Special aspects of estimating the probability of fire occurrence on diesel locomotives of various types

Igor B. Shubinsky, *ZAO IBTrans, Moscow, Russia, e-mail: igor-shubinsky@yandex.ru*
Olga B. Pronevich, *JSC NIIAS, Moscow, Russia, e-mail: O.Pronevich@vniias.ru*
Anna D. Danilova, *JSC NIIAS, Moscow, Russia, e-mail: A.Danilova@vniias.ru*

Igor B. Shubinsky

Olga B. Pronevich

Anna D. Danilova

Abstract. Aim. *The paper considers the problem of estimating the probability of fire occurrence on diesel locomotives of various types and the ways to solve it. The problem arises due to JSC RZD locomotive fleet special aspects. Thus, the operating fleet presents diesel locomotives designed and constructed in the 20th century as well as in 21st century, and this accounts for different causes of fire owing to design differences. The biggest contribution to differences in fire numbers on new and old type locomotives is made by the construction of a diesel engine as well as the fire resistance of cables. Researches show that substantial difference in fire statistics for diesel locomotives of various types for the same period of observation are caused by the volume of operating diesel locomotive fleet. For instance, volumes of operating fleet for some types of diesel locomotives amount to thousand units (loco-days), while for other types they make up just a couple of hundreds. This raises questions about whether a period of observation and a volume of operating fleet are enough for estimating the probability and what methods should be used to estimate it. Furthermore, we need an interval estimation of probability which is caused by reliability considerations, by getting "the worst scenario". Again, this is influenced by above differences in types of diesel locomotives. The paper also analyzes the necessity of estimating "the worst scenario" and problems arising in reference with its calculation. To solve the problem of enhancing the reliability of calculations is to calculate the upper boundaries of probabilities. In this case some types of diesel locomotives will have a lower boundary of probability rather than "the worst scenario" as interval estimation. The necessity of such estimation is specified for diesel locomotives of specific designs with materials complying with modern standards in terms of reliability and fire resistance or having scarce statistics for applying approximation methods of calculation because of limited operating fleet.* **Methods.** *Researches into statistics of fires on diesel locomotives of types 2TE10, 3TE10, 2TE116, 2M62, TEP70, ChMEZ, TEM2 made by the authors began with application of a "classic" statistics tool, i.e. check of statistical hypotheses about a law of distribution of a random value "fire" belonging to known discrete laws. While at this, a minimum amount of tests was defined for making sure that targeted estimates of probability are of certain reliability. The condition of a diesel locomotive during operation is not stationary, so a classic estimation of probability of fire occurrence would lead to uncertainty in applying the results of estimation for the purposes of planning and prediction. To evaluate "the worst scenario", we used both precise and approximate methods for defining confidence boundaries based on "double approximation". Further, to enable transition from estimation of probability of fire occurrence on diesel locomotives of a certain type to estimation of fire probability for certain units, a sufficient amount of rolling stock was researched. The authors have found that the amount of operating fleet should be not less than 610 loco-days to ensure the precision of probability calculation with an error not exceeding ε. The authors have also identified the method and the necessity of separately estimating the probability of fire on locomotives with operating fleet less than 610 loco-days.* **Results. Conclusions**. *In fine, for each type of locomotive we have defined a law of random value occurrence, calculated interval estimates of probability of fire occurrence considering an amount of operating fleet. Tools of statistical analysis for calculating probabilities of fire occurrence on diesel locomotives of various types have been also identified. We have determined methods for calculating interval probability estimates taking into account an available amount of observations with an error not exceeding a specified value ε at the level of $0{,}2p_j^*$. This research and related calculations have enabled us to obtain one of the primary elements for estimating a fire risk, i.e. the probability of fire on diesel locomotives of various types.*

Keywords: *fire risk, probability, diesel locomotive.*

For citation: *Shubinsky, I.B., Pronevich, O.B., Danilova, A.D. Special aspects of estimating the probability of fire occurrence on diesel locomotives of various types // Dependability. 2016. Issue No. 4. P. 24-29. DOI: 10.21683/1729-2646-2016-16-4-24-29*

