



MODELING SYNCHRONOUS AND ASYNCRHROUNOUS CORRELATIONS THROUGH STATISTICAL METHODS IN ROMANIAN TOURISM

SECAREANU, C[onstantin]; ANDREI, D[aniela] R[uxandra] & DRIDEA STANESCU, C[atrinel] R[aluca]

Abstract: *Tourism represents a social and economic phenomenon, specific for modern civilization. It can be developed like an interdependent economic activity with all arias, having a multiplier role regarding these. Due to its complex content and its dynamic character, it encompasses a large material and human potential, with important implications on society and economic evolution. In this article we intend to identify and quantify some essential correlations between a series of variable statistics of tourism activity.*

Keywords: *tourism, statistical correlation, regression function*

1. INTRODUCTION

Investigating the causal links of dependency between variables which reflect the evolution of social and economic phenomenon's can be done by a series of techniques and methods among which are the statistical correlation and regression. These methods allow various utilizations – including touristic activity – for analysis, theoretic generalizations and creating predictions.

Temporal correlations met in the case of chronological series present a few particular aspects. First of all the type of correlation must be identified:

The internal correlation or autocorrelation, through which a connection is established between the values of a statistical variable registered in a period (y_t) and the values of the same quantified variable for a previous period (y_{t-1}), as an effect of an inertial evolution:

$$y_t = f(y_{t-1}) + \varepsilon \quad (1)$$

Same-time correlation, through which the factorial variable x exerts its influence over the resulting variable y mostly in the same time period t .

$$y_t = f(x_t) + \varepsilon \quad (2)$$

The asynchronous correlation or time-delay correlation, through which the effect takes place at a certain time interval from the cause. Thus, the resulting variable for t period is dependent on the factorial variable registered in period $t-1$.

$$y_t = f(x_{t-1}) + \varepsilon \quad (3)$$

Measuring the probable correlations that should exist between the average number of employees, the existing accommodation capacity on the „hotels” sector, are realized as follows.

2. ANALYSIS OF THE CORRELATION BETWEEN THE AVERAGE NUMBER OF EMPLOYEES AND THE EXISTING ACCOMMODATION CAPACITY

The first identified and measured correlation is the synchronous one that should exist between the average number of employees and the existing accommodation capacity in the “hotels” sector.

The link between the two characteristics is as follows:

$$S = f(CE) + \varepsilon \quad (4)$$

Where:

S = the average number of employees as a resulting variable (effect variable);

CE = the existing accommodation capacity. Factorial variable (causal variable);

ε = residual variable, which quantifies the influences of the other factors considered non-essential.

Year	Average number of employees - thousands of persons -	Existing accommodation capacity - thousands of places -
	y_t	x_t
2004	33.8	166
2005	34.8	170
2006	34.4	173
2007	36.5	174
2008	38.9	182

Tab.1. Evolution of Average number of employees and existing accommodation capacity (Source: The National Statistics Institut, 2010)

To verify the existence, direction and shape of the connection between the two characteristics we use one of the simplest (elementary) methods of quality analysis, for example the graphic method.

Taking into consideration the order of the dots on the graphic we can appreciate that between the two variables there is a direct statistical link and the mathematical function that best fits this connection is the linear function. Thus, with the increase of the factorial variable CE , the resulting variable S also increases with a sole exception in the year 2006.

To measure the dependency between the two variables we utilize the regression method for the single factor linear model:

$$Y_t = a + b \cdot x_t \text{ or } S_t = a + b \cdot CE_t \quad (5)$$

Where a, b = parameters of the regression function

So the regression function that best adjusts the link between the existing accommodation capacity (x_t) and the average number of employees (Y_t adjusted values) is:

$$Y_t = -20.67 + 0.326 \cdot x_t, r = + 0.935 \quad r^2 = 87.4\% \quad (6)$$

The regression coefficient $b = 0.326$ shows by how much the resulting variable Y is changed when the factorial characteristic x is modified by one unit; in our case, by increasing the existing accommodation capacity by 1,000 places, the average number of employees will increase by 0.326 thousand persons, so 326 employees.

The value of the correlation coefficient ($r = + 0.935$) confirms the existence of a high intensity linear correlation between the two variables.

Also, the determination coefficient ($r = 87.4\%$) indicates that other factors not included in the model account for only 12.6%.

