THE SYNTHETICAL INDICATORS OF VARIABILITY AN APPROACH IN TOURISM

Associate Professor PhD Elisabeta R. ROŞCA
„Stefan cel Mare" University of Suceava, Romania
elisabetar@usv.ro

Abstract:
The paper „The Synthetical Indicators of Variability. An Approach in Tourism" develops the variability of frequencies distributions and the statistical methods of analysis of it. Among these are included the estimation and the analysis of synthetical indicators of variability which are: the average linear deviation, the standard deviation, the dispersion and the coefficient of variation. These indicators measures the degree of variability of individual items of frequencies distributions in comparison with an indicator of central tendency, as a rule the average and sometimes the median. In the analysis of dispersion, the standard deviation is finding the application also in the Chebyshev's theorem and in the empirical rule, enunciated in the paper. In the case of multidimensional distributions the total dispersion results as a sum of dispersions (the group dispersion and the dispersion between groups) under the form of rule of addition the dispersions. The paper presented a concise study of tourism in Romania in 2005, using the dispersion indicators in the analysis of travel duration, in the case of internal travels of the residents for vacations and business, on the touristic zones.

Keywords: variability, standard deviation, dispersion, coefficient of variation, the rule of addition the dispersions.

JEL Classification: C10

INTRODUCTION

The variability in statistics is determined by the influence of essential and random factors on the elements of a population, which according with the intensity of their action determine a bigger or a smaller deviation of the individual values in comparison with the central tendency. As the complexity degree grows the variation of individual values is bigger and the correct using of the indicators of central tendency in the substantiation of decisions needs the verification of the stability and the signification of values registered by these. The estimation and the analysis of the individual values variation from the typical values which characterize the central tendency give the possibility to solve some problems of statistics knowledge extremely useful in the substantiation the decisions among which we include: the analysis of the homogeneity degree of data from which have estimated the central tendency indicators and the verification of the representation of these as typical values of series, the comparison in time and/or in space of more distributions by independent and/or interdependent characteristics, the selecting of the factors with significant influence which determine the structures in the population, the separation of the action of essential factors from the action of random factors, the indentification of the way in which the essential factors change their action from a group to another, the concentration of the individual values of the characteristics and their variability in comparison with typical values, the application of different tests of mathematical statistics.

The analysis of the total degree of individual values variability of the variable in comparison with the indicators of central tendency is materialized in the estimation and the analysis of synthetical indicators of variation in a unidimensional/multidimensional distribution which are: the average linear deviation, the standard deviation, the dispersion, the coefficient of variation and the relationship among dispersions in the case of multidimensional distributions.

THE SYNTHETICAL INDICATORS OF VARIATION IN THE UNIDIMENSIONAL DISTRIBUTIONS

An unidimensional statistical series presents the registered values of one variable in the statistical population or a succession of values for a quantitative/qualitative variable $\{x_i\}, i = 1 \div n$

which characterized a population. The synthetical indicators of variation computed in such kind of distributions are: the average linear deviation, the standard deviation, the dispersion and the coefficient of variation (Biji, E. M., et al., 2000), (Isaic – Maniu, Al., et al., 1999).

The average linear deviation - $\bar{d}_{\bar{x}}$ is computed as a simple or weighted arithmetic average of the deviations the terms of series from their central tendency taken in absolute value. The central tendency can be characterized using the average or the median (Badia, J., et al., 1997), (Korka, M., et al., 2005).

