MATH 640 Final Project

Ordinal Regression with Laplacian and Ridge Priors

Georgetown University
Washington, DC

Georgetown University
Washington, DC

Georgetown University
Washington, DC

1 Introduction

In recent years, there has been an increase in the amount of assistance for health security (AHS) provided by different organizations to various countries around the world. The efficacy of this aid can be measured using Bayesian analysis, specifically ordinal regression (Boyce et al., 2021). This study aims to observe the impact that assistance for health security has on workforce capacity through ordinal regression using both Laplacian and Ridge priors. Through Bayesian analysis, the goal is to determine the effect of AHS on a workforce's ability to address public health risks.

In their work, Boyce, Meyer, Kraemer, and Katz studied a variety of different health security capacities, including a country's workforce capacity. According to the WHO, workforce capacity is a measure of the preparedness of a country's workforce to handle public health hazards, including having sufficient human resources capabilities to administer International Health Regulations or sufficient training programs in epidemiology for its workforce (WHO, 2019). Boyce, Meyer, Kraemer, and Katz conclude that assistance for health security has a positive impact on workforce capacity 1 year after it is expended (Boyce et al., 2021).

James Albert and Siddhartha Chib developed ordinal regression models as a way to represent multinomial data. They suggest this method as comparable to using a probit link (Meyer, 2022). In their 1993 publication, *Bayesian Analysis of Binary and Polychotomous Response Data,* they demonstrate a way of determining the posterior distribution of $\vec{\beta}$, a vector of unknown parameters as part of the binary regression model (Albert and Chib, 1993). For data Y_i with at least two categories, independent latent variables Z_i can be introduced. This project aims to show that these latent variables follow a truncated Normal distribution, and that this information can be used to develop a Gibbs sampler and find the posterior distribution of $\vec{\beta}$. To penalize the coefficients, both Laplacian and Ridge priors are explored.

2 Methods

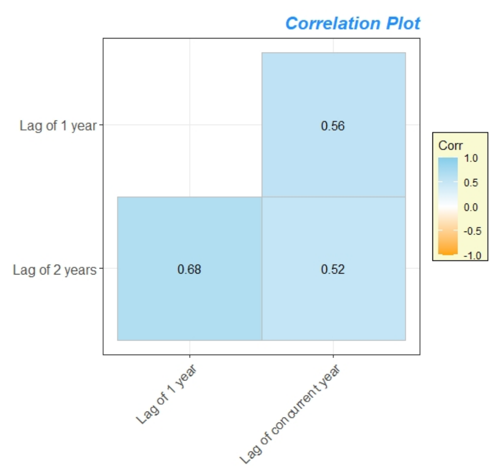
Suppose y_i can take on the ordered values $j = 1, ..., J$. The latent variable representation is then

$$y_i = \begin{cases} 1 \text{ if } \gamma_0 \le z_i \le \gamma_1 \\ ... \\ j \text{ if } \gamma_{j-1} \le z_i \le \gamma_j \\ ... \\ J \text{ if } \gamma_{J-1} \le z_i \le \gamma_J \end{cases}$$

where $z_i \sim \mathcal{N}(x_i'\beta, 1)$. The ordinal likelihood can then be represented as a mixture of Normals:

$$\mathcal{L}(\vec{Y}|\vec{\beta}, \vec{Z}, X) \propto \left\{ \sum_{j=1}^{J} 1(\gamma_{j-1}) < z_i \le \gamma_j) 1(y_i = j) \right\} \times \exp\left[-\frac{1}{2}(z_i - x_i'\vec{\beta})^2 \right] \tag{1}$$

To address the high correlation present among the independent variables all with values above 0.5 (as seen in the correlation plot below) and obtain stable estimates, the model coefficients ($\vec{\beta}$) will be penalized using both Laplacian and Ridge priors. For the Laplacian case, assume $\beta_j \sim Laplace(0, \sigma)$. This can be represented as a mixture model by letting $\beta_j \sim \mathcal{N}(0, \frac{4\sigma^2}{\alpha_j})$ where $\sigma^2 \sim IG(a, b)$ and $\alpha_j \sim IG(1, 0.5)$ for fixed, non-informative hyper-parameters a and b. For the Ridge case, assume $\beta_j \sim \mathcal{N}(0, \tau^2)$ with hyper-prior $\pi(\tau^2) \propto (\tau^2)^{-1}$. See the Appendix for calculations of conditional posterior distributions for both the Laplacian and Ridge cases.



Leveraging a mixed-model representation for the Laplacian distributed coefficients allows us to work with recognizable conditional posterior distributions. This means that a Gibbs sampler can be implemented instead of a Metropolis-Hastings Algorithm, which can be at least marginally more complex to implement. Model performance between the two methods will be assessed in the context of parameter convergence, auto-correlation between samples, and independent variable significance.

