MATHEMATICAL MODELING AND COMPUTING, Vol. 10, No. 4, pp. 1014–1022 (2023)

# Simulation of statistical mean and variance of normally distributed data $N_X(m_X, \sigma_X)$ transformed by nonlinear functions $g(X) = \cos X$, $e^X$ and their inverse functions $g^{-1}(X) = \arccos X$, $\ln X$

Kosoboutskyy P. S., Karkulovska M. S.

*Lviv Polytechnic National University,*
*12 S. Bandera Str., 79013, Lviv, Ukraine*

(Received 14 January 2023; Revised 15 November 2023; Accepted 16 November 2023)

This paper presents analytical relationships for calculating statistical mean and variances of functions $g(X) = \cos X$, $e^X$, $g^{-1}(X) = \arccos X$, $\ln X$ of transformation of a normally $N_X(m_X, \sigma_X)$ distributed random variable.

**Keywords:** *statistical mean; variance; transformation; normal distribution; random variable.*

**2010 MSC:** 62-07, 62xx **DOI:** 10.23939/mmc2023.04.1014

## 1. Introduction

Regardless of the principle of averaging the physical parameters $a$ of the statistical system, the problem of deviation of the values of the means of physical quantities $\Delta\overline{a}$ from their mean values $\overline{a}$, i.e., fluctuations, remains relevant. However, a fluctuation measure can be equal to zero $\Delta\overline{a} = \overline{a - \overline{a}} = 0$, since deviations of the given value towards larger and smaller values away from the mean are equally likely. Therefore, the mean square of the difference $\overline{(\Delta a)^2} = \overline{(a-\overline{a})^2} = \overline{(a)^2} - (\overline{a})^2$, i.e., quadratic fluctuation (dispersion) of a random variable, is considered to be a measure of fluctuations. Then, if in a physical process the square $a^2$ is viewed as its intensity, as in the case of a wave, then $\overline{a} = 0$ and the average energy will be calculated as $\overline{W} = k\int_V \overline{a^2}dV$, we will have the mean value $\overline{W} \neq 0$ and the measure of its fluctuation $\overline{\Delta a^2}$.

The above mentioned proves how important an accurate statistical description of systems is. The mean value of the energy of the system of $N$ particles is equal to $\overline{W_N} = N\overline{W_1}$. Then, given $N \to \infty$ the relative fluctuation of the additive quantity $\varepsilon_W = \frac{1}{\sqrt{N}}\sqrt{\overline{W^2}/\overline{W_1}^2 - 1}$ goes to zero, which is proved by the experimental fact that the equilibrium state of the system is the most stable. Therefore, an error analysis must be necessarily performed for statistical algorithms of experimental data processing, both obtained by direct measurements and in the process of further mathematical transformations. Given an arbitrary transformation function, direct and inverse problems of errors are distinguished. In terms of the direct problem, an estimate of the variance and standard deviation of the transformation function is obtained based on the specified errors of the argument. The calculation of the variance of the errors of the arguments based on the specified variance of the error of the transformation function refers to the inverse problem.

Corresponding mathematical transformations can be quite complex; therefore, the error transfer method is often used [1–5]. It is especially important for normally $N_X(m_X, \sigma_X)$ distributed random variables (RV) $X$. A normal distribution is the most common. It is the limiting law that other distribution laws tend towards. It is obtained by summing a considerably large number of independent (or weakly dependent) RVs and the larger the number of the summed random variables, the more precise it is. A normal distribution is typical of all RVs, whose deviations from the mean values are caused by a large set of random factors, each of which is individually insignificant [1, 6–8].

A normal distribution is stable and capable of self-reproduction, therefore it is successfully used in physical modelling, for example, laser cooling of atoms [9], phenomena of quantum mechanics [10];

 

powerful computing resources for computer simulation of the dynamics of atomic-molecular compounds have been created [11]; a specific area of statistical optics has been developed [12], etc. Thus, it is subject to comprehensive research [13]. In the continuous RV model, a normal distribution is an infinitely divisible distribution with finite variance and probability density

$$f_X(x) = \frac{C_X}{\sqrt{2\pi}\sigma_X} \exp\left\{-\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right)^2\right\}, \quad -\infty < X < +\infty, \quad C_X = 1. \tag{1}$$

From the perspective of physical modelling, the range of dispersion of experimental values of BB is limited, therefore, a truncated normal distribution is considered. Thus, this approach is used to model the reliability of physical and technical systems [14], physical processes of charge transfer in electronic devices [15,16]. The corresponding theoretical model was built a long time ago, in the works of Einstein and Smolukhovsky [17,18]. In fact, they proposed one of the first algorithms for mathematical data processing for the purpose of estimating the distribution function and probability density of empirical dependencies based on a sample of experimental data. The connection between Brownian motion and the Gaussian distribution is evidenced by the well-known Fokker–Planck equation (or Kramers equation) in physics [19,20].

