Kuzhda T. Retail sales forecasting with application the multiple regression [Електронний ресурс] / T. Kuzhda // Соціально-економічні проблеми і держава. — 2012. — Вип. 1 (6). — С. 91-101. — Режим доступу до журн. : http://sepd.tntu.edu.ua/images/stories/pdf/2012/12ktibrm.pdf.

УДК 06.51.41
JEL Classification: C53

Тетяна Кужда

Тернопільський національний технічний університет імені Івана Пулюя

ПРОГНОЗУВАННЯ ОБСЯГІВ ПРОДАЖУ ПРОДУКЦІЇ НА ОСНОВІ БАГАТОФАКТОРНОЇ РЕГРЕСІЙНОЇ МОДЕЛІ

Анотація. *В статті описано метод багатофакторного регресійного моделювання, теоретичний підхід до побудови регресійних моделей, порядок розрахунку кількісного прогнозу залежної змінної під впливом декількох незалежних змінних. Застосовано теоретичний матеріал до прогнозування обсягів продажу продукції під впливом очікуваного доходу споживачів та витрат на рекламну діяльність. Здійснено перевірку отриманої багатофакторної регресійної моделі на статистичну надійність та значущість та розраховано прогноз обсягів продажу продукції на наступний період.*

Ключові слова: *регресійний аналіз, залежна та незалежна змінні, багатофакторна регресійна модель, статистична надійність та значущість, екстраполяція трендів, прогноз обсягів продажу продукції.*

Татьяна Кужда

ПРОГНОЗИРОВАНИЕ ОБЪЕМОВ ПРОДАЖ ПРОДУКЦИИ НА ОСНОВАНИИ МНОГОФАКТОРНОЙ РЕГРЕССИОННОЙ МОДЕЛИ

Аннотация. *В статье описано метод многофакторного регрессионного моделирования, теоретический подход к построению регрессионных моделей, порядок расчета количественного прогноза зависимой переменной под влиянием нескольких независимых переменных. Использовано теоретический материал к прогнозированию объемов продаж продукции под влиянием ожидаемого дохода потребителей и затрат на рекламную деятельность. Осуществлена проверка полученной многофакторной регрессионной модели на статистическую надежность и значимость, рассчитан прогноз объемов продаж продукции на следующий период.*

Ключевые слова: *регрессионный анализ, зависимая и независимая переменные, многофакторная регрессионная модель, статистическая надежность и значимость, экстраполяция трендов, прогноз объемов продаж продукции.*

Tetyana Kuzhda

RETAIL SALES FORECASTING WITH APPLICATION THE MULTIPLE REGRESSION

Abstract. *The article begins with a formulation for predictive learning called multiple regression model. Theoretical approach on construction of the regression models is described. The*

key information of the article is the mathematical formulation for the forecast linear equation that estimates the multiple regression model. Calculation the quantitative value of dependent variable forecast under influence of independent variables is explained. This paper presents the retail sales forecasting with multiple model estimation. One of the most important decisions a retailer can make with information obtained by the multiple regression. Recently, a changing retail environment is causing by an expected consumer's income and advertising costs. Checking model on the goodness of fit and statistical significance are explored in the article. Finally, the quantitative value of retail sales forecast based on multiple regression model is calculated.

Keywords: *regression analysis, dependent and independent variables, multiple regression model, goodness of fit and the statistical significance, trend extrapolation, retail sales forecast.*

Introduction. Regression analysis includes many techniques for modeling and analyzing several variables, when the focus is on the relationship between a dependent variable and one or more independent variables. Regression analysis is also used to understand which among the independent variables are related to the dependent variable, and to explore the forms of these relationships. Regression analysis can be used to infer causal relationships between the independent and dependent variables. Variables which are used to explain other variables are called explanatory (or independent) variables. A dependent variable is what you measure in the forecast. The dependent variable responds to the independent variable. It is called dependent because it "depends" on the independent variable [1].

Regression modeling is the process of construction forecasting models based on the relationship between a dependent variable and independent variables to make the future forecast. Regression modeling is a kind of multifactor forecasting. The basis of regression modeling is the construction of regression models.

