Markov model of evaluation of learning results

V.I. Serbin
Astrakhan State University,
414056, Astrakhan, Tatishcheva, 20A,
e-mail V.I.Serbin@mail.ru

Abstract

The effectiveness of the management of the learning process depends on the adequacy and objectivity of evaluations of the learning outcomes, for which it is necessary to determine latent learning parameters, such as the training level of the trainer and the difficulty of the training and testing tasks. However, the existing meth-ods of processing test results based on the Rasch model allow us to determine not the latent parameters of training themselves, but only the relationships between them. In connection with this, it is proposed to use the Markov model of training, which in-cludes three states: the state of knowledge transfer, the state of the training, and the state of knowledge control using testing. The author shows that application of the time spent by trainees to solve tasks in the framework of this model provides addi-tional information on the knowledge, abilities, and skills of trainees, assess the values of latent parameters and, based on this, better manage the learning process. The paper presents a mathematical apparatus used within the framework of the proposed model, including formalized methods of information theory.

Key words: training system, knowledge, abilities, skills, latent parameter, training, testing, mathematical model, Markov model, differential entropy, Rasch model.

1 Introduction

The goal of any education system is the management of the learning process. To solve this problem, it is necessary to create a mathematical model that can ade-quately describe and analyze, and then control the learning process [1]. The result of the operation of the education system should be an increase in the level of knowledge, abilities and skills of the trainee. One of the important management problems in modern education system is the control and evaluation of acquired knowledge, abilities and skills. The categories of knowledge, abilities and skills in an automated teaching system in pedagogical work are usually determined in the following way.

Knowledge is the elements of information, in which learning is understood as facts, concepts, rules, algorithms, heuristics, the basic laws of the subject area, as well as decision-making strategies in this area, allowing people to solve specific production, scientific and other tasks. However, knowledge does not mean that they will be used.

In: Marco Schaerf, Massimo Mecella, Drozdova Viktoria Igorevna, Kalmykov Igor Anatolievich (eds.): Proceedings of REMS 2018 – Russian Federation & Europe Multidisciplinary Symposium on Computer Science and ICT, Stavropol – Dombay, Russia, 15–20 October 2018, published at http://ceur-ws.org

Abilities are usually understood as the abilities necessary to perform a function in the work. A necessary, but not sufficient condition for the manifestation of abilities is the possession of appropriate knowledge.

Skills are the ability to quickly and easily perform the actions necessary during the work. Speaking about the skill, it is necessary to indicate a certain normative level, the achievement of which is necessary for the effective performance of the work.

A rather detailed review of these methods is given in [2, 3]. Their analysis shows that the control of training is carried out through testing, and its results are estimated in dichotomous, and at best, in ordinal scales. In this case, the time spent on testing is either not taken into account at all, or is used only as a threshold value - to stop the testing process. Examples of such multimedia training systems are described in [4]. The Rasch model, used to process test results and obtain estimates, allows us to find only the relationship between independent latent learning parameters: the complexity of tasks and the level of training of trainees - but these values themselves remain unknown.

The main goal of this work is the development of a method for determining latent training parameters. A study has been made of the connection between the time-based learning model and the Rasch model, and describes the algorithm for managing the learning process based on the results of training and testing, taking into account the time spent.

Consider a mathematical model of learning that allows you to manage the work of the learning system in different modes. This model describes the behavior of the components of the learning system as a function of time.

The object of management in this model is the learning process. The purpose of management is to improve the quality of education.

To solve the problem of managing the learning process, we present the model of the learning system as a system with several discrete states and continuous time. We describe this model as a random Markov process consisting of the following states: the state of knowledge transfer, the state of the solution of training tasks, or training, and the state of control of knowledge, or testing.

Figure 1 shows the Markov model of a training system of three states.

