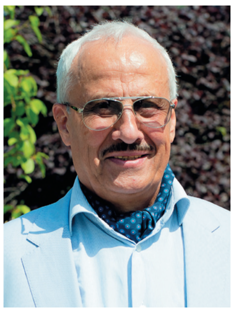
Original article
https://doi.org/10.21683/1729-2646-2019-19-4-17-23

Method of normalization of dependability indicators of railway transport facilities

Igor B. Shubinsky[1], Evgeny O. Novozhilov[1*]
[1]JSC NIIAS, Russian Federation, Moscow
[*]eo.novozhilov@vniias.ru



Igor B. Shubinsky

Evgeny O. Novozhilov

Abstract. Aim. *The results of evaluation of a technical system's (facility's) factual state allow making a decision on a further life (operation continuation, maintenance assignment, decommissioning and a facility's replacement etc.). Under the conditions of resource limits, it is vital to identify most "problematic" facilities that require primary investments. The aim of the research is to develop a method of normalization of dependability indicators whose application is intended to improve targeted investment allocation for maintenance of facilities, which allows fulfilling the requirement of uninterruptible transportation under the conditions of resource scarcity.* **Methods.** *The research uses methods of system analysis, probability theory, mathematical statistics, and correlation analysis. It proposes approximation of a time series of factual values related to a dependability indicator by a three-parameter gamma distribution based on a scarcity function $q(x)$.* **Findings.** *The research has considered the criteria of choice of railway transport facilities requiring the enhancement of dependability for the cases of unavailability and availability of a normalized dependability indicator. It has been shown that if introducing normalization of indicators one should take into account non-similar maintenance conditions for facilities in different enterprise units, which are determined by differences in climatic factors, technical capabilities for maintenance and repair, staffing levels, grades of tear and wear of facilities, requirements for their productivity. The research has analyzed the conditions of association of a service supplier's and user's requirements for normalization of a dependability indicator value. It has been demonstrated that it is reasonable to establish a single threshold normalized value x of a dependability indicator, in which case a normalized value x for the attribute x shall comply with the requirements of a service user as well as a service supplier. In the case of a single threshold value, the risk $Q = P\{x > x\}$ of noncompliance of an indicator with the specified requirements is in fact split between a service user and a service supplier according to their agreement.* **Conclusions.** *The paper proposes a method of normalization of a dependability indicator based on statistical data assuming that in general this indicator may be evaluated for a certain period of observance as acceptable for a service user. For to choose and justify the normalized value of a dependability indicator, the authors have studied the relations between a service supplier and a service user, have analyzed statistics using the method of estimation of empirical sufficiency of a raw data series as well as approximation of an ordered initial series by a three-parameter gamma distribution. The paper provides an example of normalizing a value of a facility failure rate indicator as per the criterion of a specified risk of its violation based on the quantiles of an obtained function of sufficiency. It has been shown that the proposed approach allows establishing a correlation between a normalized value and a risk of its violation via a function of sufficiency, which can be obtained on the basis of existing statistical data on a facility's dependability for the past periods. This correlation makes it possible to guarantee the ensuring of compliance of factual and normalized indicator values with a specified risk level for a facility working in normal mode.*

Keywords: *dependability indicator, dependability normalization, service supplier's risk, service user's risk, scarcity function, three-parameter gamma distribution, distribution quantile.*

For citation: *Shubinsky IB, Novozhilov EO. Method of normalization of dependability indicators of railway transport facilities // Dependability 2019;4; 17-23 p. https://doi.org/10.21683/1729-2646-2019-17-23.*

