POSSIBILIDADE DE AUMENTAR O PODER DE ENSAIO DO QUI-QUADRADO EM PEQUENAS AMOSTRAS POR MEIOS DE TRANSIÇÃO PARA A ANÁLISE DO SEU ESPECTRO DISCRETO

POSSIBILITY TO INCREASE THE CHI-SQUARE TEST POWER ON SMALL SAMPLES BY MEANS OF TRANSITION TOWARDS ANALYZING OF IT'S DISCRETE SPECTRUM

ВОЗМОЖНОСТЬ УСИЛЕНИЯ МОЩНОСТИ ХИ-КВАДРАТ КРИТЕРИЯ НА МАЛЫХ ВЫБОРКАХ ЗА СЧЕТ ПЕРЕХОДА К АНАЛИЗУ ЕГО ДИСКРЕТНОГО СПЕКТРА

VOLCHIKHIN, Vladimir¹; IVANOV, Alexander²; GAZIN, Alexei^{3*}

^{1,2,3} Lipetsk State Pedagogical P. Semenov-Tyan-Shansky University, Lipetsk, Russian Federation

* Correspondence author

e-mail: professor-gazin@yandex.ru

Received 20 June 2019; received in revised form 06 November 2019; accepted 23 November 2019

RESUMO

O objetivo do trabalho é aumentar a potência do teste qui-quadrado, analisando o espectro dos estados de saída dessa forma discreta de teste. A recusa de uma hipótese de continuidade do espectro dos estados de saída de uma molécula qui-quadrado permite observar 22 linhas espectrais de diferentes amplitudes de probabilidade em uma amostra de 11 experimentos. Propõe-se obter um modo de suporte para a superposição quântica dos estados de saída da molécula qui-quadrado através da seleção aleatória de 11 experimentos da população geral de 21 experimentos. Isso torna possível obter cerca de 200.000 estados diferentes da molécula, o que é suficiente para estimar os estados de 8 níveis da amplitude de probabilidade de 22 linhas espectrais de saída. No caso do histograma com 6 intervalos, a hipótese de um espectro contínuo do teste do qui-quadrado para uma amostra de 11 experimentos não nos permite distinguir entre leis normais e uniformes de distribuição de dados com uma probabilidade de confiança aceitável para a prática. A situação muda drasticamente se levarmos em conta a natureza discreta do espectro dos estados da molécula qui-quadrado. É possível obter retratos digitais significativamente diferentes da distribuição de dados normal e uniforme. Suas diferenças são de cerca de 40 bits conforme a distância de Hamming. A molécula qui-quadrado pode ser interpretada como um certo neurônio quadrado com uma estrutura Gammerstein. A simulação das equações de Schrödinger em um computador comum possui complexidade computacional exponencial em relação ao número de elétrons (graus de liberdade ou experimentos). Por esse motivo, a reprodução de cálculos quânticos em 22 qubits e mais dentro da estrutura da mecânica quântica em um computador comum é tecnicamente difícil de executar. Uma situação completamente diferente surge com o apoio da superposição quântica da molécula qui-quadrado. O suporte da superposição quântica da molécula qui-quadrado possui complexidade computacional linear. Como consequência, é possível implementar cálculos quânticos de 22 qubit, 222 qubit, 2222 qubits e superiores em um computador comum.

Palavras-chave: teste qui-quadrado, amostras pequenas, espectro discreto de estados do teste qui-quadrado para amostras pequenas.

ABSTRACT

The purpose of the work is to increase the chi-square test power by analyzing the spectrum of output states of this test discrete form. The refusal from a hypothesis of a spectrum continuity of a chi-square molecule output states allows you to observe 22 spectral lines of different probability amplitudes on a sample of 11 experiments. It is proposed to obtain a mode of support for the quantum superposition of the output states of the chi-square molecule through the random selection of 11 experiments from the general population of 21 experiments. This makes it possible to obtain about 200,000 different states of the molecule, which is quite sufficient to estimate the 8-level states of the probability amplitude of 22 output spectral lines. In the case of the histogram with 6 intervals, the hypothesis of a continuous spectrum of the chi-square test for a sample of 11 experiments does not allow us to distinguish between normal and uniform laws of data distribution with a confidence probability acceptable for practice. The situation changes dramatically if the spectrum discrete nature of the chi-square molecule states is taken into account. It is possible to obtain significantly different digital

