Renewed empirical formulas for estimating Weibull distribution parameters

D.G. Asatryan [1,2]
[1] Institute for Informatics and Automation Problems of NAS RA, 0014, Armenia, Yerevan, P. Sevaki street, 1;
[2] Russian-Armenian University, 0052, Armenia, Yerevan, O. Emini street, 123

Abstract

The empirical formulas proposed in the literature for estimating the parameters of a two-parameter Weibull distribution, obtained using the equations of the moment method, are considered. It is noted that the formulas used to estimate the shape parameter take the form of various types of dependences on the coefficient of variation of the distribution. By modeling the empirical formulas selected for analysis, a comparative analysis of their errors relative to accurate numerical solutions of the moment method equations was carried out. A renewed empirical formula for the shape parameter is proposed. An approach to estimating the scale parameter is proposed, in which the empirical formula of the latter is reduced to the product of the standard deviation of the distribution by a power function of the coefficient of variation with an exponent equal to -1.027. The results of applying the updated empirical formulas to numerical data obtained by modeling a random sample from the Weibull distribution are presented. It is shown that the accuracy of the proposed empirical formulas is quite high.

Keywords: Weibull distribution, shape parameter, scale parameter, coefficient of variation, empirical formula, accuracy.

Introduction

The Weibull distribution has been widely used in various fields of science and technology in solving problems of assessing reliability and determining the parameters of products, devices, and phenomena, describing the distribution of particle sizes, when working with digital images, and many others. Therefore, the development of methods for assessing the adequacy of the model and the parameters of the Weibull distribution are of significant interest to researchers involved in measurements and subsequent processing of experimental data in areas of various profiles.

When solving any applied problems using statistical methods, the question arises about the applicability of the mathematical model under consideration for studying the properties and patterns in the measured data. In this case, one of the main actions is the choice of a method for statistical estimation of distribution parameters based on a set of data. The choice of method is based on the volume of data being processed, available computing resources, the required accuracy of estimates, and the possibility of obtaining an adequate interpretation of the final results.

Therefore, many methods for estimating the parameters of the Weibull distribution have been proposed in the literature. In addition to the classical methods of maximum likelihood and the method of moments, there are methods based on the least squares method, Bayesian, approximate, empirical, and graphical methods, etc. For example, in [1] eleven are considered, in [2] ten, and in [3] there are seven methods, including hybrid ones and their comparative analysis is carried out on specific experimental material. Neural networks are also used for this purpose.

The most common are the maximum likelihood method and the method of moments, which have some attractive properties of both a theoretical nature and possibilities for practical application. However, as is known, both methods lead to the solution of the corresponding systems of transcendental equations, which inevitably slows down the process of parameter estimation. Therefore, empirical estimation methods based on the method of moments, which, of course, are less accurate, but are very simple to implement, have become widely used. At the same time, in practice, empirical formulas are often used to estimate parameters not from the original measured data, but from available estimates of their average and root-mean-square values. The latter circumstance gives these empirical formulas additional attractiveness, since it facilitates the use of old data, their transfer to other specialists, and the conduct of comparative studies. Apparently, this also explains the increased interest of researchers observed in the scientific and technical literature, especially in the field of studying the characteristics of wind flows [1-14].

Note that various simple empirical formulas have been proposed in the literature for estimating, mainly, the shape parameter of the Weibull distribution. To estimate the scale parameter, there are only isolated cases of proposals using empirical formulas. Therefore, in articles, next to the empirical formula for the shape parameter, the original transcendental equation of the moment method for the scale parameter is often placed.

In this work, by modeling the empirical formulas selected for analysis, a comparative analysis of their errors relative to the accurate numerical solutions of the moment method equations was carried out. A renewed empirical power-type formula for the shape parameter is proposed and a new empirical formula for the scale parameter is obtained.

