Logistic Regression vs. Bayesian Logistic Regression in Breast Cancer Identification

Morgan Huang, Yuxiang Feng, Issey Sone, Tongyu Wu, Enze Zhao

April 2025

1 Abstract

Early detection of breast cancer is crucial for improving survival rates, but equally important is the ability to confidently rule out malignancy to avoid unnecessary anxiety or intervention. This study compares frequentist and Bayesian logistic regression approaches for predicting malignancy using the Wisconsin Breast Cancer dataset. The Bayesian approach incorporates prior knowledge and produces full posterior distributions, allowing for uncertainty quantification. Both models performed well, and demonstrated high accuracy, but the Bayesian model achieved higher sensitivity and was able to correctly identify all malignant test cases. Furthermore, it identified additional significant predictors due to the stabilizing effect of weakly informative priors. These findings suggest that Bayesian logistic regression offers certain advantages in cancer detection tasks due to its robustness and interpretability.

2 Introduction

Breast cancer is a disease where cells in the breast tissue divide uncontrollably, and can form tumors. It is one of the leading causes of mortality, currently causing one in six deaths worldwide (Barrios, 2022). Breast cancer remains one of the most frequently diagnosed cancers and continues to pose a significant public health challenge (Wilkinson & Gathani, 2022), however when breast cancer is detected early, survival rates exceed 90%(American Cancer Society, 2022). Traditional diagnostic methods such as mammography can miss 10-30% of cancers, highlighting the need for improved analytical techniques to increase diagnostic accuracy (American Cancer Society, 2023).

Statistical modeling has become a popular and effective method in medical applications. Logistic regression is a widely used method for binary classification problems and has been used extensively to model the probability of malignancy in breast tissue. Although its popularity, the frequentist approach to logistic regression can be limited, particularly when dealing with small sample sizes or imbalanced datasets. Bayesian logistic regression offers an alternative by incorporating prior knowledge into the analysis and providing a full probabilistic framework for inference. Also, such an approach may yield more robust parameter estimates and a clearer understanding of model uncertainty. (Hosmer et al., 2013)

This study aims to compare the performance of both the Frequentist and Bayesian approach to logistic regression. Both models were trained using the Wisconsin Breast Cancer dataset on Kaggle. The overall goal is to identify the differences between the two models in terms of model prediction, interpreability of the coefficients, and the significance of each coefficient.

2.1 Explanatory Data Analysis

The data set contains 699 observations, with 444 of them diagnosed as benign and 239 of them diagnosed as malignant. The data contains nine predictor variables: Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei, Bland Chromatin, Normal Nucleoli, and Mitoses. In addition, each variable in the data is categorical, with predictor categories with values as integers between 1 and 10. The response variable, on the other hand, is a binary variable, where a value of 2 represents a benign tumor and a value of 4 represents a malignant tumor.

Before we perform multivariate analysis, we handle the missing values in the dataset. Notably, we see that there are 16 missing data out of the 699 total data points, all of which are present in the Bare Nuclei predictor. Since the number of missing data points is relatively low, we simply discard the data instead of using imputation techniques.

Now, to understand the distribution and variability among these predictors, we graph histograms of each predictor with respect to the tumor diagnosed. For the observations with benign tumors, we can see that

the distribution for each predictor is heavily right skewed as most of the observations lie near 0 to 1. For the malignant cases, the distribution is slightly more uniform to left skew.

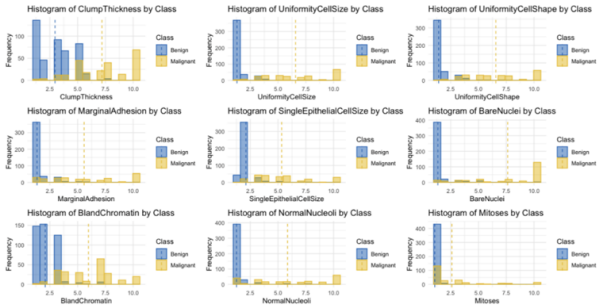


Figure 1: Histograms

In addition, box plots of each predictor with respect to the type of tumor diagnosed are graphed to further examine the variability in the data and to identify any extreme values that could influence the modeling process. We found that for benign tumors, the distribution is very close to values around 0-1, while the variability of malignant tumors are more spread out. Notice that there are a couple of outliers when the tumor is benign.

