Clayton Bass
Due 11/19/21

MATH 50 Final Project
Professor Levien

Predicting Resting Blood Pressure Based on Health and Demographic Metrics

Introduction

Blood pressure is a crucial metric to track and predict in those who have pre-existing health conditions (such as major, common problems like obesity, diabetes, cancer, chronic cigarette smoking, etc.) (Oparil et al.). Intuitively, adequate blood pressure is important in the human body since blood must provide sufficient oxygen and nutrients to our muscles and organs that are working hard throughout the day. But if this blood pressure becomes too high at rest (due to many risk factors like obesity, an inactive lifestyle, cigarette smoking, diabetes, etc.), then blood vessels can become damaged, and the risk of major cardiovascular events like heart attack or stroke rises significantly, per a published paper by Oparil et al. from 2018 (Oparil et al.). Blood pressure is technically considered too high when the systolic pressure (blood pressure during a heartbeat) is above 130 mmHg, or when the diastolic pressure (blood pressure between beats) is above 80 mmHg (Oparil et al.).

The goal of this project is to find out—using both frequentist and Bayesian linear regression models and assumptions—what risk factors might predict this dangerous bodily phenomenon (and the degree to which this phenomenon occurs depending on the values of such predictors). This is a crucial question: if we can figure out the main risk factors of excessive blood pressure at rest—and further quantify this relationship—then in the future it will be easier to modify one's lifestyle and medications to mitigate this harmful process (Oparil et al.). If this response cannot be controlled (or is ignored), then major cardiovascular diseases and events will likely only continue to worsen and become more prevalent in this country as exercise and dietary habits fail to improve; the ultimate cost to humanity posed by ignorance of the importance of blood pressure management is dauntingly large (Oparil et al.).

As someone who has struggled with inborn heart problems (and wants to control all preventable cardiovascular risk factors in the future), the project serves to provide practical insight into how one may become a healthier person in general. It will be important to communicate some of these pieces of information to classmates, friends, and family. Too many people of young age assume they are totally unaffected by these health risks (since it is true that these issues are more common in older populations), but in reality, any of these health problems—particularly diabetes and high blood pressure—can strike at any age, even in one's teens, so managing one's health from early in life is critical to maximizing lifespan and quality of life overall.

This project uses the UCI Heart Disease Dataset, which contains 303 patients' key health indicators, ranging from basic cardiovascular features (like resting blood pressure, cholesterol levels, diabetes history) to age and sex (Janosi et al.). While there are approximately 76 health metrics provided, only about 14 are actually usable in the dataset, and we narrow down this list even more in order to have only the most robust and interpretable predictors, since we have only 303 patients' data.

Methods

We perform a mixture of frequentist (ordinary least-squares) and Bayesian regression as two different ways of assessing the same problem throughout this paper. I placed the .data file in a GitHub repository as downloaded from the UCI Heart Disease Dataset; the raw file can be found at `https://raw.githubusercontent.com/claytonbass/MATH-50-Final-Project/main/processed.cleveland.csv`. The initial analysis includes a frequentist linear regression model that uses all of the relevant provided predictors. We ran an ordinary least-squares regression model (with intercept included) on the entire dataset, labeling the columns as follows (as we will refer to them from now on): ['age','sex','chest pain type','resting bp','cholesterol','fasting blood sugar > 120 mg/dl','resting ECG','maximum heart rate','exercise-induced angina','ST depression in exercise','slope of peak exercise ST segment', '# vessels colored (fluoroscopy)','heart defect','diagnosis of heart disease']; this is via statsmodels.OLS. We then examined the coefficients and their p-values in a table to determine which were significant and which were not. It became clear that the model would be minimally interpretable with so many coefficients and possible interactions, and so we made the model more parsimonious.