When the deviations of the individual values are computed and analysed in comparison with the average, the average linear deviation is determined with the formula:

- in the case of simple series:

$$\bar{d}_{\bar{x}} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n} \qquad (1)$$

- in the case of grouped data:

$$\bar{d}_{\bar{x}} = \frac{\sum_{i=1}^{k} |x_i - \bar{x}| n_i}{\sum_{i=1}^{k} n_i} \qquad (2)$$

In some statistical analysis is interested the average linear deviation of the individual values from the median $\bar{d}_{M_e}$ (Biji, E. M., et al., 2002). This is determined using the following formula:

- in the case of simple series:

$$\bar{d}_{M_e} = \frac{\sum_{i=1}^{n} |x_i - M_e|}{n} \qquad (3)$$

- in the case of grouped data:

$$\bar{d}_{M_e} = \frac{\sum_{i=1}^{k} |x_i - M_e| n_i}{\sum_{i=1}^{k} n_i} \qquad (4)$$

The average linear deviation presents the disadvantage that it didn't take into account that the greatest deviations in absolute value have a greatest influence on the degree of variation of a variable, in comparison with the smaller deviations. In the same time, from the algebraic point of view, it isn't recommended to give up in an arbitrary way at the sign of values from which is computed an average. It is a less used indicator in the analysis of frequencies distributions because it didn't characterized the theoretical distributions and also because it is expressed in the measure unit of variable and so it isn't useful in the dispersions of variables comparison (Biji, M., Biji, E. M., 1979).

Standard deviation - σ is computed as a square average of the all deviations of the variants of distribution from their arithmetic average.

The formula are:

- in the case of simple series:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}} \qquad (5)$$

- in the case of grouped data:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{k} (x_i - \bar{x})^2 n_i}{\sum_{i=1}^{k} n_i}} \qquad (6)$$

At the standard deviation computed as a square average by the amounting to square it gives more importance to the greatest deviations in absolute value. As a consequence the standard deviation will be always bigger than the average linear deviation, computed for the same data series. It is estimated that for a distribution with clear tendency of normality the average linear deviation is equal with 4/5 of the standard deviation value ($\bar{d}_{\bar{x}} \cong 4/5 \cdot \sigma$).

The standard deviation is a basic indicator which is used in the analysis of variation at the estimation of selection errors, in the calculus of correlation. Near by the average, the standard deviation represented a parameter which defines different types of theoretical distributions and serves especially at the interpretation of the curve of normal distrib ution.

Applied in according with the *Chebyshev's theorem*, the standard deviation becomes useful for the establishing the proportion in which the data of a population or a sample borders in a certain interval of variation. In according with Chebyshev's theo rem, being given any series of data and a number $k \geq 1$, the part from these data which are between k standard deviation of the average is at least $1-\frac{1}{k^2}$. Because this theorem is valid for any distri bution of observed data it is applied so for samples as for populations. The theorem is used to determine the minimum value which represents the part of observed data situated in the interval $(\bar{x}-k\cdot\sigma_W, \bar{x}+k\cdot\sigma_W)$, where σ_w represents the standard deviation in sample. This interval of values, which has as central value the average, contains a number of data so bigger as k is bigger. For example, the interval $(\bar{x}-3\sigma_W, \bar{x}+3\sigma_W)$ contains a number of values bi gger than the interval $(\bar{x}-2\sigma_W, \bar{x}+2\sigma_W)$. In according with the Chebyshev's theorem, this part of data situated in the interval $(\bar{x}-k\cdot\sigma_W, \bar{x}+k\cdot\sigma_W)$ is at least $1-\frac{1}{k^2}$ (Keller, G., et al., 1988).

So, if we applied the theorem for some specific values of k we obtain:

- for $k=1$, the information delivered is unsignificant, that the part of data situated in the interval $(\bar{x}-\sigma_W, \bar{x}+\sigma_W)$ is at least 0;
- for $k=2$ we find at least $1-\frac{1}{k^2}=\frac{3}{4}$ from the data are situated in the interval $(\bar{x}-2\sigma_W, \bar{x}+2\sigma_W)$.

In the Table no. 1 we find the correspondence between some values of k and the proportion in which the observed data are situated in the interval established by Chebyshev's theorem.

