STATISTICAL ANALYSIS OF THE PROFESSIONAL COMPETENCE STUDY ON THE EXAMPLE OF THE FIELD "APPLIED INFORMATION SCIENCE IN ECONOMICS"

Irina Ilinichna Eremina*, Aigul Ginatullovna Faizullina

Kazan Federal University, Naberezhnye Chelny, Republic of Tatarstan, Russian Federation,
Candidate of Pedagogical Sciences, Associate Professor.

Kazan Federal University, Naberezhnye Chelny, Republic of Tatarstan, Russian Federation, Senior Lecturer.

Email: ereminaii@yandex.ru

Received on 13-07-2016
Accepted on 10-08-2016

Abstract:

The article analyzes the results of forming the professional competence of future IT professionals of the Applied Information Science in Economics at the Naberezhnye Chelny Institute of the Kazan Federal University (NCHI KFU). The mathematical modeling methods, main provisions of mathematical statistics, optimization methods, apparatus of multidimensional scaling and qualimetric estimation, theories of high school education and bringing-up process management were selected as the methods of the study. A comprehensive simulation model for forming the professional competence of future IT professionals of Applied Information Science in Economics in view of the interaction of the main technological processes based on the IDEF0 technology (function modeling (IDEF=ICAM Definition), ICAM= Integrated Computer Aided Manufacturing) and integrated estimation of functioning efficiency of the professional competence formation system for future IT professionals of Applied Information Science in Economics characterized by the possibility of estimating on several indicators of adequacy and sufficiency of the analyzed competence are represented. In analyzing the data new results of scientific and practical importance in the field of mathematical modeling and education were obtained - mathematical diagnostics models for professional competence formation results.

Key words: Competence, Professional competence, Statistical methods, Level of professional competence formation, qualimetry, Assessment, Mathematical modeling methods, Modeling and Time series prediction methods, adaptive prediction techniques.

1. Introduction

In the market economy the current requirements of employers and society significantly increase to the IT industry professionalism level, but the actual level of their professional competence does not always meet these requirements.
In this connection search for new ways is relevant to improve the effectiveness of the institutions implementing the higher education programs aimed at enhancing the professional competence quality of future IT professionals. Scientific justification of the qualimetric approach of the HE system in the competence-based approach is one of these ways. Relevance of the problem is determined by insufficient development of the methods for assessing the professional competence formation level of IT professionals of Applied Information Science in Economics. Purpose of this study is a scientific justification of the analysis of professional competence formation results using statistical methods on the example of the field "Applied Information Science in Economics" at the Naberezhnye Chelny Institute of the Kazan Federal University (NCHI KFU).

2. Methods

Considering the aspects of qualimetric estimation of the professional competence formation results of future IT professionals of the Applied Information Science in Economics, it should be noted that the main problems are associated with formation of the information base of studies and lack of its processing methods. Solution to these problems can be found in the mathematical modeling theory, use of the apparatus of multidimensional scaling and qualimetric estimation, as well as the main provisions of mathematical statistics.

So, analytics with such tools as the Jonckheere's S criterion of tendencies, Wilcoxon T test, Pearson's chi-squared test ($\chi^2$), Fischer angular transformation (Fisher's criteria $\varphi$) and many other statistical methods allows to estimate the current professional competence formation level of IT professionals of the Applied Information Science in Economics.

Solution to one of the problems of the study is to design a comprehensive simulation model for forming the professional competence of future IT professionals of the Applied Informatics in Economics in view of the interaction of the main technological processes based on the IDEF0 technology (function modeling (IDEF=ICAM Definition), ICAM= Integrated Computer Aided Manufacturing). In order to solve the problem of the study, a comprehensive simulation model based on the IDEF0 standard will be considered. Applying a simulation model to show forming the professional competence of future IT professionals of the Applied Information Science in Economics, graphical results will be obtained. Interpretation of the results of modeling aims to shift from the information obtained as a result of computer simulation with a model to the conclusions relating to the process of forming an original object. The presented elementary block represents a functional element of the system, which converts the input data stream to the output one when there is control and a certain mechanism is used.
The elementary block represents a functional element of the system, which converts the input data stream to the output one when there is control and a certain mechanism is used. For the proposed model, the process of forming the professional competence of future IT professionals will be specified as the “activity”. The initial level of forming the professional competence of a future IT professional will be specified as the “input data stream” (Input), which means, in the industrial sector, any material (raw materials) processed or converted during such activity. Accordingly, the output data stream (Output - what the activity results in) means the required formation level of graduate’s professional competence. The mechanism (Mechanism) performing the activity means a teacher, and the education institution’s educational and material resources will be defined as the resources needed to perform the activity properly. The call arrow (Call) is designed to refer to another model of activity. The contextual diagram of the module of forming the professional competence of future IT professionals of the Applied Informatics in Economics is shown in Figure 2.
The comprehensive simulation model for forming the professional competence of future IT professionals of Applied Informatics in Economics represents a set of factors:

