

Ольга Іванівна КОРИЦЬКА

аспірантка кафедри обліку і аудиту,
Львівський національний університет імені Івана Франка
пр. Свободи, 18, м. Львів, 79008, Україна
E-mail: k.olja005@gmail.com
Телефон: +380966514861

ЯКІСНИЙ АНАЛІЗ МНОЖИНИ ДАНИХ У ПРОЦЕСІ ОЦІНЮВАННЯ СОЦІАЛЬНО-ЕКОНОМІЧНИХ ЯВИЩ

Корицька, О. І. Якісний аналіз множини даних у процесі оцінювання соціально-економічних явищ [Текст] / Ольга Іванівна Корицька // Економічний аналіз : зб. наук. праць / Тернопільський національний економічний університет; редкол. : В. А. Дерій (голов. ред.) та ін. – Тернопіль : Видавничо-поліграфічний центр Тернопільського національного економічного університету “Економічна думка”, 2014. – Том 16. – № 1. – С. 62-68. – ISSN 1993-0259.

Анотація

Вступ. Дослідження соціально-економічних явищ передбачає проведення моніторингу та збору інформації за різними показниками, які їх характеризують. Аналіз соціально-економічних явищ пов'язаний з обробкою великих інформаційних масивів даних. Сукупність останніх для оцінювання характеризується різноманітністю та неточністю. Інформація формується багатьма експертами, може містити сумнівні, не оброблені (упущені) дані. Для ефективного та достовірного оцінювання соціально-економічних явищ доцільно здійснити якісний аналіз множини даних. Це забезпечить формування однорідної сукупності інформації, вилучення атипових даних з аномально великим додатнім чи від'ємним значенням, підвищення достовірності одержаних результатів дослідження.

Мета. Розглянути і розвинути методика якісного аналізу множини даних у процесі оцінювання соціально-економічних явищ.

Результати. Здійснено огляд найбільш поширених методів оцінювання генеральних сукупностей даних. Звертається увага на недосконалість методики якісного аналізу множини даних у процесі оцінювання соціально-економічних явищ. Якісний аналіз множини даних містить аналіз упущеної та сумнівної інформації. Рекомендовано способи обробки упущеної та виявлення сумнівної інформації. При наявності упущеної інформації застосовують «зведення до нульових значень», умовне кодування, використання методу максимальної правдоподібності, головних компонент. При наявності сумнівної інформації проводиться її ідентифікація та оцінка. Ідентифікація здійснюється за критеріями Ф. Граббса, Г. Мура, Н. Смирнова, Г. Тітьєна. На основі оцінки даних сумнівна інформація усувається із подальшої обробки або модифікується. Запропоновано варіанти модифікації інформації з використанням методів робастної статистики. До останніх належать методи Ч. Вінзора, А. Пуанкаре, Дж. Тьюкі, П. Хубера.

Якісний аналіз множини даних у процесі оцінювання соціально-економічних явищ передбачає виявлення та обробку атипової інформації. Нами репрезентовано узагальнену схему якісного аналізу множини даних. Запропонована методика може бути застосована у вивченні різних видів економічної діяльності. Подальші наукові дослідження передбачають апробацію методики для оцінювання ефективності промислового виробництва України.

Ключові слова: критерії оцінки; робастні методи статистики; соціально-економічні явища; сукупність даних; «сумнівні» спостереження; упущена інформація; якісний аналіз.

Ольга Ивановна КОРИЦКАЯ

КАЧЕСТВЕННЫЙ АНАЛИЗ СОВОКУПНОСТИ ДАННЫХ В ПРОЦЕССЕ ОЦЕНИВАНИЯ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

Аннотация

Введение. Исследование социально-экономических явлений предусматривает проведение мониторинга и сбора информации по различным показателям, которые их характеризуют. Анализ социально-экономических явлений связан с обработкой больших информационных массивов данных. Совокупность последних для оценивания характеризуется разнородностью и неточностью. Информация предоставляется многими экспертами, может содержать сомнительные, не обработанные (упущенные)

© Ольга Іванівна Корицька, 2014

данные. Для эффективного и достоверного оценивания социально-экономических явлений целесообразно осуществить качественный анализ массива данных. Это обеспечит формирование однородной совокупности информации, извлечение атипичных данных с аномально большим положительным или отрицательным значением, повышение достоверности полученных результатов исследования.

