**Practical No: 01**

**OBJECT**: A developer of food for pig wish to determine what relationship exists among ‘age of pig’ when it starts receiving a newly developed food supplement, the initial weight of the pig and the amount of weight it gains in a week period with the food supplement. The following information is the result of study of eight piglets.

|  |  |  |  |
| --- | --- | --- | --- |
| Piglet number | Initial weight(pounds)(x1) | Initial age(weeks)(x2) | Weight gain(y) |
| 1 | 39 | 8 | 7 |
| 2 | 52 | 6 | 6 |
| 3 | 49 | 7 | 8 |
| 4 | 46 | 12 | 10 |
| 5 | 61 | 9 | 9 |
| 6 | 35 | 6 | 5 |
| 7 | 25 | 7 | 3 |
| 8 | 55 | 4 | 4 |

1. Find the regression equation of Y on X1 and X2. And interpret the value of regression coefficient.
2. How much gain in weight of a pig in a week can we expect with the food supplement if it were 9 weeks old and weighed 48 pounds?
3. Calculate Total Sum of Square, Sum of Square due to regression and Sum of Square due to Error.
4. Calculate the standard error of estimate.
5. Calculate the Multiple correlation coefficient.
6. Calculate the Coefficient of multiple determination and Adjusted Coefficient of determination. Also interpret your result.
7. Calculate the predicted and residual values.
8. Test the significance of regression coefficient and overall fit of the regression equation.

**WORKING EXPRESSION:**

Multiple regression equation is algebraic relationship between one dependent variable and two or more independent variable associated with dependent variable.

Let us assume two independent variable x1 and x2 associated with dependent variable y. the multiple regression equation of variable x1, x2 and y is given by:

y = a + b1x1 + b2x2 ----------- (i)

Where symbols have their usual meaning.

Value of a,b1 and b2 are obtained by solving three normal equations obtained by method of least square. The equations are:

y = na + b1x1 + b2x2 ------------ (ii)  
∑

Yx1 = ax1 + b1x12 + b2x1x2  --------- (iii)

Yx2 = ax2 + b1x1x2 + b2x22  ----------(iv)

In regression analysis, total variation is divided into explained variable (SSR) and unexplained variable (SSE) i.e.,

Total variation = Explained variation + Unexplained variation

TSS = SSR + SSE ---------- (v)

TSS = (y –ӯ)2  = y2 – n ӯ2

SSE = (y – ŷ)2 = y2 – a y – b1 x1y – b2 x2y

SSR = TSS -SSE

Standard error of estimate is the square root of variance computed for sample data.

Se =

Where symbol have their usual meaning,

Where, SSE = sum of square due to error

k = number of independent variables in regression model

n = number of observations.

Coefficient of determination is the ratio of sum of square due to regression to the total sum of square. It is denoted by R2 and is given by, R2 =

It is suggested that the adjusted R2 should be used in place of R2 in multiple regression model. Adjusted R2 simply a R2 adjusted by its degree of freedom and reflects both the number of independent variable and sample size used in model. It is given by, Adjusted R2 = 1 - [1-R2]

Practical No: 01

Name: Aakash Shrestha

Roll No.: 02

Subject: Statistics

Date: 2080/04/07

Faculty: BSc. CSIT 3rd Semester

**OUTPUT:**

DATASET NAME DataSet4 WINDOW=FRONT.

REGRESSION

/DESCRIPTIVES MEAN STDDEV CORR SIG N

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA CHANGE

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT Weight\_gain

/METHOD=ENTER Initial\_weight Initial\_age

/SAVE PRED RESID.

**Regression**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Descriptive Statistics** | | | | | | | |
|  | | Mean | | Std. Deviation | | N | |
| weight gain of pig (Y) | | 6.5000 | | 2.44949 | | 8 | |
| Initial weight of pig (X1) | | 45.2500 | | 11.69554 | | 8 | |
| Initial age of pig (X2) | | 7.3750 | | 2.38672 | | 8 | |
| **Correlations** | | | | | | | | |
|  | | | weight gain of pig (Y) | | Initial weight of pig (X1) | | Initial age of pig (X2) | |
| Pearson Correlation | weight gain of pig (Y) | | 1.000 | | .514 | | .794 | |
| Initial weight of pig (X1) | | .514 | | 1.000 | | .017 | |
| Initial age of pig (X2) | | .794 | | .017 | | 1.000 | |
| Sig. (1-tailed) | weight gain of pig (Y) | | . | | .096 | | .009 | |
| Initial weight of pig (X1) | | .096 | | . | | .484 | |
| Initial age of pig (X2) | | .009 | | .484 | | . | |
| N | weight gain of pig (Y) | | 8 | | 8 | | 8 | |
| Initial weight of pig (X1) | | 8 | | 8 | | 8 | |
| Initial age of pig (X2) | | 8 | | 8 | | 8 | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | |
| R Square Change | F Change |
| 1 | .939a | .881 | .834 | .99907 | .881 | 18.539 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Summaryb** | | | |
| Model | Change Statistics | | |
| df1 | df2 | Sig. F Change |
| 1 | 2a | 5 | .005 |

