Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC EQUIPMENT OF POWER SUPPLY SYSTEMS

E.M.Farhadzadeh, Y.Z.Farzaliyev, A.Z.Muradaliyev
●
Azerbaijan Scientific-Research and Design-Prospecting Institute of Energetic
AZ1012, Ave. H.Zardabi-94, e-mail:fem1939@rambler.ru

ABSTRACT

The result of comparison of criteria, which statistics characterize differing properties of random variables of sample, depends on the importance of these properties. In turn, the importance of properties can essentially change for modeled analogues of sample

I. INSTRUCTION

It is known, that the basic requirements shown to the decision of numerous operational problems in electro power systems (EPS), is maintenance of reliability of work and decrease in operational expenses [1]. The bright example are problems of the organization of maintenance service and repair (MS&R) electric equipments. In turn, maintenance of reliability provides an opportunity of comparison estimations parameters of reliability (PR) concrete electric equipment, i.e. transition from average PR to parameters of individual reliability. Average PR electric equipments are important and traditionally used, for example, at comparison PR of schemes of projected switching centre, at an estimation of size of a reserve of capacity in EPS. At calculation PR of the concrete equipment on known, simple enough, to formulas and algorithms, experts meet essential difficulties. PR calculated on population (i.e. the average estimations) unsuitable for the decision of operational problems. In addition, data on refusals and defects of the concrete equipment are so poor, that or do not allow to calculate PR, or accuracy of estimations appears unacceptable. Therefore, maintenance of reliability of work in practice carried out, unfortunately, mainly at an intuitive level. The certain contribution to the decision of this problem is brought by the automated information systems providing information support of dispatching personnel EPS. But the objective estimation of parameters is still actually on individual reliability.

II. FEATURES OF STATISTICAL DATA

Necessary to note, that at the analysis of reliability of equipment EPS classification of statistical data of operation on one, and sometimes and to the two attributes, set by nameplate data and data of conditions of operation, it is spent. For example, in [2] are resulted PR electric equipments of a various class of a voltage. Are occasionally resulted PR electric equipment grouped as or to purpose, a design, service life, a manufacturer and other attributes. Classification of statistical data more than to two attributes does not practice. The reason for that is the variety of versions of attributes (VA) and decrease in accuracy of estimations PR (increase in width of a confidential interval). Decrease in accuracy occurs within the limits of the assumption of conformity of statistical data to casual sample of some general population.
Actually:
1. The statistical data describing reliability of equipment EPS (data on non-working conditions), depend on the big number of passport and operational data (installation sites, a class of a voltage, a design, service life, etc.) and consequently cannot be considered neither as analogue of general population, nor as final sample of homogeneous data. In the mathematician such data it is accepted

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

to name multivariate. Unfortunately, analytical methods of the analysis of multivariate data developed only for the assumption of conformity of distribution of random variables to some to one, to mainly normal law of distribution. It at all does not correspond to the real histograms of distribution constructed on statistical data of operation of electric equipment. As an example on fig.1 histograms of duration of emergency switching-off are resulted (τ_a) power units 300MVt [3]. The first histogram characterizes distribution according to operation of eight power units for the period 1992-2006 years. The second histogram characterizes distribution P*(τ_a) all power units for 2005 year. The number of cases of emergency switching-off for this sample has decreased with 634 up to 48. On the third histogram distribution P* is shown $P^*(\tau_a)$ for the first power unit in 2005year. 2.
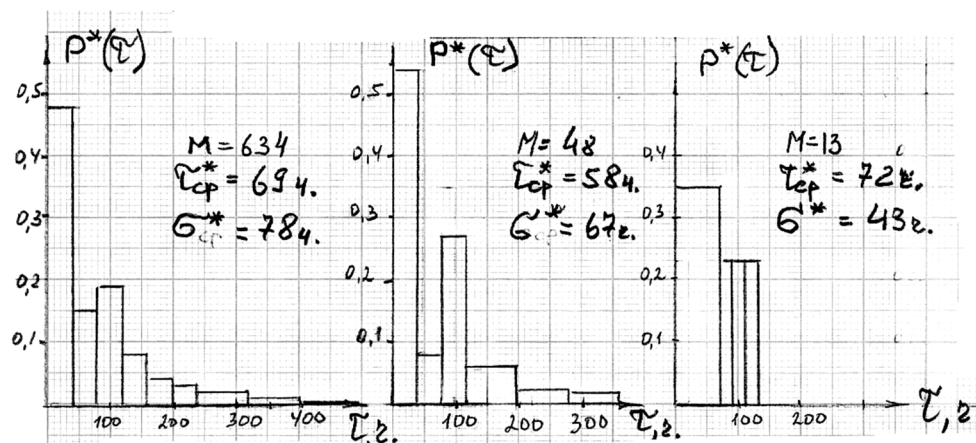

Fig.1 Histograms of duration of emergency switching-off of power units 300MVt

Comparison of character of change of these histograms and laws of change of normal distribution confirms small probability of conformity P*(τ_a) to one concrete and, in particular, to the normal law of distribution.

3. At classification of multivariate statistical data on set VA, selective data taken from final population of multivariate data <u>not casually</u>. For example, all switches with rated voltage 110Kv not casually get out. We shall specify this feature. Not casual sample:
- consists of random variables;
- number of random variables of sample n_v is casual, changes in time, for example, increases;
- features of distribution in an interval of change random variables final population of multivariate data depends from VA;

3. The type of the law of distribution of final population of multivariate statistical data not only is not known. It regularly casually varies in process of accumulation of statistical data

4. <u>The interval of change random variable in sample of final population of multivariate statistical data on set VA is no more, than an interval of change random variable in the most final population</u>. We shall remind, that for general population of a random variable the average quadratic deviation always is less, than the average quadratic deviation for any on number of representative sample and with reduction of number of random variables in sample, n_v an estimation of an average quadratic deviation increases. These features allow concluding, that <u>application of classical methods of the analysis samples from general population for the analysis samples from final population of multivariate data it is necessary to be careful</u>.

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

III. ABOUT SET STATISTIC, DESCRIBING RANDOM VARIABLES OF SAMPLE

The most objective approach to the decision of the statistical problems arising at classification of multivariate data is application of computer modeling possible samples and checks of assumptions (hypotheses) about expedient classification of data on everyone VA. Difficulties arise at an estimation of expediency of classification of data. As matter of fact - this know problem about a finding significant VA. We spent the decision of this problem within the limits of methodology of the theory of check of statistical hypotheses. In mathematical statistics, it considered two types of the problems connected with comparison of functions of distribution:

1. Check of the assumption that sample of random variables X casually taken from general population of random variables with the set type of distribution $F_\Sigma(X)$.

2. Check of the assumption of uniformity two or several sample the random variables casually taken from same general populations with the known continuous law of distribution $F_\Sigma(X)$.

The estimation of expediency of classification of multivariate statistical data offered to be carried out by comparison of statistical functions of distribution (s.f.d.). Final population of multivariate data $F_\Sigma^*(X)$ with s.f.d. samples $F_V^*(Y)$, of this population. Comparison $F_\Sigma^*(X)$ also $F_V^*(Y)$ theoretically carried out on number of numerical characteristics of a random variable a vertical divergence of distributions $F_\Sigma^*(X)$ and $F_V^*(Y)$, which we shall designate as $\Delta(Y)$ also we shall calculate under the formula:

$$\Delta(Y_i) = F_\Sigma^*(Y_i) - F_V^*(Y_i) \tag{1}$$

where: $1 < \Delta(Y_i) < 1$ c $i = 1, n_v$

According to the established practice, these numerical characteristics we shall name statistics and we shall designate S (Δ). To S(Δ) concern:

1. The greatest vertical divergence between $F_\Sigma^*(X)$ and $F_V^*(Y)$. It is calculated on following algorithm:

    1.1. n_v realizations $\Delta(Y)$ are placed in ascending order;

    1.2. Absolute values of the first and n_v -th values ranking of some $\Delta(Y)$ are compared and the greatest value is defined Δ_m;

    1.3. It also is the greatest vertical deviation $F_\Sigma^*(X)$ and $F_V^*(Y)$ with the sign;

2. Average value of a vertical divergence. It is calculated under the formula:

$$M^*[\Delta(Y)] = n_v^{-1} \sum_{i=1}^{n_v} |\Delta(Y_i)| = \Delta_{av}^* \tag{2}$$

It is necessary to note, that

$$M^*[\Delta(Y)] \neq n_v^{-1} \sum_{i=1}^{n_v} \Delta(Y_i) \tag{3}$$

As under this formula, average value of a random variable $\Delta(Y)$, instead of an average deviation is calculated. Distinction between formulas (2) and (3) shown, when among realizations $\Delta(Y)$ there are both positive, and negative sizes.