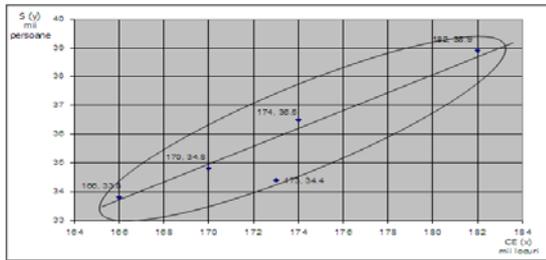


Fig. 1. Link between existing accommodation capacity and the average number of employees

3. ANALYSIS OF THE EXISTING CORRELATION BETWEEN THE EXISTING ACCOMMODATION CAPACITY AND THE VOLUME OF INVESTMENTS

The second studied correlation is the one between the existing accommodation capacity – CE (as a resulting variable) and the volume of investments (INV), as the main factor of increasing the capacity in the hotel sector.

The link between the two characteristics is:

$$CE = f(INV) + \varepsilon \tag{7}$$

Undoubtedly between the two characteristics there should be a very tight statistical connection. Using the same logic in the previous example we should identify a function:

$$CE_t = f(INV_t) \tag{8}$$

Through which the volume of investments in year t influences the change of the accommodation capacity in the same year t. But investments that take place in one year will barely influence the existing accommodation capacity (revised at 31st of July of that year). Usually the accommodation capacity in year t is influenced by the volume of investments in year t-1. Thus we can observe a time-delay correlation:

$$CE_t = f(INV_{t-1}) \tag{9}$$

Year	Existing accommodation capacity	Volume of investments
	- thousands of places - y_t	- millions of lei - x_{t-1}
2005	170	470
2006	173	479
2007	174	557
2008	182	749
2009	185	997

Tab. 2. Evolution of the existing accommodation capacity and the volume of investments, (Source: The National Statistics Institut, 2010)

The order of the dots on the graphic we can appreciate that between the two variables there is a direct statistical link and the mathematical function that best fits this connection is the linear function.

Thus, the increase of the investments volume in year t-1 led to an increase in accommodation capacity in years.

Next we apply the regression method to determine the single factor linear model that quantifies the link between the two variables:

$$Y_t = 158,98 + 0,0274 \cdot x_{t-1} \quad r = 0,962 \quad r^2 = 92,5\% \tag{10}$$

The value of the regression coefficient $b = 0,0274$ shows by how much the resulting variable Y is modified, in case that the factorial characteristic is modified by 1 unit; thus, by increasing the investment volume in year t-1 by 1 million lei, the existing accommodation capacity will increase in the next year (t) by 0.0274 thousand places, so with 27.4 places.

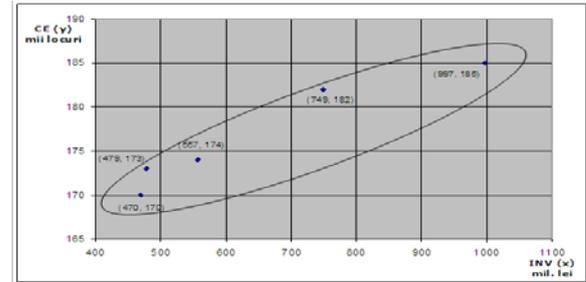


Fig. 2. The link between existing accommodation capacity and volume of investments

At the same time, 92.5% of the accommodation capacity variation is determined by the volume of investments, the other (not included) factors (accommodation structures being shut down, changing the type of activity etc.) having a residual influence of approximately 7.5%. The correlation method gathers in its current state the whole complex of identification for causal links between phenomena and investigation of statistical connections, leading to synthetic numeric representations: regression functions, regression coefficients, correlation coefficients, determination coefficients, estimation errors etc.

4. CONCLUSION

The identified and quantified links between different statistical values specific to the tourism activity offer the possibility of creating unitary analysis and previsions, coherent and correlated, to form the base of a realistic development strategy.

By identifying and quantifying some causal and relatively constant relations between the analyzed variables, the presented models allow creating the basis for predictions regarding the investment volume necessary to increase the accommodation capacity, as well as properly calculating the required work force in this sector. It's also important to mention that the availability of statistical data at more detailed levels (i.e. Types of accommodation structures) would lead to obtaining regression functions with more representative parameters.

5. REFERENCES:

Johnson R., Gouri K .Bhattacharyy(2009), *Statistics: Principles and Methods*, 6th edition, John Wiley and Sons Publishing, VoineaguVirgil, TitanEmilia(2004), *Sondaje si anchete*, EdituraFundatiei,,Andrei Soguna,Constanta
 SecareanuConstantin,GruiescuMihaela,AndreiRuxandra, (2010) *Statistica*, EdituraScrisulRomanesc, Craiova,
 Rezultate si performante ale intreprinderilor de comert si servicii, Institutul National de Statistica, Bucuresti, 2005-2010,
 Anuarul statistic al Romaniei- 2010, Institutul National de Statistica, Bucuresti.