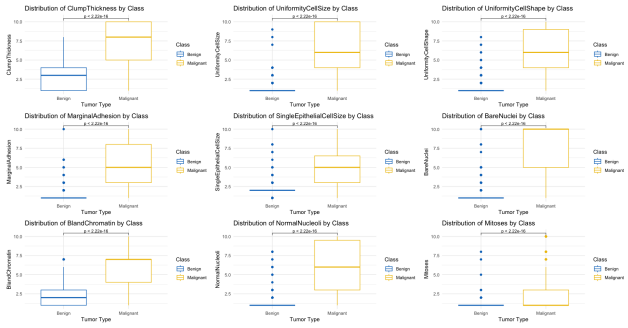


Figure 2: Boxplots

We can further justify our visual analysis using the χ^2 test of independence, which tests whether the values of each predictor variable are independent of the type of tumor diagnosed. In other words, we would like to test if the categorical distribution of a predictor variable given that the tumor is benign differs from the distribution of the variable given that the tumor is malignant. Through these tests, we obtain a p-value of $p < 0.001$ for each predictor variable in our dataset, thus concluding that there is a dependency between the type of tumor diagnosed and each predictor.

Furthermore, a correlation analysis was performed and visualized with a heat map to assess pairwise relationships among predictor variables. The heatmap identifies any strong correlations that could indicate multicollinearity, which is an important consideration when selecting variables for logistic regression modeling. Overall, the variables exhibit a general positive correlation, with a notably high correlation of 0.91 observed between Uniformity Cell Size and Uniformity Cell Shape.

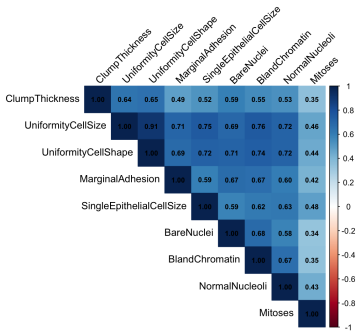


Figure 3: Correlation Matrix

3 Methodology

3.1 Logistic Regression

In classification problems, a common statistical model to use is logistic regression. In this section, we focus on binary classification problems, where we encode the truth value of a response variable Y as $Y = 1$ if Y is true and $Y = 0$ if Y is false, given some input value X. We say that $\pi(x)$ denotes the probability of "success" of Y given $X = x$, or the probability that $Y = 1$ given $X = x$, and is thus the parameter we would like to perform inference with.

Now, to perform inference, we must define our model, similar to ordinary linear regression. Now, if there are multiple predictors(categorical or continuous) in our model, as in our case, we define the model as follows:

$$\text{logit}(\pi(x)) = \log(\frac{\pi(x)}{1-\pi(x)}) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

where α and $\{\beta\}_{i=1}^{k}$ are parameters we will try to estimate, and there are k predictors $X_1, \ldots, X_k$. Usually, if a categorical predictor has multiple categories, we should create dummy variables for each category, but since our data assumes that there is an order to the categories(from levels 1 to 10), we can leave the model as is (Nahhas & Ramzi W., 2025).

