Tikrit Journal of Administrative and Economic Sciences, Vol. 21, No. 72, Part (1): 488-503
**Doi:** www.doi.org/10.25130/tjaes.21.72.1.25

**Tikrit Journal of Administrative and Economics Sciences**
**مجلة تكريت للعلوم الإدارية والاقتصادية**
**EISSN: 3006-9149**
**PISSN: 1813-1719**

# Estimating and Enhancing Parameters for Linear Regression Model with Multivariate t-Distribution using PSO

**Mohammed Zuhair Khaleel***
Ministry of higher Education and Scientific Research

***Keywords:***
Least Square Error, Maximum Likelihood Estimator, Money, Particle Swarm Algorithm, distribution.

**ARTICLE INFO**

***Article history:***

| | |
|---|---|
| Received | 08 May. 2025 |
| Received in revised form | 22 May. 2025 |
| Accepted | 25 May. 2025 |
| Available online | 31 Dec. 2025 |



*Corresponding author:
**Mohammed Zuhair Khaleel**
Ministry of higher Education and Scientific Research

**Abstract:** Linear regression is one of the significant subjects which the researchers have been addressed in their researches. It is based on the assumption that random errors follow a normal distribution, traditional and Bayesian estimation. These methods were used for estimating parameters for linear regression model. Then, the two estimation methods can be compare based on measures (unbiasedness, mean square error) by analyzing real data if available, or using the simulation method to determine the appropriate method for estimation. As for this work, it focuses on the topic of parameters estimation with linear regression model when random errors have a multivariate t-distribution, using Maximum-Likelihood-Estimator (MLE) and Least-Square-Error (LSE) methods. Furthermore, the Particle-Swarm-Optimization (PSO) is useful for optimizing these methods by minimization these error methods. These methods are applied on real data (money supply and factors affected on it) for the years (2011-2021). The real data has been taken from Iraqi Central Bank. The Mean Square Error (MSE) result showed that a PSO method has less MSE as compared to two methods (MLE and LSE). The evidence to verify these MSE results evaluated in this research as the following: MLE=0.5949, Alternating Least Square approach (ALS)=0.20035, PSO=0.1859. PSO has also optimized the parameters of the linear regression model. It was concluded from MSE result findings that these results have become more accurate with PSO approach. The t-distribution is chosen because it used for small sample numbers or unknown variations.

Doi: www.doi.org/10.25130/tjaes.21.72.1.25

# تقدير وتحسين معاملات نموذج الانحدار الخطي مع توزيع-t لمتعدد المتغيرات باستخدام PSO

**محمد زهير خليل**
**وزارة التعليم العالي والبحث العلمي**

**المستخلص**
ان الانحدار الخطي هو أحد المواضيع المهمة التي تناولها الباحثون في أبحاثهم. واستند إلى افتراض أن الأخطاء العشوائية تتبع التوزيع الطبيعي والتقدير التقليدي والتقدير البايزي. وقد استخدمت هذه الطرق لتقدير معلمات نموذج الانحدار الخطي. ومن ثم يمكن مقارنة طريقتي التقدير بناءً على المقاييس (عدم التحيز، ومتوسط مربع الخطأ) من خلال تحليل البيانات الحقيقية إن وجدت، أو باستخدام طريقة المحاكاة لتحديد الطريقة المناسبة للتقدير. أما بالنسبة لهذا العمل، فإنه يركز على موضوع تقدير معلمات نموذج الانحدار الخطي عندما يكون للأخطاء العشوائية توزيع -t لمتعدد المتغيرات، باستخدام طريقتي مقدر الامكان الاعظم (MLE) وخطأ المربعات الصغرى (LSE). علاوة على ذلك، تم استخدام تحسين سرب الجسيمات (PSO) لتحسين هذه الطرق بواسطة تقليل أخطاء هذه الطرق. وتطبق هذه الطرق على البيانات الحقيقية (المعروض النقدي والعوامل المؤثرة عليه) للسنوات (2011-2021). تم الحصول على البيانات الحقيقية من البنك المركزي العراقي. أظهرت نتيجة MSE أن طريقة PSO لها MSE أقل مقارنة بطريقتين (MLE وLSE). تم تقييم الأدلة للتحقق من نتائج MSE هذه في هذا البحث على النحو التالي: MLE = 0.5949، (ALS) 0.20035= طريقة المربعات الصغرى المتناوبة، PSO = 0.1859. كما قامت PSO بتحسين معلمات نموذج الانحدار الخطي. تم الاستنتاج من نتائج MSE أن هذه المعلمات أصبحت أكثر دقة مع طريقة PSO. تم اختيار توزيع-t لأنه يستخدم لأعداد عينات صغيرة أو اختلافات غير معروفة.
**الكلمات المفتاحية:** خطأ المربعات الصغرى، مقدر الاحتمال الأقصى، المال، خوارزمية سرب الجسيمات التوزيع.

