аналізі економічного розвитку та рівня життя це дозволить міжнародним організаціям та урядам більш об'єктивно приймати рішення щодо напрямку та розмірів інвестування окремих територій, що буде сприяти підвищенню ефективності їх використання і, як наслідок, підвищенню добробуту населення.

1. Грішнова О. А. Людський розвиток [Текст]: навч. посіб. / О. А. Грішнова – К.: КНЕУ, 2006. – 308 с. 2. Данилишин Б. Україна в міжнародних рейтингах сталого розвитку [Текст] / Б. Данилишин, О. Веклич // Економіка України. – 2008. – № 7. – С. 13-23. 3. Лібанова, Е. М. Людський розвиток регіонів України: методика оцінки і сучасний стан [Текст] / Е. М. Лібанова, Н. С. Власенко, О. С. Власюк та ін. К.: 2002. – 110 с. 4. Мірошниченко, О. Ю. Статистичне оцінювання рівня життя населення [Текст] / О. Ю. Мірошниченко // Автореферат дис. канд. екон. наук – К., 2008. – 20 с. 5. Ревенко, А. Україна у світових координатах людського розвитку [Текст] / А. Ревенко // Дзеркало тижня. – 2001. – № 28 (28 липня). – С. 1, 10. 6. Шишкін, В. С. Людський розвиток: оцінка та основні напрями підвищення його рівня [Текст] / В. С. Шишкін // Автореферат дис. – К., 2004. – 20 с. 7. Звіти з людського розвитку на сайті Програми розвитку Організації Об'єднаних Націй [Електронний ресурс] – Режим доступу: <http://hdr.undp.org>. 8. Результати досліджень Міжнародного інституту розвитку менеджменту (Institute of management development, IMD) [Електронний ресурс]. – Режим доступу: <http://www.imd.ch>. 9. Всемирный экономический форум: рейтинг глобальной конкурентоспособности 2010-2011 [Електронний ресурс] – Режим доступу: <http://gtmarket.ru/news/state/2010/09/09/2668>. 10. The Happy Planet Index 2.0. [Електронний ресурс]. / Abdallah S., Thompson S., Mishaelson J., Marks N., Steuer N. / London, New economics foundation, 2009, 62 p. – Режим доступу: <http://gtmarket.ru/news/state/2009/07/09/2090>.

Надійшла до редколегії 01.07.11

О. Гладун, д-р. екон. наук, ст. наук. співроб., В. Хвалинська, Інститут демографії та соціальних досліджень імені М.В. Птухи НАН України

го аналізу доволі часто неможливо однозначно наперед визначити групувальні ознаки та інтервали.

Тому при проведенні багатьох досліджень практично немає альтернативи застосуванню багатовимірної класифікації, перед якою стоять дві відмінні задачі: 1) формування груп однорідних об'єктів; 2) віднесення об'єкта в одну з наперед визначених груп.

При вирішенні першої задачі розбиття сукупності об'єктів на групи повинно здійснюватись таким чином, щоб у підсумку об'єкти однієї групи були максимально подібні між собою, а об'єкти різних груп – максимально відмінні. При вирішенні другої задачі необхідно забезпечити виконання умови, щоб об'єкт, який відноситься до однієї з груп, був максимально подібним до об'єктів своєї

© Гладун О., Хвалинська В., 2011