ᲡᲐᲥᲐᲠᲗᲕᲔᲚᲝᲡ ᲛᲔᲪᲜᲘᲔᲠᲔᲑᲐᲗᲐ ᲔᲠᲝᲕᲜᲣᲚᲘ ᲐᲙᲐᲓᲔᲛᲘᲘᲡ ᲛᲝᲐᲛᲑᲔ, ტ. 7, №3, 2013
BULLETIN OF THE GEORGIAN NATIONAL ACADEMY OF SCIENCES, vol. 7, no. 3, 2013

Economics

On the Multiple Linear Regression in Marketing Research

Nugzar Todua[*], **Petre Babilua**[**] **and Teona Dochviri**[§]

[*] *Faculty of Economics and Business, I. Javakhishvili Tbilisi State University, Tbilisi*
[**] *Faculty of Exact and Natural Sciences, I. Javakhishvili Tbilisi State University, Tbilisi*
[§] *Faculty of Applied Mathematics and Computer Sciences, I. Javakhishvili State University, Tbilisi*

(Presented by Academy Member Avtandil Silagadze)

ABSTRACT. Marketing research holds an important place in modern market economy. Marketing information on the main economic indexes is used in planning, control and prediction of business. Modern business companies widely use social webs. In current economic crisis, social webs turned out to be the most popular raising the awareness of brands and popularizing them. They form the company's image, gain new customers for them, encourage them to achieve success with the least expanses. Georgian companies actively followed this way and created their profiles on social webs. Georgian customers are also actively involved in social networks. They take part in different games offered by these companies, express their ideas about various products and services. Communication has become more interactive. Regression and correlation analyses play the key role in studying the problems related to this topic. A new method is proposed for application of multi-dimensional linear regression model in marketing research. Namely, a multidimensional linear regression model is constructed using linear regression models with one predictor and, vice versa. Linear models with one predictor are constructed using a multi-dimensional linear model. Numerical examples are also given in the present work. *©2013 Bull. Georg. Natl. Acad. Sci.*

Key words: *marketing information, predictor, regression, least-square method, linear model.*

1. The subject of marketing research consists in the analysis of business undertakings, reclamation and product sales. Results of this analysis are used in the end to promote effective business activities [1,2]. There are a great number of systems of receiving and processing the marketing information on socio-political and economic indexes. In the investigation of these systems, essential use is made of statistical methods, in particular of correlation and regression analysis, which is used in planning, control, monitoring and prediction of the monetary and fiscal policy of macroeconomics [3-7].

One of the important problems is the study of the behavior of dependent marketing (economic) variables (indexes) relative to independent variables (predictors). This problem can be solved with the aid of various regression models.

A decision on the choice and quantity of endogenous and exogenous indexes depends on a concrete problem. For a practical application of a regression model it is necessary to carry out experiments and receive statistical data, which, in its turn, is connected with various expenses.

Let us, for example, assume that there is information on some dependent marketing variable relative to two predictors. It is interesting to construct a two-dimensional regression model of this dependent variable without additional experimentation. This is especially important when there is a need to receive observation data in the case of expensive technological and engineering complex experiments.

In this paper we construct the multi-dimensional linear regression model by using one-dimensional linear regression models and, vice versa, by using the multi-dimensional linear model we construct one-dimensional models.

2. Let us consider some variable *Y* whose behavior depends on *k* predictors $x_1,\dots,x_k$.

Assume that we have *n* data of observations *Y* separately from the predictors $x_1,\dots,x_k$:

$$\left(x_{1i},Y_{1i}\right),\dots,\left(x_{ki},Y_{ki}\right),\quad i=1,\dots,n. \tag{1}$$

We consider two variants of construction of the multi-dimensional linear model by using data (1).

Variant 1. Equations of a simple linear regression for each predictor x_j relative to Y_j, $j=1,\dots,k$, are written by the following scheme:

$$x_{ji}=b_{j0}+b_{j1}Y_{ji},\quad j=1,\dots,k;\quad i=1,\dots,n, \tag{2}$$

where for each *j* we afterwards replace Y_j by the values of other variables $Y_1,\dots,Y_j,Y_{j+1},\dots,Y_k$,. Thus we obtain the following scheme:

1) for x_1 we have

$$\begin{aligned} x_{1i}^{(1)}&=b_{10}+b_{11}Y_{1i},\\ x_{1i}^{(2)}&=b_{10}+b_{11}Y_{2i},\\ &\dots\dots\dots\dots ,\\ x_{1i}^{(k)}&=b_{10}+b_{11}Y_{ki}, \end{aligned}$$

and so on;

k) for x_k we have

$$\begin{aligned} x_{ki}^{(1)}&=b_{k0}+b_{k1}Y_{ki},\\ x_{ki}^{(2)}&=b_{k0}+b_{k1}Y_{1i},\\ &\dots\dots\dots\dots ,\\ x_{ki}^{(k)}&=b_{k0}+b_{k1}Y_{(k-1)i}, \end{aligned}$$

where $i=1,\dots,n$.

