Building a mathematical model for assessing the hazard of emergency situations and the formation of types of causes for rolling stock derailments

M.A. Novikov [1] and *L.V.* Martynenko[1*]

[1] Irkutsk State Transport University, 664074 Irkutsk, Russia

Abstract. The technique provides an analytical criterion for selecting the one set of numerical data most involved in an emergency out of several ones which are the relative deviations of the main factors from the critical ones in transport. The criterion is based on finding the theoretical left boundary of the interval of each set of deviations. The sets under study are reduced to normal distribution laws using a well-known method. The analysis uses basic numerical characteristics similar to those used in probability theory and mathematical statistics. The proposed method is based on finding the central theoretical moments up to the third order inclusive. Explanatory demonstration examples are provided.

1 Introduction

Part of the accidents may take place due to only one significant cause occurring in a certain direction, r €1,2,.. k, but at the same time, in other directions j ≠ R there may be quite small values of δ_{ji} (i € {1, 2, .. k $_j$}). The main task is to process the numerical values δ_{ri} comparing data sets. This is how a unit of measurement is introduced in [Rudanovsky]

$$d(C_i, K_j) = \sqrt{\sum_1^{m_j}(p_n - k_n)^2/k_n^2]/m_j} \tag{1}$$

where p_n, k_n are the values of the parameters of the same name, respectively, the convergence C_i and kernels of class K_j;

m_j – the number of deviations in one class.

The average value of deviations in the R-th set of deviations. On average, it characterizes the attitude towards accidents, but may be far from the truth when δ_{ri0} is equal or very close zero, and all others are δ_{rj} $(j \neq i_0)$ and reach quite large values. When $\delta_{r1} = \delta_{r2} = \ldots\ldots \delta_{rkr} = \delta_{r0}$, it turns out that $L_{r1} = \delta_{r0}$, and this, in the initial approximation, gives, albeit a rough, estimate of the direction being studied. When comparing with other directions, such averaged values will primarily be involved. A better understanding of this can be clearly felt by comparing the two directions G_1 and G_2. Involvement in an emergency situation is numerically expressed by the formula [Rudanovsky].

* Corresponding author: liuba.martinenko@yandex.ru

$$P_{11} = \frac{1/L_1}{1/L_{11}+1/L_{21}+1/d_3} = \frac{L_{21}}{L_{11}+L_{21}};\ P_{21} = \frac{1/L_{21}}{1/L_{11}+1/L_{21}} = \frac{L_{11}}{L_{11}+L_{21}} \qquad (2)$$

where L_1 is the average of the first direction, L_2 is the average of the second direction;

P_R is a probability of involvement in the accident of the R-th direction (R=1,2) . If the deviations in each group (direction) are sufficiently close to each other, these probabilistic characteristics will give a fairly accurate forecast. But with a large scatter of deviation values in each direction, these characteristics q_1 and q_2 should be considered only as preliminary, giving a rough estimate.

The subject area being studied completely coincide with a discrete random variable (Gmurman) and with a single refinement δ_{ri} $(P = 1,2,\ldots K\ ;\ i = 1,2,\ldots,n_R)$ take only positive values. The analogue of L_{R1} in mathematical statistics is $x_{\varrho 2}$ "sample average in the ϱ-th group" of variation series $x_{R1}, x_{R2}, \ldots, x_{Ri}, x_{RnR}$. All known methods of mathematical statistics are based on methods of probability theory. Therefore, in the further presentation we will adhere to the terminology and methods of mathematical statistics and probability theory. Together with the metric of the average $\delta_{ri,}$ a quadratic metric is introduced,

$$L_{R2} = \sqrt{\frac{1}{n_R}\sum_{i=1}^{n_R}\delta^2{}_{Ri}} \qquad (3)$$

It is also average, and for $\delta_{r1} = \delta_{r2} = \ldots = \delta_{RnR} = \delta_{r0}$, takes the values δ_{r0}. It is considered as an alternative to L_{R1}, and for it the probabilities of involvement in the accident are also considered using similar formulas

$$P12 = \frac{L_{R2}}{L_{R1}+L_{R2}};\ P22 = \frac{L_{R1}}{L_{R1}+L_{R2}} \qquad (4)$$

