Bayesain Project

Ammar Hasan

18 February 2019

Contents

Introduction

This report summaries the Bayesian analysis of the Reisby data-set. The Reisby data-set is based on a 5 week (first placebo) psychiatric study which investigates response of depressed patients to IMI. The Bayesian analysis will investigate how the drug affects depression.

Data Setup and Cleaning

Data

The data is loaded from a .Rdata file containing the Riesby data-set introduced in the introduction.

Data Exploration

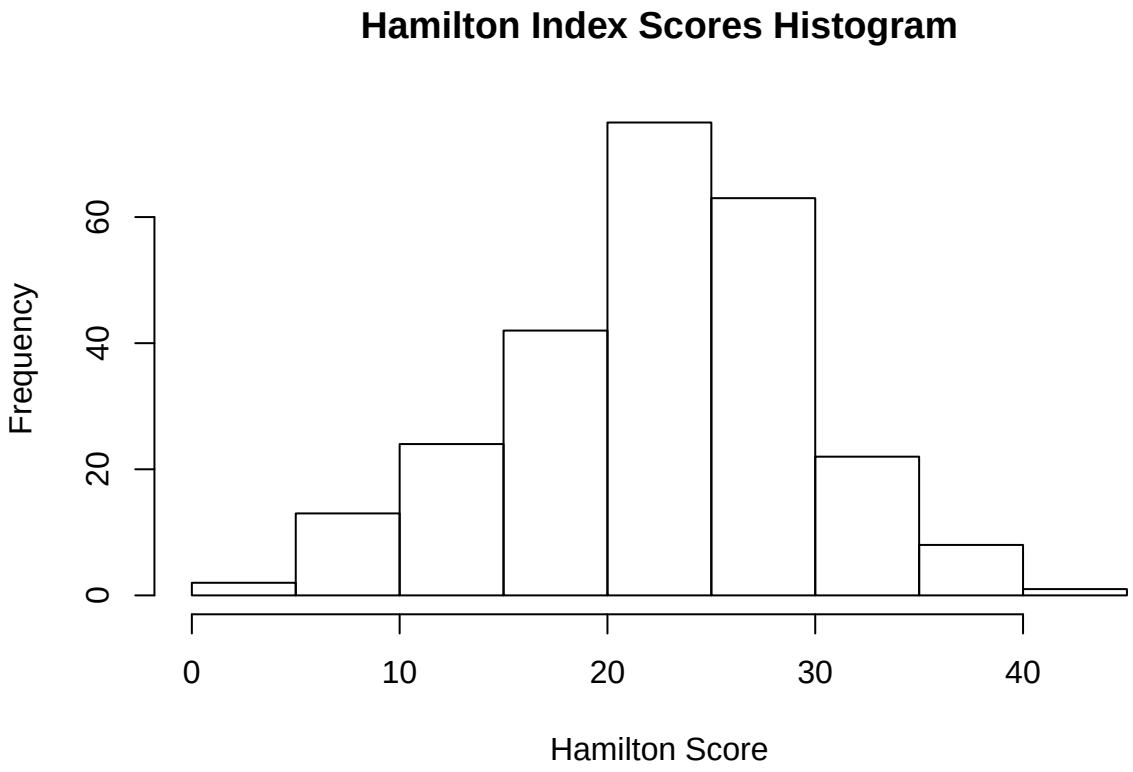
Correlation (Pairs Plot)

	hd	week	lnimi	lndmi	female	reactive_depression
hd	1.0000000	-0.3322674	-0.1165766	-0.3521130	-0.0624613	-0.0561811
week	-0.3322674	1.0000000	0.0357592	0.1239101	-0.0118326	0.0325663
lnimi	-0.1165766	0.0357592	1.0000000	0.2102979	0.0859277	-0.0365988
lndmi	-0.3521130	0.1239101	0.2102979	1.0000000	0.0945810	-0.1001591
female	-0.0624613	-0.0118326	0.0859277	0.0945810	1.0000000	0.1158473
reactive_depression	-0.0561811	0.0325663	-0.0365988	-0.1001591	0.1158473	1.0000000

The strongest correlations are negative and weak-moderate, and occur between Hamilton index with week and DMI, and rather obvious as the increase in blood concentration of the antidepressant would alleviate depression over weeks of treatment. Most other correlations are weak to very weak.

