

Applied Stats Lab 9

18/03/24

Lip cancer

Here is the lip cancer data that was used in the lecture.

- `aff.i` is proportion of male population working outside in each region
- `observe.i` is observed deaths in each region
- `expect.i` is expected deaths, based on region-specific age distribution and national-level age-specific mortality rates.

```
observe.i <- c(
  5,13,18,5,10,18,29,10,15,22,4,11,10,22,13,14,17,21,25,6,11,21,13,5,19,18,14,17,3,10,
  7,3,12,11,6,16,13,6,9,10,4,9,11,12,23,18,12,7,13,12,12,13,6,14,7,18,13,9,6,8,7,6,16,4,6,
  17,5,7,2,9,7,6,12,13,17,5,5,6,12,10,16,10,16,15,18,6,12,6,8,33,15,14,18,25,14,2,73,13,14,
  12,10,3,11,3,11,13,11,13,10,5,18,10,23,5,9,2,11,9,11,6,11,5,19,15,4,8,9,6,4,4,2,12,12,11,
  8,12,11,23,7,16,46,9,18,12,13,14,14,3,9,15,6,13,13,12,8,11,5,9,8,22,9,2,10,6,10,12,9,11,
  9,11,11,0,9,3,11,11,11,5,4,8,9,30,110)

expect.i <- c(
  6.17,8.44,7.23,5.62,4.18,29.35,11.79,12.35,7.28,9.40,3.77,3.41,8.70,9.57,8.18,4.35,
  4.91,10.66,16.99,2.94,3.07,5.50,6.47,4.85,9.85,6.95,5.74,5.70,2.22,3.46,4.40,4.05,5.74,
  16.99,6.19,5.56,11.69,4.69,6.25,10.84,8.40,13.19,9.25,16.98,8.39,2.86,9.70,12.12,12.94,
  10.34,5.09,3.29,17.19,5.42,11.39,8.33,4.97,7.14,6.74,17.01,5.80,4.84,12.00,4.50,4.39,1,
  6.42,5.26,4.59,11.86,4.05,5.48,13.13,8.72,2.87,2.13,4.48,5.85,6.67,6.11,5.78,12.31,10,
  2.52,6.22,14.29,5.71,37.93,7.81,9.86,11.61,18.52,12.28,5.41,61.96,8.55,12.07,4.29,19.4,
  12.90,4.76,5.56,11.11,4.76,10.48,13.13,12.94,14.61,9.26,6.94,16.82,33.49,20.91,5.32,6,
  12.94,16.07,8.87,7.79,14.60,5.10,24.42,17.78,4.04,7.84,9.89,8.45,5.06,4.49,6.25,9.16,1,
  9.57,5.83,9.21,9.64,9.09,12.94,17.42,10.29,7.14,92.50,14.29,15.61,6.00,8.55,15.22,18.4,
  18.37,13.16,7.69,14.61,15.85,12.77,7.41,14.86,6.94,5.66,9.88,102.16,7.63,5.13,7.58,8.0,
  18.75,12.33,5.88,64.64,8.62,12.09,11.11,14.10,10.48,7.00,10.23,6.82,15.71,9.65,8.59,8,
  12.31,8.91,50.10,288.00)

aff.i <- c(0.2415,0.2309,0.3999,0.2977,0.3264,0.3346,0.4150,0.4202,0.1023,0.1752,
```

0.2548,0.3248,0.2287,0.2520,0.2058,0.2785,0.2528,0.1847,0.3736,0.2411,
0.3700,0.2997,0.2883,0.2427,0.3782,0.1865,0.2633,0.2978,0.3541,0.4176,
0.2910,0.3431,0.1168,0.2195,0.2911,0.4297,0.2119,0.2698,0.0874,0.3204,
0.1839,0.1796,0.2471,0.2016,0.1560,0.3162,0.0732,0.1490,0.2283,0.1187,
0.3500,0.2915,0.1339,0.0995,0.2355,0.2392,0.0877,0.3571,0.1014,0.0363,
0.1665,0.1226,0.2186,0.1279,0.0842,0.0733,0.0377,0.2216,0.3062,0.0310,
0.0755,0.0583,0.2546,0.2933,0.1682,0.2518,0.1971,0.1473,0.2311,0.2471,
0.3063,0.1526,0.1487,0.3537,0.2753,0.0849,0.1013,0.1622,0.1267,0.2376,
0.0737,0.2755,0.0152,0.1415,0.1344,0.1058,0.0545,0.1047,0.1335,0.3134,
0.1326,0.1222,0.1992,0.0620,0.1313,0.0848,0.2687,0.1396,0.1234,0.0997,
0.0694,0.1022,0.0779,0.0253,0.1012,0.0999,0.0828,0.2950,0.0778,0.1388,
0.2449,0.0978,0.1144,0.1038,0.1613,0.1921,0.2714,0.1467,0.1783,0.1790,
0.1482,0.1383,0.0805,0.0619,0.1934,0.1315,0.1050,0.0702,0.1002,0.1445,
0.0353,0.0400,0.1385,0.0491,0.0520,0.0640,0.1017,0.0837,0.1462,0.0958,
0.0745,0.2942,0.2278,0.1347,0.0907,0.1238,0.1773,0.0623,0.0742,0.1003,
0.0590,0.0719,0.0652,0.1687,0.1199,0.1768,0.1638,0.1360,0.0832,0.2174,
0.1662,0.2023,0.1319,0.0526,0.0287,0.0405,0.1616,0.0730,0.1005,0.0743,
0.0577,0.0481,0.1002,0.0433,0.0838,0.1124,0.2265,0.0436,0.1402,0.0313,
0.0359,0.0696,0.0618,0.0932,0.0097)