To evaluate them, let us compare $\mathrm{L_{R1}}$ and $\mathrm{L_{R2}}$. It is easy to see that

$$R_R = L^2{}_{12} - L^2{}_{11} = 1/n^2{}_R \sum_{j=1}^{n_{R-1}} \sum_{j=i+1}^{n_R} (\delta_{ri} - \delta_{Rj})^2 \geq 0$$

In mathematical statistics [Gmurman], the value R_R is called "variance in the R-th group" and is denoted as

$$D_R = 1/n_R \sum_{i=1}^{n_R} (x_{Ri} - \overline{x}_R)^2 \qquad (5)$$

Dispersion characterizes the "scatter" of the squared distances of deviations x_{ri} relative to its center – the "sample average" $\bar{x}_R$. Undoubtedly, at a greater distance from $\bar{x}_R$, the dispersion will take on larger values.

2 Methods and materials

Let's check the feasibility of formulas (1), assuming in particular $D_2 > D_1$ for $L_1 = L_2 = L_0$ that the existing relations $L_{11} = L_{21} = L_0$ are valid, and then according to formulas (1) it should be? In fact, due to the greater scattering of data, both larger and smaller L_0, the possibility of accidents increases due to $\delta_{2j} < L_0$. Consequently, an increase in dispersion only increases the propensity for the accident effect. Therefore, common sense refutes the validity of formulas like (1), which are not true in principle. First of all, this is explained not by coordinate or metric space, but by the existence of variation series $x_{R1}, x_{R2}, \ldots\ldots x_{RnR}$. Let us give a brief confirming example of a small sample size, where for a simple calculation all deviations are increased by 10 times. Example 1.

Let the sets of deviations $G_1 = \{7;5;3\}$, $G_2 = \{9;4;2\}$ be given. Easily calculated sample means

$$L_{11} = \frac{1}{3}\ (7+5+3) = 5 = \frac{1}{3}(9+4+2) = L_{21}$$

The variances of the sets are:

$$D_1 = \frac{1}{3}[(7\text{-}5)^2 + (5\text{-}5)^2 + (3\text{-}5)^2] = \frac{8}{3};\ D_2 = \frac{1}{3}[(9\text{-}5)^2 + (4\text{-}5)^2 + (2\text{-}5)^2] = \frac{26}{3}$$

For formulas (2)

$$P_{12} = \frac{\sqrt{101}}{\sqrt{83}+\sqrt{101}};\quad P_{22} = \frac{\sqrt{83}}{\sqrt{83}+\sqrt{101}}$$

But in the second set G_2 there is an element $\delta_{23} = 2$, the smallest of those presented. It is it that causes the tendency to accident. Based on the above reasoning, it follows: with equal sample means $L_{11} = L_{21}$ of two sets of deviations, the set with the largest dispersion is more involved in the accident. Here, the process of establishing priority for an accident can be called a two-stage process. At the first stage, the involvement in the accident is determined preliminary, albeit roughly, using formulas (1.1). The second stage involves analysis using variance. A more general statement is true when, with a smaller or equal sample mean of one set of deviations and a greater dispersion of this set, the mentioned set will be more involved in the accident.

Here, in the first stage, the first set of deviations (where the sample mean is smaller) dominates. According to formulas 1.1, the first set is more prone to an accident. At the second stage, due to greater dispersion, the first set of sound reasoning also tends to crash.

A quite interesting situation is the case when, at the first and second stages, equal attitudes towards accidents arise.

Let's look at a simple example 2.

