Regression in Stan

Read in the cancer data

```
#load("~/Dropbox/My Documents/Teaching/772 Statistics/772 Spring 2020/Classes/Week 8-Linear Multiple Re
load("~/Google Drive crowston@syr.edu/Courses/IST 772 Crowston/Week 8/cancer.RData")
```

```
mycancer<-subset(cancer,select=-c(Geography,
                                  MedianAge,AvgHouseholdSize,PctEmployed16_Over,PctSomeCol18_24,
                                  popEst2015,avgDeathsPerYear,MedianAgeMale, PctMarriedHouseholds,
                                  PctEmpPrivCoverage,PctPublicCoverage,PctHS25_Over))
```

Bayesian regression using BayseFactor

```
library(BayesFactor)
```

```
## Loading required package: coda
```

```
## Loading required package: Matrix
```

```
## ************
## Welcome to BayesFactor 0.9.12-4.2. If you have questions, please contact Richard Morey (richarddmore
##
## Type BFManual() to open the manual.
## ************
```

```
cancer.mcmc <- lmBF(deathrate ~ ., data=mycancer, posterior=TRUE, iterations=10000)
```

```
## Warning: data coerced from tibble to data frame
```

```
summary(cancer.mcmc)
```

```
##
## Iterations = 1:10000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 10000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##                          Mean        SD  Naive SE Time-series SE
## mu                  1.787e+02 3.650e-01 3.650e-03      3.597e-03
## incidenceRate       2.173e-01 7.141e-03 7.141e-05      7.141e-05
## medIncome          -1.505e-04 5.370e-05 5.370e-07      5.370e-07
## MedianAgeFemale    -1.540e-01 8.044e-02 8.044e-04      7.913e-04
## PercentMarried     -1.712e-01 7.665e-02 7.665e-04      7.789e-04
## PctBachDeg25_Over  -1.690e+00 1.043e-01 1.043e-03      1.043e-03
## PctUnemployed16_Over  6.166e-01 1.540e-01 1.540e-03      1.540e-03
## PctPrivateCoverage -3.074e-01 6.254e-02 6.254e-04      6.254e-04
## BirthRate          -8.740e-01 1.925e-01 1.925e-03      1.925e-03
## sig2                4.079e+02 1.054e+01 1.054e-01      1.071e-01
## g                   1.449e-01 9.480e-02 9.480e-04      9.480e-04
##
```

```
## 2. Quantiles for each variable:
##
##                             2.5%        25%        50%        75%
## mu                     1.779e+02  1.784e+02  1.787e+02  1.789e+02
## incidenceRate          2.033e-01  2.125e-01  2.173e-01  2.221e-01
## medIncome             -2.556e-04 -1.872e-04 -1.504e-04 -1.136e-04
## MedianAgeFemale       -3.126e-01 -2.077e-01 -1.543e-01 -9.942e-02
## PercentMarried        -3.214e-01 -2.219e-01 -1.706e-01 -1.204e-01
## PctBachDeg25_Over     -1.894e+00 -1.760e+00 -1.690e+00 -1.619e+00
## PctUnemployed16_Over  3.138e-01  5.136e-01  6.156e-01  7.196e-01
## PctPrivateCoverage   -4.293e-01 -3.502e-01 -3.075e-01 -2.656e-01
## BirthRate             -1.254e+00 -1.004e+00 -8.734e-01 -7.444e-01
## sig2                   3.878e+02  4.009e+02  4.077e+02  4.148e+02
## g                      5.366e-02  8.933e-02  1.209e-01  1.708e-01
##                            97.5%
## mu                     1.794e+02
## incidenceRate          2.313e-01
## medIncome             -4.636e-05
## MedianAgeFemale        5.763e-06
## PercentMarried        -2.411e-02
## PctBachDeg25_Over     -1.488e+00
## PctUnemployed16_Over  9.201e-01
## PctPrivateCoverage   -1.842e-01
## BirthRate             -5.028e-01
## sig2                   4.290e+02
## g                      3.831e-01
```

Bayesean regression using Stan

A straightforward stan model to do regression. mu is the intercept; beta are the regression weights. beta * x is a matrix multiplication that, added to mu, gives a vector of predicted values. The model is the predicted values with some error, parametrized by a normal distribution with standard deviation sigma. I added priors for the parameters to give the model a hint about where to go.

```
data {
  int<lower=1> N;
  int<lower=1> K;
  vector[N] y;
  matrix[N,K] x;
}

parameters {
  real mu;
  vector[K] beta;
  real<lower=0> sigma;
}

model {
  mu ~ cauchy(0,10);        //prior for the intercept following Gelman 2008
  beta ~ cauchy(0,2.5);     //prior for the slopes following Gelman 2008

  y ~ normal(x * beta + mu, sigma);
}
```

Put data in a list for stan.

```
fit_data<-list(N=length(mycancer$deathrate), K=length(mycancer) - 1,
               y=mycancer$deathrate, x=subset(mycancer, select=-deathrate))
```

Sample from the model. NB. This takes a long time to run.