Цель. Рассмотреть и развить методику анализа совокупности данных в процессе оценивания социально-экономических явлений.

Результаты. Осуществлен обзор наиболее распространенных методов оценки генеральной совокупности данных. Обращается внимание на несовершенство методики анализа совокупности данных в процессе оценивания социально-экономических явлений. Качественный анализ множества данных включает анализ упущенной, сомнительной информации. Рекомендуются способы обработки упущенной информации и выявления сомнительной информации. При наличии упущенной информации применяют «сведение к нулевым значениям», условное кодирование, использование метода максимального правдоподобия, метод главных компонент. При наличии сомнительной информации осуществляется ее идентификация и оценка. Идентификация совершается по критериям Ф. Граббса, Г. Мура, Н. Смирнова, Г. Титьена. На основе оценки данных «сомнительная» информация исключается из дальнейшей обработки или модифицируется. Предложены варианты модификации информации с использованием методов робастной статистики. К робастным методам статистики относятся методы Ч. Винзора, А. Пуанкаре, Дж. Тьюки, П. Хубера.

Качественный анализ множества данных в процессе оценивания социально-экономических явлений предусматривает выявление и обработку атипичной информации. Нами представлена обобщенная схема анализа совокупности данных. Предложенная методика может быть применена в изучении различных видов экономической деятельности. Дальнейшие научные исследования предполагают апробацию методики для оценивания эффективности промышленного производства Украины.

Ключевые слова: критерии оценки; робастные методы статистики; совокупность данных; «сомнительные» наблюдения; упущенная информация; качественный анализ.

Olga Ivanivna KORYTSKA

PhD Student,
Department of Accounting and Auditing,
Ivan Franco National University of Lviv,
Svoboda str., 18, Lviv, 79008, Ukraine
E-mail: k.olja005@gmail.com
Phone: +380966514861

QUALITATIVE ANALYSIS OF DATA SETS IN THE STUDY OF SOCIAL AND ECONOMIC PROCESSES

Abstract

Introduction. Socio-economic phenomena research provides monitoring and gathering of information according to various indicators that characterize them. Analysis of socio-economic phenomena which is associated with handling large amounts of data information has been carried out. A set of evaluation data is characterized by heterogeneity and inaccuracy. Information is formed by many experts, so it can contain questionable and unprocessed data. For efficient and reliable evaluation of socio-economic phenomena it is necessary to carry out a qualitative analysis of data sets. This will ensure the formation of a homogeneous set of data, extracting of atypical data with abnormally large positive or negative value, improving of the reliability of the results of the study.

Purpose. The aim of the article is to consider and develop a method of qualitative analysis of data sets in the process of evaluation of the socio-economic phenomena.

Results. The overview of the most common methods for assessing general data set has been done. Attention is drawn to imperfect methods of qualitative analysis of data sets in the process of evaluation of the socio-economic phenomena. Qualitative analysis of data sets containing an analysis of lost and uncertain information has been carried out. Methods for lost and uncertain information treatment are recommended. If there is a loss of information then it should be used "reduce to zero", conditional coding, method of maximum likelihood of principal components. If there is questionable information its identification and evaluation should be held. Identification is carried out according to the criteria F. Hrabbs, G. Moore, N. Smirnov, G. Tityen. On the basis of assessment of data the questionable information should be eliminated from further processing or should be modified. The variants of information modification using robust statistics methods are proposed. The latter includes methods of Ch. Vinzor, A. Poincare, Jh. Tukey, P. Huber. Qualitative analysis of the data sets in the process of evaluation of the socio-economic phenomena involves the detection and treatment of atypical information. We have represented a generalized

scheme of qualitative analysis of data sets. The technique can be applied in the study of various economic activities. Further research will involve testing methods for evaluation of the efficiency of industrial production in Ukraine.

Keywords: criteria for evaluation; robust statistical methods; set of data; «doubtful» observation; missed information; qualitative analysis.

JEL classification: C13, C81, E20

Вступ

У процесі аналізу сукупності даних для оцінювання соціально-економічних явищ використовується різноманітна інформація, одержана з багатьох джерел, у різні часові періоди, багатьма експертами. Зазвичай такі дані не є згруповані, недостатньо формалізовані, містять сумнівну інформацію, з можливими упущеннями та неточностями. Важливо вчасно здійснити попередній якісний аналіз множини даних із урахуванням особливостей інформації.