1. Empirical formulas for evaluating the Weibull parameters

The density of the two-parameter Weibull distribution has the form

$$f\left(x;\lambda,\eta\right)=\frac{\eta}{\lambda}\left(\frac{x}{\lambda}\right)^{\eta-1}exp\left[-\left(\frac{x}{\lambda}\right)^{\eta}\right], \qquad (1)$$

where $\eta > 0$ – shape parameter, and $\lambda > 0$ – scale parameter.

Estimation of distribution parameters (1) by the moment method is carried out using the formulas for mathematical expectation and dispersion, which are equal, respectively

$$\mu = \lambda\Gamma\left(1+1/\eta\right),$$
$$\sigma^2 = \lambda^2\left[\Gamma\left(1+2/\eta\right)-\Gamma^2\left(1+1/\eta\right)\right]^2.$$

From here, we obtain two transcendental equations for the parameters η and λ

$$\Gamma\left(1+2/\eta\right)/\left(\Gamma^2\left(1+1/\eta\right)\right)-1=\gamma^2, \qquad (2)$$

$$\lambda = \mu/\Gamma\left(1+1/\eta\right), \qquad (3)$$

where $\gamma = \sigma/\mu$ – coefficient of variation of the distribution, and $\Gamma(\cdot)$ – is the Gamma function.

The accuracy of solving equations (2) and (3) is limited by the magnitude of the error in the numerical estimation of the Gamma function, which, with eight terms of approximation by orthogonal polynomials, can reach 3×10^{-7} [15] and depends, of course, on the number of iterations with the inevitable use of successive approximation algorithms. In this case, the convergence of the process is guaranteed by the monotonic dependence of the estimates of parameters η and λ on the coefficient of variation γ [16]. Considering these circumstances, in this work we will take the numerical values of parameter estimates obtained by the moment method as accurate and use them when analyzing the accuracy of empirical formulas.

It should be noted that most of the empirical formulas for η proposed in the literature are based on the representation of the power-law dependence of the estimate of the parameter η on the coefficient of variation $\gamma = \sigma/\mu$ of the form

$$\hat{\eta} = \hat{\gamma}^{-\alpha}, \qquad (4)$$

where $\hat{\gamma} = \hat{\sigma}/\hat{\mu}$ is calculated from a sample of initial data. It is noteworthy that the literature offers different values for the coefficient α. The most common value is $\alpha = 1.086$ [1–9], in [10] the value $\alpha = 1.0695$ was proposed, and in [11] $\alpha = 1.0638$.

There are also formulas of other types, for example, in [12] a formula equivalent to $\hat{\eta} = 0.9862\hat{\gamma}^{-1.0983}$ is recommended, in [13] several options are given – $\hat{\eta} = 1.0461\hat{\gamma}^{-1.049}$, $\hat{\eta} = 1.271/\hat{\gamma}$ and $\hat{\eta} = 1.1/\hat{\gamma}$, and in [8] the formula $\hat{\eta} = 1.2/\hat{\gamma}$ was proposed.

The difference in the above formulas is apparently due to the difference in the methods for approximating

solutions to equations (2) and (3) in the expected range of estimates of the parameter η. It is natural to expect that the accuracy of these formulas will also differ. Therefore, attention is paid to this issue in the literature. For example, in [13] the accuracy of modified versions of the formulas proposed there was studied depending on the established interval.

Note that, unfortunately, empirical formulas for estimating the scale parameter λ are rarely found in the literature. We only have access to the formula

$$\hat{\lambda} = \hat{\mu}/\left(1+0.276\hat{\eta}^{-0.776}\right), \qquad (5)$$

proposed in [12], and the formula

$$\hat{\lambda} = \hat{\mu}\left(0.568+0.433/\hat{\eta}\right)^{-1/\hat{\eta}}, \qquad (6)$$

given in [6].

Thus, there is a need to find a new, simpler, and accurate empirical formula for estimating the scale parameter λ.

2. Shape parameter definition area

The accuracy of the empirical formula depends on the approximation method used in the corresponding interval of values of the shape parameter. These values, in turn, depend on the statistical characteristics of the physical quantity under study in each specific field of technology. Let us look at examples.

Analysis of experimental data on the study of wind speed, described in articles [1–14], indicates values of the shape parameter from several decimal fractions to several units.

