

П.С. КОСОБУЦЬКИЙ

Національний університет "Львівська політехніка"  
(Львів 79012; e-mail: petkosob@gmail.com)**СТАТИСТИЧНИЙ АНАЛІЗ НОРМАЛЬНО  
РОЗПОДІЛЕНИХ ДАНИХ ІЗ ОБМЕЖЕНИМ  
ІНТЕРВАЛОМ РОЗСІЯННЯ ЗНАЧЕНЬ,  
ПЕРЕТВОРЕНИХ ПРЯМИМИ  $g(x) = x^2; \cos x; a^x$   
ТА ОБЕРНЕНИМИ ДО НИХ ФУНКЦІЯМИ**

УДК 539

*Робота присвячена теоретичному аналізу коректного застосування моделі неперервної нормально розподіленої випадкової величини при обґрунтуванні так званих формул перенесення похибок в задачі статистичного опрацювання експериментальних даних. Звернуто увагу на роль обмеження інтервалу розсіяння значень випадкової величини, підданої нелінійним прямим  $g(X)$  перетворенням елементарними функціями  $X^2; a^X$  та  $\cos X$ , і оберненими до них  $g^{-1}(X) = \sqrt{X}, \arccos X, \log_a X$ . Досліджено закономірності статистичного усереднення даних, одержаних шляхом розкладу функцій перетворення в ряд Тейлора. Для підтвердження правомірності одержаних результатів використано метод оптимізації квадратичного функціонала.*

*Ключові слова:* нормальний розподіл, математичне сподівання, дисперсія, перенесення похибок, перетворення випадкової змінної елементарними функціями.

**1. Вступ**

Відомо, що алгоритм опрацювання експериментальних даних обов'язково включає оцінку статистичної похибки  $\Delta x$  відновленої кривої або перенос похибок [1], як пряму задачу теорії похибок. Суть його полягає у тому, що за заданими похибками аргументу, одержують оцінку дисперсії  $D_X$  і стандартного відхилення (середньо квадратичне відхилення)  $\sigma_X$  функції перетворення  $\varphi(X)$  випадкової змінної (ВЗ)  $X$ . Встановлення дисперсії похибок аргументів за заданою дисперсією похибки функції, відноситься до оберненої задачі. Відомо також, що якщо ВЗ  $X$  перетворюється функцією  $\varphi(X)$ , то справджуються наближені співвідношення [2, 3]

$$m_{\varphi(X)} \cong \varphi(m_X), \quad D_{\varphi(X)} \cong \left( \frac{d\varphi}{dx} \right)^2 \Big|_{x=m_X} D_X^2. \quad (1)$$

© П.С. КОСОБУЦЬКИЙ, 2022

Як впливає із (1), математичне сподівання  $m_X$  стійкіше до перетворень, тому більш інформативною оцінкою  $\Delta x$  перетворення функцією  $\varphi(X)$  є  $\sigma_X = +\sqrt{D_X}$ .

Перенос похибок перетворень особливо актуальний для нормально  $N_X(m_X, \sigma_X)$  розподіленої ВЗ  $X$  із функцією щільності [1]:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right)^2}. \quad (2)$$

Це пов'язано з тим, що розподіл (2) граничний, до якого прямують усі решта розподіли, якщо кількість експериментальних вимірювань за однакових умов зростає. Функція (2) нелінійна, має максимум в точці  $m_X$ , тому для обчислення  $m_X, \sigma_X$  часто застосовують наближений розклад в степеневий ряд Тейлора в інтервалі, межі якого з достатньою точністю визначаються за правилом "трьох сигм":  $m_X \pm 3\sigma_X$ .

Алгоритми обчислення похибок наближеними методами переважно складні в практичній реалізації, тому інтерес становлять аналітичні співвідношення, так звані “формули перенесення похибок (ФПП)”. Саме такі формули були запропоновані автором [5–7] для елементарних функцій прямих  $g(X) = X^2; \cos X; a^X$  і обернених  $g^{-1}(X) = \sqrt{X}; \arccos X; \log_a X$  перетворень нормально розподіленої ВЗ  $X$ . Незважаючи на те, що стосовно алгоритму обґрунтування ФПП були висловлені критичні зауваження [8, 9], поява робіт [5, 4] свідчить про необхідність знову повернутись до дослідження відповідної статистичної задачі, чому присвячена дана робота. Одержані в роботі висновки, піддані перевірці методом одновимірної оптимізації квадратичного функціонала дисперсії  $D_X$ .