Regression models are used to predict one variable from one or more other variables. Regression models provide the scientist with a powerful tool, allowing predictions about future events to be made with information about past or present events. In order to construct a regression model, both the information which is going to be used to make the prediction and the information which is to be predicted must be obtained from a sample of objects or individuals. The relationship between the two pieces of information is then modeled with a linear transformation. Then in the future, only the first information is necessary, and the regression model is used to transform this information into the predicted. In other words, it is necessary to have information on both variables before the model can be constructed [1, 6]. Regression models are one of the most famous examples of economic and statistical models used in the forecasting of socio-economic processes. Construction of the regression models includes the following stages:

1) Selection of an object to forecast. The objects of socio-economic forecasting are the economic processes (for example, inflation, demand, supply, exchange rate, etc.), any indicator describing the company activity (for example, production, price, profit, income, sales, costs, etc.), any indicator describing the national economics (for example, gross domestic product, gross investment, national income, government spending, export, import, external debt, etc.), any indicator describing the social processes (for example, wage, bonus fund, incentive fund, overtime payments, employment and unemployment, emigration and immigration, etc.). An object of forecasting is a dependent variable.

2) Selection of the factors (independent variables) that explains the changes in the socio-economic processes. The factors should be in the causal link to the object of forecasting and all factors must be quantitatively measured and significant. For example, company's retail sales depend on expected consumer's income and advertising costs. In this example, company's retail sales are the object of forecasting (or dependent variable); the expected consumer's income and advertising costs are factors or independent variables.

3) Data collection is the process of obtaining useful information on key quantitative characteristics of socio-economic processes. Statistical information necessary to forecasting can be obtained from primary and secondary data sources. Data processing is any process that summarizes,

analyzes or otherwise converts data into usable information. Information base of forecasting based on regression models is the several interrelated time series with a feedback relationship.

4) Selection of the mathematical dependence between the factors or independent variables and dependent variable. Regression models can be described by the following types of dependencies: linear, power, logarithmic, etc. In linear regression, data are modeled using linear functions, and unknown model parameters are estimated from the data.

Linear regression is an approach to modeling the relationship between two or more independent variables (*X*) and a single dependent variable (*Y*). The case of one explanatory variable is called simple regression model. More than one explanatory variable is multiple regression models.

On practice is widely used the more general multiple regression model. General multiple regression model can have multiple explanatory variables. Multiple regression model is a flexible method of data analysis that may be appropriate whenever a quantitative variable (the dependent variable) is to be examined in relationship to any other factors (expressed as independent variables). For example, a multiple regression model might examine average salaries (dependent variable) as a function of age, education, gender and experience (independent variables).

Multiple regression requires a large number of observations. The number of periods must substantially exceed the number of independent variables you are using in regression. The absolute minimum is that you have five periods [1, 6].

The forecast linear equation that estimates the multiple regression model look like (1):

$$Y = b_0 + b_1 \cdot X_{1t} + b_2 \cdot X_{2t} + ... + b_m \cdot X_{mt}; \quad (1)$$

where Y – is called the exogenous variable, response variable, measured variable, or dependent variable. The decision as to which variable in a data set is modeled as the dependent variable and which are modeled as the independent variables may be based on a presumption that the value of one of the variables is caused by, or directly influenced by the other variables;

X_{1t}, X_{2t},... X_{mt} – are called endogenous variables, explanatory variables, input variables, predictor variables, or independent variables at period *t*;

b_0, b_1, b_2... b_m – are the regression coefficients; b_0 – measures the changes in *Y* with respect to random factors that are not included in the regression model; b_1 – measures the changes in *Y* with respect to X_{1t}; b_2 – measures the changes in *Y* with respect to X_{2t}; … b_m – measures the changes in *Y* with respect to X_{mt}.

To find the regression coefficients (b_0, b_1, b_2, b_m) need to calculate the system of normal equations. The calculation formulas are complex. For multiple regression, it is almost imperative to use computer software (Data Analysis) to the prediction equation. Corresponding to the multiple regression equation, software finds a forecast equation by estimating the model parameters using sample data.

5) Checking the model on the goodness of fit and the statistical significance based on statistical coefficients. If a model is reliable and statistical significant, the forecast will be accurate.

6) Calculation of the independent variables forecasts is the process of prediction the independent variables under influence of a time factor. To find the quantitative values of independent variables forecasts we can use the forecasting based on trend extrapolation.

7) Calculation of the forecast based on regression modeling is the process of prediction the quantitative value of dependent variable under influence of independent variables.

