

## MODELING OF DISTRIBUTION FUNCTIONS IDENTIFYING STUDENTS IN THE PROCESS OF LEARNING

Romanov V.P., Shiryaeva N.A.

*National Research University "MIET", Moscow,*

*e-mail: romanov.valeri@gmail.com, shiryaeva.nataliya@gmail.com*

Non-classic probabilistic-statistical method was applied for student knowledge assessment. According to it each student is identified with a distribution function (probability density), which defines the probability to find him in an elementary unit of information space. The efficiency of implementation of two-level functions for the modeling of true distribution functions is shown. Graphing of two-level model distribution functions is based on the grades (expectations), which students received in testing on examination using a point grading system. It is found that measurement scale boundaries markedly effect the distribution functions. Specifically, when expectation value converges to the upper or the lower scale bound, probability density theoretically tends to infinity. However, taking account of an expectation value measurement error can bound probability density. Distribution functions allow (knowing the grade, that student received in testing) to calculate the probability to receive any other grade, f.e. students, who received 4 points and 3 points, could have received 5 points with probability 0,2 and 0,8 respectively.

**Keywords:** student, point grading method, grade, randomness, probabilistic-statistical method, distribution function, probability density, expectation

Student knowledge evaluation methods are based on the modeling views on what a student is and how he behaves in the process of gaining knowledge. Two essentially different approaches to the student behavior modeling in the process of gaining knowledge exist (from now on "the process of gaining information" is understood by the phrase "student behavior in the process of gaining knowledge"): first (traditional) approach is based on classical determinacy of student behavior [1]; second approach is based on non-classical determinism (determinism, realized through randomness) [2].

Traditional student model allows to locate student's position pointwise in the information space, which involves totality of the results of human semantic activity. It is the principle which forms the basis for the point grading method of student knowledge evaluation [3]. Individual knowledge is being evaluated with grades (points), i.e. numbers, and that is pointwise location of student's position in the information space. Certainly, knowledge evaluation has accuracy (involves error). For the 5-point grading system absolute error is 0,5 point, whereas relative error, divided by the end of scale, is  $0,5/5 = 0,1$ . To reduce relative error one can use 10-point or 100-point scale, which can reduce relative error to 0,05 and 0,005 points accordingly. Such reduction of relative error, of course, increase the accuracy of student knowledge evaluation. However, traditional model is restricted, since it puts out of account random character of student behavior in the process of gaining knowledge.

Probabilistic-statistical student model takes into account random factor in the process of gaining information. According to this model student is identified with a distribution function (probability density) in the informa-

tion space [4, 5]. This comes from the fact that individual knowledge, being a product of consciousness, is of random character, since the determinism of consciousness is realized through randomness, for it is derived from basically random behavior of psycho-somatic state. In this case non-classical probabilistic statistical method of research is used to describe individual behavior [5]. It should be noted, that traditional student behavior model could be obtained through passing to the limit within the framework of the probabilistic-statistical model. That way, if dispersion of a distribution function, which identify a student, tends to zero, distribution function tends to Dirac delta function (probability density tends to infinity), that defines location of an individual in the information space pointwise. Thus, it is arguable, that probabilistic-statistical model contains traditional student behavior model as its extreme event. The goal of this research is to model distribution functions, which identify students in the process of gaining knowledge.

### Materials and research methods

If the point grading system is used for student knowledge evaluation, it is assumed that the measurement range (the maximum scale value) corresponds to the volume of information, that a student can learn during particular course unit or during the whole course. In informatics it is common to measure information in "bits", whereas in pedagogics — in "points" (grades). One can always set necessary quantitative relation between "bits" and "points". The grade (points), which student receives as a result of knowledge assessment, is actually an expectation (average value), it defines student's location pointwise in the information

space, putting out of account random character of its knowledge.

To eliminate this shortcoming, one should turn from the point grading method of student's knowledge assessment to the probabilistic-statistical method, according to which, as it was mentioned earlier, within the information space a student is identified with the distribution function, that accounts for the random character of its knowledge. Such transfer is possible, since the maximum volume of information (the maximum point of a measurement scale), that a student can learn, and an expectation (a grade, which student receive as a result of knowledge assessment) are known for each type of activity. However, knowing it is not enough to graph a "true" distribution function. Hence, let us employ approximation method, which allows to construct so-termed model distribution functions. i.e. functions whose characteristics are close to characteristics of true distribution functions. As a model function we will take level function [6]. In zeroth approximation such function can be presented by one-step (rectangular) function, which corresponds to the uniform distribution of probability density within the scale. But such function can be used to model only symmetric true distribution functions, whose expectation, median and mode are the same and are equal to  $\sigma = 0,5\sigma_{\max}$ , where  $\sigma$  is coordinate. It can't be used to model asymmetric true distribution functions. In zeroth approximation two-step function can be used as a model distribution function to solve this problem.