Received on 02.10.2019 / Revised on 22.11.2019 / For printing 14.12.2019

Introduction

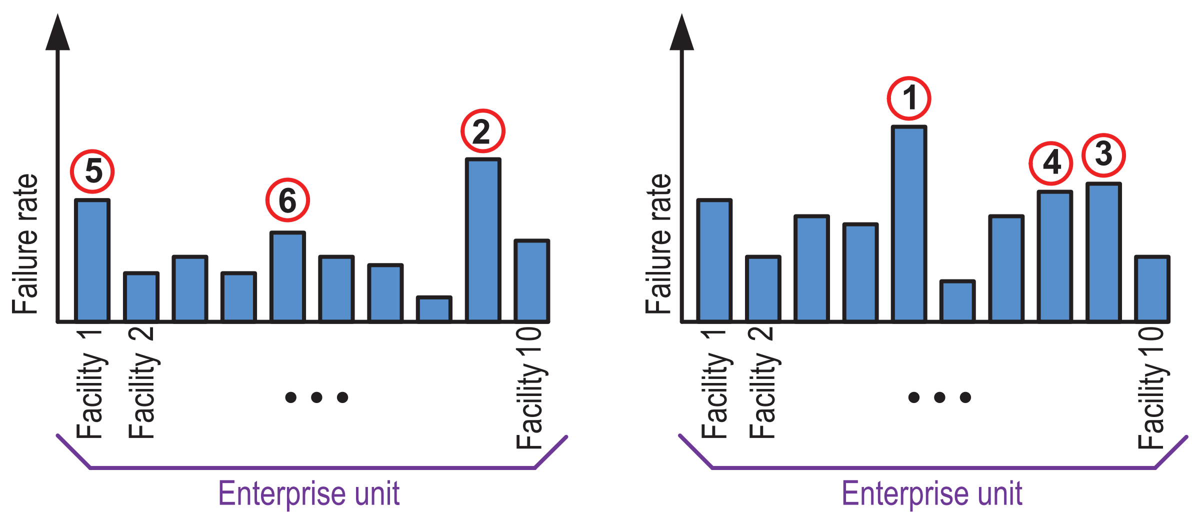
For any technical system one of the important tasks is the normalization of dependability indicators (for example, acceptable value of availability, reliability and maintainability) [1, 2]. Dependability normalization is the specification (in technical or other documentation) of quantitative or qualitative requirements for dependability. Therefore, normalization sets acceptable limits for changes of a controlled characteristic.

A dependability indicator is a characteristic (as a rule, quantitative) of one or several properties comprising the dependability of a technical system (facility). The values of dependability indicators can be normative or factual. They can be determined by calculation methods, on the basis of maintenance data or by extrapolation. Factual values of dependability indicators during the process of operation of a technical system are obtained based on the analysis of statistical data on a system's failures and time to its recovery.

As far as normative values of dependability indicators, they are as a rule specified in a quantitative way at the design stage of a facility. For most facilities one applies a normalization probabilistic approach when one normalizes and ensures a required economically justified level of probabilistic dependability indicators that is afterwards controlled by dependability tests and kept by a maintenance system. The exclusion is safety critical facilities with catastrophic failure consequences, whose failures are not acceptable (this paper doesn't consider such facilities since they belong to the field of functional safety).

