Professor Marin ANDREICA, PhD
The Bucharest Academy of Economic Studies
Lecturer Dan NICOLAE, PhD Candidate
"Titu Maiorescu" University, Bucharest
Assistant Madalina ANDREICA, PhD Candidate
The Bucharest Academy of Economic Studies
Lecturer Liliana TODOR, PhD Candidate
Commercial Academy, Satu-Mare

WORKING OUT FORECASTS ON THE BASIS OF STATISTICAL EXTRAPOLATION PRINCIPLE*

The paper presents a series of new statistical methods which can be used in order to working out some extrapolation-based forecasts.

The study begins with the deterministic, discrete and probabilistic methods. At the same time, two new antithetical concepts, the expanding and constricting of mean are introduced. The linguistic and algebraic models allow the use of fuzzy sets.

The methods proposed to this end are simple but provides an accuracy ε_i, according to the number of significant decimals, whereby the numbers of the used informatic system are represented.

The idea of researching the model where the sequence is approaching to mean, has allowed the estimation of the symmetry/asymmetry coefficients as ratio of the sequence possible combinations.

Key words: *forecasting, statistical extrapolation, fuzzy sets, expanding-compressing mean, finite difference equations, chaos.*

JEL classification: C 53, P 47

MSC classification: 91B82; 03E72

---

* **La prévision**

Une prévision peut être définie comme un ensemble de probabilités associées à un ensemble d'événements futurs (Fischhoff, 1994). Cette prévision est basée sur un ensemble d'informations disponibles à l'instant t où elle a été effectuée. Cet ensemble noté Ω_t (l'indice temporel t correspond à l'instant t) représente les données disponibles, les connaissances et les théories concernant le phénomène que l'on souhaite prévoir. La prévision au temps t d'horizon h de la variable Y, considérée comme une variable aléatoire, peut donc s'écrire (Granger et Newbold, 1977) sous la forme d'une fonction de distribution conditionnelle (à l'ensemble d'informations Ω_t) : $G(y) \equiv Prob(Y_{t+h} \leq y|\Omega_t)$

1. FORECAST BASED ON SOME STATIC STATISTICAL DATA

1.1. Deterministic Methods

These methods are used to the situations where several levels of a parameter $x_1, x_2, \ldots, x_n$ are known. For instance, *x* can be the time required to fulfilling a work, the material consumption to manufacture a product, the fuel consumption to achieving a unit distance etc. It is necessary to set which is the future level. In this case, we calculate the arithmetic mean:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} \tag{1}$$

where *n* = number of measurements made on *x* parameter.

In the future, we suppose that parameter $\bar{x}$ will be achieved (with a certain feasible rounding from the point of view of the accuracy required by the forecasted phenomenon).

1.2. Probabilistic Methods

These methods operate with certain statistical parameters. The most frequently used is the mean square deviation, namely:

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

If we know a certain distribution, then, we use the statistical tables. Knowing the sequence of the likelihood coefficients *k*, we determine a probability P ($x \geq \bar{x} + k\sigma$), respectively P ($x \leq \bar{x} + k\sigma$), if *x* is maximized and a probability P($x \geq \bar{x} - k\sigma)$, respectively P($x \leq \bar{x} - k\sigma)$ if *x* is minimized.

If the distribution function is not known, then we apply Cebâşev's formula:

$$P(\bar{x} - k\sigma \leq x \leq \bar{x} + k\sigma) \geq 1 - \frac{1}{k^2} \tag{3}$$

where *k* = the likelihood coefficient.

1.3. Fuzzy Method Used to Mean Expanding

There are cases when mean expanding is useful. For instance, it is possible that the mean not to be found in the excellence domain of the variable under research.

In this case, after the mean value $\bar{x}$ is determined, the membership degree

$\mu(\bar{x})$ = 1 is assigned to it. In order to find an expanded mean, we proceed to "expanding" figure 1, for instance, on interval [0.95;1]. We calculate a value $x^{'}$ with the property that:

$$\mu(x^{'}) = 0.95 \tag{4}$$

If this value is single, then the expanded mean is the interval [x', $\bar{x}$] or [$\bar{x}$, x']. If two values, $x^{'}$ and $x^{''}$ would exist, such that:

$$\mu(x^{'}) = \mu(x^{''}) = 0.95 \tag{5}$$

then, the expanded mean is given by the interval [x', $\bar{x}$] U [$\bar{x}$, x''] = [x',x''].

1.4. Linguistic Methods

In order that a linguistic procedure to be applied, it is necessary that the continuous parameter *x* (or the discrete parameter, but with a large number of restricted values) to be subject to a granulation operation. We consider the maximum and the minimum levels and among them, certain levels which are denominated by using words (verbs, adverbs) etc. as for instance:

- we shall use the following gradations if five levels must be introduced between a minimum weight piece and a maximum weight one:
- very light - for minimum
- light - for immediately next level;
- average - for immediately next level;
- heavy - for immediately next level;
- very heavy - for immediately next level.

If the frequency of gradations is symmetrical, in order to find the average weight, we allow for the weights X_{min} and X_{max} and we divide the level [X_{min} , X_{max}] into equal parts (because four intervals exist among the five levels).

An interval *I* will have the length:

$$I = \frac{x_{\max} - x_{\min}}{4} \tag{6}$$

In this case, $\bar{x}$ is determined by the relation:

$$\bar{x} = x_{\min} + 2I = x_{\max} - 2I \tag{7}$$

Practically, $\bar{x}$ is lower rounded and we obtain for mean $\bar{x}$ an interval under the form :

$$(\bar{x} - r, \bar{x} + r) \tag{8}$$

where r = the rounding value.

It is possible that this rounding not to be symmetrical, therefore the obtained interval is:

$$[\bar{x}-r', \bar{x}+r''] \qquad r' \neq r' \tag{8'}$$

where : r'= rounding to left; r''= rounding to right.

If the number of gradations is even, then, the number of intervals is odd, therefore the mean is already represented in an interval. If the frequency of gradations is not symmetrical, then each gradation is weighted with the interval centre (attached to each gradation). Thus, an average gradation representing an interval is obtained.

