ISSN 2224-5227

**2015 • 2**

**ҚАЗАҚСТАН РЕСПУБЛИКАСЫ**
**ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫНЫҢ**

# БАЯНДАМАЛАРЫ

# ДОКЛАДЫ

**НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК**
**РЕСПУБЛИКИ КАЗАХСТАН**

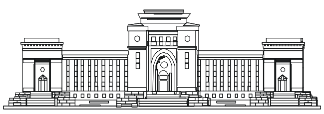
# REPORTS

**OF THE NATIONAL ACADEMY OF SCIENCES**
**OF THE REPUBLIC OF KAZAKHSTAN**

ЖУРНАЛ 1944 ЖЫЛДАН ШЫҒА БАСТАҒАН

ЖУРНАЛ ИЗДАЕТСЯ С 1944 г.

PUBLISHED SINCE 1944



ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫНЫҢ
**БАЯНДАМАЛАРЫ** **2015 • 2**

**«Доклады Национальной академии наук Республики Казахстан» ISSN 2224-5227**
Собственник: Республиканское общественное объединение «Национальная академия наук Республики Казахстан» (г. Алматы)
Свидетельство о постановке на учет периодического печатного издания в Комитете информации и архивов Министерства культуры и информации Республики Казахстан **№5540-Ж**, выданное 01.06.2006 г.
Периодичность: 6 раз в год. Тираж: 3000 экземпляров
Адрес редакции: 050010, г.Алматы, ул.Шевченко, 28, ком.218-220, тел. 272-13-19, 272-13-18
http://nauka-nanrk.kz. reports-science.kz
Адрес типографии: ИП «Аруна», г.Алматы, ул.Муратбаева, 75

**REPORTS** **2015 • 2**

OF NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN



**Reports of the National Academy of Sciences of the Republic of Kazakhstan.**
**ISSN 2224-5227**
Owner: RPA "National Academy of Sciences of the Republic of Kazakhstan" (Almaty)
The certificate of registration of a periodic printed publication in the Committee of Information and Archives of the Ministry of Culture and Information of the Republic of Kazakhstan N 5540-Ж, issued 01.06.2006
Periodicity: 6 times a year
Circulation: 3000 copies
Editorial address: 28, Shevchenko str., of.219-220, Almaty, 050010, tel. 272-13-19, 272-13-18,
http://nauka-nanrk.kz / reports-science.kz
Address of printing house: ST "Aruna", 75, Muratbayev str, Almaty

**REPORTS OF THE NATIONAL ACADEMY OF SCIENCES
OF THE REPUBLIC OF KAZAKHSTAN**
ISSN 2224-5227
Volume 2, Number 300 (2015), 12– 18

# About adaptation criteria of interaction of foundations with natural environment

**Gumenyuk V.V.**

Kazakh Leading Academy of Architecture and Civil Engineering,
International Educational Corporation
v.gumenyuk@kazgasa.kz

**Key words:** foundation, an external environment, adaptation criteria, crossed over parameters, conformity fields, load, and deformation.

**Abstract.** The article discusses definition of criteria for adaptation of two interacting systems, the foundation and external environment. It has been shown that the main criteria of adaptation are reliability assessment and durability of this system. It allows to measure degree of conformity of two interacting systems, an area of favorable values of crossed over parameters. The system which is considered interaction of foundation and environment on their crossing has random variables and functions. At the same time, the number of conformity fields equal to the number of parameters at the crossing. Adaptation criteria vary from zero to one. The substantiation of calculation formulas for calculating adaptation criteria is given. The adaptation criteria tend to a maximum of two main parameters of interaction - loads and deformations at any constructive solutions of foundation.

**УДК 624.04:519.87**

# О критериях адаптации при взаимодействии фундаментов с естественной средой

**В.В. Гуменюк**

Казахская головная архитектурно-строительная академия,
Международная образовательная корпорация, г. Алматы
v.gumenyuk@kazgasa.kz

**Ключевые слова:** фундамент, внешняя среда, критерий адаптации, пересекающиеся параметры, области соответствия, нагрузка, деформация.

**Аннотация.** В статье рассматривается определение критериев адаптации двух взаимодействующих систем, фундамента и внешней среды. Показано, что критерий адаптации является основным критерием оценки надежности и долговечности этой системы. И позволяет измерять степень соответствия двух взаимодействующих систем, область благоприятных значений пересекающихся параметров. Рассматриваемая система взаимодействия фундамента и внешней среды на своем пересечении имеют случайные величины и функции. При этом число областей соответствия равно числу параметров на пересечении. Критерий адаптации изменяется от нуля до единицы. Дается обоснование расчетных формул для вычисления критерия адаптации. При любых конструктивных решениях фундамента критерий адаптации стремится к максимуму по двум главным параметрам взаимодействия – нагрузкам и деформациям.

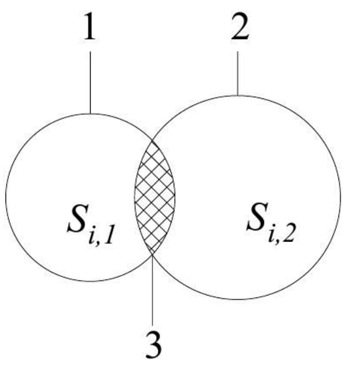
## 1. Введение

Для проектирования любых сооружений необходимым этапом при расчетах является определение значения критериев адаптации при взаимодействии сооружения, фундамента на котором оно будет функционировать с внешней средой:

- поверхностью земли;
- воздушной, водной и любой иной составляющей среды.

В этом весьма актуальном направление в науке по теории сооружений не обнаружены какие-либо основополагающие работы. В этой связи сделана первая попытка в разработке начала теоретических работах в этой области.

Рассмотрим взаимодействие двух систем (рисунок 1) только по одному параметру $S_i$ .




1 – область допустимых значений параметра $S_i$ для системы $S_1$; 2 – область допустимых значений параметра $S_i$ для системы $S_2$ ; 3 – область соответствия параметра $S_i$ для систем $S_1$ и $S_2$ .

Рисунок 1 – Область соответствия при пересечении двух систем по одному параметру

Если значение параметра $S_i$ находится в области 1, то система $S_2$ будет в состоянии полного или частичного бездействия, так как такое значение параметра $S_i$ не может быть реализовано системой $S_2$ .