Fig. 1 describes qualitatively a true distribution function and corresponding to it two-step model distribution function. To construct a model distribution function, it is essential to find the relation between median  $\sigma_{\text{med}}$ , coefficients  $a$  and  $b$  and expectation  $\langle \sigma \rangle$ .

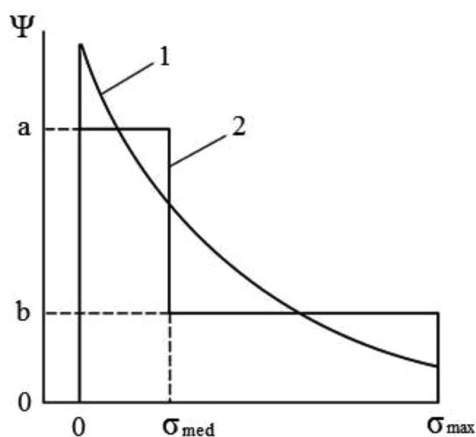


Fig. 1. Distribution function: 1 – true; 2 – model

To determine these coefficients we use an assumption, that probability to find a student within the whole information space is equal to 1 – equation (1), and also use characteristic of a median, which divides distribution function into the two equal parts with respect to the probability to find a student in the information space – equation (2):

$$a\sigma_{\text{med}} + b(\sigma_{\text{max}} - \sigma_{\text{med}}) = 1, \quad (1)$$

$$a\sigma_{\text{med}} = b(\sigma_{\text{max}} - \sigma_{\text{med}}). \quad (2)$$

By solving together (1) and (2), we will have:

$$a = 0,5 / \sigma_{\text{med}} \text{ and } b = 0,5 / (\sigma_{\text{max}} - \sigma_{\text{med}}). \quad (3)$$

Expectation will be calculated as an average of the coordinate, and taking account of (1) and (2) it will be as follows:

$$\langle \sigma \rangle = 0,5[a\sigma_{\text{med}}^2 + b(\sigma_{\text{max}}^2 - \sigma_{\text{med}}^2)]. \quad (4)$$

Due to the fact that, when we define student's knowledge completeness an expectation is being measured, let us determine the link between expectation and median by solving together (3) and (4):

$$\langle \sigma \rangle = 0,5\sigma_{\text{med}} + 0,25\sigma_{\text{max}}. \quad (5)$$

From the analysis of (3) and (5) it follows, that if  $\sigma_{\text{med}} = 0$  parameters  $a$ ,  $b$  and  $\langle \sigma \rangle$  take the following form:

$$a = 0,5\delta(\sigma - 0), \quad b = 0,5 / \sigma_{\text{max}},$$

$$\langle \sigma \rangle = 0,25\sigma_{\text{max}}.$$

where  $\delta(\sigma - 0)$  is Dirac delta function, equal to zero when  $\sigma \neq 0$  and equal to infinity when  $\sigma = 0$ .

Between the limits of zero and maximum scale value integral of delta function is equal to 1. The same is true for  $\sigma_{\text{med}} = \sigma_{\text{max}}$ :

$$a = 0,5 / \sigma_{\text{max}}, \quad b = 0,5\delta(\sigma - \sigma_{\text{max}}),$$

$$\langle \sigma \rangle = 0,75\sigma_{\text{max}}.$$

In practice the use of expression (5) is restricted by the measurement error of expectation and median. Suppose that median measurement error is  $\Delta\sigma$ , since the width of the left step of a model distribution function is the same as a median value ( $d = \sigma_{\text{med}}$ ), minimum width of the left step of a distribution function is  $d_{\text{min}} = \Delta\sigma$ , and its maximum width is  $d_{\text{max}} = (1 - d_{\text{min}})$ . Hence follow the restrictions on the possible values of median, which can be written as  $d_{\text{min}} \leq \sigma_{\text{med}} \leq (\sigma_{\text{max}} - d_{\text{min}})$ . Then, from this inequation and equation (5) it follows that

$$0,5(0,5\sigma_{\text{max}} + d_{\text{min}}) \leq \langle \sigma \rangle \leq 0,5(1,5\sigma_{\text{max}} - d_{\text{min}}).$$

Let us write down this inequation in dimensionless form

$$0,5(0,5 + d_{\min}^*) \leq \langle \sigma \rangle^* \leq (0,75 - 0,5d_{\min}^*), \quad (6)$$

where  $\langle \sigma \rangle^* = \langle \sigma \rangle / \sigma_{\max}$  and  $d_{\min}^* = d_{\min} / \sigma_{\max}$  is a dimensionless relative error and the minimum width of the left step of model distribution function respectively.

