

# A Jackknife Approach to Error-Reduction in Nonlinear Regression Estimation

Obiora-Ilouno H. O.<sup>1,\*</sup>, Mbegbu J. I.<sup>2</sup>

<sup>1</sup>Nnamdi Azikiwe University, Awka, Nigeria

<sup>2</sup>Department of Mathematics University of Benin, Benin City Edo State, Nigeria

**Abstract** The problems involving the use of jackknife methods in estimating the parameters of non linear regression models have been identified in this paper. We developed new algorithms for the estimation of nonlinear regression parameters. For estimating these parameters, computer programs were written in R for the implementation of these algorithms. We adopted the Gauss-Newton method based on Taylor's series to approximate the nonlinear regression model with the linear term, and subsequently employ least square method iteratively. In the estimation of the nonlinear regression parameters, the results obtained from numerical problems using the Jackknife based algorithm developed yielded a reduced error sum of squares than the analytic result. As the number of observations deleted in each resampling stage increases, the error sum of squares reduces minimally. This reveals the appropriateness of the new algorithms for the estimation of nonlinear regression parameters and in the reduction of the error terms in nonlinear regression estimation.

**Keywords** Non- linear Regression, Jackknife Algorithm, Delete -d, Gauss -Newton

## 1. Introduction

The Jackknife is a resampling technique use for estimating the bias and standard error of an estimator and provides an approximate confidence interval for the parameter of interest. The principle behind jackknife method lies in systematically recomputing the statistic leaving out one or more observation(s) at a time from the sample set thereby generating n separate samples each of size n-1 or n-d respectively. From this new set of replicates of the statistic, an estimate for bias and the variance of the statistic can be calculated[3]. [5] used linear regression analysis to examine the relationship between the Fish Age (FA) as response variable, Total Length (TL) and the Otolith Length (OL) as predictor variables. They examined dependence of FA on TL and OL using bootstrap and Jackknife algorithm. [5] used bootstrap to study linear regression model of the form

$$Y_i = \beta(x_i) + \varepsilon_i \text{ with error term } \varepsilon_i^* \text{ being independent.}$$

In [5], the bootstrap errors  $\varepsilon_i^*$  are drawn with replacement from the set of estimated residuals. Response values  $Y_i^*$  are constructed as the sum of an initial estimate for  $\beta(x_i)$  and  $\varepsilon_i^*$ . [6] used bootstrap method to investigate the effects of sparsity of data for the binary regression model. He

discovered that the bootstrap method provided robust (accurate) result for the sparse data. [4] considers regression and correlation models, and obtain the bootstrap approximation to the distribution of least squares estimates. [2] provides algorithm for data analysis and bootstrap for the construction of confidence sets and tests in classical models which involves exact or asymptotic distribution. [14] proposes a class of weighted jackknife variance estimators for the ordinary least squares estimator by deleting any fixed number of observations at a time. He observed that the weighted jackknife variance estimators are unbiased for homoscedastic errors. [11] proposes an unbiased ridge estimator using the jackknife procedure of bias reduction. They demonstrated that the jackknife estimator had smaller bias than the generalized ridge estimator (GRE). [1] proposes Modified jackknife ridge regression estimator (MJR) by combining the ideas of GRR and JRR estimators. In their article, they proposed a new estimator named generalized jackknife ridge regression estimator (GJR) by generalizing the MJR. Their result showed that the new proposed estimator (GJR) is superior in the mean square error (MSE) than the generalized ridge regression estimator in regression analysis.

## 2. Materials and Method

Given a model of the form

$$Y = f(X_1, X_2, \dots, X_k, \theta_1, \theta_2, \dots, \theta_j) + \varepsilon \quad (2.1)$$

where the  $\theta$ 's are the parameters,  $X$ 's are the predictor variables and the error term  $\varepsilon \sim N(0, \sigma^2)$  independently

\* Corresponding author:

obiorailouno@yahoo.com (Obiora-Ilouno H. O.)

Published online at <http://journal.sapub.org/ajms>

Copyright © 2013 Scientific & Academic Publishing. All Rights Reserved

identically distributed and are uncorrelated.

$X$ 's, then, we can write

Equation (2.1) is assumed to be intrinsically nonlinear.

