# Symptom and syndrome analysis of categorial series, logical principles and forms of logic

Nina Alexeyeva and Andrey Alexeyev and Polina Gracheva
and Ekaterina Podkhalyuzina and Konstantin Usevich

*Abstract*— **The calculation of variables of one metering type by the variables of others metering types as the process leads to the forms of logic which are described by means of the collineation group of the projective geometry. Missing metering types are replaced with the appropriate logic principles. The logic of sufficiency (preferences) on the basis of orders and the logic principle of duality is considered in detail on an RNA connectivity analysis example. The minimal sum of diagonal elements in the cross-tabulation of two finitely-linear combinations (symptoms) of fragments which differ by a shift on the given quantity of symbols is chosen as a relation between two dichotomizing series. If this relation is zero then correct classification takes place. In this paper the probability of random classification with given number of errors is estimated. Logic principle of duality allows to distinguish between weak and strong statistically significant relations.**

## INTRODUCTION

One of the important problems in statistical data analysis is the problem of contrast enhancement when the statistical criteria alone are insufficient for a conclusion. Sometimes it is necessary to weaken the influence of insignificant factors and to strengthen influence of the most significant. In particular, in series factor analysis this problem is solved by means of synchronization of direct and dual problems. It appears that the use of duality for such problems may be proved by the form of logic used in the statistical data analysis.

## I. COMBINATORIC CONCEPTION OF LOGICAL PRINCIPLES AND FORMS OF LOGIC

Forms of logic are derived from four metering types: durations, sections, orders and preferences [1]. Pair-wise interaction of metering types lead to six logical principles: duality (durations and sections), differentiation (orders and preferences), exhaustion (sections and orders), uncertainty (durations and preferences), projectivity (durations and orders) and optimality (sections and preferences).

From a combinatorial standpoint, metering types taken as four points form an affine geometry $EG(2,2)$, with six connecting lines corresponding to logical principles. Infinitely remote parallel lines' intersections transform $EG(2,2)$ into a projective geometry $PG(2,2)$. In particular, an intersection of parallel lines corresponding to the principles of duality and differentiation may be viewed as an additional type of system parameters metering. Likewise, principles of exhaustion and uncertainty lead to an additional type of model structure measurement, and projectivity and optimality lead to an additional type of phenomena character measurement. The seven points in $PG(2,2)$ correspond to seven metering types: six primary plus a supplementary principle of gear that associates a model, a system and a phenomenon.

The form of logic for computing the one of the metering types arises from three other metering types, or a metering type and a logical principle of interaction between the remaining two types. Let us consider the form of logic used to derive statistical conclusions. The usual goal of such conclusions is calculation of preferences (sufficiency). Obviously, statistical data analysis strategy is based on the structure of the input data.

If the input data has a meaning of duration, e.g. survival time, relapse-free period time or some other quantitative characteristics, then the most adequate logical principle for calculation of preferences is the principle of exhaustion. That is the interaction of the lacking metering types of sections and orders. When the difference is explained in terms of mean values, it is usually sufficient to exhaust various factors that effect the relevant characteristics jointly or separately.

If the input data has a meaning of sections, when the diversity cannot be expressed as relations between mean values, the logical principle of projectivity is an adequate one, being an interaction of the lacking metering types of durations and sections [1]. In that case, the principal objects of research are synonymous distributions [2] that allow for various models, and different optimality criteria lead to a range of nominative parameters. Nevertheless the projective approach that matches the classes of adequate models allows to obtain sufficiently full notion of the processes being studied.

The order-like data such as categorical series, e.g. texts, genotypes, behavioural structures, etc., call for the logical principle of duality (durations and sections). Other logical principles may also be used, but the use of certain statistical procedures for primary objects, as well as dual ones, allows to greatly increase the contrast due to resonance.