According to (5), median-expectation dependence in dimensionless form may be expressed as

$$\sigma_{\text{med}}^* = 2 \langle \sigma \rangle^* - 0,5, \quad (7)$$

where  $\sigma_{\text{med}}^* = \sigma_{\text{med}} / \sigma_{\max}$ .

Expression (7) is true within the range

$$d_{\min}^* \leq \sigma_{\text{med}}^* \leq (1 - d_{\min}^*). \quad (8)$$

If we use inequation (6) for calculation of distribution functions, coefficients  $a$  and  $b$  should be dimensionless. For this purpose, we will multiply these coefficients by  $\sigma_{\max}$ . Then according to (3) we have

$$a^* = 0,5 / \sigma_{\text{med}}^* \text{ and } b^* = 0,5 / (1 - \sigma_{\text{med}}^*), \quad (9)$$

where  $a^* = a \sigma_{\max}$  and  $b^* = b \sigma_{\max}$ .

Within the following range of expectation values

$$d_{\min}^* \leq \langle \sigma \rangle^* \leq 0,5(0,5 + d_{\min}^*). \quad (10)$$

Let us assume, that the width of the left step of distribution function is equal to constant ( $d_{\min}^* = \text{const}$ ) for all possible values of the expectation. Hence, when expectation changes, only heights of the left and the right steps can change. Such configuration of a model distribution function removes restrictions on the median value. After simple transformations, using equations (3) and (4), for  $a^*$ ,  $b^*$  and  $\sigma_{\text{med}}^*$  we will have

$$a^* = (1 + d_{\min}^* - 2 \langle \sigma \rangle^*) / d_{\min}^*,$$

$$b^* = (2 \langle \sigma \rangle^* - d_{\min}^*) / (1 - d_{\min}^*),$$

$$\sigma_{\text{med}}^* = 0,5d_{\min}^* + 0,5(1 - d_{\min}^*)b^* / a^*. \quad (11)$$

Proceeding similarly within the following range of the values of expectations

$$(0,75 - 0,5d_{\min}^*) \leq \langle \sigma \rangle^* \leq (1 - d_{\min}^*) \quad (12)$$

and assuming, that the width of the right step of model distribution function remains invariant ( $(1 - d_{\min}^*) = \text{const}$ ), we will get

$$a^* = (2 \langle \sigma \rangle^* - d_{\min}^*) / (1 - d_{\min}^*),$$

$$b^* = (1 + d_{\min}^* - 2 \langle \sigma \rangle^*) / d_{\min}^*,$$

$$\sigma_{\text{med}}^* = (1 - 0,5d_{\min}^*) - 0,5(1 - d_{\min}^*)a^* / b^*. \quad (13)$$

From the joint analysis of (11) and (13) it follows, that the expressions for  $a^*$  and  $b^*$  from (11) explicitly agree with the expressions for  $b^*$  and  $a^*$  from (13) (crosswise).

Let us visualize the dependence of  $a^*$ ,  $b^*$  and  $\sigma_{\text{med}}^*$  on  $\langle \sigma \rangle^*$ . To calculate the required parameters one should know the value of  $d_{\min}^*$ , which agrees with the relative error  $\delta\sigma$ . When evaluation of student's knowledge is being made with the use of the point grading scales the relative error, divided by the end of scale, usually lies within the limits of 5% to 10%. This corresponds to  $\delta\sigma$  values in the limits of 0,05 to 0,1. Then for definiteness in calculation of specified parameters we should use the relative error equal to  $\delta\sigma = 0,05$  ( $d_{\min}^* = \delta\sigma = 0,05$ ), it significantly simplifies the expressions, which relate these parameters with the expectation within all three tolerance regions of  $\langle \sigma \rangle^*$ . Let us write these relations in a final form.