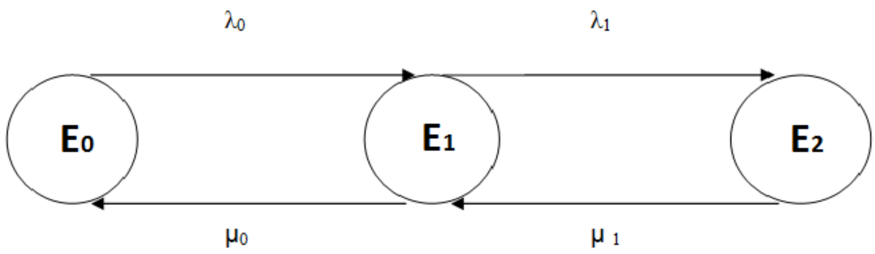

Figure 1: Markov model of a training system of three states

The Markov model of a learning system consists of three states:
E_0 -– state of knowledge transfer
E_1 -– state of training
E_2 -– state of testing
λ_0 -– transition intensity from E_0 to E_1
μ_0 -– transition intensity from E_1 to E_0
λ_1 -– transition intensity from E_1 to E_2
μ_1 — transition intensity from E_2 to E_1

Let the probability that at time t the system is in the state E_i is equal to $P_i(t), i = 0, 1, 2$. Then the system of Kolmogorov equations describing the behavior of the system in time will have the form

$$\begin{cases} \frac{dP_0(t)}{dt} = -\lambda_0 P_0(t) + \mu_0 P_1(t) \\ \frac{dP_1(t)}{dt} = \lambda_0 P_0(t) - (\mu_0 + \lambda_1) P_1(t) + \mu_1 P_2(t) \\ \frac{dP_2(t)}{dt} = \lambda_1 P_1(t) - \mu_1 P_2(t) \end{cases}$$

Here time $t \geq 0$ and probabilities must satisfy the normalizing condition $P_0(t) + P_1(t) + P_2(t) = 1$.

This system is a Markov process of birth and death with three states with continu-ous time and constant intensities.

The paper considers three models of training:

1. The model of learning from one state, which allows you to evaluate the solution to the learner's task or series of tasks in the state of training or testing;

2. The model of learning from two states: active learning (transfer of knowledge and training) and knowledge control;

3. Model of learning from three states: knowledge transfer, training and knowledge control.

2 The model of learning from one state

Figure 2 shows the Markov model of a learning system from one state.

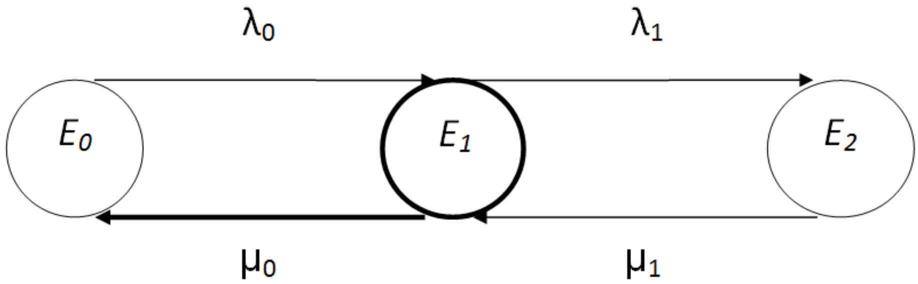

Figure 2: Markov model of a training system of one states

In this model it is assumed that $\lambda_0 = \lambda_1 = \mu_1 = 0$

In [5], the process of solving the problem is presented as a random process and a formula is obtained for calculating the estimate, depending on the time. Let the random variable T be the time of the solution of the problem. Then the probability of solving a problem within a time not exceeding a value t, or a distribution function of a random variable T is

$$F_\mu(t) = 1 - e^{-\mu t}$$

and the estimate for solving the problem for a time not less than t is [5]

$$p_\mu(t) = 1 - F_\mu(t) = e^{-\mu t}$$

When $t = 0$ this estimate is equal 1, the rate of change of the estimate decreases with time, the evaluation curve itself is exponential and monotonically decreasing, and $t \to \infty$ tends asymptotically to 0, which indicates a slow-asymptotic nature of the process, and the random variable T obeys the exponential law.