Table no. 1. Chebyshev's theorem for variable values of k

k	**Interval**	**Part of observed data included in interval**
1	$(\bar{x}-\sigma_w, \bar{x}+\sigma_w)$	at least 0
2	$(\bar{x}-2\sigma_w, \bar{x}+2\sigma_w)$	at least $\frac{3}{4}$
2,5	$(\bar{x}-2{,}5\sigma_w, \bar{x}+2{,}5\sigma_w)$	at least $\frac{21}{25}$
3	$(\bar{x}-3\sigma_w, \bar{x}+3\sigma_w)$	at least $\frac{8}{9}$

Source: Keller, G., Warrack, B., Bartel, H., 1988, *Statistics for Management and Economics. A Systematic Approa ch*, Wadsworth Publishing Company, Belmont, California, 58 -59

The importance of Chebyshev's theorem results from the fact that it is applied in any distribution of data, indifferent of their distribution. As a consequence, the value $1-\frac{1}{k^2}$ represents a minimum level for the observed data included in the interval $(\bar{x}-k\sigma_w, \bar{x}+k\sigma_w)$ and in reality in interval could exist much more data. With other words, as long as the part of observed data included in the interval $(\bar{x}-k\sigma_w, \bar{x}+k\sigma_w)$ varies from a distribution to another is recommended to establish a minimum value of this part which is correct for any distribution. For a certain distribution, the part of values situated in the interval can be much bigger.

The empirical rule is valid for the sample data, if their distribution is normal. In according with the empirical rule, if the observed data of a sample is normal distributed then:

- the interval ($\bar{x}-\sigma_w, \bar{x}+\sigma_w$) contains about 68% from the registered data;
- the interval ($\bar{x}-2\sigma_w, \bar{x}+2\sigma_w$) contains about 95% from the registered data;
- the interval ($\bar{x}-3\sigma_w, \bar{x}+3\sigma_w$) contains all the registered data.

In the Table no. 2 are presented, in according with the empirical rule and Chebyshev's theorem the percentages of data included in the interval.

Table no. 2. Proportions of including the observed data in the interval

Interval	Percentages by the empirical rule	Percentages by the Chebyshev's theorem
($\bar{x}-\sigma_w, \bar{x}+\sigma_w$)	68%	at least 0%
($\bar{x}-2\sigma_w, \bar{x}+2\sigma_w$)	95%	at least 75%
($\bar{x}-3\sigma_w, \bar{x}+3\sigma_w$)	100%	at least 89%

Source: Keller, G., Warrack, B., Bartel, H., 1988, *Statistics for Management and Economics. A Systematic Approach*, Wadsworth Publishing Company, Belmont, California, 60-61

The standard deviation, as the average linear deviation, is expressed in the same measure units as the analysed variable. From this reason, these indicators are used at the comparison the degree of variation only for the distributions which refers at the same statistical characteristic. In the comparison of the degree of variation for two or more different variables is used the coefficient of variation. This is a relative measure of the dimension of relative deviation in comparison with the average or, with other words, is a relative measure of dispersion.

The dispersion can be measured as a difference between two values selected from a population of observed data. The easiest methods of measure the dispersion are: the range, the interfractile range and the quartile deviation (Levin, I. R., 1987).

The range is the difference between the biggest and the smallest observed value. The indicator is easy to understand and determine, but as a measure of dispersion has a limited character because it didn't take into consideration the most part of the observed values and ignores the nature of their variation, but is influenced by the extreme values. Because is computed from only two values, the range can change from a sample to another in a given population, even if the individual items between the two extremes are in the most part similar.

The interfractile range is a measure of variation between two fractile in a frequencies distribution or, with other words, is the difference between the values of two fractiles. These can have distinct name, which depend by the number of equal parts in which they divide the data. So, the fractiles which divide the data in ten equal parts are named *deciles*, those which divide the data in four equal parts are named *quartiles*, those which divide the data in one hundred equal parts are named *percentiles*.

