Binomial distribution of railway superstructure elements

Vadim Korolev[1*]

[1]Russian University of Transport (MIIT), Chasovaya str. 22/2, 125190, Moscow, Russia

Abstract. Statistical methods of control of point components and quality of point products are an integral part of modern point production, and their application is closely connected with solving tasks of increasing labour productivity and general enterprise culture, saving materials, and reducing defects. The paper presents a derivation of basic formulas for a binomial distribution of elements of the railway track superstructure. Calculation of theoretical distribution is made. The interrelation between probability theory, statistics, and quality control of point products is considered.

1 Introduction

In order to be able to establish the extent to which a particular point-setting device's properties have changed, manufacturers or customers of point-setting devices often set some standards, under which the quality of the product is considered satisfactory. These norms (standards) usually define not only the desired properties of the signalling gear, but also upper and lower limits of their possible variation, at which the quality of the signalling gear can still be regarded as satisfactory. These upper and lower limits are called tolerances or set limits.

Many characteristics that determine the quality of railway track elements can be measured [1-3]. Such characteristics include, for example, geometric shapes and dimensions of components, chemical composition of steel, MTBF, etc.

Usually such quantitative characteristics of railway track elements are continuous random values; they, generally speaking, can have any value lying within certain defined limits [4, 5, 6]. The frequency distribution of the probabilities of each of these random variables, when the production process is controlled, is often approximately normal, although sometimes somewhat asymmetrical.

2 Research methods

When the track elements of an infinite population can take only two values (say 0 or 1), the distribution of sample sums or sample averages is often binomial. Thus, for example, if benign track elements are defined by 0 and defective track elements by 1, then the sample sum is equal to the number of defective track elements d in the sample, and the average value

* Corresponding author: shishkinaira@inbox.ru

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

of a sample of railway superstructure elements is equal to the proportion of defective railway superstructure elements *P* [7-9]. Since the control charts used to determine the number or percentage of defective track elements are usually derived from more or less periodic samples from a general population that can be considered indefinitely large, a binomial distribution is particularly common in production process control [10, 11]. This distribution is also used in quality control of a batch of point products when the value of N compared to the value of n is sufficiently large.

The theoretical distribution of the number of defective railway track elements in samples containing n railway track elements and taken from an infinite general population relating to the production process and characterised by the average proportion of defective railway track elements P is determined by a binomial decomposition (*P*+*Q*)*n*, where *Q* refers to the proportion of good quality railway superstructure elements, i.e. *P*+ *Q* = 1 [12-14]. If the sample volume is 4 (i.e., if the sample contains four track elements), then

$$(P+Q)^4 = Q^4 + 4PQ^3 + 6P^2Q^2 + 4P^3Q + P^4$$

The consecutive terms of the series determine the probability of obtaining the number of defective track elements equal to the index of degree P. In a more general way, we can write:

$$(P+Q)^n = Q^n + nPQ^{n-1} + \frac{n(n-1)}{2\cdot 1}P^2Q^{n-2} + \frac{n(n-1)(n-2)}{3\cdot 2\cdot 1}P^3Q^{n-3} + .. + P^n$$

Clearly, this expression can be converted to:

$$\boldsymbol{P}(d) = \binom{n}{d}P^dQ^{n-d} = \binom{n}{d}P^dQ^g\ , \quad (1)$$

where *d* denotes the number of bad or defective track elements in the sample, and *g* - the number of good or good quality track elements in the sample. Thus,

$$d + g = n.$$

This formula can be explained quite simply. The probability that any *d* The railway superstructure elements are defective, constitutes P^d.

The probability that any *n-d* The number of track structure elements that are found to be of good quality is equal to Q^{n-d}. Consequently, the probability that any *d* The track structure elements will be defective and others will be defective *n-d* The following are the main elements of the railway track structure P^dQ^{n-d}. However, *d* defective and *n-d* good railway track elements can be obtained $\binom{n}{d}$ by allowances. Thus, the likelihood of receiving *d* defective parts of the track superstructure and *n-d* of good railway track elements is:

$$\binom{n}{d}P^dQ^{n-d}$$

It is easy to see that in the case of *P* = *Q* = 0.5, expression (1) is converted to

$$\boldsymbol{P}(d) = \binom{n}{d}P^n$$

Thus, in this particular case, in order to determine the probability distribution, it is necessary to take the binomial coefficients and multiply by P^n.

