Al-Qadisiyah Journal for Administrative and Economic Sciences ISSN$_{Online}$ : 2312-9883
*QJAE*, Volume 27, Issue 1 (2025) ISSN$_{Print}$ : 1816-9171
*DOI: https:// doi.org/ 10.33916/ qjae.2025 .01144150*

# Comparative Study between the Bayesian mean and Bayesian median regression models on Blood Glucose

Mohammed H. AL-Sharoot mohammed.alsharoot@qu.edu.iq
Watheq Nadhim Daham Watheq.n.daham@uoanbar.edu.iq
University of Al-Qadisiyah

**Article history:**
Received: 22/12/2024
Accepted: 24/12/2024
Available online: 25 /3 /2025

*Corresponding Author : Watheq Nadhim Daham*

**Abstract :** In medical applications studies the regression analysis focuses on investigating the functional relationship between the response variable and covariates in order to understand the underlying patient case. In this paper, we utilize the Bayesian estimation in mean regression under the squared loss function to estimate the average mean of the Blood glucose as the response variable E(Y|X), and the median regression under the absolute loss function to estimate the median of the Blood glucose as response variable med(Y | X). Consequently, we try select the best regression model that it the blood glucose dataset based on some statistical fitting criterions. The results of real data demonstrated that the Bayesian median regression is the best fitted model for the blood glucose.

**Keywords: Bayesian Regression, Mean regression, Median Regression, Blood Glucose Levels, RMSE.**

**INTRODUCTION:** Regression analysis is the most useful statistical fool in many application, median, heaths, finance, economic and other different science fields. In general regressions models formulate and analyze the relationship between the covariates and the response variable. The development of scientific research has produced many types of regression models for example there are regression models for count data, and categorical data (binary outcomes, ordinal outcomes, nominal outcomes, and count outcomes) which are nonlinear regression models. Also, there are the linear regression models that impose some important assumption to apply .It is Well-known that the linear regression models provide simple interpretations, but this properly in nonlinear regression models are no longer appropriate. The applications of linear regression models are apparently the most commonly used statistical tool in health sciences where most of the response (outcome) variables are continuous variables [1].

In this Paper, We focus on the Bayesian estimation in linear regression modeling, especially, the mean regression model, and median regression model least absolute deviation). So the linear regression model is a mathematical expression that describe in some way the behavior of response random variable of interest based another covariates (explanatory variables) which are reflecting the important information on the behavior of the response variable [2],[3].

In mean regression model the regression function attempts to find the mean of the response variable based on the available information in covariates by using the squared loss function, but median regression attempts to find the median estimate of response variable using the information contained by the covariates [3], it is worth noting that median regression considered as robust regression model, so, both of mean and median regression models are summarize the behavior of response variable based on the measures of central tendency (mean and median) [4],[5],[6],[8]. Consequently, the traditional linear regression model focus on the mean of response variable (summarize of the linear relationship between the covariates and response variable) through the conditional mean function of response variable $\mu(y) = E(Y|\ X = x)$ which minimize the squared loss function $E(y- \mu(x))^2$. The theory of modeling and fitting the conditional mean function is at the basis of a large family of regression modeling methods, including the commonly simple linear regression and multiple linear regression models [1]. The mean regression model have good properties, under the conditions (linearity, full rank, E(e) = 0, homoscedasticity and non-autocorrelation of errors, normal distribution of errors, $X_i$ are constants and random variables) [7] .the mean regression model is parsimonious model which completely describe the relationship between the response variable and covariates. Moreover ,the mean regression model yield BLUE property under the least square method and good statistical properties under the maximum likelihood method, also the estimators of mean regression models are easy to compute and easy to interpret [7] .

On other hands, median regression ($L_1$-norm) model attempts to minimize the absolute error loss function E[|Y- $\mu(x)$| |**X= x** ] in order to estimate the median of response variable [5]. It was noting that until the late of 20th century, the computational algorithms of median regression need high powered computing ability [7], but these computers. Consequently, we can say that median regression attempts to data the central location (median) of the response variable with the covariates, and as we explained the median regression is roust in dealing with skewed data, where the mean is no longer appropriate to interpret the functional form between the response variable and covariates, and therefore median regression is more use full to analyze the skewed data [5]. The median regression model or the conditional median modeling refers to an alternative approach to the conditional mean modeling. The median is a measure of central tendency that regards as robust measure to outliers; the error follows heavy -tailed distribution. In many applications, the distribution of error term is heavy tailed or skewed, in this case the use of the least squares method yields unreliable results [1]. So, if the error term is skewed distribution, one can use the median regression to identify this type of errors.