## 2. Основні результати роботи

На рис. 1, приведені гістограми розподілів значень експериментальних значень сталої ґратки  $a$ , квадрата її значення  $a^2$  [5], кута елементарної ґратки  $\alpha$ ,  $\cos X$  [6], значень аргументів для функцій  $e^x$  та натурального логарифма [7]. Бачимо із рис. 1, що всупереч твердженням [5–7], вибірки експериментальних даних, що використовувались для апробації ФПП, не підпорядковуються нормальному розподілу. На рис. 2, наведені результати одновимірної оптимізації функціонала  $D_X$  як функції  $m_X$  для прямого  $g(X) = X^2$  і оберненого  $g^{-1}(X) = \sqrt{X}$  перетворень

$$D(x, \alpha) = \int_{-\infty}^{+\infty} (x - \alpha)^2 f(x) dx. \quad (3)$$

Результати, що представлені на рис. 1 і рис. 2, систематизовані у вигляді табл. 1 і табл. 2. Із них можна зробити висновок про те, що якщо обчислені методом оптимізації і обчислені за ФПП [5–7], значення  $m_Y$  більше-менше узгоджуються між собою, що не протирічить (1), то для середньо квадратичного відхилення спостерігаються розбіжності.

## 3. Дискусія

В [5–7] для апробації ФПП не використовувались експериментальні дані із нормальним розподілом ймовірностей, що не дозволило при складанні ФПП виявити дискусійність формальної заміни індексів в розв’язках біквадратного рівняння.

Тому обґрунтуємо перші два моменти перетворення нормально розподіленої ВЗ  $X$  функціями  $X^2$ ,  $\cos X$ ,  $\exp X$  та оберненими до них.

ФПП [5–7], прямі та обернені перетворення застосовувались до одної і тої ж RV  $X$  з параметрами  $m_X, \sigma_X$ . Так, в роботі [5] для обчислення середньо квадратичного відхилення ВЗ  $X$  із нормальним розподілом, перетворених функцією  $Y$  квадратного радикала, використане квадратне рівняння для дисперсії  $\sigma_Y^2 = \overline{Y^2} - \bar{Y}^2$ . В цьому випадку, із середнього [10]:

$$\overline{X^4} = 3\sigma_X^4 + 6m_X^2\sigma_X^2 + m_X^4, \quad (4)$$

в інтервалі значень  $X \in (-\infty < X < +\infty)$  аргументу  $X$  маємо два рівноправні розв’язки:

$$\sigma_X^2 = m_{X^2} \pm \sqrt{m_{X^2}^2 - \frac{1}{2}\sigma_{X^2}^2}. \quad (5)$$

Оскільки за означенням  $\sigma_X^2 = m_{X^2} - m_X^2$ , то в (5) береться розв’язок із знаком мінус і до нього застосовується схема заміни індексів

$$X^2 \rightarrow X, \quad X \rightarrow \sqrt{X} \quad (6)$$

в результаті чого отримуються ФПП у вигляді співвідношень (5) в [5]).

Але за умовою задачі, множиною значень нормально розподіленої ВЗ  $X$  є необмежений інтервал  $X \in (-\infty, +\infty)$  і саме для нього обчислені статистичні середні і дисперсія, тоді як множина значень аргументу функції  $g^{-1}(X) = \sqrt{X}$  обмежена напівобмеженим інтервалом  $[0 \leq X < +\infty)$ . Тому, алгоритм перетворення (6) некоректний. Дійсно, інтегрування функції  $\sqrt{X}$  в області значень

$$[0 \leq X, m_X, \sigma_X < +\infty) \quad (7)$$

супроводжується появою спеціальної неаналітичної (табульованої) функції похибок  $\text{erf}(\xi) = \frac{2}{\sqrt{\pi}} \times \int_0^\xi e^{-\xi^2} d\xi$  [11, 12]. Тоді, внаслідок обмеження інтервалу значень ВЗ  $X \in (0, x)$ , стала нормування  $C_X$  є рівною  $C_X = \frac{2}{(1 + \text{erf}(\frac{m_X}{\sqrt{2}\sigma_X}))}$  і статистичні середні будуть дорівнювати:

$$\begin{aligned} \bar{X} &= m_X + \sigma_X \Lambda, \quad \overline{X^2} = m_X^2 + \sigma_X^2 + \sigma_X m_X \Lambda, \\ \Lambda &= \frac{\sqrt{2/\pi}}{1 + \text{erf}(\frac{m_X}{\sqrt{2}\sigma_X})} e^{-m_X^2/2\sigma_X^2}. \end{aligned} \quad (8)$$

