Data Analysis Project: Nurse's Blood Pressure Study

Emilie Campos

2019-03-07

Introduction

Motivation and Goals

Increased blood pressure can lead to a multitude of health risks so understanding the factors that contribute to higher blood pressure can help health professionals intervene. Here, we will attempt to understand how family history of hypertension, menstrual cycle phase, and whether a person is working affects their blood pressure. Additionally, we will want to investigate whether there is an interaction between family history and working.

Data Description

The responses to this study are repeated measures on 203 female nurses of their systolic blood pressure (SBP) (in mmHG). Each nurse is also asked whether neither, one, or both of their parents have a history of high blood pressure. At each measurement, they are also asked what phase of their menstrual cycle they are in and whether they are working or not. The number of readings varies from person to person. The descriptions of each variable are given in Table 1 and each covariate is summarized in Table 2.

Methods

The model that will be used is the random effects model which takes the form

$$y_{ij} = \alpha_1 + \alpha_2 \times \text{One Parent}_{ij} + \alpha_3 \times \text{Two Parents}_{ij} + \alpha_4 \times \text{Working}_{ij} + \alpha_5 \times \text{Luteal Phase}_{ij} + \beta_i + \epsilon_{ij}$$

where $i = 1, ..., 203$ and $j = 1, ..., J_i$, J_i are the number of measurements on nurse i. The β_i's are modeled as $N(0, D)$ and $\epsilon_{ij} \overset{\text{iid}}{\sim} \text{N}(0, \sigma^2)$. Since we are taking repeated measurements on the same individual, the β_i's introduce correlation between the observations within each participant. We will need to develop a prior distribution for each of the regression parameters, namely α_{1-5} and β_i.

All of the prior information was gathered from literature. Regarding the average blood pressure during the follicular phase of the menstrual cycle, a study published in Applied Psychophysiology and Biofeedback on the effects of daily activities on Ambulatory Blood Pressure During Menstral Cycle in Normotensive Women [3] found that the average SBP was 116.82, standard deviation 2.05. In that same study, it was determined that the average blood pressure in the luteal phase was 119.26 so the regression coefficient for luteal phase will have a prior mean of 2.44, the difference between the luteal and follicular means. In a study published in the journal Hypertension on the effects of parental hypertension on longitudinal trends in blood pressure [1] found that parental hypertension is associated with an increase of 6.46 mmHG on average in women, standard error 0.83. We will assume this is true for one or both parents and incorporate this into the prior as the prior for both predictors. A study published in the Scandinavian Journal of Work, Environment and Health [2] found that the SBP was 4 mmHG higher on average when working versus leisure time for hospital workers in Stockholm county. Since this study was performed in a different country, we will use a small precision to reflect our uncertainty that this is true in the US.

In addition to the model described above, we are also interested in the interaction of family history and work so a second model will be used

$$y_{ij} = \alpha_1 + \alpha_2 \times \text{One Parent}_{ij} + \alpha_3 \times \text{Two Parents}_{ij} + \alpha_4 \times \text{Working}_{ij} + \alpha_5 \times \text{Luteal Phase}_{ij}$$
$$+ \alpha_6 \times \text{Working}_{ij} \times \text{One Parent}_{ij} + \alpha_7 \times \text{Working}_{ij} \times \text{Two Parents}_{ij} + \beta_i + \epsilon_{ij}.$$

Here, we will want to see if the coefficients α_6 and α_7 are significant which would indicate a differential effect. The normal model will be

Intercept	$\alpha_1 \sim \text{No}(116.82, 1/2.05)$
One Parent	$\alpha_2 \sim \text{No}(6.46, 1/0.83)$
Two Parents	$\alpha_3 \sim \text{No}(6.46, 1/0.83)$
Working	$\alpha_4 \sim \text{No}(4, 0.1)$
Luteal Phase	$\alpha_5 \sim \text{No}(2.44, 0.01)$
One Parent X Working	$\alpha_6 \sim \text{No}(0, 0.01)$
Two Parents X Working	$\alpha_7 \sim \text{No}(0, 0.01)$
Individual	$\beta_i \sim \text{No}(0, \tau_b)$
Tau_b	$\tau_a \sim \text{Gamma}(1, 1)$
ϵ_i	$\epsilon_{ij} \sim \text{No}(0, \tau_e)$
Tau_e	$\tau_e \sim \text{Gamma}(1, 1)$

