Examine perfect fit problem of hierarchical Poisson regression

Miao Cai miao.cai@slu.edu

2019-07-25

Contents

1 Statistical model

$$
\begin{aligned}
Y_i &\sim \text{Poisson}(\lambda_i) \\
\log \lambda_i &= \beta_{0,i} + \beta_1 x_1 + \beta_2 x_2 \\
\beta_{0,i} &\sim N(\mu_0, \sigma_0^2)
\end{aligned}
$$

2 Simulate data

```r
set.seed(123)
N = 50 # total obs
K = N # groups
mu0 = 3
sigma0 = 1
b0 = rnorm(N, mu0, sigma0)

x1 = rnorm(N, 1, 2)
x2 = rgamma(N, 10, 10)

b1 = 1; b2 = 0.2

lambda = exp(b0 + b1*x1 + b2*x2)

Y = rpois(n = N, lambda = lambda)
```

3 Stan code

```
data {
  int N;
  int K;
  int Y[N];
  int id[N];
  real x1[N];
  real x2[N];
}
parameters{
  real mu0;
  real sigma0;
  vector[K] b0;
  real b1;
  real b2;
}
model {
  vector[N] tmp;
  for(i in 1:N){
    tmp[i] = b0[id[i]] + b1*x1[i] + b2*x2[i];
  }

  Y ~ poisson_log(tmp);
  b0 ~ normal(mu0, sigma0);
  mu0 ~ normal(0, 5);
  sigma0 ~ gamma(1, 1);
  b1 ~ normal(0, 10);
  b2 ~ normal(0, 10);
}
```

4 Fit the model

```
## 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
```

The warning message above highly suggests that there are identifiability issue here in this model since stan cannot sample much effective samples from this model. The column `n_eff` stands for the number of effective sample size (ESS). ESS in this model are very very low compared with the number of iterations I used, typically around 30 out of 1000, which means only around 3% of the samples were effective.

This warning message is identical with the following two posts demonstrating the identifiability issues using `Stan`:

- Identifying non-identifiability by Martin Modrak
- Richard McElreath: Statistical Rethinking – 8.4.4. Non-identifiable parameters, page 261

```
## Inference for Stan model: 08Perfect_fit_Poisson.
## 1 chains, each with iter=1000; warmup=500; thin=1;
## post-warmup draws per chain=500, total post-warmup draws=500.
##
##             mean se_mean   sd      10%      50%      90% n_eff Rhat
## mu0         2.63    0.11 0.54     1.95     2.64     3.33    23 1.04
## sigma0      0.99    0.01 0.11     0.85     0.98     1.13   261 1.00
## b0[1]       2.00    0.09 0.44     1.42     2.01     2.58    22 1.03
## b0[2]       2.58    0.10 0.50     1.94     2.59     3.25    24 1.03
## b0[3]       4.06    0.14 0.66     3.23     4.08     4.93    23 1.03
## b0[4]       2.76    0.12 0.57     2.00     2.75     3.52    25 1.05
## b0[5]       2.80    0.10 0.47     2.20     2.79     3.42    24 1.03
## b0[6]       4.33    0.15 0.71     3.39     4.30     5.27    23 1.05
## b0[7]       2.19    0.11 0.72     1.26     2.16     3.06    46 1.01
## b0[8]       1.41    0.11 0.53     0.74     1.41     2.09    25 1.03
## b0[9]       1.99    0.09 0.44     1.44     1.99     2.58    25 1.04
## b0[10]      2.35    0.12 0.58     1.60     2.35     3.12    23 1.04
## b0[11]      3.86    0.10 0.48     3.22     3.84     4.49    22 1.04
## b0[12]      3.15    0.11 0.54     2.48     3.15     3.84    25 1.02
## b0[13]      2.83    0.11 0.53     2.19     2.82     3.53    25 1.03
## b0[14]      2.79    0.09 0.52     2.14     2.81     3.49    37 1.01
```

```
## b0[15]       2.13     0.07 0.57      1.37      2.15      2.83    59 1.01
## b0[16]       4.28     0.14 0.65      3.45      4.28      5.13    22 1.03
## b0[17]       2.97     0.14 0.65      2.12      2.94      3.81    22 1.04
## b0[18]       0.63     0.12 0.62     -0.19      0.63      1.46    29 1.03
## b0[19]       3.27     0.13 0.62      2.47      3.22      4.09    23 1.04
## b0[20]       2.10     0.16 0.80      1.02      2.09      3.15    24 1.05
## b0[21]       1.65     0.15 0.79      0.65      1.65      2.67    28 1.02
## b0[22]       2.01     0.09 0.92      0.82      2.04      3.21   113 1.02
## b0[23]       1.71     0.10 0.50      1.02      1.68      2.36    24 1.05
## b0[24]       2.17     0.09 0.52      1.50      2.17      2.78    36 1.02
## b0[25]       1.81     0.07 0.53      1.09      1.84      2.46    63 1.00
## b0[26]       0.77     0.16 0.75     -0.23      0.76      1.79    22 1.04
## b0[27]       3.36     0.13 0.65      2.54      3.34      4.21    24 1.03
## b0[28]       2.70     0.13 0.72      1.77      2.70      3.63    28 1.01
## b0[29]       1.50     0.11 0.56      0.80      1.50      2.24    25 1.03
## b0[30]       3.85     0.09 0.43      3.31      3.84      4.44    23 1.03
## b0[31]       2.81     0.14 0.66      1.98      2.80      3.72    23 1.03
## b0[32]       2.55     0.07 0.36      2.07      2.56      3.01    24 1.05
## b0[33]       3.26     0.11 0.57      2.55      3.24      3.99    25 1.02
## b0[34]       3.47     0.11 0.53      2.76      3.45      4.17    22 1.04
## b0[35]       3.55     0.12 0.58      2.81      3.53      4.31    24 1.03
## b0[36]       3.43     0.10 0.48      2.80      3.41      4.07    22 1.04
## b0[37]       3.22     0.12 0.57      2.46      3.20      3.97    23 1.05
## b0[38]       2.60     0.07 0.35      2.13      2.61      3.07    23 1.04
## b0[39]       2.33     0.06 0.34      1.89      2.35      2.76    30 1.04
## b0[40]       2.18     0.15 0.70      1.26      2.14      3.10    23 1.04
## b0[41]       1.59     0.19 0.91      0.43      1.55      2.82    22 1.04
## b0[42]       2.56     0.08 0.39      2.05      2.55      3.06    24 1.05
## b0[43]       1.65     0.05 0.32      1.23      1.65      2.05    37 1.03
## b0[44]       4.76     0.10 0.50      4.16      4.74      5.42    25 1.02
## b0[45]       3.98     0.10 0.53      3.24      3.97      4.66    26 1.05
## b0[46]       1.40     0.10 0.58      0.64      1.39      2.12    37 1.04
## b0[47]       2.17     0.16 0.82      1.06      2.16      3.25    25 1.05
## b0[48]       2.28     0.12 0.62      1.45      2.26      3.08    25 1.05
## b0[49]       3.42     0.08 0.39      2.95      3.42      3.91    25 1.03
## b0[50]       2.44     0.12 0.72      1.52      2.44      3.36    33 1.02
## b1           1.00     0.01 0.09      0.88      0.99      1.12    80 1.03
## b2           0.59     0.10 0.47     -0.02      0.60      1.17    24 1.02
## lp__    177132.05     0.40 5.24 177125.07 177132.50 177138.20   174 1.00
```

```
##
## Samples were drawn using NUTS(diag_e) at Thu Jul 25 10:24:02 2019.
## 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).
```

5 Compare true and estimated parameters

id	β_0	$\hat{\beta}_0$	Δ
1	2.440	1.997	0.442
2	2.770	2.582	0.187
3	4.559	4.065	0.494
4	3.071	2.765	0.306
5	3.129	2.798	0.331
6	4.715	4.328	0.387
7	3.461	2.190	1.271
8	1.735	1.407	0.328
9	2.313	1.986	0.327
10	2.554	2.351	0.203
11	4.224	3.857	0.367
12	3.360	3.149	0.211
13	3.401	2.827	0.574
14	3.111	2.793	0.318
15	2.444	2.126	0.318
16	4.787	4.276	0.511
17	3.498	2.969	0.529
18	1.033	0.631	0.402
19	3.701	3.269	0.432
20	2.527	2.101	0.426
21	1.932	1.654	0.278
22	2.782	2.012	0.770
23	1.974	1.706	0.268
24	2.271	2.169	0.102
25	2.375	1.807	0.568
26	1.313	0.767	0.546
27	3.838	3.363	0.475
28	3.153	2.699	0.454
29	1.862	1.500	0.362
30	4.254	3.850	0.404
31	3.426	2.809	0.618
32	2.705	2.550	0.155
33	3.895	3.264	0.632
34	3.878	3.467	0.411
35	3.822	3.545	0.277

36	3.689	3.430	0.258
37	3.554	3.220	0.334
38	2.938	2.605	0.333
39	2.694	2.335	0.359
40	2.620	2.178	0.442
41	2.305	1.595	0.711
42	2.792	2.555	0.237
43	1.735	1.647	0.087
44	5.169	4.757	0.412
45	4.208	3.985	0.223
46	1.877	1.398	0.479
47	2.597	2.174	0.423
48	2.533	2.278	0.256
49	3.780	3.424	0.356
50	2.917	2.444	0.473

There are noticeable difference (Δ) between the true (β_0) and estimated ($\hat{\beta}_0$) random intercepts.