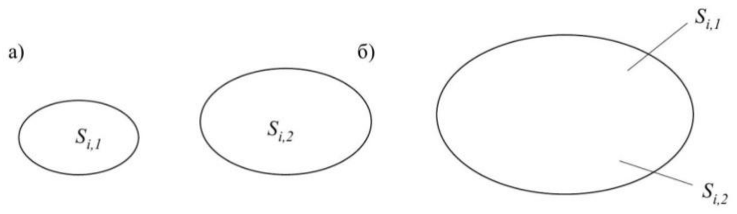
Если же значение $S_i$ находится в области 2, то в аналогичном $S_1$ состоянии находится система $S_1$. Когда значение параметра $S_i$ попадает в область 3, то системы $S_1$ и $S_2$ находятся в благоприятном и в рабочем состоянии. Из этой модели видно, что если $S_1$ и $S_2$ управляемые системы, то при попадании значения $S_i$ в область 1 система $S_2$ стремится вогнать его в область 3 или 2, если они конкурирующие. Если же значение $S_i$ находится в области 2, то уже система $S_1$ стремиться вогнать этот параметр в область 3 или 1. Если у двух систем общая система работы, то они ( $S_1$ и $S_2$ ) стремятся одновременно удерживать параметр $S_i$ в области 3.

*Таким образом, область благоприятных значений пересекающихся параметров двух систем называется областью соответствия этих параметров.*

Из этого определения следует, что для двух взаимодействующих систем этих областей соответствия ровно столько, сколько параметров на пересечении. В дальнейшем задача состоит в том, чтобы найти способ, позволяющий измерять степень соответствия взаимодействующих систем.

2. Критерий адаптации при взаимодействии фундаментов и оснований с естественной средой

До перехода к определению критерия, измеряющего степень соответствия пересекающихся систем, вновь обратимся к рисунку 1. Так как имеются три области, то существуют различные вероятности попадания значения параметра в эти области. Рассмотрим два следующих крайних случая, представленных на рисунке 2.




а) – параметры $S_{i,1}$ и $S_{i,2}$ не имеют общих значений;

б) – параметры $S_{i,1}$ и $S_{i,2}$ имеют одинаковые тождественные области задания.

Рисунок 2 – крайние случаи состояния пересекающегося параметра $S_i$

В первом случае область соответствия параметров $S_{i,1}$ и $S_{i,2}$ отсутствует, а во втором – тождественно совпадает с 1 и 2. Так как критерий, измеряющий степень соответствия систем $S_1$ и $S_2$, основывается на вычислениях вероятности попадания в область 3, то из рисунка 2 видно, что в этих двух крайних случаях значение критерия $J$ выразится: $J = 0$ для первого случая и $J = 1$ для второго.

*Следовательно, во всех других случаях критерий $J$ измеряется от 0 до 1, т.е.* $0 \le J \le 1$.

*Критерий, позволяющий измерять степень соответствия двух взаимодействующих систем, будем называть критерием адаптации.*

Это название наиболее точно отвечает физическому смыслу нового критерия, так как всякая искусственная или живая система с целью выживания всегда стремится увеличить область соответствия своих параметров окружающим системам или адаптироваться к окружающей среде.

Дальнейшая наша задача состоит в обосновании самых простейших расчетных формул для вычисления критерия адаптации.

Рассмотрим различные исходы пересечения $S_1$ и $S_2$. Пересечения систем $S_1$ и $S_2$ по детерминированным параметрам для наших исследований интереса не представляют, так как эти параметры выражают только вязи между системами.

Как следует из рисунка 1, вероятность попадания в область 1-3 существует только в том случае, когда на пересечении ряд или хотя бы один параметр будет представлен случайной величиной или случайной функцией. Вследствие того, что рассматриваемые нами управляемые искусственные и естественные системы вероятностные, на их пересечении всегда имеются случайные величины и функции.

Параметр $s_1$ для системы $S_1$ может иметь значение $0 \le s_1 \le \hat{s}_1$, где $\hat{s}_1$ - верхнее допустимое значение параметра $s_1$.

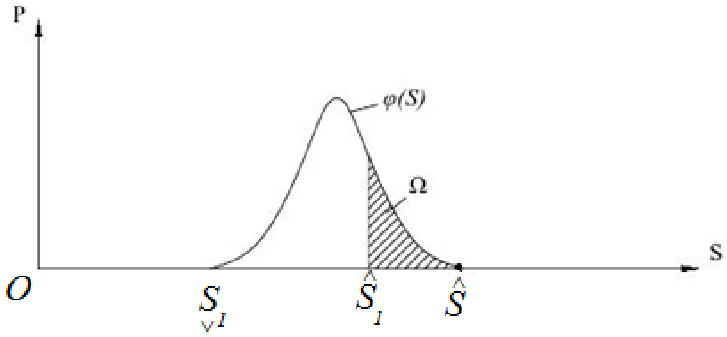
Из рисунка 3 ясно, что пока значение параметра $s$ в интервале $\underset{\vee}{s} - \hat{s}_1$, система $S_1$ удовлетворительно функционирует в среде $S_2$. Однако, как только среда принимает значение параметра $s$ в интервале $s \in \left\{\underset{\vee}{s} - \hat{s}_1\right\}$, система $S_1$ находится в неудовлетворительном или отказовом состоянии, либо «погибает». Вычислим вероятность $(P_П)$ попадания значения параметра $s$ в область $\Omega = \{\hat{s}_1 - \hat{s}\}$:

$$P_П = \int_{\Omega} \varphi(s_1) ds_1 \qquad (1)$$

По определению критерия адаптации его величина для этого случая, очевидно, выразится:

$$J = 1 - P_{П} \qquad (2)$$

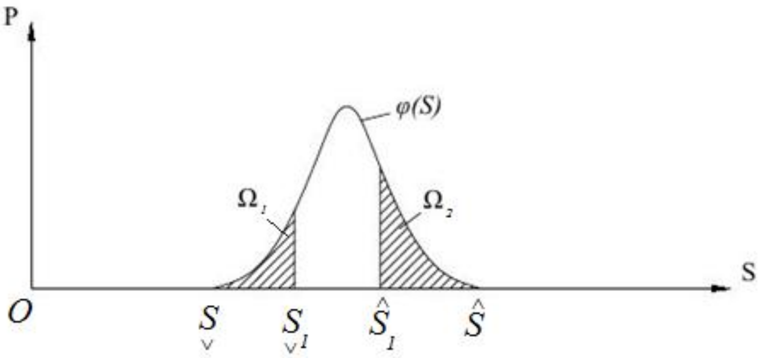
Из (2) следует, что если $\hat{s}_1 \geq \hat{s}$, то $P_{П} = 0$ и критерий адаптации равен 1, а если $\hat{s}_1 \leq \underset{\vee}{s}$, то $P_{П} = 1$ и критерий адаптации $J$ системы $S_1$ в среде $S_2$ равен 0. Во всех случаях, очевидно, $1 > J > 0$. Из (2) ясно, что в целях «выживания» системы $S_1$ в среде $S_2$ необходимо постоянно максимизировать величину критерия адаптации.




