Bayesian Study of Sleep Disorder Diagnosis

Link to code

Rose Eluvathingal Muttikkal
Skye Jung
Grace Zhang
Fall 2023

1. Problem Description

Predicting sleep disorder diagnoses is essential because timely identification can lead to early intervention, reducing the risk of associated chronic conditions and improving overall health and productivity. With the prevalence of sleep disturbances among students and working professionals—often due to stress and other factors—reliable prediction models can play a crucial role. They can alert individuals to potential disorders before serious consequences develop, help foster better sleep habits, and ultimately enhance the quality of life. We will use Bayesian approaches to predict the occurrence of sleep disorders based on lifestyle and sleep quality measurements.

We aim to accomplish three goals - describe diagnosis predictors uncertainty, evaluate regression parameters for predictors, and assess model predictions.

2. Data Description

The dataset, consisting of 374 observations and 13 variables, was collected from Kaggle. Before building models, we pre-processed the data by checking the data types, exploring the different categories under each categorical variable, and checking for missing and NA values.

Next, we categorized all object variables and transformed all variables. We used z-scale to standardize all numeric variables to have the same scales with a mean of zero and a standard deviation of one. In addition, one of the numeric variables, *Blood_Pressure*, was transformed into a categorical variable based on the categorization of blood pressure used by the American Heart Association [1]. We did this by splitting the variable into two categorical variables - *Systolic_BP* and *Diastolic_BP*. We then re-defined blood pressure based on mmHg values and created a new categorical variable called *BP_Category*, which comprises four categories: normal, elevated, hypertensive stage 1, and hypertensive stage 2. Then, we converted the response variable, *Sleep_Disorder*, into a binary variable, where 1 indicates having a sleep disorder and 0 indicates not having a sleep disorder. We also cleaned the *BMI_Category* variable by replacing `Normal Weight` with `Normal` to make all elements have a consistent name. Lastly, we performed one-hot coding for all four categorical predictors, and an updated dataset was finalized to use for model construction.

3/4. Probability Model & Approach

We used Bayesian logistic regression with a No-U-Turn Sampler (NUTS) to predict sleep disorder diagnoses. With a binary response variable, the presence or absence of a disorder, logistic regression seemed like a natural choice. Choosing NUTS as our sampling method was ideal for our task due to its efficiency with numerous predictors in high-dimensional datasets. Once the foundation of our models was determined, we used three different methods: a main effects model, a simpler model, and a Bayesian model averaging (BMA) approach.

Our main effects model, Figure 1a, was a comprehensive approach, incorporating all 11 predictors: gender, age, occupation, sleep duration, sleep quality, physical activity level, stress level, BMI category, blood pressure level, heart rate, and daily steps. We decided to use a normal prior distribution with a mean of 0 and a standard deviation of 10. A normal distribution is widely used in cases where we lack strong prior knowledge. Normal distributions are symmetric, centered around the mean, and decrease as you go in either direction from the mean. A normal distribution is a convenient choice, especially for modeling our project problems where we had no prior knowledge of sleep diagnoses. A sizable standard deviation of 10 was chosen to reflect our uncertainty about the true values of the regression coefficients and allow the data to have a stronger influence. Our use of NUTS was justified as we included all 11 predictors in our main effects model, as NUTS is efficient in sampling high-dimensional parameter spaces.

To evaluate our main effects model, we used several methods to ensure the reliability of our model and results. Firstly, our prior predictive check revealed that the observed distribution of our response variable closely mirrored the prior predictive distribution, as shown in the bar charts in Figure 2. This similarity is crucial as it justifies our choice of the prior, as the observed data distribution resembling the prior predictive distribution suggests that our priors are appropriately calibrated and not an unrealistic assumption for our model. Choice or prior can significantly influence the results the model gives, so it is

Rose Eluvathingal Muttikkal
Skye Jung
Grace Zhang
Fall 2023

very important that our prior was proved appropriate. Furthermore, our trace plots analysis, from Figure 3, displayed stationary behavior, and the distributions of the sampled values overlapped neatly. Stationarity in trace plots implies that our sampling process is stable over iterations, and the overlapping distributions indicate that the sampler is exploring the parameter space effectively. Our convergence checks summary, Figure 4, also outputs R-hat values of 1.0 for all parameters, indicating that the variance within each chain of samples is similar to the variance across the chains. This means that our sampling process is adequately exploring the entire parameter space, and our model seems to be reaching a true representation of the posterior distribution.

Our Bayesian p-value plot, Figure 5, showed that the blue line, representing the observed data, was mainly within the gray bands, which denote the range of values predicted by the model, around the value of 1. This suggests that our model predictions are consistent with the observed data. Our posterior predictive check further affirmed our model's validity. Figure 6, shows our observed values lined up very closely with the posterior predictive means. This indicates our model's strong predictive accuracy. It demonstrates that our model, given our data and prior choices, can generate predictions that closely resemble the actual observed outcomes. This not only proves our model's performance but also its ability to predict outcomes outside of the data, which is very crucial in the context of sleep disorder diagnoses, where accurate predictions can be impertinent.

