ZIPF'S LAW APPEARANCE IN THE RUSSIAN CITIES[1]

Svetlana RASTVORTSEVA
World Economy Chair – Belgorod State National Research University, Russia, http://www.bsu.edu.ru
SRastvortseva@gmail.com

Inna MANAEVA
World Economy Chair – Belgorod State National Research University, Russia, http://www.bsu.edu.ru
in.manaeva@yandex.ru

Abstract

The understanding of concentration processes about resources, population, enterprises in some regions and in the cities is very significant for economists and policy-makers. It's caused by the worldwide urbanization trend and local trend of economic activity agglomeration that increase the regional development differentiation within the country. Issues of economic activity locations and space distribution are solved by scientists over the past two centuries. Recent works show the increasing interest of economists to the Zipf's law testing in the regional system and the rank-size distribution of the cities. Research aims are to test the Zipf's law in the Russian cities and to test the hypothesis that the Russian Zipf coefficients depends on the size of the geographical territory of the Federal District.Methodology. In the paper it's used least square method for tasting the Zipf's law in Russian cities in general and separately for the federal districts. There is 1,123 Russian cities panel (cities with over 1,000 people population in 2014).Results. The Zipf's law is confirmed in the territory of Russia in general. According to the Federal Districts the Zipf coefficients range from -0.65 (Far Eastern Federal District) to -0.9 (the Urals and the North Caucasian Federal Districts). Equitability of cities hierarchy in the Ural and the North Caucasus Federal Districts dues to the fact that there are 139 cities located in the 1,789 thous. km2 in the Urals and 56 cities in the 170 thous. km2 in the Caucasus. In the Far East the city location is very disperse - 66 cities in the area of 6000 thous. km2 (Zipf coefficient - 0.65). Conclusions. Testing of the Zipf's law for the Russian cities in general shows that it's valid for the small (8,600 – 15,300 peoples) and large cities (66,700 – 331,000 peoples). For cities panel with population exceeds 100 thous. people. The Zipf's law is not valid for cities of more than 1 million people. (exception – the city of St. Petersburg). The result of the study is the confirmation of the hypothesis that the Zipf coefficient depends on the size of the Federal District.

Keywords: location theories, the Zipf's law, the city, the rank-size distribution, the cities of Russia

JEL classification: R12

1. Introduction

Understanding of concentration processes for resources, population, enterprises at specific territories and more often in cities is of significant importance at implementation of social and economic policy. It is associated and world-wide tendency to urbanization and agglomerative factors arisen in staging area of economical activity and reinforcing differentiation of regional development within the country. Economic activity is considered as not only location of industrial production, service sector and other types of activity, but also population, labour resources, investments and even prosperity (expressed for e.g. in salary level). It shall be noted that traditionally in allocation theory households and agricultural sector are studied within land use and do not belong to economic activity (Beckmann, 1999, p. 61).

[1] This work is supported by RFBR - Russian Foundation for Basic Research, project No. 15-36-20012

When business solution on production location in one or other region (city) is taken the factors which should provide high profitability in the future move to the forefront. We can put perspectives of further extension of activity, availability of labour resources of the required level and at a low price, acceptable level of taxes and possibility of subsidies acquisition, level of infrastructure development, in particular transport one, availability of social objects to such factors. Regional governments, for its part shall predict dynamics and qualitative parameters of the above factors for organization of corresponding events on promotion of the territory attraction.

Issues on the economic activity allocation in territorial area (city, region, country) are being solved by scientists within two hundred years. Very seldom economics contain synthetic relationships which can be considered as laws. Zipf's law for cities is one of the most evident dependences in economics and public-social sciences in total (Brakman et al., 1999). Modern studies demonstrate interest of economists to Zipf's law and the cities rank-size distribution.

The purpose of the survey, which results are indicated in the present article, is check of Zipf's law implementation in the Russian cities, proof or disproof of hypothesis that in Russia Zipf coefficient depends on size of geographical territory of the federal district.