## II. CLASSIFICATION OVER THE FINITE PROJECTIVE SPACES OF DICHOTOMIC VARIABLES

Object classification based on a set of $m$ dichotomic variables without any assumptions on their independence is carried out according to the procedure, which calculates the object rate for every possible combination of values [4]. The new object defined by the combination of values belongs to the population that has the maximal rate of objects with such a combination. Correctness of the classification may be measured by the uncertainty coefficients. The drawback

of such method is that for each population at least $2^m$ observations have to be made.

The information described by a set of variables may be compressed into a smaller number of similar variables. This may be achieved by constructing of a finite projective space of $m$ dichotomic variables and finding the most informative subspace, from a standpoint of classification problem [3].

### A. Symptoms and syndromes

Let us consider the components of a random dichotomic vector $X = (X_1, \ldots, X_m)^T$ as $m$ points of $S_{m-1} = PG(m-1, 2)$ projective geometry [5]. The projective geometry is invariant with respect to its collineation group. This means that the same space may be built based on different combinations of certain hidden variables, that are derived as linear combinations over $F_2$ field of $X_i$, $i = 1, \ldots, m$ components. In the field of statistical analysis of medico-biological data the term "point in a projective geometry" has been replaced by a more adequate term "symptom" [6], [7], [3].

*Definition 1:* A symptom is $X_\tau = \sum_{j=1}^{m} a_j X_j (mod\ 2)$, where $\tau = (t_1, \ldots, t_k) \subseteq (1, 2, \ldots, m)$, $a_i \in F_2$ and are not simultaneously equal to zero,

$$a_j = \begin{cases} 1, & \text{if} \quad j \in \tau, \\ 0, & \text{otherwise}\ . \end{cases}$$

Components $X_i$, $i = 1, \ldots, m$, of a vector $X$ with one non-zero coefficient $a_i = 1$, and $a_j = 0, j \neq i$ are trivial symptoms. The symptom with all coefficients equal to zero may be viewed as a degenerate one $X_\emptyset$, which is zero with probability 1. Non-degenerate symptoms, such that none of the sympoms is a finitely-linear combination of other symptoms are called *finitely-linear independent*.

In the terminology of dichotomic variables the projective subspaces $S_r$, $r \leq m-1$, are called *syndromes* of order $k$. Symptom $X_{\tau_1}$ is a zero-order syndrome. The syndrome of the first order $S_1$ (projective line) consists of three symptoms (points) $X_{\tau_1}, X_{\tau_2}, X_{\tau_3}$, such that $X_{\tau_3} = X_{\tau_1} + X_{\tau_2} (mod\ 2)$, $\tau_3 = \tau_1 \ominus \tau_2$. [1] Since $X_{\tau_1} = X_{\tau_2} + X_{\tau_3} (mod\ 2)$ and $X_{\tau_2} = X_{\tau_1} + X_{\tau_3} (mod\ 2)$ because $x + x (mod\ 2) = 0$, then $S_1$ may be derived from any pair of the three symptoms. Higher-order syndromes are built in an inductive manner:

$$S_r = (S_{r-1}, X_\tau, S_{r-1} + X_\tau (mod\ 2)),$$

such that $X_\tau \notin S_{r-1}$. In the syndrome $S_r$ the originating *basic* symptoms $X_\tau$ have ranks equal to powers of two. It is easy to verify that the syndrome given by a set of basic symptoms $X_{\mu_1}, \ldots, X_{\mu_{r+1}}$, such that $X_{\mu_i} \in S_r$, is equivalent to $S_r$. It's automorphism group of order $\prod_{i=0}^{r} (2^{r+1} - 2^i)$ may be presented as a group of non-singular matrices of order $r+1$ over $F_2$ field. The number of syndromes $S_{r-1}$ (or subspaces of order $r \leq m$) in $S_{m-1}$ is equal to

$$N_m^r = \prod_{k=0}^{r-1} \frac{2^m - 2^k}{(2^r - 2^k)}. \tag{1}$$

The proof is given in [5]. In particular, when $r = 1$, one has $N_m^1 = \frac{2^m - 1}{2^1 - 1} = 2^m - 1$ symptoms $S_0$ in a $S_{m-1}$ syndrome, and with $r = m$ one has $N_m^m = 1$, which corresponds to the uniqueness of a $S_{m-1}$ syndrome.