Our simpler model, Figure 1b, was a simplified model that focused on the significant factors that we identified from our main effects model: age, sleep quality, and stress level. We decided to use the same prior, Normal distribution with a mean 0 and standard deviation of 10, with the same justification as our main effects model. The simpler model aimed to create a model that balanced simplicity (lower number of predictors) and predictive power, utilizing only found-significant predictors. We again utilized the NUTS sampler for its efficiency in exploring the parameter space, while still ensuring reliable estimation even in a simpler model framework.

Evaluating our simpler model, we used the same evaluation methods as used for the main effects model. Our evaluation results mirrored the results of the main effects model, while also displaying the validity and reliability of our second approach. The prior predictive checks once again showed a close alignment between the observed data distribution and the prior predictive distribution. The trace plots displayed stationary behavior, and the distributions of the sampled values were well-overlapped. The R-hat values were consistently 1.0, suggesting good convergence. Similarly, in the Bayesian p-value analysis, the observed data were mostly within the range, indicating a good fit. Finally, the posterior predictive checks displayed the observed values were lined up with the posterior predictive means, proving the model's predictive accuracy. These consistent results across all evaluation methods not only confirm our prior choice is appropriate and is sampling well but also confirm the predictive power of both our main effects model and simpler model.

To compare our two models, we utilized their Watanabe-Akaike Information Criterion (WAIC) and Leave-One-Out (LOO) values. The WAIC values for the main effects and simpler models are -119.833132 and -125.041895, respectively. The simpler model's lower WAIC suggests it has a slightly better fit than the main effects model. This is further supported by the LOO values, where again, the simpler model (-125.057167) outperforms the main effects model (-120.026625). Regarding the Area Under the Curve (AUC) values, the main effects model's is 0.9, and the simpler model's is 0.7. These AUC values, which assess the models' ability to distinguish between the binary classes (presence or absence of a sleep disorder), suggest that both models perform substantially better than random guessing. However, the higher AUC of the main effects model indicates its superior performance in classification tasks. These values suggest that while the simpler model exhibits a slightly better fit and predictive accuracy as indicated by WAIC and LOO, the main effects model displays greater classification capability, as shown by its higher AUC value.

Our final approach is Bayesian model averaging, which combines predictions from both our main effects and simpler models, using the Weighted Akaike Information Criterion (WAIC) values as weights. This approach fits our objective of predicting sleep disorder diagnoses, as it handles model uncertainty and synthesizes insights from both our comprehensive and simplified models. By using WAIC values to determine the weights in our combined model, we can ensure that we have a balance of precision/good fit

to the data and model simplicity. This approach allows us to benefit from the broad view of the main effects model, capturing the full effects of all the predictors in the model, while also utilizing the simplicity of the simpler model.

5. Results

To comprehensively analyze the relationship between predictor variables and the response, we employed a range of techniques, including pair plots, a correlation matrix, a heatmap, stacked bar graphs, and box plots. Strong correlations among various factors such as age, sleep duration, sleep quality, physical activity, daily steps, stress level, and heart rate were seen in the heatmap (Figure 7). Despite observing these correlations, we decided to include all predictor variables in our main effects model to evaluate the significance of all predictors. We used stacked bar graphs to explore the association between categorical predictors and the binary response variable. The stacked bar plots highlighted certain occupations that exhibited a higher prevalence of sleep disorders (Figure 8a). In addition, individuals classified within the overweight and obese BMI categories, those with hypertension stages 1 and 2, and females displayed a higher proportion of sleep disorders compared to other groups (Figure 8b-d). Using box plots, we examined the relationship between numeric predictors and the binary response. There were noticeable differences in median values between individuals with and without sleep disorders. Specifically, we observed a higher median age, lower median sleep duration and physical activity levels, higher median heart rate, and lower median daily steps among those experiencing sleep disorders (Figure 9).

Within the main effects model encompassing 11 predictors, only three exhibited statistical significance. This outcome may be due to the correlations observed among various predictors in the heatmap analysis. Nevertheless, it was somewhat unexpected given our initial exploratory data analysis, which revealed relationships between more than three predictor variables and the response variable. The significant predictors identified were age, stress level score, and sleep quality score. The simpler model only uses these significant predictors. The mean values for the regression parameters were similar across both models. However, the simpler model had a smaller 95% highest density interval (HDI), suggesting reduced uncertainty compared to the main effects model. The regression parameter coefficients in both models indicated certain trends. Specifically, the estimated odds of experiencing a sleep disorder were found to increase with age, decrease with higher sleep quality scores, and surprisingly, also decrease with higher stress level scores. This unexpected finding contradicted our initial assumption that increased stress levels would elevate the likelihood of experiencing a sleep disorder. This paradoxical result could be partly explained by insights from the boxplot analysis, which displayed a lower median stress level among those with sleep disorders, and a similar range of stress level scores observed across both groups. It's essential to note that this interpretation might be unique to our dataset and may not be universally applicable. For a detailed interpretation of the regression parameters, please refer to Table 1, where coefficient values are interpreted for a one-standard-deviation increase or decrease in the z-scaled numeric variables.

