Algoritma: Jurnal Matematika, Ilmu pengetahuan Alam, Kebumian dan Angkasa
Volume. 3 Nomor. 6 November 2025
e-ISSN: 3046-5427; p-ISSN: 3032-6230; Hal 68-84
DOI: https://doi.org/10.62383/algoritma.v3i6.780
Tersedia: https://journal.arimsi.or.id/index.php/Algoritma

Using the Gamma Regression Model in Diagnosing Neonatal Jaundice among Newborns in Diyala Province

Gheada Ibrahim Sheab[1*], Lubna Jaafar Hussein Ibrahim[2], Nawal Muhammad Yaqoub Khalil[3], Sura Anwar Jameel[4]
[1] College of Administration and Economics, University of Diyala, Iraq
[2,3] University of Diyala, Ba'aqubah, Iraq
[4] Middle Technical University, Baqubah Technical Collage, Iraq
Email: gh.gheadaa@uodiyala.edu.iq[1], lubnamoustf24@uodiyala.edu.iq[2], nawal.muhammad@uodiyala.edu.iq[3], sura_anwar@mtu.edu.iq[4]*

**Author Correspondence: gh.gheadaa@uodiyala.edu.iq*

Abstract. This paper applies the Gamma Regression Model to determine the probability of newborns in the Diyala Province developing neonatal jaundice, or hyperbilirubinemia. This disease is prevalent, and it may be severe due to the large amount of unconjugated bilirubin in the blood. The aim of the study was to test the relationship between the levels of bilirubin and a set of independent variables, including the weight of birth, gestational age, and the proportion of red blood cells (PVC). 67 worth of data regarding neonatal cases was collected, and the outcome was that the model fitted well. The findings further indicated that whereas the influence of PVC was significant and positive on the bilirubin level, lower gestational age and less weight at birth had significant negative influence. As the results of the study indicate, the Gamma Regression Model is an effective tool to assess medical data and predict critical scenarios, which assists a clinician with the timely and accurate decision-making.

Keywords: Newborns; Bilirubin; Jaundice; Gamma Regression; Birth weight.

1. INTRODUCTION

Jaundice is one of the most common health complications that infants have to cope with. It is characterized by yellowish eye and skin color caused by the elevated concentrations of bilirubin in the blood. Bilirubin levels that are high may result in significant issues such as brain damage and seizures in case they are not addressed immediately though mild jaundice is normal and usually not a serious problem. When the unconjugated bilirubin level increases in the blood, it is termed as neonatal jaundice, or hyperbilirubinemia. This is the most widespread type of jaundice in newborns and it is generally caused by the immaturity of the liver that is unable to effectively get rid of bilirubin. Therefore, determining blood bilirubin levels, evaluating the severity of the condition, and starting the right treatment all depend on precise diagnostic techniques. In this situation, the Gamma Regression Model is crucial for data analysis and determining the likelihood that newborns will experience neonatal jaundice. Doctors and researchers can use this statistical model to examine the connection between bilirubin levels and a number of risk factors, including gestational age, birth weight, delivery method, and other pertinent variables. This makes it possible to identify high-risk cases more precisely and guarantees prompt intervention with the right kind of treatment.

Naskah Masuk: 07 September 2025; Revisi: 21 Oktober 2025; Diterima: 12 November 2025; Terbit: 21 November 2025

Neonatal jaundice diagnosis and the severity is important as it is predetermined by the presence of unconjugated bilirubin in high blood concentrations that may result in severe complications in case of their unremedied presence. The Gamma Regression Model is a powerful statistical instrument to analyse the bilirubin-related data in newborns and consider numerous independent variables or potential risk factors.

Provided the values of these independent variables, this model can enable researchers and clinicians to estimate the probability of neonatal jaundice. Moreover, it enhances the understanding of correlation between bilirubin levels and other risk factors, which can help medical practitioners make more effective and timely treatment decisions to reduce the probability of incurring severe complications.Therefore, Gamma Regression Model is an effective method of diagnosing neonatal jaundice since it helps to identify the high-risk cases more precisely and ensure timely effective treatment. The research problem is that it is crucially required to have a statistical model, which would help in the early diagnosis of neonatal jaundice and predict the severity. Due to elevated rates of unconjugated bilirubin, neonatal jaundice constitutes a health hazard. This condition may cause brain damage and seizures, which otherwise would not occur. (Ertan et al., 2023)

This research project will study the data on bilirubin concentration and a set of potential risk factors with the help of the Gamma Regression Model of the powerful statistical tools, e.g., gestational age, birth weight, and the delivery method. The objectives are to estimate the probability of neonatal jaundice and to detect those cases which are at a high risk and require urgent medical intervention.

The research is significant in that it discusses neonatal jaundice which is one of the health problems that affect newborn babies seriously. To minimize the risk of the occurrence of serious side effects such as seizures and damage to the brain, the correct and timely diagnosis of this disease is necessary. The Gamma Regression Model enables doctors to be more accurate when making treatment decisions and offering care on time because it allows them to determine the risk of the newborn developing jaundice based on the relevant risk factors.

