THE ESTIMATION OF THE QUALITY OF LIFE OF THE POPULATION OF THE REPUBLIC OF ARTSAKH BY ECONOMETRIC METHODS*

Irena Arutyunyan, Anahit Petrosyan

Keywords: Econometric Models, Regression Techniques, Correlation, Coefficient Of Correlation, Determination Coefficient

Abstract: The article considers regression techniques (only methods) for assessing the quality of life of the population of the Republic of Artsakh. In the preliminary data processing possible and subsequent types of correlations between the considered social and economic indicators for assessing the quality of life of the population of the Republic of Artsakh are revealed. After the detection of correlations, the regression analysis apparatus is used to confirm the type of dependencies and to calculate the free term of regression equations and coefficients for independent variables. In the case of significance of independent variables ($p \leq 0.05$), they are included in the model and, if the adjusted coefficient is >75%, the models are considered adequate, and recommended for practical use. The strong point of the article is the following: this is the first study of the quality of life of the population of the of Artsakh depending on social and economic indicators.

The weak point of the article is that this study is conducted is conditions in conditions of insufficient information.

1. Introduction

There are different approaches to assessing the quality of life [1].

Longevity characterizes the ability to live a long and healthy life, which is a natural life choice and one of the basic universal human needs. The basic indicator of longevity is life expectancy, which is the average life expectancy at birth. This indicator, which is calculated separately for the male and female population, is calculated on the basis of the conditional generation, which is made up of different ages population died in a given year.

2. Problem Statement

The section of applied statistics devoted to the restoration of relationship is called regression analysis.

Regression analysis is the most popular method for predicting processes and systems. The problem of regression is to find estimates of unknown parameters and to develop adequate models [1].

2.1. Multiple Linear Regression

A regression with two or more explanatory variables is called a multiple regression. Rather than modeling the mean response as a straight line, as in simple regression, it is now modeled as a function of several explanatory variables.

* Հոդվածն ընդունվել է 02.10.19:
Հոդվածն տպագրության է երաշխավորել խմբագրական խորհրդի անդամ, տ.գ.դ., պրոֆեսոր Հ. Մնացականյանը:16.10.19:

$$Y = \sum_{i=1}^{n} a_i x_i + b_0 + \varepsilon \quad (1)$$

Here is a_i - regression coefficients, b_0 - the free member (if used), the eighth, which is the cause of why different things are done, which, in turn, turns out to be the normalization of normality with the zero vector mat. expectations and correlation matrices.

The use of regression assumes that the remnants of the model (errors in accordance with the observations of the model) are independent random variables having the same distribution law with zero expectation and constant variance [2].

There can be more than one independent variable. At first sight, the residual sum of squares (SS) can be used as an indicator of data deviations from the model. The smaller the number, the better the approximation, and the model better describes the real data. However, this reasoning is only suitable for models with the same number of parameters.

The quality of the multiple linear regression model is estimated by the determination coefficient, F-statistics (Fisher statistics), sum of squares of RSS residuals, standard regression error (SEE). In the case of multiple regression, the adjusted determination coefficient (adjusted) should also be used, which is used when observations or variables are excluded or added to the model.

An important indicator of the quality of the linear regression model is the verification of Gauss - Markov requirements for residuals. All Gauss-Markov conditions are satisfied in the qualitative model of linear regression:

2.2. Conditions

1. the mathematical expectation of residues is zero for all observations ($\varepsilon(e_i) = 0$);
2. theoretical dispersion of residues is constant (equal to constant) for all observations ($\sigma^2(e_i) = \sigma^2 (e_i)$, i = 1,..., n);
3. there is no systematic relationship between the residues in any two observations;
4. no relationship between residuals and explanatory (independent) variables.

The data was collected from the Statistical Service website, 1995-2017. Social economic, housing stock, health care, education, finance, etc. This data is presented in Table 1.

Before fitting our regression model we want to investigate how the variables are related to one another. We can do this graphically by constructing scatter plots of all pair-wise combinations of variables in the data frame.

Table 1. Main social-economic indicators.