The data for this model was obtained from *Financial Assistance for Health Security: Effects of International Financial Assistance on Capacities for Preventing, Detecting, and responding to Public Health Emergencies* (Boyce et al., 2021), which also includes context and justification for all data exclusions. In summary, the resulting data includes AHS information for 59 countries, where exclusions were made due to issues such as data contamination (e.g. countries impacted by the West African Ebola Epidemic) and lack of current year funding data. Given this data, the model is intended to assess the probability of a country's workforce capacity assessment score falling into four ordered categories, where the four ordered categories are created by binning the capacity assessment score into quartiles. Additionally, the independent variables are the lagged disbursed AHS funds. Note that this variable refers to health security funding disbursed in the year of, 1 year prior, and 2 years prior to the year of the evaluation.

A sufficiently staffed, equitably distributed workforce is vital in responding effectively to public health emergencies. This model aims to give context on when disbursed funds provide the most benefit to workforce capacity. This context would then have the potential to encourage action on the part of policymakers and those directly responsible for disbursing funds.

3 Simulation Study and Application

Given all conditional posterior distributions were recognizable for both models, Gibbs samplers were employed. Parameter convergence was assessed using several diagnostics, including trace plots, the Geweke convergence diagnostic, autocorrelation function plots, and running mean plots. 20,000 samples were taken from each posterior distribution, with the first 10,000 samples treated as burn-in and discarded. For inferential purposes, 95% credible intervals were used for each lag parameter.

Laplace Priors On All Coefficients

Parameter	Median	LB*	UB**	Geweke
Intercept (β_0)	0.4226	0.0576	0.7839	1.4579
Lag 0 (β_1)	0.0051	-0.0063	0.0168	0.9741
Lag 1 (β_2)	0.0040	-0.0123	0.0198	0.4375
Lag 2 (β_3)	-0.0023	-0.0309	0.0275	0.0046

*=2.5$^{\text{th}}$ percentile; **=97.5$^{\text{th}}$ percentile

Laplace Priors On Predictor Coefficients

Parameter	Median	LB	UB	Geweke
Intercept (β_0)	0.4278	0.0538	0.8022	0.2585
Lag 0 (β_1)	0.0053	-0.0063	0.0165	-0.4499
Lag 1 (β_2)	0.0040	-0.0123	0.0202	1.1104
Lag 2 (β_3)	-0.0023	-0.0326	0.0270	-0.0761

Ridge Priors

Parameter	Median	LB	UB	Geweke
Intercept (β_0)	0.4227	0.0598	0.7839	-1.3518
Lag 0 (β_1)	0.0054	-0.0062	0.0170	0.5693
Lag 1 (β_2)	0.0040	-0.0119	0.0201	0.1308
Lag 2 (β_3)	-0.0026	-0.0319	0.0271	-0.2161

The posterior probabilities of the β coefficients for both models were also observed. The probabilities are summarized in the table below.

Posterior Probabilities, Pr($\beta > 0$), based on β priors

Parameter	Laplace V1	Laplace V2	Ridge
Intercept (β_0)	0.9875	0.9873	0.9885
Lag 0 (β_1)	0.811	0.8126	0.8169
Lag 1 (β_2)	0.6842	0.6885	0.6879
Lag 2 (β_3)	0.4411	0.4407	0.4333

V1: Laplace priors on β coefficients; V2: flat prior on β_0

In all three cases, the 95% credible interval for the coefficients of lagged disbursed funds include 0; however, the concurrently disbursed funds' coefficient had a posterior probability above 0.8, meaning the majority of posterior draws hinted at the recent funds being positively related to the probability of a country increasing their workforce capacity score.

For the Gibbs sampler developed using the Laplacian prior, no significant difference in convergence or inference is observed when the intercept is also penalized as opposed to having the non-informative prior of $p(\beta_0) \propto 1$. In both cases, the trace plots of the βs show no discernible patterns and appear to converge. When the intercept is also penalized, the absolute value of the Geweke statistics is much closer to the threshold of 1.96 for β_0. The running mean plots for all βs stabilize after the 3000th draw. For σ^2, on the other hand, in both cases convergence is not established based on the trace plots and the running mean plots, as well as the high Geweke statistics of -4.973 and -2.802 .