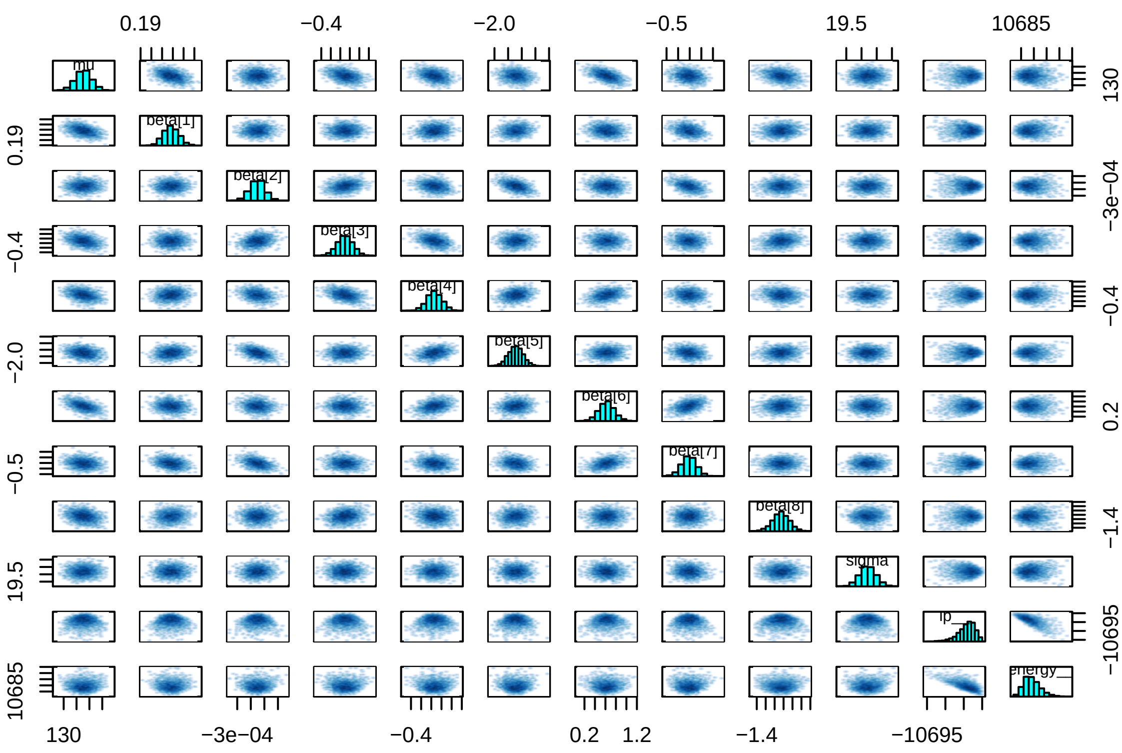
```
#install.packages("rstan")
library(rstan)
```

```
## Loading required package: StanHeaders

## Loading required package: ggplot2

## rstan (Version 2.19.3, 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)

##
## Attaching package: 'rstan'

## The following object is masked from 'package:coda':
##
##     traceplot
```

```
options(mc.cores = parallel::detectCores())

reg_fit <- sampling(reg_model, data=fit_data, control=list(max_treedepth=30))
```

Diagnositics. Without the priors, the model doesn't converge.

```
pairs(reg_fit)
```



