

УДК 331:330.46

К. В. ГОРБАТЮК

Хмельницький національний університет

**ПОБУДОВА ІНТЕРВАЛІВ ЕНТРОПІЇ З МЕТОЮ ВИЗНАЧЕННЯ ФУНКЦІЙ
НАЛЕЖНОСТІ НЕЧІТКИХ ПОКАЗНИКІВ ВИРОБНИЧИХ ПРОЦЕСІВ**

Розглянуто проблему визначення функції належності показників нечіткої природи, що використовуються для характеристики виробничих процесів при створенні комплексних систем нормування праці. Досліджено застосування методу побудови ентропійного інтервалу в поєднанні з бутстреп-методом до вибірок малого обсягу для визначення центрального значення та відхилень від нього. Проведено порівняльний аналіз різних методів отримання інтервалів невизначеності для оцінювання функцій належності нечітких показників виробничих процесів.

Ключові слова: нормування праці, робастні методи оцінювання, інтервал ентропії, бутстреп-метод, бутстреп-вибірка, бутстреп-експеримент, бутстреп-інтервал, неунімодальність, джеккнайф-метод, джеккнайф – інтервал, інтервал невизначеності.

Проблема створення комплексних систем нормування, які б охоплювали процеси праці основних, обслуговуючих робітників і фахівців, належить до числа першочергових проблем розвитку теорії і практики нормування праці у даний час. В умовах ринкової економіки база знань про процеси праці для підприємств у закінченому вигляді повинна містити інформацію про умови і необхідні вимоги, що забезпечать досягнення тієї чи іншої виробничої мети. Для цього потрібно забезпечити якісний збір, обробку та використання максимально можливого обсягу наявної у виробничій системі статистичної інформації про перебіг виробничих процесів. Крім того, необхідно налагодити використання зібраної інформації для отримання похідних характеристик процесів виробництва для побудови адекватних прогнозів та ефективного планування.

Постановка завдання

У роботах [1–2], присвячених економіко-математичному моделюванню процесів праці зазначено, що моделювання показників, які характеризують процеси виробництва та праці, при використанні тільки середніх характеристик є не достатньо адекватним.

Середня оцінка має властивості незміщеності, ефективності і спроможності лише для обмеженої кількості видів розподілів, а для якісної обробки інформації в умовах малої її кількості та невизначеності законів розподілу необхідно застосовувати спеціальні стійкі (робастні) методи оцінювання. В зазначених роботах пропонується ввести теорію відхилень у практику нормування праці, в межах якої для всіх показників, з яких складаються норми праці, враховуються не тільки середні значення, а й дозволені допуски на помилку, що включає припущення про ймовірнісний характер розглянутих показників.

Якщо відійти від уявлення використовуваних в нормуванні праці величин як випадкових з невідомими ймовірнісними характеристиками та перейти до нечітких величин, то можна повною мірою скористатися перевагами апарату теорії нечітких множин [3]. Але тут виникає інша задача – побудови функцій належності нечітких показників [3], яку може бути вирішено за допомогою методів оцінювання центру розподілу та відхилень від нього у статистичних сукупностях, що представляють досліджувані величини. Ця задача еквівалентна задачам, що вирішуються в рамках теорії відхилень у практиці нормування праці. Визначення нормативної величини допущеного відхилення від норми праці, норми трудомісткості, запланованого обсягу виробництва пов'язане з отриманням оцінки центрального

значення величини та оцінки відхилення від центрального значення. Для розв'язання таких задач також найбільш доцільним є використання стійких (робастних) оцінок [1].

В даний час сформувалося два підходи до одержання стійких оцінок. Перший, найбільш часто уживаний у навчальній і науковій літературі, заснований на припущенні про те, що наявна вибірка «забруднена» «грубими» помилками, які необхідно спочатку виявити, а потім певним методом їх усунути або підправити, після чого обчислювати шукану оцінку. Для виявлення «грубих» значень у вибірці зазвичай рекомендують застосовувати критерії Смірнова-Граббса і Тит'єна-Мура, а для обробки даних – підходи Пуанкаре, Винзора, Хубера, Хампеля, Мешалкіна й ін.

