УДК 519.86: 330.34

К. В. ГОРБАТЮК
Хмельницький національний університет

СТАТИСТИЧНІ МЕТОДИ ПОБУДОВИ ФУНКЦІЙ НАЛЕЖНОСТІ НЕЧІТКИХ ЕКОНОМІЧНИХ ПОКАЗНИКІВ

В статті розглянуто проблему побудови функцій належності нечітких економічних показників з використанням статистичного підходу до аналізу й інтерпретації наявної невизначеності в кількісних параметрах досліджуваних процесів. Обґрунтовано необхідність застосування непараметричних методів статистичного аналізу, методів стійкого оцінювання центру розподілу та методів оцінювання відхилень від центрального значення для формалізованого опису нечіткості в кількісних економічних показниках. Запропоновано методику використання стійких оцінок при побудові функцій належності нечітких економічних показників. Означені напрямки подальших досліджень стосовно нечіткого моделювання показників в економічних задачах.

Ключові слова: нечіткі економічні показники, функції належності нечітких економічних показників, непараметричні методи статистичного аналізу, методи стійкого оцінювання центру розподілу.

K. V. GORBATYUK
Khmelnytsky National University

STATISTICAL METHODS FOR MEMBERSHIP FUNCTIONS BUILDING OF FUZZY ECONOMIC INDICATORS

The aim of the research - to explore the possibility of using robust statistical methods in fuzzy modeling of economic indicators. This work presents the problem of constructing membership functions of fuzzy economic performance using statistical approach to the analysis and interpretation of uncertainties in the quantitative parameters of the studied processes. Investigated the nature of problems arising in the analysis of uncertainty in economic indicators. The peculiarities of the objects, which are investigated, such as: limited amount of experimental data required for analysis and processing; poor reproducibility of results of observations; the heterogeneity of the objects of observation; no means of checking the assumptions of normal theory; the practical impossibility of maintenance requirements that apply to random samples when conducting observations. The necessity of the use of nonparametric methods of statistical analysis, methods of robust estimation of center of distribution and methods of evaluating deviations from the Central values of the indicator for a formalized description of fuzziness in quantitative economic data. Investigated existing approaches of robust estimation of center of distribution and deviations from the Central values of the studied parameters. The technique of using the obtained stable estimates when constructing the membership functions of fuzzy economic performance. Identified areas for further research on fuzzy modeling of economic tasks.

Keywords: fuzzy economic indicators, membership functions of fuzzy economic indicators, non-parametric methods of statistical analysis, methods of robust estimation of center of the distribution.

Вступ

Невизначеність ринкового середовища висуває особливі вимоги до моделювання економічних показників в різних сферах господарської діяльності, оскільки вибір оптимальної стратегії економічного розвитку проходить в умовах наявності багатоальтернативних ситуацій, кожна з яких являє собою послідовність можливих сценаріїв розвитку в залежності від запланованих управлінських рішень. Облік факторів невизначеності та неповноти інформації є необхідною умовою моделювання економічних систем, в яких значною мірою якість функціонування більшості процесів залежить від прийнятих управлінських рішень.

Традиційний шлях урахування факторів невизначеності на основі імовірнісного та статистичного моделювання часто виявляється недоцільним через неможливість достатнього теоретичного обґрунтування застосованих методів та відсутність повної вихідної інформації. Нечіткі підходи є альтернативою загальноприйнятим кількісним методам аналізу систем та мають такі основні переваги: по-перше, дозволяють разом з числовими змінними використовувати нечіткі величини та так звані «лінгвістичні» змінні; по-друге, дають можливість відношення між змінними описувати за допомогою нечітких висловлювань; по-третє, дозволяють представляти складні відношення за допомогою нечітких алгоритмів. Такі підходи дають наближені, але у той же час, більш прості у використанні та ефективні засоби опису поведінки економічних систем, що не піддаються точному математичному аналізу.