Ω- область недопустимых состояний параметра $s$ для системы $S_1$

Рисунок 3 – пересечение случайной величины с детерминированной границей

Рассмотрим систему $S_1$ в среде $S_2$ с одним пересекающимся параметром $s$, с двух сторон, ограниченным для $s_1$ в виде $\underset{\vee}{s_1} \leq s_1 \leq \hat{s}_1$ (рисунок 4).




$\Omega_1, \Omega_2$ – область недопустимых состояний параметра $s$ для системы $S_1$.

Рисунок 4 – Пересечение случайной величины с границами $\underset{\vee}{s_1}$ и $\hat{s}_1$

Для этого случая значение критерия адаптации вычисляется по формуле:

$$J = 1 - \left[ \int_{\Omega_1} \varphi(s)\, ds + \int_{\Omega_2} \varphi(s)\, ds \right] \qquad (3)$$

Более сложные случаи взаимодействия систем по пересекающимся параметрам в виде случайных величин и случайных функций рассмотрены в известных работах академика Е.И. Рогова[1-3].

Переходя теперь к взаимодействующим системам “фундамент” – “внешняя естественная среда” отметим следующие наиболее характерные случаи.

Расчет осадок поверхности неоднородного основания от действия вертикальной сосредоточенной силы $P$ осуществляется по формуле:

$$W(r)=\frac{(1-\nu^2)\cdot P}{\pi\cdot E_0}\cdot\left[\frac{1}{r}-\frac{1}{2}\int_0^\infty\frac{3e(R)/E_0-1}{e(R)/E_0}\cdot\frac{z^3}{R^5}dz\right], \qquad )$$

где функция степени неоднородности

$$e(R)=\frac{E_0}{3}\left[1+\alpha\cdot m\cdot D\cdot\int_0^1\frac{u^m}{1+Du}du\right], \qquad (5)$$

$$R=\sqrt{z^2+r^2}, \quad r=\sqrt{x^2+y^2}, \; D=(\lambda R)^n, \; m=3/n.$$

$\alpha=(E_\infty-E_0)/E_0$ - характеризует степень упругой неоднородности основания по глубине; $m=3/n$ – скорость изменения модуля деформации при увеличении глубины основания; $\lambda$ – эмпирический параметр, имеющий размерность, обратную длине; $0\le n\le\infty$. При $\lambda=0$ упругое основание становится однородным, $E(z)=E_0$; при $\lambda\to\infty$ модуль деформации $E(z)\approx E$. На глубине $z=1/\lambda$, $E(z)$ дает среднее значение $E_{cp/}=(E_0+E_\infty)/2$

При такой формулировке функции влияния первое слагаемое в (4) представляет собой классическое решение Буссинеска для однородного упругого полупространства, а второе (интегральное) слагаемое определяет влияние неоднородности деформационных свойств грунта по глубине.

Здесь сложная функция $W(r)$ представляется случайной на всем времени существования фундамента, и она пересекается с детерминированной границей достижимой критической величины осадок.

Вертикальное перемещение поверхности основания фундамента, заданное в виде также случайной функция в каждой точке $(x_i, y_i)$

$$\delta_{ij}=\iint_{F_j}\omega(x_i,y_i,\xi,\eta)d\xi d\eta, \qquad (6)$$

где $\omega(x,y,\xi,\eta)$ – функция влияния, определяемая по формуле (4).

Эта функция пересекается с детерминированной границей

$\hat{\delta}_{ij}$, т.е. $\delta_{ij}\le\delta^0_{ij}$ (7)

где $\delta^0_{ij}$ – допустимая величина перемещения основания фундамента.

Функция распределения реактивных давлений под фундаментом в виде

$$W=\frac{2(1-\nu^2)P}{\pi E_0\alpha}\delta(0,\alpha,A,m) \qquad (8)$$

также является случайной и условие работоспособности системы задается детерминированной границей

$W\le \underset{\vee}{W}$ (9)

где $\underset{\vee}{W}$ - предельно допустимая осадка.

Здесь необходимо решить уравнение

$$tg\,\psi = i_r = \frac{1-\nu^2}{2E_0} K \frac{P \cdot e \cdot \cos\varphi}{a^3} \qquad (\ )$$

относительно угла и получить случайную функцию, которая также пересекается с детерминированной границей $\varphi_0$ т.е. $\varphi \le \varphi_0$ ,где $\varphi_0$ допустимый из условия устойчивости угол.

3. Заключение. В статье приведены результаты по обоснованию критериев адаптации оснований и фундаментов любых конструкций, к условиям среды при их проектировании, сооружении и всего периода эксплуатации. Дальнейшая задача в решении этой проблемы достижения максимального значения комплексного критерия адаптации, состоит в разработке аналитических случайных функций в явном виде и вычислительных процедур.

**ЛИТЕРАТУРА**

[1] *Рогов Е.И., Рогов С.Е., Рогов А.Е.* Теория геотехнологий. Алматы: FORTRESS.- 2010. - 355с.
[2] *Рогов А.Е., Рыспанов Н.Б.* Математические основы геотехнологий. Алмыты: FORTRESS. - 2007. - 368с.
[3] *Рогов А.Е.* Имитационное математическое моделирование. Алматы: FORTRESS. - 2007. - 96с.

**REFERENCES**

[1] *Rogov E.I., Rogov S.E., Rogov A.E.* Theory of geotechnologies. Almaty: FORTRESS.- 2010. – 355 p. (in Russ.).
[2] *Rogov A.E., Ryspanov N.B.* Mathematical foundations of geotechnology. Almaty: FORTRESS. - 2007. – 368 p.
[3] *Rogov A.E.* Simulation mathematical modeling. Almaty: FORTRESS. - 2007. - 96 p. (in Russ.).