To evaluate the uncertainty associated with diagnosing sleep disorders based on predictive factors within our model, we examined the 95% HDI for the probability of experiencing a sleep disorder derived from the output of our Bayesian regression models for the significant predictors (Figure 10). When comparing the probabilities between the main effects and a simpler model, we observed a narrower 95% HDI in the simpler model, suggesting lower overall uncertainty. This aligns with our earlier observations of a narrower 95% HDI for the regression parameters from the simpler model.

Across both the main effects and simpler models, we noted heightened uncertainty at the tails. Regarding predictor trends, both models revealed similar associations: as age increases, the probability of having a sleep disorder also rises. Examining sleep quality scores, we observed a declining probability of having a sleep disorder with increasing scores up to 7.5, beyond which the probability exhibited an upswing. Similarly, analyzing stress levels, we identified a range between scores of 4.5 and 5.5 characterized by a low probability of a sleep disorder. Preceding this range, the probability of having a sleep disorder decreased with increasing stress levels, whereas in the post-range, the probability increased with higher stress scores. The 95% HDI for the probability of having a sleep disorder uncovered a more

Rose Eluvathingal Muttikkal
Skye Jung
Grace Zhang
Fall 2023

intricate interplay between these significant predictors and response variables than what was discernible solely from interpreting regression coefficients.

Additionally, we evaluated the Bayesian Model Averaging (BMA) approach by comparing the HDI and means across all three models (Figure 11). Our comparison indicated similar HDI widths and means among the three models, suggesting a well-balanced performance of the BMA.

6. Conclusions

Our team set out to predict the probability of having a sleep disorder based on lifestyle and sleep quality measurements utilizing three distinct Bayesian models. We were able to successfully meet our outlined objectives. We determined the uncertainty surrounding sleep disorder diagnoses for lifestyle and sleep quality measurements. We assessed the significance of regression parameters within our models, and interpreted the parameters within the context of the problem. Additionally, we evaluated model predictions employing information criterion, ROC curve analysis, and accuracy metrics. The resulting models offer a framework for accurately estimating the probability of experiencing a sleep disorder, factoring in age, sleep quality score, and stress level score while accounting for uncertainty. However, it's important to note certain limitations in our analysis, notably the absence of interaction terms and nonlinearity incorporation within the model structure.

Bayesian Study of Sleep Disorder Diagnosis

Link to code

Rose Eluvathingal Muttikkal
Skye Jung
Grace Zhang
Fall 2023

7. References

1. American Heart Association. (2023) *Understanding blood pressure readings*, https://www.heart.org/en/health-topics/high-blood-pressure/understanding-blood-pressure-readings
2. PYMC Community. *Model averaging*, *PyMC*. https://www.pymc.io/projects/examples/en/latest/diagnostics_and_criticism/model_averaging.html
3. Tharmalingam, L. (2023) *Sleep health and lifestyle dataset*, *Kaggle*. https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset

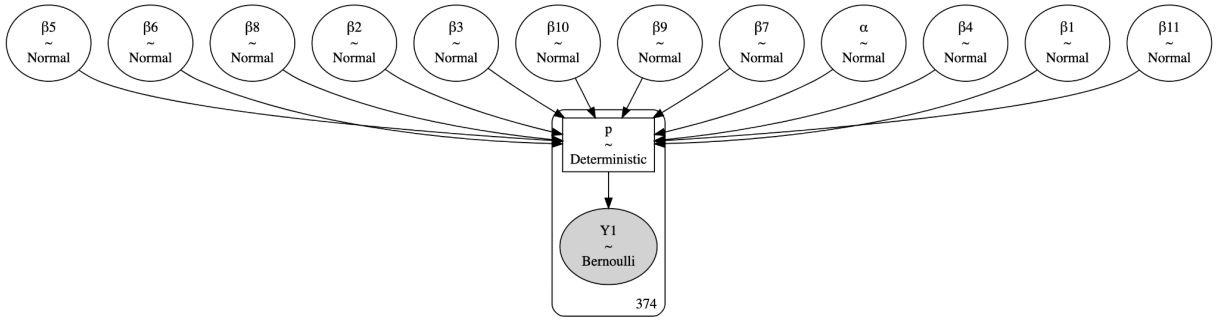
Bayesian Study of Sleep Disorder Diagnosis