In inference, it is important that we are able to interpret the parameters being estimated. We shall ignore the interpretation of α and instead focus on interpreting the other parameters $\beta_1, \ldots, \beta_k$. Now, there are two possible interpretations for a parameter β_i, where $i = 1, 2, \ldots, k$, of which we shall use the odds interpretation. Recall that the odds of success is given by the following, using our model definition:

$$\frac{\pi(x)}{1-\pi(x)} = \exp(\alpha + \sum_{i=1}^{k} \beta_i X_i) = e^{\alpha} \prod_{i=1}^{k} (e^{\beta_i})^{x_i}$$

Therefore, we can say that for a single parameter β_i, the odds of success $Y = 1$ are multiplied by a factor of $\exp(\beta_i)$ when increasing the predictor X_i by 1, holding all other predictors constant.

3.1.1 Inference

We now focus on performing statistical tests on the model parameters $\beta_1, \ldots, \beta_k$, again, ignoring the α parameter. Now, by using maximum likelihood methods, we can obtain the MLE $\hat{\beta}_i$ for parameter β_i. Note that there is no closed form solution for $\hat{\beta}_i$, however the loss function is still convex. Thus, there is a unique solution for $\hat{\beta}_i$.

After obtaining the estimate $\hat{\beta}_i$, suppose we want to see whether or not the i^{th} predictor is statistically relevant in our model. That is, is the i^{th} predictor necessary to predict the probability of success, $Y = 1$? We can conduct a Wald test to answer this question, with the null hypothesis as $H_0 : \beta_i = 0$ against the alternative hypothesis as $H_a : \beta_i \neq 0$. Using theorems regarding the MLE, we derive its test statistic as:

$$z = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \sim N(0, 1)$$

Therefore, we can construct a confidence interval with significance level of α for $\hat{\beta}_i$ as follows:

$$\hat{\beta}_i \pm z_{\alpha/2} SE(\hat{\beta}_i)$$

Now, say we want to test if the data we obtained came from a more complex model or from a more simple model. Then, we can use the likelihood ratio test to compare both likelihoods derived from their corresponding models. We define the null hypothesis as $H_0 : \beta_j = \beta_{j+1} = \cdots = \beta_n = 0$ against H_a : at least one $\beta_p \neq 0$, for $p = j, \ldots, n$, where $j, n \in [1, k], j < n$. The test statistic is defined as:

$$D = -2(L_0 - L_a) \sim \chi^2_{(df)}$$

where L_0 is the log likelihood of the reduced model under H_0, and L_a is the log likelihood of the full model. The degrees of freedom is calculated as the difference in the degrees of freedom in the full model and the degrees of freedom in the reduced model.

3.1.2 Multicollinearity

Multicollinearity arises when two or more predictor variables in a regression model are highly correlated. The presence of multicollinearity can have a negative impact on the analysis because it undermines the statistical significance of individual predictors and inflates the standard errors of the estimated coefficients.

Using Correlation to Assess Multicollinearity

As seen in our correlation matrix in Figure 3, several predictor variables exhibit moderate to high correlations, with the strongest correlation of $r = 0.91$. This indicates a potential multicollinearity issue that must be addressed during model selection.

However, pairwise correlation analysis only detects linear dependence between two variables at a time. Multicollinearity can still be present due to complex inter-dependencies among three or more variables. Therefore, we use correlation analysis as a preliminary step and follow up with more comprehensive diagnostics.

Variance Inflation Factor (VIF)

To detect more complex forms of multicollinearity, we calculate the Variance Inflation Factor (VIF) for each predictor:

$$\text{VIF}(X_j) = \frac{1}{1 - R_j^2}$$

where R_j^2 is the coefficient of determination obtained by regressing predictor X_j on all other predictors. A VIF value greater than 5 (or 10, by some conventions) is typically considered indicative of high multicollinearity.

3.1.3 Model Selection

After adding all main effects and interaction terms to the model(terms in which we introduce a dependency between the relationship between the response and a predictor, and another predictor), the model may become too complex to interpret, and thus introduces a multitude of problems(high standard errors for predictor estimates due to multicollinearity). We can use likelihood ratio testing to solve this problem, but if we were to aim for strong predictive performance, we need to introduce other methods. Note that we do model selection after inference to avoid dealing with challenges in post-selection inference.