## 2. Results and discussion

This study is a continuation of the research [21]. In this paper, the problem of constructing statistical models of direct transformations of RV by functions $g(X) = e^X$, $\cos X$ RV $X$, of $N(m_X, \sigma_X)$ type formulated to investigate the possibility of error transfer when transformations are performed by inverse functions $g^{-1}(X) = \ln X$, $\arccos X$. To do this, we consider case of transformations RV into two

$$X \to \left\{ \begin{array}{c} g(X) = Q \to g^{-1}(Q), \\ g^{-1}(X) = W \to g(W) \end{array} \right\} \to X, \tag{2}$$

in each of which it is possible to consider restrictions of set of allowed values of argument $X$ depending on the type of variable conversion $X$ by functions (1). According to the authors, the choice of declared functions of RV transforms is relevant primarily for the application in physical problems of the decomposition of trigonometric functions in a series with a basis of exponential functions

### 2.1. Case 1

$$X \to \exp X \to \ln \exp X \to X \tag{3}$$

In case of conversion

$$Y = \exp X \tag{4}$$

a nonnegative RV has a logarithmically normal distribution if its natural logarithm

$$\ln Y = X$$

is distributed normally. Then the law of distribution of the RV $Y$ (4) is logarithmically normal:

$$f(y) = \sqrt{\frac{p}{\pi}}\frac{1}{y} \exp\left\{-p(\ln y - m_X)^2\right\}.$$

Applying functions $Y = \exp(X)$ and the inverted to them for converting to dimension units is incorrect. Therefore instead of $X$, we introduce a new variable in relative units:

$$X = \sigma_X \Theta.$$

Because $\sigma_X > 0$, the set of allowed values $\Theta$ remains the same as for $X \in (-\infty, +\infty)$ and became the normalization constant $C_\Theta = C_X = 1$ and the statistical average $g(\Theta) = \exp\Theta$ will be equal to:

$$\overline{\exp\Theta} = \frac{C_\Theta e^{-\frac{m_X^2}{2\sigma_X^2}}}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^{+\infty} e^{\theta} e^{-\frac{\theta^2}{2} + \frac{m_X}{\sigma_X}\theta} dx = \frac{e^{-\frac{m_X^2}{2\sigma_X^2}}}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-r\theta^2 + (1+s)\theta} d\theta = \exp\left(\frac{1}{2} - \frac{m_X}{\sigma_X}\right), \quad \begin{cases} r = \frac{1}{2}, \\ s = \frac{m_X}{\sigma_X}. \end{cases}$$

Using the table integral (2.3.15) in [13], we calculate $\overline{(\exp\Theta)^2}$:

$$\overline{(\exp\Theta)^2} = \frac{C_\Theta\, e^{-\frac{m_X^2}{2\sigma_X^2}}}{\sqrt{2\pi}\sigma_X}\int_{-\infty}^{+\infty} e^{2\theta}e^{-r\theta^2+s\theta}dx = \frac{e^{-\frac{m_X^2}{2\sigma_X^2}}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-r\theta^2+(2+s)\theta}d\theta = \exp\left(1-\frac{m_X}{\sigma_X}\right).$$

Thus, the variance $D_{\exp\Theta}$ of statistically independent RVs $\Theta$, $\cos\Theta$, $(\cos\Theta)^2$ will be equal to:

$$D_{\exp\Theta} = \overline{(\exp\Theta)^2} - \left(\overline{\exp\Theta}\right)^2 = e^{\left(1-\frac{m_X}{\sigma_X}\right)} - e^{\left(1-2\frac{m_X}{\sigma_X}\right)} = e^{\left(1-\frac{m_X}{\sigma_X}\right)}\left(1-e^{-\frac{m_X}{\sigma_X}}\right).$$

Calculating the statistical average $\overline{g^{-1}(\Theta)} = \overline{\ln\exp\Theta}$:

$$\overline{\ln\exp\Theta} = \frac{C_\Theta e^{-\frac{m_X^2}{2\sigma_X^2}}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\ln e^{\theta}e^{-r\theta^2+s\theta}d\theta = \frac{e^{-\frac{m_X^2}{2\sigma_X^2}}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\theta\, e^{-r\theta^2+s\theta}d\theta = \frac{m_X}{\sigma_X} = \overline{\Theta}.$$

To calculate the conversion variance $g^{-1}(\Theta) = \ln(\exp\Theta)$:

$$D_{\ln\exp\Theta} = \frac{C_\Theta}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left(\ln\exp\theta - \overline{\ln\exp\Theta}\right)^2 e^{-r\theta^2+s\theta}d\theta = \overline{(\ln\exp\Theta)^2} - \left(\overline{\ln\exp\Theta}\right)^2$$

we find the statistical mean $\overline{(\ln\exp\Theta)^2}$:

$$\overline{(\ln\exp\Theta)^2} = \frac{C_\Theta}{\sqrt{2\pi}}e^{-\frac{m_X^2}{2\sigma_X^2}}\int_{-\infty}^{+\infty}(\ln e^{\theta})^2e^{-r\theta^2+s\theta}d\theta = \frac{e^{-\frac{m_X^2}{2\sigma_X^2}}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\theta^2e^{-r\theta^2+s\theta}d\theta =$$
$$= e^{-\frac{m_X^2}{2\sigma_X^2}}\frac{\partial}{\partial s}\left[\frac{s}{2r}\exp\left(\frac{s^2}{4r}\right)\right] = \left[\frac{1}{2r}+\left(\frac{q}{2r}\right)^2\right] = 1+\left(\frac{m_X}{\sigma_X}\right)^2.$$