Link to code

Rose Eluvathingal Muttikkal
Skye Jung
Grace Zhang
Fall 2023

8. Appendix

a)



b)

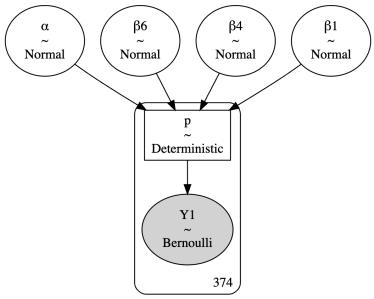
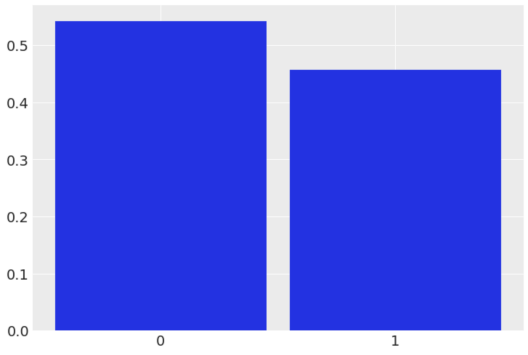


Figure 1: a) Main effects model, incorporating all 11 predictors; b) simpler model, featuring significant predictors from the Main Effects Model

a)



b)

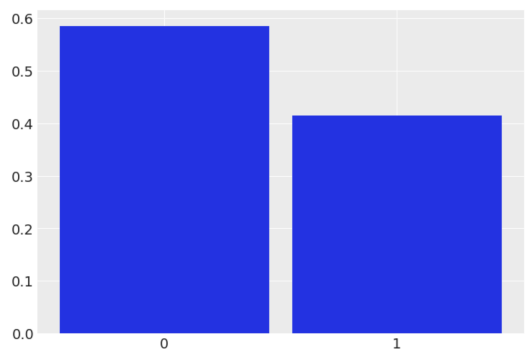


Figure 2: a) Bar chart displaying prior predictive distribution for main effects; b) bar chart displaying observed response variable

Bayesian Study of Sleep Disorder Diagnosis

Link to code

Rose Eluvathingal Muttikkal
Skye Jung
Grace Zhang
Fall 2023

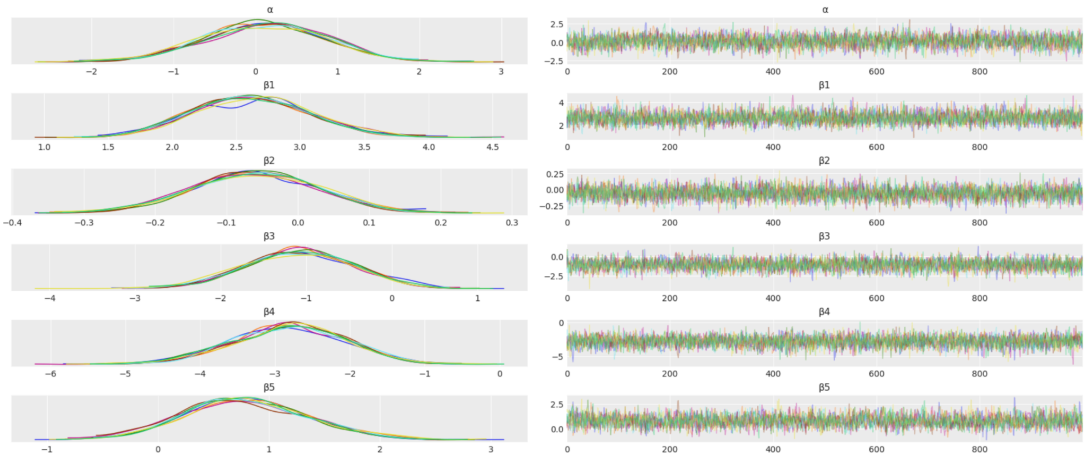


Figure 3: Trace plots of the main effects model

a)

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
α	0.114	0.748	-1.264	1.498	0.011	0.008	4260.0	5200.0	1.0
β1	2.627	0.464	1.800	3.517	0.007	0.005	4986.0	5145.0	1.0
β2	-0.059	0.090	-0.227	0.112	0.001	0.001	5932.0	5526.0	1.0
β3	-1.080	0.634	-2.222	0.147	0.009	0.006	5577.0	4909.0	1.0
β4	-2.815	0.759	-4.232	-1.426	0.010	0.007	5730.0	4870.0	1.0
β5	0.797	0.559	-0.229	1.872	0.009	0.007	3685.0	4260.0	1.0
β6	-2.083	0.653	-3.335	-0.920	0.010	0.007	4557.0	5109.0	1.0
β7	-0.507	0.592	-1.593	0.613	0.008	0.006	4865.0	5674.0	1.0
β8	0.684	0.447	-0.204	1.490	0.007	0.005	3794.0	4310.0	1.0
β9	-0.582	0.600	-1.751	0.500	0.010	0.007	3717.0	4365.0	1.0
β10	-0.012	0.413	-0.792	0.751	0.005	0.004	6952.0	5920.0	1.0
β11	-0.633	0.659	-1.815	0.676	0.009	0.007	4916.0	5533.0	1.0

b)

