Al-Kitab Journal for Pure Sciences
ISSN: 2617-1260 (print), 2617-8141(online)
https://isnra.net/index.php/kjps

Bayesian prediction of the stock price rate in the Iraq stock market based on the symmetric heavy tails regression model

Sarmad A. Salih*, Raed Sabeeh Karyakos, Ilham M. Yacoob, Sarah Ghanim Mahmood

Dep. of Mathematics, (E.P.S.) College, University of Al – Hamdaniya, Iraq.

Corresponding Author: sarmadsalih@uohamdaniya.edu.iq

Citation: Salih SA, Karyakos RS, Yacoob IM, Mahmood SG. Bayesian prediction of the stock price rate in the Iraq stock market based on the symmetric heavy tails regression model. Al-Kitab J. Pure Sci. [Internet]. 2024 May. 28 [cited 2024 May. 28];8(1):125-135. Available from: https://doi.org/10.32441/kjps.08.01.p11 .

Keywords: Bayesian prediction, Regression model, generalized multivariate modified Bessel distribution.

Article History

Received	11 Apr.	2024
Accepted	21 May.	2024
Available online	28 May.	2024

Abstract:

In this paper, we investigate the estimation of generalized modified Bessel regression model by using the Bayesian techniques under the assumption that the scale parameter and shape parameters are known. We use the informative priors for estimating of model. Then we derive a prediction distribution of the future response variable _Y_f by using informative priors for predictive future.

Our work applied our results to real data which represent the Iraqi market for securities having taken monthly data for the services sector and of Baghdad sector of Iraq for public transport for the year 2018, as the stock variable rate response variables affecting it are closing price variable, the stock turnover variable. Through the study shows that the explanatory variables are the most important influence on the stock price rate variables through variance inflation factor, the estimated model was appropriate for the data studied.

Keywords: Bayesian prediction, Regression model, generalized multivariate modified Bessel distribution.

التنبؤ البيزي لمعدل سعر السهم في سوق العراق للاوراق المالية بالاعتماد على نموذج الانحدار ذات النهايات الثقيلة

سرمد عبدالخالق صالح*، رائد صبيح قرياقوس، الهام متي يعقوب، سارة غانم محمود

قسم الرياضيات ، كلية التربية للعلوم الصرفة، جامعة الحمدانية ، العراق

sarmadsalih@uohamdaniya.edu.iq, raed_sabeeh@uohamdaniya.edu.iq, ilhammatta@uohamdaniya.edu.iq, sarahghanim@uohamdaniya.edu.iq

الخلاصة:

يهدف البحث الى تقدير نموذج انحدار بسل المحور المعمم بالأسلوب البيزي تحت افتراض ان تكون معلمة القياس ومَعلمات الشكل معلومة. إذ تم استعمال المعلومات السابقة الخبرية لتقدير النموذج. ومن ثم قمنا باشتقاق التوزيع التنبؤي لمتجه مشاهدات متغير الاستجابة المستقبلية Y_f بالاعتماد على معلومات سابقة خبرية لغرض التنبؤ بالمستقبل.

وطبقت النتائج التي تم التوصل إليها على بيانات حقيقية تتعلق بسوق العراق للأوراق المالية إذ اخذت بيانات شهرية مدرجة ضمن قطاع الخدمات والمتمثلة بقطاع بغداد العراق للنقل العام لسنة ٢٠١٨، إذ كان متغير الاستجابة هو معدل سعر السهم والمتغيرات المؤثرة عليه هي متغير سعر الاغلاق ، متغير معدل دوران السهم. ومن خلال الدراسة تبين بان المتغيرات التوضيحية هي من اهم المتغيرات المؤثرة على معدل سعر السهم من خلال عامل تضخم التباين وان النموذج المقدر كان ملائماً للبيانات المدروسة.

الكلمات المفتاحية: التنبؤ البيزي، نموذج الانحدار ، توزيع بسل متعدد المتغيرات المحور المعمم.