Question 1

The `expected.i` variable represents the expected number of deaths in each region, based on the region-specific age distribution and national-level age-specific mortality rates. Hence it is calculated as average over all ages of the population per age multiplied by the age specific mortality rate. In other words it “predicts” how many deaths would occur if the region experienced the average mortality rate observed nationally, adjusted for the age distribution of the region’s population.

When a particular area has an expected deaths value of 16, it means that, given the age distribution of that area’s population and the national age-specific mortality rates, one would expect 16 deaths from lip cancer in that area under average national conditions. This expected value serves as a baseline to compare against the actual observed number of deaths (`observe.i`) in the region.

Question 2

```
library(here)
library(ggplot2)
library(tidyverse)
library(rstan)
library(fdrtool)
library(bayesplot)
library(loo)
library(tidybayes)
```

```
stan_data <- list(N = 195, aff_i = aff.i, expect_i = expect.i, observe_i = observe.i)
mod1 <- stan(data = stan_data,
             file = here("stan", "Lip_Model1.stan"),
             iter = 3000,
             seed = 3105)
```

Running /Library/Frameworks/R.framework/Resources/bin/R CMD SHLIB foo.c

using C compiler: 'Apple clang version 14.0.3 (clang-1403.0.22.14.1)'

using SDK: 'MacOSX13.3.sdk'

clang -arch arm64 -I"/Library/Frameworks/R.framework/Resources/include" -DNDEBUG -I"/Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/S

In file included from <built-in>:1:

In file included from /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/S

In file included from /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/R

In file included from /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/R

/Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/RcppEigen/include/Eigen,

namespace Eigen {

~

/Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/RcppEigen/include/Eigen,

namespace Eigen {

~

;

In file included from <built-in>:1:

In file included from /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/S

In file included from /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/R

/Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/RcppEigen/include/Eigen,

#include <complex>

~~~~~~

3 errors generated.

make: \*\*\* [foo.o] Error 1

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 1).

Chain 1:

Chain 1: Gradient evaluation took 2.6e-05 seconds

Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.26 seconds.

Chain 1: Adjust your expectations accordingly!

Chain 1:

Chain 1:

Chain 1: Iteration: 1 / 3000 [ 0%] (Warmup)

Chain 1: Iteration: 300 / 3000 [ 10%] (Warmup)

Chain 1: Iteration: 600 / 3000 [ 20%] (Warmup)

Chain 1: Iteration: 900 / 3000 [ 30%] (Warmup)

Chain 1: Iteration: 1200 / 3000 [ 40%] (Warmup)

Chain 1: Iteration: 1500 / 3000 [ 50%] (Warmup)

Chain 1: Iteration: 1501 / 3000 [ 50%] (Sampling)

Chain 1: Iteration: 1800 / 3000 [ 60%] (Sampling)

Chain 1: Iteration: 2100 / 3000 [ 70%] (Sampling)

Chain 1: Iteration: 2400 / 3000 [ 80%] (Sampling)

Chain 1: Iteration: 2700 / 3000 [ 90%] (Sampling)

Chain 1: Iteration: 3000 / 3000 [100%] (Sampling)

Chain 1:

Chain 1: Elapsed Time: 0.424 seconds (Warm-up)

Chain 1: 0.427 seconds (Sampling)

Chain 1: 0.851 seconds (Total)

Chain 1:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 2).

Chain 2:

Chain 2: Gradient evaluation took 8e-06 seconds

Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.08 seconds.

Chain 2: Adjust your expectations accordingly!

Chain 2:

Chain 2:

Chain 2: Iteration: 1 / 3000 [ 0%] (Warmup)

Chain 2: Iteration: 300 / 3000 [ 10%] (Warmup)

Chain 2: Iteration: 600 / 3000 [ 20%] (Warmup)

Chain 2: Iteration: 900 / 3000 [ 30%] (Warmup)

Chain 2: Iteration: 1200 / 3000 [ 40%] (Warmup)

Chain 2: Iteration: 1500 / 3000 [ 50%] (Warmup)

Chain 2: Iteration: 1501 / 3000 [ 50%] (Sampling)

Chain 2: Iteration: 1800 / 3000 [ 60%] (Sampling)

Chain 2: Iteration: 2100 / 3000 [ 70%] (Sampling)

Chain 2: Iteration: 2400 / 3000 [ 80%] (Sampling)

Chain 2: Iteration: 2700 / 3000 [ 90%] (Sampling)

Chain 2: Iteration: 3000 / 3000 [100%] (Sampling)  
Chain 2:  
Chain 2: Elapsed Time: 0.428 seconds (Warm-up)  
Chain 2: 0.428 seconds (Sampling)  
Chain 2: 0.856 seconds (Total)  
Chain 2:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 3).

Chain 3:  
Chain 3: Gradient evaluation took 7e-06 seconds  
Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.07 seconds.  
Chain 3: Adjust your expectations accordingly!  
Chain 3:  
Chain 3:  
Chain 3: Iteration: 1 / 3000 [ 0%] (Warmup)  
Chain 3: Iteration: 300 / 3000 [ 10%] (Warmup)  
Chain 3: Iteration: 600 / 3000 [ 20%] (Warmup)  
Chain 3: Iteration: 900 / 3000 [ 30%] (Warmup)  
Chain 3: Iteration: 1200 / 3000 [ 40%] (Warmup)  
Chain 3: Iteration: 1500 / 3000 [ 50%] (Warmup)  
Chain 3: Iteration: 1501 / 3000 [ 50%] (Sampling)  
Chain 3: Iteration: 1800 / 3000 [ 60%] (Sampling)  
Chain 3: Iteration: 2100 / 3000 [ 70%] (Sampling)  
Chain 3: Iteration: 2400 / 3000 [ 80%] (Sampling)  
Chain 3: Iteration: 2700 / 3000 [ 90%] (Sampling)  
Chain 3: Iteration: 3000 / 3000 [100%] (Sampling)  
Chain 3:  
Chain 3: Elapsed Time: 0.425 seconds (Warm-up)  
Chain 3: 0.421 seconds (Sampling)  
Chain 3: 0.846 seconds (Total)  
Chain 3:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 4).

Chain 4:  
Chain 4: Gradient evaluation took 6e-06 seconds  
Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.06 seconds.  
