CHAPTER - VI

FACTORS AFFECTING MILK PRODUCTION: A REGRESSION ANALYSIS

Milk production is a complex process and can be conceived as a function of several variables. Production of milk depends on various factors like feeding, breeding and management of the animals. Apart from these, other factors like age at first calving, season of calving, advancement of lactation, number of lactations, human labour, age of the animal, value of the animal, etc., will also influence milk production. The knowledge of relative importance of the resource inputs influencing milk production is essential for the dairy farmer for introducing desirable changes in his operation at the micro level.

In this connection an attempt is made in this chapter to examine the average milk yields of buffaloes and inputs used by different categories of farmers in the selected area and also to examine the factors affecting milk production i.e. the relative importance of inputs used in milk production. The average milk yields, on an average, worked out to 4.50 litres of milk per day per animal. The study also examines the relative importance of various inputs used in milk production. The data reveals that the green fodder and concentrates form critical inputs in all the groups of farms is used extensively by all the groups of farmers.
For this purpose, Production function technique is used as an analytical tool to assess the contribution of various inputs to milk production. In the production function analysis, choice and specification of the variables are of considerable importance. The fitted model will be a biased one if a relevant variable is omitted or an unwarranted variable is included.

Regarding the management aspects, human labour hours (both family and hired) spent on animal for its maintenance are taken as a proxy. With regard to feeds, green fodder, dry fodder and concentrates are considered for the analysis. Among the other variables, the number of lactations, age of the animal and value of the animal are considered as the factors affecting milk yield and all the remaining variables are excluded from the analysis due to paucity of data and measurement problems.

6.1 Studies on Determination of milk production

Sambasiva Rao¹ (1985), study indicates that quantities of dry fodder, green fodder and concentrates, number of lactations, labour hours and the age of the animals had together explained more than three-fourths of the variations in the milk yield among marginal, small, medium, large and big farmers. It was also observed that the marginal value product of labour in all size groups except marginal farms was less than the factor cost. The marginal value product of dry fodder was greater than unity for the marginal and large farmers, less than one for the big farmers. In the case of green fodder and
concentrates the marginal values were greater than unity for all size groups. Thus, there was a possibility of increasing milk yield by further use of green fodder and concentrates.

Virender Singh and K.N. Rai\textsuperscript{2} (1998) examined the economics of productions and marketing of Buffalo Milk in Haryana. The study found that the feed and fodder cost was the most important item of the total maintenance cost in Zone - II. The net profit per day of a milch buffalo was very low due to the higher maintenance cost and low milk yield of milch buffalo on each herd size-group in each zone of the state. The net profit from milk production per buffalo per day was observed to be higher in the case of small size-group due to higher milk yield of milch buffaloes in this size-group as compared to medium and large herd size-groups in both the zones. However, all the herd size-groups in each zone were operating above the break-even levels. But the average yield in most of the farms was very near to the break-even point making them vulnerable to fall below with minor change in milk yield and its prices/price of feed and fodder. Price of milk was found to be the most important factor influencing volume of milk business significantly, besides production level. The establishment of milk co-operative societies in the rural areas had positive impact on the marketed surplus of milk.

Shantanu Kumar and Uma Sah\textsuperscript{3} (2000) explained various parameters of dairy development. Bovine density, cattle(buffaloes ratio, crossbred population, number of cooperative societies and
producer members per society and milk procured per society per day, A.I. routes per 1,000 breedable bovine population, cattle feed production and productivity of milch animals were found as an important decisive indicators responsible for imbalance in dairy development in different regions.

Surya Murthi⁴ (2001) has expressed that milk production can be improved substantially with minimum cost though crossbreeding in case of cows and selective (upgrading) breeding in case of buffaloes, scientific and least-cost effective feeding practices, prevention and control of disease measures and judicious management apart from assured market for milk. The steps taken to improve dairy farming will provide not only constant and regular gainful employment but also assured income to farmers/landless labourers in rural areas, which, in turn, improves the standard of living of people in rural areas.