Другий підхід, більш реалістичний, розглядає будь-яке значення у вибірці як рівномірне, що відбиває властивості вибіркової сукупності. Цей підхід для невеликих вибірок реалізується за допомогою методів «розмноження» вибірки: бустреп-метода (bootstrap) і джеккнайф-метода (jackknife) або метода «складного ножа» [1, 4].

В цій статті представлені дослідження з побудови довірчого інтервалу методом ентропії. Такі дослідження виявилися необхідними зважаючи на явні переваги цього методу в умовах застосування комп'ютерної техніки.

Результати та їх обговорення

Згідно з рекомендаціями по побудові гістограми частот для вибірки обсягу 1000 елементів, кількість інтервалів на ній може бути рівною 9, 11, 13, 15, 17 або 19.

Проведемо дослідження залежності довжини інтервалу ентропії від кількості інтервалів відповідної гістограми (для середніх, отриманих за методом бутстрепа вибіркам початкової сукупності [1] (бутстрепа-вибірки (табл. 1)).

Наведену залежність можна проілюструвати графічно за допомогою діаграм (рис. 1, 2). За отриманими результатами можна сказати, що при збільшенні кількості інтервалів на гістограмі спостерігається загальна тенденція на зменшення довжини інтервалу ентропії.

Наприклад, відмінність між значеннями 1-го і 6-го досвіду складає приблизно 0,5 %.

Таблиця 1. Залежність довжини та меж інтервалу ентропії від кількості інтервалів в гістограмі

Номер досліджу	Кількість інтервалів	Довжина інтервалу ентропії	Інтервал ентропії		Кількість значень сукупності в інтервалі	
			нижня межа	верхня межа	в од.	у %
1	9	0,4620	0,279	0,741	959	96,1
2	11	0,4582	0,265	0,723	963	94,5
3	13	0,4601	0,273	0,733	973	94,9
4	15	0,4582	0,263	0,721	963	96,8
5	17	0,4555	0,271	0,726	963	94,5
6	19	0,4594	0,275	0,734	973	94,9
Середнє значення		0,4589	0,271	0,729	966	96,9

Але за самими значеннями меж інтервалу ентропії можна сказати, що певної залежності не спостерігається. Коливання складають від 1 до 9 %.

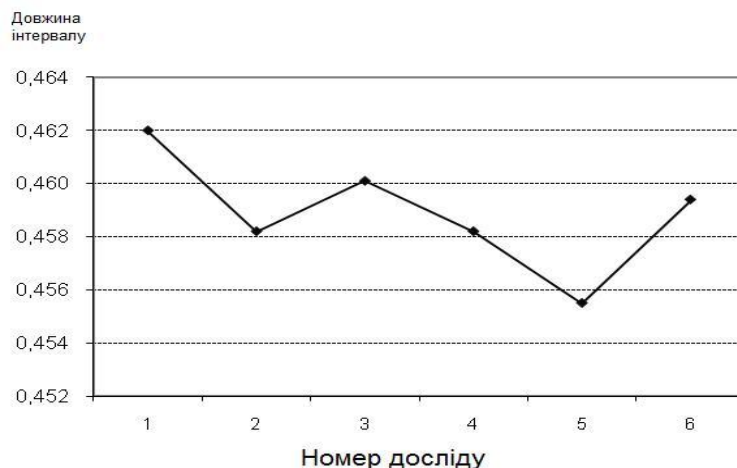


Рис. 1. Залежність довжини інтервалу ентропії від кількості інтервалів на гістограмі частот