In any statistical analysis, it is of the utmost importance that we understand the variables underlying the problem as well as possible. Some of these variables are rather confusing and require advanced medical training to truly (and fully) understand and explain at a level that others will understand. Therefore, I decided to consider a subset of these initial predictors that I felt not only better suited the limited number of data points but that I would also be able to explain better given my current medical knowledge. I next ran an ordinary least-squares regression on the predictors of age, sex, cholesterol, blood_sugar (which takes a value of 1 if greater than 120 mg/dL, indicating diabetes), cholesterol age, sex age, the latter two of which are interaction terms. We used only this set of predictors for a few reasons (even though age and blood_sugar are found to be the only two statistically significant predictors here). Age is known to be probably the most reliable risk factor for the development of high blood pressure (Oparil et al.). Furthermore, having problematic fasting blood sugar is a critical factor in the development of high blood pressure, because a value of 1 for this binary indicator variable indicates the presence of diabetes in most cases (where high blood sugar levels damage blood vessels, reducing their elasticity and inducing a response where resting blood pressure is higher than it should be) (Oparil et al.). It is intuitive that these two predictors were included. Men are known to

have higher resting blood pressure than women, so we included sex for this reason (Oparil et al.). However, women generally develop higher resting blood pressure when they are more elderly as compared with men, so it is possible that these two effects cancel – this is precisely the reason that we have included the sex * age interaction term. Lastly, we explain the interaction term between cholesterol and age – cholesterol levels tend to rise with age no matter one's health status (and it is often inevitable that cholesterol levels rise a bit above the optimal level when one is more elderly), but the combined effect of old age (and therefore inelastic, damaged blood vessels) with higher cholesterol levels that physically narrow the blood vessels creates a vicious cycle that often contributes to high resting blood pressure (Oparil et al.).

After examining the coefficients and their p-values for this model, we examined our modelling assumptions to see if anything went wrong in our analysis. We examined a scatterplot of our residuals plotted against the average y-prediction (defined as the true value of y minus the residual), in order to determine whether there were any trends in the residuals that might indicate dependence of the residuals on the particular value of y in question. We also compared the distribution of residuals in a histogram to see whether they were normally distributed according to the noise parameter in our fitted model, running the Lilliefors test using statsmodels.stats.diagnostic.kstest_normal on our residuals to see if we can reject the null hypothesis that our residuals are sampled from a normal distribution.

We then shifted to Bayesian linear regression methods to see if any additional insights could be gained. Since we have only about 300 records in this dataset, it made sense to pursue Bayesian linear regression as well since we may not be remarkably confident in the parameters we are predicting, and instead quantifying our uncertainty in these parameters by examining their posterior distributions may be more rewarding. Since this model contained a total of six predictors (including the two interactions terms we discussed), it would be very challenging to develop individual prior distributions on each predictor's coefficient. Therefore, we use a more empirical method whereby we take the coefficients generated from the ordinary least-squares regression, using the point estimates as our prior means for each coefficient, and specifying the standard deviation as 10 times the standard error on each coefficient from the OLS model. Our linear regression model is, in general, given by $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_1 x_3 + \beta_6 x_1 x_2 + \epsilon$. Note x_1 is the age of the patient in years, x_2 is the sex of the patient (0 = female, 1 = male), x_3 is blood cholesterol of the patient in mg/dl, x_4 is an indicator of blood sugar being 120mg/dl fasting (meaning diabetes, where 0 means below and 1 means above). We also have two coefficients for our two interaction terms that have already been discussed. If we used anything less than 10 as our multiplicative factor for the standard deviation on the priors, then we would get an error where MCMC cannot proceed via PyMC3 because the gradients are way too small (so we were way too confident in our prior distribution). We then used ArviZ's plot_pair function in order to analyze the posterior correlations between our coefficients to see if any correlations appeared significant (given the credible interval for each parameter).

In order to make our model more interpretable and to see how a more parsimonious model with fewer predictors would compare (out-of-sample) to the model with six predictors, we developed a model that used only two predictors to predict resting blood pressure: age (in years), and a binary indicator of fasting blood pressure being above 120 mg/dL (meaning diabetes). We performed exactly the same process as above for this two-predictor model: we fitted and analyzed residual distributions for an ordinary least-squares model, and then ran a Bayesian regression model where we examined posterior correlations using ArviZ plot_pair.

Finally, in order to compare our models' efficacy and accuracy in unseen data, we used leave-one-out cross-validation via ArviZ.loo to see which model, between our six-predictor and two-predictor model, had a lower out-of-sample error on the test sets (collectively); this out-of-sample accuracy is generated using our Bayesian regression models. We also performed another form of cross-validation where we trained our two-predictor ordinary least-squares model on only 200 of the patients but tested on the remaining patients (about 100 of them) to see how well we could predict resting blood pressure (knowing that our model may struggle with individual prediction, even though it still tells us very important information about the relationships between age, diabetes status, and resting blood pressure). We plotted the predictions versus the true values in a scatterplot to see how well the predictions aligned. Given our final two-predictor model, we then went through the process of interpreting each coefficient for a practically useful final product.

