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.2025.018994

Analysis of Infant Obesity by Using Bayesian Regression Analysis

Mohammed H. AL-Sharoot Ammar H. Taha
mohammed.alsharoot@qu.edu.iq amarht@ntu.edu.iq
University of Al-Qadisiyah

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

Corresponding Author : Ammar H. Taha

Abstract : The phenomenon of obesity in newborns, though relatively rare, poses significant health challenges, necessitating the identification of key contributing factors to aid in health awareness and preventive strategies. This study employs advanced regression techniques, including Lasso, Bayesian Lasso, Adaptive Lasso, and Bayesian Adaptive Lasso, to analyze 13 predictors influencing infant obesity. These methods address challenges such as multicollinearity, sparsity, and high dimensionality, with Bayesian Adaptive Lasso emerging as the most effective. By achieving superior predictive accuracy and interpretability, this method excels in isolating non-influential predictors while providing credible intervals for uncertainty quantification. The findings highlight the robustness of Bayesian approaches, particularly Bayesian Adaptive Lasso, in analyzing medical data, providing a foundation for future interventions in neonatal health.

Keywords: Lasso , Bayesian Lasso, Adaptive Lasso , Bayesian Regression , newborn Obesity Analysis

INTRODUCTION: Obesity in newborns, though uncommon, presents significant challenges to neonatal health and development. Understanding its underlying factors is crucial for raising awareness and designing effective preventive measures, World Health Organization. (2020). Many medical phenomena, including infant obesity, can be studied using regression models to examine the relationships between predictor variables and outcomes. However, traditional regression techniques often face limitations, especially with small sample sizes and high-dimensional datasets that are common in medical research. To overcome these challenges Bayesian regression has proven as effective methodology to parameter estimation. Bayesian regression is a statistical technique that combines the prior information about the parameters of linear regression model and updates these prior based on observed data. So, Bayes' rule is used to combine prior distributions with data (likelihood) function to obtain the posterior distributions for model of interested parameters, see Gelman et al. (2008) and Spiegelhalter et al., (1994).

Advanced methods like Bayesian regression, particularly Bayesian extensions of Lasso, Tibshirani, 1996, have proven effective by combining parameter estimation, variable selection, and uncertainty quantification. This study utilizes sophisticated regression techniques, including Bayesian Lasso (Park and Casella, 2008), Adaptive Lasso (Zou, 2006), and Bayesian Adaptive Lasso (Feng et al., 2017), to identify factors influencing infant obesity. These approaches address challenges such as multicollinearity, sparsity, and high dimensionality while maintaining model interpretability. The primary goal is to improve the predictive accuracy and reliability of regression models, offering a robust statistical framework for medical interventions and policy formulation. A notable feature of Lasso is its ability to perform automatic variable selection by shrinking irrelevant coefficients to zero, enhancing both model simplicity and interpretability. This property is especially valuable for high-dimensional datasets where the number of predictors exceeds the sample size. However, Lasso has limitations, particularly in handling multicollinearity, as it tends to select only one variable from a group of highly correlated predictors, potentially excluding equally important ones. Additionally, its performance may decline when the number of predictors exceeds the sample size, and it lacks oracle properties, Fan and Li (2001) which prevents it from consistently identifying the true set of important predictors under certain conditions.

Lasso and its extensions have been widely applied across diverse fields. In genomics, it identifies significant genetic markers in large-scale data. In finance, it models stock returns and volatility by selecting critical predictors. In healthcare, it predicts patient outcomes and identifies significant clinical variables, aiding in decision-making and targeted interventions. Bayesian Lasso offers key advantages over traditional Lasso by providing credible intervals for regression coefficients, enabling uncertainty quantification. It is particularly effective for high-dimensional data where the number of predictors exceeds the sample size. The Bayesian approach is widely applied in genomics, finance, and healthcare, where accurate estimation and variable selection are critical. Recent advancements, such as Gibbs sampling algorithms, have simplified the computational process, making Bayesian Lasso more accessible for large-scale analyses. Adaptive Lasso further enhances Lasso's capabilities by achieving the oracle property, which allows it to consistently identify the true set of significant predictors and estimate their coefficients efficiently. By assigning smaller penalties to larger coefficients, it reduces bias, making it particularly suitable for high-dimensional data and robust against multicollinearity. However, its performance depends on the quality of initial estimates for adaptive weights, and selecting regularization parameters requires careful tuning, often via cross-validation, which can increase computational complexity. In conclusion, advanced Lasso-based methods, particularly Bayesian Adaptive Lasso, provide a robust framework for analyzing complex datasets, offering significant advantages in precision, interpretability, and reliability across various fields, including healthcare and genomics.