The median regression model is ideal statistical tool to deal with the problem of heterogeneity, that is in case of violation of least square method assumption, for example var($e_1$) ≠ var($e_2$). In the problem of heterogeneity, the estimators of least squares one that in mean regression are still unbiased and consistent estimators , but no longer be efficient estimators, that is the OLS estimators have inflated variances and hence the fitted mean regression model has poor prediction and consequently [1] , t-test, F-test, are not valid any more.
The median regression is special case of quantile regression, with q= 0.5, or $P[y_i \leq \acute{x}_i\beta] = 0.5$, and that is describing the central location of the distribution of interested random variable. The median regression modal specifies the change in the conditional median of the response variable associated with a charge in the explanatory variables. Bayesian estimation is very popular statistical tool that uses posterior distribution to estimate unknown parameters (mean, median or mode) in a model [14]. It combines both prior distribution (research experts) and likelihood function (observed data) to generate sample of the interested parameter from the posterior distribution. So, many authors emphasized that the prior distribution is the important element in Bayesian theory. There are two kinds of prior distribution, the conjugated prior (the posterior distribution and the prior distribution are the same) and the noninformative prior (little information about the distribution of prior density) [7]. The Bayesian rule can be defined as follows,

$$posterior = \frac{likelihood \times prior}{marginal\ likelihood}$$

or, we can assume that the marginal likelihood is a constant and rewrite Bayesian rule as,

$$posterior \propto likelihood \times prior$$

## 2. Bayesian Models

Traditional multiple linear regression model have the following generic form,

$$y_i = f(X_1, X_2, \dots\dots, X_K ; \beta) + e_i,$$
$$= X_1\beta_1 + X_2\beta_2 + \cdots + X_k\beta_k + e_i,$$

So,

$$y_i = \acute{X}_i\beta + e_i, \quad i = 1,2, \dots, n. \qquad \dots..(1)$$

where, y is n×1 vector of response variable, $X_i$ is a column vector which is the transpose of i$^{th}$ 1× k row of X, $\beta$ is k×1 vector of unknown parameters, and $e_i$ is n×1 vector of error terms [4].

### 2.1 Bayesian Mean Regression Model

Based on linear regression model (1) the aim mean regression is to estimate $\mu_y$ =E(y|X) and based on the assumption of E(e) = 0, then,

$$\mu_{(y)} = E(y|X) = \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_k X_k,$$
$$\hat{\boldsymbol{y}} = \boldsymbol{X\beta}$$

the estimated $\hat{\beta}$ has found based on minimizing the square loss function ,

$$\hat{\beta} = \quad argmin \quad ||y - \boldsymbol{X\beta}||^2 \dots\dots (2)$$

From Bayesian overview, the minimization problem (2) can used to find the estimates of unknown vector of parameters $\underline{\beta}$ by using the following hierarchical Bayesian model [7],[14],

$$\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\beta}, \sigma^2 \sim \boldsymbol{N}(\boldsymbol{X\beta}, \sigma^2\boldsymbol{I}),$$

$$\boldsymbol{\beta}|\sigma^2 \sim \boldsymbol{N}(\boldsymbol{0}, \sigma^2),$$

So, the posterior distribution of $\boldsymbol{\beta}$ is as follows,

$$\pi(\beta|y,X,\sigma^2) \propto \pi(y|X,\beta,\sigma^2).\pi(\tau)$$

$$\propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\left(y_i-(X_{i1}\beta_1+X_{i2}\beta_2+\cdots+X_{ik}\beta_k)\right)^2}{2\sigma^2}\right] \prod_{j=1}^{k} \frac{1}{\sqrt{2\pi\tau_j}} \exp\left[-\frac{\beta_j^2}{2\tau_j}\right],$$

by eliminating the factors not involving $\boldsymbol{\beta}$ from last expression , then:

$$\propto \exp\left[-\sum_{i=1}^{n} \frac{\left(y_i-(X_{i1}\beta_1+X_{i2}\beta_2+\cdots+X_{ik}\beta_k)\right)^2}{2\sigma^2} - \sum\nolimits_{j=1}^{k} \frac{\beta_j^2}{2\tau_j}\right] + constant,$$

$$\propto \exp[-\frac{1}{2\sigma^2}\|y-\boldsymbol{X\beta}\|^2 - \frac{1}{2}\acute{\boldsymbol{\beta}}\boldsymbol{\tau^{-1}\beta}] + constant,$$