Then

$$D_{\ln\exp\Theta} = \frac{C_\Theta}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left(\ln\exp\theta - \overline{\ln\exp\Theta}\right)^2 e^{-\frac{(\sigma_X\theta-m_X)^2}{2\sigma_X^2}}d\theta = \overline{\Theta^2} - \left(\overline{\Theta}\right)^2 = 1+\frac{m_X^2}{\sigma_X^2}-\frac{m_X^2}{\sigma_X^2} = 1.$$

Thus, the statistical averages and the variance of the transformation functions according to the algorithm (3) are related to the parameters of the original RV $m_X$, $\sigma_X$ system of equations:

$$\begin{cases}\overline{\exp\Theta} = e^{\frac{1}{2}-\frac{m_X}{\sigma_X}},\\ D_{\exp\Theta} = e^{1-\frac{m_X}{\sigma_X}}\left(1-e^{-\frac{m_X}{\sigma_X}}\right),\end{cases} \quad\text{and}\quad \begin{cases}\overline{\ln\exp\Theta} = \overline{\Theta},\\ D_{\ln\exp\Theta} = 1.\end{cases}$$

## 2.2. Case 2

$$X \to \ln X \to \exp\ln X \to X. \tag{5}$$

Unlike the transformation model (3), the model (5) at the beginning of the variable conversion scenario $X$, the natural logarithm function limits its set of allowed values to half a limited interval $[0,+\infty)$. Then according to (2.3.15) in [13], for $n=0$ the normalization constant in this interval $C_{\Theta_{[0,+\infty)}} = C_{X,[0,+\infty)}$, so the statistical mean of the logarithmic transformation of a variable $\Theta_{[0,+\infty)}$ is equal to:

$$\overline{\ln\Theta_{[0,+\infty)}} = \sqrt{\frac{2}{\pi}}K(m_X,\sigma_X)\int_0^{+\infty}\ln\theta\, e^{-\frac{\theta^2}{2}+\frac{m_X}{\sigma_X}\theta}d\theta = \sqrt{\frac{2}{\pi}}K(m_X,\sigma_X)\int_0^{+\infty}\ln\theta\, e^{-r\theta^2+s\theta}d\theta,$$
$$K(m_X,\sigma_X) = \frac{e^{-\frac{m_X^2}{\sigma_X^2}}}{1+\operatorname{erf}\left(\frac{m_X}{\sqrt{2}\sigma_X}\right)}. \tag{6}$$

The integral (6) is not expressed by elementary functions, so we calculate it with the approximation method $X \to \sqrt{X} \to \left(\sqrt{X}\right)^2 \to X$:

$$\overline{\ln\Theta_{[0,+\infty)}} \cong \ln\overline{\Theta_{[0,+\infty)}} - \frac{1}{2}\frac{\sigma_\Theta^2}{\left(\overline{\Theta_{[0,+\infty)}}\right)^2}, \tag{7}$$

where $\sigma_{\Theta_{[0,+\infty)}}=\sqrt{D_{\Theta_{[0,+\infty)}}}$. Average $\overline{(\Theta_{[0,+\infty)})^2}=\frac{\overline{(X_{[0,+\infty)})^2}}{\sigma_X^2}$, where

$$\overline{(X_{[0,+\infty)})^2}=\frac{\sqrt{\frac{2}{\pi}}K(m_X,\sigma_X)}{\sigma_X}\int_0^{+\infty}x^2e^{-px^2-qx}dx=K(m_X,\sigma_X)\frac{\partial^2}{\partial q^2}\frac{1}{K(q,p)}$$
$$=\sigma_X^2\left(1+\frac{m_X^2}{\sigma_X^2}\right)+\sqrt{\frac{2}{\pi}}m_X\sigma_XK(m_X,\sigma_X).$$

Average

$$\overline{\Theta_{[0,+\infty)}}=\frac{\overline{X_{[0,+\infty)}}}{\sigma_X}=\frac{m_X}{\sigma_X}+\sqrt{\frac{2}{\pi}}K(m_X,\sigma_X),$$

so if $\Theta_{[0,+\infty)}$, $\overline{\Theta_{[0,+\infty)}}$, $\overline{(\Theta_{[0,+\infty)})^2}$ are statistically independent, the variance $D_\Theta$ is equal to

$$D_{\Theta_{[0,+\infty)}}=\overline{(\Theta_{[0,+\infty)})^2}-\left(\overline{\Theta_{[0,+\infty)}}\right)^2=1+\frac{m_X^2}{\sigma_X^2}+\sqrt{\frac{2}{\pi}}\frac{m_X}{\sigma_X}K(m_X,\sigma_X)-\left(\frac{m_X}{\sigma_X}+\sqrt{\frac{2}{\pi}}K(m_X,\sigma_X)\right)^2$$
$$=1-\sqrt{\frac{2}{\pi}}\frac{m_X}{\sigma_X}K(m_X,\sigma_X)-\frac{2}{\pi}K^2(m_X,\sigma_X)=\sigma^2_{[0,+\infty)}.$$

Then according to (7),

$$\overline{\ln\Theta_{[0,+\infty)}}=\ln\overline{\Theta_{[0,+\infty)}}-\frac{1}{2\left(\overline{\Theta_{[0,+\infty)}}\right)^2}\left(1+\sqrt{\frac{2}{\pi}}\frac{m_X}{\sigma_X}K(m_X,\sigma_X)\right),$$

where

$$\ln\overline{\Theta_{[0,+\infty)}}=\ln\frac{\overline{X_{[0,+\infty)}}}{\sigma_X}=\ln\overline{X_{[0,+\infty)}}-\ln\sigma_X.$$

