8th International Scientific Conference
Science and Society 2015

Yarnykh E.A., Oleynikov B.I., Tsylina I.T.

STATISTICAL METHODS FOR THE ANALYSIS OF DEMOGRAPHIC PROCESSES WITHIN THE REGION

Yarnykh E.A., Professor of Plekhanov Russian University of Economics, Russia
Oleynikov B.I., Associate Professor of Plekhanov Russian University of Economics, Russia
Tsylina I.T., Associate Professor of Plekhanov Russian University of Economics, Russia

Abstract
This work represents the assessment of the interaction of factors giving rise to changes in the structure, dynamics and tendencies regarding the main demographic processes within the region. It contains the analysis of interrelation within a complex of resultative indicators and exposures by means of carrying out the correlation and regression analysis of sociodemographic factors.
Keywords: demographic situation, correlation and regression analysis, factors, birth rate, housing per capita ratio, migration gain, reproduction of population.

Introduction
The demographic situation taking shape within the country, as well as the types of the reproduction of population and their modification over a long period of time at the regional level imply the need for a more profound assessment of ongoing processes in order to obtain a research-backed basis when managing the demographic situation. The identification of consistent patterns regarding reproductivity at the level of particular regions is an important factor and condition for the development of key directions of the regional demographic policy.

The general development of the country's demographic situation is determined by the result of the interaction of various social phenomena and processes running in its certain constituent entities. Due to that, the multivariate statistical analysis of demographic development at the regional level is of significant

theoretical and practical interest, both from the viewpoint of the development of particular regions and of the country as a whole.

Irkutsk Region occupies the sixth position in the country by its total area and has essential raw materials and scientific potential. Nevertheless, despite a set of vital preconditions, there is also a series of negative demographic trends in the region. This circumstance makes it necessary to carry out the objective quantitative analysis of demographic processes within the said territory.

This work represents the correlation and regression analysis of the interference of factors with the dynamics of the main demographic processes within Irkutsk Region.

Research Materials and Methods

The solution of the task aimed at revealing the current features of the reproduction of population in Irkutsk Region makes it necessary to carry out profound analysis of reasons causing the appearance of the said processes. This may be accomplished by means of the correlation and regression analysis, which will enable to make an in-depth quantitative assessment of the impact of various factors on the demographic condition of Irkutsk Region, as well as to show interrelations within the a complex of resultative indicators and exposures.

The correlative component of the analysis in question is aimed at the establishment of links and their closeness between various characteristics. The regression part of the analysis represents real numerical analysis of factors, on which basis it might be possible to construct a mathematical model reflecting the dependence of any given resultative indicator on one or more exposures. With the use of the obtained formula, particular tendencies may be checked for reliability and fairness.

The objective of this work is to determine the interrelation and interference of sociodemographic factors with the amount of changes in the reproduction of population in a particular region.

Results and Discussion

The population change is directly and fundamentally affected by fertility dynamics. In order to carry out the analysis, the birth rate (Y_1) being the main indicator was selected among principal criteria impacting the formation of the population size.

There are many factor characteristics determining the condition of the above components of the population size. For instance, among important factor features there may be the age pattern of the population of Irkutsk Region, as well as the

percentage and dynamics of female population in the region's total population. At the same time, considering the realities of modern life, the determining factors affecting the birth and death rate patterns in the region are undoubtedly the economic realities of the population's life-sustaining activity. In this regard, the assessment of dependency between the factors determining the population's material welfare and birth/death rates is of utmost interest.

In our opinion, the exposures in a varying degree affecting the birth rate being one of the key elements for the population change are:

X_1 – proportion of employed population (in % as compared to economically active population), ‰;

X_2 – ratio of population with incomes under the minimum subsistence level, %;

X_3 – proportion of population under and over employable age, %;

X_4 – morbidity rate (proportion of those fallen ill over the year as compared to total population), %;

X_5 – ratio of average monthly per capita income to the minimum subsistence level, units;

X_6 – housing (sq. m. per capita);

X_7 – marriage rate per mille;

X_8 – migration gain rate.

Table 1 represents initial data on the values of selected indicators over 10 years – from 2004 to 2013.

Table 2 represents calculated values of pair correlation coefficients involving resultative birth rate indicators and exposures.

As a result of the calculations carried out to define the values of pair correlation coefficients, it was revealed that among decisive factors to the maximum extent affecting the birth rate in Irkutsk Region (strong correlative relationship) there are the following factor features:

- Housing per capita ratio – the highest dependence coefficient of 0,907;
- Migration gain rate – 0,892.
- Morbidity level – 0,888;
- Marriages per capita – 0,751.

The significant impact of these factors at the present stage of development is quite logical:

- housing per capita is one of the indicators of stabilization of the situation in life and it becomes the condition for families' replenishment with children;

- considering the migration structure, as well as cultural and mental specifics of the main migration groups, the impact of this factor is also big;
- morbidity level is also one of the features of stabilization of citizens' social arrangement;
- the fact that the matrimony rate in the region impacts the birth rate is also certain:

It is also worth mentioning such a factor as the ratio of population with incomes under the minimum subsistence level in conjunction with the remaining factors. A high degree of correlation can be observed between housing and marriage rate per capita.