Introduction

Calculation of probability of an undesirable event is vital part of risk analysis. For analysis of fire risk, an undesirable event is a case of fire. The paper tackles analysis of fire cases as random events and estimation of probability of their occurrence. Probability of fire occurrence is a mathematical value of possibility of occurrence of necessary and sufficient conditions for fire breakout (catching fire) [1]. Conditions for fire break-out presents a set of direct fire causes or, in terms of probability theory, elementary outputs, whose set favors occurrence of fire event. That said, the set of elementary outputs will differ for different objects. The most obvious example is such: on diesel locomotives 12% of fires happen due to failures of a turbo charger. For electrical locomotives a share of fires due to this cause is 0%, since electrical locomotives do not have turbo chargers. Diesel locomotives of various types also have design differences. However, design is not the only criterion of difference between types of diesel locomotives. Not lesser influence on an amount of fires is made by a volume of diesel locomotives operation or, in other words, operating fleet. Therefore, as a random event of fire on different types of locomotives is caused by different sets of elementary outputs, the probability of fire on a locomotive of each type should be calculated separately. This paper deals with consideration of special aspects of available fire statistics, choice of statistical tools for estimating fire probabilities with each type of locomotives taken into account.

The paper uses a mathematical apparatus for calculating probabilities of fire on diesel locomotives of such types as 2TE10, 3TE10, 2TE116, 2M62, TEP70, ChMEZ, and TEM2 (based on statistics for the period of 2008-2015). The process of calculating probability includes consideration and description of the process of selection of a distribution law, check of compliance of chosen distribution conditions.

The second part of the paper deals with enhancement of reliability of calculation results. This task was solved by estimating upper boundaries of probabilities of catching fire on various types of diesel locomotives. The calculation allowed estimating the weight of probability as a value above which the weight of probability will not rise with a high reliability. The estimate of upper boundaries of probability is used not so widely as an explicit estimation of probability. The task solving was complicated by the fact that for some types of diesel locomotives the sample was small, and that was shown as a small operating fleet and as the fact that it was impossible to apply classical formula for estimating upper boundaries of parameters of known distributions. A precise method of calculation was used for estimating such probabilities.

Selection of distribution

For estimating the probability of fire occurrence it is necessary to select a mathematical model of estimation. In this case estimation of probability was made by using a Bernoulli distribution (binominal trial model) [2]. A Bernoulli distribution is called a consequence of tests satisfying to certain conditions. Table 1 provides the analysis of conditions for application of a Bernoulli distribution for analyzing fire cases as to empirical data (observations).

According to the accepted model, a trial is assumed as a month wherein we witness cases of fire or no fire on locomotives that have N volume of operating fleet.

Using a Bernoulli formula, the occurrence probability of exactly k of successes $P_n(k)$ is equal to[1]:

$$P_n(k) = C_k^n p_j^{*k} (1-p_j^*)^{n-k} = \frac{n!}{k!(n-k)!} p_j^{*k} (1-p_j^*)^{n-k} \quad (1)$$

Furthermore, we have analyzed other known discrete distribution laws, including a Poisson distribution. Below one can find checking the extent to which observed statistics is described by the chosen law.

Check of statistical hypotheses

For checking assumptions related to definition of compliance of a sample with a certain distribution law, a mechanism of checking statistical hypotheses is used.

A statistical hypothesis is called an assumption about a type of distribution law or values of unknown parameters of a distribution law in the population.

Below is the procedure of checking a statistical hypothesis about compliance of fire occurrence with a binominal distribution law:

1. Statement of a null and alternative hypothesis.

A hypothesis stated about compliance with a binominal distribution law with confident likelihood α is called a null hypothesis H_0: $\{P=P_0\}$. As a result of statistical check, a null hypothesis is either accepted (assumed as true) or rejected (assumed as false).

In addition to a null hypothesis, we define an alternative hypothesis H_1: $\{P \neq P_0\}$ that will be accepted in case a null hypothesis is rejected – a hypothesis about incompliance with binominal distribution:

2. Selection of a criterion for checking a hypothesis and calculation of its observed value.

For checking hypotheses, one applies special criteria which are calculated using observed data (based on a sample) and comply with one of the standard distribution laws (Student's, χ^2 etc.) [8]. This paper uses Pierson's criterion χ^2.

