Projects: Concepts in Bayesian Inference

April 5, 2019

DEADLINE: 3/6/2019

Dose-response model

The first project concerns determining the dose-response relationship of a possible toxic product. Diethylene Glycol Dimethyl Ether (DYME), also referred to as diglyme, bis(2-methoxyethyl) ether is a high-volume industrial chemical with diverse applications. It is used to make industrial solvents, cosmetics, protective coatings, solvents in chemical synthesis, and is used in manufacturing of textile dyes. Price *et al.* (1987) describe a study in which timed-pregnant CD-1 mice were dosed by gavage with DYME in distilled water. Dosing occurred during the period of major organogenesis and structural development of the foetuses (gestational age 6 through 15). Relating the dose of DYME to the incidence of malformations in foetuses gives the following results:

Dose d	Number of foetuses N	Number of malformations y
0	282	67
62.5	225	43
125	290	193
250	261	250
500	141	141

Questions 1-3 should be based on conjugate analyses, and using analytical results. Questions 4-10 should be based on MCMC.

1. What is the prevalence of malformations per dose group? Assume that in a similar experiment, 30 foetuses out of 224 did have malformations, but no historical information is available for exposed animals. Use a conjugate prior. Give posterior summary measures (per dose group). (Use analytical results)

2. Predict how many malformed foetuses would be expected in a future experiment with 25 animals per dose group. (Using analytical results)

3. What is the contour probability for $H_0 : \theta = 0.25$ (per dose group).

4. Assume that the likelihood of the experiment is specified by

$$\begin{aligned} y &\sim \text{binomial}(N, \pi) \\ \text{logit}(\pi) &= \alpha + \beta d. \end{aligned}$$

   Here β is the parameter of interest. Take vague priors for α and β and write this model with Bayesian software. Take 2 MCMC chains with different starting values, and check convergence with the appropriate techniques. (To improve convergence, standardise the covariate)

5. Sensitivity analysis is the practice of understanding the variation and uncertainty of the posterior inferences as a result of a different prior or model used in the analysis. For example, you might want to compare different priors, and compare the results of the separate analyses. Perform a sensitivity analysis by changing the prior distribution for the dose effect into a t-distribution with 4 degrees of freedom.

6. Summarise all results graphically and with the usual Bayesian posterior statistics. What do you conclude from these?

7. Plot the posterior dose-response relationship together with the observed probabilities of a malformation per dose.

8. Calculate the difference between prevalences for dose group i with the control group ($i = 2, 3, 4, 5$). At which dose-level is there an adverse effect (as compared to no exposure)?

9. What is the probability that the prevalence in dose group i ($i = 2, 3, 4, 5$) is larger than in the control group? Use the step-function to calculate this probabiltiy.

10. A safe level of exposure can be defined as a dose corresponding to a very small increase in excess risk of q, e.g. $q = 0.01$. This is called the Benchmark dose (BMD) d^* and can be obtained by solving the equation

$$r(d^*) = \frac{P(d^*) - P(0)}{1 - P(0)} = q$$

with $P(d)$ the probability of an adverse effect at dose level d. For a logistic regression with a linear dose model, the BMD is given by

$$\text{BMD} = \frac{\text{logit}(q^*) - \alpha}{\beta}$$

with $q^* = q(1 - P(0)) + P(0)$. Determine the posterior estimate of the safe level of exposure for DYME corresponding with an excess risk of $q = 0.01$.