Suppose that we extract samples of 3 railway track elements from the general population, 40% of the railway track elements that are defective [15, 16]. Substituting numerical values into formula (1) *n* and *P*, retrieve:

$$\boldsymbol{P}(d) = \binom{3}{d}(0.4)^d(0.6)^g$$

The results of the calculations are shown in table 1.

Table1. Calculating the binomial distribution ($n = 3$, $P = 0.4$)

p	d	g	$\binom{3}{d}$	$(0.4)^d$	$(0.6)^g$	$P(d) = \left[\binom{3}{d}(0.4)^d(0.6)^g\right]$	Integral probability
0.000	0	3	1	1.000	0.216	0.216	0.216
0.333	1	2	3	0.400	0.360	0.432	0.648
0.667	2	1	3	0.160	0.600	0.288	0.936
1.000	3	0	1	0.064	1.000	0.064	1.000

The first column is generally not very necessary in this table, but we have included it to show that the distribution can rather be seen as a distribution of the share of defective track elements p, than as a distribution of the number of defective track elements in the sample *d*.

The figures given in the column entitled $(0.4)^d$, are obtained by successive multiplication of the numbers starting from the first line of this column, and the numbers in the column entitled $(0.6)^g$, obtained in a similar way, but multiplied consecutively starting from the last row of the column.

The relative frequencies, whose numerical values are given in the last two columns of Table 1, are shown graphically in Figure 1. In this figure, instead of a diagram representing a line, the diagram is deliberately plotted as separate columns, as the distribution is discrete. The different shading of the bars in this figure shows how the integral distribution is obtained [17-21].

If *n* is large enough, it is better to use logarithms to make calculations easier. Taking the logarithm of both parts of formula 1, we get:

$$lg[\boldsymbol{P}(d)] = lg\binom{n}{g} + d \cdot lg\, P + g \cdot lg\, Q$$

Where

$$lg\binom{n}{g} = lg\, n! - lg\, d! - lg\, g!$$

And

$$lg[\boldsymbol{P}(d)] = lg\, n! - (lg\, d! + lg\, g! \; - \; d \cdot lg\, P - g \cdot lg\, Q) \tag{2}$$

 https://doi.org/10.1051/e3sconf/202340206013

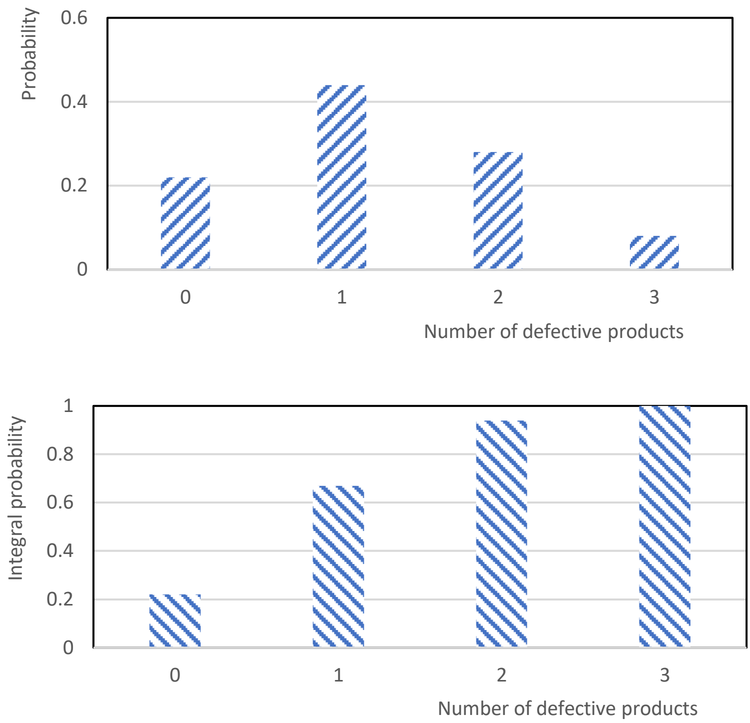

Fig. 1. Probability distribution and integral probability distribution for a binomial law (n = 3, P = 0.4)

In this expression $lg\, n!$ is a constant (for the case we have considered $lg\ 3! = lg\ 6 = 0.77815$). The sequence of probability calculations when using logarithms is shown in Table 2.