Examine the results.

```r
summary(reg_fit)
```

```
## $summary
##                  mean      se_mean           sd          2.5%
## mu       1.453420e+02 1.179632e-01 6.196327e+00  1.330842e+02
## beta[1]  2.183873e-01 1.269352e-04 7.112157e-03  2.052157e-01
## beta[2] -1.520827e-04 9.703540e-07 5.506764e-05 -2.607725e-04
## beta[3] -1.508916e-01 1.443067e-03 8.304081e-02 -3.169470e-01
## beta[4] -1.705810e-01 1.342111e-03 7.535005e-02 -3.200968e-01
## beta[5] -1.689399e+00 2.427790e-03 1.066995e-01 -1.899711e+00
## beta[6]  6.210651e-01 2.756236e-03 1.482007e-01  3.344004e-01
## beta[7] -3.078871e-01 1.285683e-03 6.304136e-02 -4.296415e-01
## beta[8] -8.601557e-01 4.241269e-03 1.888780e-01 -1.250780e+00
## sigma    2.020192e+01 4.372150e-03 2.635809e-01  1.969520e+01
## lp__    -1.068425e+04 5.947169e-02 2.313353e+00 -1.068988e+04
##                   25%           50%           75%         97.5%    n_eff
## mu       1.412011e+02  1.453362e+02  1.494685e+02  1.574779e+02 2759.152
## beta[1]  2.134941e-01  2.182289e-01  2.232415e-01  2.328229e-01 3139.342
## beta[2] -1.894059e-04 -1.512734e-04 -1.158855e-04 -4.445364e-05 3220.568
## beta[3] -2.055173e-01 -1.507974e-01 -9.539046e-02  1.050142e-02 3311.388
## beta[4] -2.202067e-01 -1.704605e-01 -1.192362e-01 -2.773971e-02 3152.030
## beta[5] -1.760307e+00 -1.687617e+00 -1.616216e+00 -1.479014e+00 1931.535
## beta[6]  5.226038e-01  6.220346e-01  7.193110e-01  9.160227e-01 2891.131
## beta[7] -3.499004e-01 -3.078497e-01 -2.665828e-01 -1.835929e-01 2404.272
## beta[8] -9.803209e-01 -8.585088e-01 -7.340854e-01 -4.919794e-01 1983.222
## sigma    2.002512e+01  2.020012e+01  2.037951e+01  2.071863e+01 3634.443
## lp__    -1.068552e+04 -1.068390e+04 -1.068259e+04 -1.068081e+04 1513.084
##              Rhat
## mu      1.0003766
## beta[1] 1.0018030
## beta[2] 1.0021268
## beta[3] 1.0002800
## beta[4] 1.0004209
## beta[5] 1.0016426
## beta[6] 1.0010618
## beta[7] 1.0011238
## beta[8] 1.0005546
## sigma   0.9991639
## lp__    1.0017157
##
## $c_summary
## , , chains = chain:1
##
##          stats
## parameter          mean           sd          2.5%           25%
##   mu       1.457085e+02 6.178298e+00  1.330842e+02  1.416551e+02
##   beta[1]  2.180957e-01 7.282832e-03  2.038142e-01  2.132093e-01
##   beta[2] -1.466460e-04 5.475659e-05 -2.540690e-04 -1.840713e-04
##   beta[3] -1.511522e-01 8.108885e-02 -3.159480e-01 -2.025135e-01
##   beta[4] -1.745450e-01 7.687582e-02 -3.241785e-01 -2.249107e-01
##   beta[5] -1.698239e+00 1.156313e-01 -1.930429e+00 -1.777754e+00
##   beta[6]  6.205272e-01 1.402994e-01  3.515563e-01  5.312380e-01
##   beta[7] -3.095908e-01 6.162162e-02 -4.211957e-01 -3.537936e-01
```

```
##   beta[8] -8.676997e-01 1.945849e-01 -1.269064e+00 -9.837245e-01
##   sigma    2.019970e+01 2.798070e-01  1.968971e+01  2.001185e+01
##   lp__    -1.068440e+04 2.341742e+00 -1.068991e+04 -1.068573e+04
##          stats
## parameter           50%           75%         97.5%
##   mu       1.458524e+02  1.495492e+02  1.577990e+02
##   beta[1]  2.179696e-01  2.230369e-01  2.319041e-01
##   beta[2] -1.456217e-04 -1.101528e-04 -4.310245e-05
##   beta[3] -1.523037e-01 -9.597310e-02  9.576896e-03
##   beta[4] -1.728924e-01 -1.211187e-01 -2.821425e-02
##   beta[5] -1.696866e+00 -1.616651e+00 -1.483673e+00
##   beta[6]  6.175533e-01  7.108579e-01  9.010978e-01
##   beta[7] -3.085051e-01 -2.690156e-01 -1.914728e-01
##   beta[8] -8.651332e-01 -7.435069e-01 -4.707615e-01
##   sigma    2.018111e+01  2.038555e+01  2.077586e+01
##   lp__    -1.068405e+04 -1.068265e+04 -1.068086e+04
##
## , , chains = chain:2
##
##          stats
## parameter          mean           sd          2.5%           25%
##   mu       1.454016e+02 6.019900e+00  1.341533e+02  1.411748e+02
##   beta[1]  2.181868e-01 6.948981e-03  2.052896e-01  2.133134e-01
##   beta[2] -1.541208e-04 5.517688e-05 -2.646248e-04 -1.918817e-04
##   beta[3] -1.517979e-01 8.152655e-02 -3.073893e-01 -2.085890e-01
##   beta[4] -1.703159e-01 7.392441e-02 -3.110111e-01 -2.192957e-01
##   beta[5] -1.685931e+00 1.032555e-01 -1.884872e+00 -1.755639e+00
##   beta[6]  6.230326e-01 1.497941e-01  3.283191e-01  5.279363e-01
##   beta[7] -3.063003e-01 6.441803e-02 -4.303957e-01 -3.480742e-01
##   beta[8] -8.625091e-01 1.771647e-01 -1.209320e+00 -9.784211e-01
##   sigma    2.019738e+01 2.529792e-01  1.969632e+01  2.003216e+01
##   lp__    -1.068405e+04 2.165349e+00 -1.068905e+04 -1.068533e+04
##          stats
## parameter           50%           75%         97.5%
##   mu       1.455087e+02  1.496217e+02  1.567388e+02
##   beta[1]  2.183920e-01  2.229025e-01  2.323491e-01
##   beta[2] -1.526381e-04 -1.189707e-04 -4.511178e-05
##   beta[3] -1.510954e-01 -9.824781e-02  7.539858e-03
##   beta[4] -1.717140e-01 -1.191325e-01 -2.773971e-02
##   beta[5] -1.687231e+00 -1.616054e+00 -1.471829e+00
##   beta[6]  6.236076e-01  7.217794e-01  9.247950e-01
##   beta[7] -3.059051e-01 -2.650695e-01 -1.797588e-01
##   beta[8] -8.548317e-01 -7.516525e-01 -4.996716e-01
##   sigma    2.019192e+01  2.037127e+01  2.069141e+01
##   lp__    -1.068376e+04 -1.068250e+04 -1.068060e+04
##
## , , chains = chain:3
##
##          stats
## parameter          mean           sd          2.5%           25%
##   mu       1.449953e+02 6.4346526388  1.325922e+02  1.404508e+02
##   beta[1]  2.182453e-01 0.0070758011  2.052862e-01  2.133563e-01
##   beta[2] -1.566085e-04 0.0000556925 -2.650883e-04 -1.928893e-04
##   beta[3] -1.520334e-01 0.0828034881 -3.175737e-01 -2.063225e-01
```

```
##   beta[4] -1.666558e-01 0.0752150736 -3.178614e-01 -2.141990e-01
##   beta[5] -1.683178e+00 0.1058883147 -1.896152e+00 -1.755891e+00
##   beta[6]  6.302422e-01 0.1573829178  3.391676e-01  5.109481e-01
##   beta[7] -3.041301e-01 0.0639291000 -4.242379e-01 -3.465032e-01
##   beta[8] -8.478482e-01 0.1959612263 -1.268243e+00 -9.666512e-01
##   sigma    2.020442e+01 0.2675300320  1.969033e+01  2.002091e+01
##   lp__    -1.068434e+04 2.4708949373 -1.069034e+04 -1.068565e+04
##          stats
## parameter           50%           75%         97.5%
##   mu       1.448465e+02  1.494471e+02  1.576058e+02
##   beta[1]  2.178679e-01  2.234008e-01  2.327910e-01
##   beta[2] -1.538505e-04 -1.189103e-04 -5.485512e-05
##   beta[3] -1.505084e-01 -9.726874e-02  3.295534e-04
##   beta[4] -1.660978e-01 -1.160786e-01 -2.883528e-02
##   beta[5] -1.679564e+00 -1.610061e+00 -1.475230e+00
##   beta[6]  6.321822e-01  7.377431e-01  9.275928e-01
##   beta[7] -3.039361e-01 -2.628039e-01 -1.746951e-01
##   beta[8] -8.455422e-01 -7.097622e-01 -4.822429e-01
##   sigma    2.020273e+01  2.038278e+01  2.075218e+01
##   lp__    -1.068389e+04 -1.068259e+04 -1.068089e+04
##
## , , chains = chain:4
##
##          stats
## parameter          mean           sd          2.5%           25%
##   mu       1.452627e+02 6.132961e+00  1.328271e+02  1.413704e+02
##   beta[1]  2.190215e-01 7.109142e-03  2.056471e-01  2.141144e-01
##   beta[2] -1.509556e-04 5.421196e-05 -2.566369e-04 -1.881187e-04
##   beta[3] -1.485832e-01 8.670573e-02 -3.209502e-01 -2.044878e-01
##   beta[4] -1.708073e-01 7.526140e-02 -3.246272e-01 -2.220873e-01
##   beta[5] -1.690246e+00 1.009901e-01 -1.883879e+00 -1.756828e+00
##   beta[6]  6.104585e-01 1.443161e-01  3.212109e-01  5.118533e-01
##   beta[7] -3.115271e-01 6.198494e-02 -4.366851e-01 -3.501708e-01
##   beta[8] -8.625660e-01 1.869102e-01 -1.227958e+00 -9.871910e-01
##   sigma    2.020616e+01 2.533686e-01  1.971436e+01  2.003537e+01
##   lp__    -1.068421e+04 2.252189e+00 -1.068955e+04 -1.068541e+04
##          stats
## parameter           50%           75%         97.5%
##   mu       1.451127e+02  1.491401e+02  1.573245e+02
##   beta[1]  2.187337e-01  2.238799e-01  2.341288e-01
##   beta[2] -1.528167e-04 -1.133410e-04 -3.790391e-05
##   beta[3] -1.488651e-01 -9.048915e-02  2.004360e-02
##   beta[4] -1.715645e-01 -1.193981e-01 -2.687656e-02
##   beta[5] -1.688306e+00 -1.623448e+00 -1.484148e+00
##   beta[6]  6.155825e-01  7.020982e-01  8.917726e-01
##   beta[7] -3.118210e-01 -2.725422e-01 -1.926567e-01
##   beta[8] -8.649962e-01 -7.326199e-01 -4.970573e-01
##   sigma    2.021250e+01  2.038158e+01  2.067951e+01
##   lp__    -1.068387e+04 -1.068256e+04 -1.068081e+04
```

A better model for regression using Stan

A more efficient stan model to do regression (a QR decomposition of the x data). It's better because the parameters to be estimated are in the same scale, so easier for the MCMC to find.

```
data {
  int<lower=0> N;   // number of data items
  int<lower=0> K;   // number of predictors
  matrix[N, K] x;   // predictor matrix
  vector[N] y;      // outcome vector
}

transformed data {
  matrix[N, K] Q_ast;
  matrix[K, K] R_ast;
  matrix[K, K] R_ast_inverse;

  // thin and scale the QR decomposition
  Q_ast = qr_Q(x)[, 1:K] * sqrt(N - 1);
  R_ast = qr_R(x)[1:K, ] / sqrt(N - 1);
  R_ast_inverse = inverse(R_ast);
}

parameters {
  real mu; // intercept
  vector[K] theta; // coefficients on Q_ast
  real<lower=0> sigma; // error scale
}

model {
  y ~ normal(Q_ast * theta + mu, sigma); // likelihood
}

generated quantities {
  vector[K] beta;
  beta = R_ast_inverse * theta; // coefficients on x
}
```

Sample from the model.

```r
#install.packages("rstan")
library(rstan)
options(mc.cores = parallel::detectCores())

qr_fit <- sampling(qr_model, data=fit_data)
```

Look at results. Note that the Rhats are all nearly 1, indicating good convergence.

```r
summary(qr_fit)
```

```
## $summary
##                   mean      se_mean           sd          2.5%
## mu        1.458747e+02 2.060855e-01 6.390318e+00  1.331551e+02
## theta[1]  3.406522e+01 2.047077e-01 6.359492e+00  2.128717e+01
## theta[2] -1.434977e+01 1.253931e-02 4.873161e-01 -1.529625e+01
## theta[3] -2.750532e+00 1.800887e-02 6.175404e-01 -3.979417e+00
## theta[4] -1.529839e+00 7.977054e-03 4.029921e-01 -2.331605e+00
## theta[5] -7.130022e+00 7.129944e-03 3.610263e-01 -7.840445e+00
## theta[6]  2.413679e+00 1.115265e-02 4.460150e-01  1.519385e+00
## theta[7] -1.897695e+00 7.017802e-03 3.723512e-01 -2.635636e+00
## theta[8] -1.741702e+00 7.625401e-03 3.809545e-01 -2.485124e+00
```

```
## sigma     2.019304e+01 4.374144e-03 2.615284e-01  1.968751e+01
## beta[1]   2.180945e-01 9.673700e-05 7.127644e-03  2.042437e-01
## beta[2]  -1.520290e-04 8.714112e-07 5.288583e-05 -2.538467e-04
## beta[3]  -1.576138e-01 1.446323e-03 8.298446e-02 -3.241412e-01
## beta[4]  -1.695506e-01 1.524771e-03 7.682900e-02 -3.230971e-01
## beta[5]  -1.694839e+00 1.959096e-03 1.011659e-01 -1.894550e+00
## beta[6]   6.220304e-01 3.810697e-03 1.535109e-01  3.198445e-01
## beta[7]  -3.076855e-01 1.146663e-03 6.105488e-02 -4.284611e-01
## beta[8]  -8.808723e-01 3.856574e-03 1.926691e-01 -1.256860e+00
## lp__     -1.067823e+04 5.352022e-02 2.277030e+00 -1.068364e+04
##                    25%           50%           75%         97.5%     n_eff
## mu        1.415253e+02  1.457072e+02  1.502284e+02  1.586629e+02  961.5014
## theta[1]  2.979665e+01  3.429991e+01  3.839122e+01  4.660795e+01  965.1092
## theta[2] -1.467784e+01 -1.434053e+01 -1.402713e+01 -1.338385e+01 1510.3384
## theta[3] -3.169268e+00 -2.746973e+00 -2.339072e+00 -1.554077e+00 1175.8658
## theta[4] -1.794554e+00 -1.526038e+00 -1.259564e+00 -7.366691e-01 2552.1613
## theta[5] -7.365434e+00 -7.131448e+00 -6.898181e+00 -6.416686e+00 2563.9257
## theta[6]  2.106672e+00  2.417289e+00  2.728259e+00  3.267545e+00 1599.3477
## theta[7] -2.141989e+00 -1.900896e+00 -1.650355e+00 -1.164058e+00 2815.1610
## theta[8] -2.003032e+00 -1.730788e+00 -1.484232e+00 -1.003409e+00 2495.8641
## sigma     2.001306e+01  2.019012e+01  2.036862e+01  2.071242e+01 3574.7971
## beta[1]   2.132997e-01  2.179790e-01  2.230069e-01  2.318999e-01 5428.8360
## beta[2]  -1.880520e-04 -1.523159e-04 -1.162235e-04 -4.849081e-05 3683.2600
## beta[3]  -2.129414e-01 -1.568231e-01 -1.011397e-01  2.700654e-03 3292.0232
## beta[4]  -2.197178e-01 -1.689062e-01 -1.174675e-01 -1.803268e-02 2538.8738
## beta[5]  -1.760405e+00 -1.695272e+00 -1.628699e+00 -1.496231e+00 2666.5916
## beta[6]   5.180455e-01  6.234254e-01  7.301166e-01  9.164528e-01 1622.8184
## beta[7]  -3.477922e-01 -3.085248e-01 -2.671044e-01 -1.869526e-01 2835.1031
## beta[8]  -1.013041e+00 -8.753520e-01 -7.506559e-01 -5.074777e-01 2495.8641
## lp__     -1.067945e+04 -1.067791e+04 -1.067663e+04 -1.067480e+04 1810.0967
##               Rhat
## mu       1.0012289
## theta[1] 1.0012294
## theta[2] 1.0004491
## theta[3] 1.0004583
## theta[4] 1.0004545
## theta[5] 1.0002715
## theta[6] 1.0016189
## theta[7] 1.0008732
## theta[8] 1.0001159
## sigma    1.0001592
## beta[1]  0.9992738
## beta[2]  1.0008910
## beta[3]  0.9993422
## beta[4]  1.0003222
## beta[5]  0.9999258
## beta[6]  1.0016847
## beta[7]  1.0008636
## beta[8]  1.0001159
## lp__     1.0004308
##
## $c_summary
## , , chains = chain:1
##
```

```
##           stats
## parameter           mean           sd          2.5%           25%
##   mu        1.458582e+02 6.230616e+00  1.337413e+02  1.418520e+02
##   theta[1]  3.406548e+01 6.191977e+00  2.163115e+01  3.012235e+01
##   theta[2] -1.434739e+01 4.769227e-01 -1.527390e+01 -1.466567e+01
##   theta[3] -2.755347e+00 6.030038e-01 -3.915950e+00 -3.152418e+00
##   theta[4] -1.525182e+00 4.097129e-01 -2.290345e+00 -1.802062e+00
##   theta[5] -7.139641e+00 3.676247e-01 -7.830700e+00 -7.380827e+00
##   theta[6]  2.426224e+00 4.560067e-01  1.558933e+00  2.101485e+00
##   theta[7] -1.903011e+00 3.598234e-01 -2.580527e+00 -2.154880e+00
##   theta[8] -1.748962e+00 3.825327e-01 -2.514047e+00 -2.015788e+00
##   sigma     2.019341e+01 2.576630e-01  1.967415e+01  2.001533e+01
##   beta[1]   2.180794e-01 6.956302e-03  2.046808e-01  2.133113e-01
##   beta[2]  -1.507086e-04 5.471368e-05 -2.552628e-04 -1.891370e-04
##   beta[3]  -1.593768e-01 8.174836e-02 -3.136013e-01 -2.144557e-01
##   beta[4]  -1.678025e-01 7.813139e-02 -3.111671e-01 -2.239322e-01
##   beta[5]  -1.696676e+00 1.029772e-01 -1.896582e+00 -1.760723e+00
##   beta[6]   6.257240e-01 1.553323e-01  3.213633e-01  5.261575e-01
##   beta[7]  -3.085423e-01 5.904823e-02 -4.193385e-01 -3.501609e-01
##   beta[8]  -8.845439e-01 1.934673e-01 -1.271488e+00 -1.019492e+00
##   lp__     -1.067826e+04 2.251253e+00 -1.068356e+04 -1.067943e+04
##           stats
## parameter            50%           75%         97.5%
##   mu        1.457738e+02  1.499742e+02  1.582546e+02
##   theta[1]  3.417500e+01  3.799888e+01  4.618881e+01
##   theta[2] -1.433653e+01 -1.402429e+01 -1.340680e+01
##   theta[3] -2.776243e+00 -2.365849e+00 -1.522549e+00
##   theta[4] -1.526313e+00 -1.246821e+00 -6.812858e-01
##   theta[5] -7.138114e+00 -6.908352e+00 -6.381947e+00
##   theta[6]  2.431739e+00  2.729762e+00  3.301442e+00
##   theta[7] -1.903833e+00 -1.657613e+00 -1.195041e+00
##   theta[8] -1.744076e+00 -1.490424e+00 -1.000965e+00
##   sigma     2.019700e+01  2.036785e+01  2.068855e+01
##   beta[1]   2.178751e-01  2.228838e-01  2.309142e-01
##   beta[2]  -1.521733e-04 -1.096522e-04 -4.455347e-05
##   beta[3]  -1.571571e-01 -1.037256e-01 -3.208894e-03
##   beta[4]  -1.668951e-01 -1.130978e-01 -1.760319e-02
##   beta[5]  -1.698278e+00 -1.631595e+00 -1.496587e+00
##   beta[6]   6.254889e-01  7.280883e-01  9.290150e-01
##   beta[7]  -3.087694e-01 -2.683732e-01 -1.934647e-01
##   beta[8]  -8.820725e-01 -7.537875e-01 -5.062419e-01
##   lp__     -1.067793e+04 -1.067666e+04 -1.067498e+04
##
## , , chains = chain:2
##
##           stats
## parameter           mean           sd          2.5%           25%
##   mu        1.458171e+02 6.503966e+00  1.329102e+02  1.414054e+02
##   theta[1]  3.414307e+01 6.456364e+00  2.143747e+01  2.991045e+01
##   theta[2] -1.436475e+01 4.896594e-01 -1.531478e+01 -1.469892e+01
##   theta[3] -2.738459e+00 6.409263e-01 -3.991109e+00 -3.180533e+00
##   theta[4] -1.512940e+00 3.862587e-01 -2.288847e+00 -1.770694e+00
##   theta[5] -7.120155e+00 3.574488e-01 -7.842280e+00 -7.354500e+00
##   theta[6]  2.400623e+00 4.587556e-01  1.514310e+00  2.062391e+00
```

```
##   theta[7] -1.880970e+00 3.655831e-01 -2.609941e+00 -2.123161e+00
##   theta[8] -1.732249e+00 3.800018e-01 -2.459144e+00 -2.007859e+00
##   sigma     2.019420e+01 2.571867e-01  1.970336e+01  2.001798e+01
##   beta[1]   2.181912e-01 7.013294e-03  2.042230e-01  2.135638e-01
##   beta[2]  -1.562555e-04 5.266992e-05 -2.565196e-04 -1.924766e-04
##   beta[3]  -1.582158e-01 8.534246e-02 -3.381049e-01 -2.133342e-01
##   beta[4]  -1.683276e-01 7.463877e-02 -3.203449e-01 -2.158305e-01
##   beta[5]  -1.693827e+00 9.921261e-02 -1.894088e+00 -1.761190e+00
##   beta[6]   6.194407e-01 1.555290e-01  3.198445e-01  5.086065e-01
##   beta[7]  -3.049609e-01 5.998072e-02 -4.244057e-01 -3.446459e-01
##   beta[8]  -8.760912e-01 1.921873e-01 -1.243721e+00 -1.015482e+00
##   lp__     -1.067818e+04 2.243695e+00 -1.068316e+04 -1.067939e+04
##           stats
## parameter           50%           75%         97.5%
##   mu        1.454551e+02  1.501646e+02  1.583726e+02
##   theta[1]  3.446324e+01  3.853951e+01  4.675830e+01
##   theta[2] -1.438057e+01 -1.404074e+01 -1.340832e+01
##   theta[3] -2.737887e+00 -2.309779e+00 -1.467966e+00
##   theta[4] -1.512080e+00 -1.237549e+00 -7.831483e-01
##   theta[5] -7.124506e+00 -6.870433e+00 -6.475477e+00
##   theta[6]  2.404830e+00  2.741301e+00  3.297225e+00
##   theta[7] -1.882795e+00 -1.641997e+00 -1.153732e+00
##   theta[8] -1.710992e+00 -1.464002e+00 -1.025083e+00
##   sigma     2.018593e+01  2.036597e+01  2.069148e+01
##   beta[1]   2.183703e-01  2.233661e-01  2.311574e-01
##   beta[2]  -1.578422e-04 -1.202307e-04 -5.392926e-05
##   beta[3]  -1.546004e-01 -1.002606e-01  8.422440e-03
##   beta[4]  -1.686011e-01 -1.165424e-01 -2.126245e-02
##   beta[5]  -1.692713e+00 -1.625527e+00 -1.501083e+00
##   beta[6]   6.186011e-01  7.323552e-01  9.186148e-01
##   beta[7]  -3.049916e-01 -2.664287e-01 -1.846730e-01
##   beta[8]  -8.653406e-01 -7.404244e-01 -5.184394e-01
##   lp__     -1.067785e+04 -1.067662e+04 -1.067473e+04
##
## , , chains = chain:3
##
##           stats
## parameter          mean           sd          2.5%           25%
##   mu        1.460697e+02 6.391671e+00  1.335583e+02  1.414472e+02
##   theta[1]  3.387274e+01 6.377796e+00  2.128599e+01  2.935863e+01
##   theta[2] -1.435245e+01 4.943088e-01 -1.528656e+01 -1.469591e+01
##   theta[3] -2.769441e+00 6.258549e-01 -4.015053e+00 -3.205328e+00
##   theta[4] -1.534589e+00 4.030120e-01 -2.336385e+00 -1.794268e+00
##   theta[5] -7.125424e+00 3.671877e-01 -7.860672e+00 -7.365480e+00
##   theta[6]  2.399965e+00 4.354637e-01  1.531770e+00  2.116694e+00
##   theta[7] -1.918972e+00 3.801917e-01 -2.693719e+00 -2.159492e+00
##   theta[8] -1.749390e+00 3.768564e-01 -2.465526e+00 -2.003032e+00
##   sigma     2.019274e+01 2.639420e-01  1.969432e+01  2.001182e+01
##   beta[1]   2.181694e-01 7.229368e-03  2.038121e-01  2.134753e-01
##   beta[2]  -1.508740e-04 5.237011e-05 -2.535187e-04 -1.853842e-04
##   beta[3]  -1.581473e-01 8.614017e-02 -3.295293e-01 -2.156503e-01
##   beta[4]  -1.694363e-01 7.745400e-02 -3.280846e-01 -2.194464e-01
##   beta[5]  -1.692432e+00 1.025099e-01 -1.886227e+00 -1.759988e+00
##   beta[6]   6.154849e-01 1.520933e-01  3.273816e-01  5.125318e-01
```

```
##   beta[7]  -3.111610e-01 6.223002e-02 -4.376868e-01 -3.497939e-01
##   beta[8]  -8.847601e-01 1.905965e-01 -1.246948e+00 -1.013041e+00
##   lp__     -1.067827e+04 2.314422e+00 -1.068379e+04 -1.067951e+04
##           stats
## parameter           50%           75%         97.5%
##   mu        1.459436e+02  1.506148e+02  1.585888e+02
##   theta[1]  3.393186e+01  3.853901e+01  4.636705e+01
##   theta[2] -1.435124e+01 -1.402707e+01 -1.337703e+01
##   theta[3] -2.749123e+00 -2.347948e+00 -1.576171e+00
##   theta[4] -1.525743e+00 -1.267745e+00 -7.395505e-01
##   theta[5] -7.132133e+00 -6.902539e+00 -6.395100e+00
##   theta[6]  2.397312e+00  2.711539e+00  3.226225e+00
##   theta[7] -1.931334e+00 -1.665351e+00 -1.173596e+00
##   theta[8] -1.737962e+00 -1.493148e+00 -1.036337e+00
##   sigma     2.018273e+01  2.037236e+01  2.073987e+01
##   beta[1]   2.179012e-01  2.228646e-01  2.328375e-01
##   beta[2]  -1.498128e-04 -1.163149e-04 -4.775019e-05
##   beta[3]  -1.600537e-01 -9.933160e-02  8.801711e-03
##   beta[4]  -1.689676e-01 -1.186817e-01 -1.457338e-02
##   beta[5]  -1.694111e+00 -1.628771e+00 -1.495549e+00
##   beta[6]   6.139544e-01  7.271403e-01  8.969400e-01
##   beta[7]  -3.133711e-01 -2.694771e-01 -1.907784e-01
##   beta[8]  -8.789808e-01 -7.551649e-01 -5.241312e-01
##   lp__     -1.067795e+04 -1.067665e+04 -1.067487e+04
##
## , , chains = chain:4
##
##           stats
## parameter          mean           sd          2.5%           25%
##   mu        1.457538e+02 6.437033e+00  1.331462e+02  1.415558e+02
##   theta[1]  3.417961e+01 6.413756e+00  2.105399e+01  2.982337e+01
##   theta[2] -1.433448e+01 4.884537e-01 -1.529686e+01 -1.465567e+01
##   theta[3] -2.738882e+00 5.998439e-01 -3.905985e+00 -3.150834e+00
##   theta[4] -1.546645e+00 4.123183e-01 -2.366081e+00 -1.817100e+00
##   theta[5] -7.134867e+00 3.518128e-01 -7.825075e+00 -7.366779e+00
##   theta[6]  2.427905e+00 4.330915e-01  1.524542e+00  2.138077e+00
##   theta[7] -1.887828e+00 3.827133e-01 -2.698310e+00 -2.139550e+00
##   theta[8] -1.736209e+00 3.846509e-01 -2.497734e+00 -1.983896e+00
##   sigma     2.019182e+01 2.675631e-01  1.966673e+01  2.000889e+01
##   beta[1]   2.179381e-01 7.313407e-03  2.044082e-01  2.129489e-01
##   beta[2]  -1.502779e-04 5.159145e-05 -2.509058e-04 -1.861817e-04
##   beta[3]  -1.547151e-01 7.853618e-02 -3.068457e-01 -2.076791e-01
##   beta[4]  -1.726360e-01 7.707028e-02 -3.239791e-01 -2.198403e-01
##   beta[5]  -1.696422e+00 1.000021e-01 -1.897844e+00 -1.759763e+00
##   beta[6]   6.274720e-01 1.509654e-01  3.075165e-01  5.247432e-01
##   beta[7]  -3.060780e-01 6.278685e-02 -4.392732e-01 -3.472156e-01
##   beta[8]  -8.780938e-01 1.945386e-01 -1.263238e+00 -1.003362e+00
##   lp__     -1.067820e+04 2.300122e+00 -1.068366e+04 -1.067950e+04
##           stats
## parameter           50%           75%         97.5%
##   mu        1.455449e+02  1.500306e+02  1.588176e+02
##   theta[1]  3.447042e+01  3.841020e+01  4.670245e+01
##   theta[2] -1.430766e+01 -1.402204e+01 -1.337581e+01
##   theta[3] -2.722406e+00 -2.339671e+00 -1.611913e+00
```

```
##   theta[4] -1.544869e+00 -1.296319e+00 -7.429155e-01
##   theta[5] -7.131330e+00 -6.904879e+00 -6.437549e+00
##   theta[6]  2.440686e+00  2.725729e+00  3.247224e+00
##   theta[7] -1.885868e+00 -1.638305e+00 -1.113540e+00
##   theta[8] -1.730788e+00 -1.482213e+00 -9.408122e-01
##   sigma     2.019214e+01  2.036732e+01  2.072623e+01
##   beta[1]   2.178383e-01  2.230110e-01  2.315675e-01
##   beta[2]  -1.502026e-04 -1.175377e-04 -4.980431e-05
##   beta[3]  -1.562505e-01 -1.014195e-01  6.045471e-04
##   beta[4]  -1.719550e-01 -1.224663e-01 -1.630214e-02
##   beta[5]  -1.694963e+00 -1.632764e+00 -1.496956e+00
##   beta[6]   6.319572e-01  7.319283e-01  9.071178e-01
##   beta[7]  -3.064184e-01 -2.654276e-01 -1.782829e-01
##   beta[8]  -8.753520e-01 -7.496344e-01 -4.758192e-01
##   lp__     -1.067793e+04 -1.067658e+04 -1.067469e+04
```