Supposing that $\overrightarrow{X_n}$ is a random sample of n volume (8 years) from the population of a continuous random value X [7]. Let then one observe a discrete random value X (number of months with fires per year) taking on i of various values $u_1, \ldots, u_i$ with positive probabilities $p_1, \ldots, p_i$, (2) and (3):

$$P\{X = u_k\} = p_k, \ k = \overline{1,i}, \quad (2)$$

[1] Here and elsewhere j index with p^* defines a type of locomotives and i index characterizes a parameter of distribution within the type.

Table 1. Compliance of conditions for application of a Bernoulli distribution and calculation data

Conditions of distribution applicability	Observations	Condition compliance
Each trial has only two outcomes: «success» and «failure»	Each trial has only two outcomes: «fire occurred» and «no fire occurred»	Compliant
Independence of trials: the result of a next trial should not depend on the results of previous experiments	Independence of trials: fire occurrence does not depend on whether a fire has occurred before or not. Fire occurrence depends on maintenance	Compliant
Probability of success should be permanent (fixed) for all trials (p^*_j).	Probability of fire occurrence is a permanent value for a given type (p^*_j), in accordance with statistical observations	Compliant

Table 2. Example of check for diesel locomotives of 2TE116 type

u_i	$n_k(\overrightarrow{x_n})$	p_k	np_k	$(n_k(\overrightarrow{x_n})-np_k)^2$	$\frac{(n_k(\overrightarrow{x_n})-np_k)^2}{np_k}$
0	0	$2{,}09\cdot10^{-5}$	$1{,}67\cdot10^{-4}$	$2{,}8\cdot10^{-8}$	$1{,}67\cdot10^{-4}$
1	0	$3{,}65\cdot10^{-4}$	$2{,}92\cdot10^{-3}$	$8{,}54\cdot10^{-6}$	$2{,}92\cdot10^{-3}$
2	0	$2{,}92\cdot10^{-3}$	$2{,}34\cdot10^{-2}$	$5{,}46\cdot10^{-4}$	$2{,}34\cdot10^{-2}$
3	1	$1{,}42\cdot10^{-2}$	0,11	0,79	6,94
4	1	$4{,}64\cdot10^{-2}$	0,37	0,39	1,07
5	1	0,11	0,86	$1{,}88\cdot10^{-2}$	$2{,}18\cdot10^{-2}$
6	1	0,18	1,46	0,22	0,15
7	2	0,23	1,82	$3{,}06\cdot10^{-2}$	$1{,}68\cdot10^{-2}$
8	0	0,21	1,66	2,75	1,66
9	1	0,13	1,07	$5{,}22\cdot10^{-3}$	$4{,}87\cdot10^{-3}$
10	1	$5{,}85\cdot10^{-2}$	0,47	0,28	0,6
11	0	$1{,}55\cdot10^{-2}$	0,12	$1{,}53\cdot10^{-2}$	0,12
12	0	$1{,}87\cdot10^{-3}$	$1{,}5\cdot10^{-2}$	$2{,}25\cdot10^{-4}$	$1{,}5\cdot10^{-2}$
	$\sum_{k=1}^{i} p_k$	1		χ^2_v	10,62
				$\chi^2_{\alpha,v}$	19,67

$$\sum_{k=1}^{i} p_k = 1. \qquad (3)$$

To check a hypothesis about a binominal distribution, p_k is calculated by formula (1).

Supposing that u_k number in a sample comes across $n_k(\overrightarrow{x_n})$ times, $k=\overline{1,i}$. Note that $\sum_{k=1}^{i} n_k(\overrightarrow{x_n}) = n$.

Then Pierson's theorem is true for *v*=*i*–1 of freedom degrees (4):

$$\chi^2_v = \sum_{k=1}^{i} \frac{(n_k(\overrightarrow{x_n})-np_k)^2}{np_k}. \qquad (4)$$

If inequation (5) is satisfied:

$$\chi^2_v \le \chi^2_{\alpha,v}, \qquad (5)$$

Then a hypothesis about compliance with a binominal distribution law is accepted with confident probability α (α=0,95).