### B. Classification randomness

Consider two populations $W$ and $\overline{W}$, identified by a variable $Y$: for individuals in $W$, $Y = 1$, otherwise $Y = 0$. The number of different combinations of $Y$ components, apart from the singular one, is equal to $2^n - 1$, where $n$ is the number of individuals. Populations are characterized by $m$ dichotomic variables. Observations of these variables $X_i = (x_{i1}, \ldots, x_{in})'$, $i = 1, \ldots, m$ form an $n \times m$ matrix $X = [X_1 | \ldots | X_m]$ over $F_2$ field of rank $r$.

The problem is to classify $Y$ based on observations in $X$. By means of parameters $a = a(m, k)$, $k < m$, over $F_2$ the observation matrix $X(n, m)$ is transformed into the matrix $\widetilde{X} = Xa$. The parameters should be chosen such that the rank $k$ is minimal, and a syndrome $S_{k-1}$ produced by the columns $\widetilde{X}$ is the most informative.

This search may be posed as a solution to an optimization problem: $\mu(Xa, Y) \longrightarrow min$, where $\mu(\cdot, \cdot)$ denotes a certain distance between matrices having different number of columns. The less the distance, the better is the classification. The problem has the most simple solution when $k = 1$, so that the number of erroneous classifications is calculated using the following distance measure:

$$||b|| = \min \left\{ \frac{1}{n} \sum_{i=1}^{n} b_i, 1 - \frac{1}{n} \sum_{i=1}^{n} b_i, \right\}, \tag{2}$$

where $b = (b_1, \ldots, b_n)' = Xa + Y (mod\ 2)$.

Let's consider a set $\mathcal{L}(X)$ of all finite linear combinations $\mathcal{L}(X) = \left\{ \sum_{j=1}^{m} a_j X_j\ (mod\ 2) \right\}$. In case of a matrix of rank $m$ the power of this set is equal to $|\mathcal{L}(X)| = 2^m - 1$ a number of all possible parameter vectors, except the singular one. The probability of a random classification, which may be calculated as the ratio:

$$p_0(r, X) = \frac{|\mathcal{L}(X) \cup (\mathcal{L}(X) + \mathbf{e})|}{2^n - 1}, \tag{3}$$

depends on rank $r$ of matrix $X$ and on the fact whether $\mathcal{L}(X)$ contains the unit vector $\mathbf{e} = (1, \ldots, 1)'$ or not. Let $X = X^\oplus$, if $\mathbf{e} \in \mathcal{L}(X)$, and $X = X^\ominus$, if $\mathbf{e} \notin \mathcal{L}(X)$.

*Theorem 1:*

$$\begin{aligned} p_0(r, X^\oplus) &= \frac{2^r}{2^n - 1}, \quad r \leq m, \\ p_0(r, X^\ominus) &= \begin{cases} \frac{2^{r+1}}{2^n - 1} & r < m, \\ \frac{2^{r+1} - 2}{2^n - 1} & r = m. \end{cases} \end{aligned} \tag{4}$$

In case of a matrix of full rank $|\mathcal{L}(X)| = 2^m - 1$, and since $\mathbf{e} \notin \mathcal{L}(X)$, then $|\mathcal{L}(X) + \mathbf{e}| = 2^m - 1$, $\mathcal{L}(X) \cap (\mathcal{L}(X) + \mathbf{e}) = \emptyset$, which leads to $|\mathcal{L}(X) \cup (\mathcal{L}(X) + \mathbf{e})| = 2(2^m - 1)$. If $\mathbf{e} \in \mathcal{L}(X)$, then $\mathcal{L}(X) \cap (\mathcal{L}(X) + \mathbf{e}) = \mathcal{L}(X) \backslash \mathbf{e}$, $|\mathcal{L}(X) \backslash \mathbf{e}| =$

[1] Symbol $\ominus$ is used as a symmetric difference: $A \ominus B = A \setminus B \cup B \setminus A = (A \cup B) \setminus (A \cap B)$.