The interquartile range and *the quartile deviation* show us how much we deviate from the median for limit a half of the observed values. For computing this distance we divide the data in four parts, each of them containing 25% from the observed values of the distribution. The quartiles are the biggest values in each of these parts and the interquartile range is the difference between the value of the first and the third quartile.

Interquartile range = $Q_3 - Q_1$

In the Figure no. 1 is illustrated, in a diagrame, the concept of interquatile distance. We mention that the distance between quartiles isn't necessary to be the same.

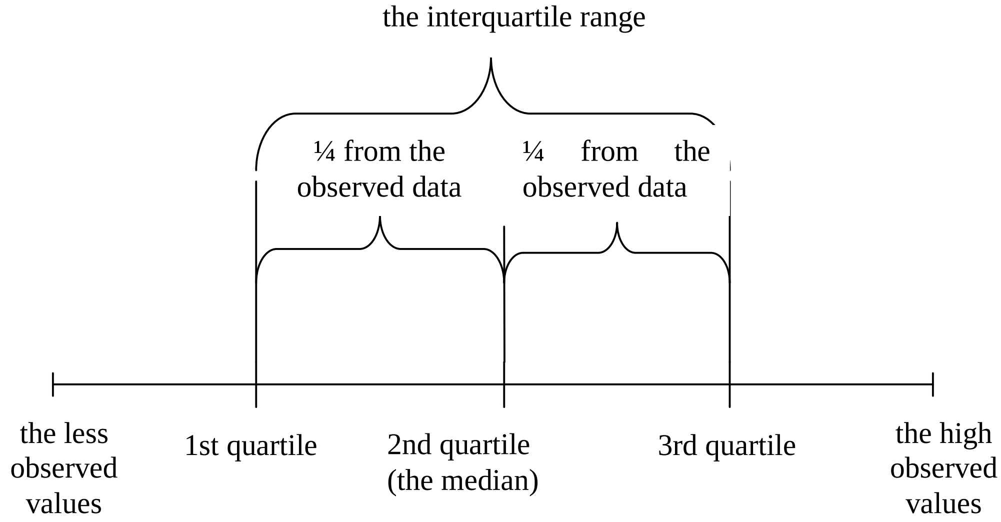

Figure no. 1. The interquartile range
Source: Levin, I. R., 1987, *Statistics for Management*, Forth Edition, Prentice – Hall, Inc., Englewood Cliffs, New Jersey, 114

The half of interquartile range is a measure named *quartile deviation*, determined with the formula:

$$\text{Quartile deviation} = \frac{Q_3 - Q_1}{2} \tag{8}$$

This indicator measure the average of distance between the first and the third quartile and is representative for the variability of all the data as it is computed like an average and it isn't chooses from a population. Also, we observe that the interquartile range and the quartile deviation are based only on two values from the population and these are not the extreme values but are values situated in the middle half of data which give they an advantage in comparison with the range which determined from the extreme values. These indicators are intermediaries for the analysis of dispersion of a frequencies distribution (Roşca, R. E., 2008).

The dispersion - σ_x^2 is computed as a simple or weighted aritmetic average of the sqares of terms deviations in comparison with their average (Ţarcă, M., 1997).

The formula are:

- in the case of simple series:

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} \tag{9}$$

- in the case of grouped data:

$$\sigma^2 = \frac{\sum_{i=1}^{k} (x_i - \bar{x})^2 n_i}{\sum_{i=1}^{k} n_i} \tag{10}$$

The dispersion of a characteristic represent at the same time, the centrat moment of second order, computing without to be necessary the establishing beforehand the individual deviations of variants from their average.