Based on the histogram of the systolic blood pressures, see Figure 1, the distribution is approximately normal but there are a few large values in the tails. This indicates that both a normal model and a t model should be explored. The t model is

Intercept	$\alpha_1 \sim \text{No}(116.82, 1/2.05, df_a)$
One Parent	$\alpha_2 \sim \text{No}(6.46, 1/0.83, df_a)$
Two Parents	$\alpha_3 \sim \text{No}(6.46, 1/0.83, df_a)$
Working	$\alpha_4 \sim \text{No}(4, 0.1, df_a)$
Luteal Phase	$\alpha_5 \sim \text{No}(2.44, 0.5, df_a)$
Individual	$\beta_i \sim \text{No}(0, \tau_b, df_b)$
Tau_b	$\tau_a \sim \text{Gamma}(1, 1)$
ϵ_i	$\epsilon_{ij} \sim \text{No}(0, \tau_e)$
Tau_e	$\tau_e \sim \text{Gamma}(1, 1)$
Df_a	$1/df_b \sim \text{Unif}(0, 0.5)$
Df_y	$1/df_y \sim \text{Unif}(0, 0.5)$

All of the models were run with 10000 iterations on 5 chains after 1000 burn-in. To reduce the amount of autocorrelation in the t model, the thinning value was set to 10.

Results

To check the convergence of each model, the autocorrelation and time series plots were inspected. Figure 2 shows that the convergence in the normal model is quick, with the autocorrelation going to zero immediately at lag 1. The t model did not approach zero until at least lag 30 even with thinning. Additionally, looking at Figure 3, the time series plot for the normal model looks like grass reflected in water whereas the time series for the t model was not at all random. Therefore we conclude that the normal model has much better convergence and we will continue with just this model. The posterior predictions based on our model are plotted in Figure 4 and highlight that normal model is doing well at predicting data with the same shape as the observed systolic blood pressures. Additionally, the distribution of the mean of each individual's observed systolic blood pressure measurements is approximately the same distribution as the expected mean, which is calculated as the sum of the intercept and the individual's beta coefficient.

The results of the normal model are given in Table 4. We inspected the interaction terms by running another model with the interactions. Each of the interaction terms were found to be insiginificant so the model without the interaction was used. The average Systolic Blood Pressure (SBP) over all individuals while resting and in the follicular phase of their menstrual cycle is 114.8 with a standard deviation of 0.86. There is sufficient evidence to conclude the effect of family history of hyptertension is not zero since the 95% confidence intervals for each of the family history indicator variables do not cross zero. Additionally, since the confidence intervals are entirely above zero, we conclude that having a parent with hypertension raises the individuals blood pressure. Having one parent with hypertension is associated with an average 3.84 mmHG increase in resting SBP, with a standard deviation of 0.74. Whereas having two parents with hypertension is associated with an average 6.86 mmHG increase in resting SBP, with a standard deviation of 0.85. Working is associated with higher blood pressure; on average SBP is 2.76 mmHG higher on working days with a standard deviation of 1.06. The only variable not associated with a change in blood pressure is the phase of the menstrual cycle, when controlling for the other covariates. The 95% confidence interval for phase crosses the value zero.

We will want to compare the Bayesian estimates to the classic OLS model results. The results of the classic OLS model are given in Table 3. The plots of each of the distributions (prior, likelihood, and posterior) are given in Figure 6. We see that for some of the parameters, each of the distributions are very similar, i.e. in the work and luteal phase parameters. However, the One Parent and Two Parent regression coefficients may be driven by the priors more strongly than by the data. This is also clear from the estimates found in the OLS model. It may be pertinent to find another expert opinion for these parameters.

References

[1] Kaneto Mitsumata, Shigeyuki Saitoh, Hirofumi Ohnishi, Hiroshi Akasaka, and Tetsuji Mirua. Effects of Parental Hypertension on Longitudinal Trends in Blood Pressure and Plasma Metabolic Profile. *Hypertension*, 60:1124–1130, 2012.

[2] Töres Theorell, Gunnel Ahlberg-hulten, Margareta Jodko, Filis Sigala, Bartolomé De Torre, Source Scandinavian, No October, Tores Theorell, Gunnel Ahlberg-hulten, and Margareta Jodko. Norwegian National Institute of Occupational Health Danish National Research Centre for the Working Environment Finnish Institute of Occupational Health Influence of job strain and emotion on blood pressure in female hospital personnel during workhours In. *Scandinavian Journal of Work, Environment & Health*, 19(5):313–318.