Multiple Model Estimation in Practice. Application above theoretical information for forecasting based on multiple regression is described in example below. Statistical data on retail sales, expected consumer's income and advertising costs within 10 months are given on table 1. We want to explain how to calculate the retail sales forecast for January based on multiple regression model.

Table 1

Statistics on retail sales, expected consumer's income and advertising costs

Months	Retail sales, thousand dollars	Expected consumer's income, thousand dollars	Advertising costs, thousand dollars
Mart	125	20,0	12,5
April	126	20,2	12,7
May	128	20,5	12,8
June	130	20,7	13,0
July	131	20,9	13,2
August	133	21,2	13,5
September	139	21,5	13,7
October	142	22,1	13,8
November	145	22,7	14,0
December	150	23,5	14,4

In this example, company's retail sales are dependent variable *Y*; the expected consumer's income and advertising costs are factors or independent variables.

To find the retail sales forecast based on regression modeling we need to use the multiple regression model (2):

$$Y = b_0 + b_1 \cdot X_{1t} + b_2 \cdot X_{2t}, \quad (2)$$

where $Y -$ is the forecast of company's retail sales, thousand dollars;

X_{1t} – is the expected consumer's income at period *t*;

X_{2t} – is the advertising costs at period *t*;

b_0, b_1, b_2 – are the regression coefficients.

The calculation of coefficients b_0, b_1, b_2 is long and laborious process. Microsoft Excel provides a lot of possibilities to forecasting based on regression modeling. Statistical data on retails sales, expected consumer's income and advertising costs within 10 months should be presented on Excel spreadsheet. Firstly, select the "Data" menu / "Data Analysis" / "Regression" (Figure 1).

