METHODOLOGICAL ASPECT OF STATISTICAL ESTIMATION AND INTER-REGIONAL DIFFERENTIATION OF SOCIAL SAFETY

© 2009 T.A. Fomicheva*

Keywords: social security, statistical evaluation, integral index of social security, social stability of the regions.

The statistical criterion, allowing to estimate the level of social safety on the territories is presented in the article. The possibilities of objective estimation of social safety are offered in the work.

The problems of social and economic safety of Russia and all its regions (though in different degree) were especially sharply designated in the light of last changes in the economy of our country. Social safety is considered as such level of development and satisfaction of social requirements of population at which its optimum existence and reproduction is provided. From the point of view of statistical analysis the following classification of social safety components: ecological, demographic, economic (including incomes, expenses and population consumption), political and public is represented.

In view of generalizing social safety indicator absence expressed in certain units of measure, its preferable substitute is a multidimensional average arithmetic, calculated by V.M. Ryabtsev's adaptive method. The choice of this kind of average is caused by the fact that investigated indicators are additive as before carrying out rationing procedure (see the formula 1) they had absolute values.

$$z = \frac{x - x_{\min}}{x_{\max} - x_{\min}}. \qquad (1)$$

Therefore, these signs can be aggregated with each other after their change only in the form of average arithmetic, but not in the geometrical one.

The resulted mathematical expression (1) is applicable in that case, when between x and (the integrated indicator) there is a direct connection. In case of their inverse relationship is calculated in the following way:

$$z = \frac{x_{\max} - x}{x_{\max} - x_{\min}}. \qquad (2)$$

Further y is established as the simple average arithmetic:

$$y = \frac{\sum z}{n}, \qquad (3)$$

Then pair correlation coefficients r_{yz} are defined under the formula:

$$r_{yz} = \frac{\overline{yz} - \bar{y} \cdot \bar{z}}{\sigma_y \cdot \sigma_z}, \qquad (4)$$

At the following stage the generalizing indicator is calculated by the scheme of the weighed average arithmetic finding:

$$y = \sum z \cdot w \quad , \qquad (5)$$

where w - the weight, received from the following ratio:

$$w = \frac{|r_{yz}|}{\sum |r_{yz}|}. \qquad (6)$$

Procedure repeats before practically full coincidence of corresponding pair correlation coefficients calculated at last and previous stages.

Values of conditionally accepted integrated indicator of social safety y have been established by us in the above-mentioned way in 14 regions of PFD (Privolzhsky federal district) during the period 2002 - 2007 (table 1).

In compliance with table 1 data, in PFD the lowest degree of social safety has been noted in 2006 in the Ulyanovsk region (the value y was 0,3157), the highest - in Republic of Tatarstan in 2005 (0,7172). The value corresponded to the average level of social safety. In 2002 Saratov region and in 2003 Republic of Bashkortostan had the most similar to the above-mentioned average values (0,4934 and 0,4931).

* Tatiana A. Fomicheva, postgraduate student of Saratov State Social and Economic University. E-mail: fomichevata@rambler.ru.

Table 1

The values of the generalizing indicator calculated by the adaptive V.M. Ryabtsev's method in PFD regions for the period 2002 - 2007

The region	Level y					
	2002	2003	2004	2005	2006	2007
Republic of Bashkortostan	0,4573	0,4931	0,4617	0,4912	0,4109	0,4492
Republic of Mary El	0,3809	0,3883	0,4105	0,4689	0,3693	0,4460
Republic of Mordoviya	0,4394	0,4549	0,4410	0,5086	0,4641	0,5485
Republic of Tatarstan	0,6091	0,6220	0,6588	0,7172	0,6642	0,7118
Udmurt Republic	0,5553	0,5755	0,5856	0,5436	0,5872	0,5806
Chuvash Republic	0,4325	0,4648	0,4721	0,5430	0,5338	0,5531
Perm territory	0,5822	0,5342	0,5107	0,4002	0,5183	0,4071
Kirov region	0,5781	0,4958	0,5177	0,4960	0,5209	0,5084
Nizhniy Novgorod region	0,6114	0,5724	0,5541	0,5479	0,6085	0,6038
Orenburg region	0,4165	0,3992	0,4125	0,3368	0,3983	0,3649
Penza region	0,3197	0,3455	0,3251	0,3360	0,3359	0,3671
Samara region	0,6910	0,6617	0,6644	0,6401	0,6670	0,6383
Saratov region	0,4934	0,4699	0,4515	0,4739	0,4144	0,4186
Ulyanovsk region	0,4292	0,3961	0,3699	0,3674	0,3157	0,4071

Definition of the Russian average level of considered social safety indicator ($\tilde{y}$) would demand enough of the considerable time and labour expenses, that is why the following decision of the given problem is represented to be the most rational.