Chain 4: Adjust your expectations accordingly!  
Chain 4:  
Chain 4:  
Chain 4: Iteration: 1 / 3000 [ 0%] (Warmup)  
Chain 4: Iteration: 300 / 3000 [ 10%] (Warmup)  
Chain 4: Iteration: 600 / 3000 [ 20%] (Warmup)  
Chain 4: Iteration: 900 / 3000 [ 30%] (Warmup)



SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 1).

Chain 1:

Chain 1: Gradient evaluation took 4e-05 seconds

Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.4 seconds.

Chain 1: Adjust your expectations accordingly!

Chain 1:

Chain 1:

Chain 1: Iteration: 1 / 3000 [ 0%] (Warmup)

Chain 1: Iteration: 300 / 3000 [ 10%] (Warmup)

Chain 1: Iteration: 600 / 3000 [ 20%] (Warmup)

Chain 1: Iteration: 900 / 3000 [ 30%] (Warmup)

Chain 1: Iteration: 1200 / 3000 [ 40%] (Warmup)

Chain 1: Iteration: 1500 / 3000 [ 50%] (Warmup)

Chain 1: Iteration: 1501 / 3000 [ 50%] (Sampling)

Chain 1: Iteration: 1800 / 3000 [ 60%] (Sampling)

Chain 1: Iteration: 2100 / 3000 [ 70%] (Sampling)

Chain 1: Iteration: 2400 / 3000 [ 80%] (Sampling)

Chain 1: Iteration: 2700 / 3000 [ 90%] (Sampling)

Chain 1: Iteration: 3000 / 3000 [100%] (Sampling)

Chain 1:

Chain 1: Elapsed Time: 0.702 seconds (Warm-up)

Chain 1: 0.69 seconds (Sampling)

Chain 1: 1.392 seconds (Total)

Chain 1:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 2).

Chain 2:

Chain 2: Gradient evaluation took 1.3e-05 seconds

Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.13 seconds.

Chain 2: Adjust your expectations accordingly!

Chain 2:

Chain 2:

Chain 2: Iteration: 1 / 3000 [ 0%] (Warmup)

Chain 2: Iteration: 300 / 3000 [ 10%] (Warmup)

Chain 2: Iteration: 600 / 3000 [ 20%] (Warmup)

Chain 2: Iteration: 900 / 3000 [ 30%] (Warmup)

Chain 2: Iteration: 1200 / 3000 [ 40%] (Warmup)

Chain 2: Iteration: 1500 / 3000 [ 50%] (Warmup)

Chain 2: Iteration: 1501 / 3000 [ 50%] (Sampling)

Chain 2: Iteration: 1800 / 3000 [ 60%] (Sampling)

Chain 2: Iteration: 2100 / 3000 [ 70%] (Sampling)

Chain 2: Iteration: 2400 / 3000 [ 80%] (Sampling)

Chain 2: Iteration: 2700 / 3000 [ 90%] (Sampling)  
Chain 2: Iteration: 3000 / 3000 [100%] (Sampling)  
Chain 2:  
Chain 2: Elapsed Time: 0.704 seconds (Warm-up)  
Chain 2: 0.693 seconds (Sampling)  
Chain 2: 1.397 seconds (Total)  
Chain 2:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 3).

Chain 3:  
Chain 3: Gradient evaluation took 1.3e-05 seconds  
Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.13 seconds.  
Chain 3: Adjust your expectations accordingly!  
Chain 3:  
Chain 3:  
Chain 3: Iteration: 1 / 3000 [ 0%] (Warmup)  
Chain 3: Iteration: 300 / 3000 [ 10%] (Warmup)  
Chain 3: Iteration: 600 / 3000 [ 20%] (Warmup)  
Chain 3: Iteration: 900 / 3000 [ 30%] (Warmup)  
Chain 3: Iteration: 1200 / 3000 [ 40%] (Warmup)  
Chain 3: Iteration: 1500 / 3000 [ 50%] (Warmup)  
Chain 3: Iteration: 1501 / 3000 [ 50%] (Sampling)  
Chain 3: Iteration: 1800 / 3000 [ 60%] (Sampling)  
Chain 3: Iteration: 2100 / 3000 [ 70%] (Sampling)  
Chain 3: Iteration: 2400 / 3000 [ 80%] (Sampling)  
Chain 3: Iteration: 2700 / 3000 [ 90%] (Sampling)  
Chain 3: Iteration: 3000 / 3000 [100%] (Sampling)  
Chain 3:  
Chain 3: Elapsed Time: 0.721 seconds (Warm-up)  
Chain 3: 0.693 seconds (Sampling)  
Chain 3: 1.414 seconds (Total)  
Chain 3:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 4).

Chain 4:  
Chain 4: Gradient evaluation took 1.3e-05 seconds  
Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.13 seconds.  
Chain 4: Adjust your expectations accordingly!  
Chain 4:  
Chain 4:  