2. Theoretical background

To obtain the stated objective, let's study available empirical studies on the given problems in foreign economic literature. F. Auerbach has proposed hypothesis of empiric dependence between size of the city (its population) and its rank in hierarchy of region or country cities (Auerbach, 1913). Zipf's law ensures that within the territory distribution per the city size complies with Pareto distribution with index equal to one. Other determination of the Zipf's law lies in the fact that if the large cities to be ranged per decreasing of their population so ratio of two cities population will be inverse to ratio of their ranks (Andreev et al., 2015). T. Rozen and M. Reznik were the first persons who have performed complex surveys basing on the 44 countries. Approbation of Zipf's law has shown that the highest index of Pareto is for Australia (1.9), the least one – for Marocco (0.8). Surveyors have stated that in Australia the index is excessive one and they put such a case to the exceptions. If Australia is removed from sampling, the leader per Pareto index is Nigeria − 1.5.

T. Cameron proposes one-stage structural model for determination of city size, approbation which has taken part in 121 cities of the USA with population exceeding 100 thous. people. Surveyor states that city size depends on a lot of factors one of which is the distance to the capitals (Cameron, 1990).

Regularities of the cities hierarchy were stated by P. Krugman (Krugman, 1996). As a result of analysis of Zipf's law action for empirical regularity and inadequacy of available cities he has entered the term "secret of the cities hierarchy". K. Gabaix states that Zipf's law for cities is a sample of agglomerating law which determines the most accurate regularity in economy. It covers almost all countries and time periods. Scientist states that Zipf's law shall be prerequisite for the law of cities growth at local level. Surveyor determines two groups of such a law explanation. Economical prerequisites and random processes. Economic explanation is based on balancing of transport expenses, positive and negative external influence, and difference in effectiveness. Main problem of such a division is that it is difficult to determine that different economical structures (the USA in 1991 and India in 1911) perform one and the same power balance. Within the given theory there are no prerequisites for the implementation of Zipf's law.

Y. Mansury suggests a model, which determines a role of spatial distribution in creation of Zipf's law (approbation in the USA). In model the distribution per cities sizes obtains permissible balance. The author extends functions of model due to the inclusion of external effects (Mansury, 2007).

Z. Xu and R. Harriss state that Zipf's law is observed in the most part of countries. Concept of cities size distribution requires accounting of inter-city relations on the process of economic growth (Xu et al., 2010).

N. Moura and M. Ribeiro call Zipf's law as demonstration of complex systems dynamics: "demographic distribution of individuals over the earth surface having sharp peaks of population concentration in cities alternating with relatively large spreads, where population

density is much low, observes power law of typical dynamics of complex systems" (Moura et al., 2006).

L. Benguigui and E. Blumenfeld-Lieberthal have developed a dynamic model for measurement of city size using Zipf's law (Benguigui et al., 2007). The given model is based on two factors: Random multiplicative growth and increasing of cities number. Authors have detected such particularities of model as adaptation to various conditions and time effect on city system.

B. Jiang and T. Jia study geo-spatial perspective of justice of Zipf's law in the cities of the USA (Jiang et al., 2011). Results of their searches are verified by the Zipf's law observance within the territory of entire country. Scientists state that cities correspond to power distribution, Zipf coefficient is varied at one.

V. Andreev and V. Lukianova has checked Zipf's law in cities of the Chuvash Republic, which stated implementation of Zipf's law for a set of city districts (Andreev, 2015). If sample will contain Cheboksary city district the Zipf's law is violated.

3. Methodology

Zipf's law relating to the distribution of cities size generally is presented by dependence between city rank (r) and its size (s), sometimes this rule is named as "rank-size". It can be represented by the following formula:

$$s = r^{-1} \quad (1)$$

Corresponding city size *s* will be equal to 1.1/2, 1/3 and etc. If city rank *r* is equal to 1.2.3. The largest city is twice larger the next per rank and tree times larger than the third one and etc. The given equation has a common nature for economic parameters in distribution of profits and sizes of companies (Alperovich, 1984).

Power-based equation of Zipf's law is presented as follows:

$$y = kx^{-\alpha} \quad (2)$$

where *x*- quantity, *k*- constanta, and α- exponent of power law. The given law is known as Pareto distribution. Italian scientist V. Pareto has stated distribution of prosperity among the regions and has determined that it is quite uneven: 20% of population owns 80% of prosperity (wealth), and 80% of population − only 20% of prosperity – rule "rich grows richer".

B. Jiang and T. Jia for determination of power function allows constructing logarithmic graph in the form of straight line (Jiang et al., 2011):

$$\ln(y) = -\alpha \ln(x) + \ln(k) \quad (3)$$

Zipf's law or Pareto distribution are equation of power law. Value of exponent α for power function is determined by method of least square, method of systematic drift is also known. Due to the fact that function of power distribution is not always compared with other types of distribution such as logarithmically normal distribution and exponential distribution, any hypothesis concerning the given law is very hard to adduce.