Проблемами дослідження сукупностей вибіркового даних, їх кількісною та якісною оцінкою у різні часи займалися С. А. Айвазян, А. М. Дубров, Л. А. Сошнікова, В. Н. Тамашевич та інші вчені [1-3].

Упровадження та використання робастних методів в оцінюванні множин даних здійснили Ч. Вінзор, Ф. Граббс, А. Пуанкаре, Н. Смирнов, Г. Тітьєн, Дж. Тьюкі, П. Хубер [4-10]. Проте науковий підхід до аналізу та удосконалення зазначених методів, з урахуванням їх можливостей у дослідженні розвитку соціально-економічних явищ є недостатньо розглянутих.

Мета статті

Мета дослідження – розвинути методику якісного аналізу множини даних на основі застосування математично-статистичних методів для оцінювання соціально-економічних явищ. Завданням є визначення підходів та методів для опрацювання сукупностей даних, які містять упущену, сумнівну інформацію чи інші дані, які потребують додаткового опрацювання.

Виклад основного матеріалу дослідження

Якісний аналіз множини даних містить: відновлення упущеної інформації, аналіз сумнівної інформації, опрацювання різних відхилень у вихідних даних. На рис. 1 відображено загальну схему якісного аналізу множини даних у процесі оцінювання соціально-економічних явищ.

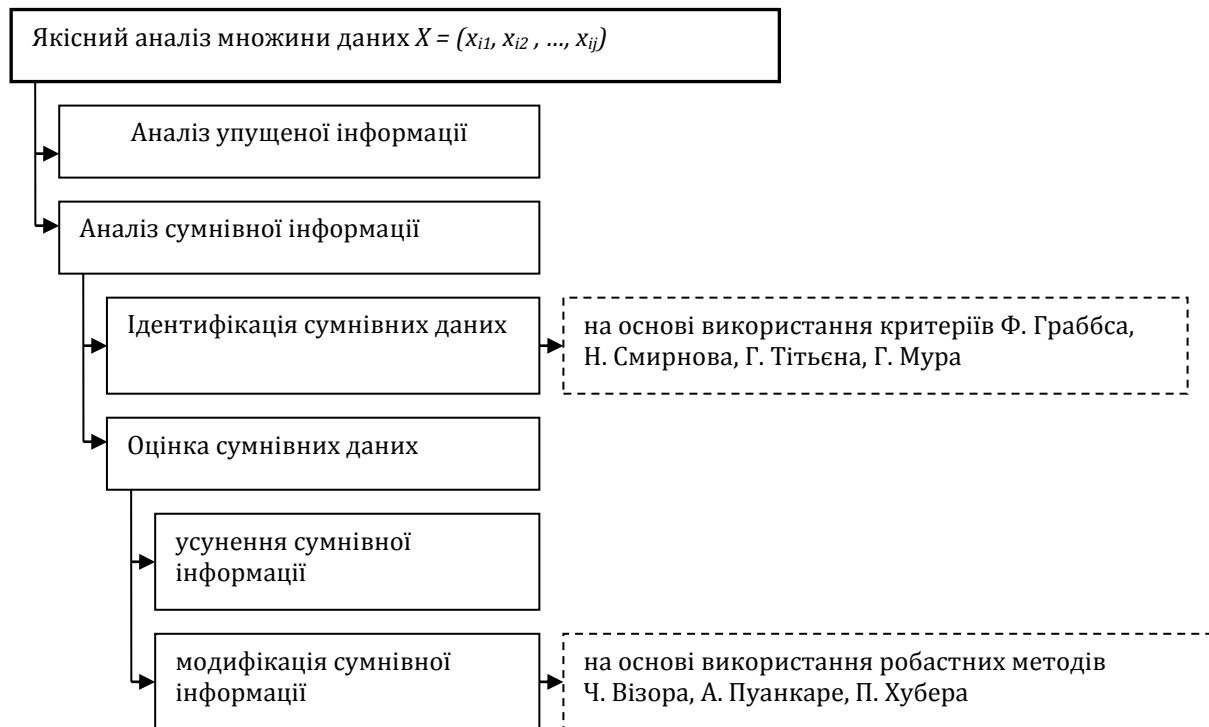


Рис. 1. Загальна схема якісного аналізу множини даних у процесі оцінювання соціально-економічних явищ (авторська розробка)

Відсутність певних значень можлива через різні причини:

- технічні (несправність досліджуваних приладів, збій у розрахунках при підготовці даних тощо)
- пов'язані зі станом об'єкта дослідження (конфіденційність інформації, зміна методики розрахунку тощо).