The input data stream means enrollee’s points in the computer science. The output data stream means a level of forming the professional competence in the field of the Applied Information Science in Economics. Based on the above IDEF0 standard methodology, one of the "mechanisms" used to form the required level of professional competence is teaching staff, which is the "production staff" (in technical terms) in the higher professional education system. The model resources means the high school’s educational and material resources, which represents a complex of material and technical means that provide training of students for the established professions and specializations in accordance with the study programs, programs of educational discipline and modern teaching method. The call of another module means an activity to be performed outside the modelled system.

The "Competence Map" in the field of the Applied Information Science in Economics and students' grades per semester grouped in individual professional competence in accordance with the study program of the field Applied Information Science in Economics serves as a form of presentation of the initial information in this work.
The Competence Map is one of the options for creating a system of qualitative estimation of professional competence. Layout of the table "Competence Map" (Table 1) is shown below.

**Table 1. Competence Map.**

<table>
<thead>
<tr>
<th>Циклы (разделы), дисциплины (модули) учебного плана ООП ВПО</th>
<th>Cycles (sections), disciplines (modules) of the study program within the principal educational program of the higher professional education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Индекс компетенции</td>
<td>Competence index</td>
</tr>
<tr>
<td>Б1 ГСО</td>
<td>Б1 General humanitarian and socio-economic disciplines</td>
</tr>
<tr>
<td>Б2 Математика и естественные науки</td>
<td>Б2 Math and science</td>
</tr>
<tr>
<td>Б3 Профессиональный</td>
<td>Б3 Professional</td>
</tr>
<tr>
<td>Б1.1 Базовая часть</td>
<td>Б1.1 Basic part</td>
</tr>
<tr>
<td>Б1.2 Вариативная часть</td>
<td>Б1.2 Variable-based part</td>
</tr>
<tr>
<td>Б2.1 Базовая часть</td>
<td>Б2.1 Basic part</td>
</tr>
<tr>
<td>Б2.2 Вариативная часть</td>
<td>Б2.2 Variable-based part</td>
</tr>
<tr>
<td>Б3.1 Базовая часть</td>
<td>Б3.1 Basic part</td>
</tr>
<tr>
<td>Б3.2 Вариативная часть</td>
<td>Б3.2 Variable-based part</td>
</tr>
<tr>
<td>Дисциплины</td>
<td>Disciplines</td>
</tr>
<tr>
<td>Модули</td>
<td>Modules</td>
</tr>
<tr>
<td>Б4 Физкультура</td>
<td>Б4 Physical education</td>
</tr>
<tr>
<td>Б5.1 Учебная</td>
<td>Б5.1 Academic</td>
</tr>
</tbody>
</table>
The information base is students’ grades per semester grouped in individual professional competence in accordance with the study program of the field Applied Information Science in Economics. The main base, which will be used for further analysis of the results of forming the professional competence of IT professionals of the Applied Information Science in Economics, is a table, which is prepared by using a comprehensive quality estimation method.

The comprehensive index of graduate’s competence is defined as a set of competence formation indicators in disciplines.