Table 2. Calculating a binomial distribution using logarithms (n = 3, P = 0.4)

d	*Q*	*lg d!*	*lg o'*	*-d lg P*	*-g lg Q*	*lg* ***P[(d)]**** $^{*)}$	***P**(d)*
0	3	0.00000	0.77815	0.00000	0.66555	-0.66555 = $\bar{1}$.33445	0.216
1	9	0.00000	0.30103	0.39794	0.44370	-0.36452 = $\bar{1}$.63548	0.432
2	1	0.30103	0.00000	0.79588	0.22185	-0.54061 = $\bar{1}$.45939	0.288
3	0	0.77815	0.00000	1.19382	0.00000	-1.19382 = $\bar{2}$.80618	0.064

$lg\ P = lg\ 0.4 = \bar{\mathbf{1}}.60206 = -0.39794;\ lg\ Q = lg\ 0.6 = \bar{\mathbf{1}}.77815 = -0.22185.$
$^{*)}$ $lg\ n! - (lg\ d! + lg\ g! - d\ lg\ P - g\ lg\ Q.$

3 Conclusion

Although the tasks assigned to production process control and quality control of arrow products have significant differences, the underlying statistical research methods are the same. However, as noted earlier, when the production process is monitored, sampling for the purpose of deciding on product quality is economically more advantageous [22-25]. If the production process is monitored, it is always possible to make a proper assessment of the quality of the products produced. Knowledge of production quality can, in turn, make it possible to select the most cost-effective quality checking plan for a batch of arrow products.

Some quality control plans involve changing the volume or frequency of sampling according to the data on the control card.

References

1. L. Benedikt Schumacher, M. Sysyn, U. Gerber, S. Fischer, Analysis of the Stressed State of Sand-Soil Using Ultrasound, Infrastructures **8(1)**, 4, (2023) https://doi.org/10.3390/infrastructures8010004
2. L. Kou, M. Sysyn, J. Liu, O. Nabochenko, Y. Han, D. Peng, S. Fischer, Evolution of Rail Contact Fatigue on Crossing Nose Rail Based on Long Short-Term Memory, Sustainability **14(24)**, 16565, (2022) https://doi.org/10.3390/su142416565
3. L. Kou, M. Sysyn, S. Fischer, J. Liu, O. Nabochenko, Optical Rail Surface Crack Detection Method Based on Semantic Segmentation Replacement for Magnetic Particle Inspection, Sensors **22(21)**, 8214, (2022) https://doi.org/10.3390/s22218214
4. V. Atapin, A. Bondarenko, M. Sysyn, et al, Monitoring and Evaluation of the Lateral Stability of CWR Track. J Fail. Anal. and Preven, **22**, 319–332 (2022). https://doi.org/10.1007/s11668-021-01307-3
5. M. Sysyn, M. Przybylowicz, O. Nabochenko, J. Liu, Mechanism of Sleeper–Ballast Dynamic Impact and Residual Settlements Accumulation in Zones with Unsupported Sleepers, Sustainability **13(14)**, 7740, (2021) https://doi.org/10.3390/su13147740
6. A. Loktev, V. Korolev, I. Shishkina, Curved Turnouts for Curves of Various Radii. In: Guda, A. (eds) Networked Control Systems for Connected and Automated Vehicles. NN 2022. Lecture Notes in Networks and Systems, vol. **509**, (2023) https://doi.org/10.1007/978-3-031-11058-0_144
7. V. Korolev, I. Shishkina, & V. Lokteva, Basic stages of creating a BIM model for transport infrastructure objects. IOP Conf. Series: Materials Science and Engineering, Vol. **918**, 012014, (2020) https://doi.org/10.1088/1757-899x/918/1/012014
8. A. Loktev, V. Korolev, & I. Shishkina, High Frequency Vibrations in the Elements of the Rolling Stock on the Railway Bridges. In IOP Conf. Series: Materials Science and Engineering, Vol. **463**, Institute of Physics Publishing, (2018) https://doi.org/10.1088/1757-899X/463/3/032019
9. V. Korolev, Switching Shunters on a Slab Base, In Advances in Intelligent Systems and Computing, vol. **1116** AISC, pp. 175–187, (2020) https://doi.org/10.1007/978-3-030-37919-3_17
10. V. Korolev, Guard Rail Operation of Lateral Path of Railroad Switch, In Advances in Intelligent Systems and Computing, vol. **1115** AISC, pp. 621–638, (2020) https://doi.org/10.1007/978-3-030-37916-2_60
11. V. Korolev, The study of rolling stock wheels impact on rail switch frogs, E3S Web of Conferences, Vol. **164**, 03033 (2020) https://doi.org/10.1051/e3sconf/202016403033
12. V. Korolev, Selecting a turnout curve form in railroad switches for high speeds of movement, In Advances in Intelligent Systems and Computing, Vol. **1258** AISC, pp. 156–172, (2021) https://doi.org/10.1007/978-3-030-57450-5_15
13. V. Korolev, Change of geometric forms of working surfaces of turnout crosspieces in wear process, In Advances in Intelligent Systems and Computing, Vol. **1258** AISC, pp. 207–218, (2021) https://doi.org/10.1007/978-3-030-57450-5_19
14. V. Korolev, Methods for Analyzing Combinations of Wheelset Sizes and Switch Elements. In: Guda, A. (eds) Networked Control Systems for Connected and