При наявності технічних причин застосовуються різні підходи: «зведення до нульових значень», умовного кодування, усереднення, використання методу максимальної правдоподібності, головних компонент [11, с. 408-415]. Найбільш простим способом обробки даних у таких випадках є заміна упущених значень ознаки її середнім арифметичним значенням, яке знаходиться на основі наявної інформації.

Статистичні процедури виділення сумнівних даних засновані на припущенні про їх однорідність. Сумнівні спостереження або цілком усуваються із подальшого дослідження, або їх внесок зменшується за допомогою вагової функції, яка спадає у міру зростання ступеня аномальності спостережень [11, с. 43].

Особливу увагу в оцінюванні множин даних звертають на спостереження, які різко виділяються із загальної сукупності. Дані з аномально великим додатнім чи від'ємним значеннями розглядаються як атипово далеко віддалені від центру розподілу (сумнівні спостереження). Інколи відхилення виникають випадково, у результаті помилкового розрахунку, механічного заповнення даних, неправильного зчитування з вимірювальних приладів або невідповідності у використовуваних прикладних моделях чи змін умов експерименту. Ідентифікація таких спостережень здійснюється з повторним розглядом та перерахунком умов їх виникнення, з можливістю виправлення аномальних значень.

Завдання аналізу даних виконується у два етапи: виявлення «сумнівних» спостережень і перевірка статистичної значущості їх відмінностей від основної сукупності інформації.

Виявлення сумнівних спостережень здійснюється за критеріями Ф. Граббса, Г. Мура, Н. Смирнова, Г. Тітьєна та інших [2; 3; 7; 8; 12; 13]. Найпоширенішим способом для розрахунку вважається Т-критерій Граббса [5]:

$$T_n = \frac{x - \bar{x}}{s}, \quad (1)$$

де \bar{x} - середнє значення, $\bar{x} = \sum_{i=1}^n \frac{x}{n}$; s - вибіркове середньоквадратичне відхилення. Одержане значення

порівнюють із табличним значенням Смирнова-Граббса: якщо $T_n > T_{кр}$, то досліджуване значення є сумнівним і належить до класу атипових відхилень.

Для виявлення нетипових спостережень у «верхній частині» ранжованого ряду даних використовують L-критерій Тітьєна-Мура [2, с. 306; 3, с. 199]:

$$L = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x}_k)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2)$$

де x_i - ознака спостереження j-го об'єкта; n - обсяг вибірки; k - кількість сумнівних спостережень; \bar{x}_k - середнє значення, розраховане за $n-k$ спостереженнями, які залишилися після вилучення k атипових

спостережень «зверху» ранжованого ряду, $\bar{x}_k = \frac{\sum_{i=1}^{n-k} x_i}{n-k}$.

Для виявлення нетипових спостережень у «нижній частині» ранжованого ряду даних використовують L'-критерій Тітьєна-Мура [2, с. 306; 3, с. 199]:

$$L' = \frac{\sum_{i=k+1}^{n-k} (x_i - \bar{x}_k)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (3)$$

де \bar{x}_k - середнє значення, розраховане за $n-k$ спостереженнями, які залишилися після вилучення k

атипових спостережень «знизу» ранжованого ряду, $\bar{x}_k = \frac{\sum_{i=k+1}^{n-k} x_i}{n-k}$.

Якщо у множині даних наявні сумнівні спостереження з найбільшими і найменшими даними, тобто розташовані «зверху» і «знизу» ранжованого ряду, для ідентифікації використовують E-критерій Тітьєна-Мура [2, с. 307; 3, с. 199]:

$$E = \frac{\sum_{i=k+1}^{n-k} (x_i - \bar{x}_k)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (4)$$

де \bar{x}_k - середнє значення, розраховане за типовими даними, після вилучення атипових найбільших (k) і найменших (k') значень. Зазначені критерії мають критичні значення для заданого рівня значущості α , при відомому обсязі вибірки n і попередньо визначених K сумнівних спостереженнях.