Assuming the sets of deviations $G_1=\{9;4;2\}$, $G_2=\{8;6;1\}$ are given. The calculations yield $L_{11} = a_1 = 5 = a_2 = L_{21}$. In addition, the “sample range” (the difference between the largest and smallest elements) is the same for sets 9-2=7=8-1. The only fact that the accident is more involved can be the single smallest deviation value for the second set $\delta_{23} = 1$ (in the first set all $\delta_{1j} \geq = 2$ $(j=1,2,3)$). This circumstance is expressed by the negative value of the “central theoretical moment [Gmurman], which in terms of mathematical statistics has the expression

$$M_3(R) = \frac{1}{n_R} \sum_{i=1}^{n_R} (x_{Ri} - \bar{x}_R)^3 \tag{6}$$

In this case

$$M_3(1) = \frac{1}{3}[(9\text{-}5)^3 + (4\text{-}5)^3 + (2\text{-}5)^3] = 12$$

$$M_3(2) = \frac{1}{3}[(8\text{-}5)^3 + (6\text{-}1)^3 + (1\text{-}5)^3] = -12$$

These properties are universal and do not depend on sample selection volumes. The simplest patterns appear equally in large and small variation series. The explanation for the sets under consideration is simple: for significantly small extreme elements δ_{Rj} set samples G_R the value $(\delta_{Ri} - \bar{a}_R)$ will be the smallest negative, and the value $(\delta_{Ri} - \bar{a}_R)^3$ makes a decisive contribution to the sum $M_3(R)$.

In relation to the case under consideration, Statement 3 holds

With equal first two central theoretical points

$$M1(1) = M1(2),\ M2(1) = M2(2) = D0$$

A set of deviations with a lower value is more related to accidents.

A more difficult question about the involvement of two sets of deviations in the accident arises in the case of a smaller sample average of one of the sets and a smaller variance for the same set.

The fundamental factor for the analyses is the central limit theorem established by A.M. Lyapunov [Gnedenko B.V., Kolmagorov A.N., Gmurman V.E.].

According to this theorem, almost all ongoing processes are described by the normal Gauss law with a probability density [Gmurman V.E]

$$f(x) = \frac{1}{\delta\sqrt{2\pi}} \overline{e} \frac{(x-a_1)^2}{2\delta^2} \qquad (7)$$

For the specified distribution, the following properties hold [Gmurman].

3 Results and discussion

Property 1.

The mathematical expectation M [X] of a continuous random variable X(X=(- ∞; + ∞)) of the normal distribution law is equal to the parameter $\bar{\bar{a}}$.

Property 2.

The variance D [X] of a continuous random variable X of the normal distribution law is equal to δ^2. The parameter $\delta = \sqrt{D[X]}$ is called "standard deviation".

Property 3. Laplace's theorem.

a continuous random variable X of the normal distribution law falling into the interval (x_1 , x_2) is calculated using the formula Φ(t2) - Φ(t2),

where $t_1 = \frac{x_1-a}{\delta}$; $t_2 = \frac{x_2-a}{\delta}$;$\Phi_{(x)}$ - Laplace integral function.

Due to the non-existence of the last definite integral in quadratures, its values are summarized in a table given in all probability theory textbooks under the heading "Appendix 2", where the argument x is tabulated in increments of 0.01, the Laplace function with an accuracy of 10^{-4}. The Laplace integral function Φ(x) is odd.

Property 4. Three sigma rule.

For a normal distribution law, the probability of a random variable falling into the interval (a -3δ; a +3δ) is 0.9973. Almost all random variables are concentrated in a central interval of length 6δ.

Next, we set the minimum value of the argument x_0 at the distribution density of the normal distribution law, in which the probability of a random variable falling into the interval (- ∞; x_0) is equal to the probability at the point x_0 itself. To do this, we bring the argument $\bar{\bar{x}}$to the unbiased variable $z = x - a$, and conduct a search in the variables z. This condition will be written $\Phi(z_0) = - z_0 f(z_0)$. Here obviously z will be a negative value, as can be seen from the graph f(z) [Gmurman]. A comparison of equality can be carried out by differentiating with respect to z the expression Φ (z) = - z_0 f(z), assuming the derivative of a definite integral with a variable upper limit at a point equal to the value of the integrand at this point [Fichtenholtz].

