УДК 519.65, 51-72

EXPERIMENTAL DATA PROCESSING WITH A SMALL SAMPLE: A COMBINATION OF PROBABILISTIC AND INTERVAL ANALYSIS METHODS

Potanina T.V., Yefimov O.V. (potaninatetiana@gmail.com, AVEfimov@kpi.kharkov.ua)
National Technical University “Kharkiv Polytechnic Institute” (Ukraine)

The applications of the interval and standard probabilistic approaches for verifying the reliability of the results of an experiment studying the mechanical properties of nuclear materials are compared.

The problem of small sample size, its statistical evaluations, and the correctness of the application of existing non-parametric tests in one or another case is one of the most frequent issues of discussion. Works that discuss and propose methods and approaches to determine the optimal size of the sample so that it is representative [1] contain contradictory arguments and proofs, and sometimes erroneous ones [2]. At the same time, there remain some fields in which, for various reasons, it is impossible to increase the amount of experimental material – medicine, psychology, biophysics, research in nuclear energy, and so on. But the need to understand, describe and determine the dependencies between factors to build a forecast remains.

The analysis of experimental data is due to objective information not being complete, and the uncertainty of the data is also complicated. Uncertainty is present in any process of research and evaluation of experimental data. It contains measurement errors, noise, round-off errors, incomplete information, and methodological errors that give rise to uncertainty. It is impossible to overcome the conditions of all these uncertainties. The analysis of numerous scientific works on the processing of experimental data shows that the interval model [3] is correct if there is uncertainty in the data and the samples have small sizes [4].

The application of the interval and standard probabilistic approaches is compared to verify the reliability of the results of an experiment to study the mechanical properties of hafnium samples.

The presence of outliers in the sample of hafnium ingots hardness values was investigated at fixed weight oxygen content: 0.03 wt. %, 0.04 wt. % and 0.05 wt. %. The situation of measurement error limiting the without reliable information about its distribution and checking the sample for consistency by the methods of interval analysis is considered. The classical statistical tests were used: "three sigma rule", "two sigma rule", Lvovsky’s criterion, "three sigma rule" for the median of the sample. The application of these methods shows conflicting results since they all assume that the sample values are subject to a Gaussian distribution. But this sample of measurements is small, and the probabilistic characteristics of the errors are unknown since it is impossible to guarantee compliance with the Gaussian distribution. The correctness of the interval analysis application for processing experimental data under conditions of uncertainty and noisy experimental data and determining anomalous points is displayed.

After removing outliers using interval analysis methods, the adjusted three subsamples determined the means (probabilistic methods) $\bar{\mathbf{x}}_{\text{kor}} = (173.857, 181.0, 192.0)$ and central actual values (interval analysis method) $\mathbf{x}_{\text{c}} = (176.24, 180.24, 193.24)$. Finally, classical statistical methods, after removing outliers from the first and second subsamples, the mean values in the samples are determined, that is $\bar{\mathbf{x}} = (172.333, 182.435, 197.222)$.

The examples considered show that if the sample of experimental data is small with gross errors of observation, the interval analysis method gives better accuracy indicators than the standard statistical method.

Next, we determine the dependence of the hardness of hafnium samples on the mass content of oxygen.

Using standard statistical approaches to experimental data processing and based on the fact that the dependence is not a linear function, the parameters of the ratio $f(x) = A \cdot e^{Bx} + C$ are determined by the method of Levenberg-Marquardt minimization [5].

Curve 1 gives a rougher estimate of the dependence of the hardness of hafnium samples on the oxygen content. Curve 2 is a more accurate estimate and is close to curve 3. That is, the construction of an

approximating line at the nodes determined for samples truncated by interval analysis methods is a more efficient approach for processing observational data.