We also give an example of the use of gradient methods in the analysis of the structural properties of an image using the Weibull distribution model for the gradient magnitude [17].

A series of calculations was carried out to estimate the shape parameter for images from various databases. Some results are shown in Tab. 1, from which we see that the values of the shape parameter in specific examples also vary from decimal fractions to several units.

Tab. 1. Examples of the shape parameter η definition area

Database	Number of images	Min	Max
TID2013	3000	0.46	1.82
Rock	1015	0.39	1.71
Brodatz	113	0.80	2.31

However, in other areas of science and technology that use models with the Weibull distribution, other values of the boundaries of the domain for determining the distribution parameters may appear that differ significantly from those given above. Therefore, in this work, all calculations were performed for a single, wide range of values of the shape parameter $0.1 \le \eta \le 10$, which corresponds to the interval $0.120 \le \gamma \le 429.8$ for the coefficient of variation. Of course, with a decrease in the interval of

considered values of the shape parameter, the accuracy of the empirical formulas will also increase.

3. Empirical formula for the scale parameter

Numerical experiments were carried out to identify useful patterns in the model data to obtain an empirical formula for the scale parameter λ. To do this, first, by dividing both sides of formula (3) by σ and denoting $\lambda\_1 = \lambda / \sigma$, we obtain

$$\lambda_1 = \frac{1}{\gamma \Gamma (1 + 1/\eta)} . \qquad (7)$$

Noting that the parameters λ_1 and η depend only on the coefficient of variation γ, a table of the values of these quantities was compiled and was calculated using formulas (2) and (7).

It was found that the correlation coefficient between η and λ_1, calculated from the data in this table, is greater than 0.9995. Since the right-hand side of (7) ultimately depends only on γ, therefore, it can also be expanded in powers of this variable and take the form (4). In this case, the estimate of the scale parameter λ will take the form $\hat{\lambda} = \hat{\lambda}_1 \hat{\sigma}$.

Using regression analysis methods, an empirical formula for the parameter λ_1 was obtained, which, after additional refinement by varying the indicator, takes the final form

$$\hat{\lambda}_1 = \hat{\gamma}^{-1.027} . \qquad (8)$$

Accordingly, the empirical formula for the scale parameter will take the form $\hat{\lambda} = \hat{\sigma} \hat{\gamma}^{-1.027}$.

4. Analysis of the accuracy of empirical formulas

To compare the accuracy of the mentioned empirical formulas, we calculated the maximum and mean-squared error (MSE) of the values given by the empirical formula relative to the "exact" parameter values in the considered interval of variation of η.

By varying the value of η with a sufficiently small step, it is shown that the optimal value of the shape parameter in (4) is $\alpha = 1.0894$, and the optimization of the formulas $\hat{\eta} = 1.271 / \hat{\gamma}$, $\hat{\eta} = 1.1 / \hat{\gamma}$ and $\hat{\eta} = 1.2 / \hat{\gamma}$ leads to the expression $\hat{\eta} = 1.173 / \hat{\gamma}$.

Thus, the updated empirical formula for the shape parameter is

$$\hat{\eta} = \hat{\gamma}^{-1.0894} . \qquad (9)$$

The results of calculations using the remaining empirical formulas mentioned are shown in Tab. 2.

As can be seen from the table, the most accurate is the empirical formulas for the shape parameter $\hat{\eta} = \hat{\gamma}^{-1.0894}$, and for the scale parameter $-\hat{\eta} = \hat{\gamma}^{-1.027}$.

It should be noted that high values of the maximum or MSE indicators of some empirical formulas, if necessary, could be reduced by narrowing the interval of variation of the shape parameter and performing an appropriate analysis of the results.