1. Introduction:

The concept of the Iraq stock market is economic market enjoy financial and administrative independence that is not related to a set party. It is managed a council consisting of (9) members representing the different economic segments the investment sector, named a council of governors. The market had placed where investors meet, as stock markets dealt with in buying and selling. It constitutes one of the channels through which money flows between individuals, institutions and different sectors, which helps in mobilizing and developing and mobilizing and preparing savings for different investment range. Trading in the market takes place according to the manual method and based on the public bidding system written on boards. A plastic board has been allocated to each listed joint stock company, arranged according to the following sectors: banking sector, insurance sector, investment sector, services sector, industrial sector, hotels sector, tourism, and agriculture sector. Sometimes a multiple regression models are estimated when random errors follow a normal distribution (ND).However, there are cases in which the (ND) of errors is not suitable. Therefore, mixture distributions are more suitable, one of these distributions use such as Generalised Multivariate Modified Bessel (GMMB) distribution. This distribution has applications in variety regions that included displaying stock market data.

The researchers [1] estimated parameters of the multivariate semi-parametric regression model when the error follows a matrix-variate (GMB) distribution using the Bayes method when previous information is not available, in addition to testing hypotheses for the model parameters through the Bayes factor criterion. The theoretical results were applied to experimental data and they concluded that the sample that was used was drawn from a population does not belong to the (GMB) community. The researchers [2] tested statistical hypotheses about the population means of a (GMMB), as well as statistical tests related to the equality of the population means, and they concluded the probability distribution of the statistics (Hotelling-T2) and (Scheffe-T2) for the two tests above respectively and when they have the same covariance matrix. The researchers [3] studied Bayes prediction of a multiple linear regression model in case of equal correlations, as it was assumed that the prior probability function of the variance parameter of the general regression model follows a Generalised Inverse Gaussian (GIG) distribution, predictive distribution was obtained that follows a (GMMB) distribution. This result differs inversely with the t-distribution obtained using the inverse chi-square distribution the prior probability function of the variance parameter, where the predictive distribution was based on interclass correlation. The research covered several sections, where the first section was a general introduction to stock market, and the (GMBR) model was described in the second section, as well as Bayesian estimation and Bayesian prediction. The third section included the application of real data related to the Iraq stock market for the purpose of Bayesian prediction based on the best Bayesian estimator. In the fourth and fifth sections, the most prominent conclusions reached by the research and future studies were presented.edings.pdf.

2. Theoretical side:

2.1 Description of the Generalized Modified Bessel Regression (GMBR) model:

The multiple linear regression models are described by the following linear equation: [4]

$$\underline{Y} = X\underline{\beta} + \underline{u} \qquad (1)$$

Where:

$\underline{Y}$: Response variable with dimension (n$\times$ 1).

X: Non-random matrix of explanatory variables with dimension (n$\times$ p), where p is number of explanatory variables.

$\underline{\beta}$: A vector of regression model parameters with dimension ($p$$\times$ 1).

$\underline{u}$: Vector of random errors has amplitude (n×1), which has a (GMMB) distribution, the probability function for $\underline{u}$ is defined to the following equation: [4]

$$f(\underline{u}) = \frac{\left(\frac{\lambda}{\psi}\right)^{\frac{n}{4}}}{(2\pi\ \sigma^2)^{\frac{n}{2}} K_v\left(\sqrt{\lambda\psi}\right)} \left[1 + \frac{\underline{u}'\underline{u}}{\psi\sigma^2}\right]^{\frac{2v-n}{4}}$$

$$\times\ K_{\frac{2v-n}{2}}\left(\sqrt{\lambda\psi\ \left(1 + \frac{\underline{u}'\underline{u}}{\psi\sigma^2}\right)}\right) \quad -\infty < u < \infty \quad , \sigma^2 > 0 \tag{2}$$