[3] Pei-shan Tsai, Carolyn B Yucha, David Sheffield, and Mark Yang. Effects of Daily Activities on Ambulatory Blood Pressure During Menstrual Cycle in Normotensive Women. *Applied Psychophysiology and Biofeedback*, 28(1):25–36, 2003.

{}

Appendix

Tables and Figures

Table 1: Variable Coding

Variable	Definition
idnum	Subject ID number
sys	Systolic blood pressure (mmHG)
one parent	Using the variable for Family history of hypertension: 1 if only one parent with a history of hyptertension, 0 otherwise
two parents	Using the variable for Family history of hypertension: 1 if both parents with a history of hyptertension, 0 otherwise
work	0 = not working, 1 = working
phase	Menstrual cycle phase: 0 = Follicular, 1 = Luteal

Table 2: Summary of Data

Category	Count	Percent
Family History		
Neither Parent	5298	0.55
One Parent	3633	0.38
Two Parents	642	0.07
Working		
Resting	4116	0.43
Working	5457	0.57
Menstrual Cycle Phase		
Follicular	4737	0.49
Luteal	4836	0.51

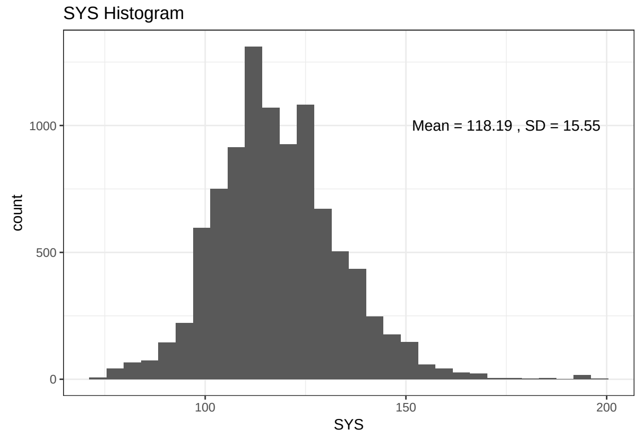

Figure 1: The distribution is approximately normal. However, there are some large values out in the tails that are not unreasonable outliers which suggests a t distribution may be suitable.

Table 3: Classic OLS Results

Label	Parameter	Mean	SD	Statistic
Intercept	alpha[1]	114.99	1.18	97.75
One Parent	alpha[2]	-0.18	1.23	-0.14
Two Parents	alpha[3]	8.50	2.37	3.59
Work	alpha[4]	3.73	1.17	3.18
Phase	alpha[5]	1.11	1.17	0.94

Table 4: Normal Model Results

Label	Parameter	Mean	SD	2.5%	97.5%
Intercept	alpha[1]	114.80	0.86	113.12	116.47
One Parent	alpha[2]	3.84	0.74	2.40	5.31
Two Parents	alpha[3]	6.86	0.85	5.19	8.53
Working	alpha[4]	2.76	1.06	0.69	4.84
Luteal Phase	alpha[5]	0.47	1.12	-1.75	2.66
sigma	sigma	13.09	0.10	12.90	13.28
sqrtD	sqrtD	8.28	0.45	7.45	9.21