$2^m-2$. It follows that $|\mathcal{L}(X)\cup(\mathcal{L}(X)+\mathbf{e})|=2(2^m-1)-(2^m-2)=2^m$.

In case of a matrix $X$ of rank $r<m$ and $\mathbf{e}\notin\mathcal{L}(X)$ it follows that $|\mathcal{L}(X)|=2^r$, because of an additional vector $\mathbf{0}$. $|\mathcal{L}(X)+\mathbf{e}|=2^r$, because of an additional unit vector $\mathbf{e}$. $\mathcal{L}(X)\cap(\mathcal{L}(X)+\mathbf{e})=\emptyset$, which leads to, $|\mathcal{L}(X)\cup(\mathcal{L}(X)+\mathbf{e})|=2(2^r)=2^{r+1}$.

If $\mathbf{e}\in\mathcal{L}(X)$ and $r<m$, then $|\mathcal{L}(X)|=2^r$, and for any $x\in\mathcal{L}(X)$ it follows that $x+\mathbf{e}\in\mathcal{L}(X)$. □

According to [10] and (1), $M^r_{nm}=N^r_m\prod\limits_{k=0}^{r-1}(2^n-2^k)$. To calculate the probability of a random classification we need now not only the number $M^r_{nm}$ of $n\times m$ matrices over $F_2$ of rank $r$, but also the $M^{r\oplus}_{nm}$ and $M^{r\ominus}_{nm}$, the number of $X^\oplus$ and $X^\ominus$ matrices, correspondingly. It is easy to prove that $M^{r\ominus}_{nm}=N^r_m\prod\limits_{k=0}^{r-1}(2^n-2^{k+1})$, $M^{r\ominus}_{nm}=N^r_m(2^r-1)\prod\limits_{k=1}^{r-1}(2^n-2^k)$. The probability of a random classification $p_0(r)$ based on matrix $X$ of rank $r$ may be derived using a formula for total probability:

$$p_0(r)=\frac{M^{r\ominus}_{nm}}{M^r_{nm}}\cdot p_0(r,X^\ominus)+\frac{M^{r\oplus}_{nm}}{M^r_{nm}}\cdot p_0(r,X^\oplus)\,, \tag{5}$$

which leads to

$$p_0(r)=\begin{cases}\frac{2^r(2^{n+1}-2^r-1)}{(2^n-1)2^n}, & r<m,\\ \frac{(2^r-1)(2^{n+1}-2^r)}{(2^n-1)2^n}, & r=m\,.\end{cases}$$

Therefore, the following statement may be proved based on the distribution of matrices of different ranks.

*Theorem 2:* Let $p_0(r)$ be similar to (5). Then the probability of a random classification for matrix of an arbitrary rank is equal to

$$p_0=\frac{1}{2^{nm}-1}\sum_{r=1}^{m}p_0(r)M^r_{nm}\,. \tag{6}$$

Let $B$ be a set of vectors $Y_\delta$, that differ from $Y$ or $Y+\mathbf{e}$ in no more than $S$ positions, so that $Y_\delta+Y\ (mod\ 2)=\delta_s$, where $|\delta_s|=S$. Obviously, the number of $Y_\delta$ is equal to $card(B)=b=\sum\limits_{s=0}^{S}C^s_n$. The probability that at least one of the vectors $Y_\delta$ is the symptom of matrix $X$ is equal to

$$P_b=1-(1-p_0)^b\approx 1-e^{-bp_0}. \tag{7}$$

### C. Random classification for Hankel matrices

Let us consider an important classification case corresponding to a Hankel matrix $\mathbf{X}(1)$ or a generalized Hankel observation matrix $\mathbf{X}(k)$ with step $k$, that is expressed using the elements of series $x_1,\dots,x_N$:

$$\mathbf{X}(k)=\begin{bmatrix}x_1 & x_{k+1} & x_{2k+1} & \dots & x_{(m-1)k+1}\\ x_2 & x_{k+2} & x_{2k+2} & \dots & x_{(m-1)k+2}\\ \dots & \dots & \dots & \dots & \dots\\ x_n & x_{k+n} & x_{2k+L} & \dots & x_{(m-1)k+n}\end{bmatrix}. \tag{8}$$