Consequently, the findings of the present study would contribute to our understanding of the relationship between bilirubin concentration and multiple risk factors, which could be the initial step in the development of new preventive interventions to reduce the occurrence of the condition in the future.

Hypothesis

Null Hypothesis (H_0): The various risk factors, including gestational age, birth weight, and red blood cell concentration (PVC), do not significantly affect a newborn's bilirubin levels.

Alternative Hypothesis (H_1): Newborn bilirubin levels and various risk factors, including gestational age, birth weight, and red blood cell concentration (PVC), are significantly correlated.

Literature Review

Unattended neonatal jaundice might severely affect infant morbidity and mortality, which is why it is one of the most widespread clinical diseases during the early childhood. The diagnosis and management of this disease in question have risen largely due to the new advancements in treatment processes, screening methods, and predictive analytics. By combining genetic, clinical, and laboratory data, machine learning techniques have become more popular for predicting congenital jaundice. With a classification accuracy of 64.34%, Hussein, Mohamad, and Qadir (2024) showed how algorithms like K-Nearest Neighbors (KNN) and Naïve Bayes could increase diagnostic accuracy. These results demonstrate how data-driven models can help with prompt clinical decision-making and interventions, which lowers the chance of complications. The predictive ability of sophisticated machine learning models in identifying infants at risk of jaundice-related mortality, especially in settings with limited resources, was also highlighted by Srivastava, Yajur, and Sujata (2024), highlighting the significance of early detection and intervention.

Additionally, there are encouraging prospects for managing neonatal jaundice with mobile health technologies. A smartphone application for jaundice screening was created and validated by Ngeow et al. (2024), who combined machine learning and the Kramer principle to improve diagnostic reliability. In addition to promoting early detection, this innovation makes remote screening easier, especially in settings with limited resources, increasing access to prompt medical care. In order to lessen dependency on intrusive blood tests, ALdabbagh and Aziz (2024) investigated the application of deep transfer learning techniques on skin images. According to their research, incorporating mobile-based image analysis could greatly increase the accessibility of jaundice severity evaluations for a range of demographics.

There has also been a lot of interest in non-invasive bilirubin measurement methods. In their evaluation of transcutaneous bilirubinometer (TCB) reading accuracy, Zhang, Wu, and Hudak (2024) found that the sternum was a more accurate measurement site than the forehead. They did, however, issue a warning that serum bilirubin testing should be done in addition to TCB, particularly at higher bilirubin levels. The effectiveness of TCB as a non-invasive

evaluation tool was reaffirmed in a more comprehensive review by Badamasi, Ihebuzo, and Nkuma-Udah (2024), which also highlighted the integration of phototherapy solutions and real-time bilirubin monitoring to improve treatment outcomes.

The management of jaundice still heavily relies on treatment methods, especially phototherapy. In their meta-analysis of the effectiveness of various blue light therapy methods, Wu and Wen (2024) offered evidence-based suggestions for clinical procedures. Their results made clear how crucial it is to standardize treatment methods in order to maximize results. Concurrently, Khan et al. (2024) examined G6PD deficiency as a significant contributing factor to neonatal jaundice, specifically in male infants, emphasizing the necessity of early intervention and routine screening in high-risk groups.

Lastly, systemic factors and wider risk factors have also been investigated. While Tampubolon et al. (2024) examined the effects of gender and nutrition on neonatal jaundice, Isaeva et al. (2024) highlighted the significance of prompt diagnosis in primary healthcare systems. These studies advocate for comprehensive approaches to neonatal care by raising awareness of the effects of maternal health and sociodemographic factors on the development of jaundice.

The development of mobile and non-invasive screening tools, the ongoing improvement of treatment protocols, and the incorporation of machine learning into clinical practice for early detection and prediction are the three emerging trends highlighted by the reviewed studies taken together. Future studies should validate mobile technologies across a range of demographics, evaluate the long-term effects of early interventions, and broaden predictive models by adding socioeconomic and environmental factors. When combined, these approaches hold the potential to enhance neonatal jaundice management's accuracy and accessibility in a variety of healthcare settings.

2. METHODOLOGY

The Gamma regression model is a statistical model that considers data, which is gamma-distributed. The premise of this model is that the dependent variable is a continuous and constant value, and it is influenced by a set of independent factors. The Gamma regression model can be applied in the analysis of information in most sectors such as medicine, economics, and engineering because it can analyze data that is skewed and do not follow a normal distribution.

This model also connects the dependent variable and the independent variables with the use of the logarithmic link function. This facilitates the interpretation of results. The Gamma regression model can be used to make a guess about the dependent variable, determine the relationship between the variables, and determine the most pertinent variables which influence the dependent variable.

