

# INCREASING PERFORMANCE THROUGH COST MANAGEMENT

Author(s)\*: Carmen Elena STOENOIU  
Position: Assoc. Prof., PhD  
University: Technical University of Cluj-Napoca  
Address: Cluj-Napoca, Memorandumului Str., No. 28, Romania  
Email: Carmen.Stoenoiu@emd.utcluj.ro  
Webpage: <http://carmen.academicdirect.ro/>

## Abstract

**Purpose** – In the present study, I proposed to conduct an analysis of the cost behavior and to establish their equation, useful for calculating the sale price and subsequently the profit.

**Methodology/approach** - Cost delimitation: material and wages was established, then the dependent variable and the independent variables were established and a regression analysis was performed.

**Findings** – By regression analysis we confirmed that there is a direct link between the dependent variable and the independent variables and we obtained the cost equation.

**Research limitations/implications** – When data distribution is not normal, regression analysis cannot be used. When data is normally distributed, the analysis is useful and can be extended by delimiting a model able to estimate a cost equation in an enterprise.

**Practical implications** – The analysis can be useful and can be used by companies to estimate costs when we know the company's past data, or when we have a normal distribution of it.

**Originality/value** – This study shows value because it involves studying cost behavior. Because in the economic activity the costs are those that ultimately determine the selling price and afterwards the profit, I consider it important to study them through statistical techniques.

**Key words:** Cost of production; Regression analysis; Cost Management

## Introduction

With the globalization of the economy, firms face growing competition on the market. Each company is trying to have a competitive advantage that allows it to produce more, more specialized or more diversified with the ultimate goal of making the most profit. Thus, the managers in their activity assign a significant share of their time to analyze and find solutions that best correspond to the profit - opportunity - cost - risk relationship (Lucas, 2003). At the base of these efforts, many times, there are often economic and financial problems which often do not answer even when there are strategies, objectives and well-defined operational plans. These problems occur most often due to insufficient information, faulty or delayed processing, distortions or processing errors (Cooper and Kaplan 1988; Shim and Sudit 1995). The source from which economic information starts is accrual accounting that uses a language that is sometimes difficult to reach for managers (Ahrens and Chapman, 2004).

With the passage of time, the effect of change as a result of globalization has also been felt in accounting that has moved from the classical or neutral form of reporting past events to a modern form that monitors, issues, and analyzes, provides additional explanations of developments or tendencies that arise at cost, income, profit, starting from general to particular or vice versa (Covaleski & all, 2003; Ahrens, & Chapman, (2004).) Estimation using specific data processing techniques an important role in the company's activity (planning and control) because for most products (new or under construction) changes in the cost components occur, changes that affect the company's profitability and its long-term survival on the market.

Linear simile regression analysis is a technique used to establish the general link between a dependent variable and independent variables (Cochran, 1976; Zellner, 1970). The advantage of developing a regression technique is that it helps to develop a model that can be used later by

managers to make decisions as a useful solution to a problem in the future (Stockton & Middle, 1986; Stout, 2015). In econometrics, there are many hypotheses and limitations related to the variables used in studies, sometimes due to simultaneous complications that produce major effects, sometimes due to intercorrelations between these variables that are difficult to control (Jensen, 1967; Cochran, 1976).

The purpose of this study is to identify the cost of production equation and to analyze the variation of these costs based on historical data taken from accounting, useful in forecasting and budgeting. The analysis assumed the link between the variables considered dependent and independent by statistical indicators: probability, variance and correlation. Using multivariate regression analysis the predictive and observed values, residual values and distances of Mahalanobis and Cook were studied.

### Material and method

In this study an analysis was made of the behavior of the production costs identified in wage and material costs versus the quantity obtained. Subsequently a statistical analysis was performed using regression analysis by delimiting the existing variables in dependent and independent variables to see if there is a link between the variables taken in the study, the shape and intensity of the binding and the establishment of the regression equation.

Table 1. Annual baseline

Month	Quantity produced	Total production costs	Wage costs	Material costs
January	2.270,00	320.000,00	153200	166800
February	2.400,00	335.000,00	161200	173800
March	2.520,00	350.000,00	164300	185700
April	2.800,00	385.000,00	154400	230600
May	2.850,00	390.000,00	165300	224700
June	2.920,00	400.000,00	181200	218800
July	2.650,00	360.000,00	165300	194700
August	2.400,00	325.000,00	135400	189600
September	3.200,00	430.000,00	171200	258800
October	3.350,00	450.000,00	175200	274800
November	3.110,00	420.000,00	169400	250600
December	2.720,00	375.000,00	162300	212700
Total	33.190	4.540.000	1.958.400	2.581.600

The regression analysis was performed using the Statistical Program (v. Soft, USA) using the data presented in Table 1.