Time	x_1	x_2	x_3	y_1	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
1995	126	36345	3103	4126	-	1640	4060	3724	341		1744
1996	131	36965	4055	7984	14221	2564	4701	3976	376		1556
1997	135	36637	5107	12261	17697	3201	7102	4354	488		1931
1998	137	33975	3630	16170	19796	4275	8249	9021	484		2514
1999	139	47015	3189	24828	20999	3893	8352	5517	205		1614
2000	134	47794	3541	27222	23149	4855	7424	7960			2211
2001	136	48862	3461	28178	23881	5904	6404	6416			2103
2002	137	49768	3403	29673	26478	8083	6734	4323	84	9262	2235
2003	137	50123	3314	33661	33884	11125	8694	4982	284	10018	3170
2004	137	51404	3278	41170	42830	18579	9139	6704	427	9167	4590
2005	138	52860	3515	51127	51379	17773	12597	7623	168	7298	6570
2006	138	54599	3560	56700	61886	24204	17602	10530	159	11914	9387
2007	139	56112	3004	68610	70791	22437	20810	14694	361	11818	11641
2008	140	58503	3464	80480	87148	25346	25836	14130	585	21108	19912
2009	141	58792	3531	88768	102339	34092	34481	21680	875	39570	21251
2010	144	58028	3445	92736	118187	42992	42769	26732	683	46418	23808

2011	145	60064	3313	98453	135499	45823	43360	20079	1008	50214	25756
2012	147	60865	2915	102777	150016	40872	37890	18249	876	74708	28756
2013	144	62420	2567	128621	168564	44339	45220	18673	206	56163	27919
2014	145	52365	1534	141193	188840	52047	49670	26303	367	59477	31485
2015	145	54079	1232	151058	209346	53541	50145	53043	675	68005	34436
2016	146	53039	1368	152707	229652	59000	45656	23485	1027	39245	33711
2017	147	53559	1417	155433	272071	97,490	47100	35960	402		34687

Table 1. Continued.

Time	x_{11}	x_{12}	x_{13}	x_{14}	y_2	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	y_3	y_4
1995	4222	2478	307		53			6477		27503	404	
1996	6179	4623	310		70			9050	22797	27768	527	
1997	8247	6317	552		93			11771	24724	29483	425	
1998	9247	6733	1352		113			14941	26499	29529	350	16392
1999	11496	8261	1081	28000	128			16189	27738	32571	291	13073
2000	11627	9417	1560	30325	176			22193	29395	32502	442	31116
2001	11132	9029	3633	52646	197			25400	29684	43688	449	35562
2002	11323	9088	5093	62205	215			28100	30068	44343	461	41139
2003	12296	9126	4152	83644	251	68	74	33306	31361	48614	619	36884
2004	14505	9915	6651	115582	314	65	74	41108	33704	50505	871	37767
2005	20246	13676	9732	133118	376	66	75	49230	34352	53886	642	33937
2006	25991	16604	7814	158511	421	66	74	56449	34720	58481	826	55314
2007	30837	19197	10178	277393	498	68	76	66862	34837	69809	1061	32933
2008	50075	30163	17334	305064	577	70	76	79086	35095	98789	1986	39909
2009	55098	33848	28995	293967	625	70	77	88249	37364	127055	1364	43593
2010	62182	38375	42229	348235	675	72	77	98677	37671	137349	1364	46389
2011	67377	41621	59437	92707	781	71	76	111530	38885	147788	1346	49383
2012	67344	38588	78927	349012	854	72	77	124513	38975	161085	1402	47340
2013	69623	41702	82162	328470	921	71	76	135719		164215	1266	61373
2014	78127	46641	97536	366651	1013	72	77	150466		170429	1156	65825
2015	83108	48671	99806	314557	1053	72	77	150741		178256	654	69214
2016	84670	50959	117736	337966	1076	71	77	155649		197929	1105	77585
2017	92099	50743	149255	444477	1065	73	78	161700		191767	825	86556

2.3. An Example of the Practical use of Linear Multiple Regression Analysis

It is necessary to identify the factors (indicators) that have the greatest impact on the quality of life of the population of the Republic of Artsakh and to find the relationship of these factors.

Dependent variables-Y1, Y2, Y3, Y4, independent variables: X1-X19.

A multiple linear regression model was used to restore the relationship. According to the results of the step-by-step analysis, independent variables (lift parameters), having (in the linear model) coefficients, slightly different from zero, in other words, little different in comparison with their dispersion, were consistently excluded from consideration. For this purpose, we used the package STATISTICA 12.592, specifically the module "Multiple regression".