Gamma Distribution

One of the continuous distributions is the Gamma distribution, which was initially described by Stacy in 1962. Positive random variables can be analyzed with great flexibility using the Gamma distribution, and the values of the variable (Y_i) fall within the interval $(0, \infty)$ Numerous medical specialties regularly use it. The probability density function (pdf) of the random variable (Y_i) is as follows if it has a two-parameter Gamma distribution (Sayed & Sabri, 2023).

$$f(y;\lambda,\alpha) = \frac{\lambda^{\alpha}}{\Gamma_{(\alpha)}}\, y^{\alpha-1} \exp(-\lambda y)\, I_{(0,\infty)} y \ , \qquad \alpha,\lambda > 0$$

Where α is a shape parameter and λ is a scale parameter, so the properties of gamma distribution as following : (Borazan & Altunay, 2025)

a. The Mean of GD

$$E(Y_i) = \frac{\alpha}{\lambda}$$

b. The Variance of GD: (Adekanmbi, 2017)

$$V(Y_i) = \frac{\alpha}{\lambda^2}$$

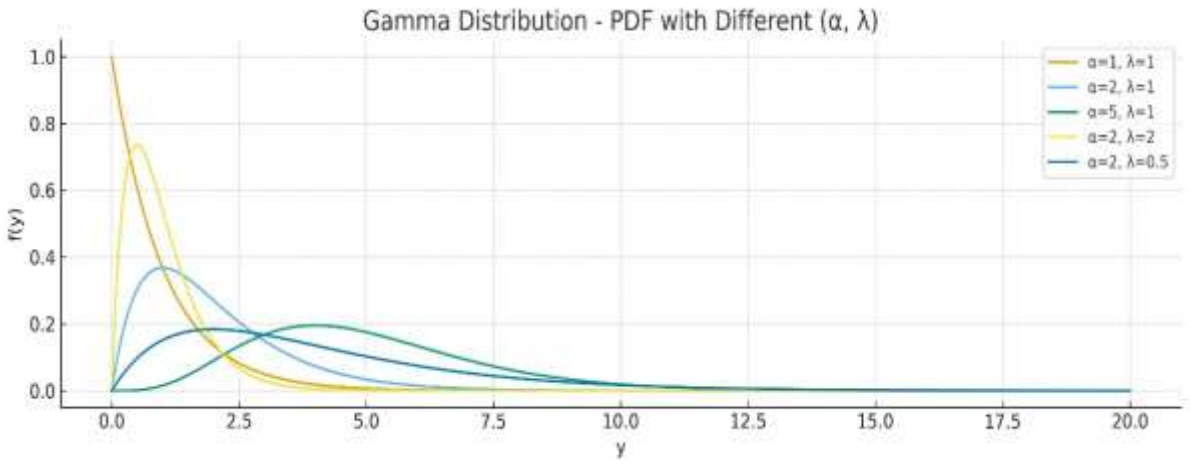

Figure 1. PDF of Gamma Distribution.

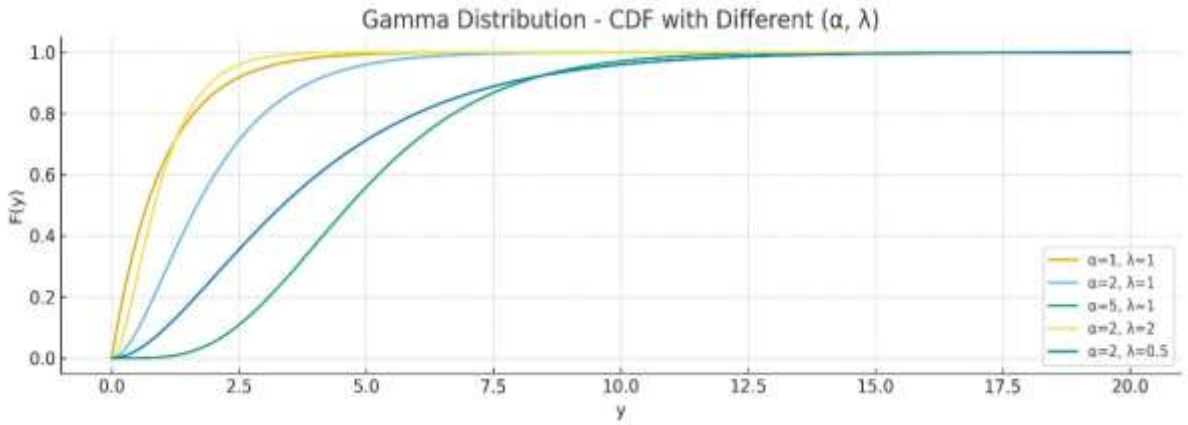

Figure 2. CDF of Gamma Distribution.

Gamma Regression

The framework of Generalized Linear Models (GLM) is thought to be extended by the Gamma Regression Model (GRM). Unlike ordinary linear regression, generalized linear models require the dependent variable's distribution to be part of the exponential family. Additionally, the link function $g(\mu_i) = \eta$, is used in place of the expected values μ_i of the random variable Y, where η is a linear predictor made up of the independent variables. Stabilizing the error term's variance is the link function's main objective. In addition, unlike linear regression, which rigidly assumes a normally distributed error term, the model's error distribution can be freely selected to be independent (Algamal et al., 2024)

The class of Gamma regression models is predicated on the idea that the dependent variable has a Gamma distribution and that a linear predictor and a link function relate the dependent variable's mean to a regression equation. The logarithmic function, the inverse function, or the canonical link could be the link function. Additionally, the model includes the shape parameter, which can either stay constant or be represented by a regression equation that combines a linear predictor with a link function like the logarithm (Aslam et al., 2024).