Regression analysis was used because it is a statistical technique that uses the data to estimate the intercept and slope of a cost equation (production cost demarcation line) and to establish the link between the variables (dependence and independence). For this, the dependent variable (y) was considered to be Material costs and the independent variable (x) the Quantity, Wage salaries and Total cost production (produced for the twelve months) taken into the study.

The regression model can be represented in terms of a matrix [Weisberg, 1985], according to Eq. (1):

$$y = x\alpha + e \quad (1)$$

Where: y - the material cost (dependent variable), where  $y = [y_1, \dots, y_n]^T$ ; x - the quantity produced for the twelve months taken into study, Wage salaries and Total cost production (independent variable); e - the variable interpreted as an error (perturbation or measurement error), where  $e = [e_1, \dots, e_n]^T$ ;  $\alpha$  - the vector of the coefficients px1 parameters of the model, where  $\alpha = [\alpha_1, \dots, \alpha_n]^T$ .

If the model is appropriate then the observed value  $y_i$  can be determined by the value  $x_i$ . The model assumes that errors are normally and independently distributed by zero and  $s^2$ , Eq. (2).

$$s^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p} \quad (2)$$

Where:  $\hat{y}$  - the value adjusted by the regression equation;  $n$  - number of observations;  $p$  - number of estimates.

Using the observed responses and predictors, unknown parameters are estimated.

In the regression analysis the following statistical indicators were calculated: R-square, P-value, standard error, ANOVA variance analysis, and then the distance Mahalanobis and the distance Cook.

R-square was calculated because it represents a possibility of statistically measuring the way in which the regression line matches the data [Comiskey, 1966]. Specifically, it measures the percentage of the independent variable variance (Total production cost, Wage salaries and Quantity) that varies and is explained by the dependent variable (Material costs). When R-square takes values close to 0, we say that there is no linear relationship between the cost of production and quantity, and when the values are close to 1, we say that there is a perfect linear relationship between the cost of production and the quantity.

The  $p$  value is for the intercept and slope and measures the probability of observing high values, as well as the estimated B coefficient when the real value is zero.

In the regression analysis, the term "standard error" was used to compute the standard regression error, i.e. to show the usual least square estimate of the standard deviation of the underlying errors. It is considered that the data-driven line by regression is the best right line to the data. Depending on the regression coefficient, the relation type (direct correlation - positive value, and inverse correlation - negative value, when  $B = 0$ , we say that variables ( $y$  and  $x$ ) are independent).

Later, a correlation method has been used which has the advantage of providing a synthetic measure of the links between the statistical variables. The indicators that measure the intensity of the link are: covariance, correlation coefficient and correlation ratio. Covariation is calculated as the simple arithmetic average of the products of the deviations of the two correlated variables,  $x$  and  $y$ , from their arithmetic meanings  $\bar{x}$  and  $\bar{y}$ , according to Eq. (3):

$$\text{Cov}(x, y) = 1/n \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (3)$$

Covariation are null if variables are independent (lack of correlation bond). Its absolute value  $\text{cov}(x, y)$  has no upper limit. As the correlation intensifies and the covariance increases. The indicator has the advantage that it is fairly easy to calculate and the disadvantage that it depends on the units in which the random variables are measured. So it is not comparable from one variable to another; high values of the indicator show a strong link.

The correlation coefficient can take values between -1 and +1, ie it satisfies the inequalities:  $-1 \leq r_{yx} \leq 1$ , and its sign, as well as the regression coefficient, signifies the type of linkage ("-" reverse link, "+" "Direct link"). The more the correlation coefficient has values closer to 1 or -1, the correlation between the  $x$  and  $y$  variables is stronger, and when approaching zero, the intensity of the link between the two variables decreases. If  $r_x = 0$ , the variables are linear or uncorrelated variables, and for equation with the unit, the functional dependence between the two variables results.