Average

$$\overline{\ln\Theta_{[0,+\infty)}}=\ln\overline{\Theta_{[0,+\infty)}}-\frac{1}{2}\frac{\sigma_X^2}{\left(\overline{\Theta_{[0,+\infty)}}\right)^2}=\ln\frac{\overline{X_{[0,+\infty)}}}{\sigma_X}-\frac{\sigma_X^2}{2}\frac{1-\sqrt{\frac{2}{\pi}}\frac{m_X}{\sigma_X}K(m_X,\sigma_X)-\frac{2}{\pi}K^2(m_X,\sigma_X)}{\left(\overline{X_{[0,+\infty)}}\right)^2}.$$

Calculating the mean $\overline{(\ln\Theta_{[0,+\infty)})^2}$ with the method:

$$\begin{cases}m_g\cong g(m_X)+\frac{1}{2}g''(m_X)\sigma_X^2,\\ \sigma_g^2\cong\left(g'(m_X)\right)^2\sigma_X^2+\frac{(g'(m_X))^2(\mu_{4X}-\sigma_X^4)}{4}+g'(m_X)g''(m_X)\mu_{3X}\cong\left(g'(m_X)\right)^2\sigma_X^2+\frac{1}{2}\sigma_X^4\left(g''(m_X)\right)^2.\end{cases}\tag{8}$$

We obtain

$$\overline{(\ln\Theta_{[0,+\infty)})^2}\cong\left(\overline{\ln\Theta_{[0,+\infty)}}\right)^2+\frac{\sigma_\Theta^2}{(\overline{\Theta})^2}\left(1-\overline{\ln\Theta_{[0,+\infty)}}\right).$$

Then if RVs $\Theta_{[0,+\infty)}$, $\ln\Theta_{[0,+\infty)}$, $\left(\ln\Theta_{[0,+\infty)}\right)^2$ are statistically independent, the variance of the transformation $\ln\Theta_{[0,+\infty)}$ is equal to:

$$D_{\ln\Theta_{[0,+\infty)}}\cong\frac{\sigma^2_{\Theta_{[0,+\infty)}}}{(\overline{\Theta_{[0,+\infty)}})^2}\left(1-\overline{ln\Theta_{[0,+\infty)}}\right)=\sigma_X^2\frac{1-\sqrt{\frac{2}{\pi}}\frac{m_X}{\sigma_X}K(m_X,\sigma_X)-\frac{2}{\pi}K^2(m_X,\sigma_X)}{\left(\overline{X_{[0,+\infty)}}\right)^2}$$
$$\times\left(1-\ln\frac{\overline{X_{[0,+\infty)}}}{\sigma_X}-\frac{\sigma_X^2}{2}\frac{1-\sqrt{\frac{2}{\pi}}\frac{m_X}{\sigma_X}K(m_X,\sigma_X)-\frac{2}{\pi}K^2(m_X,\sigma_X)}{\left(\overline{X_{[0,+\infty)}}\right)^2}\right).$$

Average conversion $\overline{\exp(\ln\Theta_{[0,+\infty)})}$ is equal to

$$\overline{\exp\ln\Theta_{[0,+\infty)}}=\sqrt{\frac{2}{\pi}}K(m_X,\sigma_X)\int_0^{+\infty}e^{\ln\theta}e^{-r\theta^2+s\theta}d\theta=\sqrt{\frac{2}{\pi}}K(m_X,\sigma_X)\int_0^{+\infty}\theta\,e^{-r\theta^2+s\theta}d\theta=\overline{\Theta_{[0,+\infty)}}$$

and the conversion variance $\exp\ln\Theta_{[0,+\infty)}$ of statistically independent values $\Theta_{[0,+\infty)}$, $\ln\Theta_{[0,+\infty)}$, $(\ln\Theta_{[0,+\infty)})^2$ is equal to

$$D_{\exp\ln\Theta_{[0,+\infty)}} = \overline{(\exp\ln\Theta_{[0,+\infty)})^2} - \left(\overline{\exp\ln\Theta_{[0,+\infty)}}\right)^2 = \overline{(\Theta_{[0,+\infty)})^2} - \left(\overline{\Theta_{[0,+\infty)}}\right)^2$$
$$= \frac{1}{\sigma_X^2}\left(\overline{(X_{[0,+\infty)})^2} - \left(\overline{X_{[0,+\infty)}}\right)^2\right).$$

Thus, the statistical averages and the variance of the conversion algorithm functions 5 are related to the parameters of the original RV by the system of equations:

$$\begin{cases} \overline{\ln\Theta_{[0,+\infty)}} = \ln\overline{\Theta_{[0,+\infty)}} - \frac{1}{2}\frac{\sigma_\Theta^2}{\left(\overline{\Theta_{[0,+\infty)}}\right)^2}, \\ D_{\ln\Theta_{[0,+\infty)}} = \frac{\sigma^2_{\Theta_{[0,+\infty)}}}{\left(\overline{\Theta_{[0,+\infty)}}\right)^2}\left(1-\overline{\ln\Theta_{[0,+\infty)}}\right), \end{cases} \quad \text{and} \quad \begin{cases} \overline{\exp\ln\Theta_{[0,+\infty)}} = \overline{\Theta_{[0,+\infty)}}, \\ D_{\exp\ln\Theta_{[0,+\infty)}} = \frac{\overline{(X_{[0,+\infty)})^2} - \left(\overline{X_{[0,+\infty)}}\right)^2}{\sigma_X^2}. \end{cases} \tag{9}$$