**ІРГЕТАСТЫҢ ТАБИҒИ ОРТАМЕН ӨЗАРА ӘРЕКЕТТЕСУІНІҢ БЕЙІМДЕЛУ ӨЛШЕМДЕРІ ТУРАЛЫ**

**Гуменюк В.В.**

Жалпы құрылыс факультеті,
Қазақ Бас сәулет-құрылыс академиясы, Халықаралық білім корпорациясы

**Кілт сөздер:** іргетас, сыртқы орта, бейімделудің өлшемі, киылысу параметрлері, сәйкестіктің облыстары, жүк, деформация.

**Андатпа.** Мақалада іргетастың сыртқы ортамен, екі әрекеттес жүйенің бейімделу өлшемдерінің анықтауы қарастырылады. Осы жүйенің негізгі сенімділік пен төзімділікті бағалау өлшемі бейімделу өлшемі болып табылады. Екі әрекеттес жүйенің сәйкестігінің дәрежесін және қиылысу параметрлерінің қолайлы мағынасын өлшеуге мүмкіндік береді.

Іргетастың сыртқы ортамен өзара әрекеттесу жүйесінің қиылысуында кездейсоқ аумақ және атқаратын қызметтері бар. Сонымен қатар, сәйкестік облыстар саны қиылыстағы параметрлерге сәйкес келеді. Бейімділік өлшемі нөлден бірге дейін өзгереді. Бейімделу өлшемін есептеу үшін есептік формуланың негіздеуі берілген. Іргетастың кез-келген құрылымдық шешімінде бейімделу өлшемі екі басты параметрлердің -жүктер және деформацияның жоғарғы шегіне дейін жетуге тырысады.

**Сведения об авторе**

Статьи «О критериях адаптации при взаимодействии фундаментов с естественной средой»

Гуменюк Валерия Владимировна – к.т.н., факультет общего строительства, Казахская головная архитектурно-строительная академия, Международная образовательная корпорация

Рабочий адрес: Республика Казахстан, 050043,

г.Алматы, ул. К. Рыскулбекова, 28

*Поступила 18.01.15 г.*

**REPORTS OF THE NATIONAL ACADEMY OF SCIENCES
OF THE REPUBLIC OF KAZAKHSTAN**
ISSN 2224-5227
Volume 2, Number 300 (2015),18 –25

UDK 004

# The effect of parallel statistical analysis of biometric data by two criteria pearson

**Akhmetov B.B., Ivanov A.I., Perfilov K.A, Funtikova Yu.V., Alibiyeva Zh.M.**
b_akhmetov@ntu.kz
International Kazakh-Turkish University named after A.Yasavi, Kazakhstan
Penza State University, Russia
Kazakh National Technical University named after K.I. Satpayev, Kazakhstan

**Key words:** multicriteria statistical analysis, a network of private Pearson parallel validation of two statistical hypotheses, biometric data, the test sample.

**Abstract.** In the transition to the use of two-criterion of statistical analysis is possible to obtain solutions with a higher reliability. Twocriterial statistical analysis on parallel inspection of two alternative statistical hypotheses about the normal and uniform distribution can reduce the likelihood of errors, which are proportional to the product of the probabilities of each particular hypothesis testing. This reduces the space requirements of the test sample several times.

УДК 004

# Эффект от параллельного статистического анализа биометрических данных двумя критериями пирсона

**Ахметов Б.Б., Иванов А.И., Перфилов К.А, Фунтикова Ю.В., Алибиева Ж.М.**
b_akhmetov@ntu.kz
Международный Казахско-Турецкий университет имени Х.А.Ясави, г. Туркестан
Пензенский государственный университет, Россия
Казахский национальный технический университет имени К.И. Сатпаева, Алматы

**Ключевые слова:** многокритериальный статистический анализ, сеть частных критериев Пирсона, параллельная проверка достоверности двух статистических гипотез, биометрические данные, тестовые выборки.

**Аннотация.** При переходе к использованию двухкритериального статистического анализа удается получать решения с более высокой достоверностью. Двухкритериальный статистический анализ по параллельной проверке двух альтернативных статистических гипотез о нормальном и равномерном распределениях позволяет снизить вероятность ошибок пропорционально произведению вероятностей проверки каждой частной гипотезы. Это позволяет снизить требования к объему тестовой выборки в несколько раз.

**Введение.** Одним из наиболее популярных при статистическом анализе данных является критерий Пирсона. Хи-квадрат критерию Пирсона полностью посвящена первая часть рекомендаций Госстандарта [1], тогда как все остальные критерии описаны во второй части рекомендаций [2]. Подробное описание критерия Пирсона в первой части рекомендаций Госстандарта [1], отражает факт высокой востребованности именно этого критерия промышленностью. Методики, построенные на использовании хи-квадрат критерия, предполагают проверку некоторой статистической гипотезы о наблюдаемом законе распределения значений $\tilde{p}(x)$. Расчеты ведутся по классической формуле:

$$\chi^2 = n \cdot \sum_{i=1}^{k} \frac{\left(\frac{b_i}{n} - \tilde{p}_i\right)^2}{\tilde{p}_i} \qquad (1)$$

где $b_i$ – число опытов, попавших i-тый интервал гистограммы, $\tilde{p}_i$ – ожидаемая теоретическая вероятность попадания в i-тый интервал гистограммы, n – число опытов в тестовой выборке, k – число столбцов гистограммы.

К сожалению, стандартные методики статистических расчетов (1) при анализе биометрических данных дают недостоверные результаты. Для того чтобы добиться вероятностей ошибок на уровне 0.05 приходится использовать порядка 100 опытов в тестовой выборке.

Главной причиной ошибок при анализе биометрических данных является недостаточный объем данных в исследуемых тестовых выборках [3, 4, 5]. Эта ситуация характерна не только для тестирования средств биометрической защиты информации. Та же самая ситуация возникает и при обработке любых биометрических данных (медицинских, спортивных, биологических). Проблеме совершенствования методик применения хи-квадрат критерия для статистической обработки нечетких биометрических данных уделяется значительное внимание журналом «Biometrics», который регулярно печатает статьи по этой тематике [6, 7, 8] начиная с 30-х годов прошлого века.

В начале 21 века наметилась тенденция решать проблему плохих данных искусственным заполнением пробелов в пустых интервалах гистограммы, так называемым «бутстрап методом» [9], который разрушает естественные корреляционные связи в существенно зависимых биометрических данных. Примерно такого же эффекта удается добиться цифровым сглаживанием гистограмм реальных данных [10]. Для той же цели можно использовать морфинг скрещивание примеров-родителей и получение от них множество примеров-потомков [11, 12].