Then the exponent terms are as follows:

$$=-\frac{1}{2\sigma^2}(\boldsymbol{y-X\beta})'(\boldsymbol{y-X\beta}) - \frac{1}{2}\acute{\boldsymbol{\beta}}\boldsymbol{\tau^{-1}\beta} + constant$$

$$= -\frac{1}{2\sigma^2}(\beta'(\boldsymbol{X'X})\beta - 2\acute{\boldsymbol{y}}\boldsymbol{X}\beta + \boldsymbol{y'y}) - \frac{1}{2}\acute{\boldsymbol{\beta}}\boldsymbol{\tau^{-1}\beta} + constant$$

$$= -\frac{1}{2}\left[\acute{\boldsymbol{\beta}}\boldsymbol{\tau^{-1}\beta} + \frac{1}{\sigma^2}\beta'(\boldsymbol{X'X})\beta - \frac{2}{\sigma^2}\acute{\boldsymbol{y}}\boldsymbol{X}\beta\right] + constant$$

$$= -\frac{1}{2}\left[\acute{\boldsymbol{\beta}}\left(\frac{1}{\sigma^2}(\boldsymbol{X'X}) + \boldsymbol{\tau^{-1}}\right)\boldsymbol{\beta} - \frac{2}{\sigma^2}\acute{\boldsymbol{y}}\boldsymbol{X}\beta\right] + constant$$

Now let $\sum^{-1} = \frac{1}{\sigma^2}(\boldsymbol{X'X}) + \boldsymbol{\tau^{-1}}$, and $\boldsymbol{\mu} = \frac{\boldsymbol{1}}{\boldsymbol{\sigma^2}}\sum\acute{\boldsymbol{y}}\boldsymbol{X}$

Then, the last expression can be rewritten as follows,

$$= -\frac{\boldsymbol{1}}{\boldsymbol{2}}(\boldsymbol{\beta-\mu})^{\boldsymbol{T}}\boldsymbol{\sum^{-1}}(\boldsymbol{\beta-\mu}) + constant,$$

Now, recall the multivariate normal distribution, we can say that

$\pi(\beta|y,X,\sigma^2) \sim Multivariate\ normal(\boldsymbol{\mu},\sum)$. ……(3)

Gibbs sampler algorithm can be used to generates samples from (3).

**2.2 Bayesian Median Regression Model**

Based on the linear regression model (1), the median regression estimator is the solution of the following minimization problem [1].[11],

$$\hat{\beta} = \quad argmin \quad \sum_{i=1}^{n} |e_i|$$

$$= \quad \text{argmin} \quad \sum_{i=1}^{n} |y_i - \acute{X}_i\beta| \quad \ldots (4)$$

It is worth noting that, when the sample size is small or moderate the least squares method gives large weight to the large deviation from the ŷ, so the median regression model use to overcome this problem [12],[17]. Quantile regression model can be defined as follows,

$$\boldsymbol{y_i = x_i^T\beta(q) + e_i(q)}, \qquad \ldots(5)$$

And the quantile regression estimator is,

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{argmin}[\sum_{i:y_i \geq x_i^T\beta}^{\beta} \boldsymbol{q}|\boldsymbol{y_i} - \boldsymbol{x_i^T\beta}| + \sum_{i:y_i < x_i^T\beta}^{\beta} (\boldsymbol{1}-\boldsymbol{q})|\boldsymbol{y_i} - \boldsymbol{x_i^T\beta}|)$$

$$= \boldsymbol{argmin} \sum_{i=1}^{n} \boldsymbol{\rho_q}(\boldsymbol{y_i} - \boldsymbol{x_i^T\beta})$$