1.5. Mean Expanding Method

In order to apply this method, we present several procedures:

a) Decimal method

The used notations will be :

Amplitude (a), error (ε), a = 2 ε , decimals number (z)

We have to verify statistically the verisimilitude of the relation:

a.z = const., true value or false value

*Problem.*_ Given a mean m $\pm\varepsilon$ where from a number of decimals *z* results, ensuring that:

$$x \in [m-\varepsilon, m+\varepsilon] \tag{9}$$

b) Mean Approaching Method

Given an ordered sequence of numbers $x_1 \leq x_2 \leq x_3 \leq \ldots \leq x_n$. Mean $\bar{x}$ is ascertained.

We determine:

$$N_s = \sum_{i/x_i\langle\bar{x}} x_i \quad \text{and} \quad N_d = \sum_{j/x_j\rangle\bar{x}} x_j \tag{10}$$

If Ns = Nd , we have a sequence *symmetry*_ .

If Ns $\rangle$ Nd , we have an asymmetry to left. ; the coefficient of the asymmetry to left is:

$$AS_s = \frac{N_s - N_d}{N_s + N_d} \tag{11}$$

If Nd $\langle$ Ns, we have a coefficient of asymmetry to right:

$$A\,S_d = \frac{N_d - N_s}{N_s + N_d} \tag{12}$$

If A $S_s = 0$ and A $S_d = 0$, a symmetry is obtained. The symmetry coefficient is:

$$S_s = 1 - A_S^S \quad \text{or} \quad S_d = 1 - A$$

If A S_s=A S_d=0, then we have a perfect symmetry coefficient:

$$S = 1$$

We take the nearest number to mean $\bar{x}$, namely :

$$\underset{h}{Min}[(x_h - \bar{x})] = x_{i1} - x \tag{13}$$

We eliminate x_i and retake the procedure of finding the nearest term against mean $\bar{x}$.

Finally, we obtain a sequence $x_{i1}, x_{i2}, \ldots, x_{in}$ called the ordered sequence of the distances against the 1 order mean . The amplitude of this sequence is : $a_1 = x_{i\,n} - x_{i\,n+1}$

We calculate iteratively a mean for one term, two terms, up to *n* terms, namely:

$$\bar{x}_{11} = x_{i1}; \quad \bar{x}_{12} = \frac{x_{i1} + x_{i2}}{2}; \quad \bar{x}_{13} = \frac{x_{i1} + x_{i2} + x_{i3}}{3} \ldots$$

$$\bar{x}_{in} = \frac{x_{i1} + x_{i2} + \ldots + x_{in}}{n} = x_1 \tag{14}$$

We reorder this sequence and we get an increasingly ordered sequence: :

$$\bar{x}_{1j_1}, \bar{x}_{2j_2}, \ldots, \bar{x}_{1j_n} \tag{15}$$

We determine again the mean of this sequence and we get a mean of means, consequently, a "2-order" mean. We retake the above procedure and we compute:

$$\bar{x}_{21} = \bar{x}_{1j_1}; \quad \bar{x}_{22} = \frac{x_{ij_1} + x_{2j_2}}{2}; \tag{16}$$

$$\bar{x}_{2j} = \frac{x_{1j_1} + x_{2j_2} + \ldots + x_{1j_n}}{n} = \bar{x}_2 \tag{17}$$

$$a_1 = Max[\bar{x}_{ij_p} - \bar{x}_{ij_r}]$$

We reorder the obtained sequence, according to the proximity to mean $\bar{x}_2$, etc.

Properties.

1) The interval of first-order means is included in the interval of i = 1order means,

namely: : $[\overline{x}_{il_1}, \overline{x}_{il_n}] \subset [\overline{x}_{i-1,k_1}, \overline{x}_{i-1,k_n}]$

2) The amplitude of i - order mean sequence has the property that:

$$\overline{a}_{i-1} \rangle \overline{a}_i$$

At limit, when n - order mean tends to infinite, the sequence amplitude tends to zero.

We notice in the considered example that :

$\overline{x}_2 = (2.5+2.5+3+3+3)/5 = 2.8$ If we order decreasingly the first - order means, we get : 3.3 ; 2.5 ; 2.5 ; 2.5 .

The second - order means are:

$$\overline{x}_{21} = 3; \overline{x}_{22} = (3+3)/2 = 3; \overline{x}_{23} = (3+3+2.5)/3 = 2.833$$
$$\overline{x}_{24} = (3+3+2.5+2.5)/4 = 2.75; \overline{x}_{25} = 2.8. \quad (18)$$

Other methods of mean expanding:

1.6. The Truncation Method
1.7. The Rounding Method

1.6. The Truncation Method

The steps of the method are:

Step 1. The value $\overline{x}_i$ is truncated and a minimum j - level truncated value is obtained, according to the j-level of truncation:

$$y_{ij}^m = T_j \overline{x}_i$$

Step 2. A unit for j-level is added, namely b^j (where b is the numeration basis) and the maximum level is obtained.

$$y_{ij}^M = y^m + b^j \quad \text{(for integer numbers)}$$

and

$$y_{ij}^M = y_{ij}^m + b^{-j} \quad \text{(for j-rank fractions)}$$

j = decimals number.

1.7. The Rounding Method.

The method consists in two steps:

Step 1. The number is lower rounded for *n* sequence, yielding:

$$x_n^m$$

Step 2. The number is upper rounded for the *n* level, yielding

$$x_n^M$$

The differences are calculated in the same way:

$$\Delta_n = x_n^M - x_n^m$$

with the property that:

$$\lim_{n\to\infty} \Delta_n = 0$$

1.8. The Amplitude Method

Let $a_1, a_2, \ldots, a_n$ be a statistical sequence. Let $\overline{x}_n$ be the average values:

$$\overline{x}_n = (a_1+a_2+\ldots+a_n)/n$$

with the property that:

$$\lim_{n\to\infty} \overline{x}_n = \overline{x}$$

We consider several *i* sequences for each *n* rank: $a_1^i, a_2^i, \ldots, a_n^i$. Several means are obtained for *n* rank, namely:

$$\overline{x_n^i} = (a_1^i + a_2^i + \ldots + a_n^i)/n \ ; \ \text{i} = 1.2\ldots\ldots,\text{m}$$