Данная статья посвящена еще одному направлению исследований, связанному с использованием двух и более статистических критериев. На данный момент известны десятки статистических критериев. Наиболее распространенные статистические критерии проверки гипотез при анализе биометрических данных приведены в таблице 1 с указанием времени их создания.

Таблица 1 – Наиболее популярные статистические критерии

| № п.п. | Название критерия и год создания | Формула |
|---|---|---|
| 1 | Хи-квадрат критерий или критерий Пирсона 1900 г. | $\int_{-\infty}^{+\infty} \frac{\{p(x) - \tilde{p}(x)\}^2}{\tilde{p}(x)} \cdot dx$ |
| 2 | Критерий Крамера-фон Мизеса 1928 г. | $\int_{-\infty}^{+\infty} \{P(x) - \tilde{P}(x)\}^2 \cdot dx$ |
| 3 | Критерий Колмогорова-Смирнова 1933 г. | $\sup_{-\infty < x < +\infty} \left\|P(x) - \tilde{P}(x)\right\|$ |
| 4 | Критерий Смирнова-Крамера- фон Мизеса 1936 г. | $\int_{-\infty}^{+\infty} \{P(x) - \tilde{P}(x)\}^2 \cdot d\tilde{P}(x)$ |
| 5 | Критерий Джини 1941 г. | $\int_{-\infty}^{+\infty} \left\|P(x) - \tilde{P}(x)\right\| \cdot dx$ |
| 6 | Критерий Андерсона-Дарлинга 1952 г. | $\int_{-\infty}^{+\infty} \frac{\{P(x) - \tilde{P}(x)\}^2}{\tilde{P}(x) \cdot \{1 - \tilde{P}(x)\}} \cdot d\tilde{P}(x)$ |

| | | |
|---|---|---|
| 7 | Критерий Купера<br>1960 г. | $\sup\limits_{-\infty\langle x\langle +\infty}\left\{P(x)-\tilde{P}(x)\right\}+\sup\limits_{-\infty\langle x\langle +\infty}\left\{\tilde{P}(x)-P(x)\right\}$ |
| 8 | Критерий Ватсона<br>1961 г. | $\int\limits_{-\infty}^{+\infty}\left\{\tilde{P}(x)-P(x)-\int\limits_{-\infty}^{x}\left[\tilde{P}(x)-P(x)\right]\cdot d\tilde{P}(x)\right\}\cdot d\tilde{P}(x)$ |
| 9 | Критерий Фроцини<br>1978 г. | $\int\limits_{-\infty}^{+\infty}\left\|P(x)-\tilde{P}(x)\right\|\cdot d\tilde{P}(x)$ |
| 10 | Дифференциальный вариант критерия Джини<br>2006 г. | $\int\limits_{-\infty}^{+\infty}\left\|p(x)-\tilde{p}(x)\right\|\cdot dx$ |

Из таблицы 1 видно, что статистические критерии создавались постепенно с 1900 года по настоящее время. Самым последним был создан дифференциальный критерий Джини [13] специально для обработки биометрических данных. Этот критерий оказался самым мощным и построен путем замены в исходном критерии Джини (строка 5 в таблице 1, [14], 1941 год) функций вероятностей на их производные (плотности распределения значений вероятности).

Очевидно, что каждый из критериев таблицы 1 исследует тестовую выборку со своей стороны. Все они дополняют друг друга. Можно попытаться создать некоторый обобщенный критерий, который будет учитывать данные всех 10 критериев таблицы 1, построенных для проверки только ПЕРВОЙ гипотезы наблюдения плотности распределения – $\tilde{p}1(x)$ или ее аналога, в виде функции вероятности – $\tilde{P}1(x)$. Более того, мы можем удвоить число статистических критериев в таблице 1, если каждый из них строить сразу для проверки ДВУХ статистических гипотез $\tilde{p}1(x)$, $\tilde{p}2(x)$ или их интегральных аналогов $\tilde{P}1(x)$, $\tilde{P}2(x)$. Последнее утверждение является далеко не очевидным. Его численному доказательству посвящена данная статья.

**Численный эксперимент для получения описания распределения значений хи-квадрат критерия на конечной тестовой выборки.** Популярность использования хи-квадрат критерия Пирсона в промышленности во многом обусловлена тем, что при $n\to\infty$ его распределение описывается через гамма функцию с m = k-1 числом степеней свободы:

$$p_{\chi^2}(n=\infty, m=k-1, x)=\frac{1}{2^{\frac{m}{2}}\cdot\Gamma\left(\frac{m}{2}\right)}\cdot x^{\frac{m}{2}-1}\cdot e^{\frac{-x}{2}} \qquad (2)$$

Аналитическое описание (2) получено Пирсоном в 1904 году и играло крайне важную роль в первой половине 20-го века, когда вычислительные возможности, используемые при статистической обработке данных были весьма и весьма ограниченными.

К сожалению, традиционное применение хи-квадрат критерия для биометрических данных дает неудовлетворительные результаты. Одной из причин является ошибка, возникающая из-за конечной тестовой выборки. Практика показывает, что при конечной тестовой выборке (например, для n=81) число степеней свободы у хи-вадрат распределения оказывается не целым (дробным) и именно из-за этого возникает значительное расхождение:

$$p_{\chi^2}(n=81, m\neq k-1, x)\neq p_{\chi^2}(n=\infty, m=k-1, x) \qquad (3)$$

Ошибку из-за конечности тестовой выборки можно учесть путем численного эксперимента. Сегодня повторить эксперимент на компьютере 1000000 раз вполне возможно, что дает значения функции распределения значений с приемлемой для практического применения погрешностью.

При организации численного эксперимента будем исходить из того, что должны проверяться две статистические гипотезы. Первая гипотеза состоит в том, что данные тестовой выборки имеют нормальный закон распределения значений. Вторая гипотеза состоит в том, что данные этой же выборки могут иметь нормальный закон распределения значений. Как следствие, при организации численного эксперимента необходимо использовать два программных генератора псевдо случайных данных, как это показано на блок-схеме рисунка 1.

