



Ю. І. Грицюк, П. Ю. Грицюк

Національний університет "Львівська політехніка", м. Львів, Україна

## МЕТОДИ УТОЧНЕНИХ ТОЧКОВИХ ОЦІНОК ПАРАМЕТРА РОЗПОДІЛУ ЙМОВІРНОСТЕЙ ВИПАДКОВОЇ ВЕЛИЧИНИ НА ПІДСТАВІ ОБМЕЖЕНОГО ОБСЯГУ ДАНИХ

Розроблено методику уточнених точкових оцінок параметра розподілу ймовірностей випадкової величини на підставі обмеженого обсягу статистичного матеріалу, що дало змогу виявити найбільш інформативний канал передачі даних і отримати його достовірну оцінку. Встановлено, що аналіз та оброблення даних здійснюються із залученням відомих методик з теорії ймовірностей та математичної статистики, де нагромаджено значний теоретичний і практичний досвід. Математичну модель, яка описує стан деякого об'єкта, процесу чи явища, подано у вигляді точкової оцінки параметра розподілу ймовірностей випадкової величини, значення якого отримують на підставі малої вибірки. Розглянуто традиційний підхід до виявлення найбільш інформативного каналу передачі даних про стан об'єкта, перебіг процесу чи явища і відсікання інших – менш достовірних. Це здійснюють за допомогою так званого механізму редукторів ступенів свободи, основний недолік якого в тому, що у відсічених каналах зв'язку може існувати деяка корисна інформація, яка не бере участі в процесі вироблення узгодженого рішення. Тому потрібно вводити механізми дискримінаторів ступенів свободи, що дадуть змогу всім каналам передачі даних брати участь в процесі підготовки рішення з вагомістю, яка відповідає найбільшому ступеню їх інформативності в поточній ситуації. Наведено ілюстративний приклад застосування розглянутих методів усереднення даних, у якому відображено результати розрахунків за ітераціями з використанням механізмів реалізації як редукторів, так і дискримінаторів ступенів свободи. Ці механізми відображають особливості реалізації ітераційних алгоритмів, характерних як для методів математичної статистики, так і для методів синергетичної системи усереднення даних.

**Ключові слова:** теорія ймовірностей та математична статистика; методи усереднення даних; інформативний канал передачі даних; механізм редукторів ступенів свободи; механізм дискримінаторів ступенів свободи; ітераційні алгоритми.

**Вступ.** При вирішенні багатьох завдань математичної статистики в розпорядженні дослідника, зазвичай, є тільки обмежений статистичний матеріал (невелика вибірка з генеральної сукупності), а точкова оцінка параметра  $\theta$  розподілу ймовірностей випадкової величини  $X$  вимагає якомога більшої точності (Gaskarov & Sharovalov, 1978). Причина в тому, що отримання кожного нового елемента вибірки, зазвичай, складний технологічний процес, пов'язаний зі значними труднощами технічного або економічного характеру. Водночас як труднощі, пов'язані з виконанням відповідних обчислень, мають значно меншу трудомісткість та низьку вартість реалізації (Lagutin, 2007).

Йдеться, наприклад, про якість програмного забезпечення (ПЗ), яка є однією із основних його характеристик у різних сферах використання інформаційних технологій (Pleskach & Zatonatska, 2011). Зазвичай, під оцінюванням якості ПЗ розуміють проведення процедури, яка визначає його властивості, призначення та ефективну роботу. Такий підхід до оцінювання ПЗ набуває особливого значення із розвитком і вдосконаленням технологій аналізу експертних даних (Morgun, 2011).

Усе це привело до потреби розроблення методів і засобів комплексного оцінювання різноманітних показників якості ПЗ, які б враховували як неточності у вхідній інформації, так і суб'єктивність оцінок експертів (Botsula & Morgun, 2008; Voronin, Ziatdinov & Kulinskiy, 2011).

Наявні методи визначення показників якості ПЗ на підставі суб'єктивних експертних оцінок, як правило, розроблені у припущенні, що їхній розподіл є нормальним (Bakhrushin & Ignahina, 2008; Bakhrushin, 2006; Brandt, 2003), оскільки багато випадкових величин підпорядковуються саме цьому закону. Припущення про нормальний розподіл експертних оцінок щодо визначення показників якості ПЗ часто виявляється також помилковим (Kobzar, 2006; Protasov, 2005). Дослідники на практиці найчастіше мають справу з вибіровими параметрами деякої випадкової величини, які визначають за невеликою кількістю значень досліджуваного показника. Їх вважають точковими оцінками параметрів розподілу і вони є випадковими. Тому, якщо розподіл експериментальних даних, зокрема й експертних оцінок, істотно відрізняється від нормального закону, то необхідно використовувати дещо інші методи та співвідношення для обчислення відповідних показників (Voro-

### Інформація про авторів:

**Грицюк Юрій Іванович**, д-р техн. наук, професор кафедри програмного забезпечення. Email: yurii.i.hrytsiuk@lpnu.ua;  
<https://orcid.org/0000-0001-8183-3466>; ResearcherID: V-3995-2017

**Грицюк Павло Юрійович**, здобувач кафедри автоматизованих систем управління. Email: pgrytsiuk1992@gmail.com

**Цитування за ДСТУ:** Грицюк Ю. І., Грицюк П. Ю. Методи уточнених точкових оцінок параметра розподілу ймовірностей випадкової величини на підставі обмеженого обсягу даних. Науковий вісник НЛТУ України. 2019, т. 29, № 2. С. 141–149.

**Citation APA:** Hrytsiuk, Yu. I., & Grytsiuk, P. Yu. (2019). The methods of the specified points of the estimates of the parameter of probability distribution of the random variable based on a limited amount of data. *Scientific Bulletin of UNFU*, 29(2), 141–149.  
<https://doi.org/10.15421/40290229>

pin, 2004, 2006, 2014).

**Аналіз останніх досліджень та публікацій.** Про застосування статистичних методів аналізу даних вперше згадується у Книзі Числа (четверта частина П'ятикнижжя Мойсея та Старого Заповіту). Основи сучасних методів аналізу даних були закладені Томасом Байєсом (байєсівський підхід, байєсівські оцінки), Данієлом Бернуллі (застосування нормального розподілу в теорії помилок, перші таблиці нормального розподілу, поділ помилок спостережень на випадкові й систематичні тощо), Карлом Гаусом і Адрієном Лежандром (метод найменших квадратів), Андрієм Колмогоровим (статистичні методи контролю за якістю, статистика Колмогорова-Смірнова, узагальнена відстань Колмогорова тощо), Вільфредо Парето (розподіл та діаграма Парето), Френсисом Гальтоном (теорія кореляції), Карлом Пірсоном (теорія кореляції, критерії згоди, метод головних компонент), Чарльзом Спірменом (техніка факторного аналізу, рангова кореляція), Рональдом Фішером (метод максимальної правдоподібності, критерії згоди тощо). Помітний внесок у розвиток цих методів зробив видатний український математик Михайло Остроградський, який у середині XIX ст. сформулював основні ідеї статистичного контролю за якістю виробництва.

