STA623 - Bayesian Data Analysis - Practical 5

Marc Henrion

2025-09-26

Practical 5

Notation

- X, Y, Z - random variables
- x, y, z - measured / observed values
- $\bar{X}$, $\bar{Y}, \bar{Z}$ - sample mean estimators for X, Y, Z
- $\bar{x}$, $\bar{y}, \bar{z}$ - sample mean estimates of X, Y, Z
- $\hat{T}$, $\hat{t}$ - given a statistic T, estimator and estimate of T
- $P(A)$ - probability of an event A occuring
- $f_X(.)$, $f_Y(.), f_Z(.)$ - probability mass / density functions of X, Y, Z; sometimes $p_X(.)$ etc. rather than $f_X(.)$
- p(.) - used as a shorthand notation for pmfs / pdfs if the use of this is unambiguous (i.e. it is clear which is the random variable)
- $X \sim F$ - X distributed according to distribution function F
- $E[X]$, $E[Y]$, $E[Z]$, $E[T]$ - the expectation of X, Y, Z, T respectively

Exercise 1

Fit the model from Practical 3, Exercise 3 using `R` and `NIMBLE`.

Use this as the data from the sampling model:

$$y = (1, 3, 2, 3, 0, 2, 6, 4, 4, 1, 1, 3, 2, 3, 1, 1, 3, 0)$$

Inspect the trace plot and plot the posterior distribution.

Compute the posterior mean and the quantile-based 95% Bayesian confidence interval.

Exercise 2

Generate the following data

```
N<-100
x<-rnorm(N)
z<-2-4*x
p<-1/(1+exp(-z))
y<-rbinom(n=N,size=1,prob=p)

dat<-list(N=N,x=x,y=y)
```

Use `R` and `NIMBLE` to fit a Bayesian logistic regression model to these data:

$$g(E[Y|X]) = \beta_0 + \beta_1 X$$

where $g(\pi) = \log(\pi/(1-\pi))$.

Compute the Gelman-Rubin convergence statistic and inspect trace plots and autocorrelations for the samples from the posterior distributions.

Compute the posterior mean, median, a 95% quantile-based confidence interval and a 95% highest posterior density confidence interval.

Compute the effective sample sizes for β_0, β_1.

[end of STA623 BDA Practical 5]