Results

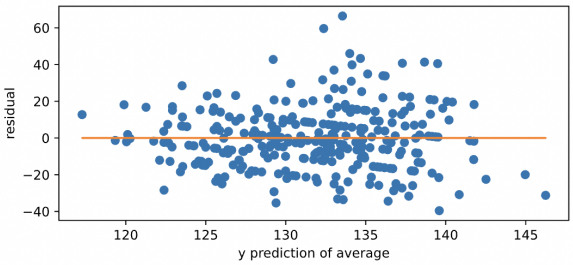
In our initial ordinary least-squares regression model with all 13 predictors included, only age and cholesterol had statistically significant coefficients (in addition to the intercept). None of the other coefficients were significant, including the more difficult-to-interpret ones (such as resting ECG/electrocardiogram, and ST depression in exercise, which both explain the degree of electrical abnormality in the heart during exercise). More info can be found in the notebook, but nothing more was done with this model.

Using our six-predictor model, we got slightly more convincing and interpretable results, although they were still not as interpretable as we would have liked. Specific coefficients and p-values can be found in the notebook, but the two significant coefficients were age, with a p-value of 0.025 and a coefficient of 1.20, and blood_sugar, with a p-value of 0.0096 and a coefficient of 7.12, with all other p-values above 0.15. It is intuitive that the two predictors of age and blood sugar are included and that they are statistically significant. The R-squared value is given by 0.173, which explains a small portion of the variation in y. However, interestingly, cholesterol and sex are not statistically significant. Men are known to have higher resting blood pressure

than women, so that aspect is surprising (and per our exploratory data analysis, it's clear that there is a fair balance of men and women in this dataset at a ratio of 2:1, although men outnumber them by a bit) (Oparil et al.). However, women generally develop higher resting blood pressure when they are more elderly as compared with men, so it is possible that these two effects cancel (Oparil et al.). This is precisely the reason that I have included the sex * age interaction term, but it is not significant, so that can be effectively excluded from the analysis. Having high cholesterol is a very well-known risk factor for heart disease, so the lack of significance of that coefficient is also confusing; however, high cholesterol is not nearly as implicated in the literature as a factor like blood sugar (diabetes), age, or sex in the development of high blood pressure (Oparil et al.). The interaction term between cholesterol and age is also not significant but is quite closer to being significant than sex * age – cholesterol levels tend to rise with age no matter one's health status (and it is often inevitable that cholesterol levels rise a bit above the optimal level when one is more elderly), but the combined effect of old age (and therefore inelastic, damaged blood vessels) with higher cholesterol levels that physically narrow the blood vessels creates a vicious cycle that often contributes to high resting blood pressure (Oparil et al.). However, the interaction term is not statistically significant in this model. There was no evidence of violation of linear regression assumptions when making the residual plots for this non-parsimonious model (this can be seen as the first and second figures in the **Appendix**).

Recall that our non-parsimonious linear regression model is, in general, given by $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_1 x_3 + \beta_6 x_1 x_2 + \epsilon$. Our final model using frequentist methods (least-squares linear regression) is given by $y = 59.395 + 1.199x_1 + 7.105x_2 + 0.168x_3 + 7.117x_4 - 0.0026x_1 x_3 - 0.156x_1 x_2 + \epsilon$, where the coefficients for x_1 and x_4 are the only ones that are significant. Note $\epsilon \sim N(0, 16.73)$. Note x_1 is the age of the patient in years, x_2 is the sex of the patient (0 = female, 1 = male), x_3 is blood cholesterol of the patient in mg/dl, x_4 is an indicator of blood sugar being 120mg/dl fasting (meaning diabetes, where 0 means below and 1 means above). We also have two coefficients for our two interaction terms that have already been discussed. It is very difficult to establish and subsequently justify specific prior distributions for each of the predictors in our model. Therefore, we set the mean for each of our prior distributions equal to the point estimate of the coefficients for each predictor from the frequentist model, and then adjust the standard deviation on the prior according to the standard error. From here, we can define our prior standard deviations as approximately 10 times the standard error (so that we are not imposing too much confidence in our prior estimates such that the result is essentially just what we would get from frequentist regression). If we do anything less than 10 as our multiplicative factor, then we get an error where MCMC cannot proceed via PyMC3 because the gradients are way too small (so we were way too confident in our prior distribution). In the posterior plot of correlations (which can be seen as the third plot in the **Appendix**), some pairs of coefficients are correlated (especially those related to the intercept term). However, the credible interval for these coefficients is small, which alleviates the concern that these correlations are a major problem.