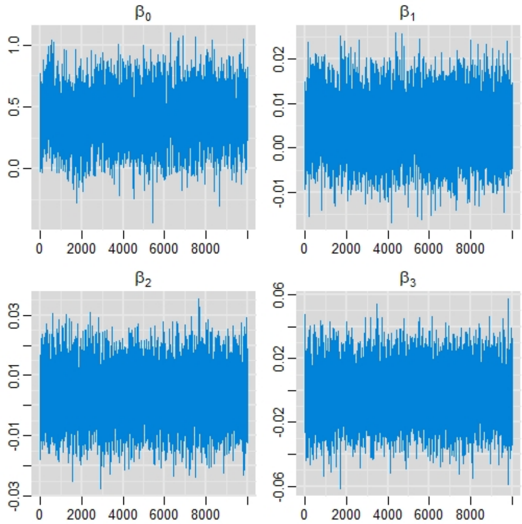


Figure 1: Trace plots of βs when intercept is penalized

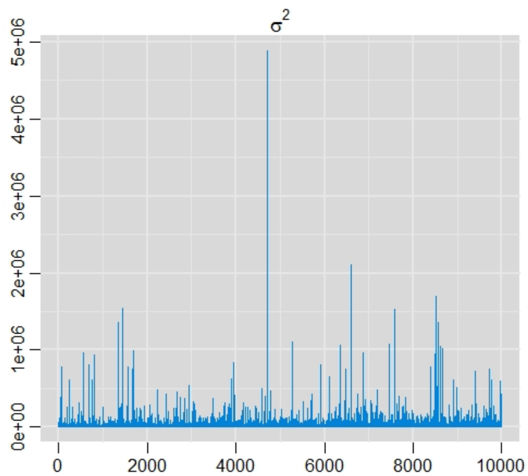


Figure 2: Trace plot of σ^2 when intercept is penalized

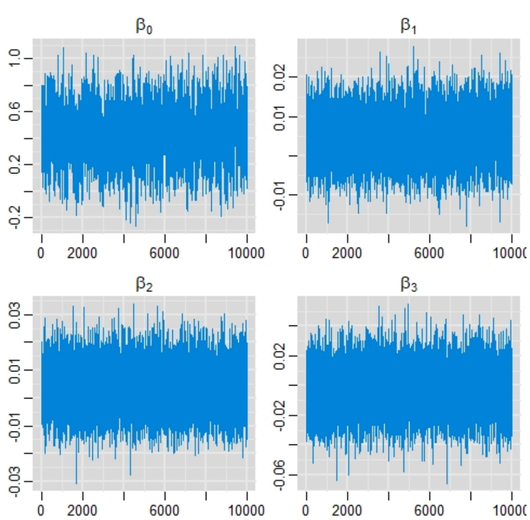


Figure 3: Trace plots of βs when intercept is not penalized

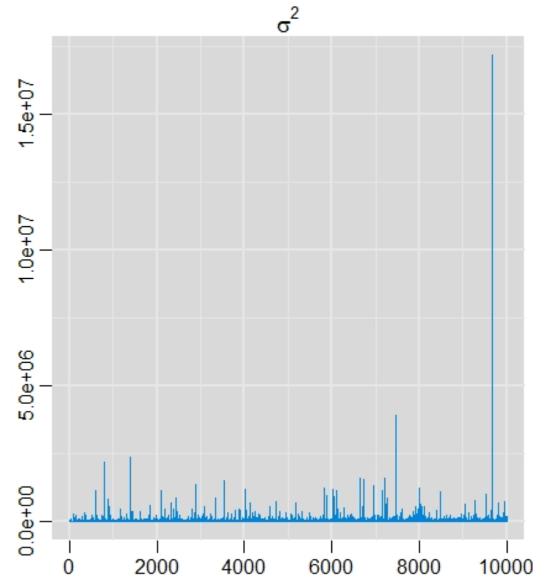


Figure 4: Trace plot of σ^2 when intercept is not penalized

For the Gibbs sampler developed using the Ridge prior, trace plots for the β parameters and the τ^2 parameter were generated. Based on several assessments, the coefficients of the Ridge model converge. However, different diagnostics give varied information about the convergence of τ^2. The Geweke diagnostic and autocorrelation function plot for τ^2 indicate that it converges and does not suffer from an autocorrelation issue (respectively), whereas the trace plot and the running mean plot do not indicate convergence.