The main indicators should be specified:

- Indicator $S_{ij} = P_{ij} \cdot V_{ij}$ means an indicator of the competence formation level in the j-th discipline of the i-th cycle;
- $P_{ij}$ means a final grade received per semester in a discipline. It is determined when a graduate is tested on the j-th discipline $P_{ij}$ of the i-th block (0-100%);
- Depending on the number of hours allocated to the study in the Federal State Educational Standards of Higher Education, the i-th block is assigned with a certain weighting factor $V_i$ (0-1.0) and j-th discipline of the i-th cycle – with $V_{ij}$ (0-1.0);

The model that will be considered was developed based on the integral, activity, competency building and person-centered approaches. The problem associated with analysis of the results of forming the professional competence of future IT professionals is connected with measurement of educational outcomes, definition of qualitative indicators and correlation with the requirements of the federal state educational standards and study programs.

The data obtained in analysis of the results of forming the professional competence are often referred to as "nonstrict" implying the ambiguity of interpretation of the obtained results. This is not quite true. Those results prove to be nonstrict that failed to have any sufficient justification and adequate statistical analysis. Proper and correct application of the statistical methods allows to draw unambiguous analysis conclusions regarding formation of the professional competence of IT professionals. The statistical methods are important in analyzing the results of formation of the professional competence.
The main difficulty in appealing to the statistical methods is not mathematical; the principle of many of them is quite clear and obvious. The problem is to process large blocks of information (usually numeric), and it is absolutely impossible to do it "manually". The usage of computers solves the problem dramatically. Given that it is possible to opt for the use of specialized packages for processing certain statistical information (e.g. Stadia, SPSS, Statgraphics, Statistica and other).

In order to analyze the results of forming the professional competence, students' grades per semester grouped in individual professional competence in accordance with the study program of the Applied Information Science in Economics were used as corresponding data.

Based on these characteristics, the mathematical methods and research tools were determined. They include the Jonckheere's S criterion of tendencies, Wilcoxon T test, Pearson's chi-squared test ($\chi^2$), Fischer angular transformation (Fisher's criteria $\phi$); Spearman's rank correlation coefficient $r_s$.

### 3. Results

One of the examples – the Spearman's rank correlation coefficient $r_s$ – will be considered in detail below. The Spearman's rank correlation makes it possible to determine the intensity, direction and significance of correlation relationship between two features or two profiles (hierarchies) of the features.

The hypotheses to be tested:

$H_0$: There is no significant correlation between the compared features (hierarchies).

$H_1$: Correlation between the compared features (hierarchies) is definitely different from zero.

Application of the method starts with a ranking in each of the sets of features - as a result transition is made from the primary data to two sets of ranks: $\{R1\}$ and $\{R2\}$. Then the rank correlation coefficient $r_s$ is calculated according to the formula:

$$r_s = 1 - \frac{6 \cdot \sum (R_{1i} - R_{2i})^2}{n(n^2 - 1)}$$

Furthermore, a critical value ($r_s)_{cr}$ (significance level $p \leq 0.05$) is determined for the available sample volume $n$; if $(r_s)_{exp} \geq (r_s)_{cr}$, an experimental hypothesis is accepted.

Restrictions on applicability of the Spearman's correlation:

1) volume of the sample, in which the features are correlated, should not be less than 5; upper limit - 40 test subjects
2) rank correlation coefficient is calculated by the formula (*), only if there are no matching values (or small number thereof) in each of the sets of ranks; if this condition is not satisfied, it is necessary to make a correction for the identical ranks - in this case \( r_s \) is calculated as follows:

- correction for identical ranks is calculated in the aggregate \( R_1 \} \) and \{ \( R_2. \):

\[
T_i = \frac{k_i}{12} \sum m_{1j} (m_{1j}^2 - 1), \quad T_i = \frac{k_i}{12} \sum m_{2j} (m_{2j}^2 - 1)
\]

where \( k_1 \) and \( k_2 \) means a number of groups of identical ranks in sets of ranks 1 and 2, respectively;

\( m_{1j} \) and \( m_{2j} \) means volumes of each of the groups of identical rank in sets 1 and 2;

- \( r_s \) is calculated with correction:

\[
r_s = 1 - \frac{6 \sum (R_{1i} - R_{2i})^2 + T_1 + T_2}{n(n^2 - 1)}
\]

Consider an example - comparison of two sets of indicators in a group. There are grades of the state examination and defense of the degree project for one group of students. Is there any significant correlation between these indicators?