Suppose we have a sample of  $n$  observations on the  $Y$  and

$$Y_i = f(X_{i1}, X_{i2}, \dots, X_{ik}, \theta_1, \theta_2, \dots, \theta_j) + \varepsilon_i ; i = 1, 2, \dots, n \quad (2.2)$$

The  $n$ -equation can be written compactly in a matrix notation as

$$Y = f(X, \theta) + \varepsilon \quad (2.3)$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{k1} \\ X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \cdots & \vdots \\ X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_j \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

and  $E(\varepsilon) = 0$

The error sum of squares for the nonlinear model is defined as

$$Q = S(\varepsilon) = \sum_{i=1}^n \{Y_i - f(X_i, \theta)\}^2 \quad (2.4)$$

Let the least square estimates of  $\theta$  by  $\hat{\theta}$ , these estimates minimize the  $S(\varepsilon)$ . The least square estimates of  $\theta$  are obtained by differentiating (2.4) with respect to  $\theta$ , equate to zero and solve for  $\hat{\theta}$ , this results in  $J$  normal equations:

$$\begin{aligned} \frac{\partial Q}{\partial \theta_p} &= \sum_{i=1}^n -2[Y_i - f(X_i, \theta)] \frac{\partial}{\partial \theta_p} (f(X_i, \theta)) \\ &- 2 \left[ \sum_{i=1}^n Y_i \left[ \frac{\partial}{\partial \theta_p} f(X_i, \theta) \right]_{\theta=\hat{\theta}} - \sum_{i=1}^n f(X_i, \hat{\theta}) \left[ \frac{\partial}{\partial \theta_p} f(X_i, \theta) \right]_{\theta=\hat{\theta}} \right] = 0 \\ &\sum_{i=1}^n Y_i \left[ \frac{\partial}{\partial \theta_p} f(X_i, \theta) \right]_{\theta=\hat{\theta}} - \sum_{i=1}^n f(X_i, \hat{\theta}) \left[ \frac{\partial}{\partial \theta_p} f(X_i, \theta) \right]_{\theta=\hat{\theta}} = 0 \\ &\sum_{i=1}^n \{Y_i - f(X_i, \hat{\theta})\} \left[ \frac{\partial}{\partial \theta_p} f(X_i, \theta) \right]_{\theta=\hat{\theta}} = 0 ; i = 1, 2, \dots, n, p = 1, 2, \dots, J \end{aligned} \quad (2.5)$$

In estimating the parameters of nonlinear regression model, we use the Gauss-Newton method based on Taylor's series to approximate equation (2.3). Now, considering the function  $f(X, \theta)$  which is the deterministic component of

$$Y_i = f(X_i, \theta) + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.6)$$

Let  $\theta^0$  be the initial approximate value of  $\theta$ . Adopting Taylor's series expansion of  $f(X_i, \theta)$  about  $\theta^0$ , we have the linear approximation

$$f(X_i, \theta) = f(X_i, \theta^0) + (\theta - \theta^0) \frac{\partial}{\partial \theta} f(X_i, \theta) \Big|_{\theta=\theta^0} \quad (2.7)$$

Substituting expressions (2.7) in (2.6) we obtain

$$Y_i = f(X_i, \theta^0) + \sum_{p=1}^J \left[ \frac{\partial}{\partial \theta_p} f(X_i, \theta) \right]_{\theta=\theta^0} (\theta_p - \theta_p^0) + \varepsilon_i \quad i = 1, 2, \dots, n, p = 1, 2, \dots, J \quad (2.8)$$

Equation (2.8) may be viewed as a linear approximation in a neighborhood of the starting value  $\theta^0$

Let

$$\begin{aligned} f_i^0 &= f(X_i, \theta^0) \\ \beta_p^0 &= \theta_p - \theta_p^0 \end{aligned}$$

$$Z_{pi}^0 = \left[ \frac{\partial}{\partial \theta_p} f(X_i, \theta) \right]_{\theta=\theta^0}, \text{ for } i=1, 2, \dots, n \text{ and } p=1, 2, \dots, J$$

Hence, equation (2.8) becomes

$$Y_i = f_i^0 + \sum_{p=1}^J Z_{pi}^0 \beta_p^0 + \varepsilon_i, i=1, 2, \dots, n \quad (2.9)$$

$$Y_i - f_i^0 = \sum_{p=1}^J Z_{pi}^0 \beta_p^0 + \varepsilon_i, i=1, 2, \dots, n \quad (2.10)$$

In a matrix form, we have

$$\begin{bmatrix} Y_1 - f_1^0 \\ Y_2 - f_2^0 \\ \vdots \\ Y_n - f_n^0 \end{bmatrix} = \begin{bmatrix} Z_{11}^0 & Z_{21}^0 & \cdots & Z_{J1}^0 \\ Z_{12}^0 & Z_{22}^0 & \cdots & Z_{J2}^0 \\ \vdots & \vdots & \cdots & \vdots \\ Z_{1n}^0 & Z_{2n}^0 & \cdots & Z_{Jn}^0 \end{bmatrix} \begin{bmatrix} \beta_1^0 \\ \beta_2^0 \\ \vdots \\ \beta_J^0 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2.11)$$

