INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES

Vol 3, Issue 9, (2013) E-ISSN: 2222-6990

A Mathematical Model on Employment versus Several Factors in Macedonia

Shpresa SYLA
Faculty of Business and Economics, SEEU
Tetovo, Macedonia
E-mail: sh.syla@seeu.edu.mk

Sadri ALIJA
Faculty of Business and Economics, SEEU
Tetovo, Macedonia
E-mail: s.aliji@seeu.edu.mk

DOI Link: http://dx.doi.org/10.6007/IJARBSS/v3-i9/196

Published Date: 12 September 2013

Abstract

In this study is made an attempt to analyze the conjunction of several indicators on the employment of citizens in the Republic of Macedonia. In this regard the logistic regression and odds ratios were used in order to see how certain factors such as gender, age, ethnicity, education, area of residence and the number of family members affect employment at citizens.

This study is based on data from a survey organized by the Center for Economic Analyses in December 2011, by 400 citizens of the Republic of Macedonia.

Results show that young males of Macedonian nationality, those with higher education, ones that live on urban areas and the families that have greater number of members are more likely to be employed.

For data processing is used *software Med Calc*.

Keywords: Employment, Several indicators, Mathematical model, Logistic regression,

1. INTRODUCTION

The total employment rate is the basic indicator of economic and social conditions in a country. The employment rate in the Republic of Macedonia according to the State Statistical Office of Macedonia has increased from 36.8% in 2004 to 41.9% in 2008 and 43.9% in 2011. Yet it is unsatisfactory if we compared with the EU27 countries where the employment rates vary between 60% -65% with increasing trend from 2000 to 2008 and a slight decrease after 2008. Employment rate in Macedonia is higher for males compared to females. In 2011, males had better success in the labor market, about 57.8% and females 39%.

It is already known that the qualitative employment and employment of qualified persons contributes to the achievement of higher labor productivity and economic growth. In

Macedonia, the highest rate of employment is achieved by individuals who have attended tertiary education, about 73.1% in 2008 and 72.4% in 2011. The employment rate of individuals with secondary level of education which meanwhile is considered as the minimum level of education required for successful integration in social activities was 58.3% in 2008 and reached 58.3% in 2011 (a constant increase from 52% in 2004). The lower rate of employment is between persons without any education or individuals that have attended primary education, which is 26.9% in 2004, 32.2% in 2008 and 34.3% in 2011.

This paper mainly relied on the analysis of the survey results using the Logistic Regression to see the relationship among several indicators on the employment of the citizens of Macedonia. Many studies related to the social sciences use logistic regression, especially odds ratios (Menard, 2001; Veenstra, 2011) indicating significant importance in interpreting the results of a statistical study. Odds ratios beside the possibility of evaluating the nature of the given sample, and through it the population, it gives us also the opportunity to determine the extent of this trend, according to changing values of predictive characteristics.

Logistic regression is part of a broad category of statistical models named General Linear Models (Agresti, 2002; Fox, 1997). As a statistical technique it is used to analyze problems in which interfere one or more variables that affect a result, the outcome is considered dependent variable. The case that will be examined in our research is the one where the result is a bi-valued variable such as: presence/absence or success/failure. This makes it arithmetically easy identifiable through two values: the outcome is marked with the value 1 with probability of success p, and the value 0 with failure probability 1-p. This type of variable is known also as Bernoulli's random variables. In our research we aim to find the most suitable model describing the relationship between bi-valued random variable and a set of independent random variables.

Logistic regression model is the result of transformation of the linear probability model. During some equivalent transformations over linear probability model transformation appear odd ratios in order the outcome to get the value 1 versus the value 0 depending on the variables (characteristics) that have influence.

It should be noted that odds ratios have special importance in the interpretation of situations. In the last section of the logistic regression according to the MedCalc software, are provided odds ratios e^{β_i} and their 95% intervals of confidence. Interpretation of changes that occur odd when changing the value of the variable X, explains the trend that the odd has for success depending on the change of variable X, when all other variables remain constant.

Thus, if a variable increase for one unit and all other variables remain constant, then the odd will change according to e^{β_i} type of factor which is called odds ratio. The X factor expresses the relative amount for which the chance of outcome increase (OR> 1) or decrease (OR <1) when the value of this variable increases by one unit.

Natural logarithm of these reports is called logistic regression model.

$$\log it(p) = \ln \frac{p}{1-p} = \alpha_i + \beta_1 X_{i1} + \ldots + \beta_k X_{ik}$$

In this paper, using logistic regression it is built a mathematical model in order to see the relationship among employment in the Republic of Macedonia and some other factors.

2. STATISTICAL MODEL FOR DATA ANALYSIS

On building the mathematical model the dependent variable "employment" is a variable with two values: with the value 1 were identified all interviewees who are employed, and 0 are identified all those who are not employed.

Whereas as independent variables (predictive) will be considered ones that indicate their relationship to the employment of citizens.

Predictive variables that will be examined in this study are:

Gender (**1**=F, **2**=M);
Age (**1**=18-25 years of age, **2**=26-33 years of age, **3**=34-41 years of age, **4**=42-49 years of age, **5**=50-57 years of age, **6**=58-65 years of age, **7**=66+ years of age);
Ethnicity (**1**= Macedonians, **2**= Albanians, **3**= Other);
Education (**1**=No education, **2**=Primary 8 years, **3**=Secondary, **4**=Undergraduate **5**=Master or Doctorate);
Living area (**1**= Rural, **2**= Urban);
Number of family members (**1**=one to two members, **2**=three to four members, **3**=five to six members, **4**=over six members).

The numerical problems are simplified using MedCalc software. The program in its implementation to logistic regression at the beginning gives general information over the survey in which are included: the volume of selection, cases in which the dependent variable is 0 and 1 with respective percentages. Then are calculated regression coefficients, the odds ratio with 95% confidence interval, it is also shown the classification table where the observed values of outcomes intersect the projected values. Finally, the model gives the precisely predicted percentages of cases, which on the other hand shows the effectiveness of the model.

3. RESULTS

We set the data in dialog table of the program MedCalc for logistic regression with *enter* method for these variables:

- **Independent**: Gender, age, ethnicity, education, living area and the number of family members.
- **Dependent**: Employment.

The results shown below are found:

Tab 1: Logistic regression

Dependent Y	Employment
Method	Enter
Sample size	400
Cases with Y=0	68 (17,00%)
Cases with Y=1	332 (83,00%)

Coefficients and Standard Errors

Variable	Coefficient	Std. Error	P
Gender	0,44841	0,31535	0,1550
Age	-0,55761	0,10606	<0,0001

Ethnicity	-0,29737	0,25581	0,2450
Education	1,18703	0,26147	<0,0001
Living Area	0,25335	0,59511	0,6703
Nr. of family members	1,06284	0,26389	0,0001
Constant	-2,7259		

Odds Ratios and 95% Confidence Intervals

Variable	Odds ratio	95% CI
Gender	1,5658	0,8439 to 2,9052
Age	0,5726	0,4651 to 0,7049
Ethnicity	0,7428	0,4499 to 1,2263
Education	3,2773	1,9631 to 5,4713
Living Area	1,2883	0,4013 to 4,1362
Nr. of family members	2,8946	1,7257 to 4,8553

Classification table (cut-off value p=0,5)

Actual group	Predicted group		Percent correct
	0	1	
Y = 0	24	44	35,29 %
Y = 1	6	326	98,19 %
Percent of cases correctly classified			87,50 %

From **Tab 1** we have:
Cases in which Y = 0, which in our case are 68 or 17% of the interviewees are not employed.
Cases in which Y = 1, which in our case are 332 or 83% of the interviewees are employed.
For this group of variables we obtain the *log it model*:

$$\log it(p) = \ln \frac{p}{1-p} = -2.726 + 0.448X_{GJ} - 0.558X_{M} - 0.297X_{E} + 1.187X_{SH} + 0.253X_{Z} + 1.063X_{NAF}$$

The gender odds ratio is 1.51 > 1, which means that a male has better chance of employment for 1.5 times than a female. Regarding the age the odds ratio is 0.57 < 1, which means that younger individuals are more likely to be employed than older ones. The ethnicity odds ratio is 0.74 < 1, which means that Albanian population has $(1-0,7)\cdot 100 = 30\%$ less chance of being employed versus Macedonian population, also (Roma, Turks and others) have 30% less chance of being employed compared with the Albanian population.

Regarding education, living area (area of residence) and the number of family members the odds ratio is 3.27 > 1 1.29 > 1 2.98 > 1, respectively. This means that the educated individuals have greater chances for 3.27 times of being employed, the population of urban areas compared to the rural ones have the advantage of 1.29 times, and finally families with more members have better chances for 2.89 times versus families with fewer members.

4. CONCLUSIONS

The results found show that the logistic regression method is a useful technique for the analysis of various problems in various fields, especially in social sciences. Also in the study

for interpreting the results are used odds ratios; in addition to assessing the nature of the trend that the sample has, depending on the changing predictive characteristic values, it becomes possible to determine the extent of this trend. Numerical Problems are simplified by the fact that today there are many software programs that offer the possibility of performing calculations. In the article is presented the implementation of MedCalc.

The data and analysis show that young Macedonian males, with higher education that live in urban areas and are member of families that have greater number of members are more likely to be employed.

Taking into account difficulties that we faced to come to the basic data from the State Statistical Office of Macedonia, we relied on the survey data from CEA. We assume that this data are accurate, but remains problematic "the sample size" of this phenomenon that represents a limitation in this study.

Bibliography

Agresti, A. (2002). *An introduction to categorical data analysis.* New York: Wiley.

Apostolova, O. (2011). "Poverty assessment and mapping of vulnerable groups, with special focus on Human Trafficking" GIZ – Skopje, Macedonia

Fox J. (2005). LogisticRegression, Maximum-Likelihood Estimation and Related Generalized Models. Zork Summer Programme in Data Analysis.

Fox, J. (1997). *Applied regression analysis, linear models and related methods.* Thousand Oaks, CA: Sage Publications.

MedCalc. (2005). MedCalc, Easy-to-use statistical software. Belgium.

Menard, S. (2001). Logistic Regression Analysis (Quantitative Applications in the Social Sciences **Volume 106**),London: Sage Publications

Pampel F.C. (2000) Logistic regression. Sage Publications

State Statistical Office, Republic of Macedonia, http://www.stat.gov.mk

Veenstra G, (2011). Race, gender, class, and sexual orientation: intersecting axes of inequality and self-rated health in Canada, Veenstra International Journal for Equity in Health.