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
α	-0.676	0.184	-1.013	-0.320	0.002	0.002	5701.0	5120.0	1.0
β1	2.541	0.255	2.104	3.051	0.004	0.003	4393.0	4527.0	1.0
β4	-3.637	0.432	-4.456	-2.826	0.007	0.005	3618.0	4278.0	1.0
β6	-1.494	0.407	-2.261	-0.738	0.006	0.005	3958.0	4304.0	1.0

Figure 4: a) Summary output of main effects model trace, displaying $\widehat{R}$ values of 1.0 and ESS values greater than 400; b) summary output of simpler model trace, displaying $\widehat{R}$ values of 1.0 and ESS values greater than 400. Summary output shows samples of both models converge well.

Bayesian Study of Sleep Disorder Diagnosis

Link to code

Rose Eluvathingal Muttikkal
Skye Jung
Grace Zhang
Fall 2023

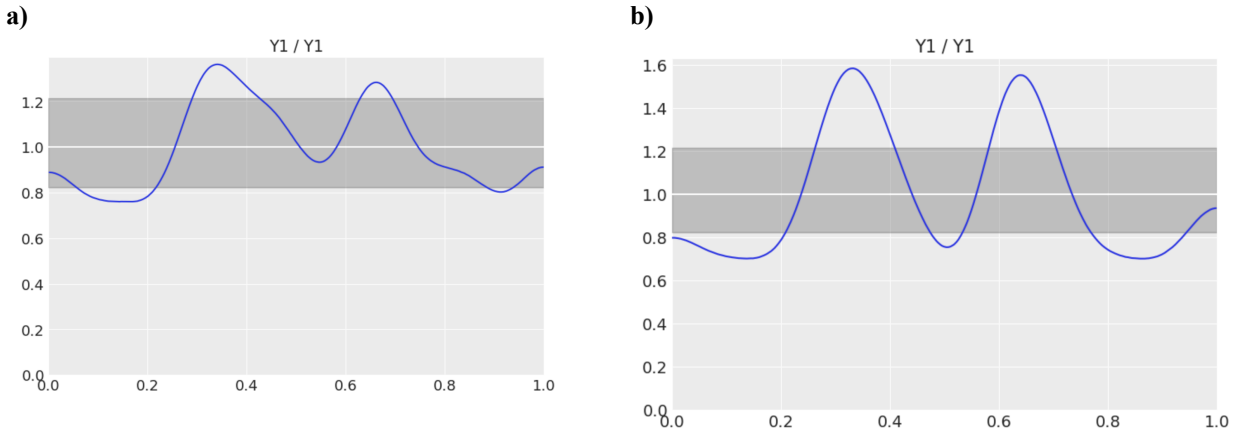


Figure 5: a) Bayesian p-value plot of the main effects model; b) bayesian p-value plot of the simpler model. The model fit for the main effect model appears to be better than the simpler model.

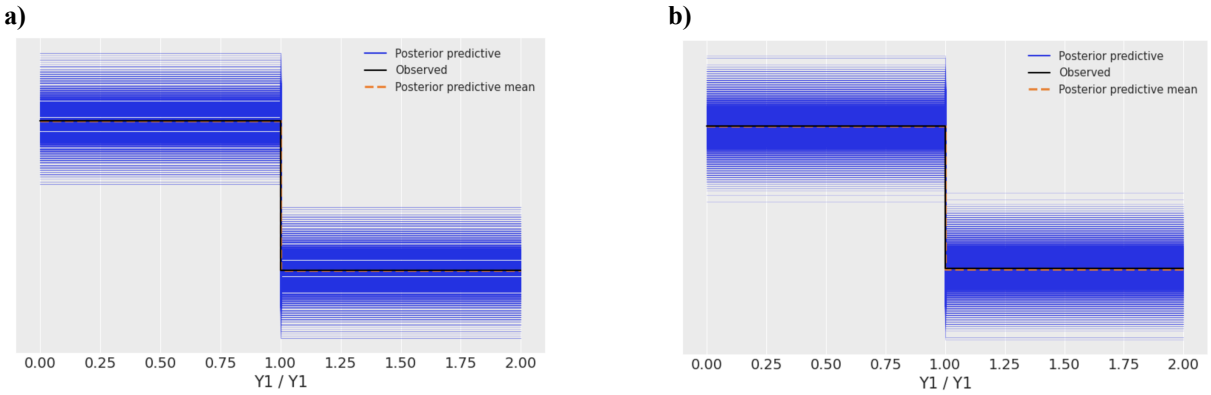


Figure 6: a) Posterior predictive plot of the main effects model; b) posterior predictive plot of the simpler model. The posterior predictive means for both models align with the observed data