Стосовно нашого дослідження, то у роботі (Vankovych, Zinko & Bozhenko, 2010) обґрунтовано метод усереднення для стохастичних систем з швидкозмінною фазою, до яких зводяться коливні процеси в істотно нелінійних системах. Також сформульовано теорему про оцінювання різниці між розв'язком точної та усередненої в першому наближенні систем на скінченному часовому інтервалі. У роботі (Mikitin, 2008) розглянуто програмну модель розрахунку середнього значення квадрата шумової напруги методом усереднення частинами. У роботі (Tuluchenko, 2008) запропоновано модифікацію методу баріцентричного усереднення, яка дає змогу формалізувати процедури вибору положення окремих обчислювальних шаблонів і обмеження їх кількості.

Однак, у їхніх працях не проведено порівняльний аналіз особливостей реалізації розрахункових алгоритмів, характерних як для методів математичної статистики, так і для методів синергетичної системи усереднення даних, не показано переваги одних і недоліки інших. Тому, як на наш погляд, видається *актуальним дослідження*, яке, використовуючи наявну інформацію про статистичні властивості досліджуваної випадкової величини, дало б змогу отримати розрахункові алгоритми для уточнених точкової оцінки параметра розподілу ймовірностей випадкової величини на підставі обмеженого обсягу статистичного матеріалу.

Запропонований у роботі (Voronin, 2004) підхід до усереднення даних дає змогу за обмеженої кількості каналів зв'язку отримувати максимально можливий обсяг доступної інформації. Автор поряд з концепцією редукторів ступенів свободи використовує так звані дискримінатори ступенів свободи, що дає можливість всім каналам зв'язку в міру їх інформативності в поточній ситуації брати участь у виробленні узгодженого рішення (Voronin, 2014). Спробуємо й ми адаптувати ці математичні методи для усереднення оцінок експертів стосовно якості ПЗ, які дають змогу виявити найбільшу кваліфікацію експерта і отримати його достовірну оцінку.

*Об'єкт дослідження* – усереднення даних при обмеженій кількості каналів зв'язку.

*Предмет дослідження* – методи та засоби уточнення точкової оцінки параметра розподілу ймовірностей випадкової величини на підставі обмеженого обсягу статистичного матеріалу.

*Метою дослідження* є розроблення методики уточнених точкових оцінок параметра розподілу ймовірностей випадкової величини на підставі обмеженого обсягу статистичного матеріалу, що дасть змогу виявити найбільш інформативний канал передачі даних і отримати його достовірну оцінку.

Для реалізації зазначеної мети потрібно виконати такі основні завдання:

- 1) навести приклади, які демонструють необхідність визначення відносного ступеня достовірності даних, отриманих різними каналами зв'язку про один і той самий стан об'єкта, перебіг процесу чи явища і потребу їх усереднення;
- 2) проаналізувати сучасні підходи до статистичного оцінювання випадкової величини, які дадуть змогу здійснити призначення певної функції апіорної щільності розподілу ймовірностей цієї величини;
- 3) навести ілюстративний приклад застосування механізмів усереднення даних, у якому відображено результати розрахунків за ітераціями з використанням механізмів реалізації як редукторів, так і дискримінаторів ступенів свободи;
- 4) зробити відповідні висновки та надати рекомендації щодо використання розробленої методики усереднення експертних оцінок якості ПЗ на підставі обмеженого обсягу статистичного матеріалу.

**1. Приклади уточнення середніх значень величин.** У сучасних програмних системах дані, що характеризують стан ( $x$ ) одного і того ж об'єкта, процесу чи явища, можна передавати декількома каналами зв'язку ( $n$ ). Проблема полягає у визначенні відносного ступеня достовірності даних, що надходять кожним каналом зв'язку в поточний момент часу, і у виробленні за допомогою механізму усереднення даних найбільш достовірної оцінки  $x^*$  дійсного стану об'єкта  $x$ , перебігу процесу чи явища при обмеженій кількості каналів зв'язку (Kartavy & Yarovaia, 2004).

Розглянемо деякі приклади, які демонструють необхідність визначення відносного ступеня достовірності даних і потребу їх усереднення (Voronin, 2004).

**Приклад 1.** Усереднення показів приладів, що мають різний клас точності (завдання поставлене в роботі (Voronin, & Ziatdinov, 2013)). При цьому кожен прилад привносить деякі збурення в результуючий показник згідно зі своїм класом точності.

**Приклад 2.** Усереднення сигналів для бістатичних радіолокацій малих небесних тіл (метод описаний в роботі (Voronin, 2004)). Для підвищення точності вимірювань при дослідженні параметрів руху таких тіл використовують бістатичну конфігурацію радіолокаційних систем. Дані від кожної з приймальних антен, рознесених на значні відстані одна від іншої, піддається обробленню й порівнянню їх між собою так, щоб результуючий сигнал був найбільш достовірним.

**Приклад 3.** Усереднення експертних оцінок з урахуванням ступеня їхньої компетентності у цьому питанні (Voronin, Ziatdinov, & Kulinskiy, 2011). Метод експертного оцінювання полягає в тому, що для визначення деякої кількісної характеристики стану об'єкта, перебігу процесу чи явища використовують висловлювання не однієї, а декількох фахівців (експертів), компетентних у предметній області знань. Передбачається, що справжнє значення невідомої кількісної характеристики знаходиться десь посе-

редині діапазону оцінок експертів, тобто їх узагальнена колективна думка є більш достовірною.

**Приклад 4.** Справжню оцінку якості ПЗ можна встановити за декількома критеріями: 1) орієнтація ПЗ на безпосереднього користувача; 2) орієнтація ПЗ на процес його виконання та отримані результати; 3) управління участю користувачів у роботі ПЗ та їх відповідальністю за нього; 4) неперервне вдосконалення функціональних можливостей ПЗ; 5) проблеми ПЗ, які залежать від користувачів; 6) проведення вимірювань продуктивності роботи ПЗ; 7) командна організація робіт з покращення якості ПЗ. Ці критерії є абсолютно правильними і для ПЗ як продукту, і для його розроблення як процесу. Дані про якість ПЗ за кожним із цих критеріїв надходять декількома каналами зв'язку (від декількох експертів), уточнюються, тобто виробляється найбільш достовірною оцінка справжньої якості ПЗ.

У наведених прикладах для усереднення даних розглядають невідому кількісну характеристику стану  $x$  деякого об'єкта, перебігу процесу чи явища як випадкову величину, відображенням закону розподілу якої є сигнал з каналу зв'язку. Для встановлення остаточної оцінки дані всіх каналів аналізують сукупно та обробляють як деякий початковий статистичний матеріал. Зазвичай, аналіз та оброблення даних здійснюють із залученням відомих методик з теорії ймовірностей та математичної

статистики (Gmurman, 2003, 2004; Turchin, 2006), де нагромаджено значний теоретичний і практичний досвід.