Основна частина

Не зважаючи на те, що перше згадування про нечіткі методи математичного моделювання з'явилося близько півстоліття тому, ця галузь наукових досліджень досі є маловивченою у нашій країні. В той же час, економічні застосування теорії нечітких множин вже утворили самостійний науковий напрям. Нечітко-множинні моделі дуже прості у використанні та дають достовірні результати навіть в умовах високої невизначеності, а нечітко-множинне моделювання в аналізі та прогнозуванні економічних явищ та процесів є одним з найбільш перспективних напрямків наукових досліджень.

Методи, що базуються на теорії нечітких множин, відносяться до методів оцінки та прийняття рішень в умовах невизначеності. Їх використання передбачає формалізацію вихідних параметрів і цільових економічних показників, що характеризується деяким ступенем невизначеності, у вигляді нечітких значень. Здійснюючи арифметичні та інші операції з такими нечіткими значеннями за правилами нечіткої математики, експерти та ОПР отримують результуючий нечітке значення для цільового показника. На основі вихідної інформації, досвіду та інтуїції, експерти часто можуть досить впевнено кількісно охарактеризувати межі (інтервали) можливих (допустимих) значень параметрів і області їх найбільш можливих (бажаних) значень.

Поняття нечіткої множини найчастіше вводиться як пара $\overline{A}=\left\{x\in X,\mu_{\overline{A}}(x)\right\}$, де $\mu_A(x)$ – функція, що визначає ступінь належності x до $\overline{A}$. Формально, функція належності задається відображенням: $\mu_{\overline{A}}(x): X\to[0,1]$, де X називається носієм нечіткої множини і за умови $\sup\limits_{x\in X}\mu_{\overline{A}}(x)=1$ – відповідна нечітка множина називається нормальною. Також у теорії нечітких множин часто використовують метод опису нечітких множин за допомогою α-рівневих множин, які є підмножинами базової множини X та описуються так: $\overline{A}[\alpha]=\left\{x\middle|\mu_{\overline{A}}(x)\ge\alpha\right\}$ для всіх значень $\alpha\in[0,1]$.

Побудова функції належності нечіткої множини або нечіткого числа (нечіткої величини) має важливе практичне та методологічне значення. Зокрема, у статті [1] подано частотний спосіб визначення функції належності, який заснований на використанні статистичної інформації про досліджувану величину, що описується нечітко. У роботі [2] здійснено аналіз існуючої літератури на предмет пошуку методів побудови функцій належності нечітких величин і зроблено висновок, що у більшості запропонованих випадків функція належності задається експертами на основі суб'єктивних суджень, тобто невипадково.

Функція належності, у загальному випадку, є формалізованим описом ступеня присутньої нечіткості на деякій множині, яка, у свою чергу, являє собою носій нечіткої величини.

Природною для функцій належності є вимога до її неперервності. Оскільки, за визначенням, функція належності нечіткого числа або нечіткої множини формалізує інтуїтивне уявлення про те, що якщо два елементи множини X лише трохи відрізняються один від одного, то значення функцій належності для цих елементів також є близькими.

Конкретний вид функцій належності визначається на основі різних додаткових припущень про властивості цих функцій (симетричність, монотонність, неперервність першої похідної тощо) з урахуванням специфіки наявної невизначеності та реальної ситуації [1].

У багатьох практичних ситуаціях функція належності повинна бути оцінена, виходячи з часткової інформації про неї (наприклад, зі значень, які вона приймає на скінченій множині базових точок x_1, ..., x_n). У цьому випадку говорять, що вона частково визначена за допомогою «пояснюючого прикладу» [2].

На практиці застосовуються методи побудови функцій належності за вибірками і на підставі апріорної інформації, в яку входять обмеження на ці функції. Якщо апріорної інформації про властивості характеристичних функцій недостатньо для побудови певних функцій, що були б «оптимальними» в якомусь значенні, доводиться вдаватися до евристичних методів знаходження цих функцій з подальшою експериментальною перевіркою «якості» вибраних функцій [2]. Основними формами функцій належності, що використовуються в теорії нечітких множин [2] та можуть бути використані для опису кількісних економічних показників, є трикутна, трапецієподібна та квазідзвоноподібна форми.