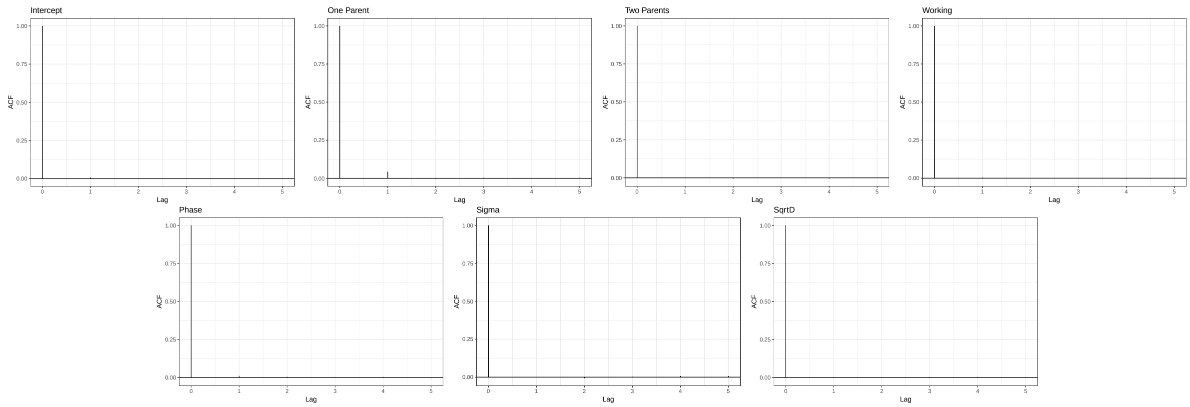

Figure 2: The autcorrelation plots for this normal model show that the convergence with the normal model is quick.

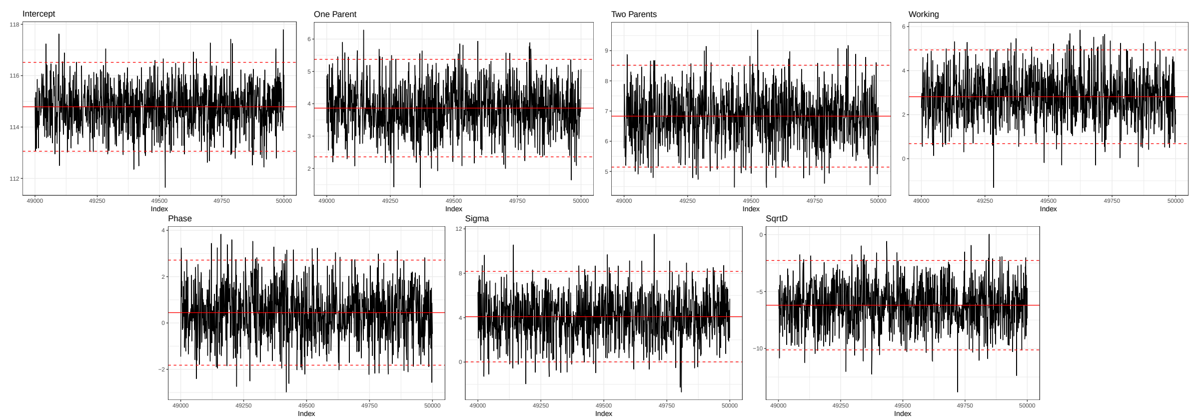

Figure 3: The plots of the time series also highlight that the convergence is quick.

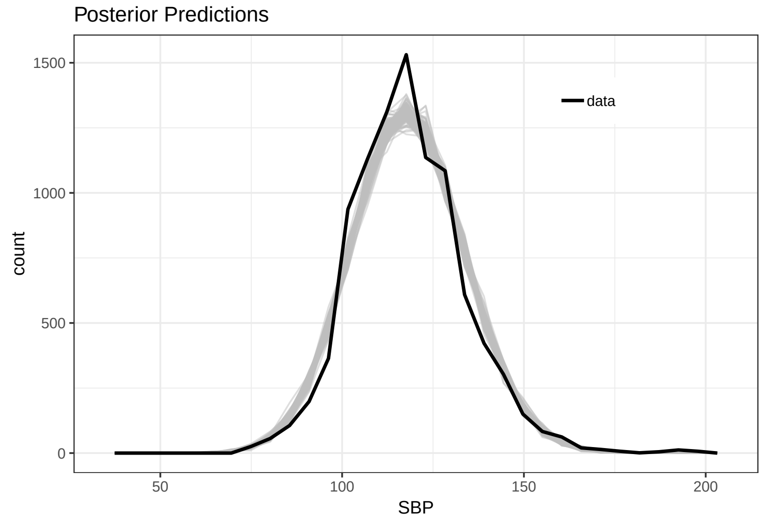

Figure 4: The posterior predicitive distributions are plotted for each individual in grey and the distribution of the systolic blood pressures from the observed data is plotted in black. The posterior predicitive distributions follow the observed data closely, implying our model is doing well at describing the data.

