DOI: 10.31319/2519-8106.1(46)2022.258342
UDK 519.22
V. Lytvynenko, Ph. D., Assosiate Professor
S. Marchenko, Ph. D., Assosiate Professor
O. Syanov, Doctor of Technical Sciences, Professor
Dnipro State Technical University, Kamianske

STATISTICAL METHODS OF PROCESSING MEASUREMENT DATA ON MEASUREMENT INSTRUMENTS

The article investigates the application of methods of mathematical statistics in the processing of data on parametric failures of measuring instruments (MI). Calculations and comparative analysis of the method of maximum likelihood and the method of least squares in estimating the parameters and confidence intervals of the functions of the distribution of probability of failure. An expert approach to estimating the parameters of the failure model based on statistical data is proposed.
***Keywords**: diffusion distributions, models of failures of measuring instruments, method of maximum likelihood, method of least squares.*

В статті досліджено застосування методів математичної статистики при обробці даних про параметричні відмови засобів вимірювань(ЗВ). Проведені розрахунки і порівняльний аналіз методу максимальної правдоподібності і методу найменших квадратів при оцінці параметрів і довірочних інтервалів функцій розподілу ймовірностей виникнення відмов. Запропонований експертний підхід оцінювання параметрів моделі відмов на основі статистичних даних.
***Ключові слова**: дифузійні розподіли, моделі відмов засобів вимірювань, метод максимальної правдоподібності, метод найменших квадратів.*

Problem's formulation

The practice of using measuring instruments for various purposes has shown that without the installation and conduct of special work to ensure metrological reliability, which is determined by hidden failures, such devices are not effective enough. The peculiarity of hidden (parametric) failures in MI is the hidden nature of their appearance. Probabilistic-physical (diffusion) models of failures are used to mathematically describe the process of hidden failures. Therefore, the task of estimating the parameters of diffusion distributions on the basis of failure statistics in the process of MI operation is relevant.

Analysis of recent research and publications

The analysis of the results of Monte Carlo modeling of samples from different types of empirical distributions of object developments conducted in [1] shows that for qualitative approximation of failure statistics it is expedient to use DM (diffusion-monotone) and DN (diffusion-nonmonotonic) distributions In this work, simulations were performed and 1,200 samples with a volume of 100 values of up to (on) failure were processed and a number of statistical tests were performed. As theoretical models of the distribution functions of the failure time were adopted: DM and DN-distribution functions, normal distribution (N), logarithmically normal (LN), Weibull (W). Distribution parameters are evaluated as sample unbiased estimates μ and ν for the case of the full test plan (NUN) [1].

As we can see, the efficiency of diffusion models is confirmed by the results of statistical modeling. Unfortunately, in real operating conditions, the specialists of the metrological service do not always have a sufficient amount and different origin of information on failure statistics. In this case, the main advantage of DM and DN-models is to take into account a priori information and the nature of the random diffusion process, which leads to the parameter exceeding the allowable limits.

Formulation of the study purpose

The purpose of the work is to study the methods of processing statistical data on metrological failures of measuring instruments and to develop methods for estimating the parameters of diffusion laws of distribution of failures based on the results of controlled operation.

Presenting main material

For research we will use statistical data on failures of induction single-phase electricity meters. The term of verification of such electricity meters according to passport data is 16 years, and operation is at least 30 years (for some models at least 40 years). The source of information on the operational characteristics of the MI is the data obtained from the results of controlled operation. Picture 1 shows the points of the empirical distribution function according to the data of controlled operation (NUr plan) of two similar groups of MI. The calculation of ordinal statistics and the corresponding values of the distribution function were performed by the Johnson method. Since the results of data processing of controlled operation of devices during the first 13 years coincide well, and then the data for the calculation of the distribution function are missing in the future we will consider only the first group of devices.

Using the statistics of failures of induction single-phase electricity meters from a sample of n = 60 (m = 22 — refused, s = 38 — not refused) we estimate the parameters of diffusion models of metrological failures by the method of maximum likelihood (MML) and compare with the results obtained for Weibull's law.