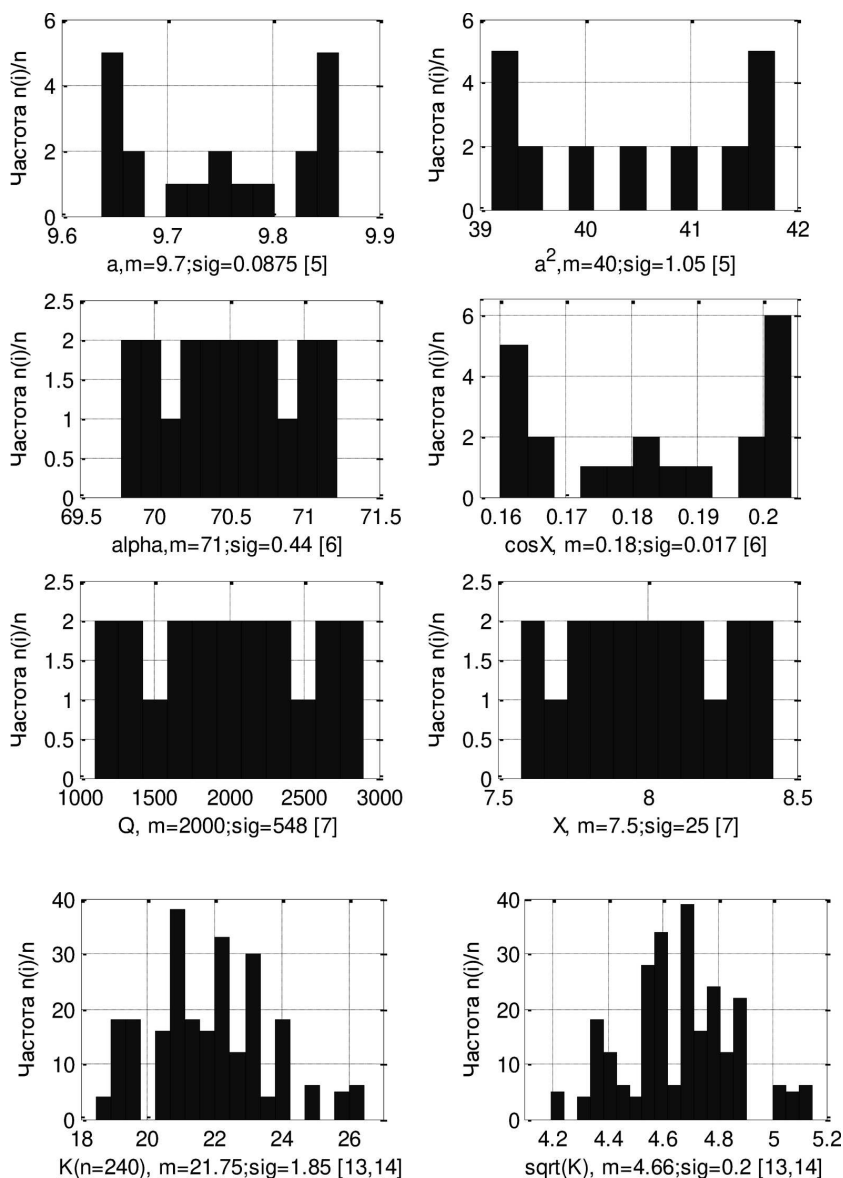


Рис. 1. Гістограми розподілів значень експериментальних даних  $a, a^2$  сталой ґратки [5], кута елементарної ґратки  $\alpha, \cos X$  [6], та значень аргументів для функцій  $e^x$  і натурального логарифма [7]

Таблиця 1. Обчислені оптимізацією дисперсії та за ФПП [5] значення  $m_Y, \sigma_Y$  функцій  $Y$  перетворень ВВ  $X$

$Y$	$X$	$X^2$	$X^2$ [5]	$\sqrt{X}$	$\sqrt{X}$ [5]
$m_Y$	21	445	446	4,65	4,58
$\sigma_Y$	2,24	89,5	94,2	0,285	0,06