The dispersion is computing with the following formula:

- in the case of simple series:

$$\sigma^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \overline{x}_2 \text{ that is } \sigma^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} x_i}{n} \right)^2 \tag{11}$$

- in the case of grouped series:

$$\sigma^2 = \frac{\sum_{i=1}^{k} x_i^2 \cdot n_i}{\sum_{i=1}^{k} n_i} - \left(\frac{\sum_{i=1}^{k} x_i \cdot n_i}{\sum_{i=1}^{k} n_i} \right)^2 \tag{12}$$

So the dispersion as the standard deviation are computed by the simplified calculus using the formula:

$$\sigma_x^2 = \frac{\sum_{i=1}^{k} \left(\frac{x_i - a}{h} \right)^2 n_i}{\sum_{i=1}^{k} n_i} \cdot h^2 - (\overline{x} - a)^2, \tag{13}$$

where:

a represents a constant choosed depending on a new origin of variation, so that to obtain less values of the variants;

h represents an over-unit coefficient with which if all the terms of distribution are decreased, the new dispersion is by k^2 times less than the dispersion of initial distribution.

In the case of simple or frequencies distributions as the dispersion value is closed by zero so the degree of variability in series is reduced. The dispersion hasn't a concrete measure unit and its dimension is direct influenced by the size order of the distribution values, being considered an intermediary indicator in the characterizing of the variability degree of this.

The coefficient of variation - *v* is computed in comparison with standard deviation and the distribution average and usually is expressed in percentages. The formula is:

$$v = \frac{\sigma}{\overline{x}} \cdot 100 \tag{14}$$

When is known only the average linear deviation, then the coefficient of variation is computing also with the formula:

$$v_{\overline{d}} = \frac{\overline{d}}{\overline{x}} \cdot 100 \tag{15}$$

The coefficient of variation can take values beginning with zero. As the value is smaller so the distribution is much homogeneous and the average is much representative. It is appreciated that in the case of a coefficient of variation over 35 – 40% the average is no more representative and the data must be separated in components by groups depending on the variation of another group variables. So this indicator is used as a test of verification in the using of the group method at the analysis of observed data.

The Shappard's correction is used to decrease the computed dispersion value for a distribution with group intervals at which the values of variable X are computed supposing that in the interval the fequencies are distributed at equal distances and their average will be equal with the simple average of the two extreme values. Shappard considers that in certain conditions the dispersion can be corrected by decreasing $\frac{1}{12}$ from the square of the interval of variation. The dispersion corrected in this mode $(\sigma^2)'$ is equal with:

$$(\sigma^2)' = \sigma^2 - \frac{h^2}{12} \tag{16}$$

where:

h represents the dimension of intreval of variation.

It conditions the application of this correction only for the series which prese nts the following characteristics:

- the frequencies distribution is continued;
- the frequencies tend to zero in the both directions;
- the number of observed cases is smaller than 1.000 units and the group is much large than those in which the interval of variation represents the twelfth part from the range of variation.

These conditions suppose a normal or easy asymmetrical frequencies distribution in which is ensured a compensation of deviations also in the inside of intervals not only on the population . So, it isn't recommended to be used for the distributions with a great degree of asymmetry, for which the application of the Sheppard's correction didn't reduce the value of error and is influenced even in the sense of increasing it. As a result, the app lication of Sheppard's correction needs prudence, by verification beforehand if the conditions of normality and volum of distribution are achieved.

DISPERSION IN MULTIDIMENSIONAL DISTRIBUTIONS. THE RULE OF ADDITION THE DISPERSIONS

How the phenomena are more complex so the degree of variation of the variables which define them is bigger. From this reason the units submissive the observation must be structured in groups depending on the variation of determined factors. In this case the total variation is t he result of combination the action of random and nonregistered factors presented inside of groups with the essential registered factors, which operate from a group to another. To measure the degree of variation determined by the combined action of the two categories of variable factors is used the analysis of variance based on the decomposition the total dispersion in the two components (Jaba, E., 1998), (Baron, T., et al., 1996).