In order to perform the regression analysis Eq. (4):

$$\text{Total production cost} = \text{Wage costs} + \text{Material costs} \quad (4)$$

The estimated linear model of the procedure is presented by Eq. (5):

$$Y = \alpha_0 X_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_{p-1} X_{p-1} + \epsilon, \quad (5)$$

This model expresses that the variable  $Y$  can be obtained as a linear combination of the variables  $X_0, X_1, \dots, X_{p-1}$  to which the error " $\epsilon$ " is added. The coefficients of regression summary (B) are:  $\alpha_1 \dots \alpha_{p-1}$ .

Mahalanobis distance is a statistical way of measuring the distance between a point P and a distribution D, being a multidimensional generalization for measuring the standard deviation distance P from the mean of D. This distance is zero if P is in the middle of D and increases as P moves away from the mean: the main axis of the component, measures the number of standard deviations from P to the average of D. The Mahalanobis distance is unique and invariant in scale and takes into account the correlations of the dataset.

Cook's distance is an estimate used in statistics in regression analysis. From a practical point of view, Cook's distance was used to indicate influential data points worth considering for validation, or to indicate areas where the initial data that led to those results should be studied more carefully.

### Results and discussions

After analyzing the evolution of the quantity and the production costs over the 12 months, Figures 1-3 were obtained.

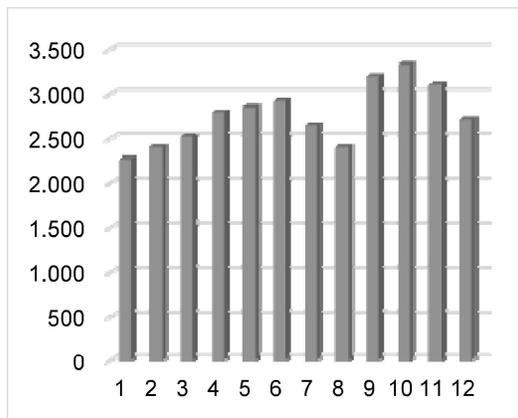


Figure 1. Distribution of the quantity produced during a calendar year

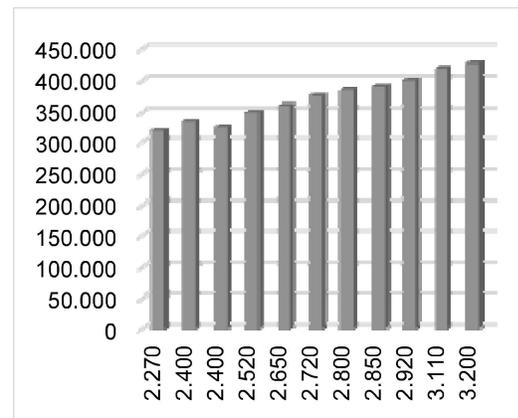


Figure 2. Distribution of total production cost over a calendar year

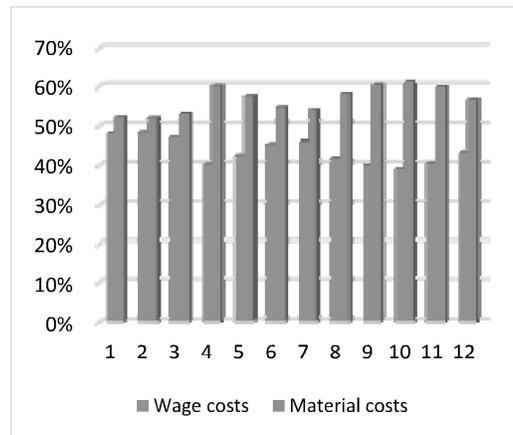


Figure 3. Distribution of wage and material costs over a calendar year

In Figure 1 it can be noticed that the lowest quantity produced is obtained in January and August and the highest in October and September.

In figure 2 it is observed that from the analysis of total production cost according to quantity produced we have the smallest quantity (2270 pieces) obtained in January and the highest (3200 pieces) obtained in October.

In figure 3 it can be noticed that the largest share is held by the material expenses followed by the salary expenditures. Also, we can see that in October we have the lowest level of wage costs (39%) and the highest level of material costs (61%). This situation is due to the highest production level (3350 pieces).

Following the statistical analysis using the multiple regression analysis, the following values were obtained in Table 2, considering the dependent variable, the material costs.

Table 2. Summary statistics

Statistic	Value
Multiple R	1.00
Multiple R <sup>2</sup>	1.00
Adjusted R <sup>2</sup>	1.00
F(3,8)	5.73e <sup>+15</sup>
p	0.00
Std.Err. of Estimate	8.71e <sup>-04</sup>

From the analysis of Table 2 we can see that the calculations resulted in R<sup>2</sup> for 1 which represents a high value close to value 1. This value shows that there is a strong linear relationship between the dependent variable and independent variables (direct link).

Following the calculations using the ANOVA variance analysis, the values given in Table 3 were obtained.

Table 3. Analysis of variance

Effect	Sums of squares	df	Mean squares	F	p-level
Regress.	1.31e <sup>+10</sup>	3	4.35e <sup>+09</sup>	5.736e <sup>+15</sup>	0.00
Residual	6.07e <sup>-06</sup>	8	7.59e <sup>-07</sup>		
Total	1.31e <sup>+10</sup>				

Table 4 presents the values of the statistical indicators obtained for the "Material costs" dependency and the "Quantity", "Total production costs" and "Wage costs" independence.

Table 4. Regression summary for dependent variable (Material costs)

N=12	Beta	Std.Err.	B	Std.Err.	t(8)	p-level
Intercept			1.45e <sup>-09</sup>	3.92e <sup>-03</sup>	0	1,00
Quantity	1.28e <sup>-13</sup>	1.15e <sup>-07</sup>	1.30e <sup>-11</sup>	1.16e <sup>-05</sup>	0	0,99
Total production costs	1,22	1.20e <sup>-07</sup>	1.00	9.86e <sup>-08</sup>	10,144,231	0,00
Wage costs	-0,34	1.33e <sup>-08</sup>	-1.00	3.88e <sup>-08</sup>	-25,764,996	0,00

As a result of the regression analysis (Table 4) it is observed that the coefficients B necessary for determining the regression equation (obtaining the equation for determining the material cost) are: the coefficient for the free term (1.45e<sup>-9</sup>), then for the quantity (1.30e<sup>-11</sup>), total production costs (1) and wage costs (-1). Regression line intercepts are interpreted as the variables taken into the study that allow the calculation of coefficients related to independent variables. The result of the regression is given by the intercept (1.45e<sup>-9</sup>), i.e. the free term, while the coefficients of the independent variables allow the slope of the regression line to be set. In other words, there is a probability that even if the first coefficient (free term) is zero, we can see an estimate of up to 1.45e<sup>-9</sup>. In the present case, this probability is very low.

Table 5 shows the current status of sweep matrix for dependent variable.

Table 5. Current Status of Sweep Matrix; DV: Material costs

Effect	Quantity	Total production costs	Wage costs	Material costs
Quantity	-227,12	236,96	-14,53	1,28e <sup>-13</sup>
Total production costs	236,96	-249,34	16,7	1,22
Wage costs	-14,53	16,7	-3,05	-0,34
Material costs	1,28e <sup>-13</sup>	1,22	-0,34	4,65e <sup>-16</sup>

The calculations resulted in the regression line indicating that the production cost equation can be written as the following relationship:

$$\text{Material costs} = 1.45e^{-9} + 1.30e^{-11} Q + \text{Total production costs} - \text{Wage costs}$$

Table 6 shows the regression residuals.

Table 6. Predicted and residual values for dependent variable: Material costs

Case no.	Observed value	Predicted value	Residual	Standard pred. val.	Standard residual	Std.Err. pred. val.	Mahalanobis distance	Deleted residual	Cook's distance
1	166800,0	166800,0	0,00	-1,40298	0,00	0,000519	2,993015	0,00	0,00
2	173800,0	173800,0	0,00	-1,19979	0,00	0,000434	1,810334	0,00	0,00
3	185700,0	185700,0	0,00	-0,85437	0,00	0,000384	1,223648	0,00	0,00
4	230600,0	230600,0	0,00	0,44895	0,00	0,000609	4,463023	0,00	0,00
5	224700,0	224700,0	0,00	0,27769	0,00	0,000288	0,285792	0,00	0,00
6	218800,0	218800,0	0,00	0,10643	0,00	0,000527	3,107936	0,00	0,00
7	194700,0	194700,0	0,00	-0,59312	0,00	0,000598	4,263207	0,00	0,00
8	189600,0	189600,0	0,00	-0,74116	0,00	0,000785	8,018147	0,00	0,00
9	258800,0	258800,0	0,00	1,26752	0,00	0,000446	1,967569	0,00	0,00
10	274800,0	274800,0	0,00	1,73195	0,00	0,000521	3,018279	0,00	0,00
11	250600,0	250600,0	0,00	1,02950	0,00	0,000373	1,101328	0,00	0,00
12	212700,0	212700,0	0,00	-0,07063	0,00	0,000339	0,747724	0,00	0,00
Minimum	166800,0	166800,0	0,00	-1,40298	0,00	0,000288	0,285792	0,00	0,00
Maximum	274800,0	274800,0	0,00	1,73195	0,00	0,000785	8,018147	0,00	0,00
Mean	215133,3	215133,3	0,00	0,00000	0,00	0,000485	2,750000	0,00	0,00
Median	215750,0	215750,0	0,00	0,01790	0,00	0,000483	2,480292	0,00	0,00

