Project 1 for Bayesian Inference
Spring 2009

In questions 1-10, construct samples by transforming output from the `runif(0,1)` command in R. Histograms and other summary statistics based on the random samples you generate should be examined to validate your solutions. In questions 11- , you can use functions other than `runif(0,1)` in order to simulate from a suitable distribution.

1. Generate a sample of length 1000 from a $\mathrm{N}(0,1)$ distribution, making use of the Box-Muller technique. Use graphical methods (e.g. histogram) to convince yourselves that you get what you want. Estimate the sample mean and the sample variance. Do a q-q plot.

2. Generate a sample of length 1000 from a $\mathrm{N}(10,10)$ distribution, making use of the Box-Muller technique. Do the same as above.

3. Simulate 1000 draws from a $\mathrm{U}(-1,1)$ distribution. Produce a histogram. Compute the sample mean and the median.

4. Simulate 1000 draws from an Exponential(2) distribution by inversion of the c.d.f. Use graphical methods (e.g. histogram) to convince yourselves that you get what you want. Next compute $\exp(x_i)$, where x_i are the draws you obtained, and convince me that these new numbers are uniformly distributed between 0 and 1.

5. Simulate 1000 draws from a $\mathrm{Gamma}(3,2)$ distribution. What is the (theoretical) mean and variance of such a distribution? Do the sample mean and variance appproximate them? Draw a histogram and compare it with the theoretical density.

6. Devise a method for simulating a Poisson(λ) variate. There are multiple ways of doing this. Can you identify at least two methods? Apply your method for $\lambda = 10$ and $\lambda = 100$. Generate a large number of such variables and use graphical methods to summarise your results.

7. Simulate a $\mathrm{Bin}(12, 0.5)$ variate. Generate a large number of such variables and use graphical methods to summarise your results.

8. Simulate throwing ten fair dice. What is the probability of getting a total of 30 or more? Compare this to the theoretical probability which you can compute numerically.

9. Simulate 1000 realisations of a pair (X, Y) of jointly normal random variables with

$$EX = 10, EY = 12, \text{var}(X) = 1, \text{var}(Y) = 9, \text{cov}(X, Y) = 2.$$

Estimate sample means, variances and covariances. Produce a histogram for $3X - 4Y$ and compare it to the actual density of $3X - 4Y$ (which is normal with a mean and variance you can find).

10. Draw a sample of length 1000 from a Beta$(2, 2)$ distribution. There are (at least) two ways to do this: (i) by simulating two variates from a gamma distribution first, and (ii) using rejection sampling. Try them both—which generates a sample of length 1000 fastest? Could you do it by inverting the c.d.f.? (Harder) Use your random sample from Beta$(2, 2)$ to generate a (smaller) random sample from Beta$(2.2, 2.2)$ by applying rejection sampling. Compare the histogram to the actual density on the same graph.

11. Implement the Metropolis-Hastings algorithm for sampling from a uniform distribution on the set 1, 2, 3, 4. By examining a sequence of outputs from the chain, verify that it converges to the desired stationary distribution. How would you modify the Markov chain in order to sample from an *arbitrary* distribution on 1, 2, 3, 4? Do the latter for, at least, the distribution that assigns probabilities $0.1, 0.2, 0.3, 0.4$ to the states $1, 2, 3, 4$, respectively. Show that what you got is right by producting a histogram.

12. Design and implement a random-walk Metropolis algorithm to simulate from a Gamma(2, 1) distribution using a Normal distribution (with mean equal to the current state) as the the proposal distribution.

    (a) By examining a time series plot of the output from the chain investigate how the mixing properties of the chain depend on the variance of the proposal distribution.

    (b) Investigate how the dynamics of your chain are affected by the choice of initial state.