Vijay Gorakh Patil⁵ (2010) had studied the production cost of milk at the farmer level in Shirpur Tehsil of Dhule District of maharashtra state. The investigator had selected fifty dairy farmers from eight villages from Shirpur Tehsil, District Dhule. The questions related to fixed and variable cost were asked. The object of the study is to estimate the cost of milk production. The total cost of milk production per farm was Rs.113.87 in which the variable cost has been 83.76 per cent (Rs.95.38) and remaining 16.24 per cent (Rs.18.49) was fixed cost. The variable cost has been the main
component of the cost of production. In variable cost, the cost of feed stuff and the cost of Labour are the main.

6.2 Specification of the Variables

6.2.1 Milk Yield:

Due attention has been paid in collecting the data on milk yield. Data on milk yield is collected at four quarterly intervals for fifteen days in each quarter. So altogether there are 460 observations. The weighted average of these 460 observations is taken by multiplying the milk yield with the corresponding market prices (The price is not constant).

6.2.2 Green Fodder:

The common green fodders are local grass, Lucerance, Jowar, guinea grass, bajra, Rhodes grass, sugarcane tops etc. but in almost all the cases local grass, pillipesara, jute is fed to the animals. Because of this, without using conversion factor (TDN Equivalent) the quantity of green fodder fed to the animals (either owned or purchased) in value terms is considered as the variable, while in the case of home produced grass, the imputed value of the grass is taken into consideration.
6.2.3 Dry Fodder:

Dry Fodder is poor in nutritional quality but it is the major item of feed. It comprises of paddy straw, wheat straw, bajra, ragi straw, kadbi, lucerne hay, local grass hay, pulses plant residues and groundnut plant residues. But it is observed in the study area that paddy straw; bajra straw, jowar straw, horse gram straw and jute straw are available to feed the animals. As all these are of the same constitution, no attempt is made to convert them into TDN equivalent. The quantity of the fodder fed per animal per day in value terms is used as a variable explaining milk production. In the case of home produced fodder, imputed value of the fodder is taken into account.

6.2.4 Concentrates:

Concentrates feeds are very essential to increase milk production because animals cannot get sufficient nutrient feed through fodders in order to meet the requirements of growth and production. The common concentrate feeds are ricebran, wheat bran, rice husk, pulses husk, groundnut cake, gingelly cake, cotton seed, mustard seed, broken rice, bajra grain, maize grain, jowar grain, gold mohur and vijaya feed etc.

6.2.5 Labour hours:

Labour hours spent on the tending of cattle has a considerable impact on milk production. It is already mentioned that feeding and management are the fundamental inputs to raise the productivity of
animals. Since management is an important aspect, maintenance of animals needs more amount of labour both family and hired. This variable includes time spent on bringing fodder from the field, chaffing, feeding the animals, milking the animal, cleaning the cattle shed etc. whenever the use of labour is on an aggregate basis, it is apportioned to different animals on the basis of the quality fed to them.

6.2.6 Age of Animal:

Age of an animal has a significant impact on milk production. It is a well known fact that as the age of a particular animal increases, the milk yield increases in every successive lactation up to a particular age, then starts decline in the milk yield. The age of animal is specified in this function, in which the animal is in milk.

6.3 Determinants of Milk Yield

The Model considered for identifying the determinants of milk yield can be expressed as follows:

\[ Y = f( X_1, X_2, X_3, X_4, X_5 ) \]

where

\[ Y \quad = \quad \text{Milk Yield per Day per Animal (Rs.)} \]
\[ X_1 \quad = \quad \text{Fodder used per animal per day (in Rs)} \]
\[ X_2 \quad = \quad \text{Green fodder used per animal per day (in Rs)} \]
\[ X_3 \quad = \quad \text{Concentrate used per animal per day (in Rs)} \]
\[ X_4 \quad = \quad \text{No of Labour hours required per day} \]
\[ X_5 = \text{Age of animal (in Years)} \]

The corresponding **Multiple Linear Regression Model** to be estimated can be specified as:

\[ Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \epsilon \]

Where \( \alpha \) is the intercept and \( \beta_1, \beta_2, \beta_3, \beta_4, \beta_5 \) are the regression coefficients to be estimated and \( \epsilon \) is the error term.

Primary data is collected from four villages, viz., Kandulapuram, Kolalapudi, Konanki and Ravipadu on 460 randomly selected individual households using a structured questionnaire.