Residual regression analysis (Table 6) allows to identify the differences recorded during the year, presenting the values observed against the estimated ones and calculating the volume of activity: minimum, maximum and average. From the analysis of Table 6 it can be seen that in the Material costs (DV), according to Mahalanobis distance, the highest value is recorded in August, which show that these month there is the highest number of standard deviations (P) from the average (D), that is, it moves away from the average. It is observed that the residue of an observed value was obtained as the difference between the observed value and the predicted value. The smallest positive residue (according to Mahalanobis distance) is in May and the highest in August. Negative residues tell us that in those months the observed values are lower than predicted values.

Analysing Table 6 we can see that at the residuals of Material costs according to Cook's distance is not identified.

### Conclusions

The regression analysis allowed the determination of the Material production cost equation, the relationship between the "Material costs" (dependent variable) and the independent variable "Quantity", "Total production costs" and "Wage costs" which is useful in cost management as a result of the relationship of linear dependence between them. According to the example taken in the study, for linear distribution, linear regression analysis can establish the type of link between variables using

existing data in the company's accounting. Thus, starting from the Material costs, we can establish the optimum production level knowing minimum, maximum and average information.

Subsequently, the residue regression analysis allowed the identification of the differences recorded during the year, presenting the values observed against the estimated ones, which is useful for analyzing the changes in the volume of activity with direct implications in the production cost (high unit costs when we have residual values big). The analysis of the residue in the volume of activity through regression analysis is useful because it then allows the managers to identify the causes that caused those changes (these can be both technical and economical).

## References

- Cook D.**,  
1977  
"Detection of influential observation in linear regression", *Technometrics*, 19 (1): 15-18.
- Comiskey E.**,  
1966  
"Cost control by regression analysis", *The accounting review*, p. 235-238.
- Cochran E.B.**,  
1976  
"Using regression techniques in cost analysis", part 2, *Int. J. Prod. Res.*, 14 (4): 489-511.
- Covaleski, M.A., Dirsmith M.W. and Samuel S.**,  
2003  
"Changes in the institutional environment and the institutions of governance: extending the contributions of transaction cost economics within the management control literature", *Accounting, Organizations and Society*, 28 (5): 417-441.
- Ahrens, T. and Chapman C.S.**,  
2004  
"Accounting for flexibility and efficiency: a field study of management control systems in a restaurant chain", *Contemporary Accounting Research*, 21 (2): 271-301.
- Stockton, D.J., Middle J.E.**,  
1982  
"An approach to improving the cost estimating", *Int. J. Prod. Res.*, 20 (6): 741-751.
- Cooper, R. and R. Kaplan.**  
1988  
"How cost accounting distorts product costs", *Management Accounting*, 69 (4): 20-28.
- Shim, E. and E.F. Sudit,**  
1995  
"How manufacturers price products", *Management Accounting*, 76 (2): 37-39.
- Lucas, M. R.**,  
2003  
"Pricing decisions and the neoclassical theory of the firm", *Management Accounting*, 14: 201-207.
- Korok R. and J. Gramlic,**  
2016  
"Reconciling Full-Cost and Marginal-Cost Pricing", *Journal of Management Accounting Research*, Spring 28 (1): 27-37.
- Weisberg, S.**  
1985  
"Applied linear regression", Wiley, New York.
- Stout D.E.**,  
2015  
"Regression analysis: Instructional resource for cost/managerial accounting", *Accounting Education: An International Journal*, 24 (1): 57-73.
- Zellner A.**,  
1970  
"Estimation of regression relationships containing unobservable independent variables", *International economic review*, 11 (3): 441-454.
- Jensen R.E.**,  
1967  
"A multiple regression model for cost control - assumptions and limitations", *The accounting review*, p. 265-273.