As (λ, ψ, v) it represents the shape parameters, and the field of its parameters is defined as follows: [5]

$$\left.\begin{array}{l} \psi > 0\ \ ; \lambda \geq 0 \quad if\ and\ only\ if\ \ v < 0 \\ \psi > 0\ \ ;\ \lambda > 0 \quad if\ and\ only\ if\ \ v = 0 \\ \psi \geq 0\ \ ;\ \lambda > 0 \quad if\ and\ only\ if\ \ v > 0 \end{array}\right\} \tag{3}$$

$K_v\ (x)$: Modified Bessel function of the third type to order (v), and is defined according to the following equation: [6]

$$K_v(x) = \frac{1}{2}\int_0^{\infty} t^{v-1} e^{\frac{-x}{2}(t+t^{-1})}\ dt \tag{4}$$

The probability distribution of $\underline{u}$ is expressed descriptively by:

$$\underline{u} \sim GMMBD_n\left(\underline{0}\,, \sigma^2 I_n\,, \lambda\,, \psi\,, v\right) \tag{5}$$

Equation (1) represents a combination in vector $\underline{u}$, which follows a (GMMB) distribution. Therefore, [7] deduced the probability distribution of the random vector Y, which follows a (GMMB) distribution. The probability function for the random response Y, is as follows:

$$f(\underline{Y}) = \frac{\left(\frac{\lambda}{\psi}\right)^{\frac{n}{4}}}{(2\pi\ \sigma^2)^{\frac{n}{2}} K_v\left(\sqrt{\lambda\psi}\right)} \left[1 + \frac{\left(\underline{Y} - X\underline{\beta}\right)'\left(\underline{Y} - X\underline{\beta}\right)}{\psi\sigma^2}\right]^{\frac{2v-n}{4}}$$

$$\times K_{\frac{2v-n}{2}}\left(\sqrt{\lambda\psi\ \left(1 + \frac{\left(\underline{Y} - X\underline{\beta}\right)'\left(\underline{Y} - X\underline{\beta}\right)}{\psi\sigma^2}\right)}\right) \tag{6}$$

Descriptively expresses this distribution $\underline{Y} \sim GMBD_n\left(X\underline{\beta}\,, \sigma^2 I_n\,, \lambda\,, \psi\,, v\right)$, for more details see [7].

2.2 Bayesian estimation of a location parameter for the (GMBR) model:

Location parameter of the (GMBR) model was estimated based on informative prior. In this section, we also find on the predictive distribution when an additional sample is available for the response variable Y and the explanatory variables that sample have no relationship to the sample that was used in the estimation process. The Bayesian prediction properties of the future response variable were also studied.

2.2.1 Estimating the location parameter using informative prior:

In this study, **[7]** concluded the estimation of the location parameter for a (GMBR) model and it was as follows:

$$\underline{\hat{\beta}}_B = \underline{\beta}_0 + \left(X'X + P_0^{-1}\right)^{-1} X'X\left(\left(X'X\right)^{-1} X'Y - \underline{\beta}_0\right) \tag{7}$$

This estimator is biased and the amount of bias is equal to:

$$bias\left(\underline{\hat{\beta}}_B\right) = \left(X'X + P_0^{-1}\right)^{-1} P_0^{-1}\left(\underline{\beta}_0 - \underline{\beta}\right) \tag{8}$$

Therefore, the mean square error matrix for $\underline{\hat{\beta}}_B$ is as follows:

$$M\left(\underline{\hat{\beta}}_B\right) = \mathrm{tr}(E\left(\underline{\hat{\beta}}_B - E\underline{\hat{\beta}}_B\right)\left(\underline{\hat{\beta}}_B - E\underline{\hat{\beta}}_B\right)' + \left(bias\left(\underline{\hat{\beta}}_B\right)\right)\left(bias\left(\underline{\hat{\beta}}_B\right)\right)')$$

$$= \mathrm{tr}\ (\sigma^2 D \times \frac{K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{\frac{-1}{2}} + \left(bias\left(\underline{\hat{\beta}}_B\right)\right)\left(bias\left(\underline{\hat{\beta}}_B\right)\right)') \tag{9}$$

$$D = \left(X'X + P_0^{-1}\right)^{-1} X'X\ \left(X'X + P_0^{-1}\right)^{-1} \tag{10}$$