2.4. Modeling Results

Social-Economic Indicators

X1 - Population size (at the end of year), thsd. Persons;
X2-Average number of employed in the economy, persons;
X3-Total number of officeially registered unemployed, persons;
Y1- Average monthly nominal wage of persons engaged in the economy, drams;
X4-Gross Domestic Product
X5-Industrial output, mln. Drams;
X6-Capital construction, mln. Drams;
X7-Humanitarian aid,

X9- Total investments,
X10- Receipts of state and regional budgets, mln. Drams;
X11-Outlays of state and regional budgets,
X12-Deficit of state and regional budgets, mln. Drams;
X13-Creditinvestments;
X14-External trade turnover, thsd. US dollars;
Y2- Money incomes of population per capita, thsd. Drams;
X15-Total population, males;
X16-Total population, females;
X17-Money expenditures and savings of population total
X18-Citizens' ownership;
X19-Passengers turnover total, thsd. passenger-kilometers;
Y3-Graduates of higher educational establishments;
Y4-Number of registered diseases with the diagnosis set for the first time, total.
The following multiple linear regression equation is obtained:

if the value of the adjusted coefficient of determination is greater than 75%, then the model adequately describes the real systems and processes and if the significance level is less than or equal to 0.05, the model is significant.

Below are the results of the regression analysis.

Table 2.Regression Summary for Variable Y2

N=15 Regression Summary for Dependent Variable: Y_1 (Irena. sta) **R=.99910982 R2=.99822043 Adjusted R2=.99688576** **F (6,8)=747.91 p<.00000 Std. Error of estimate: 2331.7**						
	b*	**Std. Err.**	**b**	**Std. Err.**	**t (8)**	**p-value**
Intercept			251369.0	96876.94	2.59472	0.031878
x4	1.15777	0.0654	0.7	0.04	17.693	0.000000
x6	0.43	0.052	1.2	0.14	8.408	0.0003
x5	-0.3716	0.0506	-0.7	0.10	-7.338	0.000081
x1	-0.257	0.0527	-2874.4	588.07	-4.887	0.001212
x16	0.101	0.0361	3303	1181.77	2.794	0.023378
x15	-0.081	0.04781	-1329.6	778.90	-1.707	0.126194
$Y1=251369+ 1.157770x_4 + 0.437384x_6 -0.371617x_5 -0.257663\ x_1 + 0.101177x_{16} -0.081753x_{15}$						

Analysis of the significance of multiple linear regression model coefficients:

Table 3. Significance levels of variables

	p-value
Intercept	0.03188
x_4	0.00000
x_6	0.00003
x_5	0.00008
x_1	0.00121
x_{16}	0.02338
x_{15}	0.12619

The following multiple linear regression equation is obtained for Y_1 adjusted:

Table 4. Regression Summary for Y1

N=15 Regression Summary for Dependent Variable: Y1 (Irena. sta) R=.99821430 R2=.99643178 Adjusted R2= 99500449 F (4,10)=698.13 p<.00000 Std. Error of estimate: 2953.2						
	b*	**Std. Err.**	**b**	**Std. Err.**	**t (10)**	**p-value**
Intercept			**426784.6**	**90553.70**	**4.71306**	**0.000825**
x4	1.158331	0.080983	0.7	0.05	14.30342	0.000000
x6	0.440356	0.057786	1.2	0.16	7.62045	0.000018
x5	-0.347013	0.062290	-0.7	0.12	-5.57091	0.000237
x1	-0.269725	0.060167	-3009.0	671.20	-4.48296	0.001174

Analysis of the significance of multiple linear regression model coefficients:

Table 5 Significance levels of variables

	p-value
Intercept	0.000825
x_4	0.000000
x_6	0.000018
x_5	0.000237
x_1	0.001174

To draw a conclusion about the quality of the model as a whole.

Answer. All indicators of model quality. The value tends to one, and the adjusted value is still high. In contrast, RSS is low and p-level is high.

The following multiple linear regression equation is obtained for Y2:

Table 6 Regression Summary for Y1

N=15 Regression Summary for Dependent Variable: Y2 (Irena. sta) R=.99962346 R2=.99924707 Adjusted R2=.99897328 F (4,11)=3649.6 p<.00000 Std. Error of estimate: 24.152						
	b*	**Std. Err.**	**b**	**Std. Err.**	**t (11)**	**p-value**
x_6	0.275191	0.048674	0.005769	0.001020	5.65381	0.000148
x_4	0.833531	0.073250	0.004288	0.000377	11.37931	0.000000
x_2	0.240890	0.022747	0.003248	0.000307	10.59001	0.000000
x_5	-0.330319	0.057808	-0.005584	0.000977	-5.71409	0.000135
Y_2= 0. 005769x_6 +0.004288x_4+ 0.003248x_2 -0. 005584x_5						

Analysis of the significance of multiple linear regression model coefficients:

Table 7. Significance levels of variables

	p-value
x_6	0.000148
x_4	0.000000
x_2	0.000000
x_5	0.000135

The following multiple linear regression equation is obtained for Y3:

Table 8. Regression Summary for Y3

N=15 Regression Summary for Dependent Variable: Y3 (Irena. sta) R=.98615741 R2=.97250644 Adjusted R2=.95188627 F (6,8)=47.163 p<.00001 Std. Error of estimate: 3628.8						
	b*	**Std. Err**	**b**	**Std. Err**	**t (8)**	**p-value**
Intercept			772680.7	157821.9	4.89590	0.001200
x_4	1.894092	0.280818	0.4	0.1	6.74490	0.000146
x_{16}	-0.475762	0.125192	-6149.7	1618.2	-3.80026	0.005235
x_1	-0.463705	0.212136	-2048.2	937.0	-2.18589	0.060306
x_3	0.452784	0.199296	8.3	3.7	2.27191	0.052733
x_2	-0.200876	0.111118	-0.9	0.5	-1.80777	0.108259
x_6	0.292791	0.200970	0.3	0.2	1.45689	0.183248
Y_3= 772680.7+ 0.4x_4 -6149.7x_{16}- 2048.2x_1+8.3x_3+0.5x_2+0.2x_6						

Analysis of the significance of multiple linear regression model coefficients:

Table 9. Significance levels of variables

	p-value
Intercept	0.001200
x_4	0.000146
x_{16}	0.005235
x_1	0.060306
x_3	0.052733
x_2	0.108259
x_6	0.183248

The following multiple linear regression equation is obtained for Y3 adjusted:

Table 10.Regression Summary for Y3

N=15 Regression Summary for Dependent Variable: Y3 (Irena. sta) R=.96900833 R2=.93897714 Adjusted R2=.92880666 F (2,12)=92.324 p<.00000 Std. Error of estimate: 4414.2						
	b*	**Std. Err.**	**b**	**Std. Err.**	**t (12)**	**p-value**
Intercept			519647.2	120894.0	4.29837	0.001035
x_4	1.348609	0.125978	0.3	0.0	10.70509	0.000000
x_{16}	-0.514912	0.125978	-6655.8	1628.4	-4.08730	0.001506
Y_3=519647.2 + 0.3x_4 -6655.8x_{16}						

Analysis of the significance of multiple linear regression model coefficients:

Table 10. Significance levels of variables

	p-value
Intercept	0.001035
x_4	0.000000
x_{16}	0.001506

The following multiple linear regression equation is obtained for Y_4:

Table 11. Regression Summary for Y4

N=15	Regression Summary for Dependent Variable: Y_4 (Irena. sta) R=.95371705 R?=.90957621 Adjusted R?=.90311736 F (1,14)=140.83 p<.00000 Std. Error of estimate: 360.07					
	b*	Std. Err.	b	Std. Err.	t (14)	p-value
x_{15}	0.953717	0.080367	15.79609	1.331090	11.86703	0.000000

Analysis of the significance of multiple linear regression model coefficients:

Table 12. Significance levels of variables

	p-value
x_{15}	0.000000

Using the Student's research, the hypothesis that the corresponding coefficient is slightly different from zero is tested, and accordingly, the variable with this coefficient has a negligible impact on the dependent variable. In turn, the column p-level displays the probability that the main hypothesis will be accepted. If the value of p-level is greater than the significance level α, then the main hypothesis is accepted, otherwise it is rejected. In our example, the significance level α=0.05 isset.

3. Conclusions

As indicators of the quality of life of the population, basic social and economic data are taken into account.

As a result of a probabilistic analysis of the data obtained for the 23 we obtained 4 dependent variables for assessing the quality of life of the population: average monthly nominal wage, per capita income, graduates of higher educational establishments, number of registered diseases with the diagnosis set for the first time, total; 7 independent variables: population size (at the end of year), thsd. Persons; average number of employed on the one hand in the economy, persons; GDP, volume of capital construction, total population, life expectancy of men and women in the Republic of Artsakh.

The undoubted advantage of the obtained models is their reliability and adequacy to real processes.

REFERENCES

1. Ayvazyan SA, Mkhitaryan B. C. Applied statistics and the foundations of econometrics.-M.: UNITY, 1998.-1000 p.
2. Dougherty K. Introduction to Econometrics. 3-ed.-M.: 2009. -465p.
3. Dubrov A. M., Mkhitaryan B. C., Troshin L. I. Multidimensional statistical methods.-Moscow: Finance and Statistics, 2003.-352p.
4. Magnus Ya. R., Katyshev PK, Peresetsky A. A. Econometrics. The initial course.-Moscow, ed. "The Case", 2005.-570p.
5. Jesse, Russell Econometrics/Jesse Russell.-Moscow: VSD, 2017.-717 p.
6. Aivazyan, S. A. Econometrics/S. A. Aivazyan, S. S. Ivanov.-M.: Market DS, 2017.-104 p.
7. Artamonov, N. Introduction to econometrics/N. In. Artamonov.-M.: mcnmo, 2015.-204 p.
8. Artamonov, N. Introduction to econometrics/N. In. Artamonov.-Moscow: mcnmo, 2016.-224 p.
9. Afanasiev, V. N. Econometrics/V. N. Afanasyev, M. M. Yuzbashev, T. I. Gulyaeva.-Moscow: Finance and statistics, 2017.-256 p.

