International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-2, December 2019

Relationship Between Year of Schooling and Household Income: Special Reference with Two Localities of Dibrugarh and Jorhat District of Assam

Bharat Bonia, Abhinanda Bordoloi, Sarat Chandra Kakaty

Abstract*: : *The study is based on the relationship between year of schooling of children and their parents' income. Gary Becker, a noble prize winning economist made a preposition in 1950s that the amount of education had a direct bearing on income. This study is based on the relationship between years of schooling of children and the income of their parents and how the socio-economic condition of a family impact on the World's Human Development. The data for the study was collected from two localities of Jorhat and Dibrugarh District of Assam, India. The data was primary and it consists of 750 people from a village and a tea tribe of 599 people. Methods used to analyse the data was Ordinary least squares.

Keywords*: *Year of Schooling, Household Income, Dibrugarh, Jorhat, Tea Tribe, Relationship, Ordinary least Square method.

I. INTRODUCTION

A. EDUCATION:

Education refers to the intellectual and moral training that systematically enlarges the spheres of knowledge, develops character as well as mental faculties and evolves a definite pattern on relationship between an individual and the society.

Education may be formal as well as informal. Both are equally essential for the society. In formal education, years of education or education years is defined as the years of schooling starting from class 4 up till attaining Doctor of philosophy (Phd).

B. ROLE OF EDUCATION:

Education has many benefits and has positive impact in our life. An educated person is an asset for any country. In today's world, human capital is considered the best national resource. On one hand, he/she can explore better opportunities for himself/herself, and on the other hand, the entire nation would get benefited from his/her work.

Revised Manuscript Received on December 30, 2019.
* Correspondence Author

Bharat Bonia*, Department of Economics, Dibrugarh University, North Lakhimpur, India. Email: bharatbonia2019@gmail.com

Abhinanda Bordoloi, Department of Statistics, Dibrugarh University, Jorhat, India. Email: abhinandabordoloi2018@gmail.com

Sarat Chandra Kakaty, Professor Dept. of Statistics, Dibrugarh University, India.

Education is the backbone of developing countries. Developing countries should understand the importance of education and develop a sound educational system. Sustainable development of economy and society is closely related to education. The better educated we are the more we will be able to help out our society.

The aim of education is not to develop one single virtue but to produce healthy normal human beings who understand the meaning of life and what it demands from them. From childhood to youth, students must be constantly reminded by educationalists and psychologists that the students are intellectually and morally brought up on co-operation, most of our troubles would be an end and life would be much healthier, richer and nobler than ever before.

C. EDUCATION AS WEALTH:

Education is the most effective weapon against poverty. The more a society is educated, the less the poor it is, since poverty take a strong hold on places where there is an abundance of illiteracy due to lack of education.

Education is the boon to the society. It helps the society to progress in the path of success. Education is the foremost wealth of a nation as it helps in socio-economic conditions, ethnic conditions etc. of a society.

Unless and until a society is not educated it cannot strive and thrive in the path of civilization. And without education in today's world no one can lead a healthy life which would be a helping hand to the nation in every sphere. So, we can say that "Education is the most important wealth of a nation".

D. EDUCATION-INCOME CONNECTION:

In recent years, education is increasingly being viewed as central to the issues of development, like productivity, income distribution, employment and knowledge as an input to production. However, policy making in education is now being informed mainly by mainstream economic theory both in the developing as well as in the developed world.

There is a lot of talk about growing inequality of incomes. There is even greater talk about how the world is brutish and unequal, and how growing inequality in incomes is the scourge of our times. But there is very little serious talk about the equalizing force that is education. Instinctively, education is the most democratic of policy or parental interventions; once it is provided, no one knows where it can lead the recipient to, but, on average, the recipient does well.

Economists love the counter-factuals; if the question is what happens to Y when X exists, they like to derive the answer by reasoning, or imagining, Y without the presence of X. Let us do that with education. If your daughter did not have any education, her life would be doomed; sure, she could live off inherited wealth, but then what would the daughter'sdaughter be able to achieve? One leading economist of the 20th century, Irving Fisher, talked about 'shirt sleeves to shirt sleeves in three generations,'i.e. that inequality did not perpetuate itself and that on an average, one came back to where one started in three generations. In this instance of the progeny having zero education, it would be shirts to shirts in one generation.

Education helps one improve one's life's circumstances and increase possibilities. It is a wealth that yields a flow of income, often an assured flow of income which is why it should be called 'Not Fool's Gold.'

Gary Becker, a Nobel Prize winning economist made proposition in 1950s: that the amount of education had a direct bearing on income.

Both rich and poor parents desire the same amount of child services. Rich people are more likely to send their children's to the best school and colleges and give best of the best education. But the poor could not effort. This is how education is related to income.

E. INTERPRETATION OF GINI CO-EFFICIENT FOR EDUCATION:

Gini co-efficient is a measure of relative inequality with values lying between 0 (complete equality, everyone gets the same education) and 100 per cent (one person gets all the education); the Gini co-efficient may exceed 100 per cent where some people have negative education. The Gini co-efficient is defined as half the mean difference divided by the mean; geometrically, it is the area between the Lorenz curve and the line of equality, divided by the area of the whole triangle.

II. FIELD OF STUDY AND DATA SOURCE:

A. Objectives of the Field Work:

Any scientific study is not perfect without any practical application. Statistics, which is essentially a branch of applied sciences with vast area of practical, cannot be exception. So, we must have practical experience in addition to our theoretical knowledge.