We suppose that available levels of the analyzed indicator can be considered as a sample of normally distributed general totality.

On the basis of social safety integrated indicator levels for six years (2002 - 2007) in 14 PFD regions (see table 1) according to Sterdgess's formula we defined the quantity of groups:

(7)

where $N = 84$.

According to the found number of intervals (=7), we had systematized the initial data in table 2.

As a result of established average arithmetic, median and modal values (see formulas 8 - 10) it is possible to draw a conclusion about the proximity of y levels distribution to the normal.

, (8)

where $\bar{y}$ - value of sample average; y_i - value of an interval middle ($i = \overline{1,7}$); m - corresponding frequencies.

(9)

Table 2

Grouping of social safety generalizing indicator values in PFD regions for the period 2002 - 2007

Groups of integrated indicator values	Frequencies	Cumulative frequencies
0,3157 - 0,3731	12	12
0,3731 - 0,4305	17	29
0,4305 - 0,4879	13	42
0,4879 - 0,5453	15	57
0,5453 - 0,6027	12	69
0,6027 - 0,6601	9	78
0,6601 - 0,7175	6	84

where M_e - median; y_0 - lower limit of the median interval; h - value of the median interval; m - the frequencies; - the sum of cumulative frequencies of the premedian interval; m_{M_e} - frequency of the median interval.

$$M_o = y_0 + h \cdot \frac{d_1}{d_1 + d_2} = 0{,}3731 + 0{,}0574 \cdot \frac{5}{5+4} = 0{,}4050, \quad (10)$$

$$d_1 = m_{M_o} - m_{M_{o-1}};\ d_2 = m_{M_o} - m_{M_{o+1}},$$

where M_o - moda; y_0 - lower limit of the modal interval; h - value of the modal interval; d_1, d_2 - the intermediate calculated values; m_{M_o} - the frequency of the modal interval; $m_{M_{o-1}}$ - the frequency of the premodal interval; $m_{M_{o+1}}$ - the frequency of the aftermodal interval.

In this case at the following investigation stage it is necessary to define the limits of a confident interval with a general average.

$P(0{,}4027 \le \tilde{y} \le 0{,}4927) = 0{,}99994,$

$P(0{,}4511 \le \tilde{y} \le 0{,}5343) = 0{,}99994.$

The calculation of the sample average error at nonrepeated sampling is carried out in the following way:

$$\mu_y = \sqrt{\frac{\sigma^2}{n} \cdot (1 - \frac{n}{N})}, \quad (11)$$

$$\sigma^2 = \frac{\sum (y_i - \bar{y})^2 \cdot m}{\sum m} = \frac{0{,}2640 + 0{,}1411 + 0{,}0143 + 0{,}0090 + 0{,}0792 + 0{,}1728 + 0{,}2310}{84} = 0{,}0109, \quad (12)$$

$$\mu_y = \sqrt{\frac{0{,}0109}{84}(1 - \frac{84}{528})} = 0{,}0104, \quad (13)$$

where μ_y - the average error of sample at nonrepeated sampling; σ^2 - variance; y_i - the value of the interval middle ($i = \overline{1,7}$); $\bar{y}$ - the value of the sample average; m - the frequencies; $N = 528$, as the number of investigated indicator values is equal to the product of its levels in 88 regions (all Russia) for 6 years.

At probability, we receive the confidential interval of general average $\tilde{y}$:

(14)

So, the following regions are corresponded to Russian average value of the social safety aggregated indicator.

It is visible from table 3 data, that PFD regions with the most typical for Russia considered indicator level are: Republic of Bashkortostan, Kirov and Saratov regions.

Along with the problem of Russian average value estimation there is a principle question in its calculation of the minimum and maximum possible values.