The hypotheses to be tested:

\( H_0 \): There is no significant correlation between the indicators.  
\( H_1 \): There is significant correlation between the indicators of the state examination and defense of the degree project.

**Table 2. Spearman’s rank correlation coefficient \( r_s \)**

<table>
<thead>
<tr>
<th>No.</th>
<th>Student</th>
<th>State examination</th>
<th>Rank of state examination</th>
<th>Defense of degree project</th>
<th>Defense rank of degree project</th>
<th>((R_{1i}-R_{2i})^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Student 1</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>Student 2</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Student 3</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>2.5</td>
<td>20.25</td>
</tr>
<tr>
<td>4</td>
<td>Student 4</td>
<td>5</td>
<td>13</td>
<td>5</td>
<td>13.5</td>
<td>0.25</td>
</tr>
<tr>
<td>5</td>
<td>Student 5</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Student 6</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2.5</td>
<td>0.25</td>
</tr>
<tr>
<td>7</td>
<td>Student 7</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>13.5</td>
<td>42.25</td>
</tr>
</tbody>
</table>
Table 3. Set of identical rank groups (state examination).

<table>
<thead>
<tr>
<th>Student</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 8</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Student 9</td>
<td>5</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Student 10</td>
<td>5</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Student 11</td>
<td>4</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Student 12</td>
<td>5</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Student 13</td>
<td>4</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Student 14</td>
<td>4</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Student 15</td>
<td>5</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>120</td>
<td>120</td>
<td>174</td>
</tr>
</tbody>
</table>

Table 4. Set of identical rank groups (defense of degree project).

<table>
<thead>
<tr>
<th>Rj</th>
<th>mj</th>
<th>m²-m</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>336</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>120</td>
</tr>
<tr>
<td>T1</td>
<td>40</td>
<td></td>
</tr>
</tbody>
</table>

After calculating the squared differences of the ranks for each student, sum of the column $(R_{11}-R_{21})^2$ was obtained. Further groups of identical ranks were determined in each set – these are groups with ranks 2, 7, 13 in the first set and numbers of values 3, 7, 5 in the groups. In the second set - groups with ranks 2.5, 8, 13.5 and numbers 4, 7, 4 according to the formulas (**). According to the formula (**), value of the Spearman's rank correlation coefficient $(r_s)_{exp}=0.67$ for $n=15$ $(r_s)_{cr}=0.52$ should be calculated. Therefore, $(r_s)_{exp}>(r_s)_{cr}$, an experimental hypothesis is accepted - there is a significant correlation based on the Chaddock's scale between indicators of the state examination and defense of the degree project.

### 4. Conclusions

As shown above, the intensity of the correlation relationship is close to a boundary between the gradations
Irina Ilinichna Eremina* et al. International Journal of Pharmacy & Technology

"remarkable" and "high", i.e. a hypothesis about the existence of correspondence between the considered parameters should be declared fair without any doubt.

Furthermore, the STATISTICA 6.0 package possibilities should be investigated for analysis of forming the professional competence results of IT professionals of the Applied Information Science in Economics.

For visualizing the analysis of forming the professional competence results of IT professionals of the Applied Information Science in Economics a line graph should be constructed. The graph shows what professional competence has high/low level of formation. ПК-17 (PC-17) has a high value, and ПК-21 (PC-21) has the lowest indicator.

![Line Plot (Book 1.xls21v*76c)](image)

**Figure 5. Professional competence formation level of IT professionals of the Applied Information Science in Economics.**

Smoothing of the level range is described below. One of the common methods of time-series trend analysis is the use of moving averages. Figure 6 depicts the original series of professional competences. Further smoothing of time series should be set forth. Due to the fact that a 4-year apprenticeship term is considered for future IT professionals of the Applied Information Science in Economics, it would be logical to choose a smoothing value equal to 4. The initial and smoothed levels should be compared, the result is shown in Figure 7. It results in a large gap between the levels.