Let us define $\Gamma^r_{nm}(k)$ as the number of general Hankel matrices with step $k$. According to [8],

$$\Gamma^r_{nm}(1)=\begin{cases}1, & r=0\\ 3\cdot 2^{2(r-1)}, & k\le r\le m-1,\\ 2^{n+m-1}-2^{2(m-1)}, & r=m.\end{cases}$$

This statement may be generalized to the case of $k>1$, in particular, when the series of elements is encoded with a pair of symbols, the case of $k=2$ is important.

According to [9] for $m>2$, $n\ge 2m-1$ we have

$$\Gamma^r_{nm}(2)=\begin{cases}1, & r=0\,,\\ 21\cdot 2^{3r-4}-3\cdot 2^{2r-3}, & 1\le r\le m-1\,,\\ 2^{n+2(m-1)}-\\ \quad -3\cdot 2^{3m-4}+2^{2m-3}, & r=m\,.\end{cases} \tag{9}$$

The numbers $\Gamma^{r\oplus}_{nm}(k)$ and $\Gamma^{r\ominus}_{nm}(k)$ represent the number of Hankel matrices $\mathbf{X}$ with step $k$, for which $\mathbf{e}\in\mathcal{L}(\mathbf{X})$ and $\mathbf{e}\notin\mathcal{L}(\mathbf{X})$, respectively. The expressions for $\Gamma^{r\oplus}_{nm}(1)$ and $\Gamma^{r\ominus}_{nm}(1)$, were derived explicitly, but the proof is not given in this work. The formulas for $\Gamma^{r\oplus}_{nm}(2)$, $\Gamma^{r\ominus}_{nm}(2)$ were obtained experimentally and were checked for a wide range of $n,m,k$.

| $r$ | $\Gamma^{r\oplus}_{nm}(1)$ | $\Gamma^{r\ominus}_{nm}(1)$ | $\Gamma^{r}_{nm}(1)$ |
|---|---|---|---|
| 0 | 0 | 1 | 1 |
| $[1;m-1]$ | $2^{2(r-1)}$ | $2\cdot 2^{2(r-1)}$ | $3\cdot 2^{2(r-1)}$ |
| $m$ | $2^{2(r-1)}$ | $2^{n+m-1}-$ $-2\cdot 2^{2(r-1)}$ | $2^{n+m-1}-$ $-2^{2(r-1)}$ |

| $r$ | $\Gamma^{r\oplus}_{nm}(2)$ | $\Gamma^{r\ominus}_{nm}(2)$ |
|---|---|---|
| 0 | 0 | 1 |
| $[1;m-1]$ | $3\cdot 2^{3r-4}-$ $-2^{2r-3}$ | $2^{3r}+2^{3(r-1)}-$ $-2^{2(r-1)}$ |
| $m$ | $3\cdot 2^{3m-4}-$ $-2^{2m-3}$ | $2^{n+2(m-1)}-$ $-3\cdot 2^{3(m-1)}+2^{2(m-1)}$ |

The expression for the probability of a random classification is derived directly from (5), where $M^r_{nm}$, $M^{r\ominus}_{nm}$ and $M^{r\oplus}_{nm}$ are replaced by $\Gamma^r_{nm}(2)$, $\Gamma^{r\ominus}_{nm}(2)$ and $\Gamma^{r\oplus}_{nm}(2)$, respectively, and the total number of matrices $2^{nm}-1$ is replaced by the total number $2^{n+2(m-1)}-1$ of Hankel matrices with step 2.