Залежно від цілей, завдань і доцільності деталізації, для опису нечітких економічних показників можуть бути використані нечіткі величини з відповідною формою функцій належності.

Імовірнісний метод побудови функцій належності оснований на схожості понять нечіткості та ймовірності [3]. Обидва ці поняття використовуються за наявності невизначеності в системах, яка є наслідком випадкових факторів, неточності нашого знання або принципової неможливості і непотрібності отримання точних рішень. Але навіть у випадку, коли невизначеність може бути подана ймовірнісною моделлю, зазвичай зручніше оперувати з нею методами теорії нечітких множин без залучення апарату теорії ймовірностей.

Багато авторів праць із теорії нечітких множин та теорії ймовірностей зазначають суб'єктивність функції належності на відміну від об'єктивності ймовірності як характеристики певного явища. Але цілком можливо задавати функцію належності так, щоб малоймовірна з точки зору теорії ймовірностей подія мала найменше значення функції належності [2]. «Нестатистична» природа побудови деяких функцій належності не означає, що в якості таких функцій не можна використовувати функції розподілу ймовірностей, нормовані певним чином. Отже, при побудові функцій належності можна виокремити два способи: статистичний та нестатистичний.

Зокрема, у [2] наведено визначення поняття нечіткої величини або нечіткого числа як нечіткої множини, яка задана на базовій множині дійсних чисел R. Також, у цій роботі визначено алгебраїчні операції над нечіткими величинами, досліджено їх властивості, показано деякі аналітичні та чисельні методи знаходження цих результатів.

Так, арифметичні операції додавання, множення, віднімання і ділення, що визначені на множині дійсних чисел, поширюються на клас нечітких чисел. Для знаходження результатів алгебраїчних операцій над нечіткими величинами використовується низка аналітичних та чисельних методів.

Дослідження природи задач, що виникають при аналізі невизначеності в економічних показниках, дозволяє стверджувати, що в статистичному розумінні особливостями об'єктів, які досліджуються, є: обмежений обсяг експериментальних даних, необхідних для аналізу і обробки; погана відтворюваність результатів спостережень; неоднорідність об'єктів спостережень; відсутність можливості перевірки передумов нормальної теорії; практична неможливість забезпечення вимог, які висуваються до випадкових вибірок при проведенні спостережень. Ці особливості диктують необхідність застосування непараметричних методів статистичного аналізу, методів стійкого оцінювання центру розподілу та методів оцінювання відхилень від центрального значення досліджуваного показника.

Якщо говорити про методи стійкого оцінювання центру розподілу, то можна зробити висновок, що на даний час сформувалося два підходи до отримання стійких оцінок. Перший підхід, який найчастіше вживається в навчальній та науковій літературі, засновано на припущенні про те, що наявна вибірка є «забрудненою» «грубими» помилками, що необхідно спочатку виявити, а потім певним методом усунути або підправити, а потім одержати шукану оцінку. Для виявлення «грубих» значень у вибірці зазвичай рекомендуються критерії Смірнова-Граббса і Титьена-Мура, а для обробки даних – підходи Пуанкаре, Вінзора, Хубера, Хампеля, Мешалкина та інші [4; 5].

Другий підхід є більш реалістичним, він розглядає будь-яке значення у вибірці як рівноправне, яке відображає властивості вибіркової сукупності. Цей підхід для невеликих вибірок реалізується за допомогою методів «розмноження» вибірки: бустреп-методу (bootstrap) і джекнайф-методу (jackknife або методу «складного ножа») [4; 5].

У методах виявлення й усунення засмічення вибірок за результатами спостереження $x_1, x_2, \ldots, x_n$ будується варіаційний ряд (упорядкованих за зростанням значень) [4]:

$$x_{(1)} < x_{(2)} < \ldots < x_{(N)}. \tag{1}$$

Для застосування критерію Смірнова-Граббса обчислюється значення:

$$T_{(N)} = \frac{x_{(N)} - \bar{x}}{s}, \tag{2}$$

де s – оцінка середньоквадратичного відхилення σ; $\bar{x}$ – оцінка математичного сподівання μ.