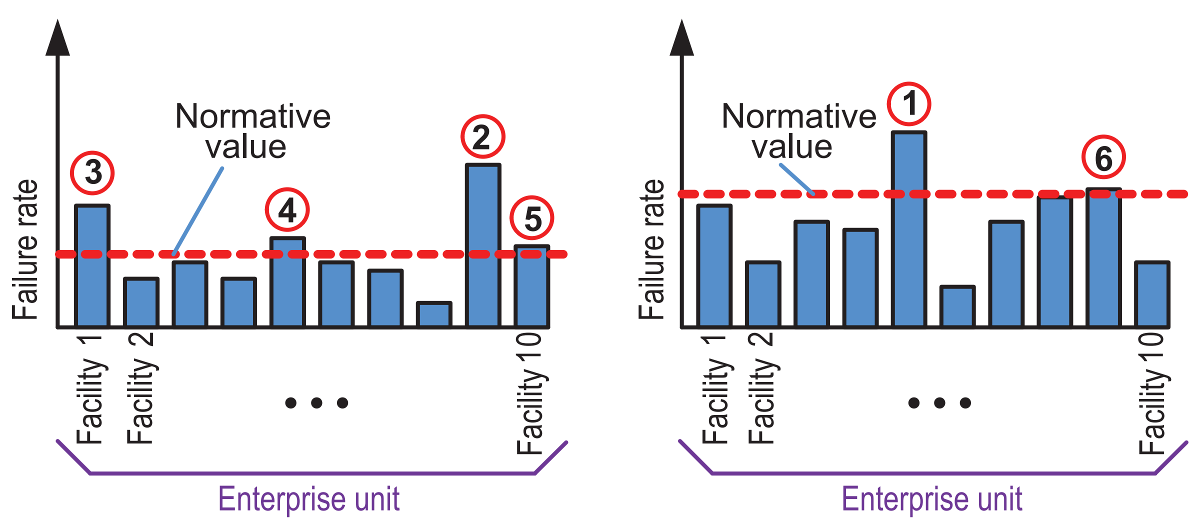
1. The goal of normalization of dependability indicators

The results of evaluation of a technical facility's factual state allow making a decision [3] on a further life (operation continuation, maintenance assignment, decommissioning and a facility's replacement etc.). Under the conditions of



a) b)

Figure 1. Example of determination of facilities' order of priority for repair assignment (without normalization).



a) b)

Figure 2. Example of determination of facilities' order of priority for repair assignment (with normalization).

resource limits, it is vital to identify most "problematic" facilities that require primary investments.

Figures 1 a and 1 b show an example of determination of priority levels of railway infrastructure facilities requiring the enhancement of dependability – for example, by assignment and execution of repair – for two enterprise units, where facilities of one type are under different operation conditions. In this example we assume that in these two enterprise units there are funds reserved for repair of 6 facilities.

Figure 1 shows that based on factual values of a dependability indicator (for example, a failure rate), that reflects the current state of the facility in operation, we can identify those facilities that require repair assignment as a priority with the size of an allocated investment taken into account. In this case, if normative values are not available, facilities are chosen by the criterion of the worst indicator value.

When introducing normalization of indicators one should take into account non-similar maintenance conditions for facilities in different enterprise units, which are determined by differences in climatic factors, technical capabilities for maintenance and repair, staffing levels, grades of tear and wear of facilities, requirements for their productivity (for example, with different sizes of train traffic). In this case facilities will be chosen for repair assignment by the criterion of an indicator's deviation to the worse side from a normative value (Fig. 2 a и 2 b).

Obviously, introduction of normative indicators considering operation conditions and other factors of enterprise units' activities improves targeted investment allocation for maintenance of facilities, which allows fulfilling the requirement of uninterruptible transportation under the conditions of resource scarcity [4].

2. User and supplier interests

In case when a technical system is involved in providing services (for example, a railway infrastructure facility ensures transportation process execution), normative values of dependability indicators shall consider relations between a supplier and a user of a service (for example, an enterprise unit in charge of the functioning of a railway infrastructure facility and an enterprise unit executing transportation process).

It is worth to note that this scheme presents an inevitable conflict between the interests of a user and a supplier of a service. From the one hand, a user is interested that there would not be any failures of a facility providing a service at all; this would allow him to execute his activities with no risk related to a facility failure (for example, a risk of train hours loss due to the failure of a railway infrastructure facility). From the other hand, a supplier is interested in reducing the costs of a service, thus increasing the operating profit, but a reduction of costs inevitably causes increased failure rates. Normalization of a facility's dependability indicators shall in essence ensure a compromise between the interests of a supplier who seeks to provide a service under the conditions of resource limits and the interests of a user who seeks to have a service of high quality with the lowest expenditures.