Chain 4: Iteration: 1 / 3000 [ 0%] (Warmup)  
Chain 4: Iteration: 300 / 3000 [ 10%] (Warmup)  
Chain 4: Iteration: 600 / 3000 [ 20%] (Warmup)





make: \*\*\* [foo.o] Error 1

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 1).

Chain 1:

Chain 1: Gradient evaluation took 3.5e-05 seconds

Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.35 seconds.

Chain 1: Adjust your expectations accordingly!

Chain 1:

Chain 1:

Chain 1: Iteration: 1 / 3000 [ 0%] (Warmup)

Chain 1: Iteration: 300 / 3000 [ 10%] (Warmup)

Chain 1: Iteration: 600 / 3000 [ 20%] (Warmup)

Chain 1: Iteration: 900 / 3000 [ 30%] (Warmup)

Chain 1: Iteration: 1200 / 3000 [ 40%] (Warmup)

Chain 1: Iteration: 1500 / 3000 [ 50%] (Warmup)

Chain 1: Iteration: 1501 / 3000 [ 50%] (Sampling)

Chain 1: Iteration: 1800 / 3000 [ 60%] (Sampling)

Chain 1: Iteration: 2100 / 3000 [ 70%] (Sampling)

Chain 1: Iteration: 2400 / 3000 [ 80%] (Sampling)

Chain 1: Iteration: 2700 / 3000 [ 90%] (Sampling)

Chain 1: Iteration: 3000 / 3000 [100%] (Sampling)

Chain 1:

Chain 1: Elapsed Time: 0.609 seconds (Warm-up)

Chain 1: 0.577 seconds (Sampling)

Chain 1: 1.186 seconds (Total)

Chain 1:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 2).

Chain 2:

Chain 2: Gradient evaluation took 9e-06 seconds

Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.09 seconds.

Chain 2: Adjust your expectations accordingly!

Chain 2:

Chain 2:

Chain 2: Iteration: 1 / 3000 [ 0%] (Warmup)

Chain 2: Iteration: 300 / 3000 [ 10%] (Warmup)

Chain 2: Iteration: 600 / 3000 [ 20%] (Warmup)

Chain 2: Iteration: 900 / 3000 [ 30%] (Warmup)

Chain 2: Iteration: 1200 / 3000 [ 40%] (Warmup)

Chain 2: Iteration: 1500 / 3000 [ 50%] (Warmup)

Chain 2: Iteration: 1501 / 3000 [ 50%] (Sampling)

Chain 2: Iteration: 1800 / 3000 [ 60%] (Sampling)

Chain 2: Iteration: 2100 / 3000 [ 70%] (Sampling)

Chain 2: Iteration: 2400 / 3000 [ 80%] (Sampling)  
Chain 2: Iteration: 2700 / 3000 [ 90%] (Sampling)  
Chain 2: Iteration: 3000 / 3000 [100%] (Sampling)  
Chain 2:  
Chain 2: Elapsed Time: 0.61 seconds (Warm-up)  
Chain 2: 0.585 seconds (Sampling)  
Chain 2: 1.195 seconds (Total)  
Chain 2:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 3).

Chain 3:  
Chain 3: Gradient evaluation took 9e-06 seconds  
Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.09 seconds.  
Chain 3: Adjust your expectations accordingly!  
Chain 3:  
Chain 3:  
Chain 3: Iteration: 1 / 3000 [ 0%] (Warmup)  
Chain 3: Iteration: 300 / 3000 [ 10%] (Warmup)  
Chain 3: Iteration: 600 / 3000 [ 20%] (Warmup)  
Chain 3: Iteration: 900 / 3000 [ 30%] (Warmup)  
Chain 3: Iteration: 1200 / 3000 [ 40%] (Warmup)  
Chain 3: Iteration: 1500 / 3000 [ 50%] (Warmup)  
Chain 3: Iteration: 1501 / 3000 [ 50%] (Sampling)  
Chain 3: Iteration: 1800 / 3000 [ 60%] (Sampling)  
Chain 3: Iteration: 2100 / 3000 [ 70%] (Sampling)  
Chain 3: Iteration: 2400 / 3000 [ 80%] (Sampling)  
Chain 3: Iteration: 2700 / 3000 [ 90%] (Sampling)  
Chain 3: Iteration: 3000 / 3000 [100%] (Sampling)  
Chain 3:  
Chain 3: Elapsed Time: 0.602 seconds (Warm-up)  
Chain 3: 0.579 seconds (Sampling)  
Chain 3: 1.181 seconds (Total)  
Chain 3:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 4).

Chain 4:  
Chain 4: Gradient evaluation took 8e-06 seconds  
Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.08 seconds.  
Chain 4: Adjust your expectations accordingly!  
Chain 4:  
Chain 4:  
Chain 4: Iteration: 1 / 3000 [ 0%] (Warmup)  
Chain 4: Iteration: 300 / 3000 [ 10%] (Warmup)

