

# Assignment 7

Sam Olson

## The Assignment

Your assignment is to conduct a Bayesian analysis of relative risk for these data. You will do so using both the Multinomial probability model and the Poisson probability model.

### 1.

Using prior distributions and prior parameter values as given in Section 4 (Bayesian Analysis), derive the joint posterior distributions for  $\boldsymbol{\theta} = (\theta_{11}, \theta_{10}, \theta_{01}, \theta_{00})$  in the multinomial model, and  $\boldsymbol{\lambda} = (\lambda_{11}, \lambda_{10}, \lambda_{01}, \lambda_{00})$  in the Poisson model.

### Answer

**Multinomial–Dirichlet** Multinomial likelihood (with  $N = \sum_{ij} y_{ij}$ ,  $\sum_{ij} \theta_{ij} = 1$ ):

$$L(\boldsymbol{\theta} \mid \mathbf{y}) = \Pr(\mathbf{Y} = \mathbf{y} \mid \boldsymbol{\theta}) = \frac{N!}{\prod_{ij} y_{ij}!} \prod_{ij} \theta_{ij}^{y_{ij}}$$

For the priors:

Dirichlet prior on cell probabilities  $\boldsymbol{\theta}$  (with support on a simplex, i.e.,  $\sum_{i,j} \theta_{ij} = 1$ ):

$$\boldsymbol{\theta} \sim \text{Dirichlet}(\gamma_{11}, \gamma_{10}, \gamma_{01}, \gamma_{00}), \quad p(\boldsymbol{\theta}) = \frac{1}{B(\boldsymbol{\gamma})} \prod_{ij} \theta_{ij}^{\gamma_{ij}-1}$$

And

$$B(\boldsymbol{\gamma}) = \frac{\prod_{ij} \Gamma(\gamma_{ij})}{\Gamma(\sum_{ij} \gamma_{ij})},$$

As given,  $(\gamma_{11}, \gamma_{10}, \gamma_{01}, \gamma_{00}) = (0.10, 0.75, 0.10, 0.75)$ .

Taken together, we have the (proportional) posterior derivations

Multinomial–Dirichlet:

$$p(\boldsymbol{\theta} \mid \mathbf{y}) \propto L(\boldsymbol{\theta} \mid \mathbf{y}), p(\boldsymbol{\theta})[3pt] \propto \left( \prod_{ij} \theta_{ij}^{y_{ij}} \right) \left( \prod_{ij} \theta_{ij}^{\gamma_{ij}-1} \right) [3pt] \propto \prod_{ij} \theta_{ij}^{y_{ij} + \gamma_{ij} - 1}, \quad \text{for } \sum_{ij} \theta_{ij} = 1$$

After collecting exponents and normalizing over the simplex (by the Dirichlet Beta function), conjugacy gives

$$\boldsymbol{\theta} \mid \mathbf{y} \sim \text{Dirichlet}(y_{11} + \gamma_{11}, y_{10} + \gamma_{10}, y_{01} + \gamma_{01}, y_{00} + \gamma_{00})$$

**Poisson–Gamma** (Independent) Poisson likelihoods:

$$Y_{ij} \mid \lambda_{ij} \stackrel{\text{ind}}{\sim} \text{Poisson}(\lambda_{ij}) \quad \Rightarrow \quad L(\boldsymbol{\lambda} \mid \mathbf{y}) = \prod_{ij} \frac{e^{-\lambda_{ij}} \lambda_{ij}^{y_{ij}}}{y_{ij}!}$$

For the priors:

Gamma priors (assuming shape–rate parametrization) on Poisson means:

$$\lambda_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij}), \quad p(\lambda_{ij}) = \frac{\beta_{ij}^{\alpha_{ij}}}{\Gamma(\alpha_{ij})} \lambda_{ij}^{\alpha_{ij}-1} e^{-\beta_{ij} \lambda_{ij}}$$

Given  $(\alpha, \beta)_{11,10,01,00} = (0.5, 5), (3, 4), (0.5, 5), (3, 4)$ .