For instance, for $n=100$, $m=10$, $k=2$ according to (7) at the $\alpha=0.05$ the critical number of errors may be estimated as $S=29$. In this case $p_0=1.6140\cdot 10^{-27}\approx 1.6156\cdot 10^{-27}=2^{-n+r+1}$.

## III. Duality and Contrast

As a practical example we use the data on the RNA nucleotide sequences of three microorganisms. Using the widely accepted terminology they are denoted as $16S\ Halobact$, $16S\ Ecoli$ and $16S\ Deionoc$. The aminoacids are encoded as $a=(0,0)$, $t=(1,1)$, $g=(1,0)$ and $c=(0,1)$. Using this encoding, we obtain three binary series $(x_1,\dots,x_n)$. Since each letter is encoded into two values, the observation matrix is a Hankel matrix with step $k=2$.

We use the following dependence measures between two binary sequences $X=(x_1,\dots,x_n)'$ and $Y=(y_1,\dots,y_n)'$: one-sided $J_{x|y}$ and two-sided $J$ uncertainty coefficients [3], and $\varrho(X,Y)=1-||X+Y(mod\ 2)||$ according to (2) or the number of mistakes $K(X,Y)=n(1-\varrho(X,Y))$.

Let consider $\check{X}=(x_1,x_1+x_2,x_2+x_3,\dots,x_{n-1}+x_n)'$ as a dual series for $X=(x_1,\dots,x_n)'$. All operations are in the field $F_2$. $X$ can be derived from $\check{X}$ via a cumulative sequence $X=\{\check{x}^{\le}_j\}=\left\{\sum\limits_{i=1}^{j}\check{x}_i\right\}$. It does not matter whether

to agglomerate dual sequences or to consider the dual to a cumulative one.

Obviously, different fragments of sequences $x_1, \dots, x_N$ and $y_1, \dots, y_N$, $N > n$, have different relations. Let us pick out some fragments that we are interested in and convert them to Hankel matrices $\mathbf{X} = \mathbf{X}(2) = [X_1|\dots|X_m]$ and $\mathbf{Y} = \mathbf{Y}(2) = [Y_1|\dots|Y_t]$ according with (8). Then we choose symptoms $X_\tau$ and $Y_\mu$ ($\tau = (\tau_1, \dots, \tau_m)$, $\mu = (\mu_1, \dots, \mu_t)$, def. 1), that have maximal dependence in terms of some measure $\varrho(X_\tau, Y_\mu)$.

Therefore, the influence of penetrant parameters and the relation between a straight and a dual sequences are taken into account.

In short this method can be called *canonical symptom analysis*. The implementation of this method requires big storage budgets and a lot of computer time. And the optimization based on the parallel algorithms is another branch of research. That is why we deal just with subsequences of length $n = 100$ and the case of $t = 1$ in this work. As it was mentioned, the critical number of mistakes in this case is $S = 29$.

For example, let $X = (x_1, \dots, x_{N_1})$ and $Y = (y_1, \dots, y_{N_2})$ be RNA nucleotides sequences $16S\ Halobact$ and $16S\ Ecoli$ coding with zeros and ones. We denote the sequences with the same length $n$ and with the beginning from $n_1$ or $n_2$ by $X(n_1) = (x_{n_1}, \dots, x_{n_1+n})'$ and $Y(n_2) = (x_{n_2}, \dots, y_{n_2+n})'$.

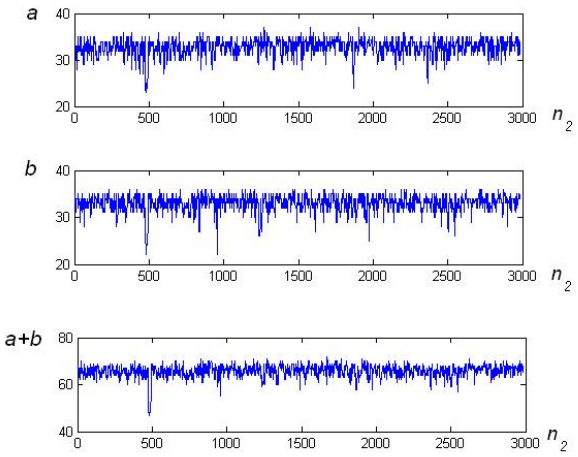


Fig. 1. Dependence of $a = K(X(n_1), Y(n_2))$, $b = K(\check{X}(n_1), \check{Y}(n_2))$ and $a + b$ on $n_2$ where $n = 100$, $n_1 = 433$.

