IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 10, Issue 4 Ver. III (Jul-Aug. 2014), PP 57-68
www.iosrjournals.org

On the Estimation and Properties of Logistic Regression Parameters

[1]Anthony Ngunyi, [2]Peter Nyamuhanga Mwita, [2] Romanus O. Odhiambo

Abstract: *Logistic regression is widely used as a popular model for the analysis of binary data with the areas of applications including physical, biomedical and behavioral sciences. In this study, the logistic regression model, as well as the maximum likelihood procedure for the estimation of its parameters, are introduced in detail. The study has been necessited with the fact that authors looked at the simulation studies of the logistic models but did not test sensitivity of the normal plots. The fundamental assumption underlying classical results on the properties of MLE is that the stochastic law which determines the behaviour of the phenomenon investigated is known to lie within a specified parameter family of probability distribution (the model). This study focuses on investigating the asymptotic properties of maximum likelihood estimators for logistic regression models. More precisely, we show that the maximum likelihood estimators converge under conditions of fixed number of predictor variables to the real value of the parameters as the number of observations tends to infinity.We also show that the parameters estimates are normal in distribution by plotting the quantile plots and undertaking the Kolmogorov -Smirnov an the Shapiro-Wilks test for normality,where the result shows that the null hypothesis is to reject at 0.05% and conclude that parameters came from a normal distribution.*

Key Words: *Logistic, Asymptotic, Normality, MRA(Multiple Regression Analysis)*

I. Introduction

Regression analysis is one of the most useful and the most frequently used statistical methods [24, 3]. The aim of the regression methods is to describe the relationship between a response variable and one or more explanatory variables. Among the different regression models, logistic regression plays a particular role. The basic concept, however, is universal. The linear regression model is, under certain conditions, in many circumstances a valuable tool for quantifying the effects of several explanatory variables on one dependent continuous variable. For situations where the dependent variable is qualitative, however, other methods have been developed. One of these is the logistic regression model, which specifically covers the case of a binary (dichotomous) response. [6] discussed an overview of the development of the logistic regression model. He identifies three sources that had a profound impact on the model: applied mathematics, experimental statistics, and economic theory. [?] also provided details of the development on logistic regression in different areas. He states that, "Sir [5] introduced many statisticians to logistic regression through his 1958 article and 1970 book, "The Analysis of Binary Data". However, logistic regression is widely used as a popular model for the analysis of binary data with the areas of applications including physical, biomedical, and behavioral sciences.

In this study, the logistic regression models, as well as the maximum likelihood procedure for the estimation of their parameters, are introduced in detail. Based on real data set, an attempt has been made to illustrate the application of the logistic regression model.

Simulation is used in the study since it involves construction of complicated integrals that do not exists in a closed form that can be evaluated. Simulation methods can be used to evaluate it to within acceptable degrees of approximation by estimating the expectation of the mean of a random sample.

II. Literature Review

The method of maximum likelihood is the estimation method used in the logistic regression models, however, two other methods have been and may still be used for estimating the coefficient . These methods are the least squares and the discriminant function analysis. The linear model approach of analysis of categorical data proposed by Grizzle et al.(1969) used estimaton based on NonLinear Weighted S(NLWS). They demostrated that logistic model can be handled by the method of maximum likelihood using an iterative reweighted least squares algorithm. The discriminant approach to estimation of the coefficients is of historical importance as popularized by [4]. [14] compared the two methods when the model is dichotomous and

[1] Dedan Kimathi University of Technology (*Kenya*)
P.O BOX 657-10100, Nyeri
[2] Jomo Kenyatta University of Agriculture & Technology (*Kenya*)
P.O BOX 62000-00200, Nairobi