Compactly, equation (2.11) becomes

$$Y - f^0 = Z^0 \beta^0 + \varepsilon \quad (2.12)$$

$$\text{where } Y - f^0 = [Y_1 - f_1^0, Y_2 - f_2^0 \dots Y_n - f_n^0]', Z^0 = \begin{bmatrix} Z_{11}^0 & \cdots & Z_{J1}^0 \\ Z_{12}^0 & \cdots & Z_{J2}^0 \\ \vdots & \cdots & \vdots \\ Z_{1n}^0 & \cdots & Z_{Jn}^0 \end{bmatrix} \beta^0 = (\beta_1^0 \cdots \beta_J^0)' \varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$$

We obtain the Sum of squares error ( $SS\varepsilon$ )

$$\begin{aligned} SS\varepsilon &= (\varepsilon' \varepsilon) = ((Y - f^0) - Z^0 \beta^0)' ((Y - f^0) - Z^0 \beta^0) \\ &= (Y - f^0)' (Y - f^0) - 2(Y - f^0)' (Z^0 \beta^0) + (Z^0 \beta^0)' (Z^0 \beta^0) \\ \frac{\partial SS\varepsilon}{\partial \beta^0} &= -2(Y - f^0)' Z^0 + 2(Z^0 \hat{\beta}^0)' Z^0 = 0 \end{aligned}$$

$$(Y - f^0)' Z^0 = Z^0 (Z^0 \hat{\beta}^0) \quad (2.13)$$

Hence,

$$\hat{\beta}^0 = (Y - f^0)' Z^0 (Z^{0'} Z^0)^{-1} \quad (2.14)$$

Therefore, the least square estimates of  $\beta^0$  is

$$\hat{\beta}^0 = (Z^{0'} Z^0)^{-1} Z^{0'} (Y - f^0) \quad (2.15)$$

Thus,  $\hat{\beta}^0 = (\hat{\beta}_1^0, \hat{\beta}_2^0, \dots, \hat{\beta}_J^0)'$  minimizes the error sum of squares,

$$S^*(\varepsilon) = \sum_{i=1}^n \left( Y_i - f_i^0 - \sum_{p=1}^J Z_{pi}^0 \hat{\beta}_p^0 \right)^2 \quad (2.16)$$

Now, the estimates of parameters  $\theta_p$  of non-linear regression (2.1) are

$$\theta_p^1 = \hat{\beta}_p^0 + \theta_p^0; p=1, 2, \dots, J \quad (2.17)$$

Iteratively, equation (2.17) reduces to

$$\theta^1 = \hat{\beta}^0 + \theta^0$$

$$\theta^2 = \hat{\beta}^1 + \theta^1$$

$$\vdots \quad \vdots \quad \vdots$$

$$\theta^r = \hat{\beta}^{r-1} + \theta^{r-1}$$

$$\theta^{r+1} = \hat{\beta}^r + \theta^r$$

Thus

$$\theta^{r+1} = \theta^r + (Z^{r'} Z^r)^{-1} Z^{r'} (Y - f^r) \quad (2.18)$$

where  $\hat{\beta}^r = (Z^{r'} Z^r)^{-1} Z^{r'} (Y - f^r)$  are the least squares estimates of  $\beta$  obtained at the

$(r+1)^{th}$  iterations. The iterative process continues until

$$\left| \frac{\hat{\theta}_p^{(r+1)} - \hat{\theta}_p^{(r)}}{\hat{\theta}_p^{(r)}} \right| < \delta$$

where  $\delta = 10^{-5}$  is the error tolerance [12][7]

After each iteration,  $S^*(\varepsilon)$  is evaluated to check if a reduction in its value has actually been achieved. At the end of the  $(r+1)^{th}$  iteration, we have

$$S^*(\varepsilon)^r = \sum_{i=1}^n \left( Y_i - f_i^r - \sum_{j=1}^p Z_{ji}^r \hat{\beta}_p^r \right)^2 \quad (2.19)$$

and iteration is stopped if convergence is achieved. The final estimates of the parameters at the end of the  $(r+1)^{th}$  iteration are:  $\hat{\theta}_1^{(r+1)}, \hat{\theta}_2^{(r+1)}, \dots, \hat{\theta}_j^{(r+1)}$ .

### 2.1. Jackknife Delete-One Algorithm for the Estimation of Non-linear Regression Parameters

Let  $W_i = (Y_i, Z_{ji})'$  vector denotes the values associated with  $i^{th}$  ( $w_1, w_2, \dots, w_n$ ) observation sets. The steps of the delete-one jackknife regression are as follows.

