П. КОСОБУЦЬКИЙ

Національний університет "Львівська політехніка" Львів Україна
(Вул. С. Бандери, 12, Львів 79012; e-mail: petkosob@gmail.com)

УДК 53.088.3

СТОСОВНО МОДЕЛЮВАННЯ МАТЕМАТИЧНОГО СПОДІВАННЯ І ДИСПЕРСІЇ ВИБІРОК ГАУСОВО РОЗПОДІЛЕНИХ ВИПАДКОВИХ ВЕЛИЧИН

*Проведений аналіз виведення запропонованих у роботах УФЖ **6**, №4, 355–362 (2017) та УФЖ **62**, №2, 184–190 (2017) правил переносу середнього і дисперсії фізичних величин, функціонально пов'язаних перетвореннями $X^2, \cos X, \sqrt{X}, \arccos X$. Показано, що обґрунтування "правил переносу похибок" не ґрунтуються на базових положеннях теорії ймовірностей і математичної статистики, а запропоноване пониження індексів $X \to \sqrt{X}; X^2 \to X$ в коренях квадратних рівнянь, покладеного в основу запису формул переносу, обмежує значення параметрів нормального розподілу m_X, σ_X.*

Ключові слова: нормальний розподіл, математичне сподівання, дисперсія, випадкові величини, правила статистичного усереднення, похибки.

1. Вступ

Нещодавно в УФЖ вийшли роботи [1], присвячені отриманню так званих правил "переносу похибок" для прямих $Y = g(X) = X^2$, $\cos X^2$ і обернених до них $Z = g^{-1}(X) = \sqrt{X}$, $\arccos X$ перетворень випадкової величини (ВВ) X, що підпорядкована нормальному закону розподілу $N(m_X, \sigma_X)$ [1]. Для їх одержання, використані елементарні арифметичні перетворення квадратних рівнянь дисперсій ВВ$X, X^2, \cos X, \arccos X$, в яких явний вигляд функцій середніх обчислювався шляхом застосування табличних інтегралів [2].

Однак в основу обґрунтування аналітичних виразів – правил переносу, не покладені базові положення теорії ймовірностей і математичної статистики [3–5]. Крім того, одержані співвідношення апробувались на експериментальних вибірках, що не задовольняли критерії нормального розподілу. Приймаючи до уваги актуальність цього підходу для статистичного усереднення результатів фізичних досліджень, що завжди несуть в собі флуктуаційну ознаку, дана робота присвячена аналізу обґрунтування формул переносу похибок [1].

2. Теоретичний аналіз

При аналізі експериментальних даних широко використовується статистичний розподіл Гауса $N(m_X, \sigma_X)$, в якому щільність ймовірності одержати значення вимірюваної величини ознаки x залежить від дисперсії $\sigma_X^2 = D_X$ і зміщення $m_X = E_X$. На основі розподілу Гауса розвинута теорія статистичних оцінок похибок [6]. Вона ґрунтується на базових закономірностях кривої щільності ймовірностей стандартного $N(0, \sigma_X)$ розподілу, як симетричність, одномодовість, з максимумом в точці з координатою E_X висотою $\frac{1}{\sqrt{2\pi\sigma_X^2}}$. На висоті ма-

[1] З експериментом пов'язують величину із випадковим числовим значенням ψ. Функція ψ дійсна, повинна задовольняти всі аксіоми ймовірності. Параметри m_Ψ, σ_Ψ – визначають не фізичну величину, а вигляд розподілу щільності ймовірностей її випадкових значень.

ксимуму $\frac{0.607}{\sqrt{2\pi\sigma_X^2}}$, ширина контуру $2\sigma_X$, а значення координат у цих точках дорівнюють $E_X \mp \sigma_X$.

Для кривої Гауса середнє збігається з медіаною і модою, а ймовірність великих відхилень Δx від m_X спадає пропорційно експоненті $\exp\left(-\frac{\Delta x^2}{2\sigma_X^2}\right)$. Це дозволяє здійснювати експрес-оцінку відповідності емпіричного розподілу нормальному, чому часто не приділяється належна увага [1], порушуючи цим самим вимоги стосовно перевірки статистичних моделей на гаусовість.