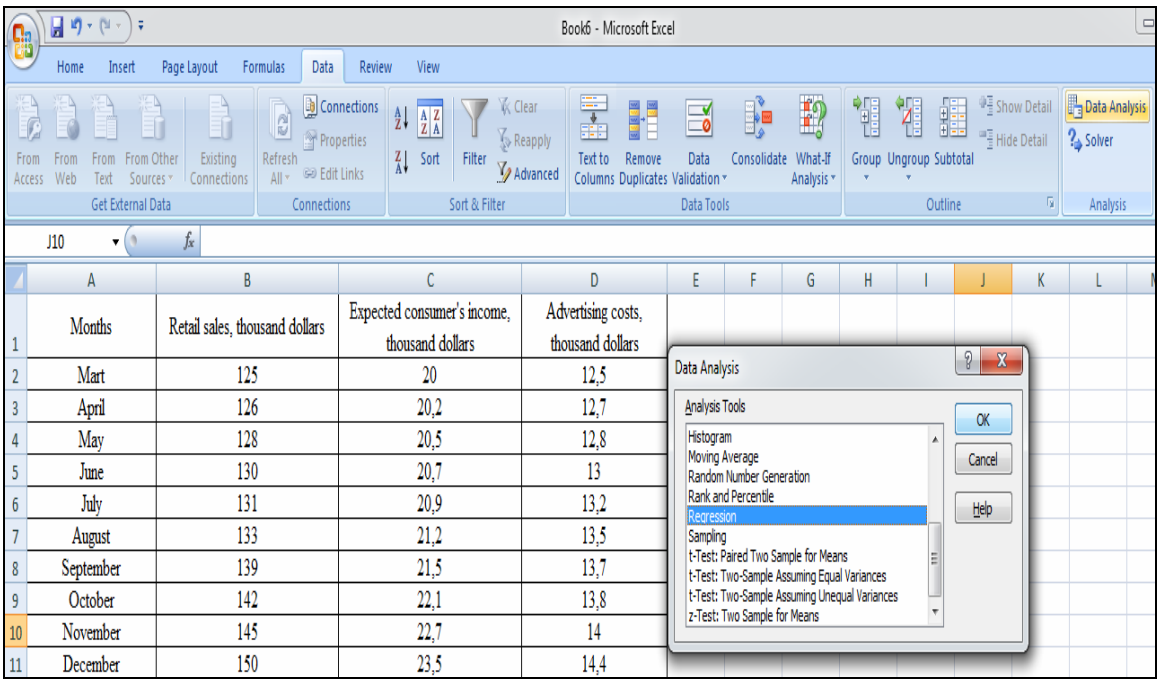


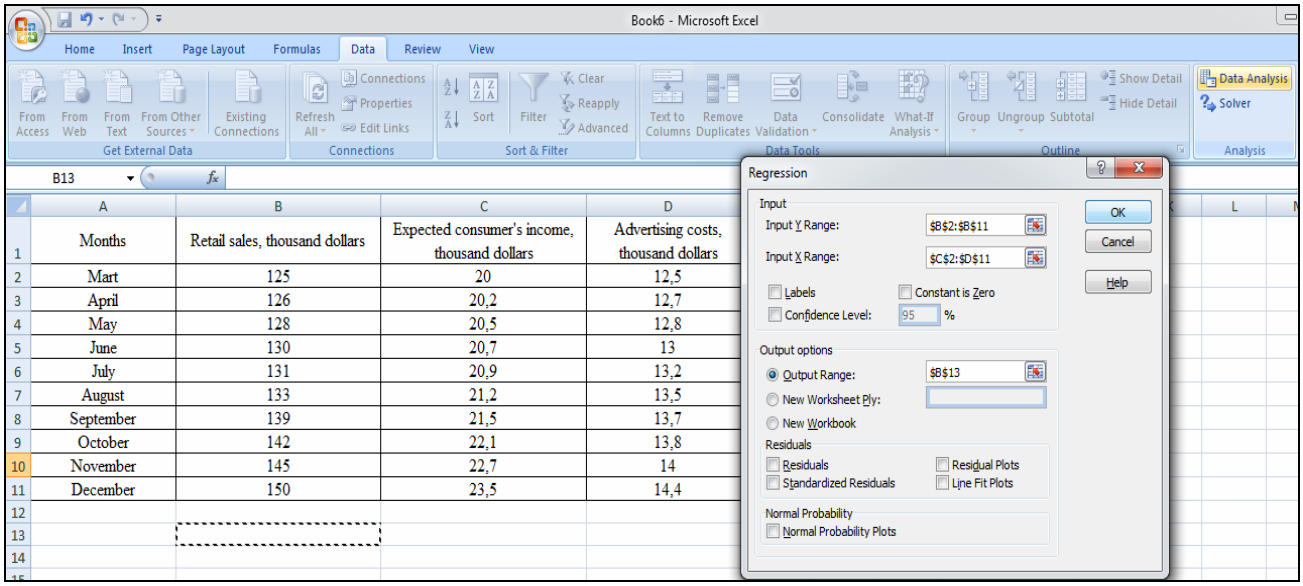
Fig. 1. "Data" menu / "Data Analysis" / "Regression"

The following window appears (Figure 2). The first box is the “Input Y Range”. Here, we tell Excel about our dependent variable (retail sales). The dependent variable must be a column. To fill “Input Y Range” need click here and enter the cell reference for the range of data on retail sales.

The next stage is to input independent variables. The independent variables must be a block of data, if the independent variables are several, or column of data, if the independent variable is one. In the dataset we are using we have two independent variables: the expected consumer’s income and advertising costs. To fill “Input X Range” need click here and enter the cell reference for the block of data on expected consumer’s income and advertising costs.

If the Confidence Level equals to 95%, you can say that you are 95% sure that the retail sales forecast will be accurate.

Next we tell Excel where we want the results to be written. To fill “Output range” enter the reference for the cell (B13) of the output table. So, finally, we click OK.



	A	B	C	D
1	Months	Retail sales, thousand dollars	Expected consumer's income, thousand dollars	Advertising costs, thousand dollars
2	Mart	125	20	12,5
3	April	126	20,2	12,7
4	May	128	20,5	12,8
5	June	130	20,7	13
6	July	131	20,9	13,2
7	August	133	21,2	13,5
8	September	139	21,5	13,7
9	October	142	22,1	13,8
10	November	145	22,7	14
11	December	150	23,5	14,4

Fig. 2. Regression window

And we get a lot of output. The regression output has three components: Regression statistics table, ANOVA table, Regression coefficients table (Figure 3). Figure 3 contains the information need to get the multiple regression model.

Quantitative values of the coefficients: b_0 is opposite “Intercept” ($b_0 = -33{,}926$); b_1 is opposite “X Variable 1” ($b_1 = 5{,}204$); b_2 is opposite “X Variable 2” ($b_2 = 4{,}328$).