Probable methods were proposed by M. Goldstein (Goldstein et al., 2004), M. Newman (Newman, 2005) basing on the Kolmogorov-Smirnov test for determination of distribution per power law. These methods are used not only for correspondence to data (or their part) per power law but also for determination insofar as these data comply with other types of distribution. Indicator is specified by the following formula (Jiang et al., 2011):

$$\alpha = 1 + n\left[\sum \ln \frac{x_i}{x_{\min}}\right]^{-1} \quad (4)$$

α means exponent to be valued, x_{min}- minimum value, which distribution per power function obtains. In Zipf's law exponent is equal to one.

Kolmogorov-Smirnov test, modified by A. Clauset (Clauset et al., 2009), allows obtaining maximum correspondence: Cities sizes were in accordance with law on power distribution. Main idea – maximum distance (δ) between data per cumulative function of density and model created:

$$\delta = \max\left|f(x) - g(x)\right|, x > x_{mim} \qquad (5)$$

where *f(x)* - cumulative function of synthetic data with value not less x_{min} , and *g(x)* - cumulative function of power law of distribution which in the best way corresponds to condition $x > x_{min.}$

4. The data

To check Zipf's law in the Russian cities, we have used data of Federal service of government statistics of 2014. Survey object is the Russian cities within federal districts and country as a whole. Sampling has included settlements with city status in 2014. In total the sampling was formed which has included 1,123 cities with population from 1 thous. people to 12108.3 thous. people. The next stage of survey fulfillment of Zipf's law was checked in the Russian cities with population exceeding 100 thous. people. Due to the fact that within the territory of federal district number of cities with population exceeding 100 thous. people. Is insignificant, for economic analysis check of Zipf's law action was performed for the Russia totally.

5. Results

Classification of Russian cities per the population is as follows: small towns – population up to 20 thous. citizens, medium cities – from 20 to 100 thous. citizens, large cities from 100 to 250 thous. citizens, big cities – from 250 thous. to 1 mls. Taking into consideration such a classification, let's review the Russian cities in 2014 (Table 1).

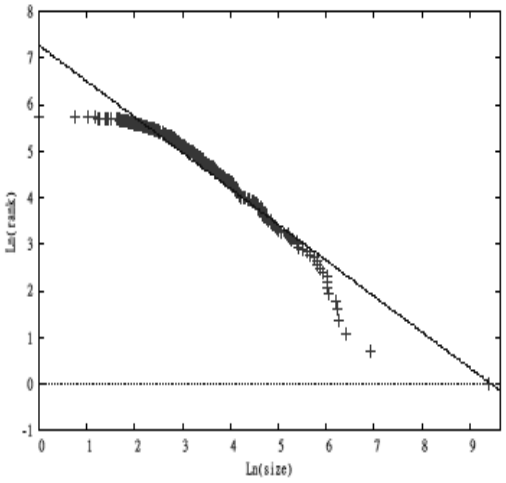
Table 1. Classification of the Russian cities per population in 2014

Federal district	Small towns, up to 20 thous. people		Medium cities, 20 – 100 thous. people		Large cities, 100 – 250 thous. people		Big cities, 250 -1000 thous.people	
	qty of units	Ratio, %	qty of units	Ratio, %	qty of units	Ratio, %	qty of units	Ratio, %
Central	139	45	124	40	27	9	17	6
North-western	84	56	53	36	4	3	7	5
Southern	19	24	43	53	9	12	8	11
North- Caucasian	7	13	35	62	9	16	5	9
Privolzhsky	71	36	95	47	15	8	17	9
Ural	32	23	86	62	11	8	10	7
Siberian	44	34	65	50	11	8	10	8
Far Eastern	30	46	26	39	6	9	4	6
Total per Russia	**426**	**38**	**527**	**47**	**92**	**8**	**78**	**7**