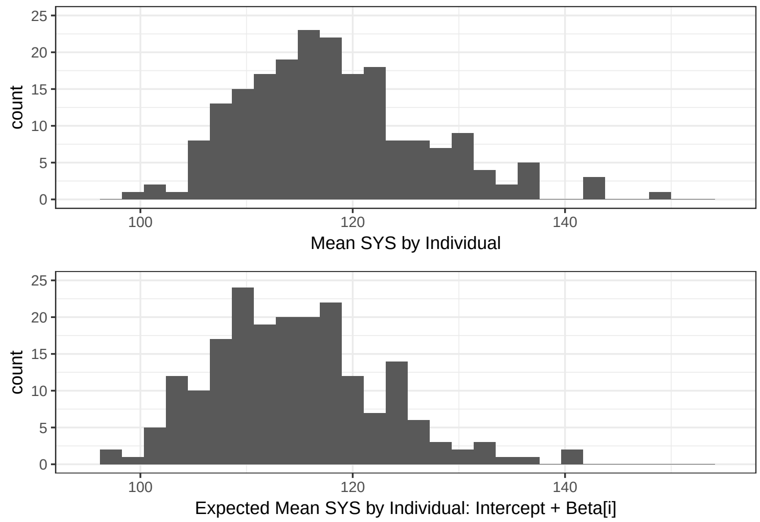

Figure 5: The mean of each individual's observed systolic blood pressure values compared to the expected mean systolic blood pressure values according to the normal model have very similar distributions.

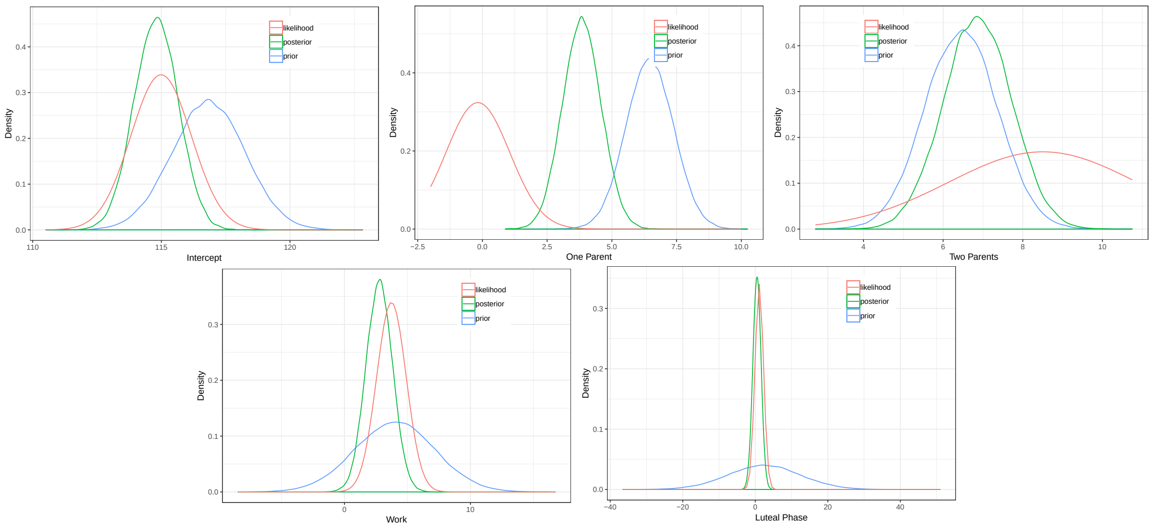

Figure 6: Each of the distributions for the parameters of interest using only the normal model

Code

Normal Model:

```
cat("
  model
    {
      # the likelihood contribution
      for(i in 1:N) {
        y[i] ~ dnorm(mu[i], tau.e)
        mu[i] <- inprod(x[i,], alpha[]) + beta[idnum[i]]
      }

      # the priors for the fixed effects
      for(j in 1:J) {
        alpha[j] ~ dnorm(m[j], prec[j])
      }

      # the priors for the random effects
      for(k in 1:K) {
        beta[k] ~ dnorm(0, tau.b)
      }

      tau.e ~ dgamma(ea, eb)
      tau.b ~ dgamma(ba, bb)

      sigma <- 1 / sqrt(tau.e)
      sqrtD <- 1 / sqrt(tau.b)
```

```
    rho <- sqrtD*sqrtD / (sigma*sigma + sqrtD*sqrtD)
  }",
  fill = TRUE,
  file = here("Data Analysis Project/dap_normalmodel.txt"))
```