We introduce new variables by means of the equalities

$$\overline{x}_{1i}=\frac{1}{k}\sum_{j=1}^{k}x_{1i}^{(j)}, \tag{3}$$

$$\dots\dots\dots\dots\dots\dots$$

$$\overline{x}_{ki}=\frac{1}{k}\sum_{j=1}^{k}x_{ki}^{(j)}, \tag{4}$$

$$\overline{Y}_i=\frac{1}{k}\sum_{j=1}^{k}Y_{ji}, \tag{5}$$

where $i=1,\dots,n$. Thus we obtain the set of data $\left(\overline{x}_{1i},\dots,\overline{x}_{ki},\overline{Y}_i\right)$, $i=1,\dots,n$, by means of which we can write the equation of multi-dimensional linear regression

$$\hat{Y}^{(1)}=b_0^{(1)}+b_1^{(1)}\overline{x}_k+\dots+b_k^{(1)}\overline{x}_k+\varepsilon_1, \tag{6}$$

where the coefficients $b_e^{(1)}$, $e=0,1,\dots,k$, are uniquely defined by the values $\overline{x}_{ji}$, $j=1,\dots,k$, $i=1,\dots,n$, by virtue of equalities (3), (4). In equation (6), $\varepsilon_1=\varepsilon_{1i}$ are normally distributed random values $E\left(\varepsilon_{1i}\right)=0$ with mathematical expectation and $D\left(\varepsilon_{1i}\right)=\sigma_1^2>0$ with unknown dispersion; also $E\left(\varepsilon_{1i}\cdot\varepsilon_{1j}\right)=0$, $i,j=1,\dots,n$, $i\neq j$. Note that the values ε_{1i} are called the errors of model (6). Let us consider the so-called remainders

$$e_{1i}=\overline{Y}_i-\hat{Y}^{(1)}\left(\overline{x}_{1i},\dots,\overline{x}_{ki}\right), \tag{7}$$

where $i=1,\dots,n$. Using remainders (7), the unknown dispersion σ_1^2 is estimated as follows:

$$s_1^2=\frac{\sum_{i=1}^{n}e_{1i}^2}{n-(k+1)}. \tag{8}$$

Variant 2. Let us write multi-dimensional linear equations of regression for each predictor x_j relative to the variables Y_j, $j=1,\dots,k$. Thus we have

$$\begin{aligned}\hat{x}_{1i}&=b_{10}^{(1)}+b_{11}^{(1)}Y_{1i}+\dots+b_{1k}^{(1)}Y_{ki},\\ &\dots\dots\dots\dots\\ \hat{x}_{ki}&=b_{k0}^{(k)}+b_{k1}^{(k)}Y_{1i}+\dots+b_{kk}^{(k)}Y_{ki},\end{aligned} \tag{9}$$

where $i=1,\dots,n$. We consider the new variables $(\hat{x}_{1i},\dots,\hat{x}_{ki},\overline{Y}_i)$, where the values $\overline{Y}_i$ are defined by (5) and construct the multi-dimensional linear model

$$\hat{Y}^{(2)}=b_0^{(2)}+b_1^{(2)}\hat{x}_1+\dots+b_k^{(2)}\hat{x}_k+\varepsilon_2. \tag{10}$$

Thus, using data (1) we have obtained two equations of multi-dimensional linear regression. Let σ_2^2 be the estimate of the unknown dispersion $D(\varepsilon_{2i})$, $i=1,\dots,n$. From two models (6) and (10) we should choose the model with the smallest estimate of dispersion s_1^2, s_2^2. We denote this model as follows:

$$\hat{Y}=b_0+b_1\tilde{x}_1+\dots+b_k\tilde{x}_k+\varepsilon, \tag{11}$$

where

$$\hat{Y}=\begin{cases}\hat{Y}^{(1)}, & s_1^2<s_2^2,\\ \hat{Y}^{(2)}, & s_1^2>s_2^2.\end{cases} \tag{12}$$

If $s_1^2=s_2^2$, then we with equal preference choose one of the two models (6), (10). Note that equation (11) enables us to carry out a complete regression analysis of the behavior of the variable *Y*: testing the hypothesis for the coefficients b_j, $j=0,1,\dots,k$, calculating the determination coefficients and so on.