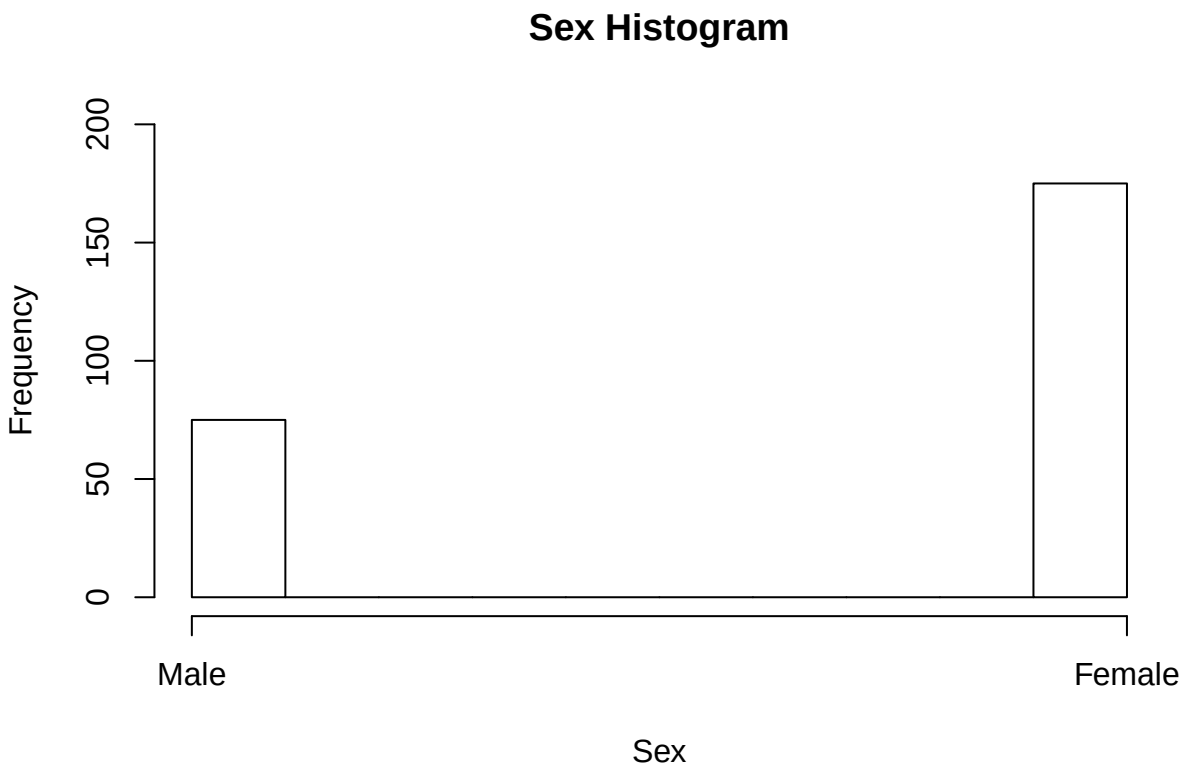
Graphical Summaries

Hamilton Scores



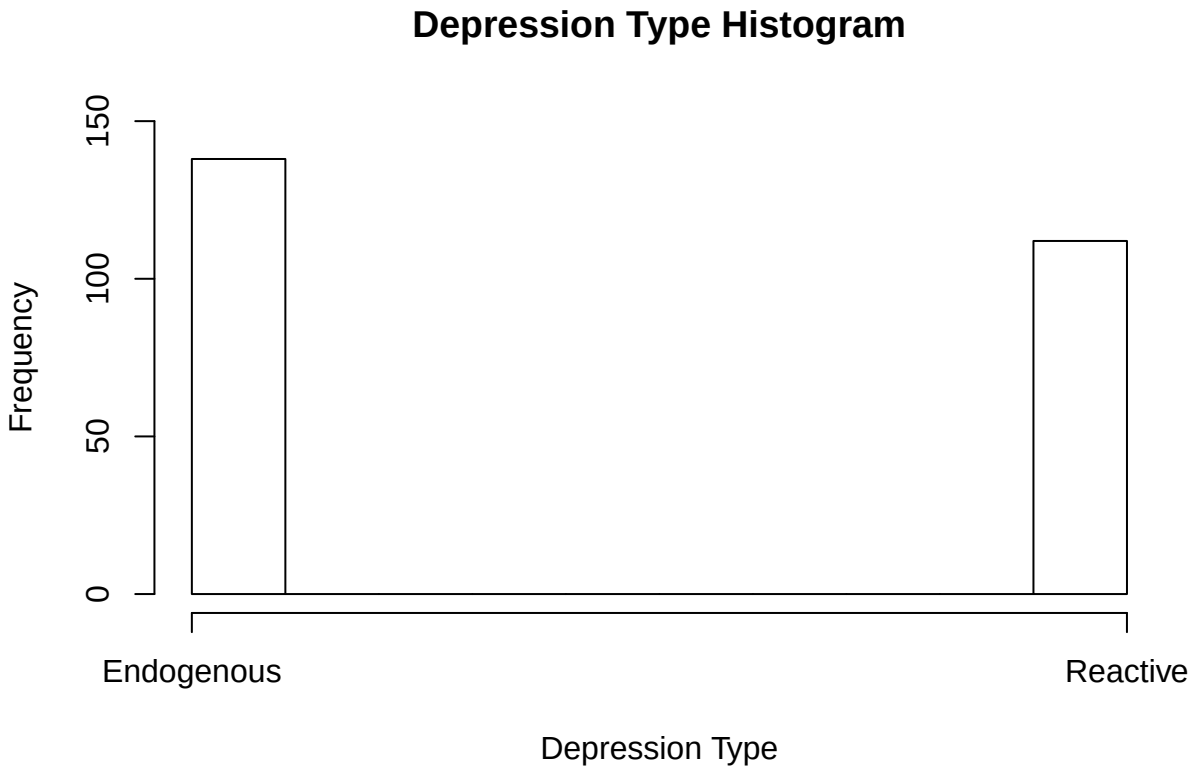
Most Hamilton scores are above 20 (i.e. moderate and severe