За відповідною таблицею критичних значень визначається C_α [4] та порівнюється зі значенням $T_{(N)}$. Якщо $T_{(N)} < C_\alpha$, то вірною є гіпотеза H_0 про те, що $x_{(N)}$ не є грубою помилкою. При $T_{(N)} > C_\alpha$ робиться висновок, що $x_{(N)}$ суттєво відхиляється від $\bar{x}$, а отже, є грубою помилкою. Існують також інші різновиди даного методу, які викладені у спеціальній літературі й узагальнені у [4].

За необхідності виключення декількох грубих помилок використовується підхід Титьена-Мура – узагальнення критерію Граббса на декілька екстремальних спостережень.

Тоді вирішальне правило для віднесення найбільших спостережень до грубих помилок засновано на статистиці

$$L_{(k)} = \frac{\sum_{i=1}^{N-k} (x_{(i)} - \bar{x}_k)^2}{\sum_{i=1}^{N} (x_{(i)} - \bar{x})^2}, \text{ де } \bar{x}_k = \frac{\sum_{i=1}^{N-k} x_{(i)}}{N-k}. \tag{3}$$

Аналогічне правило діє для k найменших спостережень. Статистика $L_{(k)}$ порівнюється з відповідним табличним значенням та робиться статистичний висновок щодо належності найбільших або найменших спостережень до грубих помилок.

Після застосування методів виявлення грубих помилок переходять до оцінки центру розподілу за вибіркою. Для цього за упорядкованим варіаційним рядом (1) обчислюється оцінка за Пуанкаре, тобто α - обрізана середня ($0 \le \alpha < 0{,}5$), яка визначається формулою

$$T(\alpha) = \frac{1}{N - 2[\alpha N]} \sum_{i=[\alpha N]+1}^{N-[\varepsilon N]} x_{(i)}, \tag{4}$$

де [αN] – ціла частина від αN, тобто найбільше ціле число, що не перевершує αN.

Оцінка, запропонована Вінзором, передбачає заміну «грубих» значень вибірки на модифіковані, так звані вінзоровані значення з усуненими помилками [4].

За сукупністю (1) середнє значення для рівня α ($0 \le \alpha < 0{,}5$) визначається формулою

$$W(\alpha) = \frac{1}{N}\left[\sum_{i=[\alpha N]+2}^{N-[\varepsilon N]-1} x_i + [\alpha N]\left(x_{([\alpha N]+1)} + x_{(N-[\alpha N])}\right)\right]. \tag{5}$$

Дана процедура відрізняється від середньої за Пуанкаре, оскільки [αN] значення не виключаються ні з лівого, ні з правого кінця варіаційного ряду. Ці значення проектуються в найближчу точку частини впорядкованого варіаційного ряду, яка залишилася. Отже, при визначенні середньої арифметичної беруть участь усі N спостережень, тобто втрати обсягу початкової вибірки не спостерігається. Величину ε (рівень засмічення) можна визначити за вибіркою, а відповідне значення α – за відповідними таблицями.

Разом із перерахованими методами робастного оцінювання, широке розповсюдження має також класичний підхід Хубера. Він нагадує процедури для послідовного «поліпшення» даних за Вінзором. При цьому використовується деяка початкова величина k, що визначається з урахуванням ступеня «засмічення» статистичної сукупності ε та задає крок модифікації різко відмінних спостережень [4].

Оцінка середньої величини за методом Хубера знаходиться за формулою:

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

де θ – стійка оцінка, визначається за допомогою ітеративних процедур; k – величина, яка допускається як відхилення від центру сукупності, приймає постійні значення з урахуванням питомої ваги грубих помилок у сукупності даних ε; n_1 – чисельність групи спостережень з інтервалу $\left(-\infty;\hat{\theta}-k\right)$; n_2 – чисельність групи спостережень з інтервалу $\left(\hat{\theta}+k;\infty\right)$; $\hat{\theta}$ – отримана на попередній ітерації оцінка показника.