Field work is a way to get coped with experience. It is also a part of curriculum in the Department of Statistics, in the 4th semester of two year M.Sc. course in Statistic

By means of field work, we are able to utilize our theoretical knowledge in practical field such as in industrial field, in agricultural field etc. Thus, field works make our mind research oriented and gives confidence to make any independent study in our future carrier. Our present study is a very modest one undertaken with limited scope, time and resources coupled with inexperience.

In the field work, data suitable for allotted topics are to be collected by students and its analysis is done by using the various statistical methods. By this process the student is trained up to carry out any practical work of statistical nature independently so that he would be able to face statistical problems in the future as statistician.

So, we can say that in undertaking field work the students are highly benefited by applying statistical methods in practical field by analyzing and interpreting the collected data from different field of study viz. gardens, industries, village, medical etc.

B. Field of Study:

The data for the study were collected from two different areas. One is situated in Dibrugarh, Assam and the other in Jorhat, Assam. The area from which data were collected in Dibrugarh is known as Barpathar. It is a village having a panchayat with a population of 750 people.

The area from which data were collected in Jorhat is known as Sangsua tea estate. It is a tea garden having a population 599 of tea tribe.

C. Source of Data:

A There are two types of data: Primary data and Secondary data:

Primary data are those which are collected for the first time by the investigator directly from the field of enquire. This type of data may be used with greater confidence, because the investigator will himself decide upon the coverage of the data and the definitions to be used and as such will have a measure control on the reliability of the data.

The secondary data are those that have already been collected by others (which may be governmental or private agencies) and are in published or in an unpublished form. In making use of secondary data, the investigator has to be particularly careful about the nature of the data, their coverage, definitions on which they are based and their degree of reliability.

The source of data for this study is primary, since, data here are collected directly from the respondent.

Data Collected:

A sample of size 300 has been collected randomly from the village Barpathar, near Dibrugarh University, Dibrugarh, Assam. The data relates to income of the households and education years of their children. Another sample of 300 were collected from a tea garden at random. The data collected relates to income of the households and education years of their children. The income of the household refers to income per month, while schooling years refer to the level of education from class 4 to Ph.D.

E. Objectives of the Study:

To study:

(a) Education Inequality that prevails in the areas under study.

(b) Establishing relationship between schooling years and income of household.

III. METHODOLOGY:

Models and Methods Used

A. MODEL- Linear Regression Model

Regression analysis is a statistical technique for investigation and modeling the relationship between variables. Applications of regression are numerous and occur inβ almost every field, including engineering, the physical and chemical sciences, economics, management, life and biological sciences, and the social sciences. In fact, regression analysis may be the most widely used statistical technique.

The simple linear regression model is-

$$y=\beta_0+\beta_1 x+\epsilon \qquad (3.1.1)$$

where the intercept β_0 and the slope β_1 are unknown constants and ϵ is a random error component. The errors are assumed to have mean zero and unknown variance σ^2. Additionally we usually assume that the errors are uncorrelated. This means that the value of one error does not depend on the value of any other error.

It is convenient to view the regressor x as controlled by the data analyst and measured with negligible error, while the response y is a random variable. That is, there is a probability distribution for y at each possible value for x. The mean of this distributions is

$$E(y/x)=\beta_0+\beta_1 x$$

And the variance is

$$Var(y/x) = Var(\beta_0+\beta_1 x+\varepsilon)=\sigma^2$$

Thus, the mean of y is a linear function of x although the variance of y doesnot depend on the value of x. Further more because the errors are uncorrelated, the responses are also uncorrelated. The parameters β_0 and β_1 are usually called regression coefficient.

For multiple regression the matrix form is given below-

$$Y=X\beta+\mathcal{E}$$

where

y= , X= ,β= , Ɛ=

Method Used

B. Ordinary Least Squares:

The method of least squares is used to estimate β_0 and β_1. That is we will estimate β_0 and β_1 so that the sum of the squares of the differences between the observations yi and the straight line is a minimum. From equation (3.1.1) we may write-

$$y_i=\beta_0+\beta_1 x_i+\mathcal{E}_i, \quad i= 1,2,\dots,n \qquad (3.2.1)$$

Equation (3.1.1) may be viewed as a population regression model while equation (3.2.1) is a simple regression model, written in terms of the n pairs of data (y_i, x_i), $(i=1,2,\dots,n)$. Thus the least squares criterion is-

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i-\beta_0-\beta_1 x_i)^2$$

The least squares estimators of β_0 and β_1, say $\hat{\beta}_0$ and $\hat{\beta}_1$, must satisfy-

$$\left[\frac{\partial S}{\partial \beta_0}\right]_{\hat{\beta}_0,\hat{\beta}_1} = 0 \qquad (3.2.2)$$

$$\left[\frac{\partial S}{\partial \beta_1}\right]_{\hat{\beta}_0,\hat{\beta}_1} = 0 \qquad (3.2.3)$$

Solving equations (3.2.2) and (3.2.3), we get-

$$\hat{\beta}_0 = \bar{y}-\hat{\beta}_1 \bar{x} \qquad (3.2.4)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - \left(\sum_{i=1}^{n} y_i\right)\left(\sum_{i=1}^{n} x_i\right)/n}{\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2/n} \qquad (3.2.5)$$

R2 Statistics- coefficient of determination

The quantity

$$R^2= SS_R/SS_T = 1- SS_{RES}/SS_T \qquad (3.2.6)$$

is called the coefficient of determination. Since SS_T is a measure of the variability in y without considering the effect of the regressor variable x and SS_{RES} is a measure of the variability in y remaining after x has been considered, R2 is often called the proportion of variation explained by the regressor because $0\le SS_{RES}\le SS_T$, it follows $0\le R^2\le 1$. Values of the R2 that are close to 1 imply that most of the variability in y is explained by the regression model.