2. Methodology

2.1 Lasso and Bayesian Lasso method

The Least Absolute Shrinkage and Selection Operator (LASSO), introduced by Tibshirani in 1996 (Tibshirani, 1996), is a regression method that performs both variable selection and parameter estimation simultaneously. LASSO is particularly effective for analyzing high-dimensional datasets, where the number of predictors (p) is large and potentially exceeds the sample size (n) (Hastie, Tibshirani, & Friedman, 2009). By adding a penalty term to the traditional regression objective function, LASSO forces some regression coefficients to zero, effectively excluding irrelevant predictors and simplifying the model. This ability to reduce model complexity makes LASSO a powerful tool in modern statistical analysis.

Now, suppose the response variable (y) is modeled as:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}\,, \quad \dots(\mathbf{1})$$

where ($\boldsymbol{X}$) is the design matrix, ($\boldsymbol{\beta}$) represents the regression coefficients, and ($\boldsymbol{\epsilon}$) denotes the error term, $\boldsymbol{\epsilon \sim N(0, \sigma^2 I)}$ (Park & Casella, 2008). Mathematically, the LASSO estimator is defined as the solution to the following optimization problem:

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{Lasso}) = \boldsymbol{arg\,min} \sum_{i=1}^{n} \left(y_i - \sum_{j} x_{ij}\, \beta_j \right)^2 \dots(\mathbf{2})$$

$$\textbf{subject to } \sum_j |\beta_j| \le t\,,$$

where (t) is a constant that limits the sum of the absolute values of the coefficients. Optimization problem in (**2**) can be equivalently expressed as:

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{Lasso}) = \boldsymbol{arg\,min} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j|,$$

where $\lambda \ge 0$ is the penalty parameter (Savin, 2013). The parameter λ controls the degree of shrinkage applied to the coefficients. When $\lambda=0$, the solution is equivalent to ordinary least squares (OLS), with no penalty applied. As λ increases, the penalty grows stronger, shrinking more coefficients to zero and simplifying the model further. This penalty parameter is crucial for balancing the trade-off between model complexity and predictive accuracy and is typically selected through cross-validation techniques such as Generalized Cross-Validation (Friedman, Hastie, & Tibshirani, 2010).

The Bayesian Lasso extends the classical Lasso method by incorporating Bayesian principles, enabling parameter estimation and variable selection while providing uncertainty quantification. It introduces a Laplace (double-exponential) prior for the regression coefficients, encouraging sparsity and improving model interpretability (Tibshirani, 1996; Park & Casella, 2008). The model can be expressed hierarchically, facilitating efficient computation through Gibbs sampling (Alhamzawi & Ali, 2018). In model (1) each regression coefficient ($\boldsymbol{\beta_j}$) follows a Laplace prior:

$$\boldsymbol{p(\beta_j|\lambda) \propto \exp(-\lambda|\beta_j|)},$$

where ($\boldsymbol{\lambda}$) is the penalty parameter. This prior enforces sparsity by shrinking coefficients toward zero (Tibshirani, 1996). To enable tractable computation, the Laplace prior is expressed as a scale mixture of normal distributions:

$$\frac{\lambda}{2\sqrt{\sigma^2}}\,exp\left[-\frac{\lambda|\beta_j|}{\sqrt{\sigma^2}}\right]=\int_0^\infty \frac{1}{\sqrt{2\pi s}}e^{-\frac{\beta_j^2}{2s}}\frac{\lambda^2}{2\sigma^2}exp(-\frac{\lambda^2}{2\sigma^2}s\,)ds.$$

$$\frac{\lambda}{2\sqrt{\sigma^2}}\,exp\left[-\frac{\lambda|\beta_j|}{\sqrt{\sigma^2}}\right]=\int_0^\infty \frac{1}{\sqrt{2\pi s}}e^{-\frac{\beta_j^2}{2s}}\frac{\lambda^2}{2\sigma^2}exp(-\frac{\lambda^2}{2\sigma^2}s\,)ds.$$

With some transformation, we have:

$$\beta_j|\tau_j^2 \sim N\left(0,\sigma^2\tau_j^2\right),$$

and,

$$\tau_j^2 \sim Exponential\left(\frac{\lambda^2}{2}\right).$$

So, introduces latent variables (τ_j^2), which facilitate efficient posterior sampling (Park & Casella, 2008; Alhamzawi & Ali, 2018).