As a result, we get $f(z) = - f(z) + \frac{z+z}{\delta^2} f(z)$. Considering that the exponential function does not vanish at the end points, we obtain $z^2 = 2\delta^2$.

This implies $z_1 = -\sqrt{2}\delta$, corresponding to the negative value of the interval (- ∞; z_1). Another value $z_2 = +\sqrt{2}\delta$ corresponds to the interval $G_2 = (z_2 ; + \infty)$, where the probability of a random variable falling into the interval G_2 is equal to the probability at point z_2 . Returning to the shifted variables, we obtain $x_1 = a - \sqrt{2}\delta;\ x_2 = a - \sqrt{2}\delta$. Thus, one can formulate

Property 5.

For a normal distribution law, the probability of getting into the interval coincides in modulus with the value of the probability function at the end point of the interval j_1 , also the probability of getting into the interval G_1 = (- ∞; $a - \sqrt{2}\delta$) coincides in modulus with the value of the probability function at the end point of the interval G_1 , also the probability of falling into the interval $G_2 = (a+\sqrt{2}\delta; +\infty)$ coincides with the value of the probability function at the starting point of the interval G_2.

In Appendix 2 of the values of the integral function $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-z^2/2}\,dz$ the largest value equal to 0.5, for an increasing function occurs at x=5.0. In practice, it is possible for a random variable x to fall into the interval (a -5δ; a +5δ) to be reliable. Modeling the initial variation series (set of deviations) using the laws of normal distribution greatly simplifies the task of determining priorities for the occurrence of emergency situations. In this case, the mathematical expectation α_R in mathematical statistics is interpreted as a "sample average", in our terminology d_{R1}; the standard deviation, both in probability theory and mathematical statistics, is interpreted as the square root of the variance D_R. Then for two sets of deviations G_1 and G_2 the parameters $a_j = d_{R1}$ are found; $\delta i = \sqrt{D_i}$ $(i = 1.2)$, and based on them, distribution laws with probability density functions are compiled

$$f_1(x) = \frac{1}{\delta_1\sqrt{2\pi}}\, e^{\frac{(x-a_1)^2}{2\delta_1^2}};\ f_2(x) = \frac{1}{\delta_2\sqrt{2\pi}}\, e^{\frac{(x-a_2)^2}{2\delta_2^2}} \tag{8}$$

G_1 using property 5.

Example 3.

Let $G_1 = \{7,6,2\}$ be given. The calculations performed result in a =5; D= $\frac{1}{3}$[(7-5)^2 +(6-5)^2 + (2-5)^2]=14/3; δ=$\sqrt{14/3}\sim$2.16; $b_1 = a - \sqrt{2}\delta = 5 - \sqrt{2} * 2.16 \sim 1.945$ the found numerical value of b_1 differs slightly (error 0.055) from the true one. Let's consider another example 4. Let the set of deviations {8,8,5,3} be given.

The calculations performed yield: a=6; D = $\frac{1}{4}$[(2^2 +2^2 +(-1)^2 +(-3)^2 + (2-5)^2]=9/2;

$b_1 = 6 - \sqrt{2}\sqrt{\frac{9}{2}} = 6 - 3 = 3$. The latter exactly coincides with the given one.

In some cases, the theoretical calculated bound $b_1 = a - \sqrt{2}\delta$ turns out to be significantly smaller in a given initial set. This only means that the values of the set under study are located to the right of the left border of the theoretical series. Although the approximate value $(a - \sqrt{2}\delta)$ characterizes the theoretical left boundary, $a + \sqrt{2}\delta$ will also characterize the right theoretical boundary of the data set under study. Thus, theoretically, the entire set of deviations G_R is within the theoretical interval $J_0 = (a - \sqrt{2}\delta; a + \sqrt{2}\delta)$. The length of this interval is equal to $2\sqrt{2}\delta$, which will be greater than the range of the set, denoted by T.