Tab. 2. Accuracy indicators of the empirical formulas under consideration

Empirical formula	Maximal error	MSE
$\hat{\eta} = \hat{\gamma}^{-1.086}$	0.152	0.048
$\boldsymbol{\hat{\eta} = \hat{\gamma}^{-1.0894}}$	0.151	0.037
$\hat{\eta} = 0.9862 * \hat{\gamma}^{-1.0983}$	0.155	0.048
$\hat{\eta} = 1.0461 * \hat{\gamma}^{-1.049}$	0.325	0.171
$\hat{\eta} = \hat{\gamma}^{-1.0638}$	0.456	0.263
$\hat{\eta} = \hat{\gamma}^{-1.0695}$	0.340	0.205
$\hat{\eta} = 1.173 / \hat{\gamma}$	0.245	0.156
$\hat{\eta} = 1.2 / \hat{\gamma}$	0.306	0.204
Formula (5)	0.225	0.136
Formula (6)	0.255	0.134
$\hat{\lambda}_1 = \hat{\gamma}^{-1.027}$	0.065	0.046

Below is a fragment of a table of exact and approximate values of the Weibull distribution parameters, calculated using the above formulas (2), (3), (8) and (9).

Tab. 3. Exact and approximate values of parameter estimates γ, η and λ_1

γ	η	η'	λ_1	$\lambda_1^{'}$
429.8	0.1	0.0014	0.0024	0.0020
2.236	0.5	0.416	0.470	0.438
1.000	1	1.000	1.051	1.000
0.523	2	2.026	2.011	1.946
0.363	3	3.015	2.892	2.831
0.229	5	4.979	4.583	4.544
0.168	7	6.977	6.255	6.246
0.120	10	10.064	8.754	8.824

Thus, we have obtained updated empirical formulas that are easy to implement for statistical estimation of Weibull distribution parameters.

5. Application to model data

To illustrate the results of applying the updated empirical formulas to numerical data, a simulation of a random sample from a Weibull distribution was carried out, implemented using the well-known inverse transformation procedure based on the use of a uniform distribution. The generation of samples was carried out for different initial values of the shape and scale parameters of the Weibull distribution. Table 4 shows examples of parameter estimates obtained by the above-mentioned "exact" and updated empirical formulas for samples with different sizes n. We see that the accuracy of the empirical formulas is quite consistent with the estimates given in Tab. 4.

Conclusion

The article discusses various empirical formulas frequently found in the literature, proposed as estimates of the parameters of the Weibull distribution. The appendices mainly consider approximate formulas for the shape parameter obtained from the equations of the moment method. For the scale parameter of similar formulas -

units. The article experimentally studies the accuracy of these formulas by comparing them with accurate values obtained by approximate methods by solving the well-known transcendental equations of the moment method for the parameters of the Weibull distribution. A renewed expression for the empirical formula for estimating the shape parameter is obtained. A new approach to estimating the scale parameter and the corresponding empirical formula is proposed. The proposed formulas can be applied to problems of statistical processing of measurement data on the properties of materials, assessing the reliability of machine parts and components, studying wind parameters, and other areas of science and technology.

Tab. 4. Numerical examples

n	η	$\hat{\eta}$	η_e	λ	$\hat{\lambda}_1$	λ_e
50	2.0	2.147	2.173	1.5	1.636	1.618
100	2.0	1.928	1.956	1.5	1.539	1.530
500	2.0	1.981	2.008	1.5	1.501	1.490
1000	2.0	2.003	2.030	1.5	1.511	1.499
50	2.5	2.608	2.627	0.5	0.536	0.527
100	2.5	2.365	2.389	0.5	0.510	0.503
500	2.5	2.472	2.494	0.5	0.500	0.493
1000	2.5	2.496	2.518	0.5	0.503	0.495

References

[1] Vega-Zuñiga S, Rueda-Bayona JG, Ospino-Castro A. Evaluation of eleven numerical methods for determining Weibull parameters for wind energy generation in the Caribbean region of Colombia. Math Model Eng Probl 2022; 9(1): 194-199. DOI: 10.18280/mmep.090124.

[2] Kapen PT, Gouajio MJ, Yemélé D. Analysis and efficient comparison of ten numerical methods inestimating Weibull parameters for wind energy potential: Application to the city of Bafoussam, Cameroon. Renew Energy 2020; 159: 1188-1198. DOI: 10.1016/j.renene.2020.05.185.