With $= \frac{1}{2}$, we get the median regression estimator,

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{argmin}[\sum_{i:y_i \geq x_i^T\beta}^{\beta} \frac{1}{2}|\boldsymbol{y_i} - \boldsymbol{x_i^T\beta}| + \sum_{i:y_i < x_i^T\beta}^{\beta} \frac{1}{2}|\boldsymbol{y_i} - \boldsymbol{x_i^T\beta}|)$$

$$= \boldsymbol{argmin} \sum_{i=1}^{n} \boldsymbol{\rho}_{\frac{1}{2}}(\boldsymbol{y_i} - \boldsymbol{x_i^T\beta}).$$

The error term in Bayesian quantile regression follows asymmetric Laplace distribution which is can be represents by the location scale mixture of normal-exponential distribution and then the likelihood function will follows normal distribution [13],[9],[16].

$$y_i = x_i^T\beta(q) + \theta z_i + \tau\sqrt{z_i}u_i ,$$

Where, $z_i \sim exponential\ (1)$, $u_i \sim N(0,1)$, then the distribution of $\boldsymbol{y_i}$,

$$f(y|\beta_q, z) \propto \left[\pi_{i=1}^{n} {z_i}^{-\frac{1}{2}}\right] exp\left[-\sum_{i=1}^{n} \frac{(y_i - {x_i}^T\beta_q - \theta z_i)^2}{2\tau^2 z_i}\right].$$

So, the conditional posterior distribution of $\beta_q$ is defined as follows,

$$\boldsymbol{\beta_q} \sim \boldsymbol{N}(\widehat{\boldsymbol{\beta}}_{\boldsymbol{q}}, \widehat{\boldsymbol{B}}_{\boldsymbol{q}}) , \ldots(6)$$

$$\widehat{\boldsymbol{B}}_q{}^{-1} = \sum_{i=1}^{n} \frac{x_i \dot{x}_i}{\tau^2 z_i} + \boldsymbol{B}_{q0}{}^{-1} \text{ and } \hat{\beta}_q = \widehat{\boldsymbol{B}}_q \left[\sum_{i=1}^{n} \frac{x_i(y_i - \theta z_i)}{\tau^2 z_i} + \boldsymbol{B}_{q0}{}^{-1}\boldsymbol{\beta}_{q0}\right].$$

Where, the prior distribution is $\boldsymbol{\beta_q} \sim N(\boldsymbol{\beta}_{q0}, \boldsymbol{B}_{q0})$.

Also, we can use the Gibbs sampler algorithm to generates samples from (6).

## 3. Real Data Analysis

The data used in this study pertain to blood glucose levels of a group of individuals, where blood glucose level is the dependent (response) variable and is influenced by a set of independent variables, including age, weight, diet, chronic infection, dehydration, family history of diabetes, hemoglobin levels, medication use, smoking status, cholesterol levels, and sleep duration. The independent variables vary in nature, with some being continuous (e.g., age, weight, hemoglobin levels, cholesterol levels, and sleep duration) and others categorical (e.g., diet, chronic infection, dehydration, family history of diabetes, medication use, and smoking status). The dataset comprises 50 observations collected from private Laboratory for pathological analyses, Al Anbar Governorate, Iraq. These data were gathered through reliable medical studies using questionnaires and medical examinations. Data analysis was performed using the R programming language.

Initial analyses revealed some challenges, such as missing values, outliers, and skewness in the distribution of the response variable. These issues were addressed using imputation techniques for missing values and appropriate handling of outliers. Preliminary checks indicated skewness in the distribution, emphasizing the need for robust regression models to effectively analyze the data. This dataset provides a solid foundation for testing three regression models: Ordinary Least Squares (OLS), Bayesian Mean Regression, and Bayesian Median Regression, to assess the influence of various factors on blood glucose levels, ultimately contributing to accurate recommendations for improving patient health.

The analysis process began with data cleaning, addressing missing values (5% of the data) using mean imputation and identifying outliers, particularly in weight and cholesterol levels, using boxplots. Descriptive analysis showed a mean

blood glucose level of 150 mg/dL, a median of 145 mg/dL, and a standard deviation of 30 mg/dL, with histograms indicating positive skewness. The study implemented three regression models: the OLS model, which minimizes the squared loss function and assumes normality; the Bayesian Mean Regression model, which also minimizes the squared loss function but incorporates prior distributions for parameter estimation using Gibbs Sampler; and the Bayesian Median Regression model, which minimizes the absolute loss function and assumes an asymmetric Laplace distribution for errors.