У повсякденній практиці кількість каналів передачі даних для подальшого їх усереднення відносно невелика, аналогічно як і у випадку малої вибірки в задачах математичної статистики (Zhuktenko & Nakonechny, 2000; Zhuktenko, Nakonechny & Savin, 2001; Tolbatov, 1994). Отже, модель стану деякого об'єкта, процесу чи явища можна подати як точкову оцінку параметра розподілу ймовірностей випадкової величини, яку отримано на підставі вибірки обмеженого обсягу (Gaskarov & Shapovalov, 1978; Tolbatov, 1994).

Стосовно прикладу 5, то зібрані в інтерактивному режимі оцінки якості ПЗ від кожного з експертів зберігають у базі даних, доступ до якої здійснюють через відповідний програмний засіб. Отримання оцінок від експертів має проходити у вигляді їхнього опитування з використанням ранжованої шкали за кожним з критеріїв. Отже, кожен з експертів має виставити відповідні оцінки, які потім заносять у відповідну таблицю. Зрозуміло, кожен з експертів буде по різному впливати на якість ПЗ за кожним критерієм і на узагальнений показник його якості.

**Табл. 1. Критерії оцінювання якості ПЗ та відповідні оцінки експертів**

№ з/п	Критерії оцінювання якості ПЗ	Експерти:					Середнє значення
		1	2	3	4	5	
		Оцінки експертів, бали					
1	Орієнтація ПЗ на безпосереднього користувача	7,9	9,6	9,2	8,7	9,9	9,06
2	Орієнтація ПЗ на процес його виконання та отримані результати	8,3	7,8	9,1	9,8	9,2	8,84
3	Управління участю користувачів у роботі ПЗ та їх відповідальністю за нього	7,1	9,8	7,6	7,9	6,9	7,86
4	Неперервне вдосконалення функціональних можливостей ПЗ	6,4	7,8	7,5	6,9	8,1	7,34
5	Проблеми ПЗ, які залежать від користувачів, мають становити до 20%	8,3	9,1	8,7	9,0	9,7	8,96
6	Проведення вимірювань продуктивності роботи ПЗ	7,4	9,8	8,6	9,2	8,4	8,68
7	Командна організація робіт з покращення якості ПЗ	8,7	9,5	9,6	7,6	9,4	8,96
	Загальна кількість балів	54,1	63,4	60,3	59,1	61,6	<b>59,70</b>

Оскільки результати цього дослідження можна застосувати не тільки для підвищення інформативності потенційних каналів передачі даних для сучасних програмних систем, але й в інших сферах наукової, виробничої та економічної діяльності, то подальше формулювання та вирішення завдань доцільно проводити в загальних термінах математичної статистики (Zhuktenko & Nakonechny, 2000; Zhuktenko, Nakonechny & Savin, 2001).

Основне завдання математичної статистики полягає в знаходженні розподілу досліджуваної випадкової величини  $X$  за даними наявної вибірки (Vogonin, 2004). В багатьох випадках вигляд розподілу величини  $X$  можна вважати відомим і завдання дослідника зводиться до визначення невідомого параметру  $\theta$  цього розподілу. Окрім цього, точкову оцінку параметра  $\theta$  розподілу ймовірностей випадкової величини  $X$  можна визначити на підставі оброблення статистичного матеріалу, який складається з сукупності експериментальних значень досліджуваної випадкової величини (Gmurman, 2003, 2004).

Отже, розглянемо неперервну дійсну випадкову величину  $X$ , щільність розподілу ймовірностей якої  $f(x|\theta)$  відома з точністю до невідомого параметра  $\theta$ .

Задано сукупність незалежних реалізацій випадкової величини  $X$  у вигляді такої множини:

$$\tilde{X} = \{x_j, j = \overline{1, n}\}, \quad (1)$$

де:  $x_j$  – значення  $i$ -ої випадкової величини;  $n$  – кількість незалежних реалізацій.

Утворену вибірку (1), яка інтерпретує сукупність даних, що підлягають усередненню, отримано  $n$  каналами їх передачі. За даними цієї вибірки потрібно визначити найкращу в деякому сенсі точкову оцінку  $\hat{\theta}$  невідомого параметра  $\theta$  розподілу ймовірностей випадкової величини  $X$ . Якщо випадкова величина  $X$  розподілена нормально з відомою дисперсією  $\sigma^2$ , то, згідно з законом нормально розподіленої випадкової величини

$$f(x|\theta) = f(x|m_x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(m_x - x)^2}{2\sigma^2}\right) \quad (2)$$

і параметром  $\theta$ , що підлягає оцінюванню, є математичне сподівання:  $\theta = m_x$ .

Згідно з інформацією, наведеною в роботі (Orlov, 2006), якість статистичних оцінок характеризується такими основними властивостями:

- 1) *спроможністю*, якщо при зростанні об'єму вибірки  $n$  вона прямує за ймовірністю до оцінюваного параметра;
- 2) *незміщеністю*, якщо її математичне сподівання дорівнює оцінюваному параметру;



3) *ефективністю*, якщо дисперсія мінімальна при заданому обсязі вибірки.

Якщо в розпорядженні дослідника є тільки та інформація, яка міститься у викладеній вище постановці завдання, то для визначення найкращих (спроможних, незміщених і ефективних) оцінок невідомого параметра  $\theta$  використовують метод максимальної правдоподібності, запропонований у 1912 р. англійським статистиком Рональдом Фішером у роботі (Sage & Mels, 1976). Тоді оцінка математичного сподівання нормально розподіленої випадкової величини має такий вигляд

$$\hat{\theta} = X_c = \frac{1}{n} \sum_{j=1}^n x_j, \quad (3)$$

де  $X_c$  – середнє значення елементів вибірки (1), що відображає оцінку математичного сподівання  $m_x$ .

Часто для малих вибірок оцінки максимальної правдоподібності, отримані за формулою (3), не забезпечують дослідника задовільною точністю (Turchin, 2006). Це вимагає розроблення більш ефективних процедур оцінювання параметра  $\theta$  з тим, щоб в конкретних програмних додатках (Bakhrushin & Ignahina, 2008; Tolbatov, 1994) достовірно відображати оцінку математичного сподівання  $m_x$ . Всі вони якоюсь мірою пов'язані з залученням додаткової інформації про статистичні властивості досліджуваної випадкової величини.

Отже, наведено деякі приклади, які демонструють необхідність визначення відносного ступеня достовірності даних з різних каналів зв'язку і потребу їх усереднення. Встановлено, що аналіз та оброблення даних здійснюють із залученням відомих методик з теорії ймовірностей та математичної статистики, де нагромаджено значний теоретичний і практичний досвід. Внаслідок аналізу та оброблення даних модель стану деякого об'єкта, процесу чи явища можна подати у вигляді точкової оцінки параметра розподілу ймовірностей випадкової величини, яку отримують на підставі вибірки обмеженого обсягу.