The situation in question is similar to the situation when a user receives a product batch from a supplier and where the unambiguity of mutual acknowledgment of a product's quality by a supplier and a user is in most cases regulated by methods of statistical acceptance tests. And the relations between a supplier and a user characterize an acceptable level of quality x_α (the maximum acceptable value of defective items share in a batch) and an unacceptable quality level x_β (the boundary of defective items share for attributing a batch as defective), where $x_\alpha \le x_\beta$ (Fig. 3). Therefore, the area of a user's interests is $x \le x_\beta$, and the area of a supplier's interests is $x > x_\alpha$; it is obvious that the two areas cross each other that being a prerequisite condition for the existence of compromise between both interests. The area of an attribute value x under x_α is "acceptance region", that above x_β

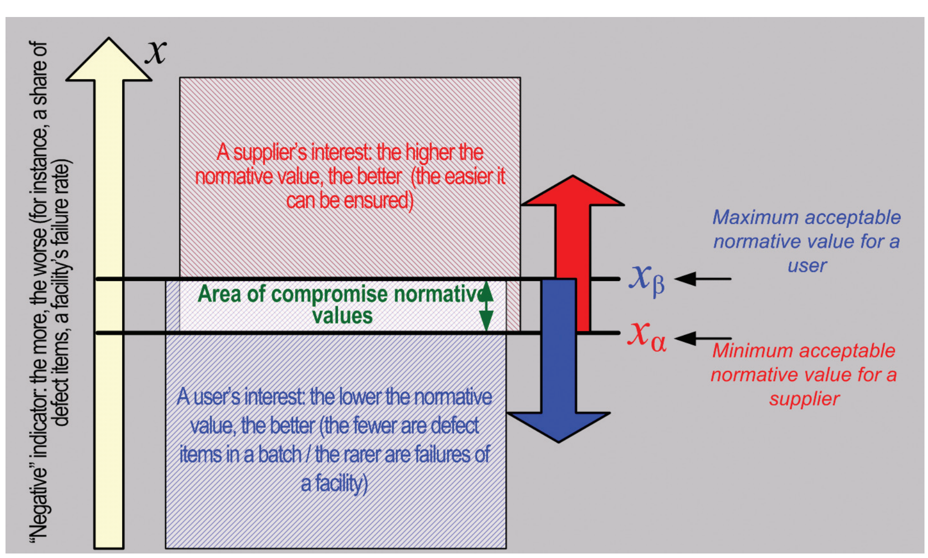


Figure 3. Areas of a service user's and a service supplier's interests.

is “unacceptance region”, and that between x_α and x_β is “uncertainty region”.

Note that the application of two points (x_α, x_β) as threshold values is a general practice for a sequential selective test [5], where a conclusion on usability (or non-usability) of a batch is made on the basis of a defective items share in a selection that is a part of a batch volume. For this process, acceptable and unacceptable levels are set using confident intervals. A probability that a share of defective items in the whole batch is not larger than x_α, when an upper confident interval of an unacceptable level is exceeded for a selection test, is a supplier’s risk; vice versa, a probability that a share of defective items in the whole batch is larger than x_β, when a lower confident interval of an acceptable level is reached for a selection test, is a user’s risk.

From Fig. 3 it follows that a service supplier can guarantee that the factual value x of an indicator will be above the threshold value x_α with a high degree of confidence (probability), for example, $P_\alpha = P\{x > x_\alpha\} > 0.95$ (supplier’s risk $Q_\alpha = 1 - P_\alpha \leq 0.05$ (GOST R ISO 8422-2011. Statistical methods. Sequential plans of selective tests as per alternative attribute); a service user expects that the factual value x of an indicator will not be higher than the threshold value x_β with a high degree of confidence (probability), for example, $P_\beta = P\{x \leq x_\beta\} > 0.9$ (user’s risk $Q_\beta = 1 - P_\beta \leq 0.1$).

