С.В. Гадецька¹, В.Ю.Дубницький², Ю.І. Кушнерук³, О.І. Ходирев²

¹ Харківський національний автомобільно-дорожній університет, Харків, Україна

² Харківський навчально-науковий інститут Державного вищого навчального закладу
“Університет банківської справи”, Харків, Україна

³ Харківський національний університет Повітряних Сил ім. І. Кожедуба, Харків

ПЛАНУВАННЯ ЕКСПЕРИМЕНТУ ПРИ РОЗВ'ЯЗАННІ ОБЕРНЕНОЇ ЗАДАЧІ ПОБУДОВИ ТОЛЕРАНТНИХ (РЕФЕРЕНСНИХ) ІНТЕРВАЛІВ

Поставлено і розв'язано для деяких окремих випадків, важливих в практичній діяльності, обернену задачу побудови толерантних інтервалів. Розв'язок отримано для планування експерименту в непараметричному випадку, а також для рівномірного розподілу, показникового розподілу, розподілу Вейбулла, нормального розподілу, логарифмічно нормального розподілу. Запропоновано чисельні методи розв'язання поставлених задач, доступних для найбільш поширених програмних продуктів. Прямою задачею побудови толерантних інтервалів в параметричному випадку названо задачу, в якій при заданому об'ємі вибірки, відомому закону розподілу і його параметрів, визначених за вибірковими даними, заданому рівні довіри (статистичній надійності) необхідно визначити межі можливих значень випадкової величини, в яких може знаходитися задана частка вибірки. За відсутності відомостей про вид закону розподілу вибірки розглянуто розв'язок задачі в непараметричному випадку.

При виконанні розрахунків чисельних прикладів прийнято наступні умови. Для частки генеральної сукупності прийнято стандартні умови: 0,75; 0,90; 0,95; 0,99; для заданої статистичної надійності прийнято значення: 0,90; 0,95; 0,99. Прийнято, що вибірка містить не менш ніж 30 спостережень. Обмеження на об'єм вибірки обумовлені тим, що, по-перше, при меншому об'ємі вибірки необхідне створення спеціалізованих програмних продуктів, по-друге, обробка даних, отриманих по вибірках меншого об'єму, не завжди має змістовний сенс. При розв'язанні задачі в непараметричному випадку отримано таблицю, яка дозволяє обрати розв'язок поставленої задачі для заданих умов. Показано, що вибірки більше, ніж 300 спостережень не дають істотних змін у розв'язку задачі.

Для всіх перерахованих розподілів визначені, для нижнього і верхнього значень меж толерантних інтервалів, об'єми вибірок, що забезпечують необхідні ймовірнісні характеристики: частку вибірки в генеральній сукупності і її статистичну надійність.

Для нормального розподілу поставлена задача вирішена для варіанту двостороннього толерантного інтервалу.

Ключові слова: толерантні інтервали, рівномірний розподіл, показовий розподіл, розподіл Вейбулла, нормальний розподіл, логарифмічно нормальний розподіл.

Вступ

Постановка проблеми. Можлива форма подання результатів клінічних досліджень аналізу крові показана в табл. 1. Цю форму використовують в практичній діяльності однієї з медичних лабораторій.

У цій таблиці наведено фактичні результати і так звані референтні (референсні) інтервали. Результати фактичних вимірювань порівнюють з допустимими межами (референтними або референсними значеннями) і після чого роблять висновки про стан об'єкту досліджень. У роботі [1] дано таке визначення референсного (референтного) значення: рефе-

рентні значення – результати спостереження множини певного типу величин, отриманих від однієї особи або групи осіб, відповідно до заявленого опису. У роботі [1] основними завданнями, що виникають при статистичній обробці референсних значень, названі наступні: розподіл (стратифікація) референсних значень на відповідні групи; визначення законів розподілу усередині кожної групи, виділення аномальних даних, визначення референсних значень. Можливі методи розв'язання кожного з цих завдань розглянуто в роботі [2]. У 1939 р. вийшла робота [3], в якій була поставлена така задача.

Таблиця 1

Результати клінічних лабораторних досліджень аналізу крові

Показник	Результат	Одиниця виміру	Референтний інтервал
Аналіз крові розгорнутий (ОАК+СОЭ)			
Швидкість осідання еритроцитів (СОЭ, ESR)	5	mm/час	Нормативні значення (метод Вестергрена)
			Діти до 17 років: до 10,0
			Чоловіки (17 – 50 років): до 15,0
			Чоловіки (старше 50 років): до 20,0
			Жінки (17–50 років): до 20,0
			Жінки (старше 50 років): до 30,0

Використовуючи визначення деякої властивості, отриманої за даними вибірових спостережень, визначити межі інтервалу, в який із заданою вірогідністю потрапляє задана частка генеральної сукупності. Такий інтервал отримав назву толерантного інтервалу. Тобто референсний і толерантний інтервали – це синонімічні визначення, з тією лише різницею, що перше визначення використовують в медично-біологічних дослідженнях, друге – в технічних. В деяких випадках толерантні межі називають допустимими, підкреслюючи при цьому рівноправ'я цих визначень [4]. У 1946 р. була опублікована робота [5], в якій було наведено фрагменти таблиць, форма яких залишилася практично незмінною до теперішнього часу. Документ [6] визначає толерантний інтервал, як інтервал, який визначено за випадковою вибіркою так, щоб із заданим рівнем довіри накривати, щонайменше, встановлену частку генеральної сукупності. Рівнем довіри в даному випадку є частка інтервалів, побудованих таким чином, що включають, принаймні, задану частку вибірки при багаток-