Bayesian Study of Sleep Disorder Diagnosis

Link to code

Rose Eluvathingal Muttikkal
Skye Jung
Grace Zhang
Fall 2023

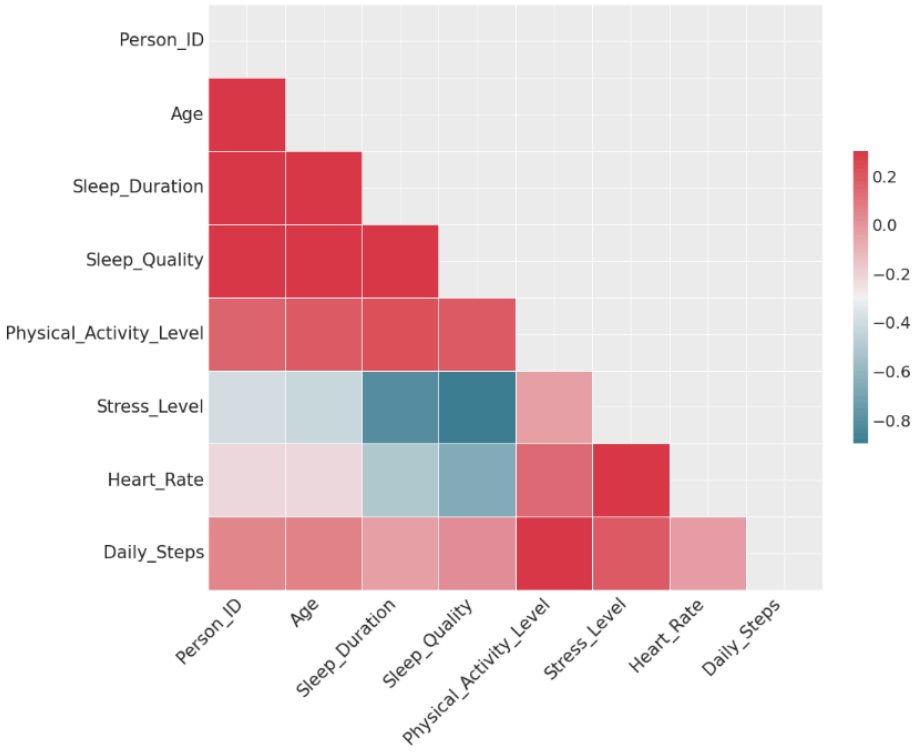
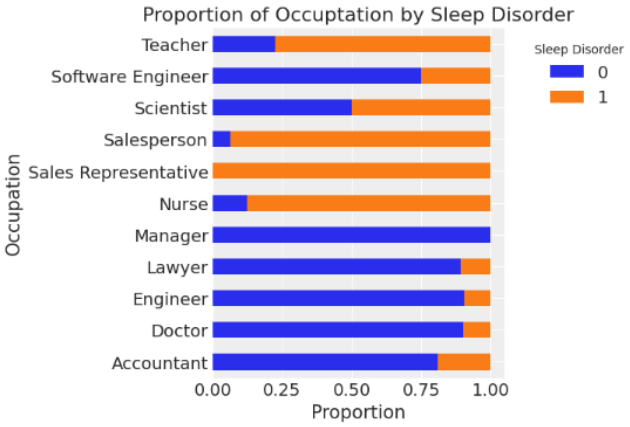
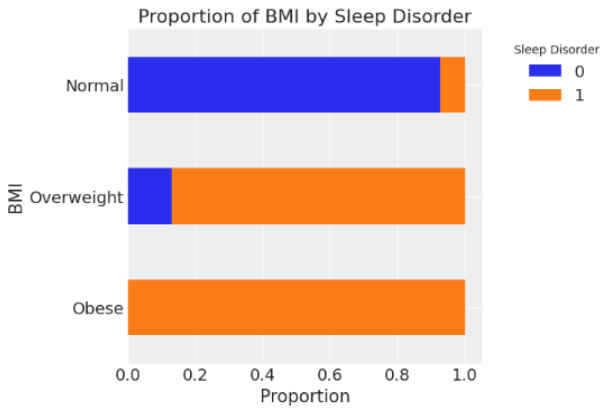


Figure 7: Heatmap displaying the correlation between all numeric variables, red indicates a strong correlation, and blue indicates a weak correlation

a)



b)