portraits of the normal and uniform data distribution. Their differences are about 40 bits as per the Hamming distance. The chi-square molecule can be interpreted as a certain square neuron with a Gammerstein structure. Simulation of Schrödinger equations on an ordinary computer has exponential computational complexity in relation to the number of electrons (degrees of freedom or experiments). For this reason, the reproduction of quantum computations on 22 qubits and more within the framework of quantum mechanics on an ordinary computer is technically difficult to perform. A completely different situation arises with the support of quantum superposition of the chi-square molecule. The support of quantum superposition of the chi-square molecule has linear computational complexity. As a consequence, it is possible to implement 22-qubit, 222-qubit, 2222-qubits, and higher quantum computations on an ordinary computer.

Keywords: *chi-square test, small samples, discrete spectrum of states of the chi-square test for small samples.*

АННОТАЦИЯ

Целью работы является повышение мощности хи-квадрат критерия за счет анализа спектра выходных состояний дискретной формы этого критерия. Отказ от гипотезы непрерывности спектра выходных состояний хи-квадрат молекулы позволяет на выборке из 11 опытов наблюдать 22 спектральных линии разной амплитуды вероятности. Предложено получать режим поддержки квантовой суперпозиции выходных состояний хи-квадрат молекулы за счет случайного выбора 11 опытов из генеральной совокупности в 21 опыт. Это позволяет получать порядка 200 000 разных состояний молекулы, что вполне достаточно для оценки 8-ми уровневых состояний амплитуды вероятности 22 выходных спектральных линий. Гистограммы с 6 интервалами гипотеза непрерывного спектра хи-квадрат критерия при выборке в 11 опытов не позволяет различать нормальный и равномерный законы распределения данных приемлемой для практики достоверной вероятностью. Ситуация кардинально меняется, если учитывать дискретный характер спектра состояний хи-квадрат молекулы. Удаётся получать значительно различающиеся цифровые портреты нормального и равномерного распределений данных. Их отличия составляют порядка 40 бит по расстоянию Хэмминга. Хи-квадрат молекулу можно интерпретировать как некоторый квадратичный нейрон со структурой Гаммерштейна. Моделирование уравнений Шредингера на обычном компьютере имеет экспоненциальную вычислительную сложность по отношению к числу электронов (степеней свободы или опытов). По этой причине воспроизведение квантовых вычислений на 22 кубит и более в рамках квантовой механики на обычном компьютере технически трудно выполнимо. Совершенно иная ситуация возникает при поддержке квантовой суперпозиции хи-квадрат молекул. Поддержка квантовой суперпозиции хи-квадрат молекулы имеет линейную вычислительную сложность. Как следствие на обычном компьютере можно реализовать квантовые вычисления на 22 кубита, 222 кубита, 2222 кубита и выше.

Ключевые слова: *хи-квадрат критерий, малые выборки, дискретный спектр состояний хи-квадрат критерия при малых выборках.*

1. INTRODUCTION

The algorithm given in (GOST R 52633.5-2011) requires about 20 examples of the "Own" image for automatic training of neural network converters "biometrics – access code". If the examples are "good", their distribution on each of the biometric parameters is close to the normal distribution law. If the data is "bad", their distribution is close to the uniform law. In order to check the normality of the distribution law of values, the classical chi-square test can be used (Kobzar, 2006; GOST R 50.1.037-2002). Unfortunately, it is possible to apply the chi-square test with a confidence probability acceptable for practice only on very large samples from 200 to 400 experiments (GOST R 50.1.037-2002).

Thus, if a sample of 11 experiments is

used, it is possible to distinguish between normal and uniform distribution laws of values by the chi-square test under the conditions of equal error probabilities of the first and second kinds $P1 \times P2 \times PEE \times 0.484$. If the sample size is increased up to 21 experiments, the values of equiprobable errors decrease to the value of $P1 \times P2 \times PEE \times 0.352$. It turns out that in the case of such small sample sizes, the confidence probability of decision making on the Pearson's chi-squared test is in the range from 0.516 to 0.848, which is not enough for practical needs. The power of the classical statistical chi-square test is low on the small sample. It is necessary to take measures which can reduce the probabilities of errors of the first and second kinds and in doing so to increase the power of the applied statistical functional on small samples.