In general, R2 always increases when a regressor is added to the model, regardless of the value of the contribution of that variable. Therefore, it is difficult to judge whether an increase in R2 is really telling us anything important. A fairer coefficient known as adjusted R2 is developed , which is considered as an alternative to R2 so far as proportion of variation explained by the regressor is concerned. Adjusted R 2 is defined as:

$$\text{Adjusted } R^2= R^2_{Adj} = 1 - \frac{SS_{Res}/(n-p)}{SS_T/(n-1)} \qquad (3.2.7)$$

C. GINI Co-Efficient:

Gini proposed a concentration ratio (ρ) defined as follows:

$$\rho = \frac{\text{AREA BDOB}}{\triangle BOC}$$

$$= \frac{\triangle BOC - \text{AREA BDOCB}}{\triangle BOC}$$

$$= 1- \frac{\text{AREABDOCB}}{\triangle BOC}$$

But $\triangle BOC=1/2\times OC\times BC=1/2$ (OC=BC=1)

$$\rho=1-2\,\text{Area BDOCB}$$

$$=1-2\int_0^1 q\,dp$$

$$=1-2\int_0^1 [1-(1-p)^{1/\delta}]\,dp$$

$$=1-2\left[p+ \frac{(1-p)^{1/\delta+1}}{1/\delta+1}\right]_0^1$$

$$=1-2[1-\delta/(1+\delta)]$$

$$=(\delta-1)/(\delta+1), \qquad \vartheta=\vartheta/(u-1)$$

$$=1/(2\vartheta-1)$$

As δ varies from 1 to ∞ ,ρ varies from 0 to1.

(Variable considered here is education years, not monthly income)

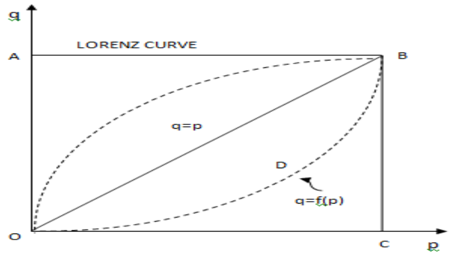

Figure: Lorentz curve(variable considered is education years)

IV. RESULTS AND ANALYSIS:

Results and figures derived from the data collated are outlined below.

Results based regression model used are:

4.1:

MODEL 1:

FOR VILLAGE,

The model fitted for the data is:

(Y)ˆ(Education years)=A+BX (monthly income)

Yˆ(Education years)=8.164+.00055793X

t value:(51.5) (14.068)

p value: <.0001 ; <.0001

F= 197.905*

R^2= .399; Adjusted R2=.397

CORRELATION BETWEEN X AND Y = .632

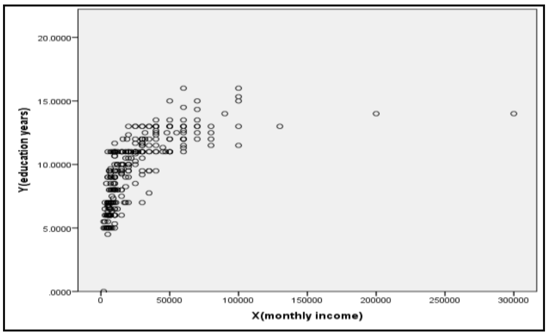

FIGURE 4.1.1: SCATTER DIAGRAM FOR X (Monthly Income) AND Y (Education Years)

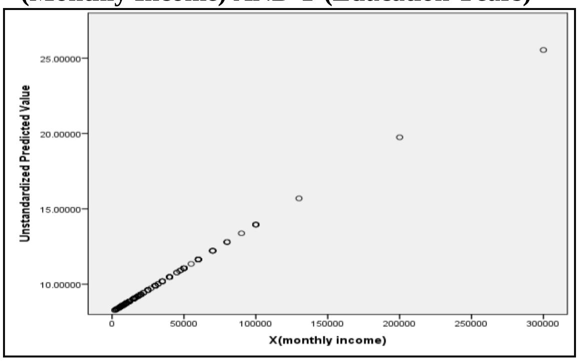

FIGURE 4.1.2.: SCATTER DIAGRAM FOR X (Monthly Income) AND PREDICTED Y (Education years)

4.2:

MODEL 2:

FOR VILLAGE,

The model fitted for the data is:

LOGY (Education years) =A+BX(Monthly Income)

(lnY)ˆ(Education years)=A+BX(Monthly income)

(lnY)ˆ(Education years)=2.078+5.917E-006X

t value:(113.7) (12.486)

p value: (<.0001) (<.0001)

F =155.906*

R2=.344; Adjusted R2=.342

CORRELATION BETWEEN LOGY AND X=.586

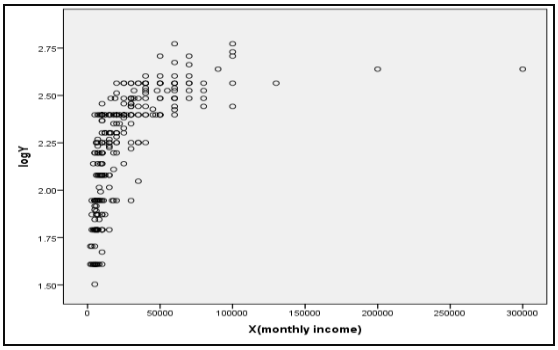

FIGURE 4.2.1.: SCATTER DIAGRAM FOR X (Monthly Income) AND LOGY (Education years)

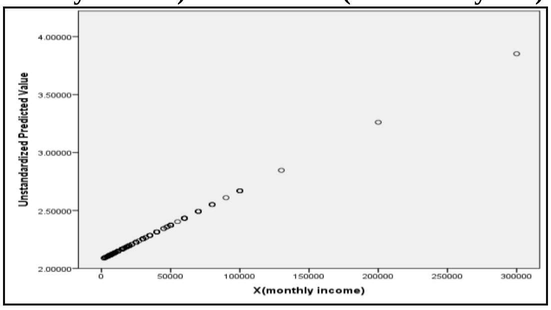

FIGURE 4.2.2.: SCATTER DIAGRAM FOR X (Monthly Income) AND PREDICTED LOGY (Education years)

4.3:

MODEL 3:

FOR TEA GARDEN,

The model fitted for the data is:

Yˆ(Education years)=A+BX (Monthly income)

Yˆ(Education years)=1.138+.001X

t value: (9.140) (30.406)

p value= (<.0001) (<.0001)

R2=.756; Adjusted R2=.755

CORRELATION BETWEEN X AND Y =.870

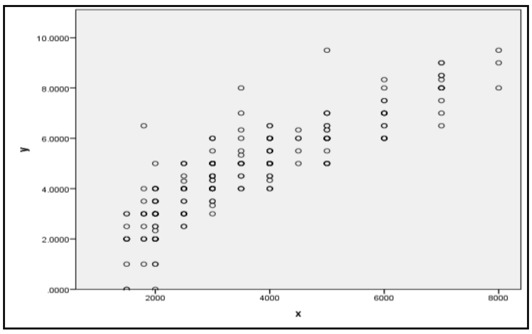

FIGURE 4.3.1. : SCATTER DIAGRAM FOR X (Monthly Income) AND Y (Education years)

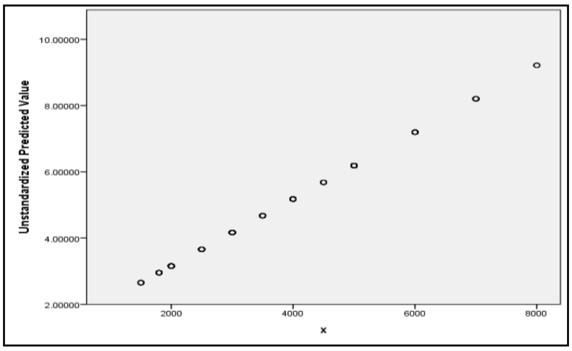

FIGURE 4.3.2.: SCATTER DIAGRAM FOR X (Monthly income) AND PREDICTED Y(Education years)

4.4:
MODEL 4:
FOR COMBINED DATA SET OF BOTH VILLAGE AND TEA GARDEN,
The model fitted for the data is:
ln$\hat{Y}$(Education years)=A+BX(Monthly income)
ln$\hat{Y}$(Education years)=5.742+9.717E-005X
t value: (48.772) (22.567)
p value: (<.0001) (<.0001)
F= 509.253*
R^2= .460; Adjusted R^2=.459
CORRELATION BETWEEN X AND Y= .678

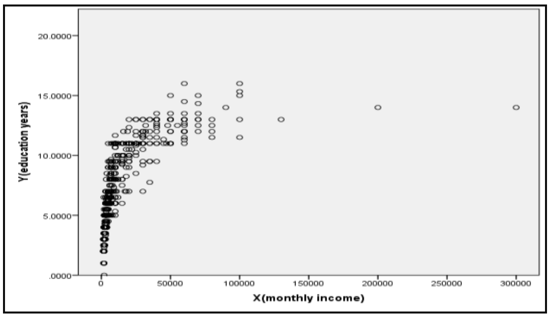

FIGURE 4.4.1.: COMBINED (Village and tea tribe) SCATTER DIAGRAM for X (Monthly income)

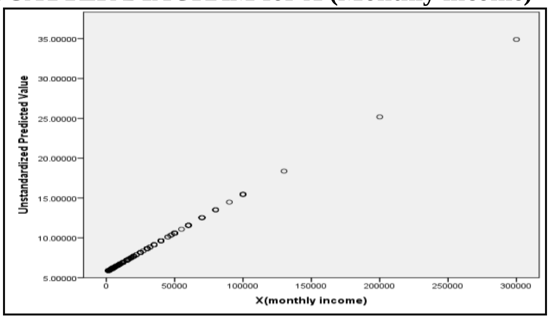

FIGURE 4.4.2.: COMBINED (Village and tea tribe) SCATTER DIAGRAM FOR X (Monthly income) AND PREDICTED Y (Education years)

4.5:
MODEL 5:
FOR VILLAGE,
The model fitted for the data is:
$\hat{Y}$(Education years)=A+BX (Monthly income)+CX2(Monthly income)
$\hat{Y}$(Education years)=7.144+.000X-3.564E-010X2
t value: (44.3) (18.4) (-11.2)
p value: (<.0001) (<.0001) (<.0001)
F=203.914
R^2=.579; Adjusted R^2=.576
CORRELATION BETWEEN X AND Y= .632

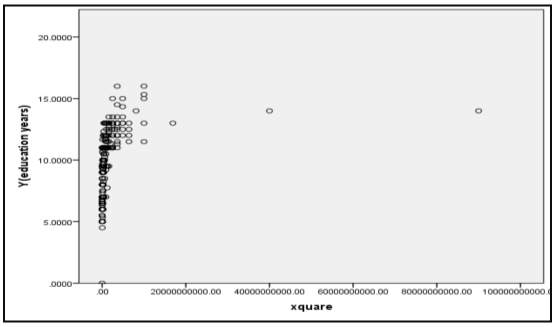

FIGURE 4.5.1.: SCATTER DIAGRAM FOR X2(Monthly income) AND Y (Education years)