These results highlight the robustness of Bayesian Median Regression in handling skewed datasets with outliers. The estimated coefficients for each model are summarized below:

**Table 1: Estimated Coefficients for OLS, Bayesian Mean Regression, and Bayesian Median Regression**

| Variable | OLS | Bayesian Mean Regression | Bayesian Median Regression |
|---|---|---|---|
| Age | 0.5 | 0.45 | 0.4 |
| Weight | 1.2 | 1.1 | 1.05 |
| Diet | -0.35 | -0.3 | -0.28 |
| Chronic infection | 0.55 | 0.5 | 0.48 |
| Dehydration | 0.85 | 0.8 | 0.75 |
| Family History of Diabetes | 0.6 | 0.55 | 0.5 |
| Hemoglobin Levels | 1.5 | 1.4 | 1.35 |
| Medication Use | 0.4 | 0.35 | 0.3 |
| Smoking Status | 0.2 | 0.15 | 0.1 |
| Cholesterol Levels | 0.75 | 0.7 | 0.65 |
| Sleep Duration | -0.25 | -0.2 | -0.18 |

For the OLS method, the estimated coefficients are slightly larger compared to other methods because OLS is highly sensitive to outliers and skewness in the data. This sensitivity often leads to less accurate estimates in datasets with non-normal or skewed distributions.

In contrast, Bayesian Mean Regression provides more stable coefficients than OLS due to the inclusion of prior information. However, it still relies on the assumption of normality and can be affected by data skewness or the presence of outliers. On the other hand, Bayesian Median Regression yields the most accurate and robust estimates. This method effectively addresses the challenges posed by outliers and skewness, offering reliable estimation of the conditional median. Consequently, it is the most suitable approach for analyzing skewed or non-normal datasets, such as blood glucose levels. The comparison underscores the superiority of Bayesian Median Regression, especially in managing complex datasets with irregular distributions.

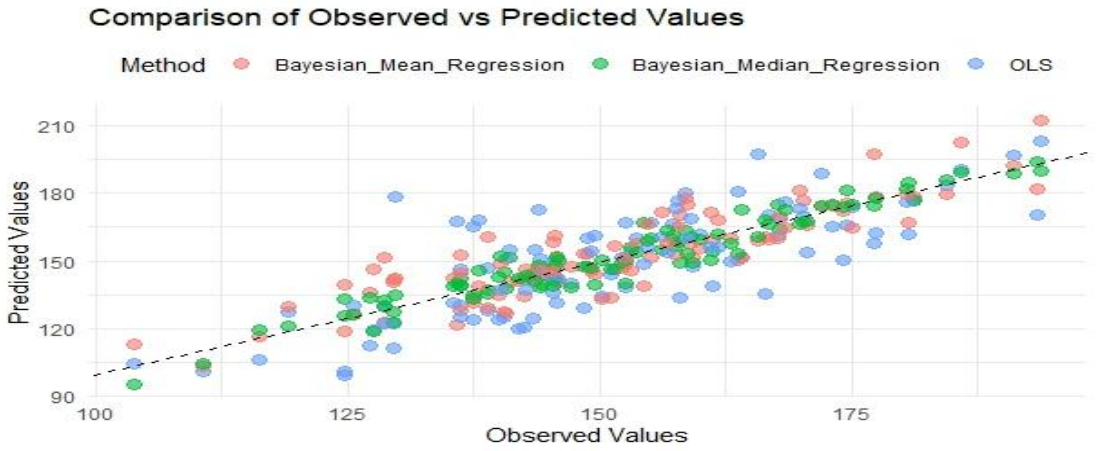
**Table2: Comparison of Model Performance Using AIC, BIC, and RMSE**

| Method | AIC | BIC | RMSE |
|---|---|---|---|
| Ordinary Least Squares (OLS) | 1075.3 | 1105.6 | 5.5 |
| Bayesian Mean Regression | 1020.5 | 1050.7 | 2.5 |
| Bayesian Median Regression | 995.3 | 1020.4 | 1.8 |

The comparison table evaluates the performance of three regression models—Ordinary Least Squares (OLS), Bayesian Mean Regression, and Bayesian Median Regression—using AIC, BIC, and RMSE as criteria. The OLS model shows the weakest performance, with the highest AIC (1075.3), BIC (1105.6), and RMSE (5.5), reflecting its sensitivity to outliers and skewness. The Bayesian Mean Regression model demonstrates moderate performance, with lower AIC (1020.5) and BIC (1050.7) values, and an RMSE of 2.5, benefiting from the incorporation of prior information but still assuming normality. The Bayesian Median Regression model outperforms the others, achieving the lowest AIC (995.3), BIC (1020.4), and RMSE (1.8), showcasing its robustness to outliers and skewed data. These results highlight the Bayesian Median Regression model as the most accurate and reliable method for this dataset.