We represent the process of solving a problem as an information processing pro-cess or as a sequence of elementary Data Transformation Operations (DTO) [6]. Then is the number of DTO performed per unit time, or the speed of processing information. The magnitude, the inverse of the intensity $\tau = 1/\mu$ is equal to the mathematical expectation of a random variable T - this is the average time of exe-cution of one DTO. We shall treat this quantity as the difficulty of the problem [7].

The distribution density of a random variable T is

$$f_\mu(t) = \mu e^{-\mu t}$$

and the differential entropy, or measure of the average information processed in the process of solving the problem, is [8]

$$H(T) = \ln(e/\mu) = \ln e\tau = \alpha$$

Then

$$\tau = e^\alpha - 1$$

If minutes are selected as the unit of time, then τ is measured in min, μ in - dto/min where dto is one DTO, and α – in logits. Values τ and corresponding to it α are continuous latent variables [9] and for their finding it is necessary to use methods of statistical estimation.

Let the problem be decided by n trainees. In [7] it was shown by the maximum likelihood method that the point estimate of the difficulty of the problem is equal to the mean time of solving a series of problems

$$\tau = \sum_{i=1}^{n} t_i \Big/ n = \bar{t}$$

The disadvantage of this model is that it takes into account only the measure of the difficulty of the task and does not take into account the level of training of the trainee.

3 The model of a training system that includes two states

Figure 3 shows the Markov model of a learning system from two states.

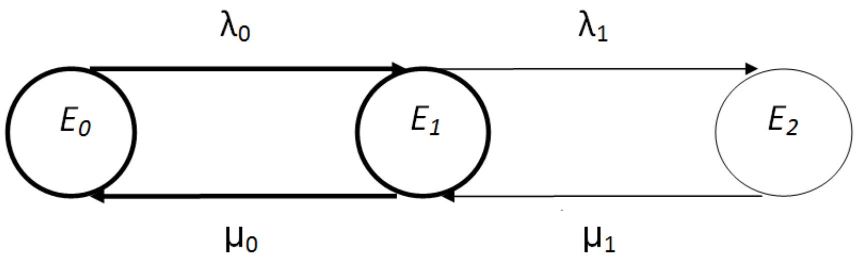

Figure 3: Markov model of a training system of two states

In this model it is assumed that $\lambda_1 = \mu_1 = 0$

This model describes the learning system as a random Markov process of the "birth-death" type, consisting of two states: the state of active learning (knowledge transfer and training) and the state of knowledge control [10, 17]. Let the random variable T be the learning time; $P_0(t)$ -– the probability of successful completion of the training process not later than in time t; $P_1(t) = 1 - P_0(t)$ -– the probability of a successful completion of the learning process not earlier than in time t, or the evaluation of training. The behavior of such a learning model is described by a system of Kolmogorov equations [10]

$$\begin{cases} \frac{dP_0(t)}{dt} = -\lambda_0 P_0(t) + \mu_0 P_1(t) \\ \frac{dP_1(t)}{dt} = \lambda_0 P_0(t) - \mu_0 P_1(t) \end{cases} \tag{1}$$

where λ_0 is the intensity of the state of active learning; μ_0 -– the intensity of the state of the solution of the problem and the control of knowledge. The normalization condition $P_0(t) + P_1(t) = 1$ for any $t \geq 0$.

Solving the system (1) for $P_0(0) = 0$ and $P_1(0) = 1$, we obtain

$$\begin{cases} P_0(t) = P_0(\infty)\left(1 - e^{-(\lambda_0+\mu_0)t}\right) \\ P_1(t) = 1 - P_0(t) \end{cases} \tag{2}$$

Values

$$P_0(\infty) = \lim_{t\to\infty} P_0(t) = \mu_0/(\lambda_0 + \mu_0), P_1(\infty) = \lim_{t\to\infty} P_1(t) = \lambda_0/(\lambda_0 + \mu_0) - - \tag{3}$$

limit values of process parameters.