### 2.3. Case 3

$$X \to \cos X \to \arccos\cos X \to X. \tag{10}$$

Transformation $g(X) = \cos X$ is implemented for multiple argument values $X \in (-\infty, +\infty)$, so average

$$\overline{\cos X} = \frac{1}{\sqrt{2\pi}\sigma_X}\int_{-\infty}^{+\infty} \cos x\, e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}\,dx = \cos(m_X)\, e^{-\frac{\sigma_X^2}{2}},$$

and dispersion

$$D_{\cos X} = \frac{1}{\sqrt{2\pi}\sigma_X}\int_{-\infty}^{+\infty}\left(\cos x - \overline{\cos X}\right)^2 e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}\,dx = \overline{(\cos X)^2} - \left(\overline{\cos X}\right)^2$$
$$= \frac{1+\cos(2m_X)\,e^{-2\sigma_X^2}}{2} - \left(\cos(m_X)\,e^{-\sigma_X^2}\right)^2,$$

which agrees with [22] if RV $X$, $\cos X$ are statistically independent.

Average conversion $\arccos\cos X$ is equal to

$$\overline{\arccos\cos X} = \frac{1}{\sqrt{2\pi}\sigma_X}\int_{-\infty}^{+\infty}\left(\arccos\cos x\right) e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}\,dx = \overline{X}$$

and the variance

$$D_{\arccos\cos X} = \frac{1}{\sqrt{2\pi}\sigma_X}\int_{-\infty}^{+\infty}\left(\arccos\cos x - \overline{\arccos\cos X}\right)^2 e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}\,dx$$
$$= \frac{1}{\sqrt{2\pi}\sigma_X}\int_{-\infty}^{+\infty}\left(x-\overline{X}\right)^2 e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}\,dx = D_X.$$

Thus, the statistical averages and the variance of the transformation functions according to the algorithm (10) are related to the original RV parameters $m_X$, $\sigma_X$ by the system of equations

$$\begin{cases} \overline{\cos X} = \cos(m_X)\,e^{-\frac{\sigma_X^2}{2}}, \\ D_{\cos X} = \frac{1+\cos(2m_X)\,e^{-2\sigma_X^2}}{2} - \left(\cos(m_X)\,e^{-\sigma_X^2}\right)^2, \end{cases} \quad \text{and} \quad \begin{cases} \overline{\arccos\cos X} = \overline{X}, \\ D_{\arccos X} = D_X. \end{cases}$$

### 2.4. Case 4

$$X \to \arccos X \to \cos\arccos X \to X. \tag{11}$$

This scenario starts with the conversion $\overline{\arccos X}$, which limits the set of allowed argument values $X$ to interval $X \in [-1, +1]$. In this case the normally distributed sample $X_1, X_2, \ldots, X_n$ with parameters $m_X$, $\sigma_X$ we renormalize the property of the function $K(m_\xi, \sigma_\xi)$ and present as a sample $V$:

$$V: \quad \frac{X_1}{m_X+3\sigma_X} = \nu_1, \quad \frac{X_2}{m_X+3\sigma_X} = \nu_2, \quad \ldots, \quad \frac{X_n}{m_X+3\sigma_X} = \nu_n. \tag{12}$$

Representation of the RV $X \in N(m_X, \sigma_X)$ as (12) by further eliminating items by value $X \geqslant m_X + 3\sigma_X$, introduces an error not exceeding 1%, but allows to correctly implement algorithm (11). The cosine function is symmetric about the $y$-axis, so from the point of view of the problem formulated in this work,

$$0 \leqslant \frac{X}{m_X + 3\sigma_X} = \frac{X_{[0,\Delta_{3\sigma_X}]}}{\Delta_{3\sigma_X}} = V \leqslant +1, \tag{13}$$

which became the rationing $C_X$ is equal to

$$C_X = \frac{1}{\int_0^{\Delta_{3\sigma_X}} \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left\{-\frac{(x-m_X)^2}{2\sigma_X^2}\right\} dx} = \frac{2}{\operatorname{erf}\left(\frac{3}{\sqrt{2}}\right) + \operatorname{erf}\left(\frac{m_X}{\sqrt{2}\sigma_X}\right)}.$$

We define the constant rationing $C_V$ from $C_X$:

$$C_X f(x)dx = C_V f(\nu)d\nu \quad \Rightarrow \quad C_V f(\nu) = \Delta_{3\sigma_X} C_X f(x).$$