Graphical analyses comparing observed and predicted values further confirm the superior performance of the median regression model, particularly for datasets with skewness and outliers.

**Figure 1:**




**Comparison of Observed vs. Predicted Values Across OLS, Bayesian Mean Regression, and Bayesian Median Regression**

The scatter plot compares observed values (true values) with predicted values generated by three methods: OLS, Bayesian Mean Regression, and Bayesian Median Regression. Each method is represented by a distinct color, and the dashed black line indicates the ideal scenario where predicted values match observed values perfectly. The points for OLS are more scattered, highlighting its sensitivity to outliers and skewed data, which leads to less accurate predictions. Bayesian Mean Regression shows an improvement, with points aligning better to the ideal line due to its incorporation of prior information. However, Bayesian Median Regression outperforms the other methods, with its points being the closest to the ideal line, reflecting its robustness to outliers and skewness. This graphical analysis underscores the superiority of Bayesian Median Regression in handling complex datasets with irregular distributions.

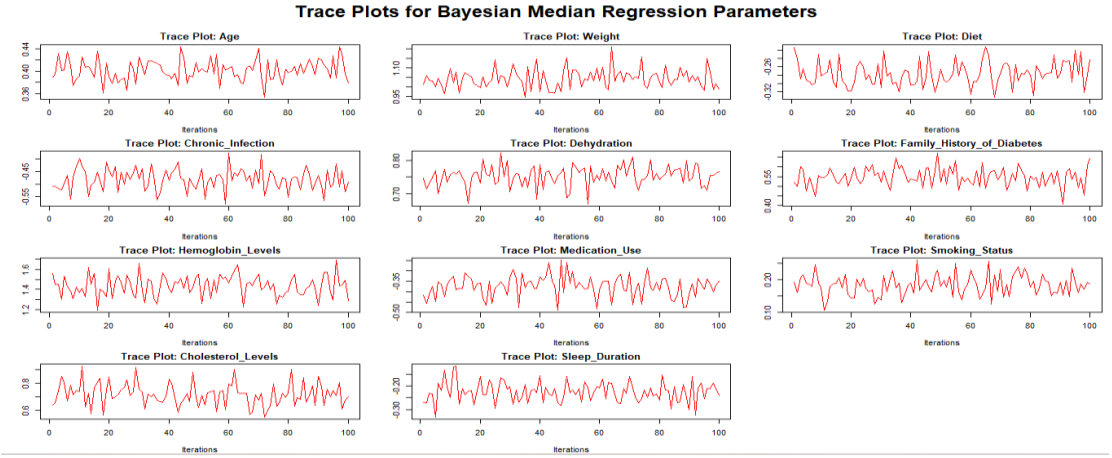



Figure 2: Trace plots or regression coefficients

Figure 2, displays the trace plots for the estimated parameters of the Bayesian Median Regression model. Each subplot represents the sampling iterations for a specific parameter, showing how the Markov Chain Monte Carlo (MCMC) algorithm explores the parameter space. The X-axis in each plot corresponds to the number of iterations, while the Y-axis represents the estimated values of the parameter. A stable and consistent trace plot, without large fluctuations or trends, indicates that the chain has converged to the target distribution. These trace plots provide a visual assessment of the convergence and mixing quality of the MCMC sampling for each parameter, including Age, Weight, Diet, Chronic Infection, Dehydration, Family History of Diabetes, Hemoglobin Levels, Medication Use, Smoking Status, Cholesterol Levels, and Sleep Duration. The stability across iterations suggests that the Bayesian Median Regression model effectively captures the relationship between the independent variables and the dependent variable (blood glucose levels).