**2. Сучасні підходи до статистичного оцінювання випадкової величини.** Дієвим засобом підвищення ефективності статистичного оцінювання якості ПЗ за різними критеріями є Байєсовський підхід (Bakhrushin & Ignahina, 2008). Він полягає в тому, що параметр  $\theta$ , який підлягає оцінюванню, розглядають як реалізоване значення випадкової величини  $\Theta$ . Всю доступну йому попередню (до проведення процедури оцінювання) інформацію дослідник зводить до форми апіорного розподілу ймовірностей випадкової величини  $\Theta$ , що характеризується апіорною щільністю розподілу ймовірностей  $f_a(\theta)$ . Ця функція вважається відомою до початку аналізу даних, отриманих експериментально. Теорема Байєса комбінує апіорну щільність розподілу ймовірностей випадкової величини ( $X$ ) і дані спостережень  $\theta$  так, щоб утворилася апостеріорна щільність розподілу ймовірностей  $f(x|\theta)$ .

Фізичний зміст теореми Байєса полягає в тому, що якщо функція  $f_a(\theta)$  – апіорна щільність розподілу ймовірностей випадкової величини, яку приписують параметру  $\theta$  перед проведенням процедури оцінювання якості ПЗ, то функція  $f(x|\theta)$  буде апостеріорною щільністю розподілу ймовірності цієї величини, яку варто приписати параметру  $\theta$  після отримання даних. Статис-

тичні оцінки експертів, обчислені на підставі апостеріорного розподілу ймовірностей, мають кращу якість, ніж оцінки максимальної правдоподібності, оскільки використовують додаткову інформацію про невідомий параметр  $\theta$  у вигляді апіорної щільності розподілу ймовірностей  $f_a(\theta)$ .

Найбільш значущим моментом Байєсового оцінювання невідомого параметра  $\theta$  є призначення певної функції апіорної щільності розподілу ймовірностей випадкової величини  $f_a(\theta)$ . Ця функція має відповідати наявній попередній інформації про форму апіорного розподілу ймовірностей випадкової величини  $\Theta$ . Тут присутні такі два моменти. З одного боку, до цієї інформації належать тільки ті відомості, які є в апіорних даних. Тому за функцію апіорної щільності розподілу ймовірностей  $f_a(\theta)$  потрібно вибирати на підставі вигоди наявності максимально можливої ентропії (в сенсі теорії Шеннона (Shannon, 1963) при заданих умовах у вигляді конкретних апіорних даних, що розглядаються як обмеження (Gaskarov & Shapovalov, 1978). З іншого боку, нехтування будь-якою об'єктивною апіорною інформацією призводить до вибору менш інформативної функції апіорної щільності розподілу ймовірностей випадкової величини, що робить статистичне оцінювання менш ефективним.

В роботі (Voronin, 2014) було запропоновано вводити функцію апіорної щільності розподілу ймовірностей з визначеним параметром  $\theta$ , який розглядають як невідому константу. Така апіорна щільність розподілу ймовірностей вносить у функцію  $f_a(\theta)$  тільки ті відомості, які є в апіорних даних, і, водночас, дає змогу використовувати об'єктивну апіорну інформацію про вид закону розподілу оцінюваного параметра.

Детальні викладки алгоритму уточнення оцінки параметра  $\theta$  наведено в роботі (Voronin & Ziatdinov, 2013), а тут наведемо тільки загальний вигляд відповідної формули:

$$\hat{\theta} = \frac{\sum_{j=1}^n x_j f(x_j|\hat{\theta}') f_a(x_j|\hat{\theta}')}{\sum_{j=1}^n f(x_j|\hat{\theta}') f_a(x_j|\hat{\theta}')}. \quad (4)$$

Наявну в цьому алгоритмі залежність  $\hat{\theta} = \varphi(x_j, j = \overline{1, n}; \hat{\theta}')$  можна розв'язати ітераційним методом (Guter & Reznikovskii, 1971), реалізація якого випливає з такої рекурентної формули

$$\hat{\theta}^{(l)} = \varphi(x_j, j = \overline{1, n}; \hat{\theta}^{(l-1)}), l \in [1, L],$$

де  $l$  – номер поточної ітерації,  $L$  – кількість ітерацій. Ітераційна процедура завершується при виконанні такої умови:  $\hat{\theta}^{(L)} - \hat{\theta}^{(L-1)} \leq \lambda_\theta$ , де  $\lambda_\theta$  – задана точність обчислення оцінки  $\hat{\theta}$  невідомого параметра  $\theta$  апіорної щільності розподілу ймовірностей випадкової величини. Для аналізу збіжності застосовують відому теорему (Guter & Reznikovskii, 1971), згідно з якою для збіжності ітераційної процедури достатньо на розглянутому інтервалі уточнення оцінки  $\hat{\theta}$  дотримуватися такої нерівності

$$\left| \frac{d\varphi(x_j, j = \overline{1, n}; \hat{\theta})}{d\hat{\theta}} \right| < 1.$$

Отже, найбільш значущим моментом Байєсового оцінювання невідомого параметра є призначення певної функції апіорної щільності розподілу ймовірностей випадкової величини. Ця функція має відповідати наявній попередній інформації про форму апіорного розподілу ймовірностей цієї величини.

**Виявлення найбільш інформативного каналу передачі даних.** У складних синергетичних системах дані про стан об'єкта, перебіг процесу чи явища, зазвичай, передають декількома каналами зв'язку (Voronin, 2006; Kolesnikov, 1994). Проблема полягає в тому, щоб визначити, які канали зв'язку передають найбільш достовірні дані. Залежно від цього потрібно об'єднати (узагальнити) отримувані дані для вироблення узгодженого рішення про дійсний стан об'єкта, перебіг процесу чи явища.

Традиційний підхід передбачає виділення одного або кількох найбільш інформативних каналів передачі даних і відсікання інших – менш достовірних. Це здійснюють за допомогою так званого механізму "редукторів ступенів свободи", описаного в роботі (Kolesnikov, 1994). Перевага такого способу не тільки в простоті його реалізації, але в багатьох випадках він єдино можливий і фізично виправданий. Однак у відсічених каналах зв'язку може існувати деяка корисна інформація, яка не братиме участі в процесі вироблення узгодженого рішення.

Для розуміння зазначеного вище відомий такий почальний приклад (Voronin, 2004). Ще стародавні китайські мислителі вважали, що будь-який вибір одного варіанта з декількох можливих є не виправданим і збитковим, оскільки відкидає всі інші. Головним поняттям китайської культури завжди був вибір, при якому можливі гілки розвитку подій не відтиналися, а з'єднувалися в єдине ціле, що сукупно визнавалося єдино правильним рішенням.

Тому, часто під час синтезу складної синергетичної системи у процедурі усереднення даних варто відмовитися від концепції домінування одного з каналів передачі даних. Це означає, що замість *редукторів ступенів свободи* потрібно вводити механізми, що дають змогу всім каналам передачі даних брати участь в процесі підготовки рішення з вагомістю, що відповідає ступеню їх інформативності в поточній ситуації. Внаслідок цього вся доступна інформація буде належно використана через так звані *дискримінатори ступенів свободи*.