ратному повторенні процедури. Кінцеву точку толерантного інтервалу називають толерантною межею. Толерантні інтервали можуть бути односторонніми і двосторонніми. Одностороннім толерантним інтервалом називають інтервал, що має верхню або нижню фіксовану природну статистичну межу. Природна межа випадкової величини – це граничне значення односторонньої межі. Двосторонній толерантний інтервал – це інтервал, що має верхню і нижню межі.

Аналіз літератури. Формальне визначення толерантних інтервалів наведено в роботах [7–8], докладні теоретичні відомості про їх властивості наведено, наприклад, в роботі [9]. В табл. 2, згідно з роботами [7–8; 13], наведено для множини об'єму n вибірових даних $(x_1, x_2, \dots, x_n)$, отриманих з генеральної сукупності X , визначення нижньої L , та верхньої U межі толерантного інтервалу і межі двостороннього інтервалу.

Таблиця 2

Визначення односторонніх і двосторонніх толерантних інтервалів

Тип толерантного інтервалу	Визначення толерантного інтервалу	
	У загальному вигляді	Для заданої щільності розподілу
Односторонній, нижній	$\Pr\{\Pr(L \leq X < +\infty) \geq P_1\} = \gamma$	$\Pr\left\{\int_{-\infty}^L f(x)dx = 1 - P_1\right\} = \gamma$
Односторонній, верхній	$\Pr\{\Pr(-\infty \leq X \leq U) \geq P_2\} = \gamma$	$\Pr\left\{\int_{-\infty}^U f(x)dx = P_2\right\} = \gamma$
Двосторонній	$\Pr\{\Pr(L \leq X \leq U) \geq P\} = \gamma$	$\Pr\left\{\int_L^U f(x)dx = P\right\} = \gamma$
Обмеження	–	$P_1 + P_2 = 1 + P$

Величину P_1 визначають згідно з умовою:

$$\int_{-\infty}^L f(x)dx = P_1. \quad (1)$$

У табл. 2 і далі прийнято, що $\Pr$ – вірогідність події, умови здійснення якої наведено в дужках, P – частка $(100\% \cdot P)$ генеральної сукупності, γ $(0 < \gamma < 1)$ – задана статистична надійність. При роз-

гляді толерантних інтервалів виділяють два випадки. Перший – параметричний, тобто такий, при якому відомо вид функції розподілу, точніше її щільність $f(x, \Theta)$, де Θ – вектор параметрів розподілу або їх оцінки $\hat{\Theta}$, отримані по вибірці об'єму n [9, С. 176]. Другий – непараметричний, тобто такий, при якому відомий тільки варіаційний ряд, побудований за даними вибірки об'єму n [9, С. 695]. У свою чергу величина толерантного інтервалу залежить від толерантного множника $k = \theta(\Theta, n, P, \gamma)$.

Визначення величини k будемо називати прямою задачею побудови толерантних інтервалів. Мета розв'язання прямої задачі – визначення меж толерантних інтервалів. При цьому значення величин P і γ обирає дослідник, вибіркові дані і їх властивості розглядають як вхідні дані, незалежні від дослідника. Для різних видів щільності розподілу її розв'язки наведено в роботі [10]. Тексти програм для розв'язання цієї задачі, написані на мові системи R, наведено в роботах [11–12]. В роботі [13] описано чисельний метод, що дозволяє визначати коефіцієнти толерантності і межі інтервалів толерантності для всіх основних типів розподілів. Застосування цього методу вимагає створення спеціального програмного забезпечення, що істотно обмежує його застосування. Найбільш відпрацьованими слід вважати методи визначення толерантних інтервалів для нормального (ISO 16269-6:2014 Statistical interpretation of data – Part 6: Determination of statistical tolerance intervals, MOD) і показникового розподілів (IEC 60605-4:2001. Equipment reliability testing – Part 4: Point estimate, confident intervals, prediction intervals and tolerance intervals MOD). Останній документ написано на “мові теорії надійності”, що ускладнює його застосування в загальному випадку.

При плануванні експериментів виникає необхідність визначити, при заданій величині толерантного інтервалу, значення деяких параметрів, що його визначають. Розв'язання такої задачі називатимемо оберненою по відношенню до задачі визначення толерантного інтервалу. При цьому вид розподілу і оцінки його параметрів будемо вважати заданими або визначеними за даними вибірки.

Слід зазначити, що в доступній авторам даного повідомлення літературі, будь-яких посилань на можливі варіанти постановки оберненої задачі і способи її розв'язання не знайдено.

Постановка завдання. Формулювання оберненої задачі визначення величини толерантного інтервалу і її розв'язання в непараметричному випадку, а також в параметричному для рівномірного розподілу показового розподілу, розподілу Вейбулла, нормального розподілу, логарифмічно нормального розподілу. Розробка чисельних методів рішення поставлених задач, доступних для найбільш поширених програмних продуктів.

