STATISTICAL APPROACHES BETWEEN THE NUMBER OF INHABITANTS PER DOCTOR AND THE LIFE EXPECTANCY, IN ROMANIA

Gabriela Opait
gabriela.opait@ugal.ro
Dunarea de Jos University of Galați

This article expresses an approache concerning the intensity of the correlation between the number of inhabitants per doctor and the life expectancy in Romania, through by means of the correlation report. The number of inhabitants per doctor is a statistical indicator which reflects the number of inhabitants who are ascribed at one doctor in medicine, indicator which is used in any country from the world.

Key words: life expectancy, number of inhabitants per doctor, correlation, intensity.

JEL Code: C1, C12, C2.

1. Introduction

This research reflects a statistical analysis regarding the intensity of the correlation between the Number of inhabitants per doctor and the Life Expectancy, in Romania, in the period 2007-2013. The purpose of the research consists in to reflect the existence of the correlation between the decrease of the Life Expectancy in Romania and the Number of inhabitants per doctor. The statistical methods applied are the „Coefficients of Variation Method", the „Least Squares Method", used for to calculate the parameters of the regression equation and the Correlation Report Method. The sections 2 presents the methodology of the „Least Squares Method" for to achieve the trend model between the Number of inhabitants per doctor and the Life Expectancy in Romania, in the period 2007-2013, in the view of the calculation of the correlation report. The state of the art in this domain is represented by the research belongs to Carl Friederich Gauss, who created the „Least Squares Method" [1].

2. The modeling of the trend between the Number of inhabitants per doctor and the Life Expectancy in Romania, in the period 2007-2013.

In the period 2007-2013, we observe the next evolution concerning the Number of inhabitants per doctor and the Life Expectancy in Romania, according to the table no. 1:

Table no. 1 The values regarding the Number of inhabitants per doctor and the Life Expectancy in Romania, between 2007-2013

YEARS	THE NUMBER OF INHABITANTS PER DOCTOR IN ROMANIA	THE LIFE EXPECTANCY IN ROMANIA (YEARS)
2007	433	73,30
2008	409	73,50
2009	404	73,60
2010	388	73,80
2011	383	74,60
2012	374	74,50
2013	385	74,45

Source: dates debated from „Romania in dates 2014", page no. 47, I.N.S.S.E., Bucharest.

First, we want to identify the trend model between the **Number of inhabitants per doctor** and the **Life Expectancy**, in Romania, in the period 2007-2013, using the table no.1.

- if we formulate the null hypothesis H_0 *:* which mentions the assumption of the existence for the model of tendency of the correlation between *Y factor = Life Expectancy* and *X factor = Number of inhabitants per doctor*, as being the function $y_{x_i} = a + b \cdot x_i$, then the parameters *a* and *b* of the adjusted linear function, can to be calculated by means of the next system [1]:

$$\begin{cases} n \cdot a + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ a\sum_{i=1}^{n} x_i + b \cdot \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \cdot y_i \end{cases}$$

Thus, $$a = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$ and $$b = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

Table no. 2 The estimate of the value for the variation coefficient in the case of the adjusted linear function, in the hypothesis concerning the linear evolution of the correlation between the Number of inhabitants per doctor and the Life Expectancy in Romania, in the period 2007-2013

YEARS	THE NUMBER OF INHABITANTS PER DOCTOR IN ROMANIA (x_i)	THE LIFE EXPEC-TANCY IN ROMANIA (y_i)	*LINEAR TREND*			
			x_i^2	$x_i y_i$	$y_{x_i} = a + bx_i$	$\lvert y_i - y_{x_i}\rvert$
2007	433	73,30	187489	31738,9	73,1034194919	0,20
2008	409	73,50	167281	30061,5	73,6705784223	0,17
2009	404	73,60	163216	29734,4	73,7887365328	0,19
2010	388	73,80	150544	28634,4	74,1668424864	0,37
2011	383	74,60	146689	28571,8	74,2850005969	0,32
2012	374	74,50	139876	27863,0	74,4976851958	0,00
2013	385	74,45	148225	28663,25	74,2377373527	0,21
TOTAL	2776	517,75	1103320	205267,25	517,750000000	1,46

If we calculate the statistical data for to adjust the linear function, we obtain for the parameters *a* and *b* the values:

$$a = \frac{1103320 \cdot 517{,}75 - 2776 \cdot 205267{,}25}{7 \cdot 1103320 - (2776)^2} = 83{,}3359118612$$

$$b = \frac{7 \cdot 205267{,}25 - 2776 \cdot 517{,}75}{7 \cdot 1103320 - (2776)^2} = -0{,}0236316221$$