Obviously :

$$\lim_{n\to\infty} \overline{x_n^i} = \overline{x}_n$$

A minimum and a maximum value corresponds to each *n* rank:

$$\overline{x_{n,m}^{\min}} = \underset{i}{Max}\, x_n^i ; i = 1,2,\ldots,m.$$

$$\overline{x_{n,m}^{\min}} = \min_i x_n^i ; i = 1,2,\ldots,m \tag{19}$$

If we pass to limit, we yield :

$$\lim_{n\to\infty} \overline{x_{n,m}^{\max}} = \overline{x_n^{Max}}$$

$$\lim_{n\to\infty} \overline{x_{n,m}^{\min}} = \overline{x_n^{\min}}$$

Graphically, these results are reported below:

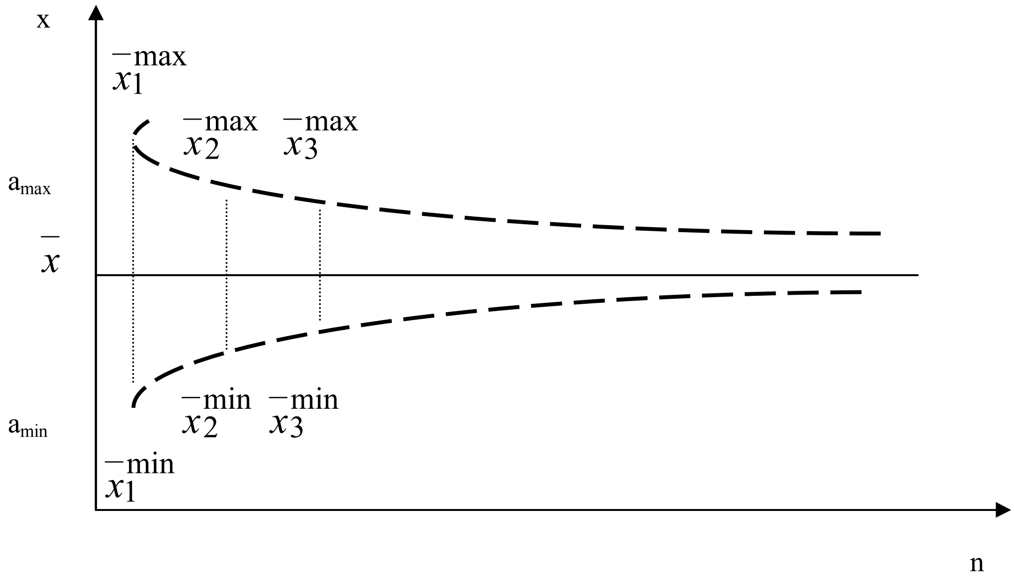

Figure 1. Graph of the amplitude method results

The maximum values of sequences $a_1^i, a_2^i, \ldots, a_n^i$ are written in the above figure.
i=1,2,…,m, where $m \to \infty$, as well as, the minimum ones :

$$\begin{aligned} &\underset{i}{Max}\, a_1^i = a_1^{Max};\ \underset{i}{\min}\, a_1^i = a_1^{\min} \\ &\underset{i}{Max}\, a_2^i = a_2^{Max};\ \underset{i}{\min}\, a_2^i = a_2^{\min} \\ &\ldots\ldots\ldots\ldots\ldots \\ &\underset{i}{Max}\, a_n^i = a_n^{Max};\ \min a_n^n = a_n^{\min} \end{aligned} \tag{20}$$

If $m \to \infty$, then:

$$a_1^{Max} = a_2^{Max} = \ldots = a_n^{Max}, a_1^{\min} = a_2^{\min} = \ldots.. = a_n^{\min} \tag{21}$$

If also $n \to \infty$ the interval ($a_{\min}$,a_{Max}) is obtained, representing the general amplitude of the sequence and ($a_{n,m}^{\min}, a_{n,m}^{Max}$) is the sequence amplitude for *n* experiences and

m sequences. If m tends to infinite, an interval $(a_n^{\min}, a_n^{Max})$ is obtained for each number of n experiences . If n tends to infinite, we get:

$$\lim_{n\to\infty} x_n^{\min} = \lim_{n\to\infty} x_n^{Max} = \bar{x}$$

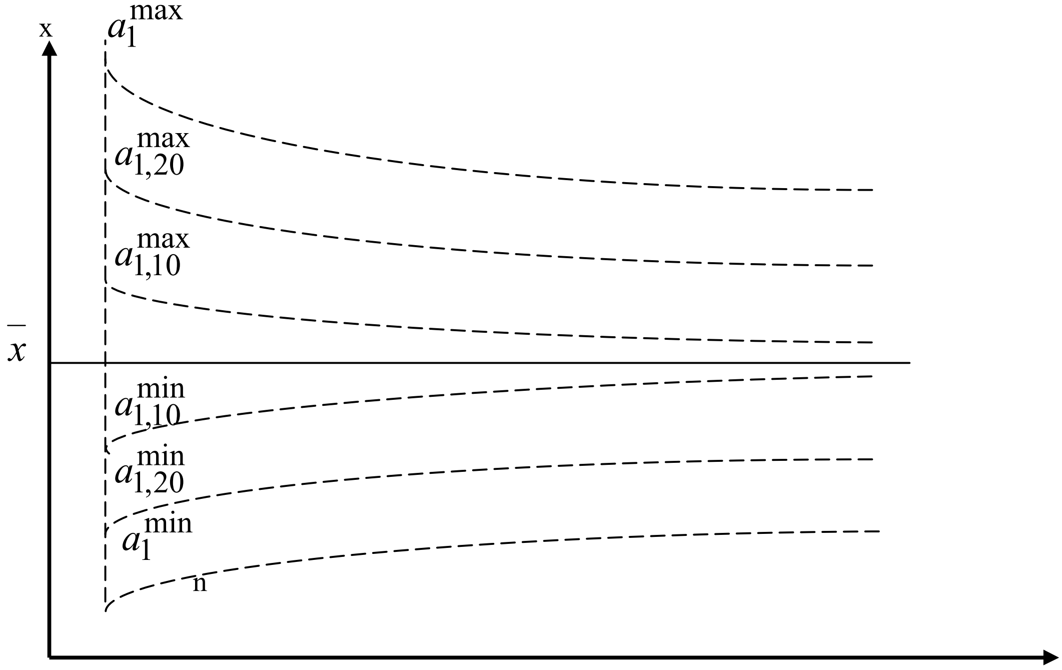

Figure 2. Graph of the amplitude method results

The following relation exists between the amplitude and the minimum, respectively the maximum value:

$$\overline{x_n^{Max}} = \bar{x} + 0.5(a_n^{Max} - a_n^{\min}) / \sqrt{n}$$

$$\overline{x_n^{\min}} = \bar{x} - 0.5(a_n^{Max} - a_n^{\min}) / \sqrt{n} \tag{22}$$