It should be noted that one of such

methods is parallel use of several statistical tests when processing one small sample. In such a way, the application of three different statistical tests (Ivanov, 2019a; Ivanov, 2019b) allows you to reduce the probabilities of errors of the first and second kinds by about 20%. The increase of the number of tests to be united up to 12 allows you to reduce the probabilities of errors of the first and second kinds by about 93% (Ivanov, 2019b). If to generalize about 60 statistical tests, then statistical estimates on small samples should have the probabilities of errors of the first and second kinds about 0.01 and less (Ivanov, 2019b), which is quite acceptable for practice.

Another direction to increase the power of statistical estimates is the transition from the practice of continuous output data analysis of the conventional chi-square test towards analyzing the spectral lines of the discrete variant of the chi-square test. The first discrete variant of the chi-square test was discovered in 1973 (Dahiya and Gurland, 1973). The authors of this discovery did not understand what to do with their new, unexpected result. It is for this reason that the title of the article (Dahiya and Gurland, 1973) was formulated with a question mark. The problem was that the mechanical and mathematical thought in 1973 was not ready for quantum calculations, qubits, quantum superposition, and quantum entanglement. All this was formulated later in the works of Yuri Manin (1980), Richard Feynman (1981), Paul Benioff (1982) and Stephen Wiesner (1983). As a consequence, the fact that the chi-square test has a “promising” discrete form, which can be considered as a certain chi-square mathematical molecule, was forgotten for 50 years. Only in 2015, the forgotten fact was discovered again (Akhmetov *et al.*, 2015a; Akhmetov *et al.*, 2015b). Under the new conditions of the XXI century, a different situation arose, because quantum-mechanical mathematics of a large number of qubits, quantum physics, and quantum chemistry had already been well studied.

It is quite possible to be on the threshold of appearing another branch of scientific understanding of reality that is the precision quantum statistics of small samples (Kulagin *et al.*, 2016; Volchikhin *et al.*, 2017a; Ivanov, 2016; Volchikhin, 2017b; Volchikhin, 2017c). Presumably, their discrete analogues can be constructed for a set of usual traditionally continuous statistical tests and for a set of usual traditionally continuous statistical moments, having spectral lines of output states as in the case of a hydrogen molecule or molecules of

other substances (Yelashev, 2015; Yelashev, 2018). Moreover, there are initially discrete classical statistical tests (Kobzar, 2006) such as

- Wilcoxon test;
- Attila-Kersinga-Tsukini test;
- Bracheya-Gastvayda-Wright test;
- Boosa test;
- Gupta test;
- Fraser test;
- Symbolic symmetry test.

These tests initially have discrete (linear) spectra of output states. However, this fact has never been used for transition to qubits and for support of the quantum superposition for extraction of additional information from the ratio of spectral lines (probability amplitudes of quantum states occurrence of statistical molecules).

Much depends on the subjective interpretation of the experimental data. It is known that in physics, electrons can be considered both as waves and particles (the postulate of Louis de Broglie); everything depends on the position of the observer and his educational and subjective preferences. Approximately the same is observed in the statistics of small samples. It is possible to consider the values of correlation coefficients of two statistical parameters as a continuous value, however, if the observer wishes, these values turn out to be discrete. Therefore, they can be considered as states of correlational molecules (Volchikhin *et al.*, 2017d; Serikov and Kachalin, 2019) with a finite number of output spectral lines for small samples.

2. MATERIALS AND METHODS

2.1. The transition from ordinary molecules of chemistry and physics towards mathematical chi-square molecules

It is known from physics that the planetary model of a hydrogen molecule generates a series of Lyman, Balmer, Paschen and Brackett spectral lines. Each of the series is generated by electron jumps from one orbit to another, as shown in Figure 1.

Assumed that for some molecule at some temperature the most probable is to find the largest number of electrons in the fourth orbit, a chi-square molecule with a discrete spectrum of

states as in the hydrogen molecule will be obtained. In this case, six orbits can be interpreted as six histogram intervals, as shown in Figure 1.