Виклад основного матеріалу

При розв'язанні поставленої задачі, відповідно до роботи [7], прийняті наступні значення змінних: частка генеральної сукупності $P = 0,75; 0,9; 0,95; 0,99$; задана статистична надійність $\gamma = 0,90; 0,95; 0,99$. Об'єм вибірки $n > 30$. Такий вибір цієї величини обумовлений тим, що при меншому об'ємі вибірки виникає необхідність в застосуванні спеціальних обчислювальних засобів, описаних в роботі [13].

Розв'язання оберненої задачі побудови толерантних інтервалів в непараметричному випадку. Пряма задача в цьому випадку формулюється так. Для вибірки, елементи якої $x_1, x_2, \dots, x_n$ утворюють варіаційний ряд, розташований за збільшенням. Необхідно визначити таке мінімальне значення величини n , щоб з надійністю γ між її крайніми елементами містилася частка P генеральної сукупності. У роботі [14] приведено розв'язок прямої задачі, а в роботі [8] наведено її розв'язок у вигляді таблиць.

Згідно роботи [14] розв'язок прямої задачі має вигляд:

$$nP^{n-1} - (n-1)P^n = 1 - \gamma. \quad (2)$$

Вирішуючи умову (2) відносно кожної змінної і розглядаючи γ як параметр, отримаємо чисельні значення її розв'язку. Результати представлено в табл. 3.

Таблиця 3

Визначення статистичних властивостей розв'язків оберненої задачі побудови толерантних інтервалів у непараметричному випадку

Об'єм вибірки, n	Частка вибірки, P	Статистична надійність, γ
30	0,75	0,9980
30	0,9	0,8163
30	0,95	0,4464
30	0,99	0,03614
300	0,75	0,9999
300	0,9	0,9999
300	0,95	0,9999

Закінчення табл. 3

300	0,99	0,8023
3000	0,75	0,9999
3000	0,9	0,9999
3000	0,95	0,9999
300	0,99	0,9999

З наведених результатів виходить, що при даній постановці завдання збільшення об'єму вибірки понад 300 спостережень не робить істотного впливу на статистичну надійність висновку.

Розв'язання оберненої задачі побудови толерантних інтервалів для рівномірного розподілу. Відповідно до роботи [15] щільність рівномірного розподілу представимо у вигляді:

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b. \quad (3)$$

Оцінками $\hat{a}, \hat{b}$ для параметрів a, b , визначеними по методу максимуму правдоподібності, згідно з

роботою [15] будуть найменше x_1 і найбільше x_n значення у вибірці об'єму n .

Нижню L і верхню U межі толерантних інтервалів визначають за умовами, наведеними в роботі [12]:

$$L = x_1 + \frac{(x_n - x_1)(1-P)}{\lambda^{1/n}}; \quad U = \frac{(x_n - x_1)P}{\lambda^{1/n}}. \quad (4)$$

В табл. 4 наведено варіанти розв'язання поставленої задачі.

Таблиця 4

Варіанти розв'язання оберненої задачі побудови верхньої і нижньої меж толерантних інтервалів для рівномірного розподілу

Границі інтервалів	Параметри толерантних інтервалів		
	γ	P	n
Нижнього, L	$\left[\frac{(x_n - x_1)(1-P)}{L - x_1} \right]^n$	$1 - \frac{(L - x_1)\gamma^{1/n}}{x_n - x_1}$	$\left[\frac{\ln \left(\frac{(x_n - x_1)(1-P)}{L - x_1} \right)}{\ln \gamma} \right]^{-1}$
Верхнього, U	$\left[\frac{(x_n - x_1)P}{U - x_1} \right]^n$	$\frac{(U - x_1)\gamma^{1/n}}{x_n - x_1}$	$\left[\frac{\ln \left(\frac{(x_n - x_1)P}{U - x_1} \right)}{\ln \gamma} \right]^{-1}$

Розв'язання оберненої задачі побудови толерантних інтервалів для показникового розподілу. Відповідно до роботи [15] щільність показникового розподілу представимо у вигляді:

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0, \lambda > 0, \bar{x} = \lambda^{-1}. \quad (5)$$

У відповідності з роботою [16] межі односторонніх інтервалів нижнього L і верхнього, U визначають за умовами:

$$L = -2 \sum_{i=1}^n x_i \cdot \frac{\ln \left(\frac{1+P}{2} \right)}{\chi_{1+\gamma; 2n}^2} = -2n\bar{x} \frac{\ln \left(\frac{1+P}{2} \right)}{\chi_{1+\gamma; 2n}^2} = -2 \frac{n}{\lambda} \cdot \frac{\ln \left(\frac{1+P}{2} \right)}{\chi_{1+\gamma; 2n}^2}; \quad (6)$$

$$L = -2 \sum_{i=1}^n x_i \cdot \frac{\ln \left(\frac{1-P}{2} \right)}{\chi_{1-\gamma; 2n}^2} = -2n\bar{x} \frac{\ln \left(\frac{1-P}{2} \right)}{\chi_{1-\gamma; 2n}^2} = -2 \frac{n}{\lambda} \cdot \frac{\ln \left(\frac{1-P}{2} \right)}{\chi_{1-\gamma; 2n}^2}. \quad (7)$$

