CS-E5710 Bayesian Data Analysis - Project work

Khoa Lai & Vinh Nguyen

November 2020

Contents

1. Introduction

Motivation: In the medical field, there are a lot of images that are taken on a daily basis from the patient around the world. Detecting abnormal signs from the image is expensive in terms of the amount of time spent and effort from the doctor's perspective. What has been done traditionally is to have a doctor looking at each of the image and appling available knowledge to evaluate whether a patient has developed any symptoms of a disease. Thanks to the advance in the Machine Learning field, nowadays computers could help in predicting whether a patient having a disease based in the assumption that the measurements are different between a patient and a normal person. In this project, we would apply what we have learned during the course in terms of Bayesian inference and its methods to classify whether a patient having breast cancer given the available measurements that have been extracted from the patient's digitized image of a fine needle aspirate (FNA) of a breast mass.

The problem:

Main modeling idea:

Exploratory Data Analysis: Illustrative Figure:

2. Data Description and Problem Analysis

The dataset, namely $BreastCancerWisconsin(Diagnostic)$, is obtained through the UCI Machine Learning Repository. According to the source, features are computed from a digitized image of a find needle aspirate (FNA) of a breast mass. They describe characteristics of the cell nuclei present in the image.

There are 10 real-valued features that are computed for each cell nucleus:
a) radius (mean of distances from center to points on the perimeter)
b) texture (standard deviation of gray-scale values)
c) perimeter
d) area
e) smoothness (local variation in radius lengths)
f) compactness (perimeter^2 / area - 1.0)
g) concavity (severity of concave portions of the contour)
h) concave points (number of concave portions of the contour)
i) symmetry
j) fractal dimension ("coastline approximation" - 1)

The target variable is to predict whether a person having breast cancer given the attributes. (0: not cancer, 1: cancer).

The data consists of binary observations $y_n \in$ {0,1} paired with 30-dimensional vectors of predictors x_n for n $\in$ 1:N.

Obtaining the dataset in suitable form requires some pre-processing steps, and here we drop the ID column from the original dataset as it does not attribute to the final result.

Since the features in the data originally are not in the same scale, we would standardize the data so that for every column k corresponding to a predictor, we have:

$$\text{mean}(x_{1:N,k} = 0)$$

and

$$\text{sd}(x_{1:N,k} = 1)$$

For the purposes of simulation, we would assume the data have a multivariate normal distribution with a positive-definite covariance matrix $\sum$,

$$x_n \sim \text{multinormal}(0, \sum)$$

Since the dataset is public, there has been several case studies and use of the dataset since it was published on 1995. The authors did not try to find any specific case studies / project that involves the same dataset to compare against, and we would conduct the experiment independently for that reason.

3. Model Description

In this project, we would use Logistic Regression to classify / predict whether a person having a breast cancer given the measurements from the image.

The model is parameterized with an intercept $\alpha \in \mathbb{R}$ and coefficient vector $\beta \in \mathbb{R}^K$. Logistic regression is a generalized linear model where the linear predictor is defined as:

$$\alpha + x_n\beta = \alpha + \textstyle\sum_{k=1}^{K} x_{n,k} * \beta_k$$

represents the log odds of y_n being equal to one.

Given the log odds, the probability that y_n is one is given by inverting the log odds function:

$$\Pr[y_n = 1] = logit^{-1}(\alpha + x_n\beta)$$

The sampling distribution is then defined to follow the log odds:

$$y_n \sim bernoulli(logit^{-1}(\alpha + x_n\beta))$$

for observations indexed by n ∈ 1:N. It is important to note that the y_n are defined by sampling according to the log odds.

Loss Functions for Evaluating predictions
Our goal is to provide predictive estimates of the probability that an unobserved outcome $\tilde{y}_n$ takes valuee 1, given predictors $\tilde{x}_n$ and training data (x,y). In mathematical notation, we want to estimate

$$\Pr[\tilde{y}_n = 1|\tilde{x}_n, x, y]$$

We will consider 2 scoring functions (log loss and square loss) for evaluating the accuracy of the model.

Log loss:
It is the log probability(density or mass) of the true result under the model. For a given target y_n and probabilistic prediction $\hat{y}_n$, the log loss is defined by

$$\text{Log Loss} = \sum_{(x,y)\in D} -ylog(y') - (1-y)log(1-y')$$

where:
- (x,y) ∈ D is the dataset containing many labeled examples, which are (x,y) pairs.
- y is the label in a labeled example. Since this is logistic regression, every value of y must either be 0 or 1.
- y' is the predicted value (somewhere between 0 and 1), given the set of features in x.

Square error:
It is a loss function that can be used in the learning setting in which we are predicting a real-valued variable y given an input variable x.

$$\text{Square Loss } (y_n, \hat{y}_n) = (y_n - \hat{y}_n)^2.$$

Given the entire test set, the loss can be expressed as:

$$\text{Square Loss } (y, \hat{y}) = \sum_{n=1}^{\hat{N}} (y_n - \hat{y}_n)^2$$

4. Prior justification

Given the logistic scale and standardized predictors, we assume weakly informative priors,

$$\alpha, \beta_k \sim normal(0, 2),$$

for k ∈ 1:K.

Model implementation

In this section, we are going to simulate some correlated predictors, and outcomes and then fit our logistic regression model with full Bayes, posterior modes, and posterior means for the purposes of evaluating predictive accuracy.

Load the packages

```
set.seed(42)
library(rstan)
```

```
## Loading required package: StanHeaders
```

```
## Loading required package: ggplot2
```

```
## rstan (Version 2.21.2, GitRev: 2e1f913d3ca3)

## For execution on a local, multicore CPU with excess RAM we recommend calling
## options(mc.cores = parallel::detectCores()).
## To avoid recompilation of unchanged Stan programs, we recommend calling
## rstan_options(auto_write = TRUE)
```

```
rstan_options(auto_write = TRUE)
options(mc.cores = parallel::detectCores())
rstan_options(auto_write = TRUE)
```

Load and pre-process the data

```
breast_data <- read.csv("wdbc.data")

# Split the data into training and testing set. Use 50% for training and 50% for testing
training_size = round(0.5 * nrow(breast_data),0)
tr_idx = sample(nrow(breast_data), training_size)
training_data = breast_data[tr_idx,]
testing_data = breast_data[-tr_idx,]

# Convert the target (originally categorical) into numerical variable
y_train = as.numeric(factor(training_data$M))
y_test = as.numeric(factor(testing_data$M))

# Subtract both the target label by 1.
# Since in R, index starts from 1 and for the logistic regression, the label needs
# to be within the interval [0,1]
y_train = y_train - 1
y_test = y_test - 1

# Remove the target and ID column from the training and testing set
x_train = subset(training_data, select=-c(X842302, M))
x_test = subset(testing_data, select=-c(X842302, M))

# Standardize the data
scaled_training = scale(x_train)
df_train = as.data.frame(scaled_training)
scaled_testing = scale(x_test)
df_test = as.data.frame(scaled_testing)

# Prepare the data
data <- list(N = nrow(df_train), J = ncol(df_train), x = df_train, y = y_train,
             N_test = nrow(df_test), K = ncol(df_test), x_test = df_test,
             y_test = y_test)
```

