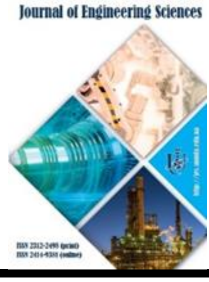
JOURNAL OF ENGINEERING SCIENCES

Volume 11, Issue 1 (2024)

Frolov M., Tanchenko S., Ohluzdina L. (2024). Parameter estimation of the Weibull distribution in modeling the reliability of technical objects. Journal of Engineering Sciences (Ukraine), Vol. 11(1), pp. A1–A10. https://doi.org/10.21272/jes.2024.11(1).a1



Parameter Estimation of the Weibull Distribution in Modeling the Reliability of Technical Objects

Frolov M.*[0000-0002-1288-0223], Tanchenko S.[0000-0002-1954-015X], Ohluzdina L.[0000-0002-1771-8612]

National University "Zaporizhzhia Polytechnic", 64, Zhukovskogo St., 69063 Zaporizhzhia, Ukraine

Article info:
Submitted: October 12, 2023
Received in revised form: February 20, 2024
Accepted for publication: February 29, 2024
Available online: March 1, 2024

***Corresponding email:**
mfrolov@zp.edu.ua

Abstract. The article discusses one of the most widely used distribution laws for reliability analysis – Weibull distribution. It describes a wide range of processes for all stages of the life cycle of technical objects, including yield stress of steel distribution and failures in the reliability theory regarding the wide range of technical objects (e.g., metal cutting tools, bearings, compressors, and wheels). A significant number of works are devoted to evaluating distribution law parameters based on empirical data in search of the most precise one, ignoring the probabilistic character of the parameters themselves. Parameters may have a relatively wide confidence range, which can be considered the parameter estimation error compared to biases of parameters estimated by different methods. Moreover, many approaches should be used for certain selection volumes, including comprehensive calculating procedures. Instead, this paper suggested and statistically confirmed a universal simplified approach. It demands a minimal set of data and connects the shape and scale parameters of the Weibull distribution with the variation coefficient as one of the leading statistical characteristics. This approach does not demand variational sequence arrangement. Nevertheless, it is supposed to be quite efficient for the engineering practice of reliability analysis. The adequacy of the results was confirmed using generated selections analysis and experimental data on cutting tool reliability. Within the achieved results, it was also demonstrated that the variation coefficient reflects not only selection stability and variable volatility degree, which are its main aim, but the cause of failure as well.

Keywords: cutting tool life, least squared estimation, maximum likelihood estimation, confidence interval, variation coefficient, bias, empirical data.

1 Introduction

One of the most essential tasks solved in engineering and industry is providing reliable, uninterrupted operation of objects designed, manufactured, or operated objects. Ones can be united by the general term "technical object". This especially applies to objects subject to friction and wear [1]. One of the most illustrative examples of such objects is metal cutting tools, which are related mainly to friction and wear [2].

Simultaneously, the cutting tool is a critical part of any CNC machining, and its reliability directly affects the efficiency and stability of the machining process in general [3]. Due to the tool wear, the machine outage reaches up to 25 % of the CNC machine outage. Moreover, the tool and retooling cost reaches up to 12 % of the cost per part [4]. If the cost of the machined part is broken due to a cutting tool failure during the machining process or failure of some vital part during exploitation will be additionally considered, the need to control and ensure the reliability of the technical object becomes apparent.

To simulate the reliability of any technical object, including prediction of the failure probability or that of the no-failure operation during a certain time, prediction of the no-failure time, determination of the mean time between failures or time to failure, and planning of the test programs, it is necessary to establish what kind of theoretical distribution law describes available empirical (selective) data, referring to the operation time to failure. In other words, it is necessary to establish a correspondence between the empirical and the theoretical distribution laws and evaluate the parameters of such laws based on the empirical data [5]. From the formal point of view, this correspondence can be established by Pearson's chi-square, Kolmogorov-Smirnov type, and other criteria. However, it is not only the formal side that is important – namely, how close the form of the theoretical distribution law is to the available empirical data, but also physical interpretation, connecting the type of theoretical distribution law and its parameters with the process mechanism [6].

Different theoretical distribution laws can be used to describe the same empirical data [7, 8]. In addition, one distribution law can be a particular case of another. In reliability analysis and modeling, it is essential to consider the period of the life cycle at which the object operates: wear-in, normal operation, or wear-out, which affects, first of all, the nature of the dependence between time and failure rate. The Weibull distribution is one of the most widely used distribution laws [8]. This law, or rather a family of distributions, describes a wide range of processes, such as yield stress of steel distribution [6, 7] and failures in the reliability theory.

2 Literature Review

The Weibull distribution probability density is described as the function of time t as follows:

$$f(t) = \frac{\alpha}{\beta^{\alpha}} t^{\alpha-1} exp\left[-\left(\frac{t}{\beta}\right)^{\alpha}\right], \quad (1)$$

where $\alpha > 0$ and $\beta > 0$ are the shape and scale parameters, respectively.

Some basic general characteristics of the Weibull distribution are shown in Table 1.

Table 1 – General characteristics of the Weibull distribution*

Name	Value
Argument	$0 \le t \le +\infty$
Mathematical expectation	$\mu = \beta \cdot \Gamma(1 + 1/\alpha)$
Variance	$\sigma^2 = \beta^2 \cdot \Gamma(1 + 2/\alpha) - \mu^2$
Failure rate	$\lambda(t) = \frac{\alpha}{\beta^{\alpha}} t^{\alpha-1}$
Median	$Me = \beta ln(2)^{1/\alpha}$
Mode	$Mo = \beta\left(\frac{\alpha-1}{\alpha}\right)^{1/\alpha}$

* $\Gamma(X)$ – Eiler's general gamma function.