Indeed, even with the most unfavorable scatter: part of the G values in the amount of R_1 is located on the left edge of the interval $J_{00} = (a - \delta; a + \sqrt{2}\delta)$ another part of the G values in the amount of n - R_1 is located in the right edge of J_{00}. Then

$$D = \frac{1}{n}\sum_{i=1}^{n}\ [R_1(-\delta)^2 + (n - R_1 + \delta)^2 \tag{9}$$

In this case, the range of the data set is $T = 2\delta < 2\sqrt{2}\delta$. The given examples 3-4 express the left theoretical boundary of the interval J_0. But there will also be similar examples demonstrating the right boundary of the interval J_0. Obviously, this will be the case for the mirror image to $G_1 = \{\delta_{11}, \delta_{12}, \ldots, \delta_{1n1}\}$ relative to the center of the set $G_2 = \{\delta_{21}, \delta_{22,\ldots}\delta_{2n1}$ in which each component δ_{2i} of the second set is removed for exactly the same distance $|\delta_{1i} - a|$ from the point $\bar{\bar{a}}$, for the set G_2 the main numerical characteristics of the mathematical expectation and dispersion will be the same:

$$\frac{1}{n_1}\sum_{i=1}^{n_1}\delta 2i = \frac{1}{n_1}\sum_{i=1}^{n_1}(2a - \delta_{1i}) = \frac{2an_1 - \sum_{i=1}^{n_1}\delta_{1i}}{n_1} = 2a - \frac{1}{n_1}\sum_{i=1}^{n_1}\delta_{1i} = 2a - a = a \tag{10}$$

$$D2 = \frac{1}{n}\sum_{i=1}^{n}\ [(2a - \delta_{1i}) - a]^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(a - \delta_{1i})^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(\delta_{1i} - a)^2 = D_1 \tag{11}$$

The difference consists only in different theoretical central points of the third order:

$$M_3(2) = \frac{1}{n_1}\sum_{i=1}^{n}\left[(2a-\delta_{1i})-a\right]^3 = \frac{1}{n_1}\sum_{i=1}^{n_1}(a-\delta_{1i})^3 = -\frac{1}{n_1}\sum_{i=1}^{n_1}(\delta_{1i}-a)^3 = -M_3(1) \quad (12)$$

Let's look at supporting example 5.

Let the set $G_1=\{10;7;4;3\}$ be given. The calculations yield $a_1=L_{11}=6$; $D_1=\frac{1}{4}[(10-6)^2+(7-6)^2+(4-6)^2+(3-6)^2]=\frac{1}{4}(16+1+4+9)=\frac{30}{4}$. Hence $\delta_1=2.74$, and the value $a-\sqrt{2}\delta=6-\sqrt{15=2{,}13}$. But this value has nothing to do with the smallest value $\delta_{13}=3$. At the same time, the value $(a+\sqrt{2}\delta)=9.87$ is quite close to the largest value of the set $\delta_{11}=10$. Consequently, the set under study is "pressed" to the upper boundary of the interval J_0. Its lower bound is easy to find using the range of the data set. In this case, T=10-3=7, and the theoretical lower bound will be equal to $(a+\sqrt{2}\delta-T)=2.87$. The last value of the theoretical smallest value corresponds to the smaller value $\delta_{1i}=3$ of the data set G_1.

The mirrored row here will be $G_2=\{12\text{-}10;12\text{-}7;12\text{-}4;12\text{-}3\}=\{2;5;8;9\}$, where also $a=6$; $D=\sqrt{\frac{15}{2}}$. The previously calculated value 2.13 approximately expresses the lower limit of the set G_2, equal to $\delta_{21}=2$. Here $M_3(1)=\frac{1}{4}[(4)^3+(1)^3+(-2)^3+(3)^3]=\frac{30}{4}$ shows the shift of G_1 to the right edge of the interval J_0. The value $M_3(2)=\frac{1}{4}[(-4)^3+(-1)^3+(2)^3+(3)^3]=-\frac{30}{4}$ explains the shift of G_2 to the left edge of J_0.