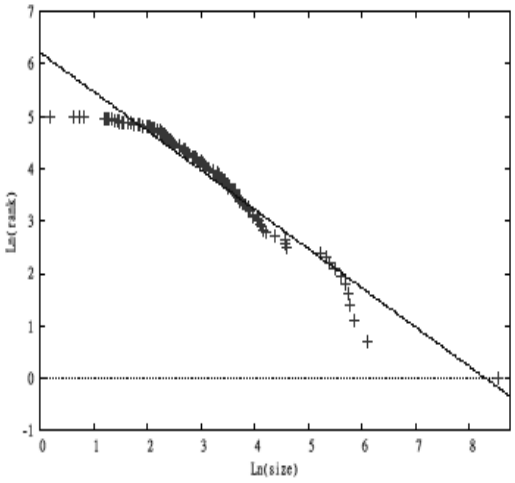
At the Russian territory the most part of cities of the medium size. In some districts small towns are prevailing: In Central federal district − 45% of total number, North-Western federal district − 56 %. As analysis results show average number of small and medium cities exceeds number of large and big ones in five times. For calculation in research method of least squares is used for distribution exponent determination.

Graphs which show regularity rank-size (Zipf's law) in cities at regional and national level in the Russia are indicated in Fig. 1.

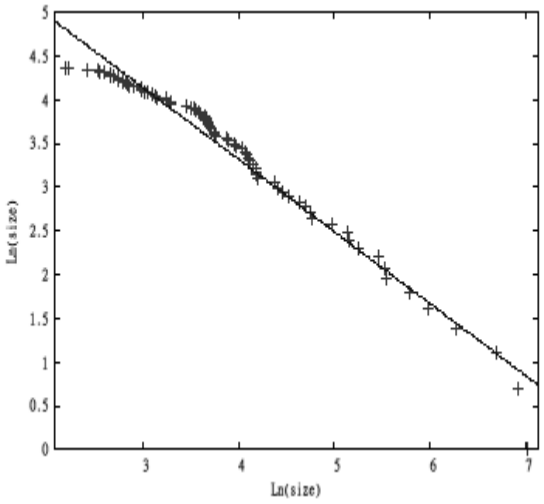
Figure 1 Dependence "rank-size", calculated for the Russian cities in total and per federal districts