Then considering that $X_{[0,\Delta_{3\sigma_X}]} = [0, \Delta_{3\sigma_X}]$ and the table integrals [24]

$$\int \operatorname{erf}(ax)\, dx = x \operatorname{erf}(ax) + \frac{x}{a\sqrt{\pi}} \exp\left(-a^2x^2\right),$$

$$\int x \operatorname{erf}(ax)\, dx = \left(\frac{x^2}{2} - \frac{1}{4a^2}\right) \operatorname{erf}(ax) + \frac{x}{2a\sqrt{\pi}} \exp\left(-a^2x^2\right),$$

we calculate the average $\overline{V} = \frac{1}{\Delta_{3\sigma_X}} \overline{X_{[0,\Delta_{3\sigma_X}]}}$:

$$\begin{aligned}\overline{V} &= \int_0^{+1} \nu\, C_V f(\nu)\, d\nu = \frac{C_X}{\Delta_{3\sigma_X}} \int_0^{\Delta_{3\sigma_X}} x \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}} dx \\ &= \frac{C_X}{2\Delta_{3\sigma_X}} \Bigg\{ \Delta_{3\sigma_X} \operatorname{erf}\left(\frac{3}{\sqrt{2}}\right) + \sqrt{2}\sigma_X \frac{3}{\sqrt{2}} \left[\operatorname{erf}\left(\frac{3}{\sqrt{2}}\right) + \frac{1}{\sqrt{\pi}} \exp\left(-\frac{9}{4}\right)\right] \\ &\quad + \frac{m_X}{\sqrt{2}\sigma_X} \left[\operatorname{erf}\left(\frac{m_x}{\sqrt{2}\sigma_X}\right) + \frac{1}{\sqrt{\pi}} \exp\left(-\left(\frac{m_X}{\sqrt{2}\sigma_X}\right)^2\right)\right]\Bigg\},\end{aligned}$$

and the mean of the square

$$\begin{aligned}\overline{V^2} &= \int_0^{+1} \nu^2\, C_V f(\nu)\, d\nu = \frac{C_X}{\Delta_{3\sigma_X}^2} \int_0^{\Delta_{3\sigma_X}} x^2 \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}} dx \\ &= \frac{C_X}{\Delta_{3\sigma_X}^2} \int_0^{(m_X+3\sigma_X)} x^2 \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}} dx \\ &= \frac{C_X}{2} \operatorname{erf}\left(\frac{3}{\sqrt{2}}\right) + \frac{C_X}{\Delta_{3\sigma_X}^2} \sqrt{2}\sigma_X m_X \Bigg\{ \left(\frac{-x+m_X}{\sqrt{2}\sigma_X}\right) \operatorname{erf}\left(\frac{-x+m_X}{\sqrt{2}\sigma_X}\right) + \frac{x}{a\sqrt{\pi}} \exp\left(-a^2x^2\right)\Bigg\}\Bigg|_0^{\Delta_{3\sigma_X}} \\ &- \frac{C_X}{\Delta_{3\sigma_X}^2} 2\sigma_X^2 \Bigg\{ \left[\frac{1}{2}\left(\frac{-x+m_X}{\sqrt{2}\sigma_X}\right)^2 - \frac{1}{4}\right] \operatorname{erf}\left(\frac{-x+m_X}{\sqrt{2}\sigma_X}\right) + \frac{x}{2\sqrt{\pi}} \exp\left[-\left(\frac{-x+m_X}{\sqrt{2}\sigma_X}\right)^2\right]\Bigg\}\Bigg|_0^{\Delta_{3\sigma_X}}.\end{aligned}$$

Then if random variables $V$, $V^2$ are statistically independent, the variance of the RV $V$ will be equal to

$$D_V = \overline{V^2} - \left(\overline{V}\right)^2 = \sigma_V^2.$$

Now, accepting a sample with parameters $\overline{V}$ and $\sigma_V$ for the initial, we calculate the statistical parameters $\overline{\cos V}$ and $\sigma_{\cos V}$ of cosine transformation by the basic sampling algorithm (11):

$$V \to \arccos V \to \cos \arccos V \to V.$$

Computing $\overline{\cos V}$

$$\overline{\cos V} = \int_0^{+1} \cos(\nu)\, C_V f(\nu)\, d\nu = \frac{C_X}{\sqrt{\pi}} \int_0^{\Delta_{3\sigma_X}} \cos\left(\frac{x}{\Delta_{3\sigma_X}}\right) \exp\left\{-\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right)^2\right\} d\frac{x-m_X}{\sqrt{2}\sigma_X}. \tag{14}$$

To use the table integrals (861.20) [23]:

$$\int_0^{+\infty} \cos(bx)\,\exp\left(-a^2x^2\right)dx = \frac{\sqrt{\pi}}{2a}\exp\left(-\frac{b^2}{4a^2}\right), \tag{15}$$

and (No 25) [24]:

$$\int_0^{+\infty} \sin(bx)\,\exp\left(-a^2x^2\right)dx = \frac{\pi}{2ia}\operatorname{erf}\left(\frac{ib}{2a}\right)\exp\left(-\frac{b^2}{4a^2}\right), \tag{16}$$

let us transform (14) to the form:

$$\overline{\cos V} = \frac{C_{X\in[0,+\infty)}}{\sqrt{\pi}}\int_0^{\Delta_{3\sigma_X}} \cos\left(\frac{m_x}{\Delta_{3\sigma_X}} + \left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right)\frac{\sigma_X\sqrt{2}}{\Delta_{3\sigma_X}}\right)\exp\left\{-\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right)^2\right\} d\frac{x-m_X}{\sqrt{2}\sigma_X}.$$