In order to move from the usual chi-square test with the usual continuous spectrum to the chi-square molecule with the discrete output spectrum, it is necessary that the mathematical expectation of the analyzed sample should always be on the boundary of the third and fourth intervals (Dahiya and Gurland, 1973; Akhmetov *et al.*, 2015a; Akhmetov *et al.*, 2015b; Kulagin *et al.*, 2016) for the histogram of 6 intervals. Calculation of the values of the chi-square molecule for the sample of 11 experiments should be performed by Eq. 1:

$$\chi^2 = 11 \cdot \sum_{i=1}^6 \frac{\left(\frac{n_i}{11} - p_i\right)^2}{p_i} \quad (\text{Eq. 1})$$

Where n_i is the number of experiments included in the i -th interval of the histogram; p_i is the probability of hitting into the i -th interval of the histogram of the theoretical (verifiable) value distribution.

In this case, the spectrum of states in Eq. 1 is always finite; it has 21 lines of a significant level of probability amplitudes on the range of values of the chi-square test from 0 to 10. This is shown in Figure 2.

The height of the spectral lines of the chi-square molecule $\psi(\chi^2)$ corresponds to the probability amplitude of occurrence of one or another discrete chi-square state. The sum of all probability amplitudes of all spectral lines over the entire range of states is singular:

$$\Psi = \int_0^{\infty} \psi(\chi^2) \cdot d(\chi^2) = 1 \quad (\text{Eq. 2})$$

Since the spectrum of states of the considered chi-square molecules is discrete and finite, the continuous integral (Eq. 2) can always be replaced by some finite sum (Eq. 3):

$$\Psi = \int_0^{\infty} \psi(\chi^2) \cdot d(\chi^2) = 1 = \sum_{i=1}^N \psi_i(\chi^2) \quad (\text{Eq. 3})$$

where N is the number of spectral lines in the finite range from 0 to $\max(\chi^2)$.

Note that the number of spectral lines N for the chi-square molecules depends on the number of experiments in the sample and the number of intervals in the histogram. So, if the number of experiments in the sample is increased from 11 to 21, the number of significant spectral lines increases from 21 to 159, as it is shown in Figure 5.

The transition from continuous spectra to discrete spectra is a cardinal change that leads to a huge increase in the sensitivity of statistical methods of data processing. Thus, in the analysis of continuous spectra the sensitivity of the method (relative error) is quite achievable at the level of 0.1%, and the transition to the analysis of discrete spectrum lines (Yelashev, 2015; Yelashev, 2018) allows you to detect extremely small values at the level of 0.0000001% of the weight composition of impurities. In this way, forensic scientists detect gold microinclusions on the balance pan, where previously criminal gold was weighed. All you have to do is to wipe the balance pan with cotton-wool and alcohol. Then you should burn the cotton-wool, and split up the light of the flame by a prism, seeing the lines of the flame spectrum. Gold microinclusions are detected because the spectral lines of gold heated by the flame become visible, and they are between the spectral lines of burning organics (alcohol and cotton-wool).

Similar statistical relations are obtained by multivariate neural network analysis of biometric data. A neural network converter "Biometrics – Access Code" can be considered as a neural network molecule (Volchikhin *et al.*, 2017c) generating a discrete spectrum of amplitudes of the output code state probabilities. Probability estimation of errors of the second kind (the erroneous omission of "Alien" as "Own"), estimated in accordance with GOST R 52633.3, reaches values less than 0.000000001.

Approximately the same error probabilities of the first and second kinds should be achieved by statistical calculations based on a deep statistical analysis of the spectral lines of chi-square molecules. In this case, the structure of the chi-square molecule can be considered as a chi-square neuron. The structure of such a neuron is shown in Figure 4.

2.2. Maintaining the neurodynamic mode of one chi-square neuron (one chi-square molecule)

It should be noted that the model of the hydrogen molecule (Figure 3) is a dynamic structure described by the Schroedinger differential equation given in (Yelashev, 2015; Yelashev, 2018; Buravlev, 2000). The equation of the chi-square molecule (Eq. 1) is static, i.e., to realize the potential advantages of discrete spectral analysis (Akhmetov *et al.*, 2015a; Akhmetov *et al.*, 2015b; Kulagin *et al.*, 2016) it is necessary to drive the static chi-square molecule (Eq. 1) into the dynamic mode of the continuous data conversion. One of the methods of organizing the dynamics for the chi-square molecule is the random extraction of many small sub-samples of 11 experiments from the general sample of 21 experiments. This method is illustrated in Figure 5.