For a truncated and censored sample n, the plausibility function in the case of diffusion distributions takes the form

$$L = \prod_{i=1}^{m} f(\tau_i) \cdot \prod_{j=1}^{s} [1 - F(t_j)], \quad (1)$$

where $f(\tau)$ — the density of the distribution of operating time to failure, and $F(t)$ — the distribution of the time of failure-free operation (values τ_i and t_j are given in Annex B, Table B.1). The vector of parameters of diffusion distributions $\Theta = [\mu, v]$ at which the function $L(X,\Theta)$, or $\ln L(X,\Theta)$ reaches the point of local optimum (maximum) is the most plausible estimate of the parameters of the theoretical distribution (X -matrix of random values of developments).

To obtain a numerical solution of the equation $\frac{\partial L(X,\Theta)}{\partial \Theta} = 0$, we use the following method. To find the maximum $L(X,\Theta)$ using the optimization procedures of the Matlab package, we find the minimum of the likelihood function with the opposite sign $(-L(X,\Theta)$. The search for the minimum was performed using an algorithm that implements the simplex Nelder-Mead method (does not require gradient calculation).

The result of the calculation for the case of DM-distribution showed that the iteration process coincides with the accuracy of the argument of the equal function 10^{-9}. The number of estimates of the objective function — 153, the number of iterations — 75. The most plausible estimates of the parameters $\tilde{\mu}$ = 22,0915 and $\tilde{v}$ = 0,7630. To verify the results of the calculation, we construct the surface (three-dimensional model) of the plausibility function Z (v, μ) and its projection on the plane of values XY = (v μ) (Fig. 1).

For convenience of calculation we will round off values of parameter of scale to tenths $\mu = 22{,}1$, and parameter of form to hundredths $v = 0{,}76$ (in fig. 2 designations mu = μ and v = v) and we will present their vector-line m = [22,1; 0.76]. We will change the value μ in the range [15; 34] with a step of 0.1, and v = 0.1… 2 with a step of 0,01.

First, we calculate the values of the elements of the matrix L, which are equal $L_{i,j} = L(\mu_i, v_j, \tau, t)$ to $i = j = 1..191$. The result will be a surface matrix of the likelihood function, which is calculated as the ratio $Z = L(\mu_i, v_j, \tau, t) / L(m, \tau, t)$. As can be seen from fig. 1 Z = 1 at a point with coordinates $X = v = 0{,}76$ and Y = $\mu = 22{,}1$ which confirms the reliability of the calculation of estimates of the parameters of the DM-distribution.

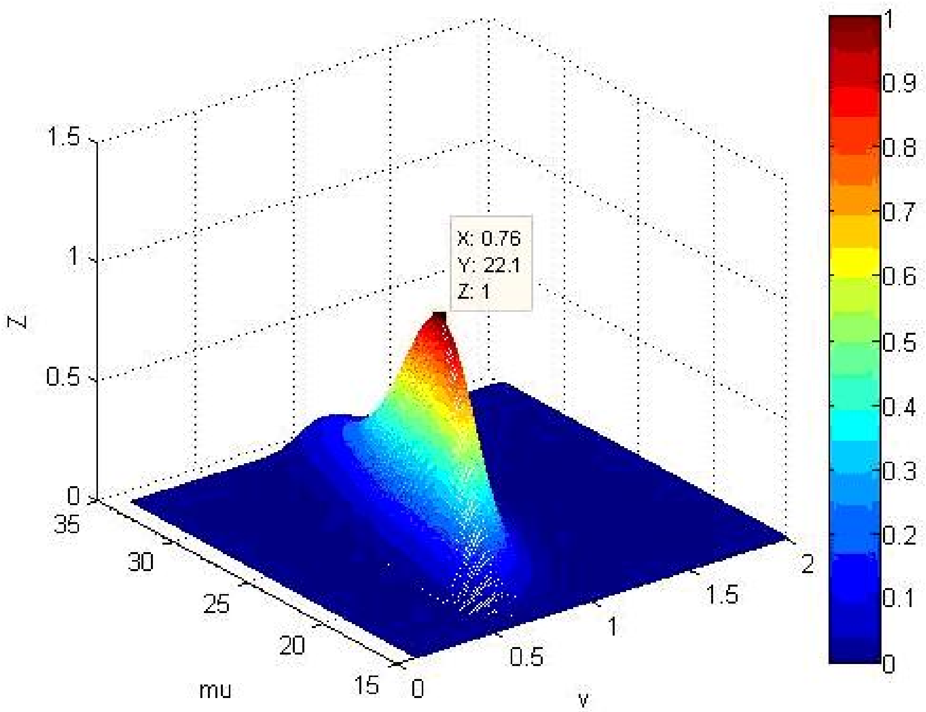

Fig. 1. Graph of the surface of the likelihood function Z(v, μ)

The interval estimation of the DM-distribution parameters according to the NUr test plan for the lower and upper confidence limits of the parameters μ_H, v_H and μ_B, v_B accordingly, is calculated using the expressions given in [1—2]

$$\mu_H = \tilde{\mu}\left[1+\frac{\tilde{v}^2 U_q}{2m}-\frac{\tilde{v} U_q}{2\sqrt{m}}\sqrt{4+\tilde{v}^2 U_q/m}\right]; \quad (2)$$

$$v_H = \tilde{v}\left[1+\frac{(1+2\tilde{v}^2)U_q^2}{4m}-\frac{U_q}{4m}\sqrt{[8m+(1+2\tilde{v}^2)U_q^2](1+2\tilde{v}^2)}\right]; \quad (3)$$

$$\mu_B = \tilde{\mu}\left[1+\frac{\tilde{v}^2 U_q}{2m}+\frac{\tilde{v} U_q}{2\sqrt{m}}\sqrt{4+\tilde{v}^2 U_q/m}\right]; \quad (4)$$

$$v_B = \tilde{v}\left[1+\frac{(1+2\tilde{v}^2)U_q^2}{4m}+\frac{U_q}{4m}\sqrt{[8m+(1+2\tilde{v}^2)U_q^2](1+2\tilde{v}^2)}\right], \quad (5)$$

where q — is the reliability of the estimate, v — the number of recorded failures (in this example m = 22), U_q — the value of the quantile of the normalized normal distribution.

Tabl. 1 presents the calculations of the values of the confidence limits for the values of the parameters of the scale μ and v in the case of DM-distribution for a number of values of reliability q. To compare the diffusion distributions of failures, similar calculations were performed under the assumption of the law DN-distribution of the occurrence of metrological failure (tabl. 2).

Estimation of the DN-distribution parameters by the MML method gave the following values: $\tilde{\mu}$ = 30,1752 and $\tilde{v}$ = 0,8644. It should be noted that the sample weighted estimates of the estimated parameters are $\tilde{\mu}_3$ = 20,8675 and $\tilde{v}_3$ = 0,5046 and belong to the confidence interval for the parameters of DM and DN-distributions at specified values of reliability q [3—4].

Table 1. Confidence limits for DM-distribution functions

q	DM	
0,95	$\mu \in$ [16,9182; 28,8467]	$\nu \in$ [0,5308; 1,0967]
0,99	$\mu \in$ [15,1667; 32,1801]	$\nu \in$ [0,4580; 1,2710]
0,999	$\mu \in$ [13,4317; 36,3344]	$\nu \in$ [0,3894; 1,4951]
0,9999	$\mu \in$ [12,1713; 40,0971]	$\nu \in$ [0,3418; 1,7033]

Table 2. Confidence limits for DN-distribution functions

q	DN	
0,95	$\mu \in$ [22,3095; 40,8142]	$\nu \in$ [0,5857; 1,2757]
0,99	$\mu \in$ [19,7189; 46,1762]	$\nu \in$ [0,5002; 1,4937]
0,999	$\mu \in$ [17,2001; 52,9381]	$\nu \in$ [0,4207; 1,7761]
0,9999	$\mu \in$ [15,4004; 59,1245]	$\nu \in$ [0,3662; 2,0404]

Consider another example of processing and analyzing data on failures of electricity meters. In this case, the number of failures r_j in the group N_j with MI in the interval is known $[0;T_j]$.

The authors of this monograph propose to calculate the values of the empirical distribution function by the formula $F_i = 1 - e^{-\lambda_i}$, where $\lambda_i = \sum_{j=1}^{i} \frac{r_j}{N_j}$ the accumulated failure rate is, and to determine the parameters of the theoretical model by the method of least squares. We will use this assessment approach to study the quality of the approximation of failure statistics using diffusion distributions.

The parameters of diffusion failure models are obtained using the least squares method (LSM) by minimizing the objective function of the form $\Psi(\Theta) = \sum_{j=1}^{k} (F_2 - F_{DM}(t_j, \Theta)) \Rightarrow \min$, where k — number of options, F_2 — the value of the empirical distribution function, $\Theta = [\mu, \nu]$ — the parameters of the theoretical model F_{DM}. The problem of nonlinear programming was solved using numerical methods for optimizing the Matlab environment (modified simplex Nelder-Mead method). As a result, received $\Theta = [11{,}5390,\ 0{,}4129]$.

Based on the analysis of the obtained empirical distribution functions and the monotonic nature of random processes of change in time of metrological characteristics MI, it is shown that the DM distribution function better approximates empirical distribution functions based on different sources of statistical information compared to Weibull distribution results (Fig. 2).

Here is one of the possible options for expert evaluation of the parameters of the theoretical model of failures based on statistics of failures of controlled operation and data on the number of failures at a fixed point in time. For example, if the limits of the confidence interval for the parameters of the DM-failure model (truncated sample) $\mu \in$ [13,4317; 36,3344] and $\nu \in$ [0,4207; 1,7761] at q = 0,999, you can choose the parameters of the DM-distribution, which would belong to a given confidence interval and do not contradict the hypothesis in the selection of the theoretical function for data obtained on the basis of the accumulated failure rate. Checking the adequacy of the statistical model in the second case can be assessed using the Kolmogorov agreement criterion and calculations of graphs of distribution functions using a PC (non-parametric statistics method).

In our example, the parameters $\mu = 13{,}4318$ and $\nu = 0{,}3895$ are within the confidence interval and the selected theoretical model does not contradict the experimental data on the criterion of Kolmogorov's agreement at the level of significance $\alpha = 0{,}05$. The maximum difference between the empirical and theoretical distribution does not exceed 0,1884.

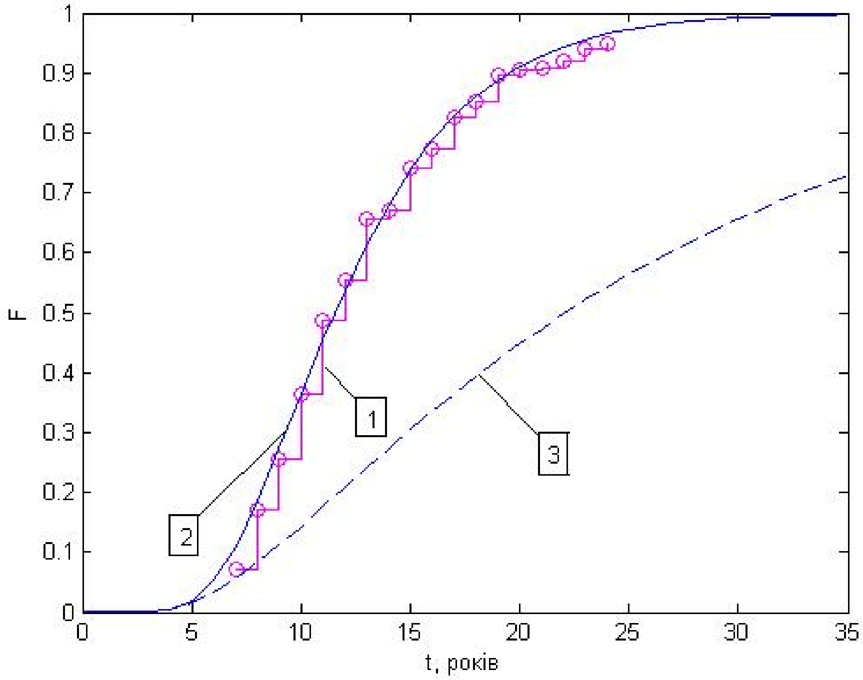

Fig. 2. Approximation of experimental data using DM-distribution: 1 — empirical function of distribution of failures; 2 — approximating curve of DM-distribution (estimation of function parameters was performed by MNC method); 3 — DM-distribution function (estimation of distribution parameters 3 was performed using the MML method)

Conclusions

The conducted research allowed to choose and substantiate the methods of statistical processing of metrological failures of MI and to set the task of complex application of these methods. The proposed set of methods increases the reliability of the procedure for processing and analysis of failure statistics in assessing the reliability of MI. The methods of mathematical statistics developed in the work expand the results of similar studies of metrological reliability of MI in operation.

References

[1] Azarskov, V.N., & Strelnikov, V.P. (2004). Nadyezhnost system upravleniia i avtomatiki [Reliability of control systems and automation]. Kyiv: NAU [in Russian]

[2] Strelnikov, V.P., & Egorov, S.V. (2008). Isledovanie svoistv statistiki primeniaemoi dlia otsenki veroiatnosti bezotkaznoi raboty [Investigation of the properties of statistics used to assess the probability of failure-free work]. Matematychni mashyny i systemy — Mathematical machines and systems, 3, 147–152 [in Russian]

[3] Lytvynenko, V.A. (2013). Analiz pokaznykiv nadiynosti sukupnosti zasobiv vumiriuvalnoi tekhiky v umovath shyrokogo promyslovogo zastosyvania[Analysis of reliability indicators of the set of measuring equipment in terms of wide industrial application]: *Extended abstract of candidate's thesis*. Kyiv:NTUU «KPI» [in Ukrainian]

[4] Lytvynenko V.A. (2012). Deiaki pytania modeliuvania protsesy ekspluatatsii i metrologichnogo obslugovuvania zasobiv vymiriuvalnoi tekhyky pry optymizatsii metrologichnogo kontroliu[Some issues of modeling the process of operation and metrological maintenance of measuring equipment in the optimization of metrological control] Matematychne modelyvanya — Mathematical modeling, 1, 70–75 [in Ukrainian]

СТАТИСТИЧНІ МЕТОДИ ОБРОБКИ ДАНИХ ПРО ВІДМОВИ ЗАСОБІВ ВИМІРЮВАНЬ

Литвиненко В.А., Марченко С.В., С'янов О.М.

Реферат

Для математичного опису процесу виникнення прихованих відмов застосовують ймовірністно-фізичні (дифузійні) моделі відмов. Тому актуальною є задача оцінка параметрів дифузійних розподілів на основі статистики відмов в процесі експлуатації ЗВ.

Мета дослідження полягає в розробці методу статистичної обробки даних папаметричних і раптових відмов. Особливість параметричних відмов полягає в прихованому характеру їх появи. Для математичного опису процесу виникнення прихованих відмов застосовують ймовірністно-фізичні (дифузійні) моделі відмов. В роботі проведені розрахунки і порівняльний аналіз методу максимальної правдоподібності і методу найменших квадратів при оцінці параметрів і довірочних інтервалів функцій розподілу ймовірностей виникнення відмов. На основі аналізу отриманих емпіричних функцій розподілу та монотонний характер випадкових процесів зміни в часі метрологічних характеристик ЗВ показано, що функція DM-розподілу краще апроксимує емпіричні функції розподілу побудовані на основі різних джерел статистичної інформації в порівнянні з результатами отриманими для розподілу Вейбулла.

Запропонований набір методів підвищує достовірність процедури обробки і аналізу статистики відмов при оцінці показників надійності ЗВ. Розвинуті в роботи методи математичної статистики розширюють результати досліджень в області метрологічної надійності. Дані методи розроблені для спеціалістів метрологічних служб підприємств при виконанні спеціальних робіт по забезпеченню метрологічної надійності в складі автоматизованих систем.

Література

1. Азарсков В.Н., Стрельников В.П. Надежность систем управления и автоматики : учеб. пособ. Киев : НАУ, 2004. 164 с.
2. Стрельников В.П., Егоров С.В. Исследование свойств статистики, применяемой для оценки вероятности безотказной работы. *Математичні машини і системи*. 2008. № 3. С. 147–152.
3. Литвиненко В.А. Аналіз показників надійності сукупності засобів вимірювальної техніки в умовах широкого промислового застосування:автореф. дис. … канд. техн. наук: 05.01.02. Київ, 2013. 20 с.
4. Литвиненко В.А. Деякі питання моделювання процесу експлуатації і метрологічного обслуговування засобів вимірювальної техніки при оптимізації метрологічного контролю. *Математичне моделювання*. 2012. Вип. 1(26). С. 70–75.