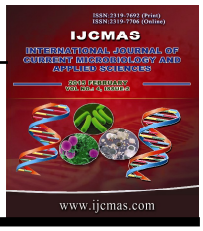
*Int.J.Curr.Microbiol.App.Sci* (2015) 4(2): 713-719

*International Journal of Current Microbiology and Applied Sciences*
*ISSN: 2319-7706* **Volume 4 Number 2 (2015) pp. 713-719**
http://www.ijcmas.com




**Original Research Article**

# Regression analysis of profit per 1 kg milk produced in selected dairy cattle farms

**K. Stankov[1]*, St. Stoyanova[2] and D. Petrov[3]**

[1]Department of Management, Faculty of Economics, Trakia University, 6000, Bulgaria
[2]Department of Biology and Aquaculture, Faculty of Agriculture, Trakia University, 6000, Bulgaria
[3]Department of Social medicine and Health management, Faculty of Medicine, Trakia University, 6000, Bulgaria
**Corresponding author*

**Keywords**

Profit,
Dairy cattle farming,
Regression and correlation analysis,
SPSS

**A B S T R A C T**

The aim of this article was to investigate the effects of several independent factors on the profit per 1 kg milk produced in Holstein-Friesian cows. The survey included 82 farms of different capacity and size from Central and Southeastern Bulgaria. Data were collected by filling in questionnaires, specially developed for the purposes of the research. The results were processed by means of licensed SPSS software. Data from regression and correlation analysis demonstrated that the different studied factors had a different effect on the profit from 1 kg milk produced. The factors with strongest influence on the profit at dairy farms from Central and Southeastern Bulgaria were the labour efficiency (r=0.982), followed by the farm size (r=0.722) and the lactation milk yield per cow (r=0.721). Farms of larger size which were characterised with higher production and labour efficiency, were more profitable.

## Introduction

Dairy cattle farming is a branch of livestock husbandry of particular significance. It has long-standing traditions in providing essential foodstuffs to consumers, which are of great importance even today.

The efficiency of agriculture is a priority issue being closely related to the viability of agricultural enterprises. For the different farms, including dairy cattle farms, increasing of profitability is particularly important in periods of financial troubles similar to those experience by national agriculture during the last years. In the long run, the chances for survival and prosperity of well-organized and profitable farms generating a higher income, are higher at the current economical background. Furthermore, the role of dairy farming for the development of the second pillar of the Common Agricultural Policy – rural

development policy, would become more and more important during the new programming period 2014 – 2020 (Stankov, 2012).

The abovementioned considerations require a thorough analysis of dairy cattle farming with emphasis on the most important product of the branch – the profit per 1 kg raw milk produced.

## Material and Methods

The following statistical analyses were used: Kolmogorov-Smirnov test of normal sample distribution and standard multiple logistic regression.

The regression analysis shows the relationships between variables, which could be determined as cause-effect relationships. The test is designed for solution of general problems – the type of the relationship, to determine the function of the relationship, quantitative determination of function's parameters. The variable whose variations have to be explained or predicted, are called dependent variables. The purpose of regression analysis is to determine how and at what extent the dependent variables change or vary as a function of changes in the fixed variable (Gatev, 1991).

Correlation analysis tests the interdependence, and the regression analysis – the causal effects among more than two variables. More than a single variable could be accepted as independent and to assess their effects on one dependent variable only. Usually, the coefficient of correlation is known as r.

It could take a value within the range from -1 to +1. When it is negative, the relationship between the two variables is reverse (in different directions) – when one variable gets bigger, the other gets smaller. With positive correlation coefficients, the two variables change together (in one direction) – when the one gets bigger or smaller, the other also gets bigger or smaller. The closer the value of r to -1 or +1, the stronger the relationship between variable. The weaker the relationship, the closer the value of r to zero. It is conditionally accepted that correlation coefficients between 0 and 0.3 suggest a weak relationship; from 0.3 to – moderate and over 0.7 – strong relationship. This is true for both positive and negative values.