Imagine the learning process as a process of information processing or as a sequence of elementary DTO [6]. Then the intensity λ_0 is the number of DTO per-formed per unit time in the state of active learning. The inverse of the intensity $\sigma_0 = 1/\lambda_0$ -– the time of execution of one DTO, will be taken as the level of the student's preparation. Intensity μ_0 is the number of DTO performed per unit of time in the state of knowledge control, or the processing speed of information. The inverse of the intensity $\tau_0 = 1/\mu_0$ is the average time of execution of one DTO. We shall treat this quantity as the difficulty of the problem [10].

In (2) the probability of a successful completion of the learning process is equal

$$P_0(t) = P_(\infty)F_{\lambda_0\mu_0}(t)$$

, to the product of the probability of completing the learning process "in general" [3] and the probability

$$F_{\lambda_0\mu_0}(t) = (1 - e^{-(\lambda_0+\mu_0)t}) -$$

complete the learning process no later than in time t.

The probability of completing the training process not earlier than in time t is equal to

$$p_{\lambda_0\mu_0}(t) = 1 - F_{\lambda_0\mu_0}(t) = e^{-(\lambda_0+\mu_0)t} = e^{-\lambda_0 t}e^{-\mu_0 t} = p_{\lambda_0}(t)p_{\mu_0}(t), \tag{4}$$

where

$$p_{\lambda_0}(t) = e^{-\lambda_0 t}$$

is the probability of completing the active learning process in a time not less than t, and

$$p_{\mu_0}(t) = e^{-\mu_0 t}$$

is the probability of completing the process of monitoring knowledge in a time not less than t.

From (4) follows the independence of the processes of active learning and knowledge control. According to the local independence axiom of Lazarsfeld [9], we will assume that the quantities of λ_0 and μ_0 are also independent. The active learning time distribution function is

$$F_{\lambda_0}(t) = 1 - p_{\lambda_0}(t) = 1 - e^{-\lambda_0 t} \ ,$$

and the density of distribution –

$$f_{\lambda_0}(t) = \lambda_0 e^{-\lambda_0 t}.$$

Then the differential entropy, or the average amount of information, is

$$H(T) = \ln(e/\lambda_0) = \ln e\sigma_0 = \beta$$

and

$$\sigma_0 = e^{\beta-1}$$

Accordingly, the distribution function of knowledge control time is

$$F_{\mu_0}(t) = 1 - p_{\mu_0}(t) = 1 - e^{-\mu_0 t} \ ,$$

distribution density –

$$f_{\mu_0}(t) = \mu_0 e^{-\mu_0 t},$$

differential entropy –

$$H(T) = \ln(e/\mu_0) = \ln e\tau_0 = \alpha$$

and

$$\tau_0 = e^{\alpha-1}.$$

If σ_0 and τ_0 are measured in min, then λ_0 and μ_0 in dto/min, and α and β in $logits$.

As a result, the limit value of the probability of successful solution of tasks can be represented by the formula

$$P_0(\infty) = \mu_0/(\lambda_0 + \mu_0) = 1/(1 + \lambda_0/\mu_0) = 1/(1 + \tau_0/\sigma_0) = 1/(1 + e^{\alpha-\beta}) \tag{5}$$

Values μ_0 , λ_0 , τ_0 , σ_0 , α and β are continuous latent variables [11], and for their finding, it is permissible that n trainers take part in the training, and each trainee solves m problems. Based on the results of solving these problems, we make up a time table of nm size. The element of the table t_{ij} is the time spent by the i-th student to solve the j-th problem. Letλ_i - the intensity of the active learning of the i-th trainee, μ_j - the intensity of the solution of the j-th problem. In [12] it was shown by the method of maximum likelihood that point estimates of the quantities λ_i and μ_j are equal to the limiting values $\lambda_i^{(k)}$ and $\mu_j^{(k)}$ by $k \to \infty$. The values of $\lambda_i^{(k)}$ and $\mu_j^{(k)}$ are found using iterative formulas

$$\begin{cases} \lambda_i^{(k+1)} & = \sum_{j=1}^{m} \left(\frac{\lambda_i^{(k)}}{\lambda_i^{(k)} + \mu_j^{(k)}} \right) \Big/ \sum_{j=1}^{m} t_{ij}, i = 1, 2, ..., n \\ \mu_j^{(k+1)} & = \sum_{i=1}^{n} \left(\frac{\mu_i^{(k)}}{\lambda_i^{(k)} + \mu_j^{(k)}} \right) \Big/ \sum_{i=1}^{n} t_{ij}, j = 1, 2, ..., m \end{cases} \tag{6}$$