Крім описаних вище методів оцінювання, існує оригінальний метод, запропонований Тьюки Д., що знаходиться, по суті, на межі між двома відміченими підходами щодо визначення стійких оцінок центру. У літературі цей метод називається бівес-методом [6].

Бівес-метод дає можливість визначити центр розподілу і величину відхилення в умовах, коли точний вид закону розподілу є невідомим, але апріорно можна зробити припущення про його симетричність. Для бівес-оцінки $(\hat{x})$ «внесок» даних числового ряду оцінюється біквадратними вагами, які обчислюються ітеративно залежно від поточних значень x_i та значення $(\hat{x})$ [7]. Шуканою величиною є $\hat{x}$:

$$\hat{x} = \left(\sum_{i=1}^{n} w_i x_i\right) : \sum_{i=1}^{n} w_i , \tag{6}$$

де x_i – значення, що спостерігається; w_i – біквадратна вага (бівес).

У (6) параметри n, x_i є відомими, а біквадратна вага обчислюється в такій послідовності:

Визначається різниця між поточними значеннями x_i та бівесом, одержаним на попередній ітерації. Для першої ітерації $\hat{x}$ вибирається довільно (наприклад, x_{cp}).

Визначається медіана абсолютних відхилень поточних значень від передбачуваного центру (усередненого значення), бівеса $\hat{x}$: $S = Me\left|x_i - \hat{x}\right|$.

Біквадратна вага визначається за формулою:

$$w_i = \begin{cases} 0, & \text{якщо } \left(\dfrac{x_i - \hat{x}}{CS}\right) > 1; \\ \left[1 - \dfrac{(x_i - \hat{x})}{CS}\right]^2, & \text{якщо } \left(\dfrac{x_i - \hat{x}}{CS}\right) \le 1. \end{cases}, \tag{7}$$

Постійна C – це параметр, за допомогою якого визначається, при яких відхиленнях вага w_i у рівнянні (7) обертається в нуль. Практично C = 6 або C = 9. Добуток CS вимірює допустиму варіацію щодо центру вимірювань, а $u_i = (x_i - \hat{x})/CS$ – відхилення в частках цієї величини.

Як свідчать численні дослідження [6; 7], бівес-оцінка володіє достатньо високою ефективністю, як для малих, так і для великих вибірок.

Крім того, за наявності вибірки невеликого обсягу застосовуються методи, що засновані на ідеї штучного збільшення обсягу початкової вибірки. Уперше ідея «розмноження» (resampling) вибірок була запропонована Кенуєм М. у 1949 р. у зв'язку з обговоренням проблеми зсуву та стійкості оцінок. Основна ідея методу полягає в оцінці ефективності, незміщеності та стійкості оцінки в тому випадку, якщо початкові дані розбити на групи й оцінити результат, який одержується після того, як будь-яку групу виключити з розгляду. Такий метод має назву «складний ніж» (jackknife), підкреслюючи тим самим універсальність методу. На думку автора, цей метод конкурує з багатьма окремими методиками, що не завжди є придатними [6].

Загальна ідея методу полягає у послідовному виключенні з початкової вибірки по одному елементу з подальшим їх поверненням; обчисленні шуканої оцінки для кожної нової, штучно одержаної вибірки та порівнянні результатів. Тоді для початкової вибірки обсягу n одержуємо стільки ж вибірок, однак обсягом (n-1) кожна, що означає і (n-1) нових значень оцінки. Якщо виключати по 2 елементи, то одержимо вже кількість вибірок $C_n^2 = n(n-1)/2$ з обсягом у (n-2) елементів кожна. У загальному випадку можна одержати C_k^2 штучних вибірок обсягу (n- k) елементів, де k – кількість елементів, які виключаються.

