# Using Iterative Reweighting Algorithm and Genetic Algorithm to Calculate The Estimation of The Parameters Of The Maximum Likelihood of The Skew Normal Distribution

**Murtadha Alaa AlKhafaji**
**Al-Taif Islamic Bank, Kut, Wasit Iraq**

alamortada.alaa1101a@coadec.uobaghdad.edu.iq

**Rabab AbdulReda AlBakri**
**University Of Baghdad, Adhamiya, Baghdad, Iraq**

rabab.saleh@coadec.uobaghdad.edu.iq

## Abstract

Excessive skewness which occurs sometimes in the data is represented as an obstacle against normal distribution. So, recent studies have witnessed activity in studying the skew normal distribution (SND) that matches the skewness data which is regarded as a special case of the normal distribution with additional skewness parameter ($\alpha$), which gives more flexibility to the normal distribution. When estimating the parameters of (SND), we face the non-linear equations problem and by using the method of Maximum Likelihood estimation (ML) their solutions will be inaccurate and unreliable. To solve this problem, two methods can be used that are: the genetic algorithm (GA) and the iterative reweighting algorithm (IR) based on Maximum Likelihood method. Monte Carlo simulation was used with different skewness levels and sample sizes, and the superiority of the results was compared. It was concluded that (SND) model estimation using (GA) is the best when the samples sizes are small and medium, while large samples indicate that the (IR) algorithm is the best .The study was also done using real data to find the parameter estimation and a comparison between the results superiority based on (AIC, BIC, Mse and Def) criteria.

**Keywords:** Skew Normal Distribution, Maximum Likelihood Method, Genetic Algorithm, Iteratively Reweighting Algorithm.

## 1- Introduction

Although the normal distribution is the most widespread distribution in statistical theory, the reason is due to its analytical beauty and simplicity of computation, but upon practical application and data analysis, we find that non-normal distributions are more widespread than normal distributions and the reason is due to the excessive skew found in the data and here the importance of studying non-normal distributions appears.

In our research we studied the skew normal distribution (SND) that shares common characteristics of the normal distribution (ND) added of the skew factor (the skewness parameter α), which gives flexibility to the family of this distribution in processing the data that contains the skewness , and when the value (α = 0) the (SND) becomes (ND).

For the estimation of the (SND) model, the maximum likelihood (ML) method was used. However, this method gives non-explicit solutions to the sample observations due to the nonlinear equations that result when the (ML) function is derived. For the purpose of improving and calculating (ML) estimates, iterative reweighing algorithm (IR) and genetic algorithm (GA) based on the (ML) functions were used.

The study was carried out by applying experimental data using Monte Carlo simulation by generating random data distributed in a standard skew normal distribution (SSND) by experimenting with different sample sizes (n=20,50,100,200,500) and studying different levels of the skewness parameter (α = 0.5,2,5). In addition to applying the model to real data and comparing the results of algorithm solutions based on criteria (Mse, Def, AIC and BIC) in order to obtain the best method for calculating the skew normal distribution model (SND).

## 2- Skew Normal Distribution (SND)

The skew normal distribution (SND) was introduced for the first time by the researcher Azzalini in 1985, and the normal distribution (ND) extends to the skew normal distribution (SND) by means of the skewness parameter (α), and the mathematical formula (pdf) for the (SND) distribution of the variable is written The location parameter (μ), the scale parameter (τ) and the shape parameter (α) are as follows [4:pp.331-332].

$$f(y, \mu, \tau, \alpha) = \frac{2}{\tau} \, \emptyset\left(\frac{y - \mu}{\tau}\right) \, \psi\left(\alpha \frac{y - \mu}{\tau}\right) \qquad y \in R \qquad (1)$$

whereas :

$\emptyset\left(\frac{y-\mu}{\tau}\right)$ : The probability density function of the normal distribution .

$\psi\left(\alpha \frac{y-\mu}{\tau}\right)$ : where equal, $\psi\left(\alpha \frac{y-\mu}{\tau}\right) = \int_{-\infty}^{\alpha \frac{y-\mu}{\tau}} \frac{1}{\sqrt{2\pi}} e^{t^2} \, dt$ ,[2:pp.172] .

And when the value of the skew parameter is (α = 0) the distribution (SND) is transformed into the normal distribution (ND), and the standard skew normal (SSND) is called the skewness parameter (α) when the location parameter is (μ = 0) and the scale parameter is (τ = 1),And he writes mathematically as follows [6:pp.4-6] :

$$f(y, \alpha) = 2 \, \emptyset(y) \, \varphi(\alpha y) \, ,$$
$$y \in R \qquad\qquad (2)$$

The following figure shows the probability density function of the (SSND) distribution and the effect of the shape parameter (α) on the shape of the distribution. Values (α=0, 3, 7, 150) were chosen. We note that the value of (α=0) transforms (SND) into the (ND), and the amount of skewness also increases on the left and right when the value of (α) increases [7:pp.908]:
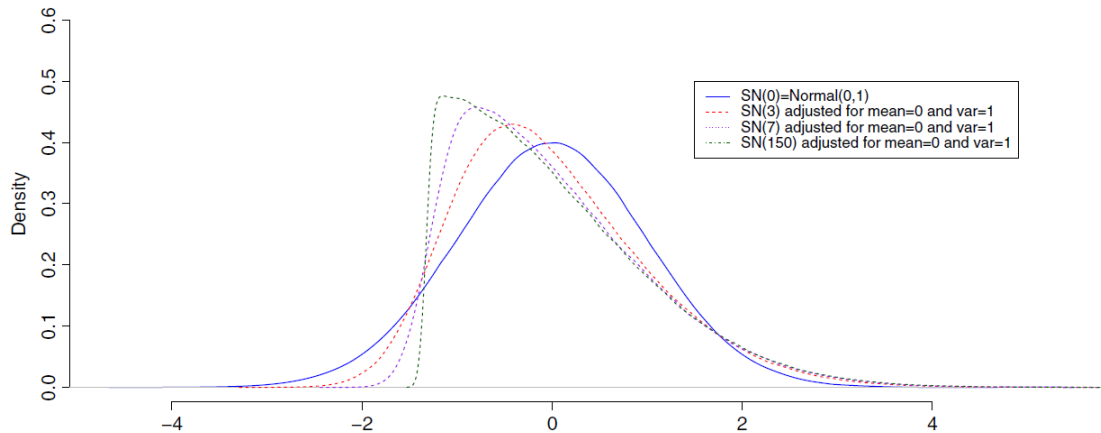


**Figure 1: Effect of parameter (α) on the (SSND).**

The moment generating function (mgf) , mean and variance of the (SSND) can be written as follows [2:pp.174]:

$$M_y^t = 2\, e^{\frac{t^2}{2}}\, \psi(\,t\,) \tag{3}$$

$$E(Y) = \delta\sqrt{\frac{2}{\pi}} \tag{4}$$

$$Var(y) = 1 - 2\frac{\delta^2}{\pi} \tag{5}$$

$\delta$ : **Represents the correlation coefficient whose magnitude is ( $|\delta| < 1$ ), where** $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$

## 3- Maximum Likelihood Estimation (ML)

It is considered one of the most prominent methods used to estimate the parameters of the statistical model with its high efficiency (ML) capabilities. Whereas, the maximum likelihood estimators are the points that make the function the Maximum possible, and the resulting estimates are called the The maximum likelihood estimators (MLE), and the maximization function can be written as the logarithm of the (SND) as follows [10:pp.6] , [6:pp.7]:

$$lnL = nln(2) - nln(\tau) - \frac{n}{2}ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i-\mu}{\tau}\right)^2 + \sum_{i=1}^{n}ln\,\psi\left(\alpha\frac{y_i-\mu}{\tau}\right) \tag{6}$$

Deriving the function with respect to the parameters of (μ, τ, α), the following nonlinear equations are obtained:

$$\frac{\partial lnL}{\partial\mu} = \sum_{i=1}^{n}\left(\frac{y_i-\mu}{\tau}\right) - \alpha\sum_{i=1}^{n}\frac{\emptyset\left(\alpha\frac{y_i-\mu}{\tau}\right)}{\psi\left(\alpha\frac{y_i-\mu}{\tau}\right)} = 0 \tag{7}$$

$$\frac{\partial lnL}{\partial\tau} = -n + \sum_{i=1}^{n}\left(\frac{y_i-\mu}{\tau}\right)^2 - \alpha\sum_{i=1}^{n}\frac{\emptyset\left(\alpha\frac{y_i-\mu}{\tau}\right)}{\psi\left(\alpha\frac{y_i-\mu}{\tau}\right)}\left(\frac{y_i-\mu}{\tau}\right) = 0 \tag{8}$$

$$\frac{\partial lnL}{\partial\alpha} = \sum_{i=1}^{n}\frac{\emptyset\left(\alpha\frac{y_i-\mu}{\tau}\right)}{\psi\left(\alpha\frac{y_i-\mu}{\tau}\right)}\left(\frac{y_i-\mu}{\tau}\right) = 0 \tag{9}$$

Where $w = \frac{\emptyset\left(\alpha\frac{y_i - \mu}{\tau}\right)}{\psi\left(\alpha\frac{y_i - \mu}{\tau}\right)}$ are nonlinear equations, by simplifying the above equations, we get the following:

$$\hat{\mu} = \frac{\sum_{i=1}^{n} y_i \hat{w}(y_i)}{\sum_{i=1}^{n} \hat{w}(y_i)} \tag{10}$$

$$\hat{\tau} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \mu)^2}{n}} \tag{11}$$

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} \left(\frac{y_i - \hat{\mu}}{\hat{\tau}}\right)}{\sum_{i=1}^{n} \hat{w}(y_i)} \tag{12}$$

## 4- Iteratively Reweighting Algorithm (IR)

It is one of the types of function maximization algorithms, and it is used in many previous studies in calculating the estimated of the maximum likelihoods parameters of the skew normal distribution (SND) [3:pp.80]. The idea of this method is to start with the initial (primary) values since the method is not very robust unless good initial values are used, and by updating the nonlinear weight function $(W_i^{(p)} = \frac{\emptyset(\alpha t_i^{(p)})}{\psi(\alpha t_i^{(p)})})$ ,and after selecting the initial values, the nonlinear weight functions (W) are updated repeatedly for the new estimates, as each iteration of the algorithm increases the maximum likelihood function (lnL) until the algorithm converges to the local maximal optimal solution Or the general (global) of the maximum likelihood function [10:pp.10-11],[3:pp.79-80].

Where:

$t_i$: which equals $(t_i = \frac{y_i - \mu^{(p)}}{\tau^{(p)}})$

P: the number of iterations (p = 1, 2, 3 ,.).

The new estimates of the SND parameters are found using the following update equations [3:pp.79-80].

$$\mu^{(P+1)} = \bar{y} - \alpha^{(p)} \overline{W}^{(p)} \tau^{(p)} \tag{13}$$

$$\tau^{(p+1)} = \frac{-\alpha^{(p)}(\sum_{i=1}^{n} W_i^{(p)} y_i - n\overline{W}^{(p)}\bar{y}) + \sqrt{\left(\alpha^{(p)}(\sum_{i=1}^{n} W_i^{(p)} y_i - n\overline{W}^{(p)}\bar{y}\right)^2 + 4n\sum_{i=1}^{n}(y_i - \bar{y})^2}}{2n} \tag{14}$$

$$\alpha^{(P+1)} = \frac{\bar{y} - \mu^{(P)}}{\tau^{(p)} \overline{W}^{(p)}} \tag{15}$$

## 5- Genetic Algorithm (GA)

It is a random search technique that depends on the natural selection mechanism and natural genetics to find the best accurate or approximate solution to the problem under study. It was invented, suggested and developed by Professor John Holland in (1975) [5:pp. 3].The (GA) has similar features to natural systems that are based on the theory of natural evolution by (Darwin) who explained that living things develop on the principle of Natural choice and living organisms, how do they preserve themselves in the presence of the external conditions surrounding them, in order to remain alive across generations,(GA) generates a randomly generated complex of possible solutions, after which each chromosome is evaluated by the evaluation function, and the selection process comes into play, by selecting the best chromosomes, and then the genetic algorithm passes the two hybridization processes during which the best chromosomes are Crossover and the mutation in which random changes occur

between the chromosome genes And it ends with the optimal solution through some iterative adjustments, where the most resolute individuals dominate over the weaker ones through the natural development mechanism [9:pp1019].

## 6- Simulation

Simulation is considered one of the most important tools for solving complex problems through which the models under study can be studied, analysed, and compared between the methods used, and the best solution can be found. We evaluated in this study the use of the Monte Carlo simulation method, which is considered the most common and used in the study and analysis of the parameters of the model under study [10:pp15-18] .

Where in our research we generated a random variable with a standard skew normal distribution (SSND) with the location parameter ($\mu = 0$),the scale parameter ($\tau = 1$) and the shape parameter ($\alpha$),the goal of the simulation is Calculation of the estimator of the parameters of the maximum likelihood for the distribution of the skew  normal depending on the method of maximum likelihood, where the two methods of genetic algorithm (GA) and iterative re-weighing algorithm (IR),to compare the preference of the two methods in calculating the model parameters, several different sizes of the sample were applied (n = 20,50,100,200,500),in addition to studying the effect of skewness on by taking different values for shape parameter ($\alpha = 0.5,2,3$),noting that if it was a parameter value The shape is equal to zero, it turns into a symmetric normal distribution, and when the parameter value increases, the skewness increases.

We compared the results extracted from the methods used by the arithmetic mean, the mean squares of error (Mse), which is the lower its value, the better it is, and the sum of the mean squares of error of the distribution parameters (Def), which measures the preference of the three parameters in the efficiency of the model as a whole. Its value is less, the better.

We note from the simulation results the following:

1- In Table (1), when ($\mu = 0$, $\tau = 1$, $\alpha = 0.5$), when calculating the estimation of the location parameter ($\mu$) and the skewness parameter ($\alpha$), we notice that the iterative weighting algorithm (IR) is better than the genetic algorithm. (GA) in all samples sizes used to obtain the lowest values of (Mse), and in calculating the scale parameter ($\tau$) when using small and medium samples sizes, (GA) was the best, as well as the superiority of the (IR) algorithm for large samples sizes.

**Table 1: Calculate model parameters for (SND) , when ( μ =0 , τ =1 , α = 0.5 )**

| N | Method | $\hat{\mu}$ | | $\hat{\tau}$ | | $\hat{\alpha}$ | | |
|---|--------|------|-----|------|-----|------|-----|-----|
| | | Mean | Mse | Mean | Mse | Mean | Mse | Def |
| 20 | IR (ML) | 0.487136 | 0.000875 | 0.260399 | 0.319772 | 0.000562 | 0.00000001 | 0.320647 |
| | GA(ML) | 0.184553 | 0.001894 | 0.232137 | 0.065394 | 0.243303 | 0.000588 | 0.067877 |
| 50 | IR (ML) | 0.400528 | 0.000232 | 0.322131 | 0.215312 | 0.000040 | 0.0000004 | 0.215545 |
| | GA(ML) | 0.196311 | 0.000977 | 0.259290 | 0.111822 | 0.263397 | 0.000697 | 0.113497 |
| 100 | IR (ML) | 0.375059 | 0.000229 | 0.270832 | 0.251487 | 0.000107 | 0.0000001 | 0.251716 |
| | GA(ML) | 0.239171 | 0.000826 | 0.307702 | 0.247803 | 0.326604 | 0.001065 | 0.249695 |
| 200 | IR (ML) | 0.389489 | 0.000056 | 0.260699 | 0.228934 | 0.000018 | 0.0000004 | 0.228991 |
| | GA(ML) | 0.234021 | 0.000687 | 0.304668 | 0.451937 | 0.319979 | 0.001023 | 0.453648 |
| 500 | IR (ML) | 0.324833 | 0.000065 | 0.237638 | 0.246605 | 0.000050 | 0.00000006 | 0.246670 |
| | GA(ML) | 0.212209 | 0.000862 | 0.274592 | 1.224832 | 0.285900 | 0.000816 | 1.226511 |

2-In Table (2), when (μ = 0, τ = 1, α = 2), we note that the (GA) algorithm was superior in estimating the location parameter (μ) as well as in estimating the scale parameter (τ) when using small samples. The medium samples sizes, as well as the medium samples sizes (IR) algorithm components are more efficient than (GA) in calculating the estimation of the shape parameter (α) in all the samples sizes used and the large samples sizes with respect to the estimation of the scale parameter (τ).

**Table 2: Calculate model parameters for (SND), when ( μ =0 , τ =1 , α = 2 )**

| n | Method | $\hat{\mu}$ | | $\hat{\tau}$ | | $\hat{\alpha}$ | | |
|---|--------|------|-----|------|-----|------|-----|-----|
| | | Mean | Mse | Mean | Mse | Mean | Mse | Def |
| 20 | IR (ML) | 0.091815 | 0.009674 | 0.184248 | 0.339320 | 0.143655 | 0.000175 | 0.349170 |
| | GA (ML) | 0.183431 | 0.001215 | 0.232137 | 0.065394 | 0.2462686 | 0.000267 | 0.066878 |
| 50 | IR (ML) | -0.34539 | 0.011461 | 0.239420 | 0.226305 | 0.070619 | 0.000001 | 0.237768 |
| | GA (ML) | 0.195159 | 0.000471 | 0.259290 | 0.111822 | 0.265894 | 0.000426 | 0.112720 |
| 100 | IR (ML) | -0.36777 | 0.006497 | 0.197035 | 0.258943 | 0.076428 | 0.000013 | 0.265454 |
| | GA (ML) | 0.238225 | 0.000401 | 0.307702 | 0.247803 | 0.329018 | 0.000841 | 0.249045 |
| 200 | IR (ML) | -0.59017 | 0.004967 | 0.192246 | 0.233589 | 0.064924 | 0.000006 | 0.238563 |
| | GA (ML) | 0.233328 | 0.000300 | 0.304668 | 0.451937 | 0.321671 | 0.000859 | 0.453097 |
| 500 | IR (ML) | -0.80667 | 0.003002 | 0.173303 | 0.249470 | 0.0613751 | 0.000004 | 0.252477 |
| | GA (ML) | 0.211862 | 0.000427 | 0.274592 | 1.224832 | 0.286742 | 0.000724 | 1.225984 |

3- In Table (3), when (μ = 0, τ = 1, α = 3), we note from the results of (Mse) that (GA) was the best in estimating the location parameter (μ) in all samples sizes because it had less Values from the (Mse) standard, as for the calculation of the scale parameter (τ), then (GA) outperformed in small and medium samples sizes. As for the large sample sizes of the parameter (τ), the (IR) algorithm was superior. All samples sizes used in estimating the skewness parameter (α).

**Table 3: Calculate model parameters for (SND), when ( μ =0 , τ =1 , α = 3 )**

| N | Method | $\hat{\mu}$ Mean | $\hat{\mu}$ Mse | $\hat{\tau}$ Mean | $\hat{\tau}$ Mse | $\hat{\alpha}$ Mean | $\hat{\alpha}$ Mse | Def |
|---|---|---|---|---|---|---|---|---|
| 20 | IR (ML) | 0.104157 | 0.009292 | 0.177810 | 0.340999 | 0.150017 | 0.000198 | 0.350490 |
|  | GA (ML) | 0.183395 | 0.001241 | 0.232137 | 0.065394 | 0.246364 | 0.000259 | 0.066895 |
| 50 | IR (ML) | -0.33826 | 0.011258 | 0.230879 | 0.227455 | 0.077883 | 0.000003 | 0.238717 |
|  | GA (ML) | 0.195077 | 0.000473 | 0.259290 | 0.111822 | 0.266075 | 0.000409 | 0.112705 |
| 100 | IR (ML) | -0.35414 | 0.006310 | 0.189576 | 0.259702 | 0.082725 | 0.0000165 | 0.266030 |
|  | GA (ML) | 0.238171 | 0.000409 | 0.307702 | 0.247803 | 0.329157 | 0.000828 | 0.249041 |
| 200 | IR (ML) | -0.58194 | 0.004890 | 0.185057 | 0.234081 | 0.071897 | 0.000008 | 0.238980 |
|  | GA (ML) | 0.233276 | 0.000302 | 0.304668 | 0.451937 | 0.321796 | 0.000848 | 0.453087 |
| 500 | IR (ML) | -0.79269 | 0.002940 | 0.166667 | 0.249767 | 0.067336 | 0.000004 | 0.252712 |
|  | GA (ML) | 0.211839 | 0.000432 | 0.274592 | 1.224832 | 0.286798 | 0.000718 | 1.225984 |

We note from the tables above Table (1,2,3) in all the difference in the size of the skewness and based on the value of the standard sum of the mean squares error of the samples (Def) for the skew normal distribution (SND), that the genetic algorithm (GA) possessed the lowest value of the standard at small samples sizes And the mean (n = 20,50,100), therefore, is considered the best in the model estimation, but when using a large sample size (n = 200,500) the iterative (IR) weighing algorithm is superior in calculating the (SND) model .

**7- Application**

The study relied on employing medical data for the variable of Diabetes in the human being measured for (250) of reviewing patients (diabetes is a chronic disease resulting from the inability of the pancreas to generate insulin in sufficient quantity), and the data were collected from the records of Al Zahra Teaching Hospital in Wasit Governorate. These data were used to estimate the parameters of the skew normal distribution (SND). In order to know the behavior of this data, descriptive statistics for the variable under study were extracted, as shown in the following table:

**Table 4: Descriptive statistics for the variable of diabetes**

| Mean | Median | Mode | Std.deviation | Variance | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| 198.64 | 155.00 | 120.00 | 103.02 | 10615.01 | 1.90 | 3.60 |
| Minimum | Maximum | Sum. | | | | |
| 620.00 | 100.00 | 49660.00 | | | | |

From the above table, it can be seen that the data shows a high degree of inconsistency. The reason is due to the high skew in the distribution of the data, as the value of the skew coefficient (1.90) for the variable under study.

To know the shape of the histogram and the bar chart of the variable of diabetes, the two charts were drawn as follows :
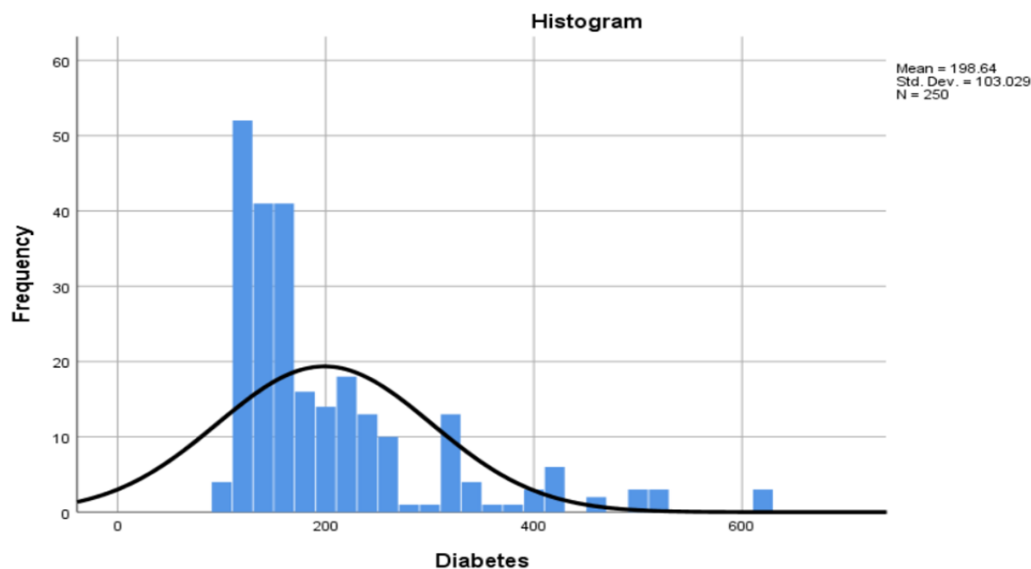
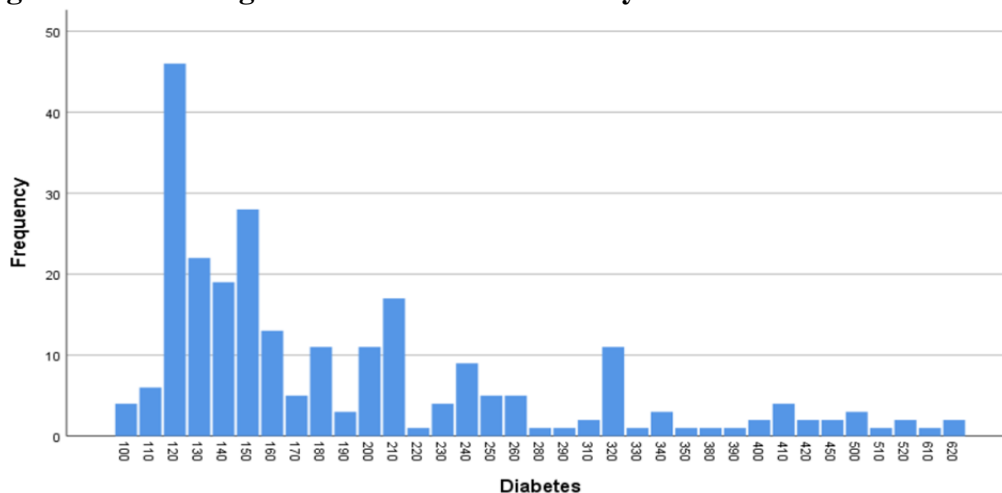**Figure 2: The histogram of the data under study.**



**Figure 3: The bar chart for the data under study.**

Through the above drawings, the behavior of the data for the studied variable can be observed and it is clearly shown that it does not distribute the normal distribution (ND) due to the existing skewness, but rather it tends to the skew normal distribution (SND). It can be concluded from Table (4) and the graphs that the data has a curvature to the right (positive skewness), and to make sure of what has been reached, we must test whether the data is distributed normally or not. To do this, the Kolmogorov-Smirnov test was employed which is based on the following hypothesis:

$H_0$: Sample data follow a normal distribution (ND).

$H_1$: The sample data do not follow a normal distribution (ND).

The result of the test statistic value was equal to (0.000), which is less than the value of the level of significance (0.05). Thus, the null hypothesis is rejected and the alternative hypothesis is accepted (the significant differences). This indicates that the data is not distributed as a normal distribution (ND), but rather as a skewed normal distribution (SND). For the purpose of calculating the estimation of the parameters of the maximum likelihood (ML) of the skew

normal distribution (SND), the program (MATLAB 2012) was used by using the iterative reweighting algorithm (IR) and the genetic algorithm (GA). :

Table 5: Calculate the values of the estimator of the parameters of (SND).

|     | $\mu$ | $\tau$ | $\alpha$ | Def | Lnl | AIC | BIC |
|-----|-------|--------|----------|-----|-----|-----|-----|
| IR  | 198.6400 | 2.0852 | 1.6284 | 4.5071 | -2.5510 | 5.1080 | 5.1186 |
|     | 3.4188 | 4.5071 | 1.0000 | | | | |
| GA  | 1.1643 | 7.3834 | 4.7373 | 18.4691 | -3.2295 | 6.4639 | 6.4727 |
|     | 3.1554 | 18.4681 | 1.0000 | | | | |

Rough Table (5), the estimation of the location, scale and shape parameter of the (SND) of the (IR) algorithm was (198.6400, 2.0852 and 1.6284) respectively, while in the genetic algorithm (GA) the estimation of the parameters was (1.1643, 7.3834 and 4.7373) on Straight. We note from the results of the table that the results of the (IR) algorithm are the best because it has the smallest values from the Akakiy Information Standard (AIC) and the Bayse Information Standard (BIC) in addition to the standard sum of average squares of error for the parameters (Def), where each of the two methods was (4.5071 and 18.4691) In a row, it can be concluded that the iterative reweighting algorithm (IR) algorithm is the most reliable and reliable in showing the best performance in representing diabetes data for a skew normal distribution (SND).

The application aspect can be summarized in terms of the method preference used (GA and IR) in the calculation of the model estimation and the parameters of the skew normal distribution (SND).
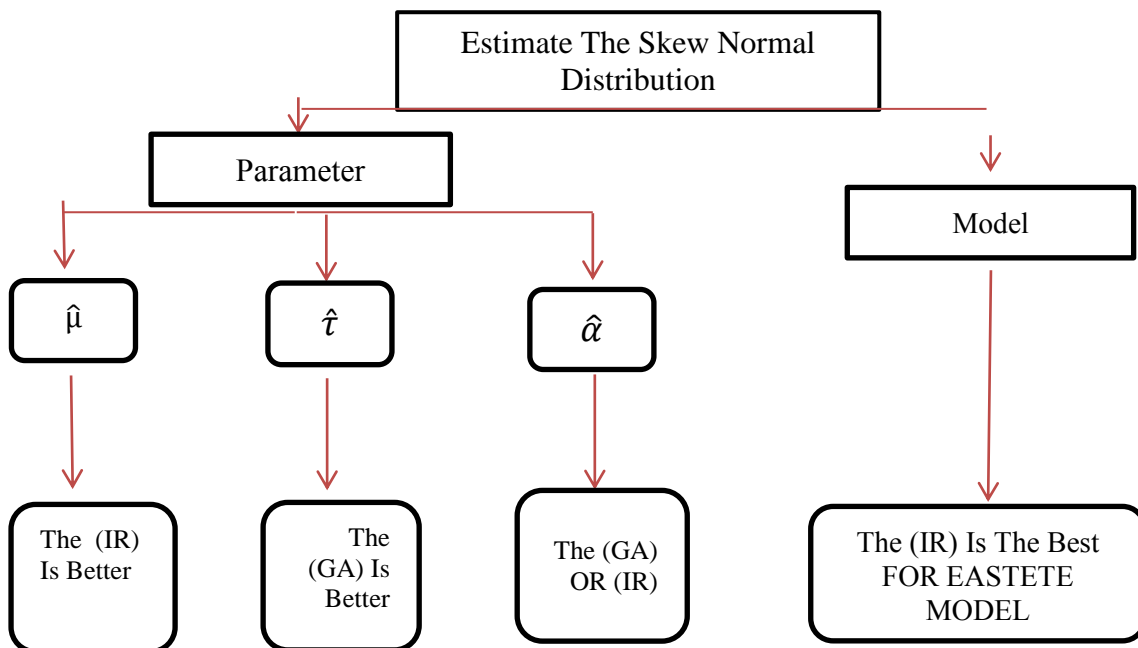


Diagram 1: shows the preference of the methods used in estimating the model and parameters from the (SND).

## 8- Conclusions

Based on the experimental and Application side, the most important conclusions can be summarized as follows:

1- Based on the simulation, we see the superiority of the results of the genetic algorithm (GA) in estimating the model skew normal distribution (SND) at samples sizes (n= 20, 50, 100) because it has the lowest value (Def) regardless of the skewness value.

2- Based on the simulation, it was concluded that the results of the iterative reweighting algorithm (IR) of the model were superior to the sample sizes (n = 200, 500) because it obtained the lowest value for the (Def) criterion.

3- Through the simulation results, we find that the (IR) algorithm was the best in calculating the estimator of the location parameter ($\mu$) when the skewness was slight (0.5), while the genetic algorithm (GA) was the best when the skewness increased from (0.5).

4-Through the simulation results, it was concluded that the (GA) method is better because it has less (Mse) when samples are small or medium, and (IR) is the best for large samples sizes in estimating scale parameter ($\tau$).

5-We notice from the simulation study that the (IR) algorithm has a clear advantage in calculating the estimator of the skewness parameter ($\alpha$) because it has less (Mse).

6-On the application side, we note the preference of the iterative reweighting algorithm (IR) in estimating the skew normal distribution model (SND) over the genetic algorithm (GA) because it has the lowest values of the calibrators (AIC, BIC and Def).

7-It can be concluded from the Application side that the (GA) algorithm is better in calculating the estimation of the location parameter ($\mu$) because it has the lowest value of (Mse) and a clear superiority of the (IR) algorithm by employing it to calculate the estimation of the scale parameter ($\tau$) because it has the lowest value of( Mse). The (IR) algorithm and the (GA) are equal in calculating the estimation of the skewness parameter ($\alpha$) based on the (Mse) criterion for having them.

## References

1- Al-Douri , Yamur , Hamodi , Hussan and Lundberg ,  Jan ( 2020 ) , "Risk-based life cycle cost analysis using a two-level multi-objective genetic algorithm" INTERNATIONAL JOURNAL OF COMPUTER INTEGRATED MANUFACTURING .

2- AZZALINI , A. ( 1985 ) , "A Class of Distributions which Includes the Normal Ones " , Scand J Statist 12: 171-178 .

3- Celik , Nuri , Senoglu , Birdal , Arslan , Olcay ( 2015 ) , "Estimation and Testing in One-Way ANOVA when the Errors are Skew-Normal " , Revista Colombiana de Estadística , Volume 38, Issue 1, pp. 75 to 91 .

4- Chiogna , Monica (2005 ) , "A note on the asymptotic distribution of the maximum likelihood estimator for the scalar skew-normal distribution " , Statistical Methods & Applications  14: 331–341 .

5- Elvira-Ortiz , David , Jaen-Cuellar , Arturo , Morinigo-Sotelo , Daniel , Morales-Velazquez , Luis , Osornio-Rios , Roque and Romero-Troncoso , Rene (

2020 ) , "Genetic Algorithm Methodology for the Estimation of Generated Power and Harmonic Content in Photovoltaic Generation" , Appl. Sci., 10, 542 .

6- Figueiredo , Fernanda , Gomes , M. (2011) , "THE SKEW-NORMAL DISTRIBUTION IN SPC " , Monitoring industrial processes with robust control statistics .

7- Gupta , Arjun , Chen , Tuhao ( 2001 ) , "GOODNESS-OF-FIT TESTS FOR THE SKEW-NORMAL DISTRIBUTION" , COMMUN. STATIST.—SIMULA., 30(4), 907–930 .

8- Gupta , Ramesh , Brown , Nicole ( 2001 ) "RELIABILITY STUDIES OF THE SKEW-NORMAL DISTRIBUTION AND ITS APPLICATION TO A STRENGTH-STRESS MODEL " , COMMUN. STATIST.THEORY METH., 30(11), 2427–2445.

9- Rodriguez-Guerrero , Marco , Jaen-Cuellar , Arturo , Carranza-Lopez-Padilla , Rene , Osornio-Rios , Roque , Herrera-Ruiz , Gilberto and Romero-Troncoso , Rene ( 2018 ) , " Hybrid Approach Based on GA and PSO for Parameter Estimation of a Full Power Quality Disturbance Parameterized Model " , IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 14, NO. 3 .

10- Yalçınkaya , Abdullah , Şenoğ̌lu , Birdal And Yolcu , Ufuk (2017),"MAXIMUM LIKELIHOOD ESTIMATION FOR THE PARAMETERS OF SKEW NORMAL DISTRIBUTION USING GENETIC ALGORITHM " ,Swarm and Evolutionary Computation.

# استخدام خوارزمية اعادة الوزن التكرارية والخوارزمية الجينية لحساب مقدر معلمات الامكان الاعظم للتوزيع الطبيعي الملتوي

| أ.م.د رباب عبد الرضا البكري | مرتضى علاء الخفاجي |
|---|---|
| جامعة بغداد ، الاعظمية، بغداد ، العراق | مصرف الطيف الاسلامي، كوت، واسط، العراق |
| **rabab.saleh@coadec.uobaghdad.edu.iq** | **alamortada.alaa1101a@coadec.uobaghdad.edu.iq** |

**مستخلص البحث:**

يعد الالتواء الزائد الموجود في البيانات هو عقبة امام التوزيع الطبيعي (ND) ، لذلك شهدت الدراسات الحديثة نشاطا في دراسة التوزيع الطبيعي الملتوي (SND) الذي يلائم البيانات الملتوية. ويعتبر الاخير هو حالة خاصة من التوزيع الطبيعي مضاف الية عامل الالتواء ($\alpha$) ، الذي يعطي مرونة أكثر للتوزيع الطبيعي . عند تقدير معلمات (SND) باستخدام طريقة الامكان الاعظم (ML) نواجه مشكلة ظهور معادلات غير الخطية ، لذا فأن حلولها تكون غير صريحة وغير دقيقة ولا يمكن الاعتماد عليها . لحل هذه المشكلة ، تم استخدام طريقة الخوارزمية الجينية (GA) و خوارزمية اعادة الوزن التكرارية (IR) اعتمادا على طريقة الامكان الاعظم . تم استخدام طريقة محاكاة مونت كارلو بمستويات مختلفة من الالتواء واحجام العينة ، وتم مقارنة افضلية النتائج وفقا للمعايير الاحصائية حيث كان حساب انموذج (SND) باستخدام (GA) هي الافضل عندما تكون حجم العينه صغير او متوسط، وعندما يكون حجم العينية كبير، فأن خوارزمية (IR) الافضل. واجريت دراسة ايضا باستخدام بيانات حقيقية لحساب تقدير المعالم والمقارنة بين النتائج إستناداً الى المعايير (AIC، BIC، Mse, Def ).

**المصطلحات الرئيسة للبحث** : التوزيع الطبيعي الملتوي، طريقة الامكان الاعظم، الخوارزمية الجينية، خوارزمية اعادة الوزن التكرارية .