A natural criterion to use for model selection is the Akaike Information Criteria(AIC), which calculates how close the fitted values are with the expected values(using KL-divergence) and penalizes a large number of parameters. We define this criteria as:

$$AIC = -2(L(\hat{\beta}|X) - k)$$

where k is the number of parameters in our model. It is obvious to see that we aim to minimize the AIC. To do this, we perform a backwards elimination algorithm, where we start with all possible predictors, and remove a variable one-by-one such that the AIC is minimized. We terminate if our algorithm cannot minimize the AIC any further. This will be our final model we will use for prediction.

3.2 Bayesian Logistic Regression

3.2.1 Bayes Theorem and Logistic Regression

Bayesian logistic regression is a probabilistic approach to modeling binary outcomes using Bayesian Inference. Its approach extends from ordinary logistic regression by treating the coefficients themselves as random variables with their own prior distribution, as opposed to an unknown quantity (Johnson et al., 2021). This allows the model to incorporate prior knowledge or beliefs and to quantify uncertainty in the parameter estimates through posterior distributions. The idea follows from Bayes Theorem:

$$P(\beta|X, y) \propto P(y, X|\beta) \cdot P(\beta)$$

where

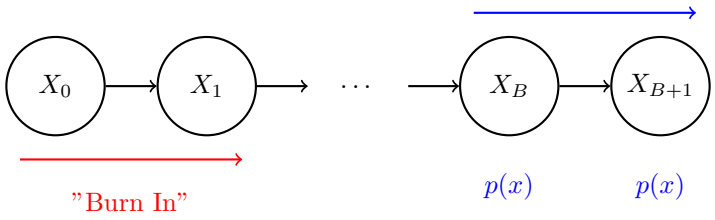
$$P(y, X|\beta) = \prod_{i=1}^{n} \sigma(\beta^T x_i)^{y_i}(1 - \sigma(\beta^T x_i)^{1-y_i})$$

Bayesian logistic regression is particularly useful when datasets are smaller, sparse, or imbalanced. This is due to regularization being built in to the model with the prior distribution. Bayesian logistic regression also gives a full posterior distribution of the model coefficients, and allows incorporation of prior belief based on existing studies.

The key challenge to Bayesian logistic regression is modeling the posterior distribution as it does not have a closed-form solution due to the non-conjugacy between the logistic likelihood and commonly used priors (eg. Gaussian Prior). To address these issues, approximation methods such as MCMC are used to get a good approximation of the posterior distribution by drawing samples from it.

3.2.2 Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) is a class of algorithms used to draw samples from a probability distribution ones that are too complex. The way it works is by constructing Markov chains that converge to an overall target distribution (Gunn-Sandell LB et al., 2023). Given a sample $x_t \sim p(x)$, it uses that to produce a new sample $x_{t+1} \sim p(x)$. The idea can be visualized below, where the "Burn-In" is the first couple of chains needed to get closer to the true posterior, and after enough samples, it will be able to draw from the true posterior $p(x)$.



Metropolis-Hastings is one of the algorithms for MCMC which generates a sequence of parameter samples that approximate the posterior distribution. Here is an overview of how the algorithm works:

1. Define the target posterior distribution up to a constant

$$p(x) = \frac{f(x)}{NC} \quad \text{(NC = Normalizing Constant)}$$

2. Initialize the chain with some value x_0, and define a proposal distribution $g(x_{t+1}|x_t)$
   Common choice:

$$g(x_{t+1}|x_t) = N(x_t, \sigma^2)$$

3. For each iteration t, draw a new candidate x_{t+1} from the proposal distribution centered at x_t
4. Compute the acceptance probability:

$$A(x_t \to x_{t+1}) = \min(1, \frac{f(x_{t+1})g(x_t|x_{t+1})}{f(x_t)g(x_{t+1}|x_t)})$$