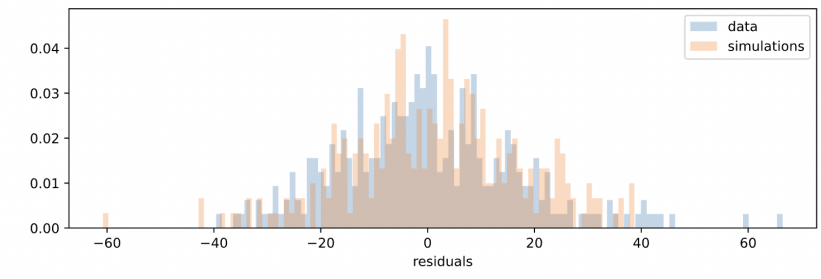
We now proceed to the parsimonious model using only age and blood sugar indicator as predictors. The frequentist model is given by the following: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon = 102.2355 + 0.5217x_1 + 7.0944x_2 + \epsilon$, where x_1 is age and x_2 is indicator of high blood sugar (0 or 1), and $\epsilon \sim N(0, 16.738)$. All the coefficients are statistically significant with a p-value rounded off to 0.000 (except for the coefficient of blood sugar, which has a p-value of 0.010). The standard errors on the intercept, age coefficient, and blood sugar coefficient are given by 5.887, 0.107, and 2.723, respectively. Although our coefficients are very significant, note that our R-squared value is only 0.1, so the majority of the variation in our blood pressure readings is not explained by our model. Even though the model is valid, it may not be ideal for making individual predictions since it's not capturing a sufficient amount of the variation in blood pressure values, which we will explore later. The R-squared value is about 0.17 for the six-predictor model, so while we explain more of the variation in y with more predictors, this is not unexpected, and the increase in R-squared is not so large that we would necessarily expect the additional four predictors to be worth the loss of model simplicity and interpretability that occurs in such a model. We test our model assumptions:



Notice that once again there does not seem to be a relationship between the y prediction of average and the residuals here. Even though we cannot immediately see any problems with the model just from the simple diagnostics that we have been running, it is still possible that biologically speaking we do not have a linear relationship between our predictors and the response variable of resting blood pressure. It is possible (although not clear or even indicated fully here) that there is a nonlinear relationship, especially since blood pressure tends to remain pretty stable through teenaged years and young adulthood, but commonly rises quite a bit during middle adulthood and to a large extent in older adulthood. In addition, the physical dynamics and relationship between high blood sugar and resulting resting blood pressure are not perfectly understood even by the most accomplished physicians and scientists, but could easily reflect a nonlinear relationship given the amount of damage that high

blood sugar exerts on blood vessel walls.

As before, we will also compare the distribution of residuals in a histogram to see whether they are normally distributed according to the noise parameter in our fitted model.



We get very similar results to before, both with respect to the p-value of the Lilliefors test and with respect to the actual visualized distribution (although here the p-value is a bit smaller, usually coming in at around 0.01 to 0.02). It's therefore hard to conclusively state that our residual samples are not generated according to a normal distribution with the given noise parameter, even though if we truly wanted to use a hard cutoff of 0.05 then we can technically reject the null hypothesis that these residual samples come from a normal distribution.

We will also do Bayesian regression in this model using the coefficients from the frequentist model as the information for our prior distribution. This is based on the following frequentist model:

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon = 102.2355 + 0.5217 x_1 + 7.0944 x_2 + \epsilon$, where x_1 is age and x_2 is indicator of high blood sugar (0 or 1), and $\epsilon \sim N(0, 16.738)$. There are no significant posterior correlations in our model if we plot the pairs of coefficients side-by-side. The sixth plot in the **Appendix** can shed more light on these relationships. One main takeaway: the credible interval for all of these coefficients (β_1 in particular) is quite small, so we should not be too worried about significant correlations between any of these posterior coefficients in practice.