The Weibull distribution was first applied in the aerospace and automotive industries. It is employed in electric power, nuclear, medical, dental, and components such as bearings, compressors, cables, and wheels. Chemical, oil, instrumentation, electronics, and railroads are the newest converts to Weibull [9]. It suggests a great variety of forms depending on the shape factor:

1) for $\alpha < 1$, the failure rate is a decreasing function, and the law describes the wear-in phase;

2) at $\alpha = 1$, the failure rate is constant, and the distribution goes exponentially and describes sudden failures – normal operation phase. However, it is possible only for the objects that are not subjected to wear [10];

3) for $\alpha > 1$, the failure rate is an increasing function, and the distribution law describes failures associated with wear;

4) if $1 < \alpha < 2$, the failure rate is a convex function bounded above;

5) if $\alpha = 2$, the failure rate is a straight line, and the distribution law coincides with Rayleigh one;

6) if $\alpha > 2$, the failure rate is the concave curve that is not bounded from above.

A traditional and most common method for estimating Weibull distribution parameters is the least squared estimation (LSE) [6], which is considered classical. Empirical data are approximated by the linear function of empirical time t_u, corresponding to interval u:

$$Y = b_o + \hat{\alpha} \cdot X; \quad (2)$$

$$\begin{cases} X = \ln(t_u); \\ b_o = -\hat{\alpha} \cdot ln(\hat{\beta}); \\ Y = ln\left[ln\left(\frac{1}{1-\hat{F}(t_u)}\right)\right], \end{cases} \quad (3)$$

where equations $\hat{\alpha}$, $\hat{\beta}$ – empirical shape and scale parameters, respectively; $\hat{F}(t_u)$ – empirical cumulative proportionate failure frequency of hitting the interval u.

Any distribution law is restored by constructing an interval variational series based on the sample of volume N. In this case, neither shape nor scale parameters are known (they should be found by estimation analysis).

Contrarily, the comparison of different estimation methods is based on simulation experiments when selection is generated with known theoretical parameters, which are then compared with ones obtained from the selection.

Many works are devoted to numerous methods for determining shape and scale parameters that differ by complexity, aiming to find the most "precise" in the given conditions. Moreover, many proposed methods are based on rather complicated calculation algorithms.

Let's consider some of them, representing a far from exhaustive list. The work [11] compares the Bayesian method with maximum likelihood estimates (MLE), indicating that the first is better for small samples.

The article [12] proposes a new method based on transforming the cumulative distribution function constructed as a mapping from the value of the random variable and its corresponding cumulative distribution probability to the scale parameter. The proposed method's accuracy gives more accurate results compared with the LSE and the MLE.

The paper [13] estimates parameters utilizing the MLE approach, the maximum product spacing (MPS) technique, and the Bayesian assessment approach. Monte Carlo Simulation was utilized to compare the three techniques above. Bayes estimators have been computed using the Lindley Approximation approach. It was found that Bayesian estimation behaves better.

Work [14] compares the following methods: the L-moment estimator (LM), MLE approach, moment estimation (MoE), LSE, the modified MLE (MMLE), modified MoE (MMoE), and the maximum product spacing (MPS), for example, for some of them the best-fit area is defined.

The LM method is almost always close to the best estimation method, including the scenario where a significant shape parameter $\alpha > 6$ is coupled with a small sample size. The MLE best fits when the shape parameter $1.5 \le \alpha \le 4.0$, even for a small sample of $n = 10$.

The MPS estimator is better than others when $0.5 < \alpha < 1.5$.

For large $\alpha > 6.0$ and sample size $n > 50$, the best method is MMLE.

Instead, the paper [15] analyzes the following methods: Justus empirical method, method of moment estimation (MoE), graphical method, energy pattern factor method (EPM), energy trend method, and MLE, using manta-ray foraging optimization (MRFO) method as metaheuristic algorithm, indicate that MLE method as well as EPM are the least successful. Paper [16] has almost the same results regarding MLE and proposes its iterative method.

The work [17] is dedicated to comparing the following estimation methods: MLE, MME, and median rank regression (MRR), and does not find any sufficient difference between the above methods.

Article [18] indicates that classical parameter estimation methods, e.g., MLE and LSE, do not provide robust estimates for large and heavily censored samples and develop their own novel parameter estimation method based on information extracted from censored observations and evaluate the accuracy and robustness of the proposed method through a numerical experiment.

The problem of the above works, including [14, 17], is that they are searching for the most “precise” method comparing the theoretical value of parameters and ones obtained from simulation (numerical experiment) based on the generated selection. Instead, in real-life tests, distribution parameters are unknown, and as with any random values, interval estimation is in the form of a confidence interval (CI) that covers unknown parameters with a given confidence probability should be considered.

Thus, the parameters can take any value inside the CI. Moreover, the confidence interval can be considered an error in determining the parameter’s actual value.

The upper and lower confidence limits of the shape parameter for confidence probability 0.95 from [6] can be found as follows:

$$\alpha_U = \hat{\alpha} \cdot exp\left(\frac{1.53}{\sqrt{N}}\right); \quad (4)$$

$$\alpha_L = \hat{\alpha} \cdot exp\left(-\frac{1.53}{\sqrt{N}}\right). \quad (5)$$

The length of the shape parameter CI:

$$D_a = \alpha_U - \alpha_L. \quad (6)$$

As for the scale parameter, its upper and lower confidence limits for confidence probability 0.95 can be defined as follows [6]:

$$\beta_U = \hat{\beta} exp\left(\frac{2.058}{\hat{\alpha}\sqrt{N}}\right); \quad (7)$$

$$\beta_L = \hat{\beta} exp\left(-\frac{2.058}{\hat{\alpha}\sqrt{N}}\right). \quad (8)$$

The length of the scale parameter CI, respectfully:

$$D_b = \beta_U - \beta_L. \quad (9)$$

The results of the shape parameter CI calculation by formulas (4)–(6), based on the data given in [17] for different methods and samples of size N, and their comparison with the biases of the recovered shape parameters relative to the theoretical value ($\alpha = 5.0$) are shown in Figure 1.

