Bayesian Inference for Modeling Low Birth Weight and Maternal Risk Factors

Authors: Johan Carlo Ilagan, Sophia Lee, Daniel O'Connor
Journal: Australian Journal of Biomedical Data Science
DOI: 10.1177/ajbds.2024.0045
Publication Date: December 10, 2024

Abstract

This paper applies Bayesian inference to model the relationship between maternal characteristics and the likelihood of low birth weight outcomes in newborns. Using data from 2,500 births across Western Australia, we estimate posterior probabilities for risk factors including maternal age, BMI, smoking, and prenatal care frequency. Results indicate significant nonlinear effects of BMI and smoking on birth weight distributions. The Bayesian approach provides robust uncertainty quantification for population health modeling.

---

1. Introduction

Low birth weight (LBW) is a major public health concern due to its association with infant morbidity and long-term developmental risks. Traditional regression techniques often fail to capture the inherent uncertainty and heterogeneity in biological data. Bayesian inference provides a flexible framework for probabilistic modeling and prior incorporation based on medical knowledge.

---

2. Methods

2.1 Dataset and Variables

Data were sourced from hospital birth records collected between 2018–2023. Predictor variables included maternal age, BMI, smoking status, and prenatal care attendance.

2.2 Model Specification

A hierarchical Bayesian logistic regression was implemented in *RStan*. Noninformative priors were used for most coefficients, and Gibbs sampling (10,000 iterations) produced posterior distributions for all parameters.

---

3. Results

Posterior analysis revealed a strong negative relationship between maternal BMI and LBW probability. Smoking increased LBW risk by approximately 2.3× (95% credible interval: 1.9–2.8). The model's posterior predictive accuracy reached 87%, outperforming frequentist baselines.

---

4. Discussion and Conclusion

Bayesian inference enhances interpretability and uncertainty quantification in biomedical modeling. The results align with prior epidemiological findings and highlight the potential of probabilistic modeling for health policy planning.
Future research will integrate genetic and socioeconomic covariates to improve causal inference on LBW outcomes.

Acknowledgements

We thank the WA Health Data Linkage Branch for data access and the University of Western Australia's Biomedical Data Science Lab for support.