Laboratory Session 05 : May 11, 2023

Exercises due : May 28, 2023

Exercise 1

- Ladislaus Josephovich Bortkiewicz was a Russian economist and statistician. He noted that the Poisson distribution can be very useful in applied statistics when describing low-frequency events in a large population. In a famous example he showed that the number of deaths by horse kick among the Prussian army follows the Poisson distribution.
- Considering the following to sets of observations taken over a fixed large time interval in two different corps:

y death soldiers	0	1	2	3	4	≥ 5
n_1 observations	109	65	22	3	1	0
n_2 observations	144	91	32	11	2	0

(a) assuming a uniform prior, compute and plot the posterior distribution for λ, the death rate over the measurement time. Determine the posterior mean, median and variance, and compute the 95% credibility interval.

(b) assuming now a Jeffreys' prior,

$$g(\lambda) \propto 1/\sqrt{\lambda} \text{ , with } \lambda > 0$$

compute and plot the posterior distribution for λ, the death rate over the measurement time. Determine the posterior mean, median and variance, and compute the 95% credibility interval.

Exercise 2

- solve Exercise 1 with a Markov Chain Monte Carlo. Build your own MCMC, using the functions introduced during lectures, or using a MCM library like JAGS.

Exercise 3

- A study on water quality of streams, a high level of bacter X was defined as a level greater than 100 per 100 ml of stream water. $n = 116$ samples were taken from streams having a high environmental impact on pandas. Out of these, $y = 11$ had a high bacter X level.
- indicating with p the probability that a sample of water taken from the stream has a high bacter X level,

(a) find the frequentist estimator for p

(b) using a $\texttt{Beta}(1, 10)$ prior for p, calculate and posterior distribution $P(p \mid y)$

(c) find the bayesian estimator for p, the posterior mean and variance, and a 95% credible interval

(d) test the hypotesis

$$H_\circ : p = 0.1 \text{ versus } H_1 : p \neq 0.1$$

at 5% level of significance with both the frequentist and bayesian approach

- a new measurement, performed one month later on $n = 165$ water samples, gives $y = 9$ high bacter X level

(e) find the frequentist estimator for p

(f) find a bayesian estimator for p, assuming both a $\texttt{Beta}(1, 10)$ prior for p, and assuming the posterior probability of the older measurement as the prior for the new one.

(g) find the bayesian estimator for p, the posterior mean and variance, and a 95% credible interval

(h) test the hypotesis

$$H_\circ : p = 0.1 \text{ versus } H_1 : p \neq 0.1$$

at 5% level of significance with both the frequentist and bayesian approach

Exercise 4

- analyze the data of Exercise 3 and solve points (b) and (c) using a MCMC with JAGS