Under the real conditions of a technical system’s operation, the task of evaluating its compliance with specified requirements of dependability is often brought down to comparison of the value of a factual dependability indicator obtained for some period of observance of statistical operational data with a normative value specified in technical or other documentation. In this case the presence of “uncertainty region” will complicate estimation making it ambiguous. That’s why technical documentation for a facility generally contains a normative value of an indicator in form of a single threshold value (for example, “mean time to failure shall be not lower than 30 000 h, maintenance inclusive).

Let a single threshold value normative value x_η be specified for a facility dependability indicator by agreement between a user and a supplier, then we will assume that for $x \leq x_\eta$ this facility complies with the requirements, and for $x > x_\eta$ it does not. It is obvious (see Fig. 3) that when transiting from two threshold levels to one it is reasonable to comply with the condition $x_\alpha < x_\eta \leq x_\beta$ (x_η belongs to the area of “compromise values”), in which case the normative value x_η for the attribute x satisfies to the requirements of both a service user and a service supplier.

In the case of a single threshold value, the risk $Q_\eta = P\{x > x_\eta\}$ of noncompliance of an indicator with specified requirements is in fact split between a user and a supplier of a service according to their agreement (for example, th exceedance of a normative value at one interval of observance is a user’s risk, while at two or more consecutive intervals of observance it is the responsibility of a supplier).

One of the ways of normalizing dependability indicators used in the global practice (in particular, in the power supply field) is the normalization based on past experience (analysis of factual data on dependability) [5]. Given the availability of such data on railway transport, we will consider a further task as a choice and justification of the value x_η using existing statistical data on the operation of a facility during some interval of observance, assuming that in general these indicators of a facility’s dependability may be evaluated for this interval of observance as acceptable for a service user.

3. Analysis of statistical data and evaluation of their sufficiency

As it was noted earlier, the factual values of dependability indicators are random values. For example, for a facility’s failure rate (number of failures per time unit) the statistics presents a time series of discrete values – for instance, this is a sequence of failure rate values per each annual interval of observance for several years.

A random value is fully defined by a distribution law, for discrete values this is a distribution series or a discrete distribution function. A distribution series (a discrete distribution function) presents a table of possible values of a random size with respective probabilities.

There are a great number of various theoretical laws of distribution (uniform, Bernoulli, Cauchy, Poisson, normal, lognormal, Gumbel, Jonson, 13 Pearson’s curved distributions etc.) [6]. However, in practice one often deals with statistical material of rather a limited volume, and it is not always possible to identify a concrete distribution law for a random value based on this volume. In such cases it is necessary to describe the behaviour of a random value by numeric characteristics.

For engineering calculations and scientific researches one uses empirical curved distributions of random values characteristics. When constructing such curves, major stages are ranking of an initial time series and estimation of its empirical sufficiency. Solving the first of these tasks presents no difficulties, whereas for the second it is necessary to take into account that some formulas for estimation of sufficiency lead to systematic errors and give different values of random errors.

Scarcity function $q(x)$ is an analog of distribution function $F(x)$ and characterizes a probability that the value of an argument exceeds a specified threshold value. [7] based on theoretical researches and results of testing defined a formula, which gives efficient, nonbiased and effective values of scarcity estimates of the i-th ($i = 1…n$) member of a discrete sample ranked in descending order (i.e., of probabilities q_i that the factual value x exceeds the value of x_i series member):

$$q_i = P\{x > x_i\} = \frac{i}{n+1}, \quad (1)$$

Where n is a number of series members.