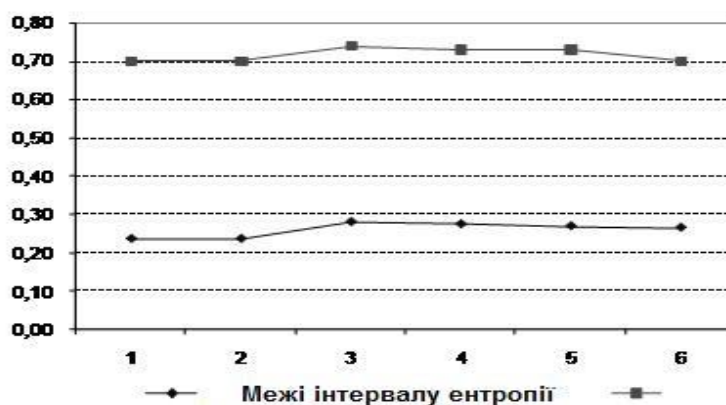


Рис. 2. Залежність значень меж інтервалу ентропії від кількості інтервалів на гістограмі

Проведені автором чисельні експерименти з різними методами оцінювання центральних значень вибірок та кількістю інтервалів гістограми дозволяють зробити висновок про відсутність певних закономірних відмінностей в оцінках меж інтервалу ентропії для бутстреп-середніх, наприклад, за бівес-оцінками і за простими середніми при різних значеннях числа інтервалів гістограми [1].

Таким чином, за відсутності чітко вираженої залежності, неможливо точно визначити яке оптимальне число інтервалів слід узяти для побудови гістограми частот і знаходження меж інтервалу ентропії. Тому можна прийняти для усіх подальших обчислень кількість інтервалів рівним 15.

Для визначення меж інтервалу ентропії запропонуємо інший алгоритм розрахунку. Позначимо: min і max – відповідно мінімальне і максимальне значення за отриманою штучною вибіркою з 1000 елементів; k – кількість інтервалів гістограми частот, побудованої за цією вибіркою; Δ_E – довжина інтервалу ентропії невизначеності для цього числа k ; h – значення зрушення меж інтервалів з довжиною Δ_E для пошуку значень меж ентропійного інтервалу ($h = \frac{(max - min) \cdot \Delta_E}{k}$).

Пошук розпочнемо з інтервалу, ліва межа якого L_1 співпадає з мінімальним значенням за даною вибіркою, а права розраховується за формулою: $R_i = L_i + \Delta_E$, $i = 1, 2, \dots$

Таким чином, виконуємо наступні кроки:

1. Розглянемо інтервал $[L_i; R_i]$ і підрахуємо кількість q_i значень з вибірки, що потрапили в нього.
2. Визначимо нове значення для лівої межі інтервалу $L_{i+1} = L_i + h$.
3. Визначимо значення правої межі інтервалу R_{i+1} за описаною вище формулою.
4. Повторюватимемо кроки 1–3 до тих пір, доки права межа інтервалу R_i не перевищить максимального значення за вибіркою.

Отриману вибірку значень q_1, q_2, q_3, \dots дослідимо на наявність максимуму і визначимо, який інтервал слід прийняти як інтервал ентропії.

У цій ситуації доводиться зіткнутися з проблемою неунімодальності отриманого розподілу частот. Числові експерименти показують, що однакова кількість значень з початкової вибірки може потрапляти в різні інтервали з довжиною інтервалу ентропії. Межі цих інтервалів розрізняються не істотно, але ця проблема не дозволяє автоматизувати пошук найкращого інтервалу серед наявних інтервалів за одним бутстреп-експериментом. Результати моделювання бутстреп-вбірок за однією і тією ж початковою вибіркою показують, що повторення бутстреп-експерименту до тих пір, доки досліджуваний розподіл частот матиме єдиний максимум, не набагато ускладнюють розрахунки, але призводять до цілком прийнятних результатів. Назвемо такий дослід формуванням бутстреп-вбірок, який призводить до розподілу частот по інтервалах довжини інтервалу ентропії з єдиним максимумом, «вдалим» дослідом. Процес отримання результатів по «вдалих» дослідіах можна повторити декілька разів та взяти середні значення меж отриманих інтервалів.