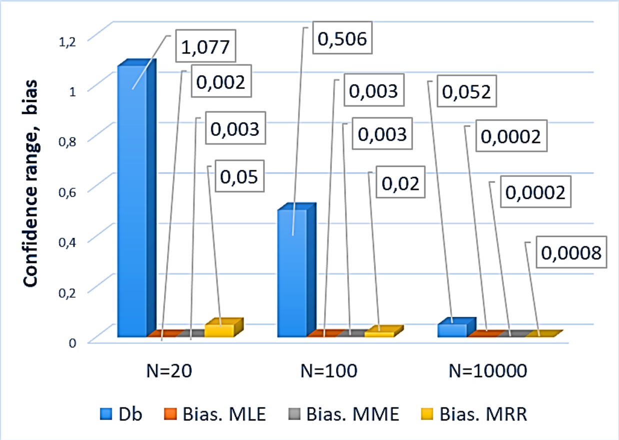


Figure 1 – Comparison of the CI length with biases of the shape parameter for different estimation methods and sample size for $\alpha = 5$ and $\beta = 90$

The same trend persists for small values of the shape and scale parameters.

Figure 2 compares the scale parameter CI length and biases for the above estimation methods [17]. The formulas (7)–(9) calculate CI length for $\alpha = 2$ and $\beta = 2.5$.

As seen from Figures 1, 2, CI lengths are orders of magnitude higher the biases obtained for different estimation methods. Confidence interval as well as biases completely depends on the sample size and it is seen in formulas (4), (5), (7), (8). Thus, from a statistical point of view, it does not matter how “precise” the estimation method is, since the parameter value is inside the confidence interval – in real life, the parameter’s theoretical values are unknown, besides simulation experiments. The method should be adequate to the conditions and accuracy defined by the CI. In addition, the above methods demand rather complicated calculation algorithms.

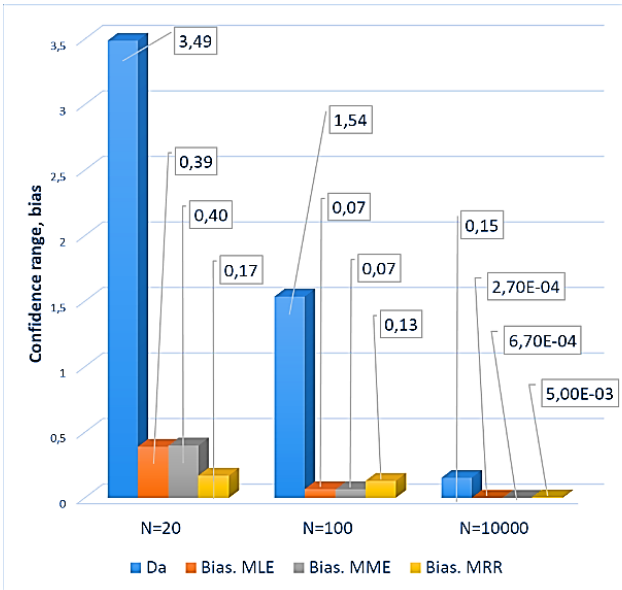


Figure 2 – Comparison of the CI length with biases of the scale parameter for different estimation methods and sample size for $\alpha = 2$ and $\beta = 2.5$

Thus, in this article, for the case when there are good reasons, based on the broad engineering practice, to assume Weibull distribution in reliability analysis, we would like to achieve the following goal: develop a universal, simple, and precise method for distribution parameters estimation with a minimal set of data, giving adequate results from the statistical point of view.

To achieve this aim, the following objectives should be solved:

1. Development of a methodology for conducting a simulation experiment, during which a sample is generated, following the Weibull distribution law, with known parameters, imitating the time to failure of a technical object.
2. Develop a methodology for estimating distribution law parameters based on basic statistical indicators of a sample; confirm its statistical adequacy and assess the possibility of its application at different sample sizes and theoretical values of distribution parameters.
3. Establishing a statistical relationship between defined distribution parameters and statistical characteristics of the sample to ensure the relevance of the methodology developed.
4. Evaluation of the obtained results by comparing them with similar ones obtained by a known method taken as a reference, including using confidence intervals.
5. To perform the transition from simulation modeling to the results of physical reliability testing of metal-cutting tools as an example of application of the methodology in real conditions for technical objects subject to wear and aging.
6. Draw conclusions from the results obtained.

3 Research Methodology

The research methodology for this article is based on the mathematical and statistical analysis of the distribution laws of random variables in the context of the reliability of technical objects. The adequacy of the resulting dependencies is checked in the following ways.

Firstly, setting up a simulation experiment. During the experiment, a random sample of size N is generated, simulating the time t_i to failure of a certain technical object. The sample obeys the Weibull distribution law with shape and scale parameters α and β, respectively.

The sample is generated using the built-in functions of Microsoft Excel:

$$\langle t\rangle = \langle\beta\rangle \cdot \{-\ln[RAND(A)]\}^{-\frac{1}{\langle\alpha\rangle}}.$$

For the generated sample, the range A, the number of partitioning intervals k, and the length of the interval Δm [19] are determined, respectively:

$$A = t_{max} - t_{min}; \qquad (10)$$

$$k = 1 + 3.32 \cdot \lg(N); \qquad (11)$$

$$\Delta m = \frac{A}{k}. \qquad (12)$$

It should be noted that the number of intervals is adjusted to an integer such that it is odd and $k \geq 5$.

Second, the boundaries of the intervals are determined – left and right, respectively, where u is the number of the interval ($u = \overline{1, k}$), as well as the middle of the interval $\bar{t}_u$:

$$t_{u-1} = t_{min} + \Delta m \cdot (u-1); \qquad (13)$$

$$t_u = t_{min} + \Delta m \cdot u; \qquad (14)$$

$$\bar{t}_u = \frac{t_{u-1} + t_u}{2}. \qquad (15)$$

After that, the number of sample elements that fall into each of the intervals is designated m_u, the relative frequency p_u, the empirical probability density $\hat{f}(\bar{t}_u)$ and the empirical cumulative failure frequency $\hat{F}(t_u)$ are determined:

$$p_u = \frac{m_u}{N}; \qquad (16)$$

$$\hat{f}(\bar{t}_u) = \frac{p_u}{\Delta m}; \qquad (17)$$

$$\hat{F}(t_u) = \hat{F}(t_{u-1}) + p_u. \qquad (18)$$

Formulas determine the distribution parameters (2) and (3) with the imposition of confidence intervals defined by formula (4).