2.2.2 Predictive distribution:

Predictive distribution is an important issue in predictive inference applied in many real-life domains. It represents the probability density function of future observations $\underline{Y}_f$ conditional on a set of current observations Y. **[8] [9]**

After merging the complete posterior distribution of $\underline{\beta}$ conditional on σ^2, τ, determined by the following equation:

$$P\left(\underline{\beta}\middle|\underline{Y}, \sigma^2, \tau\right) = \frac{\left|X'X + P_0^{-1}\right|^{\frac{1}{2}}}{(2\pi\sigma^2\tau)^{\frac{P}{2}}}\ e^{-\frac{1}{2\sigma^2\tau}\left[\left[\left(\underline{\beta}-\underline{\hat{\beta}}_B\right)'\left(X'X+P_0^{-1}\right)\left(\underline{\beta}-\underline{\hat{\beta}}_B\right)\right]\right]} \tag{11}$$

With the probability density function of $\underline{Y}_f$ conditional by the random variable τ and defined according to the following equation:

$$f\left(\underline{Y}_f \middle| \underline{\beta}, \sigma^2, \tau\right) = \frac{1}{(2\pi\sigma^2\tau)^{\frac{n_f}{2}}} e^{-\frac{1}{2\sigma^2\tau}\left[\left(\underline{Y}_f - X_f\underline{\beta}\right)'\left(\underline{Y}_f - X_f\underline{\beta}\right)\right]} \quad (12)$$

We obtain that the predictive distribution of $\left(\underline{Y}_f \middle| \underline{Y}, \tau\right)$ is as follows:

$$f\left(\underline{Y}_f \middle| \underline{Y}, \tau\right) = \int_{\underline{\beta}} f\left(\underline{Y}_f \middle| \underline{\beta}, \sigma^2, \tau\right) \times P\left(\underline{\beta} \middle| \underline{Y}, \tau, \sigma^2\right) \, d\underline{\beta} \quad (13)$$

$$f\left(\underline{Y}_f \middle| \underline{Y}, \tau\right) = \frac{\left|X'X + P_0^{-1}\right|^{\frac{1}{2}} e^{-\frac{\left(\underline{Y}_f - X_f\hat{\underline{\beta}}_B\right)'\left(\underline{Y}_f - X_f\hat{\underline{\beta}}_B\right)}{2\sigma^2\tau}}}{(2\pi\sigma^2\tau)^{\frac{n_f}{2}} \left|X'X + P_0^{-1} + X'_f X_f\right|^{\frac{1}{2}}} \quad (14)$$

Since $\hat{\underline{\beta}}_B$ was previously defined in equation (7), and integrating equation (14) relative to the random variable τ, the predictive distribution of $\underline{Y}_f$ as follows:

$$f\left(\underline{Y}_f \middle| \underline{Y}\right) = \frac{\left|X'X + P_0^{-1}\right|^{\frac{1}{2}} \left(\frac{\lambda}{\psi}\right)^{\frac{n_f}{4}}}{\left|X'X + P_0^{-1} + X'_f X_f\right|^{\frac{1}{2}} (2\pi\sigma^2)^{\frac{n_f}{2}} \, K_v\left(\sqrt{\lambda\psi}\right)} \left[1 + \frac{\Re}{\psi\sigma^2}\right]^{\frac{2v - n_f}{4}}$$

$$\times \; K_{\frac{2v - n_f}{2}}\left(\sqrt{\lambda\psi\left(1 + \frac{\Re}{\psi\sigma^2}\right)}\right) \quad (15)$$

Where $\Re = \left(\underline{Y}_f - X_f\hat{\underline{\beta}}_B\right)' \left(\underline{Y}_f - X_f\hat{\underline{\beta}}_B\right)$.