3. Average quadratic deviation $\Delta(Y)$. Are calculated under the formula:

$$\sigma^*[\Delta(Y)] = \sqrt{\frac{\sum_{i=1}^{n_v} [\Delta_{av}^* - |\Delta(Y_i)|]^2}{(n_v - 1)}} = \Delta_{ck}^* \tag{4}$$

By analogy with p.2

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

$$\sigma^*[\Delta(Y)] \neq \sqrt{\frac{\sum_{i=1}^{n_v}[\Delta_{av}^* - \Delta(Y_i)]^2}{(n_v - 1)}} \qquad (5)$$

4. Scope of dispersion of a random variable $L_V^*(\Delta)$. It is calculated under the formula:

$$L_V^*(\Delta) = \Delta_{max} - \Delta_{min} \qquad (6)$$

This list could be continued. But also it is enough resulted numerical characteristics for an illustration of the mechanism of comparison efficiency of criteria check assumptions of character a divergence s.f.d. $F_\Sigma^*(X)$ and $F_V^*(Y)$

IV. A QUESTION ON EFFICIENCY OF CRITERIA

According to the established practice efficiency of criteria is characterized by function of capacity of criterion $W[S(\Delta)]$. In turn $W[S(\Delta)] = 1 - \beta[S(\Delta)]$, where- $\beta[S(\Delta)]$ - an error II type for statistics $S(\Delta)$. The essence of considered criteria is same and is reduced to comparison of empirical value $S_9(\Delta)$ with boundary value of distribution $R[S_{\alpha(\Delta)}] = \alpha[S(\Delta)]$, where $\alpha[S(\Delta)]$ - an error I type.

It is accepted, for the fixed value $\alpha[S(\Delta)]$ to consider criterion as more effective if its function of capacity has the greatest value. So that to compare with efficiency of criteria it is enough to construct dependences $W[S(\Delta)]$ from $\alpha[S(\Delta)]$ and to compare $W[S(\Delta)]$ for $0 < \alpha[S(\Delta)] < 1$

The algorithm of construction of this dependence reduced to following calculations:

1. Construction s.f.d. realizations of statistics $S_1(\Delta)$ for initial assumption H_1 according to which distributions $F_\Sigma^*(X)$ also $F_V^*(Y)$ differ casually. We shall designate this distribution as $F^*[S_1(\Delta)]$. The sequence of calculations, features of imitating modeling of realizations representative sample, results of calculations for of some n_v is resulted in [4] on an example of statistics of the greatest vertical deviation Δ_m;

2. Here the sequence of construction s.f.d. is resulted. Realizations of statistics $S_2(\Delta)$ for assumption H_2 according to which distributions $F_\Sigma^*(X)$ also $F_V^*(Y)$ differ not casually.

3. It is systematized realizations $F^*[S_1(\Delta)]$ and $F^*[S_2(\Delta)]$ at $S_1(\Delta) = S_2(\Delta)$. As quantile distributions $F^*[S_1(\Delta)]$ are not equal quantile distributions $F^*[S_2(\Delta)]$, performance p.3. It appears impossible. The analysis of realizations quantile these distributions after ranking shows, that distinction of some realizations $S_1(\Delta)$ and $S_2(\Delta)$ takes place not less than in the fourth category of their quantitative estimations. If to neglect this difference, the number of equal realizations $S_1(\Delta)$ and $S_2(\Delta)$ reaches 10%. Unfortunately, this quantity is often not enough for the full characteristic of dependence $W[S(\Delta)] = \varphi\{\alpha[S(\Delta)]\}$. The decision of this problem found the basis of the assumption of linear character of change s.f.d. Intervals between quantile distributions.

Considering, that the number quantile estimated in hundreds, the size of an entered error calculations corresponding quantile probabilities appears less accuracy of calculation quantile.
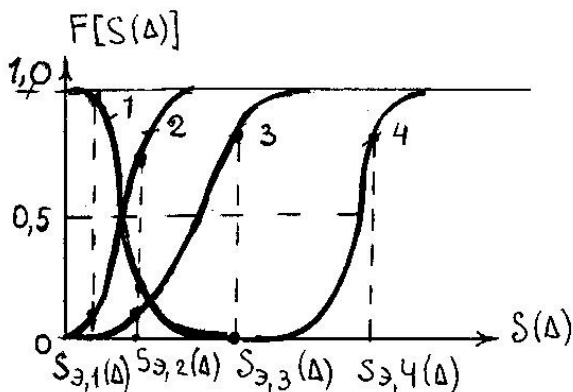
Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September
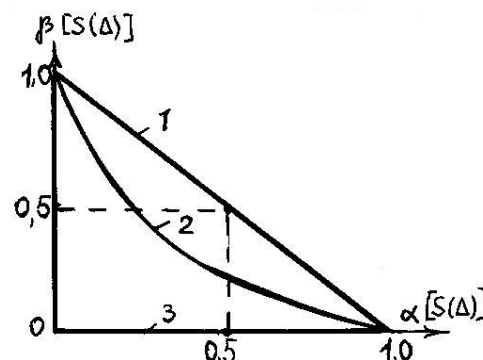
Fig.2. Curves   **Fig.3.** Typical dependences

1. $R_1[S(\Delta)] = 1-F_1[S(\Delta)]$; $M_1[S(\Delta)]$
2. $F_2[S(\Delta)]$; $M_2[S(\Delta)] = M_1[S(\Delta)]$
3. $F_3[S(\Delta)]$; $M_3[S(\Delta)] = 2.0M_1[S(\Delta)]$
4. $F_4[S(\Delta)]$; $M_4[S(\Delta)] \gg M_1[S(\Delta)]$

$\beta[S(\Delta)] = \varphi\{\alpha[S(\Delta)]\}$
Are constructed according to $R_1[S(\Delta)]$ and
1. $F_2[S(\Delta)]$; 2. $F_3[S(\Delta)]$; 3. $F_4[S(\Delta)]$

Let's consider features of application of this approach to the sample analysis from final population of multivariate data on set VA. On fig.2 typical functions of statistic distribution, describing a divergence $F_\Sigma^*(X)$ are resulted and $F_V^*(Y)$. As simplification, s.f.d. F * [S (Δ)] are represented by continuous functions of distribution. Three variants of sample distributions are shown. Curves 2 and 4 characterize limiting parities s.f.d. Final population of multivariate data $F_\Sigma^*(X)$ and s.f.d. the second and the fourth sample $F_{V,2}^*(X)$ and $F_{V,4}^*(Y)$.

The parity $F_\Sigma^*(X)$ also $F_{V,2}^*(Y)$ characterizes a case, when functions of distribution $[1-R_1[S(\Delta)]]$ and $F_2[S(\Delta)]$ are practically identical, and a parity $F_\Sigma^*(X)$ and $F_{V,4}^*(Y)$ - a case, when a divergence $[1-R_1[S(\Delta)]]$ and $F_4[S(\Delta)]$ it is not casual. A parity of functions of distribution $R_1[S(\Delta)]$ and $F_3[S(\Delta)]$ borrows intermediate position.