Позначимо $\hat{x}_{заг}$ – бівес-оцінка для всієї вибірки; нехай x_j – аналогічний результат після відкидання одного елементу. Тоді, маючи в своєму розпорядженні різниці $x_j = k\hat{x}_{заг} - (k-1)\hat{x}_j$, одержимо «псевдозначення» чисел, які використовуються як початкові дані для подальших розрахунків середньої «складного ножа» [6]:

$$x^* = \frac{1}{k}\left(x_1 + x_2 + ... + x_j + ... + x_k\right), \quad (j = 1, k). \tag{8}$$

Отже, основна ідея методу полягає у використанні «псевдозначень» як незалежних результатів для отримання необхідного статистичного висновку.

Ще один непараметричний метод «розмноження» вибірок має назву бутстреп-методу [6]. Це один із методів отримання оцінок характеристик генеральної сукупності за вибірками малого обсягу.

Нехай в результаті спостереження за випадковою величиною X отримано таку наступну невелику вибірку значень (у вигляді вектора):

$$x_e = (x_{e1}, x_{e2}, \ldots, x_{en}). \tag{9}$$

За схемою даного методу створюється допоміжна вибірка x_b, так звана бутстреп-вибірка [6]. Ця нова вибірка повинна бути багато в чому подібна до експериментальної вибірки x_e, а процес її створення – до процесу отримання інформації з експерименту. Таким вимогам цілком задовольняє n-мірний вектор x_b:

$$x_b = (x_{b1}, x_{b2}, \ldots, x_{bn}), \tag{10}$$

де кожний елемент вектора x_{bi} випадковим чином вибирається з елементів x_e. Іншими словами $x_{bi} = x_{ej}$, причому j є випадковим цілим числом із діапазону [1; n].

За вибіркою обчислюється середнє значення або будь-яка інша оцінка (наприклад, бівес-оцінка). Позначимо її x_{mb} та назвемо бутстреп-середнім:

$$x_{mb} = \sum x_{bi} / n. \tag{11}$$

Значення x_{mb} може дещо відрізнятися від оцінки x_{me} за експериментальною вибіркою (9):

$$x_{me} = \sum x_{ei} / n. \tag{12}$$

Цей факт знову імітує поведінку випадкової величини X в експерименті. Якби можна було ще раз повторити весь експеримент, то він не обов’язково міг би дати те ж саме значення x_{me}, що і в першому експерименті.

До штучно створеної вибірки величин x_{mb} можна застосовувати всі відомі методи репрезентативної статистики, оскільки обсяг вибірки є великим. Зокрема, можна обчислити середнє значення або будь-яку іншу оцінку за всіма величинами x_{mb}, побудувати гістограму, за якою визначити центр розподілу тощо.

Для отримання величини оцінки відхилення традиційно застосовується метод побудови довірчих інтервалів [6]:

$$x_ц \pm t_{p,\alpha}\sqrt{\mu_2}, \tag{13}$$

де $x_ц$ – центр розподілу; t – значення статистичного критерію Стьюдента для відомої кількості ступенів свободи і заданого рівня довірчої ймовірності (α = 0,90; 0,95; 0,99); μ_2 – другий центральний момент або дисперсія:

$$\mu_2 = S^2 = \int_{-\infty}^{+\infty}\left(x - x_ц\right)^2 p(x)\, dx. \tag{14}$$

Дисперсія, виражаючи ступінь розсіювання випадкової величини, є центрованою характеристикою, а, отже, безпосередньо залежить від коректності визначення центру розподілу. Крім того, сам метод отримання оцінки відхилення (13) орієнтований на симетричні розподіли, тобто дає можливість одержати інтервал невизначеності, симетричний, відносно передбачуваного центру $x_ц$. Якщо ж розподіл не має бажаної симетричності (що на практиці зустрічається частіше), то і застосування методу є некоректним. Тому застосування підходу (13) без відповідного обґрунтування є сумнівним.

Розглянемо деякі окремі методи побудови довірчих інтервалів, які відповідають першому підходу, а, зокрема, класичний довірчий інтервал, бівес-інтервал та джекнайф-інтервал [6].