Таким чином, дисперсія ВЗ  $X \in (0, x)$  при даному перетворенні буде дорівнювати:

$$D_X = \sigma_X^2 + (m_X - \bar{X})^2 + \sigma_X(2(m_X - \bar{X}) - m_X)\Delta. \tag{9}$$

Отже, в обмеженому інтервалі (7), статистичне середнє  $\bar{X}^4 \neq 3\sigma_X^4 + 6m_X^2\sigma_X^2 + m_X^4$  і задача (4)–(5) з

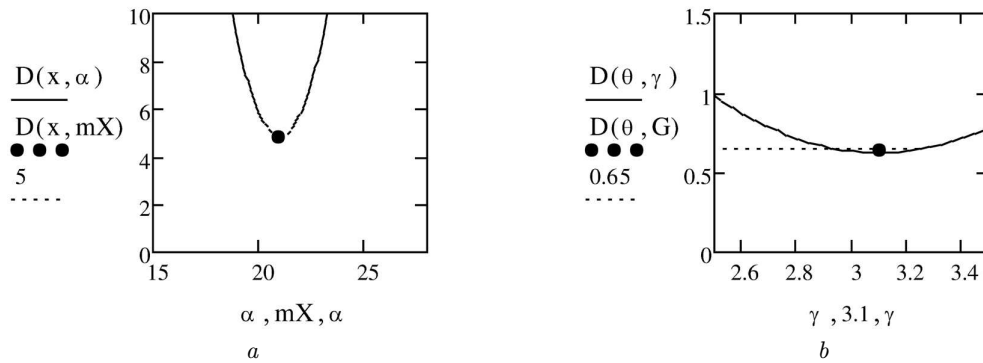


Рис. 2. Ілюстрація визначення середнього і дисперсії функцій перетворень  $g(X) = X$  (a) та  $g^{-1}(X) = \sqrt{X}$  (b) методом оптимізації функціонала (3)

Таблиця 2. Визначені із рис. 2 та обчислені із ФПП [5–7] значення  $\frac{\sigma_Y}{m_Y}$

$Y$	$X$	$X^2$	$\sqrt{X}$	$\cos(X \frac{\pi}{180})$	$\alpha \cos(X \frac{\pi}{180})$	$e^{X/5}$	$\ln(X/5)$
$m_Y$	22	474	4,65 4,68 (7) [5]	0,93	1,18 1,183	83	1,466 1,535 (13) [7]
$\sigma_Y$	1,9	86,5	0,2 0,245 (7) [5]	0,013	0,042 0,001 (11) [6]	33,3	0,0868 0,0037 (13) [7]
$V = \frac{\sigma_Y}{m_Y}$	0,086	0,18	0,043 0,05	0,014	0,036 0,001	0,4	0,06 0,0024

подальшою заміною індексів (6), точного аналітичного розв'язку не має.

В принципі, параметри  $m_{\sqrt{X}}, \sigma_{\sqrt{X}}$  можна обчислити, використовуючи табличні інтеграли (2.3.15(6)) [13, 14]. Але їх застосування вимагає інші спеціальні циліндричні функції. Тому наведемо ще статистичне обґрунтування формули для обчислення середньоквадратичного відхилення для перетворення  $\sqrt{X}$ , для чого прийемо до уваги рівність  $(\sqrt{X})^2 = X$ :

$$D_{\sqrt{X}} = (\overline{(\sqrt{X})^2}) - (\overline{\sqrt{X}})^2 = m_X + \Lambda \sigma_X - (\overline{\sqrt{X}})^2. \quad (10)$$

Згідно з розкладом (3), для нормального розподілу справджується наближення  $\overline{\sqrt{X}} \cong \sqrt{m_X} - \frac{1}{4} \frac{\sigma_X^2}{m_X^{3/2}}$ , тому

$$\sigma_{\sqrt{X}} = m_X + \Lambda \sigma_X - \left( \sqrt{m_X} - \frac{1}{4} \frac{\sigma_X^2}{m_X^{3/2}} \right)^2. \quad (11)$$

Для параметрів ВЗ  $X$   $m_X \cong 21$  та  $\sigma_X \cong 1,85$ , середньоквадратичне відхилення  $\sigma_{\sqrt{X}} \cong 0,77$ , тоді як визначене із гістограми на рис. 1,  $\sigma_{\sqrt{X}} \cong 0,2$ .

