Pakistan Journal of Nutrition 9 (4): 380-386, 2010
ISSN 1680-5194

Constructing the Best Regression Model for Maiwa Variety

F. Busari Abdullahi, Abubakar Usman and A.T. Cole
Department of Mathematics/Computer Science,
Federal University of Technology, Minna, Niger State, Nigeria

Abstract: As difficult as it can be to determine the plant attribute that contributes most to better yield of cereal crop named Maiwa. We use multivariate regression model to determine the contribution of Plant height (X_1); Number of leaves (X_2); Number of tillers (X_3) and Leaf's area in square feet (X_4). Four multivariate regression models were developed by dropping each attribute. A data set collected from the Institute of Agricultural Research (IAR) Ahmadu Bello University, Samaru-Zaria was used for the analysis. Using each of the models to assess the contribution of each attribute, it was discovered that the Multivariate regression model that has the best fits of the data set, when covariates are dropped one after the other is $Y = 0.02371 - 0.003111X_2 + 0.001759X_3 - 0.002503X_4$. Thus, plant height ($X_1$) is an irrelevant plant attribute for the variety-Maiwa.

Key words: Regressor, predictor, goodness-of-fit, multivariate, yield, tiller

INTRODUCTION

The statistical technique that is used to establish the existence of linear relationship between the dependent variable and the independent variables is the Regression Analysis. If there is a single independent or predicator variable is referred to as simple linear regression, while if it involved more than one independent or predictor variables we have the case of Multivariate regression or multiple regression analysis.

In many crops, especially arable crops, yield depend on some plants attributes such as plant height, number of leaves, stalk thickness, spacing variates etc. These plants attributes are referred to as the independent variables, covariates, predictors, or regressors; while the yield is the corresponding dependent variables or responses. Each of these regressors contributes to the variation in the yield of the variety.

Although the contribution varies from one crop to another, while in some crop it causes high variation in other the variation is insignificant.

According to Ogunremi (1970) the pod number or unit is an important independent variable that determines the yield in pod producing crops. Gilbert Tukers (1976) worked on sunflower and discovered that the number of heads of plant is very significant factor that determines the yield.

In fiber crops such as Kenaf, it is the plant height that determines the yield significant (Baker, 1970), while in cotton it is the number of bolls that determines the yield (Gardner and Tucker, 1976). Generally foods are produced by the leaves through photosynthesis and stored in the plant roots for plant consumption, while water required is transported through the stem from the roots to the leaves for food manufacturing. Therefore, leaves stem and roots are significant plant parts to plant attributes; hence, there is a relationship between these plant attributes and the yield.

If there is no correlation between two or more co varieties there is a possibility of having a good-fit, while if there is multicollinearity or relationship between these is the possibility of sources of variation. We shall in this paper considered the conditions of good fit in relation to the variables or attributes as proposed by Rizzi Laura (2008) and Gerald Keller and Brian Warrack (2003).

MATERIALS AND METHODS

Here we shall among other things distinguish between simple linear regression and multivariate regression model; furthermore discussed the assessing of model.

Regression analysis: A statistical technique that is used to establish the existence of linear relationship between the dependent variable and the independent variables is known as Regression analysis. It is also used to predict the value of one variable; this technique requires developing a Mathematical equation called Model.

According to Gerald Keller and Brian Warrack (2003): In developing a model it is necessary to known the nature of the relationship between dependent or response variables and each of the independent or predictor variables. This could be done by method of either Deterministic or Probabilistic models. The first model is not realistic because it does not assume the randomness of the variables involved or other external factors. The second includes the random component which measures the error of the deterministic component. The random component accounts for measureable and immeasurable variables that are not part of the model.

Corresponding Author: Abubakar Usman, Department of Mathematics/Computer Science, Federal University of Technology, Minna, Niger State, Nigeria