As follows from fig.2:

As $R_1[S_{Э,1}(\Delta)] \gg F_2[S_{Э,1}(\Delta)]$, $H \Rightarrow H_1$
$R_1[S_{Э,2}(\Delta)] \ll F_2[S_{Э,2}(\Delta)]$, $H \Rightarrow H_2$
$R_1[S_{Э,3}(\Delta)] \ll F_3[S_{Э,3}(\Delta)]$, $H \Rightarrow H_2$ (7)
$R_1[S_{Э,2}(\Delta)] > F_3[S_{Э,3}(\Delta)]$, $H \Rightarrow H_1$

For these parities are constructed and represented on fig.3 dependences $\beta[S(\Delta)] = \varphi\{\alpha[S(\Delta)]\}$. In particular, a curve 1 according to $R_1[S(\Delta)]$ and $F_2[S(\Delta)]$, a curve 2 according to $R_1[S(\Delta)]$ and $F_3[S(\Delta)]$ a curve 3-on data $R_1[S(\Delta)]$ and $F_4[S(\Delta)]$.

V. EXPERIMENTAL RESEARCHES

In practice for statistical check of the assumption that sample of random variables X is casually taken from general population of random variables with the set law of distribution $F_\Sigma(X)$, the greatest distribution was received Kolmogorov's based on statistics Dn [4] with the criterion. This criterion concerns to group nonparametric. In other words, this criterion with success can be used as for comparison $F_\Sigma(X)$ and $F_v^*(X)$, $F_\Sigma^*(X)$ and $F_v^*(X)$. Formulas and tables for application of this criterion are resulted in many monographers and manuals. And practically in all these sources the inaccuracy of a finding of size of the greatest vertical divergence of distributions $F_\Sigma(X)$

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

and $F_v^*(X)$ is marked as maximal value absolute sizes of observable values Δ. However, in one of many seen managements on mathematical statistics <u>the reason of this mistake</u> is not underlined.

The analysis of statistics Δ_m

Realizations of statistics Δ_m, were calculated on following algorithm [5].

- Pay off $\qquad\qquad \Delta_i = (\xi_i - i/n_v); \; i=1,n_v$
- Are defined $\qquad\quad \Delta_{m,1} = \min \{\Delta_1, \Delta_2,\ldots, \Delta_i, \ldots, \Delta_{nv}\}$

$\qquad\qquad\qquad\qquad \Delta_{m,1} = \max \{\Delta_1, \Delta_2.., \Delta_i, \ldots, \Delta_{nv}\}$ $\qquad\qquad\qquad$ (8)
- If $\qquad\qquad\qquad |\Delta_{m,1}| > |\Delta_{m,2}|$, that $\Delta_m = \Delta_{m,1}$
- Differently $\qquad\quad \Delta_m = \Delta_{m,2}$

Here ξ- random numbers with uniform expansion in an interval [0,1], simulating true values quantile uniform distribution. S.f.d. $F^*(\Delta_m)$ constructed on 25000 realizations Δ_m and for of some n_v are resulted on fig.4. Importance of these researches consist first of all that with sufficient accuracy for practice borders of change could be established Δ_m, describing the greatest vertical divergence s.f.d. $F_\Sigma(\xi)$ and $F_v^*(\xi)$ with the set significance value. And by that to have an opportunity to estimate character of a divergence any s.f.d. $F_\Sigma(X)$ and $F_v^*(X)$.
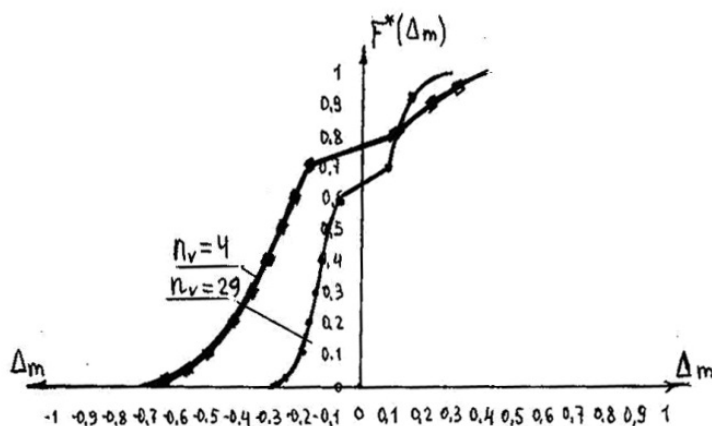

Fig.4. Statistical functions of distribution $F^*(\Delta_m)$ for n_v=4 both 29
and number of iterations N=25000

It is established, that quantile distributions $F^*(\Delta_m) = \alpha \leq 0,1$ for $n \geq 2$ are equal on size and opposite on a sign quantile distributions $F(D_n) = 2\alpha$. In other words D_n though characterizes the greatest divergence of expansions $F_\Sigma(X)$ and $F_v^*(X)$, but there is no the greatest vertical divergence of these distributions. Distributions Δ_m it is original. Here Δ_m it considered simply as some numerical value, ранжированное in ascending order. If Δ_m to consider as the greatest vertical divergences s.f.d. $F_\Sigma(\xi)$ and $F_v^*(\xi)$, function $F^*(\Delta_m)$ - not is s.f.d. The reason for that presence of positive and negative values Δ_m. And than more negative value Δ_m, i.e. than divergence $F_\Sigma(\xi)$ and $F_v^*(\xi)$ there is more, that probability of acceptance of hypothesis H_1 ($F_\Sigma(\xi)$ and $F_v^*(\xi)$ miss casually) there will be more, and a probability of event Δ_m=0 essentially more zero. Laws of change of distribution of positive and negative values Δ_m for n_v=4 are resulted on fig. 5, and a parity of their number for of some n_v - in table 1.

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
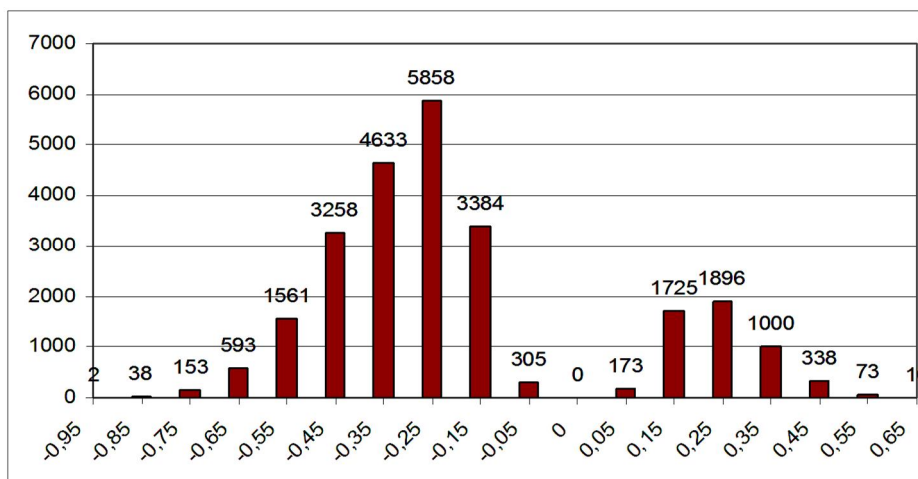(Vol.8) 2013, September

Fig.5. The histogram of distribution of the greatest vertical distribution $F_\Sigma(\xi)$ and $F_v^*(\xi)$

Table 1.

Data on a parity of positive and negative values Δ_m

Number of random variables in sample	2	4	7	11	16	22	29	150
Relative number of negative values Δ_m	0,87	0,79	0,73	0,68	0,65	0,63	0,61	0,55
Parity of negative and positive values Δ_m	6,7	3,8	2,7	2,1	1,9	1,7	1,6	1,2

As follows from table 1, with increase n_v a parity of negative and positive values Δ_m decreases, but at $n_v=150$ it still is not equal to unit. In table 2 experimental and settlement values quantile distributions $F_v^*(\Delta_m)$ for of some n_v and probabilities $R_v^*(\Delta_m) = \left[1 - F_v^*(\Delta_m)\right] = \alpha$ are resulted. We shall remind that experimental values received by imitating modeling on the computer [5], and settlement value-under the formula:

$$\Delta_{m,(1-0.5\alpha)}^{settl} = -\left[\Delta_{m,0.5\alpha}^{\exp er} + n_v^{-1}\right]$$ (9)

Table 2

Experimental and settlement values quantiles distributions $F_v^*(X)$ for of some n_v and probabilities $R_v^*(\Delta_m) = \left[1 - F_v^*(\Delta_m)\right]$.