**1st range.** By applying  $d_{\min}^* = 0,05$  in (10) and (11), we have

$$0,05 \leq \langle \sigma \rangle^* \leq 0,275, \quad a^* = 21 - 40 \langle \sigma \rangle^*,$$

$$b^* = (2 \langle \sigma \rangle^* - 0,05) / 0,95,$$

$$\sigma_{\text{med}}^* = 0,025 + 0,475 b^* / a^* \quad (14)$$

**2nd range.** By applying  $d_{\min}^* = 0,05$  in (6)–(9), we have

$$0,275 \leq \langle \sigma \rangle^* \leq 0,725, \quad a^* = 1 / (4 \langle \sigma \rangle^* - 1),$$

$$b^* = 1 / (3 - 4 \langle \sigma \rangle^*),$$

$$\sigma_{\text{med}}^* = 2 \langle \sigma \rangle^* - 0,5. \quad (15)$$

**3rd range.** By applying  $d_{\min}^* = 0,05$  in (12) and (13), we have

$$0,725 \leq \langle \sigma \rangle^* \leq 0,95,$$

$$a^* = (2 \langle \sigma \rangle^* - 0,05) / 0,95,$$

$$b^* = 21 - 40 \langle \sigma \rangle^*,$$

$$\sigma_{\text{med}}^* = 0,975 - 0,475 a^* / b^*. \quad (16)$$

Fig. 2 visualizes calculated dependencies of the dimensionless coefficients  $a^*$ ,  $b^*$  and  $\sigma_{\text{med}}^*$  from the values of dimensionless expectation.

It is apparent, that coefficients  $a^*$  and  $b^*$  are mirror symmetric with respect to the line, that goes through the point with coordinates (0,5; 0) and is parallel to the vertical axe, whereas the median is characterized by rotational symmetry. When the median is rotated by 180° with respect to the point with coordinates (0,5; 0,5) it corresponds to itself. Bound-

aries of the measurement scale answer for such remarkable characteristics of the concerned parameters.

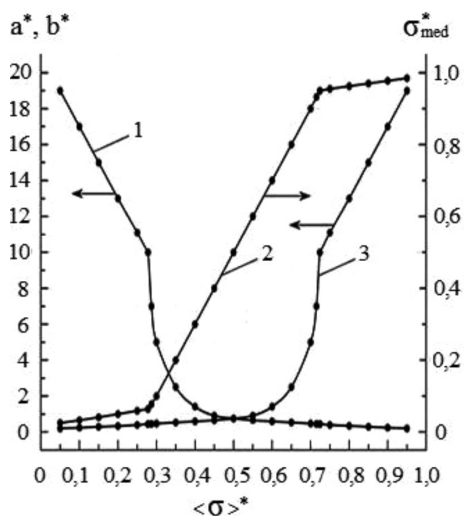


Fig. 2. Dependency of the dimensionless parameters of model distribution function from the values of dimensionless expectation: 1 – parameter  $a^*$ ; 2 – parameter  $b^*$ ; 3 – median  $\sigma_{med}^*$

**Research results and discussion**

Equations (14) – (16), and also graphics, represented in Fig. 2, could be used to calculate model distribution functions in the dimensionless coordinates. Fig. 3 shows the model distribution functions, which were calculated using equations (14) – (16) for the various values of dimensionless expectation.

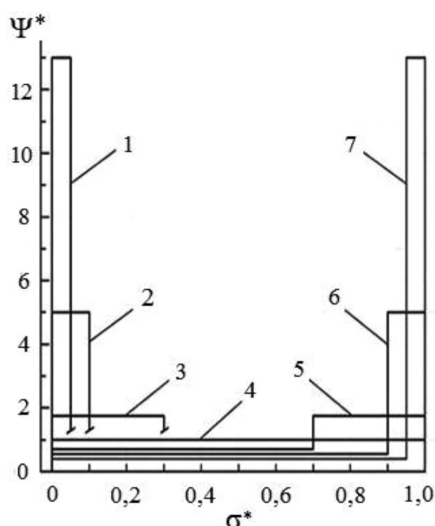


Fig. 3. Distribution functions in the dimensionless coordinates for the various values of dimensionless expectation: 1 –  $\langle \sigma \rangle^* = 0,2$ ; 2 –  $\langle \sigma \rangle^* = 0,3$ ; 3 –  $\langle \sigma \rangle^* = 0,4$ ; 4 –  $\langle \sigma \rangle^* = 0,5$ ; 5 –  $\langle \sigma \rangle^* = 0,6$ ; 6 –  $\langle \sigma \rangle^* = 0,7$ ; 7 –  $\langle \sigma \rangle^* = 0,8$