Підходи до опрацювання «екстремальних (сумнівних)» спостережень загалом поділяються на два види:

- вилучення аномальних спостережень із подальшої обробки;
- використання робастних методів обробки (модифікації) «екстремальних» даних.

Якщо значення неможливо виправити, такі спостереження, зазвичай, вилучаються із подальшої обробки як нетипові. С. А. Айвазян зазначає: «Історично було зауважено, що при оцінці центру розподілу бажано відкидати зліва і справа частки крайніх членів варіаційного ряду і взяти середнє арифметичне тих членів, які залишилися» [1, с. 341]. У 1809 році К. Ф. Гаус підійшов до створення нормального закону розподілу через вирішення задачі пошуку розподілу, для якого середнє арифметичне значення незалежних змінних \bar{x} є оцінкою максимальної правдоподібності для параметра зміщення [13]. Для опису розподілів, близьких до Гаусового, традиційно використовують середні арифметичні значення і середньоквадратичні відхилення. Проте такий підхід не завжди виправданий, оскільки таке оцінювання втрачає оптимальні властивості при відхиленні розподілів від нормального закону. Тому актуальними є методи, які порівняно менш ефективні, ніж Гаусовий розподіл, проте більш стійкі (робастні) до відхилень від нормального закону.

Робастність (від англ. robust – міцний, здоровий, дужий; robust estimation – грубе, стійке оцінювання) – нечутливість до незначних відхилень від припущень [10, с. 9].

Робастні методи оцінювання дають можливість одержати достатньо надійні оцінки статистичної сукупності з урахуванням невизначеності її закону розподілу і наявності значних відхилень від центру розподілу. Вперше термін «робастність» ввів Дж. Е. П. Бокс [14, с. 7]. Вивчення та впровадження робастних методів здійснили Ч. Вінзор, А. Н. Колмогоров, А. Пуанкаре, В. Роджерс, Дж. Тьюкі, П. Хубер та інші вчені [4; 7; 9; 10; 15].

Наступним етапом після достовірного виявлення екстремальних відхилень є застосування робастних методів для їх обробки.

До робастних методів належать [1, с. 361; 10, с. 51-78]:

- оцінки типу максимальної правдоподібності;
- лінійні комбінації порядкових статистик;
- оцінки масштабу;
- графічні методи.

Такі методи оцінювання передбачають для кожного виявленого атипового значення виділення окремо «істинного» значення ознаки і її помилки: $x = x_{icm} + \zeta$, де ζ - викривлений («сумнівний») елемент. А також здійснення модифікації даних таким чином, щоб ξ одержав нормальний закон розподілу з нульовим математичним сподіванням. Як наслідок, для сукупності спостережень X з атиповими значеннями, сума $\sum \zeta \rightarrow 0$, а оцінки \hat{x} наближаються до істинних значень параметрів вибіркової сукупності.

А. Пуанкаре запропонував формулу для розрахунку середнього арифметичного значення за «урізаною» сукупністю (де атипові значення відкидаються і замінюються на середні) [3, с. 200-201]:

$$T(a) = \frac{1}{n-2k} \sum_{i=k+1}^{n-k} x_i, \quad (5)$$

де k – кількість сумнівних спостережень; $k \leq an$ - ціла частина від an ; n – обсяг вибіркової сукупності за винятком екстремальних значень; a – деяка функція наявності атипових значень вибірки (ξ), табличне значення.