Total dispersion - σ_y^2 is computed on the base of all individual deviations from the average of population using the formula:

$$\sigma_y^2 = \frac{\sum_{j=1}^{m} \left(y_j - \bar{y}\right)^2 n_{\bullet j}}{\sum_{j=1}^{m} n_{\bullet j}} \qquad (17)$$

If we consider that the individual values appear under the influence of all the factors and the average of population supposes that all the factors ar e constant these means that the total dispersion measures all the variation of studied variable Y, produced by the essential and nonessential, registerd and nonregistered factors. How the value of this indicator is bigger as the chara cter of the population is more heterogeneous and it is influenced by a great number of essential and nonessential, constant and variable factors, presented from an interval to another or in the backgroud of the same interval.

The dispersion of group - σ_i^2 is computed on the base of all deviations of variants of a group in comparison with their average, weighted with the group frequencies $\left(\sum_{j=1}^{m} n_{\bullet j} = n_{i\bullet}\right)$, using the formula:

$$\sigma_i^2 = \frac{\sum_{j=1}^{m} \left(y_j - \bar{y}_i\right)^2 n_{ij}}{\sum_{j=1}^{m} n_{ij}} \qquad (18)$$

This indicator ensures the measurement of the influence degree of variable factors which operate inside of each group and how it has a bigger value as the group is more heterogeneous and it is approached by the variation degree of population. To synthetize this variation in one single indicator computed on the all population we resort to the average.

The average of group dispersion - $\bar{\sigma}_y^2$ is computed as a simple or weighted aritmetic average of all the group dispersions determined in the analysed population. If the intervals are equal as volum is used the simple aritmetic average and if the intervals are unequal is computed a weighted aritmetic average.

The average of group dispersions is computed in two variants which are:

- as an average of group dispersions, using the formula:

$$\bar{\sigma}_y^2 = \frac{\sum_{i=1}^{k} \sigma_i^2 n_{i\bullet}}{\sum_{i=1}^{k} n_{i\bullet}} \tag{19}$$

- as a sum of deviation squares inside of groups from the average of the group weighted with the groups frequency, using the formula:

$$\bar{\sigma}_y^2 = \frac{\sum_{i=1}^{k}\sum_{j=1}^{m} \left(y_j - \bar{y}_i\right)^2 n_{ij}}{\sum_{i=1}^{k}\sum_{j=1}^{m} n_{ij}} = \frac{\sum_{i=1}^{k} \sigma_i^2 n_{i\cdot}}{\sum_{i=1}^{k} n_{i\cdot}} \tag{20}$$

The dispersion among groups - $\delta_{y/x}^2$ is computed on the base of deviations of group averages from the average of population and it measures the influence degree of the group factor on the variation of analysed variable.

The dispersion among groups is computed using the formula:

$$\delta_{y/x}^2 = \frac{\sum_{i=1}^{k} \left(\bar{y}_i - \bar{y}\right)^2 n_{i\bullet}}{\sum_{i=1}^{k} n_{i\bullet}} \tag{21}$$

Among the three indicators exist the following relationship: *dispersion in population is equal with the average of group dispersions at which is added the dispersion among the groups of population*, that is:

$$\sigma_y^2 = \bar{\sigma}_y^2 + \delta_{y/x}^2 \tag{22}$$

This formula is known under the name of *the rule of addition the dispersions*.

The relationships between the standard deviation computed on population and the standard deviation among groups are achieved through the realtions between the corresponded dispersions, because the standard deviation in population is equal with the square root from the sum of the average of group dispersions and dispersion among groups, that is:

$$\sigma_y = \sqrt{\bar{\sigma}_y^2 + \delta_{y/x}^2} \tag{23}$$

The indicators of variation are computed also using the relative frequencies, with the condition to respect the quantitative relatioships between absolute and relative frequencies, to take into account the fact that the absolute value of a percentage of the relative frequencies is the same for all the elements of population.

