M. VINCZE
B. SAMOCHIŞ
D. STEGĂROIU

Procedure for Measuring the Influence of Factors în Functional Type Relations

"Can we conclude at the same time from cause to effect and from effect to cause?"

Aristotle — Organon

The evaluation of factor influence on the variation of the resultant phenomenon, on the basis of a statistical sample, is made by coefficients of determination; such evaluations do require however difficult computations and procedures differentiated according to the nature of relation by which there is shown the constraint between the phenomenon under investigation and its factors [3]. In the present work there is shown a method for the computation of the coefficients of determination of general practicability for the class of differentiable functions, in which forms of functional relations important from the viewpoint of economic analysis are included.

The method is based on the individualization of a simple procedure of the decomposition of the variance of dependent variable (s_y^2) on account of the variance of each argument.

If $y = f(x_1, x_2, ..., x_n)$ represents the form of functional relation which shows the link between the dependent variable y and the arguments x_1, x_2, ..., x_n, then the following relation can be deduced [1]):

$$s_y^2 \approx f'^2_{x_1}(M)\, s^2_{x_1} + f'^2_{x_2}(M)\, s^2_{x_2} + ... + f'^2_{x_n}(M)\, s^2_{x_n} + ... \\ + 2f'_{x_1}(M)\, f'_{x_2}(M)\, s_{x_1x_2} + ... + 2f'_{x_{n-1}}(M)\, f'_{x_n}(M)\, s_{x_{n-1}x_n} \qquad (1)$$

[1]) See appendix 1.

where: $M = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ namely in the partial derivatives there are replaced the mean values of the stochastic variables $x_1, x_2, ..., x_n$, $s^2_{x_i}$ = the variances of the variables and $s_{x_i x_j}$ = their covariance.

By this formula there is achieved a practical procedure to measure the influence of various factors on the variance of the resulting variable.

The evaluation of the validity of the formula (1) has also been made through the application of the same forms of functional constraints for which adequate computation formulas have been derived in the work [3].

Thus for $y = \sum_{i=1}^{n} x_i$ according to the relation (1) it results:

$$s_y^2 \approx s^2_{x_1} + s^2_{x_2} + ... s^2_{x_n} + 2s_{x_1 x_2} + ... + 2s_{x_{n-1} x_n}$$

for:

$$y = \prod_{i=1}^{n} x_1$$

$$s_y^2 \approx \bar{x}_2 ... \bar{x}_n s^2_{x_1} + ... + \bar{x}_1 ... \bar{x}_{i-1} \bar{x}_{i+1} ... \bar{x}_n s^2_{x_i} + ... \bar{x}_1 ...$$

$$... \bar{x}_{n-1} s^2_{x_n} + 2\bar{x}_1 \bar{x}_2 \bar{x}_3^2 ... \bar{x}_n^2 s_{x_1 x_2} + ... + 2\bar{x}_{n-1} \bar{x}_n \bar{x}_1^2 ... \bar{x}_{n-2}^2 s_{x_{n-1} x_n}$$

analogous for:

$$y = \frac{x_1}{x_2}$$

$$s_y^2 \approx \frac{\bar{x}_1^2}{\bar{x}_2^2} \left[\frac{s^2_{x_1}}{\bar{x}_1^2} + \frac{s^2_{x_2}}{\bar{x}_2^2} - 2 \frac{s_{x_1 x_2}}{\bar{x}_1 \bar{x}_2} \right]$$

The results obtained through the application of the contemplated formula (1) to the three forms of functional relations are identical with those inferred in the mentioned work.

Owing to the economic importance of certain relations, such as the Cobb-Douglas, functions, the quadratic function and the radical sign function, there shall be further presented the procedure of decomposing the variance of the dependent variable for each of them.