Table 1
Initial Data on Birth Rate Exposures in Irkutsk Region over 2004 – 2013

Period	Crude birth rate, ‰	Proportion of employed population as compared to economically active population, %	Ratio of population with incomes under the minimum subsistence level, %	Ratio of population not falling within the category of able-bodied citizens, %	Morbidity rate (proportion of those fallen ill over the year as compared to total population), %	Ratio of average monthly per capita income to the minimum subsistence level, units	Housing (sq. m. per capita)	Marriage rate per mille, ‰)	Net migration rate (per 10000 people), ‰)
	Y_1	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
2004	12,4	89,7	29,0	34,8	84,27	2,00	19,6	6,1	-69,65
2005	11,9	90,2	21,3	35,1	83,39	2,39	20	7,2	-75,76
2006	12,3	91,5	18,9	35,8	86,71	2,57	20,3	7,8	-71,14
2007	13,8	92	18,4	36,2	89,05	2,61	20,6	9,3	-46,83
2008	15,0	91,5	16,8	36,7	87,11	2,78	20,8	8,3	-37,98
2009	15,2	89,2	18,5	37,9	91,19	2,69	21,1	8,6	-45,29
2010	15,2	89,9	18,1	38,3	90,72	2,67	21,4	8,8	-58,61
2011	15,3	90,9	19,2	39,4	92,05	2,63	21,7	9,7	-28,02
2012.	15,9	92,2	16,8	40,3	92,06	2,72	22,1	8,6	-29,9
2013	15,7	90	17,0	41,3	94,63	2,50	22,3	9,2	-35,3

The proportion of population over and under employable age (i.e. children and pensioners), in compliance with obtained results, also has a negative impact on the birth rate dynamics. The correlation level making up -0,22 demonstrates that the more children and pensioners there are in the region, the lower childbirth figures tend to be obtained. Such a correlation is caused by the fact

that these population groups take no part in the regional reproduction.

Table 2
Matrix of Pair Correlation Coefficients by Selected Indicator Sets

	Y_1	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Y_1	1								
X_1	0,069	1							
X_2	-0,653	-0,414	1						
X_3	0,877	0,042	-0,608	1					
X_4	0,888	0,024	-0,638	0,944	1				
X_5	0,656	0,425	-0,933	0,477	0,542	1			
X_6	0,907	0,127	-0,711	0,988	0,947	0,599	1		
X_7	0,751	0,284	-0,797	0,716	0,830	0,757	0,788	1	
X_8	0,892	0,318	-0,595	0,799	0,801	0,597	0,823	0,766	1

Table 3 represents indicators of significance testing with regard to correlation coefficients basing on the use of Student's criterion. As can be seen from the data obtained in Table 3, the resulting correlation coefficients are essential for exposures characterizing the population's occupational level and the ratio of population with incomes under the minimum subsistence level because the calculated value of Student's t-criterion is higher than its critical value.

Pursuant to the data from Table 3, it is appropriate to include exposures X_6 (housing per capita), X_7 (number of marriages per capita) and the time parameter into the regression equation.

Using the stepwise regression procedure, it is necessary to include exposure X_6 - housing per capita ratio - into the regression model as it has the highest correlation coefficient of 0,907. On the basis thereof, the one-factor regression model will be as follows:

$$\hat{Y}_1 = -18{,}48 + 1{,}56X_6$$

The computed coefficient of determination (0,71) demonstrates that 71 % of variation in the resultative indicator are caused by the variation in the exposure of X_6, and, accordingly, 29 % - by a random component.

Table 3

Indicators of significance testing of linear correlation coefficients (basing on Student's criterion) by results of correlation and regression analysis in Irkutsk Region for the population's birth rate

Factor components	Root-mean square error	Student's criterion value	Student's criterion critical value	Test significance	Number of degrees of freedom
X_4	0,22	4,2	2,20	0,04	11
X_6	0,24	2,9	2,20	0,04	11
X_7	0,25	1,6	2,20	0,04	11
X_8	0,24	1,9	2,20	0,04	11

To make the regression model more precise and to decrease the impact of a randomly varied factor characteristic, it is necessary to introduce the factor variable of X_8 having the highest partial correlation coefficient. Thus, a two-factor regression model will be as follows:

$$\widehat{Y_1} = -3{,}31 + 0{,}93X_6 - 0{,}0385X_8$$

The multiple coefficient of determination equaling to 0,86 demonstrates that 87 % of general variations in the resultative characteristic are caused by the variation in factor features of X_6 and X_8.

As it is possible to treat a model as a reliable one if the coefficient of determination is over or equal to 0,7, then this model can be viewed as essential. As the time parameter was also included into the analysis, the resulting equation will be as follows:

$$\widehat{Y_1} = -3{,}34 + 0{,}93X_6 - 0{,}0385X_8 + 0{,}02t$$

A high value of the multiple coefficient of determination (0,86) demonstrates that 82 % of general variations in the resultative variable are described by means of factor characteristics making up part of the displayed model. The regression coefficients are statistically significant according to Student's criterion (2,55; 12,21), and the model is significant according to Fisher's ratio test (9,67).

Conclusions

The obtained regression model indicates that birth rate dynamics in Irkutsk Region depends significantly on the population's quality of life indicators.

References:

[1] Yarnykh E.A. System of Indicators for the Characteristics of Population's Living Standards. Proceedings of the International Applied Research Conference "Values and Interests of Present-Day Society." - M., RGTEU, 2011.

[2] V. S. Mkhitaryan, V. P. Sirotin, M. Yu. Arkhipova, T. A. Dubrova. Mathematical and Statistical analysis of Social and Economic Processes. Inter-University Collection of Research Papers. Issue 10 / Executive Editor: T. A. Dubrova; under the general editorship of V. S. Mkhitaryan; Issue 10. M. : MESI, 2013.