Значення  $\sigma_{\sqrt{X}} \cong 0,2$  граничне, до якого прямує середньоквадратичне відхилення  $\sigma_{\sqrt{X}}$  (6) при  $m_X \rightarrow \infty$ , оскільки за законом великих чисел, середнє арифметичне сходиться за ймовірністю до свого математичного сподівання.

Обмеження на множину значень аргументу стосуються також функції натурального логарифма  $\ln X$  та  $a \cos X$ -перетворення. Крім цього, у випадку степеневого перетворення  $e^X$ , при обчисленні середніх  $\overline{\exp X}$  і  $(\overline{\exp X})^2$ , в [7] у формулі (13) на с. 737, була допущена некоректність:

$$1 + \frac{\sigma_X^2}{m_X^2} m_{\ln X} = \frac{1}{2} \ln \left( \frac{m_X^4}{\sigma_X^2 + m_X^2} \right). \quad (12)$$

Суть її полягає в тому, що середнє  $\overline{\exp X}$  для нормально розподіленої ВЗ  $X$  не дорівнює інтегралу

$$\overline{\exp X} \neq \frac{1}{\sqrt{2\pi}\sigma_X} \int_0^{+\infty} e^x e^{-\frac{(x-m_X)^2}{2\sigma_X^2}} dx. \quad (13)$$

В представленні (13), підінтегральний вираз складається добутку двох степеневих функцій, в одній

з яких степінь не виражена у відносних одиницях, як це має місце для експоненти Гауса. Тому для забезпечення коректності обчислень, нормально розподілену ВЗ  $X$ , треба виразити у відносних одиницях:

$$X \rightarrow \frac{X}{\sigma_X}, \quad (14)$$

де прийнято до уваги, що  $\sigma_X \neq 0$ . Тоді статистичні середні  $\overline{\exp X}$  і  $(\overline{\exp X})^2$  в інтервалі  $[0 \leq X < +\infty)$  будуть дорівнювати:

$$\begin{aligned} \overline{\exp\left(\frac{X}{\sigma_X}\right)} &= \frac{C_X}{\sqrt{2\pi}\sigma_X} \int_0^{+\infty} e^{\frac{x}{\sigma_X}} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}} dx = \\ &= \frac{C_X}{2} \left(1 + \operatorname{erf}\left(\frac{m_X + \sigma_X}{\sqrt{2}\sigma_X}\right)\right) \exp\left(\frac{m_X}{\sigma_X} + \frac{1}{2}\right), \end{aligned} \quad (15)$$

$$\begin{aligned} \left[\overline{\exp\left(\frac{X}{\sigma_X}\right)}\right]^2 &= \frac{C_X}{\sqrt{2\pi}\sigma_X} \int_0^{+\infty} e^{\frac{2x}{\sigma_X}} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}} dx = \\ &= \frac{C_X}{2} \left(1 + \operatorname{erf}\left(\frac{m_X + 2\sigma_X}{\sqrt{2}\sigma_X}\right)\right) \exp\left[2\left(\frac{m_X}{\sigma_X} + 1\right)\right]. \end{aligned} \quad (16)$$

Крім цього зауважимо, що одночасне застосування функцій  $\cos X$  і  $a \cos X$  перетворень ВЗ  $X$ , також вимагає представлення аргументу  $X$  у відносних одиницях, наприклад, у вигляді  $X \rightarrow \frac{X}{\max(X)}$  враховуючи, що область зміни аргументу функції  $[-1, +1]$ . У відповідності до стандарту для нормально розподілених ВЗ  $X$ , в даному випадку можна обмежитись наближенням  $\max(X) \cong m_X + 3\sigma_X$ .

Аналогічний підхід справджується у випадку перетворення тригонометричними функціями нормально розподіленої випадкової змінної. Однак, при цьому треба врахувати, що функція  $Y = \cos X$  при  $0 \leq X \leq \pi$  взаємно обернена функції  $Y = a \cos X$  в тому розумінні, в якому взаємно обернені степенева і логарифмічна при  $X > 0$  та функції піднесення в квадрат і взяття квадратного кореня при  $X \geq 0$ . Тоді в інтервалі  $[0; \pi]$ , перетворення  $a \cos$  за числовим значенням дорівнює куту, косинус якого дорівнює  $X$ , то із визначення функцій  $\cos X$  і  $a \cos X$  випливає, що  $\cos(a \cos X) = X$  при  $X \leq 1$  та  $a \cos(\cos X) = X$  при  $0 \leq X \leq \pi$ .