If symmetry does not exist, the relation becomes:

$$\overline{x_n^{Max}} = \bar{x} + \beta(\overline{x_n^{Max}} - \overline{x_n^{\min}}) / \sqrt{n} \tag{23}$$

$$\overline{x_n^{\min}} = \bar{x} - \alpha(\overline{x_n^{Max}} - \overline{x_n^{\min}}) / \sqrt{n}$$

where:

$$\alpha = (\bar{x} - \overline{x_n^{\min}}) / (\overline{x_n^{Max}} - \overline{x_n^{\min}});$$

$$\beta = (\overline{x^{Max}} - \overline{x}) / (\overline{x^{Max}} - \overline{x_n^{\min}});$$

We take the maximum amplitude:

$$A = \underset{k}{Max}(a_k^{Max} - a_k^{\min}) \quad ,k\in \{1,2,\ldots,n\} \tag{24}$$

In case of symmetry, we have:

$$\overline{x_n^{Max}} = \overline{x} + 0,5A/\sqrt{n}; \tag{25}$$
$$\overline{x_n^{\min}} = \overline{x} - 0,5A/\sqrt{n};$$

For a mean included into the limits of a tolerance, $(-\varepsilon,\varepsilon)$ we have:

$$0.5A/\sqrt{n} = \varepsilon \Rightarrow A = 2\varepsilon\sqrt{n}$$

Assuming that *n* experiences have been made, a mean included in interval $(\overline{x}-\varepsilon,\overline{x}+\varepsilon)$ results.

If only one experiment would be made, then the mean would be included between $\overline{x} - A/2$ and $\overline{x} + A/2$. Therefore, we should use a function *f(x)* which to transform the interval ($\overline{x}-\varepsilon,\overline{x}+\varepsilon$) into the expanded interval ($\overline{x} - A/2, \overline{x} + A/2$), namely :

$$f_{\varepsilon 1}(\overline{x}_{a/2} - \varepsilon) = \overline{x} - A/2$$
$$f_{\varepsilon 1}(\overline{x}_{a/2} + \varepsilon) = \overline{x} + A/2 \tag{26}$$

If *f(x)* is a linear function of the form: f(x) = ax + b, we have:

$$a(\overline{x} - \varepsilon) + b = \overline{x} - A/2$$
$$a(\overline{x} - \varepsilon) + b = \overline{x} + A/2 \tag{27}$$

hence : $2\varepsilon a = A \rightarrow a = A/2\varepsilon$

The transformation function becomes:

$$f_{\varepsilon,A/2} = \frac{A}{2\varepsilon} - \frac{A-2\varepsilon}{2\varepsilon}\overline{x} = \frac{Ax-(A-2\varepsilon)\overline{x}}{2\varepsilon} = \overline{x} + \frac{A}{2\varepsilon}(x-\overline{x}) \tag{28}$$

This function transforms the tolerance mean ε, into a tolerance mean A/2, namely, into a statistically non-processed value (value of the initial sequence $[a_1^i, a_2^i, ..., a_n^i]$).

In conclusion, in order to characterize a statistical sequence, there are necessary:

- Amplitude A
- Mean $\overline{x}$.

In order to simulate a process with values *y* having amplitude A and mean $\overline{x}$, with an accuracy of $\pm\varepsilon$, we take the values $x \in [\overline{x} - \varepsilon, \overline{x} + \varepsilon]$, using the relation:

$$y = f_{\varepsilon, A/2}(x) = \overline{x} + \frac{A(x - \overline{x})}{2\varepsilon} \quad (29)$$

We notice that $f_{\varepsilon, A/2}(\overline{x}) = \overline{x}$

If we allow for $\varepsilon < B/2 < A/2$, then we have:

$$z = f_{\varepsilon, B/2}(x) = \overline{x} + \frac{B}{2\varepsilon}(x - \overline{x}); B = 2\varepsilon\sqrt{k} \Rightarrow \frac{B}{A} = \frac{\sqrt{k}}{\sqrt{n}} \Rightarrow k = \frac{B^2}{A^2} n$$

It follows that the expanding can be made either in $(\overline{x} - \varepsilon, \overline{x} + \varepsilon)$, or in $(\overline{x} - A/2, \overline{x} + A/2)$ that is exactly the statistical sequence or in a sequence of expanded means $[\overline{x} - B/2, \overline{x} + B/2] \subset [\overline{x} - A/2, \overline{x} + A/2]$

Another operation that can be applied to a sequence of means is the **mean constricting** (namely, the opposite of expanding , presented below).

We consider a set of statistical sequences:

$$\{a_1^i, a_2^i, ..., a_n^i, i = 1,2,...,m\}$$

We calculate the mean for each sequence:

$$\overline{x_n^i} = \frac{a_1^i + a_2^i + ... + a_n^i}{n} \quad (30)$$

with tolerance ε_1.

Considering the amplitude A = $\underset{n,1}{Max}(a_n^i) - \min(a_n^i)$ (31)

but also of the means:

$$A_n = \underset{i}{Max}\,\overline{x_n^i} - \min \overline{x_n^i} \quad (32)$$

Obviously, $A_n \langle A$

We calculate the mean of means (2-order mean)

$$\overline{x_n^2} = \frac{\overline{x_1^{(1)1}} + \overline{x_2^{(1)2}} + \ldots + \overline{x_n^{(1)n}}}{n} \quad (33)$$

with a tolerance of $\varepsilon_2 < \varepsilon_1$

The process is continued until we obtain the 3-order mean, with the tolerance $\varepsilon_3 < \varepsilon_2$ etc. The transformation function is:

$$f_{\varepsilon_1,\varepsilon_2}(x) = \overline{x} + \frac{\varepsilon_2}{\varepsilon_1}(x - \overline{x}) \quad (34)$$

It follows that the constricting operation is the opposite of the expanding operation.