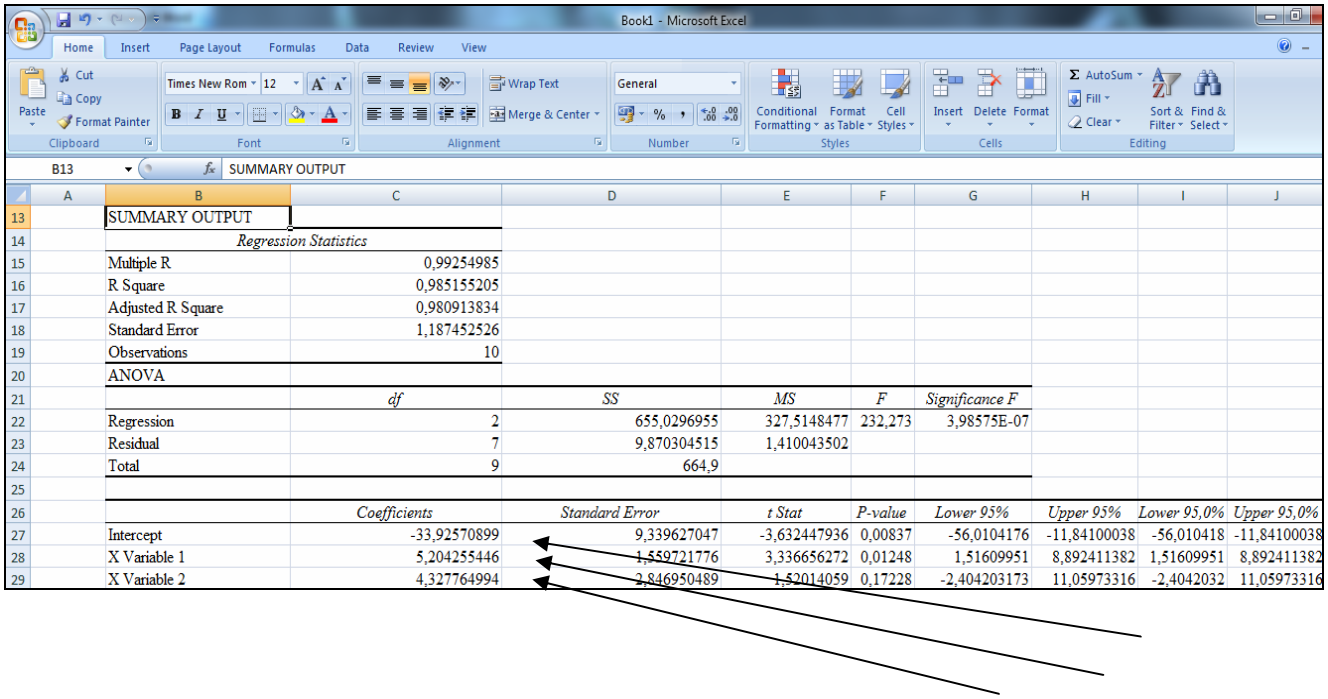
The multiple regression model need to forecast the retail sales (Y) for January is:

$$Y = -33{,}926 + 5{,}204 \cdot X_{1t} + 4{,}328 \cdot X_{2t}. \quad (3)$$

We have the multiple regression model (3) need to forecast the retail sales, but quantitative value of the forecast using Data Analysis we can not get.

The next stage is checking the multiple regression model (3) on the goodness of fit and the statistical significance. And after checking the model, we can calculate quantitative value of the retail sales forecast.

Statistical goodness of fit for the multiple regression model can be determined by the following statistical coefficients: the correlation coefficient (r), the coefficient of determination (R^2) and adjusted coefficient of determination (AR^2).