```

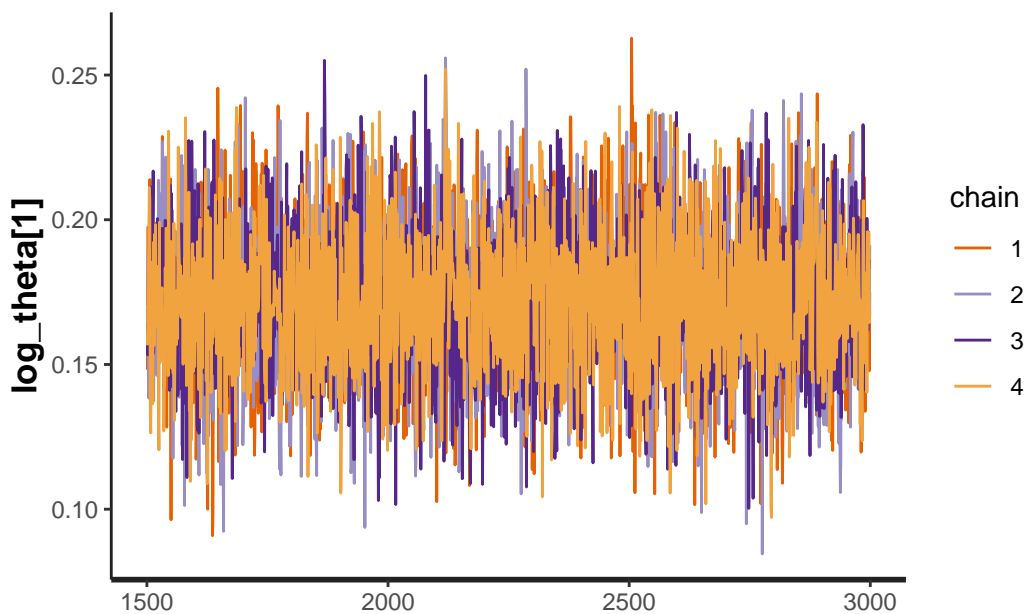
Chain 4: Iteration: 600 / 3000 [ 20%] (Warmup)
Chain 4: Iteration: 900 / 3000 [ 30%] (Warmup)
Chain 4: Iteration: 1200 / 3000 [ 40%] (Warmup)
Chain 4: Iteration: 1500 / 3000 [ 50%] (Warmup)
Chain 4: Iteration: 1501 / 3000 [ 50%] (Sampling)
Chain 4: Iteration: 1800 / 3000 [ 60%] (Sampling)
Chain 4: Iteration: 2100 / 3000 [ 70%] (Sampling)
Chain 4: Iteration: 2400 / 3000 [ 80%] (Sampling)
Chain 4: Iteration: 2700 / 3000 [ 90%] (Sampling)
Chain 4: Iteration: 3000 / 3000 [100%] (Sampling)
Chain 4:
Chain 4: Elapsed Time: 0.607 seconds (Warm-up)
Chain 4:           0.578 seconds (Sampling)
Chain 4:           1.185 seconds (Total)
Chain 4:

```

```

traceplot(mod1, pars = "log_theta[1]")

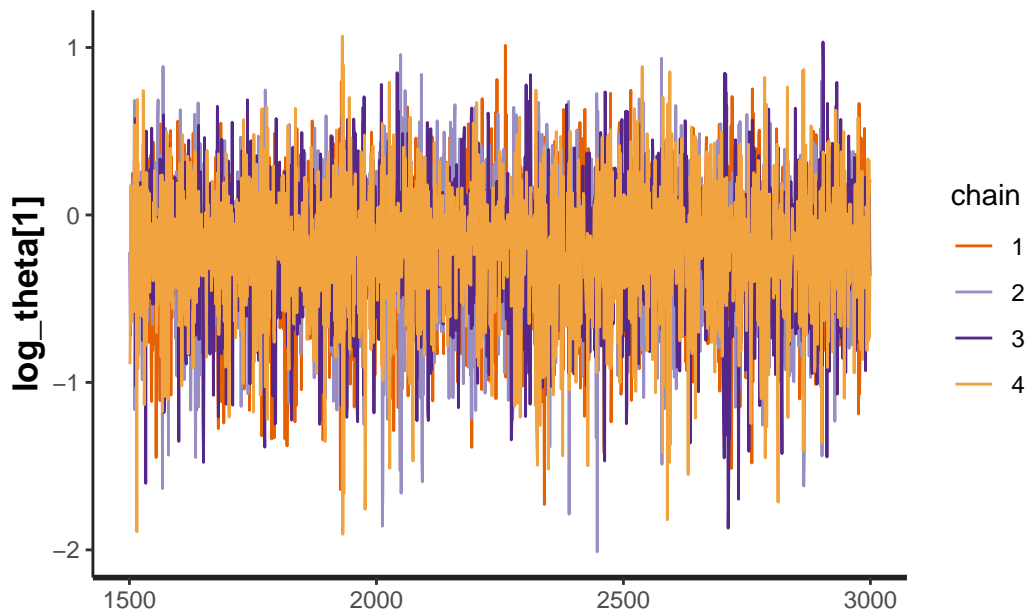
```



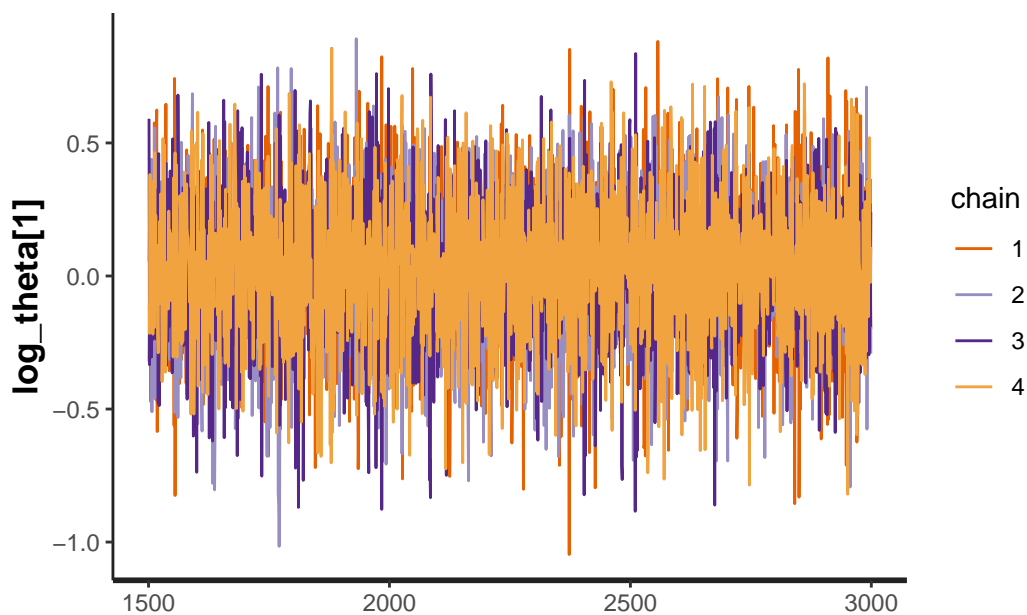
```