5. Sample $u \sim U[0, 1]$
6. If $A(x_t \to x_{t+1}) > u$ then accept the new candidate and x_{t+1} is the next state in the chain

Over many iterations, the Markov chain produces samples that approximate the posterior, allowing estimation of expectations, variances, and credible intervals for parameters. The convergence of the algorithm can be visualized below where the orange line represents the approximating distribution and blue represents the true distribution:

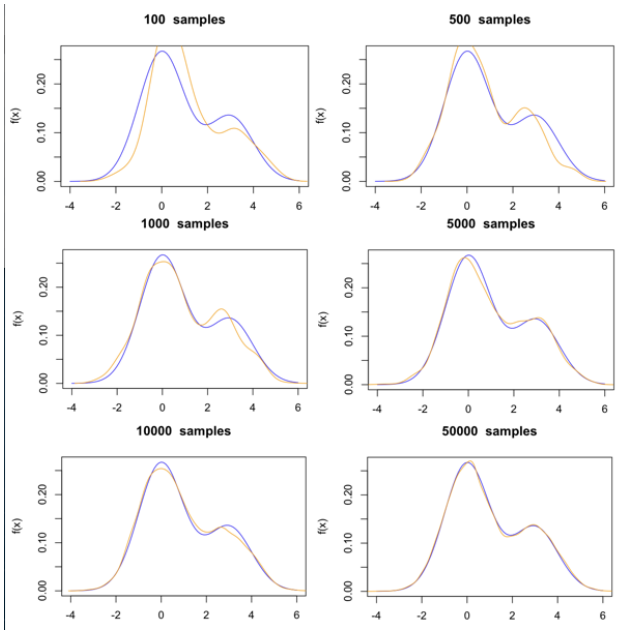


Figure 4: Convergence of Metropolis-Hastings algorithm.

Source: Chdrappi. *Metropolis Algorithm Convergence Example*. 1 Apr. 2013. *Using R; FOSS Statistical Software*, `https://commons.wikimedia.org/wiki/File:Metropolis_algorithm_convergence_example.png`. Accessed 23 Mar. 2025.

The overall "Burn-In" is the first 100 to 500 samples where it tries to approximate the true distribution. Once it reaches 10000 to 50000 samples, it has converged to the true distribution.

3.2.3 Bayesian Logistic Regression Setup

Recall, the posterior distribution for logistic regression follows from Bayes Theorem:

$$P(\beta|X, y) \propto P(y, X|\beta) \cdot P(\beta)$$

In Bayesian logistic regression, the choice of prior distributions plays a crucial role in shaping the model's behavior, especially when data is limited or noisy. Priors act as a form of regularization, helping to prevent overfitting by incorporating domain knowledge or assumptions about parameter scales. Well-chosen priors, such as weakly informative distributions such as the Cauchy or normal, can guide the model toward more stable and interpretable solutions without overpowering the data. This is particularly important in sensitive applications like cancer detection, where model robustness and interpretability are essential. A common prior choice for logistic regression is a Cauchy prior:

$$\beta_j \sim \text{Cauchy}(0, \text{s})$$

First suggested by Andrew Gelman, specifically, he suggested a Cauchy(0, 2.5) prior for logistic regression(Gelman et al. 2008). The Cauchy distribution is heavy-tailed, which allows large coefficients when the data strongly supports them, but also shrinks small and weakly identified coefficients toward zero, effectively acting as a form of regularization. It can be thought of as "Most effects are small. but I occasionally allow for large coefficients". Figure 5 shows the heavy-tailed property of the Cauchy distribution, as the density is higher on the tails compared to the normal.