Наприклад, при $n = 100$, $\sum_{i=1}^n x_i = 3600$ $P=0,9$;

$\gamma=0,95$ отримаємо, що $L=1,120$ і $U = 89,477$. Це означає, що в генеральній сукупності, з якої була отримана вибірка, 90% даних з вірогідністю 0,95 перевершають величину, що дорівнює 1,120 і така ж частка вибірки з вірогідністю 0,95 не перевершить величину, яка дорівнює 89,447. Для вирішення зворотного завдання введемо наступні позначення:

$$E_L = \frac{-L}{2 \sum_{i=1}^n x_i} \chi_{1+\gamma; 2n}^2; \quad (8)$$

$$E_U = \frac{-U}{2 \sum_{i=1}^n x_i} \chi_{1-\gamma; 2n}^2. \quad (9)$$

Прирівнюючи отримані умови чисельникам дробу у виразах (6) і (7) і, розв'язуючи отримані вирази відносно P , отримаємо, що:

$$P_L = 2 \exp E_L - 1; \quad (10)$$

$$P_U = 1 - 2 \exp E_U. \quad (11)$$

З умов (8–11) витікає, що на розв'язання оберненої задачі істотний вплив робить величина χ^2 залежна від об'єму вибірки при фіксованому рівні довірчої вірогідності γ . В процесі підготовки до проведення експерименту цю величину приймають постійною і рівною одному з перелічених раніше значень. Необхідний об'єм вибірки можна встановити або підбором, використовуючи будь яку з процедур обчислення функції χ^2 , або використовувати одну з її апроксимацій. Останнє переважніше, бо при цьому одночасно можна змінювати і суму початкових даних. Досвід авторів показав, що такий підхід особливо зручний при проведенні обчислень в реальному масштабі часу. Розглянемо приклади такої апроксимації, використовуючи відомості, наведені в роботі [15]. Припустимо, що об'єм вибірки дорівнює величині n . Тоді число степенів свободи, відповідно до умов (6) та (7), дорівнює величині $f=2n$. Величина $\gamma=0,9$. У випадку, коли виконується умова: $30 < f < 100$, використовуємо апроксимацію вигляду (12):

$$\tilde{\chi}_{f,1-\gamma}^2 = f \left[1 - \frac{2}{9f} + u_z \sqrt{\frac{2}{9f}} \right]^3; \quad (12)$$

де прийнято, що $z = (1+\gamma)/2$, u_z – квантиль рівня z стандартного нормального розподілу, $\tilde{\chi}_{f,1-\gamma}^2$ – значення величини $\tilde{\chi}_{f,1-\gamma}^2$, яке визначено з використанням апроксимуючого виразу:

$$\tilde{\chi}_{f,1-\gamma}^2 = f + u_z \sqrt{2f} + \frac{2}{3} (u_z^2 - 1) + \frac{1}{9\sqrt{2f}} (u_z^2 - 7u_z). \quad (13)$$

Тоді при тому ж значенні величини γ , як і в першому прикладі, отримаємо, що $\tilde{\chi}_{f,1-\gamma}^2 = 286,5978$ при табличному значенні $\chi_{f,1-\gamma}^2 = 298,6106$. Тобто відносна помилка $\delta = 4\%$. Підставляючи вирази (12) або (13) в умови (8) або (9) і, розв'язуючи отримані рівняння відносно n , дістанемо можливість моделювати межі толерантних інтервалів залежно від передбачуваного об'єму вибірки.

Розв'язання оберненої задачі побудови толерантних інтервалів для розподілу Вейбулла. Згідно з роботою [15] щільність розподілу Вейбулла має вигляд:

$$f(x) = \frac{b}{a} \left(\frac{x}{a} \right)^{b-1} \exp \left(- \left(\frac{x}{a} \right)^b \right). \quad (14)$$

Відповідно до роботи [16] інтервали для розподілу Вейбулла мають вигляд:

$$L = \left[-2 \sum_{i=1}^n x_i^b \cdot \frac{\ln \left(\frac{1+P}{2} \right)}{\chi_{1+\gamma; 2n}^2} \right]^{1/b}; \quad (15)$$

$$U = \left[-2 \sum_{i=1}^n x_i^b \cdot \frac{\ln \left(\frac{1-P}{2} \right)}{\chi_{1-\gamma; 2n}^2} \right]^{1/b}. \quad (16)$$

Зведемо цю задачу до вже розв'язаної для показникового розподілу за допомогою заміни змінних:

$$W_L = L^b = \left[-2 \sum_{i=1}^n z_i \cdot \frac{\ln \left(\frac{1+P}{2} \right)}{\chi_{1+\gamma; 2n}^2} \right]; \quad (17)$$

$$W_U = U^b = \left[-2 \sum_{i=1}^n z_i \cdot \frac{\ln \left(\frac{1-P}{2} \right)}{\chi_{1-\gamma; 2n}^2} \right]. \quad (18)$$

В умовах (17) та (18) прийнято, що $z_i = x_i^b$.