Book1 - Microsoft Excel

B13 SUMMARY OUTPUT

	A	B	C	D	E	F	G	H	I	J
13		SUMMARY OUTPUT								
14		*Regression Statistics*								
15		Multiple R	0,99254985							
16		R Square	0,985155205							
17		Adjusted R Square	0,980913834							
18		Standard Error	1,187452526							
19		Observations	10							
20		ANOVA								
21			*df*	*SS*	*MS*	*F*	*Significance F*			
22		Regression	2	655,0296955	327,5148477	232,273	3,98575E-07			
23		Residual	7	9,870304515	1,410043502					
24		Total	9	664,9						
25										
26			*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
27		Intercept	-33,92570899	9,339627047	-3,632447936	0,00837	-56,0104176	-11,84100038	-56,010418	-11,84100038
28		X Variable 1	5,204255446	1,559721776	3,336656272	0,01248	1,51609951	8,892411382	1,51609951	8,892411382
29		X Variable 2	4,327764994	2,846950489	1,52014059	0,17228	-2,404203173	11,05973316	-2,4042032	11,05973316

Fig. 3. The regression output: Regression statistics table, ANOVA table, Regression coefficients table

Coefficient of determination (R^2) – is a measure to assess how well the multiple regression model explains and predicts future outcomes. It is expressed as a value between 0 and 1. A value of one indicates a perfect fit, and therefore, a very reliable multiple regression model for future forecasts. A value of zero, on the other hand, would indicate that the multiple regression model fails to accurately forecast the dataset [3, 5].

The following points are accepted guidelines for interpreting the coefficient of determination: values between 0 and 0,3 indicate a weak positive linear relationship; values between 0,3 and 0,7 indicate a moderate positive linear relationship; values between 0,7 and 1 indicate a strong positive linear relationship.

The correlation coefficient (r), is a measure of the strength of the relationship between two or more independent variables (X) and a single dependent variable (Y).

One of ways to find this coefficient is the following: correlation coefficient (r) is the square root of the coefficient of determination (4):

$$r = \sqrt{R^2} \tag{4}$$

The correlation coefficient takes on values ranging between +1 and -1. The following points are accepted guidelines for interpreting the correlation coefficient: 0 indicates no linear relationship; +1 indicates a perfect positive linear relationship; -1 indicates a perfect negative linear relationship; values between 0 and 0,3 (0 and -0,3) indicate a weak positive (negative) linear relationship; values between 0,3 and 0,7 (-0,3 and -0,7) indicate a moderate positive (negative) linear relationship; values between 0,7 and 1 (-0,7 and -1) indicate a strong positive (negative) linear relationship [4, 5].

In a multiple linear regression model, adjusted coefficient of determination (AR^2) measures the share of the variation in the dependent variable accounted by the explanatory variables. Adjusted coefficient of determination is generally considered to be a more accurate goodness-of-fit measure than the coefficient of determination. The adjusted R^2 will always be less than or equal to the coefficient of determination (R^2). Adjusted coefficient of determination is particularly useful in the feature selection stage of model building [4, 5].

Adjusted coefficient of determination (R-Square) is computed using the following formula (5):

$$Adjusted \;\; R^2 = 1 - \frac{(1 - R^2) \cdot (n - 1)}{(n - k - 1)}, \qquad (5)$$

where R^2 – is the coefficient of determination;

n – is the number of observations (or periods);

k – is the number of independent variables.

To find the correlation coefficient and coefficient of determination we need to interpret Regression statistics table (Figure 3).

Table 2

Regression statistics

		Explanation
Multiple R	0,99254985	Correlation coefficient
R Square	0,985155205	Coefficient of determination
Adjusted R Square	0,980913834	Adjusted coefficient of determination
Standard Error	1,187452526	Standard Error is a measure of error in prediction
Observation	10	Number of observations used in the regression

Correlation coefficient can be calculated by the formula (4):

$$r = \sqrt{0{,}985155205} \approx 0{,}99$$

Correlation coefficient r=0,99 may be interpreted as follows: approximately 99% (0,99*100%) of the variation in the dependent variable (retail sales) can be explained by the multiple regression model (3).

If the coefficient of determination is greater than 0,7, as it is in this case, there is a good fit to the data. The coefficient of determination 0,985 means approximately 98,5% (0,985*100%) of the variation in the dependent variable (retail sales) can be explained by the independent variables (the expected consumer's income and advertising costs).

Adjusted coefficient of determination by the following formula (5):

$$Adjusted \;\; R^2 = 1 - \frac{(1 - 0{,}985155205) \cdot (10 - 1)}{(10 - 2 - 1)} \approx 0{,}981$$

Adjusted coefficient of determination 0,981 means approximately 98,1% (0,981*100%) of the variation in the dependent variable (retail sales) can be explained by the independent variables (the expected consumer's income and advertising costs).