2. FORECASTING ON THE BASIS OF CERTAIN DYNAMIC STATISTIC DATA

2.1. Principle of Conserving an Invariant for the Analyzed Domain

We start from the analysis of a dynamic sequence $x_1,x_2,...,x_n$. We search for finding a function $F(x_1,x_2,\ldots,x_k)$ where k<n (usually, k<<n) with the property that it is at least approximately conserved, namely:

$$F(x_p,x_{p+1},x_{p+2},\ldots,x_{p+k}) \cong C \quad (35)$$

where : C = constant value

$$p+k \leq n, (\forall)k \in \{0,1,2,...,n-p\}$$

In this case, we can assume the hypothesis that also for the future, the sequence keeps its property that function $F(x_p,x_{p+1},\ldots,x_{p+k})$ is an invariant. This hypothesis allows to state the principle of conserving an invariant, according to the studied period. For instance, we find for a dynamic sequence that:

$$F(x_p, x_{p+1}) = \frac{x_{p+1}}{x_p} \cong C \tag{36}$$

Obviously, it follows:

$$x_{p+1} = Cx_p \tag{37}$$

If we would know the value x_0 from the basic period, then we could apply the principle of conserving the invariant C, also for the past, consequently:

$$\frac{x_1}{x_0} = C \tag{38}$$

hence $x_1 = Cx_0$

We have in the second period:

$$x_2 = Cx_1 = C^2 x_0 \tag{39}$$

Therefore, a geometric property whose rate is C, follows: The last term will be:

$$x_n = C^n x_0 \tag{40}$$

Applying the principle of conserving invariant C that is an index of growth between two consecutive periods, we can calculate for the first forecasting period:

$$x_{n+1} = C^{n+1} x_0 \tag{41}$$

We shall have for the (n + k) rank forecasting period:

$$x^{n+k} = C^{n+k} x_0 \tag{42}$$

Usually, the experts prefer to choose an uneven *n*. In order the obtained forecasting to be credible, we must have:

$$k = \frac{n-1}{2} \tag{43}$$

Another example to illustrating the principle of conserving the invariant is given by the following simple function:

$$\mathrm{F}(x_{p+1}, x_{p+2}, x_{p+3}) = \frac{x_{p+3}}{x_{p+1} + x_{p+2}} = C \tag{44}$$

We get:
$$x_{p+3} = C(x_{p+1} + x_{p+2}) \tag{45}$$

A particular case is C = 1. It follows:

$$x_{p+3} = x_{p+1} + x_{p+2} \tag{46}$$

A finite differential equation was obtained that can be solved simply by substituting:

$$x_p = ar^p, a \neq 0 \tag{47}$$

We obtain:

$$ar^{p+3} = ar^{p+1} + ar^{p+2} \tag{48}$$

Simplifying by ar^{p+1}, it results: $r^2=r+1$, namely a II-grade equation under the form:

$$r^2-r-1=0$$

The solutions of this equation are:

$$r_1 = \frac{1+\sqrt{1+4}}{2} = \frac{1+\sqrt{5}}{2} \cong 1.615 \tag{49}$$

$$r_2 = \frac{1-\sqrt{5}}{2} \cong \frac{-1.23}{2} = -0.615$$

The first solution is a relation very frequently met in nature, called the golden section.

The general solution of finite differential equation is :

$$x_n = ar_1^n + br_2^n \quad r_1 \neq r_2 \tag{50}$$

In order to determine the constants *a* and *b*, we should know both the initial value x_0, and a previous value x_{-1} or the pair of the level from the basic period

x_0 and the level from the first period. In this last case, the equation system was obtained:

$$\begin{aligned} x_0 &= a+b \\ x_1 &= ar_1+br_2 \end{aligned} \tag{51}$$

By solving this equation system, (taking *a* and *b* as known) we obtain:

$$a = \frac{x_1 - r_2 x_0}{r_1 - r_2} \text{ and } b = \frac{x_1 - r_1 x_0}{r_2 - r_1} = \frac{r_1 x_0 - x_1}{r_1 - r_2} \tag{52}$$

In conclusion,, the general solution of the finite differential equation is:

$$x_n = \frac{(x_1 - r_2 x_0) r_1^n + (r_1 x_0 - x_1) r_2^n}{r_1 - r_2} \tag{53}$$

Accordingly, we can calculate the adjusted values $\bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n$ if these values are getting near to the experimental values, namely:

$$\left|\bar{x}_2 - x_2\right| \le \varepsilon; \left|\bar{x}_3 - x_3\right| \le \varepsilon, \ldots, \left|\bar{x}_n - x_n\right| \le \varepsilon, \tag{54}$$

then, one extrapolates:

$$\bar{x}_{n+1}, \bar{x}_{n+2}, \ldots, \bar{x}_{n+p} \tag{55}$$

2.2. Mobile Means Method under Deterministic Hypothesis

We consider a sequence of numbers representing statistical data for periods: x_1, x_2,…, x_n. These data are going to be **adjusted** by using a mobile mean for 3 periods (or for 5 , 7 etc periods). So, we take a mobile mean for 3 periods. An extra term is calculated at each iteration (term $\tilde{x}_{n+1}^{(1)}$ for the first iteration, term $x_{n+2}^{(2)}$ for the second iteration etc.). We obtain at the first iteration:

$$x_1^{(1)} = x_1;\ x_2^{(1)} = \frac{x_1 + x_2 + x_3}{3};\ x_3^{(1)} = \frac{x_2 + x_3 + x_4}{3};\dots;$$
$$x_{n-1}^{(1)} = \frac{x_{n-2} + x_{n-1} + x_n}{3};\ x_n^{(1)} = x_n \tag{56}$$

We apply the following relation to calculate $x_{n+1}^{(1)}$,

$$x_n^{(1)} = \frac{x_{n-1}^1 + x_n^{(1)} + x_{n+1}^{(1)}}{3} \tag{57}$$

hence

$$\widetilde{x}_{n+1}^{(1)} = 2x_n^{(1)} - x_{n-1}^{(1)} = 2x_n - x_{n-1}^{(1)} \tag{58}$$