Bayesian Study of Sleep Disorder Diagnosis

Link to code

Rose Eluvathingal Muttikkal
Skye Jung
Grace Zhang
Fall 2023

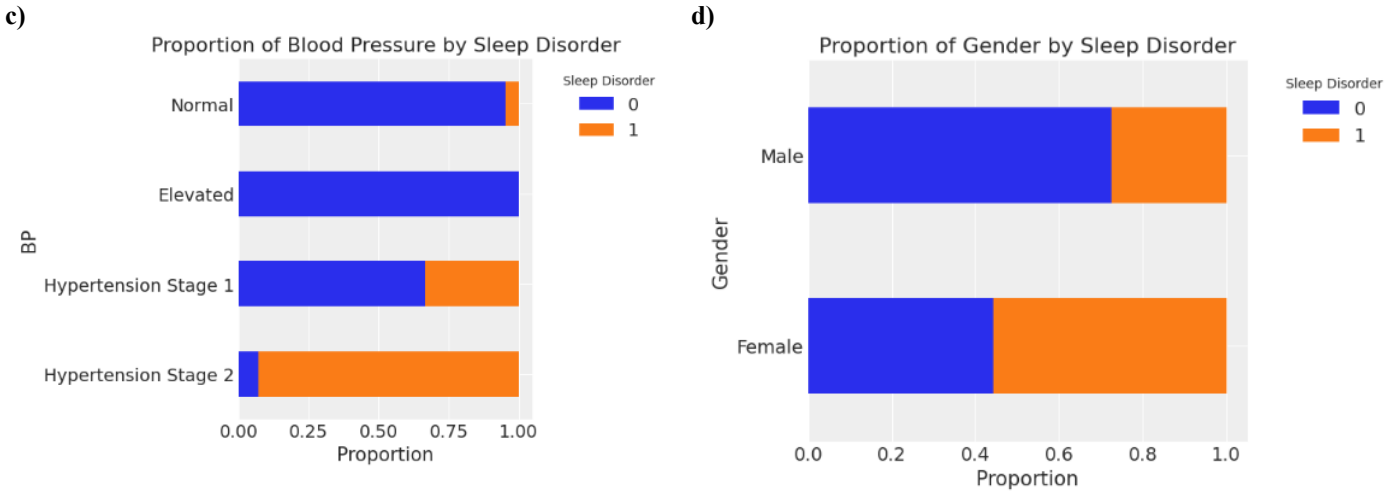


Figure 8: a) Vertically stacked bar graph displaying the proportion of occupation by sleep disorder; b) vertically stacked bar graph displaying the proportion of body mass index (BMI) category by sleep disorder; c) vertically stacked bar graph displaying the proportion of blood pressure (BP) category by sleep disorder; d) vertically stacked bar graph displaying the proportion of gender by sleep disorder

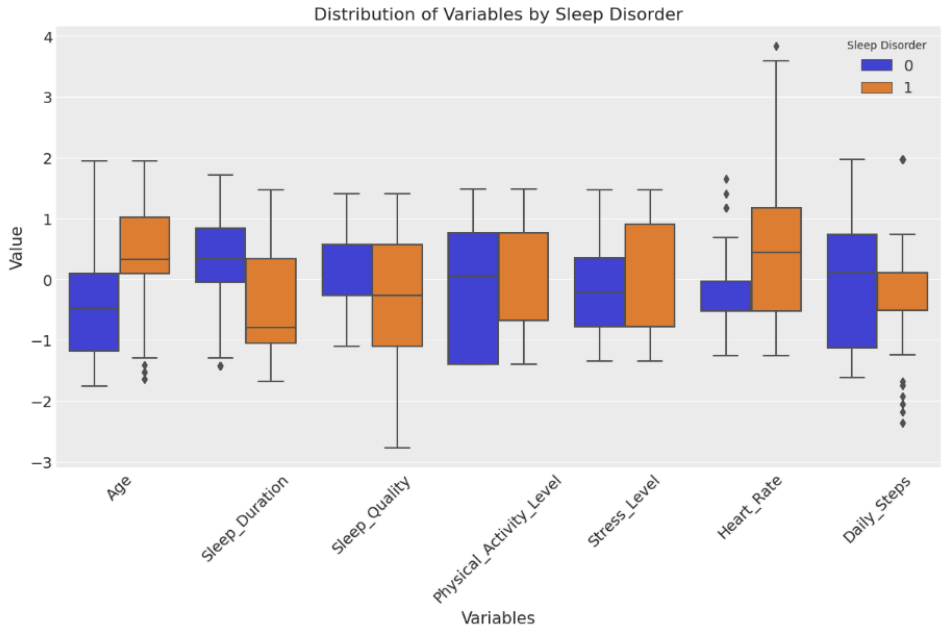


Figure 9: Boxplots of all z-scaled numeric predictors by sleep disorder category. Differences between median values for the response are present for most numeric predictors.