Using data (1) we can construct models with a different number of predictors. The total number of such models is equal to

$$C_n^1+\dots+C_n^n=\sum_{i=1}^{n}C_n^i=2^n-1.$$

3. Let us assume that we have data of *y* observations of the dependent variable *Y* relative to the predicators $x_1,\dots,x_k$:

$$(x_{1i},\dots,x_{ki},Y_i),\quad i=1,\dots,n. \tag{13}$$

Then we can derive the equation of multi-dimensional linear regression

$$Y=a_0+a_1x_1+\dots+a_kx_k. \tag{14}$$

We wish to write equations of one-dimensional linear regressions of the variable *Y* for the predictors $x_1,\dots,x_k$ separately. For this, we first write the equation of one-dimensional linear regressions for each predictor relative to all other predictors. Thus we obtain

$$\begin{cases}x_2=a_{01}^{(2)}+a_{11}^{(2)}x_1,\\ x_3=a_{01}^{(3)}+a_{11}^{(3)}x_1,\\ \dots\dots\dots\dots\\ x_k=a_{01}^{(k)}+a_{11}^{(k)}x_1,\end{cases} \tag{15}$$

$$\begin{cases}x_1=a_{02}^{(1)}+a_{12}^{(1)}x_2,\\ x_3=a_{02}^{(3)}+a_{12}^{(3)}x_2,\\ \dots\dots\dots\dots\\ x_k=a_{02}^{(k)}+a_{12}^{(k)}x_2,\end{cases} \tag{16}$$

$$\dots\dots\dots\dots\dots\dots$$

$$\begin{cases}x_1=a_{0k}^{(1)}+a_{1k}^{(1)}x_k,\\ x_2=a_{0k}^{(2)}+a_{1k}^{(2)}x_k,\\ \dots\dots\dots\dots\\ x_{k-1}=a_{0k}^{(k-1)}+a_{1k}^{(k-1)}x_k.\end{cases} \tag{17}$$

If we substitute these values of predictors into equation (14) in each separate case, then, after simple calculations, we obtain *k* equations of linear regression relative to the predictors $x_1,\ldots,x_k$:

$$Y_1 = a_0 + \sum_{\substack{j=1\\ j\neq 1}}^{k} a_{01}^{(j)} + \left(a_1 + \sum_{\substack{j=1\\ j\neq 1}}^{k} a_{11}^{(j)} \right) x_1 + \varepsilon_1,$$

$$Y_2 = a_0 + \sum_{\substack{j=1\\ j\neq 2}}^{k} a_{02}^{(j)} + \left(a_2 + \sum_{\substack{j=1\\ j\neq 2}}^{k} a_{12}^{(j)} \right) x_2 + \varepsilon_2, \qquad (18)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$Y_k = a_0 + \sum_{\substack{j=1\\ j\neq k}}^{k} a_{0k}^{(j)} + \left(a_k + \sum_{\substack{j=1\\ j\neq k}}^{k} a_{1k}^{(j)} \right) x_k + \varepsilon_k .$$

Finally, we note that in an analogous manner we can construct linear regression models for the variable Y relative to any subset of predictors from the set of predictors $x_1,\ldots,x_k$.

Let us now consider the numerical examples for the case of two predictors $(k = 2)$.

Example 1. The marketing manager must investigate the effect of inflation (the variable x_1), and the other manager must investigate the effect of advertising expenses (the variable x_2) on the company's income (the variable Y). It is assumed that there are statistical data of observations for 10 months:

$$x_1 = 7,6,8,9,8,9,6,5,4,4;$$
$$Y_1 = 12,12,11,10,12,11,14,14,15,14;$$
$$x_2 = 22,20,21,18,20,24,23,25,24,22;$$
$$Y_2 = 14,11,12,10,12,15,14,16,16,14.$$

Construct models (6) and (10).

Solution. Computer-aided calculations yield the following results:

$$Y^{(1)} = 5.41 - 0.47\overline{x}_1 + 0.49\overline{x}_2 , \qquad (19)$$

$$Y^{(2)} = 3.82 - 0.44\hat{x}_1 + 0.55\hat{x}_2 . \qquad (20)$$

Example 2. Consider the equation (20). Using formulas (18), construct the models for Y_1 and Y_2.