The correspondence of the empirical distribution to the theoretical one is checked by a nonparametric best-to-fit test of the Kolmogorov–Smirnov type, adjusted for parameter estimates in the theoretical law instead of their actual values [20–22]. For a confidence probability of

0.95, the hypothesis about the form of the distribution law is not rejected if the condition is met:

$$\sqrt{N} \cdot sup\left|\hat{F}(t_u) - F(t_u)\right| \leq 0.895, \qquad (19)$$

where $F(t_u)$ is the value of the integral distribution function of the Weibull law, which is defined in the following way:

$$F(t) = 1 - exp\left[-\left(\frac{t}{\beta}\right)^{\alpha}\right]. \qquad (20)$$

Consequently, it needs to be determined whether the theoretical parameter values fall within the confidence range defined based on the proposed equations. At the same time, the LSE method will be taken as a reference one.

Finally, the results obtained according to the developed dependencies with the results of full-scale real-life independent experiments on studying the metal-cutting tool life and its reliability should be compared.

4 Results

The variation coefficient [7] is a significant indicator specifying stability or the variable volatility degree of the selection and, hence, is connected with the time to failure. The following equation can define it:

$$V = \frac{S}{\bar{T}}, \qquad (21)$$

where S is the standard deviation of the time to failure and $\bar{T}$ is the mean time to failure (MTTF).

For Gamma distribution [7], which also has form and scale parameters r and λ_0, respectively, their values are directly depend on the variation coefficient:

$$r = \frac{1}{V^2}; \qquad (22)$$

$$\lambda_0 = \frac{r}{\bar{T}}. \qquad (23)$$

Since the conformity of failures to the Gamma distribution corresponds to the cumulative failure scheme, physically, r is the number of damages leading to failure, and λ_0 is the number of damages per unit time to failure.

A similar dependence can be obtained for the Weibull distribution using equations from Table 1 of general characteristics. 1, replacing them with the corresponding statistical estimates. As a result, the following outcome is obtained:

$$V = \sqrt{2 \cdot \alpha \cdot \frac{\Gamma\left(\frac{2}{\alpha}\right)}{\Gamma^2\left(\frac{1}{\alpha}\right)} - 1}, \qquad (24)$$

where $\Gamma(X)$ is the general Gamma function, the computational solution of which can be obtained by in-built functions of Microsoft Excel soft:

$$\langle\Gamma(\mathrm{X})\rangle = EXP[GAMMALN(\langle X\rangle)]. \qquad (25)$$

As seen from equation (25), the variation coefficient is the function of the shape parameter only and, thus, vice-versa. The reverse function of equation (25) – shape factor from variation coefficient – can be approximated by the following empirical regression:

$$\hat{\alpha} = b_0 V^{b_1} \qquad (26)$$

For the wide range of α – namely from 0.2 to 25, the following regression was found, having a very high determination factor $R^2 = 0.999968$, which is the empirical estimation of the shape parameter:

$$\hat{\alpha} = V^{-1.09}. \qquad (27)$$

Since the estimate criterion for the adequacy is the fact that the actual value belongs to the confidence interval, for the range of $\alpha \leq 2.0$ that covers the portion of the Weibull distribution shape parameter domain where it describes wear-in, sudden failures, and gradual wear-out [7], equation (27) can be further simplified and take the following form:

$$\hat{\alpha} \approx V^{-1}. \qquad (28)$$

Notably, when using formula (28), finding an estimate of the shape parameter within the confidence interval is limited by the sample size. The condition for which can be obtained from equations (5), (27), and (28) for confidence probability 0.95:

$$N \leq \frac{295}{\ln^2(V)}. \qquad (29)$$

The maximum sample size varies widely, having a breakpoint at $V = 1$: from 81 at $V = 0.15$ to 26474 at $V = 0.9$ and from 32474 at $V = 1.1$ to 614 at $V = 2$.

This is sufficient for research on many technical objects and metal-cutting tools. A more accurate approximation (28) should be used in other cases.

The graph of the dependence of the maximum sample size on the coefficient of variation is shown in Figure 3.

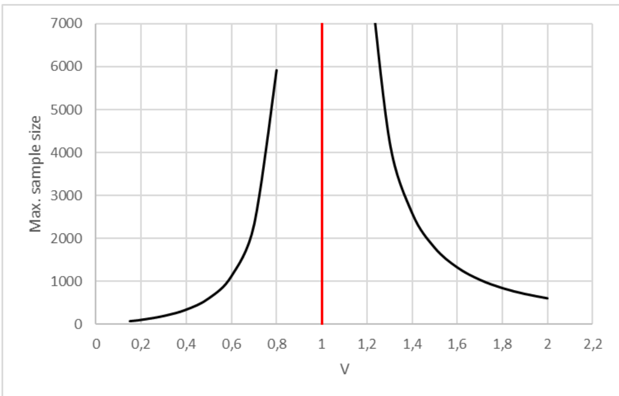


Figure 3 – Maximum sample size vs variation coefficient

As an example, for bearings for which the well-known estimate of the shape parameter is $\alpha = 1.5$, the maximum sample size will be estimated at 1700 pcs from the above reasons.

The scale factor can be found using general characteristics from Table 1, namely math expectation and median. Since the MTTF $\bar{T}$ is a statistical estimate of math expectation μ, from Table 1 and equation (27), the scale factor can be found as in equation (30):

$$\hat{\beta} = \frac{\bar{T}}{\Gamma(1+V^{1.09})}. \qquad (30)$$

The second approach is to find the scale factor, knowing the median of the selection:

$$\hat{\beta} = \frac{Me}{(\ln 2)^{V^{1.09}}}. \qquad (31)$$

5 Discussion

Discussion of the obtained results is based on analysis of sets of 30 samples with the size of $N = 100$ elements each according to the above methodology. Typical results are shown below.

