Homework 5

ISyE 6420

Spring 2022

Use of unsolicited electronic and printed resources is allowed except the communication that violates Georgia Tech Academic Integrity Rules (e.g., direct communication about solutions with a third party, use of HW-solving sites, and similar). [1]

1. Blood Volume in Infants. The total blood volume of normal newborn infants was estimated by Schücking (1879) who took into account the addition of placental blood to the circulation of the newborn infant when clamping of the umbilical cord is delayed. Demarsh et al. (1942) further studied the importance of early and late clamping. For 16 babies in whom the cord was clamped early the total blood, as a percentage of weight, on the third day is listed below:

脐带夹紧

13.8	8.0	8.4	8.8	9.6	9.8	8.2	8.0
10.3	8.5	11.5	8.2	8.9	9.4	10.3	12.6

For 16 babies in whom the cord was not clamped until the placenta began to descend, the corresponding figures are listed below:

胎盘

10.4	13.1	11.4	9.0	11.9	16.2	14.0	8.2
13.0	8.8	14.9	12.2	11.2	13.9	13.4	11.9

Using WinBUGS, find the posterior distribution for difference in mean blood percentages for the two procedures. Assume gamma likelihoods with different shape/rate parameters for the two procedures. For both the shape and rate parameters use noninformative gamma priors, say $\mathcal{G}a(0.001, 0.001)$.

Does the 95% Credible Set for the difference of means contain 0? Comment.

Hint: Starter file is `babies0.odc`. From the posterior simulations of parameters in gamma likelihoods you will need to calculate the means for the two groups as well as their difference.

2. Can Skull Variations in *Canis lupus L.* Predict Habitat? Data set described below provides skull morphometric measurements on wolves (*Canis lupus L.*) coming from two geographic locations: Rocky Mountain (`0`) and Arctic (`1`). Original source of data

[1] Course Material for ISyE6420 by Brani Vidakovic is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

is Jolicoeur (1959) [2], but subsequently many authors used this data to illustrate various multivariate statistical procedures.

The goal of Jolicoeur's study was to determine how the location and gender affect the skull shape among the wolf populations. There were 9 predictor variables:

```
x1 = palatal length;
x2 = postpalatal length;
x3 = zygomatic width;
x4 = palatal width outside the first upper molars;
x5 = palatal width inside the second upper molars;
x6 = width between the postglenoid foramina;
x7 = interorbital width;
x8 = least width of the braincase;
x9 = crown length of the first upper molar.
```

These 9 measurements (Columns 3-11) and gender (Column 2: male=0, female=1) are associated to one of the two locations (Column 1: Rocky Mountain = 0, Arctic = 1)

```
0 0 4.96 4.09 5.55 3.19 1.25 2.59 2.00 1.73 0.72
0 0 5.04 4.37 5.94 3.17 1.33 2.75 2.07 1.70 0.73
0 0 4.96 4.25 5.98 3.37 1.37 2.72 1.94 1.80 0.70
0 0 4.92 4.29 5.55 3.27 1.34 2.68 1.90 1.72 0.72
0 0 4.96 4.21 5.63 3.22 1.34 2.60 1.93 1.67 0.70
0 0 5.04 4.33 5.63 3.17 1.30 2.56 1.83 1.58 0.72
0 1 4.57 4.02 5.16 3.02 1.24 2.56 1.79 1.54 0.66
0 1 4.72 4.06 5.12 2.96 1.19 2.51 1.75 1.62 0.67
0 1 4.57 4.06 4.92 2.94 1.24 2.46 1.63 1.74 0.67
1 0 4.61 3.90 5.28 3.28 1.37 2.68 1.60 1.46 0.68
1 0 4.53 3.94 5.87 3.19 1.30 2.63 1.86 1.59 0.70
1 0 4.61 4.17 5.59 3.23 1.28 2.60 1.77 1.50 0.72
1 0 4.61 3.98 5.67 3.24 1.29 2.66 1.78 1.63 0.75
1 0 4.61 4.06 5.87 3.26 1.38 2.77 1.90 1.72 0.70
1 0 4.69 3.98 5.63 3.21 1.34 2.72 1.97 1.62 0.74
1 0 4.53 4.02 5.75 3.20 1.33 2.61 1.88 1.65 0.72
1 0 4.61 3.94 5.67 3.20 1.46 2.63 1.63 1.48 0.70
1 0 4.49 4.02 5.55 3.31 1.25 2.67 1.88 1.49 0.68
1 0 4.33 3.70 5.20 3.03 1.19 2.44 1.65 1.59 0.71
1 1 4.41 3.70 5.28 3.13 1.26 2.49 1.77 1.68 0.70
1 1 4.29 3.58 5.24 3.07 1.20 2.44 1.78 1.62 0.67
1 1 4.41 3.90 5.47 3.04 1.29 2.65 1.85 1.61 0.72
1 1 4.41 3.90 5.24 3.09 1.28 2.58 1.74 1.34 0.69
1 1 4.45 3.82 5.75 3.32 1.39 2.70 2.01 1.72 0.68
```