We get at the second iteration:

$$x_1^{(2)} = x_1^{(1)};\ x_2^{(2)} = \frac{x_1^{(1)} + x_2^{(1)} + x_3^{(1)}}{3};\ x_3^{(2)} = \frac{x_2^{(2)} + x_3^{(2)} + x_4^{(2)}}{3};\dots;$$
$$x_n^{(2)} = \frac{x_{n-1}^{(1)} + x_n^{(1)} + \widetilde{x}_{n+1}^{(1)}}{3};\ \widetilde{x}_{n+1}^{(2)} = \widetilde{x}_{n+1}^{(1)} \tag{59}$$

and $\widetilde{x}_{n+2}^{(2)} = 2\widetilde{x}_{n+1}^{(1)} - \widetilde{x}_n^{(1)}$

Between iterations $p-1$ p, we shall have the relation:

$$x_1^{(p)} = x_1^{(p-1)};\ x_2^{(p)} = \frac{x_1^{(p-1)} + x_2^{(p-1)} + x_3^{(p-1)}}{3};$$
$$x_3^{(p)} = \frac{x_2^{(p-1)} + x_3^{(p-1)} + x_4^{(p-1)}}{3};\dots;\widetilde{x}_{n+p-1}^{(p)} = \widetilde{x}_{n+p-1}^{(p-1)}; \tag{60}$$
$$\widetilde{x}_{n+p}^{p} = 2x_{n+p-1}^{(p-1)} - x_{n+p-2}^{(p-1)}$$

We test if:

$$\left| x_i^{(p)} - x_i^{(p-1)} \right| \leq \varepsilon,\ \forall i \in \{1,2,...,n+p\} \tag{61}$$

If this test is fulfilled, then the iterations are continued until the test is fulfilled for the first (n+p) periods.

Remarks

1) The mobile means method uses a conservative principle, where the invariant is (the mean for 3 periods) under the form:

$$\frac{x_{i-1} + x_i + x_{i+1}}{x_i} = C$$

For C = 3, it follows that the adjusted sequence has the property that:

$$\widetilde{x}_i = \frac{\widetilde{x}_{i-1} + \widetilde{x}_{i+1}}{2},$$

if an arithmetic progression is obtained.

2) If we consider the mobile mean for 5 periods, we have the following relation between two iterations:

$$x_i^{(k)} = \frac{x_{i-2}^{(k-1)} + x_{i-1}^{(k-1)} + x_i^{(k-1)} + x_{i+1}^{(k-1)} + x_{i+2}^{(k-1)}}{5},\ i \in \{1,\ 2,\ ...,\ n+k\}$$

The invariant property of the adjusted sequence will be that term $\widetilde{x}_i$ is an arithmetic mean of two anterior and two posterior terms, namely:

$$\widetilde{x}_i = \frac{\widetilde{x}_{i-2} + \widetilde{x}_{i-1} + \widetilde{x}_{i+1} + \widetilde{x}_{i+2}}{4}$$

3) Similarly to the extensions for the future, extensions for past can be also made (excepting the analyzed periods), namely the terms $\widetilde{x}_0$, $\widetilde{x}_{-1}$, $\widetilde{x}_{-2}$ etc. can be determined (one term at each iteration).

2.3. Tolerance Mobile Means Method

If we use one of the methods of expanding the mean, then two values, a minimum value and a maximum one are obtained at each iteration.

At the same time, we can use a method specific to mobile means.

To this end, at each term $\tilde{x}_i^{(p+k)}$, we allow for the minimum and the maximum values taken for the (p+k) iterations, i.e.:

$$x_{i_{MIN}}^{(p+k)} = ROT\left[Min\left[x_1^{(1)}, x_2^{(2)}, \ldots\ldots, x_i^{(p+k)}\right]\right] \tag{62}$$

and

$$x_{i_{MAX}}^{(p+k)} = ROT\left[Max\left[x_1^{(1)}, x_2^{(2)}, \ldots\ldots, x_i^{(p+k)}\right]\right] \tag{63}$$

Consequently, a forecasting between two confidence limits is obtained. With a probability that can be defined relatively simple, the real trajectory will be in the band described by the curve of the minimum and maximum values.

3.CENTROIDS METHOD

This method consists in three stages:

Stage I. We divide the set of pairs (t_i, x_i) i=$\overline{1,n}$ into three groups: two equal groups and a smaller group (in order to verify the principle of conserving an invariant).

Stage II. The calculation of the centroid coordinates of the two equal groups and setting the law of adjusting the points of the given sequence.

Stage III. Verifying the hypothesis that the principle of conserving an invariant is true according to a feasible deviation (error) ε_a.

Stage IV. We find the forecasting for the *n+1* period. We apply the sliding principle (we eliminate 1 period and introduce the *n+1* period. The procedure is continued until the forecasting for the proposed *T* horizon is obtained.

The method can be applied either to working out a rigid (deterministic) forecasting or to working out a flexible forecasting (with tolerance).

3.1.The Centroids Method in Deterministic Variant

We consider that the sequence $(t_i,\ x_i),\ i \in \{1,\ 2,\ \ldots,\ n\}$ has an uneven number of points, for instance *n=11* points. We divide the points graphic visualization into three groups:

- Two equal groups, for instance group I with five points, and group II, also with five points (in order to get the same confidence in the centroids which we use to determine the law of adjusting);
- One group of a single point that will be used to verifying the acceptance of the adjusting law is left.

We calculate the coordinates of the first centroid, namely:

$$t_{g_1}^{(1)} = \frac{t_1 + t_2 + t_3 + t_4 + t_5}{5}$$
$$x_{g_1}^{(1)} = \frac{x_1 + x_2 + x_3 + x_4 + x_5}{5} \tag{64}$$

We notice that in case when the intervals are equal to 1 year we get:

$$t_{g_1}^{(1)} = \frac{t_1 + t_1 + 1 + t_1 + 2 + t_1 + 3 + t_1 + 4}{5} = \frac{5t_1 + 10}{5} = t_1 + 2 = t_3 \tag{65}$$

namely, the period in the middle of the first group.

