ISSN NO. 2320-5407

Journal Homepage: -www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI: 10.21474/IJAR01/7851
DOI URL: http://dx.doi.org/10.21474/IJAR01/7851

RESEARCH ARTICLE

COMPARISON BETWEEN NORMAL METHODS AND WITH GENETIC ALGORITHM IMPROVED METHODS IN THE ESTIMATION OF THE BINARY LOGISTIC REGRESSION MODEL PARAMETERS.

Sarah adel mazloum al-ridini and prof.dr. Rabab abdulrida saleh al-bakri.
Department ofStatistics ,University ofBaghdad, College ofAdministration and Economics, Baghdad, Iraq.

Manuscript Info

Manuscript History
Received: 09 August 2018
Final Accepted: 11 September 2018
Published: October 2018

Keywords:-
Binary Logistic Regression Model, Genetic Algorithm , Maximum Likelihood,Minimum Chi-Square,Weighted Least Squares,Bayes.

Abstract

In this research suggested, one of the most important models of nonlinear regression models extensive use in the modeling of applications statistical, in terms of heart disease which is the binary logistic regression model.

Some standard methods have been proposed and employed after modifying them by using the genetic algorithm approach in estimation to suit the estimation of the parameters of this of nonlinear regression models, and then making a comparison between two types of the important estimation methods including the standard estimation methods which included the maximum likelihood method, minimum chi-square method, weighted least squares, bayes method, and improved estimation methods developed which by the researcher which included genetic algorithm method depending on the technique estimates MLE, genetic algorithm method depending on the technique estimates MCSE, genetic algorithm method depending on the technique estimates WLSE, genetic algorithm method depending on the technique estimates BE, to choose the best method of estimation by assuming a number of models during simulation and by using the statistical criteria Mean Squares Error (MSE) for estimators for the purpose of comparing the preference of model parameters estimation methods.

Generally, improved estimate methods excellence on the normal methods in estimating parameters, as well the (Wls)method is found to be the best one in the first place one among the standard estimation methods, and (Mcs.GA)method is the best among the important estimation methods for the purpose of estimating the parameters for binary logistic regression model because it has less (MSE) for estimators compared to other methods.

Introduction:-

Is a binary logistic regression model of nonlinear models to describe the relationship between the dependent variable binary value which takes two values are zero for the probability of occurrence of a particular event correct one for the probability of that event independent variables and the occurrence take descriptive or quantitative values. Only then do the nonlinear relationships become linear relationships between the dependent variable and the ... meters Is

Corresponding Author:-Sarah adelmazloum al-ridini.
Address:- Department ofStatistics ,University ofBaghdad, College ofAdministration and Economics, Baghdad, Iraq.

methods (MLE, MCSE, WLSE, BE). In this methods, it is aimed to find the optimum parameter values (the best result estimation) that maximize the likelihood of observing of an individual. To do this, first order derivatives are taken on the model parameters. Because the model parameters are not in a linear form after the derivation step, we could not obtain optimum parameter values, analytically. Therefore, we have used Newton-Raphson (NR) algorithm as a method to solve the problem of the nonlinearand we used method to minimize the chi- square to find the best estimate of the model by reducing the sum squares error random to extract the first derivative of the function as possible and equal to zero also a relationship is not linear as in the method (MLE) Newton Raphson algorithm[5] and we use to solve them as well as method we used the weighted least squares that do not need to initial values shall be written and we use to solve them as well as also method we used bayes in the estimation of parameters of using gibbs sampler which look at the parameters of the logistic model as random variables and because of Newton Raphson algorithm enjoyed in methods(MCS, ML) Of the constraints and assumptions ,If a good starting point could not be determined for each parameter, the solution may get stuck on local optimums and only regional optimums could be found[14] and method (WLS) be inefficient by not having less variation and also fail method (B) in finding the best estimate of the parameter, there has been a need for developing an alternative method that can solve this problems. Thus, due to the restrictions of the iterative methods, We use the genetic algorithm (GA) approach an improved method to resolve to contain has some important properties such as not requiring a differentiable objective function and other supporting information. GA is also able to find global optimum points without being stuck with local optimums [1], successfully, form the best parameter set that would maximize the log-likelihood function and minimizes the error to as little as possible. It is obviously difficult to estimate the parameters of logistic regression methods normal model when the number of parameters $(P + 1)$ as it emerged the need or the development of her[5]. In this study , we gave some information about the theory of the LR model to propose and employ the normal methods(Mle, Mcse, Wlse, Be) after being improved by following the GA approaches to estimate binary logistic regression model,used for that program MATLAB and Monte Carlo simulation with Subsequently, we held discussion on whether GA could be used instead of the traditional normal methods. and after comparison we get recommendations from our findings from the methods used in the estimation .