Мета роботи [1] – запропонувати аналітичні співвідношення, які б дозволили оцінити параметри $E_{\sqrt{X}}$, $D_{\sqrt{X}}$,$E_{\arccos X}$, $D_{\arccos X}$ ВВ, підданих оберненим до прямих перетворень $Y = g(X) = X^2$, $\cos X$ перетворенням $Y = g^{-1}(X) = \sqrt{X}$, $\arccos X$, через задані параметри E_X, D_X із випадковими значеннями нормально $N(m_X, \sigma_X)$ розподіленої ВВ X. Такі одержані ним у вигляді співвідношень:

$$E^4_{\sqrt{X}} = E^2_X - D_X/2, \tag{1a}$$

$$D_{\sqrt{X}} = E_X - \sqrt{E^2_X - D_X/2}. \tag{1b}$$

З метою аналізу статистичної правомірності схеми пониження індексів, застосованої автором [1] при одержанні співвідношень (1), проаналізуємо алгоритм їх обґрунтування в дещо іншому викладі.

В теорії ймовірностей і математичній статистиці, фундаментальним поняттям ВВ є дисперсія, як міра інтенсивності флуктуацій, що дорівнює середньому квадрата без квадрата середнього, для якої строго виконується нерівність

$$D > 0. \tag{2}$$

Якщо вихідна ВВ X перетворена квадратично $Y = X^2$, то за умови статистичної незалежності ВВ, їх коваріація дорівнює нулю і рівняння дисперсії набирають вигляду:

$$D_X = E_{X^2} - E^2_X, \tag{3a}$$
$$D_{X^2} = E_{X^4} - E^2_{X^2}. \tag{3b}$$

Ці рівняння справджуються для довільних ВВ незалежно від закону їх розподілу. Тому за аналогією (3), для перетворень ВВ за законами $\arccos X, \sqrt{X} \Leftarrow X \Rightarrow X^2, \cos X$, можна скласти систему рівнянь:

$$D_{\sqrt{X}} + E^2_{\sqrt{X}} = E_X, \tag{4a}$$
$$D_X + E^2_X = E_{X^2}, \tag{4b}$$
$$D_{X^2} + E^2_{X^2} = E_{X^4}, \tag{4c}$$
$$D_{\arccos\ X} + E^2_{\arccos X} = \overline{(\arccos X)^2}, \tag{4d}$$
$$D_X + E^2_X = \overline{X^2}, \tag{4f}$$
$$D_{\cos\ X} + E^2_{\cos\ X} = \overline{(\cos\ X)^2}. \tag{4e}$$

Рівняння (4) – базові для обґрунтування аналітичних співвідношень оцінок дисперсії і математичного сподівання ВВ, справджуються для незалежних ВВ [2] і не залежать від типу розподілу ймовірностей. В літературі поширена думка, що задача розв'язується тоді, коли інтеграли $E_{\sqrt{X}}$, E_X, E_{X^2}, E_{X^4}, $\overline{(\arccos\ X)^2}$, $E_{\cos X}$, як праві частини (4), виражаються через елементарні функції.

Автор [1] обчислив інтегральні середні $\overline{X^2}$, $\overline{X^4}$, $\overline{\cos X}$, для вихідної ВВ X, розподіленої нормально $N(m_X, \sigma_X)$, для якої випадкові значення параметрів змінюються в межах

$$-\infty < m_X < +\infty, \quad \sigma > 0. \tag{5}$$

Такі обчислення здійснені з використанням табличних інтегралів (3.462.2) та (3.896.2) [2] (аналогічне інтегрування можна зробити частинами [7]). В результаті чого ним одержані вирази:

$$\begin{aligned} &E_X = m,\\ &E_{X^2} = \sigma^2_X + m^2_X,\\ &E_{X^4} = 3\sigma^4_X + 6\sigma^2_X m^2_X + m^4_X =\\ &= E^2_{X^2} + 2\sigma^2_X(E_{X^2} + m^2_X),\\ &E_{\cos X} = \exp\left(-\sigma^2_X/2\right)\cos m_X. \end{aligned} \tag{6}$$