An example of calculation of check for locomotives of 2TE116 type is given in Table 2.

Let us assume that α=0,95. In our case a number of freedom degrees *v*=*i*–1=12–1=11 (*i* is a number of various values u_i). Hence $\chi^2_{\alpha,v} = 19,67$. Therefore, inequation (5) is satisfied, and we have confirmed a hypothesis about compliance of fire occurrence with a binominal distribution law at this level. This means that selection of a hypothesis is in line with experimental data.

Estimation of sufficiency of a testing volume

For each type of diesel locomotives, the sufficiency of a testing volume is estimated so that estimated probabilities would have certain reliability. Let us exemplify it by diesel locomotives of 2TE116 type. Let us state this task as follows: how many trials should be made in order to define an unknown parameter of a binominal distribution with an error not exceeding a specified value ε [5]. Let us accept an error ε at the level of $0,2p_j^*$.

The volume of a sample is calculated by formula (6):

$$n = \frac{u_\alpha^2}{\varepsilon^2} p_j^*(1 - p_j^*), \quad (6)$$

Where u_a is a quantile of standard normal distribution, α is confident probability, p_j^* is a parameter of binominal distribution.

For α=0,9, u_a=1,645. Substitute values in formula (6), we have (7):

$$n = \frac{1,645^2}{(0,2*0,495)^2} 0,495(1 - 0,495) = 69, \quad (7)$$

This testifying to the sufficiency of a trial volume as the number of months based on which a conclusion was made about p_j^* amounted to n=96.

Estimation of probability of fire occurrence on diesel locomotives of 2TE10, 2TE116, ChMEZ, TEM2 types

As it was already said, estimation of fire probability on diesel locomotives of 2TE10, 2TE116, ChMEZ, TEM2 types was made by using a binominal distribution. In this case a Bernoulli trial consisted in observation of at least one fire per month. The probability of fires occurring per year was calculated by formula (8):

$$P(B)=1-P(k=0), \quad (8)$$

Where $P(B_0)$ is a probability that no month will witness any fires to be calculated by formula (9):

$$P(B_0) = C_k^n p_г^{*k}(1 - p_г^*)^{n-k} = \frac{n!}{k!(n-k)!} p_г^{*k}(1 - p_г^*)^{n-k} = = \frac{12!}{0!12!} p_г^{*0}(1 - p_г^*)^{12}, \quad (9)$$

Where $p_г^*$ is a parameter of binominal distribution considering upper boundaries calculated depending on the type of statistics. An upper confident boundary corresponds to such parameter of binominal distribution $p_г^*$, for which it is unlikely to obtain that distribution parameter $p_j^*(p_j^* = \frac{n_{months\ with\ fire}}{n})$, which we got by experience or even smaller parameter of distribution. This means that a parameter of binominal distribution definitely does not exceed an upper value, and consequently we consider the worst variant.

Estimation of a binominal distribution parameter

Calculation of upper boundaries of a distribution parameter $p_г^*$ [6,7] was made by formula (10):

$$p_г^* = p_j^* + u_\alpha \left(\frac{p_j^*(1 - p_j^*)}{n} \right)^{1/2}, \quad (10)$$

Where u_a is a quantile of standard normal distribution, p_j^* is a j-th binominal distribution parameter, $p_j^* = \frac{n_{months\ with\ fire}}{n}$, n is a number of months, α is confident probability, i.e. the probability of p^* being in the interval constructed for it.

For confident probability α=0,95, u_a = 1,96.

Calculation by this formula is approximate and can be applied for rather a big volume of samples. Application is actually related to "double approximation": a law of distribution of $p_г^*$ parameter estimation is substituted by a normal distribution law, and an approximate value is calculated instead of a precise value. For small and medium volumes of samples, application of formula (10) can cause substantial errors. Therefore, application of this formula is just a first approximation.

Based on a resulting upper boundary of a distribution parameter and information about operating fleet, we calculated an upper boundary of fire probability by formula (9).