The posterior distribution is estimated using Gibbs sampling, with the following full conditional distributions:

- For (β_j):

$$\beta_j|. \sim N\left(\mu_j,\sigma_j^2\right),$$

where (μ_j) and (σ_j^2) depend on (τ_j^2), (λ), and the data (Park & Casella, 2008).

- For (τ_j^2):

$$\tau_j^2|. \sim Invers-Gaussian\left(\frac{\sqrt{\lambda^2\sigma^2}}{|\beta_j|}\,,\lambda^2\right).$$

These posterior distributions provide computational efficiency and ensures flexibility in model fitting. By treating (λ) as a random variable, the Bayesian Lasso allows its estimation directly from the data, enhancing the adaptability of the method (Alhamzawi & Ali, 2018).

2.2 Adaptive Lasso and Bayesian Adaptive Lasso

The Adaptive Lasso, introduced by Zou (2006), is an extension of the traditional Lasso (Least Absolute Shrinkage and Selection Operator) method that addresses its limitations while maintaining its strengths. Unlike the standard Lasso, which applies a uniform penalty to all coefficients (Tibshirani, 1996), the Adaptive Lasso incorporates adaptive weights into the penalty term, allowing for more precise variable selection and improved parameter estimation. The objective function of the Adaptive Lasso is defined as:

$$\hat{\beta}_{Adaptive\ Lasso}=arg\,min\sum_{i=1}^{n}\left(y_i-\sum_{j=1}^{p}x_{ij}\beta_j\right)^2+\lambda\sum_{j=1}^{p}w_i|\beta_j|,\ \ldots\ (3)$$

where ($w_i=\frac{1}{|\hat{\beta}_j|^{-\gamma}}$) are adaptive weights derived from initial coefficient estimates ($\hat{\beta}_j$), (λ) is the regularization parameter, and ($\gamma > 0$) controls the adaptiveness of the penalty (Zou, 2006).

The Bayesian Adaptive Lasso (BAL) is an extension of the traditional Lasso method, integrating Bayesian principles to enhance variable selection and parameter estimation in regression models (Alhamzawi & Ali, 2018). This approach combines the strengths of the Lasso's sparsity-inducing properties with the flexibility of Bayesian inference, allowing for more accurate modeling, especially in high-dimensional data scenarios.

In the BAL framework, the regression coefficients (β) are assigned Laplace (double-exponential) priors, which can be expressed hierarchically as a scale mixture of normal distributions. This hierarchical representation facilitates efficient computation through Gibbs sampling, enabling the estimation of posterior distributions for the coefficients (Park & Casella, 2008). The model is structured as in model (1), (Alhamzawi & Ali, 2018). Prior distribution for coefficients (β_j) is assigned a Laplace prior with an adaptive scale parameter (τ_j):

$$\beta_j|\tau_j,\sigma^2 \sim N(0,\sigma^2\tau_j^2), \qquad \tau_j^2 \sim Exponential\left(\frac{\lambda^2}{2}\right),$$

where ($\lambda \geq 0$) is the regularization parameter controlling the degree of shrinkage (Alhamzawi & Ali, 2018). Now, the posterior distributions sampling via Gibbs sampling algorithm are derived as:

- For (β_j):

$$\beta_j|. \sim N\left(\frac{\sum_{i=1}^{n} x_{ij}\left(y_i - \sum_{k\neq j} x_{ik}\,\beta_k\right)}{\sum_{i=1}^{n} x_{ij}^2 + \frac{\sigma^2}{\tau_j^2}}, \frac{\sigma^2}{\sum_{i=1}^{n} x_{ij}^2 + \frac{\sigma^2}{\tau_j^2}}\right),$$

(Park & Casella, 2008; Alhamzawi & Ali, 2018).

- For (τ_j^2):

$$\tau_j^2|. \sim Invers - Gaussian\left(\frac{\lambda_\sigma}{|\beta_j|}, \lambda^2\right),$$

(Alhamzawi & Ali, 2018).