The Cobb-Douglas function:

$y = Kx^{a_1} x^{a_2} \cdots x^{a_n}$, computing the partial derivatives we get:

$$\begin{aligned} f'_{x_1} &= k a_1 x_1^{a_1 - 1} x_2^{a_2} ... x_n^{a_n} \\ &..... \\ f'_{x_n} &= k a_n x_1^{a_1} x_2^{a_2} ... x_n^{a_n - 1} \end{aligned} \qquad (2)$$

hence:

$$s_y^2 \approx k^2 \bar{x}_1^{a_1} \bar{x}_2^{a_2} ... \bar{x}_n^{a_n} \left[a_1^2 \frac{s^2_{x_1}}{\bar{x}_1} + ... + a_i^2 \frac{s^2_{x_i}}{\bar{x}_i} + ... + \right.$$

$$\left. + a^2 \frac{s^2_{x_n}}{\bar{x}_n} + 2a_1 a_2 \frac{s_{x_1 x_2}}{\bar{x}_1 x_2} + ... + 2a_{n-1} a_n \frac{s_{x_{n-1}} x_n}{\bar{x}_{n-1} \bar{x}_n} \right]$$

The quadratic function:

$$y = b_1x_1^2 + b_2x_2^2 + 2b_3x_1x_2 + b_4x_1 + b_5x_2$$
$$f'_{x_1} = 2b_1x_1 + 2b_3x_2 + b_4 \tag{3}$$
$$f'_{x_2} = 2b_2, x_2 + 2b_3x_1 + b_5$$

hence:

$$s_y^2 \approx (2b_1\bar{x}_1 + 2b_3\bar{x}_2 + b_4)^2 s_{x_1}^2 + (2b_2\bar{x}_2 + 2b_3\bar{x}_1 + b_5)^2 s_{x_2}^2 +$$
$$+ 2(2b_1x_1 + 2b_3\bar{x}_2 + b_4)(2b_2\bar{x}_2 + 2b_3\bar{x}_1 + b_5)\, s_{x_1x_2}$$

The radical sign function:

$$y = b_1x_1 + b_2x_2 + b_3\sqrt{x_1x_2} + b_4\sqrt{x_1} + b_5\sqrt{x_2} \tag{4}$$
$$f'_{x_1} = b_1 + \frac{b_3\sqrt{x_2}}{2\sqrt{x_1}} + \frac{b_4}{2\sqrt{x_1}}$$
$$f'_{x_2} = b_2 + \frac{b_3\sqrt{x_1}}{2\sqrt{x_2}} + \frac{b_5}{2\sqrt{x_2}}$$

Hence:

$$s_y^2 \approx \left(b_1 + \frac{b_3\sqrt{\bar{x}_2}}{2\sqrt{\bar{x}_1}} + \frac{b_4}{2\sqrt{\bar{x}_1}}\right)^2 s_{x_1}^2 + \left(b_2 + \frac{b_3\sqrt{\bar{x}_1}}{2\sqrt{\bar{x}_2}} + \frac{b_5}{2\sqrt{\bar{x}_2}}\right)^2 s_{x_2}^2 +$$
$$+ 2\left(b_1 + \frac{b_3\sqrt{\bar{x}_2}}{2\sqrt{\bar{x}_1}} + \frac{b_4}{2\sqrt{\bar{x}_1}}\right)\left(b_2 + \frac{b_3\sqrt{\bar{x}_1}}{2\sqrt{\bar{x}_2}} + \frac{b_5}{2\sqrt{\bar{x}_2}}\right) s_{x_1x_2} \tag{5}$$

Disposing of a computation formula [1] which allows the decomposition of the variance of dependent variable y for a wide class of functions, it remains to define the very relation according to which the coefficient of determination is computed.

Pointing out the characteristic that, in case of some functional constraints between variables, the coefficient of total determination is always equal to one, it results that the problem consists only in establishing the coefficients of partial determination.

In quantifying the influence of each separate factor on the variance of the resulting variable, we may start from the following assumption:

$$\text{I.} \sum_{i=1}^{n} d_{y\cdot x_i} = 1\,(5) \text{ and II. } \sum_{i=1}^{i} d_{y\cdot x_i} + \sum_{i,j=1}^{n} d_{y\cdot x_ix_j} = 1^{2)} \tag{6}$$

---

[2]) There has been used the simplified notation:

$$d_{y\cdot x_i} = d_{yx_i\cdot x_1x_2\ldots x_{i-1}x_{i+1}\ldots x_n}$$
$$d_{y\cdot x_ix_j} = d_{yx_i\cdot x_jx_1x_2\ldots x_{i-1}x_{i+1}\ldots x_{j-1}x_{j+1}\ldots x_n}$$