Hence, the coefficient of variation for the adjusted linear function is:

$$v_I = \left[\frac{\sum_{i=1}^{n} \left| y_i - y_{x_i}^I \right|}{n} : \frac{\sum_{i=1}^{n} y_i}{n} \right] \cdot 100 = \frac{\sum_{i=1}^{n} \left| y_i - y_{x_i}^I \right|}{\sum_{i=1}^{n} y_i} \cdot 100 = \frac{1,46}{517,75} \cdot 100 = 0,2819893\%$$

- in the situation of the alternative hypothesis H_1 *: which specifies the supposition of the existence for the model of tendency of the correlation between *Y factor = Life Expectancy* and *X factor = Number of inhabitans per doctor*, as being the quadratic function $y_{x_i} = a + b \cdot x_i + cx_i^2$, the parameters *a*, *b* şi *c* of the adjusted quadratic function, can to be calculated by means of the system:

$$\begin{cases} n \cdot a + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i \\ a\sum_{i=1}^{n} x_i + b \cdot \sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i^3 = \sum_{i=1}^{n} x_i \cdot y_i \\ a \cdot \sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^4 = \sum_{i=1}^{n} x_i^2 \cdot y_i \end{cases}$$

Table no. 3 The estimate of the value for the variation coefficient in the case of the adjusted quadratic function, in the hypothesis concerning the parabolic evolution of the correlation between the Number of inhabitants per doctor and the Life Expectancy in Romania, in the period 2007-2013

YEARS	THE NUMBER OF INHABI-TANTS PER DOCTOR IN ROMANIA (x_i)	THE LIFE EXPEC-TANCY IN ROMA-NIA (y_i)	*PARABOLIC TREND*					
			x_i^2	x_i^3	x_i^4	$x_i^2 \cdot y_i$	$y_{x_i} = a + bx_i + cx_i^2$	$\left| y_i - y_{x_i} \right|$
2007	433	73,30	187489	81182737	35152125121	13742943,7	73,3896198170	0,09
2008	409	73,50	167281	68417929	27982932961	12295153,5	73,4422512470	0,06
2009	404	73,60	163216	65939264	26639462656	12012697,6	73,5427829320	0,06
2010	388	73,80	150544	58411072	22663495936	11110147,2	74,0720322314	0,27
2011	383	74,60	146689	56181887	21517662721	10942999,4	74,3022813570	0,30
2012	374	74,50	139876	52313624	19565295376	10420762,0	74,7945602500	0,29
2013	385	74,45	148225	57066625	21970650625	11035351,25	74,2064754940	0,24
TOTAL	2776	517,75	1103320	439513138	175491626396	81560054,65	517,75000	1,31

If we calculate the statistical data for to adjust the quadratic function, we obtain for the parameters *a, b* and *c* the next values

$$\begin{cases} 7 \cdot a + 2776 \cdot b + 1103320 \cdot c = 517,75 \\ 2776 \cdot a + 1103320 \cdot b + 439513138 \cdot c = 205267,25 \\ 1103320 \cdot a + 439513138 \cdot b + 175491625396 \cdot c = 81560054,65 \end{cases}$$

$$a = \frac{\begin{vmatrix} 517{,}75 & 2776 & 1103320 \\ 205267{,}25 & 1103320 & 439513138 \\ 81560054{,}65 & 439513138 & 175491625396 \end{vmatrix}}{\begin{vmatrix} 7 & 2776 & 1103320 \\ 2776 & 1103320 & 439513138 \\ 1103320 & 439513138 & 175491625396 \end{vmatrix}} = 183{,}7323673169$$

$$b = \frac{\begin{vmatrix} 7 & 517{,}75 & 1103320 \\ 2776 & 205267{,}25 & 439513138 \\ 1103320 & 81560054{,}65 & 175491625396 \end{vmatrix}}{\begin{vmatrix} 7 & 2776 & 1103320 \\ 2776 & 1103320 & 439513138 \\ 1103320 & 439513138 & 175491625396 \end{vmatrix}} = -0{,}5222981444$$

$$c = \frac{\begin{vmatrix} 7 & 2776 & 517{,}75 \\ 2776 & 1103320 & 205267{,}25 \\ 1103320 & 439513138 & 81560054{,}65 \end{vmatrix}}{\begin{vmatrix} 7 & 2776 & 1103320 \\ 2776 & 1103320 & 439513138 \\ 1103320 & 439513138 & 175491625396 \end{vmatrix}} = 0{,}0006177021$$