To determine the left edge of the theoretical interval J_0 the formula will be used

$$b=\begin{cases}(a-\sqrt{2}\delta)\ if\ M_3\le 0\\ (a+\sqrt{2}\delta-T)\ if\ M_3>0\end{cases} \quad (13)$$

Finally, we can formulate the main criterion for the relation to the accident rate of two sets of deviations G_1, G_2, for which the main characteristics up to the third order will be:

$$a_1=L_{11},\ \delta_1=\sqrt{D_1}=\sqrt{{L_{21}}^2-{L^2}_{11}};\ M_3(1)=\frac{1}{n_1}\sum_{i=1}^{n_1}(\delta_{1i}-a_1)^3$$

span $G_1=T_1$; $a_2=L_{21}$; $\delta_2=\sqrt{D_2}$; $M_3(2)=\frac{1}{n_2}\sum_{i=1}^{n_2}(\delta_{2i}-a_2)^3$; the range $G_2=T_2$ and the values of b_1, b_2 - the theoretical left boundaries of the sets are calculated using formulas 1.3.

Statement 4.

Of the two sets of deviations G_1 and G_2, the set for which there is the smallest value of the values b_1, b_2 has a greater relation to accident rates.

To demonstrate it, let's use example 6.

Let the sets of deviations $G_1=\{7;5;3\}$, $G_2=\{8;8;2\}$ be given.

Calculation of the main characteristics of the sets is obtained: $a_1=5$; $D_1=\frac{14}{3}$;

$\delta_1\sim1.633$; $M_3(1)=0$; $a_2=6$; $D_2=8$; $\delta2\_\sim2.831$; $M_3(2)=\frac{1}{3}(2^3+2^3-4^3)=-12$. From here we find from the formulas $b_1=a_1-\sqrt{2}\delta_1\sim2.69$; $b_2=a_2-\sqrt{2}\delta_2\sim1.996$.

G_2 causes the greatest propensity for an accident. By a simple comparison of the original ones with the smallest deviation $\delta_{23}=2$, one can conclude from common sense that the second set of data has a tendency to accidents.

Statement 4 does not necessarily express a percentage. If b_1 and b_2 are equal, the involvement of two sets in the accident should be considered equal, and the final determination of the cause of the accident does not relate to numerical characteristics.

Let's consider example 7.

Let the following sets be given: $G_1=\{7;6;2\}$, $G_2=\{8;8;2\}$, which can be visually compared. The main numerical characteristics here will be the following: $a_1=5$; $D_1=\frac{14}{3}$; $\delta_1=2.16$; $M_3(1)=\frac{1}{3}(8+1-27)=-16$. Then $b_1=a_1-\sqrt{2}\delta_1\sim1.95$; $b_2=a_1-\sqrt{2}\delta_1=2$

The approximate calculations of the values of b_1 and b_2 are the same here, only with an accuracy of 0.05 the value of b_1 is less than b_2. Therefore, the first set can be considered more relevant to accident rates. Generally speaking, both sets of deviations involve the same smallest deviations $\delta_{13}=\delta_{23}=2$. Only in this case, in the set G_2, the remaining deviations are more distant from zero than the deviations G_1.

In general, we can note the consistency of statement 4 with the previous three. Using the same scheme, it is possible to select sets of deviations G_1, G_2, G_3 that are more involved in the accident, for each of which the main numerical characteristics are found: sample average; dispersion and corresponding standard deviation; central theoretical point of the third order; sample range for $M_3>0$. The quantities b_1, b_2, b_3 are calculated using formulas 1.3. Then it's true

Statement 5 (the main criterion for selecting the set involved in the accident).

Of the sets of deviations G_1, G_2, G_3, the set corresponding to the smallest value of the numbers b_1, b_2, b_3 has a greater relation to the accident rate .

Let us give an explanatory example 8.