Distribution function  $\Psi^*(\sigma^*) = \Psi(\sigma)\sigma_{max}^*$  and coordinate  $\sigma^* = \sigma/\sigma_{max}$  are represented in the dimensionless form. Model distribution function  $\Psi^*(\sigma^*)$  is actually a universal function. It has the same form in all point grading measurement systems, it allows to study the behavior aspects of such function and to translate them next onto the model distribution functions in any point grading measurement system, f.e. in 20-point or 100-point grading system. It is achieved by changing the scale factor of the coordinate axes, i.e., multiplying all the numbers on the Y-axis by the maximum value of the selected scale ( $\sigma = \sigma^* \cdot \sigma_{max}$ ) and dividing all the numbers on X-axis by the maximum value of the selected scale ( $\Psi(\sigma) = \Psi^*(\sigma^*)/\sigma_{max}$ ).

As an example, Fig. 4 shows model distribution functions, identifying students, who have got 3, 4 or 5 points on examination.

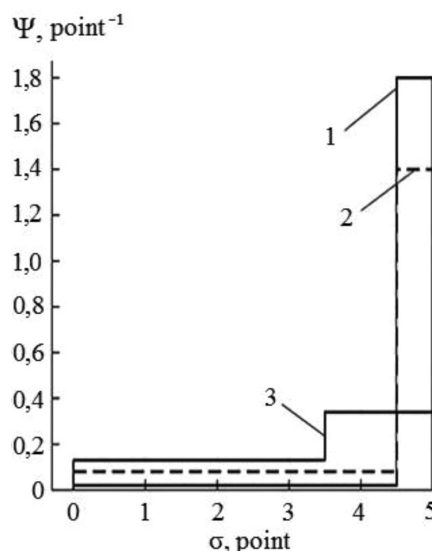


Fig. 4. Distribution function, identifying students, who have got the grades: 1 – 5 points; 2 – 4 points (dashed line); 3 – 3 points

Analysis of data, represented in Fig. 4, shows, that distribution functions 1, 2 and 3 overlaps within the range of 4,5 to 5 points. Since the error of 0,5 points, common to the 5-point measurement system, defines the lower limit for the expectation values obtained during the knowledge evaluation, the range of 4,5 to 5 points corresponds to “5” (the result of examination). Then, integration of the distribution function (probability density) over the coordinate in the range of 4,5 to 5 points specifies the probability, that the student will get “5”. In our case the student, who received “5”, received this grade due to the fact, that the probability to learn the educational mate-

rial amounts to 0,9. Students, who received “4” and “3”, could have got “5” with probability 0,8 and 0,2 respectively.

### Conclusions

1. In place of model distribution functions, which approximate empirical distribution functions, identifying students in the process of learning, it is convenient to use two-step functions, that can be calculated on the base of the information about the grades of individuals, which they received after testing using the point grading measurement system.

2. When expectation converges to the boundaries of the measurement scale, probability density of distribution functions tends to infinity. However, probability density could be bounded, if expectation value measurement error is taken into account.

3. Model distribution functions, calculated in the dimensionless coordinates, have the same form in all the point grading measurement systems. This allows to study the behavior aspects

of such functions and then to translate them to the model distribution functions in any point grading measurement system.

4. Model distribution functions allow, knowing the grade, that student received in testing, to calculate the probability to get any other grade.

### References

1. Mikheev V.I. Modeling and methods of measurement theory in pedagogy. M.: KRASAND, 2010. 224 p.
2. Romanov V.P., Shiryayeva N.A. Role of random factor in the process of student's knowledge completeness measurement // Modern high technologies. 2017. № 6. P. 157–163.
3. Podlasij I.P. Pedagogy: Theory and technology of education. M.: VLADOS, 2007. 575 p.
4. Sokolova N.A. Experimental research of students behavior in the process of learning by probabilistic-statistical method // Modern problems of science and education. 2013. № 4. URL: <http://www.science-education.ru/ru/article/view?id=9729> (date of access: 25.05.2020).
5. Romanov V.P., Shiryayeva N.A. Probabilistic-statistical method of scientific research: Application in psychological pedagogy. M.: De'Libri, 2018. 136 p.
6. Korn G.A., Korn T.M. Mathematical handbook for scientists and engineers. St. Petersburg: Lan, 2003. 832 p.