Розв'язання оберненої задачі побудови толерантних інтервалів для нормального розподілу. Розглянемо варіант розв'язання цієї задачі для випадку, коли оцінками параметрів нормального розподілу служать середнє значення випадкової величини $\bar{x}$ і середнє квадратичне відхилення s , визначені за вибіркою об'єму n . У цьому випадку нижня і верхня межі двостороннього толерантного інтервалу визначають, згідно з роботами [7–14], за правилом:

$$L = \bar{x} - ks, \quad U = \bar{x} + ks. \quad (19)$$

Використовуючи величину $s/\bar{x} = v$, відому як коефіцієнт варіації (коефіцієнт однорідності), представимо коефіцієнт толерантності k у вигляді:

$$k = \frac{\bar{x} - L}{v\bar{x}} = \frac{U - \bar{x}}{v\bar{x}}. \quad (20)$$

Позначимо $L/\bar{x} = h$ і $U/\bar{x} = g$ і представимо умову (20) в безрозмірному вигляді:

$$k = \frac{1-h}{v} = \frac{g-1}{v}. \quad (21)$$

Умова (20) дозволяє представити коефіцієнт толерантності як функцію коефіцієнта однорідності,

який в умовах масового виробництва слугує характеристикою технологічної культури. Зміна значення толерантного множника залежно від характеристики

технологічної культури виробництва і відносного значення нижньої і верхньої толерантної межі показано на рис. 1 і рис. 2.

Рис. 1. Зміна значення толерантного множника $k1$ залежно від характеристики технологічної культури виробництва v і відносного значення нижньої толерантної межі h

Рис. 2. Зміна значення толерантного множника $k2$ залежно від характеристики технологічної культури виробництва v і відносного значення нижньої толерантної межі h

Толерантний множник k визначають по таблицях, приведених в роботах [7–8]. У роботі [18] наведено наближену формулу, що дозволяє визначати толерантний множник:

$$k = u_{1+P/2} \left(1 + \frac{u_\gamma}{\sqrt{2n}} + \frac{5u_\gamma^2 + 10}{12n} \right). \quad (22)$$

Величина $u_{(\cdot)}$ дорівнює відповідному квантилю нормально розподіленої випадкової величини.

З приведених результатів видно, що при об'ємі вибірки $n < 30$ для визначення толерантних множників переважним буде використання їх табличних значень.

Таблиця 5

Толерантні множники для визначення двостороннього толерантного інтервалу*)

Об'єм вибірки	Толерантний множник k	
	Табличне значення	Розрахункове значення
20	2,564	2,819
30	2,413	2,505
40	2,334	2,315
50	2,284	2,27
60	2,248	2,238

Закінчення табл. 5

70	2,222	2,214
80	2,202	2,195
90	2,185	2,18
100	2,172	2,167
200	2,102	2,1
300	2,073	2,072
400	2,057	2,056
500	2,019	2,045

^{*)} Розрахунки виконано при $P = 0,90$ і $\gamma = 0,99$.

Розв'язання оберненої задачі можна отримати, розв'язуючи рівняння (22) відносно змінних u або p чисельними методами. Використовуємо як приклад

дані, які наведено в роботі [14, С. 256]. Вплив зміни імовірнісних характеристик толерантних множників на об'єм вибірки наведено в табл. 6.

Таблиця 6

Вплив зміни імовірнісних характеристик толерантних множників на об'єм вибірки

Статистична надійність, γ	Частка вибірки, P	Об'єм вибірки, n
0,99	0,89	1514
0,95	0,94	773
0,90	0,97	482

Результати, наведені в цій таблиці, дозволяють планувати об'єм вибірки, необхідний для досягнення необхідних характеристик толерантних множників.

Розв'язання оберненої задачі побудови толерантних інтервалів для логарифмічно нормального розподілу. Використовуємо для цього розв'язок аналогічної задачі для нормального розподілу. Відповідно до роботи [16] зробимо заміну вибірки $(x_1, x_2, \dots, x_n)$ на вибірку $(\ln x_1, \ln x_2, \dots, \ln x_n)$. Тоді умова (19) буде перетворена в умову:

$$\ln(L) = \bar{x}_{\ln} - k s_{\ln} \quad \ln(U) = \bar{x}_{\ln} + k s_{\ln}; \quad (23)$$

де $\bar{x}_{\ln}$, $s_{\ln}$ – відповідно середнє значення і середнє квадратичне відхилення, визначене для вибірки $(\ln x_1, \ln x_2, \dots, \ln x_n)$.

$$L = \exp(k(\bar{x}_{\ln} - k s_{\ln})); \quad U = \exp(\bar{x}_{\ln} + k s_{\ln}). \quad (24)$$

Відповідно до умови (22) спосіб визначення толерантного множника залишається таким же, як і для попередньої задачі.

Висновки

1. Поставлено і розв'язано для деяких окремих випадків, важливих в практичній діяльності, обернену задачу побудови толерантних інтервалів.

2. Розв'язок отримано для планування експерименту в непараметричному випадку, а також для рівномірного розподілу, показникового розподілу, розподілу Вейбулла, нормального розподілу, логарифмічно нормального розподілу.

