STATISTICAL METHODS USED IN THE ANALYSIS AND FORECAST OF THE TOURISM ACTIVITY AFFECTED BY SEASONALITY

Constantin Secăreanu,
Daniela Firoiu*

Abstract

A lot of economic activities (tourism, commerce, transport, agriculture, constructions etc.) witness a different evolution of the performance indicators throughout the year, as there are seasonal variations (oscillations) on shorter intervals of time (trimesters, months, weeks, even days). These periodic variations, on time lapses shorter than a year, get repeated with relative accuracy from one period to another.

The objective of this paper is to analyze the possibilities of quantification of the phenomenon of tourism seasonality, through the utilization of two methods: the seasonal coefficients and the decomposition of the chronological series to its main components.

Keywords: Tourism, seasonality, statistical methods.

JEL Classification: L83, C10, C16

1. Introduction

In tourism, seasonal variation concretizes in a greater or lesser concentration of tourist flows in certain periods of the year, as a result of the impact of:

- natural factors (the succession of seasons, climate conditions);
- social factors (legal holidays, days off, the structure of the school and university years etc.)

The implications of seasonality on the tourism activity reflect in:

- the overuse or the incomplete use of the material basis and of the workforce;
- the effects on prices, costs and, implicitly, on profitability;
- the effects on the quality of tourism services and, further, on the consumer degree of satisfaction.

In what follows, we will analyze the seasonality of the tourism activity in Romanian mountainous resorts, based on the indicator "number of tourists accommodated in tourism lodging structures". The necessary information for the period 2007-2009, by months, has been taken from the monthly statistical Bulletins published by the National Institute of Statistics (http://www.insse.ro).

* Constantin Secăreanu is Professor at the Romanian American University in Bucharest
e-mail: constantin. secareanu@yahoo.com
Daniela Firoiu is Professor of Tourism Economics at the Romanian American University in Bucharest
e-mail: danafiroiu@yahoo.com

2. The analysis of the seasonality of the tourism activity with the help of seasonality coefficients

To characterize the seasonality of the analyzed phenomenon we use the **seasonality coefficients** (KS_L), which are calculated as percentage ratio between the average level of each month registered on a period of a few years ($\overline{X}_L$) and the general monthly average ($\overline{X}_G$):

$$KS_L = \frac{\overline{X}_L}{X_G} * 100$$

The average of each month over a period of a few years ($\overline{X}_L$) is obtained as an arithmetic mean of the levels registered in the same month over the period (in years) that is being considered.

The general average ($\overline{X}_G$) is calculated as a monthly average of the whole period (in our case three years, respectively thirty-six months). The general average can be obtained in a few ways:

- as an arithmetic mean of the monthly medium levels of each year;
- as an arithmetic mean of the medium levels of each month registered over the three years;
- as a ratio between the total number of tourists registered over the three years and the total number of months of the period.

The interpretation of the seasonal coefficient is the following:

- if $KS_L > 100\%$, then the level registered in month L is above the general monthly medium level of the period, the seasonal factor having a favorable influence;
- if $KS_L < 100\%$, then the level registered in month L is under the general monthly medium level of the period, the seasonal factor having an unfavorable influence;

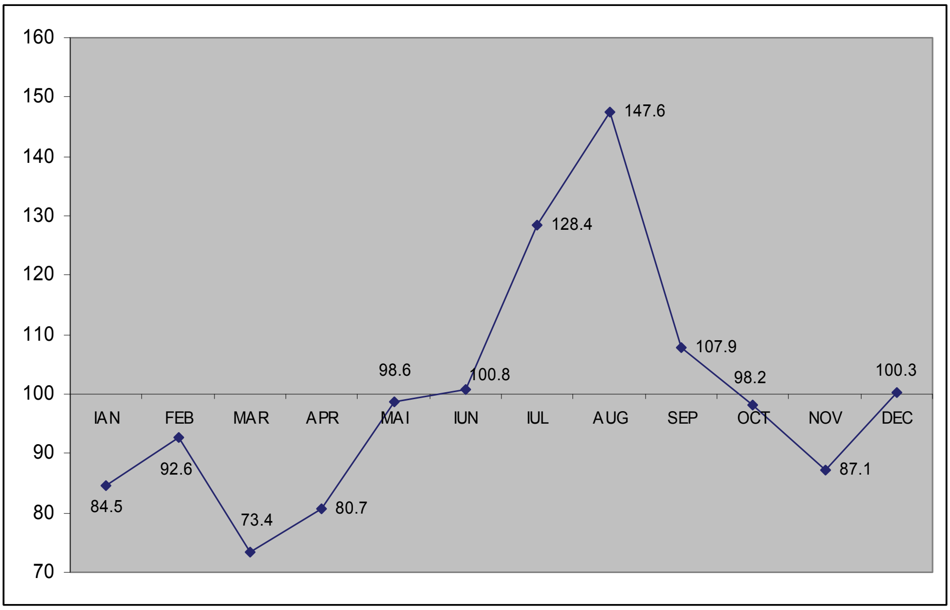
The seasonal coefficients have the property that their sum must be equal to their number (twelve for the monthly analysis, four for the trimestrial analysis), multiplied by a hundred.