depression dominates against mild and normal)

Sex



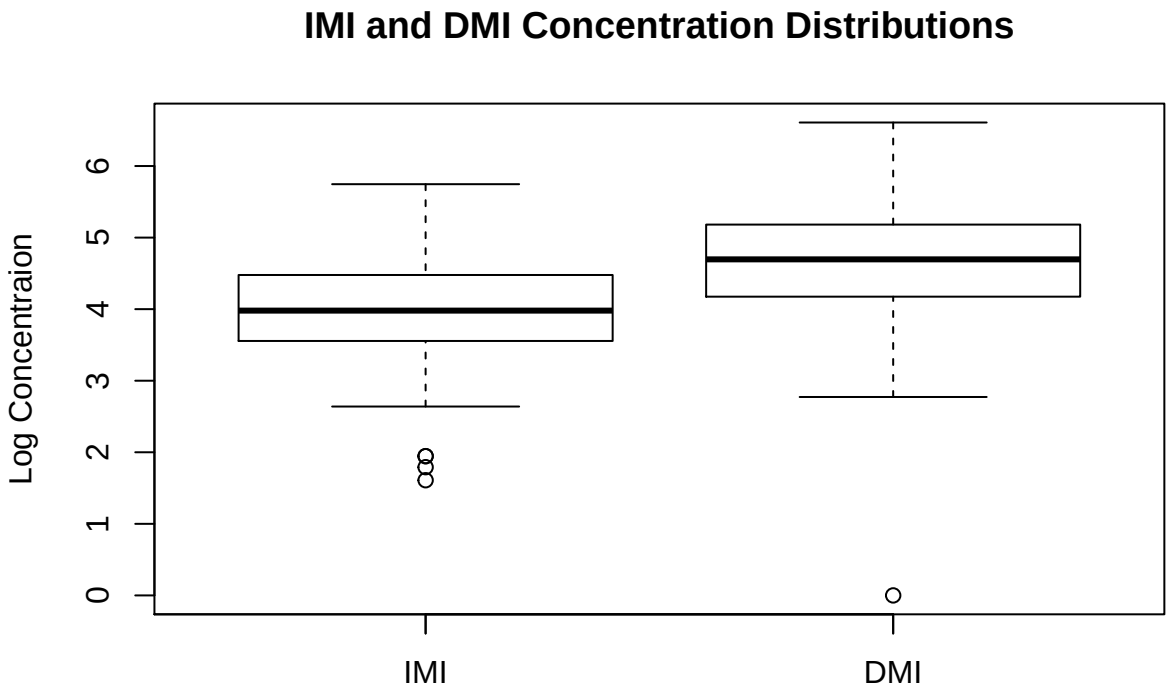
The female test subjects are overwhelmingly higher than the males (nearly double!).