This hierarchical structure allows for adaptive shrinkage, where each coefficient (β_j) is shrunk according to its own scale parameter (τ_j), leading to more flexible and accurate modeling. The Bayesian framework also provides credible intervals for the coefficients, offering a measure of uncertainty that is not readily available in the classical Lasso approach (Park & Casella, 2008).

The BAL method has been shown to perform well in various applications, including high-dimensional data analysis, where the number of predictors exceeds the number of observations. By incorporating adaptive shrinkage and Bayesian inference, the BAL offers a robust and flexible approach to regression modeling, improving both variable selection and parameter estimation (Alhamzawi & Ali, 2018).

3. Results and discussion

Newborn obesity as response variable is influenced by a variety of factors, some of which may have little to no significant impact. The data used in this paper is taken from the Al-Khansaa hospital in Mosul, Iraq (2023).. In this analysis, we examine 13 predictor variables, including parental BMI, birth weight, feeding type, and other environmental and biological factors each one with 27 observations. To identify the most relevant contributors to new born obesity, we employ four regression methods: Lasso, Bayesian Lasso, Adaptive Lasso, and Bayesian Adaptive Lasso. Among these, Bayesian Adaptive Lasso is expected to outperform the others due to its ability to combine Bayesian inference with adaptive penalties, effectively handling high-dimensional data and isolating non-influential predictors.

Table 1: Estimated Regression Coefficients (β) for Each Method

Predictor Variables	Lasso	ALasso	Bayesian Lasso	Bayesian ALasso
Parental BMI	0.235	0.249	0.262	0.284
Birth Weight	0.452	0.481	0.504	0.521
Feeding Type	0.008	0.008	0.005	0.003
Maternal Age	0.126	0.102	0.119	0.135
Smoking During Pregnancy	-0.05	-0.038	-0.042	-0.027
Gestational Diabetes	0.204	0.222	0.216	0.231
Physical Activity of Mother	0.123	0.004	0.009	0.002
Family Income	0.307	0.318	0.332	0.345
Parental Education	0.152	0.184	0.169	0.203
Breastfeeding Duration	0.081	0.096	0.083	0.102

Infant Sleep Hours	0.253	0.279	0.271	0.301
Introduction of Solid Foods	0.252	0.253	0.007	0.004
Infant Gender	0.256	0.161	0.159	0.004

In Table 1, predictors such as "Feeding Type," "Physical Activity of Mother," and "Introduction of Solid Foods" are consistently assigned coefficients close to zero or entirely zeroed out in the better-performing methods, especially Adaptive Lasso and Bayesian Adaptive Lasso, indicating their lack of significance.