THE USING OF VARIATION INDICATORS IN THE ANALYSIS OF TOURISTIC TRAVEL DURATION

In the case of sample data the group is achieving by the necessities of analysis, by the groups achieved in a previous similar research or by the aim and necessities of present analysis. In Romania, the average travel duration has situated, in the period 1990 – 2005, under 4 days, registered even an easy tendency of decrease. On the development regions, the less value was registered in Region of North-East, where the computed averages are under 3 days and in the Region Bucharest – Ilfov, where the average travel duration is even under 2 days, in the considered period. However the two mentioned regions have remarkable touristic resources and the forms of tourism which characterized these regions are cultural tourism, which has an itinerant character, scientific, transit, religious, affairs tourism a.s.o., all of them having reduced travel duration. We mention the regions South – East (which include the littoral tourism) and South – West Oltenia (a rich zone in cultural touristic objectives), where the average travel duration is over the average of the country (Zaharia, M., Gogonea, R. M., 2005).

The indicator the number of overnight stays on the touristic zones in 2005 has a higher value than in the case of balnear tourism (about 5,3 millions overnight stays), in the residence of district tows (5,2 millions overnight stays), in the littoral zone (4,0 millions overnight stays), in the mountainous zone (2,0 millions overnight stays) a.s.o. The average travel duration was 3,2 days on the country level, in 2005, with variations from 8,2 days in average in the case of balnear tourism, 5,6 days in the case of littoral tourism, 2,4 days in the case of mountainous tourism a.s.o. The Romanian touristic statistic offers a group of internal touristic travels of residents for holidays and business, on the touristic zones by the travel duration, presented for 2005 in Table no. 3.

Table no. 3. Internal travels of residents for holidays and business on touristic zones, by the duration of travel in Romania, in 2005

Travel duration (overnight stays)	The number of internal travels (millions)	Touristic zones					
		Littoral	Mountainous	Health resorts	Danube Delta	Diverse circuits	Others zones
1-3	5,17	0,13	0,83	0,05	0,01	0,06	4,08
4 -7	2,29	0,49	0,57	0,04	0,02	0,02	1,15
8 -14	0,99	0,36	0,19	0,05	0,01	0,02	0,37
15 – 28	0,22	0,04	0,05	0,02	-	-	0,11
29 and over	0,07	-	0,01	-	-	-	0,06
Total	8,74	1,02	1,65	0,16	0,04	1,0	5,77

Source: *Anuarul statistic al României – ediţie electronică*, INS Bucureşti, 2007

We observe that by elimination from the analysis “other zones”, in the mountainous zone are registered the most touristic travels, followed by the littoral and “diverse circuits”. In the mountainous zone the most travels have a duration travel by 1-3 overnight stays, in the littoral zones by 4 – 7 overnight stays, in the health resorts by 8 – 14 overnight stays a.s.o. (Biji, E. M., et al., 2006).

The average of internal travels duration of residents in Romania in 2005 was by 4,7 overnight stays and the average linear deviation was by 3,19 overnight stays, which means that in average the individual values deviate from the level of average with 3,19 overnight stays. The standard deviation is 4,87 overnight stays and the dispersion of the individual values from the average is 23,7. The coefficient of variation has values over 40% as computed by using the standard or the average linear deviation. By applied the Chebyshev’s theorem as it is presented in the Table no. 1 we observe that at least ¾ from the observed data are situated in the interval [-5,04; 14,44], at least 21/25 from the data are situated in the interval [-7,475; 16,88] and at least 8/9 from the data are situated in the interval [-9,91; 19,31]. The empirical rule can't be used in this case because the distribution is asymmetrical.