10. Berndt, Ernst, the Practice of econometrics. Classic and modern/Ernst Berndt.-M.: Unity-Dana, 2016.-848 p.
11. Verbik, Marno Guide to modern econometrics/Marno Verbik.-M.: Scientific book, 2016.-616 p.
12. Gladilin, A. V. Workshop on econometrics/the A. V. Gladilin, A. N. Gerasimov, E. I. Gromov.-Moscow: Phoenix, 2016.-336 p.
13. Dayitbegov, D. M. Computer technology for data analysis in econometrics/D. M. Dayitbegov.-Moscow: IL, 2015.-592 p.
14. Kolemaev, V. A. Econometrics/V. A. Kolemaev.-M.: INFRA-M, 2016.-160 p.
15. Kochetygov, A. A. Fundamentals of econometrics/A. A. Kochetygov, L. A. Tolokonnikov. -. Moscow: Publishing center "March", 2015.-352 p.
16. www. stat-nkr. Am.

ԱՄՓՈՓԱԳԻՐ

Արցախի բնակչության կյանքի որակի գնահատումը էկոնոմետրիկական մեթոդներով

Իռենա Հարությունյան, Անահիտ Պետրոսյան

Բանալի բառեր՝ էկոնոմետրիկ մոդելներ, ռեգրեսիայի մեթոդներ, կոռելացիայի գործակից, հարաբերակցություն, դետերմինացիայի գործակից:

Հոդվածում դիտարկվում է ռեգրեսիոն տեխնիկան (միայն մեթոդներ) ՝ Արցախի Հանրապետության բնակչության կյանքի որակը գնահատելու համար: Տվյալների նախնական մշակման գործընթացում բացահայտվում են Արցախի Հանրապետության բնակչության կյանքի որակի գնահատման համար դիտարկված սոցիալական և տնտեսական ցուցանիշների միջև փոխհարաբերությունները: Համադրումների հայտնաբերումից հետո, ռեգրեսիայի վերլուծության ապարատը օգտագործվում է հաստատելու կախվածության տեսակը և հաշվարկում է անկախ փոփոխականների համար ռեգրեսիայի հավասարումների և գործակիցների վստահելի հավանականու-թյունը: Անկախ փոփոխականների նշանակալիության մակարդակի՝ ($p \le 0.05$) դեպքում դրանք ներառվում են մոդելի մեջ, իսկ եթե ճշգրտված գործակիցը> 75% է, ապա մոդելները համարվում են համարժեք և առաջարկվում են գործնական օգտագործման համար: Հոդվածի ուժեղ կետը հետևյալն է. Սա արցախյան բնակչության կյանքի որակի առաջին ուսումնասիրությունն է ՝ կախված սոցիալական և տնտեսական ցուցանիշներից:
Հոդվածի թույլ կետն այն է, որ այս ուսումնասիրությունն իրականացվում է անբավարար տեղեկատվության պայմաններում պայմաններում:

РЕЗЮМЕ

Оценка качества жизни населения Республики Арцах эконометрическими методами

Ирена Арутюнян, Анаит Петросян

Ключевые слова: эконометрические модели, методы регрессии, корреляция, коэффициент корреляции, коэффициент детерминации.

В статье рассматриваются регрессионные методы (только методы) оценки качества жизни населения Республики Арцах. При предварительной обработке данных выявлены возможные и последующие типы взаимосвязей между рассмотренными социально-экономическими показателями для оценки качества жизни населения Республики Арцах. После

обнаружения корреляций используется аппарат регрессионного анализа для подтверждения типа зависимостей и вычисления свободного члена уравнений регрессии и коэффициентов для независимых переменных. В случае значимости независимых переменных ($p \leq 0,05$) они включаются в модель и, если скорректированный коэффициент составляет> 75%, модели считаются адекватными и рекомендуются для практического использования. Сильной стороной статьи является следующее: это первое исследование (эконометрическими методами) качества жизни населения Арцаха в зависимости от социально-экономических показателей. Слабым местом статьи является то, что данное исследование проводится в условиях недостаточной информации.