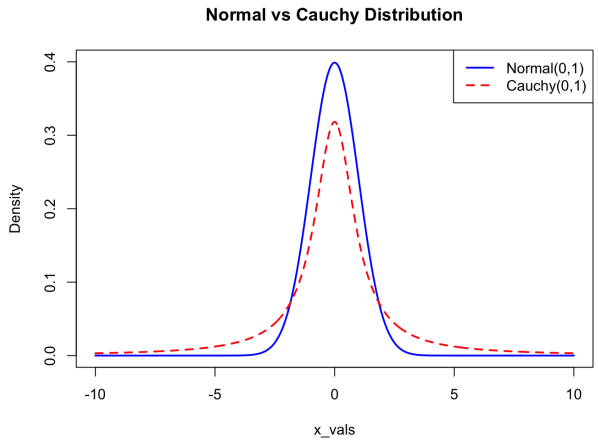


Figure 5: Normal vs. Cauchy Distribution

This can be particularly beneficial in the context of cancer-detection, where a weakly-informative prior can help identify which predictors truly contribute to cancer-detection. Although the Cauchy prior is often recommended for its heavy tails and weak informativeness, it can lead to slower convergence and potential numerical instability in MCMC sampling. Because the Cauchy prior allows for very large coefficient values, the posterior distribution can have long, flat tails, making it harder for samplers to explore efficiently. As a result, it's essential to monitor convergence diagnostics and ensure that the posterior samples are stable. As for the training, the STAN and brms package in R were used to train the model. The set up involved four Markov chains, 2000 iterations per chain with 1000 for the warmup and 4000 post-warmup samples.

3.3 Bayesian inference and interpretation

Inference about model parameters in a Bayesian perspective is based on their posterior distributions. As opposed to classical methods using p-values and null hypothesis testing, in a Bayesian perspective, direct probabilities that a coefficient is greater or less than a certain value are calculated. For example, in Bayesian logistic regression, we can quantify the probability that a predictor has a positive effect by calculating the posterior probability:

$$P(\beta_j > 0|\text{data})$$

If this probability is close to 1, it provides strong evidence that a predictor increases the odds of the outcome. Just like in classical hypothesis testing, we also set some significant level. In our case, after training the model, if any probabilities were less than 95% they were considered not statistically significant by Bayesian standards. One of the advantages to this approach would be the interpretation being more intuitive, rather than rejecting or failing to reject a null hypothesis, we directly assess the probability of practical importance. The interpretation follows: there is an X% probability that an increase in β_j is associated with an increase in odds of malignancy.

4 Results

To identify an optimal set of predictors, we used backward stepwise selection based on AIC, starting with all main effects and two-way interactions. Predictors were removed iteratively to minimize AIC, stopping when no further reduction was possible. The final model yielded an AIC of 108.99, a substantial improvement from the initial model's AIC of 136.12. Notably, two predictors—*Uniformity of Cell Size* and *Single Epithelial Cell Size*—were excluded from the final model. Several interaction terms were retained, as visualized in Figure, indicating their contribution to the model's explanatory power.

```
Clump_thickness:Marginal_adhesion
Uniformity_of_cell_shape:Marginal_adhesion
Uniformity_of_cell_shape:Bare_nuclei
Marginal_adhesion:Bare_nuclei
Bare_nuclei:Mitoses
Bland_chromatin:Normal_nucleoli
Bland_chromatin:Mitoses
```

Figure 6: Interaction Terms Included

As for the Bayesian model, all coefficients were fitted into the model, and the posterior distribution of the predictors can be seen below in figure 7. The strength and significance of each predictor can be seen, with certain predictors having their distribution to the right of zero, indicating those were overwhelmingly significant coefficients.