## 1. Introduction

Statistical modeling is a way of using statistical analysis methods to find, examine, evaluate, and correct patterns of data. The stage of placing the data within a studied scientific model it is one of the most important stages of statistical analysis on which the rest of the stages depend. This stage is known as the “data modeling” stage, which known probability distributions are used to represent the data and predict them (Al-Birmani et al., 2015). The development of the topic of “linear” and “non-linear” regression is based on the development of statistical methods used in the process of estimating and testing the parameters of the linear and non-linear regression model, in two directions, the first one is related to the classical school and the second related to the Bayesian school. The difference between them is that the traditional school considers the parameters to be estimated as constant values and not random variables, while the Bayesian school considers the parameters to be estimated to be random variables and not fixed values (Nieto et al., 2024) and (Denis et al., 2018). Many traditional and Bayesian estimation methods have emerged for regression models when the random errors are normally distributed. Among the traditional estimation methods were the least squares method. Regarding the Bayesian approaches, they vary based on the primary probability function that can be maintained. These functions include (a non-informative prior distribution. An informative prior distribution a natural conjugate prior distribution a prior distribution based on previous samples. When faced with the difficulty of representing prior information in the form of an initial probability function, the empirical Bayesian method is commonly used (Safiul et al., 1990), (Visco et al., 1988) and (Kazem et al., 1988). PSO is one of the most commonly used algorithms for optimization (Khanesar et al., 2007). A statistical method called regression established a connection between dependent and independent variables (Lang et al., 2020).

The contribution work presented by applying the t-distribution of the multivariate random error, i.e. linear-regression-model-suffers from the problem of deviation from the normal distribution. By using two methods, it can be estimating parameters of a linear model with the t-distribution of the multivariate random error. The least-squares (LS) approach and the maximum-likelihood approach (MLE) used for estimating parameters of simple linear regression model. Then improving the solution using artificial intelligence algorithms, represented by optimization algorithm is called Particle Swarm method. PSO, ALS, and LSE were used with the t-distribution to increase parameter estimation accuracy in the presence of unknowable variances and small data size.

The organization of this work includes the following: Section 2 is research problem, Section three is research aim, Section four is reference review, Section five is methodology, Section six is particle swarm algorithm, Section seven is results and discussion, Section eight is conclusion, and finally section nine is future work.

2. **Research Problem:** The main problem is to develop an accurate method to estimate and optimize the parameters of the “linear’ “regression” model by multivariate- t-distribution.

3. **Research Aim:** The research aim of this research problem can explained as below:

❖ Estimating "regression model" by multivariate-t-distribution using maximum likelihood method (MLE) and the least squares (LS) method. Then study the properties of a multivariate t- distribution and estimate its parameters.

❖ Improving parameters of the model using artificial intelligence algorithm (PSO). PSO is combined with the MLE approach to optimize the likelihood formula because it is complex, nonlinear, and lack of solution. By avoiding local of optima and enhancing robustness for parameter estimation, PSO offered global-searching capacity.

❖ Comparison between estimation methods using the comparison criterion Mean Square Error (MSE). Finally, PSO, ALS, and LSE were used with the t-distribution to increase parameter estimation accuracy in the presence of unknowable variances and small data size.

4. **Reference Review:** (Liu et al., 2008) introduced and elaborated on two parameter changes multivariate t-regression models featuring Power Exponential (PE) random errors. The initial model assumes independent observations, while the second model considers dependent observations. Furthermore, it was utilized genetic approach to estimate model parameters and conduct predictor subset selection under both model types. (Lopez, 2011) presented method for estimating the joint-distribution for responses within a regression model with a uni-variate. Drawing inspiration from Beran's (1981) conditional Kaplan-Meier estimator, the proposed estimator extended multivariate utilized in unrestricted scenarios. The authors have been established asymptotic depictions of an integral by the estimated distribution formula. Their methodology was found applications in censored regression and density estimation tasks. (Safiul et al., 1990) introduced the concept of prediction distributions, which represent the distributions of future responses based on data obtained from an informative experiment. The researchers have been derived these prediction distributions from the model featuring a multivariate t-Student for t-distribution, leveraging the structural relationships within prototype. Notably, the prediction distributions were identified as multivariate-t-variables that remain independent of those of the error distribution. (Al-Birmani et al., 2015) conducted a comparative analysis between genetic algorithms and neural