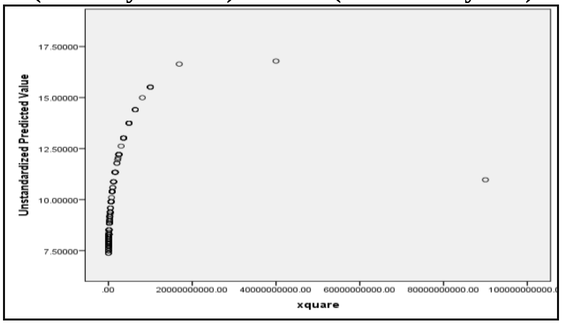

FIGURE 4.5.2.: SCATTER DIAGRAM FOR X2(Monthly income) AND PREDICTED Y (Education years)

4.6.:
MODEL 6:
FOR TEA GARDEN,
The model fitted for the data is:
The model we have fitted is-
$\hat{Y}$(Education years)=A+BX (Monthly income) + CX2 (Monthly income)$\hat{Y}$(Education years)=.282+.002X-6.538E-008X2
t value: (1.010) (10.343) (-3.754)
p value: (<.313) (< .0001) (< .0001)
F=527.551*
R^2=.780; Adjusted R^2=.779
CORRELATION BETWEEN X AND Y=.877:

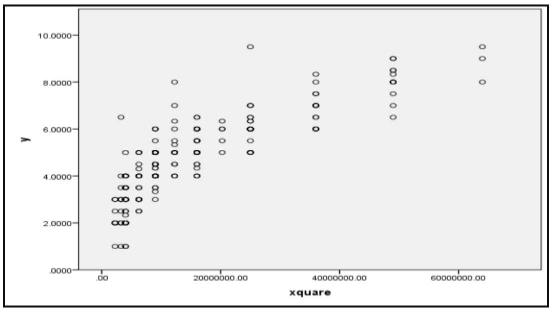

FIGURE 4.6.1.: SCATTER DIAGRAM FOR X2(Monthly income) AND Y(Education years)

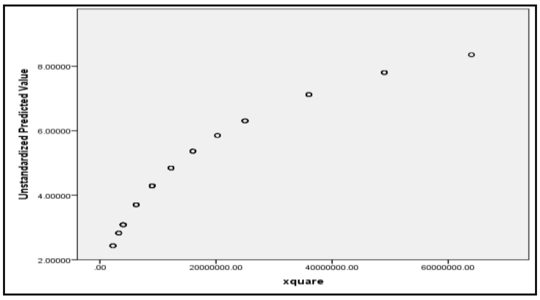

FIGURE 4.6.2.: SCATTER DIAGRAM FOR X2 (Monthly income) AND PREDICTED Y (Education years)

4.7:

MODEL 7:

FOR COMBINED DATA SET OF BOTH VILLAGE AND TEA GARDEN,

The model we have fitted is-

Y῾(Education years)=A+BX (Monthly income)+CX2(Monthly income)

Y῾(Education years)= 4.923+.000X-5.883E-010X2

t value : (46.8) (30.931) (-18.022)

p value: < .0001; <.0001; <.0001

F=554.897*

R2=.650; Adjusted R2=.649

CORRELATION BETWEEN X AND Y=.678

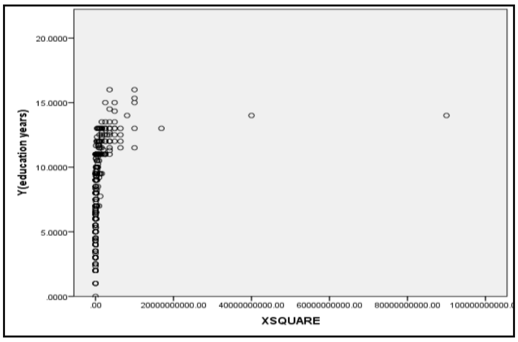

FIGURE 4.7.1.: SCATTERDIAGRAM FOR X2 (Monthly income) AND Y(Education years)

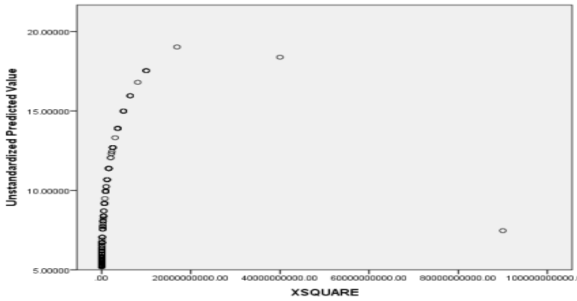

FIGURE 4.7.2.: SCATTER DIAGRAM FOR X2 (Monthly income) AND PREDICTED Y(Education years)

4.8:

MODEL 8:

FOR VILLAGE,

The model we have fitted is-

Y῾(Education years)=A+BX(Monthly income)+CX2(Monthly income)+DX3(Monthly income)

Y῾(Education years)=6.344+.000X-1.544E-009+3.277E-015

t value: (33.869) (15.715) (-9.027) (7.057)

p value: (<.0001) (<.0001) (<.0001) (<.0001)