Проведені числові експерименти показують, що отримані значення меж інтервалів, яким відповідають максимальні частоти, відрізняються один від одного незначно (відмінність в сотих долях) для різних дослідів формування бутстреп-вбірок. Необхідно відмітити, що в інтервал ентропії потрапляє не менше 95 % значень з отриманої бутстреп-методом сукупності для 1000 бутстреп-повторень за початковою вибіркою. Таким чином, інтервал ентропії забезпечує рівень довірчої вірогідності для оцінки середнього значення по початковій вибірці не менше 95% для цих експериментів над початковою вибіркою. Повторимо процедуру оцінювання інтервалів невизначеності для центру розподілу заданої генеральної сукупності різними способами за отриманими з неї вибірками обсягів $n = 12, 16, 20, 25, 50$ (табл. 2).

Таблиця 2. Межі довірчих інтервалів для центрального значення генеральної сукупності за вибірками різного обсягу

Обсяг вибірки, n	Бутстреп-інтервал			Джеккнайф-інтервал			Класичний інтервал		
	ниж.	верх.	довжина	ниж.	верх.	довжина	ниж.	верх.	довжина
5	0,280	0,732	0,452	0,112	0,676	0,565	0,111	0,787	0,677
8	0,384	0,769	0,385	0,298	0,749	0,450	0,289	0,757	0,468
12	0,480	0,763	0,283	0,420	0,755	0,335	0,413	0,712	0,300
16	0,463	0,700	0,237	0,433	0,690	0,258	0,432	0,691	0,259
20	0,479	0,681	0,202	0,463	0,673	0,210	0,458	0,667	0,209
25	0,449	0,637	0,187	0,432	0,628	0,196	0,436	0,625	0,190
50	0,514	0,651	0,137	0,506	0,649	0,143	0,508	0,644	0,136

Найширшим інтервалом виявився джекнайф-інтервал. Бутстреп-інтервали за бівес-оцінками дещо зміщені вліво відносно бутстреп-інтервалів за середніми значеннями. Відмінність не значна, але зважаючи на лівобічне зміщення центру розподілу бутстреп-середніх за окремими дослідями (рис. 3), можна запропонувати спосіб розрахунку бівес-оцінок як метод визначення центрального значення за бутстреп-вибірками в запропонованому вище алгоритмі пошуку значень меж інтервалу ентропії за вибірками малого обсягу. Нижня межа джекнайф-інтервала практично співпадає зі значеннями нижньої межі класичних інтервалів, побудованих по вибірках різного обсягу.

На рис. 3 представлена діаграма залежності значень меж довірчих інтервалів від обсягу досліджуваної вибірки. На діаграмі чітко простежується залежність між значеннями меж інтервалів невизначеності, побудованими за різними початковими вибірками і обсягом цих вибірок. Причому, вже для початкової вибірки обсягу 20 відмінність між значеннями верхньої і нижньої межі довірчих інтервалів, побудованих різними методами, складає не більше 0,01 долі їх значень.

Для бутстреп-інтервалів в цих дослідях характерне зміщення меж праворуч відносно меж усіх інших інтервалів, що в принципі логічно, оскільки центральне значення генеральної сукупності дійсно лежить правіше вибірових середніх за цими вибірками.

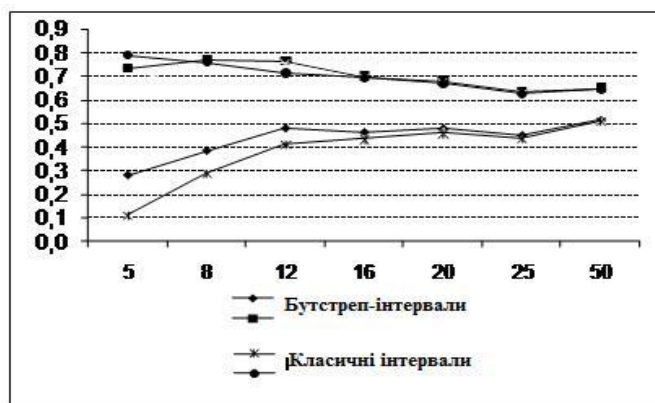


Рис 3. Залежність значень меж довірчих інтервалів, побудованих по вибірках різного обсягу