Replacement (13) covers at least 99% of the data, as we apply (15) and (16) and transform the upper bound of integration as

$$\Delta_{3\sigma_X} \to +\infty. \tag{17}$$

Then

$$\overline{\cos V} = \frac{C_{X\in[0,+\infty)}}{\sqrt{\pi}}\int_0^{\infty}\cos\left(\frac{m_x}{m_X+3\sigma_X} + \frac{t\,\sigma_X\sqrt{2}}{m_X+3\sigma_X}\right)e^{-t^2}dt = \frac{C_{X\in[0,+\infty)}}{2}\left\{\cos\left(\frac{m_x}{m_X+3\sigma_X}\right)\right.$$

$$\left. - \sin\left(\frac{m_x}{m_X+3\sigma_X}\right)\frac{\sqrt{\pi}}{i}\operatorname{erf}\left(i\frac{\sigma_X\sqrt{2}}{2(m_X+3\sigma_X)}\right)\right\}\exp\left\{-\left(\frac{\sigma_X\sqrt{2}}{2(m_X+3\sigma_X)}\right)^2\right\}, \tag{18}$$

where providing (17) can be put to (2.3.15) in [13].

According to ( [24]) tabulated values of the expression

$$\frac{\sqrt{\pi}}{i}\operatorname{erf}\left(i\frac{\sigma_X\sqrt{2}}{2(m_X+3\sigma_X)}\right)\exp\left\{-\left(\frac{\sigma_X\sqrt{2}}{2(m_X+3\sigma_X)}\right)^2\right\} = \frac{\sqrt{\pi}}{2i}\operatorname{erf}(i\xi)\exp(-\xi^2). \tag{19}$$

In the interval $0 < \xi < 0.924$, value (19) increases monotonically, reaching maximum 0.541 at $\xi = 0.924$, and then it goes down monotonously, going to the limit $\xi \to \infty$ to zero. This means that in specific calculations the expression (19) can be replaced by the average in the range of values $\xi$.

Let us compute with (17) average $\overline{\cos V^2}$, applying the conversion

$$\overline{\cos V^2} = \overline{\frac{1}{2} + \frac{\cos 2V}{2}} = \frac{1}{2}\left(1 + \overline{\cos 2V}\right).$$

Then

$$\overline{\cos 2V} = \frac{C_X}{\sqrt{\pi}}\int_0^{\infty}\cos\left(\frac{2m_X}{\Delta_{3\sigma_X}} + \frac{2t\,\sigma_X\sqrt{2}}{\Delta_{3\sigma_X}}\right)e^{-t^2}dt$$

$$= \frac{\left\{\cos\left(\frac{2m_X}{\Delta_{3\sigma_X}}\right) - \sin\left(\frac{2m_X}{\Delta_{3\sigma_X}}\right)\frac{\sqrt{\pi}}{i}\operatorname{erf}\left(i\frac{\sigma_X\sqrt{2}}{\Delta_{3\sigma_X}}\right)\right\}\exp\left\{-\left(\frac{\sigma_X\sqrt{2}}{m_X+3\sigma_X}\right)^2\right\}}{\operatorname{erf}\left(\frac{3}{\sqrt{2}}\right) + \operatorname{erf}\left(\frac{m_X}{\sqrt{2}\sigma_X}\right)},$$

and the variance of the cosine transform is calculated as

$$D_{\cos V} = \overline{(\cos V)^2} + \left(\overline{\cos V}\right)^2. \tag{20}$$

Part (2.6.17) (18) was taken into account by the average value of the coefficient before the sine function: $0.1 \cdot \sin\left(\frac{2m_X}{m_X+3\sigma_X}\right)$.

Taking into account (13), we make an estimation of $\overline{\arccos V}$ in approximation (8):

$$\overline{\arccos V} \cong \arccos \overline{V} + \frac{\overline{V}\sigma_V^2}{\left(1-\overline{V}^2\right)^{3/2}} \tag{21}$$

and conversion $\overline{(\arccos V)^2}$

$$\overline{(\arccos V)^2} \cong \left(\overline{\arccos V}\right)^2 + 2\left(\frac{1}{1-\overline{V}^2} + \frac{\overline{\arccos V}\,\overline{V}}{\left(1-\overline{V}^2\right)^{3/2}}\right)\sigma_V^2.$$

Then

$$D_{\arccos V} = \overline{(\arccos V)^2} - \left(\overline{\arccos V}\right)^2 = \left[\frac{2}{1-\overline{V}^2} + \frac{2\,\overline{\arccos V}\,\overline{V}}{\left(1-\overline{V}^2\right)^{3/2}}\right]\sigma_V^2 = \sigma^2_{\arccos V}. \tag{22}$$

We apply the model of the transformation of the variable $X \to \sqrt{X} \to (\sqrt{X})^2 \to X$ for values $X \in [0, +\infty)$ to algorithm

$$\ln X \leftarrow X \to \exp X. \tag{23}$$

$$\begin{cases} D_{\ln X_{[0,+\infty)}} = \overline{(\ln X_{[0,+\infty)})^2} - \left(\overline{\ln X_{[0,+\infty)}}\right)^2, \\ D_{X_{[0,+\infty)}} = \overline{(X_{[0,+\infty)})^2} - \left(\overline{X_{[0,+\infty)}}\right)^2, \\ D_{\exp X_{[0,+\infty)}} = \overline{(\exp X_{[0,+\infty)})^2} - \left(\overline{\exp X_{[0,+\infty)}}\right)^2. \end{cases}$$

The transformation algorithm (23) is similar to algorithm (5). Therefore, statistical averages $\overline{\exp X_{[0,+\infty)}}$, $\overline{\ln X_{[0,+\infty)}}$ and variances $D_{\exp X_{[0,+\infty)}}$, $D_{lnX_{[0,+\infty)}}$ the transformations in (23) will be equal (9).