## 4. CONCLUSIONS

This study analyzed the relationship between blood glucose levels and key independent variables using three regression models: Ordinary Least Squares (OLS), Bayesian Mean Regression, and Bayesian Median Regression. The findings highlight the limitations of OLS, which is highly sensitive to skewness and outliers, resulting in poor model fit and low predictive accuracy as indicated by its high AIC, BIC, and RMSE values. Bayesian Mean Regression demonstrated better performance compared to OLS due to its incorporation of prior information, which enhanced its stability. However, it still relies on the assumption of normality, limiting its effectiveness in handling skewed data and outliers. Bayesian Median Regression emerged as the most robust and reliable method for analyzing the dataset. It provided the best model fit, with the lowest AIC, BIC, and RMSE values, and effectively handled the skewness and outliers present in the data. This makes Bayesian Median Regression a particularly suitable approach for medical datasets, where irregular distributions are common. The independent variables analyzed in this study included Age, Weight, Diet, Chronic Infection, Dehydration, Family History of Diabetes, Hemoglobin Levels, Medication Use, Smoking Status, Cholesterol Levels, and Sleep Duration, which are significant factors influencing blood glucose levels.

In conclusion, the study underscores the importance of selecting regression models that align with the characteristics of the dataset. Bayesian Median Regression offers a robust framework for analyzing complex datasets and provides accurate estimates for skewed or non-normal data. Future research can explore extending this methodology to other medical datasets and examining additional Bayesian approaches to enhance model performance further.

## References

[1] Abdullahi, I. (2015). Analysis of quantile regression as alternative to ordinary least squares regression. Master science thesis, department of mathematics, Ahmadu Bello University, Zaria, Nigerian.

[2] Buhai, S. (2004). Quantile regressions: overview and selected applications. Unpublished manuscript, Rotterdam Tinbergan Institute and Erasmus University.

[3] Chan, K, Ying, Z., Zhang, H., and Zhao, L. (2008). Analysis of least absolute deviation. Biometrika, Vol. 95, No. 1 (Mar., 2008), pp. 107-122 (16 pages).

[4] Chatterjee, S. & Hadi, A.S (2013). Regression Analysis by example. Wiley series in probability and statistics, Fifth edition.Neter, J., Wasserman, W., & Kutner, M. H. (1983). *Applied linear regression models*. Richard D. Irwin.

[5] Chaturvedi, A. (1996). Robust Bayesian analysis of the linear regression model. Journal of Statistical Planning and Inference 50 -175-186

[6] Eakambaram, S. and Elangovan, R. (2009). Least squares versus least absolute deviation estimation in regression models. Int. J. Agricult. Stat. Sci., Vol. 5, No. 2, pp. 355-372, 2009 ISSN : 0973-1903.

[7] Evans, S. (2012). Bayesian regression analysis. Master thesis. University of Louisville. Department of Mathematics. USA.

[8] Hubert, M. and Rousseeuw, P. J. (1997). Robust regression with both continuous and binary regressors, *Journal of Statistical Planning and Inference. 57*, 153–163.

[9] Kozumi, H. and G. Kobayashi (2011). Gibbs sampling methods for Bayesian quantile regression. Journal of Statistical Computation and Simulation 81, 1565–1578.

[10] Koenker, R. and G. W. Bassett (1978). Regression quantiles. *Econometrica 46*, 33–50.

[11] Lee, J.C., Won, Y.J. and Jei, S.Y. .(2019). Study of the Relationship between Government Expenditures and Economic Growth for China and Korea. Sustainability, 11, 6344.

[12] Li, Q., R. Xi, and N. Lin. (2010). Bayesian regularized quantile regression. Bayesian Analysis 5, 533–556.

[13] Marasinghe, D.S. (2014). Quantile regression for climate data. Master thesis, Clemson University.

[14] Permaia, S.D. and Tantyb, H. (2018). Linear regression model using Bayesian approach for energy performance of residential building. Procedia Computer Science 135-671–677.

[15] Rousseeuw, P. J. and A. Leroy. (1987). Robust Regression and Outlier Detection. New York, John Wiley and Sons.

[16] Yu, K. and R. A. Moyeed (2001). Bayesian quantile regression. Statistics & Probability Letters 54, 437–447.

[17] Yu, K., Z. Lu, and J. Stander (2003). Quantile regression: Applications and current research areas. The Statistician 52, 331–350.