traceplot(mod2, pars = "log_theta[1]")

```



```
traceplot(mod3, pars = "log_theta[1]")
```



We fitted three different models in Rstan (Check the stan file for the Code). After checking the traceplots for various  $\theta_i$ 's we see that all chains have converged nicely for all three model's (here we showed only the traceplot of the  $\theta_1$ ).

### Question 3

First we extract the predicted theta values.

```
mod1_posterior_samples = extract(mod1)$log_theta
mod1_mean_theta = apply(mod1_posterior_samples, 2, mean)

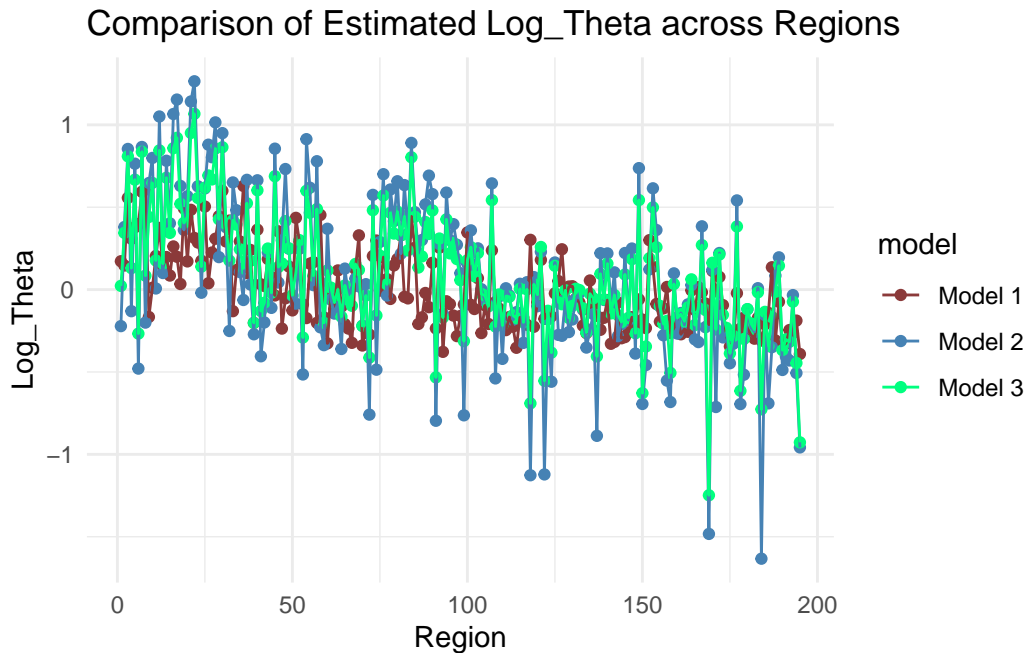
mod2_posterior_samples = extract(mod2)$log_theta
mod2_mean_theta = apply(mod2_posterior_samples, 2, mean)

mod3_posterior_samples = extract(mod3)$log_theta
mod3_mean_theta = apply(mod3_posterior_samples, 2, mean)

regions = 1:length(mod1_mean_theta)
data_model1 = data.frame(region = regions, mean_theta = mod1_mean_theta, model = 'Model 1')
data_model2 = data.frame(region = regions, mean_theta = mod2_mean_theta, model = 'Model 2')
data_model3 = data.frame(region = regions, mean_theta = mod3_mean_theta, model = 'Model 3')