4 Connection with the Rash model

In the one-parameter model of Rasch's training [13], the equation determining the probability that the trainee with the level of preparation s will perform the task of difficulty d , the so-called success function,

$$P(d, s) = 1/(1 + d/s) \tag{7}$$

or

$$P(d, s) = 1/(1 + e^{\delta-\theta}), \tag{8}$$

where $d = e^{\delta}$ (or $\delta = \ln d$), $s = e^{\theta}$ (or $\theta = \ln s$).

Up to notation fomula (7) coincides with the formula for $P_0(\infty)$, and [3] formula (8) – with the formula (5). Thus, the education system model from two states can be considered as the development of the Rasch model, and the Rasch success function should be considered equal to the limit value of the probability of successful solution of tasks.

5 Comparative analysis of models

Let the training be conducted in three stages, and the education system can be in one of three states: the transfer of knowledge, the development of abilities through training and testing skills through testing. Abilities suggest that the learner simply knows how to solve the problem, and skills – that the problem should be solved as quickly as possible. In this case, the maximum permissible or desirable time of solving the problems can be specified, in their totality and/or for each task.

To test knowledge, a set of control questions on the topic of my study is used. The development and testing of skills is carried out in the course of training under the guidance of the teacher with the obligatory solution of all training tasks. Testing skills in the testing process presupposes an independent solution for the learner of a set of test tasks. They should be arranged in order of increasing difficulty and, in addition, be adequate to the levels of training of trainees. In the case of training, the complexity of tasks should not be lower than the level of training of trainees, because as a result, this level should be increased. The complexity of test tasks cannot be higher than the level of training of trainees, since these tasks include tasks that the learner must already be able to solve, and when checking skills, the speed of problem solving is important. The choice of the relationship between the level of complexity of tasks and the level of training of trainees at the stages of training and testing is an independent task. These relationships can be selected, for example, by one of the iterative methods.

Let n students participate in the learning process. After each stage of the training, you should check the result with the help of test tasks. Let these tasks consist of m questions or tasks. Based on the results of the tests, we will compile two tables consisting of n rows and m columns: table A and table T. Table A contains the results by the dichotomous principle: $a_{ij} = 1$ if the i-th trainee correctly performed the j-th job and, 0 otherwise. Table T contains a set t_{ij} – i.e. times spent by the i-th trainee for the execution of the j-th task.

Let us consider the peculiarities of processing the obtained results in the Rasch model and in the model of the training system from two states.

The Rasch model for determining latent parameters is one of the sections of the theory of latent analysis of Lazarsfeld [14], whose goal was to process the results of sociological surveys. Based on the results of such surveys, tables were constructed based on the dichotomous principle. Therefore, only the table A is used in the Rasch model, which is called the response matrix. Before processing, rows and columns containing all zeros and all units are deleted from it. After this, the table is processed - this uses the logistic one-parameter model of Rasch in the modification of Yu.M. Neumann [15].

Thus, the Rasch model is not suitable for evaluating the results of the training, in the course of which the learner solves all problems, and the results table A consists of one units.

In the case of testing, in which skills are assessed, the time of solving the problem becomes a decisive factor in obtaining an assessment, because skills require not only and not so much the ability to correctly solve problems (as reflected in table A), but also solve them in an acceptable time (as reflected in table T). Therefore, you can expect table A of the test results to contain many rows and columns of only one unit. This makes the Rasch model less suitable for evaluating the test results than our model of the two-state training system.

6 Managing the learning process

Based on the results obtained, the following procedure for operating the training system is proposed.

Before starting the training with the help of an expert group, tasks are selected, problems solved and their complexity determined [16]. Then, students are tested and the levels of their training are determined. After that, the transfer of knowledge, training and testing to determine the level of knowledge based on the use of test tasks.