Now, we perform leave-one-out cross-validation on our six-predictor and two-predictor Bayesian regression models to see which is better with unseen data. The information criterion used to decide the better model assigned a score (denoted "loo," see the fourth plot in the **Appendix** for the screenshot) of -1288.8 to the non-parsimonious model and -1289.11 to the parsimonious model, meaning that our parsimonious model is just barely less effective in prediction on unseen data. It's likely more valuable to use the more interpretable and simpler model that has coefficients that are not only statistically significant but also relatively straightforward in their interpretation. So, this means that I am likely basing my final conclusions on coefficients on the Bayesian and frequentist regression models that use only age and diabetes indicator (an indicator variable for blood sugar above 120) as predictors of resting blood pressure.

We also performed cross-validation where we trained on only 200 of the patients but test on about 100 of the remaining patients to see how well we can predict resting blood pressure (knowing that our model may struggle with individual prediction, even though it still tells us very important information about the relationships between age, diabetes status, and resting blood pressure). See the fifth plot in the **Appendix** for more info. As was somewhat expected given our very low R-squared value, our model struggled with individual predictions. We may be seeing the consequences of such great variation in resting blood pressure that cannot be explained by our model that uses only age and diabetes status. The mean squared error is given by 266; if we take the absolute value of every error, then the mean error is given by about 12.87, so we can expect to be about 13 mmHg off in our blood-pressure prediction on average, which does not seem terribly large. However, this does indicate that our model could use some work when it comes to individual prediction.

Finally, we can interpret the individual coefficients in our model to determine the relationship between our predictors and resting blood pressure.

Recall that our parsimonious model uses the following coefficients and error term:

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon = 102.2355 + 0.5217 x_1 + 7.0944 x_2 + \epsilon$, where x_1 is age and x_2 is indicator of high blood sugar (0 or 1), and $\epsilon \sim N(0, 16.738)$. We will provide interpretations of the two slope coefficients (since the intercept term does not make sense on its own anyway, as it is not really useful to examine the blood pressure of someone who is age 0, meaning a newborn infant, as high blood pressure in infants is remarkably rare, along with diabetes – although the intercept value of 102.2355 as the resting blood pressure for this age group actually does makes sense).

$\beta_1 = 0.5217$ means that, when comparing two people who have the same diabetic status (as either diabetic or non-diabetic), every increase in age by one year will result, on average, in an increase of 0.5217 in blood pressure. This intuitively makes sense – extrapolating a bit, for every 10 years gained in age, we would expect about a 5 mmHg average increase in blood pressure. This is in line with what is commonly known about rising resting blood pressure throughout life.

Perhaps more importantly and remarkably, $\beta_2 = 7.0944$ means that, comparing two people of the same age, diabetics have, on average, resting blood pressure that is 7.0944 mmHg higher than non-diabetics. This is striking: diabetes is clearly a strong risk factor for high blood pressure, and age-adjusted comparisons of blood pressure clearly indicate here that someone with diabetes (or simply excessively high fasting blood sugar for whatever reason) on average has significantly higher resting blood pressure

than someone with normal fasting blood sugar. This highlights the importance of diabetes management in the treatment of high blood pressure: having persistently elevated blood sugar can severely damage the blood vessels, causing their elasticity to decrease and consequently for resting blood pressure to rise (since inelastic blood vessels cannot expand to accommodate increased blood flow, thereby causing harmful high blood pressure). In addition, it is remarkable that age is such a striking risk factor for high blood pressure: it is almost inevitable that one's blood pressure will rise significantly over the course of many years, so these findings highlight that it is very important to keep track of blood pressure, especially as one ages, and also to ensure that blood sugar levels are kept in check to keep blood pressure secondarily in check.

Recall, though, that our model has an R-squared value of about 0.1, so while it may be a valid model that is truly picking up on relationships between age, diabetes status, and resting blood pressure, the model itself may struggle with making individual predictions since it fails to explain a majority of the variation in the response variable (it explains only a small portion of it), so there are likely some other parameters elsewhere out there that we are missing, or perhaps the relationship between these predictors and resting blood pressure is not truly linear.

Discussion

In summary, I answered the question that I set out to answer to a large degree, given that in my two-predictor model, the final statistically significant coefficients had a very clear interpretation and helped shed light on the relationship between age, fasting blood sugar, and resting blood pressure, and even shed some light on the relative range in which blood pressure values should lie in very young people.