So, the coefficient of variation for the adjusted quadratic function has the value:

$$v_{II} = \left[\frac{\sum_{i=1}^{n} \left| y_i - y_{x_i}^{II} \right|}{n} : \frac{\sum_{i=1}^{n} y_i}{n} \right] \cdot 100 = \frac{\sum_{i=1}^{n} \left| y_i - y_{x_i}^{II} \right|}{\sum_{i=1}^{n} y_i} \cdot 100 = \frac{1{,}31}{517{,}75} \cdot 100 = 0{,}25\%$$

- in the case of the alternative hypothesis H_2 *:* which describes the assumption of the existence for the model of tendency of the of the correlation between *Y factor = Life Expectancy* and *X factor = Number of inhabitants per doctor*, as being the exponential function $y_{x_i} = ab^{x_i}$, then the parameters *a* and *b* of the adjusted exponential function, can to be calculated by means of the next system:

$$\begin{cases} n \cdot \lg a + \lg b \cdot \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \lg y_i \\ \lg a \cdot \sum_{i=1}^{n} x_i + \lg b \cdot \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \cdot \lg y_i \end{cases}$$

Table no. 4 The estimate of the value for the variation coefficient in the case of the adjusted quadratic function, in the hypothesis concerning the exponential evolution of the correlation between the Number of inhabitants per doctor and the Life Expectancy in Romania, in the period 2007-2013

YEARS	THE NUMBER OF INHABITANTS PER DOCTOR IN ROMANIA (x_i)	THE LIFE EXPECTANCY IN ROMANIA (y_i)	*EXPONENTIAL TREND*				
			$\lg y_i$	$x_i \lg y_i$	$\lg y_{x_i} = \lg a + x_i \cdot \lg b$	$y_{x_i} = ab^{x_i}$	$\lvert y_i - y_{x_i} \rvert$
2007	433	73,30	1,865103975	807,5900210	1,863954290	73,10621343	0,19
2008	409	73,50	1,866287339	763,3115217	1,867286402	73,66927604	0,17
2009	404	73,60	1,866877814	754,2186370	1,867980592	73,78712550	0,19
2010	388	73,80	1,868056362	724,8058684	1,870202000	74,16551213	0,37
2011	383	74,60	1,872738827	717,2589709	1,870896190	74,28415542	0,32
2012	374	74,50	1,872156273	700,1864460	1,872145732	74,49819183	0,00
2013	385	74,45	1,871864702	720,6679103	1,870618514	74,23667535	0,21
TOTAL	2776	517,75	13,08308529	5188,039375			1,45

Consequently, if we calculate the statistical data for to adjust the exponential function, we obtain for the parameters *a* and *b* the values:

$$\begin{cases} 7 \cdot \lg a + 2776 \cdot \lg b \cdot = 13,08308529 \\ 2776 \cdot \lg a + 1103320 \cdot \lg b = 5188,039375 \end{cases}$$

$$\lg a = \frac{\begin{vmatrix} 13,08308529 & 2776 \\ 5188,039375 & 1103320 \end{vmatrix}}{\begin{vmatrix} 7 & 2776 \\ 2776 & 1103320 \end{vmatrix}} = 1,924071144$$

$$\lg b = \frac{\begin{vmatrix} 7 & 13,08308529 \\ 2776 & 5188,039375 \end{vmatrix}}{\begin{vmatrix} 7 & 2776 \\ 2776 & 1103320 \end{vmatrix}} = -0,000138838$$

Accordingly, the coefficient of variation for the adjusted exponential function has the next value:

$$v_{\exp} = \left[\frac{\sum_{i=1}^{n} \left| y_i - y_{x_i}^{\exp} \right|}{n} : \frac{\sum_{i=1}^{n} y_i}{n} \right] \cdot 100 = \frac{\sum_{i=1}^{n} \left| y_i - y_{x_i}^{\exp} \right|}{\sum_{i=1}^{n} y_i} \cdot 100 = \frac{1,45}{517,75} \cdot 100 = 0,2800579\%$$

We apply the coefficients of variation method as criterion of selection for the best model of trend.
We notice that:

$$v_{II} = 0,25\% < v_{\exp} = 0,2800579\% < v_I = 0,2819893\%$$

***So, the path followed by Y factor is a parabolical trend** of the shape* $y_{x_i} = a + b \cdot x_i + cx_i^2$*, with other words it confirms the hypothesis* H_1 .