#### 4. Висновки

Дана робота є продовженням раніше виконаних автором досліджень [8, 9], в якій додатково загострена увага на проблемах, якими супроводжується некоректне застосування моделей теорії ймовірностей та математичної статистики із обмеженим інтервалом розсіяння значень для статистичного аналізу експериментальних даних в фізичних системах з флуктуаціями. Інші проблеми, які виникають при цьому, детально описані в роботі [15].

1. Д. Худсон. *Статистика для фізиків* (Мир, 1970).
2. З. Павловський. *Введение в математическую статистику* (Статистика, 1967).
3. И.М. Кодолов, С.Т. Худяков. *Теоретические основы вероятностных методов в инженерно-экономических задачах. Функциональные преобразования случайных величин и случайные функции* (МАДИ, 1985).
4. В.І. Романенко, Н.В. Корніловська. Про точність розрахунку перенесення похибок за аналітичними формулами для оберненого перетворення. *УФЖ* **64**, 215 (2019).
5. Г.Г. Роде. Перенос похибок та середніх вимірів фізичної величини для елементарних функцій  $x^2$  та  $\sqrt{x}$ . *УФЖ* **62**, 184 (2017).
6. Г.Г. Роде. Перенос похибок та середніх вимірів фізичної величини для елементарних функцій  $\cos(x)$  та  $\arccos(x)$ . *УФЖ* **61**, 355 (2016).
7. Г.Г. Роде. Перенос похибок та середніх вимірів фізичної величини для елементарних функцій  $a^x$  та  $\log_a x$ . *УФЖ* **65**, 369 (2019).
8. П.С. Кособуцький. Стосовно моделювання математичного сподівання і дисперсії вибірок гаусово розподілених величин. *УФЖ* **62**, 823 (2017).
9. П.С. Кособуцький. Аналітичні співвідношення обчислення математичного сподівання і середньої квадратичної похибки стандартно розподіленої випадкової величини, підданої перетворенню  $\sqrt{X}$ . *УФЖ* **63**, 215 (2018).
10. І.І. Горбань. *Теорія ймовірностей і математична статистика для наукових працівників і інженерів* (НАН України, ІППММіС, 2003).
11. E. Ng, M. Geller. A table of integrals of the error functions. *J. Research of the National Bureau of Standards-B. Math. Sci. B* **73** (1), 1 (1969).
12. From Web Resource: Table of Integrals. 2014 From <http://integral-table.com>
13. А.П. Прудников, Ю.А. Брычков, О.И. Маричев. *Интегралы и ряды. Элементарные функции* (Наука, 1981).
14. From Web Resource: <http://mathworld.wolfram.com/GaussianIntegral.html>
15. T. Lang. Twenty errors even you can find in biomedical research articles. *Croatian Medical J.* **45** (4), 361 (2004).

Одержано 28.10.19

*P. Kosobutsky*

STATISTICAL ANALYSIS  
OF NORMALLY DISTRIBUTED DATA  
WITH A LIMITED SCATTERING INTERVAL  
OF VALUES CONVERTED BY DIRECT  
 $g(x) = x^2$ ;  $\cos x$ ;  $a^x$  AND INVERSE FUNCTIONS

The work is devoted to the theoretical analysis of the correct application of the model of a continuous normally distributed random variable in the substantiation of the so-called error transfer formulas in the problem of a statistical processing of experimental data. Attention is paid to the role of limiting the

scattering interval of values of a random variable subjected to nonlinear direct  $g(X)$  transformation by elementary functions  $X^2$ ;  $a^X$  and  $\cos X$ , as well as the inverse  $g^{-1}(X) = \sqrt{X}$ ,  $\arccos X$ ,  $\log_a X$  to them. The regularities of the statistical averaging of the data obtained by the disorder of the Taylor transform functions are studied. To confirm the validity of the obtained results, the method of quadratic functional optimization is used.

*Keywords:* normal distribution, mathematical expectation, variance, random variables, calculation and transfer of errors, transformation of a random variable with elementary functions.