Since the predictive distribution of $\underline{Y}_f$ is not a common probability distribution, the Bayesian prediction for the response variable Y is found by the following formula:

$$E\left(\underline{Y}_f \middle| \underline{Y}\right) = \int_0^{\infty} \int_{\underline{\beta}} E\left(\underline{Y}_f \middle| \underline{\beta}, \sigma^2, \tau\right) \times \, P\left(\underline{\beta} \middle| \underline{Y}, \tau, \sigma^2\right) P(\tau) \, d\underline{\beta} \, d\tau$$

$$E\left(\underline{Y}_f \middle| \underline{Y}\right) = X_f \left(\underline{\beta}_0 + \left(X'X + P_0^{-1}\right)^{-1} X'X \left(\left(X'X\right)^{-1} X'Y - \underline{\beta}_0\right)\right) \quad (16)$$

3. Applied side:

The Iraqi market during the period between 1992 and 2003 was known as the Baghdad stock market. It was instituted according to Law (24) for the year 1991. This market was a government exchange that was able to list 113 companies (Iraqi and mixed stock), and was able to attract in the last year It has an annual trading rate of over (17,500,000) dollars. The Iraq Stock market was established on (18 April 2004). Temporary Law (74) was issued to establish two important institutions in the capital sector, which Iraq Stock market and Iraq stock Commission. [10]

3.1 Determine basic variables and preparing data:

In this side, practical application will be made on data related to the Iraq Stock market, as monthly data related to the services sector, represented by the Iraq Baghdad for General Transportation sector for the year 2018, will be studied, as the effect of both (Stock turnover rate X1, Closing price X2) which represents the explanatory variables on (Stock price rate Y) which represents the response variable, and **Table 1** shows the measurements for the study variables measured in Iraqi dinars.

Table 1: Measurements of the variables of closing price and stock turnover rate affecting the stock price rate for the year 2018

		.Response V	Elanatory V.'sxp	
No.	Month	Stock price rate	Stock turnover rate	Closing price
1	January	56.710	0.66	51.000
2	February	58.920	0.59	54.100
3	March	55.480	0.45	52.800
4	April	43.460	0.80	42.850
5	May	45.170	0.31	48.000
6	June	40.920	0.08	42.000
7	July	41.320	0.05	41.750
8	August	43.430	0.08	47.000
9	September	45.860	0.02	44.010
10	October	52.210	0.34	57.050
11	November	54.470	0.24	51.380
12	December	43.300	0.76	41.000

Source: http://www.isx-iq.net/isxportal/portal/homePage.html

Before using the (GMBR) model to represent the relationship between variables, the data is time series or not. This was done by drawing Auto-correlation function for stock price rate; it became clear from the drawing that the data is white noise. **Figure 1** shows a plot auto-correlation function for stock price rate variable:

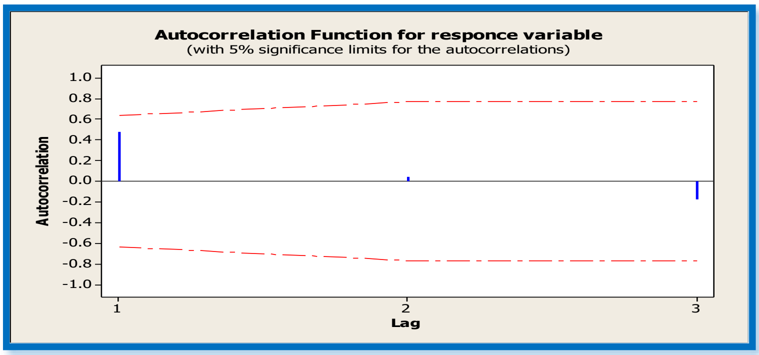

Figure 1: Plot of Auto-correlation function for stock price rate variable

In order to determine the suitability of the data to the model used, the study data was tested through goodness of fit and based on the default shape parameters (λ = 5, ψ = 2, v = 3). The value of the Kolmogorov-Smirnov test extracted (0.2140) for this model was lowerhan the tabulate value $D_{tab}(0.05,12) = 0.2420$, which indicates that the data fit to model used.