The dependent variable is marked also with Y, and independent - with $X_i$, also called factors. In the general case, let us have *k* independent variables. The linear combination of all impendent variables is $b_1X_1 + b_2X_2 + \ldots + b_kX_k$, where $b_i$ are the correlation coefficients. The regression equation is of the following type:

$$Y = b_1X_1 + b_2X_2 + \ldots + b_kX_k + a,$$

where a – regression constant

The number of regression coefficients in the equation is equal to that of impendent coefficients plus the constant.

A correlation involving more than two dependent variables at a time, is called multiple correlation. Multicollinearity exists when the impendent variables are highly correlated (r=0.9 and higher).

## Result and Discussion

### Multiple regression analysis of the profit per 1 kg milk produced

In this study, the effects of some factors on the profit per 1 kg milk produced was investigated in Holstein-Friesian cows,

which is considered among the most highest-production dairy cattle breeds in the world and it the most suitable for rearing under the conditions in Bulgaria. In the study, 82 farms of various capacity and size from Central and Southeastern Bulgaria are included.

The profit of an enterprise (P) is defined as dependent variable and marked with Y, and independent variables and the influencing factors, marked with X. In this study, the included factors were:

***$X_1$ – Farm size*** *(number of cows)*

***$X_2$ – Productivity per employed person*** (AWU/man-hour) *

***$X_3$ – Feed cost per cow*** (BGN)

***$X_4$ – Milk yield per cow*** (kg)

***$X_5$ – Average purchase price of milk*** (BGN)

* AWU/man-hour) – annual working unit/ hour

These factors were chosen as they had the most significant effect in the formation of farm's profit and cows productivity.

The farm size and the concentration of production have an essential impact on farm organisation and management, and the administration of all processes related to milk production (Todorov, 2003).

The productivity per employed person is directly associated with milk production, milk yield, milk quality and the general condition of cows, their health in particular. The organisation of labour and production is very important, as well as the qualification of farm workers and specialists.

The feed cost per cow is included in the regression analysis as the feeding of animals has a direct effect on productivity and milk yields, and also constitutes a significant part of milk production costs.

The milk yield of a cow is one of most important factors influencing the profitability, because it influences significantly the overall produce of an enterprise and last but not least, the quality of milk also influences the overall economic status of the farm.

The average purchase price of milk determine the production threshold of the work of farmers after accounting for production expenses, the determination of the production cost of 1 kg milk and the formation of farm profitability (Stankov, 1997).

Standard multiple regression analysis was used to predict the farm profit (Y), based on the influence of the following factors:

- Farm size ($X_1$)
- Productivity per employed person ($X_2$)
- Feed cost per cow ($X_3$)
- Milk yield per cow ($X_4$)
- Average purchase price of milk ($X_5$).

The preliminary analyses ensured that there was no interference of the assumptions for normal distribution, linearity, multicollinearity and homoscedasticity.

The obtained correlation coefficients of the different factors vs the profit were as followed:

**$r_1$ = 0.722 p=0.0001**
**$r_2$ = 0.982 p=0.0001**
**$r_3$ = -0.070 p=0.266**
**$r_4$ = 0.721 p=0.0001**
**$r_5$ = 0.188 p=0.045**

The coefficients were indicative about a direct relationship; only the coefficient of correlation between the profit per 1 kg milk and feed costs per cow was statistically insignificant – **$r_3$ = -0.070, p=0.266>0.05,** and could be ignored.

The multiple correlation coefficient R is a measure of the association between the dependent variable and linear combination of independent factors. R=0.986, and the coefficient of determination $R^2$ is 0.973, which, multiplied by 100 explains what proportion of the variance of the dependent variable could be predicted on the basis of the independent variables (factors) of the model. In other words, 97.3% of the profit is explained by the model. The high value showed that the selected impendent variables (factors) in the model determined the profit.

The ANOVA table presents the results from the dispersion analysis. The dispersion parameter F=539.081 has a level of significance p= 0.0001, which means that the model is statistically significant.

Two types regression coefficients are calculated – non-standardised and standardised. The former (raw coefficients) in this study were as followed:

$b_1$ = 0.040
$b_2$ = 1.272
$b_3$ = - 0.184
$b_4$ = 0.045
$b_5$ = - 61.059
a = 1011.703

The following regression equation is obtained:

$Y = 0.040*X_1 + 1.272*X_2 – 0.184*X_3 + 0.045*X_4 – 61.059*X_5 + 1011.703$

The regression equation was statistically significant as the calculated level of significance was:

sign**. F = 0.0001 < 0.05**

The coefficients of $X_1$ and $X_5$ were statistically insignificant, i.e. the effect of these variables on Y was negligible and they could be omitted from the model.