Binary Logistic Regression Model(BLRM)

Binary logistic regression used to estimate the influence of some explanatory variable (X) to the response variable (Y) that is binary or dichotomous. Said to be binary or dichotomous because it has two possible values that is 0 and 1. Form a binary logistic regression equation used in this study is [3]:

$$\pi(X_i) = \frac{e^{\underline{X_i'\beta}}}{1 + e^{\underline{X_i'\beta}}} \ldots (1)$$

If the equation (1) is transformed by the natural logarithm, it will obtain the linear regression equation as follows[2]:

$$Z = ln\left(\frac{\pi(X_i)}{1 - \pi(X_i)}\right) = \underline{X_i'\beta} \quad , \quad i = 1,2, \ldots, n \ldots (2)$$

Binary Logistic Regression Model Parameters Estimation

Maximum Likelihood Method (MLM)

In binary logistic regression model, one of which is distributed according to the Bernoulli distribution, which has two levels, zero and one according to equation (3) as follows[11]:

$$P(Y_i/X_i) = [\pi(X_i)]^{Y_i}[1 - \pi(X_i)]^{1-Y_i} \ldots (3)$$

In order to obtain parameter estimates by Maximum Likelihood Method , the limits in equation (3) are multiplied by the sample of their size (N) based on (X) of the explanatory variables and the variable of the Y response according to the following mathematical formula[1]:

$$P(Y/X) = \prod_{i=1}^{N}[\pi(X_i)]^{Y_i}[1 - \pi(X_i)]^{1-Y_i} \ldots (4)$$

Since $\pi(X_i)$and $P(Y/X)$ are upon the parameters and our goal is to estimate the parameters unknown, we can determine theby Maximum Likelihood function$l(\beta)$ to detect dependence.

$l(\beta) = P(Y/X) \ldots (5)$, $l(\beta)$:Represent probability of Y observed .

therefore the Maximum Likelihood function in binary logistic regression is [8]:

$$l(\beta) = \prod_{i=1}^{N}[\pi(X_i)]^{Y_i}[1-\pi(X_i)]^{1-Y_i} \dots (6)$$

And by taking the Log on both sides to get the MLE to facilitate the solution

$$ln[l(\beta)] = \sum_{i=1}^{N} Y_i ln[\pi(X_i)] + (1-Y_i)ln[1-\pi(X_i)] \dots (7)$$

And to compensate for $\pi(X_i)$ and $1-\pi(X_i)$ to get this equation[4]:

$$ln[l(\beta)] = \sum_{i=1}^{N}\left[Y_i ln\left(\frac{e^{\acute{X}_i\underline{\beta}}}{1+e^{\acute{X}_i\underline{\beta}}}\right) + (1-Y_i)ln\left(1-\frac{e^{\acute{X}_i\underline{\beta}}}{1+e^{\acute{X}_i\underline{\beta}}}\right)\right] \dots (8)$$

$$\acute{L}\left(\underline{\beta}\right) = \frac{\partial \acute{ln}[l(\beta)]}{\partial \underline{\beta}} = \sum_{i=1}^{N}[Y_i\acute{X}_{ij} - \pi(X_i)\acute{X}_{ij}] = 0 \dots (9)$$

$$\acute{\acute{L}}\left(\underline{\beta}\right) = -\sum_{i=1}^{N}\acute{X}_{ij}\left[\pi(X_i)\left(1-\pi(X_i)\right)\right]X_{ij} \dots (10)$$

Therefore, the Newton Raphson algorithm to find the estimated $(\hat{\beta})$ values of Maximum Likelihood function $l(\beta)$in the logistic model will be in (m + 1) of the frequencies as follows[3]:

$$\hat{\beta}^{(m+1)} = \acute{\hat{\beta}}^{(m)} + (\acute{X}V_m X)^{-1}\acute{X}(Y-P_m) \dots (11)$$

Where Y response variable ,P_m Represents the probability values for the occurrence of the response variable for the frequency m , X represent matrix of explanatory variables ,V_m represent variances matrix $[\pi_i(1-\pi_i)]$.