In our case:

$$\sum_{L=1}^{12} KS_L = 1.200$$

The signification of the seasonal coefficients results very clearly from their graphical representation:

Figure no.1 The seasonal monthly coefficient for the number of tourists accommodated in lodging units in the mountains



The main conclusions that arise from the analysis of the seasonality of the tourism activity in the mountainous resorts are synthesized as follows:

- the seasonal factor had a favorable influence in the months of August (147.6%), July (128.4%) and September (107.9), so in the third trimester of the year:
- in the months of July and December the number of accommodated tourists was at the level of the general monthly average (100.8%, respectively 100.3%); what is surprising is the reduced influence of the seasonal factor in December;
- the seasonal factor has influenced unfavorably the tourism activity in the other three trimesters, the most reduced levels of the seasonal coefficient being registered in::
  - the first trimester, especially in March (73.4% - the most reduced level) and in January (84.5%),
  - the second trimester, especially in April (80.7%);
  - the fourth trimester, especially in November (87.1%).

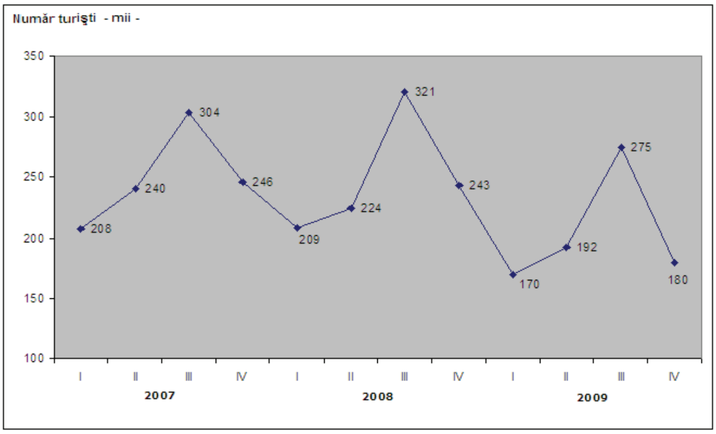
The analysis over a period of three years reflects the fact that at the level of the tourism activity in the mountainous area a series of seasonal periodic variations (monthly and trimestrial) can be identified, variations that get repeated relatively constantly.

The method of the seasonality coefficients offers a synthetic but static image of the periodic oscillations (variations), without achieving a connection between the evolution of the respective phenomenon in time and the influence of the seasonal factors. In order to quantify the seasonal component in a dynamic framework, through which we could identify change as well, the transformation and the development of the phenomenon in time, we resort to the decomposition of the chronological series to its main components.

3. The modeling of the seasonality of the tourism activity based on the analysis of the main components of a chronological series

In order to analyze the seasonality of the number of tourists accommodated in the lodging units of the resorts situated in the mountainous area, the trimestrial data for the period 2007-2009 has been used. The visualization of the seasonal variations is achieved through the graphical representation of this trimestrial chronological series.

Figure no. 2 The evolution of the number of tourists accommodated in the lodging units in the mountainous area



Even without any calculations, we can identify from this chart a series of tendencies:

- the variations are repeated annually, so every four trimesters;
- every year of the period the minimum level is registered in the first trimester, and the maximum level is reached in the third trimester;
- the evolution trend for the whole period 2007-2009 is a reduction of the phenomenon.