In carrying out the training, we will adopt the Yerkes-Dodson law as the axiom, which states that "as the intensity of motivation increases, the quality of activity changes along a bell-shaped curve: first it rises and then gradually decreases" [16]. Therefore, the solution of the sequence of problems ordered in increasing difficulty is fulfilled. First, the tasks are solved, the difficulty of which coincides with the current level of training of trainees, then the level of training of trainees is assessed. If it is increased, then the following sequence of tasks with a higher difficulty value is selected and the training continues. The process is terminated when, at the next stage, the solution of tasks does not lead to an increase in the level of training of trainees. It is assumed that the "maximum" result of training is achieved.

After the training, testing is performed with the help of tasks, the difficulty of which is equal to the current level of training of trainees and the final level of training of trainees is determined.

The results of the research presented in this article will be used by the author in developing the automated training system at the rate of mathematical logic and the theory of algorithms.

7 The model of a training system that includes three states

Since the work of any training system is divided, as a rule, into two stages, we present the work of the model of training from three States in the form of two interrelated processes: first, the process of knowledge transfer and training, and then the process of training and testing.

In the process of transferring knowledge and training, the system operates in the states E_0 and E_1. (see Figure 4), it is assumed that $\lambda_1 = \mu_1 = 0$.

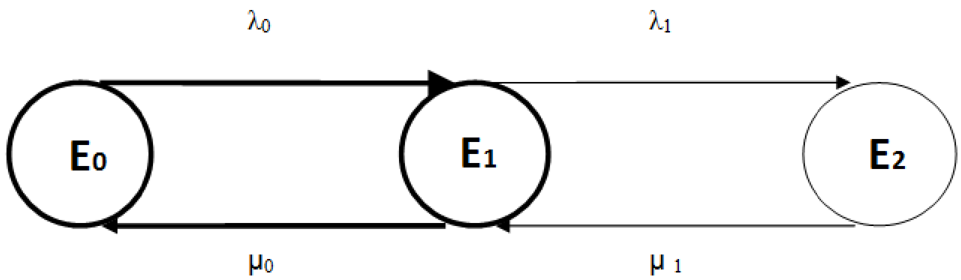

Figure 4: Markov model of knowledge transfer and training process

In the process of training and testing, the system operates in the states E_1 and E_2. (see Figure 5), it is assumed that $\lambda_0 = \mu_0 = 0$.

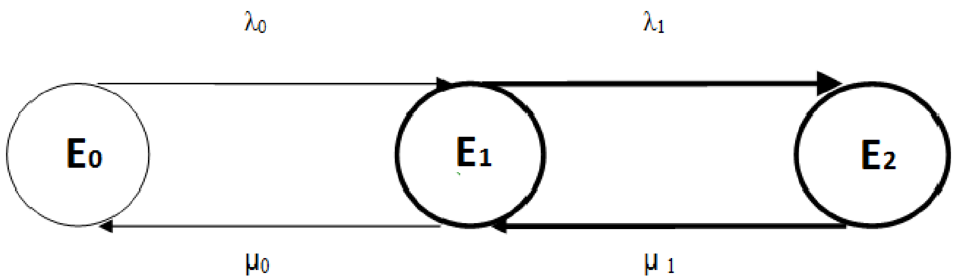

Figure 5: Markov model of the process of training and testing

8 Conclusions

1. Using the Rasch model, in which the evaluation of learning outcomes is carried out by processing responses to test tasks presented in dichotomous and ordinal scales, it is impossible to find the values of all latent learning parameters, but only the ratio between these parameters.

2. The learning model described in the work, taking into account the time spent on the tasks, allows you to find more objective values of the assessments for the trainees, determine their level of preparation and the complexity of the tasks.

3. Ultimately, this allows you to more effectively manage the learning process.

References

[1] Novikov D. A. *Teoriya upravleniya organizatsionnymi sistemami.* Moskva: MPSI, 2005. – 584 S.