Multicollinearity between explanatory (independent) variables was tested based on criterion called the Variance Inflation Factor (VIF). Where both values were less than 10, which indicates that there is no problem of multicollinearity, as shown in the following **Table 2**:

Table 2: Variance inflation factor values for explanatory variables

Variables	VIF
Stock turnover rate	1.0081
Closing price	1.0081

The sample data was divided into two parts, the first consisting of a random sample of size (n=10). This data was used for estimation and the last two observations were used for the purpose of prediction. The parameter vector β was estimated for the sample data with a size of (n=10) by the classical method to choose the initial values, these values were as follows:

$$\underline{\beta}_0 = \begin{bmatrix} 5.9985 \\ 0.9921 \end{bmatrix}, \quad {\sigma_0}^2 = 6.0151, \ a_0 = 25 \ , b_0 = 120 \quad , P_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, P_0 = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$

In this aspect, the location parameter for the (GMBR) model will be estimated based on the shape parameters (λ=5, ψ=2, v=3) and future values will be predicted. The following **Table 3** shows the location parameter estimation using the Bayesian method.

Table 3: Location parameter estimation for the generalized modified Bessel regression model using Bayesian method

Estimator	Bayesian method	
	Informative prior	
	$P_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$P_0 = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$
$\hat{\underline{\beta}}$	$\begin{bmatrix} 6.0024 \\ 0.9752 \end{bmatrix}$	$\begin{bmatrix} 6.2739 \\ 0.9581 \end{bmatrix}$
$MSE(\hat{\underline{\beta}})$	2.6571	4.6210

Table 4 shows the real and estimated values of the stock price rate by using the Bayesian estimator based on informative prior.

Table 4: Real and estimated values for the stock price rate

	Real values	Estimated values
1	56.710	52.5637
2	58.920	54.7972
3	55.480	53.7163
4	43.460	47.3048
5	45.170	47.3010
6	40.920	39.8218
7	41.320	40.9038
8	43.430	42.9570
9	45.860	41.9761
10	52.210	55.9836

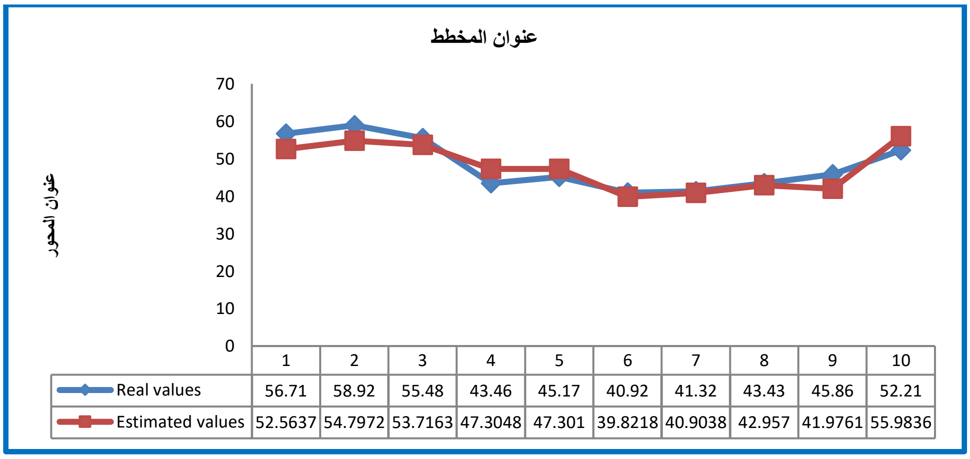

Figure 2: Behavior of the real and estimated values of the stock price rate

Figure 2 shows the real estimated values for **Table 4**. We note from the above figure that the estimated values of the response variable vector have the same pattern as the real values, which indicates that the estimated model was appropriate to the study data.

3.2 Bayesian prediction:

In this study, future values will be predicted based on the best Bayesian estimator. The predictive value represents the predictive mean defined in equation (16).

Table 5: Real and Predictive Values

No.	Real values	Predictive values
1	54.470	50.0124
2	43.300	46.2560

Figure 3 shows the behavior of the real values of the stock price rate variable that was not used in the estimation process and the predictive values.