Класичний довірчий інтервал для генеральної середньої за малою вибіркою будується за формулою:

$$\overline{x} - \Delta_{м.в.} \le \overline{x_0} \le \overline{x} + \Delta_{м.в.}, \tag{15}$$

де $\Delta_{м.в.} = \dfrac{t_{\gamma,n-1}s}{\sqrt{n-1}}$ – гранична помилка малої вибірки; $t_{\gamma,n-1}$ – критичне значення розподілу Стьюдента при заданій довірчій імовірності γ і кількості ступенів вільності (n-1); $\overline{x}$ – вибіркове середнє; s – стандартне відхилення за вибіркою.

При цьому передбачається, що випадкова величина X підпорядкована нормальному закону розподілу.

Для розглянутої вище (6) бівес-оцінки можна одержати оцінку відхилення та побудувати стійкий довірчий інтервал за даними вибірки, що досліджується [8].

Процедура розрахунку дисперсії використовує бівес-оцінку $\hat{x}$, а для дисперсії, що оцінюється, застосовується вираз:

$$\hat{var}\ \hat{x} = \frac{n \cdot \sum_{i=1}^{n_1} (x_i - \hat{x})^2 (1 - u_i^2)^4}{\left[\sum_{i=1}^{n_1} (1 - u_i^2)(1 - 5u_i^2)\right]\left[\sum_{i=1}^{n_1} (1 - u_i^2)(1 - 5u_i^2) - 1\right]}, \tag{16}$$

де $\sum_{i=1}^{n_1}$ – означає підсумовування за $u_i \le 1$.

Для отримання довірчих меж для $\hat{x}$ використовується *t*-критерій Стьюдента з кількістю ступенів свободи $v = 0{,}7(n-1)$, що приводить до довірчого інтервалу:

$$\hat{x} \pm t_v \sqrt{\hat{var}\ \hat{x}}. \tag{17}$$

Ця апроксимація добре працює при $n \ge 8$. Для вибірок меншого обсягу використання даної оцінки довірчого інтервалу є недоцільним, що пов'язано зі значеннями $t_{v.}$.

За одержаною оцінкою центрального значення вибірки за методом «складного ножа» визначається значення оцінки дисперсії [8]:

$$D^* = \left[\sum_{j=1}^{k} x_j^2 - \frac{1}{k}\left(\sum_{j=1}^{k} x_j\right)^2\right] \Big/ k(k-1). \tag{18}$$

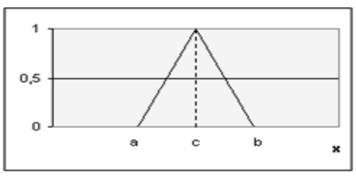
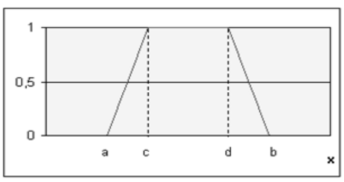
Оцінка дисперсії використовується для побудови довірчих інтервалів за допомогою критерію Стьюдента (надалі будемо використовувати термін джекнайф-інтервал):

$$x^* \pm t_{k-1} \sqrt{D^*}. \tag{19}$$

На основі описаних вище оцінок можу бути визначені параметри найбільш вживаних в теорії нечітких множин функцій належності. Запропонуємо методику застосування комбінацій статистичних оцінок для встановлення значень параметрів функцій належності трикутної та трапецієподібної форм, як найбільш вживаних в економічних дослідженнях, а також найбільш простих у використанні на практиці (табл. 1).