Синергетичний принцип усереднення даних (Voronin, 2006; Kolesnikov, 1994) має багато спільного з ідеями математичної статистики (Orlov, 2006). Якщо синергетичну концепцію поєднання даних застосовують для оцінювання дійсного стану об'єкта, перебігу процесу чи явища за наявної сукупності даних, то математична статистика вивчає методи оцінювання моментів розподілу ймовірностей випадкових величин за наявною сукупністю елементів вибірки. Спільність проблем обох теорій робить завдання дослідження синергетичних особливостей математичної статистики актуальною як для синергетики, так і для розвитку статистичних методів оброблення даних.

Розглянемо завдання уточненого статистичного оцінювання математичного сподівання  $m_x$  випадкової величини  $X$ , розподіленої за нормальним законом з апостеріорною щільністю  $f(x|\theta) = f(x|m_x)$ , заданою формулою (2), за результатами випадкової вибірки (1), якщо відома дисперсія  $\sigma^2$ . Апіорна інформація містить оцін-

ку математичного сподівання  $X_c$ , розподілену також за нормальним законом з відомою дисперсією  $\sigma^2/n$ .

Після нескладних математичних перетворень, наведених у роботі (Voronin, 2006), з виразу (4) було отримано алгоритм обчислення оцінки математичного сподівання  $X_c$ , для реалізації якого потрібно організувати таку ітераційну процедуру:

$$X_c^{(l)} = \frac{\sum_{j=1}^n x_j \text{Exp}\left(-\frac{(x_j - X_c^{(l-1)})^2 (n+1)}{2\sigma^2}\right)}{\sum_{j=1}^n \text{Exp}\left(-\frac{(x_j - X_c^{(l-1)})^2 (n+1)}{2\sigma^2}\right)}, \forall l = 1, 2, 3, \dots, \quad (5)$$

де  $l$  – номер поточної ітерації. Як нульове наближення у цій процедурі доцільно прийняти оцінку максимальної правдоподібності (3), а саме:

$$X_c^{(0)} = \frac{1}{n} \sum_{j=1}^n x_j.$$

Ітераційна процедура (5) під час застосування її до задач усереднення даних (приклади 1–4) відображає механізм реалізації *дискримінаторів ступенів свободи*.

Результати багатьох тестів (Voronin, 2004, 2006; Voronin & Ziatdinov, 2013) показують, що довірчий інтервал, який відповідає уточненій оцінці (5), менше довірчого інтервалу оцінки максимальної правдоподібності (3). Найбільший вигравш в ефективності виходить для малих вибірок, оскільки зі збільшенням кількості вимірювань відносний внесок апіорної інформації при отриманні оцінок поступово стає меншим, а байєсівська оцінка (5) й оцінка максимальної правдоподібності (3) асимптотично збігаються (Orlov, 2006). Тому обчислювати уточнену оцінку параметра доцільно, в основному, при малих обсягах вибірки випадкової величини.

Важливою властивістю апіорної щільності розподілу ймовірностей випадкової величини є те, що вона не повинна бути власною щільністю, тобто не обов'язково її інтеграл має дорівнювати одиниці (Bakhrushin & Ignahina, 2008). У деяких випадках вважаються цілком виправданими спроби використання псевдобайєсівських оцінок, при отриманні яких замість відсутньої апіорної щільності розподілу ймовірностей оцінюваного параметра вводять іншу щільність. Заслугує уваги можливість використання для відображення апіорної щільності розподілу ймовірностей випадкової величини для однієї з так званих потенційних функцій (Yzerman, Braverman & Rozonoer, 1970), окремим випадком яких є нормальний закон розподілу (2). Альтернативами також можуть слугувати ще й такі функції:

$$f_1(\alpha) = \frac{\alpha}{|X_c - x|}, f_2(\beta) = \frac{\beta}{(X_c - x)^2}, f_3(\gamma, \delta) = \frac{\gamma}{1 + \delta(X_c - x)^2}, \quad (6)$$

де  $\alpha, \beta, \gamma, \delta$  – деякі константи чи поточні значення випадкової величини. Будь-яка з потенційних функцій (6) характерна тим, що вона монотонно збуває з віддаленням від значення  $X_c$ , тобто є симетрично-парною щодо оцінки  $X_c$ . Якщо відомо тільки те, що оцінюваний параметр розподілений в генеральній сукупності симетрично, то доцільно отримати уточнену оцінку  $X_c$ , вибравши для апіорної щільності розподілу ймовірностей досить просту потенційну функцію.

Іноді для скорочення обсягу обчислень доцільно навмисно замінити відомий (наприклад, нормальний) закон розподілу іншою, більш простою потенційною функцією (Voronin, 2004). Наприклад, якщо випадкова

величина розподілена за рівномірним законом, то оцінка її середнього значення підпорядковується нормальному закону розподілу. Однак, якщо за апіорну щільність розподілу ймовірностей вибрати функцію  $f_1(\alpha)$  з потенційних функцій (6), то прийдемо до такого простого ітераційного алгоритму обчислення оцінки математичного сподівання  $X_c$ :

$$X_c^{(0)} = \frac{1}{n} \sum_{j=1}^n x_j; X_c^{(l)} = \sum_{j=1}^n \frac{x_j}{|X_c^{(l-1)} - x_j| \sum_{j=1}^n \frac{1}{|X_c^{(l-1)} - x_j|}}; \quad (7)$$

$$l \in [1, L], |X_c^{(l)} - X_c^{(l-1)}| \leq \lambda_x.$$

Цей алгоритм при його застосуванні до процедури розв'язання задачі усереднення даних відображає механізм реалізації *редукторів ступенів свободи*.

Наведена вище методика усереднення даних на підставі обмеженого обсягу статистичного матеріалу передбачає індивідуальний підхід до кожної реалізації випадкової величини – зважування відповідно до апостеріорної ймовірності її появи. Тобто, вона дає змогу усунути втрати цінної інформації при обчисленні шуканих оцінок параметра з малої вибірки (Gaskarov & Sharov, 1978).

Якщо за апіорну щільність розподілу ймовірностей випадкової величини вибрати функцію  $f_2(\beta)$  чи  $f_3(\gamma, \delta)$  з потенційних функцій (6), то прийдемо до дещо інших простих ітераційних алгоритмів обчислення достовірної оцінки математичного сподівання  $X_c$ , а саме:

$$X_c^{(0)} = \frac{1}{n} \sum_{j=1}^n x_j; X_c^{(l)} = \sum_{j=1}^n \frac{x_j}{(X_c^{(l-1)} - x_j)^2 \sum_{j=1}^n \frac{1}{(X_c^{(l-1)} - x_j)^2}}; \quad (8)$$

$$l \in [1, L], |X_c^{(l)} - X_c^{(l-1)}| \leq \lambda_x;$$

$$X_c^{(l)} = \sum_{j=1}^n \frac{x_j}{(1 + x_j(X_c^{(l-1)} - x_j))^2 \sum_{j=1}^n \frac{1}{(1 + x_j(X_c^{(l-1)} - x_j))^2}}; \quad (9)$$

$$l \in [1, L], |X_c^{(l)} - X_c^{(l-1)}| \leq \lambda_x.$$

Автор роботи (Voronin, 2004) вважає, що для отримання достовірної оцінки математичного сподівання  $X_c$  за допомогою організації ітераційних процедур (8) і (9) мають обов'язково взаємодіяти між собою всі елементи вибірки на кожній ітерації. Аналогічно синергетична система усереднення даних передбачає процедуру, що характеризується самоврядуванням і самоорганізацією