Let the sets of deviations be given: $G_1=\{7;5;3\}$; $G_2=\{7;7;4\}$; $G_3=\{10;4;4\}$.

The main numerical characteristics of the sets are as follows: $a_1=5$; $D_1=8/3$; $\delta\sim1.633$; $M_3(1)=0$; $a_3=6$, $D_3=8$; $\delta_3\sim2.83$; $M_3(3)=+12$; $T_3=6$. Using formulas (2) we calculate: $b_1=5-\sqrt{2}*1.633\sim2.69$; $b_2=6-\sqrt{2}*1.414\sim4$; $b_3=6+\sqrt{2}*2.83-T_3\sim4$.

According to statement 5, the smallest of the numbers is: 2.69; 4; 4 is b_1, which determines the accident propensity of the first set of deviations. And in fact, the smallest deviation $\delta_{13}=3$ leads to the same conclusion.

4 Conclusion

Finally, it should be noted that in mathematical statistics the main numerical characteristics $\bar{\bar{a}}$, D , δ, M_3 ... are determined the more accurately, the larger the sample volume provided is. This is explained by some general property of the available random variables. In sets of deviations, all quantities δ_{ri} (R € { 1;2,….., K ; j € {1,2,… n_R }) do not depend on each other, and therefore the main numerical characteristics do not depend on the number of factors considered. But the analysis methods are similar.

Particular attention should be paid to the positive influence of the methods of probability theory and mathematical statistics on the analysis of numerical data.

The influence of dispersion on a greater propensity to an emergency situation has been established. For two or more sets of deviations, a criterion for the involvement of one or another set in the accident was identified. It is found by comparing the values that determine the theoretical left boundary of the sets, calculated from the main numerical characteristics of the given sets of deviations. The main criterion for involvement in an accident has been compiled for two and three sets of deviations, formulated by statements 4 and 5. The possibility of compiling involvement in an accident using basic numerical characteristics has been shown.

References

1. A. N. Baibakov, et.al, Automatic control of geometric parameters of wheelsets in train running time

2. M.F. Verigo, Interaction of track and rolling stock in small radius curves and the fight against lateral wear of rails and wheel flanges (Moscow, PTKB TsPMPS, 1997)
3. M.F. Verigo, A.Ya. Kogan, Interaction between track and rolling stock (Moscow, Transport, 1986)
4. A.Z. Venediktov, et.al, ZhDM **9** (2003)
5. S.V. Vershinsky, V.N. Danilov, V.D. Khusidov, Dynamics of a car: Textbook for universities of railway transport (Moscow, Transport, 1991)
6. I.I. Galiev, V.A. Nikolaev, B.B. Sergeev, E.A. Samokhvalov, D.Yu. Lux, News of Transsib **4(16),** 102-110 (2013)
7. L.V. Martynenko, Transport research procedure **158**, 130-134
8. Yu.S. Romen, Vestnik VNIIZhT **78(3),** 149 - 154 (2019) https://doi.org/10.21780/2223-9731-2019-78-3-149-154
9. L.V. Martynenko, Transport research procedure **63,** 465-471 (2022)
10. V.E. Gmurman, Probability theory and mathematical statistics (Moscow, Higher . school, 1999)
11. L.V. Martynenko, Transport research procedure **389** (2023) https://doi.org/10.1051/e3sconf /202338905044
12. B.V. Gnedenko, Course in Probability Theory (6th edition, revised and supplemented (Moscow, Nauka, 1988)
13. L.V. Martynenko, Transport research procedure **363** (2022) https://doi.org/10.1051/e3sconf /202236301059
14. B.V. Gnedenko, Course in Probability Theory (Moscow, Nauka, 1988)
15. L.V. Martynenko, Ensuring the safety of rolling stock movement when entering transition curves **402** (2023) https://doi.org/10.1051/e3sconf/202340206010
16. A.N. Kolmagorov, I.G. Zhurbenko, A.V. Prokhorov, Introduction to probability theory (Moscow, Publishing House MTsNMO, 2015)