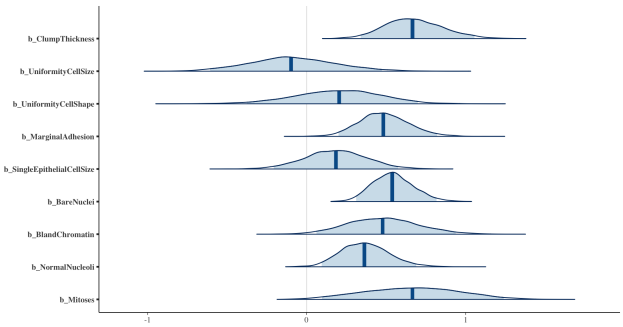


Figure 7: Distribution of posterior coefficients

We evaluated the performance of our logistic regression and Bayesian logistic regression models using an 80-20 train-test split with a decision threshold of 0.5. Both models showed a strong performance on the dataset, as shown by the confusion matrix below:

Logistic Regression

Prediction	**Benign**	**Malignant**
Benign	86	3
Malignant	2	44

Bayesian Logistic Regression

Prediction	**Benign**	**Malignant**
Benign	86	0
Malignant	2	47

Logistic regression achieved 97.62% training accuracy and 96.30% testing accuracy. Its sensitivity was 93.62% and its specificity was 97.73%. The Bayesian logistic regression model achieved a training accuracy of 97.08% and a test accuracy of 98.52%. It reached a 100% sensitivity and the same specificity of 97.73%.

These results suggest that the Bayesian LR model performs slightly better than the Ordinary LR model, especially in identifying malignant cases. In medical diagnosis, higher sensitivity is important to reduce the number of false negatives, since diagnosing a benign tumor as malignant is costly. Note that the decision threshold can be fine tuned to adjust sensitivity and specificity.

5 Discussion

The comparison between frequentist and Bayesian logistic regression reveals several key insights about model performance and interpretation in the context of breast cancer diagnosis. While both models achieved high accuracy, the Bayesian model offered slightly better performance, particularly in sensitivity, as it correctly identified all malignant cases in the test set. Note that sensitivity is crucial in detecting cancerous tumors, as a false negative carries a high cost.

In terms of statistically significant predictors, we used a significance level of $\alpha = 0.05$. Using posterior probabilities for Bayesian LR(Logistic Regression) and Wald Tests for Ordinary LR, both models identified

several shared predictors, including *ClumpThickness*, *MarginalAdhesion*, *BareNuclei*, *BlandChromatin*, and *NormalNucleoli*. However, some differences were observed. The Bayesian logistic regression model additionally highlighted *Mitoses* as a significant predictor, while the ordinary logistic regression model instead selected *UniformityOfCellShape*. This difference may be due to the Cauchy prior's ability to shrink weak signals while allowing strong ones to emerge, stabilizing estimates even in the presence of multicollinearity.

Figure 8: Significant predictors for each model

Posterior predictive checks further confirmed the Bayesian model's good fit, with predicted values closely reflecting the binary nature of the data. These results suggest that Bayesian logistic regression is not only effective for prediction but also offers a more nuanced understanding of uncertainty, which is particularly valuable in clinical decision-making contexts. The distribution of replicated results (y_{rep}) was U-shaped, indicating that most predictions were near 0 or 1, which aligns with the binary nature of the task. The similarity between y and y_{rep} suggests that the model fits the data well. It's also worth noting that all Rhat values were 1, and effective sample sizes were high, indicating good convergence and stable posterior estimates.