First of all, the results of the analysis of the correlation relationship between the values of the shape parameter, which varies within wide limits (from 0.4 to 6.0), namely the theoretical one, determined by the LSE method and by equation (27), as well as the coefficient of variation V, were obtained. Figure 4 shows the correlation matrix indicating the correlation coefficients R and the nature of the relationships.

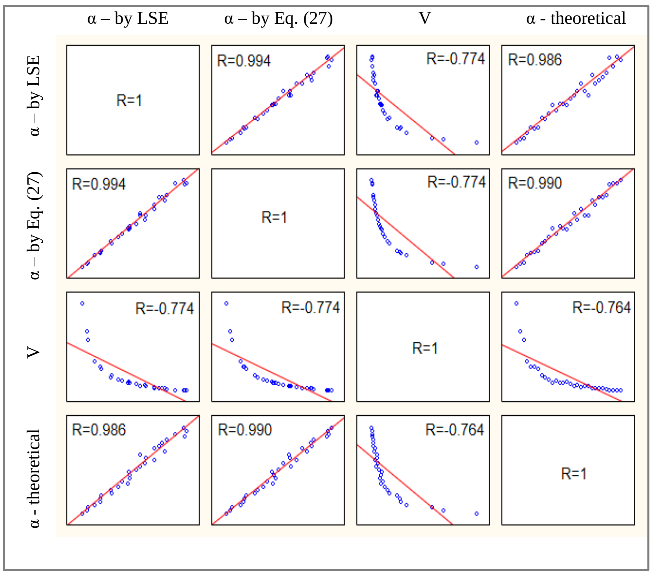


Figure 4 – Correlation matrix between results of different approaches to the form parameter α estimation and variation coefficient V

The provided data demonstrate the strength of the relationship from significant to very significant, which indicates the presence of an actual strong relationship between the values of the variation coefficient and the shape parameter, as well as the equivalence in this sense of the reference and proposed methods.

For generated samples with predetermined theoretical parameters – varying shape parameter and scale parameter $\beta = 100$, their empirical shape parameters were estimated by the LSE method and equations (27) and (28). Belonging to the Weibull distribution for all the samples was confirmed by both Pearson chi-square and Kolmogorov–Smirnov type goodness-of-fit criteria described above. The theoretical values, values obtained by equations (27) and (28), as well as those estimated by the LSE method, happened to be inside the exact confidence intervals, except those obtained by equation (28) for $\alpha > 2.5$ (that was indicated above).

Some estimation results are presented in Figure 5, where theoretical shape factors are shown as the sample identification.

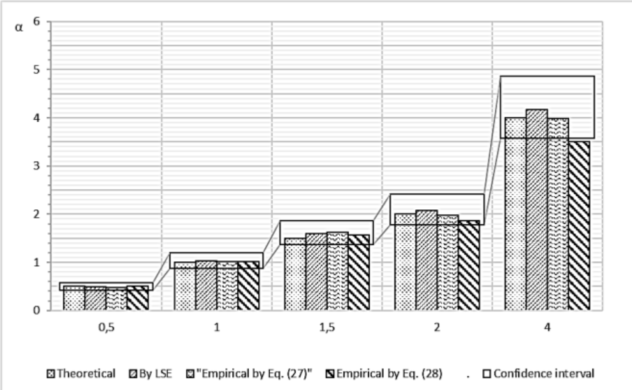


Figure 5 – Comparison of the form parameter estimation methods within CI

Thus, equation (27) can be applied to estimate the shape parameter and equation (28) with the above limitations. The accuracy of the shape parameter estimates will be analyzed based on its deviation from the theoretical value to the CI width:

$$E_\alpha = \frac{|\hat{\alpha} - \alpha|}{D_\alpha}. \qquad (32)$$

The share of the best results is shown in Figure 6.

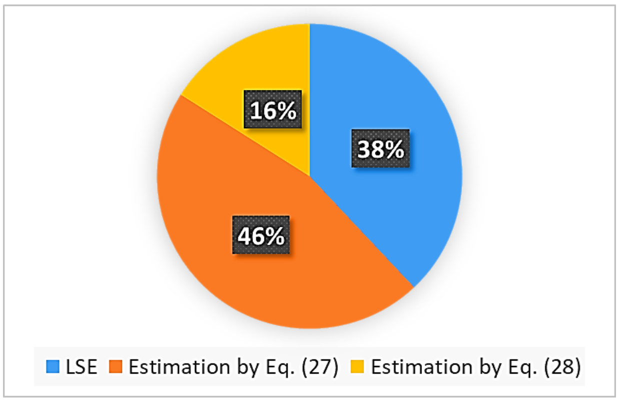


Figure 6 – The most accurate shape for parameter estimation approaches

It can be seen that estimates based on the coefficient of variation give the best result in a much greater number of cases, while shape parameter changes in a wide range from 0.4 to 6.2. However, when the shape parameter increases and takes the value $\alpha > 7$, the number of estimates determined by equation (28) that fall outside the confidence interval increases.

Correlation matrices will be constructed to analyze the approaches mentioned above for the scale parameter estimation (characteristic time), with one of them featuring a wide variation in the theoretical value of the scale parameter ranging from 5 to 280. The shape parameter remains constant at $\alpha = 2$ (Figure 7a). For the other matrix, the scale parameter will be held constant ($\beta = 120$). The shape parameter varies from 0.4 to 6.2 (Figure 7b).