In that sense, the project went well. However, given the complexity of this topic (from both a medical and a statistical perspective), there are a few different areas for improvement, or at least exploration, in future study. The fact that sex and cholesterol levels were not significant predictors (and that age and sex, along with age and cholesterol, were not significant interaction terms) was a bit surprising to me. It is known that women have much lower resting blood pressure than men on average, and high cholesterol is a significant risk factor for high blood pressure as well (Oparil et al.). However, the cholesterol levels in this dataset were actually uniformly quite high, and there is a 2:1 ratio of men to women in the dataset, so there are some factors that might be preventing these variables from exerting their full effect. Furthermore, the dataset is based strictly on a sample of about 300 people from Cleveland. Perhaps there is a need for data that is from a larger sample from a more diverse geographic and cultural landscape so that we get many more lifestyles and risk factors incorporated into our predictive model. Or perhaps we have redundant predictors, where cholesterol and diabetes risk are so intertwined and co-occurring that it does not make sense to have both in our model, and the interaction effect of age and cholesterol is not striking enough to be a statistically significant variable. However, simply because the statistical cutoff of $p = 0.05$ was not met for these variables does not completely exclude them from analysis. Even though the non-parsimonious model was difficult to interpret on its own, it's important to consider that sex and cholesterol are important metrics to track when treating and screening for high blood pressure (Oparil et al.).

This project has shown me just how complex research in the medical field can quickly become. Even when simplifying our model down to just age and diabetes status when trying to predict resting blood pressure, it still became clear that the model, while not complex, was struggling with individual prediction because human resting blood pressure is remarkably complicated and affected by a multitude of factors. In fact, since this resting blood pressure was taken upon admission to the hospital, the actual value provided may not have been indicative or reflective of the patient's underlying health conditions. This is for a variety of reasons. First, some patients experience (understandably) extreme anxiety upon admission to the hospital, and so their blood pressure can spike unreasonably (Oparil et al.); therefore, some of the blood pressure values may be artificially inflated in this dataset. Finally, many healthcare practitioners do not perfectly follow the proper method of taking blood pressure – oftentimes patients are stressed, have just had coffee, are standing and talking with the physician or nurse, and are wearing sleeves on their arms; all of these factors can collectively raise blood pressure (Oparil et al.). Therefore, lots of the variation in blood pressure seen is not explained by our model likely because there is a remarkably complex, yet-to-be-fully-understood relationship between genetics, lifestyle, and immediate situational circumstances (including blood pressure measurement methodology) that combine to make an individual blood pressure reading (Oparil et al.). Therefore, in the future, it may be very interesting to create a predictive model of blood pressure that can leverage predictor variables such as whether the patient has recently had caffeine, whether the patient was standing or talking during the collection, whether the patient was wearing sleeves, whether the patient was in acute distress (or was known to be overly comforted when in medical settings), and the particular device used to measure blood pressure. In reality, it is not practical for such a dataset to be collected on a large population, but I would assume that this would help solve some of the problems inherent in this model. In addition, for future methodologies and research avenues, simple logistic regression to predict whether or not someone has high blood pressure (as a binary indicator) is another potentially useful project to pursue, assuming a relatively simple model and set of predictors. It can be more interpretable to think in terms of binary outcomes rather than a continuous spectrum of blood pressures, where the difference between a systolic blood pressure of 122 mmHg vs. 126 mmHg may not be immediately clear (or have much practical meaning in reality). This provides an exciting, open-ended new direction for research into an important health-related topic at the intersection of medicine and statistics.

References

Janosi, Andras, M.D., et al. "Heart Disease Dataset." UCI Machine Learning Repository, UCI, archive.ics.uci.edu/ml/datasets/heart+disease. Accessed 27 Sept. 2021.