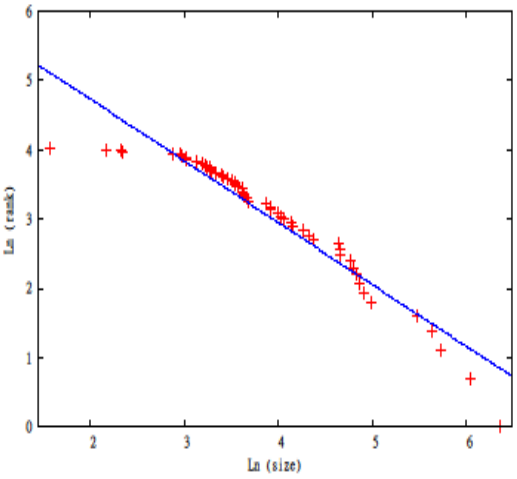
a) Central federal district



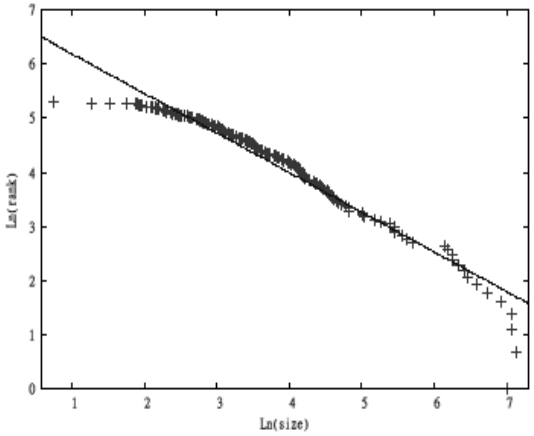
b) North-western federal district



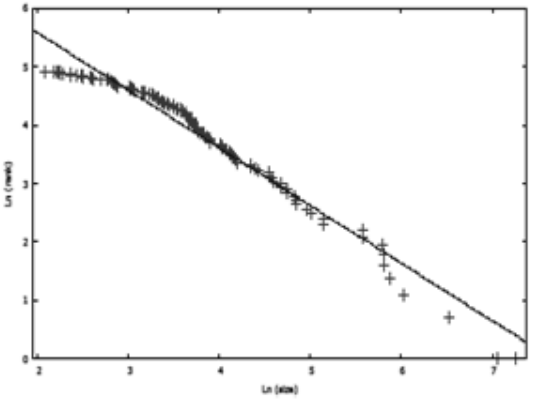
c) Southern federal district



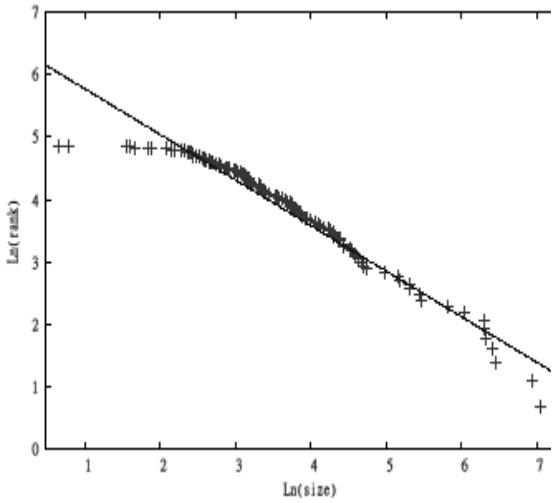
d) North-Caucasian federal district



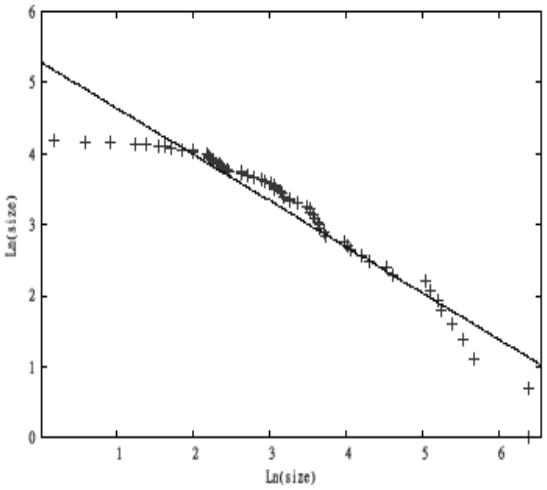
e) Privolzhsky federal district



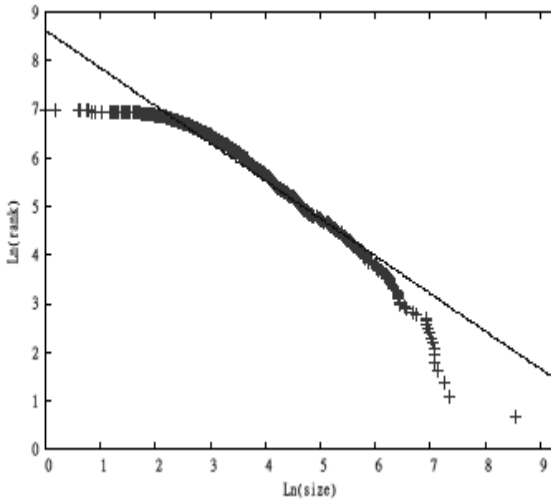
f) Ural federal district



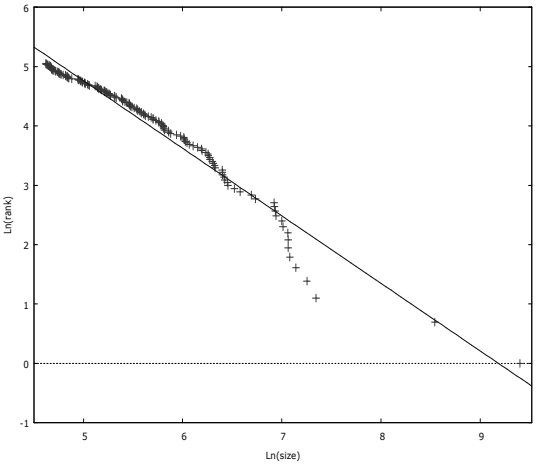
g) Siberian federal district



h) Far Eastern federal district



i) Russian cities, total



g) Russian cities with population exceeding 100 thous. people

Conventional symbols: + - observed; − - linear.

In Central federal district distribution of population corresponds to Zipf's law except for very small towns (1,000-6,000 people) and nine largest cities (408,500-1,014,600 people). It is noticeable that Moscow falls within the scope of Zipf's law. For such cities as Voronezh, Yaroslavl, Ryazan, Lipetsk, Tula, Kursk, Tver, Ivanovo, Bryansk, it is necessary to increase population.