Given randomly drawn sample of size  $n$  from a population and label the elements of the vector  $W_i = (Y_i, Z_{ji})'$  as the vector  $Y_i = (y_1, y_2, \dots, y_n)'$  be the response variables,  $Z_{ji} = (z_{j1}, z_{j2}, \dots, z_{jn})'$  is the matrix of dimension  $n \times k$  for the predictor variables, where  $j = 1, 2, \dots, k$  and  $i = 1, 2, \dots, n$ .

1. Omit first row of the vector  $W_i = (Y_i, Z_{ji})'$  and label remaining  $n-1$  observation sets  $Y_i^{(J)} = (y_2^{(J)}, y_3^{(J)}, \dots, y_n^{(J)})'$  and  $Z_{ji}^{(J)} = (z_{j2}^{(J)}, z_{j3}^{(J)}, \dots, z_{jn}^{(J)})'$  as the first delete-one Jackknife sample ( $W_1^{(J)}$ )

2. Calculate the least square estimates for nonlinear regression coefficient from the first jackknife sample;  $\hat{\beta}^0 = (Z'Z)^{-1} Z'(Y - f)$ .

3. Compute  $\hat{\theta}^1 = \hat{\theta}^0 + \hat{\beta}^0$  using the Gauss-Newton method, the  $\hat{\theta}^1$  value is treated as the initial value in the first approximated linear model.

4. We return to the second step and again compute  $\hat{\beta}'s$ . At each iteration, new  $\hat{\beta}'s$  represent increments that are added to the estimates from the previous iteration according to step 3 and eventually find  $\hat{\theta}^2$ , which is  $\hat{\theta}^2 = \hat{\theta}^1 + \hat{\beta}^1$  up to  $\theta^{r+1} = \theta^r + \hat{\beta}^r$ .

5. Stopping Rule; this iteration process continues until

$\left| \frac{\hat{\theta}_p^{(r+1)} - \hat{\theta}_p^{(r)}}{\hat{\theta}_p^{(r)}} \right| < \delta$ , where  $\delta = 10^{-5}$ , for the values of  $\hat{\theta}_p^{(1)}, \hat{\theta}_p^{(2)}, \dots, \hat{\theta}_p^{(r+1)}$  from the first delete-one Jackknife estimates  $\hat{\theta}^{(J1)}$ .

6. Then, omit second row of the vector  $W_i = (Y_i, Z_{ji})'$  and label remaining  $n-1$  sized observation sets  $Y_i^{(J)} = (y_1^{(J)}, y_3^{(J)}, \dots, y_n^{(J)})'$  and  $Z_{ji}^{(J)} = (z_{j1}^{(J)}, z_{j3}^{(J)}, \dots, z_{jn}^{(J)})'$  as ( $W_2^{(J)}$ ) and repeat steps 2 to 5 above for the estimate of regression coefficients  $\hat{\theta}^{(J2)}$ . Similarly, omit each one of the  $n$  observation sets and estimate the non linear regression coefficients as in the step 2 to 5 above for  $\hat{\theta}^{(Ji)}$  alternately, where  $\hat{\theta}^{(Ji)}$  is Jackknife regression coefficient vector estimated after deleting of  $i^{th}$  observation set from  $W_i$ .

7. Obtain the probability distribution  $F(\hat{\theta}^{(J)})$   $\hat{\theta}^{(J1)}, \hat{\theta}^{(J2)}, \dots, \hat{\theta}^{(Jn)}$  of Jackknife estimates  $\hat{\theta}^{(J1)}, \hat{\theta}^{(J2)}, \dots, \hat{\theta}^{(Jn)}$ .

8. Calculate the jackknife regression coefficient estimate which is the mean of the  $F(\hat{\theta}^{(J)})$  distribution [9] as;

$$\hat{\theta}^{(J)} = \frac{\sum_{r=1}^n \hat{\theta}^{(Ji)}}{n} = \bar{\theta}^{(Ji)} \quad (2.21)$$

### 2.2. Jackknife Delete-d Algorithm for Estimation of Non-Linear Regression

Let  $W_i = (Y_i, Z_{ji})'$  vector denotes the values associated with  $i^{th}$  ( $w_1, w_2, w_3, \dots, w_n$ ) observation sets. Draw a random sample of size  $n$  from the observation set (population) and label the elements of each vector  $W_i = (Y_i, Z_{ji})'$  as the vector  $Y_i = (y_1, y_2, \dots, y_n)'$  be the response variables, and  $Z_{ji} = (z_{j1}, z_{j2}, \dots, z_{jn})'$  be the matrix of dimension  $n \times k$  for the predictor variables, where  $j = 1, 2, \dots, k$  and  $i = 1, 2, \dots, n$ .