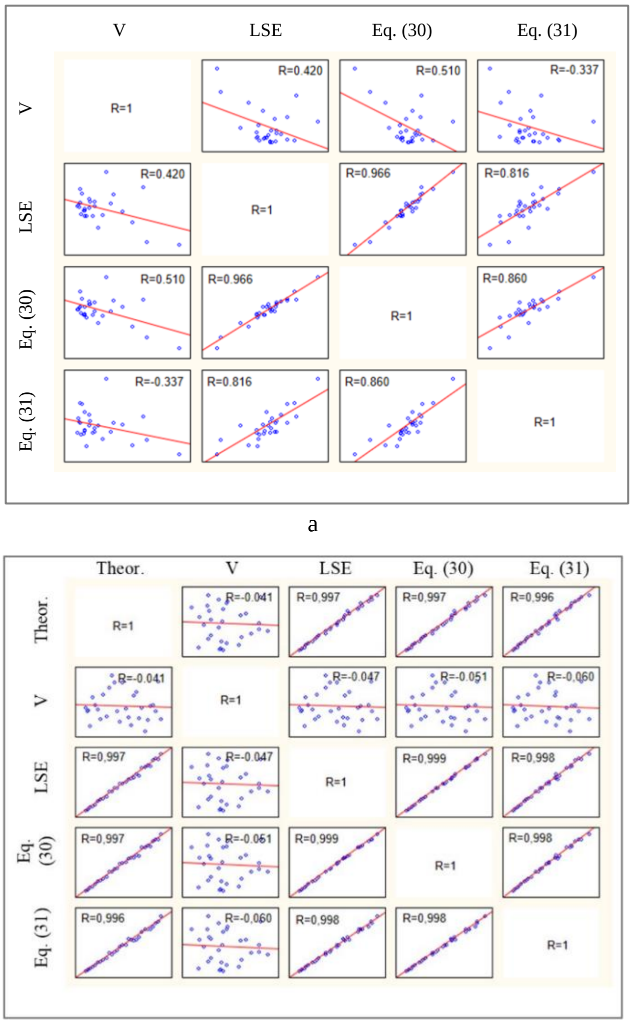


Figure 7 – Correlation matrix between results for different approaches to estimate the scale parameter β and variation coefficient V: a – α = const; b – β = const

As can be seen from Figure 7a, there is a highly significant strength of relationships between the theoretical value of the scale parameter and its estimates by the LSE reference method, by equation (30) based on the MTTF, and by equation (31) based on the sample median time to failure. The relationship between the considered estimation methods is also very significant. There is no relationship between the variation coefficient and the considered estimation methods, which is quite natural due to the presence of exclusively random fluctuations in the values of the empirical variation coefficient when α = const.

As for the strength of the relationship when β = const and α is variable, Figure 7b shows a very significant strength of the relationship between the values of the scale parameters obtained by different estimation methods. The influence of the variation coefficient is present, as predicted by equations (30) and (31), and its strength is estimated from moderate to noticeable, which is also explained by a significant random component.

Comparison of different scale parameter estimates with confidence intervals also shows that they, together with the theoretical value, are within the limits of the intervals.

Figure 8a shows the results of estimations with variable shape parameter β and a constant scale parameter $\alpha = 2$. Figure 8b shows the estimates within the confidence intervals with a constant scale parameter $\beta = 120$ and variable shape parameter α.

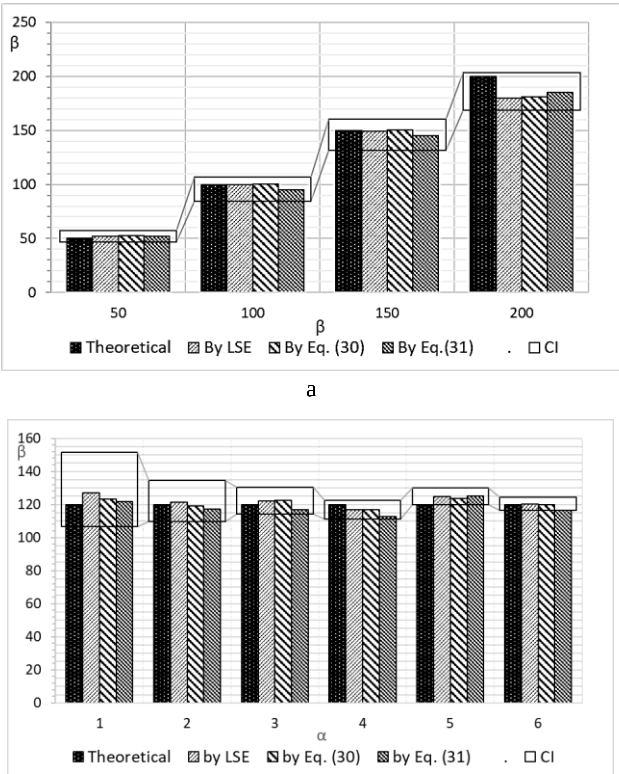


Figure 8 – Comparison of the scale parameter estimation methods within CI: a – α = const; b – β = const

The above gives grounds for asserting that equations (30) and (31) can be used for estimation based on empirical data within a wide range of the estimated parameters.

The accuracy of the assessment for different methods will be determined by a criterion similar to (32):

$$E_\beta = \frac{|\hat{\beta}-\beta|}{D_b}. \quad (33)$$

The weighted mean values are shown in Figure 9.

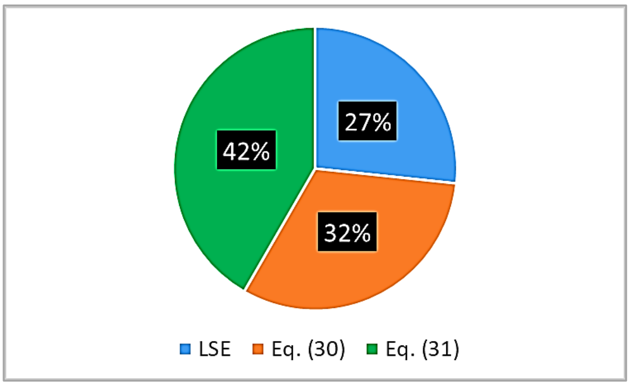


Figure 9 – The most accurate scale for parameter estimation approaches

As can be seen, the proposed approaches give a better result in more cases than the reference one.