Among the many empirical uses of the Gamma regression model is the construction of actuarial frameworks in insurance firms. It is also widely used in medical research. For example, it is used to analyze hospital admissions resulting from rare diseases, where time series observations are limited due to the rarity of these events. The model's name comes from the response variable (Y_i) having a Gamma distribution and values limited to the interval $(0, \infty)$ (Seifollahi et al., 2023).

Gamma Regression Models

Let $Y_i \sim G(\mu_i, \alpha)$, where $i = 1,2, \dots, n$ are independent random variables, and α denotes the shape parameter, which in this case is assumed to be constant. Accordingly, the gamma regression model is expressed in terms of the mean of the response variable Y as follows (Asar & Algamal, 2022):

$$\eta_i = g(\mu_i) = x_i'\beta$$

where $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$ is the vector of unknown regression parameters, $x_i = (x_{i1}, \dots, x_{ip})'$ is the vector of p independent variables, η_i represents the linear predictor, and $g(\cdot)$ is the link function. In the gamma regression framework, three types of link functions are commonly used (Adekanmbi, 2017): the logarithmic link function $g(\mu) = \log(\mu)$, the identity link function $g(\mu) = \mu$, and the inverse link function $g(\mu) = 1/\mu$.

In cases where the shape parameter α is not constant but instead depends on a linear predictor, the gamma regression model can be extended to jointly model both the mean and shape parameters of the gamma distribution. Thus, if $Y_i \sim G(\mu_i, \alpha_i)$, the model can be specified as (Dauda et al., 2023).

$$\eta_{1i} = g(\mu_i) = x_i'\beta$$
$$\eta_{2i} = h(\alpha_i) = z_i'\gamma$$

where $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$ and $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_k)'$ are the vectors of regression parameters associated with the mean and shape, respectively. Here, $g(\mu)$ denotes the link function for the mean (typically the identity link), while $h(\alpha)$ represents the link function for the shape parameter (commonly the logarithmic link). The linear predictors η_{1i} and η_{2i} correspond to the systematic components of the mean and shape models, with $x_i = (x_{i1}, \dots, x_{ip})'$ denoting the vector of p covariates and $z_i = (z_{i1}, \dots, z_{ik})'$ denoting the vector of k covariates.

Assumptions of the Model

The following are some of the fundamental presumptions that underpin the Gamma regression model:

1) A Gamma distribution, a continuous probability distribution frequently used to characterize positively skewed data, is assumed to be followed by the dependent variable in the model.
2) After applying the logarithmic link function, it is assumed that the relationship between the independent and dependent variables becomes linear.

3) The dependent variable's value for a particular observation is unaffected by the dependent variable's values for other observations, according to the Gamma regression model's assumption of observational independence.
4) Since multicollinearity can make it difficult to estimate the model parameters, the model makes the assumption that the independent variables are not highly collinear.
5) Before using the Gamma regression model, these assumptions need to be confirmed;

otherwise, the results could be deceptive or erroneous. It might be required to alter the data or use different models that offer a better fit if any of these presumptions are broken (Fidelis et al., 2024).

Estimation Methods

Several approaches have been proposed by researchers for estimating the parameters of the Gamma Regression Model (GRM). In this study, we focus on one of the most widely used and important methods, namely the Maximum Likelihood Method (MLM).

Maximum Likelihood Method

The maximum likelihood method is a fundamental estimation technique due to its wide range of applications in statistical modeling. This method possesses several desirable inferential properties, including consistency, asymptotic efficiency, and near-unbiasedness in many cases. For the Gamma distribution, the likelihood function can be written as follows (Hwang & Jeon, 2024).

$$L = \prod_{i=1}^{n} f(y_i; \mu_i, \alpha_i) \dots (2.12)$$

$$L = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha_i)} \left(\frac{\alpha_i}{\mu_i}\right)^{\alpha_i} y_i^{\alpha_i - 1} \exp\left(-\frac{\alpha_i}{\mu_i} y_i\right)$$

Taking the logarithm yields the log-likelihood function:

$$\log L = \sum_{i=1}^{n} \left\{-\log \Gamma(\alpha_i) + \alpha_i \log\left(\tilde{i}^{\,i} \frac{y_i}{\mu_i}\right) - \log(y_i) - \frac{\alpha_i}{\mu_i} y_i\right\}$$

where $\mu_i = x_i'\beta$ and $\alpha_i = \exp(z_i'\gamma)$. The score functions are obtained as:

$$\frac{\partial L}{\partial \beta_j} = \sum_{i=1}^{n} \left(-\frac{\alpha_i}{\mu_i}\right)\left(1 - \frac{y_i}{\mu_i}\right) x_{ij}, j = 1, \dots, p \dots (2.15)$$

$$\frac{\partial L}{\partial \gamma_k} = \sum_{i=1}^{n} \left\{-\alpha_i \left[\frac{d}{d\alpha_i} \log \Gamma(\alpha_i) - \log\left(\frac{\alpha_i y_i}{\mu_i}\right) - 1 + \frac{y_i}{\mu_i}\right]\right\} z_{ik}, k = 1, \dots, r$$