[2] Danilova C. D. *Adaptivnaya nechetkaya model otsenivaniya rezultatov avtomatizirovannogo testirovaniya s razdeleniem zadaniy po urovnyam usvoeniya* dis. kand. tekhn. nauk. – Ulan-Ude: VSHTU, 2005. – 122 S.

[3] Zaytseva L. V., Prokofyeva N. O. *Modeli I metody adaptivnogo kontrolya znaniy* Educational Technology & Society. – 2004. – 7(4). – S. 265 – 277.

[4] Solovov A. V. *Elektronnoe obucheniye: Problematika, Didaktika, Tekhnologiya.* Samara, Novaya tekhnika, 2006.

[5] Serbin V. I. *Metody otsenki znaniy v podsisteme treninga obuchayushchey sistemy* Vestnik Astrakhanskogo Gosudarstvennogo Tekhnicheskogo Universiteta, 1(36), 2007. – Astrakhan: Izd-vo AGTU, 2007. – S.247-251.

[6] Serbin V. I. *Opredelenie parametrov informatsionnykh protsessov v avtomatizirovannoy obuchayushchey sisteme.* Vestnik Astrakhanskogo Gosudarstvennogo Tekhnicheskogo Universiteta. Seriya Upravlenie, vychislitelnaya tekhnika i Informatika. - 2012. - T. 1(53). S. 57-61.

[7] Serbin V. I. *Metody raschota parametrov v avtomatizirovannoy obuchayushchey sisteme.* Prikaspiyskiy Zhurnal – Upravlenie I vysokie tekhnologii. - 2012. - iyun.- T. 2(18). - S. 66-70.

[8] Kudryashov B. D. *Teoriya informatsii. Kodirovanie diskretnykh istochnikov.* Sankt-Peterburg. PITER, 2009. – 315 S.

[9] Lazarsfeld P. *Latentno-strukturnyi analiz i teoriya testov. Matematicheskie metody v sotsialnykh naukakh.* Moskva: Progress, 1973.

[10] Laptev V. V., Serbin V. I. *Uchet vremeni pri otsenivanii rezultatov avtomatizirovannogo kontrolya.* Izvestiya Volgogradskogo Tekhnicheskogo Unversiteta. Seriya Aktualnye problemy upravleniya vychislitelnoy tekhniki i informatiki v tekhnicheskikh sistemakh. 11 2010.

[11] Kochenkov A. I. *Tolstova Yu.N. Idei latentno-strukturnogo analiza Lazarsfelda v sovremennoy sotsiologii.* Sotsiologiya: metodologiya, metogy, matematicheskie modeli. – 2003 - 16. – S. 127-149.

[12] Serbin V. I. *Metod rascheta parametrov avtomatizirovannoy obuchayushchey sistevy iz dvukh sostoyaniy.* V mire nauchnykh otrpytiy. – 2012. – E. 1(25). – S. 153-166.

[13] Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests.* Copenhagen, Denmark: Danish Institute for Educational Research, 1960.

[14] Khlebnikov V. A. *Teoreticheskie osnovy obyertivnogo izmereniya uchebnych dostizheniy uchashchikhsya.* Moskva: Federalnoe gosudarstvennoe uchrezhdenie Federalnyi tsentr testirovaniya, 2005. – 127 S.

[15] Neyman Yu. M, Khlebnikov V. A. *Vvedenie v teoriyu modelirovaniya i parametrizatsiyu pedagogicheskogikh testov.* Moskva: Prometey, 2000. – 169 s.

[16] Kharchenko M. A. *Korrelyatsionnyi analiz. Uchebnoye posobie dlya vuzov.* Voronezh: Izdatelstvo Voronezhskogo Gosudarstvennogo Universiteta. 2008.

[17] Serbin V. I. *Markovskaya model upravleniya protsessom obucheniya.* XII Vserossiyskoe soveshchanie po problemam upravleniya VSPU-2014. Moskva, 16-19 iyanya 2014g.: Trudy [Elektronnyi resurs] Moskva: Institut Problem Upravleniya im. V.A. Trapeznikova RAN, 2014, s. 9489-9497.