$R_v^*(\Delta_m)$	Δ_m	Number of random variables in sample (Π_1)						
		2	4	6	11	40	90	150
0.025	experiment	0,343	0,377	0,358	0,302	0,185	0,131	0,104
	settlement	0,342	0,373	0,356	0,298	0,183	0,130	0,103
0.05	experiment	0,285	0,319	0,303	0,260	0,164	0,116	0,092
	settlement	0,285	0,317	0,302	0,262	0,162	0,116	0,092
0.1	experiment	0,184	0,240	0,244	0,216	0,140	0,100	0,079
	settlement	0,184	0,244	0,244	0,218	0,139	0,100	0,079
0.2	experiment	0,060	0,160	0,171	0,160	0,112	0,091	0,065
	settlement	0,061	0,165	0,171	0,164	0,116	0,081	0,064
0.3	experiment	-0,239	-0,173	-0,127	-0,097	0,089	0,067	0,053
	settlement	-0,027	0,105	0,125	0,128	0,094	0,068	0,055

Given tables 2 show, that the formula (9) precisely enough displays interrelation of boundary values of an interval of change of statistics Δ_m provided that 0,25. We shall enter into consideration three statistics based on random variables of an absolute vertical divergence of distributions $F_\Sigma(X)$ and $F_v^*(X)$:

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

− The greatest value of an absolute divergence. The algorithm of calculation looks like:
  - calculated $\qquad \Delta_i = (\xi_i - i/n_v); \quad i=1,n_v$
  - defined $\qquad B_v = \max \{\Delta_1, \Delta_2.., \Delta_i, \dots \Delta_{nv}\}$ $\qquad\qquad$ (10)

− Average value of an absolute divergence $M^*_{v,j}(\Delta)$ with j=1,N, where N – number of iterations. It is calculated under the formula (3)

− Average quadratic value of an absolute divergence $\sigma^*_{v,j}(\Delta)$. It is calculated under the formula (5)

The analysis of statistics B_v

$\qquad$ Distribution $F^*(B_v)$ has essential advantage in comparison $F^*(\Delta_m)$. It characterizes distribution of size of the greatest deviation of functions of distribution $F_\Sigma(X)$ and $F^*_v(X)$ without taking into account a sign on a deviation, i.e. it is considered equivalent both positive, and negative value of a deviation Δ_m. In table 3 are resulted quantile distributions B_v for of some values n_v and probabilities $F^*(B_v)$. If with them to compare to data table.2 it is easy to notice essential distinction of their critical values. So at $F^*(D_n)=F^*(B_v)=F^*(\Delta_m)=0,05$ and $n_v=4$ corresponding quantile will be equal $D_n=0.624$, $B_v=0.570$ and $\Delta_m=0,319$. Thus, the essence of a mistake at practical applications of criterion of Kolmogorov consists more often that statistics B_v is compared not to critical value of distribution $R^*(B_v)$, and with critical value of statistics of Kolmogorov D_n. If to sum up the aforesaid it is necessary to note, <u>each of entered in consideration statistics, for example, D_n, B_v or Δ_m, at check of the assumption it should be compared to the critical values, calculated on distributions, accordingly, $F^*(D_n)$, $F^*(B_v)$ and $F^*(\Delta_m)$</u>. In the illustrative purposes according to table 3 on fig.6 statistical distributions $R^*(B_v)=1-F^*(B_v)$ for $n_v=4$, 22 and 150 are resulted. As one would expect with increase n_v critical values B_v decrease. Character of distribution $R^*(B_v)$ changes also
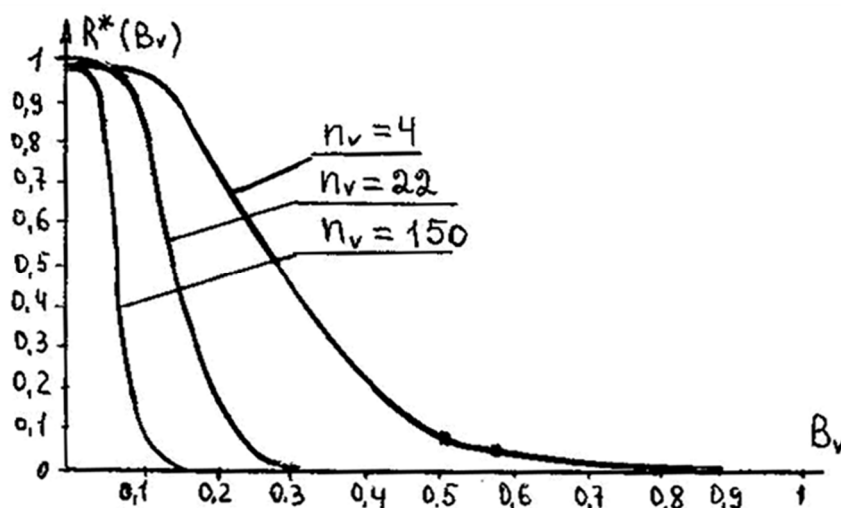

Fig.6 Statistical distributions $R^*(B_v)=1-F^*(B_v)$ for of some n_v

$\qquad$ In table 4 the factors of the equation $B_v = A \cdot n_v^{-b}$ calculated according to table 3 and factor determination R^2 are resulted.

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

Table 4

Factors of the equation of regress

F*(B_v)	a	b	R^2
0.9	1.079	0.459	0.9998
0.9	0.942	0.453	0.9997
0.8	0.774	0.439	0.9986
0.7	0.668	0.430	0.9982
0.6	0.590	0.422	0.9985
0.5	0.518	0.412	0.9975
0.4	0.447	0.396	0.9956
0.3	0.384	0.382	0.9922
0.2	0.317	0.360	0.9862
0.1	0.236	0.321	0.9829

As an example on fig.7 laws of change of a curve $B_v = A \cdot n_v^{-b}$ for F*(B_v)=0,95 (for a significance value are shown α=0,05 and 0,50)
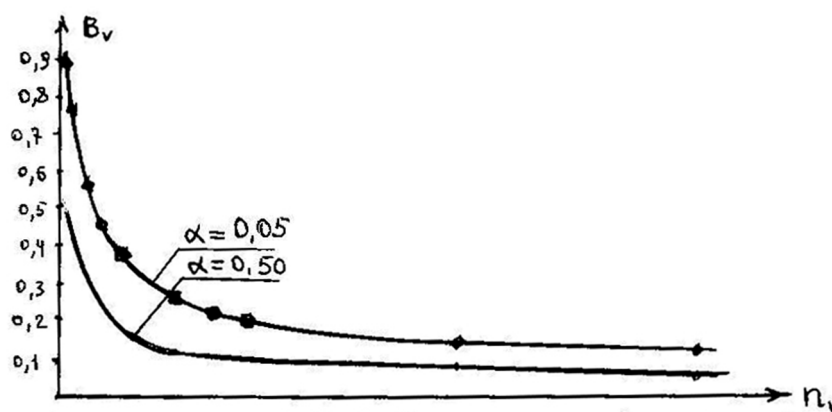

Fig.7 Laws of change of absolute size of the greatest vertical deviation
from number of sample units n_v at α=0,05 and 0,05

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

Table 3

Quantile distributions of statistics for of some values n_v and probabilities $F^*(B_v)$.