Minimum Chi - Square Method (MCSM)

It is one of the common methods used in the estimation and is based on the reduction of the chi square statistic of Pearson's known according to the following mathematical formula[9]:

$$x^2 = R(\beta) = \sum_{i=1}^{N}\frac{(\theta_i-\varepsilon_i)^2}{\varepsilon_i} \dots (12)$$

θ_i: represent observer value at i level , ε_i: represent expected value at i level , :$R\left(\underline{\beta}\right)$ Pearson's statistic .

Then in binary logistic regression model[8] :

$$R\left(\underline{\beta}\right) = \sum_{i=1}^{N}\frac{(Y_i-\pi_i)^2}{\pi_i} + \frac{[(1-Y_i)-(1-\pi_i)^2]}{1-\pi_i} \dots (13)$$

The equation (14) can be reduce to :

$$R\left(\underline{\beta}\right) = \sum_{i=1}^{N}\frac{(Y_i-\pi_i)^2}{\pi_i(1-\pi_i)} \dots (14)$$

After compensating the value $\pi_i, 1-\pi_i$in the equation above we get :

$$R\left(\underline{\beta}\right) = \sum_{i=1}^{N}\left[\left(Y_i-\frac{e^{\acute{X}_i\underline{\beta}}}{1+e^{\acute{X}_i\underline{\beta}}}\right)^2\frac{\left(1+e^{\acute{X}_i\underline{\beta}}\right)^2}{e^{\acute{X}_i\underline{\beta}}}\right] \dots (15)$$

And by simplified equation (16) we get to :

$$R\left(\underline{\beta}\right) = \sum_{i=1}^{N}\left[{Y_i}^2 e^{\acute{X}_i\underline{\beta}} + (1-Y_i)^2 e^{\acute{X}_i\underline{\beta}} - 2Y_i(1-Y_i)\right] \dots (16)$$

And then find the first and second derivatives and then equal to zero we get the following equations:

$$\frac{\partial R\left(\underline{\beta}\right)}{\partial \beta_j} = \sum_{i=1}^{N} \acute{X}_{ij} \left[(1-Y_i)^2\left(\frac{\pi_i}{1-\pi_i}\right) - (Y_i)^2\left(\frac{1-\pi_i}{\pi_i}\right)\right] \dots (17)$$

$$\frac{\partial^2 R\left(\underline{\beta}\right)}{\partial \beta_j \acute{\beta}_j} = \sum_{i=1}^{N} X_{ij}\, \acute{X}_{ij} \left[(1-Y_i)^2\left(\frac{\pi_i}{1-\pi_i}\right) + (Y_i)^2\left(\frac{1-\pi_i}{\pi_i}\right)\right] \dots (18)$$

Then we can get to $\hat{\beta}$ by using newton raphson method for iterative m as following [10]:

$$\hat{\beta}^{(m+1)} = \hat{\beta}^{(m)} - \left[\acute{L}(\hat{\beta})^{(m)}\right]^{-1}\left[\acute{L}(\hat{\beta})^{(m)}\right] \dots (19)$$

Where $\acute{L}(\hat{\beta})^{(m)} = \frac{\partial R(\underline{\beta})}{\partial \beta_j}$; $\acute{L}\left(\underline{\hat{\beta}}\right)^{(m)} = \frac{\partial^2 R(\underline{\beta})}{\partial \beta_j \acute{\beta}_j}$

Weighted Least Square Method (WLSM)