In order to construct a linear trend, at least two methods can be used in the STATISTICA package: graphical method - by means of an option for constructing time-series graphs; analytical alignment - where the module means are used.
The first method will be applied to construct a linear trend in the package. It will result in a graph of formed professional competences and corresponding linear trend. According to the data given in the graph of Figure 6, a linear trend equation is displayed at the top part of the graph (Y=0.7070-0.0048x). But this method does not represent any information about the model’s statistical significance, and it is useful only in the exploratory analysis. The second method - construction of a trend by analytic alignment – should be considered. Before proceeding to construction of the linear trend, it is necessary to make a remark that t can be set in two ways, from the beginning and from the center of the series. New (time) variables $t_1$ and $t_2$ should be formed for the original series. A trend in the form of a first-degree polynomial (straight line) should be constructed. The PC formation level should be selected as a dependent variable, and the moment of time from the beginning of the series $t_1$ serves as an independent variable.

A trend in the form of a first-degree polynomial (straight line) will be constructed. The PC formation level was selected as a dependent variable, and the moment of time from the beginning of the series $t_1$ was selected as an independent variable. The trend results in the form of a polynomial contain the estimated model parameters and main indicators of the regression construction adequacy.
The package formed a table containing estimation model parameters and main indicators of the regression construction adequacy.

<table>
<thead>
<tr>
<th>N=21</th>
<th>Beta</th>
<th>Std. Err. of Beta</th>
<th>B</th>
<th>Std. Err. of B</th>
<th>t(19)</th>
<th>p-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.707701</td>
<td>0.223152</td>
<td>3.171307</td>
<td>0.005036</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>-0.061699</td>
<td>0.226990</td>
<td>-0.284792</td>
<td>0.017772</td>
<td>-0.269060</td>
<td>0.739079</td>
</tr>
</tbody>
</table>

Figure 10. Estimation results of regression dynamic model with a dummy variable t1

The second column shows \( \beta \)-coefficients, and the third column compiles an average error of \( \beta \)-coefficients. The fourth column contains the regression equation parameters, and the fifth one includes an average error of the equation parameters.

Now detection and removal of autocorrelation should be started. The graphical method for detecting autocorrelation will be used.

According to the data presented in the graph, no deviation trend is observed, respectively, the absence of autocorrelation can be assumed.

The Durbin-Watson Test will be used.

This is due to the fact that the correlation coefficient is within the range

\[ 0 \leq |\rho_{t,t-1}| \leq 1. \]
According to the table of critical values of these statistics for n=21 and k=2, the following parameters are valid: lower limit $d_l = 1.13$ and upper limit $d_u = 1.54$; hence it appears that the actual value falls within $2 < \text{DW} < 4 - d_l$.

The Breush–Godfrey test will be applied. Results of the trend estimation $\hat{y}_t = a_0 + a_1 t^2$ should be used, and values of the disturbance term $\varepsilon_t$ should be obtained. The calculation automatically results in the STATISTICA 6.0 package. Due to the fact that the test implementation requires estimation of the equation $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$, a new variable $\varepsilon_{ut}$ should be generated. Furthermore, a shift of one level forward should be made. Then using the Multiple Regression module, the model estimation results $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$ will be obtained.

According to the model construction results $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$, the received parameter $\rho$ is not statistically significant, therefore, the absence of correlations can be asserted.

![Figure 14. Model estimation results $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$](image)

5. Summary

Testing the qualimetric estimation effectiveness of the professional competence formation results of IT professionals of the Applied Information Science in Economics by comparing the experimental group’s test results and grades obtained by students at the end of studying a discipline showed that the indicator of professional competence is different from the average value of the student’s rating points. Moreover, it demonstrates residual (rather than current) knowledge possessed by a graduate at the time of graduation. The estimation tools were standardized, and the estimation technology’s reliability and validity were defined. Experimental work on the implementation of the qualimetric professional competence estimation model and technology allowed to adjust the training content on forming the professional competence of the IT professionals of the Applied Information Science in Economics, to increase the student’s motivation and to provide an objective, valid and valuable level representation of the educational activity results connected with the professional training of students and graduates. Statistical methods of the study data analysis proved effectiveness of the results.
Acknowledgements

The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

6. References


Irina Ilinichna Eremina* et al. International Journal of Pharmacy & Technology


Corresponding Author:

Irina Ilinichna Eremina*,