Формула для розрахунку середнього арифметичного значення за «урізаною» сукупністю Ч. Вінзора є [1, с. 343]:

$$W(a) = \frac{1}{n} \left(\sum_{i=k+2}^{n-k-1} x_i + k(x_{k+1} + x_{n-k}) \right) \quad (6)$$

Згідно з підходом П. Хубера для модифікації використовується деяка величина K , яка визначається з урахуванням ступеня наявності атипових значень сукупності вибірки ξ і встановлює їхній крок модифікації [2 с. 289-294; 3 с. 201-202]:

$$\hat{\Theta} = \frac{1}{n} \left(\sum_{|x_i - \Theta| < k} x_i + (n_2 + n_1)k \right), \quad (7)$$

де $\hat{\Theta}$ - стійка оцінка, визначена на основі ітеративних процедур; k - допустима величина відхилення атипових значень від центру розподілу, має постійні значення з урахуванням питомої ваги атипових значень у сукупності ξ ; n_1 - кількість спостережень із сукупності, які відрізняються найменшими значеннями: $x_i < \Theta - k$, чи значення в інтервалі $(-\infty; \Theta - k; n_2)$ - кількість спостережень із сукупності, які відрізняються найбільшими значеннями: $x_i > \Theta + k$, чи значення в інтервалі $(\Theta + k; \infty)$. У процесі розрахунків, як $\hat{\Theta}$, може використовуватись середнє арифметичне значення чи медіана за даними вибірки. Наступним етапом є розподіл вибіркової сукупності на три частини. В одну частину потрапляють ознаки, які залишаються без змін ($|x_i - \Theta| < k$), в інші дві частини (для $x_i > \Theta + k$ і $x_i < \Theta - k$) потрапляють атипові модифіковані значення, розраховані як $x_i - k$ і $x_i + k$. На кожному етапі серед ознак, які залишилися без змін («істинних») та модифікованих, здійснюється нова оцінка середнього значення Θ і процедура повторюється. Ітерації повторюються до моменту, коли усі спостереження потраплять в інтервал $|x_i - \Theta| < k$. Недоліком запропонованого методу є втрата оптимальних властивостей зі збільшенням питомої ваги атипових значень у сукупності ξ .

Висновки та перспективи подальших досліджень

Якісний аналіз множини даних є передумовою ефективною оцінки об'єкта дослідження. Перелік методів для його здійснення залежить від сукупності інформаційного масиву, яким оперує дослідник. Зважаючи на наявність пропущеної, сумнівної інформації, використовуються математично-статистичні методи, запропоновані багатьма ученими. Доцільність їх застосування визначає дослідник.

При обробці сукупності вихідних даних, які розподілені не за нормальним законом розподілу, вони дають можливість модифікації пропущеної, сумнівної інформації і подальшого її використання для оцінювання соціально-економічних явищ. Запропонована схема якісного аналізу множини даних є універсальною і може використовуватись у дослідженні різних видів економічної діяльності на макро- та мікрорівнях здійснення господарської діяльності. Застосування комбінованого підходу у виборі представлених методів ідентифікації та обробки «сумнівної» інформації забезпечить більш достовірні результати оцінювання.

Перспективами подальших наукових досліджень є апробація методики якісного аналізу сукупності показників для оцінювання ефективності промислового виробництва України.

Список літератури

1. Айвазян, С. А. Прикладная статистика: Основы моделирования и первичная обработка данных : Справочное издание [Текст] / С. А. Айвазян, И. С. Енюков, Л. Д. Мешалкин. - М.: Финансы и статистика, 1983. - 471 с.
2. Дубров, А. М. Многомерные статистические методы [Текст]: учебник / А. М. Дубров, В. С. Мхитарян, Л. И. Трошин. - М.: Финансы и статистика, 2003. - 352 с., ил.
3. Сошникова, Л. А. Многомерный статистический анализ в экономике: учеб. пособие для вузов [Текст] / Л. А. Сошникова, В. Н. Томашевич, Г. Уебе, М. Шеффер / под ред. проф. В. Н. Томашевича. - М.: ЮНИТИ-ДАНА, 1999. - 598с.
4. Low moments for small samples: a comparative study of order statistics [Text] / C. Hasings, F. Mosteller, J. V. Tukey, C. P. Winsor, 1947.
5. Frank, E. Grubbs Sample Criteria for Testing Outlying observations [Text] / Frank E. Grubbs // Ann. Math. Statist, 1950. - Vol. 21. - No. 1. - P.27-58.
6. Пуанкаре, А. Наука и метод [Текст] / А. Пуанкаре. - СПб. 1910.
7. Смирнов, Н. В. Оценка максимального члена в ряду наблюдений [Текст] / Н. В. Смирнов // Доклады АН СССР, 1941. - Т. 33. - № 5. - С. 346-349.
8. Tietjen, G. Some Grubb's type statistics for the detection of several outliers [Text] / G. Tietjen, H. Moore. - Technometrics, 1972.
9. Robust estimates of location: survey and advances [Text] / D. F. Andrews, P. J. Bickel, F. R. Hampel, P. J. Huber, W. H. Rogers, J. W. Tukey, Princeton University Press. - 1972.
10. Хьюбер, Дж. П. Робастность в статистике. Пер. с англ. [Текст] / Дж. П. Хьюбер. - М.: Мир, 1984. - 304с., ил.
11. Бондіна, Н. М. Комп'ютеризація спеціалізованих середовищ: навч. посіб. [Текст] / Н. М. Бондіна, А. І. Поворознюк, О. М. Шейн. - Х.: НТУ "ХПІ", 2013. - 378 с.
12. Сравнительный анализ критериев грубых ошибок [Электронный ресурс] режим доступа: <http://www.pisali.ru/zalvladimir/98176/> (дата обращения: 15.04.2014). - Название с экрана.