Run the model

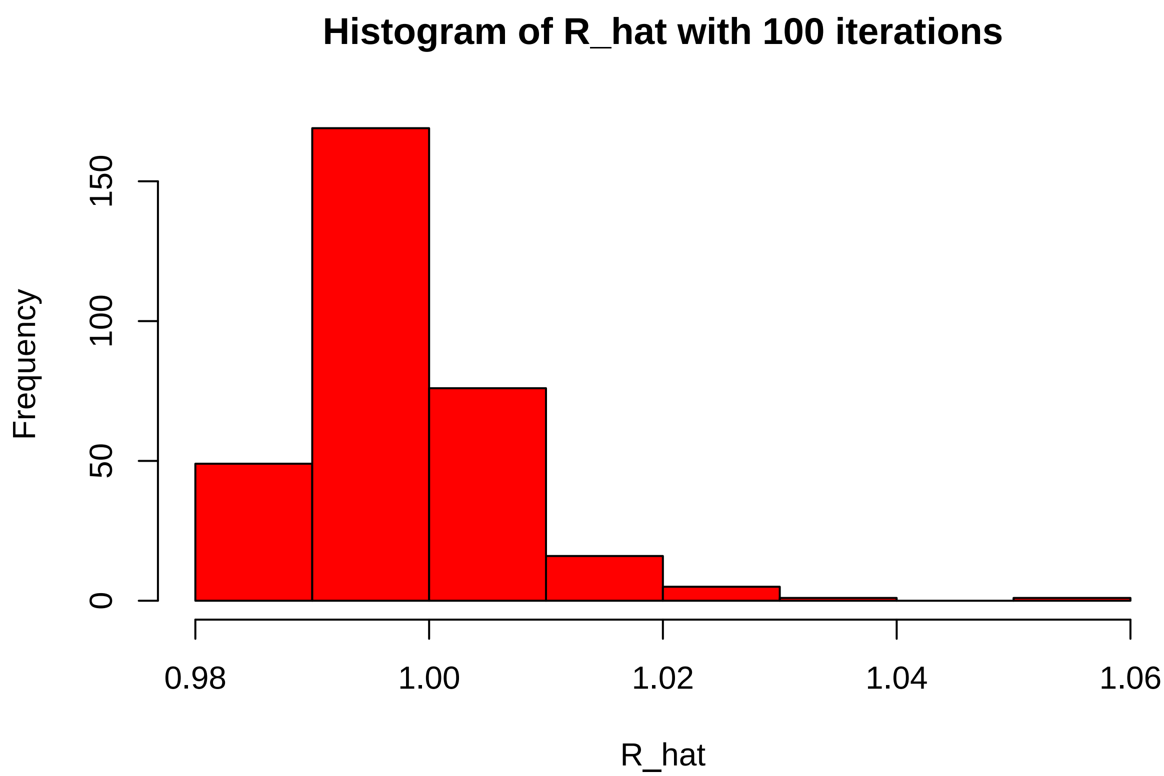
```
# Fit the model with 100 iterations
fit_logistic_regression = stan(file = "logistic_regression.stan", data = data, iter = 100)
```

```
## Warning: The largest R-hat is NA, indicating chains have not mixed.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#r-hat
```

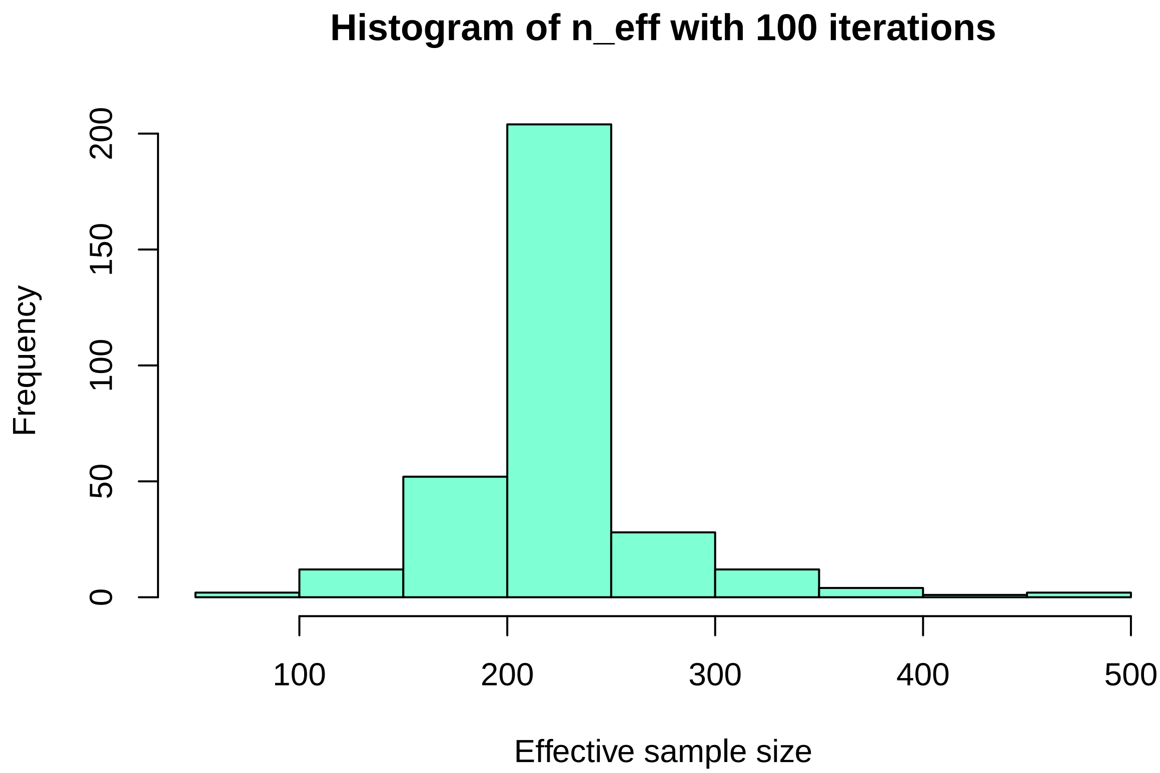
```
## Warning: Bulk Effective Samples Size (ESS) is too low, indicating posterior means and medians may be
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#bulk-ess

## Warning: Tail Effective Samples Size (ESS) is too low, indicating posterior variances and tail quant
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#tail-ess
```

```r
# Extract the R_hat column
r_hat_100 = summary(fit_logistic_regression)$summary[,10]
hist(r_hat_100, main = "Histogram of R_hat with 100 iterations",
     xlab = "R_hat", col = "red")
```