Table 1. Initial time series of a facility's failure rate

Observance year	2008	2009	2010	2011	2012	2013	2014	2015	2016
Failure rate, x_i, 1/year	34	37	24	17	12	9	13	43	36

Table 2. Ranked time series with scarcity estimates and modulus coefficients

Item No., i	1	2	3	4	5	6	7	8	9
Failure rate, x_i, 1/year	83	76	37	34	24	14	13	12	9
Mod. coeff., k_i	2.4735	2.2649	1.1026	1.0132	0.7152	0.4172	0.3874	0.3576	0.2682
Scarcity, q_i	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

Table 3. Example of an approximated time series with scarcity estimates and modulus coefficients (C_v = 0.8 и C_{sv} = 1.4)

Item No., i	-	-	1	2	3	4	5	6	7	8	9	-	-
q_i	0.01	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
Mod. coeff., m_i	2.01	1.8	1.66	1.47	1.31	1.16	1.01	0.855	0.69	0.511	0.305	0.182	0.055
Failure rate, y_i, 1/year	50.25	45	41.5	36.75	32.75	29	25.25	21.38	17.25	12.78	7.625	4.55	1.375

Let us consider the algorithm comprising the ranking of an initial time series, the estimation of its empirical sufficiency and the approximation by a theoretical distribution law using the statistical data on failures of primary railway telecommunications network facilities for the years of 2008–2016 (Table 1, the data submitted by the Central telecommunications station – JSC RZD branch).

1) An initial series is ranked in order of descending of an indicator's values. Instead of observance years we introduce conditional numbers of a ranked series' members (1, 2, 3, …).

2) For each member of a ranked series we calculate values q_i of scarcity function using formula (1).

3) Then we calculate mathematical expectation $\overline{x}$ of series members.

4) For each member of a ranked series we calculate a modulus coefficient equal to a relation of a series member's value to a series' mathematical expectation.

As a result, we have Table 2.

5) For refinement of values of distribution quantiles (q), especially at levels lower than 0.2, that are of practical interest, we make approximation of a series (Table 2) using one of the theoretical distribution laws. As an example, let us consider approximation by a three-parameter gamma distribution [8] that has been in particular applied in hydrological calculations [9], calculations of construction resources for random flows of loads [10] and calculations of structures' service life under random load flows [11].

Using modulus coefficients k_i from Table 2 we calculate the coefficient C_v of series variation and relation C_{sv} of a series' asymmetry coefficient to a series' variation coefficient:

$$C_v = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i}{\overline{x}}-1\right)^2},\ C_{sv} = \frac{n\sum_{l=1}^{L}\left(\frac{x_i}{\overline{x}}-1\right)^3}{(n-1)(n-2)C_v^4}, \quad (2)$$

Where if in (2) we obtain the value $C_v < 0.1$, then before calculating C_{sv}, as well as for further usage, we assume that $C_v = 0.1$ (as for series with a very small variation it is complicated to define distribution quantiles). After calculation the value C_v is approximated to multiplicity 0.1 (0.1; 0.2; 0.3 …), while the value C_{sv} is approximated to multiplicity 0.5 (0; ±0.5; ±1.0; ±1.5; …) to allow the application of existing table distribution function values since their analytical calculation is very complicated.

Using table values of three-parameter gamma distribution functions [9] for a specified scarcity probability q_i, we define the ordinate m_i in form of a modulus coefficient (the mentioned tables contain values of a distribution functions for various values C_v and most widely-spread relations C_s / C_v).

In the example in question for values $C_v = 0.5$ and $C_{sv} = 0$ obtained using formulas (2), we have a series of values of function ordinates as modulus coefficients $m_i(p_i)$, including additional values at the boundaries of a function (Table 3). In order to obtain quantitative values y_i of failure rate that will be exceeded with the probability q_i, we should multiply modulus coefficients m_i by the value $\overline{x}$ of mathematical expectation of a ranked series from Table 2 (the results are summarized in Table 3).

The estimation of approximation reliability was made using a coefficient of an empirical linear correlation $x_i(q_i)$ and a function chosen as per this method $y_i(q_i)$ (for i = 1…9). We obtained the value of a linear correlation coefficient as 0.974, which is close to 1, thus confirming the closeness of the chosen function to the initial series with a high reliability.

Fig. 4 present a graph of an empirical series (points) and an approximating function of a three-parameter gamma distribution (solid line).