Тоді на підставі (6), система рівнянь (4) перетвориться до вигляду:

$$\begin{aligned} &D_{\sqrt{X}} = E_X - E^2_{\sqrt{X}} \overset{Y=\sqrt{X}}{\Leftrightarrow}\\ &\overset{Y=\sqrt{X}}{\Leftrightarrow} D_X = E_{X^2} - E^2_X \overset{Y=X^2}{\Leftrightarrow}\\ &\overset{Y=X^2}{\Leftrightarrow} D_{X^2} = 2D_X(E_{X^2} + E^2_X),\\ &D_{\arccos X} = E_{\cos X} - E^2_{\arccos X} \overset{Y=\arccos X}{\Leftrightarrow}\\ &\overset{Y=\arccos X}{\Leftrightarrow} D_X = E_{X^2} - E^2_X \overset{Y=\cos X}{\Leftrightarrow}\\ &\overset{Y=\cos X}{\Leftrightarrow} D_{\cos X} = \frac{1}{2}\left[1 - \exp\left(-D_X\right)\right] \times\\ &\times \{1 - \exp\left(-D_X\right)\cos E_X\}. \end{aligned} \tag{7}$$

[2] Для незалежних ВВ коваріація дорівнює нулю.

що дозволяє для прямих $Y = X^2$ і $Y = \cos X$ перетворень ВВ X записати рівняння:

$$D_X = E_{X^2} - E_X^2, \tag{8a}$$

$$D_{X^2} = 2D_X(E_{X^2} + E_{X^2}^2), \tag{8b}$$

$$D_X = E_{X^2} - E_X^2, \tag{8c}$$

$$2D_{\cos X} - 1 = \left(\frac{E_{\cos X}}{E_X}\right)^4 \cos 2E_X - 2E_{\cos X}^2, \tag{8d}$$

розв'язки (8) яких мають вигляд [1]:

$$\begin{aligned} &D_X = E_{X^2} - \sqrt{E_{X^2}^2 - D_{X^2}/2}, \\ &E_X^4 = E_{X^2}^2 - D_{X^2}/2; \end{aligned} \tag{9}$$

для (8a)–(8b), та

$$\begin{aligned} &D_X = \ln\left(\frac{\cos^2 E_X}{E_{\cos X}^2}\right), \\ &E_{\cos X} = E_X \exp(-D_X/2). \end{aligned} \tag{10}$$

для (7c)–(7d).

Оскільки для нормально $N(m_X, \sigma_X)$ розподіленої вихідної ВВ X, задача обчислень середніх $\overline{\sqrt{X}}$ і $\overline{\arccos X}$

$$\overline{\sqrt{X}} = C \int\limits_{-\infty}^{+\infty} \sqrt{x} \exp\left(-\frac{z^2}{2}\right) dx$$

та

$$\overline{\arccos\ x} = C \int\limits_{-\infty}^{+\infty} \arccos\ x \cdot \exp\left(-\frac{z^2}{2}\right) dx, \tag{11}$$

автором [1] не була розв'язана, для знаходження $E_{\sqrt{X}}$,$D_{\sqrt{X}}$ та $E_{\arccos X}$, $D_{\arccos X}$, в рівняннях (7) (або розв'язках (10) та (11)), ним була запропонована схема зниження індексів за алгоритмом:

$$X \to \sqrt{X}; \quad X^2 \to X. \tag{12}$$

Тоді із (9) можна скласти систему рівнянь:

$$D_{\sqrt{X}} = E_X - E_{\sqrt{X}}^2, \tag{13a}$$

$$D_X = 2D_{\sqrt{X}}(E_X + E_{\sqrt{X}}^2), \tag{13b}$$

розв'язками якої є радикали (1). Автор [1] не дав алгоритму (12) статистичного обґрунтування.