Endogenous vs Reactive (Depression Type)



Most Depression cases in the test population are Endogenous (i.e. not a reaction to an environmental event).

DMI and IMI Concentrations



Generally tight distribution *especially around 25% to 75%) for both with little outlines with both distribution looking nearly identical sans the shift up with DMI. It seems that after being processed as DMI, the concentration of the antidepressant in the blood increases.

Data Cleaning and Preparation

Weeks

There are going to be two weeks variables, the initial week number for Autoregressiona and Gaussian Proccess, and another or on whether it is a placebo week or not for linear regression.

Standardisation and Predictor Separation

The response (hamilton index) and predictors (everything else) are separated for the modelling stage. Two sets of predictors (x) are used, one for linear regression (uses placebo indicator for weeks) and one for other time series based methods (uses normal weeks). All non indicator variables are standardised using a to ensure a fair impacts between variables and easier uninformative prior selection.

Models

Linear Regression

In this method a relationship between a predictor x and a response y is establised through a Linear Function with a coefficient the predictors and the intercept.

Basic Multiple

Multiple Linear Regression is a simple extension of Linear regression where multiple predictors each with their own coefficients are introduced.

Modelling and Diagnostics

‘‘‘

Summaries

Term Interactions

Modelling and Diagnostics

Summaries

Not everyone was measured every week.

Hierarchical

Modelling and Diagnostics

Summaries

Auto Regression

AR(1)

Modelling and Diagnostics

Summaries

AR(2)

Modelling and Diagnostics

Summaries

Gaussain Process

Modelling and Diagnostics

Summaries

Conclusion

Appendix

Abbrevations

IMI - antidepressant drug imipramine

DMI - desmethylimipramlne (Processed IMI)

AR - Auto Regressive Models

AR(1) - Auto Regressive Models (1st Degree)

AR(2) - Auto Regressive Models (2nd Degree)

Code

Data

Data Loading

```
load("Reisby.RData")
Reisby = as.data.frame(Reisby)
```

Correlation (Pairs Plot)

```
kable(cor(Reisby[,-1])) # no id
```

Hamilton Scores Histogram

```
hist(Reisby$hd, main = "Hamilton Index Scores Histogram",
     xlab = "Hamilton Score")
```

Sex Histogram

```
# Replace x axis here with male/female axis

hist(Reisby$female, main = "Sex Histogram",
     xlab = "Sex", xaxt = "n", ylim = c(0, 200))
axis(1, at=0:1, labels=c("Male","Female"))
```

Depression Type Histogram

```
hist(Reisby$reactive_depression, main = "Depression Type Histogram",
     xlab = "Depression Type", xaxt = "n", ylim = c(0, 150))
axis(1, at=0:1, labels=c("Endogenous","Reactive"))
```

DMI and IMI Boxplot

```r
boxplot(Reisby$lnimi, Reisby$lndmi, main = "IMI and DMI Concentration Distributions",
        ylab = "Log Concentraion")
axis(1, at=0:2, labels=c("IMI","IMI", "DMI")) # uses 3 labels due to strange bug
```

Weeks Cleaining

Models

Basic Multiple Linear Regression

```r
data_list = list(x = lin_x_scaled,
                 y = y_scaled,
                 n = nrow(lin_x_scaled),
                 p = ncol(lin_x_scaled))

# JAG model (as a string)

model_string = "
model {
  beta0 ~ dnorm(-0.8, 1/50)

  for (i in 1:p) {

  }

  beta1 ~ dnorm(8, 1/10)
  tau ~ dexp(10)
  for (i in 1:N) {
    y[i] ~ dnorm(beta0 + beta1*x[i], tau)
  }
}
"
# normal for b0 and b1 and exponential for the precision

# model construction and sampling

model_string_conn = textConnection(model_string)

#model = jags.model(model_string_conn, data = data_list, n.chains = 4)

#samples = coda.samples(model,
 #                      variable.names = c("beta0","beta1","tau"),
  #                     n.iter = 1000)
```