To evaluate second-order conditions, the Hessian matrix-containing all possible second-order partial derivatives-is employed (Bossio, 2015). Using the Fisher Information Matrix, the variance-covariance matrix of the maximum likelihood estimators is derived as (Yasin et al., 2024)

$$I(\beta) = \begin{bmatrix} -E\left(\frac{\partial^2 L}{\partial \beta_k \partial \beta_j}\right) & -E\left(\frac{\partial^2 L}{\partial \gamma_k \partial \beta_j}\right) \\ -E\left(\frac{\partial^2 L}{\partial \gamma_k \partial \beta_j}\right) & -E\left(\frac{\partial^2 L}{\partial \gamma_k \partial \gamma_j}\right) \end{bmatrix}$$

It is observed that the Fisher Information Matrix is block-diagonal, where one block corresponds to the mean regression parameters (β) and the other to the shape regression parameters (γ). This structure implies that the maximum likelihood estimators $\hat{\beta}$ and $\hat{\gamma}$ are asymptotically independent. Consequently, estimation of GRM parameters cannot generally be achieved using standard closed-form solutions; instead, iterative algorithms are required. One such method is the Fisher Scoring Algorithm, an iterative procedure analogous to the Newton-Raphson method or Iteratively Weigh、 Least Squares (IWLS). This algorithm replaces the Hessian with its expected value and updates the estimates of β and γ iteratively until convergence (Gonçalves et al., 2023).

The updating rules are given as:

$$\hat{\beta}^{(k+1)} = \left(X' W_1^{(k)} X\right)^{-1} X' W_1^{(k)} Y$$
$$\hat{\gamma}^{(k+1)} = \left(Z' W_2^{(k)} Z\right)^{-1} Z' W_2^{(k)} Y \quad \ldots (2.25)$$

where $W_1^{(k)}$ and $W_2^{(k)}$ are diagonal weight matrices with elements defined in terms of μ_i and α_i. The algorithm proceeds as follows:

1. Set the iteration counter $v = 0$.
2. Initialize parameter values for β and γ.
3. Update β using Equation (2.25).
4. Update γ using Equation (2.27).
5. If $\left|\hat{\beta}^{(k+1)} - \hat{\beta}^{(k)}\right| < 0.00001$ and $\left|\hat{\gamma}^{(k+1)} - \hat{\gamma}^{(k)}\right| < 0.00001$, stop; otherwise, set $v = v + 1$ and return to Step 3 (Bossio, 2015).

This iterative Fisher scoring approach provides consistent and efficient estimates for the Gamma Regression Model parameters.

3. RESULTS AND DISCUSSIONS

Descriptives of Data and Testing

Unquestionably, a key component of achieving trustworthy and dependable results in any research project is the quality of data appropriate for the applied aspect on the one hand, and its accuracy on the other. The variables of newborn bilirubin levels and several risk factors, such as gestational age, birth weight, and packed red blood cells (PVC), were among the data gathered from Baqubah Teaching Hospital. The dataset, which included 67 observations, was chosen to be used to carry out the study's methodology since it contained the necessary relationships. The response variable (Y) represented the newborns' bilirubin levels, while the independent variables were gestational age (x1), birth weight (x2), and packed red blood cells (PVC) (x3). Plotting the response variable's histogram and using a goodness-of-fit test revealed that the response variable did not fit the normal distribution but did fit the Gamma distribution well.

Table 1. Goodness-of-Fit Test Results for Different Distributions of the Response Variable (Bilirubin Levels).

Distribution	Sample Size	Statistic	P-Value
Normal	67	0.18245	0.04215
Exponential	67	0.14322	0.08742
Weibull	67	0.12187	0.16359
Lognormal	67	0.11034	0.24891
Gamma	67	0.09716	0.52035

The Goodness-of-Fit Test results for the response variable, neonatal bilirubin levels, under various probability distributions are shown in Table (1). The adequacy of fit was evaluated using the Kolmogorov–Smirnov test, where the p-value denotes the level of significance and the statistic measures the maximum deviation between the empirical and theoretical distribution functions. With a p-value of 0.04215 and a test statistic of 0.18245 for the normal distribution, these results fall below the standard cutoff point of $\alpha = 0.05$. This finding implies that the bilirubin levels are not distributed normally. the test statistic for the exponential distribution was 0.14322, and the p-value was 0.08742. Even though the p-value is higher than the normal distribution's, it still suggests a less satisfactory fit than the others. With a p-value of 0.16359 and a lower test statistic (0.12187), the Weibull distribution demonstrated moderate suitability and a better fit. With a p-value of 0.24891 and a statistic of 0.11034, the lognormal distribution further enhanced the fit and suggested an even greater correspondence with the data.

Lastly, the highest p-value (0.52035) and the smallest test statistic (0.09716) indicated that the Gamma distribution had the best fit. This demonstrates unequivocally that the Gamma distribution of bilirubin levels in neonates is more appropriate than the other distributions that were tested.

The Gamma distribution is the best model for the response variable (bilirubin levels), according to the Goodness-of-Fit results. This offers the statistical underpinnings for using the Gamma regression model in the analysis that follows.