The transition from simulation to the results of real-life experiments will be carried out by using the independent experimental data obtained for a wide range of metal cutting tools, the operating conditions of which are significantly different, such as cutters, drills, multi-flute drills, screw taps, and threading dies – totally 43 lots. The data reflect the dependence of the γ-percent time-to-failure factor G upon the following variation coefficient:

$$G = \frac{T_\gamma}{\bar{T}}, \quad (34)$$

where T_γ is γ-percent time-to-failure when the probability of failure-free operation is equal to γ. For the Weibull distribution, it is equal to

$$T_\gamma = \beta \cdot [-ln(\gamma)]^{1/\alpha}. \quad (35)$$

According to the experiment's conditions, the no-failure probability was ensured at $\gamma = 0.9$.

After applying equations (27), (30), (34), and (35), a theoretical model of the above G factor can be obtained:

$$G = \frac{[-\ln(\gamma)]^{V^{1.09}}}{V^{1.09}\Gamma(V^{1.09})}. \quad (36)$$

A comparison of the empirical data and simulation results by equation (36) is represented in Figure 10.

Correlation analysis of variation coefficient G, empirical data, and ones obtained by theoretical model equation (36) shown in Figure 11 indicates sufficient strength of the relationship between all these factors.

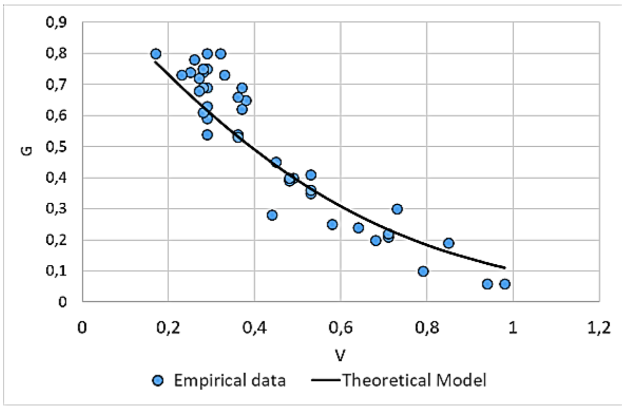


Figure 10 – Dependence of the γ-percent time-to-failure factor G upon the variation coefficient

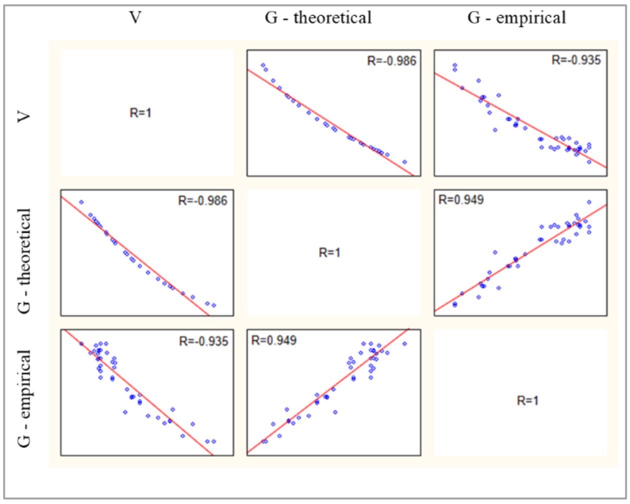


Figure 11 – Correlation matrix between variation coefficient V, theoretical and empirical values of the G-factor

Most importantly, the analysis of the theoretical model according to the Fisher criterion demonstrates its adequate reflection of empirical data with a confidence probability of 0.98, the estimated value of the Fisher criterion is $F = 2.29$, while the critical value is $F_{0.02,17,26} = 2.66$, i.e., $F < F_{0.02,17,26}$.

Thus, the validity of using the proposed approach to estimating the Weibull distribution parameters is confirmed by both simulation and real-life physical experiments on the reliability of the metal-cutting tools.

6 Conclusions

In literary sources, numerous methods of estimating the parameters of the Weibull distribution are proposed to find the most accurate one, many of which, among other things, require complex calculation procedures. The accuracy of parameter estimation is determined by proximity to the theoretical value of the generated sample, while the estimated parameters are random values, and in real-life conditions, the theoretical value is unknown and is not considered.

Simultaneously, the width of the confidence interval, which can be considered the parameter determination error, can be defined and is much broader than the estimation accuracy based on the generated sample. Thus, the search for the most accurate method solely based on the above indicator loses its meaning.

For the case when there are strong reasons, based on the broad engineering practice, to consider the subjection of the technical object failure probability to the Weibull distribution, a universal approach is proposed for estimating the distribution parameters based on typical indicators from descriptive statistics - variation coefficient and the MTTF or the median of the sample and which requires a minimal set of data and for which calculation procedures are as simple as possible. The proposed approach can be used for a wide range of sample sizes and parameter values. The values of the parameters obtained by the proposed approach are in the confidence interval and the theoretical value and determined by the LSE method, which is the most widespread and was taken as the reference. Thus, they can be used on equal terms.

The results show a strong correlation between the parameters of the Webul distribution and the variation coefficient as one of the main statistical indicators.

Since, as is known, the value of the form parameter of the Weibull distribution reflects the physical nature of the technical object failure. Thus, it can be concluded that the variation coefficient reflects not only the stability and degree of variability of the sample, which consists of the time to failure, but also the physical nature of the failures themselves during different stages of the life cycle of a technical object.

Statistical analysis of the proposed approach showed that it reflects a consistent pattern of a sample subject to the Weibull distribution and can be used to determine parameters within wide limits. The limitations of the simplified approach are also shown.

It is also shown that the approach for determining the distribution parameters by the variation coefficient gives more accurate results in a more significant number of cases than the reference one. Simultaneously, determining the scale factor by the selection median gives a more accurate result in the more significant number of cases than MTTF.

The possibility of using the proposed approach to estimate the parameters of the Weibull distribution is confirmed by simulation tests on generated samples and by independent results of experimental reliability studies of a wide range of metal-cutting tools that work in very different conditions. The proposed approach provides a result adequate to the results of experimental studies.