Графіки розв'язків (13b) – це параболи з одним максимумом із координатами $E_X^2, 0$ для (1a) та $E_X^2, D_{\sqrt{X}} = E_X$ (1b). Це означає, що на значення величин E_X, D_X накладаються обмеження

$$2E_X^2 \geqslant D_X, \tag{14}$$

і в формулах (1) не можна використовувати нормально $N(m_X, \sigma_X)$ розподілену вихідну ВВ X із математичним сподіванням в інтервалі

$$m_X \leqslant \frac{\sigma_X}{\sqrt{2}}. \tag{15}$$

Аналогічне обмеження існує для дисперсії $D_{\arccos X}$ [1]:

$$\begin{aligned} &\sqrt{2}\sigma \leqslant 1 - E_y \leqslant 1 - E_y^2 \to \\ &\to \sqrt{2}\sigma \leqslant 1 \leqslant 1 + E_y \to \sigma \leqslant 1/\sqrt{2}. \end{aligned} \tag{16}$$

Обмеження (14)–(16) заперечують початковій умові задачі про випадковість вибору в інтервалі (5) параметрів E_X, D_X нормально $N(m_X, \sigma_X)$ розподіленої вихідної ВВ X.

Наведемо деякі ймовірнісно-статистичні аргументи, що свідчать про некоректність застосування схеми пониження індексів. Нелінійні перетворення, в тому числі типу (12), змінюють вигляд функцій розподілу щільності ймовірностей перетворених ВВ [3–5]. Автор [1] провів звичне обчислення середніх типу E_X, D_X, E_{X^2}, D_{X^2}, E_{X^4} для ВВ $X \in N(m_X; \sigma_X^2)$, на підставі яких склав рівняння типу (4), не провівши аналіз статистичних закономірностей перетворених ВВ.

Насправді обґрунтування рівняння дисперсії перетвореної ВВ:

$$\begin{aligned} &D_Y = \int\limits_{-\infty}^{\infty} \left(y - \overline{Y}\right)^2 f_Y(y) dy = \\ &= \int\limits_{-\infty}^{\infty} \left(y^2 + \overline{Y}^2 - 2y\overline{Y}\right) f_Y(y) dy = \\ &= \int\limits_{-\infty}^{\infty} y^2 f_Y(y) dy + \overline{Y}^2 \int\limits_{-\infty}^{\infty} f_Y(y) dy - 2\overline{Y} \int\limits_{-\infty}^{\infty} y f_Y(y) dy = \\ &= \overline{Y^2} + \overline{Y}^2 \int\limits_{-\infty}^{\infty} f_Y(y) dy - \left(2\overline{Y}\right)^2, \end{aligned} \tag{17}$$

з подальшим можливим розв'язком, вимагає аналізу множини існування коренів функції перетворення, встановлення областей її монотонності та, в разі потреби, коректування умови нормування функції $f_Y(y)$ розподілу щільності ймовірностей перетвореної ВВ нормуючим множником у доданку $\overline{Y}^2 \int_{-\infty}^{\infty} f_Y(y)dy$ (17). Наприклад, функція нелінійного перетворення $y = x^2$ двозначна в інтервалі $(-\infty, +\infty)$ і має два корені $X_1 = -\sqrt{Y}$ в області значень $x < 0$ і $X_2 = +\sqrt{Y}$, тоді як типу радикала $y = \sqrt{x}$ має один корінь лише в області значень $x > 0$, тощо.