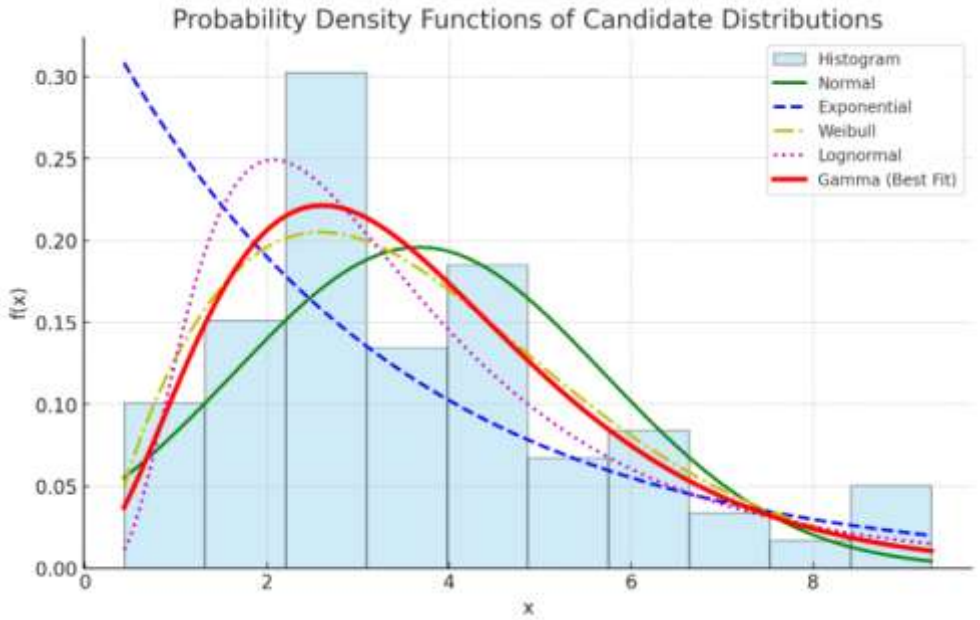

Figure 3. Histogram of Bilirubin Levels with Fitted Probability Density Functions of Candidate Distributions.

The histogram of the response variable (newborns' bilirubin levels) and the fitted probability density functions of five potential distributions—Normal, Exponential, Weibull, Lognormal, and Gamma are shown in figure (3). The Weibull and Lognormal distributions offer a closer alignment with the data, whereas the Normal and Exponential distributions exhibit pronounced departures. The best overall fit, however, is shown by the Gamma distribution (red solid line), which closely resembles the histogram's shape throughout the whole data set. The Gamma distribution is confirmed to be the best model for the bilirubin levels by this graphical evidence, which also supports the goodness-of-fit test results.

Table 2: Descriptive Statistics of Study Variables.

Variables		N	Minimum	Maximum	Mean	Std. Deviation
Dependent Variable	y	67	5.5	16.8	9.670	3.5031
Covariates	x1	67	1.3	3.0	2.264	0.5629
	x3	67	0.35	0.81	0.5775	0.16834
	x2	67	2	7	4.09	1.323

Table (2) provides a summary of the descriptive statistics for the continuous variables used in this investigation. The outcome of interest that the regression model seeks to forecast or explain is represented by the dependent variable, y. With a mean of 9.670 and a standard deviation of 3.5031, the observed values of y fall between 5.5 and 16.8, suggesting that they are somewhat distributed around the mean.

Potential predictors that could affect the dependent variable are the independent variables (x1, x2, and x3). In particular, x1 has a mean of 2.264 and a standard deviation of 0.5629, ranging from 1.3 to 3.0. While x2 ranges from 2 to 7, with a mean of 4.09 and a standard deviation of 1.323, variable x3 displays values between 0.35 and 0.81, with a mean of 0.5775 and a standard deviation of 0.16834.

Greater dispersion is indicated by higher standard deviations, which show how variable the values are around their respective means. The total number of observations used in the analysis is represented by the sample size (N = 67). Before performing regression analysis, these descriptive statistics offer crucial information about the properties and distributions of the variables.

Goodness-of-Fit Test for the Gamma Regression Model

The Deviance test is used to evaluate the model's goodness-of-fit. The null hypothesis, which holds that the fitted model accurately depicts the data, is compared to the alternative hypothesis, which holds that the model does not fit the data well, using the Chi-square statistic. Table (3) provides a summary of the findings.

Table 3. Model Significance and Goodness-of-Fit Statistics.

Goodness of Fit[a]			
	Value	df	Value/df
Deviance	0.238	63	0.004
Scaled Deviance	67.040	63	
Pearson Chi-Square	0.236	63	0.004
Scaled Pearson Chi-Square	66.545	63	
Log Likelihood[b]	-54.055-		
Akaike's Information Criterion (AIC)	118.111		
Finite Sample Corrected AIC (AICC)	119.094		
Bayesian Information Criterion (BIC)	129.134		
Consistent AIC (CAIC)	134.134		

It is evident from Table (3) that the data is well-fitted by the Gamma regression model. The model is in good agreement with the observed data, as demonstrated by the small values of Deviance and Pearson Chi-Square when divided by their respective degrees of freedom. Furthermore, the comparatively low values of information criteria like AIC and BIC point to a suitable trade-off between simplicity and model fit. The fitted model performs better than a straightforward reference model, as indicated by the negative Log Likelihood. Overall, the results confirm that the Gamma regression model fits the data under study, even though the AICC and CAIC values are marginally higher because of the small sample size.