In Northern-Western federal district small towns with population from 1,200 to 5,800 people, large cities – from 187,300 to 448,500 people fall within the scope of Zipf's law. Saint-Petersburg falls within the scope of Zipf's law at that its population can be even less.

In Southern federal district there is a following situation: small towns (9,000-18,000 people) do not correspond to the distribution of Zipf's law. Large cities − Rostov-on-Don (1,109,800 people) and Volgograd (1,018,000 people) − it is reasonable to increase population.

Due to the North-Caucasian federal district it shall be noted that such large cities as Makhachkala (578,300 people), Stavropol (419,800 people), Vladikavkaz (307,300 people), big cities - Pyatigorsk (146,000 people) and Khasavyurt (135,300 people) shall increase their population. Small towns with population from 4,800 people to 10,400 people in North-Caucasian district do not fall within the Zipf's law.

Within the territory of Privolzhsky federal district exception from Zipf's law is small towns with population from 2,100 to 5,800 people, large cities – from 1,096,700 to 1,263,900 people (Nizhny Novgorod, Kazan, Samara, Ufa).

In Ural federal district bright discrepancy to Zipf's law is detected in Ekaterinburg (1,412,300 people) and Chelyabinsk (1,169,400 people). Increasing of population is required for such cities as Tyumen (679,900 people), Magnitogorsk (414,900 people), Nizhny Tagil (357,300 people), Surgut (332,300 people). Population of small towns which do not correspond to rank-size law is varied from 8,000 to 16,100 people.

IN Siberian federal district Zipf's law covers small towns (1,900-8,600 people) and large Novosibirsk (1,547,900 people), Omsk (1,166,100 people), Krasnoyarsk (1,035,500 people), Barnaul (632,800 people), Irkutsk (613,000 people).

At the territory of Far Eastern federal district Zipf's law covers small towns with population from 9,400 to 11,800 people, in medium from 58,800 to 94,300 people, and in large one − Artem (102,400 people). For all the rest cities of the federal district Zipf's law is not implemented.

At check of Zipf's law in total for the Russian cities the following statements were detected. The law covers small towns with population from 8,600 to 15,300 people, large – from 66,700 to 331,000 people.

Basing on the analysis results of cities with population exceeding 100 thous. people. The following conclusion can be made: Zipf's law does not cover cities with population exceeding mln. people (except for Saint-Petersburg).

Results of Zipf's law analysis within the Russian cities are indicated in Table 2.

Table 2. Results of Zipf's law analysis in the Russian cities per federal districts in 2014

Federal district	Qty of cities in federal district, units	Territory area, thous. km^2	Minimum population of cities included in federal district, thous.people	Maximum population of cities included in federal district, thous.people	Coefficient of Zipf for district city, coeff.
Central	307	650	1	12108.3	- 0.76
North-western	148	1677.9	1.2	5132	- 0.7
Southern	79	421.3	9	1109.8	- 0.82
North- Caucasian	56	170	4.8	578	- 0.9
Privolzhsky	198	1037	2.1	1096.7	- 0.73
Ural	139	1789	8	1169.4	- 0.9
Siberian	130	5114.8	1.9	1547.9	- 0.7
Far Eastern	66	6000	1.2	603.2	- 0.65
Total per Russia	1123	16860	1	12108.3	- 0.77

We have studied exposure of Zipf's law separately in each Russian federal district in 2014. This allows determining the following particularities. In total Zipf's law covers the entire territory of Russia. Per districts, Zipf coefficient is varied within the range from –0.65 (Far Eastern federal district) to –0.9 (Ural and North-Caucasian federal districts). We suppose that evenness of hierarchy of cities included in Ural and North-Caucasian federal districts is determined to the fact that in Ural federal district 139 cities are located at the territory of 1789 thous. km^2, in Caucasian - 56 cities − at 170 thous. km^2. In Far East cities are located rare - 66 cities at the territory 6000 thous. km^2 (Zipf coefficient − -0.65).

As a result of the analysis on Russian cities sampling with population exceeding 100 thous. people, we have obtained Zipf coefficient -1.13, that indicates even distribution of cities to be studied at the Russian territory in total. High values of R^2 (app. 0.9) indicate presence of strong binding: "rank-size".