References

1. Tsyganov, V.V., Sheyko, S. (2023). Features of engineering the wear-resistant surface of parts with the multicomponent dynamic load. *Wear*, Vol. 494–495. 204255. https://doi.org/10.1016/j.wear.2022.204255
2. Tsyganov, V. V., Mokhnach, R. E, Sheiko, S. P. (2021). Increasing wear resistance of steel by optimizing structural state of surface layer. *Steel Transl.*, Vol. 51, pp. 144–147. https://doi.org/10.3103/S096709122102011X
3. Gao, T., Li, Y., Huang, X., Li, H. (2022). Turning tool life reliability analysis based on approximate Bayesian theory. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, Vol. 236(5), pp. 696–704. https://doi.org/10.1177/1748006X211043753
4. Lin, F., Zhang, H., Zhou, Y., Zhang, Z., Zhang, L. (2021) tool reliability modeling based on gamma process in multiple working conditions. *In: 2021 Global Reliability and Prognostics and Health Management, PHM-Nanjing 2021*. Nanjing, China, pp. 1–6. https://doi.org/10.1109/PHM-Nanjing52125.2021.9612765
5. Bracke, S., Radetzky, M., Rosebrock, C. (2021). Reliability engineering data analytics as the base of operations for maintenance planning: A cutting tool case study. *IFAC-PapersOnLine*, Vol. 54(1), pp. 1260–1265. https://doi.org/10.1016/j.ifacol.2021.08.151
6. Abernethy, R. B. (2006). *The New Weibull Handbook*. Dr. Robert. Abernethy (5th ed.), North Palm Beach, FL, USA.
7. Frolov, M. (2019). Variation Coefficient and some distribution laws in the context of cutting tools and other technical objects reliability modeling. *In: Advances in Design, Simulation and Manufacturing. DSMIE 2019*. Lecture Notes in Mechanical Engineering, pp 13–22. Springer, Cham. https://doi.org/10.1007/978-3-319-93587-4_2
8. Gaddafee, M., Chinchanikar, S. (2020). An Experimental investigation of cutting tool reliability and its prediction using Weibull and gamma models: A Comparative assessment. *Materials Today: Proceedings*, Vol. 24(2), pp. 1478–1487. https://doi.org/10.1016/j.matpr.2020.04.467
9. Kim, J. S., Yum, B.-J. (2008). Selection between Weibull and lognormal distributions: A comparative simulation study. *Computational Statistics & Data Analysis*, Vol. 53(2), pp. 477–485. https://doi.org/10.1016/j.csda.2008.08.012
10. Zhang, X., Wang, Y., Lu, D.(2021) Parameter optimization estimation based on mixed exponential Weibull distribution. *In: Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery. ICNC-FSKD 2020. Lecture Notes on Data Engineering and Communications Technologies*, Vol 88, pp. 1679–1686. Springer, Cham. https://doi.org/10.1007/978-3-030-70665-4_182
11. Alshenawy, R., Feroze, N., Tahir, U., Al-Alwan, A., Ahmad, H. H., Ali, R. (2022). On suitability of modified Weibull extension distribution in modeling product lifetimes and reliability. *Advances in Mechanical Engineering*, Vol. 14(11), pp. 1–16. https://doi.org/10.1177/16878132221136688

12. Xie, L., Wu, N., Yang, X. (2023). A minimum discrepancy method for Weibull distribution parameter estimation. *International Journal of Structural Stability and Dynamics*, Vol. 23(08), 2350085. https://doi.org/10.1142/S0219455423500852
13. Aljeddani, S. M. A., Mohammed, M. A. (2023). Estimating the power generalized Weibull Distribution's parameters using three methods under Type-II Censoring-Scheme. *Alexandria Engineering Journal*, Vol. 67, pp. 219–228. https://doi.org/10.1016/j.aej.2022.12.043
14. Akram, M., Hayat, A. (2014). Comparison of estimators of the Weibull distribution. *J Stat Theory Pract*, Vol. 8, pp. 238–259. https://doi.org/10.1080/15598608.2014.847771
15. Bulut, A., Bingöl, O. (2023). Analysis and comparison of Weibull parameters for wind energy potential using different estimation methods: A Case study of Isparta province in Turkey. *Electric Power Components and Systems*, Vol. 51(16), pp. 1829–1845. https://doi.org/10.1080/15325008.2023.2210574
16. Yang, X., Xie, L., Zhao, B., Kong, X., Wu, N. (2022). An iterative method for parameter estimation of the three-parameter Weibull distribution based on a small sample size with a fixed shape parameter. *International Journal of Structural Stability and Dynamics*, Vol. 22(12), 2250125. https://doi.org/10.1142/S0219455422501255
17. Nielsen, M. A. (2011). *Parameter Estimation for the Two-Parameter Weibull Distribution. BYU ScholarsArchive*. Brigham Young University, Provo, UT, USA.
18. Jiang, R. (2022). A novel parameter estimation method for the Weibull distribution on heavily censored data. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, Vol. 236(2). pp. 307–316. https://doi.org/10.1177/1748006X19887648
19. Jamieson, A. (1989). Computer search for the ideal histogram. *Annual Quality Congress Transactions*, Vol. 43, pp. 277–283.
20. Inglot, T., Ledwina, T., Ćmiel, B. (2019). Intermediate efficiency in nonparametric testing problems with an application to some weighted statistics. *ESAIM: Probability and Statistics*, Vol. 23, pp. 697–738. https://doi.org/10.1051/ps/2018022
21. Wang, F., Xu, R., Zhong, Z. (2011). Low complexity Kolmogorov-Smirnov modulation classification. *In: 2011 IEEE Wireless Communications and Networking Conference, WCNC 2011*. IEEE, Cancun, Mexico, pp. 1607–1611. https://doi.org/10.1109/WCNC.2011.5779375
22. Chen, X. X., Ge, S. L., Lin, M. (2011). An analysis of statistical techniques applying to multi-feature similarity comparison between Corpora. *Applied Mechanics and Materials*, Vol. 66–68, pp. 2323–2329. https://doi.org/10.4028/www.scientific.net/AMM.66-68.2323