For to reflect the intensity of the parabolical correlation between the Number of inhabitants per doctor and the Life Expectancy in Romania, in the period 2007-1013 we use the correlation report noted with η:

Table no. 5 The calculation of the value for the correlation report in the case of the parabolical correlation between the Number of inhabitants per doctor and the Life Expectancy in Romania, in the period 2007-2013

YEARS	THE NUMBER OF INHABITANTS PER DOCTOR IN ROMANIA (x_i)	THE LIFE EXPECTANCY IN ROMANIA (y_i)	$x_i^2 y_i$	$(y_i)^2$	$x_i y_i$
2007	433	73,30	13742943,7	5372,89	31738,9
2008	409	73,50	12295153,5	5402,25	30061,5
2009	404	73,60	12012697,6	5416,96	29734,4
2010	388	73,80	11110147,2	5446,44	28634,4
2011	383	74,60	10942999,4	5565,16	28571,8
2012	374	74,50	10420762,0	5550,25	27863,0
2013	385	74,45	11035351,25	5542,80	28663,25
TOTAL	2776	517,75	81560054,65	38296,75	205267,25

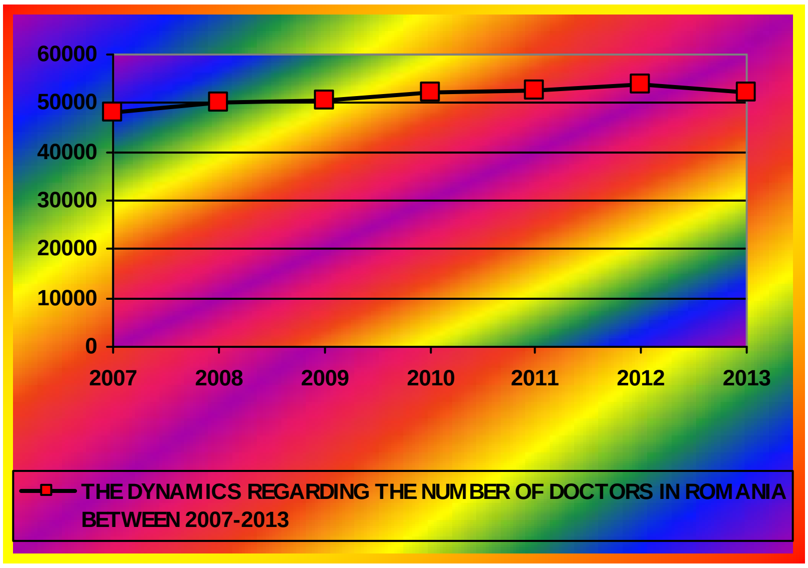
$$\eta = \sqrt{\frac{a\sum_{i=1}^{n} y_i + b\sum_{i=1}^{n} x_i y_i + c\sum_{i=1}^{n} x_i^2 y_i - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}}{\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}}} =$$

$$= \sqrt{\frac{(183{,}7323673169)\cdot 517{,}75 + (-0{,}5222981444)\cdot 205267{,}25 + (0{,}000617702)\cdot 81560054{,}65 - \frac{(517{,}75)^2}{7}}{38296{,}75 - \frac{(517{,}75)^2}{7}}} = 0{,}937249605$$

In conclusion, because the value of the correlation report tends to 1, there is a very strong intensity between the Number of inhabitants per doctor and the Life Expectancy in Romania, in the period 2007-2013.

Table no. 6 The calculation in relative sizes and absolute sizes regarding the dynamics of the doctors in Romania, in the period 2007-2013

YEARS	DOCTORS IN ROMANIA	$I_{t/t-1}^{Doctors}$ (%)	$\Delta_{t/t-1}^{Doctors}$
2007	48.199	-	-
2008	50.267	104,29	2.068
2009	50.386	100,24	119
2010	52.204	103,61	1.818
2011	52.541	100,65	337
2012	53.681	102,17	1.140
2013	51.993	96,86	-1688



Th type no. 1 The dynamics concerning the number of doctors in Romania in the period 2007-2013

4. Conclusions

We can to synthesize that, there is a very strong correlation of parabolical type between the Number of inhabitants per doctor and the Life Expectancy in Romania. Also, the decrease of the Life Expectancy in Romania in 2013 is strong influenced by the decrease of the number of doctors with 3,14 %, respectively 1.688 doctors in 2013 comparative with 2012.

References

[1]. Gauss C. F. - „Theoria Combinationis Observationum Erroribus Minimis Obnoxiae", Apud Henricum Dieterich Publising House, Gottingae, 1823.

[2]. Kariya T., Kurata H. - „Generalized Least Squares", John Wiley&Sons Publishing House, Hoboken, 2004.

[3]. Rao Radhakrishna C., Toutenburg H., Fieger A., Heuman C., Nittner T., Scheid S. - „Linear Models: Least Squares and Alternatives", Springer Series in Statistics, Springer-Verlag New York – Berlin Heidelberg Publishing House, New-York, 1999.

[4]. Wolberg J. - „Data Analysis Using the method of Least Squares: Extracting the Most Information from Experiments", Springer-Verlag New York – Berlin Heidelberg Publishing House, New-York, 2006.