Significance Test of the Gamma Regression Model

The Chi-square test is used to assess the Gamma regression model's overall significance, as indicated in Table (4).

Table 4. Omnibus Test.

Omnibus Test[a]		
Likelihood Ratio Chi-Square	df	Sig.
239.185	3	0.000

The results of the Omnibus Test, sometimes referred to as the likelihood ratio test, are shown in Table (4) and can be explained as follows:

With three degrees of freedom, the Likelihood Ratio Chi-Square value is 239.185, which is comparatively high. The corresponding p-value (Sig.) is 0.000, significantly less than the

standard significance level of 0.05. This finding suggests that at least one of the model's parameters—that is, the independent variables' estimated coefficients—deviates noticeably from zero. To put it another way, the model has explanatory power overall. The null hypothesis, according to which all model coefficients (apart from the intercept) are equal to zero, is rejected by the likelihood ratio test. This demonstrates that the dependent variable and the independent variables in the model have a statistically significant relationship. Consequently, this result is encouraging and shows that the used Gamma regression model fits the data with statistical and overall significance.

Estimation of Model Parameters

The Maximum Likelihood Estimation (MLE) method was used to estimate the parameters of the Gamma regression model. Table (5) displays the estimated coefficients and the significance tests for them.

Table 5. Parameter Estimates of the Gamma Regression Model.

Parameter Estimates							
Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	1.499	0.0600	1.382	1.617	624.877	1	0.000
x2	-0.036-	0.0057	-0.047-	-0.025-	38.770	1	0.000
x1	-0.088-	0.0158	-0.119-	-0.056-	30.530	1	0.000
x3	1.825	0.0520	1.723	1.927	1229.922	1	0.000
(Scale)	0.004[a]	0.0006	0.003	0.005			

From Table (5), the estimated regression model can be expressed as follows:

$$\widehat{y_i} = \mathrm{e}^{(x_i\underline{\widehat{\beta}})} = exp(1.499 - 0.088X_1 - 0.036\mathrm{X}_2 + 1.825X_3)$$

To test the significance of the parameters, the following hypotheses were formulated:

$$H_0: \beta_j = 0 \quad vs \quad H_1: \beta_j \neq 0 \quad , j = 1,2,3$$

The estimated coefficients can be interpreted as follows:

The expected value of the dependent variable when all independent variables are zero is represented by the intercept (β_0). It has a highly significant estimate of 1.499 ($p < 0.001$). the dependent variable is negatively impacted by the coefficient of x2 (birth weight), which is -0.036. In particular, the expected bilirubin level in neonates drops by 0.036 units for every unit increase in birth weight. There is statistical significance in this effect.

The coefficient for x1 (Child Age) is -0.088, indicating a negative correlation with bilirubin levels. A statistically significant drop of 0.088 units in bilirubin occurs for every unit increase in age. with a positive coefficient of 1.825, x3 (Red Blood Cell Percentage, PVC) shows the biggest and most significant effect. The expected bilirubin level rises by 1.825 units for every unit increase in x3. This effect is very important.

Furthermore, the estimated scale parameter, which represents the variance of the model errors, is 0.004. In order to shed light on the accuracy of the estimates, the table additionally displays the 95% Wald confidence intervals for every parameter.

4. CONCLUSIONS

In the current study, risk factors for neonatal jaundice (hyperbilirubinemia) in the Diyala Governorate were investigated using the Gamma regression model. The Deviance test, AIC, and Chi-square statistics all support the findings, which show how well the model fits the data and how robust and dependable it is at capturing changes in bilirubin levels.

Several predictors had significant effects on bilirubin levels, according to the analysis. Red blood cell percentage (PVC) demonstrated the strongest positive influence, making it the most important factor among those examined, while lower birth weight and younger age were linked to higher bilirubin concentrations.

Additionally, the model was successful in anticipating high-risk neonatal jaundice cases, promoting prompt clinical interventions, and lowering the risk of serious complications. These results imply that the Gamma regression model can be a useful instrument for data-driven decision-making in neonatal care, enabling early monitoring and diagnosis.

All things considered, the research shows that the Gamma regression model is a reliable and useful clinical framework for comprehending the causes of neonatal jaundice and directing evidence-based medical decisions.