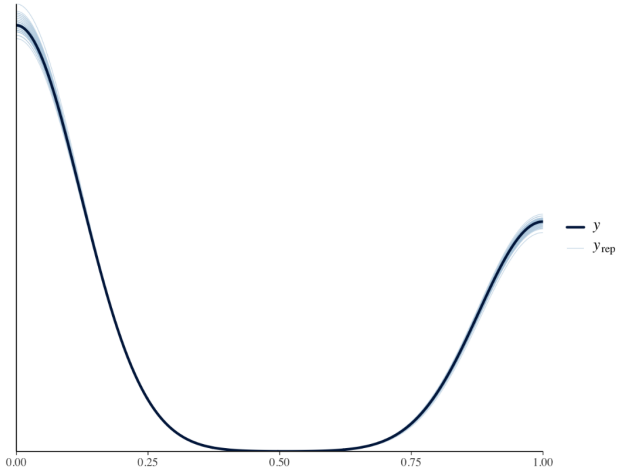


Figure 9: Post posterior check for 50 draws

5.1 Future Directions

To improve the ordinary Logistic Regression model, in the future we can use several techniques in addressing multicollinearity in our data. For example, we can use Principal Component Analysis(PCA) to combine correlated variables into several principal components and then fit our model. Or, we can instead use Ridge or Lasso Regression. In the presence of multicollinearity, the standard errors of some of the predictors will become very large, meaning that the coefficients may change drastically when small changes in the data appear. This can greatly affect the performance of the model on new data, and the coefficients become uninterpretable. Ridge regression aims for generally lower coefficient estimates, which thus reduces the variance of the model at the cost of some bias (Bishop, 2006). On the other hand, Lasso regression can perform variable selection, by forcing some coefficients to equal 0. However, it may be undesirable since it may choose a highly correlated but unimportant predictor, over a more important correlated predictor.

6 Conclusion

This study compared logistic regression and Bayesian approaches to logistic regression for breast cancer classification. Both models demonstrated strong performance, with accuracy above 95% on the test set. Even though predictors were selected using backwards AIC, the model was fitting interaction terms and resulted in a model with greater complexity. The use of weakly informative priors in the Bayesian model helped stabilize coefficient estimates and also makes the coefficients more interpretable. There were slight differences to the importance of each predictor between both of the models, and this was likely due to the weakly informative prior in the Bayesian model. In a medical setting where the cost of misclassifications is high, the Bayesian approach is more suited as it gave full posterior distributions, and was able to account for a certain level of uncertainty as well. Future work can be done by incorporating different priors or trying different regularization methods, transforming the data in certain ways like PCA before training the model, and using other diagnostic tools to ensure the model is completely robust.

7 References

1. Barrios, C. H. (2022). Global challenges in breast cancer detection and treatment. The Breast, 62, S3–S6. https://doi.org/10.1016/S0960-9776(22)00212-1

2. Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.

3. Canada, Public Health Agency of. (2024, December 30). Breast cancer. Government of Canada. https://www.canada.ca/en/public-health/services/chronic-diseases/cancer/breast-cancer.html

4. Chdrappi. (2013, April 1). Metropolis algorithm convergence example. Using R; FOSS Statistical Software. https://commons.wikimedia.org/wiki/File:Metropolis_algorithm_convergence_example.png

5. 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. https://doi.org/10.1214/08-AOAS191

6. Gunn-Sandell, L. B., Bedrick, E. J., Hutchins, J. L., Berg, A. A., Kaizer, A. M., & Carlson, N. E. (2023). A practical guide to adopting Bayesian analyses in clinical research. Journal of Clinical and Translational Science, 8(1), e3. https://doi.org/10.1017/cts.2023.689

7. Johnson, A. A., Ott, M. Q., & Dogucu, M. (2021). Bayes rules! An introduction to applied Bayesian modeling. CRC Press. https://www.bayesrulesbook.com/chapter-13

8. Nahhas, R. W. (2025). Introduction to regression methods for public health using R. Chapman & Hall/CRC.

9. Survival rates for breast cancer. (n.d.). American Cancer Society. Retrieved March 17, 2025, from https://www.cancer.org/cancer/types/breast-cancer/understanding-a-breast-cancer-diagnosis/breast-cancer-survival-rates.html

10. Limitations of mammograms: How accurate are mammograms? (n.d.). American Cancer Society. Retrieved March 17, 2025, from https://www.cancer.org/cancer/types/breast-cancer/screening-tests-and-early-detection/mammograms/limitations-of-mammograms.html

11. Wilkinson, L., & Gathani, T. (2022). Understanding breast cancer as a global health concern. The British Journal of Radiology, 95(1130), 20211033. https://doi.org/10.1259/bjr.20211033