Correlation and Regression Analysis to study the Variables Affecting the Technological Gapin Sericulture Production

Dakhani N.T.[1], Sananse S.L.[2*], Hake D.A.[3]

{[1]Research Scholar, [2]Associate Professor} Department of Statistics, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (MS), 431004, INDIA.

[3]Sericulture Development Officer, Sericulture Dept., Govt. of Maharashtra, Aurangabad, Maharashtra, INDIA.

[*]Corresponding Address:

dr.sananse5@gmail.com

Research Article

Abstract: The correlation and Regression analysis has many applications in agricultural science.Here the these techniques wereused to estimate the technological gap in adoption of sericulture practices by the farmers and to identify important personal and socio economic characters which affect the technological gap in production. Thirteen independent variables were selected for the analysis. The correlation analysis indicated that the dependent variable "technological gap" exhibited negative and significant relationship with education, social participation, and experience in sericulture farming, cosmopoliteness, risk preference, extension contact, knowledge, economical motivation, annual income and area under mulberry cultivation of the sericulturists. This clearly indicated that due to increase in the levels of these variables, the technological gap in production reduced significantly. The multiple correlation and regression indicated that that a set of 13 independent variables under study had explained 89.80 percent variation in composite technological gap whereas remaining 10.20 percent variation might be due to the factors not included in the study.

Key Words: Correlation, Regression, Technological Gap, Socio economic variables.

1.0 Introduction

India stands second among the silk producing countries of the world. India produces all the four types of natural silks and enjoys the monopoly of producing mega silk in the world. In India, sericulture is recognized as an instrument for social and economic development and hence occupies an important place in the development plans of the country. Maharashtra is considered as a non-traditional state of sericulture. Sericulture in Maharashtra did not gain importance due to alternate cropping pattern like Sugarcane, Cotton and other cash crops (Jadhav 1999). The sericulture enterprise is characterized by low investment and as such small and marginal farmers dominate the enterprise. While the income from agricultural crop production is seasonal, sericulture provides a year round income, which is an important incentive for the small farmers to take up sericulture. Technology in sericulture is changing at on increasing rate which is being noticed through the improvement in the levels of productivity.Adoption of new technologies in sericulture happens to be inevitable as to make the enterprise more productive, However, there still exists a notable gap in their diffusion at the field level. This emanates as a result of various factors such as social, economical, psychological etc. Beed district in Maharashtra stands second in area under mulberry cultivation and cocoon production. However, the utility of sericulture technologies are yet to be well taken up by the sericulturists in the Beed district. Hence, the present study was carried out in Beed district with the main objective to study the technological gap in adoption of recommended sericulture practices.

2.0 Research Methodology

2.1 Sampling Technique:

The present study was conducted in Beed district of Maharashtra, where sericulture is becoming more popular farm enterprise among the farming community in the district. A sample 93 sericulturists were selected as respondents for the study, using a probability proportionate sampling (PPS) method. The selected respondents were personally interviewed using a schedule designed for the purpose.

2.2 Statistical Methods Used:

A) Correlation Analysis: Was applied to test the association of the Independent with Dependent Variables.

B) Multiple Correlation Analysis: Was carried out to estimate multiple correlation (R) and Coefficient of multiple determination R^2

C) Multiple Linear Regression Analysis: Was carried out to know the functional relationship of the Independent Variables with Dependent Variables.

A) Correlation Coefficient

For assessing the relationship between personal and socio-psychological characteristics, as well as the parameters of economic feasibility with dependent variables technological Gap, the **KARL Pearson's Co-efficient of Correlation,** "r" was worked out as under:

$$r_{xy} = \frac{\Sigma xy - 1/n \Sigma x \Sigma y}{\sqrt{\Sigma x^2 - \frac{(\Sigma x)^2}{n}} \sqrt{\Sigma y^2 - \frac{(\Sigma y)^2}{n}}}$$

Where,

r = Coefficient of Correlation
X = Score of Independent Variables
Y = Score of Dependent Variables
n = Number of Respondents.

B) Multiple Correlation Analysis

The *multiple* correlation coefficients generalize the standard coefficientof correlation. It is used in multiple regression analysis toassess the quality of the prediction of the dependent variable. Itcorresponds to the squared correlation between the predicted andthe actual values of the dependent variable. It can also be interpretedas the proportion of the variance of the dependent variableexplained by the independent variables. In statistics, the coefficient of **multiple correlation** is a measure of how well a given variable can be predicted using a linear function of a set of other variables. It is measured by the square root of the coefficient of determination, but under the particular assumptions that an intercept is included and that the best possible linear predictors are used, whereas the coefficient of determination is defined for more general cases, including those of nonlinear prediction and those in which the predicted values have not been derived from a model-fitting procedure. The coefficient of multiple correlation takes values between zero and one; a higher value indicates a better predictability of the dependent variable from the independent variables, with a value of one indicating that the predictions are exactly correct and a value of zero indicating that no linear combination of the independent variables is a better predictor than is the fixed mean of the dependent variable.The coefficient of multiple correlation, denoted *R*, is a scalar that is defined as the Pearson correlation coefficient between the predicted and the actual values of the dependent variable.

Estimation of Multiple Correlation (R) and Coeff. of Multiple Determination R^2:

The square of the coefficient of multiple correlation can be computed using the vector*c* of cross-correlations between the predictor variables (independent variables) and the target variable (dependent variable), and the correlation matrixR_{xx} of inter-correlations between predictor variables. It is given by

$$R^2 = c' R_{xx}^{-1} c,$$

where*c* ' is the transpose of *c*, and R_{xx}^{-1} is inverse of the matrix R_{xx}.

If all the predictor variables are uncorrelated, the matrix R_{xx} is the identity matrix and R^2 simply equals *c'c*, the sum of the squared cross-correlations with the dependent variable. If there is cross-correlation among the predictor variables, the inverse of the cross-correlation matrix accounts for this.The squared coefficient of multiple correlation can also be computed as the fraction of variance of the dependent variable that is explained by the independent variables, which in turn is 1 minus the unexplained fraction. The unexplained fraction can be computed as the sum of squared residualsthat is, the sum of the squares of the prediction errors divided by the sum of the squared deviations of the values of the dependent variable from its expected value.

Multiple Linear Regression Analysis:

General linear model:

In the more general multiple regression model, there are *p* independent variables:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i,$$

wherex_{ij} is the i^{th} observation on the j^{th} independent variable, and where the first independent variable takes the value 1 for all *i* (so β_1 is the regression intercept).

The least squares parameter estimates are obtained from *p* normal equations. The residual can be written as

$$e_i = y_i - \hat{\beta}_1 x_{i1} - \cdots - \hat{\beta}_p x_{ip}.$$

The **normal equations** are

$$\sum_{i=1}^{n} \sum_{k=1}^{p} X_{ij} X_{ik} \hat{\beta}_k = \sum_{i=1}^{n} X_{ij} y_i, \ j = 1, \ldots, p.$$

In matrix notation, the normal equations are written as

$$(\mathbf{X}^\top \mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}^\top \mathbf{Y},$$

where the *ij* element of *X* is x_{ij}, the *i* element of the column vector *Y* is y_i, and the *j* element of $\hat{\boldsymbol{\beta}}$is $\hat{\beta}_j$. Thus *X* is *n*×*p*, *Y* is *n*×1, and $\hat{\beta}$is *p*×1. The solution is

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y}.$$

The functional relationship of dependent variables with independent variables was studied by fitting Multiple Linear Regression Equation as below:

$$\mathbf{Y} = \mathbf{a} + \mathbf{b_1 x_1} + \mathbf{b_2 x_2} + \mathbf{b_3 x_3} + \mathbf{b_4 x_4} + \mathbf{b_5 x_5} + \mathbf{b_6 x_6} + \mathbf{b_7 x_7} + \mathbf{b_8 x_8} + \mathbf{b_9 x_9} + \mathbf{b_{10} x_{10}} + \mathbf{e}$$

Where ;

Y_1 = Dependent variable i.e. technological gap.
X_1 = Age
X_2 = Education
X_3 = Family size
X_4 = Social participation
X_5 = Experience in sericulture
X_6 = Cosmopoliteness
X_7 = Risk preference
X_8 = Extension contacts
X_9 = Knowledge about recommended sericulture practices.
X_{10} = Economic motivation
X_{11} = Size of land holding
X_{12} = Annual income
X_{13} = Area under mulberry cultivation
a = intercept.
$b_1 \ldots b_n$ = The partial regression co-efficient for the respective variable.
e = Error term

2.3 Estimation of Technological Gap

Composite or overall technological gap (Total Gap):

Thus the technological gaps observed in adoption of five major sericulture practices was summed and composite or overall technological gap index was computed. This was termed as composite or overall technological gap for the purpose of present investigation.

The composite technological gap for each respondents was calculated as:

Composite or overall technological gap =[∑Gap indices (I)] / N

Where,

I = Gap index in a particular package of practices.

N = Total number of practices.

Thus the composite or overall technological gap index was calculated by summing the gap indices of each selected practice and dividing it by total number of selected practices.

3.0 Results and Discussion

3.1 Socio-personal and socio-economic characteristics:

The respondents were categorized into different groups based on their age, family type,land holding and education and are presented in Table1.

It is revealed from the analysis presented in Table 1 that Majority of the respondents were middle aged (40.90 percent) followed by old age and young age.Majority (71.00 percent) of the respondents were having medium family size followed by large family size and small family size. With respect land holding of the sericulturists,40.86 percent of the respondents were medium farmers and 40.00 percent were big farmers.Majority (35.50 percent) of the respondents were illiterate followed by primary level 25.8 percent , secondary level 22.6 percent and 10.8 percent were higher secondary level. Very few were graduate and post graduate.

Table 1: Distribution of respondents according to their age , family type, land holdingand education

Sr. No.	Characteristics	Frequency (N = 99)	Percentage %
1	***Age***		
	Young age	21	22.6
	Middle age.	38	40.9
	Old age.	34	36.6
2	***Family Type***		
	Small family	11	11.8
	Medium family	66	71.0
	Large family	16	17.2
3	***Land holding***		
	Marginal farmers	04	4.3
	Small farmers.	13	13.98
	Medium farmers.	38	40.86
	Big farmers	38	40.00
4	***Education***		
	Illiterate	33	35.5
	Primary level	24	25.8
	Secondary level	21	22.6
	Higher secondary level	10	10.8
	Graduate level	03	3.2
	Post graduate level	02	2.2

3.2 Correlation Analysis

The correlation between the personal, socio-economic and psychological characters of the sericulturists and composite technological gap in adoption of recommended sericulture practices in Beed district is presented in Table 2. It is clearly evident that dependent variable "technological gap" exhibited position and significant relationship with age of the sericulturists and negative significant relationship with education, social participation, and experience in sericulture farming, cosmopoliteness, risk preference, extension contact, knowledge, economical motivation, annual income and area under mulberry cultivation of the sericulturists. This clearly indicated that due to increase in these variables the technological gap in production reduced significantly. Therefore these variables having negative correlation with technological gap were important variables to reduce the gap in production. The independent variable family size is positively related with technological gap, but it is non-significant. While the independent variable size of land holding is negatively related with technological gap, however it is non-significant.

Table 2: Correlation coefficient between composite technological gap and personal, socio-economic and psychological characteristics of the sericulturists.

Sr. No.	Independent Variables	Correlation Coefficient
1.	Age	0.368 **
2.	Education	-0.840 **
3.	Family size	0.091 ^{NS}
4.	Social participation	-0.576 **
5.	Experience in sericulture	-0.860 **
6.	Cosmopoliteness	-0.792 **
7.	Risk preference	-0.777 **
8.	Extension contact	-0.834 **
9.	Knowledge	-0.386 **
10.	Economical motivation	-0.933 **
11.	Size of land holding	-0.047 ^{NS}
12.	Annual income	-0.285 **
13.	Area under mulberry cultivation	-0.193 *

*, ** : Statistically significant at 5% and 1% level of significance.

NS : Statistically Non-Significant.

3.3 Multiple Correlation and Regression Analysis:

To find out the influence of different personal, socio-economic and psychological characters on composite technological gap in adoption of recommended sericulture practices, multiple repression analysis using linear model was carried out. The results are presented in Table 3. It was observed from the analysis that a set of 13 independent variables under study had explained 89.80 percent variation in

composite technological gap whereas remaining 10.20 percent variation might be due to the factors not included in the study. It was found that out of thirteen independent variables, regression coefficient for two variables namely education and knowledge were significant at 5% level of significance whereas risk preference and economic motivation are significant at 1% level of significance.

Table 3: Multiple regression analysis of different independent variables with dependent variable composite technological gap.

Sr. No.	Independent Variables	Regression coefficient	t-Calculated Value
1.	Age	0.023 ^{NS}	0.675
2.	Education	-1.257 *	2.201
3.	Family size	0.278 ^{NS}	1.380
4.	Social participation	-0.093 ^{NS}	0.235
5.	Experience in sericulture	-0.078 ^{NS}	0.335
6.	Cosmopoliteness	-0.168 ^{NS}	1.122
7.	Risk preference	-0.355 **	2.619
8.	Extension contact	-0.026	0.147
9.	Knowledge	-0.220 *	2.043
10.	Economical motivation	-1.675 **	6.879
11.	Size of land holding	-0.083	0.830
12.	Annual income	-2.49	0.047
13.	Area under mulberry cultivation	-0.339	0.419
	Coeff. Of Multiple Determination R^2 = 0.898 **	**F Calculated Value = 53.66**	

*, ** = Statistically significant at 5% and 1% level of significance respectively.
NS – Statistically non-significant.

Regression coefficient for five independent variables namely age, family size, social participation, experience in sericulture, and cosmopoliteness is non-significant. Extension contact, size of land holding, annual income and area under mulberry cultivation has no effect on composite technological gap in adoption of recommended sericulture practices.

References

1. Geeta, G.S. Srinivasa, G. and VijayaPrakash, N.B. (2003) "Importance of characters of farmers in adopting and sustaining bivoltine (CSR) race-a case study.'Recent Researches of Indian Sericulture Industry P.P.169 -171.
2. Jaishankar and Dandin S.B. (2004) "Socio-Economic Attributes in the Adoption of Improved sericlutural Technologies by farmers in KolarDistrit, Karnataka."Indian J.Seric; Vol. 43 (2) P.P.194-199.
3. Kotikhane,R.R.(1986) "A study on technological gap inadoption of irrigated farming systems under Purna Command Area of Maharashtra State." Ph.D.Thesis Submitted to Marathwada Agricultural University Parbhani.
4. Lakshmanan, S., and Geetha Devi, R.G. (2006)" Adoption of sericluturalTechnoloies in Mandyadistrit of Karnataka: Multivariate Logit Model Approach,".Abstracts: National Conference on New strategies in Research and Development of Sericulture – Indian Prospective, March (2006), Bangalore Universit, Bangalore P.P. : 117.
5. Meenal, M. and Rajan, R.K. (2007) "Impact of Socio-economic characters of sericulturists on knowledge, adoption and cocoon production in Tamil Nadu."IndianJ.Seric., 46 (1) : 49-51.
6. Mallikarjuna, B., Sariful Islam, Srikantaswamy, K. (2009) "A study on knowledge and adoption of bivoltine sericulture technologies". Karnataka J.Agric.Sci., 22 (5) : 1113-1115.
7. Qadri, S.M., Mani, A, Masilamani, S., Krishnamoorthy, T.S. Sakthivel, N., Vijayakumar, R. and Thirunavukkarasu, T. (2003) "Socio-economic profile of sericluturists in potential districts of Tamil Nadu."Abstracts : National conference on Tropical Sericluture for Global competitiveness, November (2003), CSR & TI, Mysore : 150-151.
8. Shankar, L., Jayaram, H.andGeethadevi, R.G. (2000) "An analysis of factors discriminating technology use among sericulturists in Karnataka". Indian J.Seric., 39 (1) : 52-59.
9. Sreenivasa, G., Dolli, S.S.Raveendra, M.andIyenger, M.N.S. (1996) "Socio-economic factors and their relation to adoption of improved sericlutural practices." Indain J. Seric., 35 (1) : 43-45.
10. Srinivasa, GyanaRanjanSahoo, Rahmathulla, V.K. and Vindhya, G.S. (2008) "Knowledge and adoption of sericulture technologies by sericluturists in Mysore district."Karnataka, (2008) State level workshop on Sericluture Management, Bhubaneshwar, April. (2008)
11. Sujatha, B. Lakshminarayana Reddy, P.SankarNaik, S., Sreenivasulu Reddy, P. and Sujathamma, P. (2008)"Overall Knowledge and adoption of recommended sericlutural practices by sericluturists in Anantpur district of Andhra Pradesh-A Case Study." Proceedings : National Seminar on Seribiotechnolgical Research In India, August (2008), Sri PadmavatiMahila -Vishvavidyalayam, Tiupati : 257-260.