Step 1: Divide the sample into "s" independent group of size  $d$ .

Step 2: Omit first  $d$  observation set from full sample at a time and estimate the nonlinear regression parameter  $\hat{\theta}^{(dd_1)}$  from  $(n-d)$  remaining observation set using the least square estimate for the nonlinear regression parameter from the first delete-d sample;  $\hat{\beta}^0 = (Z'Z)^{-1} Z'(Y - f)$ .

Step 3: Compute  $\hat{\theta}^1 = \hat{\theta}^0 + \hat{\beta}^0$  using the Gauss-Newton method, the  $\hat{\theta}^1$  value is assumed as the initial value in the first approximation.

Step 4: Repeat the second step and again compute  $\hat{\beta}'s$ .

At each iteration, new  $\hat{\beta}'s$  represent increments that are added to the estimates  $\hat{\theta}^1$  from the previous iteration according to step 3 and eventually obtain  $\hat{\theta}^2 = \hat{\theta}^1 + \hat{\beta}^1$  up to  $\hat{\theta}^{r+1} = \hat{\theta}^r + \hat{\beta}^r$ ,

and consequently

$$\hat{\theta}^{r+1} = \hat{\theta}^r + (Z^{r'} Z^r)^{-1} Z^{r'} (Y - f^r)$$

Step 5: Stopping Rule; the iteration process continues until

$$\left| \frac{\hat{\theta}_p^{(r+1)} - \hat{\theta}_p^{(r)}}{\hat{\theta}_p^{(r)}} \right| < \delta, \text{ (where } \delta = 10^{-5} \text{ is the tolerance}$$

magnitude)

and the parameters  $\hat{\theta}_p^{(1)}, \hat{\theta}_p^{(2)}, \dots, \hat{\theta}_p^{(r+1)}$  are computed

from  $(n-d)$  delete-d samples  $W^{(dd_1)}$

Step 6: Omit second d observation set from full sample at a time and estimate the nonlinear regression parameters  $\hat{\theta}^{(dd_2)}$  from remaining  $(n-d)$  observation set based on the delete-d sample; and repeat step 3 to step 5 for the second delete-d sample.

Step 7: Alternately omit each d of the n observation set and estimate the parameters as  $\hat{\theta}^{(dd_k)}$  where  $\hat{\theta}^{(dd_k)}$  is the jackknife regression parameter vector estimated after deletion of  $k^{th}$  d observation set from full sample, for  $k = 1, 2, \dots, s$ ; where  $s = \binom{n}{d}$ , and  $1 < d < n-1$ ; where d is an integer.

Step 8: Obtain the probability distribution  $F(\hat{\theta}_p^{(dd)})$  of nonlinear regression parameter estimates  $\hat{\theta}_p^{(dd_1)}, \hat{\theta}_p^{(dd_2)}, \dots, \hat{\theta}_p^{(dd_s)}$ .

Step 9: Calculate the nonlinear regression parameter estimate

$$\bar{\hat{\theta}}_p^{(dd)} = \frac{\sum_{k=1}^s \hat{\theta}_p^{(dd_k)}}{s}, \quad p = 1, 2, \dots, J \quad (2.22)$$

(see[9])

Standard error for the nonlinear regression parameters  $\hat{\theta}_p^{(dd)}$  is

$$Se(\hat{\theta}_p^{(dd)}) = \left\{ \frac{n-d}{\binom{n}{d}} \sum_{k=1}^s \left( \hat{\theta}_p^{(dd_k)} - \bar{\hat{\theta}}_p^{(dd)} \right)^2 \right\}^{\frac{1}{2}},$$

Where

$$\bar{\hat{\theta}}_p^{(dd)} = \frac{\sum_{k=1}^s \hat{\theta}_p^{(dd_k)}}{\binom{n}{d}}, \quad p = 1, 2, \dots, J \quad (2.23)$$

#### The computer program in R for Jackknife

#x is the vector of independent variable

#theta is the vector of parameters of the model

#This function calculates the matrix of partial derivatives

F=function(x,theta)

{

output=matrix(0,ncol=2,nrow=length(x))

for(i in 1:length(x)) output[i,]=c(exp(theta[2]\*x[i]),theta[1]\*x[i]\*exp(theta[2]\*x[i]))

output

}

#This function calculates the regression coefficients using the Gauss-Newton Method

gaussnewton=function(y,x,initial,tol)

{

theta=initial

count=0

eps=y-(theta[1]\*exp(theta[2]\*x))

SS=sum(eps\*\*2)

diff=1

while(tol<diff)

{

S=SS

ff=F(x,theta)

theta=c(theta+solve(t(ff)%\*%ff)%\*%t(ff)%\*%eps)

eps=y-(theta[1]\*exp(theta[2]\*x))

SS=sum(eps\*\*2)

diff=abs(SS-S)

count=count+1

if(count==100) break

pp=c(theta,SS)

#at each iteration

}

pp

}

#This part of the code does the Delete d jackknife

jack=function(data,p,d,initial)

#p is the no of cols in the data.