Weighted Least Square method is a commonly used method, in which the parameters of the binary logistic regression model are estimated as following:

$$W_i = \pi_i(1-\pi_i) \dots (20)$$

WhereW_i is the variance matrix of the weights selected for level in parameters are estimated are $(\beta_\circ, \beta_1, \dots, \beta_p)$ by using WLS method upon find the values that make the difference between the observed response and the estimated response, it is possible to minimize the sum of the error [8] :

$$sse_i = \sum w_i\left(Z_i - \hat{Z}_i\right)^2 \dots (21)$$

$Z_i$represent of the linear relationship resulting from taking the natural logarithm in the equation (2).

$$Z_i = ln\frac{\pi_i}{1-\pi_i}$$

$$sse_i = \sum w_i\left(Z_i - \hat{\beta}_\circ - \hat{\beta}_1 x_{1i} - \dots - \hat{\beta}_p x_{pi}\right)^2 \dots (22)$$

Using partial differentials for $(\beta_\circ, \beta_1, \dots, \beta_p)$ and output equivalence to zero, we obtain the following:

$$\beta_\circ \sum w_i + \beta_1 \sum w_i x_{1i} + \dots + \beta_p \sum w_i x_{pi} = \sum w_i Z_i \dots (23)$$

$$\beta_\circ \sum w_i x_{1i} + \beta_1 \sum w_i x^2{}_{1i} + \dots + \beta_p \sum w_i x_{1i} x_{pn} = \sum w_i x_{i1} Z_i \dots (24)$$

$$\vdots$$

$$\beta_\circ \sum w_i x_{pi} + \beta_1 \sum w_i x_{pi} x_{1i} + \dots + \beta_p \sum w_i x^2{}_{pi} = \sum w_i x_{ip} Z_n \dots (25)$$

Using the matrix method, these equations are written as follows:

$$\underline{Z} = \begin{bmatrix} ln\frac{\pi_1}{1-\pi_1} \\ ln\frac{\pi_2}{1-\pi_2} \\ \vdots \\ ln\frac{\pi_n}{1-\pi_n} \end{bmatrix} = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix}, W = \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & w_n \end{bmatrix}$$

Where

$$\acute{X} W \underline{\beta} = \acute{X} W \underline{Z} \dots (31)$$

$$\acute{X}WX = \begin{bmatrix} \sum w_i & \sum w_i x_{i1} & \sum w_i x_{i2} & \dots & \sum w_i x_{ip} \\ \sum w_i x_{i1} & \sum w_i x^2{}_{i1} & \sum w_i x_{i1} x_{i2} & \dots & \sum w_i x_{i1} x_{np} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \sum w_i x_{ip} & \sum w_i x_{ip} x_{i1} & \sum w_i x_{ip} x_{i2} & \dots & \sum w_i x^2{}_{ip} \end{bmatrix}, \acute{X}WZ = \begin{bmatrix} \sum w_i z_i \\ \sum w_i x_{i1} z_i \\ \vdots \\ \sum w_i x_{ip} z_n \end{bmatrix}$$

Where Z represent logit of order $(n \times 1)$, W represent weighted matrix of order $(n \times n)$ and the final formula for parameter estimation is [13] :

$$\underline{\hat{\beta}} = (X'WX)^{-1} X'WZ \qquad \dots (26)$$

Bayes Method (BM)

The Bayes technique consists of three functions Prior probability density function , Likelihood function and Posterior probability density function According to technique, Posterior probability density function of binary logistic regression model parameters can be written according to the following mathematical formula [7]:

$$\pi(\beta \backslash Y, Z) = f(Y \backslash \beta, Z) f(\beta \backslash Z) \dots (27)$$

$$\pi(\beta \backslash Y, Z) = l(\beta)\pi(\beta) \dots (28)$$

Where $\pi(\beta \backslash Y, Z)$ is Posterior probability density function for parameters.

$l(B) = f(Y \backslash \beta, Z)$represent Likelihood function and $\pi(\beta) = f(\beta \backslash Z)$ is Prior probability density function for parameters .