[3] Guenoukpati A, Salami AA, Kodjo MK, Napo K. Estimating Weibull parameters for wind energy applications using seven numerical methods: Casestudies of three coastal sites in West Africa. Int J Renew Energy Dev 2020; 9(2): 217-226. DOI: 10.14710/ijred.9.2.217-226.

[4] Justus CG, Hargraves WR, Mikhail A, Graber D. Methods for estimating wind speed frequency distributions. J Appl Meteorol Clim 1978; 17(3): 350-353. DOI: 10.1175/1520-0450(1978)017<0350:MFEWSF>2.0.CO;2.

[5] Ramírez PJ, Carta A. Influence of the data sampling interval in the estimation of parameters of the Weibull wind speed probability density distribution a case study. Energy Conv Manag 2005; 46: 2419-2438. DOI: 10.1016/J.ENCONMAN.2004.11.004.

[6] Mohammadi K, Mostafaeipour A. Using different methods for comprehensive study of wind turbine utilization in Zarrineh, Iran. Energy Conv Manag 2013; 65: 463-470. DOI: 10.1016/J.ENCONMAN.2012.09.004.

[7] Costa Rocha PA, Sousa RC, Andrade CF, Silva MEV. Comparison of seven numerical methods for determining Weibull parameters for wind energy generation in the northeast region of Brazil. Appl Energy 2012; 89: 395-400. DOI: 10.1016/j.apenergy.2011.08.003.

[8] Indhumathy D, Seshaiah CV, Sukkiramathi K. Estimation of Weibull parameters for wind speed calculation at Kanyakumari in India. Int J Innov Res Sci Eng Technol 2014; 3(1): 8341-8345.

[9] Mohammadi K, Alavi O, Mostafaeipour A, Goudarzi N, Jalilvand M. Assessing different parameters estimation methods of Weibull distribution to compute wind power density. Energy Conv Manag 2016; 108: 322-335. DOI: 10.1016/j.enconman.2015.11.015.

[10] Kelly M, Troen I, Jørgensen HE. Weibull-k revisited: "Tall" profiles and height variation of wind statistics. Boundary-Layer Meteorol 2014; 152(1): 107-124. DOI: 10.1007/s10546-014-9915-5.

[11] Robinson EY. Estimating Weibull parameters for materials. JPL Tech Memo 33-580. Pasadena, CA, USA: Jet Propulsion Lab; 1972. 62 p.

[12] Kanji O. A simple estimation method of Weibull modulus and verification with strength data. Appl Sci 2019; 9: 1575. DOI: 10.3390/app9081575.

[13] Azad AK, Rasul MG, Alam MM, Ameer Uddin SM, Sukanta Kumar M. Analysis of wind energy conversion system using Weibull distribution. Procedia Eng 2014; 90: 725-732. DOI: 10.1016/j.proeng.2014.11.803.

[14] Chang TP. Performance comparison of six numerical methods in estimating Weibull parameters for wind energy application. Appl Energy 2011; 88(1): 272-282. DOI: 10.1016/j.apenergy.2010.06.018.

[15] Abramowitz M, Stegun IA, eds. Handbook of mathematical functions with formulas, graphs and mathematical tables. National Bureau of Standards; 1972. 1060 p.

[16] Asatryan DG. Comparative study of three approaches for estimating the Weibull distribution parameters. Vestnik RAU 2018; 2: 5-16.

[17] Asatryan DG. Gradient-based technique for image structural analysis and applications. Computer Optics 2019; 43(2): 245-250. DOI: 10.18287/2412-6179-2019-43-2-245-250.

Author's information

David Geghami Asatryan, (b. 1940). Doctor of sciences (Engineering), professor, works as a Head of Center for Critical Technologies of the Russian-Armenian University and as a leading scientific worker at the Institute for Informatics and Automation Problems of National Academy of Sciences of Armenia. Research interest: signal and image processing. E-mail: *dasat@iiap.sci.am*