F=174.877*

R2= .639; Adjusted R2=.63

CORRELATION BETWEEN X AND Y= .632

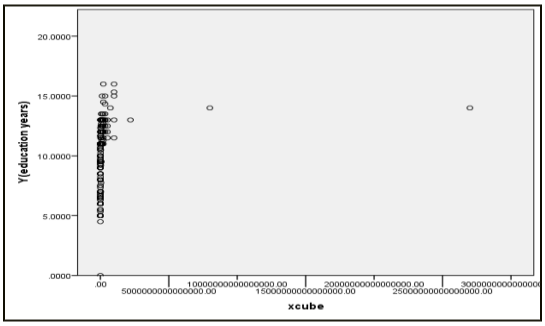

FIGURE 4.8.1.: SCATTER DIAGRAM FOR X3 (Monthly income) AND Y(Education years)

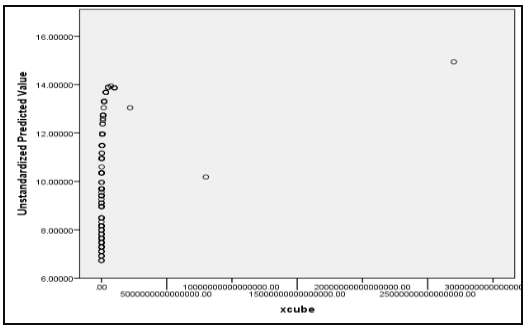

FIGURE 4.8.2: SCATTER DIAGRAM FOR X3 (Monthly income) AND PREDICTED Y(Education years)

4.9:

MODEL 9:

FOR TEA GARDEN,

The model we have fitted is-

Y῾(Education years)=A+BX (Monthly income)+CX2(Monthly income)+DX3(Monthly income)

Y῾(Education years)=-1.696+.003X-4.996E-007

t value: (-2.5) (6.1) (-3.7) (3.2)

p value= (<.011) (<.0001) (<.0001) (<.001)

R2=.788; Adjusted R2=.786

F=366.944*

CORRELATION BETWEEN X AND Y=.87

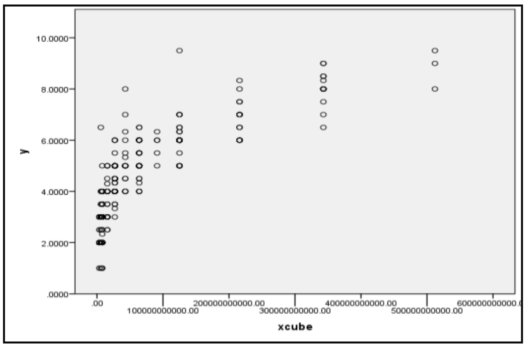

FIGURE 4.9.1.: SCATTER DIAGRAM FOR X3 (Monthly income) AND Y(Education years)

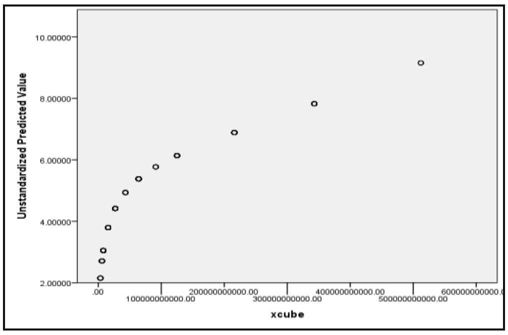

FIGURE 4.9.2.: SCATTER DIAGRAM FOR X3 (Monthly income) AND PREDICTED Y (Education years)

4.10:
MODEL 10:
FOR VILLAGE,
The model fitted for the data is:
Y=AX^ϑ
logY=logA+ϑlogX
(logY)̂=logA +ϑlogX
(logY)̂= -.293+.260logX
t value: (-2.69) (20.31)
p value=(<.007) (<.0001)
R2 = .646 ; Adjusted R2=.645
F=543.125*
CORRELATION BETWEEN logX AND logY= 0.803

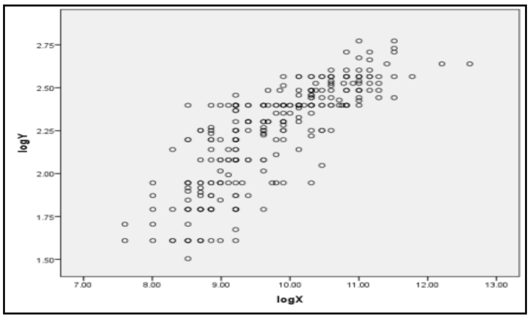

FIGURE 4.10.1.: SCATTER DIAGRAM FOR LOGX(MONTHLY INCOME) AND LOGY(EDUCATION YEARS)

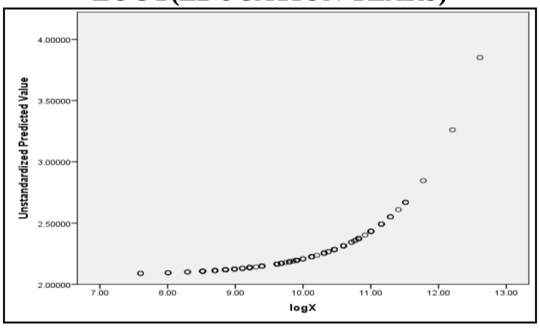

FIGURE 4.10.2.: SCATTER DIAGRAM FOR LOGX(MONTHLY INCOME) AND PREDICTEDLOGY(EDUCATION YEARS)

4.11.:
MODEL 12:
FOR TEA GARDEN,
The model fitted for the data is:
Y=AX^ϑ
logY=logA+ϑlogX
(logY)̂=logA +ϑlogX
(logY)̂= -5.155+.821logX
t value: (-20.8) (26.862)
p value=(<.0001) (<.0001)
R2 = .710; Adjusted R2=.709
F=721.581*
CORRELATION BETWEEN logX AND logY= 0.843

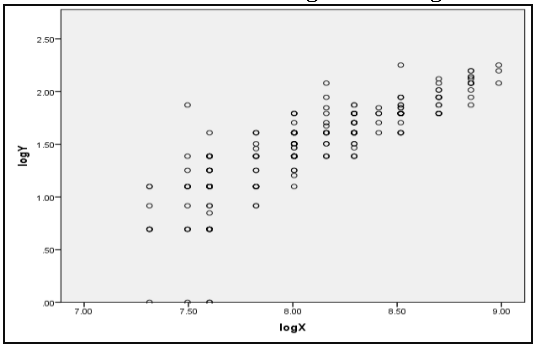

FIGURE 4.11.1.: SCATTER DIAGRAM FOR LOGX(MONTHLY INCOME) AND LOGY(EDUCATION YEARS)

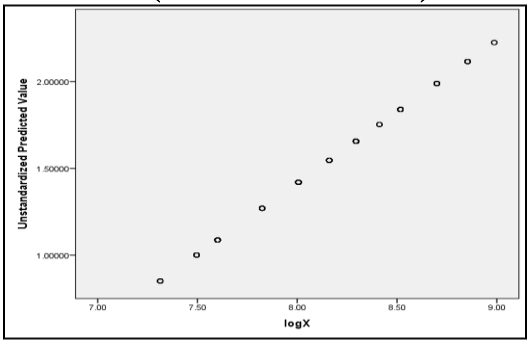

FIGURE 4.11.2.: SCATTER DIAGRAM FOR LOGX(MONTHLY INCOME) AND PREDICTED LOGY(EDUCATION YEARS)

4.13.:
MODEL 12:
FOR COMBINED DATA SET OF BOTH VILLAGE AND TEA GARDEN
The model fitted for the data is:

Y=AX^ϑ
logY=logA+ϑlogX
(logY)̂=logA +ϑlogX
(logY)̂= -1.858+.417logX
t value: (-21.9) (44.004)
p value=(<.0001) (<.0001)
R2 = .764; Adjusted R2=.764
F=721.581*

CORRELATION BETWEEN logX AND logY=.843

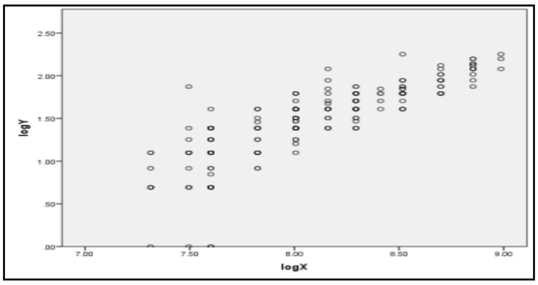

FIGURE 4.12.1.: SCATTER DIAGRAM FOR LOGX(MONTHLY INCOME) AND LOGY(EDUCATION YEARS)

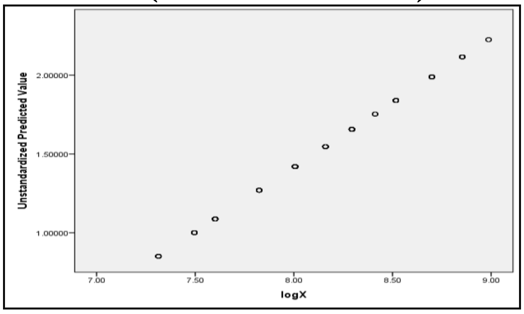

FIGURE 4.12.2.: SCATTER DIAGRAM FOR LOGX (MONTHLY INCOME) AND PREDICTED LOGY (EDUCATION YEARS)

V. ANALYSIS:

What can be concluded from the study go as follows-

FOR VILLAGE,

Upon observing MODEL 1 we see that 39.9% of total variation in explained variable Y (education years) is explained by the explanatory variable X(monthly income of household).

p value suggests that the variable X is significant. Again, F value, which indicates overall significance of the regression model, is highly significant in the model(indicated by F)

Considering the MODEL 2 we observe that 34.4% of the total variation in explained variable i.e. log of education years is explained by the explanatory variable X (monthly income of household).

p value suggests that the variable X is significant. Again, F value, which indicates overall significance of the regression model, is highly significant in the model(indicated by F)

We have seen in MODEL 5 that 57.9% of total variation in explained variable Y is explained by the explanatory variable X

Here, p value suggests the variable X is significant and F value indicates the overall significance of the regression model, and here it is seen that it is highly significant in the model(indicated by F).

It has been observed in MODEL 8 that 63.9% of total variation in explained variable Y is explained by the explanatory variable X.

Here, p value suggests that the variable X is significant and F value indicates the overall significance of the regression model, for this model it is highly significant in the model (indicated by F).

Observing upon the MODEL 11 we have seen that 64.6% of total variation in explained variable Y is explained by the explanatory variable X.

Here, the p value suggests the variable X is significant and F value indicates the overall significance of the regression model, and here we see that it is highly significant in the model.

Among the models used, model 11(power curve) appears to be best one as evidenced by high R2.

VI. FOR TEA GARDEN:

We have seen in MODEL 3 that 75.6% of total variation in explained variable Y is explained by the explanatory variable X

Here, p value suggests the variable X is significant and F value indicates the overall significance of the regression model, and here it is seen that it is highly significant in the model(indicated by F)

Upon observing MODEL 6 we see that 78% of total variation in explained variable Y (education years) is explained by the explanatory variable X(monthly income of household).

p value suggests that the variable X is significant. Again, F value, which indicates overall significance of the regression model, is highly significant in the model (indicated by F)

It has been observed in MODEL 9 that 78.8% of total variation in explained variable Y is explained by the explanatory variable X.

Here, p value suggests that the variable X is significant and F value indicates the overall significance of the regression model, for this model it is highly significant in the model(indicated by F)

Observing upon the MODEL 12 we have seen that 71% of total variation in explained variable Y is explained by the explanatory variable X.

Here, the p value suggest the variable X is significant and F value indicates the overall significance of the regression model, and here we see that it is highly significant in the model (indicated by F).

Among the models used, MODEL 9(viz. cubic curve) appears to be the best one as evidenced by high R2.

FOR COMBINED DATA SET OF VILLAGE AND TEA GARDEN:

By observing MODEL 4 we have seen that 46% of total variation in explained variable logY is explained by the explanatory variable logX.

Here, the p value suggest the variable logX is significant and F value indicates the overall significance of the regression model, and here we see that it is highly significant in the model(indicated by F)

We have seen in MODEL 7 that 65% of total variation in explained variable logY is explained by the explanatory variable logX.

Here, p value suggests the variable logX is significant and F value indicates the overall significance of the regression model, and here it is seen that it is highly significant in the model (indicated by F)

Observing upon the MODEL 10 we have seen that 74.2% of total variation in explained variable logY is explained by the explanatory variable logX.

Here, the p value suggest the variable logX is significant and F value indicates the overall significance of the regression model, and here we see that it is highly significant.

Considering the MODEL 13 we observe that 76.4% of the total variation in explained variable i.e. log of education years is explained by the explanatory variable log X (monthly income of household).

p value suggests that the variable logX is significant. Again, F value, which indicates overall significance of the regression model, is highly significant in model(indicated by F).

Among the models used MODEL 13(viz. power curve) appears to be the best one as evidenced high R2.

VII. ANALYSIS BASED ON GINI INDEX:

Gini indices for education inequality that prevail in the village and tea garden considered are-

G(Village)= .0313

G(Tea garden)= .0062(about zero)

The figures indicate that education inequality decreases both in village and tea garden under study, which is a good indication and bode well for the society and the nation.

VIII. CONCLUSION:

What has been revealed from the study is that there is close relationship between education years and income of the household. This very idea "A DIRECT AND CLOSE RELATIONSHIP BETWEEN EDUCATION YEARS AND INCOME OF THE HOUSEHOLD" came into the fore from the study of Nobel Laureate GARRY BECKER in 1950.

From our study it is seen that the mean years of schooling of the village children is 9.626 which is greater than 7.4 which is the mean years of schooling of children in the World as given in HUMAN DEVELOPMENT REPORT 2010 , which is a good indicator.

For tea garden, the mean years of schooling of children is 4.646 which is less than the mean years of schooling of children in the world. The situation needs to be addressed and improved.

From this we can conclude that the direct relation between education years and income is true. As we have seen in the sample taken from tea garden that the monthly income of the household is in between 1500- 8000, where the monthly income of the village household is between 5000- 1,00,000. As the incomes of the household rises the education years of their children rises, which gives a clear picture from the mean years of schooling of both the groups.

Education attainment A comparative cross-country education attainment is given below-

COUNTRIES	MEAN YEARS OF SCHOOLING 2010
NORWAY	12.6
RUSSIA	8.8
BRAZIL	7.2
CHINA	7.5
SRI LANKA	8.2
SOUTH AFRICA	8.2
INDIA	4.4
PAKISTHAN	4.9
BANGLADESH	4.8
WORLD	7.4

(Courtesy: "EDUCATION AND ECONOMICS" authored by SAUMEN CHATTOPADHYAY)

REFERENCES

1. . Becker, Gary S, and Barry R. Chiswick, (1996): "Education and the Distribution of Earings".The American Economic Review56, no.1/2:358-369.
2. Bhalla.Surajit.S, (2017): "The New Wealth of Nations", Simon & Schuster.
3. Chattopadhyay, Saumen (2012): "Education and Economics".
4. Montgomery D.C., Peck E.A., Vining G.G. (2001): "Introduction to Linear Regression Analysis", John Wiley & Sons, Inc., Third edition.
5. Gupta, S.C and Kapoor, V.K (2007): "Fundamentals of Applied Statistics", Sultan Chand& Sons.

AUTHORS PROFILE

Bharat Bonia was born in Narayanpur, state of Assam on March 30, 1995. He received his Bachelor degree in Economics from North Lakhimpur College, Assam, India. He received a Masters in Economics from Dibrugarh University, Assam, India. His areas of interest and research include Public Policy, Environmental Economics and Development economics. He has more than 1 years of teaching experience.

Abhinanda Bordoloi, Pursuing phd in Gauhati university,Assam, India. She has Passed out M. Sc in 2018 from Dibrugarh University, Assam, India Area of interest in research: Biostatistics and epidemiology, Demography, Econometrics Aim of research : To serve the mankind.

Sarat Chandra Kakaty , was born in Hawajan, state of Assam. He received his B.Sc degree in Statistics from Guwahati University. He has completed M.Sc degree in Statistics from Guwahati University. He has completed Ph.D from Dibrugarh University, Assam, India. He has published articles in reputed journals and conferences in the domain of Statistics. He is presently working at Dibrugarh University, Dibrugarh, Assam, India as Professor. His area of interest is Stochastic processes.