networks for estimating the mediator's location. The authors identified a specific method and they noted that sample sizes increased. A genetic algorithm and neural network were generating new generations. Furthermore, the determination of the optimal method from the comparison between genetic algorithms, neural networks, and proposed techniques were explained. The coefficients of the multivariate t-regression model for time-varying and the features were estimated using smoothing splines. Many strategies of alignment for analytical prospective are studied by researchers in (Su et al., 2012) and (Kunle et al., 2014). The multivariate of t-distribution is described in (Roth et al., 2012). (Retno et al., 2018) explored the Bayesian method as a means of estimating parameters within multivariate-t-multiple-regression. The prior of Jeffrey-distribution was considered a suitable when information about parameters is lacking. By combining Jeffreys' prior with sample information, the posterior distribution was derived, which was then utilized for parameter estimation. The research aimed to estimate these parameters using the Bayesian method with Jeffreys' prior distribution. In 2020, Lang et al. (Lang et al., 2020) studied statistical methodologies which were aimed to probabilistically forecasting wind profiles along airport approach paths for the next hour. Accurate predictions of wind profiles were crucial for enhancing safety and optimizing air traffic management, particularly. In scenarios, sudden shifts in wind direction necessitate timely rerouting of landing aircraft. The researchers improved distributional regression trees and forests to predict vertical wind profiles by utilizing a multivariate t-normal distribution. These models were designed to provide probabilistic forecasts for both wind speed and direction by simultaneously modeling the components of the wind vector across various height levels of a measurement tower. The resultant tree-based models effectively capture nonlinear effects and interactions, automatically selecting relevant covariates associated with changes in any parameter of the potentially high-dimensional multivariate t-normal distribution utilized. Extending these models to multivariate t-distributional regression forests further enhances predictive performance through regularization and smoothing of covariate effects. (Nieto et al., 2024) discussed the utility of projected distributions in analyzing circular and directional data. While various multivariate t-distributions can be applied to create projected models, these distributions were typically parametric. They were essential for defining directional-

Doi: www.doi.org/10.25130/tjaes.21.72.1.25

directional regression models. Additionally, the multivariate-t-linear-regression by Polya-tree-errors was defined, and then projected to create a linear-regression-directional-model. The characterization of posterior for all-the models is obtained through their filled restricted distributions. The performance of these models was demonstrated through both simulated and real datasets.

5. **Methodology:** This section included the following:

**5-1 Multivariate of t-distribution**: A (single-variable) addition of a normal probability distribution from one dimension to the highest dimension is the multivariate t-normal distribution. In other words, if every linear combination of k's components had a univariate normal distribution, then a random vector would be a k-variable normally distributed. However, the multivariate t-central limit theorem is primarily responsible for its significance (Roth et al., 2012). A probability density function for the multivariate-t-distribution is as the following: If a random value (x) takes form as multivariate-t-distribution, then it was probability density function (pdf) is as follows: (Roth et al., 2012), (Retno et al., 2018) and (Rens et al., 2021).

$$f(x) = \frac{g(V_\circ)}{(\sigma^2)^{\frac{n}{2}}}[V_\circ + (x-\mu)'(x-\mu)/\sigma^2]^{-(n+V_\circ)/2}, \quad \sigma, V_\circ > 0, -\infty < x < \infty \tag{1}$$

The variables were represented as follows:

$f(x)$ : the probability density function (pdf).

$g(V_\circ)$: this was a formula of $V_\circ$ that can be defined in the Eq. (2).

$\sigma$ : scalar variance.

$V_\circ$: possible variable which was represented the degrees freedom in statistic.

n: observations number.

x: The variable of density formula was defined.

$\mu$: the distribution mean.

$(x-\mu)'$: vector's transpose, and $(x-\mu)$ : x and $\mu$ were vectors and they indicated to multivariate distribution. $[V_\circ + (x-\mu)'(x-\mu)/\sigma^2]$ : this term adjusted the distribution scale depended on x , and $V_\circ$.

$$g(V_\circ) = V_\circ^{\frac{V_\circ}{2}}\left(V_\circ + \frac{n}{2}\right)/\pi^{\frac{n}{2}}\ \Gamma\left(\frac{V_\circ}{2}\right) \tag{2}$$

Γ: was denoted Gamma function. It was generalized the factorial function to non-integer values. The following Figure 1 showed “probability density function”. (pdf) of the distribution.