[2] Jolicoeur, P. (1959). Multivariate geographical variation in the wolf *Canis lupus L. Evolution*, **13**, 3, 283–299.

Data here are given in inches.

```
1 1 4.21 3.82 5.39 3.07 1.21 2.43 1.77 1.47 0.65
```

What is the probability that a female wolf with measures $x3 = 5.28$ and $x7 = 1.78$, comes from Arctic habitat? Set a Bayesian logistic regression with two predictors, $x3$ and $x7$, and ignore other variables. A starter file is `wolves0.odc`.

Hint: If all predictor variables are used, the classical logistic regression will fail since 0's and 1's are perfectly separated and iterative reweighed least squares algorithm produces an error. Bayesian logistic regression with all variables will work (since it behaves as iterative penalized least squares), but the model is not stable. Take relatively high precision in priors for regression coefficients, say `dnorm(0, 0.01)`. Expect large standard deviations for the regression coefficients (and wide 95% credible sets).

3. Micronuclei. The Micronuclei (MN) Assay procedure involves breaking the DNA of lymphocytes in a blood sample with a powerful dose of radiation, then measuring the efficiency of its ability to repair itself. Micronuclei are fragments of DNA that have not healed back into either of two daughter nuclei after irradiation. The MN assay entails scoring the number of micronuclei; the higher the number, the less efficient is the subject's DNA repair system.

The dose response of the number of micronuclei in cytokinesis-blocked lymphocytes after in-vitro irradiation of whole blood with x-rays in the dose range 0-4 Gy was studied by Thierens et al, 1991.[3]

The data provided in table are from one patient (male, 54 y.o.) and represent the frequency of micronuclei numbers for six levels of radiation, 0, 0.5, 1, 2, 3, and 4 Gy.

		Number of micronuclei						
		0	1	2	3	4	5	6
Dose (in Gy)	0	976	21	3	0	0	0	0
	0.5	936	61	3	0	0	0	0
	1	895	94	11	0	0	0	0
	2	760	207	32	1	0	0	0
	3	583	302	97	12	6	0	0
	4	485	319	147	35	11	2	1

(a) Fit a Poisson regression in which the number of micronuclei is the response (y) and dose is a covariate (x).

(b) What is the average number of micronuclei for dose of 3.5 Gy?

Hint: The composite data from the table are entered to WinBUGS in form of individual pairs (x, y). There are 6000 pairs in DATA in the starter file `micronuclei0.odc`. Since 6000

[3]Thierens, H., Vral, A., and de Ridder, L. (1991). Biological dosimetry using the micronucleus assay for lymphocytes: interindividual differences in dose response. *Health Phys.*, **61**, 5, 623–630.

0.027, 0.067, 0.116, 0.274, 0.556, 0.778

pairs slow down WinBUGS, do not simulate more than 3,000 instances from corresponding posteriors. The burn in of 500 is fine.