The terms of this chronological series show a high degree of variability, according to the action of various factors that are active throughout the period. Consequently, the chronological series can be decomposed in three components (Secăreanu and Gruiescu, 2010, p. 208):

The general tendency or the trend (y_T) constitutes the main component of the evolution line, formed as a result of the action of the essential factors, with permanent influence and which give the phenomenon its direction of development.

A. **Periodic oscillations (variations)** (y_S) that are systematic and repetitive, usually produced under the influence of natural factors – the climate, which manifest every trimester. Sometimes the periodic oscillations manifest as well as an effect of social and organizational factors such as: legal holidays, the regimen of days off, the structure of the school and university years etc.

B. **Accidental deviations** or **the residual component** (y_R) manifest as irregular deviations from the line of systematic evolution, as an effect of the action of some accidental (random) factors.

For the decomposition of the chronological series in the three components, we will resort in what follows to the **additive model** (Voineagu and Țițan, 2004, p. 123), according to which the level of the characteristic (the number of accommodated tourists) each trimester (y_t) equals the sum of the trend (y_T), the seasonal variations (y_S) and the residual component (y_R):

$$y_t = y_T + y_S + y_R$$

A. To determine the long-term tendency (**trend**) we will use the method of **moving averages.**

The moving averages ($\overline{y}_t$) result by progressively replacing the real terms y_t, with partial averages calculated from the terms of the series. The number of terms out of which the moving averages are calculated equals the number of terms at which a complete oscillation occurs. In our case there are twelve terms, corresponding to the four trimesters of the three years. Since a complete oscillation happens annually, thus after four terms (corresponding to the four trimesters), then the moving averages will be calculated out of four terms.

To calculate the moving averages from an even number of terms (four terms) we must go through two stages (see Table no. 1):

a. The temporary moving averages ($\overline{y}_t$) are calculated. The temporary moving averages are situated in-between the two terms with a middle position in the group of the four terms that participate in the calculation of the average, a reason why these must be centered.

b. The centered moving averages ($\overline{\overline{y}}_t$) are calculated as an arithmetic mean of two temporary moving averages.

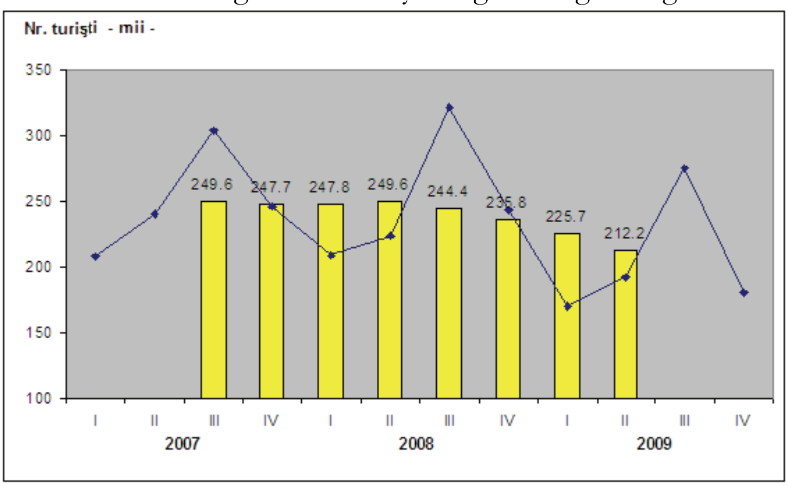
We may notice that determining the tendency through the method of the moving averages leads to a certain loss of information. Thus, for the first and the last terms of the chronological series we cannot calculate adjusted values (moving averages). In this case the real series (y_t) comprises twelve terms, whereas the adjusted series ($\overline{\overline{y}}_t$) is made up of only eight terms.

Table no. 1 The calculation of moving averages (thousands of tourists)

Year	Trimester	Period	Number of tourists y_t	Moving averages: Temporary $\bar{y}_t$	Moving averages: Centered $\bar{\bar{y}}_t$
2007	I	1	208		
	II	2	240		
				249.5	
	III	3	304		249.6
				249.8	
	IV	4	246		247.7
				245.7	
2008	I	5	209		247.8
				250.0	
	II	6	224		249.6
				249.3	
	III	7	321		244.4
				239.5	
	IV	8	243		235.5
				231.5	
2009	I	9	170		225.7
				220.0	
	II	10	192		212.2
				204.3	
	III	11	275		
	IV	12	180		

The following graphical representation clearly illustrates the way in which the moving averages ($\bar{\bar{y}}_t$) allow the calculation of the tendency (trend) under the influence of the objective factors, by eliminating seasonal variation and accidental (residual) factors, ensuring the leveling of the chronological series.