#For ex., if p=4 then there is 1 dept. var and 3 indept vars

{

```

n=length(data[,1]) #the sample size
z=matrix(0,ncol=p,nrow=n)
u=combn(n,d)
output=matrix(0,ncol=p+1,nrow=ncol(u))
y=data[,1]
x=data[,2:p]
for (i in 1:(ncol(u))) # is the number of iterations
{
dd=c(u[,i])
yn=y[-dd]
xn=x[-dd]
logreg=gaussnewton(yn,xn,initial,tol)
coef=logreg
output[i,]=c(coef) #store the regression coefficients
}
output
}
#Then to run the code use the following
y <- c(data)

```

```

x <- c(data)
data=cbind(y,x)
initial=c(initial)
expo=jack(data,p,d,initial)
#Run the following to view the jackknife results
theta_0=mean(expo[,1])
theta_0
theta_1=mean(expo[,2])
theta_1
SSE=mean(expo[,3])
SSE

```

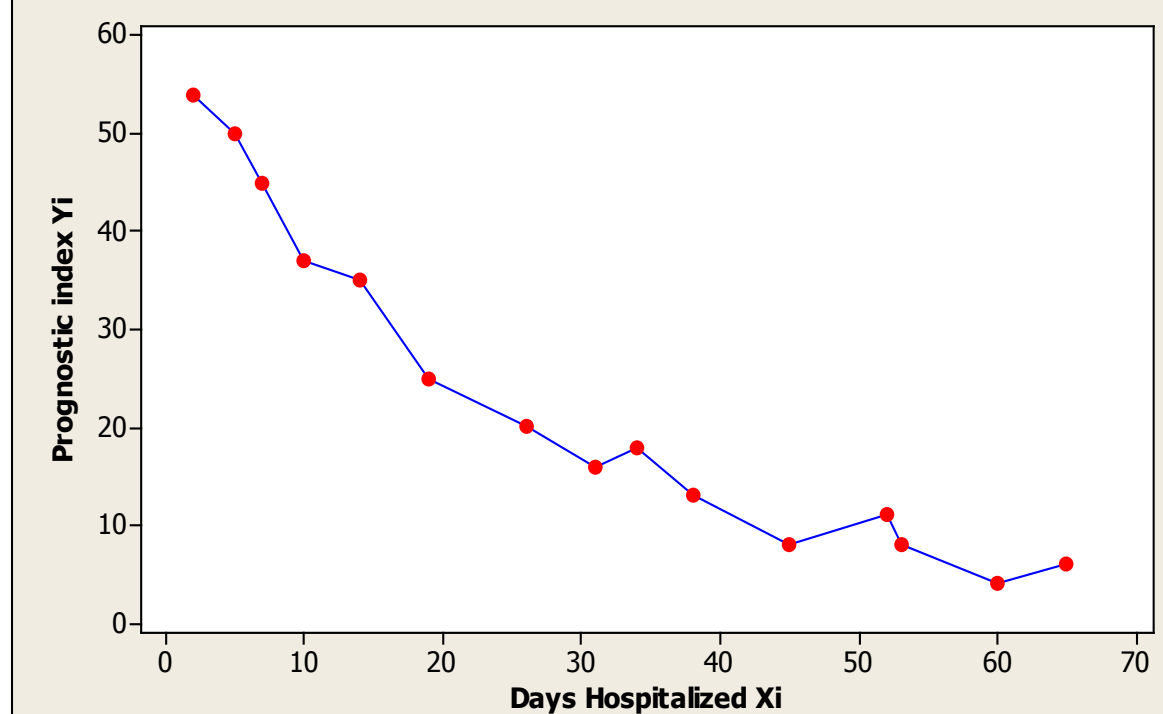
#### Problem:[8]

A hospital administrator wished to develop a regression model for the predicting the degree of long-term recovery after discharge from the hospital for fifteen severely injured patients. The predictor variable to be utilized is number of days of hospitalization (x), and the dependent variable is a prognosis index values. Data collected are shown in the Table below.

**Table 1.** Data for Severely Injured Patients Example[8]

Patient	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	2	5	7	10	14	19	26	31	34	38	45	52	53	60	65
Y	54	50	45	37	35	25	20	16	18	13	8	11	8	4	6

### scatter plot and fitted nonlinear regression function (severely injured patients)



**Figure 1.** ScatterPlot and Fitted Nonlinear Regression Function - Severely Injured Patients Example

## 3. Results and Discussion

Table shows the result for the parameters estimates and the error sum of squares obtained in each iteration.