The **Multiple Linear Regression Model** as specified above is estimated by the method of **Ordinary Least Squares** for the entire sample and also for four subsamples of individual villages. As there is a possibility of **multicollinearity** problem, the correlation matrix of independent variables was calculated (Table – 6.1) and it reveals that none of the independent variables have significant correlation between them thus validating, OLS estimation of the impact of each independent variable on the dependent variable. But on inspection of OLS residuals graphically and by White's test, it is found that the **heteroscedasticity problem** is present in the OLS regression. Therefore, **OLS procedure with correction for heteroscedasticity is**
chosen to estimate the specified model. The regression results are tabulated below and the inferences drawn is discussed subsequently.

Table – 6.1

Correlation Coefficients

<table>
<thead>
<tr>
<th></th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1.0000</td>
<td>0.0660</td>
<td>0.1719</td>
<td>-0.0401</td>
<td>0.1153</td>
</tr>
<tr>
<td>$X_2$</td>
<td></td>
<td>1.0000</td>
<td>0.1975</td>
<td>0.1165</td>
<td>0.0708</td>
</tr>
<tr>
<td>$X_3$</td>
<td></td>
<td></td>
<td>1.0000</td>
<td>-0.0820</td>
<td>0.0749</td>
</tr>
<tr>
<td>$X_4$</td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>0.0048</td>
</tr>
<tr>
<td>$X_5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Using the observations 1 – 460

Under the null hypothesis that the variables are correlated

5% critical value (two-tailed) = 0.0914 for n = 460
Model - 1:

6.4 Heteroscedasticity -corrected OLS

Table – 6.2
Regression Coefficients

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>coefficient error</th>
<th>standard</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>144.601</td>
<td>16.21</td>
<td>8.917</td>
<td>0.000 ***</td>
</tr>
<tr>
<td>X₁</td>
<td>0.592</td>
<td>0.17</td>
<td>3.383</td>
<td>0.000 ***</td>
</tr>
<tr>
<td>X₂</td>
<td>0.191</td>
<td>0.08</td>
<td>2.313</td>
<td>0.021 **</td>
</tr>
<tr>
<td>X₃</td>
<td>0.963</td>
<td>0.22</td>
<td>4.325</td>
<td>0.000 ***</td>
</tr>
<tr>
<td>X₄</td>
<td>2.400</td>
<td>1.64</td>
<td>1.458</td>
<td>0.145</td>
</tr>
<tr>
<td>X₅</td>
<td>5.09</td>
<td>2.06</td>
<td>2.462</td>
<td>0.0142 ***</td>
</tr>
</tbody>
</table>

No of observations (n = 460)

Dependent variable: Y

** Indicates significant at 5% level,

*** Indicates significance at 1% level.

Statistics based on the weighted data:

R-squared          0.675226  Adjusted R-squared 0.665041
F(5, 454)          7.38      P-value(F) 0.000

Excluding the constant, p-value is highest for (X⁴)
Regression results for the total sample (Model-1), shows that the estimated model was able to explain about 66 per cent of the variations in the dependent variable. Except for the variable, number of labour hours required per day ($X_4$), all the remaining independent variables have statistically significant positive impact on the dependent variable, milk yield ($Y$). The overall significance of the model is validated by the value of F statistic. Out of all the independent variables, concentrates used per animal per day ($X_3$) turned out to be statistically the most significant positive impact on milk yield. In terms of magnitude, age of the animal has the highest impact on milk yield. Similarly the fodder used per animal per day ($X_1$) and the green fodder used per animal per day ($X_2$) have considerable positive impact on milk yield as indicated by the positive sign of the estimated coefficients. However number of labour hours spent per day ($X_4$) does not have significant impact on the dependent variable. These results are in tune with the results of earlier studies.
Model - 2:

### 6.5 Heteroscedasticity - corrected OLS:

**Table – 6.3**

Regression Coefficient of Kandulapuram

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>coefficient error</th>
<th>standard</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>81.13</td>
<td>4.14</td>
<td>19.59</td>
<td>0.000 ***</td>
</tr>
<tr>
<td>X₁</td>
<td>0.14</td>
<td>0.062</td>
<td>2.38</td>
<td>0.018 ***</td>
</tr>
<tr>
<td>X₂</td>
<td>0.26</td>
<td>0.055</td>
<td>4.72</td>
<td>0.000 ***</td>
</tr>
<tr>
<td>X₃</td>
<td>0.41</td>
<td>0.024</td>
<td>17.20</td>
<td>0.000 ***</td>
</tr>
<tr>
<td>X₄</td>
<td>1.43</td>
<td>0.44</td>
<td>3.20</td>
<td>0.001 ***</td>
</tr>
<tr>
<td>X₅</td>
<td>0.45</td>
<td>0.46</td>
<td>0.97</td>
<td>0.332</td>
</tr>
</tbody>
</table>

Using Sampling observations for Village, Kandulapuram (n = 136)

Dependent variable: Y

*** Indicates significance at 1% level

Excluding the constant, p-value is highest for variable (X₅)

Statistics based on the weighted data:

- R-squared: 0.801
- Adjusted R-squared: 0.794
- F(5, 130): 105.11
- P-value(F): 0.000

Excluding the constant, p-value is highest for variable (X₅)
The model specified performed well for the data on Kandulapuram village which consists of 136 observations. The adjusted R-squared value indicates that independent variables of the model account for about 79 per cent of the variations in the dependent variable. The overall significance of the model is validated by the value of F statistic. Except for the age of the animal ($X_5$) all the remaining independent variables have shown statistically significant positive impact on milk yield. Among the independent variables, Concentrate used per day ($X_3$) was statistically the most significant and in terms of magnitude of impact, number of labour hours required per day ($X_4$) is the dominant factor. Green fodder and other fodders used are also influencing the milk yield of animals positively.
Model - 3:

6.6 Heteroscedasticity-corrected OLS

Table – 6.4

Regression Coefficient of Kolalapudi

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>coefficient error</th>
<th>standard error</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>81.76</td>
<td>3.96</td>
<td>20.62</td>
<td>0.008 ***</td>
</tr>
<tr>
<td>X₁</td>
<td>0.240</td>
<td>0.09</td>
<td>2.64</td>
<td>0.009***</td>
</tr>
<tr>
<td>X₂</td>
<td>0.14</td>
<td>0.03</td>
<td>4.66</td>
<td>0.000***</td>
</tr>
<tr>
<td>X₃</td>
<td>0.287</td>
<td>0.06</td>
<td>4.27</td>
<td>0.000***</td>
</tr>
<tr>
<td>X₄</td>
<td>0.630</td>
<td>0.50</td>
<td>1.23</td>
<td>0.214</td>
</tr>
<tr>
<td>X₅</td>
<td>0.806</td>
<td>0.34</td>
<td>2.32</td>
<td>0.022 **</td>
</tr>
</tbody>
</table>

using sampling observations for village Kolalapudi (n = 105)

Dependent variable: Y

** Indicates significant at 5% level,

*** Indicates significance at 1% level

Statistics based on the weighted data:

R-squared 0.637723  Adjusted R-squared 0.619

F(5, 100) 6.23  P-value(F) 0.000

Statistics based on the original data:

Excluding the constant, p-value is highest for variable (X₁)
The model specified has a moderate fit for the data on village Kolalapudi with a sample size of 106 observations. The independent variables of the model could explain about 62 per cent of variations in the milk yield of animal. Fodder used per animal ($X_1$), green fodder used per animal ($X_2$) and concentrate used ($X_3$) have displayed statistically significant positive impact on milk yield. As the case with the earlier results number of labour hours required per day ($X_4$) does not have significant impact on milk yield. The age of animal ($X_5$) seems to have strong positive impact on milk yield as per the regression results. Even though the model could moderately explain the variations in the dependent variable, the overall fit of the model is quite good as indicated by the p-value of the F statistic.
Model - 4:

6.7 Heteroscedasticity - corrected

Table – 6.5
Regression Coefficients of Konanki

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>coefficient error</th>
<th>standard</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>104.93</td>
<td>4.45</td>
<td>23.54</td>
<td>0.000 ***</td>
</tr>
<tr>
<td>X(_1)</td>
<td>0.18</td>
<td>0.07</td>
<td>2.47</td>
<td>0.022 **</td>
</tr>
<tr>
<td>X(_2)</td>
<td>0.09</td>
<td>0.05</td>
<td>1.82</td>
<td>0.07</td>
</tr>
<tr>
<td>X(_3)</td>
<td>0.56</td>
<td>0.15</td>
<td>3.74</td>
<td>0.000 ***</td>
</tr>
<tr>
<td>X(_4)</td>
<td>0.98</td>
<td>0.74</td>
<td>1.32</td>
<td>0.189</td>
</tr>
<tr>
<td>X(_5)</td>
<td>1.84</td>
<td>0.69</td>
<td>2.66</td>
<td>0.009 ***</td>
</tr>
</tbody>
</table>

using sampling observations for village Konanki (n = 100)

Dependent variable: Y

** Indicates significant at 5% level,

*** Indicates significance at 1% level

Statistics based on the weighted data:

R-squared 0.604   Adjusted R-squared 0.592
F(5, 100) 5.13   P-value(F) 0.0003

Excluding the constant, p-value is highest for variable (X\(_4\))
For the village Konanki the sample data consists of 106 observations and the model fit is appropriate as the F statistic is highly significant with about 60 per cent of the variations in dependent variable are captured by the independent variables. Both fodder used per day ($X_1$) and concentrate used ($X_3$) have significant positive impact on milk yield. However green fodder ($X_2$) and number of labour hours required per day ($X_4$) does not show significant impact on milk yield. The impact of age of the animal on milk yield is quite significant even in this sample.
**Model - 5:**

**6.8 Heteroscedasticity - corrected OLS:**

**Table – 6.6**

**Regression Coefficients of Ravipadu**

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>coefficient</th>
<th>standard</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>204.93</td>
<td>20.97</td>
<td>9.77</td>
<td>0.000 ***</td>
</tr>
<tr>
<td>X₁</td>
<td>1.54</td>
<td>0.26</td>
<td>5.82</td>
<td>0.000 ***</td>
</tr>
<tr>
<td>X₂</td>
<td>0.37</td>
<td>0.14</td>
<td>2.61</td>
<td>0.01***</td>
</tr>
<tr>
<td>X₃</td>
<td>1.34</td>
<td>0.56</td>
<td>2.39</td>
<td>0.018***</td>
</tr>
<tr>
<td>X₄</td>
<td>1.60</td>
<td>0.95</td>
<td>1.68</td>
<td>0.096</td>
</tr>
<tr>
<td>X₅</td>
<td>-0.83</td>
<td>1.79</td>
<td>-0.46</td>
<td>0.64</td>
</tr>
</tbody>
</table>

using observations for village Ravipadu (n = 119)

Dependent variable: Y

*** Indicates significance at 1% level

Statistics based on the weighted data:

R-squared 0.697495  Adjusted R-squared 0.67870

F(5, 109) 9.231804  P-value(F) 0.000

Excluding the constant, p-value is highest for variable (X₅)
The model fit for Ravipadu village with a sample size of 115 observations is reasonably good with an adjusted R-squared value of about 68 percent and highly significant F statistic. For this sample also fodder ($X_1$), green fodder ($X_2$) and concentrate used ($X_3$) are the major explanatory variables for milk yield and concentrate used per animal ($X_3$) being the most dominant factor as indicated by the magnitude of the coefficient. However both number of hours required per day ($X_4$) and the age of animal ($X_5$) does not show significant impact on milk yield.

6.9 Summary

The regression results for the total sample and also for subsamples reveal that on an average 60 to 80 percent of the variations in the value of Milk Yield Per Day Per animal could be explained by fodder used per animal per day, Green fodder used per animal per day, Concentrate used per animal per day, and the age of the animal. Of these explanatory variables Fodder used per animal, green fodder used per animal per day are highly significant with positive impact and in terms of magnitude of impact on dependent variable, the concentrate used per animal per day is the highest. However the explanatory variable, Number of Labour hours required per day did not show any significant impact on the dependent variable. The model fitted for the various samples are valid as indicated by their respective p-values of the F statistic. The model predictability power can be further augmented by adding new explanatory variables but that could lead to the problem of multicollinearity.