And the condition distribution of all parameters is uniform distribution As in the following mathematical formula [6]:

$$\beta_j \beta_{j-p}, Y, Z \sim U(a_j, b_j) \qquad ; j = 0,1, \dots, P \dots (29)$$

Where $U(a_j, b_j)$ is uniform distribution for interval (a , b)

$$a_j = \max\left[\frac{1}{X_{ip}} \log\left(\frac{Z_i}{1 - Z_i}\right) - \sum_{j \neq p}^{N} BX_{ij}\right] \dots (30)$$

$$b_j = \min\left[\frac{1}{X_{ip}} \log\left(\frac{Z_i}{1 - Z_i}\right) - \sum_{j \neq p}^{N} BX_{ij}\right] \dots (31)$$

And we apply Gibbs sampling for conditional distribution for $Z_i$and by iterative can be generate random number form uniform distribution as following [6]:

$$Z_i \backslash \beta, Y \sim U \left\{ \begin{array}{ll} \left(0, \dfrac{e^{\underline{X}'_i \underline{\beta}}}{1 + e^{\underline{X}'_i \underline{\beta}}}\right) & if Y_i = 1 \\ \left(\dfrac{e^{\underline{X}'_i \underline{\beta}}}{1 + e^{\underline{X}'_i \underline{\beta}}}, 0\right) & if Y_i = 0 \end{array} \right\} ; i = 1,2, \dots, N \dots (32)$$

$\beta^{(i)}$ is generated by using the initial values $Z^{(i)}$ in equation (29) to complete one frequency and then in equation(32) to generate ($Z^{(1)}$) and after (n) of the frequencies the conditional distributions will be updated to generate $Z^{(n)}$,$\beta^{(n)}$ until we reach the stability stage where thesample distribution of the sample is approaching the exponential rate of the original post-joint distribution $P(\beta \backslash Z, Y)$ the number of times the new samples are generated infinitely $(n \to \infty)$ [15]. We obtain the estimated (β) values by taking the extracted sample (generated) from the subsequent distribution according to the following mathematical formula [2]:

$$\overset{(i)}{\beta} \sim \hat{\pi}_i(\beta \backslash Y) = \frac{1}{N} \sum_{j=1}^{N} P(\beta / Y, \overset{(i-1)}{Z_j}) \dots (33)$$

Genetic Algorithms (GA)

genetic algorithms (GA) is based on Darwinian's theory of survival of the fittest. Genetic algorithms (GA) may contain a chromosome, a gene, set of population, fitness, fitness function, breeding, mutation and selection. Genetic algorithms (GA) begin with a set of solutions represented by chromosomes, called population [1]. Solutions from one population are taken and used to form a new population, which is motivated by the possibility that the new population will be better than the old one. Further, solutions are selected according to their fitness to form new solutions, that is, offsprings. The above process is repeated unti some condition is satisfied. Algorithmically, the basic genetic algorithm (GA) is outlined as below [12][16]:

Step I [Start] Generate random population of chromosomes, that is, suitable solutions for the problem.
Step II [Fitness] Evaluate the fitness of each chromosome in the population.
Step III [New population] Create a new population by repeating following steps until the new population is complete.

1. [Selection] Select two parent chromosomes from a population according to their fitness. Better the fitness, the bigger chance to be selected to be the parent.
2. [Crossover] With a crossover probability, cross over the parents to form
3. new offspring, that is, children. If no crossover was performed, offspring is the exact copy of parents.
4. [Mutation] With a mutation probability, mutate new offspring at each locus.
5. [Accepting] Place new offspring in the new population.

Step IV [Replace] Use new generated population for a further run of the algorithm.
Step V [Test] If the end condition is satisfied, stop, and return the best solution in current population.
Step VI [Loop] Go to step 2.

The genetic algorithms performance is largely influenced by crossover and mutation operators.

Stage s of Building Simulation Experiment

Building the simulation experience involves four important stages :

Stage One - Determination of default values

During this Stage, default values are set , This stage is one of a phases are dependent upon the rest of the stages of simulation, have been chosen default values as follows:

Select Total Views

It is determined by the total number of views through six different sizes of the samples to determine, and thus the total number of views(N=50,100,150,200,250,300) respectively.