**Табл. 2. Результати розрахунків за ітераціями з використанням механізму реалізації редукторів ступенів свободи**

Ітерація	Експерти та їхні оцінки:					Сума	Позначення	Середнє знач.	$\lambda_x =$	0,005
	1	2	3	4	5					
	<b>5,0</b>	<b>6,5</b>	<b>4,3</b>	<b>5,2</b>	<b>6,0</b>	27,000	$X_c[0]=$	5,400		
1	0,400	1,100	1,100	0,200	0,600					
	2,500	0,909	0,909	5,000	1,667	10,985				
	1,138	0,538	0,356	2,367	0,910	5,309	$X_c[1]=$	5,309	0,091	не виконано
2	0,309	1,191	1,009	0,109	0,691					
	3,237	0,840	0,991	9,177	1,447	15,692				
	1,031	0,348	0,272	3,041	0,553	5,245	$X_c[2]=$	5,245	0,064	не виконано
3	0,245	1,255	0,945	0,045	0,755					
	4,078	0,797	1,058	22,106	1,325	29,363				
	0,694	0,176	0,155	3,915	0,271	5,211	$X_c[3]=$	5,211	0,034	не виконано
4	0,211	1,289	0,911	0,011	0,789					
	4,735	0,776	1,097	89,437	1,268	97,314				
	0,243	0,052	0,048	4,779	0,078	5,201	$X_c[4]=$	5,201	0,010	не виконано
5	0,201	1,299	0,901	0,001	0,799					
	4,977	0,770	1,110	1105,01	1,251	1113,116				
	0,022	0,004	0,004	5,162	0,007	5,200	$X_c[5]=$	<b>5,200</b>	0,001	виконано
	0,200	1,300	0,900	0,000	0,800	0,000	$\leftarrow$ min			

відповідно до поставленої мети. Тут складні процеси розвиваються за допомогою колективної взаємодії компонент. Узгодження компонент дає змогу використовувати резервні можливості синергетичної системи і тим самим значно підвищує ступінь емерджентності (системний ефект).

Отже, механізм редукторів ступенів свободи є традиційним підходом до виявлення найбільш інформативного каналу передачі даних про стан об'єкта, перебіг процесу чи явища і відсікання інших – менш достовірних (Voronin, Ziatdinov & Kulinskiy, 2011). Його основний недолік у тому, що у відсічених каналах зв'язку може існувати деяка корисна інформація, яка не братиме участі в процесі вироблення узгодженого рішення. Тому потрібно вводити механізми дискримінованих ступенів свободи, які дають змогу всім каналам передачі даних брати участь в процесі підготовки узгодженого рішення з вагомністю, що відповідає найбільшому ступеню їх інформативності в поточній ситуації.

**3. Ілюстративний приклад застосування механізмів усереднення даних.** Нехай кількість каналів передачі даних становить  $n \geq 3$  (кількість ступенів свободи в синергетичній системі усереднення даних). Вхідні дані подамо у вигляді одновимірного масиву  $\tilde{X} = \{x_j, j = \overline{1, n}\}$ , де  $x_j, j = \overline{1, n}$  – значення деякої випадкової величини  $X$ , отриманої  $j$ -им каналом зв'язку (компонента системи усереднення даних).

Залежно від практичних вимог (див. приклади 1–4) необхідно виявити найбільш інформативний канал передачі даних і отримати достовірну оцінку  $x^*$  випадкової величини  $X$ .

Нехай масив вхідних даних має такий вигляд

$$\tilde{X} = \{x_j, j = \overline{1, 5}\} = \{5,0, 6,5, 4,3, 5,2, 6,0\}.$$

Умова зупинки ітераційної процедури становить

$$|X_c^{(l)} - X_c^{(l-1)}| \leq \lambda_x = 0,005.$$

Оцінка нульової ітерації має таке значення:

$$X_c^{(0)} = \frac{1}{5} \cdot (5,0 + 6,5 + 4,3 + 5,2 + 6,0) = 27,0/5 = 5,40.$$

**Варіант 1.** Розглянемо завдання визначення найбільш інформативного каналу передачі даних і застосуємо механізм реалізації редукторів ступенів свободи, вираженого ітераційним алгоритмом (7). Результати розрахунків за ітераціями наведено в табл. 2.



Отже, внаслідок реалізації ітераційного алгоритму (7) отримано достовірну оцінку випадкової величини, яка становить  $x^* = X_c^{(n)} = X_c^{(5)} = 5,20$ , а найбільш інформативний канал передачі даних потрібно визначити за таким виразом:

$$x^* = \min \{ |X_c^{(j)} - x_j|, j = \overline{1, n} \}. \quad (10)$$

У нашому прикладі найбільш інформативним виявився 4-ий канал зв'язку.

На рисунку наведено відповідні результати розрахунку, виконані за допомогою відповідного програмного додатку.

Рис. Програмний додаток для виконання відповідних розрахунків

**Варіант 2.** Розглянемо завдання визначення найбільш достовірної оцінки випадкової величини  $X_c$  за допомогою механізму реалізації дискримінованих ступенів свободи, вираженого ітераційним алгоритмом (5). При цьому передбачається, що випадкова величина  $X_c$  роз-

поділена в генеральній сукупності нормально з відомою дисперсією.

Результати розрахунків за ітераціями наведено в табл. 3.

Табл. 3. Результати розрахунків за ітераціями з використанням механізму реалізації дискримінованих ступенів свободи

Ітерація	Експерти та їхні оцінки:					Сума	Позначення	Середнє знач.	$\lambda_c =$	0,005	
	1	2	3	4	5					$ X_c[l] - X_c[l-1] $	
	5,0	6,5	4,3	5,2	6,0	27,000	$X_c[0]=$	5,400			
1	0,644	4,872	4,872	0,161	1,450						
	0,525	0,008	0,008	0,851	0,235	1,626					
	2,625	0,050	0,033	4,426	1,408	8,542	$X_c[1]=$	5,253	0,147	не виконано	
2	0,257	6,264	3,655	0,011	2,249						
	0,773	0,002	0,026	0,989	0,106	1,895					
	3,866	0,012	0,111	5,142	0,633	9,765	$X_c[2]=$	5,152	0,101	не виконано	
3	0,093	7,317	2,923	0,009	2,896						
	0,911	0,001	0,054	0,991	0,055	2,012					
	4,556	0,004	0,231	5,152	0,332	10,275	$X_c[3]=$	5,108	0,044	не виконано	
4	0,047	7,805	2,627	0,034	3,206						
	0,954	0,000	0,072	0,966	0,041	2,034					
	4,772	0,003	0,311	5,025	0,243	10,353	$X_c[4]=$	5,090	0,017	не виконано	
5	0,033	8,001	2,516	0,048	3,332						

	0,968	0,000	0,081	0,953	0,036	2,037				
	4,838	0,002	0,348	4,954	0,214	10,357	$X_c[5]=$	5,084	0,007	не виконано
6	0,028	8,079	2,472	0,055	3,382					
	0,972	0,000	0,084	0,947	0,034	2,038				
	4,861	0,002	0,363	4,924	0,204	10,354	$X_c[6]=$	<b>5,081</b>	0,003	виконано
	0,081	1,419	0,781	0,119	0,919	0,081	$\leftarrow$ min			

Примітка:  $n = 5$ ;  $2\sigma^2 = 1,490$ ;  $(n+1)/2\sigma^2 = 4,027$ .