If repetitive small sub-samples are not used, it is possible to get

$$C_{21}^{11} = \frac{21!}{11!(21-11)!} = 352716 \quad (\text{Eq. 4})$$

a series of 11 experiments (Eq. 4). Such a large number of series of 11 experiments can be successively fed to the comparators of the chi-square neuron, providing a mode of chaotic neurodynamics, in its turn generating the appearance of spectral lines. Obviously, it is possible to feed to the input of the chi-square neuron both data with the normal distribution law of values and data with a uniform distribution of values. As a result, different spectra will be obtained, and it is possible to compare these spectra, maintaining a quantum superposition in neurodynamics.

2.3. Calculations performed in the quantum superposition support mode

One of the methods of processing the chi-square molecule spectra is the separation of the 22 most significant lines in terms of the probability amplitude by condition (Eq. 5):

$$\frac{\psi_{\text{норм}}(\chi^2) + \psi_{\text{равн}}(\chi^2)}{2} \geq 0.01 \quad (\text{Eq. 5})$$

At least one of the probability amplitudes

of normal or uniform distributions should be significant, as shown in Figure 6.

In this situation, 22 spectral lines of the chi-square molecules with a normal nuclear and a uniform nuclear will be described by the table of states 1.

The data will be processed by comparing the probability amplitudes of the chi-square molecule spectra with 8 levels of comparators corresponding to the vector of reference values {0.01, 0.02, 0.03, ..., 0.08}. During quantization, the value "0" will be assigned if the detected probability amplitude is below the threshold. If the probability amplitude is above the threshold, the code state "1" is assigned.

This transformation results in obtaining digital portraits of the chi-square molecule states in the form of 176 bits. Figure 7 shows two digital portraits of the chi-square molecule upon its excitation by data with normal and uniform laws of value distribution.

It should be noted that only significant spectral lines were taken into account in the formation of digital portraits. As a consequence, the total probability of all considered states is less than one (Eq. 6):

$$\begin{cases} \Psi_{\text{норм}} = \sum_{i=1}^{22} \psi_i(\chi^2) = 0.899 \\ \Psi_{\text{равн}} = \sum_{i=1}^{22} \psi_i(\chi^2) = 0.912 \end{cases} \quad (\text{Eq. 6})$$

3. RESULTS AND DISCUSSION:

Remaining in the traditional continuous models about the distribution of chi-square test values for the sampling of 21 experiments there is only one decision with an error probability of the first and second kinds $P1 \approx P2 \approx PEE \approx 0.352$. In case of moving over to discrete representations of data for many sub-samples of 11 experiments, there is the opportunity to make 22 decisions. There is the opportunity to make our own decision on each of the most significant probability amplitudes of 22 spectral lines.

In this case, equal error probabilities of the first and second kinds are described in the first approximation by the following ratio (Eq. 7):

$$P_{EE,i} \approx 1 - \frac{h_i}{s_i + h_i} \quad (\text{Eq. 7})$$

where i is the number of the spectral line on

which the decision is made; h_i is the number of states "*" in the Hamming portrait (Figure 7) for the i -th spectral line, s_i is the number of states "1" in the Hamming portrait (Figure 7) for the i -th spectral line.

Estimates of error probabilities in decisions made according to one spectral line are given in Table 1.

As can be seen from Table 1, approximately half of the spectral lines give worse decisions than the decisions made within the framework of the "continuous" spectrum of chi-square test states. These decisions are marked by background fill. The second half of the spectral lines gives the decisions with very good separability of the normal distribution hypothesis and the uniform distribution hypothesis in the study sample. Moreover, analysis of the spectral lines with the numbers 3, 11, 12, 13, 21, 22 should lead to almost error-free decision-making. The probabilities of detecting different spectral lines are given in Table 2.

4. CONCLUSIONS:

It should be highlighted that all the above mentioned leads to a significant complication of calculations performed in comparison with classical calculations. Formally, it is possible to consider that instead of one chi-square test calculation for a sample of 21 experiments, it will be necessary to perform about 10,000 calculations for samples in 11 experiments. Taking into account the 22 most significant spectral lines, the states of the neural network molecule, arising with a probability of 0.912 (obtained by solving simultaneous Eq. 6), are taken into account. A sample of 10,000 calculations is quite sufficient for a reliable estimation of the probability amplitude of each of the 22 spectral lines.

