The selection of interval boundaries of input and output data method for obtaining complete probabilistic characteristics

Yuly Bay[1], *Anatoliy* Shmoilov[1], *Alexander* Gusev[1], and *Igor* Razzhivin[1,*]

[1]National Research Tomsk Polytechnic University, 634050 Tomsk, Russia

Abstract. The results of development and experimental investigations of the selection of interval boundaries of input and output data method are considered in the paper. The article presents an improved algorithm for obtaining a laws of probability of functional. The experimental part contains examples of finding the full probabilistic characteristics for sums of independent random variables.

1 Introduction

The electric power industry, as well as many others, frequently requires the assessment of risks of overvoltage [3, 9], accidental losses and destruction caused either by real values exceeding standard operating limits or by reduction of service life parameters relative to regulations [4]. This requires the knowledge of total probabilistic characteristics of the parameters of interest in the form of laws of probability.

It is considered rational to address mode parameters (and through them the electrical quantities during failures) through the arguments of active and reactive powers of load buses and active powers and voltages of generator buses. In this case, according to research [2], total probabilistic characteristics in the form of probability distribution functions (PDFs) and probability densities (PDs) fully comply with normal laws of probability (LPs). It is difficult to obtain representative statistics for mode parameters and electrical quantities during breakdowns; however, when the arguments' LPs and mode parameters' functional dependencies (FDs) are known, it is natural to design a method to form the LPs of the mode parameters' functional dependencies according to the LPs of the arguments of the FDs. Such method was designed and called the selection of interval boundaries of input and output data (SIBD). One of its modifications is presented in [5].

2 SIBD Logic

The principle of the SIBD method relies on the logical assumption that if all FD arguments are set as equally probable, the result of FD transformation will have the same probability [8]. This procedure can be implemented through the PDF and the PD of each argument and

---

[*] Corresponding author: lionrash@tpu.ru

the result of FD transformation. In this case, the PDF as a non-decreasing function of a random object guarantees the unambiguity of all input and output data, while the PD of this data ensures the correlation of all specified data (arguments and LDs) in the form of the equality of probability of assuming values from the intervals linked respectively with one of its boundaries to the arguments of non-decreasing one-dimensional PDFs. Consequently, second boundaries are also arguments of the specified PDFs that are determined on the basis of the equality of probability of assuming of value by all data (argument as input data and FD as output data), calculated according to the PDs of each of these data. The boundaries of the arguments and the obtained FD transformation result are rigidly fixed by quantiles of orders that are determined by the same values of their PDFs and the same probability of assuming values from these intervals. At the same time, the order of quantiles or argument's PDFs on the other (right) boundary of the intervals is also determined by adding equal probabilities of assuming of values from the intervals to equal values of PDFs on the previous (first) boundary of these intervals; while the result of the FD is determined by adding or subtracting the same probability depending on the increase or the decrease of the FD.

Further, the sequence of calculations to obtain the values of the PD of the mode parameters' functional dependence is given for the proposed fundamental validation of the SIBD method [8].

3 SIBD Algorithm

1. Preparation of types and parameters of one-dimensional PDFs and PDs of random variables $X_1, \ldots, X_i, \ldots, X_n$, i.e. $f_1(x_1), \ldots, f_i(x_i), \ldots, f_n(x_n)$ and $c_1(x_1), \ldots, c_i(x_i), \ldots, c_n(x_n)$

2. Selection of the quantity and the values of PDF quantile orders 0, $p_1, \ldots, p_j, \ldots, 1$

3. Forming the RV values as quantiles of the selected orders and values of the RVs of the functional dependence $Y = \varphi(X_1, \ldots, X_i, \ldots, X_n)$, as well as the RV PD from these quantiles 0, $p_1, p_j, : 1$. For example quantile 0:

$$x_{10} = f_1^{-1}(0), \ldots, x_{i0} = f_i^{-1}(0), \ldots, x_{n0} = f_n^{-1}(0), \; y_0 = \varphi(x_{10}, \ldots, x_{i0}, \ldots, x_{n0}),$$

$$c_1(x_{10}), \ldots, c_i(x_{i0}), \ldots, c_n(x_{n0});$$

4. The basic components of the joint PDFs and PDs of random FD values or FD quantile orders as probabilities of combining all independent RVs as quantiles of one order. For example quantile 0:

$$0 \cdots 0 \cdots 0 = f_{1 \ldots i \ldots n}(x_{10}, \ldots, x_{i0}, \ldots, x_{n0}) = f_1(x_{10}) \cdots f_i(x_{i0}) \cdots f_{in}(x_{n0}),$$

$$c_{1 \ldots i \ldots n}(x_{10}, \ldots, x_{i0}, \ldots, x_{n0}) = c_1(x_{10}) \cdots c_i(x_{i0}) \cdots c_{in}(x_{n0});$$

5. Exhaustive search, selection, and summing of joint PDFs and PDs of independent RVs as components of the functional dependence $Y = \varphi(X_1, \ldots, X_i, \ldots, X_n)$

Indices of the orders of quantiles $(j, \ldots, k, \ldots, l) = 0, p_1, \ldots, p_{j, \ldots, k, \ldots, l}, \ldots, 1$

Exhaustive search of quantile-values of RVs and calculation of the FD values compatible with the PDF and PD $y_{j, \ldots, k, \ldots, l} = \varphi(x_{1j}, \ldots, x_{ik}, \ldots, x_{nl})$, so for quantile 0

if $y_{j,\ldots,k,\ldots,l} = y_0$, then $c_{1\ldots i\ldots n}(x_{1j},\ldots,x_{ik},\ldots,x_{nl}) = c_1(x_{1j})\cdots c_i(x_{ik})\cdots c_n(x_{nl})$ is calculated and first occupies the cell $c_{1\ldots i\ldots n}(x_{10},\ldots,x_{i0},\ldots,x_{n0})$, and then, during the exhaustive search, is added to the content of this cell, etc.