3. Запропоновано чисельні методи розв'язання поставлених задач, доступних для найбільш поширених програмних продуктів.

4. Прямою задачею побудови толерантних ін-

тервалів в параметричному випадку названо задачу, в якій при заданому об'ємі вибірки, відомому законі розподілу і його параметрів, визначених за вибірко-вими даними, заданому рівні довіри (статистичній надійності) необхідно визначити межі можливих значень випадкової величини, в яких може знаходитися задана частка вибірки. За відсутності відомостей про вид закону розподілу вибірки розглянуто розв'язок задачі в непараметричному випадку.

5. При виконанні розрахунків чисельних прикладів прийнято наступні умови. Для частки генеральної сукупності прийнято стандартні умови: 0,75; 0,90; 0,95; 0,99; для заданої статистичної надійності прийнято значення: 0,90; 0,95; 0,99.

6. Прийнято, що вибірка містить не менш ніж 30 спостережень. Обмеження на об'єм вибірки обумовлені тим, що, по-перше, при меншому об'ємі вибірки необхідне створення спеціалізованих програмних продуктів, по-друге, обробка даних, отриманих по вибірках меншого об'єму, не завжди має змістовний сенс.

7. При розв'язанні задачі в непараметричному випадку отримано таблицю, яка дозволяє обрати розв'язок поставленої задачі для заданих умов. Показано, що вибірки більше, ніж 300 спостережень не дають істотних змін в розв'язку задачі.

8. Для всіх перерахованих розподілів визначено нижні і верхні межі толерантних інтервалів, об'єми вибірок, що забезпечують необхідні ймовірнісні характеристики: частку вибірки в генеральній сукупності і її статистичну надійність.

9. Для нормального розподілу поставлена задача вирішена для варіанту двостороннього толерантного інтервалу.

Список літератури

1. Кривенко М.П. Модели для представления и обработки референсных значений / М.П. Кривенко // Информатика и её применение. – 2015. – Том 9. – Выпуск 2. – С. 63-74.
2. Burtis C.A. Tietz fundamentals of clinical chemistry and molecular diagnostics / C.A. Burtis, D.E. Bruns. – St. Louis, USA, Elsevier Health Sciences, 2014. – 1104 p.
3. Shewhart W.A. Statistical Method from the Viepoint of Quality Control. ch 2 / W.A. Shewhart. – Washington, 1939. – P. 50-79.
4. ГОСТ 22015–76. Качество продукции. Нормирование и статистическая оценка качества металлических материалов и изделий по механическим характеристикам. – Москва: Издательство стандартов, 1976. – 31 с.
5. Wald A. Tolerance Limits for a Normal Distribution / A. Wald // Ann. Mat. St. – 1946. – Vol. 17. – P. 208-215.
6. ISO 3534-1-2019 Статистические методы. Словарь и условные обозначения. Часть 1. Общие статистические термины и термины, используемые в теории вероятностей. – Москва: Издательство стандартов, 2019. – 22 с.
7. Большев Л.Н. Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. – Москва: Наука, 1983. – 416 с.
8. Мюллер П. Таблицы математической статистики / П. Мюллер, П. Нойман, Р. Шторм. – Москва: Финансы и статистика, 1982. – 278 с.
9. Кендалл М. Статистические выводы и связи / М. Кендалл, А. Стьюарт. – Москва: Наука, 1973. – 899 с.
10. Meeker W.Q. Statistical Intervals. A Guide for Practitioners and Researchers / W.Q. Meeker, G.J. Hahn, L. Escobar. – New Jersey: John Wiley & Sons, 2017. – 605 p.
11. Herve M. Aide-memoire de statistique applique a la biologie. Construire son etude et analyser les resultats a l'aide du logiciel R [Electronic resource] / M. Herve. – Available at: www.cran.r-project.org/doc/contrib/Herve-Aide-memoire-statistique.pdf.
12. Мастицкий С.Э. Статистический анализ и визуализация данных с помощью R [Электронный ресурс] / С.Э. Мастицкий, В.К. Шитиков. – 2014. – Режим доступа: <http://r-analytics.blogspot.com>.
13. Сухорученков Б.И. Анализ малой выборки. Прикладные статистические методы / Б.И. Сухорученков. – Москва: Вузовская книга, 2010. – 384 с.
14. Дунин-Барковский И.В. Теория вероятностей и математическая статистика в технике (общая часть) / И.В. Дунин-Барковский, Н.В. Смирнов. – Москва: Государственное издательство технико-теоретической литературы, 1955. – 556 с.
15. Вадзинский Р.Н. Справочник по вероятностным определениям / Р.Н. Вадзинский. – Санкт-Петербург: Наука, 2001. – 295 с.
16. Могильний Д.В. Толерантні границі для деяких законів розподілення, застосовуваних в теорії надійності / Д.В. Могильний // Збірник наукових праць Київського інституту Військово-Повітряних сил. – 1998. – Вип. 2. – С. 55-58.
17. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников / А.И. Кобзарь. – Москва: ФИЗМАТЛИТ, 2006. – 816 с.