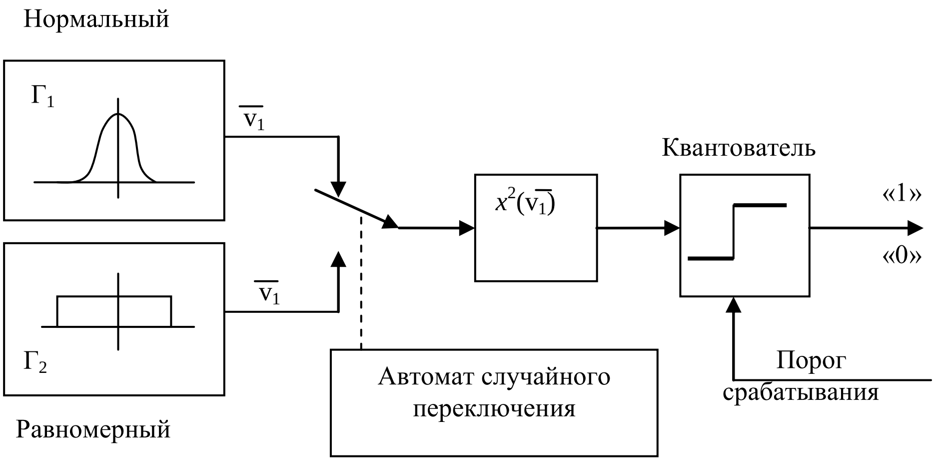



Рис. 1 – Блок-схема организации численного эксперимента по оценке мощности одномерного критерия хи-квадрат

Каждый из генераторов случайных данных $\Gamma_1$ (нормальные данные) и $\Gamma_2$ (данные с равномерным законом распределения) случайным образом подаются на вход вычислителя значения хи-квадрат критерия (1). Далее значения хи-квадрат критерия должны сравниваться с некоторым порогом квантователя. Если значение хи-квадрат менее порога, то принимается решение о нормальности исследуемых входных данных. Если значение хи-квадрат критерия (1) оказывается выше или ниже порога, то принимается решение о наибольшей справедливости одной из гипотез.

**Проверка первой гипотезы о нормальном законе распределения значений для конечной тестовой выборки.** Будем исходить из того, что по критерию хи-квадрат требуется распознать ситуацию появление данных, соответствующих серии из 81 отсчетов, полученных от нормального генератора $\Gamma_1$. Для этой цели будем вычислять математическое ожидание тестовой выборки – E(x) и ее среднеквадратическое отклонение $-\sigma(x)$. Далее будем строить гистограмму, состоящую из k=9=$\sqrt{81}$ столбцов, равномерно покрывающих интервал от минимального значения – ($E(x)-3\cdot\sigma(x)$) до максимального значения – ($E(x)+3\cdot\sigma(x)$). При этом значения критерия хи-квадрат будем вычислять следующим образом:

$$\chi^2(\Phi)=81\cdot\sum_{i=1}^{9}\frac{\left(\frac{b_i}{81}-\frac{1}{\sigma(x)\sqrt{2\pi}}\int\limits_{x_i}^{x_{i+1}}\exp\left\{\frac{-(E(x)-u)^2}{2\cdot(\sigma(x))^2}\right\}du\right)^2}{\frac{1}{\sigma(x)\sqrt{2\pi}}\int\limits_{x_i}^{x_{i+1}}\exp\left\{\frac{-(E(x)-u)^2}{2\cdot(\sigma(x))^2}\right\}du} \tag{4}$$

где пределы интегрирования $x_1$, $x_2$, …, $x_{10}$ – это границы равномерных интервалов, на которых строится гистограмма частот появления данных в тестовой выборке.

На рисунке 2 приведены кривые гистограмм распределения значений хи-квадрат критерия для данных, полученных от двух программных генераторов.

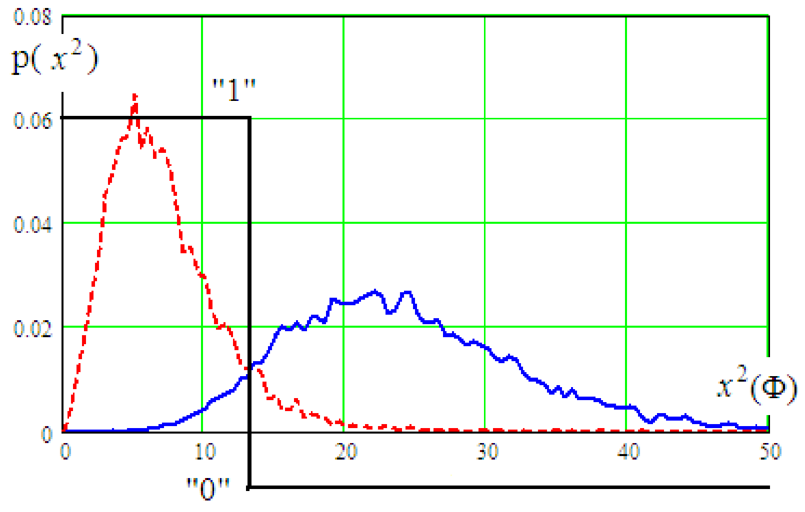



Рис. 2 – Выделение данных с нормальным законом распределения значений (пунктирная линия) при проверке первой гипотезы

Из рисунка 2 видно, что компаратор, принимающий решение об обнаружении входной нормальной последовательности должен давать состояние «1» в интервале значений от 0 до 14. Порог переключения компаратора в состояние «0» – 14. В этом случае вероятности ошибок первого и второго рода оказываются одинаковыми $P_1 = P_2 = P_{EE} = 0.054$.

**Проверка второй гипотезы о равномерном законе распределения значений для конечной тестовой выборки.** Будем исходить из того, что по критерию хи-квадрат требуется распознать ситуацию появление данных, соответствующих серии из 81 отсчетов, полученных от генератора данных с равномерным законом – $\Gamma_2$. Для этой цели будем находить max(x) и min(x) в тестовой выборке. Далее будем строить гистограмму, состоящую из k=9=$\sqrt{81}$ столбцов, равномерно покрывающих интервал от min(x) до max(x). При этом значения критерия хи-квадрат будем вычислять следующим образом:

$$\chi^2(\text{const}) = 81 \cdot \sum_{i=1}^{9} \frac{\left(\frac{b_i}{81} - \frac{1}{9}\right)^2}{\frac{1}{9}} \tag{5}$$

где границы интервалов гистограммы находятся следующим образом:

$$x_i = \min(x) + \frac{(\max(x) - \min(x)) \cdot i}{10} \tag{6}$$

На рисунке 3 приведены кривые гистограмм распределения значений хи-квадрат критерия для данных, полученных от двух программных генераторов.