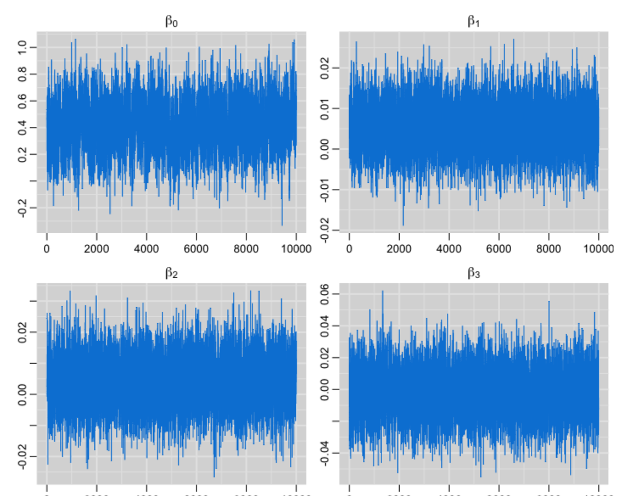


Figure 5: Trace plots for β parameters

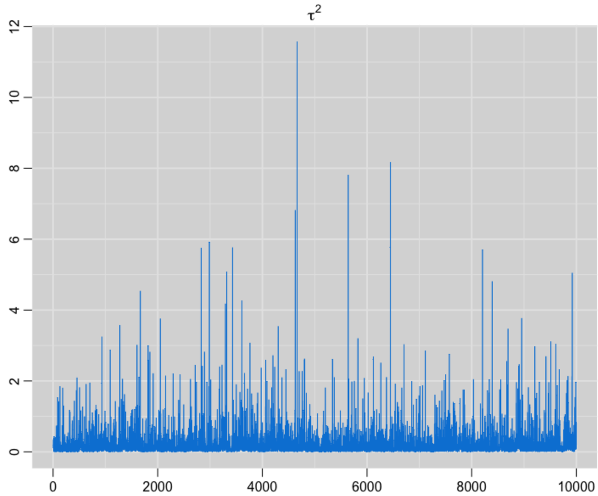


Figure 6: Trace plots for τ^2 parameter

Thinning was not performed as part of this analysis because the results of the autocorrelation plots demonstrated no significant autocorrelation issues. Only the draws for β_0 had an autocorrelation above the desired threshold up to 30 lags, which, however, were below 0.2. Overall, the three models display similar performance both in terms of parameter convergence and inference. Notably, the median of the posterior draws for all coefficients are extremely close to each other, if not exactly the same up to the number of significant digits. See the Appendix for other diagnostics used to assess convergence, including running mean plots and autocorrelation function plots.

4 Discussion

Given the lagged disbursed AHS funds were not significant (i.e. their 95% credible intervals contained 0) in both the Laplace and Ridge case, this signals that controlling for other variables available in the data may improve the predictive power of the model and allow additional information to be gleaned from the disbursed AHS funds coefficients. In *Financial Assistance for Health Security: Effects of International Financial Assistance on Capacities for Preventing, Detecting, and responding to Public Health Emergencies* (Boyce et al., 2021), which implements a Ridge prior and includes other variables in addition to the disbursed AHS funds, this is the case (namely, at least one of the disbursed AHS fund variables is significant and has a positive effect on workforce capacity). Incorporating the same additional independent variables in the Laplacian prior model is a potential area of further work.

Because the β components converge for both models, the samples are, in fact, draws from the posterior in both cases. For the Ridge regression sampler, only the β components converged, and the τ^2 component did not converge. Although all components would converge ideally, it is most critical to this study that the β components converge. These results point to the opportunity for further work that can be done to extract more information from the model coefficients. For instance, a future study could involve adding independent variables to help model performance.

Initially, each of the Gibbs samplers were run with a penalty on the intercept. Considering that the intercept is a vector comprised of 1s, it is unnecessary to penalize it, and the final samplers were changed so that there was no penalty on the intercept. This adjustment altered the results of the sampler, further indicating that the β components reach convergence.

Given recent events, such as the global COVID-19 pandemic, understanding when financial assistance is beneficial for certain capacities is critical if the global community is looking to better identify and address public health emergencies. In the case of workforce capacity specifically, this study unfortunately finds that countries facing a high burden of diseases are generally the ones without a sufficiently staffed workforce. Although these models did not result in conclusive findings related to disbursed AHS funds and workforce capacity, it warrants additional research given that donor financial assistance is a relatively easy way of aiding in a response.

To encourage positive change, both the evidence supporting that change and the desire of those who have the ability to make change need to be present. The JEE is a critical way through which a country's preparedness to prevent, detect, and respond to public health risks can be measured, and having this measurement framework is important in helping countries identify areas of strength and weakness. However, a measurement framework alone is not enough to encourage positive change. Leveraging the advanced tools available to us to understand when financial assistance disbursal is most beneficial for certain capacities can help countries more efficiently leverage the financial assistance available to them.

References

J. H. Albert and S. Chib. 1993. Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88(422):669–679.

M. R. Boyce, M. J. Meyer, J. D. Kraemer, and R. Katz. 2021. Financial assistance for health security: Effects of international financial assistance on capacities for preventing, detecting, and responding to public health emergencies. *International Journal of Health Policy and Management*, pages 1–8.

M. J. Meyer. 2022. Math 640 notes set ii.

WHO. 2019. Workforce development / human resources core capacity for health security: Findings from spar & jeewho sear member states.

A Derivation of the full posterior (Laplace prior)

Assuming the existence of a latent continuous random variable $Z_i \sim N(\vec{x_i}'\vec{\beta}, 1)$, such that $Y_i = j$ if $\gamma_{j-1} < Z_i \leq$