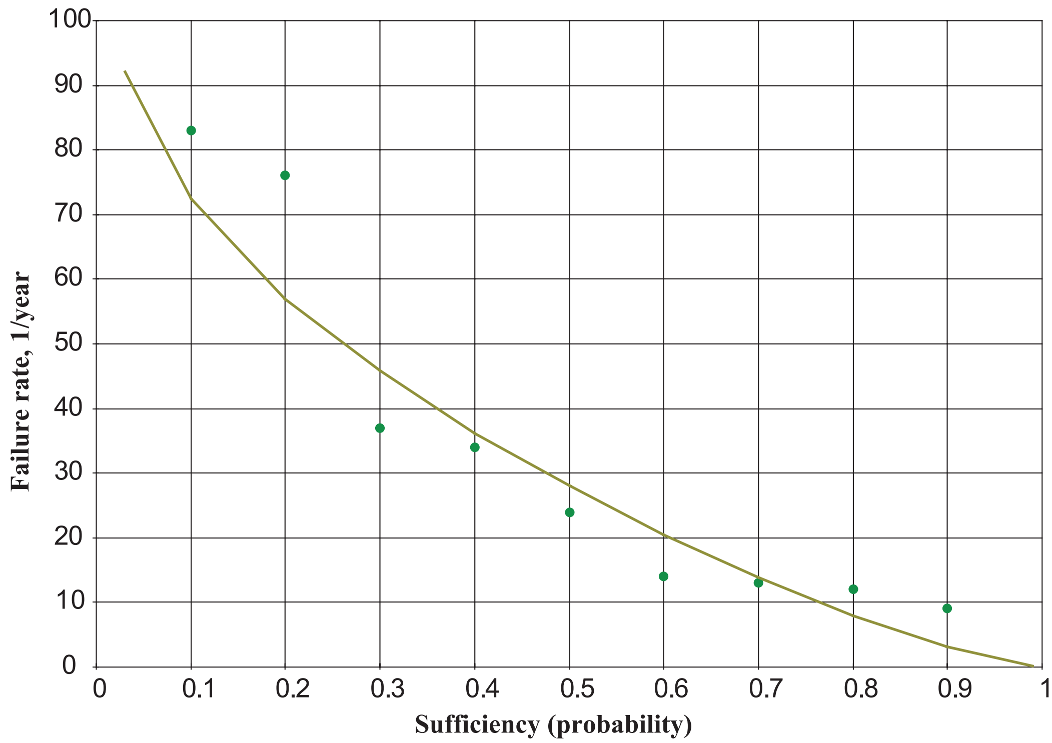


Figure 4. Graph of an empirical series (points) and approximation function (solid line).

4. Choice and justification of a normative indicator value

The results of sufficiency estimation obtained above (see Table 3) can be applied for defining a threshold value x_η for a specified level of risk Q_η agreed between a supplier and a user of a service or vice versa for estimating risk Q_η based on a specified value x_η.

Let us consider a case when for a specified risk level of noncompliance with a normative value (for example, $Q_\eta = 0.1$) we have to define a normative value x_η of a dependability indicator (in our case it is a facility's failure rate).

Let us estimate the quantile of a sufficiency function that corresponds to a specified risk ($q_i = Q_\eta = 0.1$). According to the data of Table 3 we have:

$$y(Q_\eta) = y(q_1 = 0.1) = 70.8 \approx 71.$$

Therefore, as an indicator's normative value we can take a failure rate equal to 71 1/year, which will be not ensured with a risk of 0.1.

In case if by agreement between a supplier and a user of a service there is a specified normative value of dependability, in a similar way based on the obtained results of sufficiency estimation (see Table 3) one can define risk of noncompliance of an indicator with specified requirements.

In any case an agreement between a supplier and a user of a service shall foresee both the specification of a normative value of all dependability indicators in question and the specification of risk levels for nonfulfillment of these normative values as well as the procedure of splitting of responsibility between a supplier and a user of a service.