Checking the model on the statistical significance based on ANOVA table (Figure 3), where (SS – is the sum of squares, the numerator of the variance; DF – is the denominator; MS – is the mean square of variance; Significance F means the statistical significance of the multiple regression model). ANOVA means an analysis of variance that consists of calculations that provide information about levels of variability within a regression model and form a basis for tests of significance.

Significance F means the statistical significance of the multiple regression model. In this example (Figure 3), the value of "Significance F" is lower than 0,05, then we can say the multiple regression model is generally acceptable and statistical significant to forecast of the retail sales ($3{,}98*10^{-7} < 0{,}05$).

Checking of each coefficient on the statistical significance based on Regression coefficients table (Figure 3), where column "Coefficient" gives the quantitative values of regression coefficients

b_0, b_1, b_2; column "Standard error" gives the standard errors (i.e. the estimated standard deviation) of regression coefficients; column "t Stat" gives the computed t-statistic (is a ratio of the departure of an estimated parameter from its notional value and its standard error); column "P-value" gives the probability value for each regression coefficient. If "P-value" is less than 0,05 (5% mistake probability), then the coefficient is statistical significant (95 % probability means the forecast based on multiple regression model is accurate), and if "P-value" is more than 0,05; the coefficient is statistical insignificant.

In this example, "P-value" for coefficient b_0 is 0,008 (lower than 0,05), "P-value" for coefficient b_1 is 0,01 (lower than 0,05), "P-value" for coefficient b_2 is 0,17 (higher than 0,05), then we can say the multiple regression model in generally is statistical significant.

Thus, the multiple regression model (3) is statistical significant, the model is useful and reliable to forecast. To find the forecast of the retail sales for January, at first, we need to calculate the quantitative values of expected consumer's income forecast and advertising costs forecast for January.

Calculation of the expected consumer's income forecast and advertising costs forecast for January is possible using the forecasting based on trend extrapolation. To do this we need to find the forecast of expected consumer's income depending on time (*t*) and the forecast of advertising costs depending on time (*t*). Firstly, we need to calculate the expected consumer's income forecast based on trend extrapolation (using a linear equation). Linear equation looks like (6):

$$x_t = a + b \cdot t, \tag{6}$$

where x_t – is the expected consumer's income forecast based on trend extrapolation (or advertising costs forecast based on trend extrapolation);

a and *b* –are the designate coefficients;

t – is the time unit.

Coefficient *b* can be calculated by the formula (7):

$$b = \frac{\sum t \cdot x - n \cdot \bar{t} \cdot \bar{x}}{\sum t^2 - n \cdot \bar{t}^2}, \tag{7}$$

where *n* – number of periods;

$\bar{t}$ – is the average value of variable *t* (time or independent variable);

$\bar{x}$ – is the average value of dependent variable *x* (average value of expected consumer's income or average value of advertising costs).

Average value of variable "*t*" can be calculated by the formula (8):

$$\bar{t} = \frac{\sum t}{n}, \tag{8}$$

where *n* – is the number of periods;

$\sum t$ - is the sum of numbers from 1 to *n*;

Average value of variable "*x*" can be calculated by the formula (9):

$$\bar{x} = \frac{\sum x}{n}; \tag{9}$$

where *n* – is the number of periods;

$\sum x$ - is the sum of statistical data for *n* periods.

Coefficient *a* can be calculated by the formula (10):

$$a = \bar{x} - b \cdot \bar{t}. \tag{10}$$

To write down a linear equation $x_{1t} = a + b \cdot t$ (where x_{1t} – is the expected consumer's income forecast) and calculate the coefficients "*b*" and "*a*" need to find: "t^2", "$x_1 * t$" on table 3.