References

1. Krivenko, M.P. (2015), "Modeli dlja predstavlenija i obrabotki referensnyh znachenij" [Models for the representation and processing of reference values], *Informatics and its application*, Vol. 9, Issue 2, pp. 63-74.
2. Burtis, C.A. and Bruns, D.E. (2014), *Tietz fundamentals of clinical chemistry and molecular diagnostics*, Elsevier Health Sciences, St. Louis, USA, 1104 p.
3. Shewhart, W.A. (1939), *Statistical Method from the Viepoint of Quality Control*, ch 2, Washington, pp. 50-79.
4. GOST 22015–76 (1976), "Kachestvo produktcii. Normirovaniye i statisticheskaya otsenka kachestva metallicheskih materialov i izdelij po mekhanicheskim kharakteristikam" [Product quality. Rationing and statistical assessment of the quality of metallic materials and products according to mechanical characteristics], Izdatel'stvo standartov, Moscow, 36 p.
5. Wald, A. (1946), Tolerance Limits for a Normal Distribution, *Ann. Math. St.*, Vol. 17, pp. 208-215.
6. ISO 3534-1-2019 (2019), "Statisticheskiye metody. Slovar' i uslovnyye oznacheniya. Chast' Obshchiye statisticheskiye terminy i terminy, ispol'zuyemyye v teorii veroyatnostey" [Statistical methods. Vocabulary and conventions. Part 1. General statistical terms and terms used in probability theory], Izdatel'stvo standartov, Moscow, 22 p.
7. Bol'shev, L.N. and Smirnov, N.V. (1983), "Tablitsy matematicheskoy statistiki" [Tables of mathematical statistics], Nauka, Moscow, 416 p.
8. Muller, P., Neumann, P. and Storm, R. (1982), "Tablitsy matematicheskoy statistik" [Tables of mathematical statistics], Finance and Statistics, Moscow, 278 p.
9. Kendall, M. and St'juart, A. (1973), "Statisticheskie vyvody i svyazi" [Statistical conclusions and connections], Nauka, Moscow, 899 p.
10. Meeker, W.Q., Hah, G.J. and Escobar, L. (2017), *Statistical Intervals. A Guide for Practitioners and Researchers*, John Wiley & Sons, New Jersey, 605 p.
11. Herve, M. (2014), *Aide-memoire de statistique applique a la biologie. Construire son etude et analyser les resultats a l'aide du logiciel R*, available at: www.cran.r-project.org/doc/contrib/Herve-Aide-memoire-statistique.pdf.
12. Mastitsky, S.E. and Shitikov, V.K. (2014), "Statisticheskij analiz i vizualizaciya dannyh s pomoshh'ju R" [Statistical analysis and data visualization using R], available at: www.r-analytics.blogspot.com (accessed 03 March 2020).

13. Suhoruchenkov, B.I. (2010), "Analiz maloj vyborki. Prikladnye statisticheskie metody" [Analysis of a small sample. Applied statistical methods], Vuzovskaja kniga, Moscow, 384 p.
14. Dunin-Barkovsky, I.V. and Smirnov, N.V. (1955), "Teoriya veroyatnostej i matematicheskaja statistika v tehnike (obshhaja chast')" [Probability theory and mathematical statistics in technology (general part)], Gosudarstvennoe izdatel'stvo tekhniko-teoreticheskoy literatury, Moscow, 556 p.
15. Vadzinskiy, R.N. (2001), "Spravochnik po veroyatnostnym raspredeleniyam" [Handbook of probability distributions], Nauka, Sankt-Peterburg, 295 p.
16. Moghyl'nyj, D.V. (1998), "Tolerantni ghranyci dlja dejaknykh zakoniv rozpodilennja, zastosovujemykh v teorii nadijnosti" [Tolerant boundaries for some distribution laws applicable in reliability theory], *Proceedings of the Kiev Institute of Military and Air Forces*, No. 2, pp. 55-58.
17. Kobzar', A.I. (2006), "Prikladnaya matematicheskaya statistika. Dlya inzhenerov i nauchnykh rabotnikov" [Applied mathematical statistics. For engineers and scientific workers], FIZMATLIT, Moscow, 816 p.

Надійшла до редколегії 17.03.2020

Схвалена до друку 12.05.2020

Відомості про авторів:

Гадецька Світлана Вікторівна

кандидат фізико-математичних наук доцент
доцент Харківського національного
автомобільно-дорожного університету,
Харків, Україна
<https://orcid.org/0000-0002-9125-2363>

Дубницький Валерій Юрійович

кандидат технічних наук старший науковий співробітник
старший науковий співробітник Харківського навчально-
наукового інституту Державного вищого навчального
закладу "Університет банківської справи",
Харків, Україна
<https://orcid.org/0000-0003-1924-4104>

Кушнерук Юрій Іонович

кандидат технічних наук доцент
доцент Харківського національного
університету Повітряних Сил ім. І. Кожедуба,
Харків, Україна
<https://orcid.org/0000-0001-5844-7137>

Ходирєв Олександр Іванович

старший викладач кафедри
Харківського навчально-наукового інституту
Державного вищого навчального закладу
"Університет банківської справи",
Харків, Україна
<https://orcid.org/0000-0001-9871-9440>

Information about the authors:

Svitlana Gadetska

PhD in Physics and Mathematics Associate Professor
Senior Lecturer of Kharkiv National
Automobile and Highway University,
Kharkiv, Ukraine
<https://orcid.org/0000-0002-9125-2363>

Valeriy Dubnitskiy

Candidate of Technical Sciences Senior Research
Senior Research Associate of Kharkiv Educational
Scientific institute SHEI
"University of Banking" Kharkiv,
Kharkiv, Ukraine
<https://orcid.org/0000-0003-1924-4104>

Youriy Kushneruk

Candidate of Technical Sciences Associate Professor
Senior Lecturer of Ivan Kozhedub Kharkiv
National Air Force University,
Kharkiv, Ukraine
<https://orcid.org/0000-0001-5844-7137>

Alexander Khodyrev

Senior Instructor
of Kharkiv Educational
Scientific institute SHEI
"University of Banking",
Kharkiv, Ukraine
<https://orcid.org/0000-0001-9871-9440>

ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА ПРИ РЕШЕНИИ ОБРАТНОЙ ЗАДАЧИ ПОСТРОЕНИЯ ТОЛЕРАНТНЫХ (РЕФЕРЕНСНЫХ) ИНТЕРВАЛОВ

С.В. Гадецкая, В.Ю. Дубницкий, Ю.И. Кушнерук, А.И. Ходырев

Поставлена и решена для некоторых частных случаев, важных в практической деятельности, обратная задача построения толерантных интервалов. Решение получено для планирования эксперимента в непараметрическом случае, а также для равномерного распределения, показательного распределения, распределения Вейбулла, нормального распределения, логарифмически нормального распределения. Предложены численные методы решения поставленных задач, доступные для наиболее распространенных программных продуктов. Прямой задачей построения толерантных интервалов в параметрическом случае названа задача, в которой при заданном объеме выборки, известном законе распределения и его параметрах, определенных по выборочным данным, заданном уровне доверия (статистической надёжности) необходимо определить границы возможных значений случайной величины, в которых может находиться заданная доля выборки. При отсутствии сведений о виде закона распределения выборки рассмотрено решение задачи в непараметрическом случае.

При выполнении расчётов численных примеров приняты следующие условия. Для доли генеральной совокупности приняты стандартные условия: 0,75; 0,9; 0,95; 0,99; для заданной статистической надёжности приняты значения: 0,90; 0,95; 0,99. Принято, что выборка содержит не менее 30 наблюдений. Ограничения на объём выборки обусловлены тем, что, во-первых, при меньшем объёме выборки необходимо создание специализированных программных продуктов, во-вторых, обработка данных, полученных по выборкам меньшего объёма, не всегда имеет содержательный смысл. При решении задачи в непараметрическом случае получена таблица, позволяющая выбрать решение поставленной задачи для заданных условий. Показано, что выборки больше, чем 300 наблюдений не дают существенных изменений в решение задачи.

Для всех перечисленных распределений определены, для нижнего и верхнего значений границ толерантных интервалов, объёмы выборок, обеспечивающие требуемые вероятностные характеристики: долю выборки в генеральной совокупности и её статистическую надёжность.

Для нормального распределения поставленная задача решена для варианта двустороннего толерантного интервала.

Ключевые слова: толерантные интервалы, равномерное распределение, показательное распределение, распределение Вейбулла, нормальное распределение, логарифмически нормальное распределение.

DESIGN OF EXPERIMENT IN INVERSE PROBLEM SOLUTION OF TOLERANCE (REFERENCE) INTERVALS BUILDUP

S. Gadetska, V. Dubnitskiy, Yu. Kushneruk, A. Khodyrev

Inverse problem of tolerance intervals buildup was stated and solved for certain particular cases important in practical activities. The solution was obtained for design of experiment in non-parametric case, as well as for uniform distribution, exponential distribution, Weibull's distribution, normal distribution, lognormal distribution. Numerical methods were proposed for solution of stated problems which are accessible for most widely applied software products. A direct problem of tolerance intervals buildup in parametric case is one where under set sample size, known distribution law and its parameters determined from sample data, preset confidence level (statistical validity) one has to determine like value limits of a random value within which the preset sampling fraction may be located. Under unknown distribution type of sample the problem solution in non-parametric case is discussed.

In calculation of numerical examples the following terms were adopted. For sampling ratio standard terms were adopted: 0,75; 0,9; 0,95; 0,99; for preset statistical validity values were adopted: 0,90; 0,95; 0,99. Sample was assumed to contain at least 30 findings. Sample size limitations were caused by the facts that, first, under a smaller sample size specialized software products would have to be developed, and second, treatment of data obtained from smaller samples does not always bear substantive sense. Under problem solution in non-parametric case we obtained a table enabling selection of stated problem solution for specified terms. It was shown, that a sample larger than 300 findings sense did not bring substantial changes to problem solution.

For all specified distributions we determined sample sizes enabling for top and bottom values of tolerance interval limits necessary probability characteristics, namely, sample proportion in population and its statistical validity.

For normal distribution the stated problem was solved for bilateral tolerance interval option.

Keywords: tolerance intervals, uniform distribution, exponential distribution, Weibull's distribution, normal distribution, lognormal distribution.