This could be seen from their levels of significance p(sign.): $p_1$ = 0.586, $p_5$ = 0.970; both higher than 0.05.

The coefficients of $X_2$, $X_3$, $X_4$ were statistically significant as could be seen from respective levels of significance from the table.

Standardized regression coefficients were:

β1 = 0.016
β2 = 0.856
β3 = -0.082
β4 = 0.161
β5 = -0.005

The highest absolute value was that of $\beta_2$ (productivity per employed person). This means that the share of this independent variable (factor) explained at a most substantial extent the dependent variable changes (the profit).

The graph shows only very little deviations from the normal distribution (ideally, all points should lie on the line, but this is rarely observed in the practice). Obviously, the calculated mathematical model is relatively good.

**Table.1** ANOVA[a]

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 3730748,112 | 5 | 746149,622 | 539,081 | 0,000[b] |
| | Residual | 105192,597 | 76 | 1384,113 | | |
| | Total | 3835940,709 | 81 | | | |

a. Dependent Variable: profit on the farm

b. Predictors: (Constant), average purchase price of milk; feed costs per cow, farm size, milk yield per cow, productivity per employed person

**Fig.1**

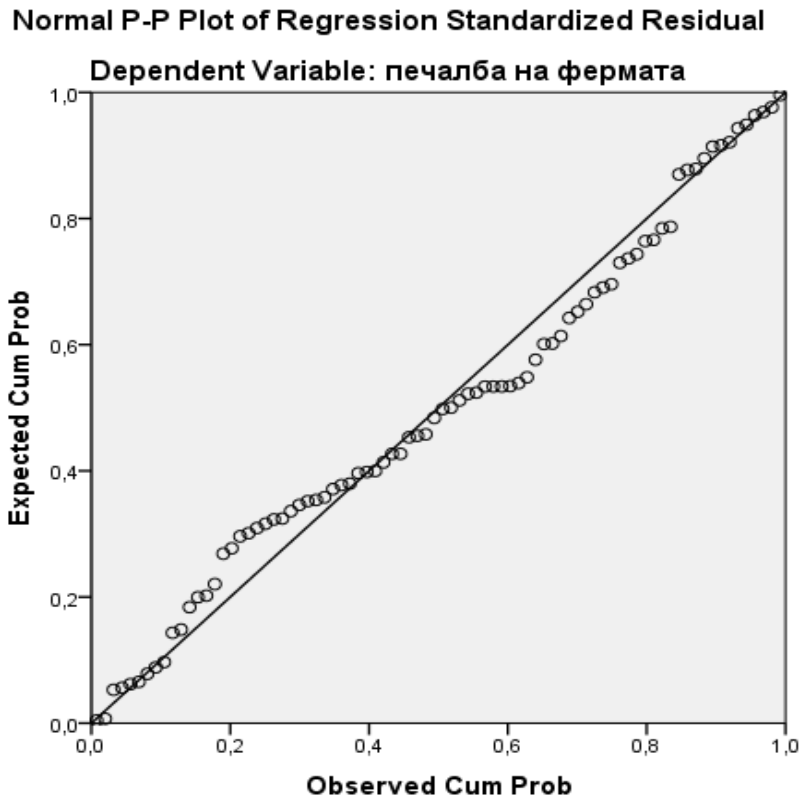
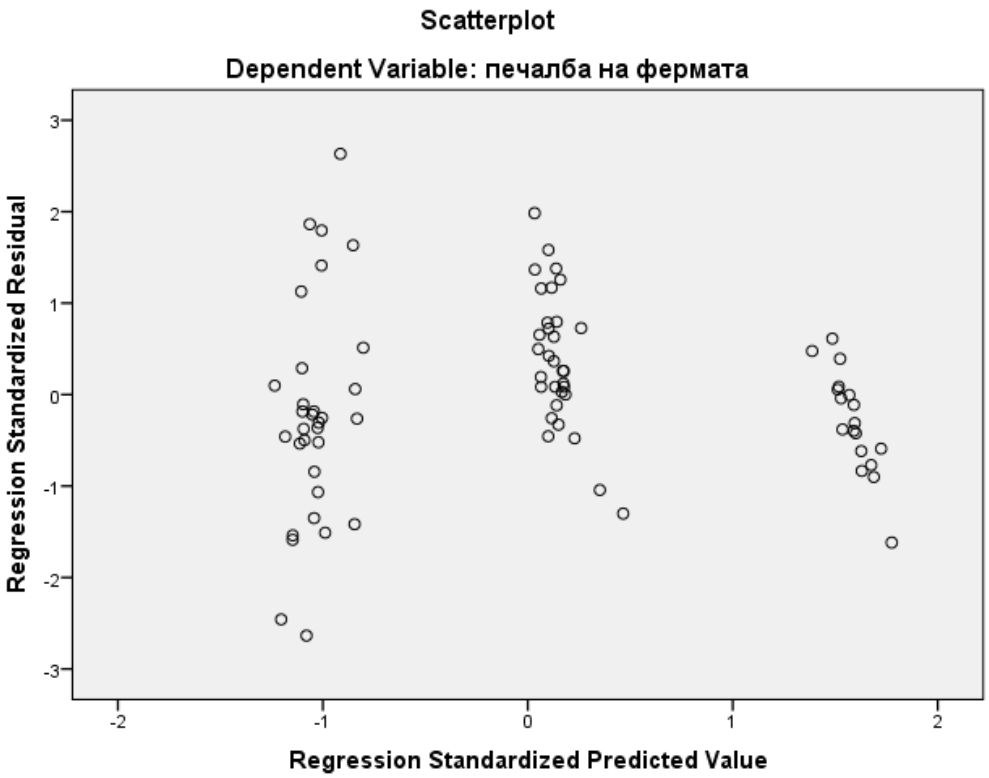