N	$F^*(B_v)$	Number of sample units (n_v)								
		2	4	7	11	22	29	40	90	150
1	0,05	0,112	0,127	0,116	0,104	0,083	0,075	0,067	0,048	0,038
2	0,1	0,157	0,154	0,136	0,120	0,094	0,084	0,075	0,053	0,042
3	0,15	0,193	0,175	0,151	0,131	0,103	0,092	0,081	0,057	0,045
4	0,2	0,223	0,191	0,164	0,142	0,110	0,098	0,087	0,061	0,048
5	0,25	0,249	0,208	0,177	0,152	0,117	0,104	0,092	0,064	0,051
6	0,3	0,274	0,222	0,189	0,160	0,124	0,110	0,097	0,067	0,053
7	0,35	0,300	0,236	0,201	0,170	0,130	0,115	0,101	0,071	0,056
8	0,4	0,324	0,250	0,213	0,179	0,136	0,121	0,106	0,074	0,058
9	0,45	0,348	0,268	0,225	0,189	0,143	0,127	0,111	0,077	0,061
10	0,5	0,376	0,286	0,236	0,198	0,150	0,132	0,116	0,080	0,063
11	0,55	0,401	0,306	0,249	0,209	0,157	0,139	0,121	0,083	0,066
12	0,6	0,426	0,326	0,262	0,219	0,164	0,145	0,127	0,087	0,069
13	0,65	0,449	0,348	0,277	0,231	0,172	0,152	0,133	0,091	0,072
14	0,7	0,473	0,370	0,294	0,244	0,181	0,160	0,139	0,095	0,075
15	0,75	0,499	0,393	0,313	0,258	0,191	0,169	0,147	0,100	0,079
16	0,8	0,548	0,421	0,334	0,276	0,203	0,179	0,155	0,106	0,083
17	0,85	0,620	0,454	0,358	0,295	0,217	0,191	0,166	0,112	0,088
18	0,9	0,683	0,497	0,391	0,322	0,235	0,206	0,179	0,122	0,096
19	0,95	0,778	0,568	0,442	0,362	0,263	0,232	0,201	0,136	0,107
20	0,99	0,902	0,689	0,538	0,440	0,320	0,283	0,240	0,164	0,129

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

As is known, average arithmetic value of random variables is the basic numerical characteristic of their center of grouping. Distinguish also a geometrical average, a harmonious average, a fashion and a median. In spite of the fact that all these numerical characteristics is united with concept of the center of grouping of random variables, each of them, so to say, «has the center» and only it and characterizes. Hence, «the center of grouping of random variables» considered as an attribute, and its versions will be the numerical characteristics noted above.

Each of VA will characterize features of distinction of distributions $F_\Sigma(X)$ and $F_v^*(X)$ peculiar only to it. Below formulas for an estimation of these numerical characteristics are resulted.

Calculation of average arithmetic value $n_v^*(\Delta)$ spent under the formula (2), calculation of average geometrical value - under the formula:

$$\sigma_v^*(\Delta) = \sqrt[n_v]{\prod_{i=1}^{n_v} |\Delta_i|} \,, \tag{11}$$

and calculation of an average harmonious – under the formula:

$$H_v^*(\Delta) = \left[\frac{1}{n_v} \sum_{i=1}^{n_v} \frac{1}{|\Delta_i|} \right]^{-1} \tag{12}$$

The estimation of a fashion spent under the histogram as average value of an interval, probability of hit in which random variables of sample the greatest.

The estimation of a median is spent by a finding $0,5n_v$ values ranging random variables of sample, if n_v even, and

$$\Delta_{med} = 0.5 \left[\left| \Delta_{\frac{n_v+1}{2}} \right| + \left| \Delta_{\frac{n_v-1}{2}} \right| \right] \tag{13}$$

Values – if n_v uneven.

According to algorithm of classification of data with a view of decrease in duration of calculations, the expediency of classification of data supervised for the sample having the greatest value of statistics. Thus considered, that if for this sample divergence $F_\Sigma(X)$ and $F_v^*(X)$ with the minimal risk of the erroneous decision it can be accepted casual the divergence of all others sample on set VA casual also.

In table 5 some results of calculation quantile distributions of statistics $M_v^*(\Delta)$ for of some n_v and probabilities $F^*[M_v^*(\Delta)]$ are resulted at N=25000

In the illustrative purposes according to table 5 on fig.8 distributions $R^*[M_v^*(\Delta)] = \{1 - F^*[M_v^*(\Delta)]\}$ for n_v=4 are resulted; 10 and 50, and on fig.9 a curve of dependence of size of statistics $M_v^*(\Delta)$ from n_v for $R^*[M_v^*(\Delta)] = \alpha = 0.05$ и 0.5.
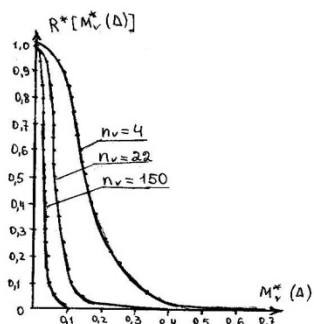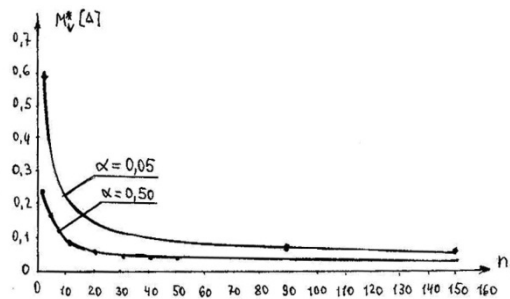

Fig.8. Laws of change
$R^*[M_v^*(\Delta)]$

Fig.9. Dependence of critical values
$M_v^*(\Delta)$ from number of sample units $_{nv}$

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

Table 5

Quantile distributions of statistics $M_v^*(\Delta)$ for of some n_v and probabilities $F^*[M_v^*(\Delta)]$

N	$F^*[M_v^*(\Delta)]$	Number of sample units (n_v)								
		2	4	7	11	22	29	40	90	150
1	0,05	0,077	0,0714	0,056	0,045	0,032	0,028	0,024	0,016	0,013
2	0,1	0,111	0,085	0,065	0,052	0,037	0,032	0,027	0,018	0,014
3	0,15	0,136	0,097	0,072	0,057	0,040	0,035	0,030	0,020	0,015
4	0,2	0,158	0,106	0,078	0,062	0,044	0,038	0,032	0,021	0,016
5	0,25	0,177	0,115	0,084	0,067	0,046	0,040	0,034	0,023	0,017
6	0,3	0,194	0,124	0,090	0,071	0,049	0,043	0,036	0,024	0,019
7	0,35	0,209	0,133	0,096	0,075	0,052	0,045	0,038	0,025	0,020
8	0,4	0,224	0,142	0,102	0,080	0,055	0,048	0,041	0027	0,021
9	0,45	0,238	0,152	0,109	0,085	0,058	0,050	0,043	0,028	0,022
10	0,5	0,250	0,163	0,116	0,090	0,062	0,053	0,045	0,030	0,023
11	0,55	0,276	0,173	0,123	0,095	0,065	0,056	0,048	0,032	0,024
12	0,6	0,303	0,185	0,131	0,101	0,069	0,059	0,051	0,033	0,026
13	0,65	0,331	0,198	0,140	0,107	0,073	0,063	0,054	0,035	0,027
14	0,7	0,361	0,213	0,150	0,115	0,078	0,068	0,058	0,038	0,029
15	0,75	0,395	0,232	0,162	0,123	0,084	0,073	0,062	0,040	0,031
16	0,8	0,433	0,254	0,176	0,134	0,091	0,078	0,067	0,043	0,034
17	0,85	0,475	0,281	0,194	0,147	0,099	0,085	0,073	0,048	0,037
18	0,9	0,526	0,314	0,217	0,164	0,112	0,095	0,081	0,053	0041
19	0,95	0,591	0,363	0,254	0,192	0,130	0,111	0,095	0,062	0,048
20	0,99	0,68	0,448	0,318	0,250	0,168	0,145	0,121	0,080	0,062

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

As follows from fig.8, s.f.d. the sums of random variables with uniform distribution in an interval [0,1] even for $n_v=150$ it is dissymmetric. And consequently, critical values quantile these distributions cannot be calculated according to average arithmetic value $M_{v;0.5}^*(\Delta)$ and average quadratic value $\sigma^*[M_v^*(\Delta)]$. The analysis shows, that the equation of interrelation $M_v^*(\Delta)$ and n_v for the fixed value $R^*[M_v^*(\Delta)] = \alpha$ (fig.9) can be presented by sedate function $M_v^*(\Delta) = An_v^{-b}$ with factor of determination $R^2 > 0.99$

In table 6 the constant factors of this equation calculated under the standard program of sedate transformation for of some values $R^*[M_v^*(\Delta)] = \alpha$ are resulted.

