TECHNICAL UNIVERSITY OF CLUJ-NAPOCA

ACTA TECHNICA NAPOCENSIS

Series: Applied Mathematics and Mechanics
Vol. 54, Issue IV, 2011

ENTRY PREDICTION IN THE TECHNICAL UNIVERSITY OF CLUJ-NAPOCA

Lia CHIOREAN

***Abstract*:** *The present paper approaches an important component of university management, which is represented by university admission. Starting from the number of places provided for the entrance examination in the Technical University of Cluj-Napoca and the extent to which they were occupied in the interval 2005-2010, prognoses regarding the number of candidates for the year 2011 were developed, using specific methods belonging to the theory of probability and mathematical statistics. The study refers only to the bachelor level admission for the entire university. The results of the prognoses were validated by comparison with the data collected for the entry in September 2011.*
Key words: *entry, prediction, statistics, regression, candidate*

1. INTRODUCTION

The present paper aims at making use of methods found in mathematical statistics to make a prediction of the number of candidates to take part in the entrance examination in the Technical University of Cluj-Napoca. The analysis of the data regarding the interval of years 2005-2010 with respect to the number of entry candidates and the number of places provides through comparative studies, syntheses and regression functions other data of interest.

The results of the prognoses were compared to the actual data collected after the 2011 entry to demonstrate the method validity as well as its advantages and disadvantages.

The conclusions of the study can bring forth solid information that could be used in defining further strategies developed by the management of the university with respect to planning future number of places as well as education marketing activities.

2. MATERIALS AND METHODS

Starting from the premise that" the objective of the strategic management in the universities lies in reaching competitive advantage in a given competitive context" [1], the university has to pay much attention to attracting students. Shattock states that: "the best guarantee for the financial security may be an academically successful university which attracts large numbers of students, educates them well and has good results in"[2].

John Naisbitt [3] has the opinion that "people survive only through his capacity to act now, in the present moment, based on former experience and thinking of potential future consequences. Assuming future, people make the present bearable and give significance to the past. Past, present and future alternatives interconnect in anticipating and predicting would-be actions."

This paper presents a number of analyses and comparative evaluations regarding the number of candidates enrolled for admission in the period under investigation and the number of places for entrance examination, as well as the extent to which these places were occupied both for budgeted places and places with tuition fees.

Fort the study of the territorial distribution of the candidates, the authors used the Pareto analysis, a classification statistical method also

known under the name of rule 80/20. This rule states that 20% of the resources generate 80% of the most important results. The rule also mentions the great lack of balance between inputs and outputs, i.e. between cause and result. The Pareto diagram is a tool that helps distinguish significant issues from possible issues that ranks information function of priority and importance and provides actual information that can be used in decision making while clearly and suggestively indicating the direction of action.

Prediction as support of the decision-making process represents a process through which the development of specific future events is anticipated and indicator evolution can be estimated. Prediction has to be but at the basis of finding out the needs, potential means and further development of the university.

The university strategy needs to be defined capitalising the data used from the prognoses related to domains of interest and data obtained from own system along an interval of time that should be as long as possible.

Prognoses are based on specific studies related to the theory of probabilities and mathematical statistics. The present study made use of the regression functions and probabilistic calculations.

The regression function is a mathematic expression deduced after applying experimental data that approximate the dependence between two or more variables. In this case, the variable is represented by the number of candidates for admission in every academic year. Given a set of known data for a number *n* of years, simple mathematical expressions are looked for as well as a related graphical representation, trying to find the curve passing nearest to the defined points. The expression deduced in this manner is called regression function. In cases where the relationship depends upon only one variable, the shape of the regression function can be determined with the method of the least squares.

Using the dedicated software "CurveExpert", for the case of a single variable, we looked for the regression function that best approximates the set of experimental data, so that the values calculated for the prediction year shall be as close as possible to reality. The software specialised in curve processing gives the possibility of selecting the function giving the best approximation from the functions obtained. It can use a variety of function types covering the majority of mathematical models found in practice and applied in the Windows operation environment. This software was developed by Daniel Hyams [4].

The approximation quality is given by the *estimation standard error* and the *correlation coefficient*.