Being based on the first assumption the coefficient of determination is defined as follows:

$$d^{(1)}_{y\cdot x_i} = \frac{f^2_{x_i}(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)\, s^2_{x_i}}{\sum_{h=1}^{n} f'^2_{x_h}\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)\, s^2_{x_n}} \tag{7}$$

$$d^{(1)}_{y\cdot x_i x_j} = \frac{2f'_{x_i}(x_1, x_2, ..., \bar{x}_n)\, f'_{xj}(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n) s_{x_i x_j}}{\sum_{h=1}^{n} f'^2_{xh}(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)\, s^2_{x_h}}$$

or accepting the second assumption:

$$d^{(2)}_{y\cdot x_i} = \frac{f'^2_{x_i}(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)\, s^2_{x_i}}{s^2_y}$$

$$d^{(2)}_{y\cdot x_i x_j} = \frac{2f'_{x_i}(\bar{x}_1, \bar{x}_2 ..., \bar{x}_n)\, f'_{x_j}(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)\, s_{x_i x_j}}{s^2_y} \tag{8}$$

Evidently that

$$d^{(1)}_{y\cdot x_i} \geqslant d^{(2)}_{y\cdot x_i} \text{ and } d^{(1)}_{y\cdot x_i x_j} \geqslant d^{(2)}_{y\cdot x_i x_j}$$

While $d^{(1)}_{y\cdot x_i}$ measures the weight of the influence of factor i in the total direct influence of factors, $d^{(2)}_{y\cdot x_i}$ emphasizes the weight of the direct influence of factor i in the total influence of factors.

In order to materialize the above said, there has been proceeded to the study, on a test sample of 61 agricultural producer co-operatives, of a number of six factors which determine the cost price of cow's milk (y) i.e.: labour force expenses (x_1); fodder expenses (x_2); other direct expenses (x_3); indirect expenses (x_4); average annual output on each fodder fed cow (x_5) and mean number of fodder fed cows (x_6).

The link between factors and the phenomenon investigated is of functional type having the form:

$$y = \frac{x_1 + x_2 + x_3 + x_4}{x_5 x_6} \tag{9}$$

The quantification of factor influence on the variation of cost price has an economic importance through the identification of those of them by help of which there may be conditioned the variation of the phenomenon under study.

In view of applying the methodology presented, there is necessary the preliminary computation of the coefficients of simple correlation between factors (table 1).

Table 1

The coefficients of simple correlation between arguments

The variables	x_1	x_2	x_3	x_4	x_5	x_6
x_1	1.000	0.714	0.332	0.777	0.518	0.834
x_2		1.000	0.367	0.785	0.375	0.823
x_3			1.000	0.290	0.178	0.303
x_4				1.000	0.489	0.779
x_5					1.000	0.361
x_6						1.000

Remark. Taking into account that $\sqrt{x_i x_j} = \sqrt{x_j x_i}$, only dissimilar correlations have been included in the table.

In order to che. the meaning of the coefficients of correlation there has been computed $\Gamma_{min} = \dfrac{\sqrt{n+36} - \sqrt{n}}{6}$, getting for $n = 61$, $\Gamma_{min} = 0.340$. By comparing with the data in the table 1, there may be ascertained that a series of values — Γ_{x_1x}, $\Gamma_{x_3x_4}$, $\gamma_{x_3x_5}$, $\gamma_{x_3x_6}$, are not characteristic. The high values of coefficients $\Gamma_{x_1x_6}$, $\Gamma_{x_3x_6}$, $\Gamma_{x_4x_6}$ are, in exchange, pointing to the existence of close linear constraints between the mentioned variables, which is logical, because the expenses depend firstly on the number of animals. On the basis of these grounds, we may conceive that x_1, x_2, and x_4 are linear functions of x_6 — which were tested by the linearity criterion and of which must be kept account when computing the partial derivative of function y in relation to x_6.

With a view to a correct economic interpretation of results, there have been computed the values $\Gamma_{yx_5} = -0.442$ and $\Gamma_{yx_6} = -0.085$ which outlines an inverse correlation between the cost price and the milk average yearly output, and the non-existence of a linear constraint of interdependence between the cost price and the number of dairy cows.