Taken together, we have the (proportional) posterior derivations

$$p(\lambda_{ij} \mid y_{ij}) \propto e^{-\lambda_{ij}} \lambda_{ij}^{y_{ij}} \times \lambda_{ij}^{\alpha_{ij}-1} e^{-\beta_{ij} \lambda_{ij}} [3pt] \propto \lambda_{ij}^{\alpha_{ij}+y_{ij}-1} \exp[-(\beta_{ij} + 1)\lambda_{ij}]$$

Recognizing this kernel as a Gamma density, conjugacy gives

$$\lambda_{ij} \mid y_{ij} \sim \text{Gamma}(\alpha_{ij} + y_{ij}, \beta_{ij} + 1)$$

## 2.

Using the data of Table 2 give posterior expected values for  $\theta$  and  $\lambda$ , and the cell probabilities that would correspond to these means.

### Answer

We have observed counts  $y_{11} = 15$ ,  $y_{10} = 50$ ,  $y_{01} = 3$ ,  $y_{00} = 84$

So our total  $N$  is given by:  $N = 152$

Let  $G = \sum \gamma_{ij} = 1.70$ ,

Then:  $N + G = 153.70$

**Multinomial–Dirichlet** Using properties of known distributions, for Dirichlet:

$$E[\theta_{ij}] = \gamma_{ij}/G$$

And

$$\text{Var}(\theta_{ij}) = \frac{\gamma_{ij}(G - \gamma_{ij})}{G^2(G + 1)}.$$

So:

$$E[\theta_{ij} | y] = \frac{y_{ij} + \gamma_{ij}}{N + G}$$

Evaluating,

$$\begin{aligned} E[\theta_{11} | y] &= \frac{15 + 0.10}{153.70} = 0.0982, \\ E[\theta_{10} | y] &= \frac{50 + 0.75}{153.70} = 0.3302, \\ E[\theta_{01} | y] &= \frac{3 + 0.10}{153.70} = 0.0202, \\ E[\theta_{00} | y] &= \frac{84 + 0.75}{153.70} = 0.5514 \end{aligned}$$

Note: We don't need to normalize the above, as we have  $\sum_{ij} \theta_{ij} = 0.098 + 0.330 + 0.020 + 0.551 = 1$  (well, ok: 0.999 but that's just due to rounding!)

**Poisson–Gamma** Again, for a known Gamma distribution (as parametrized):

$$E[\lambda_{ij}] = \alpha_{ij}/\beta_{ij},$$

And

$$\text{Var}(\lambda_{ij}) = \alpha_{ij}/\beta_{ij}^2.$$

So, we have:

$$E[\lambda_{ij} | y] = \frac{\alpha_{ij} + y_{ij}}{\beta_{ij} + 1}$$

So

$$E[\lambda_{11}] = \frac{15.5}{6} = 2.5833$$

$$E[\lambda_{10}] = \frac{53}{5} = 10.6$$

$$E[\lambda_{01}] = \frac{3.5}{6} = 0.5833$$

$$E[\lambda_{00}] = \frac{87}{5} = 17.4$$

To compare on the probability scale, we need to normalize:

$$\tilde{\theta}_{ij} = \frac{E[\lambda_{ij} \mid y]}{\sum_{kl} E[\lambda_{kl} \mid y]} = (0.0829, 0.3401, 0.0187, 0.5583)$$

### 3.

Evaluate relative risk  $RR$  at the expected values of the posterior distributions you derived in exercise 1 and compare to the observed or sample-based  $RR$  based only on the observed cell counts. Briefly explain why this is not the posterior expected value of  $RR$  (although we hope it is not far off).