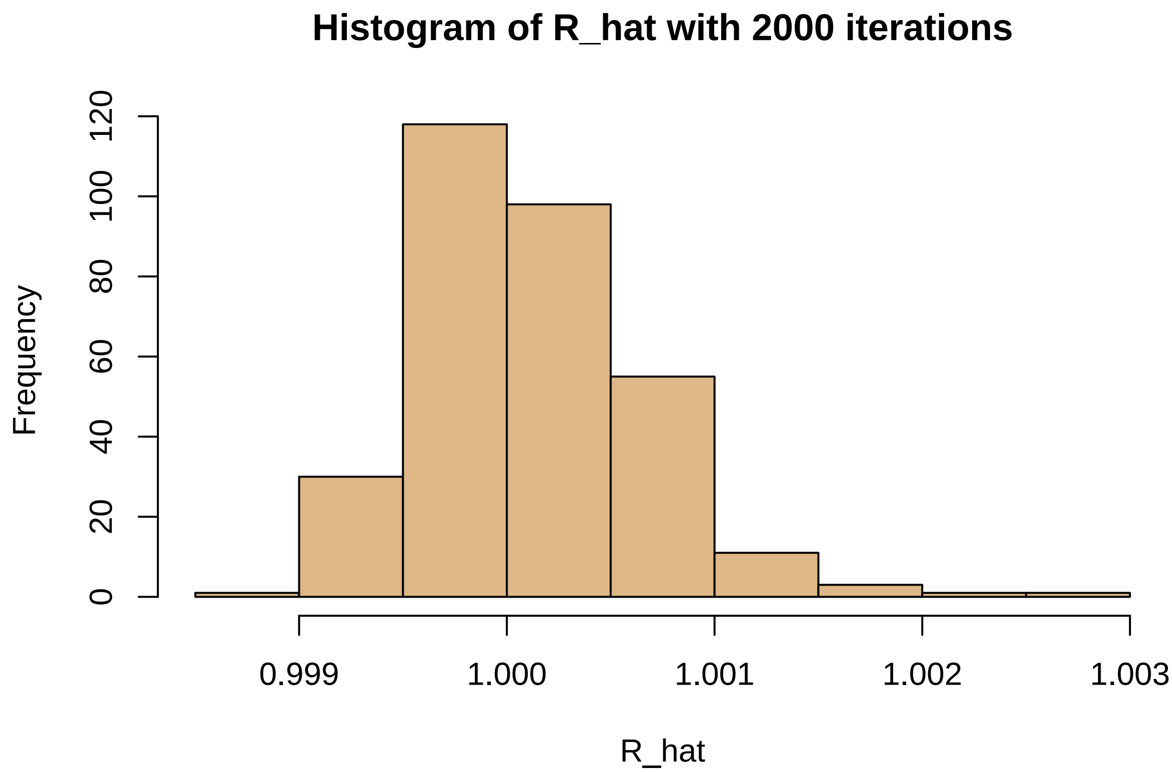



```r
# Extract the n_eff column
n_eff_100 = summary(fit_logistic_regression)$summary[,9]
# Plot the histogram of n_eff
hist(n_eff_100, main = "Histogram of n_eff with 100 iterations",
     xlab = "Effective sample size", col = "aquamarine")
```



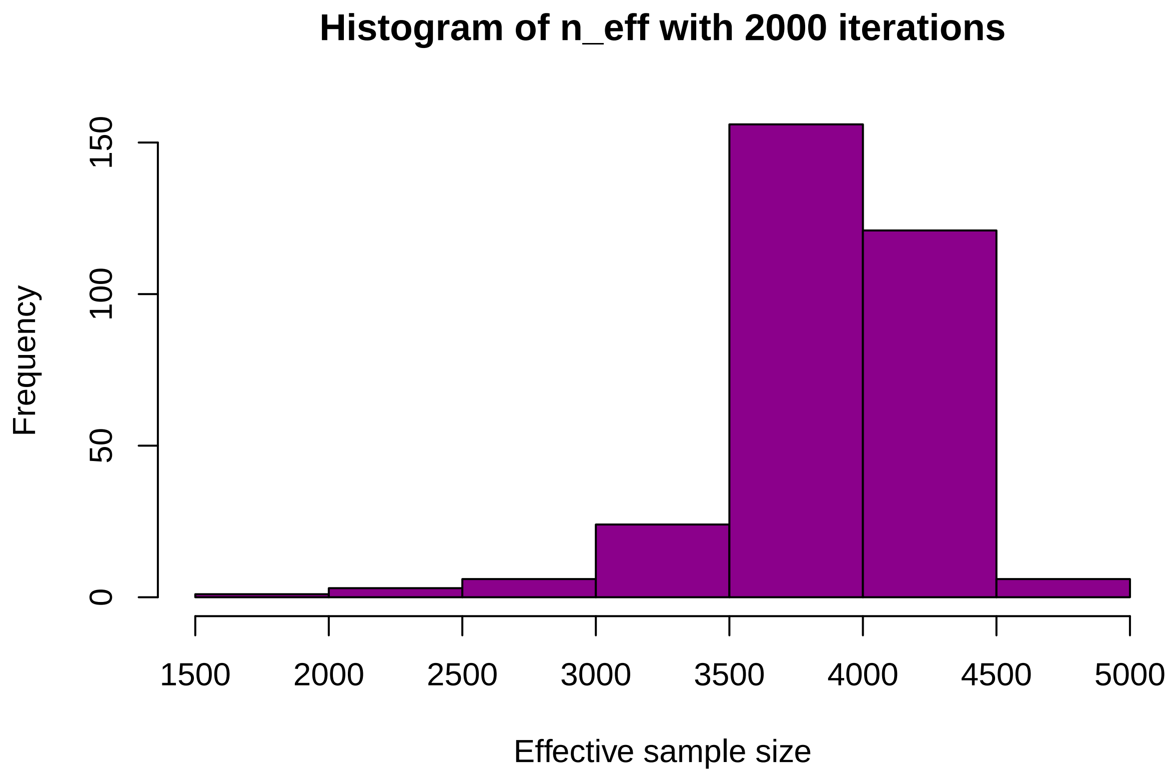

```r
# Fit the model with 2000 iterations
fit_logistic_regression_2000 = stan(file = "logistic_regression.stan", data = data, iter = 2000)
```

```r
# Extract the R_hat column
r_hat_2000 = summary(fit_logistic_regression_2000)$summary[,10]
hist(r_hat_2000, main = "Histogram of R_hat with 2000 iterations",
     xlab = "R_hat", col = "burlywood")
```




```r
# Extract the n_eff column
n_eff_2000 = summary(fit_logistic_regression_2000)$summary[,9]

# Plot the histogram of n_eff
hist(n_eff_2000, main = "Histogram of n_eff with 2000 iterations",
     xlab = "Effective sample size", col = "darkmagenta")
```



5. Stan code

Here's a Stan program implementing the logistic regression model:

```
writeLines(readLines("logistic_regression.stan"))
```

```
## data {
##   int < lower =0 > N; # Number of observations in training set
##   int < lower =0 > J; # Number of columns in training set
##   matrix [N,J] x; # Training dataset
##   int<lower=0, upper=1> y[N]; # Target
##
##   int N_test; # Number of observations in testing set
##   int <lower=0> K; # Number of columns in testing set
##   matrix[N_test, K] x_test; # Testing dataset
##   int y_test[N_test];
## }
##
## parameters {
##   real alpha;
##   vector [J ] beta ;
## }
##
## model {
##   alpha ~ normal(0,2);
##   beta ~ normal(0,2);
```

```
##   y ~ bernoulli_logit(alpha + x * beta);
## }
##
## generated quantities {
##   vector[N_test] E_y_test = inv_logit(alpha + x_test * beta);
##   real log_loss = -bernoulli_logit_lpmf(y_test | alpha + x_test * beta);
##   real sq_loss = dot_self(to_vector(y_test) - E_y_test);
## }
```

The program includes not only the data and model declaration, but also the predictive distributions for square error and log loss. The true values of the test cases are provided to the program, but they are not used during training.

6. Explanation on how the Stan code was run

The experiment was run with 4 chains, first with 100 iterations and then increased to 2000 iterations.

```r
# Evaluate the output after fitting
fit_logistic_regression_2000
```

```
## Inference for Stan model: logistic_regression.
## 4 chains, each with iter=2000; warmup=1000; thin=1;
## post-warmup draws per chain=1000, total post-warmup draws=4000.
##
##                  mean se_mean   sd   2.5%    25%    50%    75%  97.5% n_eff Rhat
## alpha            0.03    0.01 0.69  -1.36  -0.43   0.02   0.49   1.42  3546    1
## beta[1]          0.08    0.03 1.77  -3.39  -1.14   0.09   1.23   3.66  4984    1
## beta[2]          1.09    0.02 1.06  -0.99   0.36   1.11   1.80   3.16  3247    1
## beta[3]          0.01    0.03 1.76  -3.47  -1.16   0.01   1.15   3.54  4467    1
## beta[4]          0.56    0.03 1.79  -2.95  -0.59   0.57   1.80   4.02  4795    1
## beta[5]          1.31    0.02 1.19  -1.02   0.52   1.29   2.13   3.66  3701    1
## beta[6]         -1.19    0.02 1.58  -4.28  -2.24  -1.17  -0.13   1.83  4358    1
## beta[7]          1.46    0.03 1.68  -1.82   0.29   1.47   2.60   4.72  3920    1
## beta[8]          1.54    0.03 1.60  -1.61   0.48   1.56   2.59   4.70  3907    1
## beta[9]         -0.54    0.02 0.94  -2.39  -1.18  -0.54   0.09   1.28  3690    1
## beta[10]        -0.71    0.02 1.35  -3.36  -1.62  -0.68   0.23   1.90  3836    1
## beta[11]         2.81    0.02 1.52  -0.11   1.77   2.77   3.86   5.86  4117    1
## beta[12]        -0.46    0.02 0.88  -2.21  -1.04  -0.44   0.13   1.24  3353    1
## beta[13]         1.46    0.02 1.55  -1.53   0.41   1.43   2.50   4.45  4161    1
## beta[14]         1.93    0.03 1.75  -1.52   0.77   1.93   3.11   5.33  4331    1
## beta[15]         0.75    0.02 0.86  -0.90   0.17   0.72   1.30   2.49  3155    1
## beta[16]        -1.13    0.02 1.30  -3.71  -2.03  -1.11  -0.25   1.40  3581    1
## beta[17]        -0.27    0.02 1.31  -2.89  -1.13  -0.26   0.60   2.29  3506    1
## beta[18]         0.66    0.02 1.24  -1.71  -0.20   0.67   1.52   3.10  3558    1
## beta[19]        -0.75    0.02 1.08  -2.82  -1.49  -0.78  -0.02   1.43  3408    1
## beta[20]        -1.40    0.02 1.33  -3.97  -2.28  -1.36  -0.49   1.09  3682    1
## beta[21]         1.18    0.03 1.84  -2.44  -0.08   1.20   2.43   4.72  4084    1
## beta[22]         1.76    0.02 1.24  -0.66   0.94   1.76   2.59   4.19  2889    1
## beta[23]         0.73    0.03 1.85  -2.87  -0.52   0.73   2.01   4.37  4638    1
## beta[24]         1.36    0.03 1.83  -2.22   0.10   1.34   2.63   4.89  4209    1
## beta[25]         0.27    0.02 1.18  -2.02  -0.50   0.27   1.04   2.67  3391    1
## beta[26]        -0.05    0.03 1.49  -2.97  -1.06  -0.04   0.97   2.90  3193    1
## beta[27]         1.92    0.02 1.55  -1.18   0.85   1.93   2.98   4.91  4108    1
```

```
## beta[28]        2.09    0.02 1.55  -0.92   1.08   2.11   3.16   5.09  3844    1
## beta[29]        2.09    0.02 0.97   0.22   1.42   2.06   2.74   4.05  2549    1
## beta[30]       -0.09    0.02 1.40  -2.79  -1.04  -0.10   0.81   2.80  3623    1
## E_y_test[1]     1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3752    1
## E_y_test[2]     0.98    0.00 0.04   0.85   0.99   1.00   1.00   1.00  3587    1
## E_y_test[3]     0.97    0.00 0.10   0.64   0.99   1.00   1.00   1.00  3368    1
## E_y_test[4]     1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4040    1
## E_y_test[5]     0.99    0.00 0.06   0.94   1.00   1.00   1.00   1.00  3409    1
## E_y_test[6]     1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3621    1
## E_y_test[7]     0.80    0.00 0.25   0.11   0.70   0.92   0.98   1.00  3951    1
## E_y_test[8]     1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3631    1
## E_y_test[9]     1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4041    1
## E_y_test[10]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3994    1
## E_y_test[11]    0.11    0.00 0.13   0.00   0.03   0.06   0.14   0.50  4097    1
## E_y_test[12]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3831    1
## E_y_test[13]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4044    1
## E_y_test[14]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4017    1
## E_y_test[15]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4011    1
## E_y_test[16]    1.00    0.00 0.01   1.00   1.00   1.00   1.00   1.00  3970    1
## E_y_test[17]    0.99    0.00 0.03   0.93   0.99   1.00   1.00   1.00  3620    1
## E_y_test[18]    1.00    0.00 0.02   0.97   1.00   1.00   1.00   1.00  3350    1
## E_y_test[19]    0.99    0.00 0.02   0.94   0.99   1.00   1.00   1.00  3583    1
## E_y_test[20]    0.91    0.00 0.16   0.37   0.90   0.98   1.00   1.00  3895    1
## E_y_test[21]    0.97    0.00 0.07   0.75   0.97   0.99   1.00   1.00  2901    1
## E_y_test[22]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3724    1
## E_y_test[23]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3999    1
## E_y_test[24]    0.00    0.00 0.01   0.00   0.00   0.00   0.00   0.03  4227    1
## E_y_test[25]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.01  3861    1
## E_y_test[26]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4014    1
## E_y_test[27]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3882    1
## E_y_test[28]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3971    1
## E_y_test[29]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3663    1
## E_y_test[30]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4087    1
## E_y_test[31]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3539    1
## E_y_test[32]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3642    1
## E_y_test[33]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.01  4084    1
## E_y_test[34]    0.57    0.01 0.46   0.00   0.01   0.86   1.00   1.00  4056    1
## E_y_test[35]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4057    1
## E_y_test[36]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3740    1
## E_y_test[37]    0.00    0.00 0.01   0.00   0.00   0.00   0.00   0.00  4008    1
## E_y_test[38]    0.05    0.00 0.08   0.00   0.01   0.02   0.05   0.28  3643    1
## E_y_test[39]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4011    1
## E_y_test[40]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.01  3813    1
## E_y_test[41]    0.03    0.00 0.08   0.00   0.00   0.00   0.02   0.26  3265    1
## E_y_test[42]    0.47    0.00 0.28   0.04   0.22   0.46   0.70   0.95  4217    1
## E_y_test[43]    0.98    0.00 0.11   0.74   1.00   1.00   1.00   1.00  3681    1
## E_y_test[44]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3797    1
## E_y_test[45]    0.50    0.00 0.31   0.02   0.21   0.50   0.78   0.98  4028    1
## E_y_test[46]    0.03    0.00 0.04   0.00   0.00   0.01   0.03   0.14  3410    1
## E_y_test[47]    0.02    0.00 0.05   0.00   0.00   0.00   0.01   0.14  4176    1
## E_y_test[48]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3722    1
## E_y_test[49]    0.68    0.00 0.26   0.11   0.51   0.76   0.90   0.99  4202    1
## E_y_test[50]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4027    1
## E_y_test[51]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4050    1
```

```
## E_y_test[52]     0.24     0.00 0.22    0.01    0.06    0.16    0.37    0.80   4708     1
## E_y_test[53]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   4051     1
## E_y_test[54]     1.00     0.00 0.03    0.98    1.00    1.00    1.00    1.00   3994     1
## E_y_test[55]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3771     1
## E_y_test[56]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   3913     1
## E_y_test[57]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   4021     1
## E_y_test[58]     0.02     0.00 0.04    0.00    0.00    0.01    0.02    0.12   4132     1
## E_y_test[59]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   4078     1
## E_y_test[60]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   4005     1
## E_y_test[61]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   3857     1
## E_y_test[62]     1.00     0.00 0.01    0.99    1.00    1.00    1.00    1.00   3817     1
## E_y_test[63]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   3916     1
## E_y_test[64]     0.20     0.00 0.17    0.01    0.07    0.14    0.28    0.63   4441     1
## E_y_test[65]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3710     1
## E_y_test[66]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3845     1
## E_y_test[67]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3736     1
## E_y_test[68]     0.00     0.00 0.01    0.00    0.00    0.00    0.00    0.03   3788     1
## E_y_test[69]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3990     1
## E_y_test[70]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3599     1
## E_y_test[71]     0.00     0.00 0.02    0.00    0.00    0.00    0.00    0.01   3681     1
## E_y_test[72]     0.00     0.00 0.01    0.00    0.00    0.00    0.00    0.00   4016     1
## E_y_test[73]     0.14     0.00 0.20    0.00    0.01    0.05    0.18    0.73   4028     1
## E_y_test[74]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3793     1
## E_y_test[75]     0.15     0.00 0.21    0.00    0.01    0.05    0.20    0.80   3723     1
## E_y_test[76]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   3554     1
## E_y_test[77]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.01   3859     1
## E_y_test[78]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   3816     1
## E_y_test[79]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   4106     1
## E_y_test[80]     1.00     0.00 0.00    0.99    1.00    1.00    1.00    1.00   3669     1
## E_y_test[81]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   3504     1
## E_y_test[82]     1.00     0.00 0.01    0.98    1.00    1.00    1.00    1.00   3789     1
## E_y_test[83]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3949     1
## E_y_test[84]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   4015     1
## E_y_test[85]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3915     1
## E_y_test[86]     1.00     0.00 0.00    0.99    1.00    1.00    1.00    1.00   3627     1
## E_y_test[87]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3913     1
## E_y_test[88]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   4017     1
## E_y_test[89]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3779     1
## E_y_test[90]     0.52     0.00 0.29    0.04    0.27    0.52    0.77    0.97   4010     1
## E_y_test[91]     1.00     0.00 0.01    0.98    1.00    1.00    1.00    1.00   3985     1
## E_y_test[92]     0.54     0.01 0.43    0.00    0.03    0.66    0.99    1.00   3797     1
## E_y_test[93]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3972     1
## E_y_test[94]     0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   3892     1
## E_y_test[95]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   4017     1
## E_y_test[96]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   3675     1
## E_y_test[97]     0.08     0.00 0.13    0.00    0.01    0.03    0.10    0.49   2962     1
## E_y_test[98]     1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   4022     1
## E_y_test[99]     0.16     0.00 0.15    0.01    0.05    0.11    0.23    0.56   3457     1
## E_y_test[100]    0.55     0.01 0.31    0.03    0.26    0.58    0.84    0.99   3862     1
## E_y_test[101]    0.01     0.00 0.02    0.00    0.00    0.00    0.01    0.05   3875     1
## E_y_test[102]    1.00     0.00 0.00    1.00    1.00    1.00    1.00    1.00   3798     1
## E_y_test[103]    0.00     0.00 0.00    0.00    0.00    0.00    0.00    0.00   4113     1
## E_y_test[104]    0.56     0.00 0.28    0.06    0.32    0.58    0.80    0.98   4052     1
## E_y_test[105]    0.01     0.00 0.04    0.00    0.00    0.00    0.01    0.12   3718     1
```

```
## E_y_test[106]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3931    1
## E_y_test[107]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3867    1
## E_y_test[108]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4016    1
## E_y_test[109]   0.00    0.00 0.01   0.00   0.00   0.00   0.00   0.02  4423    1
## E_y_test[110]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4051    1
## E_y_test[111]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3885    1
## E_y_test[112]   0.01    0.00 0.02   0.00   0.00   0.00   0.01   0.05  3728    1
## E_y_test[113]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3964    1
## E_y_test[114]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3919    1
## E_y_test[115]   0.02    0.00 0.07   0.00   0.00   0.00   0.01   0.21  3979    1
## E_y_test[116]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4025    1
## E_y_test[117]   0.00    0.00 0.01   0.00   0.00   0.00   0.00   0.02  3846    1
## E_y_test[118]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4015    1
## E_y_test[119]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3413    1
## E_y_test[120]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3897    1
## E_y_test[121]   0.02    0.00 0.05   0.00   0.00   0.00   0.01   0.16  3411    1
## E_y_test[122]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4044    1
## E_y_test[123]   0.03    0.00 0.10   0.00   0.00   0.00   0.01   0.35  3711    1
## E_y_test[124]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4172    1
## E_y_test[125]   1.00    0.00 0.00   0.99   1.00   1.00   1.00   1.00  4087    1
## E_y_test[126]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3996    1
## E_y_test[127]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3974    1
## E_y_test[128]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4041    1
## E_y_test[129]   0.22    0.00 0.21   0.01   0.06   0.14   0.32   0.77  4074    1
## E_y_test[130]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4047    1
## E_y_test[131]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4011    1
## E_y_test[132]   0.00    0.00 0.01   0.00   0.00   0.00   0.00   0.02  3566    1
## E_y_test[133]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3838    1
## E_y_test[134]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4021    1
## E_y_test[135]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3620    1
## E_y_test[136]   0.51    0.01 0.37   0.00   0.11   0.52   0.90   1.00  3586    1
## E_y_test[137]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3577    1
## E_y_test[138]   0.98    0.00 0.05   0.85   0.99   1.00   1.00   1.00  3726    1
## E_y_test[139]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3991    1
## E_y_test[140]   0.00    0.00 0.01   0.00   0.00   0.00   0.00   0.03  3942    1
## E_y_test[141]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3720    1
## E_y_test[142]   0.03    0.00 0.09   0.00   0.00   0.00   0.01   0.29  3448    1
## E_y_test[143]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4016    1
## E_y_test[144]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4151    1
## E_y_test[145]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4052    1
## E_y_test[146]   0.02    0.00 0.12   0.00   0.00   0.00   0.00   0.33  3540    1
## E_y_test[147]   0.32    0.00 0.25   0.02   0.11   0.26   0.49   0.87  3999    1
## E_y_test[148]   0.00    0.00 0.01   0.00   0.00   0.00   0.00   0.01  3911    1
## E_y_test[149]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4092    1
## E_y_test[150]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3633    1
## E_y_test[151]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4033    1
## E_y_test[152]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4093    1
## E_y_test[153]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4100    1
## E_y_test[154]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3970    1
## E_y_test[155]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4017    1
## E_y_test[156]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4081    1
## E_y_test[157]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3947    1
## E_y_test[158]   0.01    0.00 0.06   0.00   0.00   0.00   0.00   0.10  3756    1
## E_y_test[159]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3313    1
```

```
## E_y_test[160]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4031    1
## E_y_test[161]   0.00    0.00 0.01   0.00   0.00   0.00   0.00   0.01  4173    1
## E_y_test[162]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4020    1
## E_y_test[163]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3525    1
## E_y_test[164]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4143    1
## E_y_test[165]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3992    1
## E_y_test[166]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3953    1
## E_y_test[167]   0.02    0.00 0.05   0.00   0.00   0.01   0.02   0.18  4007    1
## E_y_test[168]   0.01    0.00 0.04   0.00   0.00   0.00   0.00   0.03  2891    1
## E_y_test[169]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3699    1
## E_y_test[170]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4030    1
## E_y_test[171]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4004    1
## E_y_test[172]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3967    1
## E_y_test[173]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4173    1
## E_y_test[174]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4091    1
## E_y_test[175]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3588    1
## E_y_test[176]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3833    1
## E_y_test[177]   0.13    0.00 0.18   0.00   0.01   0.05   0.16   0.69  3871    1
## E_y_test[178]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4052    1
## E_y_test[179]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4015    1
## E_y_test[180]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3903    1
## E_y_test[181]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4024    1
## E_y_test[182]   0.01    0.00 0.03   0.00   0.00   0.00   0.00   0.06  3441    1
## E_y_test[183]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.01  3754    1
## E_y_test[184]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3942    1
## E_y_test[185]   0.41    0.00 0.28   0.02   0.17   0.36   0.63   0.93  3710    1
## E_y_test[186]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3718    1
## E_y_test[187]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4064    1
## E_y_test[188]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4011    1
## E_y_test[189]   0.02    0.00 0.03   0.00   0.00   0.01   0.02   0.11  4396    1
## E_y_test[190]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3901    1
## E_y_test[191]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3808    1
## E_y_test[192]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3924    1
## E_y_test[193]   0.00    0.00 0.02   0.00   0.00   0.00   0.00   0.01  2528    1
## E_y_test[194]   0.99    0.00 0.07   0.86   1.00   1.00   1.00   1.00  3582    1
## E_y_test[195]   0.03    0.00 0.07   0.00   0.00   0.01   0.03   0.23  4397    1
## E_y_test[196]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3714    1
## E_y_test[197]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3989    1
## E_y_test[198]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3830    1
## E_y_test[199]   0.00    0.00 0.02   0.00   0.00   0.00   0.00   0.00  3981    1
## E_y_test[200]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  2460    1
## E_y_test[201]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3667    1
## E_y_test[202]   0.35    0.00 0.25   0.02   0.14   0.30   0.54   0.89  4831    1
## E_y_test[203]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3533    1
## E_y_test[204]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  2441    1
## E_y_test[205]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3715    1
## E_y_test[206]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3797    1
## E_y_test[207]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4079    1
## E_y_test[208]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3820    1
## E_y_test[209]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3579    1
## E_y_test[210]   0.04    0.00 0.08   0.00   0.00   0.01   0.03   0.28  3674    1
## E_y_test[211]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.01  4112    1
## E_y_test[212]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4041    1
## E_y_test[213]   0.51    0.00 0.27   0.05   0.28   0.51   0.73   0.94  3847    1
```

```
## E_y_test[214]   0.97    0.00 0.08   0.73   0.98   1.00   1.00   1.00  3622    1
## E_y_test[215]   0.01    0.00 0.02   0.00   0.00   0.00   0.00   0.05  3758    1
## E_y_test[216]   0.01    0.00 0.06   0.00   0.00   0.00   0.00   0.12  4033    1
## E_y_test[217]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.01  4043    1
## E_y_test[218]   0.08    0.00 0.19   0.00   0.00   0.00   0.03   0.78  3376    1
## E_y_test[219]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3924    1
## E_y_test[220]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3843    1
## E_y_test[221]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4077    1
## E_y_test[222]   1.00    0.00 0.03   0.97   1.00   1.00   1.00   1.00  4026    1
## E_y_test[223]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4012    1
## E_y_test[224]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.01  3967    1
## E_y_test[225]   0.99    0.00 0.02   0.94   0.99   1.00   1.00   1.00  4119    1
## E_y_test[226]   0.01    0.00 0.01   0.00   0.00   0.00   0.01   0.04  4069    1
## E_y_test[227]   0.01    0.00 0.01   0.00   0.00   0.00   0.01   0.04  3794    1
## E_y_test[228]   0.09    0.00 0.15   0.00   0.01   0.03   0.10   0.55  4143    1
## E_y_test[229]   0.12    0.00 0.15   0.00   0.02   0.06   0.16   0.58  4538    1
## E_y_test[230]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4055    1
## E_y_test[231]   0.03    0.00 0.07   0.00   0.00   0.01   0.02   0.21  3541    1
## E_y_test[232]   0.46    0.01 0.34   0.01   0.12   0.42   0.80   0.99  4097    1
## E_y_test[233]   0.01    0.00 0.02   0.00   0.00   0.00   0.01   0.05  3260    1
## E_y_test[234]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4007    1
## E_y_test[235]   1.00     NaN 0.00   1.00   1.00   1.00   1.00   1.00   NaN    1
## E_y_test[236]   0.00    0.00 0.01   0.00   0.00   0.00   0.00   0.02  3995    1
## E_y_test[237]   0.10    0.00 0.21   0.00   0.00   0.01   0.08   0.81  3928    1
## E_y_test[238]   0.30    0.00 0.27   0.01   0.07   0.21   0.48   0.90  4374    1
## E_y_test[239]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4024    1
## E_y_test[240]   0.46    0.01 0.34   0.01   0.13   0.42   0.79   0.99  3595    1
## E_y_test[241]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.01  3806    1
## E_y_test[242]   0.26    0.00 0.28   0.00   0.03   0.15   0.43   0.94  3907    1
## E_y_test[243]   0.01    0.00 0.02   0.00   0.00   0.00   0.01   0.05  3958    1
## E_y_test[244]   0.01    0.00 0.07   0.00   0.00   0.00   0.00   0.16  3259    1
## E_y_test[245]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4006    1
## E_y_test[246]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.01  4275    1
## E_y_test[247]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3406    1
## E_y_test[248]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4014    1
## E_y_test[249]   0.97    0.00 0.07   0.78   0.98   1.00   1.00   1.00  4073    1
## E_y_test[250]   0.00    0.00 0.01   0.00   0.00   0.00   0.00   0.03  4130    1
## E_y_test[251]   0.03    0.00 0.11   0.00   0.00   0.00   0.01   0.38  3413    1
## E_y_test[252]   0.02    0.00 0.03   0.00   0.00   0.01   0.02   0.11  4096    1
## E_y_test[253]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4015    1
## E_y_test[254]   0.10    0.00 0.14   0.00   0.01   0.05   0.13   0.52  3365    1
## E_y_test[255]   0.00    0.00 0.01   0.00   0.00   0.00   0.00   0.01  3944    1
## E_y_test[256]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4036    1
## E_y_test[257]   0.02    0.00 0.06   0.00   0.00   0.00   0.01   0.18  4087    1
## E_y_test[258]   0.99    0.00 0.04   0.91   0.99   1.00   1.00   1.00  4011    1
## E_y_test[259]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3671    1
## E_y_test[260]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  3957    1
## E_y_test[261]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4073    1
## E_y_test[262]   0.01    0.00 0.01   0.00   0.00   0.00   0.00   0.04  4164    1
## E_y_test[263]   0.11    0.00 0.19   0.00   0.01   0.03   0.12   0.71  4416    1
## E_y_test[264]   0.02    0.00 0.04   0.00   0.00   0.00   0.01   0.12  3594    1
## E_y_test[265]   0.01    0.00 0.07   0.00   0.00   0.00   0.00   0.13  3634    1
## E_y_test[266]   1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4018    1
## E_y_test[267]   0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4018    1
```

```
## E_y_test[268]    0.05    0.00 0.08   0.00   0.01   0.02   0.06   0.31  3900    1
## E_y_test[269]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3581    1
## E_y_test[270]    0.45    0.01 0.31   0.01   0.16   0.42   0.74   0.97  3583    1
## E_y_test[271]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3902    1
## E_y_test[272]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4005    1
## E_y_test[273]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  4098    1
## E_y_test[274]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4053    1
## E_y_test[275]    0.22    0.00 0.31   0.00   0.01   0.05   0.34   0.98  4114    1
## E_y_test[276]    0.00    0.00 0.02   0.00   0.00   0.00   0.00   0.00  3823    1
## E_y_test[277]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3998    1
## E_y_test[278]    0.09    0.00 0.12   0.00   0.01   0.04   0.11   0.45  3920    1
## E_y_test[279]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  3970    1
## E_y_test[280]    0.00    0.00 0.00   0.00   0.00   0.00   0.00   0.00  2451    1
## E_y_test[281]    0.08    0.00 0.13   0.00   0.01   0.02   0.09   0.48  4019    1
## E_y_test[282]    0.24    0.00 0.25   0.00   0.04   0.13   0.37   0.87  4180    1
## E_y_test[283]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4043    1
## E_y_test[284]    1.00    0.00 0.00   1.00   1.00   1.00   1.00   1.00  4019    1
## log_loss        33.39    0.16 9.20  19.12  26.57  32.27  38.77  55.08  3240    1
## sq_loss          8.73    0.03 1.88   5.54   7.39   8.60   9.94  12.73  3738    1
## lp__           -32.25    0.09 3.93 -40.75 -34.69 -32.08 -29.37 -25.44  1726    1
##
## Samples were drawn using NUTS(diag_e) at Sat Dec  5 11:29:50 2020.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

The output from Stan produces the mean, standard error mean, standard deviation and the central posterior interval for parameter values α^m, β^m, drawn from the posterior. With only 'rN' training examples with correlated parameters, there is a great deal of uncertainty in the posteriors for coefficients and hence for predictions and for loss versus the test set.

Explanation on the Stan code:

There are 4 components attributing to the Stan model:
- Data:
K: the number of observations in the training set
N: the number of columns in the training set
x: the training dataset, in matrix format
y: the training target

- Parameters:
alpha: The intercept
beta: The weight of all features

- Model:
alpha ~ normal(0,2)
beta ~ normal(0,2)

7. Convergence Diagnostics

$\hat{R}$ **diagnostic:** First, I tried running the model with 100 iterations, and evaluating the $\hat{R}$ produced in the summary of the model indicates that the chains are far from convergence. The $\hat{R}$ values for numerous β are > 1, so therefore approximate convergence has not been reached. Then, I tried to increase the number of iterations to 1000, and checking the $\hat{R}$ value indicates that both of the $\hat{R}$ for $\alpha and \beta$ are ≈ 1, which shows that the chains have converged and the draws from posterior distribution follow the same distribution of the real data.

The Effective Sample Size (ESS): The Effective Sample Size (ESS) of a parameter sampled from an MCMC is the number of effectively independent draws from the posterior distribution that the Markov chain is equivalent to. From the histogram above, we see that with 2000 iterations, the effective sample size for most of the parameter ranging from 1000 to 5000, with its peak is within the [2000, 3000] interval, which is quite small compared to the total sample size of 8000 data points. In addition, due to the definition of the effective sample size, these sample size are independent of each other.

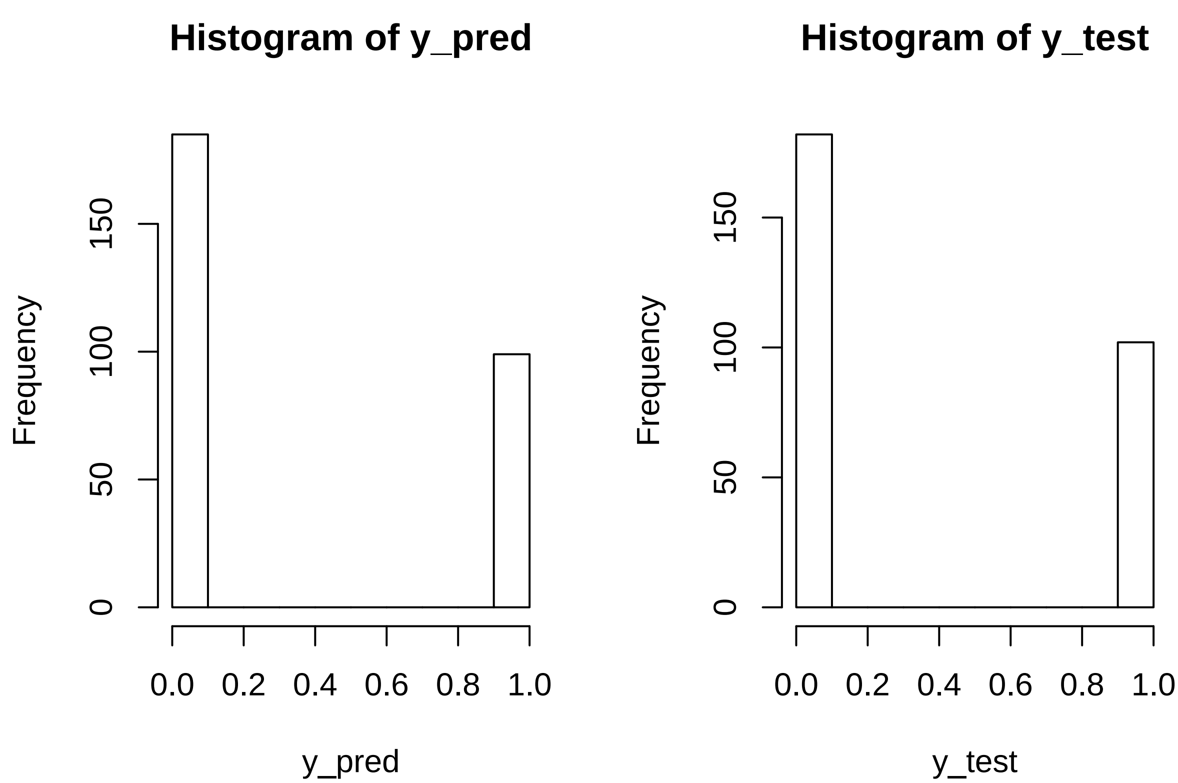
8. Posterior predictive checking

```r
# Extract the y_pred from the model
y_pred = round(extract(fit_logistic_regression_2000)$E_y_test[284,],0)

par(mfrow=c(1,2))

# Plot the histogram of the predicted label
hist(y_pred)

# Plot the histogram of the real label
hist(y_test)
```



Above are 2 histograms that show the frequency of both labels 0 and 1 for the testing data. The first histogram is obtained through feeding the testing data into the logistic regression model and using the learned parameter of the model to get the predicted label. The second histogram is the frequency of the labels of the real testing data.

From the graph above, we see that the distribution of the predicted label is roughly the same as the distribution for the true label of the test data. The graph shows that it is likely that the model has learned the distribution of the data through the training, and indicates that we might have already obtained a good model. However, more tests / evaluation methods would be made in the next part in order to evaluate whether the parameter is sufficiently fit before coming to conclusion.

9. Model comparison

10. Predictive performance assessment

We will consider full Bayesian inference, where our estimate is derived from:

$$p(\tilde{y}_n|\tilde{x}_n, x, y) = \int p(\tilde{y}_n|\tilde{x}_n, \alpha, \beta) * p(\alpha, \beta|x, y) d(\alpha, \beta).$$

```
N_test = nrow(df_test)
y_test_hat_bayes <- rep(0, N_test)

# Extract the predicted label for the test set
for (n in 1:N_test) {
  y_test_hat_bayes[n] <- mean(extract(fit_logistic_regression_2000)$E_y_test[,n])
}

# Extract the log loss and square loss
log_loss = mean(extract(fit_logistic_regression_2000)$log_loss)
square_loss = mean(extract(fit_logistic_regression_2000)$sq_loss)
cat("Log loss:", round(log_loss,2), "Square loss:", round(square_loss,2))
```

```
## Log loss: 33.39 Square loss: 8.73
```

To evaluate the accuracy of the model, we would take a look at how many samples have been misclassified:

```
# We set a threshold to be 0.5, that is which sample < 0.5 would be classified as 0, and >= 0.5 would b
y_test_hat_bayes = round(y_test_hat_bayes, 0)

# Count the number of 0 and 1 in the prediction
occurence_count_pred = table(y_test_hat_bayes)
cat("The occurence of 0 and 1 in the predicted label:", occurence_count_pred, "\n")
```

```
## The occurence of 0 and 1 in the predicted label: 182 102
```

```
# Count the number of 0 and 1 in the real data
occurence_count_real = table(y_test)
cat("The occurence of 0 and 1 in the real label:", occurence_count_real, "\n")
```

```
## The occurence of 0 and 1 in the real label: 182 102
```

```
# Calculate the accuracy of the model
misclassification_count = (occurence_count_pred[1] - occurence_count_real[1]) +  (occurence_count_pred[
cat("The number of misclassification is ", misclassification_count)
```

```
## The number of misclassification is  0
```

11. Sensitivity Analysis