Calculation of fire probability and restrictions for operating fleet

The probability of fire occurrence per year on one of operating fleet locomotives was calculated by using a theorem of probabilities multiplication [2] (events being independent) by formula (11):

$$P(AB) = P(A)P(B), \quad (11)$$

Where P(A) is the probability of choosing one locomotives from operating fleet, $P(A) = \frac{1}{N}$, where N is operating fleet per month; $P(B)$ is the probability that there will be fire during a year.

Table 3 presents the main values of operating fleet.

Table 3. Value of operating fleet of various types

Type	Operating fleet (loco-days), N
2TE116	630
2TE10	1041
3TE10	153,75
2M62	237
TEP70	338
ChMEZ	2332
TEM2	1118
Average operating fleet	835

Then the average value

$$P(A)=\frac{1}{N}=\frac{1}{835*30}=3,99\cdot 10^{-5}.$$

Let us estimate constraints for operating fleet. We shall consider the task: what volume of operating fleet is necessary to define probability $P(A)$ with an error not exceeding a given value ε. The task is solved by equation (12) [9]:

$$N=\frac{u_{\alpha}^{2}}{\varepsilon^{2}}\frac{P(A)(1-P(A))}{(30)^{2}} \qquad (12)$$

For α=0,9, u_a=1,645.

Value ε will be set as 0,35$P(A)$. Then N will take a value:

$$N=\frac{1,64^{2}}{(30)^{2}(0,35*11,98\cdot 10^{-4})^{2}}11,98\cdot 10^{-4}\left(1-11,98\cdot 10^{-4}\right)=610.$$

Therefore, a minimum requisite volume of operating fleet is the value of 610 (loco-days).

Estimation of fire occurrence probability for 2M62, TEP70, 3TE10 diesel locomotives

2M62, TEP70, 3TE10 diesel locomotives presented the following features different from those of 2TE10, 2TE116, ChMEZ, TEM2 diesel locomotives:

- small operating fleet not exceeding the value of 610 and substantially smaller in numbers than other types;
- small numbers of fires during the whole observed period.

In relation to it, the calculation for 2M62, TEP70, 3TE10 diesel locomotives differed from the calculation for 2TE10, 2TE116, ChMEZ, TEM2 diesel locomotives.

For demonstration of necessity of other estimate, we'll give values obtained in calculation by method used for 2TE10, 2TE116, ChMEZ, TEP70 diesel locomotives. Table 4 presents values for an initial parameter of binominal distribution p^*, Δ of a binominal distribution parameter $\left(u_{\alpha}\left(\frac{p_j^*(1-p_j^*)}{n}\right)^{1/2}\right)$, and a parameter of binominal distribution with upper boundaries $p_г^*$ taken into account.

Table 4. Values for calculation of probability for ChMEZ, TEP70, 2TE116, 2TE10

Type	p_j^*	Δ	$p_г^*$	Operating fleet
2M62	0,041667	0,05653	0,007535	236,8919
TEP70	0,0625	0,06848	0,01748	337,975
2TE116	0,495	0,097	0,592	630
2TE10	0,693	0,089	0,783	1041,3

As seen from Table 4, Δ of 2M62 and TEP70 diesel locomotives is comparable to or exceeds a binominal distribution parameter, whereas the value Δ of 2TE10, 2TE116 diesel locomotives is considerably less than that of a binominal distribution parameter. The value Δ of 2M62 and TEP70 diesel locomotives comparable to the parameter considerable increases it when calculating a binominal distribution parameter with upper boundaries taken into account, while it gives a negative value when calculating with lower boundaries taken into account. Furthermore, for small and medium volumes of samples, as said before, substantial errors can be caused by the application of formula (10). And in this case a sample is influenced by a small operating fleet.

Therefore, the calculation of $p_г^*$ for 2M62, TEP70, 3TE10 diesel locomotives due to their difference from other types of locomotives was made by using a precise estimate [3,5].

Precise definition of confident boundaries was carried out by the following formulas. A binominal distribution parameter with upper boundaries taken into account was calculated by formula (13):

$$p^*_в==\frac{m}{nR_2}, \qquad (13)$$

Where m is a number of months with fires, n is a number of months.