**Fig.2**




The multiple regression analysis is very sensitive to extreme values (very high or very low). Extreme values are those ≥ 3.3 or ≤ -3.3. The graph shows the range of data, i.e. there are no extreme values to be deleted. Also, the rectangles corresponding to the three types of farms could be observed, so no deviations from the hypotheses are present.

The respective standardised (beta) coefficients were calculated. Most commonly, it serves for comparison of the effects of the different factors when they are two or more. The factor with the highest beta coefficient has the strongest influence.

The analysis of data from the regression analysis showed that the factors with strongest influence on the profit of dairy farms from Central and Southeastern Bulgaria were the labour efficiency (r=0.982), followed by farm size (r=0.722) and the lactation milk yield per cow (r=0.721).

The magnitude of the correlation was strong, suggesting that the increase in the absolute values of these three independent factors would result in increased profit at the respective farms. The average purchase prices of milk with correlation coefficient of 0.188 i.e. r<0.30, are of little significance for the profit, and feed costs, being statistically insignificant, could be ignored.

The high multiple correlation coefficient R (R=0.986) and coefficient of determination (0.973) demonstrated that the selected independent variables (factors) were closely associated to the profit in the present model.

Similar high values of the independent variable labour force productivity are obtained after calculation of non-standardised and standardised regression coefficients, which means that this factor is with the highest relative proportion in profit formation. Although with lower regression coefficient, the positive role of lactation milk yield for the profit at the farm should be also acknowledged.

## References

Chobanov, G., 1999. Statistics. Manual for Economics Students. p. 134.

Gatev, K., Koseva, D., Spasov, A., 1991. General Theory of Statistics. p. 274.

IBM SPSS Statistics 21 Brief Guide.

Rangelova R., 2010. A New and Useful Statistics Textbook.

SPSS Base 16.0 User's Guide. SPSS - Short Manual.

Stankov, K., 6, 2012. The role of dairy cattle farming in rural development. Management and sustainable development. Volume 38, ISSN 1311 – 4506.

Stankov, W. 1997. Management and organization of the agricultural enterprise (Agricultural management). PSSA. p. 151.

Todorov, N. 2003. Dairy cattle stock nutrition and breeding. Zemizdat LLC. p. 125 - 128.

http://www.quantitativeskills.com/sisa/statistics/corrhlp.htm