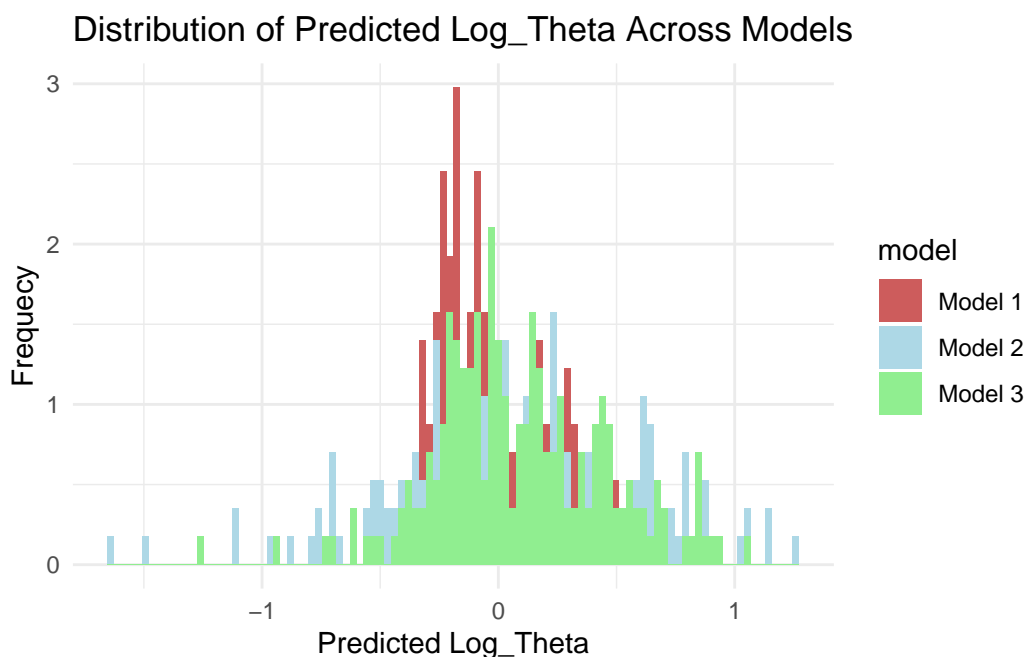
combined_data = rbind(data_model1, data_model2, data_model3)

ggplot(combined_data, aes(x = region, y = mean_theta, color = model)) +
  geom_point() +
  geom_line(aes(group = model)) +
  theme_minimal() +
  xlab("Region") +
  ylab("Log_Theta") +
  ggtitle("Comparison of Estimated Log_Theta across Regions") +
  scale_color_manual(values = c("Model 1" = "indianred4", "Model 2" = "steelblue", "Model 3" = "darkgreen"))
```



First we plot the differences in  $\theta_i$ 's across regions and across models. It seems that the  $\log\_theta$  values seem to be more frequently positive in the first 100 regions, while the other regions's predicted  $\log\_theta$  values seem to drop below 0 more at an increased frequency. This trend seems to hold true for all three models.

```
ggplot(combined_data, aes(x = mean_theta, fill = model)) +
  geom_histogram(aes(y = after_stat(density)), position = "identity", alpha = 1, bins = 10) +
  scale_fill_manual(values = c("Model 1" = "indianred", "Model 2" = "lightblue", "Model 3" = "lightgreen")) +
  labs(x = "Predicted Log_Theta", y = "Frequency", title = "Distribution of Predicted Log_Theta") +
  theme_minimal()
```



Secondly we plot a histogram to compare the frequency of the predicted `log_theta` values across the three models. It appears that the predicted values are more spread out for model 2 and 3. The distribution for model 1 seems a lot more dense.

#### Question 4

```
log_lik1 <- extract(mod1)$log_lik
log_lik2 <- extract(mod2)$log_lik
log_lik3 <- extract(mod3)$log_lik

loo1 <- loo(log_lik1, save_psis = TRUE)
loo2 <- loo(log_lik2, save_psis = TRUE)
loo3 <- loo(log_lik3, save_psis = TRUE)

loo_compare(loo1, loo2, loo3)
```

|        | elpd_diff | se_diff |
|--------|-----------|---------|
| model1 | 0.0       | 0.0     |
| model3 | -484299.1 | 27150.6 |
| model2 | -676184.8 | 37783.5 |



According to these results, Model 1 is substantially better in terms of predictive accuracy than Models 3 and 2. The large negative differences in ELPD for Models 3 and 2 suggest they have much lower predictive performance compared to Model 1. The standard errors are relatively small in comparison to the differences in ELPD, reinforcing the conclusion that Model 1 is the preferred model based on LOO-CV.