Solution. Computer-aided calculations yield the following results:

$$Y_1 = 18.26 - 0.8\hat{x}_1 , \qquad (21)$$

$$Y_2 = 12.97 + 0.75\hat{x}_2 . \qquad (22)$$

ეკონომიკა

მრავლობითი წრფივი რეგრესიის შესახებ მარკეტინგულ კვლევაში

ნ. თოდუა[*], პ. ბაბილუა[**], თ. დოჭვირი[§]

[*] ი. ჯავახიშვილის სახ. თბილისის სახელმწიფო უნივერსიტეტი, გამოყენებითი მათემატიკისა და კომპიუტერულ მეცნიერებათა ფაკულტეტი, თბილისი

[**] ი. ჯავახიშვილის სახ. თბილისის სახელმწიფო უნივერსიტეტი, ზუსტ და საბუნებისმეტყველო მეცნიერებათა ფაკულტეტი, თბილისი

[§] ი. ჯავახიშვილის სახ. თბილისის სახელმწიფო უნივერსიტეტი, ეკონომიკისა და ბიზნესის ფაკულტეტი, თბილისი

(წარმოდგენილია აკადემიკოს ა. სილაგაძის მიერ)

მარკეტინგულ კვლევას მნიშვნელოვანი ადგილი უკავია თანამედროვე საბაზრო ეკონომიკაში. ძირითადი ეკონომიკური მახასიათებლების შესახებ მარკეტინგული ინფორმაცია გამოიყენება, ფირმების ბიზნესის დაგეგმვაში, მართვაში, კონტროლსა და პროგნოზირებაში. მაგალითად კომპანიების მიერ სოციალური ქსელების გამოყენება საკმაოდ გავრცელებულია თანამედროვე ბიზნესში. მსოფლიოში მიმდინარე ეკონომიკური კრიზისის ფონზე, სოციალური საიტები ყველაზე რეიტინგული და პოპულარული აღმოჩნდა. სოციალური ქსელები ხელს უწყობენ ბრენდის ცნობადობის შექმნას და მის პოპულარიზაციას, კომპანიის იმიჯის ფორმირებას, ახალი მომხმარებლების შეძენას, მათ წახალისებას და უმცირესი დანახარჯებით (ზოგჯერ ფინანსური დანახარჯების გარეშეც) წარმატებული შედეგების მიღწევას. ასეთ ტენდენციას აქტიურად აყვნენ ქართული კომპანიები, რომლებმაც სოციალურ ქსელში საკუთარი პროფილები შექმნეს. ეს იწვევს მომხმარებელთა აქტიურ ჩართვას. ქართული კომპანიების გააქტიურებას მომხმარებელიც შესაბამისად აჰყვა. ისინი მონაწილეობენ კომპანიების მიერ შემოთავაზებულ გათამაშებებში, გამოთქვამენ საკუთარ აზრს სხვადასხვა პროდუქტისა და მომსახურების შესახებ, კომუნიკაცია გახდა უფრო ინტერაქტიული. ამ საკითხების შესწავლაში არსებითი როლი კორელაციურ და რეგრესიულ ანალიზს ეკუთვნის. ნაშრომში შემოთავაზებულია მარკეტინგულ კვლევაში მრავლობითი წრფივი რეგრესიული მოდელის გამოყენების ახალი მეთოდი. კერძოდ, ერთპრედიქტორიანი წრფივი რეგრესიული მოდელებით აგებულია მრავლობითი წრფივი რეგრესიული მოდელი და პირიქით, მრავლობითი წრფივი მოდელით აგებულია ერთპრედიქტორიანი წრფივი მოდელები. მოტანილია აგრეთვე საილუსტრაციო რიცხვითი მაგალითები.

REFERENCES

1. *N. Todua, B. Mgebrishvili* (2009), Principles of Marketing. Tbilisi (in Georgian).
2. *P. Kotler, G. Armstrong* (2009), Principles of Marketing. New Jersey.
3. *P. Doyle, P. Stern* (2006), Marketing Management and Strategy. Prentice Hall New York.
4. *P. Newbold, W.L. Carlson, B. Thorne* (2007), Statistics for Business and Economics. New Jersey.
5. *M. Eddowes, S. Robin* (1997), Metody priniatiia reshenii. M. (in Russian).
6. *C. Dougherty* (2001), Vvedenie v ekonometriku. M. (in Russian).
7. *N. K. Malhotra* (2008), Marketing Research: An Applied Orientation. New Jersey.

Received May, 2013