Table 6

Estimations of constant factors of the equations of regress and factor of determination

N	$R^*[M_v^*(\Delta)] = \alpha$.	Factors of regress		R2
		AND	IN	
1	0.05	0.81	0.58	0.9961
2	0.1	0.71	0.58	0.9949
3	0.2	0.57	0.58	0.9941
4	0.3	0.48	0.57	0.995
5	0.4	0.41	0.56	0.996
6	0.5	0.35	0.55	0.998
7	0.6	0.30	0.54	0.998
8	0.7	0.26	0.53	0.998
9	0.8	0.22	0.52	0.999
10	0.9	0.16	0.49	0.999

So that to find critical value of statistics $M_v^*(\Delta)$, for example, $\alpha=0,5$ and $n_v=5$ it is necessary to calculate $M_{v,0.5}^*(\Delta) = 0.81/5^{0.58}$ only and if we shall compare with empirical value of statistics $M_{v;Э}^*(\Delta)$ with $M_{v;0.5}^*(\Delta)$ at $M_{v;Э}^*(\Delta) < M_{v;0.5}^*(\Delta)$ it is possible to approve, that sample with a high probability is homogeneous with final population of multivariate data. In other words, classification of data on set VA is inexpedient.

Analysis of statistics $\sigma_v^*(\Delta)$.

At the analysis of statistical data of operation, EPS a degree of dispersion n_v realizations of a random variable Δ concerning the center of grouping $M_v^*(\Delta)$ it is characterized by an average statistical deviation $\sigma_v^*(\Delta)$ more often. The factor of a variation, and size of scope of dispersion calculated under formula $L^*(\Delta)=(\Delta_{max}-\Delta_{min})$ is less often used it is not applied. Practice of classification of multivariate data shows, that sample of random variables X on significant VA is concentrated to some interval $[X_j; X_{j+n_v}]$, which according to recommended algorithm is located in the top part of an interval of change final population of multivariate data since, $M_\Sigma^*(\Delta) < M_v^*(\Delta)$ and it is essential less it.

In table 7 some results of calculation quantile distributions of statistics $\sigma_v^*(\Delta)$ for of some n_v and probabilities $F^*[\sigma_v^*(\Delta)]$ with step 0,05 for number of iterations N=25000 are resulted. In the illustrative purposes on fig.10 are resulted in the form of continuous curves s.f.d. $F^*[\sigma_v^*(\Delta)]$ for of some n_v. On fig.11 the curve changes $\sigma_v^*(\Delta) = \varphi(n_v)$ received under table 7 and the standard program of sedate transformation of statistical data are resulted
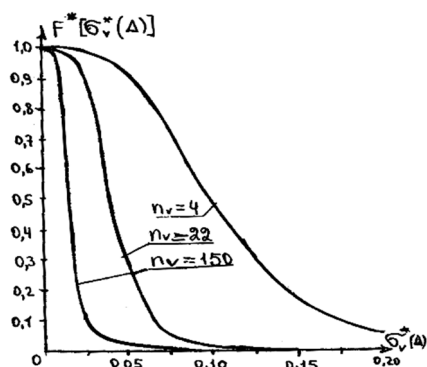
Fig.10 Character of change s.f.d.
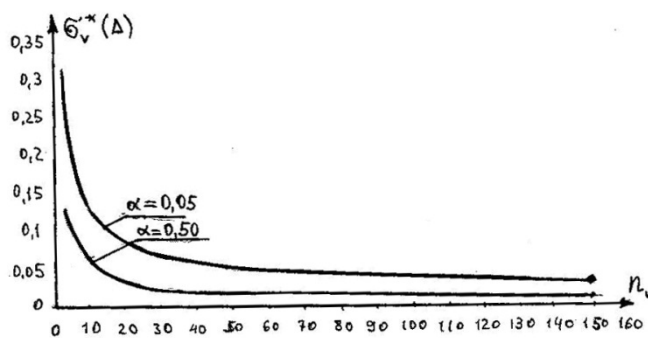$F^*[\sigma_v^*(\Delta)]$ depending on n_v

Fig.11. Laws of change critical
values of statistics $\sigma_v^*(\Delta)$ depending on n_v

In table 8 factors of the equation the regresses calculated under the standard program of sedate transformation for of some values $R^*[\sigma_v^*(\Delta)]$ are resulted

Table 8

Estimations of constant factors of the equations of regress and factor of determination

N	$R^*[\sigma_v^*(\Delta)] = \alpha$.	Factors of regress		R^2
		AND	IN	
1	0.05	0,428	0,54	0,997
2	0.1	0,385	0,55	0,996
3	0.2	0,322	0,545	0,996
4	0.3	0,276	0,54	0,998
5	0.4	0,237	0,53	0,999
6	0.5	0,197	0,50	0,999
7	0.6	0,163	0,48	0,995
8	0.7	0,129	0,46	0,991
9	0.8	0,094	0,43	0,996
10	0.9	0,061	0,40	0,992

VI. SOME RESULTS COMPARISON OF CRITERIA

Results of the analysis of laws of change s.f.d. статистик B_v, $M_v^*(\Delta)$ also $\sigma_v^*(\Delta)$ have allowed estimating probabilities $F^*[B_{v,\ni}]$, $F^*[M_v^*(\Delta_\ni)]$ and $F^*[\sigma_v^*(\Delta_\ni)]$, where the index «e» designates "experimental" value of probability of display of each of statistics. And as these of statistics characterize those or other properties of casual values of a vertical divergence of distributions $F_\Sigma(X)$ and $F_v^*(X)$, the probability of display of statistics will characterize, as a matter of fact, the importance of this property. In other words, an attribute of divergence $F_\Sigma(X)$ and $F_v^*(X)$ is the vertical distance between these distributions, and versions of an attribute – statistics.

By comparison, of this statistic the question on that, first, is of interest, probabilities of display of each of statistics, calculated on the same sample of general population how much essentially differs. Some results of calculations are resulted in table 9.

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

Table 7

Quantile distributions of statistics $\sigma_v^*(\Delta)$ for of some n_v and probabilities $F^*[\sigma_v^*(\Delta)]$ on N=25000

N	$F^*[\sigma_v^*(\Delta)]$	Number of sample units (n_v)								
		2	4	7	11	22	29	40	90	150
1	0,05	0,012	0,038	0,037	0,031	0,023	0,020	0,017	0,012	0,009
2	0,1	0,025	0,050	0,044	0,036	0,026	0,023	0,019	0,013	0,010
3	0,15	0,037	0,059	0,048	0,039	0,028	0,025	0,021	0,014	0,011
4	0,2	0,049	0,066	0,053	0,042	0,030	0,026	0,023	0,015	0,012
5	0,25	0,062	0,073	0,056	0,045	0,032	0,028	0,024	0,016	0,012
6	0,3	0,075	0,079	0,060	0,048	0,034	0,030	0,025	0,017	0,013
7	0,35	0,089	0,085	0,064	0,051	0,036	0,031	0,027	0,018	0,014
8	0,4	0,103	0,091	0,067	0,053	0,038	0,033	0,028	0,019	0,014
9	0,45	0,118	0,096	0,071	0,056	0,040	0,034	0,029	0,019	0,015
10	0,5	0,130	0,102	0,075	0,059	0,041	0,036	0,031	0,020	0,016
11	0,55	0,150	0,107	0,079	0,062	0,044	0,038	0,032	0,021	0,017
12	0,6	0,168	0,114	0,084	0,065	0,046	0,040	0,034	0,022	0,017
13	0,65	0,186	0,120	0,088	0,069	0,048	0,042	0,036	0,024	0018
14	0,7	0,204	0,128	0,094	0,073	0,051	0,044	0,038	0,025	0,019
15	0,75	0,224	0,137	0,100	0078	0,054	0,047	0,040	0,026	0,021
16	0,8	0,244	0,147	0,107	0,083	0,058	0,050	0,043	0,028	0,022
17	0,85	0,267	0,159	0,116	0,090	0,062	0,054	0,046	0,030	0,024
18	0,9	0,292	0,175	0,127	0,098	0,068	0,059	0,051	0,033	0,026
19	0,95	0,320	0,199	0,145	0,112	0,078	0,068	0,058	0,038	0,030
20	0,99	0,347	0,244	0,178	0,140	0,097	0,084	0,072	0,048	0,037