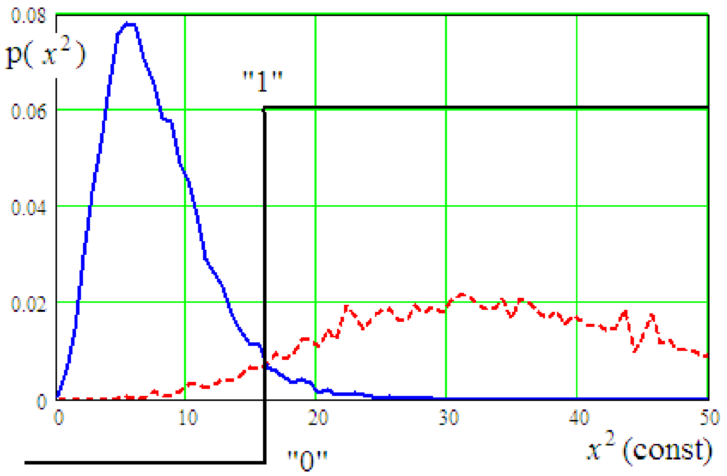



Рис. 3 – Выделение данных с нормальным законом распределения значений (пунктирная линия) при проверке второй гипотезы

Из рисунка 3 видно, что компаратор, принимающий решение об обнаружении входной нормальной последовательности должен давать состояние «1» в интервале значений от 17 и выше. Порог переключения компаратора в состояние «0» – 16. В этом случае вероятности ошибок первого и второго рода оказываются одинаковыми $P_1 = P_2 = P_{EE} = 0.054$.

**Обобщенный критерий хи-квадрат, учитывающий параллельную проверку двух гипотез.** Хи-квадрат критерий, построенный под поверку первой гипотезы (4) и хи-квадрат критерий, построенный под проверку второй гипотезы (5) – это две разных нелинейных функций преобразования, имеющих два выходных компаратора, настроенных по разному. Так как эти критерии дополняют друг друга, обобщим их с использованием логической функции «или»:

$$\begin{cases} p(x) = \dfrac{1}{\sigma(x)\sqrt{2\pi}} \cdot \exp\left\{\dfrac{-(E(x)-x)^2}{2\cdot(\sigma(x))^2}\right\}, \text{ если } (\chi^2(\Phi) \le 14) \wedge (\chi^2(\text{const}) \ge 16); \\ p(x) = \text{const}, \text{ если } (\chi^2(\Phi) \ge 14) \vee (\chi^2(\text{const}) \le 16). \end{cases} \quad (7)$$

Практика показала, что для обобщенного решающего правила (7) по сравнению с более простыми решающими правилами происходит значительное снижение вероятностей ошибок первого и второго рода $P_1 = P_2 = P_{EE} \approx 0.0025$. То есть, частные критерии хи-квадрат (4) и (5) обладают высоким уровнем согласованности принятия ими верных решений, при этом их ошибочные решения, оказываются слабо коррелированны.

**Заключение.** Если пользоваться хи-квадрат критерием по стандартным методикам [1], то для тестовой выборки в 81 отсчет мы получаем вероятности ошибок на уровне 0.054. Однако как только мы переходим к учету ошибок, возникающих из-за конечности тестовой выборки (3) и применяем обобщенный хи-квадрат критерий (7), то вероятность ошибок снижается примерно в 20 раз. Столь существенное снижение вероятностей ошибок может быть достигнуто только при размерах тестовой выборки в 800 отсчетов, это эквивалентно 10-ти кратному снижению требований к размерам тестовой выборки.

Также следует обратить внимание на то, что данная статья посвящена только критерию хи-квадрат, однако все что в статье изложено, оказывается справедливо и для других статистических критериев таблицы 1. За счет использования обобщенного критерия, учитывающего 10 или 20 частных статистических критериев, видимо, удастся добиться более чем 10-ти кратного снижения требований к размерам тестовых выборок при статистической обработке биометрических данных.

ӘОК 004

## ПИРСОННЫҢ ЕКІ КРИТЕРИІМЕН БИОМЕТРИЯЛЫҚ ДЕРЕКТЕРДІ ПАРАЛЛЕЛЬДІ СТАТИСТИКАЛЫҚ ТАЛДАУДЫҢ САЛДАРЫ

**Ахметов Б.Б., Иванов А.И., Перфилов К.А, Фунтикова Ю.В., Алибиева Ж.М.**
b_akhmetov@ntu.kz
Х.А. Яссауи атындағы Халықаралық Қазақ-Түрік университеті, Туркестан қ.
Пенза мемелекеттік университеті, Ресей
Қ.И. Сәтбаев атындағы Қазақ Ұлттық техникалық университеті, Алматы қ.

**Кілт сөздер:** көпкритерийлі статистикалық талдау, жеке Пирсон критерилерінің желісі, екі статистикалық гипотезаның дұрыстығын параллельді тексеру, биометриялық деректер, тестілік іріктеулер.

**Аңдатпа.** Екі критерийлі статистикалық талдауды қолдануға өту кезінде дұрыстығы тым жоғары шешімдер алуға мүмкіндіктер ашылады. Қалыпты және біркелкі тарату екі альтернативалы статистикалық гипотезеларын параллельді тексеру бойынша екі критерийлі статистикалық талдау әрбір жекеше гипотезалардың тексеру ықтималдылықтарының туындысына пропорционал қателер ықтималдылығын төмендетуге мүмкіндік береді. Бұл тестілік іріктеулер көлеміне қойылатын талаптарды бірнеше рет төмендетуге жағдайлар туғызады.

### ЛИТЕРАТУРА

[1] Р 50.1.037-2002 Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть I. Критерии типа $\chi^2$. Госстандарт России. Москва-2001 г., 140 с.

[2] Р 50.1.037-2002 Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть II. Непараметрические критерии. Госстандарт России. Москва-2002 г., 123 с.