Transformation

$$\arccos X \leftarrow X \to \cos X \tag{24}$$

is also more correctly to implement for the renormalized variable (12). Then the statistical average $\overline{\arccos V}$ will be equal to (21), and the variance $D_{\arccos V}$ to (22). Statistical average $\overline{\cos V}$ and dispersion $D_{\cos V}$ of the cosine transform in (24) is described by formulas (18) and (20).

## 3. Conclusions

In this article, the statistical models of the mean and variances of functional transformations in a straight line $g(X) = \cos X$; $\exp X$ and inverted $g^{-1} = \arccos X$; $\ln X$ elementary functions of normally distributed random variables are developed. Moreover, the statistical models are constructed for RV samples with a semi-limited volume, and the substantiation of the transition to a two-sided confidence interval limit for the transformation functions stated in the task of this work is carried out. The obtained relationships represent an opportunity to apply the error transfer to functional transformations of experimental data with random values.


[1] Hudson D. J. Lectures on elementary statistics and probability. Geneva, CERN (1963).

[2] Ku H. H. Notes on the Use of Propagation of Error Formulas. Journal of Research of the National Bureau of Standards. Section C: Engineering and Instrumentation. **70C** (4), 263–273 (1966).

[3] Bevington P., Robinson D. K. Data Reduction and Error Analysis for the Physical Sciences. McGraw-Hill, Boston (2002).

[4] Taylor J. R. An Introduction to Error Analysis. Univ. Sci. Books, Sausalito, CA (1997).

[5] Wikipedia. Propagation of uncertainty. https://en.wikipedia.org/wiki/Propagation of uncertainty.

[6] Suhir E. Applied Probability for Engineers and Scientistics. McGraw-Hill Companies (1997).

[7] Papoulis A. Probability, Random Variables, and Stochastic Processes. McGraw-Hill (1991).

[8] Bohm G., Zech G. Introduction to Statistics and data Analysis for Physics. Verlag Deutsches Elektronen-Synchrotron (2010).

[9] Samorodnitsky G., Taqqu M. S. Stable Non-Gaussian Random Processes. New York, Chapman and Hall (1994).

[10] Wichmann E. Quantum Physics. Berkeley Physics Course. Vol. 4. McGraw-Hill Book Company (1971).

[11] Leach A. Molecular Modelling: Principles and Applications. Prentice Hall (2001).

[12] Goodman J. W. Statistical Optics. John Wiley and Sons. Inc. (2015).

[13] Patel J. K., Read C. B. Handbook of the Normal Distribution. New York, Dekker (1982).

[14] Marvin R., Arnljot H. System Reliability Theory: Models, Statistical Methods, and Applications. John Wiley and Sons. Inc. (2004).

[15] Liang T., Jia X. Z. An empirical formula for yield estimation from singly truncated performance data of qualified semiconductor devices. Journal of Semiconductors. **33** (12), 125008 (2012).

[16] Gu K., Jia X., You T., Liang T. The yield estimation of semiconductor products based on truncated samples. International Journal of Metrology and Quality Engineering. **4** (3), 215–220 (2013).

[17] Einstein A. On the motion of particles suspended in a fluid at rest, required by the molecular-kinetic theory of heat. Collection of articles: Leningrad, ONTI-GROTL (1936).

[18] Einstein A., Smolukhovsky M. Brownian motion. Collection of articles. Leningrad, ONTI-GROTL (1936).

[19] Risken H. The Fokker–Planck quation. Berlin–Heidelberg, Springer (1989).

[20] Gihman I., Skorochodov A. The Theory of Statistic Processes I. Springer, Berlin (2004).

[21] Kosobutskyy P. S., Karkulovska M. S. Simulation of statistical mean and variance of normally distributed random values, transformed by nonlinear functions $\sqrt{|X|}$ and $\sqrt{X}$. Mathematical Modeling and Computing. **9** (2), 318–325 (2022).

[22] Rode G. G. Propagation of measurement errors and measured means of a physical quantity for the elementary functions $\cos X$ and $\arccos X$. Ukrainian Journal of Physics. **61** (4), 345–352 (2016).

[23] Dwight H. Tables of Integrals and other Mathematical data. New York, The Macmillan Company (1961).

[24] Abramowitz M., Stegun I. A. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. New York, Dover (1972).

## Моделювання статистичних середнього та дисперсії нормально $N_X(m_X, \sigma_X)$ розподілених даних, перетворених нелінійними функціями $g(X) = \cos X$, $e^X$ та оберненими до них $g^{-1}(X) = \arccos X$, $\ln X$

Кособуцький П. С., Каркульовська М. С.

*Національний університет "Львівська політехніка", вул. С. Бандери, 12, 79013, Львів, Україна*

Обгрунтовані аналітичні співвідношення обчислення статистичних середніх і дисперсії функцій $g(X) = \cos X$, $e^X$, $g^{-1}(X) = \arccos X$, $\ln X$ перетворення нормально $N_X(m_X, \sigma_X)$ розподіленої випадкової величини.

**Ключові слова:** *статистичне середнє; дисперсія; перетворення; нормальний розподіл; випадкова величина.*