SC7/SM6 Bayes Methods

Third problem Sheet for classes in week 7

1. We considered a version of ABC related to the rejection algorithm. Consider the following MCMC-ABC algorithm, targeting \( \pi(\theta|y) \) (approximately) using the statistics \( S(y) \), distance \( d(S, S') \) and threshold \( \rho \). The observation model is \( p(y|\theta) \) and the prior is \( \pi(\theta) \).

Suppose \( X_t = \theta \).

Step 1. Simulate \( \theta' \sim q(\theta'|\theta) \) and \( y' \sim p(y'|\theta) \).

Step 2. If \( d(S(y), S(y')) < \rho \) then accept \( \theta' \) (set \( X_{t+1} = \theta' \)) with probability

\[
\alpha(\theta'|\theta) = \min \left\{ 1, \frac{\pi(\theta')q(\theta'|\theta')}{\pi(\theta)q(\theta'|\theta)} \right\}
\]

and otherwise reject \( \theta' \) (set \( X_{t+1} = \theta \)).

(a) Show this algorithm targets \( \pi(\theta|d(S(Y), S(y)) < \rho) \) (\( y \) is fixed here and \( Y \sim p(\cdot|\theta) \)).

(b) Suppose we have \( y_i \sim \text{Poisson}(\Lambda) \), \( i = 1, 2, ..., n \) with \( n = 5 \). Prior \( \lambda \sim \Gamma(\alpha = 1, \beta = 1) \). Give the ABC-MCMC algorithm targeting \( \pi(\lambda|y) \) (approximately). Take \( S(y) = \bar{y}, d(\bar{y}', \bar{y}) = |\bar{y}' - \bar{y}| \) and \( \rho = 0.5 \).

(c) (Optional) The code in the R-file `ProblemSheet3incomplete.R` is missing the code for step 2 above. Give the missing code.

2. Consider an MCMC algorithm targeting \( \pi(\theta) \propto \theta^{-1/2}/(1 + \theta^2) \) with \( \theta > 0 \) a scalar parameter. In the following \( \nu \) is a fixed parameter of the MCMC and \( t(\nu) \) denotes the student-t distribution with \( \nu \) degrees of freedom.

(a) Calculate the acceptance probability for the following proposal: \( u \sim t(\nu), \theta' = \theta^u \).

(b) Comment briefly on how you would decide a value for \( \nu \).

(c) (optional) Implement this and check you have the right answer!

3. The skew-normal distribution\(^1\) has density \( Q(y; \mu, \sigma, \xi) \) and is obtained from the normal by skewing it with a positive weight \( \xi > 0 \). The skewing is negative for \( 0 < \xi < 1 \), positive for \( \xi > 1 \) and absent for \( \xi = 1 \), ie \( N(y; \mu, \sigma) = Q(y; \mu, \sigma, 1) \).

The Shoshoni data \( y = (y_1, ..., y_{20}) \) give the values of 20 scalar width-to-length ratios of beaded rectangles used by the Shoshoni Indians. They are available here: [www.statsci.org/data/general/shoshoni.html](http://www.statsci.org/data/general/shoshoni.html).

You can see them and an example of the skew-normal in `ProblemSheet3incomplete.R`.

Consider using Bayesian inference and RJ MCMC to carry out model selection and model averaging over skewed and normal models for the Shoshoni data.

\(^1\) If you would like to see the density and further detail see Fernandez & Steel “Bayesian Modeling of Skewness and Fat Tails”, JASA, 1998
(a) Suppose the prior probability for normal (model \( m = 1 \)) or skew-normal (model \( m = 2 \)) is \( 1/2 \). Write down the joint posterior distribution \( \pi(\theta, m|y) \) for the model index \( m = 1, 2 \) and parameters \( \theta = (\mu, \sigma, \xi) \) in as much detail as you can, though without eliciting priors for the parameters.

(b) Give a reversible jump MCMC algorithm targeting \( \pi(\theta, m|y) \). You can omit the fixed dimension updates.

(c) (Optional) The code in the R-file `ProblemSheet3incomplete.R` is missing the expressions for the numerator and denominator in the birth and death updates. All the functions you need are available in the lines above. Give the missing code and use it to estimate the Bayes factor mentioned above.

4. In a lottery, a fair coin is tossed until a head comes up. If it takes \( n \) tosses then you receive a reward \( r = £2^n \). Suppose you are given a choice between accepting a certain fixed reward \( r = r^* \) or taking the random lottery reward \( r = 2^n \).

(a) Suppose your utility function is linear and increasing in \( r \), \( U(r) = c + mr \) with \( m > 0 \).
Show that, under the expected utility hypothesis, (ie choose the action maximising expected utility) you would choose the lottery, however large \( r^* \) is.

(b) Write down a suitable concave utility function for which there exists finite \( r_0 \) such that you would take the fixed reward for every \( r^* \geq r_0 \).

5. Consider two urns. In the first urn there are 50 black balls and 50 red balls. In the second urn there are 100 balls, the number of each color unknown. Suppose the proportion of back balls in the second urn is equal \( \phi \).

Jane’s \( \phi \)-prior, \( \pi(\phi) \), satisfies \( E(\phi) = 1/2 \) and \( \phi \) is is not almost surely constant, ie the prior is not simply \( \pi(\phi) = \delta(\phi - 1/2) \). Jane is offered a choice of urn and color and two balls are drawn (with replacement) from the chosen urn. Jane receives a £1 reward for each ball matching her chosen color. Her utility function is \( U(0) = 0, U(1) = v, U(2) = 1 \) with \( 1/2 < v < 1 \).

Jane is offered red from the first urn or black from the second.

(a) Show that the expected utility of choosing the second urn given \( \phi \) is
\[
E(U|\phi) = 2\phi(1 - \phi)v + \phi^2.
\]

(b) Show that Jane should choose the first urn.

(c) Suppose Jane had been offered black from the first urn or red from the second. Show that Jane should again choose the first urn.

(d) In what sense does this offer a resolution of the Ellsberg paradox?

6. (optional) The Savage axioms (as formulated by DeGroot) characterise coherent prior preference for events stated in terms of inequalities, so that \( A \leq B \) implies \( B \) is at least as likely as \( A \).
(a) Write down the first three axioms (see Lecture notes).

(b) Show that preference relations satisfying the axioms have the following properties:

i. If $A \leq B$ then $A^c \geq B^c$ (where $A^c$ is the complement of $A$ etc);

ii. If $A_1 \cap D = B_1 \cap D = \emptyset$ then $A_1 \cup D \geq B_1 \cup D$ iff $A_1 \geq B_1$.

iii. The order is transitive, ie, if $A \leq B$ and $B \leq C$ then $A \leq C$. 