Select default parameter values

The following default values were used for the parameters of the binary logistic regression model, note that this is value done identified from to estimate data the real a method OLS as shown at table(1)

Table 1:-The Default Values Parameters in the Binary Logistic Regression Model

Para.	$\beta_\circ$	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
	-1.03	-0.073	0.095	-0.055	-0.000415	0.013	-0.002	0.816	-0.000039	0.001	0.004

Stage Two- Data Generation

At this stage, the values of the explanatory variables and the variables of the response variable are generated, where ten explanatory variables are generated using the monte-carlo method through the uniform distribution, with regard to the generation of variable response values under the assumption that it follows the Bernoulli distribution period [0,1].

The rejection method was used to generate two-step binary response variable:

1. Determining the probability value of the success case (occurrence of the phenomenon)
2. Calculate the number of random numbers so that they are smaller or equal to the probability value of success.

Stage Three - Finding Estimations

It is at this stage estimate the parameters of model only binary logistic regression given in the formula(1)according to the methods, and also according to employment Genetic algorithm with the usual estimation methods, that it mentioned, these methods are:

1. Normal estimation methods (MLE , MCSE, WLSE, BE).
2. Improved estimation methods (MLE.GA, MCSE.GA, WLS.GA , BE.GA).

Stage Four- Comparison of estimation methods

After the estimates are found in the third stage, at this stage the compared between methods are the estimation parameters binary logistic regression model, that bast one of the important statistical standards workers , which is the mean squares error for the model is studied, the aime which is intended achieve is getting estimated which gives less mean squares a possible error [11].

$$MSE = \frac{1}{R}\sum_{i=1}^{N} MSE_i \ldots (34)$$

Represent the frequency of the experiment, the simulation was repeated R=1000 to obtain the results.

Analysis of Simulation Results

The results of the simulation will be presented and analyzed to arrive at the best methods for estimating the parameters of the binary logistic regression model based on the statistical scale mean squares error (MSE)for capabilities model. It has been obtained all the results of the simulation program written Bast workers in a language MATLAB , and the following results will be presented in the tables that will be analyzed according to the sequence of the tables(2),(3),(4),(5),(6),(7) as follows:

Table 2:-Mean Squares Error (MSE) for Estimators Binary Logistic Regression in the case Methods Normal and Genetic at the Sample size 50

Sample size	Methods	Normal	Genetic	Best
N=50	Ml	0.231229552871543	0.005659256776011	Genetic
	Mcs	0.215989821797131	0.000970068091206	Genetic
	Wls	0.019099986421551	0.007367729974521	Genetic
	Bayes	0.896063717954890	0.003344028384662	Genetic
	Best	Wls	Mcs.GA	

for the table(2)

When sample size $(N = 50)$ of ratio for the default values for the parameters and supposed model, It has outperformed (Ml.GA) method on (Ml) method and(Mcs.GA) method on (Mcs) method and(Wls.GA) method on (Wls) method and(B.GA) method on (B) method, Also note that the (Mcs.GA) method of ranked first compared to improved estimation methods and that the (Wls) method of ranked first compared to normal estimation methods.

Table 3:-Mean Squares Error (MSE) for Estimators Binary Logistic Regression in the case Methods Normal and Genetic at the Sample size 100

Sample size	Methods	Normal	Genetic	Best
N=100	Ml	0.195214313074604	0.002374785723625	Genetic
	Mcs	0.155371787851206	0.000468182866024	Genetic
	Wls	0.017568718160916	0.003554196161797	Genetic
	Bayes	0.687385040696479	0.000836777244857	Genetic
	Best	Wls	Mcs.GA	

for the table(3)

When sample size $(N = 100)$ of ratio for the default values for the parameters and supposed model, It has outperformed (Ml.GA) method on (Ml) method and(Mcs.GA) method on (Mcs) method and(Wls.GA) method on (Wls) method and(B.GA) method on (B) method, Also note that the (Mcs.GA) method of ranked first compared to improved estimation methods and that the (Wls) method of ranked first compared to normal estimation methods.

Table 4:-Mean Squares Error (MSE) for Estimators Binary Logistic Regression in the case Methods Normal and Genetic at the Sample size 150

Sample size	Methods	Normal	Genetic	Best
N=150	Ml	0.010072144667444	0.002319596320434	Genetic
	Mcs	0.038498319485976	0.000312568716347	Genetic
	Wls	0.023542680540880	0.002325482871412	Genetic

	Bayes	0.458073693010931	0.000494265654543	Genetic
	Best	Mle	Mcs.GA	

for the table(4)
When sample size ($N = 150$) of ratio for the default values for the parameters and supposed model, It has outperformed (Ml.GA) method on (Ml) method and(Mcs.GA) method on (Mcs) method and(Wls.GA) method on (Wls) method and(B.GA) method on (B) method, Also note that the (Mcs.GA) method of ranked first compared to improved estimation methods and that the (Ml) method of ranked first compared to normal estimation methods.

Table 5:-Mean Squares Error (MSE) for Estimators Binary Logistic Regression in the case Methods Normal and Genetic at the Sample size 200

Sample size	Method	Normal	Genetic	Best
N=200	Ml	0.105158079021595	0.001984539745217	Genetic
	Mcs	0.073204095012192	0.000241971930941	Genetic
	Wls	0.006287675940011	0.001940312308433	Genetic
	Bayes	0.472422431130471	0.000601257242288	Genetic
	Best	Wls	Mcs.GA	

for the table(5)
When sample size ($N = 200$) of ratio for the default values for the parameters and supposed model, It has outperformed (Ml.GA) method on (Ml) method and(Mcs.GA) method on (Mcs) method and(Wls.GA) method on (Wls) method and(B.GA) method on (B) method, Also note that the (Mcs.GA) method of ranked first compared to improved estimation methods and that the (Wls) method of ranked first compared to normal estimation methods.

Table 6:-Mean Squares Error (MSE) for Estimators Binary Logistic Regression in the case Methods Normal and Genetic at the Sample size 250

Sample size	Methods	Classic	Genetic	Best
N=250	Ml	0.118088251341391	0.000941999509518	Genetic
	Mcs	0.084338380647715	0.000186929282195	Genetic
	Wls	0.015775518249187	0.001125336417390	Genetic
	Bayes	0.479378222933252	0.000707549950305	Genetic
	Best	Wls	Mcs.GA	

for the table(6)
When sample size ($N = 250$) of ratio for the default values for the parameters and supposed model, It has outperformed (Ml.GA) method on (Ml) method and(Mcs.GA) method on (Mcs) method and(Wls.GA) method on (Wls) method and(B.GA) method on (B) method, Also note that the (Mcs.GA) method of ranked first compared to improved estimation methods and that the (Wls) method of ranked first compared to normal estimation methods.

Table 7:-Mean Squares Error (MSE) for Estimators Binary Logistic Regression in the case Methods Normal and Genetic at the Sample size 300

Sample size	Methods	Normal	Genetic	Best
N=300	Ml	0.082075989273375	0.001095041889349	Genetic
	Mcs	0.071375768123519	0.000160345430057	Genetic
	Wls	0.005883737386357	0.000775013208108	Genetic
	Bayes	0.287783785987223	0.000146223487213	Genetic
	Best	Wls	Bayes.GA	

for the table(7)
When sample size ($N = 300$) of ratio for the default values for the parameters and supposed model, It has outperformed (Ml.GA) method on (Ml) method and(Mcs.GA) method on (Mcs) method and(Wls.GA) method on (Wls) method and(B.GA) method on (B) method, Also note that the (B.GA) method of ranked first compared to improved estimation methods and that the (Wls) method of ranked first compared to normal estimation methods.

Conclusion:-

Through the experimental side and all the default values for the parameters and the values of the explanatory variables and the regression models that have been assumed logistics and in terms of the number of times to own estimation methods less (MSE) Of the logistic regression model estimators has reached the researcher to the following:

1. Excellence improved estimate methods on the normal methods in estimating parameters of the binary logistic regression model for all sample sizes assumed,this shows improved performance (BLR) model estimatorsmethod when you use of the (GA) in finding the best parameter estimator.
2. Excellence method $(Mcs.GA)$ Which was developed by the researcher on all improved estimation methods, as it has achieved ranked first of preference in estimate the parameters of the binary logistic regression model, for most of samplessizes and assumed model, in other than$(N = 300)$occupied$(B.GA)$ methodranked first.
3. Excellence method (Wls)on all normal estimation methods, as it has achieved ranked first of preference in estimate the parameters of the binary logistic regression model, for most of samplessizes and assumed model,in other than$(N = 150)$occupied(Mle) methodranked first.
4. Value decreases (MSE) for estimators binary logistic regression model using estimate methods all normal and improved by increasing the size of the sample , and this corresponds to the statistical theory.