Figure no. 3. Determining the tendency using moving averages



The graphical representation confirms that between 2007-2009 a decrease in the number of tourists accommodated in the lodging units of the mountainous area was registered.

B. **Determining seasonal variations** presupposes going through the following stages:

a. We eliminate from the real values of the chronological series (y_t) the trend component (y_T), determined through the method of moving averages:

$y_t - y_T = y_S + y_R$

where:

$y_T = \overline{\overline{y}}$ (calculated in Table no. 1.)

b. For each trimester (season) we calculate the seasonal deviation as an average of the deviations obtained in the first stage. Through the calculation of the averages we eliminate most of the residual (random) variations.

c. We calculate seasonal deviation through the correction of the averages obtained at the previous stage, taking into account the condition that the sum of all deviations should be null.

The results reflect that the seasonal factor deviates the number of tourists accommodated in the lodging units of the mountainous area as follows:

- In the first and the second trimesters, under the trend line with forty-seven thousand tourists, respectively twenty-two thousand tourists.
- In the third and the fourth trimesters, above the long-term tendency with sixty-six thousand tourists, respectively three thousand tourists

C. **Determining the residual (random) component** – y_R is achieved based on the relation:

$y_R = y_t - y_T - y_S$

The final results synthesize the influence of every component on the evolution of the phenomenon: the trend, the seasonal component and the residual component.

Table no. 2 The components of the chronological series (thousands of tourists)

Year	Trimester	No. of tourists (actual data) y_t	Tendency (centered moving averages) $y_T = \overline{\overline{y}}$	Seasonal component (seasonal deviations) y_S	Residual (random) component $y_R=y_t-y_T-y_S$
2007	I	208	-	-	-
	II	240	-	-	-
	III	304	250	+66	-12
	IV	246	248	+3	-5
2008	I	209	248	-47	+8
	II	224	250	-22	-4
	III	321	244	+66	+11
	IV	243	235	+3	+5
2009	I	170	226	-47	-9
	II	192	212	-22	+2
	III	275	-	-	-
	IV	180	-	-	-

4. Conclusions

The initial information offered an unclear image regarding the evolution of the phenomenon in time: great variations from one trimester to another, more or less explicable variations.

By applying the model we have presented, a series of important conclusions have been identified:

1. The number of tourists accommodated in the mountainous area between 2007-2009 is decreasing: from two hundred and fifty thousand tourists in the third trimester of 2007, up to two hundred and twelve thousand tourists in the second trimester of 2009.

2. The seasonal factor has a different influence: negative in the first and the second trimesters (-47 thousand tourists, respectively -22 thousand tourists) and positive in the third and the fourth trimesters (+66 thousand tourists, respectively + 3 thousand tourists).

3. The residual (random) component had a different influence, between -12 thousand tourists in the third trimester – 2007 and +11 thousand tourists in the third trimester – 2008, the third trimester being the one with the greatest seasonal deviation.

A specific situation is met in the fourth trimester, when – although the absolute values of the residual component are reduced – these have higher values than the seasonal component. Thus, in the fourth trimester – 2007, the seasonal variation is +3 thousand tourists, whereas the random component is -5 thousand tourists; in the fourth trimester – 2008 the increase of the number of tourists by three thousand is due to the seasonal factor, whereas the accidental (random) factors contributed to the increase of the number of tourists by five thousand. Consequently, the seasonal factor determined by the increase of the number of tourists during the winter holiday period was countered by a series of accidental elements (for instance: unfavorable climate conditions, unattractive domestic tourism offers and, implicitly, the orientation of the population on external tourism activities etc.).

Various methods of analysis for seasonality allow not only the diagnosis of past periods, but also the achievement of realistic forecasts on the evolution of the phenomena that are influenced by seasonal variations.

Bibliography

Secareanu, Constantin and Gruiescu Mihaela. 2010. Statistics – Syntheses, tests and applications. Craiova: Scrisul Romanesc publishing house.

Voineagu, Vergil and Titan, Emilia. 2004. Surveys and Inquires. Constanta: „Andrei Saguna" Foundation Press.

***http://www.insse.ro