Examine the regression coefficients

```r
qrOuts <- as.data.frame(qr_fit)
names(qrOuts)[11:18]<-names(mycancer)[2:9]
sapply(qrOuts[,c(1,11:18,10)], function(x) quantile(x,c(0.025,0.5,0.975)))
```

```
##             mu incidenceRate     medIncome MedianAgeFemale PercentMarried
## 2.5%  133.1551     0.2042437 -2.538467e-04    -0.324141191    -0.32309708
## 50%   145.7072     0.2179790 -1.523159e-04    -0.156823128    -0.16890622
## 97.5% 158.6629     0.2318999 -4.849081e-05     0.002700654    -0.01803268
##       PctBachDeg25_Over PctUnemployed16_Over PctPrivateCoverage  BirthRate
## 2.5%          -1.894550            0.3198445         -0.4284611 -1.2568602
## 50%           -1.695272            0.6234254         -0.3085248 -0.8753520
## 97.5%         -1.496231            0.9164528         -0.1869526 -0.5074777
##          sigma
## 2.5%  19.68751
## 50%   20.19012
## 97.5% 20.71242
```

Bayesean regression using brms

brms is a Bayesean regression library that uses Stan on the backend.

```r
library(brms)
```

```
## Loading required package: Rcpp
```

```
## Registered S3 method overwritten by 'xts':
##   method     from
##   as.zoo.xts zoo
```

```
## Loading 'brms' package (version 2.11.1). Useful instructions
## can be found by typing help('brms'). A more detailed introduction
## to the package is available through vignette('brms_overview').
```

```r
brm_fit <- brm(deathrate ~ ., data=mycancer, file="brm_cancer_regression")
```

Look at the results

```r
summary(brm_fit)
```

```
##  Family: gaussian
```

```
##   Links: mu = identity; sigma = identity
## Formula: deathrate ~ incidenceRate + medIncome + MedianAgeFemale + PercentMarried + PctBachDeg25_Over
##    Data: mycancer (Number of observations: 3047)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Population-Level Effects:
##                     Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## Intercept             145.84      6.36   133.54   158.27 1.00     2933
## incidenceRate           0.22      0.01     0.20     0.23 1.00     3781
## medIncome              -0.00      0.00    -0.00    -0.00 1.00     3590
## MedianAgeFemale        -0.16      0.08    -0.31     0.01 1.00     2863
## PercentMarried         -0.17      0.08    -0.33    -0.02 1.00     2026
## PctBachDeg25_Over      -1.69      0.10    -1.90    -1.50 1.00     2771
## PctUnemployed16_Over    0.62      0.15     0.32     0.91 1.00     2220
## PctPrivateCoverage     -0.31      0.06    -0.42    -0.19 1.00     2552
## BirthRate              -0.87      0.19    -1.25    -0.49 1.00     2454
##                     Tail_ESS
## Intercept               2961
## incidenceRate           3170
## medIncome               3055
## MedianAgeFemale         2977
## PercentMarried          2671
## PctBachDeg25_Over       2825
## PctUnemployed16_Over    2847
## PctPrivateCoverage      2992
## BirthRate               2215
##
## Family Specific Parameters:
##       Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma    20.20      0.27    19.67    20.74 1.00     3271     2558
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```