The method considered in the paper allows defining a relation between a value of a dependability normative indicator and a risk of its nonfulfillment by objective criteria based on factual capabilities of operated facilities that are estimated as per existing statistical data for the past periods.

Conclusions

The paper has considered a method of normalization of a dependability indicator based on statistical data assuming that in general this indicator may be evaluated for a certain period of observance as acceptable for a service user.

For to choose and justify the normalized value of a dependability indicator, the authors have studied the relations between a service supplier and a service user, have analyzed statistics using the method of estimation of empirical sufficiency of a raw data series as well as approximation of an ordered initial series by a three-parameter gamma distribution. The paper provides an example of normalizing a value of a facility failure rate indicator as per the criterion of a specified risk of its violation based on the quantiles of an obtained function of sufficiency.

The research has demonstrated that the proposed approach allows establishing a correlation between a normalized value and a risk of its violation via a function of sufficiency, which can be obtained on the basis of existing statistical data on a facility's dependability for the past periods. This correlation makes it possible to guarantee the ensuring of compliance of factual and normalized indicator values with a specified risk level for a facility working in normal mode.

References

[1] Dolganov AI, Sakharov AV. On the assignment of dependability level. Dependability 2018;18(3):18-21.

[2] Kuvashov YA, Novozhilov EO. Method of evaluation of the railway track's availability for traffic operations. Dependability 2017;17(2):17-23.

[3] Semenov SS, Poltavsky AV, Maklakov VV, Krianev AV. Overview of decision-making techniques used in the development of complex engineering systems. Dependability 2014;3:85-96.

[4] Gapanovich VA, Shubinsky IB, Rozenberg EN, Zamyshlyaev AM. System of adaptive management of railway transport infrastructure technical maintenance (URRAN project). Dependability 2015;2:14-22.

[5] Rudenko YuN. O podkhodakh k normirovaniyu pokazateley nadezhnosti elektrosnabzheniya potrebiteley [On the approaches to the standardization of the dependability indicators of electric power supply to consumers]. Proceedings of the Academy of Sciences of the USSR. Energy and transportation 1975;1:14-23 [in Russian].

[6] Litvinenko RS, Pavlov PP, Idiyatullin RG. Practical application of continuous distribution laws in the theory of reliability of technical systems. Dependability 2016;16(4):17-23.

[7] David H. Order statistics. Moscow: Nauka; 1979.

[8] Vadzinsky RN. Spravochnik po veroyatnostnym raspredeleniyam [Handbook of probability distribution]. Saint Petersburg: Nauka; 2001 [in Russian].

[9] Sikan AV. Metody statisticheskoy obrabotki gidrometeorologicheskoy informatsii [Methods of statistical processing of hydrometeorological information]. Saint Petersburg: RSHU; 2007 [in Russian].

[10] Gusev AS. Soprotivlenie ustalosti i zhivuchesti konstruktsiy pri sluchaynykh nagruzkakh [Fatigue strength and survivability of structures under random load]. Moscow: Mashinostroenie; 1989 [in Russian].

About the authors

Igor B. Shubinsky, Doctor of Engineering, Professor, Deputy Director of Integrated Research and Development Unit, JSC NIIAS, Moscow, Russian Federation, phone: +7 (495) 786-68-57, e-mail: igor-shubinsky@yandex.ru

Evgeny O. Novozhilov, Candidate of Engineering, Head of Unit, JSC NIIAS, Moscow, Russian Federation, phone: +7 495 967 77 02, e-mail: eo.novozhilov@vniias.ru

The authors' contribution

Igor B. Shubinsky reviewed and analyzed the state of the art of the problem under consideration, defined the theoretical aspects of the paper, applied mathematical methods.

Evgeny O. Novozhilov analyzed the existing approaches to the normalization of dependability indicators. He proposed an algorithm of dependability indicators normalization, performed an example calculation.