Таблиця 1

Основні види функцій належності нечітких множин

Функція належності	Оцінки параметрів функції належності
Трикутна $\mu_2(x,a,b,c) = \begin{cases} 0, \text{ якщо } x \le a; \\ \frac{x-a}{c-a}, \text{ якщо } a < x \le c; \\ \frac{b-x}{b-c}, \text{ якщо } c < x < b; \\ 0, \text{ якщо } x \ge b. \end{cases}$ 	**Параметр *c*:** оцінка Пуанкаре, Вінзора, Хубера або Тьюки (бівес-оцінка); метод «складного ножа» або бутстреп-метод; **Параметри *a* та *b*:** класичний довірчий інтервал, бівес-інтервал та джекнайф-інтервали (13), (15), (17), (19), де в якості центрального значення береться відповідне вибране значення ***c***
Трапецієподібна $\mu_3(x,a,b,c,d) = \begin{cases} 0, \text{ якщо } x \le a; \\ \frac{x-a}{c-a}, \text{ якщо } a < x < c; \\ 1, \text{ якщо } c \le x \le d; \\ \frac{b-x}{b-d}, \text{ якщо } d < x < b; \\ 0, \text{ якщо } x \ge b. \end{cases}$ 	**Параметри *c* та *d*:** класичний довірчий інтервал, бівес-інтервал та джекнайф-інтервали (13), (15), (17), (19), де в якості центрального значення береться одна з оцінок (Пуанкаре, Вінзора, Хубера або Тьюки (бівес-оцінка); або оцінка за методом «складного ножа» або бутстреп-методом) **Параметри *a* та *b*:** Відповідно *мінімальне* та *максимальне* значення досліджуваної економічної величини

Таким чином, зазначені вище методи складають основу математичного апарату обробки статистичних даних щодо економічних показників та організації інформаційної бази з такими показниками для забезпечення їх якісного аналізу та прогнозування з врахуванням присутньої невизначеності нечіткої природи.

Висновки

Отже, запропонована вище методика побудови функцій належності нечітких економічних показників дозволяє отримати нечіткі уявлення про кількісні характеристики різних економічних об'єктів з врахуванням існуючої невизначеності нечіткої природи. Даний підхід дає можливість використовувати наявну статистичну інформацію про динаміку економічних показників та отримувати формалізований опис присутньої в них невизначеності. Методи стійкого оцінювання центру розподілу та відхилення від нього дають основні значення оцінок для побудови функцій належності нечітких економічних показників, що повинні бути доповнені змістовним аналізом цих показників для обґрунтованого вибору форми функцій належності та точних значень їх параметрів. Побудовані нечіткі економічні показники, в свою чергу, можуть бути використані в подальших алгоритмах обробки інформації для отримання нових знань про перебіг різноманітних процесів в економічних об'єктах та динаміку станів їхніх підсистем.

Література

1. Алексеев А. В. Интерпретация и определение функций принадлежности нечетких множеств. Методы и системы принятия решений / А. В. Алексеев. – Рига, 1979. – С. 42–50.

2. Алтунин А. Е. Модели и алгоритмы принятия решений в нечетких условиях : монография / А. Е. Алтунин, М. В. Семухин. – Тюмень : Издательство Тюменского государственного университета, 2000. – 352 с.

3. Кандель А. Нечеткие множества, нечеткая алгебра, нечеткая статистика / А. Кандель, У. Дж. Байатт // Труды американского общества инженеров-радиоэлектроников. – 1978. – Т. 66. № 12. – С. 37–61.

4. Дубров А. М. Многомерные статистические методы : учебник / А. М. Дубров, В. С. Мхитарян, Л. И. Трошин. – М. : Финансы и статистика, 2000. – 352 с.

5. Мостеллер Ф. Анализ данных и регрессия : в 2 т. / Ф. Мостеллер, Д. Тьюки. – М., 1982. – 252 с.

6. Эфрон Б. Нетрадиционные методы многомерного статистического анализа : сб. статей / Б. Эфрон ; [пер. с англ. / предисловие Ю. П. Адлера, Ю. А. Кошевника]. – М. : Финансы и статистика, 1988. – 263 с.

7. Игумнов Б. Н. Кибернетические основы построения экономических систем для предприятий : [уч. пособие.] / Б. Н. Игумнов, Т. П. Завгородняя. – Хмельницкий : ТУП, 2000. – 344 с.

8. Шурыгин А. М. Прикладная стохастика : робастность, оценивание, прогноз / Шурыгин А. М. – М. : Финансы и статистика, 2000. – 224 с.

Надійшла 10.09.2015; рецензент: д. е. н. Григорук П. М.