Отже, отримано значення випадкової величини  $x^* = X_c^{(7)} = X_c^{(6)} = 5,081$ , що є найбільш достовірною оцінкою при заданих припущеннях. Водночас, найбільш інформативним каналом передачі даних, визначеним за (10), виявився 1-ий канал.

Проведений вище порівняльний аналіз особливостей реалізації ітераційних алгоритмів (5) і (7), характерних як для методів математичної статистики, так і для методів синергетичної системи усереднення даних, показав деякі переваги одних, так і недоліки інших. Результати аналізу можна застосовувати для підвищення достовірності отримання статистичної оцінки випадкової величини, яку обчислюють за наявності малої вибірки експертних оцінок чи експериментальних даних, а також для визначення оцінок дійсного стану об'єкта, перебігу процесу чи явища у синергетичних системах усереднення даних при обмеженій кількості каналів зв'язку.

## Висновки

Розроблено методику уточнених точкових оцінок параметра розподілу ймовірностей випадкової величини на підставі обмеженого обсягу статистичного матеріалу, що дало змогу виявити найбільш інформативний канал передачі даних і отримати його достовірну оцінку. За результатами дослідження можна зробити такі основні висновки.

1. Наведено деякі приклади, які демонструють необхідність визначення відносного ступеня достовірності даних, отриманих різними каналами зв'язку про один і той самий стан об'єкта, перебіг процесу чи явища і потребу їх усереднення. Встановлено, що аналіз та оброблення даних здійснюють із залученням відомих методик з теорії ймовірностей та математичної статистики, де нагромаджено значний теоретичний і практичний досвід. Запропоновану модель, яка описує стан деякого об'єкта, процесу чи явища, подано у вигляді точкової оцінки параметра розподілу ймовірностей випадкової величини, значення якого отримують на підставі малої вибірки.

2. Проаналізовано сучасні підходи до статистичного оцінювання випадкової величини, найпоширенішим з яких є Байєсовський підхід. Встановлено, що найбільш значущим моментом Байєсового оцінювання невідомого параметра є призначення певної функції апіорної щільності розподілу ймовірностей випадкової величини. Ця функція має відповідати наявній попередній інформації про форму апіорного розподілу ймовірностей цієї величини.

3. Розглянуто традиційний підхід до виявлення найбільш інформативного каналу передачі даних про стан об'єкта, перебіг процесу чи явища і відсікання інших – менш достовірних. Це здійснюють за допомогою так званого механізму редукторів ступенів свободи, основний недолік якого в тому, що у відсічених каналах зв'язку може існувати деяка корисна інформація, яка не бере участі в процесі вироблення узгодженого рішення. Тому потрібно вводити механізми дискримінації ступенів свободи, що дадуть змогу всім каналам передачі даних брати участь в процесі підготовки рішення з вагомністю, яка відповідає найбільшому ступеню їх інформативності в поточній ситуації.

4. Наведено ілюстративний приклад застосування розглянутих методів усереднення даних, у якому відображено результати розрахунків за ітераціями з використанням механізмів реалізації як редукторів, так і дискримінації ступенів свободи. Ці механізми відображають особливості реалізації ітераційних алгоритмів, характерних як для методів математичної статистики, так і для методів синергетичної системи усереднення даних.

5. Зроблено відповідні висновки та надано рекомендації щодо використання розробленої методики уточнених точкових оцінок параметра розподілу ймовірностей випадкової величини на підставі обмеженого обсягу статистичного матеріалу.

5. Зроблено відповідні висновки та надано рекомендації щодо використання розробленої методики уточнених точкових оцінок параметра розподілу ймовірностей випадкової величини на підставі обмеженого обсягу статистичного матеріалу.

## Перелік використаних джерел

- Bakhrushin, V. E. (2006). *Data Analysis: a tutorial*. Zaporozhye: PG "Humanities", 128 p. [In Ukrainian].
- Bakhrushin, V. E., & Ignahina, M. A. (2008). Application of statistical methods in processing the results of production control in metallurgy of semiconductors. *System Technology*, 3(56), Vol. 1, 3–7. [In Russian].
- Botsula, M., & Morgun, I. (2008). The problem of quality examination of distance courses. *Scientific Journals of Vinnytsia National Technical University*, 4, 1–7. Retrieved from: <http://nbuv.gov.ua/e-journals/vntu/2008-4/2008-4.files/uk/08mpbcme.uk.pdf>. [In Ukrainian].
- Brandt, Z. (2003). *Data analysis: Statistical and Computational Methods for Scientists and Engineers*. Moscow: Mir, AST, 686 p. [In Russian].
- Gaskarov, D., & Shapovalov, V. I. (1978). *Small sample*. Moscow: Statistics, 248 p. [In Russian].
- Gmurman, B. E. (2004). *Guide to solving problems of the theory of probability and mathematical statistics*. Moscow: Higher School, 404 p. [In Russian].
- Gmurman, V. E. (2003). *Probability theory and mathematical statistics*. Moscow: Higher School, 479 p. [In Russian].
- Gryciuk, Yu. I., & Grytsyuk, M. Yu. (2016). Valuation models of anthropogenic factors during the appearance fires in the storage of petroleum products. *Proceedings of the Forestry Academy of Sciences of Ukraine*, 14, 198–209. <https://doi.org/10.15421/411628>
- Grytsyuk, Yu. I., & Leshkevych, I. F. (2017). The Problems of Definition and Analysis of Software Requirements. *Scientific Bulletin of UNFU*, 27(4), 148–158. <https://doi.org/10.15421/40270433>
- Guter, R. S., & Reznikovskii, P. T. (1971). *Programming and computational mathematics*. Moscow: Science. Vol. 2, 273 p. [In Russian].
- Hrytsiuk, Yu. I. (2018). *Analysis of Software Requirements: Tutorial*. Lviv: Publishing House of Lviv Polytechnic, 460 p. Retrieved from: <https://192.168.253.4/Research/TrainingAidsEdit.aspx?id=11750>. [In Ukrainian].
- Hrytsiuk, Yu. I., & Buchkovska, A. Yu. (2017). Visualization of the Results of Expert Evaluation of Software Quality Using Polar Diagrams. *Scientific Bulletin of UNFU*, 27(15), 137–145. <https://doi.org/10.15421/40271025>
- Hrytsiuk, Yu. I., & Dalyavskyy, V. S. (2018). Using Petal Diagram for Visualizing the Results of Expert Evaluation of Software Quality