The estimation standard error is determined with the relationship:

$$S = \sqrt{\frac{\sum_{i=1}^{n}[y_i - f(x_i)]^2}{n - n_c}} \qquad (1)$$

where:

n - the number of experimental points;

x_i, y_i – the coordinates of the experimental points;

f – the shape of the regression function under study;

n_c - the number of function coefficients.

The value of the correlation coefficient yields from the relationship:

$$R = \sqrt{\frac{S_t - S_r}{S_t}} \qquad (2)$$

where:

$$S_t = \sum_{i=1}^{n}(y - y_i)^2 \qquad (3)$$

y being the arithmetical average of the ordinates of the experimental points y_i, and S_r the fraction numerator from relationship (1).

The standard error (S) defines the spreading of the experimental points around the regression function line. The more the points come closer to the diagram, the more the standard error reaches zero as value.

The correlation coefficient (R) represents a more accurate measure of the regression function quality, found when the standard deviation is relatively large. The more its value nears 1, the more the regression function is seen to better approximate the set of experimental points.

After introducing the experimental data, the "CurveExpert" program displays the graph of the regression function and the user has more options to choose from:

- to visualise the mathematical expression, the coefficients of the regression function and intermediate results;
- to make calculations using the determined relationship;
- to save the coordinates of the calculated points from the line of the regression function.

The user can request the performance of the calculations for all the types known to the software and then, can select the line and the numerical data corresponding to the studied functions to be visualised. In case none of the regression function types known to the software fits, the user has the possibility of defining the own type of function.

After fin ding with the regression method the estimate value for the current academic year, a correlation coefficient was applied by making the ratio between the potential candidates from the current academic year and the preceding academic year, that is value v.

Going on with the extrapolation of the dynamic series with value (v) previously found and using the student distribution, we calculated the size of the prediction field, its minimum and maximum values respectively:

$$\max = v + t_q \cdot \frac{\sigma}{\sqrt{n}} \; ; \quad \min = v - t_q \cdot \frac{\sigma}{\sqrt{n}}, \quad (4)$$

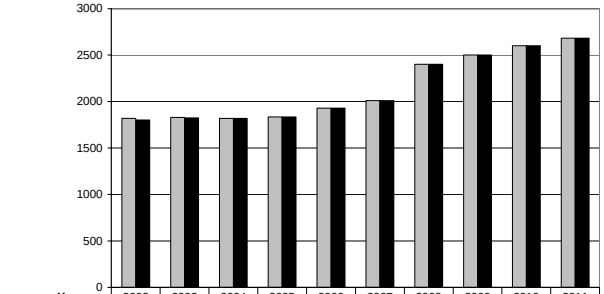
where:

- tq is the corresponding value from the table regading student distribution for=0,05 and a number of degrees of freedom of f=n-1,
- n is the number of experimental points;
- $\sigma = \sqrt{\frac{(dif_1 - \bar{u})^2 + ... + (dif_n - \bar{u})^2}{n-1}}$; (5)
- $\bar{u} = \frac{\sum_{i=1}^{n} dif_i}{n}$; (6)
- dif_i is the difference between the experimental value and the calculated one, by means of the regression function at point i.

The Student distribution is normally used to check the statistical hypotheses based on results obtained from small volume samples, $n \leq 30$, or when the approximation with one single normal distribution does not fit.

3. RESULTS AND DISCUSSION

The analysis made in the Technical University of Cluj-Napoca with respect to the degree of enrolling students after the admission, highlights a special performance of the university concerning the places financed by the budget. In Figure 1, it is evident that the degree of enrolment is 100% all along the period of time under investigation, with a very insignificant difference for the years 2002 and 2003.



Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Number of places	1820	1830	1820	1835	1930	2010	2400	2500	2600	2683
Admitted	1802	1824	1820	1835	1930	2010	2400	2500	2600	2683

Fig.1. Degree of budget places enrolment in the period 2002-2011

As for the tuition fee places, they were occupied to various extents. The analysis demonstrates that in the years 2003, 2007 and 2008 most places with tuition fee were occupied, while, the minimum number of tuition fee places is found in 2010, with a percentage of 11.87% (Figure 2).

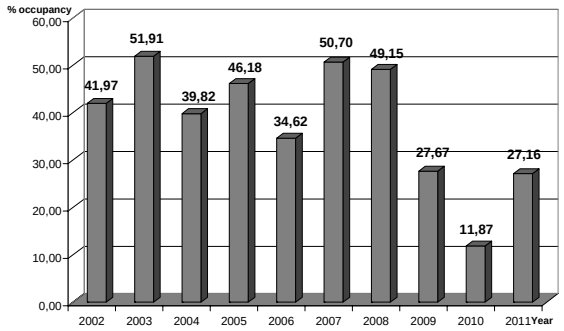


Fig.2. Degree of tuition fee enrolment in the period 2002-2011

Interesting to notice the distribution as per counties; the admission candidates come from all the 42 counties in the country, in various percentages, most options originating in the counties that neighbour Cluj County.

As there is a number of significant counties providing a large number of admission candidates, we tried to use a Pareto diagram for this purpose (ABC). In Figure 3, the diagram for year 2010 is given, in this respect. A similar distribution was found for the rest of the years in question.

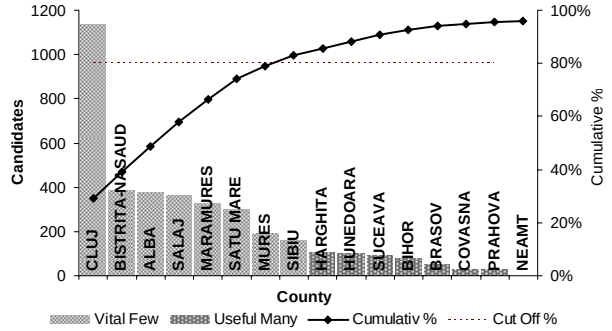


Fig.3. Diagram ABC for the admission candidates in 2010/counties of origin

Analysing the Pareto diagram plotted for year 2010 (fig.3), one can notice that the majority of the admission candidates (82,93%) come form eight candidates: Cluj, Bistrița-Năsăud, Alba, Sălaj, Maramureş, Satu-Mare, Mureş and Sibiu, representing 21,05% of the candidates' county of origin.

Except for the county of Cluj, with a normal first position, counties Bistrița-Năsăud, Sălaj, Satu-Mare and Alba are also included here, as extensions of the Technical University of Cluj-Napoca can be found there. In this way, it is demonstrated that the strategy of establishing extensions in these places contributed to enrolling many students from these areas.

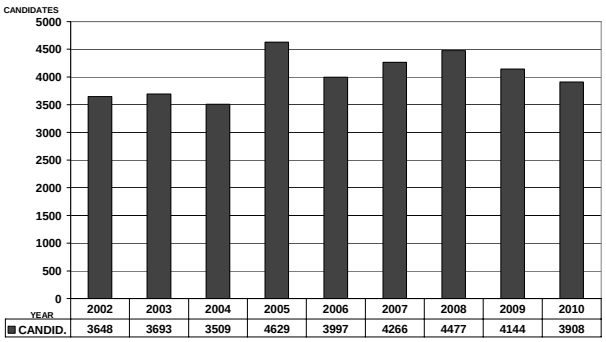


Fig.4. Evolution of admission candidates' number for the Technical University of Cluj-Napoca

With the statistical data regarding the number of candidates in the admission in the Technical university of Cluj-Napoca in the previous years (Figure 4) and the CurveExpert software, we looked for estimate results for the tendency of becoming candidates for the admission in our university, taking into account both an optimistic and a pessimistic scenario.

On the basis of these data, the CurveExpert software led to a regression function, of Gauss type $y = ae^{\frac{-(x-b)^2}{2c^2}}$, which graphically represented most faithfully the given points (Figure 5).

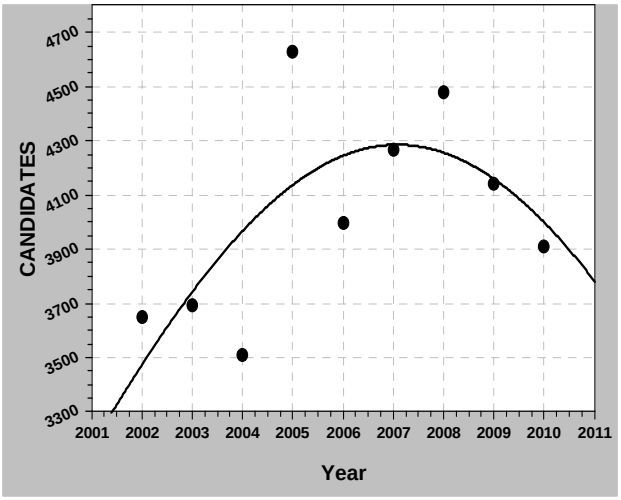


Fig. 5. The regression function corresponding to the evolution of the number of admission candidates in the Technical University of Cluj-Napoca

The calculated coefficients are:

a = 4285,4841; b = 2007,0741; c = 7,8421526;

The standard error is: 317,1626039;

The correlation coefficient is: 0,6971313.

The value calculated for the year 2011 is y=3780,76.

By means of the CurveExpert software, the calculated values for every point on the line are highlighted, so that the differences dif_i between the experimental and calculated values can be obtained (Table 1).

In order to calculate the width of the prediction field, relationships (5), (6), respectively (4) are applied, with a correction coefficient of 0,9456. This coefficient was calculated dependent on the number of potential candidates, according to the statistical data related to the number of children born between 1985 and 2003. For the year 2011, a

number of 260.393 children born in the year 1992 was considered and reported to the number of children born in 1191, i.e. 275.275.

Tab.1 Calculated values for the number of candidates

Year	Candidates	Calculated	Diff.
2002	3648	3476,1	171,90000
2003	3693	3744,5	-51,50000
2004	3509	3968,56	-459,56000
2005	4629	4138,19	490,81000
2006	3997	4245,48	-248,48000
2007	4266	4285,29	-19,29000
2008	4477	4255,72	221,28000
2009	4144	4158,18	-14,18000
2010	3908	3997,35	-89,35000

For q = 0,05 and a number of degrees of freedom $f = n - 1 = 8$, one obtains from the student distribution table, t_q=2,306. With these data, one finds the width of the prediction field calculated for the year 2011: **max.=3787,** and **min.=3365**.

If one takes into account the number of places offered for the year 2011/2012, and defined by the University Senate according to the demands of the faculties in the university and approved by the Ministry of Education, Research, Youth and Sports, namely 3563 (of which 2683 places financed from the government budget and 880 financed by the students themselves), one will notice that the number of places is situated below the average value predicted, though very close to it, which could lead to a probability of occupying the places for admission in a proportion of 53,08 % from admission places offered (Figure 6).

The probability of occupy the admission places (PO) is calculated with relationship 7:

$$PO = \frac{max - cs}{L} \cdot 100, \qquad (7)$$

where: max, is the maximum value of the prediction field;

cs – the number of places offered;

L – the width of the calculated prediction field.

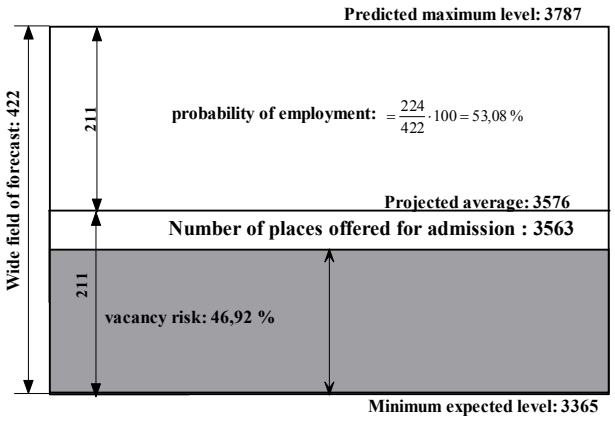


Fig.6. The probability to occupy the admission places in the admission session 2011/2012

If one takes into consideration only the admission number for the budget places, it is visible that the number of predicted candidates exceeds the figure offered, meaning that all the places will be covered in the 20111 admission

After the admission in September 2011, it was shown that predictions were right, all the budget places were fully occupied, the tuition fee places being covered in a proportion of 27,16 %, which is beyond expectations.

At the level of faculties, the prediction can be used as an instrument only if the admission conditions remained constant for many years in line, and this is the case of the Faculty of Automation and Computer Science.

For the rest of the faculties, the predictions do not exhibit the same relevance because of various reasons. The most significant cause lies in not observing the rule that past trends extend in the future, interruption elements appearing due to the fact that they gave up an entrance examination based on a written test in favour of documents provided by the candidates as well as to the conservative opinions of some of the faculties with respect to the predicted number of places. Among these, one can enlist the Faculty of Architecture and Town Planning, where there has been a high level competition every year, though the number of places offered has remained relatively small, and the candidates main concern lies not in occupying the places, but in getting a good entrance score or obtaining a budget place or a tuition fee place, as a candidate.

4. CONCLUSIONS

Statistical processing cannot take into account all the variables that influence candidate decisions. In this kind of prediction, the study of the budget places degree of occupation seems to be nearer to reality as in the case of tuition fee places subjective factors and financial means of candidates need to be taken into account especially in relation to present-day economic situation in the country.

Prediction cannot, then, take into account unpredictable conditions such as changing of entrance examination requirements, introduction of new specialisations, changes in legislation, changes in advertising educational programs, economic crises, overturning of circumstances, as well as relevance of considering only the fist option expressed by the candidates.

Other aspects that were not considered are the baccalaureat scores, which actually have quite a big influence for the candidates' decision. However, experience has proved that there is no direct connection between the two components mentioned above. Hence, the conclusion could be that better trained highschool graduates make for the university, in our case those who pass their final highschool examination during the first session of exams.

But it was proved that the concrete situation produced after the entrance examination for the academic year 2011/2012 came very close to the predictions presented here. The budgeted places were occupied in full, those with tuition fee in a proportion of 27,16 %, a figure that exceeded expectations.

The aim of the prognoses should consist in stimulating those entities whose prediction and forecasting is negative, to enable them develop strategies that could cancel negative trend producing factors; in the case of entities that possess a positive trend, the prediction could help them plan future number of places and capacity of education devices, in correlation with market demand.

Our university strategy consisted in attracting as many students as possible and of occupying first the budget places every year. At the end of the admission session, in case some faculties were left with vacant places, these vacant places were redistributed according to the students' demands to other faculties.

The study here concerns only Bachelor's level. The study may not have the same relevance for Master's students, first because the latter do not have a history as the second bologna cycle began only in 2009 in the Technical University of Cluj-Napoca.

5. BIBLIOGRAPHY

[1] Nicolescu, O. (coord.), Strategia universităţii. Metodologie şi studii de caz, Bucureşti, Editura Economică, 2007.

[2] Shattock, M., Managing successful universities, Society for research Higher education & Open University, New York, 2003.

[3] John Naisbitt, Megatendințe, Bucureşti, Editura Politică, 1989.

[4] http://www.curveexpert.net

PROGNOZA ADMITERII ÎN UNIVERSITATEA TEHNICĂ DIN CLUJ-NAPOCA

***Rezumat*:** Lucrarea abordează o componentă importantă a managementului universitar, admiterea în universitate. Pornind de la informațiile referitoare la numărul de locuri scose anual la concursul de admitere în Universitatea Tehnică din Cluj-Napoca şi gradul lor de ocupare din perioada 2005-2010, prin metode specifice din teoria probabilităților şi statistica matematică, s-au efectuat prognoze cu privire la numărul de candidați care se vor prezenta la admitere în anul 2011. Studiulde față s-a ocupat doar de ciclul de licență, la nivel de universitate. Rezultatele prognozelor au fost validate prin comparația acestora cu datele culese în urma admiterii încheiate în luna septembrie 2011.

Chiorean Lia, engineer, University chief secretary, The Technical University of Cluj-Napoca, The Rector's Office, E-mail: Chiorean.Lia@staff.utcluj.ro, Office telephone: 0264-202209, Address: Calea Dorobanților nr.89, bl.X3, Sc.1, Ap.8., cod:400609, Cluj-Napoca, Home telephone: 0264-410876