CONCLUSIONS

We ascertain that the analysis of characteristic of variability in the freque ncies distributions consists in the estimation of the synthetical indicators of variability, which measures the degree of variation of the individual values from an indicator of central tendency as a rule the average and sometimes the median. The average linear deviation has some disadvantages connected by no taking into consideration of the bigger influences of bigger deviations in absolute value on the variation degree of a characteristic and also connected by the expression in the unit of measure of characteristic, which make that the indicator can't be used in the comparison of dispersions of different variables in frequencies distributions. More useful in the analysis of dispersion is the standard deviation used in the estimation of sample errors, in co rrection calculus, as a parameter which defines the theoretical distributions and serves at the interpretation of the normal curve of distribution. The standard deviation is finding the application also in Chebyshev's theorem, at the establishing the proportion in which the data of a population or a sample is appointing in an interval of variation, indifferent how is the form of data distribution or in empirical rule, where has the same role but in the case of normal distributions. How the value of the disp ersion indicator is less that shows a smaller degree of variability. It didn't have a concrete unit of measure and so it isn't used in the comparison of frequencies distributions in the absolute expression, but in the relative expression under the form of coefficient of variation which value as it grows shows a bigger degree of nonhomogeneity of distribution and of nonrepresentative of the average. In the multidimensional distributions the total dispersion is determined as a sum of the average of group disp ersions which results from the action of nonessential factors and of the dispersion among groups, determined by the action of essential factors. Used in the touristic activity, the analysis of the internal residents distribution by the touristic duration travel in Romania, in 2005 has emphasized a high degree of nonhomogeneity, a nonrepresentative average by this group and the necessity of group again on other intervals of variation to obtain the increasing degree of homogeneity in distribution.

BIBLIOGRAPHY:

1. Badia, J., Bastida, R., Haït, J. R., 1997, *Statistique sans mathématique*, Ellipses, Édition Marketing S.A., Paris
2. Baron, T., Biji, E. M., Tövissi, şi alţii, 1996, Statistică teoretică şi economică, Editura Didactică şi Pedagogică R. A. Bucureşti
3. Biji, M., Biji, E. M., 1979, *Statistică teoretică*, Editura Didactică şi Pedagogică, Bucureşti
4. Biji, M., Biji, E. M., Lilea, E., Anghelache, C., 2002, *Tratat de statistică*, Editura Economică, Bucureşti
5. Biji, E. M., Lilea, E., Roşca, R. E., Vătui, M., 2000, *Statistică aplicată în economie*, Editura Universal Dalsi, Bucureşti
6. Biji, E. M., Lilea, E., Vătui, M., 2006, *Unele probleme privind analiza statistică a distribuţiei pe regiuni de dezvoltare a activităţii turistice*, în volumul „Dezvoltarea regională în contextul integrării în Uniunea Europeană", Editura Economică, Bucureşti, 199 - 203
7. Isaic – Maniu, Al., Mitruţ, C., Voineagu, V., 1999, *Statistica pentru managementul afacerilor*, Editura Economică, Bucureşti
8. Jaba, E., 1998, *Statistică*, Editura Economică, Bucureşti

9. Keller, G, Warrack, B., Bartel, H., 1988, *Statistics for Management and Economics. A Systematic Approach*, Wadsworth Publishing Company, Belmont, California
10. Korka, M., Begu, L. S., Tuşa, E., Manole, C., 2005, *Bazele statisticii pentru economişti*, Editura Tribuna Economică, Bucureşti
11. Levin, I. R., 1987, *Statistics for Management*, Forth Edition, Prentice – Hall, Inc., Englewood Cliffs, New Jersey
12. Roşca, R. E., 2008, *Analysis of Repartition Form in Distributions of Frequencies*, în The Annals of the "Stefan cel Mare" University Suceava, Fascicle of the Faculty of Economics and Public Administration, Section 4. Statistics, data processing (informatics) and mathematics, Year 8, No. 8, Editura Universităţii Suceava, 191 - 201
13. Ţarcă, M., 1997, *Tratat de statistică aplicată*, Editura Didactică şi Pedagogică R. A. Bucureşti
14. Zaharia, M., Gogonea, R. M., 2005, *Circulaţia turistică pe destinaţii turistice în zonele geografice ale României*, în volumul „Teorie şi practică în dezvoltarea regională", Editura Economică, Bucureşti, 185 – 199
15. * * *Anuarul statistic al României*–ediţie electronică, INS Bucureşti, 2007