As soon as these aspects are clarified, there may be proceeded to the establishment of the coefficients of determination. For this purpose, there have been computed the partial derivatives of function $y = f(x_1, x_2, x_3, x_4, x_5, x_6)$, resulting:

$$f'_{x_1} = f'_{x_2} = f'_{x_3} = f'_{x_4} = \frac{1}{x_5 x_6}$$

$$f'_{x_5} = \frac{x_1 + x_2 + x_3 + x_4}{x_5^2 x_6}$$

$$f'_{x_6} = \frac{(b_1 + b_2 + b_4)\, x_5 x_6 - (x_1 + x_2 + x_3 + x_4)\, x_5}{x_5^2 x_6^2}$$

where b_1, b_2, b_4 are the coefficients of regression of the equations:

$$x_1 = b_1 x_6 + a_1 + e_1$$

$$x_2 = b_2 x_6 + a_2 + e_2$$

$$x_4 = b_4 x_6 + a_4 + e_4$$

In order to compute the coefficients of determination there has been used the formula (7), the total influences of factors being thus quantified, without taking also into account the means by which these influences are exerted, a procedure justified by the very nature of the problem analysed in the work. The results obtained are included in table 2.

Table 2

The coefficients of determination

$dy \cdot x_i$		I		$dy \cdot x_i x_j$			
i	j	x_1	x_2	x_3		x_5	x_6
x_1	0.3124		0.4004	0.0210	0.2586	0.3218	0.1694
x_2	0.2522			0.2090	0.2348	0.2094	0.1502
x_3	0.0032				0.0098	0.0112	0.0062
x_4	0.0889					0.1621	0.0844
x_5	0.3096						0.0729
x_6	0.0336						

There has to be remarked the large weight of the influence of average output (30,9%), of fodder expenses (25,2%) as well as of the labour force expenses (31,2%) on the variation of the cost price for milk, on the whole accounting for 87,3%. The degree of affectation of labour force expenses is relatively high, owing firstly to the considerable labour force employment in the conditions of a reduced mechanization, as well as to the amplitude of the variation of the remuneration of labour force in the agricultural co-operatives within the test sample.

The procedure suggested ensures a simple course, of general practicability for the separation and measuring of factor influence in the case of relations which are modelling the most important economic phenomena.

Appendix 1

Let $y = f(x_1, x_2, ..., x_n)$ a function of n variables differentiable in the field D.

The overall function increase in a point $M(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ where $M \in D$, may be expressed by the following form [1]:

$$\Delta y_M \approx f'_{x_1}(M)\Delta x_1 + f'_{x_1}(M)\Delta x_2 + ... + f'_{x_n}(M)\Delta x_n \tag{1}$$

We regard a system of N values in the vicinity of point M, belonging to the field D:

$$\begin{array}{l} x_{11}x_{21} \dots x_{j1} \dots x_{n1} \\ \dots\dots\dots\dots\dots\dots\dots \\ x_{1i}x_{2i} \dots x_{ji} \dots x_{ni} \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ x_{1N}x_{2N} \dots x_{jN} \dots x_{nN} \end{array} \tag{2}$$

so that by replacing these points in (1) the relation should remain valid, that is:

$$\Delta y_{i_M} \approx f'_{x_1}(M)\Delta x_{1_i} + f'_{x_2}(M)\Delta x_{2_i} + ... + f'_{xj}(M)\Delta x_{ji} + ... +$$

where:

$$+ f'_{xn}(M)\,\Delta x_{ni}, \qquad i = \overline{1,\ N} \tag{3}$$

$$\Delta x_{ji} = x_{ji} - \bar{x}_j \qquad j = \overline{1,\ n} \tag{4}$$

$$\Delta y_{i_M} = y_i - \bar{y} \qquad i = \overline{1,\ N}$$

noting by

$$\bar{y} \approx f(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$$

$$y_i = f(\bar{x}_{1i}, \bar{x}_{2i}, ..., x_{ni})$$

Squaring both parts of relations (3) and totalling the units in columns there is obtained:

$$\sum_{i=1}^{N} \Delta^2 y_{i_M} \approx f'^2_{x_1}(M)\sum_{i=1}^{N}\Delta^2 x_{1i} + ... + f'^2_{x_j}(M)\sum_{i=1}^{N}\Delta^2 x_{ji} + ... +$$

$$+ f'^2_{x_n}\sum_{i=1}^{N}\Delta^2 x_{n_i} + 2f'_{x_1}(M)f'_{x_n}(M)\sum_{i=1}^{N}\Delta x_{1i}\Delta x_{2i} + ... +$$

$$+ 2f'_{x_{n-1}}(M)f'_{x_n}(M)\sum_{i=1}^{N}\Delta x_{(n-1)i}\Delta x_{ni} \tag{5}$$