Similarly, we calculate the centroid coordinates for the second group of points. We get:

$$t_{g_2}^{(1)} = t_8;$$
$$x_{g_2}^{(2)} = \frac{x_6 + x_7 + x_8 + x_9 + x_{10}}{5} \tag{66}$$

The straight line representing the law of adjusting the sequence points will be:

$$\frac{t - t_3}{t_8 - t_3} = \frac{x - x_{g_1}^{(1)}}{x_{g_2}^{(1)} - x_{g_1}^{(1)}} \tag{67}$$

We verify the 11-th year , what means to replace t with t_n in the above equation and we obtain the adjusted ordinate line $x_{11}^{(1)}$ (for the 11-th year), i.e.:

$$x_{11}^{(1)} = x_{g_1}^{(1)} + (x_{g_2}^{(1)} - x_{g_1}^{(1)})\frac{t_{11} - t_3}{t_8 - t_3} = x_{g_1}^{(1)} + \frac{8}{5}(x_{g_2}^{(1)} - x_{g_1}^{(1)}) = \frac{8}{5}x_{g_2}^{(1)} - \frac{3}{5}x_{g_2}^{(1)} \tag{66}$$

This value should be very close to that of the given sequence, namely to x_{11}. Consequently, if:

$$\left| x_{11}^{(1)} - x_{11} \right| \le \varepsilon_a, \tag{68}$$

We take into consideration the hypothesis where the centroids method is validated.

Further on, we apply the sliding principle. To this end, point (t_1, x_1) is eliminated and point (t_{11}, x_{11}) is introduced. The centroids method is used again. The following coordinates are obtained:

$$t_{g_1}^{(2)} = t_4;\ x_{g_1}^{(2)} = \frac{x_2 + x_3 + x_4 + x_5 + x_6}{5}; \tag{69}$$

$$t_{g_2}^{(2)} = t_9;\ x_{g_2}^{(2)} = \frac{x_7 + x_8 + x_9 + x_{10} + x_{11}}{5}; \tag{70}$$

The new equation of the adjusted straight line is written.

$$\frac{t - t_{g_1}^{(2)}}{t_{g_2}^{(2)} - t_{g_1}^{(2)}} = \frac{x - x_{g_1}^{(2)}}{x_{g_2}^{(2)} - x_{g_1}^{(2)}} \tag{71}$$

In order to get the first forecasted point, we replace in this straight line, the time *t* with the 12-th year. It follows:

$$x_{12}^{(2)} = x_{g_1}^{(2)} + (x_{g_2}^{(2)} - x_{g_1}^{(1)})\frac{t_{12} - t_4}{t_9 - t_4} = x_{g_1}^{(2)} + \frac{8}{5}(x_{g_2}^{(2)} - x_{g_1}^{(1)}) = \frac{8}{5}x_{g_2}^{(2)} - \frac{3}{5}x_{g_2}^{(2)} \tag{72}$$

We proceed analogously, until the proposed *p* forecasting horizon is obtained: the third point is replaced with point $(12,\ x_{12}^{(2)})$ etc.

3.2.Tolerance Centroids Method

In order to apply this method, it is necessary that the terms of the statistical data sequence $x_1,\ x_2,\ \ldots,\ x_n$ to have two values:

- A minimum value $x_1^{\min},\ x_2^{\min},\ \ldots,\ x_n^{\min}$;
- A maximum value $x_1^{\max},\ x_2^{\max},\ \ldots,\ x_n^{\max}$.

If we have only a deterministic sequence of data, then, the two sequences can be constructed, by using a simple relation:

$$x_i^{\min} = x_i - \frac{\varepsilon_a}{2},\ x_i^{\max} = x_i + \frac{\varepsilon_a}{2},\ i = \{1,\ 2,\ ...,\ n\} \tag{73}$$

Further on, the method of mobile centroids is applied for the minimum values sequence and a forecasted sequence of the minimum values and analogously, of the maximum values is obtained.

Finally, the usual rounding is made for both sequences.

Case Study

A comparative case study is presented between advantages/disadvantages of applying the conventional statistics and new statistical methods to work out forecasts in the statistics of the invisible.

	ADVANTAGES	DISADVANTAGES
Conventional statistics application	- Known methods whose accuracy is proved for about three centuries are applied	- requires a high volume of observations - ignores the Universe pulsating feature (alternance between expanding and constricting)
Applying the (non-conventional) new statistics methods	- takes into account the Universe pulsating feature - reduced volume of observations; extremely, a single observation can be used; - (measurements determination)	- Methods unknown by practitioners and even by theoreticians are applied - experience does not exist, (especially for non-forecasted factors) - lower accuracy if we do not know certain factors (but, it is possible that this disadvantage to be scumbled by experts' plausible hypotheses)

CONCLUSIONS

In the present conditions of the computing technique development, the exigencies as regards the quality of primary statistical data processing results have increased, which represent the first-ordinal mean and for which a tolerance ε_i is accepted. Further on, we calculate the second ordinal mean, by using the first processed data, we obtain a higher result, as the experts have acquired a higher volume of knowledge and experience, fact that allows to diminish the tolerance from ε_1 to ε_2. Consequently an *expanding* process of mean takes place, but with a decreasing and antithetic speed. If a compressing process of mean would take place, then, the tolerances would be increasing. For a long period of time, a pulsating alternance of expanding – compressing the mean takes place, recording a recession of increasing – decreasing the tolerance.

The class of the methods presented in antithesis, but with possibilities of co-operation, harmonization represents the generative type methods and will be used in the statistics theory of the invisible, fact that take shapes also in the theory devoted to chaos [11].

A principle of the new non-conventional statistics is that if there is a single measurement, we can consider that it represent the "mean" ($\bar{x}$) obtained by compressing.

REFERENCES

[1] **Albu, L. L.(2008),** ***A Model to Estimate the Composite Index of Economic Activity in Romania***, in Romanian Journal of Economic Forecasting (ISI – Thompson Scientific Master Journal List), Bucharest, Romania, no.2., vol .9;

[2] **Andreica, M. (2006) ,** ***A Model to Forecast the Evolution of the Structure of a System of Economic Indicators,*** in Romanian Journal of Economic Forecasting (ISI – Thompson Scientigic Master Journal List), Bucharest, Romania, no.1., vol .7;

[3] **Aluja , J. Gil (1994),** ***La selection de inversiones en base a criterios diversificades (Investment Selection Base on Diversified Criteria),*** Annals of Royal Academy of Economic And Financial Sciences, Academic Year 1993/1994, pp 129-157;

[4] **Andreica, M., Stoica, M., Luban F. (1998),** ***Metode cantitative in management, (Management Quantitative Methods),*** Economica Publishing House, Bucharest;

[5] **Bellman, R.L. , Zadeh L.A. (1970),** ***Decision-making in a Fuzzy Environment,,*** Management Science, 17, No.4;

[6] **Bulz, N., Stefanescu, V., Mariano, L.B., Botezatu M., Stoica, M.,** ***Parteneriat creativ de bunastare (Creative Partnership of Welfare),*** AISTEDA Publishing House, Bucharest, Romania;

[7] **Dragota, V., Dumitrescu, D., Ruxanda, G., Ciobanu, A,. Brasoveanu, I., Stoian A., Lipara, C. (2007),*****Estimation of Control Premium The Case of Romanian Listed Companies,*** in Journal of Economic Computation and Economic Cybernetics Studies and Research, Bucharest, Romania, Archive Page 2007, ISSUES 3- 4 (ISI - Thompson Scientific Master Journal List).