|  |
| --- |
| a. Predictors: (Constant), Initial age of pig (X2), Initial weight of pig (X1) |
| b. Dependent Variable: weight gain of pig (Y) |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 37.009 | 2 | 18.505 | 18.539 | .005b |
| Residual | 4.991 | 5 | .998 |  |  |
| Total | 42.000 | 7 |  |  |  |

|  |
| --- |
| a. Dependent Variable: weight gain of pig (Y) |
| b. Predictors: (Constant), Initial age of pig (X2), Initial weight of pig (X1) |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | -4.192 | 1.888 |  | -2.220 | .077 |
| Initial weight of pig (X1) | .105 | .032 | .501 | 3.247 | .023 |
| Initial age of pig (X2) | .807 | .158 | .786 | 5.097 | .004 |

|  |
| --- |
| a. Dependent Variable: weight gain of pig (Y) |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Residuals Statisticsa** | | | | | |
|  | Minimum | Maximum | Mean | Std. Deviation | N |
| Predicted Value | 4.0747 | 10.3087 | 6.5000 | 2.29936 | 8 |
| Residual | -1.07467 | 1.40931 | .00000 | .84437 | 8 |
| Std. Predicted Value | -1.055 | 1.656 | .000 | 1.000 | 8 |
| Std. Residual | -1.076 | 1.411 | .000 | .845 | 8 |

|  |
| --- |
| a. Dependent Variable: weight gain of pig (Y) |

**RESULTS:**

1. The multiple regression model of y on x1 and x2 is,

Y = -4.192 + 0.105X1 + 0.807X2

Interpretation:

b1 = 0.105 means on average weight(pound) is increased by 0.105 when weight gain is increased by 1 holding the age constant.

b2 = 0.807 means average age is increased by 0.807 when weight gain is increased by 1 holding the weight(pound) constant.

1. When x1 = 48 and x2 = 9

Y = -4.192 + 0.105X1 + 0.807X2

= -4.192 + 0.105\*48 + 0.807\*9

= 8.111

Therefore, expected gain in weight of a pig in a week is 8.111.

1. Total sum of Square (T.S.S) = = 42.000

Sum of square due to error (S.S.E) = = 4.991

Sum of square due to regression (S.S.R) = 37.009

1. Standard error of estimate (Se) = = 0.99907
2. Multiple correlation coefficient (R) = 0.939
3. Coefficient of multiple determination (R²) =0.881

Interpretation: It means 88.1% of total variation in weight gain (y) can be explained by the variation in weight gain and age.

Adjusted R square: 0.834

Interpretation: This indicates this regression equation can represent 83.4% of the variation in the dependent variable (response variable) is explained by the independent variables.

VII. The predicted and residual value are:

|  |  |  |
| --- | --- | --- |
| S. N | Predicted | Residual |
| 1 | 6.34885 | 0.65115 |
| 2 | 6.09869 | -0.09869 |
| 3 | 6.59069 | 1.40931 |
| 4 | 10.30870 | -0.30870 |
| 5 | 9.46171 | -0.46171 |
| 6 | 4.31651 | 0.68349 |
| 7 | 4.07467 | -1.07467 |
| 8 | 4.80019 | -0.80019 |

VIII. As from ANOVA Table we can clearly see the value of Fcal = 18.539 also

Ftab = Fα(k,n-k-1) = F0.05(2,3) = 9.55

**Decision:** Since, Fcal = 16.280 > Ftab = 9.55. So, we reject Ho Which means there is no significant linear relationship between regression equation.

**CONCLUSION:**

Hence, we have obtained the regression model equation and interpret it. We have calculated sum of square due to error, Total sum of square, sum of square due to regression, standard error of estimate. We have calculated and interpreted the value of coefficient of multiple determination and Adjusted R2. We have calculated the predicted value and residual value from variable view and at last we have test the significance of regression equation using p-value.