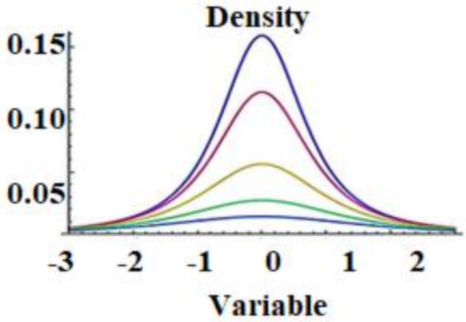


Figure (1): Pdf function of the distribution (Roth et al., 2012), (Retno et al., 2018). and (Rens et al., 2021).

**5-2 Regression analysis:** Regression analysis is a fundamental pillar of statistics. It was one of the important methods in applied statistics. It used to study all the phenomena involved in solving many problems. It was widely used in statistical tools, and to provide a simple method to determine the relationship between variables. By dealing with various formulas of mathematical models, it can described relationships between variables. So that models of these relationships can be used for the purpose of description, prediction, estimation, and other statistical conclusions (Kazem et al., 1988). The linear models were being addressed, which were written in the form of a linear model. It has two types: the simple linear regression model, which was represented by a dependent and non-dependence variables and the multiple (general) linear regression models, which it was considered the most common through its use in most research and studies and its application in many fields. It was used in medical, financial, agricultural fields (Kazem et al., 1988). In this research, the (parameters for linear-regression-model) would be estimated when an error is multivariate t- distributed.

**5-3 Linear-regression-model (Visco et al., 1988)**: There is a linear-relationship between a dependent single variable ($Y_i$), independent single variable ($X_i$), and random error term ($U_i$) as in Equation (3). So that this model is univariate-regression as in of Equations (4) and (5). But, random error term ($U_i$) is come from a multivariate-distribution. This method used for accounting variance structure through the observations.

$$Y_i = \beta_0 + \beta_1 X_i + U_i \quad (3)$$

Since: $\beta_0$: constant limit parameter, $\beta_1$: is the marginal slope parameter, $U_i$: Random error term which has a Multivariate t- distribution and $U_i \sim (\text{Multivariate})(MV)(0\,, \left[\frac{V_\circ \sigma^2}{V_\circ - 2}\right] I_n)$, When the error term has an MV distribution, the mean is $E(U_i) = 0$. Thus, the estimated formula of the model is as follows:

$$\hat{Y} = E(Y) = b_0 + b_1 X_i + E(U_i) \quad (4)$$

$$E(Y) = b_0 + b_1 X_i \quad (5)$$

Since the error distribution is considered as one of the most crucial terms, through which the process of estimating regression model parameters is determined, when the expectation of the random error was zero.

**5-4. Estimation Methods**: Since the multivariate-distribution analysis of the linear regression model's random error revealed that it was not distributed normally, after that, it could be explained the multivariate distribution and the model of it. It became necessary (for estimating parameters of this model), which would be done through two methods:

**5-4-1. Adjusted Least Square (ALS) Method (Kazem et al., 1988):** The basis of the least squares method was based on calculating the values of the unknown features of the regression model, which make the summation of random squares errors at its least limit. The values of the features calculated in this way are called least squares estimators. Therefore, this work was done to use the modified least squares method in the case of the (MV) error distribution, and to obtain the adjusted Least Squares estimates for the model by the multivariate distribution, it would be as follows:

- Adjusted least squares estimates for (model by the (MV)) distribution)

Since the error term for the dependent variable ($Y_i$) for the model (MVSLM) is as follows:

$$U = Y - E(Y) \quad (6)$$

The Eq. (7) becomes:

$$U_i = Y_i - \beta_0 - \beta_1 X_i \quad (7)$$

**Doi: www.doi.org/10.25130/tjaes.21.72.1.25**

Or it can be written as follows:

| | |
|---|---|
| $U_i = Y_i - X\beta$ | (8) |

Therefore, the sum of squares of errors is:

| | |
|---|---|
| $\sum_{i=1}^{n} {U_i}^2 = \sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2$ | (9) |

By taking the differential of Eq. (9) for both the fixed term parameter and the marginal slope $(\beta_0, \beta_1)$, respectively:

| | |
|---|---|
| $\frac{\partial \sum_{i=1}^{n} {U_i}^2}{\partial \beta_0} = -2 \sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)$ | (10) |

| | |
|---|---|
| $\frac{\partial \sum_{i=1}^{n} {U_i}^2}{\partial \beta_1} = -2 \sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)X_i$ | (11) |