After the exhaustive search of all joint RV values and the summing, the cells' content will be nothing other than the desired PD values of the functional dependence $Y = \varphi(X_1,\ldots,X_i,\ldots,X_n)$, e.g. $c_y(y_{p_j})$ is the content of the cell

$$c_{1,\ldots,i,\ldots,n}(x_{1p_j},\ldots,x_{ip_j},\ldots,x_{np_j}) + c_{1,\ldots,i,\ldots,n}^{\varphi(x_{1p_j},\ldots,x_{ip_j},\ldots,x_{np_j})}(x_{1p_j},\ldots,x_{ip_k},\ldots,x_{np_l}).$$

Qualitative representations of the beginning and the end of the exhaustive search process provide the possible charts of the PD od the functional dependence.

3 Full probabilistic characteristics for sums of independent random variables

1.   Sums of continuous (three, four) uniform random variables

With an increase in the number of RVs of the uniform density involved in the convolution, the resulting form of the LP will tend to form the normal LP [2].

In the example, we define the PD of a quantity $Z = \varphi(X + Y + V)$ and $Z = \varphi(X + Y + V + W)$. Since we still consider the quantities RVs independent, so $f(x,y,v) = f_1(x) * f_2(y) * f_3(v)$ and $f(x,y,v,w) = f_1(x) * f_2(y) * f_3(v) * f_4(w)$ respectively the laws of uniform density are also given on the section (0.1).

2.   Convolution of n uniform densities.

Carrying out the experiment and increasing the number of CAs to 4x the same result is observed.
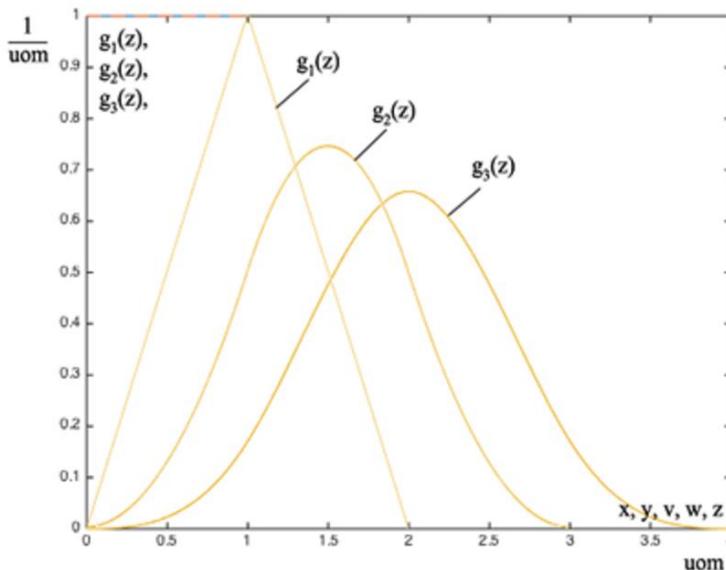
Fig. 1. PDs of the sum of three and four uniform random variables.

4 Conclusions

According to the obtained data, comparing with the already available results obtained in classical ways, it can be said with certainty that the PDs found by the SIBD method are correct.

Summation of densities with equal values of the FDs is quite easy and simple method of obtaining full probabilistic characteristics. Unlike standard methods, the method does not require integration, transformation of the original PF. Also, unlike modifications of the Monte Carlo method [1, 7], the SIBD does not require an increase in the number of tests according to the number of random arguments to obtain a statistical representation of the result with the required accuracy [6, 10].

In addition, the SIBD uses not a sample, but a entire assembly (using the method of full enumeration), which allows to capture all areas of the received total probabilistic characteristics without distortion.

References

1. A. Genz, J. Statistics. Computing, **14** (3), 251 (2003)
2. J. Hsu, *Multiple Comparisons* (London: Chapman and Hall, 1996)
3. S. Stavitsky, N. Palukhin, J. Kobenko, E. Riabova, EPJ Web Conf. **110**, 119591 (2015)
4. M. Andreev, Y. Borovikov, N. Ruban, 11th Int. Forum Strategic Technol. 7884223 (2016)
5. Y. Bai, A. Shmoilov, *Analysis of the possibilities of the development of the selection of interval boundaries of input and output data method for obtaining probabilistic characteristics of the functional dependencies of the tasks of the electric power industry* (Proceedings of the VII International Scientific and Technical Conference "Electrotechnical Energy Converters", Tomsk, 2015) [In Russian]
6. A. Suvorov, A. Gusev, A. Sulaymanov, O. Suslova, MATEC Web Conf. **91**, 01052 (2017)
7. A. Gusev, A. Suvorov, A. Sulaymanov, ICIEAM (2016)
8. Y. Bai, A. Shmoilov, *Full probabilistic characteristics of multidimensional Functional dependencies tasks of Electric Power Industry* (Proceedings of the First International Scientific Conference of Young Scientists "Electrical Engineering. Power engineering. Mechanical engineering »*, Novosibirsk, 2014) [In Russian]
9. M. Andreev, Y. Borovikov, N. Ruban, A. Khlebov, MATEC Web Conf. **91**, 010484 (2016)
10. R. Ufa, A. Vasilev, A. Gusev, A. Suvorov, 11th Int. Forum Strategic Technol. 7884204 (2017)