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

Table 9

Parities $R^*[M_v^*(\Delta)]$ and $R^*[\sigma_v^*(\Delta)]$, calculated for the same samples with $n_v=4$

N	ξ_1	ξ_2	ξ_3	ξ_4	$R^*[M_v^*(\Delta)]$	$R^*[\sigma_v^*(\Delta)]$
1	0,399	0,363	0,688	0,524	0,32	0,64
2	0,945	0,781	0,225	0,848	0,81	0,43
3	0,429	0,488	0,724	0,682	0,68	0,36
4	0,921	0,812	0,913	0,432	0,29	0,69
5	0,778	0,459	0,402	0,1	0,25	0,77

The examples resulted in table 10 testify that probabilities $R^*[M_v^*(\Delta)]$ and $R^*[\sigma_v^*(\Delta)]$ can essentially differ. The reasons of such distinction are known. Average value of realizations of a vertical deviation s.f.d. $F_\Sigma(X)$ and $F_v^*(X)$ there can be small enough, and their average quadratic deviation – greater and on the contrary. In other words, examples tables 10 testify that <u>comparison of efficiency of criteria of check of hypotheses differing статистик not always is justified</u>. First, because the result of comparison depends on distribution $F_v^*(X)$, i.e. the result of comparison not is a rule, and secondly because statistics of criteria can have various physical sense, for example $M_v^*(\Delta)$ and $\sigma_v^*(\Delta)$.
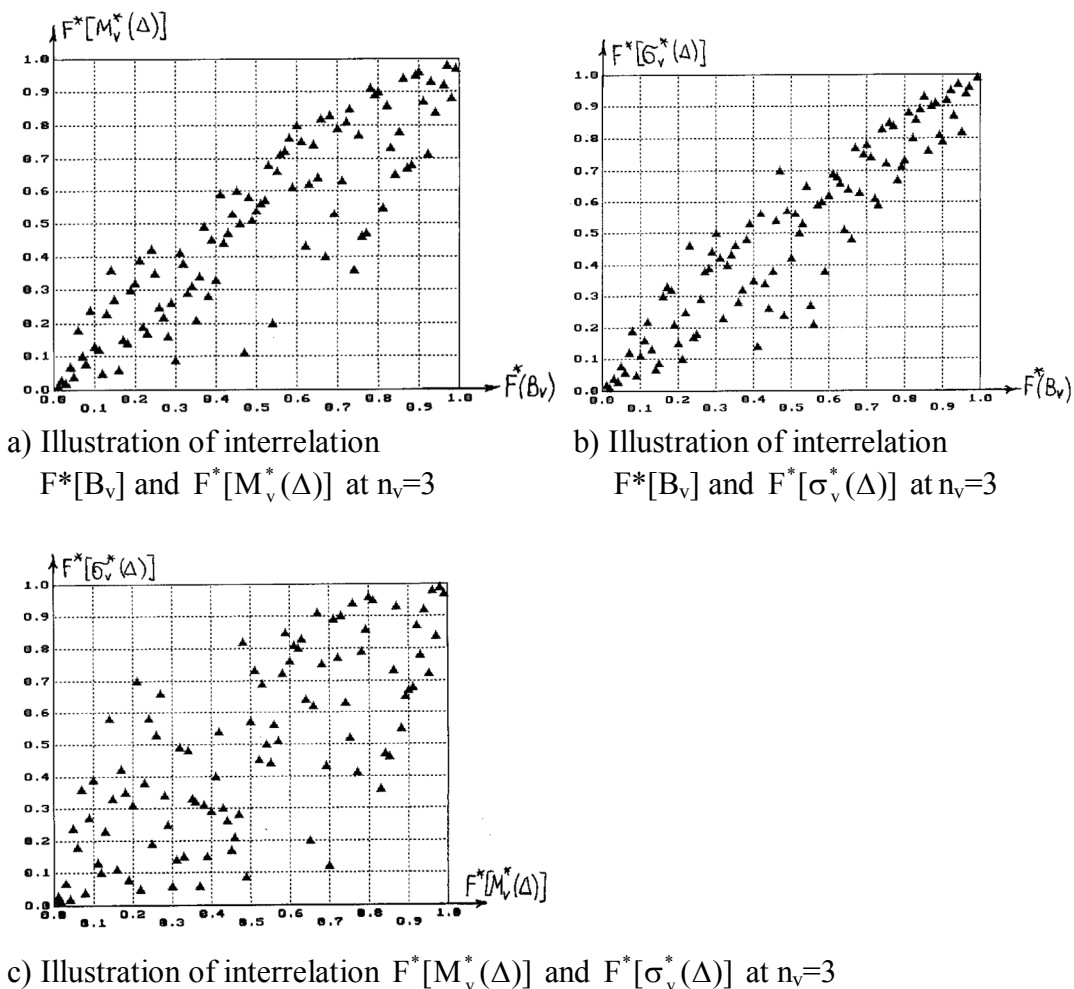

a) Illustration of interrelation
$F^*[B_v]$ and $F^*[M_v^*(\Delta)]$ at $n_v=3$

b) Illustration of interrelation
$F^*[B_v]$ and $F^*[\sigma_v^*(\Delta)]$ at $n_v=3$

c) Illustration of interrelation $F^*[M_v^*(\Delta)]$ and $F^*[\sigma_v^*(\Delta)]$ at $n_v=3$

Fig.12. Correlation field of interrelation of probabilities of display of realizations statistic

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

Moreover, as they can be independent as well as $M_v^*(\Delta)$ and $\sigma_v^*(\Delta)$, casual character of divergence $F_\Sigma(X)$ from $F_v^*(X)$ by criterion with statistics $M_v^*(\Delta)$ yet does not mean, that divergence $F_\Sigma(X)$ from $F_v^*(X)$ by criterion with statistics $\sigma_v^*(\Delta)$ it will appear also casual. In the illustrative purposes on рис.10 the correlation field of interrelation of probabilities of display of realizations statistics к B_v, $M_v^*(\Delta)$ and $\sigma_v^*(\Delta)$, calculated for the same samples from $n_v=3$ random variables is resulted.

Calculations spent in following sequence:

- For each sample from n_v random variables ξ, In regular intervals distributed in an interval [0,1], realizations B_v are calculated, $M_v^*(\Delta)$ and $\sigma_v^*(\Delta)$. Calculations are spent N time, where N-number of iterations. Results of calculations brought in the table and which form is shown on fig.13.