At the beginning of the 20th century, at the time of Pearson, a 10,000-fold complication of calculations was impossible. Today, at the beginning of the 21st century, calculations of such complexity are easily performed on ordinary desktop computers. Calculation of one Pearson's chi-squared test takes about 0.01 seconds, so the calculation of 10,000 such tests takes about 3 minutes of machine time. The costs are quite acceptable for many practical cases of statistical data processing of biometrics, biology, medicine, pharmacology, and economics

It seems that the brightest applications

may be pharmaceutical applications. The development of new drugs takes several years and costs dozens of millions of dollars. Market launch of a new drug is possible only after its sufficiently complete statistical testing, which is carried out on several hundred patients with the necessary pathology. Such testing can take several years, which increases the cost of a new drug (the payback period increases as well). Reduction of the required number of patients, on which a new drug is tested, by several times, should finally lead to a significant reduction in its cost while reducing the payback period of the new drug.

The fundamentally important point is that the computational costs of solving Schrödinger equations grow exponentially as the number of degrees of freedom or electrons in the simulated molecule grows. For the very reason, it is inexpediently to model Schrödinger equations using software methods on conventional computers. A completely different situation arises for the chi-square molecule, for which the quantum superposition of the output states is supported in a programmatic manner. Equations of mathematical molecules (chi-square molecules, correlation molecules, and excess molecules) have linear computational complexity. That is, the problem of analyzing the ratio of 22 spectral lines (22 quantum states or 22 qubits), which is considered in this article, will become 2-fold more complicated and will take about 6 minutes of machine time to analyze 44 spectral lines (44 quantum states).

The fundamental results obtained at the end of the last century for quantum-mechanical computing devices are quite applicable in practice for chi-square molecules, excess molecules, and correlation molecules. There is no sense in waiting for the appearance of full-fledged quantum computers for solving statistical problems because the results quite suitable for practical use can be obtained by software methods of quantum superposition support on ordinary desktop computers. Our article was meant to be a description of one simplest example of a similar approach implementation.

5. REFERENCES:

1. Akhmetov, B. B., Ivanov, A. I., Serikova, N. I., Funtikova Yu.V. Discrete nature of the distribution law of chi-square test for small test samples. *Bulletin of the*