REFERENCES

ALdabbagh, B., & Aziz, M. H. (2024). Neonatal jaundice severity detection from skin images using deep transfer learning techniques. Journal of Electronics, Electromedical Engineering, and Medical Informatics. https://doi.org/10.35882/jeeemi.v7i1.576

Algamal, Z. Y., Salih, A. M., & Khaleel, M. A. (2024). A new ridge-type estimator for the gamma regression model. Iraqi Journal for Computer Science and Mathematics, 5(1). https://doi.org/10.52866/ijcsm.2024.05.01.006

Asar, Y., & Algamal, Z. (2022). A new two-parameter estimator for the gamma regression model. Statistics Optimization & Information Computing, 10(3), 750–761. https://doi.org/10.19139/soic-2310-5070-822

Aslam, M. Z., Amin, M., Mahmood, T., & Akram, M. N. (2024). Shewhart ridge profiling for the Gamma response model. Journal of Statistical Computation and Simulation, 94(8), 1715–1734. https://doi.org/10.1080/00949655.2023.2299354

Badamasi, S., Ihebuzo, N. G., & Nkuma-Udah, K. I. (2024). Exploring the fundamentals of bilirubinometry: Methods of assessment, timely monitoring, and intervention in neonatal jaundice with integrated phototherapy solutions—A review. International Journal of Engineering and Science. https://doi.org/10.9790/1813-1310104117

Borazan, M., & Altunay, S. A. (2025). Gamma regression for the poverty rate in Türkiye. Ekoist Journal of Econometrics and Statistics, 42, 22–35. https://doi.org/10.26650/ekoist.2025.42.1523478

Dauda, K. A., Lamidi, R. K., Dauda, A. A., & Yahya, W. B. (2023). A new generalized gamma-Weibull distribution with applications to time-to-event data. bioRxiv. https://doi.org/10.1101/2023.11.18.567670

Ertan, E., Erkoç, A., & Akay, K. U. (2023). A new Liu-type estimator for the gamma regression model. Communications in Statistics - Simulation and Computation, 1–18. https://doi.org/10.1080/03610918.2023.2220999

Fidelis, C. R., Ortega, E. M. M., Prataviera, F., Vila, R., & Cordeiro, G. M. (2024). Reparametrized generalized gamma partially linear regression with application to breast cancer data. Journal of Applied Statistics, 51(15), 3248–3265. https://doi.org/10.1080/02664763.2024.2337086

Gonçalves, J. H. D., Gomes, J. J. F., Rubio, L., & Ramos, F. R. (2023). A generalized log gamma approach: Theoretical contributions and an application to companies' life expectancy. Mathematics, 11(23), 4792. https://doi.org/10.3390/math11234792

Hussein, M. M. F., Mohamad, S. H., & Qadir, B. H. (2024). Machine learning approaches to classify and predict congenital jaundice. Khazain for Economic and Administrative Sciences. https://doi.org/10.69938/keas.2401025

Hwang, E., & Jeon, C. (2024). Nonnegative GARCH-type models with conditional gamma distributions and their applications. Computational Statistics & Data Analysis, 198, 108006. https://doi.org/10.1016/j.csda.2024.108006

Isaeva, G. A., Seisebaeva, R. Z., & Nurgalieva, J. J. (2024). Optimization of diagnosis of jaundice syndrome in infants under one year at primary health care level: Literature review. Reproduktivnaâ Medicina. https://doi.org/10.37800/rm.3.2024.36-41

Khan, I., Nawaz, S., Ullah, I., et al. (2024). Frequency, severity & outcome of G6PD deficiency among male newborns presenting to neonatal unit with neonatal jaundice. South Eastern European Journal of Public Health. https://doi.org/10.70135/seejph.vi.2350

Ngeow, A. J. H., Moosa, A. S., Tan, M. G., et al. (2024). Development and validation of a smartphone application for neonatal jaundice screening. JAMA Network Open. https://doi.org/10.1001/jamanetworkopen.2024.50260

Sayed, A. I. A., & Sabri, S. R. M. (2023). Generalized gamma distribution based on the Bayesian approach with application to investment modelling. International Journal for Simulation and Multidisciplinary Design Optimization, 14, 10. https://doi.org/10.1051/smdo/2023011

Seifollahi, S., Bevrani, H., & Kamary, K. (2023). Inequality restricted estimator for gamma regression: Bayesian approach as a solution to multicollinearity. arXiv. https://doi.org/10.48550/arxiv.2303.05120

Srivastava, M., Yajur, & Sujata. (2024). Prediction of neonatal mortality from jaundice using machine learning. CRC Press. https://doi.org/10.1201/9781003477280-10

Tampubolon, C. H., Wija, I. B. E. U., & Pasaribu, T. R. (2024). Description of risk factors affecting neonatal jaundice. International Journal of Science and Healthcare Research. https://doi.org/10.52403/ijshr.20240424

Wu, R., & Wen, L. (2024). Meta-analysis of the efficacy of different blue light therapy methods for neonatal jaundice. Journal of Maternal-Fetal & Neonatal Medicine. https://doi.org/10.1080/14767058.2024.2430649

Yasin, H., Purhadi, N., & Choiruddin, A. (2024). Spatial clustering based on geographically weighted multivariate generalized gamma regression. MethodsX, 13, 102903. https://doi.org/10.1016/j.mex.2024.102903

Zhang, E., Wu, T. J., Hudak, M. L., et al. (2024). Optimal site for applying transcutaneous bilirubinometer as an outpatient screening tool for neonatal jaundice: A comparison between the sternum and forehead. Frontiers in Pediatrics. https://doi.org/10.3389/fped.2024.1446524