-
13. Gauss, C. F. *Theoria motus corporum coelestium*. Hamburg, 1809. [Electronic resource] mode of access: http://books.google.com.ua/books?id=ORUOAAAAQAAJ&pg=PA3&hl=uk&source=gbs_toc_r&cad=4#v=onepage&q&f=false (data viewed: 16.04.2014).
 14. Box, G. E. P. A basis for the selection of a response surface design [Text] / G. E. P. Box, N. R. Draper. – J. Amer. Statist. Ass., v. 54, № 287, p. 622—654.
 15. Колмогоров, А. Н. Теория вероятностей и математическая статистика [Текст] / А. Н. Колмогоров. – М.: Наука, 1989. – 623 с.

References

1. Aivazian, S. A., Enukov, I. S., Meshalkin, L. D. (1983). *Applied statistics: Basics modeling and analysis of spatial data [Prykladnaia statystyka: Osnovy modelirovaniia i pervichnaia obrabotka danykh]*. Moscow: Finance and statistic.
2. Dubrov, A. M., Mhitarian, I. S., Troshun, L. I. (2003). *Multivariate statistic methods [Mnogomernye statisticheskie metody]*. Moscow: Finance and statistic.
3. Soshnikova, L. A., Tomashevuch, V. N., Uebe, H., Sheffer, M. (1999). *Multivariate statistic analysis in economics [Mnogomernyi statisticheskii analiz v ekonomike]*. Moscow: YUNITI-DANA.
4. Hasings, C., Mosteller, F., Tukey, J. W., Winsor, C.P. (1947). *Low moments for small samples: A comparative study of order statistics*.
5. Grubbs, F. E. (1950). *Sample Criteria for Testing Outlying observations*.
6. Puankare, A. (1910). *Method and science [Nauka i metod]*. Sankt-Peterburg.
7. Smirnov, N. V. (1941). *Estimation of the maximum term in the series of observations [Otsenka maksimalnogo chlena v riadu nabludenii]*.
8. Tietje, n G., Moore, H. (1972). *Some Grubb's type statistics for the detection of several outliers*. *Technometrics*.
9. Andrews, David F; Peter, J Bickel; Frank, R. Hampel; Peter J. Huber; W H Rogers & John W Tukey (1972). *Robust estimates of location: survey and advances*. Princeton University Press.
10. Huber, G. P. (1984). *Robust statistics [Robastnost v statistike]*. Moscow: Mir.
11. Bondina, N. M., Povoroznuk, A. I., Shein, O. M. (2013). *Computerization of specialized media [Komputeruzatsia spetsializovanykh seredovyshch]*. Kharkiv: NTU: HPY.
12. *Comparative analysis of the criteria blunders [Sravnitelnyi analiz kriteriev grubych oshybok]*, from : <http://www.pisali.ru/zalvladimir/98176/>.
13. Gauss, C. F. (1809). *Theoria motus corporum coelestium*. Gamburg.
14. Box, G. E. P., Draper, N. R. (n. d.). *A basis for the selection of a response surface design*. J. Amer. Statist. Ass.
15. Kolmogorov, A. N. (1989). *Probability theory and mathematical statistics [Teoria veroiatnostei i matematicheskaia statystyka]*. Moscow: Nauka.

Стаття надійшла до редакції 13.05.2014 р.