- National Academy of Sciences of the Republic of Kazakhstan, **2015a**, 1, 17-25.
2. Akhmetov, B., Ivanov, A., Gilmudinov, A., Bezyaev A., Funtikova Y. The Family of Chi-Square Molecules Pearson: Software-Continuum Quantum Accelerators of High-Dimensional Calculations. *International Conference on New Directions in Multidisciplinary Research & Practice (NDMRP-15-265)*. April 22, Istanbul, Turkey. **2015b**.
 3. Dahiya, R. C., Gurland J. How many classes in the Pearson chi-square test? *Journal of the American Statistical Association*, **1973**, 68(343), 707-712.
 4. GOST R 50.1.037-2002. 2001. *Applied statistics. Rules of check of experimental and theoretical distribution of the consent. Part I. Tests of χ^2 type*. Gosstandart of Russia.
 5. GOST R 52633.3-2011. *Information protection. Information protection technology. The high reliability biometric protection means endurance testing from matching attacks*. Gosstandart of Russia. **2011**.
 6. GOST R 52633.5-2011 (The state standard of the Russian Federation 52633.5-2011). **2011**. *Information protection. Information protection technology. The neural net biometrics-code converter automatic training*. Gosstandart of Russia.
 7. Ivanov, A. I., Bannykh, A. G., Kupriyanov, E. N., Lukin, V. S., Perfilov, K. A., Savinov, K. N. Collection of Artificial Neurons Equivalent to Statistical Tests for Their Joint Use in Testing the Hypothesis of Normality of Small Samples of Biometric Data. *Proceedings of the All-Russian Scientific and Technical Conference "Information Technologies Security"*, April 24; Penza, Russia, **2019a**, 156-164.
 8. Ivanov, A. I., Kupriyanov, E. N., Tureev, S. V. Neural network generalization of the classical statistical tests for processing the small samples of the biometric data. *Reliability*, **2019b**, 2, 22-27.
 9. Ivanov, A.I. Multivariate neural network processing of biometric data with program reproduction of quantum superposition effects. *Penza Electrical Engineering Institute Publishing House*. **2016**.
 10. Kobzar, A. I., ed. *Applied Mathematical Statistics. For Engineers and Scientists*. Fizmatlit. **2006**.
 11. Kulagin, V. P., Ivanov, A. I., Gazin, A. I., Akhmetov, B. B. Cyclical continuum-quantum calculations: increase of the chi-square test power on small samples. *Analytics*, **2016**, 5(30), 22-29.
 12. Serikov, A. V., Kachalin, S. B. Correlation Molecule with Elliptical Quantizers for Calculations on Small Training Samples. *Proceedings of the 1-st All-Russian Scientific and Technical Conference*, April 24; Penza, Russia, **2019**, 123-129.
 13. Volchikhin, V. I., Ivanov, A. I. Neural network molecule: solution of the inverse biometric problem through the software support of the quantum superposition at the outputs of the artificial neuron network. *Mordovia University Bulletin*, **2017a**, 27(4), 518-523.
 14. Volchikhin, V. I., Ivanov, A. I., Bezyaev, A. V., Kupriyanov E.N. Neural network analysis of small samples of biometric data using the chi-square test and the Anderson-Darling test. *Engineering technologies and systems*, **2019**, 29(2), 2005-2017.
 15. Volchikhin, V. I., Ivanov, A. I., Gazin, A. I., Bannih, A. G. Conditions of obtaining the discrete kurtosis spectrum of statistical distributions of biometric data for small samples. *Journal of Computational and Engineering Mathematics*, **2017b**, 4(4), 53-63.
 16. Volchikhin, V. I., Ivanov, A. I., Pashchenko, D. V., Akhmetov, B. B., Vyatchanin, S. E. Opportunities of creating a cyclic continuum-quantum chi-square machine for testing statistical hypotheses on small samples of biometric and other nature data. *Bulletin of higher education institutions. Volga region. Technical sciences. Penza State University*, **2017c**, 1, 5-15.
 17. Volchikhin, V. I., Ivanov, A. I., Serikov, A. V., Serikova, Y. I. Quantum superposition of the states discrete spectrum of the mathematical correlation molecule for small samples of biometric data. *Mordovia University Bulletin*, **2017d**, 27(2), 230-243.
 18. Yelashev, M. A. Atomic and Molecular Spectroscopy. *Atomic Spectroscopy, Editorial URSS*, **2018**.
 19. Yelashev, M. A. Atomic and Molecular Spectroscopy. *Molecular Spectroscopy, Editorial URSS*, **2015**.
 20. Buravlev Yu. M. Atomic-emission Spectrometry of Metals and Alloys. *Donetsk: DonSU*, **2000**.

Table 1. Estimates of error probabilities in decisions made according to one spectral line

i	PEE
1	0.5
2	0.5
3	0.0
4	0.25
5	0.75
6	0.20
7	1.0
8	0.2
9	0.25
10	0.8
11	0.0
12	0.0
13	0.0
14	0.43
15	0.8
16	0.725
17	0.8
18	0.6
19	0.8
20	0.25
21	0.0
22	0.0

Table 2. The probability amplitudes of the spectrum No. 2 lines

Spectrum data of the chi-square molecular for 11 experiments, 6 histogram intervals, normal value distribution law of the program pseudo-random number generator																																												
D	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22																						
χ^2																																												
ϵ	0.0269	3.19	0.0233	3.47	0.0136	4.23	0.0403	4.33	0.0335	4.54	0.0557	4.60	0.1977	4.78	0.0423	4.85	0.0320	5.03	0.0920	5.06	0.0255	5.65	0.0452	5.82	0.0069	6.16	0.0372	6.40	0.0483	6.62	0.0749	6.92	0.0299	7.20	0.0257	8.48	0.0163	8.54	0.0267	9.28	0.0035	9.58	0.0017	10.96
Spectrum data of the chi-square molecular for 11 experiments, 6 histogram intervals, uniform value distribution law of the program pseudo-random number generator																																												
D	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22																						
χ^2																																												
ϵ	0.0104	3.19	0.0085	3.47	0.0076	4.23	0.0108	4.33	0.0239	4.54	0.0110	4.60	0.2071	4.78	0.0107	4.85	0.0137	5.03	0.0650	5.06	0.0062	5.65	0.0429	5.82	0.0094	6.16	0.0742	6.40	0.0359	6.62	0.1563	6.92	0.0454	7.20	0.0514	8.48	0.0604	8.54	0.0398	9.28	0.0108	9.58	0.0103	10.96