[3] Ахметов Б.С., Иванов А.И., Фунтиков В.А., Безяев А.В., Малыгина Е.А. Технология использования больших нейронных сетей для преобразования нечетких биометрических данных в код ключа доступа. // Монография, Казахстан, г. Алматы, ТОО «Издательство LEM», 2014 г. -144 с. (http://portal.kazntu.kz/files/publicate/2014-06-27-11940.pdf)

[4] Ахметов Б.С., Волчихин В.И., Иванов А.И., Малыгин А.Ю. Алгоритмы тестирования биометрико-нейросетевых механизмов защиты информации // Казахстан, Алматы, КазНТУ им. Сатпаева, 2013 г. – 152 с. ISBN 978-101-228-586-4, http://portal.kazntu.kz/files/publicate/2014-01-04-11940.pdf

[5] Ахметов Б.С., Надеев Д.Н., Фунтиков В.А., Иванов А.И., Малыгин А.Ю. Оценка рисков высоконадежной биометрии. // Монография. – Алматы: Из-во КазНТУ им. К.И. Сатпаева, 2014 г. – 108 с.

[6] Cochran W. G. Some Methods of Strengthening the Common χ2 Tests // Biometrics, 1954. – V. 10. – P. 417-419.

[7] Gilbert R.J A sample formula for cuterpolating tables of χ2 // Biometrics, 1977. – V. 33. – P. 383-385.

[8] Pearson E.S. Note on an approximation to the distribution of non-central χ2 // Biometrics, 1959. – V. 46. – P. 364-366.

[9] Болл Р.М., Коннел Дж.Х., Панканти Ш., Ратха Н.К., Сеньор Э.У. Руководство по биометрии. Москва: Техносфера, 2007. – 368 с., ISBN 978-594836-109-3

[10] Ахметов Б.С., Иванов А.И., Серикова Н.И., Фунтикова Ю.В. Алгоритм искусственного повышения числа степеней свободы при анализе биометрических данных по критерию согласия хи-квадрат. Вестник национальной академии наук республики Казахстан. №5, 2014 г. с. 28-34.

[11] Akhmetov B.S., Ivanov A.I., Kachalin S.V., Seilova N.A., Doszhanova A.A. Addition fuzzy biometric data morphingre production examples of parents in several generations o examples descendants. Wulfenia Jornal vol 21, No.7; jun 2014. Klagenfurt, Austria, ISSN: 1561-882x: office@multidisciplinarywulfenia.org

[12] Bakhytzhan Akhmetov, Alexander Ivanov, Alexander Malyghin, Sergey Kachalin & Nurgul Seilova // Morph-Reproduction Examples of Parents in Several Generations of Examples Descendants // International Conference on Global Trends in Academic Research (ICMRP-December 17-18, 2014) at Kuala Lumpur, Malaysia

[13] Малыгин А.Ю., Волчихин В.И., Иванов А.И., Фунтиков В.А. Быстрые алгоритмы тестирования нейросетевых механизмов биометрико-криптографической защиты информации / Пенза-2006 г., Издательство Пензенского государственного университета, 161 с.

[14] Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. М.: ФИЗМАТЛИТ, 2006 г., 816 с.

### REFERENCES

[1] P 50.1.037-2002 Recommendations about standardization. Applied statistics. Rules of check of a consent of skilled distribution with the theoretical. Part I. Criteria like χ2. Gosstandart of Russia. Moscow-2001, 140 p. (in Russ.).

[2] P 50.1.037-2002 Applied statistics. Rules of check of a consent of skilled distribution with the theoretical. Part II. Nonparametric criteria. Gosstandart of Russia. Moscow-2002, 123 p. (in Russ.).

[3] Akhmetov B.S., Ivanov A.I., Funtikov V.A., Bezyaev A.V., Malygina E.A. Tekhnologiya of use of big neural networks for transformation of indistinct biometric data to an access key code. // The monograph, Kazakhstan, Almaty, LEM Publishing House LLP, 2014-144 with. (http://portal.kazntu.kz/files/publicate/2014-06-27-11940.pdf) (in Russ.).

[4] Akhmetov B.S., Volchikhin V.I., Ivanov A.I., Malygin A.Yu. Algorithms of testing of biometriko-neural network mechanisms of information security // Kazakhstan, Almaty, KAZNTU of Satpayev, 2013 - 152 p. ISBN 978-101-228-586-4, http://portal.kazntu.kz/files/publicate/2014-01-04-11940.pdf (in Russ.).

[5] Akhmetov B.S., Nadeev D. N., Funtikov V.A., Ivanov A.I., Malygin A.Yu. Otsenka of risks of highly reliable biometrics.//Monograph. – Almaty: Publishing house of KAZNTU of K.I. Satpayev, 2014 - 108 p. (in Russ.).

[6] Cochran W. G. Some Methods of Strengthening the Common χ2 Tests // Biometrics, 1954. - V. 10. - P. 417-419.

[7] Gilbert R.J A sample formula for cuterpolating tables of χ2 // Biometrics, 1977. - V. 33. - P. 383-385.

[8] Pearson E.S. Note on an approximation to the distribution of non-central χ2 // Biometrics, 1959. - V. 46. - P. 364-366.

[9] Ball R. M., Connell J. H., Pank anti-Sh., Ratkh N. K., Senior E.U. Rukovodstvo on biometrics. Moscow: Technosphere, 2007.-368 p, ISBN 978-594836-109-3. (in Russ.).

[10] Akhmetov B.S., Ivanov A.I., Serikova N.I., Funtikova Yu.V. Algoritm of artificial increase of number of degrees of freedom in the analysis of biometric data on criterion of a consent a chi-square. Bulletin of national academy of Sciences of the Republic of Kazakhstan. No. 5, 2014 P. 28-34. (in Russ.).

[11] Akhmetov B.S., Ivanov A.I., Kachalin S.V., Seilova N.A., Doszhanova A.A. Addition fuzzy biometric data morphingre production examples of parents in several generations o examples descendants. Wulfenia Jornal vol 21, No. 7; jun 2014. Klagenfurt, Austria, ISSN: 1561-882x: office@multidisciplinarywulfenia.org

[12] Bakhytzhan Akhmetov, Alexander Ivanov, Alexander Malyghin, Sergey Kachalin & Nurgul Seilova // Morph-Reproduction Examples of Parents in Several Generations of Examples Descendants//International Conference on Global Trends in Academic Research (ICMRP-December 17-18, 2014) at Kuala Lumpur, Malaysia

[13] Malygin A.Yu., Volchikhin V.I., Ivanov A.I., Funtikov V.A. Fast algorithms of testing of neural network mechanisms of biometriko-cryptographic information security / Penza-2006, Publishing house of the Penza state university, 161 p. (in Russ.).

[14] Kobza player A.I. Applied mathematical statistics. For engineers and scientists. M.: FIZMATLIT, 2006, 816 p. (in Russ.).

Поступила 17.02.2015 г.