60 cases with $K(X(433), Y(n_2))$ number of mistakes, that are less then the critical line 30 ($4.02\%$), where found during the research of sequences $X(n_1)$ and $Y(n_2)$ ($n_1 = 433$). 66 cases with $4.42\%$ can be considered in the case of dual sequences. A combination of these two characteristics let us divide this cases into two groups: strong and weak ones, according with their common or separate appearance (picture 1). Thus, the most significant measures can be found. For example,

$$K(Halobac(433), Ecoli(475)) = 23,\ p = 5.64 \cdot 10^{-5},$$
$$K(Ecoli(1), Deionoc(3)) = 8,\ p = 2.22 \cdot 10^{-16},$$
$$K(Halobac(407), Deionoc(423)) = 17,\ p = 1.34 \cdot 10^{-8}.$$

The significance levels $p$ for the given parameters calculated using Hankel matrix properties remain unchanged. A well-known fact about common fragments identifying microorganisms' RNA was chosen to illustrate the duality logical principle in action. This example shows that the statistical criteria coupled with the principle of low-probability events is not enough to determine the relations. In order to reduce the number of statistically significant factors to be interpreted one may choose only the factors that are stable with respect to the dual problem.

## REFERENCES


[1] N.Alexeyeva, A.Alexeyev, E.Verbitskaya, E.Krupitsky. "Final geometry and a logic principle of projectivity in the statistical analysis of the medical data", *in Bulletin of the international statistical institute -LXII (2007)*, 2009, pp. 3021-3024.

[2] *Alexeyeva N.P., A.O.Alexeyev A.O.* "Synonymy of power gamma-distributions in the statistical model of muscles", *in Simulation. Proceedings of the 5th St.Petersburg Workshop on Simulation. Edited by S.M.Ermakov, V.B.Melas and A.N.Pepelyshev*, 2005. pp. 39-43.

[3] N. Alexeyeva, P. Gracheva, B. Martynov, I. Smirnov. "The finitely geometric symptom analysis in the glioma survival", *In The 2nd International Conference on BioMedical Engineering and Informatics (BMEI'09) China*, 2009. 10.1109/BMEI.2009.5305560.

[4] A.A.Afifi, S.P.Asen. Statistical Analysis. *A computer Oriented Approach. Academic Press. New York · San Francisco · London;* 1979.

[5] Marshall Hall, Jr. *Combinatorial Theory.* Waltham (Massachusetts), Blaisdell Publishing; 1967.

[6] Alexeyeva N.P., Alexeyev A.O. "About a role finite geometries in the correlation analysis of binary attributes", *in Mathematical models. The theory and appendices. Iss.4. Edited by prof. Chirkov M.K.* , Saint Petersburg, 2004. pp. 102-117. (in Russian).

[7] Alexeyeva N.P. *Combinatorial analysis of two basic forms of hidden periodicity in categorial sequences*, Vestnik St. Petersburg University: Mathematics, v. 40, no. 3, 2007, pp. 193-200.

[8] J. Cherly. *Exponential sums and rank of persymmetric matrices over* $F_2$. arXiv:0711.1306v1 [math.NT], 8 Nov 2007.

[9] J. Cherly. *Exponential sums and rank of double persymmetric matrices over* $F_2$. arXiv:0711.1937v1 [math.NT], 13 Nov 2007.

[10] T. Migler, K.E. Morrison, M. Ogle. How Much Does a Matrix of Rank k Weigh? *Mathematics Magazine*, v. 79, no. 4, October 2006, pp. 262-271.