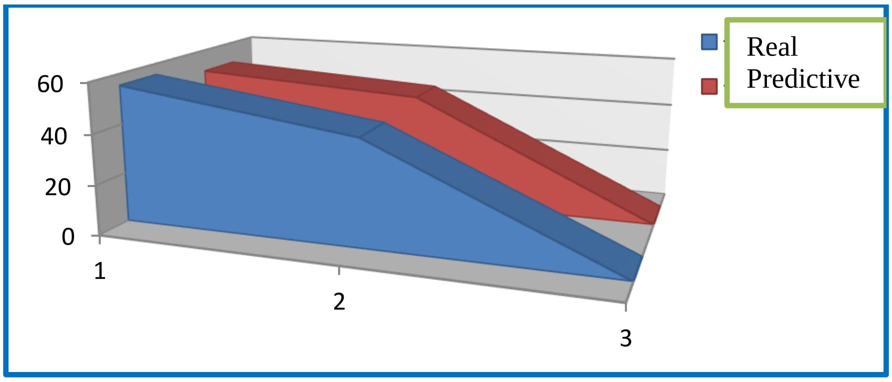

Figure 3: Behavior of the real and predictive values of the stock price rate variable

4. Conclusions:

The researchers arrived at the most important theoretical and practical conclusions, including:

1. The property of linearity is achieved in the (GMB) distribution, and this property similar to case the (ND) and the student-t distribution.

2. When estimating the location parameter using the Bayesian method, the best estimator was the assumed identity matrix.

3. From **Figure 2**, the estimated generalized modified Bessel regression model was appropriate to the study data.

4. Based on the best Bayesian estimator, there is a correspondence between the behavior of the real and predicted values through the (R^2) criterion, the value of which reached (0.91).

5. Recommendations:

1. We recommend expanding the generalized modified Bessel regression model to a multivariate regression model, as this model can be used in principal components analysis and non-normal factor analysis models.

2. We recommend analyzing the generalized modified Bessel regression model when the scale and shape parameters are unknown.

3. We recommend using one of the modern artificial intelligence algorithms and comparing it with the Bayesian method of the generalized modified Bessel regression model.

6. References:

[1] Salih SA., Aboudi EH. Bayesian Inference of a Non normal Multivariate Partial Linear Regression Model. Iraqi Journal of Statistical Sciences. 2021;18(2):51-64.

[2] Thabane L, Drekic S. Testing linear Hypothesis with a generalized multivariate modified Bessel error variable. IIQP Research Report. 2002.

[3] Thabane L, Kibria B. M. Bayesian Prediction for the linear model with Equi-Correlated Responses: An Application of the generalized multivariate modified Bessel distribution. Journal of Applied Statistical Science. 2007:328-335.

[4] Thabane L, Drekic S. Hypothesis testing for the generalized multivariate modified Bessel model. Journal of Multivariate Analysis. 2003 Aug 1; 86(2):360-374.

[5] Butler RW. Generalized inverse Gaussian distributions and their Wishart connections. Scandinavian journal of statistics. 1998 Mar; 25(1):69-75.

[6] Hu W. Calibration of multivariate generalized hyperbolic distribution using the EM algorithm, With applications in risk management, Portfolio optimization and portfolio credit risk. The Florida State University; 2005.

[7] Saieed HA, Salih SA. Estimation and Statistical tests For the Parameters of Multiple Linear Generalized Modified Bessel Regression Model. Iraqi Journal of Statistical Sciences. 2013; 13 (25):257-274.

[8] Jefferys H. The Theory of probability. OuP Oxford; 1998 Aug 6.

[9] Saieed HA, Alobedi SAS. Bayesian analysis of multiple linear generalized modified Bessel regression model. Iraqi Journal of Statistical Sciences. 2014; 14 (26):98-116.

[10] Hamza HK, Abdulhamid GR. The Iraq stock market, its inception- analysis and evaluated of its indicators. Kufa Studies Center Journal. 2012; 1 (24): 255-284.