By equating both Eq. (10) and Eq. (11) to zero and simplifying, the estimated formula for $(\beta_0, \beta_1)$ as follows:

| | |
|---|---|
| $\hat{b}_0 = \bar{Y} - b_1 \bar{X}$ | (12) |

| | |
|---|---|
| $\hat{b}_1 = \sum_{i=1}^{n} X_i Y_i - n\bar{X}\bar{Y} / \sum_{i=1}^{n} {X_i}^2 - n\bar{X}^2$ | (13) |

**5-4-2. Maximum Likelihood Method (Visco et al., 1988):** A statistical method of estimating population characteristics from a sample is done by choosing parameter values that maximize the probability that this sample belongs to the-population-from which it was drawn. The maximum likelihood method as considered one of the general and estimation methods, which was used in many research and applications, as it worked on the basis of maximizing the probability density function (Visco et al., 1988). The maximum likelihood method was the most common method for obtaining estimates for regression model parameters when the error was multivariate t-distributed. The random errors in this model are not normally distributed, but rather have a common distribution, which was multivariate t, i.e.:

| | |
|---|---|
| $f(U\|V_\circ,\sigma) = \frac{g(V_\circ)}{(\sigma^2)^{\frac{n}{2}}}[V_\circ + U'U/\sigma^2]^{-(n+V_\circ)/2}$ | (14) |

whereas:

| | |
|---|---|
| $g(V_\circ) = V_\circ^{\frac{n}{2}}\left(V_\circ + \frac{n}{2}\right)/z^{\frac{n}{2}}\Gamma(V_\circ/2)$ | (15) |

Where, $z = \sqrt{2\pi}$. The function (14) represented the multivariate distribution t of the random error vector $U$ which represented the mean and variance respectively:

| | |
|---|---|
| $E(U) = 0$ | (16) |

| | |
|---|---|
| $E(U'U) = (V_\circ\sigma^2/(V_\circ - 2)).trace(\sigma^2.I) = (V_\circ.\sigma^2.p)/(V_\circ - 2)$ | (17) |

$p$: denoted the elements number for $U$, while $trace$ is matrix trace. It is necessary to know the fact that the multivariate distribution function. It was defined in (14), belongs to the (family of continuous mixed normal distributions), which was defined as follows:

| | |
|---|---|
| $f(U\|.) = \int_0^\infty f_N(U\|\tau)f(\tau\|.)d\tau$ | (18) |

whereas:

| | |
|---|---|
| $f_N(U\|\tau) = (2\pi\tau)^{-\frac{n}{2}}exp[-U'U/2\tau^2],\ -\infty < U_i < \infty, i = 1,2,\dots,n$ | (19) |

$f(\tau|.)$: The probability density function represents the parameter (τ), which is defined on the interval$(0 < \tau < \infty)$. If the function $f(\tau|.)$ is chosen as an inverse gamma distribution, the function be as follows:

| | |
|---|---|
| $f(\tau\|V_\circ) = \tau^{-(V_\circ+2+1)}exp[-V_\circ S^2/2\tau],\quad 0 < \tau < \infty$ | (20) |

The term $(V_\circ S^2)$ denoted as quantity in prior of variance for linear model as follows:

| | |
|---|---|
| $V_\circ S^2 = (Y_i - X\beta)'(Y_i - X\beta)$ | (21) |

| | |
|---|---|
| $\hat{\beta} = (X'X)^{-1}(X'Y)$ | (22) |

| | |
|---|---|
| $V_\circ = n - k$ | (23) |

By integrating the formula, the probability function of vector random errors for t-distribution is derived as:

$$f(Y|\beta, V_\circ, \sigma) = g(V_\circ) / (\sigma^2)^{\frac{n}{2}} [V_\circ + (Y - X\beta)'(Y - X\beta)/\sigma^2]^{(n+V_\circ)/2} \quad (24)$$

$g(V_\circ)$: constant normalization, $n$: observations number, $V_\circ$: degree for freedom, and $\sigma^2$:variance. By maximizing Eq. (24), the maximum potential function of joint-likelihood represented as follows:

$$L = \prod_{i=1}^{n} f(Y_i|\beta, V_\circ, \sigma) \quad (25)$$

For the purpose of estimating the parameters of Eq. (24), it is necessary to convert it to linear form by taking the logarithm of both sides and replacing Eq. (7) with what is equal to Eq. (24) to be estimated (β) is as follows:

$$\ln f = \ln g(V_\circ) - (n/2)\ln \sigma^2 - (n + V_\circ)/2 \, . \ln[V_\circ + ((Y - X\beta)'(Y - X\beta)/\sigma^2)] \quad (26)$$

Eq. (26) called as log likelihood formula in the Eq. (24). Taking the differential of Eq. (26) with respect to the parameter (β) we get:

$$\frac{\partial \ln L}{\partial \beta} = 0 = -\left((n + V_\circ)/2\right).1/[V_\circ + ((Y - X\beta)'(Y - X\beta)/\sigma^2)].1 /\sigma^2(-2X'Y + 2X'X\beta) \quad (27)$$

By equaling Eq. (27) to zero, we get:

$$\hat{\beta} = (X'X)^{-1}(X'Y) \quad (28)$$

6. **Particle Swarm Algorithm:** The basic procedures of the Particle Swarm algorithm can be summarized in several steps explained as follows (Khanesar et al., 2007):

A. Setting particles swarm $i = 1,2, \dots, n$ and the setting of the position and velocity of the particles.

B. Evaluate the fitness function of all swarm particles using their current position.

C. Update each individual particle swarm for its best performance:

If $f(X_{i,j}^t) < f(P_{best,i}^t)$ then $P_{best,i}^t = X_{i,j}^t$

**Doi:** www.doi.org/10.25130/tjaes.21.72.1.25

D. Update all particles in the swarm for their best global performance:

If $f(P^t_{best,i}) < f(G^t_{best,i})$ then $G^t_{best,i} = P^t_{best,i}$

E. Update the particle velocity Eq. (29):

$$V^{t+1}_{i,j} = w\,V^t_{i,j} + c_1\,Rand^t_{1,j}\left(P^t_{best,i} - X^t_{i,j}\right) + c_2\,Rand^t_{2,j}\left(G^t_{best,i} - X^t_{i,j}\right) \quad (29)$$

F. Update Particle position Eq. (30)

$$X^{t+1}_{i,j} = X^t_{i,j} + V^{t+1}_{i,j} \quad (30)$$

G. Go to step 2, until the conditions for the algorithm are finished (Khanesar et al., 2007).

**7. Results and discussion:** In this work, real data was relied about the money supply, as there are many variables that affect this variable. In this research, the one variable was taken by consulting some specialists in this field and relying on sources. A money samples were chosen from Central Bank of Iraq according to the available data for the years (2011-2021). That is, for a period of 10 years, as the average for 12 months was taken for all variables and treated as a time series. The dependent variable $(Y)$ represents the money supply and the variable $(X_1)$ represents government deposits. Table 1 and Figure 2 show the data on the money supply and the most important factors affect it.

Table (1): The real data of the money supply and the variables affected it

| Years | Y | $X_1$ |
|---|---|---|
| 2011 | 63597078 | 1171316.25 |
| 2012 | 73775147.08 | 2644099.333 |
| 2013 | 83963663.5 | 3625013.083 |
| 2014 | 89717148.92 | 5104757.25 |
| 2015 | 88763709.08 | 8414022.667 |
| 2016 | 89102140.83 | 8643589.083 |
| 2017 | 90435351.08 | 8750859.5 |
| 2018 | 91602765.5 | 9260048.25 |
| 2019 | 98870119.17 | 9674173.083 |
| 2020 | 110677320.8 | 9075179.583 |
| 2021 | 132759848.9 | 9269803.167 |

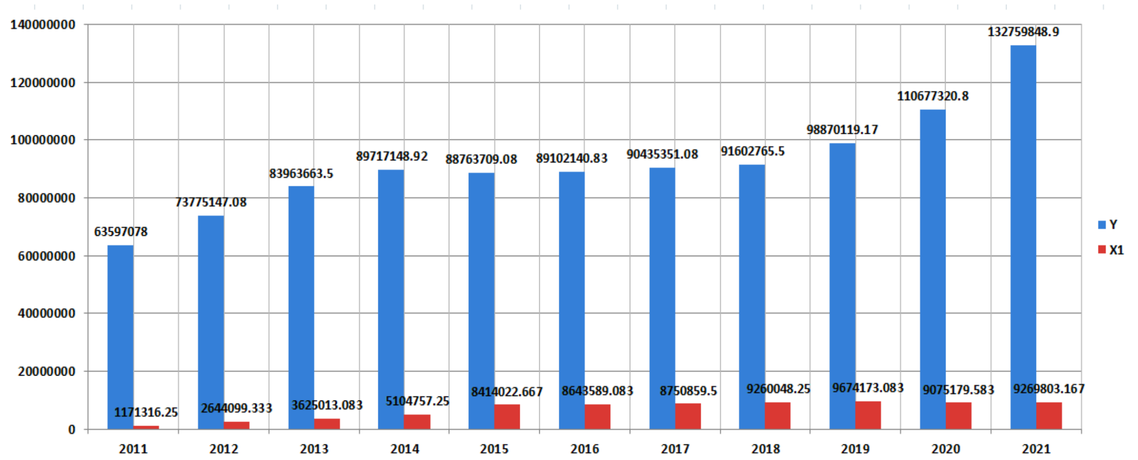


Figure (2): The real data of Y and $X_1$ VS years

Since the distribution of (Y) is the same as the error distribution. The data of this work is found out whether the data for the money supply (Y) follows multivariate t-distribution, since the result of distribution of (Y) is the same results as the error distribution, the data of the dependent variable was tested for real data using a program (Easy Fit) to test goodness of fit, and it was found that the money supply (Y). Table (2) shows Kolmogorov–Smirnov test results which are indicated that they accepted t-distribution as good fitting for our data. This table test did not rejected null of hypothesis.

Table (2): (Kolmogorov_Smirnov choice)

| | | | | | |
|---|---|---|---|---|---|
| **(Size of sample)** | 11 | | | | |
| **(Statistic)** | 1.0 | | | | |
| **(P-Value)** | 0.57 | | | | |
| **(Rank)** | 60 | | | | |
| **(α)** | 0.2 | 0.1 | 0.05 | 0.02 | 0.01 |
| **(Critical Value)** | 0.3082 | 0.3524 | 0.3912 | 0.4367 | 0.4677 |
| **(Reject?)** | No | No | No | No | No |

After the real data was defined and described, the program was used for estimation parameters for multivariate-regression-model using the real data. The MSE is best one of PSO as compared with two estimation methods of MLE and ALS. The following results were obtained as seen in Table 3:

Table (3): Regression estimator comparisons among the three methods (MLE, ALS, and PSO) for real data.

| **MLE** | | | **ALS** | | | **Particle Swarm** | | |
|---|---|---|---|---|---|---|---|---|
| $\hat{\beta}_0$ | $\hat{\beta}_1$ | **MSE** | $\hat{\beta}_0$ | $\hat{\beta}_1$ | **MSE** | $\hat{\beta}_0$ | $\hat{\beta}_1$ | **MSE** |
| 3.9725 | 0.9725 | 0.5949 | 3.9986 | 4.3343 | 0.20035 | 0.8650 | 0.6027 | 0.1859 |

$$\hat{Y}_i = 0.8650 + 0.6027X_1 \quad (31)$$

From Table 3 it has been noted if through the regression equation (31) above that the one explanatory variable (current deposits) has a change and a direct relationship with the dependent variable (money supply). A single unit change in current deposits will result in a change in the money supply of approximately 60% as explained in Table (2). In addition, the process of predicting the dependent variable, which is (money supply), can be performed by using equation by substituting the values of the explanatory variable, whether it is for the following month or the following year.

8. **Conclusion:** Three estimation methods were used in the research for real data taken for 10 years from 2011 to 2021. Two estimation methods called MLE and LSE. Third artificial intelligence estimation method named PSO. The MSE was computed for these three methods. From results of Tables (1-3) and due to t-distribution specifications that explained in introduction section, it can showed that result of MSE was decreasing with the increasing in sample size which indicates that, the validity of the statistical theory. The results were also showed MSE is the best way to estimate the mixed gamma regression model using the Particle Swarm method. Because the data are skewed and positive, larger deviations are penalized more severely when the errors are squared (as in MSE), which is consistent with the characteristics of the gamma distribution. Through the (Kolmogorov-Smirnov) test, it has been noticed that the dependent variable ($Y$) follows the (Multivariate t) (Mt) distribution, which is the same as the distribution of the random error in the regression equation. Through the Table 3 the MSE equals to (0.1859) in PSO method and it was best error as compared to two other methods: MLE and LSE. It has been noticed that the results of the practical application showed that good estimation of the model parameters. In addition to that, it has been concluded that the independent variable has a positive effect and a direct change with the response variable. The t-distribution is much flatter and fatter tails than the normal distribution.
9. **Future work:** For the future work, we seek another artificial intelligence approach to optimize the parameters of linear regression and the dataset it will be changed such as in the following references: (Rens et al., 2021) - (Zahraa et al., 2025).

## References:

1. Al-Birmani, F. A. H. J. (2015). A comparison between genetic algorithm and neural networks in estimating the position of the mediator for nonparametric multivariate t-regression models (Doctoral dissertation, College of Administration and Economics, University of Baghdad, Iraq)
2. Nieto-Barajas, L. E. (2024). Multivariate and regression models for directional data based on projected Pólya trees. Statistics and Computing, 34(1), 25.
3. Denis, D. J. (2018). SPSS data analysis for univariate, bivariate, and multivariate statistics. John Wiley & Sons.P:224.
4. Safiul Haq, M., & Khan, S. (1990). Prediction distribution for a linear regression model with multivariate Student-t error distribution. Communications In Statistics-Theory and Methods, 19(12), 4705-4712.
5. Visco, I. (1988). Advanced Econometric Methods. Journal of Applied Econometrics, 3(2), 165-168.
6. Kazem H., Amore., & Al-Dulaimi, M. (1988). Introduction to linear regression analysis (Doctoral dissertation, College of Administration and Economics, University of Baghdad, Iraq). P:30-34
7. Khanesar, M. A., Teshnehlab, M., & Shoorehdeli, M. A. (2007). A novel binary particle swarm optimization. Mediterranean conference on control & automation, IEEE, 1-6.
8. Lang, M. N., Mayr, G. J., Schlosser, L., Simon, T., Stauffer, R., & Zeileis, A. (2020). Multivariate distributional regression forests for probabilistic nowcasting of wind profiles. In 35th International Workshop on Statistical Modelling, 142.
9. Liu, M., & Bozdogan, H. (2008). Multivariate regression models with power exponential random errors and subset selection using genetic algorithms with information complexity. European Journal of Pure and Applied Mathematics, 1(1), 4-37.
10. Lopez, O. (2011). Nonparametric estimation of the multivariate distribution function in a censored regression model with applications. Communications in Statistics-Theory and Methods, 40(15), 2639-2660.
11. Su, X., Yan, X., & Tsai, C. L. (2012). Linear regression. Wiley Interdisciplinary Reviews: Computational Statistics, 4(3), 275-294.
12. Kunle, A. M., Olowe, S. O., & Oluwafolakemi, F. O. (2014). Impact of foreign direct investment on Nigeria economic growth. International Journal of Academic Research in Business and Social Sciences, 4(8), 234.
13. Roth, M. (2012). On the multivariate t distribution, Linköping University Electronic Press. P:42-47
14. Retno, S., Saputro, Widyaningsih, A. F., & Affan, P. R. C. (2018). Parameter estimation of multivariate multiple regression model using Bayesian with non-informative Jeffreys prior distribution. Journal of Physics: Conference Series, IOP Publishing, 1022(1).
15. Rens v., de Schoot, Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G.& Yau, C. (2021). Bayesian statistics and modelling. Nature Reviews Methods Primers, 1(1), 1.
16. R. M., Leadbetter, Watson, G. S., &. (1963). On the estimation of the probability density, I. The Annals of Mathematical Statistics, 34(2), 480-491.

Doi: www.doi.org/10.25130/tjaes.21.72.1.25

17. Rafael H., Lopez, de Cursi, J. E. S., & Lemosse, D. (2011). Approximating the probability density function of the optimal point of an optimization problem. Engineering Optimization, 43(3), 281-303.
18. Rand Z, Khaleel, M. Z., Khaleel, & Khaleel, H. Z. (2024). Proposed Neural Network Model to Estimate Hemorrhagic Fever Iraqi People Statistics. Applied Sciences Research Periodicals, 2(7), 10-15.
19. Rand Z, Khaleel, H. Z., Khaleel, Abdullah Al-Hareeri, A. A., Mahdi Al-Obaidi, A. S., & Humaidi, A. J. (2024). Improved Trajectory Planning of Mobile Robot Based on Pelican Optimization Algorithm. Journal Européen des Systèmes Automatisés, 57(4).
20. Roaa Ziyad Kadir, & Thaeir Ahmad Alsamman. (2024). Adopting Industry 5.0 Technologies Using the Technical Acceptance Model (TAM): An analytical Study at The University of Mosul. Tikrit Journal of Administrative and Economic Sciences, 20(68, part 2), 118–135. https://doi.org/10.25130/tjaes.20.68.2.7
21. Sundas Ali Khalifa. (2025). Strategies for Enhancing Cybersecurity and Impact on Customer Trust and the Performance of Iraqi Banks. Tikrit Journal of Administrative and Economic Sciences, 21(69, part 1), 1–17. https://doi.org/10.25130/tjaes.21.69.1.1
22. Zahraa Faris Younis Abdel Rahman, & Raghad Mohammed Yahya. (2025). Describing and diagnosing the dimensions of enlightened leadership at the level of humanities and scientific colleges: An analytical study of the opinions of a sample of teachers at the University of Mosul. Tikrit Journal of Administrative and Economic Sciences, 21(69, part 1), 30–45. https://doi.org/10.25130/tjaes.21.69