На завершення відзначимо таке. Співвідношення (1) узгоджуються із відомим в статистиці наближеним методом оцінок перетворених ВВ X, що ґрунтується на теоремі 2.18 [9, с.77]:

$$E_{\sqrt{X}} \cong \sqrt{E_X}, \quad D_Y \cong D_X \left(\frac{dy}{dx}\right)^2 \bigg|_{X=E_X}, \qquad (18)$$

лише в граничному випадку $E_Y \gg D_Y$. Так, якщо $Y = \sqrt{X}$, то $E_{\sqrt{X}} \cong \sqrt{m_X}$, $D_{\sqrt{X}} \cong \frac{\sigma_X^2}{4m_X}$. Аналогічно для інших перетворень: $Y = X^2$ – $E_{X^2} \cong m_X^2$, $D_{X^2} \cong 4E_X^2 D_X = 4m_X^2 \sigma_X^2$ і $Y = \cos X$ – $E_{\cos X} \cong \cos(E_X) = \cos(m_X)$, $D_{\cos X} = \sin^2 E_X \, D_X = \sin^2 m_X \, \sigma_X^2$.

Випадок $E_Y \gg D_Y$ має малу практичну цінність. В процесі обробки ВВ криву розподілу $f_Y(y)$ переважно центрують в окіл вибіркового середнього. Це дозволяє шляхом оптимізації статистично незалежних параметрів розподілу, для нормального двопараметричного E_Y,D_Y, застосовувати розроблені стандартні алгоритми побудови статистичної моделі досліджуваної фізичної системи.

3. Висновки

1. Схема зниження індексів запропонована автором [1] без статистичного обґрунтування. Одержані співвідношення оцінок середнього нормально розподіленої випадкової величини узгоджуються з відомим у літературі наближеним методом лише у граничному випадку $E_Y \gg D_Y$.

2. Сформульована автором [1] задача актуальна, має практичну цінність і оскільки не розв'язана належним чином, то потребує подальшого вирішення.

1. Г.Г. Роде. Перенос похибок та середніх вимірів фізичної величини для елементарних функцій $\cos X$ та $\arccos X$. *УФЖ* **61**, 355 (2016); Перенос похибок та середніх вимірів фізичної величини для елементарних функцій x^2 та $\sqrt{x}$. *УФЖ* **62**, 184 (2017).
2. И.С. Градштейн, И.М. Рыжик. *Таблицы интегралов, сумм, рядов и произведений* (Физматгиз, 1963).
3. J. Mande. *The Statistical Analysis of Experimental Data* (Dover, 1964).
4. E. Suhir. *Applied Probability for Engineers and Scientistics* (McGraw-Hill, 1997)
5. T. Koski. *Lecture Notes. Probabiity and Random Processes at KTN for sf2940 Probability Theory* (KTN Royal Institute of Technology, 2017).
6. М.М. Дороховець. *Опрацювання результатів вимірювань* (Видавництво Національного університету Львівська політехніка, 2007) [ISBN: 978-966-553-640-6].
7. Л.М. Малярець, А.В. Ігначкова, Л.Д. Широкорад. *Теорія ймовірностей і математична статистика у вправах, прикладах та задачах* (Вид. ХНЕУ, 2010).
8. А.П. Прудников, Ю.А. Брычков, О.И. Маричев. *Интегралы и ряды. Элементарные функции* (Наука, 1981).
9. В. Павловський. *Введение в математическую статистику* (Статистика, 1967).

Одержано 26.06.17

P. Kosobutskyy

ON THE SIMULATION
OF THE MATHEMATICAL EXPECTATION
AND VARIANCE OF SAMPLES
FOR GAUSSIAN-DISTRIBUTED RANDOM VARIABLES

S u m m a r y

The derivation of propagation rules for the mean and the variance of physical quantities functionally connected by the transformations X^2, $\cos X$, $\sqrt{X}$, and $\arccos X$, which were proposed in Ukr. J. Phys. **61**, 345 (2016) and Ukr. J. Phys. **62**, 184 (2017), has been analyzed. It is shown that the substantiation of the "error propagation rules" was not based on the fundamentals of probability theory and mathematical statistics. Moreover, the proposed reduction of indices, $X \to \sqrt{X}$ and $X^2 \to X$, in the roots of the square equations forming a basis for the propagation formulas restricts the values of the normal distribution parameters m_X and σ_X.