Table 3

Results of calculations

Months	Expected consumer's income (x_1), thousand dollars	*t*	t^2	$x_1 * t$
Mart	20,0	1	1	20
April	20,2	2	4	40,4
May	20,5	3	9	61,5
June	20,7	4	16	82,8
July	20,9	5	25	104,5
August	21,2	6	36	127,2
September	21,5	7	49	150,5
October	22,1	8	64	176,8
November	22,7	9	81	204,3
December	23,5	10	100	235
∑	**213,3**	**55**	**385**	**1203**

Average value of time (*t*) by the formula (8):

$$\bar{t} = \frac{\sum t}{n} = \frac{55}{10} = 5,5 .$$

Average expected consumer's income (x_1) by the formula (9):

$$\bar{x}_1 = \frac{213,3}{10} = 21,33$$

Coefficient b by the formula (7):

$$b = \frac{1203 - 10 \cdot 5,5 \cdot 21,33}{385 - 10 \cdot (5,5)^2} \approx 0,361$$

Coefficient a by the formula (10):

$$a = 21,33 - 0,361 \cdot 5,5 = 19,34$$

Linear equation looks like:

$$x_{1t} = a + b \cdot t = 19,34 + 0,361 \cdot t$$

Forecast of expected consumer's income for January based on trend extrapolation:

$$x_{1t} = 19,34 + 0,361 \cdot 11 \approx 23,311 \ \textit{thousand dollars}.$$

To write down a linear equation $x_{2t} = a + b \cdot t$ (where x_{2t} – is advertising costs forecast) and calculate the coefficients "*b*" and "*a*" need to find: "t^2", "$x_2 * t$" on table 4.

Table 4

Results of calculations

Months	Advertising costs (x_2), thousand dollars	t	t^2	$x_2 * t$
Mart	12,5	1	1	12,5
April	12,7	2	4	25,4
May	12,8	3	9	38,4
June	13,0	4	16	52
July	13,2	5	25	66
August	13,5	6	36	81
September	13,7	7	49	95,9
October	13,8	8	64	110,4
November	14,0	9	81	126
December	14,4	10	100	144
∑	**133,6**	**55**	**385**	**751,6**

Average value of time (t) by the formula (8):

$$\bar{t} = \frac{\sum t}{n} = \frac{55}{10} = 5,5$$

Average advertising costs (x_2) by the formula (9):

$$\bar{x}_2 = \frac{133,6}{10} = 13,36$$

Coefficient b by the formula (7):

$$b = \frac{751,6 - 10 \cdot 5,5 \cdot 13,36}{385 - 10 \cdot (5,5)^2} \approx 0,203$$

Coefficient a by the formula (10):

$$a = 13,36 - 0,203 \cdot 5,5 = 12,24$$

Linear equation looks like:

$$x_{2t} = a + b \cdot t = 12,24 + 0,203 \cdot t$$

Forecast of advertising costs for January based on trend extrapolation:

$$x_{2t} = 12,24 + 0,203 \cdot 11 = 14,473 \text{ thousand dollars.}$$

Retail sales forecast for January based on multiple regression model (formula 3):

$$Y = -33,926 + 5,204 \cdot X_{1t} + 4,328 \cdot X_{2t} =$$
$$= -33,926 + 5,204 \cdot 23,311 + 4,328 \cdot 14,473 \approx 150,024 \text{ } \textit{thousand dollars}.$$

Thus, the retail sales forecast for January based on multiple regression model equals to 150,024 thousand dollars.

Conclusion. The multiple regression model was effective for forecasting retail sales under influence of expected consumer's income and advertising costs. It can be applied for forecasting other business data. Using such models for forecasting retail sales can assist company managers in planning and making decisions more effectively.

References:

1. Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*, 3rd Ed. Mahwah, NJ: Lawrence Erlbaum Associates.
2. Rogers, David S. "A Review of Sales Forecasting Models," *International Journal of Retail and Distribution Management*, MCB University Press, Vol. 20, Issue 4, 1992.
3. Mining-Long Lee & R. Kelley Pace "Spatial Distribution of Retail Sales", *The Journal of Retail Estate Finance and Economics*, Springer, Vol. 31(1), pages 53-69, August, 2005.
4. Lundholm, Russell J. and McVay, Sarah E., "*Forecasting Sales: A model and some evidence from the retail industry*" (January, 2004).
5. Wassana Suwanvijit, Chamnein Choonpradub, Nittaya McNeil "Statistical Model For Short-Term Forecasting Sparkling Beverage Sales In Southern Thailand", *International Business & Economics Research Journal*, Vol.8, №9, September 2009.
6. Samawi, H.M., Ababneh, F.M., On regression analysis using ranked set sample, *Journal of Statistical Research*. 35 (2001), 93-105.

Рецензія: д.е.н., проф. Кирич Н. Б.

Received: March, 2012

1st Revision: April, 2012

Accepted: May, 2012