Department of Physics
Computational Inference
Assignment 1 (due 15 August by close of play)

1. Consider the joint pdf $\pi(x, y, \theta)$ over x, y, and θ. Using Bayes' rule, product factoring, etc, establish the following identities between the pdfs over various conditional and marginal distributions:

   (a) $\pi(\theta|y) = \dfrac{\pi(x, \theta|y)}{\pi(x|\theta, y)}$ (This is useful because the RHS may be evaluated at *any* x for which the denominator is non-zero.)

   (b) $\pi(x, \theta|y) = \pi(x|\theta, y)\,\pi(\theta|y)$, i.e., we can condition the usual product factoring.

   (c) $\pi(\theta|y) = \dfrac{\pi(x, y|\theta)\,\pi(\theta)}{\pi(x|\theta, y)\,\pi(y)}$ (Again, RHS may be evaluated at any x for which denominator is non-zero. Useful when full conditional over θ is a GMRF.)

2. Assume the $x_1, x_2, \ldots$ are the sequence of states generated by a Markov process (perhaps an MCMC) that is ergodic with respect to some distribution $\pi(x)$. Let $y = g(x)$ for some function g of x, with corresponding pdf $\pi(y)$ (given by the change of variable formula).

   (a) Show that $g(x_1), g(x_2), \cdots$ are a the sequence of states generated by a Markov process that is ergodic with respect to $\pi(y)$.

   (b) Show that if the n-step distribution over x_i converges in distribution, then so does the n-step distribution over y_i.

3. Consider sampling from the univariate standard normal, N(0,1). In each case generate 10000 samples, and compute IACT for state variable.

   (a) Using the code mcgauss (link on paper website), or otherwise, plot IACT versus window size when using random walk Metropolis with a normal proposal. What is the optimal IACT and window size?

   (b) Repeat the previous question but with a uniform proposal. (Isn't that interesting).

   (c) Download the IA2RMS code (link on paper website), use that to sample from N(0,1), and give the IACT. Can you determine how many evaluations of the N(0,1) pdf have been made (to ensure that IA2RMS is not cheating)?

   (d) Download Marko Laine's mcmcrun code (link on paper website) and use the adaptive Metropolis option (see sample code on the paper website) to sample from N(0,1). Does it achieve the optimum you found in the first part?

4. MatLab code to carry out inference by MCMC for the diffusion coefficient can be found on the paper's web page https://coursesupport.physics.otago.ac.nz/wiki/pmwiki.php/ELEC446

   (a) Run that code (or python equivalent) and produce a plot of the posterior pdf over D conditioned on the (fake) measured data. Give a best value, and measure of uncertainty.

(b) Plot the integrated autocorrelation time versus proposal window, and acceptance rate versus proposal window for this MCMC (for fixed fake data). What is the optimal IACT, window size, acceptance rate?

(c) Repeat the sampling (generate 10000 samples) using AM and IA2RMS. How do these algorithms compare for efficiency?

(d) Using whichever you think is the best sampling algorithm available to you, modify the formulation and code to allow the final time `Tmax` to be uncertain, uniformly distributed in the range $[1.9, 2.1]$. In particular, perform joint inference over D and the (nuisance parameter) `Tmax`, and plot the posterior pdf over D conditioned on the (fake) measured data. Give a best value, and measure of uncertainty. How does the uncertainty compare to the result when `Tmax` was treated as certain?