Automated Vehicles. NN 2022. Lecture Notes in Networks and Systems, vol **509**, (2023) https://doi.org/10.1007/978-3-031-11058-0_143

15. I. Shishkina, Determination of Contact-Fatigue of the Crosspiece Metal, In Advances in Intelligent Systems and Computing, Vol. **1115** AISC, pp. 834–844, (2020) https://doi.org/10.1007/978-3-030-37916-2_82
16. I. Shishkina, Hardening features for high manganese steel cores of crosspieces along the way. E3S Web of Conf., **164**, 14020, (2020) https://doi.org/10.1051/e3sconf/202016414020
17. I. Shishkina, Change of geometric and dynamic-strength characteristics of crosspieces in the operation, Advances in Intelligent Systems and Computing, Vol. **1258**, AISC, pp. 146-155, (2021) https://doi.org/10.1007/978-3-030-57450-5_14
18. I. Shishkina, Wear peculiarities of point frogs. Advances in Intelligent Systems and Computing, Vol. **1258**, AISC, pp 197-206, (2021) https://doi.org/10.1007/978-3-030-57450-5_18
19. I. Shishkina, Requirements to Check Rails of Railroad Switches. Lecture Notes in Networks and Systems, Vol. **247**, (2022) https://doi.org/10.1007/978-3-030-80946-1_18
20. A. Mikhal′chenkov, V. Komogortsev, A. Gutsan, I. Shishkina, Nature of metal drop displacement during deposit welding elements of railway transport, Transportation Research Procedia, Vol. **63**, pp. 2627-2635, (2022) https://doi.org/10.1016/j.trpro.2022.06.303
21. A. Loktev, I. Shishkina, I. Ulanov, M. Savulidi, N. Klekovkina, A. Kuznetsov, Justification of the parameters of a flexible grillage to strengthen the soft foundations of the subgrade of high-speed highways, Transportation Research Procedia, Vol. **63**, pp. 825-835, (2022) https://doi.org/10.1016/j.trpro.2022.06.079
22. A. Loktev, I. Ulanov, I. Shishkina, M. Savulidi, N. Klekovkina, A. Kuznetsov, Determination of settlement parameters of highway embankment and base consolidation time depending on soil characteristics, Transportation Research Procedia, Vol. **63**, pp. 946-955, (2022). https://doi.org/10.1016/j.trpro.2022.06.093
23. A. Loktev, V. Korolev, I. Ulanov, M. Savulidi, N. Klekovkina, A. Kuznetsov, Models of deformation behavior and analytical methods for determining settlement of weak soils, Transportation Research Procedia Vol. **63** pp. 817-824 (2022) https://doi.org/10.1016/j.trpro.2022.06.078
24. A. Loktev, V. Korolev, I. Ulanov, M. Savulidi, N. Klekovkina, A. Kuznetsov, Theoretical approaches for modeling and calculating the consolidation of a composite weak bottom, Transportation Research Procedia, Vol. **63** pp. 938-945, (2022) https://doi.org/10.1016/j.trpro.2022.06.092
25. A. Savin, A. Loktev, V. Korolev, I. Shishkina, Technology for Recovery of Defective Rail Bars, In: Bieliatynskyi A., Breskich V. (eds) Safety in Aviation and Space Technologies. Lecture Notes in Mechanical Engineering, (2022) https://doi.org/10.1007/978-3-030-85057-9_37