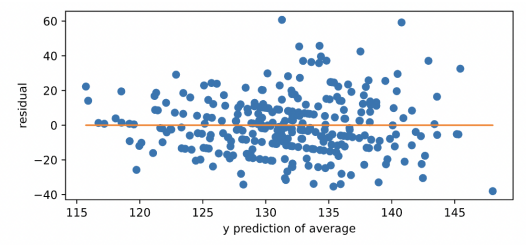
Oparil, Suzanne, et al. "Hypertension." Nat Rev Dis Primers. NIH, www.ncbi.nlm.nih.gov/pmc/articles/PMC6477925/. Accessed 12 Oct. 2021

Colab Notebook

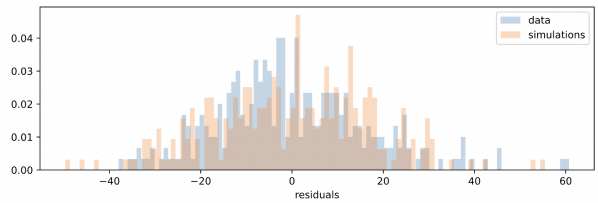
Here is the link to the Colab notebook used to generate the results and figures in this project: https://colab.research.google.com/drive/109Bv46AUlqbGmLhsdXJpCdRNcUbI8BES?authuser=1#scrollTo=VLOzs2hx3REx

Appendix

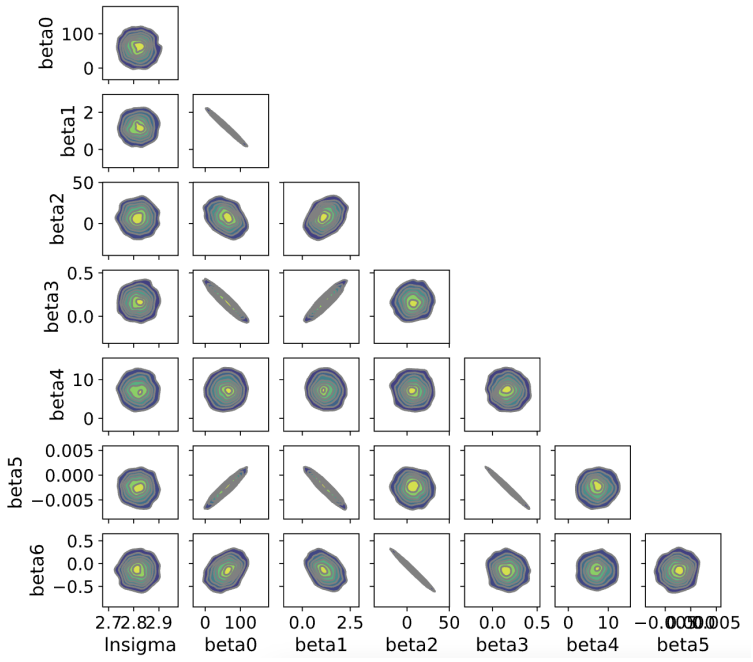
Here is the first residual plot for our non-parsimonious ordinary least-squares regression (for the residual distribution versus the mean y prediction):



Here is the second residual plot for our non-parsimonious ordinary least-squares regression (for our comparison of the distribution of residuals in a histogram to see whether they are normally distributed according to the noise parameter in our fitted model):



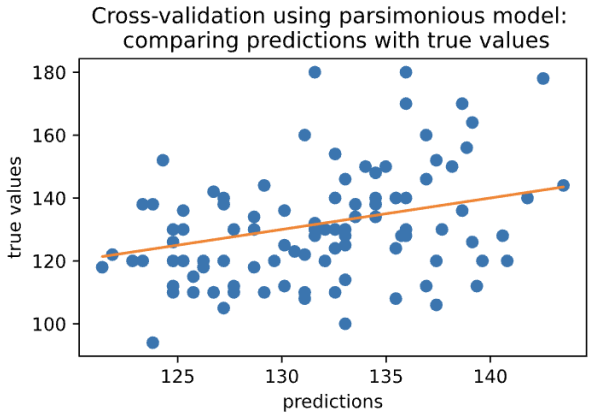
Here is the third plot of posterior correlations for the non-parsimonious model:



Here is the fourth screenshot of the leave-one-out cross-validation result:

	rank	loo	p_loo	d_loo	weight	se	dse	warning	loo_scale
non-parsimonious	0	-1288.835294	8.755829	0.000000	0.5089	14.391575	0.000000	False	log
parsimonious	1	-1289.110603	4.562225	0.275309	0.4911	14.904932	3.447643	False	log

Here is the fifth plot that shows how well our parsimonious model does on unseen data:



Here, the sixth plot shows us our posterior correlations for the parsimonious model: