ON THE VALUE OF THE EVALUATION OF THE COEFFICIENT OF VARIATION IN PROCESSING STATISTICAL DATA ON THE FAILURE OF MECHANICAL AND CHEMICAL SYSTEMS

Prof. Kerimova Lala[1], Ph.D Goshgar Aliyev[2], Esmira Ahmadova[1], Turkana Gahramanli[1], Zemfira Aliyeva[3], Xumara Gahramanli[4]

Article History: Received: 03/03/2023 Revised: 05/04/2023 Accepted: 13/05/2023

[1]Western Caspian University, Department of İnformation Technologies, Azerbaijan
[2]Institute of Catalysis and Inorganic Chemistry named after Acad. M. Nagiyev, Azerbaijan
[3]Mingachevir State University, Department of Information Technology
[4]Azerbaijan State Pedagogical University, Department of Computer Sciences, Azerbaijan

Email:Professor_lk@mail.ru,hemproblem@mail.ru,aesmiranq@gmail.com, turkana.gahramanli927@mail.ru, eliyeva.zemfira@mail.ru,khumara.gahramanli.93@mail.ru

Doi: 10.31838/ecb/2023.12.5.387

When processing statistical data, the standard deviation σ(x) is often taken as the main numerical characteristic of the dispersion of a random variable x. Since the scattering of a random variable x is associated with the deviation x-a of this variable from its center of distribution a=M(x), then it can also be characterized by a coefficient of variation V equal to the ratio of the standard deviation to the mathematical expectation of the random variable

$$V=\sigma(x)/M(x)$$

The coefficient of variation shows how large the dispersion of a random variable is compared to its mathematical expectation. Unlike the standard deviation, the coefficient of variation, being a dimensionless indicator, also makes it possible to compare individual samples obtained under different conditions.

When analyzing the reliability of mechanical systems, the following distribution laws of the probability of random variables are most often encountered: truncated normal, Weibull, exponential; somewhat less often - Rayleigh's law and gamma - distribution. Among the listed distribution laws, due to its universality, the Weibull distribution occupies an important place. The density of this distribution (for t> 0) has the form:

f (t)= a·m·tm-1 exp(-atm)

Where, a is the distribution parameter that determines the scale, m is the parameter that determines the skewness and kurtosis of the distribution.

In particular, the Weibull distribution gives a fairly good approximation both for sudden (m≤1) and gradual (m>1) failures.

The coefficient of variation in the Weibull distribution is determined by the following relationship [1].

$$V=\frac{\sqrt{Q\left(1+\frac{2}{m}\right)-\left[Q\left(1+\frac{1}{m}\right)\right]^{2}}}{Q\left(1+\frac{1}{m}\right)}$$

Where Q(x) $=\int_0^{\infty} z^{x-1}e^{x}dz$ - is the gamma function, the values of which are tabulated.

As can be seen, dependence (2) establishes a relationship between the Weibull distribution parameter m and the coefficient of variation V and does not depend on time t. Therefore, it can be used to estimate the distribution parameters according to statistical data, based on the sample value of the coefficient of variation

$$V=\sigma/T.$$

Using the known value of V from the nomogram (see figure) or by calculating the dependence (2), the value of the parameter m of the Weibull distribution is found. The second parameter a, taking into account the known value of the mean time to failure T, is determined from the formula:

$$T=\frac{Q\left(1+\frac{1}{m}\right)}{a^{Vm}} \qquad a=\left[\frac{Q\left(1+\frac{1}{m}\right)}{T}\right]^{m} \tag{3}$$

The most typical for the Weibull distribution are the values of the parameter m=1; 2; 3.5 (see figure). With the parameter m=1, the Weibull distribution becomes exponential. For m ≥ 3.5, the Weibull distribution can be approximately considered normal. The region m=2 is typical for the Rayleigh distribution.

With the indicated values of the Weibull distribution parameter m, according to dependence (2) we have: at m=1, V=1; at m=2, V=0.52; at m=3.5, V=0.34.

As can be seen, in the zone of predominance of gradual failures, the scattering of a random variable decreases significantly. The estimate of the coefficient of variation for the case of a truncated normal distribution is quite fully considered in [2]. It is shown that in the range of changes in the ratio b0=a0/σ0 (truncated normal distribution parameters) from 0.5 to 4, the coefficient of variation changes from 1 to 0.25.

For Rayleigh distribution

$$V=\frac{\sigma_x}{m_x}=\frac{\sqrt{2-\pi/2}}{\sqrt{\pi/2}} \tag{4}$$

That is, in this case, the value of the coefficient of variation is V≈0.52 (which coincides with its value given above with the parameter m=2 of the Weibull distribution).

For the gamma distribution, the coefficient of variation is determined [1]:

$V=\sqrt{2/k}$,

Where k is the distribution parameter characterizing the skewness and kurtosis.

For k>1, the failure rate increases, fork<1 it decreases. When k<1, the gamma distribution is considered a convenient characteristic of the time of occurrence of system failures during the run-in period. However, in the region of positive values of k, the coefficient of variation has a value of V>1 (at k=0.9), which indicates a significant scattering of statistical data.

The gamma distribution has a certain relationship with the Weibull distribution. Thus, in the range of values of the gamma distribution parameter k>1, it was established [3] that if each of the elements of the system is characterized by a gamma distribution, but the parameters of these distributions fluctuate somewhat from element to element, then with a sufficiently large number of elements, the distribution of elements, the distribution of the failure-free time works are well approximated by the Weibull law. In addition, the gamma distribution also has exponential distributions in the composition [1].

For equipment that differs in design, manufacturing technology or operating conditions, the sample values of the coefficient of variation can naturally differ. To assess the significance of the difference between the Gamma distribution of V1 and V2, the Student's t-test is used [4].

$$t=\frac{|V_1-V_2|}{\sqrt{\frac{(V_1)^2}{2n_1}+\frac{(V_2)^2}{2n_2}}} \qquad (6)$$

Where, n_1 and n_2 are the volumes of the cutouts.

At t>3, the difference in the values of V_1 and V_2 is not significant.

Along with this, the above values of the coefficient of variation, characteristic of the considered distribution laws, make it possible to make a decision on the need to check the quality and homogeneity of statistical information, in cases of a noticeable discrepancy between the sample values of V and the expected ones.

Systematization of the accumulated statistical data on failures of oilfield equipment allows us to formulate some recommendations on the characteristic values of the coefficient of variation V and the area of their use in relation to mechanical systems. Cases with values V≤0.3÷0.35, obtained from the operation data, are quite good in terms of the stability of the developments of the research objects.

At V=0.35÷0.5, the obtained results can be considered satisfactory. Values V>0.5 indicate, first of all, the possibility of a predominance of sudden failures in the flow, which for mechanical systems is a consequence of the undeveloped design, instability of manufacturing quality, or violations of the equipment operation modes.

As an example, the table shows the results of processing statistical data on failures of the U8-6M mud pump, breaking them down into two subsystems (samples), taking into account the categories of failures according to the nature of their manifestation [5].

	m	V
Exponential law	1,00	1,00
	1,10	0,90
	1,20	0,84
	1,30	0,78
	1,40	0,72
	1,50	0,68
	1,75	0,58
	2,00	0,52
Law Reel	2,25	0,47
	2,50	0,43
	2,75	0,39
	3,00	0,36
	3,25	0,33
Normal law	3,50	0,31

No. subsystem (sample)	**MTBF T**	**Standard deviation, h σ**	**The coefficient of variation V**	**Coefficient Law of distribution of time between failures of variation**
I	51,46	50,92	0,99	Exponential
II	687,50	184,65	0,27	Normal

The first sample included mainly parts and assembly units of the hydraulic part of the pumps, which have a low level of reliability and are characterized by sudden failures due to hydraulic erosion of the working surfaces of parts with a high intensity of development. The second sample included parts and assembly units of the drive part of the pumps, which have a high level of reliability and are characterized by gradual failures.

As can be seen from the table, the obtained laws of distribution of the time between failures of subsystems and the values of the coefficients of variation correspond to the noted categories of failures, and the evaluation of the Student's criterion according to dependence (6) confirms that the difference in the values of the coefficients of variation of the subsystems is significant t=7.662.

Thus, the results of estimating the coefficient of variation not only make it possible to speed up the analysis of statistical data, but also to obtain a preliminary conclusion on the nature of failures, the maturity of the design and manufacturing technology.

References

Штовба С.Д. Проектирование нечетких систем средствами MATLAB. – М.: Горячая линия – Телеком, 2007. – 288 с.

Лапач С.Н., Радченко С.Г. Регрессионный анализ в условиях неоднородности факторного пространства //

Aliev G.S., Aliev E.F., Akhmedova I.V., Gulieva G.A., Ibragimova L.A. Physico-chemical methods for the study of oil sludge. Eurasian Union of Scientists. Series: biological and chemical sciences, No. 4 (97)/2022, volume 1, pp. 6-9.Бабаев С.Г.. Васильев Ю.А. повышение надежности оборудования, применяемого для бурения на нефть и газ. М., Машиностроение ,1972, 162 с.

G.S.Aliyev, R.N.Najiyeva. Calculation method for variable kinetic and diffusion equation coefficients for adsorption of sulfanol in the nonlinear isotherm region. Azerbaijan Chemical Journal, 2019, №3, p. 47-52.

Katsev P.G. Statistical methods for studying cutting tools. M., Mashinostroenie, 1974, 231 p.

Kostochkin V.V. reliability of aircraft engines and power plants. M., Mashinostroenie, 1976, 248 p.

Lapach S.N., Radchenko S.G. The main problems of building regression models // Mathematical machines and systems. - 2012. - No. 4. – S. 125–133.

Morozov E.A. Investigation of the properties of a hard-alloy inner coating obtained by laser cladding. Sovremennye problemy nauki i obrazovaniya. - 2015. - No. 2-2.

V.A.Majidzade; G.S. Aliyev; S.P. Javadov; A.Sh.Aliev, S.D.Dadashova; DB Tagiev, “Mathematical modeling and optimization of the formation of functional thin films”, Matem. Modeling, 2022, Vol 34, No. 6, pp.111-119.Г. С. Алиев, Х. М. Рустамли, Х. Ш. Гаджиахмедзаде, “Расчет кинетических и диффузионных коэффициентов процесса адсорбции поверхностно-активных веществ в нефтеносных пористых породах”, *Матем. моделирование*, 34:11 (2022), 35–47

Лапач С.Н., Радченко С.Г. Основные проблемы построения регрессионных моделей // Математичні машини і

Морозов Е.А. Исследование свойств твердосплавного внутреннего покрытия, полученного лазерной наплавкой //