N	B_v,	$M_v^*(\Delta)$	$\sigma_v^*(\Delta)$	$F^*(B_v)$	$F^*[M_v^*(\Delta)]$	$F^*[\sigma_v^*(\Delta)]$

Fig.13. Table of initial data

- Ranking of realizations B_v of the table and by way of increase in numerical values B_v is spent. Together with $B_{v,i}$ move and corresponding $B_{v,i}$ values $M_{v,i}^*(\Delta)$ and $\sigma_{v,i}^*(\Delta)$;

- Pays off $F_i^*[B_v] = \dfrac{1}{N}$ with i=1,N and are brought in a column $F^*(B_v)$ of the table A

- In the table B, the similar table A, ranking of realizations $M_v^*(\Delta)$ is spent and further under the formula $F_i^*[M_v^*(\Delta)] = \dfrac{1}{N}$ corresponding $M_{v,i}^*(\Delta)$ probabilities $F_i^*[M_v^*(\Delta)]$ are calculated

- For each value of statistics $M_v^*(\Delta)$ from the table A there is a value equal to it in the table B and corresponding value of probability $F^*[M_v^*(\Delta)]$ which is brought in a column $F^*[M_v^*(\Delta)]$ of the table A

- In the table B ranking of realizations $\sigma_v^*(\Delta)$ is spent and further under the formula $F_i^*[\sigma_v^*(\Delta)] = \dfrac{1}{N}$ corresponding $\sigma_{v,i}^*(\Delta)$ probabilities $F_i^*[\sigma_v^*(\Delta)]$ are calculated

- For each value of statistics $\sigma_v^*(\Delta)$ from the table A there is a value equal to it in the table B and corresponding value of probability $F^*[\sigma_v^*(\Delta)]$ which is brought in a column $F^*[\sigma_v^*(\Delta)]$ of the table A

As follows from the resulted figures the essential interrelation between probabilities $F^*[B_v]$ and $F^*[M_v^*(\Delta)]$ or $F^*[\sigma_v^*(\Delta)]$ observed. This interrelation has evident physical interpretation: with growth B_v grow, on the average, $M_v^*(\Delta)$ and $\sigma_v^*(\Delta)$.

Figure 12 full enough characterizes weak interrelation between $M_v^*(\Delta)$ and $\sigma_v^*(\Delta)$. Therefore and the answer to a question on, whether is enough to check up character of divergence $F_\Sigma(X)$ from $F_v^*(X)$ only on one statistics B_n it appears ambiguous, and the priority is given expediency of attraction to the decision of all statistic.

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

VII. RECOGNITION EXPEDIENCY OF CLASSIFICATION
OF MULTIVARIATE DATA

Above-stated testifies to necessity of check of the assumption expediency of classification of multivariate data by the criteria reflecting the basic properties of random variables of a vertical divergence of distributions $F_{\Sigma}^{*}(X)$ and $F_{V}^{*}(Y)$. Conditions of check of possible assumptions recommended by authors look like:

If $S_e(\Delta) < S_{0.5}(\Delta)$,   that H_1
If $S_e(\Delta) > S_{0.05}(\Delta)$,   that H_2 (14)
If $S_{0,5}(\Delta) < S(\Delta) > S_{0,05}(\Delta)$,   that H_3

where $S(\Delta)$ – one of possible statistics a random variable Δ; $S_{0.05}(\Delta)$ and $S_{0.5}(\Delta)$ – critical values of statistics with a significance value, accordingly 0,05 and 0,5; H_1, H_2 and H_3 – assumptions, accordingly, about casual character of a divergence $F_{\Sigma}^{*}(X)$ and $F_{V}^{*}(Y)$ and inexpediency of classification of data; about not casual divergence $F_{\Sigma}^{*}(X)$ and $F_{V}^{*}(Y)$ and expediency of classification of data; expediency of an estimation and comparison of risk of the erroneous decision for H_1 and H_2.

Algorithm of an estimation of expediency of classification of multivariate data we shall consider on a following example. Let sample with $n_v=4$ is set: {0,151, 0,341, 0,259, 0,120}. Random numbers are received by program way, are called pseudo-casual, and have uniform distribution in an interval [0;1]. The basic assumption: random numbers of sample have uniform distribution in an interval [0;1]. For check of this assumption that is identical to the assumption of inexpediency of classification of final population of statistical data, we shall calculate realizations of values of a vertical divergence of distributions $F_{\Sigma}(X)$ from $F_{v}^{*}(X)$. It is easy to be convinced, that they are equal: $\Delta_1 = -0{,}349$; $\Delta_2 = -0{,}660$; $\Delta_3 = -0{,}498$; $\Delta_4 = -0{,}130$. Results of calculations $\Delta_{m,E}$; $B_{v,E}$; $M_{v,E}^{*}(\Delta)$ and $\sigma_{v,E}^{*}(\Delta)$ under formulas 8,10; 2 and 4 are resulted in table 10. Here probabilities $R^*(\Delta_m)$; $R^*(B_{v,E})$; $R^*[M_{v,E}^{*}(\Delta)]$ and $R^*[\sigma_{v,E}^{*}(\Delta)]$ are resulted

Table 10

Results of calculations at an estimation of expediency of classification of statistical data

N	Statistics	Estimations статистик	$R^*[S_v(\Delta)]$	The decision
1	$\Delta_{m,E}$	-0,660	0,02	H_2
2	$B_{v,E}$	0,660	0,02	H_2
3	$M_{v,E}^{*}(\Delta)$	0,407	0,03	H_2
4	$\sigma_{v,E}^{*}(\Delta)$	0,224	0,27	H_3

As follows from table 10, three from four criteria testify that the set sample is unpresentable, and the lead classification is expedient (H_2). And only size of an average quadratic deviation of realizations of vertical divergence $F_{\Sigma}(X)$ from $F_{v}^{*}(X)$ has probability 0,27 that value of a error II type testifies about necessities of the account of this property and attraction to check of hypothesis H_1.

Thus, one criterion testifies to expediency of classification of data, others – on the contrary, testify to uniformity of compared data. Thus, it is necessary to answer a question: whether there are properties of compared random variables on which they cannot be acceptance homogeneous? In our example is. This question and character of the answer are natural.

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliev A.Z. – PRINCIPLES OF CLASSIFICATION RELIABILITY STATISTICAL DATA OF THE ELECTRIC
EQUIPMENT OF POWER SUPPLY SYSTEMS

RT&A # 03 (30)
(Vol.8) 2013, September

CONCLUSION

1. The expediency of classification of final population of multivariate data, in other words, presence significant VA established based on the theory of check of statistical hypotheses.
2. Criterion of check of a hypothesis is the condition non-ascendance empirical value of statistics of its critical value. As statistics, Kolmogorov's D_n statistics is most often used. However random variables of vertical divergence $F_\Sigma(X)$ from $F_v^*(X)$ are characterized also:

   - the greatest on absolute size and constant on a sign a vertical divergence Δ_m;
   - the greatest on absolute value of divergences B_v;
   - average arithmetic value of an absolute divergence $M_v^*(|\Delta|)$
   - average quadratic value of an absolute divergence $\sigma_v^*(|\Delta|)$;

   This list could be continued. But the main thing here is that each of considered above statistics characterizes, distinct from other properties of random variables Δ property and has the distribution $R^*[S_i(\Delta)]$. The importance of properties is defined by a parity $B_v = A \cdot n_v^{-b}$, which calculated on concrete sample. The $M_n^*(\Delta)$ is less the importance VA above and on the contrary.

3. To compare with these curves on capacity, certainly, it is possible. As a result, of comparison we shall be convinced that the greatest capacity has criterion which statistics characterizes properties of random variables the samples having among all other properties the greatest importance. For following sample with considerable probability, the importance of properties of random variables can essentially change.

REFERENCE

1. Dyakov A.F., Isamuhamedov Y.Sh. Modern a condition of electric power industry of Russia and factors of decrease in reliability of electro supply. Methodical questions of research of reliability of greater systems of power. Problems of reliability of systems of power in market conditions. Baku, 2013, 7-15 p.
2. Power transformers. The help book. S.D.Lizunova, A.K.Lohanina. M: Energyatomizdat, 2004 г.-616 p.
3. Farhadzadeh E.M., Muradaliyev A.Z., Farzaliyev Y.Z. Increase of accuracy of an estimation and reliability of comparison of parameters of individual reliability of power units of a state district power station. M.: the Electricity №9, 2008, c.10-17
4. Ryabinin I.A. Bas of the theory and calculation of reliability ship electro power system. "Shipbuilding", 1971, 453 p.
5. Farhadzadeh E.M., Muradaliyev A.Z., Farzaliyev Y.Z. Decrease in risk erroneous classification the multivariate statistical data describing the technical condition of the equipment of power supply system. Journal: «Reliability: Theory&applications», R&RATA (Vol.8 No.2 (29)), 2013, June, USA, p.55-64