Acknowledgment:-

I would like to express my sincere gratitude to my supervisor**Dr.Rabab Abdul- Rida**for his suggestion this work, his patience, great support, generous help and continuous guidance throughout the course of study. I would like to thankhead of statistics department**Dr.Qutaiba Nabil**for their assistance and support. Also I would like to thank all the members of the department ofStatisticsat the College ofAdministration and Economics,University Baghdadand in particularly**Dr.Sabah Manfi**for their kind help and support.

Reference:-

1. Akkus ,Ö. , Demir , E. , (2016), "Comparison Some Classical And Meta-Heuristic Optimazation Techniques in The Estimation Of The Logit Model Parameters", International Journal of Advanced Research, *ISSN*, pp.1026-1042.
2. Chatterjee, S., Hadi, A., (2012) ,"Regression Analysis By Example", John Wiley & sons.
3. Cramer, J. S., (2003),"Logit Models From Economics and other fields", Cambridge University Press cape Town, New York, *ISBN*, pp.33-45.
4. Czepiel, S. A., (2002) ,"Maximum Likelihood Estimation Of Logistic Regression Models : Theory and Implementation" , Retrieved from , *https://czep.net/stat/mlelr.pdf*, pp.1-23.
5. Demir , E. , Akkus ,Ö., (2015), " An Introductory Study on How the Genetic Algorithm Works in the Parameter Estimation of Binary Logit Model", International Journal of Sciences : Basic and Applied Research, *ISSN* , pp.162-180.
6. Groenwald , P., Mokgatlhe, L., (2005) ,"Bayesian Computation for Logistic Regression", Computaion Statistics & Data Analysis 48 , University of Bostswana , pp. 857-868.
7. Henry , A. ,(2013), "Bayesian Logistic Regression Modelling Via Markov Chain Monte Carlo Algorithm ", University of Cape Coast , Ghana , pp. 193 – 197.
8. Hosmer, D., Lemeshow, S. ,Sturdivant , R. ,(2013)," Applied Logistic Regression",3rd edition ,New York: *willey http : // ihmsi.org*.
9. Hussain, J. N., Nassir, A. J., (2015)," Cluster Analysis as a Strategy Of Grouping to Construct Goodness – Of – Fit Tests when the Continuous Covariates Present in the Logistic Regression Models", *BJMCS*, pp. 1-16.
10. Matilainen K. & ,(2013)," Employing A Monti Carlo Algorithm in Newton-Type Methods For Maximum Likelihood Estimation of Genetic Parameters", *PLOS ONE, e80821, Doi:10:1371 / journal*, pp.1-7.
11. Menard ,S.,(2002),"Applied Logistic Regression Analysis",2nd Edition Thousand Oaks Edition Thousand Oaks , CA : Sage Publications , Series Quantitative Applications in the Social Sciences , pp. 1 – 111 .
12. Misra, A. ,(2013)," Portfolio optimization of commercial Banks- An Application of Genetic Algorithm", European Journal of Business and management, *ISSN*, pp. 120-129.
13. Muller, Marlene , (2004) ," Generalized Linear Models", Fraunhofer Institute for Industrial Mathematics *(ITWM*) , (Germany) ,*WWW. Marlenmuller . ed / publication / hand book CS. Pdf.*
14. Sukono, Sholahuddin, A. & et al., (2014)," Credit Scoring for Coopera of Financial Services Using Logistic Regression Estimated by Genetic Algorithm", Appied Mathematical Sciences, pp. 45-57.

15. Zager, S., & Karim, R., (1991),"Generalized Linear Models with Random Effects : A Gibbs Sampling Approach", *JASA* , pp. 79-86.
16. *Introduction to Genetic Algorithm cs.felk.cvut.cz/~xobitko/ga//*.