Dividing each term by N it results:

$$\frac{\sum_{i=1}^{N}\Delta^2 y_{i_M}}{N} \approx f''^{2}_{x_n}(M)\frac{\sum_{i=1}^{N}\Delta^2 x_{1i}}{N} + \ldots + f''^{2}_{x_n}(M)\frac{\sum_{i=1}^{N}\Delta^2 x_{n_i}}{N} +$$

$$+ 2f'_{x_1}(M)f'_{x_2}(M)\frac{\sum_{i=1}^{N}\Delta^x 1i\Delta^x 2i}{N} + \ldots +$$

$$+ 2f'_{x_{n-1}}(M)f'_{x_n}(M)\frac{\sum_{i=1}^{N}\Delta^x(n-1)\,i\Delta^x ni}{N}. \tag{6}$$

Considering:

$$\frac{\sum_{i=1}^{N}\Delta^2 y_{iM}}{N} = \frac{\sum_{i=1}^{N}(y_i - \bar{y})^2}{N} \approx s_y^2; \quad \frac{\sum_{i=1}^{N}\Delta^2 xji}{N} =$$

$$= \frac{\sum_{i=1}^{N}(x_{ji} - \bar{x}_j)^2}{N} = s_{xj}^2, \; j = \overline{1,\, n}$$

and

$$\frac{\sum_{i=1}^{N}\Delta^x{}_{ji}\,\Delta^x{}_{ki}}{N} = \frac{\sum_{i=1}^{N}(x_{ji} - \bar{x}_j)(x_{ki} - \bar{x}_k)}{N} = s_{x_j x_k} \tag{7}$$

$$j = \overline{1,\, n}$$

$$k = \overline{1,\, n}$$

$$j \neq k$$

It reaches to the formula:

$$s_y^z \approx f'_{x_1}(M)\, s_{x_1}^2 + \ldots + f'^{2}_{x_n}(M)\, s_{x_n}^2 + 2f'_{x_1}(M)\cdot f'_{x_2}(M)\, s_{x_1 x_2} + \ldots +$$

$$+ 2f'_{x_{n-1}}(M)\; f'_{x_n}(M)\; s_{x_{n-1}x_n} \tag{8}$$

where: $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ — stands for the mean values of stochastic variables $x_1, x_2, \ldots, x_n$

$s_{x_1}^2, \ldots, s_{x_n}^2$ — stands for the variances of variables $x_1, \ldots, x_n$

$s_{x_j x_k}$ — stands for the covariances of variables $x_j, \ldots, x_k$

Of the way the formula (8) is inferred it is remarked that the approximations are the more exact so as the deviations of values compared to the mean are smaller.

REFERENCES

1. FIHTENHOLT M.G.: *Curs de calcul diferențial și integral.* Vol. II — Translated from Russian language, București, Ed. tehnică, 1964.
2. MARINESCU I., MOINEAGU G., NICULESCU R., RANCU N., URSEANU V. *Elemente de statistică matematică și aplicațiile ei.* Ed. științifică, 1966.
3. MARIA VINCZE, SAMOCHIȘ B., TĂRĂU GH., STEGĂROIU D.: *Procedee statistice de separare și măsurare a influenței factorilor.* Revista de statistică nr. 10/1969.

PROCEDES POUR MESURER L'INFLUENCE DES FACTEURS SUR LES RELATIONS DU TYPE FONCTIONNEL

L'estimation de l'influence des facteurs sur la variation d'un phenomène résultatif par les coefficients de détermination, recquiert des calculs difficiles, obligeant de recourir à des procédés différenciés selon la forme de la relation de dépendance.

Dans le présent article, les auteurs exposent une méthode générale d'estimer les coefficients de détermination au cas des relations du type fonctionnel, utilisant la décomposition de la dispersion de la variable dépendante à l'aide des dispersions des variables indépendantes.

СПОСОБ ИЗМЕРЕНИЯ ВЛИЯНИЯ ФАКТОРОВ В ОТНОШЕНИЯХ ФУНКЦИОНАЛЬНОГО ТИПА

Оценка влияния факторов на вариацию отдельного результативного явления при помощи определяющих коэффициентов требует трудоемких расчетов и применения методов, дифференцированных по форме связи зависимости.

Авторы указывают в настоящей статье общий метод оценки определяющих коэффициентов в случае отношений функционального типа, основывающийся на разбивке дисперсии зависимой переменной за счет дисперсий независимых переменных.