The parameter R_2 is calculated by formula (14):

$$R_2=\frac{m(2n-m+\frac{1}{2}\chi_{\alpha})}{n\chi_{\alpha}}, \qquad (14)$$

Where χ_α is a quantile of chi-squared distribution with k=2(m+1) of freedom degrees, α is confident probability accepted at the level of 0,95.

A binominal distribution parameter with upper boundaries taken into account is calculated by formula (15):

$$p^*_н==\frac{m}{nR_1}, \qquad (15)$$

Where m is a number of months with fires, n is a number of months.

The parameter R_2 is calculated by formula (16):

$$R_1=\frac{m(2n-m+1+\frac{1}{2}\chi_{1-\alpha})}{n\chi_{1-\alpha}}, \qquad (16)$$

Where $\chi_{1-\alpha}$ is a quantile of chi-squared distribution with k=2m of freedom degrees, α is confident probability accepted at the level of 0,95.

Because of a small volume of operating fleet, as a resulting parameter of binominal distribution for calculating a probability for 2M62, TEP70, 3TE10 diesel locomotives, we accepted a binominal distribution parameter with upper boundaries taken into account and calculated by formula (15). Further calculation was made by analog with the calculation for 2TE10, 2TE116, ChMEZ, TEP70 by formulas (1), (11).

Conclusion

1. The paper has defined tools of statistical analysis for calculating probabilities of fire catching on diesel locomotives of various types. It has demonstrated the necessity of application of various statistical tools for calculating probabilities of fire catching on diesel locomotives of various types, with special aspects of various types taken into account: design, operating fleet. The paper has defined groups of types whose probabilities it is possible to estimate through estimating a binominal distribution parameter: 2TE10, 2TE116, ChMEZ, TEM2. For 2M62, TEP70, 3TE10 diesel locomotives, the probability of fire catching has been estimated by precisely defining confident boundaries.

2. The paper has defined a number of observation sufficient for estimating an unknown parameter of binominal distribution with an error not exceeding a given value ε at the level of $0,2p_j^*$.

References

1. GOST 12.1.033-81 SSBT. Fire safety. Terms and definitions.

2. A.V. Pechkin, O.I. Teskin, G.M. Tsvetkova et al./ Theory of probabilities: Textbook for higher school 2nd Ed. Publishing house of Bauman MGTU. 2001. – 456 p.

3. Pyatnitsky A. M./Estimation of event probability by its occurrence rate.

4. Ventsel E.S./Theory of probabilities: Textbook for higher school. – 6th Ed. st. – M.: Higher school, 1999. – 576 p.

5. Ya. B. Shor / Statistical methods of analysis and quality and reliability management: M.: Sovetskoe radio, 1962. – 553 p.

6. GOST ISO 11453-2005 Statistical methods. Statistical representation of data. Check of hypotheses and confident intervals for proportions.

7. V.B. Goryainov, I.V. Pavlov, G.M. Tsvetkova et al./ Mathematical statistics: Textbook for higher school. Publishing house of Bauman MGTU, 2008. – 424 p.

8. R 50.1.033-2001 Recommendations for standardization. Applied statistics. Rules of agreement of experimental distribution with theoretical distribution. Part 1. Chi-squared criteria.

9. R.G. Piotrovsky, K.B. Bektaev, A.A. Piotrovkaya / Mathematical linguistics. Textbook for pedagogical institutes. – M.: Higher school, 1977. – 383 p.

About the authors

Igor B. Shubinsky, Doctor of Engineering, Professor, Director of ZAO IBtrans, Moscow, Russia, phone: +7 (495) 786 68 57, e-mail: igor-shubinsky@yandex.ru

Olga B. Pronevich, Head of Division, JSC NIIAS, Moscow, Russia, phone: +7 (495) 967 77 05, ext. 516, e-mail: O.Pronevich@vniias.ru

Anna D. Danilova, Lead Specialist, JSC NIIAS, Moscow, Russia, phone: +7 (495) 967 77 05, ext. 516, e-mail: A.Danilova@vniias.ru

Received on 10.09.2016