In this connection one of the decisions of the stated problem is the "artificial" introduc-

Table 3

Values of social safety integrated indicator in PFD regions for the period 2002 - 2007 similar to Russian average level

The region	2002	2003	2004	2005	2006	2007
Republic of Bashkortostan	0,4573	0,4931	0,4617	0,4912	-	-
Republic of Mary El	-	-	-	0,4689	-	-
Republic of Mordovia	-	0,4549	-	0,5086	0,4641	-
Udmurt Republic	-	-	-	-	-	-
Chuvash Republic	-	0,4648	0,4721	-	0,5338	-
Perm region	-	0,5342	0,5107	-	0,5183	-
Kirov region	-	0,4958	0,5177	0,4960	0,5209	0,5084
Saratov region	0,4934	0,4699	0,4515	0,4739	-	-

The note: the values which are not in confidential interval, are noted by symbol "-".

tion of two regions (one - with the "best", another - with the "worst" parameters of social safety) for finding the lowest and highest permissible levels of the aggregated indicator y in the regions by V.M. Ryabtsev's adaptive method. The first point in algorithm of the adaptive method is rationing initial private signs. Accordingly, all " " in case of the analysis of the region with the "best" parameters have appeared equal to 1, with "worst" - to 0. Further according to the formula (3) if every , the rating indicator was defined also equal to 1, if every $z = 0$, then $y = 0$. The next step is the reception of pair correlation coefficients by formula (4). As mean-square deviation of constant value is equal to zero, in the situation when all z are equal, . Thus, calculation r_{zy} is impracticable because of its denominator equal to zero. Hence, the impossibility of the given variant application of the boundary values establishment is obvious.

So the question of finding the maximum and minimum levels of the social safety investigated indicator is solved. As the affinity of considered distribution is similar to normal, so the rule of three sigma applies on it, according to which it is practically authentic, that all values of the random variable X, having the normal law of distribution, are concluded in the interval $(a - 3\sigma, a + 3\sigma)$.

The average distribution value is unknown, therefore as "a" the lower limit of a confidential interval is taken, that is 0,4511 and the top limit (0,5343) for calculation of the least and greatest possible levels of social safety integrated indicator (see the formula 15).

So,

$$(a - 3\sigma, a + 3\sigma),$$

$$(0{,}4511 - 3 \cdot 0{,}1044, \quad , \quad , 0{,}8475), \qquad (15)$$

where the borders of the last interval estimate the limits of practical values of the social safety generalizing indicator.

At monitoring regional rating the interval analysis is preferable which allows to classify more exactly its essential and insignificant changes. According to the rule of three sigma the following gradation of the degree of social safety is represented logically (table 4).

This principle allows to define the weight of the change of aggregated indicator levels.

The following division of the fields of the quality of objects functioning is proved.

As we see, the classification offered by us (table 4) does not contradict the above-mentioned characteristic resulted in table 5.

Thus, the regions, where in the period from 2002 to 2007 social intensity was observed, are: Republic of Bashkortostan, Mary El, Mordovia, as well as Orenburg, Penza, Saratov and Ulyanovsk regions. Socially stable for the same period are Republic of Tatarstan, Udmurtiya, Chuvashiya, Perm region, as well as Kirov, Nizhniy Novgorod and Samara regions.

The distinctive feature of the presented statistical social safety criterion is the application

Table 4

Classification of social safety levels

The intervals of the aggregated social safety indicator values	The characteristic of the levels
0,1 - 0,3	social crisis
0,3 - 0,5	social intensity
0,5 - 0,7	social stability
0,7 - 0,9	social well-being

Table 5

Classification of the fields of objects functioning quality

Intervals of indication values	The characteristics of levels
0,000 - 0,167	critically low
0,167 - 0,333	low
0,333 - 0,500	satisfactory
0,500 - 0,667	good enough
0,667 - 0,833	high
0,833 - 1,000	the highest

of V.M. Ryabtsev adaptive method for calculating social safety integrated indicator which advantages are the definition of standardized parameters, weights proceeding from mathematical calculations, and also the possibility of expanding the investigated signs range and their change depending on time requirements.

Rajzberg V.A., Lozovskij L.S., Starodubtseva E.B. The Modern economic dictionary. M., 2006.

Ryabtsev V.M. Competitiveness of the Russian regions: methodology of an estimation and the comparative analysis. Samara, 2002.

The statistics theory: the Textbook / Under the editorship of prof. R.A.Shmojlovoj. 4th issue. M., 2006.

Prokofiev V.A. Methodological questions of formation of rating indicators of objects // Balance. 2006.

Received for publication on 12.08.2009