**Table 2.** Analytical Result for the Estimated Parameters ( $\theta_0$  and  $\theta_1$ ) for data in Table 1

ITERATIONS	$\theta_0$	$\theta_1$	$SS\mathcal{E}$
0	$\theta_0^{(0)} = 56.66512$	$\theta_1^{(0)} = -0.03797$	56.08298
1	$\theta_0^{(1)} = 58.558$	$\theta_1^{(1)} = -0.03953$	49.46378
2	$\theta_0^{(2)} = 58.60549$	$\theta_1^{(2)} = -0.0395$	49.4593
3	$\theta_0^{(3)} = 58.60653$	$\theta_1^{(3)} = -0.03$	49.4593

**Table 3.** Summary results of the Analytical, Jackknife delete-1 and Jackknife delete-d techniques and their Least Squares Criterion Measure

	Analytic	Delete-1	Delete-2	Delete-3	Delete-4	Delete-5
$\theta_1$	58.6065	58.599	58.5892	58.5735	58.5501	58.5147
$\theta_2$	-0.0395	-0.03958	-0.03957	-0.03956	-0.03955	-0.03953
$SS\mathcal{E}$	49.4593	45.7369	42.007	38.2684	34.5174	30.75165

Delete-6	Delete-7	Delete-8	Delete-9	Delete-10	Delete-11
58.46	58.37	58.2415	58.04314	58.7724	57.61277
-0.03949	-0.03944	-0.03936	-0.03924	-0.03924	-0.03888
26.9677	23.16296	19.33635	15.49097	11.63559	7.78429

## Discussion

The least squares criterion measure ( $SS\mathcal{E}$ ) for the starting values has been reduced in the first iteration and also further reduced in the second, third iterations respectively. The third iteration led to no change in either the estimates of the coefficient or the least squares ( $SS\mathcal{E}$ ) criterion measure. Hence, convergence is achieved, and the iterations end. Table 3 shows the results of the analytical and the Jackknifes computation. The fitted regression functions for both analytical and Jackknifes delete -1 computation are:

$$\hat{Y} = 58.6065 \exp(-0.0395X) \quad \text{and}$$

$$\hat{Y} = 58.5996 \exp(-0.0395X) \quad \text{respectively.}$$

The sums of squares error for the analytical and Jackknifes computation are also shown in the table 3. Also, as the number of  $d$  observations deleted in each resampling stage increases, the error sum of squares reduces minimally.

## 4. Conclusions

We have described the Jackknife algorithm in estimation of the parameters of nonlinear regression model

implementation in exponential regression model. The results obtained as shown in Tables 2 and 3 indicate that the Jackknife methods produced a minimum error sum of squares than the analytical method. We also observe that as the number of  $d$  observations deleted in each resampling stage increases, the error sum of squares reduces minimally. Hence, the Jackknife techniques yielded approximately the same inference as the analytical method with a better reduced error sum of squares.

## Appendix

### Delete One Jackknife Result for the Estimates of Parameters $\theta_0$ and $\theta_1$

```

y <- c(54,50,45,37,35,25,20,16,18,13,8,11,8,4,6)
x <- c(2,5,7,10,14,19,26,31,34,38,45,52,53,60,65)
data=cbind(y,x)
initial=c(56.66,-0.03797)
expo=jack(data,2,1,initial)
#Run the following to view the jackknife results
theta_0=mean(expo[,1])
theta_0
[1] 58.59964

```

```
theta_1=mean(expo[,2])
theta_1
[1] -0.03958257
SSE=mean(expo[,3])
SSE
[1] 45.73693
> expo
```

```
      [,1]      [,2]      [,3]
[1,] 58.72515 -0.03967516 49.42362
[2,] 57.69838 -0.03906341 44.59006
[3,] 58.40287 -0.03950187 49.05152
[4,] 59.16565 -0.03961937 42.57768
[5,] 58.45181 -0.03973249 47.48378
[6,] 58.66917 -0.03904419 41.73934
[7,] 58.54728 -0.03931089 48.45668
[8,] 58.48971 -0.03921048 47.87337
[9,] 58.92722 -0.04049873 40.86874
[10,] 58.60389 -0.03957958 49.45880
[11,] 58.36243 -0.03902207 45.56932
[12,] 59.04047 -0.04053060 35.99040
[13,] 58.70595 -0.03980039 48.74806
[14,] 58.44419 -0.03925252 47.21911
[15,] 58.76037 -0.03989685 47.00351
```