Bayesian Study of Sleep Disorder Diagnosis

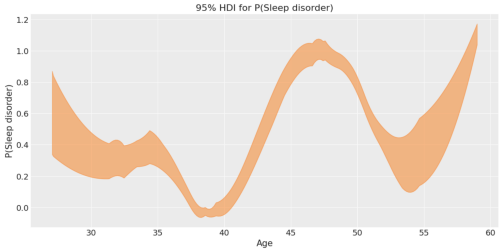
Link to code

Rose Eluvathingal Muttikkal
Skye Jung
Grace Zhang
Fall 2023

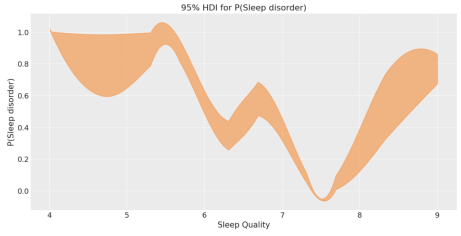
Table 1: The mean, 95% HDI, and interpretation of significant regression parameters, age, sleep quality score, and stress level score for the main effects and simpler models

Model	Predictor	Mean	95% HDI	Interpretation
Main Effects	Age (β_1)	2.5	1.8 to 3.3	The estimated odds for having a sleep disorder increases by exp(2.5) = 12.2 for 1 STD increase in age, when controlling for other predictors
Main Effects	Sleep Quality Score (β_4)	-3.2	-4.6 to -1.8	The estimated odds for having a sleep disorder decreases by exp(3.2) = 24.5 for 1 STD increase in sleep quality score when controlling for other predictors
Main Effects	Stress Level (β_6)	-2	-3.2 to -.077	The estimated odds for having a sleep disorder decreases by exp(2) = 7.3 for 1 STD increase in stress level score, when controlling for other predictors
Simpler	Age (β_1)	2.5	2.1 to 3.1	The estimated odds for having a sleep disorder increases by exp(2.5) = 12.2 for 1STD increase in age, when controlling for other predictors
Simpler	Sleep Quality Score (β_4)	-3.6	-4.5 to -2.8	The estimated odds for having a sleep disorder decreases by exp(3.6) = 36.5 for 1 STD increase in sleep quality score, when controlling for other predictors
Simpler	Stress Level Score (β_6)	-1.5	-2.3 to -.074	The estimated odds for having a sleep disorder decreases by exp(1.5) = 4.5 for 1 STD increase in stress level score, when controlling for other predictors

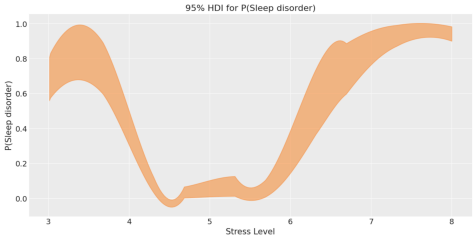
a)



b)



c)



Bayesian Study of Sleep Disorder Diagnosis

Link to code

Rose Eluvathingal Muttikkal
Skye Jung
Grace Zhang
Fall 2023

Figure 10: 95% HDI for the probability of having sleep disorder for significant predictors, a) age, b) sleep quality score, c) stress level score for main effects model; 95% HDI for the probability of having sleep disorder for significant predictors, d) age, e) sleep quality score, f) stress level score for simpler model

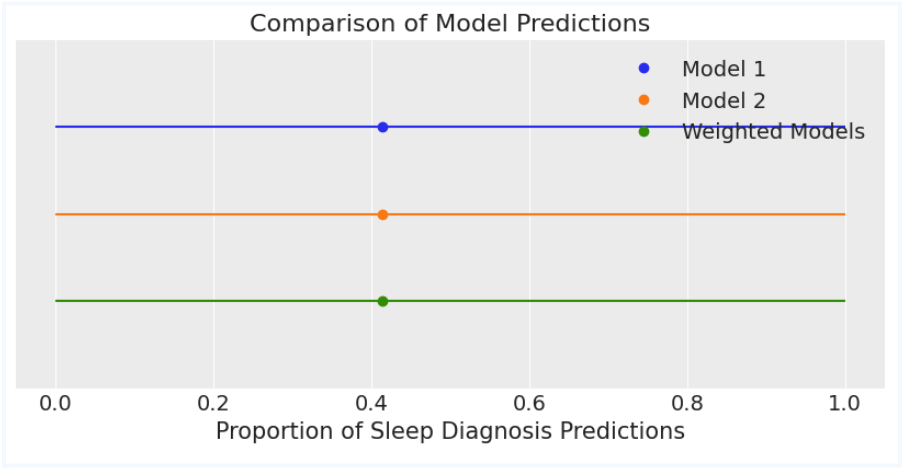


Figure 11: Comparison of HDI and mean of three models: main effects, simpler, and BMA model