6. Conclusion

In the present article we have checked implementation of Zipf's law at regional and national level in Russia in 2014. Survey has included settlements with "city" status with population in 2014 from 1000 people to 12108.3 thous. people. Individually analysis on sampling of cities with population exceeding 100 thous. people was performed.

We have detected that Zipf's law is implemented within the Russian territory in various degree. Analyzing federal districts separately we have determined that the present law does not cover small towns with population from 1,000 people to 18 thous. people and large from 135.3 thous. people to 5,132 thous. people. It shall be noted that Moscow in Central federal district corresponds to Zipf's law. At check of Zipf's law totally for the Russian cities we have detected that the law covers small (8,600 people -15,300 people) and large (66,700 people – 331,000 people) cities. The most interesting is sampling of cities with population exceeding 100,000 people. In this group Zipf's law does not cover cities with population exceeding 1 mln. people (except for Saint-Petersburg).

Within Russia, Zipf coefficient is varied within the range from -0.64 (Far Eastern federal district) to 0.9 (Ural and North-Caucasian federal districts). At analysis of sampling of cities with population exceeding 100 thous. people Zipf coefficient was - 1.13, which indicates evenness of cities hierarchy in this sampling. Result of the surveys performed was verification of hypothesis upon the presence of dependence of Zipf coefficient on sizes of geographical territory of federal district.

References

1. Alperovich, G. (1984), "The size distribution of cities: On the empirical validity of the rank-size rule". Journal of Urban Economics, 16, 232-239
2. Andreev, V., Lukiyanova, V. (2015), "Proverka zakona Gibrata dlya naselennykh punktov Chuvashskoi respubliki" [Check Gibrat law for the settlements of the Chuvash Republic]. Vestnik Chuvashskogo universiteta = Bulletin of Chuvash University, 1, 154–161.
3. Auerbach, F. (1913), "Das gesetz der bevolkerungskoncentration". Petermanns Geographische Mitteilungen, 59, 74-76.
4. Beckmann, M. (1999), Lectures on location theory. Berlin: Springer
5. Benguigui, L., Blumenfeld-Lieberthal, E. (2007), "A dynamic model for city size distribution beyond Zipf's law". Physica A: Statistical Mechanics and Its Applications,384, 613-627.
6. Brakman, S., Garretsen, H., Marrewijk, C., & Berg, M. (1999), "The Return of Zipf: Towards a Further Understanding of the Rank-Size Distribution". Journal of Regional Science, 31(9), 182-213.
7. Cameron, T. A. (1990), "One-stage structural models to explain city size". Journal of Urban Economics, 27(3), 294-307.
8. Clauset, A., Shalizi, C., Newman, M. (2009), "Power-Law Distributions in Empirical Data". SIAM Rev. SIAM Review, 51, 661-703.
9. Goldstein, M., Morris, S., Yen, G. (2004), "Problems with fitting to the power-law distribution". The European Physical Journal B Eur. Phys. J. B, 41, 255-258.
10. Goula M., Ladias C., Gioti-Papadaki O., Hasanagas N. (2015), "The spatial dimension of environment-related attitudes: does urban or rural origin matter?" Regional Science Inquiry, Vol. VII, (2), pp. 115-129
11. Jiang, B., Jia, T. (2011), "Zipf's law for all the natural cities in the United States: A geospatial perspective". International Journal of Geographical Information Science,25(8), 1269-1281
12. Krugman, P. (1996), "Confronting the Mystery of Urban Hierarchy". Journal of the Japanese and International Economies, 10(4), 399-418

13. Mansury, Y., Gulyás, L. (2007), "The emergence of Zipf's Law in a system of cities: An agent-based simulation approach". Journal of Economic Dynamics and Control, 31, 2438-2460.

14. Moura, N., Ribeiro, M. (2006), "Zipf's law for Brazilian cities". Physica A: Statistical Mechanics and Its Applications, 367, 441-448

15. Newman, M. (2005), "Power laws, Pareto distributions and Zipf's law". Contemporary Physics, 46(5), 323-351.

16. Rastvortseva S.N., Chentsova A.S. (2015), "Regional specialization and geographical concentration of industry in Russia". Regional Science Inquiry, Vol. VII, (2), pp. 97-106

17. Xu, Z., Harriss, R. (2010), "A Spatial and Temporal Autocorrelated Growth Model for City Rank--Size Distribution". Urban Studies, 47(2), 321-335