Figure 1. The planetary model of a hydrogen molecule generating a series of spectrum lines

Figure 2. Position of spectral lines of the chi-square molecule in 11 experiments and 6 histogram intervals for normal data distribution

Figure 3. Position of spectral lines of the chi-square molecule in 21 experiments and 6 histogram intervals for normal data distribution

Figure 4. The structure of the chi-square neuron with input quantization of data by comparators of the histogram interval separation

Figure 5. Creating a set of small sub-samples by random extraction of 11 experiments from a large general sample of 21 experiments (neurodynamics)

Figure 6. Eight levels quantizing the amplitude of the spectral line of the chi-square molecule normal and uniform spectra

Digital portrait of normal data distribution

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
0.08	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"
0.07	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"0"
0.06	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"0"
0.05	"0"	"0"	"0"	"0"	"0"	"1"	"1"	"0"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"0"
0.04	"0"	"0"	"0"	"1"	"0"	"1"	"1"	"1"	"0"	"1"	"0"	"0"	"0"	"1"	"1"	"1"	"0"	"0"	"0"	"0"	"0"	"0"
0.03	"0"	"0"	"0"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"0"	"0"	"0"	"1"	"1"	"1"	"1"	"0"	"0"	"0"	"0"	"0"
0.02	"1"	"1"	"0"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"0"	"0"	"0"	"1"	"1"	"1"	"1"	"1"	"0"	"0"	"0"	"0"
0.01	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"0"	"0"	"0"	"1"	"1"	"1"	"1"	"1"	"1"	"0"	"0"	"0"

Digital portrait of uniform data distribution

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
0.08	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"0"
0.07	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"0"
0.06	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"1"	"0"	"0"	"0"	"1"	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"0"
0.05	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"1"	"0"	"0"	"0"	"1"	"0"	"1"	"0"	"1"	"1"	"0"	"0"	"0"
0.04	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"1"	"0"	"1"	"0"	"1"	"0"	"1"	"1"	"1"	"1"	"1"	"0"	"0"
0.03	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"1"	"0"	"1"	"0"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"0"	"0"
0.02	"0"	"0"	"0"	"1"	"0"	"1"	"0"	"0"	"0"	"1"	"0"	"1"	"0"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"0"	"0"
0.01	"1"	"1"	"0"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"0"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"1"	"1"

Portrait of Hamming distances between normal and uniform distribution

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
0.08	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	0.3800	"0"	"0"	"0"	"0"	"0"	0.3800	"0"	"0"	"0"	"0"	"0"	"0"
0.07	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	0.3800	"0"	"0"	"0"	0.3800	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"0"
0.06	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"	"0"	"1"	"0"	"0"	"0"	0.3800	"0"	"1"	"0"	"0"	"0"	"0"	"0"	"0"
0.05	"0"	"0"	"0"	"0"	"0"	0.3800	"1"	0.3800	"0"	"1"	"0"	"0"	"0"	0.3800	"0"	"1"	"0"	0.3800	0.3800	"0"	"0"	"0"
0.04	"0"	"0"	"0"	0.3800	"0"	0.3800	"1"	0.3800	0.3800	"1"	"0"	0.3800	"0"	"1"	0.3800	"1"	0.3800	0.3800	0.3800	0.3800	"0"	"0"
0.03	"0"	"0"	"0"	0.3800	0.3800	0.3800	"1"	0.3800	0.3800	"1"	"0"	0.3800	"0"	"1"	"1"	"1"	"1"	"1"	"1"	0.3800	"0"	"0"
0.02	0.3800	0.3800	"0"	0.3800	"1"	0.3800	"1"	0.3800	0.3800	"1"	0.3800	0.3800	"0"	"1"	"1"	"1"	"1"	"1"	"1"	0.3800	"0"	"0"
0.01	"1"	"1"	0.3800	"1"	"1"	"1"	"1"	"1"	"1"	"1"	0.3800	0.3800	0.3800	"1"	"1"	"1"	"1"	"1"	"1"	"1"	0.3800	0.3800

Figure 7. Digital portraits of the chi-square molecule upon its excitation by data with normal and uniform distribution of values