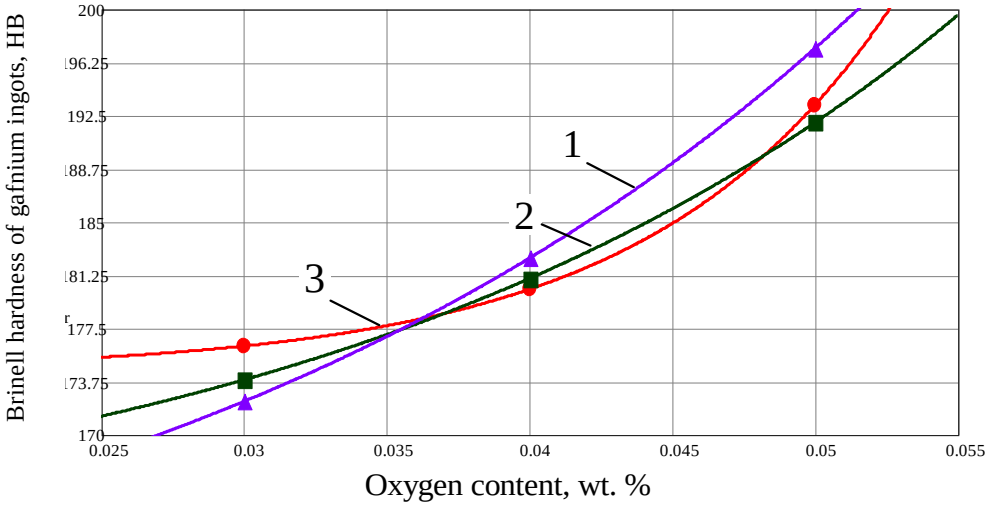

Fig. 1. The dependence of the Brinell hardness of hafnium samples on the weight oxygen content:
1 – approximation by mean values of the sample;
2 – approximation by corrected mean values of the sample; 3 – approximation by central actual values.

References

[1] V. Olefir, V. Bosniuk, "Розрахунок обсягу вибірки як наріжний камінь планування наукового дослідження," in *Visnyk of the Lviv University. Series Psychological science.* Issue 9. pp. 186–195, 2021, doi:10.30970/PS.2021.9.24.

[2] Richard Horton, "Offline: What is medicine's 5 sigma?," *The Lancet*, Vol. 385, Issue 9976, p. 1380, 2015, doi:10.1016/S0140-6736(15)60696-1.

[3] *Introduction to interval analysis*, R.E. Moore, R.B. Kearfott, M.J. Cloud, Society for Industrial and Applied Mathematics, Philadelphia, 2009.

[4] *Materials and decision support systems in the nuclear power industry*, O. Yefimov, M. Pylypenko, T. Potanina, at al., LAPLAMBERT Academic Publishing, 2020.

[5] *Methods for nonlinear least squares problems. Technical Report. Informatics and Mathematical Modeling*, K. Madsen, N.B. Nielsen, O. Tingleff, Technical University of Denmark, 2004.

УДК658.784:510.6

МОДЕЛЮВАННЯ НАДІЙНОСТІ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ НА ОСНОВІ АЛГЕБРИ АЛГОРИТМІВ І НЕЧІТКОЇ ЛОГІКИ

Ракитянська Г.Б. (rakit@vntu.edu.ua), Прус Б.В. (bohdan.prus.vntu@gmail.com)
Вінницький національний технічний університет (Україна)

Запропоновано модель надійності ПЗ на основі алгебри алгоритмів і нечіткої логіки. Логіко-алгоритмічна модель надійності ПЗ дозволяє проектувати алгоритми з необхідними рівнями безпомилковості (ризику) та витрат на основі експертно-експериментальної інформації про надійнісно-часові характеристики етапів та стадій життєвого циклу ПЗ.

Постановка задачі

Основною метою розробки програмного забезпечення (ПЗ) є надання продукту найкращої якості за умови обмеженого часу (вартості). Функціональна надійність є основним критерієм, що потребує моделювання надійності на різних етапах розробки. Висновки, отримані на основі традиційних моделей зростання надійності [1, 2], не дають розуміння практикам куди спрямувати