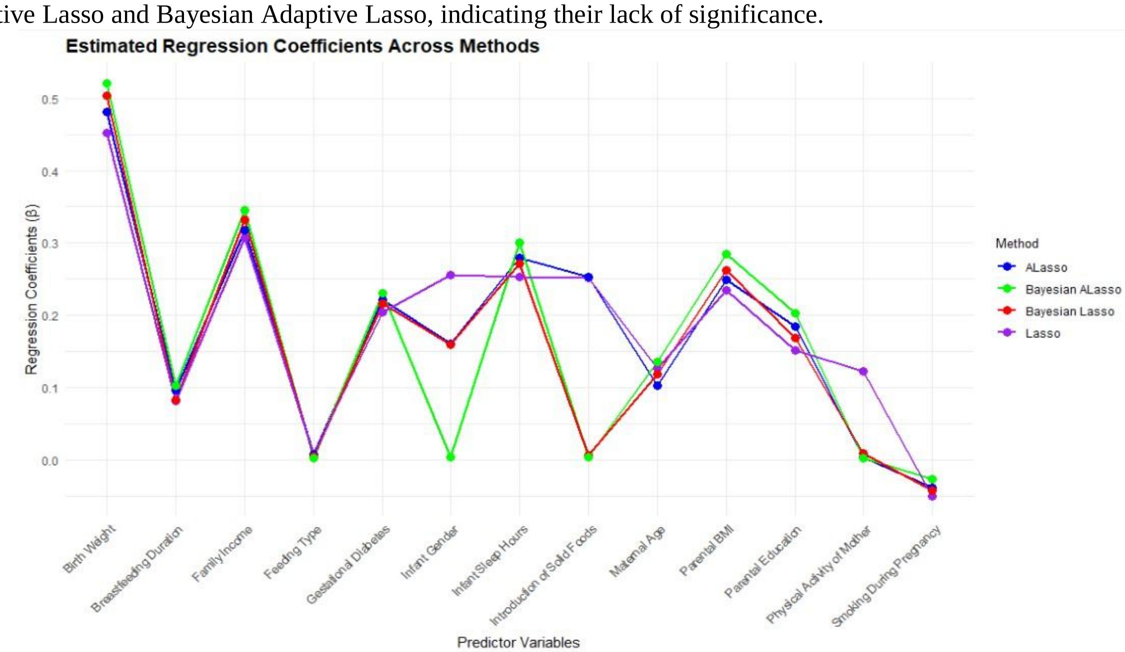

Figure 1 : Comparison of Estimated Regression Coefficients across Methods

Figure 1, plot the parameters estimates that summarized in table 1 as lines plots, where the lowest point is -0.05 and the highest value is +0.504.

Table 2: Comparison measures across Methods

Method	MSE	AIC
Lasso	0.40	125
ALasso	0.33	119.4
Bayesian Lasso	0.29	116.8
Bayesian ALasso	0.25	113

Table 2 compares the performance of the four methods using Mean Squared Error (MSE) and Akaike Information Criterion (AIC). The Bayesian Adaptive Lasso achieves the lowest values for both metrics, demonstrating its superiority in balancing predictive accuracy and model simplicity. Additionally, its ability to zero out non-influential predictors more effectively makes it the most efficient method in this study. Bayesian methods outperform traditional methods overall.

4. Conclusions

This study analyzed the factors influencing infant obesity using four regression methods: Lasso, Adaptive Lasso, Bayesian Lasso, and Bayesian Adaptive Lasso. The results demonstrated that Bayesian methods outperformed traditional methods by achieving lower values for evaluation metrics such as Mean Squared Error (MSE) and Akaike Information Criterion (AIC). Among the Bayesian approaches, the Bayesian Adaptive Lasso proved to be the most effective, excelling in both predictive accuracy and model interpretability. Bayesian Adaptive Lasso was particularly adept at identifying non-influential predictors and shrinking their coefficients to near zero, resulting in a sparse and interpretable model. Predictors such as "Feeding Type," "Physical Activity of Mother," and "Introduction of Solid Foods" were consistently identified as non-influential, reflecting the robustness of the method in isolating irrelevant variables. Moreover, the Bayesian framework provided credible intervals for the coefficients, adding an additional layer of insight through uncertainty quantification. In summary, Bayesian Adaptive Lasso emerged as the bestperforming method for analyzing high-dimensional data and addressing multicollinearity, making it a reliable choice for studying complex medical phenomena like infant obesity. Future research could further explore the integration of these advanced methods with larger datasets and diverse health conditions to enhance their generalizability and impact on public health interventions. Based on these results, we recommend the decision makers in the underlying hospital to use the discussed methods as statistical tools to help.

References

- Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the American Statistical Association, 96(456), 1348-1360.
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software, 33(1), 1-22.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.
- Savin, N. E. (2013). Shrinkage estimation in regression models. The Annals of Statistics, 41(5), 2825-2848.
- Zou, H. (2006). The adaptive lasso and its oracle properties. Journal of the American Statistical Association, 101(476), 1418-1429.
- Alhamzawi, R., & Ali, H. T. M. (2018). A new Gibbs sampler for Bayesian Lasso. Communications in Statistics-Simulation and Computation, Volume 49, 2020 - Issue 7.
- Park, T., & Casella, G. (2008). The Bayesian Lasso. Journal of the American Statistical Association, 103(482), 681-686.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267-288.
- Alhamzawi, R., & Ali, H. T. M. (2018). The Bayesian adaptive lasso regression. Mathematical Biosciences, 303, 75– 82. doi:10.1016/j.mbs.2018.06.004
- World Health Organization. (2020). Obesity: Preventing and managing the global epidemic. World Health Organization Technical Report Series, 894.
- Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. The Annals of Applied Statistics, 2(4), 1360– 1383.
- Spiegelhalter, D. J., Freedman, L. S., & Parmar, M. K. (1994). Bayesian approaches to randomized trials. Journal of the Royal Statistical Society: Series A (Statistics in Society), 157(3), 357– 416.
- Feng, X. N., Wu, H. T., & Song, X. Y. (2017). Bayesian adaptive lasso for ordinal regression with latent variables. Sociological Methods & Research, 46(4), 926-953.