[8] **Doval, E., Stoica, M., Nicolae, R.D., Teodoru, G., Ungureanu, G. (2007),** ***Knowledge to Understand the Natural Language Subtlety in Business Environment,*** in Computational Intelligence Applied to New Digital Economy and Business Ecosystems, Proceeding of the XIV Congress of International Association for Fuzzy-Set Management and Economy SIGEF, Universitaria Publishing, pp. 690-699;
[9] **Dubois, D.M , Sabatier , Ph. (1998),** ***For a Naturalist Approach to Anticipation from Catastrophe Theory to Hyperincursive Modelling,****, in* CASYS International Journal of Computing Anticipatory System (IJICAS), International Conference CASYS'98 on Computing Anticipatory Systems, Liège, Belgium , D.M. Dubois Ed), August 10 -14;
[10**] Kauffmann, A., Aluja, J.G. (1990),** ***Las matematicas del azar y de la incertidumbre***, Editorial Ceura, Madrid;
[11] **Mitrut, C., Constantin, D., Dimian, G., Dimian, M. (2007), *,Indicators and Methods for Characteristing Regional Specialization and Concentration,*** in Journal of Economic Computation and Economic Cybernetics Studies and Research, Bucharest, Romania; ISSUES 3- 4 , (ISI - Thompson Scientific Master Journal List);
[12] **Negoita, C.V. (2003),** ***Vag (Vague),*** Paralela 45 Publishing House, Pitesti;
[13] **Nicolae, D, Andreica, M. (2008),** ***Modeles Structureles Utilisez pour Optimiser les Investissements*** , Simpozionul International - Dezvoltarea Sociala si Performanta Economica, a XII –a Sesiune de Comunicari Stiintifice, Academia Comerciala Satu Mare, Satu Mare, Romania, 20-21 Iunie; (Included in EBSCO international data base);
[14] **Nicolae D., Lupasc, I. (2008),** ***Modelarea matematica a conceptului de dezvoltare pe termen lung si sustenabila (Mathematical Modeling of the Long-term and Sustainable Development Concept),*** Journal Studii si cercetari de calcul economic si cibernetica economica, No. 1, ASE Publishing House, Bucharest;
[15] **Odobleja, St. (1938),** ***Psychologie consonnantiste*** **,** Meloine Publishing House, Paris ;
[16] **Osmatescu , P. (1997),** ***Basis of the Subtle Spaces and Algebraical Structures*****,** in Scripta scientiarum mathematicarum, tomul I, fasciculul I., Chisinau Publishing House, pp. 96-209;
[17] **Pau V., Stoica, M., Nicolae, D. (2008),** ***Mathematical Model of Economic Forecasting and Analysis of The Market Competitive Mechanism,*** Proceedings of the 10th IBIMA Conference on Innovation and Knowledge Management in Business Globalization , Kuala Lumpur, Malaysia, 30 June – 2 July, (ISI classification);
[18] **Paun , Gh. (1983),** ***An Impossibility Theorem for Indicators Aggregation, in Fuzzy Sets and Systems,*** North Holland Publishing Company, 9 pp. 205-210;
[19]**Ruxanda, Gh., Botezatu, A., (2008),** ***Spurious Regression and Cointegration. Numerical Example Romania's M2 Money Demand,*** Romanian Journal of Economic Forecasting, no. 3, Bucharest; (ISI classification).
[20**] Serban, D. Mitrut, C., Christache, S. (2008),** ***Statistical Tests to Evaluate the Level of Migration Intention from Village Town in Romania,*** in Journal of

---

Economic Computation and Economic Cybernetics Studies and Research, (ISI - Thompson Scientific Master Journal List) , vol I-II, ISSUES 1-2;

[21] **Stoica M., Andreica M., Nicolae, D., Cantau, D. (2006),** ***Methods and Models of Forecasting*** Economic, University Publishing House, Bucharest;

[22] **Stoica, M., Doval, E. (2003),** ***Aplicatii ale operatorului de act in economia firmei (Applications of the Operator of Act in Firm's Mangement)*** , (In memoriam of the Romanian mathematician, Petre Osmatescu) in ARA Journal ,Montreal University, vol. 25;

[23] **Stoica, M., Andreica, M., Nicolae, D., Andreica, R. (2008),** ***Multimile subtile si aplicatiile lor (Subtle Sets and their Applications),*** Cibernetica Publishing House, Bucharest;

[24] **Stoica, M., Andreica, M., Nicolae, D., Cantau, D. (2006),** ***Metode si modele de previziune economica (Methods and Models of Economic Forecasting),*** Universitara Publishing House, Bucharest;

[25] **Stoica, M., Nicolae D, Doval E.,Lupasc I., Negoescu Gh.(2008)** ***Asymmetric Modelling of Sustainable Development Decision-Making for Urban Agglomerations Effective Management,*** Proceedings of the 14th International Congress of Cybernetics and Systems of WOSC , Wroclaw, Poland, September 9-12;

[26] **Stoica, M., Nicolae, D., Ungureanu, M.A.,,Andreica, A., Andreica, M. (2008),** ***Fuzzy Sets*** **and** ***their Aplications***, Proceedings of the 9th WSEAS International Conference on Mathematics & Computers in Business & Economics (MCBE'08), Bucharest, Romania, June 24 – 26, **(ISI classification).**

[27] **Zadeh, L.A. (1974),** ***Fuzzy Logic and its Applications to Approximate Reasoning,*** in Information Processing 3, IFIP Congress, 5-10 August, Stockolm;

[28] **Zadeh, L.A. (1965),** ***Fuzzy Sets,*** Info&Cth., Vol. I, pp. 338-353.