- lity. *Scientific Bulletin of UNFU*, 28(9), 95–104. <https://doi.org/10.15421/40280919>
- Hrytsiuk, Yu. I., & Nemova, E. A. (2018). Peculiarities of Formulation of Requirements to the Software. *Scientific Bulletin of UNFU*, 28(7), 135–148. <https://doi.org/10.15421/40280727>
- Hrytsiuk, Yu. I., & Zhabych, M. R. (2018). Risk Management of Implementation of Program Projects. *Scientific Bulletin of UNFU*, 28(1), 150–162. <https://doi.org/10.15421/40280130>
- Kartavy, V., & Yarovaya, V. (2004). *Mathematical Statistics*. Kyiv: Professional, 484 p. [In Ukrainian].
- Kobzar, A. I. (2006). *Applied Mathematical Statistics*. Moscow: Fizmatlit, 816 p. [In Russian].
- Kolesnikov, A. A. (1994). *Synergetic theory of management*. Moscow: Energoatomisdat, 344 p. [In Russian].
- Lagutin, M. B. (2007). *Transparent mathematical statistics*. Moscow: Binom, 472 p. [In Russian].
- Mikitin, J. P. (2008). Programming model averaging method Noise. *Bulletin of the National University "Lviv Polytechnic". Series: Computer Science and Information Technology*, 629, 21–24. [In Ukrainian].
- Morgun, I. (2011). The method of peer review software quality. *Software Engineering: mater. Intern. nauk. and practical. Conf. graduate students*, 2(6), 33–37. Vinnytsia. Retrieved from: <http://jrn1.nau.edu.ua/index.php/IPZ/article/view/3086>. [In Ukrainian].
- Orlov, A. I. (2006). *Applied Statistics*. Moscow: Exam, 671 p. [In Russian].
- Pleskach, V. L., & Zatonatska, T. (2011). *Information systems and technology in enterprises: textbook*. Kyiv: Knowledge. 718 p. Retrieved from: [http://pidruchniki.com/1194121347734/informatika/analiz\\_yakosti\\_programnogo\\_zabezpechennya#42](http://pidruchniki.com/1194121347734/informatika/analiz_yakosti_programnogo_zabezpechennya#42). [In Ukrainian].
- Protasov, K. V. (2005). *Statistical analysis of experimental data*. Moscow: Mir, 142 p. [In Russian].
- Sage, E., & Mels, J. (1976). *Estimation theory and its application in communication and management*. Moscow: Communication, 496 p. [In Russian].
- Shannon, K. (1963). *Work on information theory and cybernetics*. Moscow: Publishing House of Foreign Literature, 829 p. [In Russian].
- Tolbatov, A. (1994). *Mathematical Statistics and task optimization algorithms and programs*. Kyiv: High School, 226 p. [In Ukrainian].
- Tuluchenko, G. Y. (2008). Geometry computing templates bars of centric averaging method. *Bulletin of the Zaporizhzhya National University*, 1, 187–190. [In Ukrainian].
- Turchin, V. (2006). *Probability and Mathematical Statistics: Concepts, examples, problem*. Dnepropetrovsk: Dniprovsky National University, 476 p. [In Ukrainian].
- Vankovych, T.-N. M., Zinko, J. A., & Bozhenko, M. (2010). An averaging method for oscillating stochastic systems with quick phase. *Bulletin of the National University "Lviv Polytechnic". Series: Dynamics, Durability and Design of Machines and Devices*, 678, 11–14. [In Ukrainian].
- Voronin, A. N. (2004). Method of interconnecting signals for bistatic radar small celestial bodies. *System analysis and management: meas. rep. 9th International. Conf.*, (pp. 113–114). Moscow: Publishing house of the Moscow Aviation Institute. [In Russian].
- Voronin, A. N. (2006). Synergistic methods of data aggregation. *Cybernetics and Systems Analysis*, 2, 24–30. [In Russian].
- Voronin, A. N. (2014). Methods of data aggregation. *Cybernetics and Systems Analysis*, 50(5), 78–84. [In Russian].
- Voronin, A. N., & Ziatdinov, J. K. (2013). *Theory and practice of multi-criteria decisions: models, methods, implementation*. Saarbrücken (Deutschland); Lambert Academic Publishing, 305 p. [In Russian].
- Voronin, A. N., Ziatdinov, J. K., & Kulinskiy, M. V. (2011). *Multicriteria task: models and methods: a monograph*. Kyiv: NAU, 348 p. [In Russian].
- Yzerman, M. A., Braverman, E. M., & Rozonoer, L. I. (1970). *The method of potential functions in machine learning theory*. Moscow: Science, 384 p. [In Russian].
- Zhlukenko, V. I., & Nakonechny, S. (2000). *Probability and Mathematical Statistics: training method. manual*. In 2 parts. Part I. Probability. Kyiv: Kyiv National Economic University, 304 p. [In Ukrainian].
- Zhlukenko, V. I., Nakonechny, S., & Savin, S. (2001). *Probability and Mathematical Statistics: training method. manual*. In 2 parts. Part II. Mathematical Statistics. Kyiv: Kyiv National Economic University, 336 p. [In Ukrainian].

**Yu. I. Hrytsiuk, P. Yu. Grytsyuk**

*Lviv Polytechnic National University, Lviv, Ukraine*

## **THE METHODS OF THE SPECIFIED POINTS OF THE ESTIMATES OF THE PARAMETER OF PROBABILITY DISTRIBUTION OF THE RANDOM VARIABLE BASED ON A LIMITED AMOUNT OF DATA**

The method of the specified points of the estimates of the parameter the probability distribution of the random variable based on the limited amount of statistical data, which allowed to identify the most informative data link and get it accurate assessment, has been developed. It is established that data analysis and processing are carried out with the help of well-known methods on probability theory and mathematical statistics, where considerable theoretical and practical experience is accumulated. The mathematical model that describes the state of an object, process or phenomenon presented as point estimates of the parameter of probability distribution of the random variable whose value is obtained based on a small sample. The traditional approach to identify the most informative data channel on the state of the object, or the process progress or phenomena and clipping other – less reliable, has been considered. This is carried out by so-called gear mechanism of reducers of degrees of freedom, which is the main drawback is that the cut off communication channels may be some useful information, which is not involved in the process of developing an agreed solution. So, it is needed to enter arrangements of discriminator degrees of freedom that allow all data channels to participate in the process preparation of solution with significance, which corresponds to the level of their information content in the current situation. An illustrative example is given of applying considered methods of averaging data that shows the results of calculations per iteration using gear mechanisms of implementation as a reducers as well as discriminators degrees of freedom. These arrangements reflect the peculiarities of the implementation of iterative algorithms for specific as for methods of mathematical statistics and methods for synergistic system of data averaging.

**Keywords:** probability theory and mathematical statistics; data averaging methods; informative data channel; gear reducers degrees of freedom; mechanism of discriminators of the degrees of freedom; iterative algorithms.