#### Answer

The relative risk is generally defined as (Equation 2):

$$RR = \frac{\Pr(D | R)}{\Pr(D | R^c)} = \frac{\theta_{11}(\theta_{01} + \theta_{00})}{\theta_{01}(\theta_{11} + \theta_{10})}$$

For our sample estimate, we have:

$$\widehat{RR} = \frac{y_{11}(y_{01} + y_{00})}{y_{01}(y_{11} + y_{10})} = \frac{15(3 + 84)}{3(15 + 50)} \approx 6.69$$

Using the posterior “plug-in” values, we have:

Multinomial–Dirichlet:

$$RR(E[\theta]) = \frac{0.0982(0.0202 + 0.5514)}{0.0202(0.0982 + 0.3302)} \approx 6.50$$

And

Poisson–Gamma:

$$RR(\tilde{\theta}) = \frac{0.0829(0.0187 + 0.5583)}{0.0187(0.0829 + 0.3401)} \approx 6.04$$

**Thoughts** All three values are close, but not exactly the same. This happens because the “plug-in” value  $RR(E[\theta | y])$  is not the same as the posterior mean  $E[RR | y]$ . This difference primarily stems from the nonlinearity of the function  $RR$ , which means Jensen’s inequality applies: Expectations do not generally “pass through” nonlinear transformations such that they ensure equality.

To be specific, the function  $g$  for relative risk is given by:

$$g(\theta) = \frac{\theta_{11}(\theta_{01} + \theta_{00})}{\theta_{01}(\theta_{11} + \theta_{10})}$$

and because  $g$  is nonlinear in  $\theta$ ,

$$E[g(\theta) | y] \neq g(E[\theta | y])$$

Though this might be repetitive of the above, taking an expectation and then applying a nonlinear function (like this ratio for  $RR$ ) does not yield the same result as applying the function first and then averaging. The difference reflects the variability in the posterior distribution of  $\theta|y$  (and also I believe the form of the function  $g$ ). Finally, the observed sample estimate  $\widehat{RR}$  also differs slightly because it is based on just the counts, while the Bayesian “plug-in” values incorporate both the data likelihood and prior information.

#### 4.

Find the posterior distributions of RR under both the multinomial and Poisson models. Give summary and 95% credible intervals. Produce some type of a graphical display of the posterior distributions. How do the results compare for the multinomial and Poisson models? How do the results compare to the initial sample version of RR in expression (2)? What would you conclude about the relation between time since vaccination and the chances of getting chickenpox in this population of school children? For example, what is your posterior probability that the risk factor of having more than 5 years since vaccination is positively related to the chance of contracting chickenpox in these students?

#### Answer

For the Multinomial-Dirichlet, we have:

$$\boldsymbol{\theta} \mid \mathbf{y} \sim \text{Dirichlet}(y_{11} + \gamma_{11}, y_{10} + \gamma_{10}, y_{01} + \gamma_{01}, y_{00} + \gamma_{00})$$

And for the Poisson-Gamma, we have:

$$\lambda_{ij} \mid \mathbf{y} \sim \text{Gamma}(\alpha_{ij} + y_{ij}, \beta_{ij} + 1)$$

Which requires normalization via:

$$p_{ij} = \frac{\lambda_{ij}}{\sum_{kl} \lambda_{kl}}$$

The form of RR is then given by:

For the Multinomial-Dirichlet:

$$RR(\boldsymbol{\theta}) = \frac{\theta_{11}(\theta_{01} + \theta_{00})}{\theta_{01}(\theta_{11} + \theta_{10})}$$

And for the Poisson-Gamma:

$$RR(\mathbf{p}) = \frac{p_{11}(p_{01} + p_{00})}{p_{01}(p_{11} + p_{10})}$$

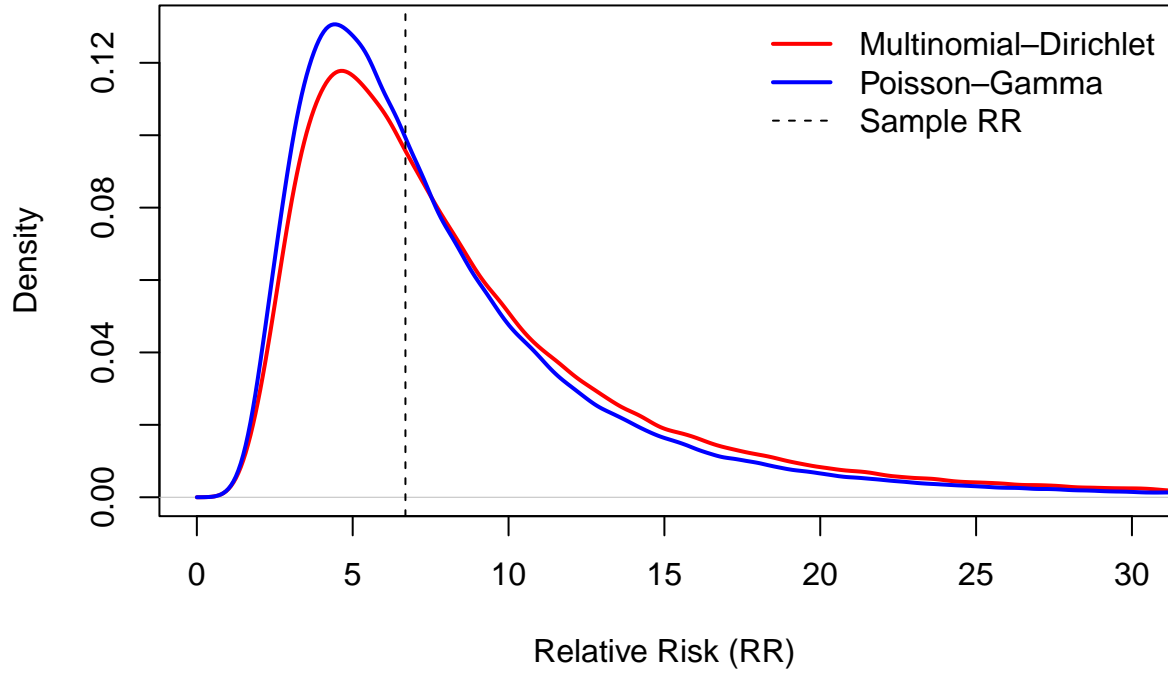
The procedure then, using the above, is to generate posterior draws for the parameters using `rdirichlet` (for the Multinomial-Dirichlet) and `rgamma` (for the Poisson-Gamma model). For each draw, an RR is computed using the corresponding sampled probabilities (and for each model). The resulting collection of RR values then is used as our “empirical” posterior distribution of RR. It is the distribution of these RR values then that are summarized; particularly, the credible interval is determined by the empirical 2.5% and 97.5% quantiles of the (simulated) RR distribution.

Special thanks to Kaiser for noting that we don’t need to do any special (rejection) sampling for this problem.

Table 1: Posterior summaries for RR under two Bayesian models

Model	Posterior Summary						
	Mean	Median	SD	2.5%	97.5%	$\Pr(RR > 1)$	$\Pr(RR > 5)$
<b>Multinomial-Dirichlet</b>	9.47	7.08	9.15	2.42	30.91	1	0.72
<b>Poisson-Gamma</b>	8.38	6.50	7.14	2.29	25.67	1	0.68

### Posterior Distributions of RR



**Thoughts** Both posteriors have high density above 1 (typically concentrated most around 5). And  $\Pr(RR > 1) \approx 1$  under both models (0.99996 for Dirichlet and 0.99990 for Gamma, just rounded up). The credible intervals are somewhat wide in the right-tail, but both contain and are “close” in center to the observed sample ( $\widehat{RR} \approx 6.69$ , vertical reference line in the density plot). Also, the “plug-in” summaries ( $RR(E[\theta]) \approx 6.50$ ) and ( $RR(\hat{\theta}) \approx 6.04$ ) show mild shrinkage relative to the MLE (observed sample RR). I think we can make the case that there is evidence (via posterior examination) that “Time Since Vaccination > 5 years” is associated with higher chickenpox risk for the population in the study (elementary school students).