#### Deleted (= 5) E Result for the Estimates of Parameters

#### $\theta_0$ and $\theta_1$

```
y <- c(54,50,45,37,35,25,20,16,18,13,8,11,8,4,6)
x <- c(2,5,7,10,14,19,26,31,34,38,45,52,53,60,65)
data=cbind(y,x)
initial=c(56.66,-0.03797)
expo=jack(data,2,5,initial)
#Run the following to view the jackknife results
theta_0=mean(expo[,1])
theta_0
[1] 58.51474
theta_1=mean(expo[,2])
theta_1
[1] -0.03953024
SSE=mean(expo[,3])
SSE
[1] 30.75165
[1,] 58.67728 -0.03959126 27.73554
[2,] 58.17681 -0.03846691 38.39546
[3,] 59.16003 -0.04070913 29.55899
[4,] 59.68595 -0.04223456 20.55436
[5,] 59.23983 -0.04121757 22.01381
[6,] 59.70170 -0.04224374 18.53772
[7,] 58.67701 -0.03992907 35.55294
[8,] 59.17548 -0.04102820 33.88313
[9,] 58.75452 -0.04008201 33.77203
[10,] 59.75422 -0.04241509 20.84062
      :      :      :      :
[2994,] 59.78918 -0.04239372 16.85996
[2995,] 59.39368 -0.04150764 18.60413
[2996,] 58.89694 -0.04038925 32.77863
[2997,] 59.84530 -0.04254249 17.11752
[2998,] 58.80620 -0.04000943 30.68688
```

```
[2999,] 59.26868 -0.04100269 29.02210
[3000,] 58.87835 -0.04014695 28.89213
[3001,] 58.37072 -0.03902467 40.60791
[3002,] 59.35796 -0.04123589 29.70771
[3003,] 59.01265 -0.04042930 27.88792
```

## REFERENCES

- [1] F. SH. Batah, T. V. Ramnathan and S. D. Gore, The Efficiency of Modified Jackknife and Ridge Type Regression Estimators: A comparison, Surveys in Mathematics and its applications 24 No.2 (2008), 157-174.
- [2] Beran, R., The Impact of the Bootstrap on Statistical Algorithms and Theory, Statist. Science, Mathematical Reviews (Maths Sci. Net) Vol. 18, No. 2, pp. 175-184, 2003.
- [3] Efron, B. and Tibshirani, R.J., An Introduction to the Bootstrap, Library of Congress Cataloging in Publication Data, CRC Press, Boca Raton, Florida, 60-82, 1998.
- [4] Freedman, D.A., Bootstrapping Regression Models, Annals of Statistical Statistics, Vol. 9, 2, pp. 158-167, 1981.
- [5] Hardle, W. and Mammen, E., Comparing Nonparametric versus Parametric Regression Fits. Annals of Statistics, Vol. 21, pp. 1926-1947, 1993.
- [6] Mohammed, Z. R., Bootstrapping: A Nonparametric Approach to Identify the Effect of Sparsity of Data in the Binary Regression Models, Journal of Applied Sciences, Network for Scientific Information, Asian, Vol. 8, No. 17, pp. 2991-2997, 2008.
- [7] Nduka, E. C., Principles of Applied Statistics 1, Regression and Correlation Analyses, Crystal Publishers, Imo State, Nigeria, 1999.
- [8] Neter, J., Kutner, M.H., Nachtsheim, C.J. and Wasserman, W., Fourth Edition, Mc-Graw Hill, U.S.A, pp 429-431, 531-536, 1996.
- [9] Obiora-Ilouno, H.O. and Mbegbu, J.I, Jackknife Algorithm for the Estimation of Logistic Regression Parameters, Journal of Natural Sciences Research, IISTE UK, Vol.2, No.4, 74-82, 2012.
- [10] Sahinler, S. and Topuz, D., Bootstrap and Jackknife Resampling Algorithm for Estimation of Regression Parameters, Journal of Applied Quantitative Method, Vol. 2, No. 2, pp.188-199, 2007.
- [11] Singh, B, Chuubey, Y. P. and Dwivedi, T.D., An Almost Unbiased Ridge Estimator. Sankhya series, 48, 342-36, 1986.
- [12] Smith, H. and Draper, N. R., Applied Regression Analysis, Third Edition, Wiley- Interscience, New York, 1998.
- [13] Venables, W. N and Smith, An Introduction to R, A Programming Environment for Data Analysis and Graphics, Version 2.6.1, pp. 1-100, 2007.
- [14] WU, C.F.C., Jackknife, Bootstrap and other Resampling Methods in Regression Analysis. The Annals of Statistics Vol. 14, No.4, 1261 -1295, 1986.