Виходячи з вищевикладеного, можна зробити висновок, що бутстреп-метод як метод «розмноження вибірок» у поєднанні з ентропійним методом побудови інтервалів невизначеності для початкових вибірок обсягу не менше 20 елементів дають результати цілком порівнянні з результатами по класичних методах. А для вибірок малого обсягу (менше 20 елементів) використання бутстреп-метода прийнятніше, з огляду на те, що для його використання не потрібно знання закону розподілу випадкової величини і призначення довірчої вірогідності.

Висновки

Отже, запропонований вище алгоритм знаходження меж інтервалу ентропії за бутстреп-середніми дозволяє автоматизувати пошук меж довірчих інтервалів за бутстреп-вибірками і забезпечує отримання усереднених оцінок для шуканих характеристик, що помітно зменшує розкид результатів за різними дослідями. Це дозволить використовувати усю наявну статистичну інформацію про процеси праці та виробничі процеси на підприємствах для побудови відповідних функцій належності нечітких величин для опису невизначеності нечіткої природи у характеристиках процесів, що мають місце на підприємствах.

Список використаної літератури

1. Горбатюк Е. В. Экономико-математические модели в нормировании труда: [монография] Т. П. Завгородняя, Е. В. Горбатюк. – Хмельницкий: НВП «Еврика» ТОВ, 2001. – 212 с.
2. Игумнов Б.Н., Завгородняя Т.П., Барский С.Н. Системы нормирования производственной деятельности. – Хмельницкий «Поділля», 1997 – 388 с.
3. Дилигенский Н. Нечеткое моделирование и многокритериальная оптимизация производственных систем в условиях неопределенности: технология, экономика, экология / Н. Дилигенский, Л. Дымова, П. Севастьянов. – М.: Машиностроение –1, 2005. – 238 с.
4. Эфрон Б. Нетрадиционные методы многомерного статистического анализа: [сб.статей: пер. с англ.] / Б. Эфрон. – М.: Финансы и статистика, 1988. – 263 с.

Стаття надійшла до редакції 17.10.2012

Построение интервалов энтропии с целью определения функций принадлежности нечетких показателей производственных процессов

Горбатюк Е.В.

Хмельницкий национальный университет

Рассмотрена проблема определения функции принадлежности показателей нечеткой природы, которые используются для характеристики производственных процессов при создании комплексных систем нормирования труда. Исследовано применение метода построения энтропийного интервала вместе с бутстрепа-методом к выборкам малого объема для определения центрального значения и отклонения от него. Проведено сравнительный анализ разных методов получения интервалов неопределенности для оценивания функций принадлежности нечетких показателей производственных процессов.

Ключевые слова: нормирования труда, робастные методы оценки, интервал энтропии, бутстреп-метод, бутстреп-выборка, бутстреп-эксперимент, бутстреп-интервал, неунимодальность, джеккнайф-метод, джеккнайф-интервал, интервал неопределенности.

A construction of entropy intervals with purpose of belonging functions determination for fuzzy performance indicators of production processes

Gorbatyuk K.

Khmelnysky National University

The problem of determining the membership functions of the fuzzy nature performance indicators used for the characteristics of production processes at the establishment of integrated systems of labour standardization. Investigated the application of the entropy interval construction method in conjunction with the bootstrap method to a small volume samples for definitions of the central values and deviations from it. The comparative analysis of different methods of obtaining uncertainty intervals for the evaluation of the membership functions of the production process fuzzy characteristics.

Keywords: valuation work, robust estimation methods, entropy interval, bootstrap method, bootstrap sampling, bootstrap experiment, bootstrap interval, nonunimodality, jackknife method, jackknife interval, the interval of uncertainty.