STA 4503, Spring 2013 — Theory Assignment #1

Due at the start of class on March 6. Please hand it in on 8 1/2 by 11 inch paper, stapled in the upper left, with no other packaging.

This assignment is to be done by each student individually. You may discuss it in general terms with other students, but the work you hand in should be your own. In particular, you should not leave any discussion with someone else with any written notes (either on paper or in electronic form).

Recall that convergence of a Markov chain Monte Carlo method can be measured by the "total variation distance" between the desired distribution and the distribution after some number of iterations. Writing π for the desired distribution, and P_x^n for the distribution of the chain after n iterations starting at state x, the total variation distance between these distributions is defined as

$$||P_x^n \; - \; \pi|| \; = \; \max_S |P_x^n(S) - \pi(S)|$$

where the maximum is over subsets of the state space for x, which in this assignment is assumed to be finite. We say that the chain is "uniformly ergodic" if, for some $a > 0$ and some $b \in (0, 1)$, for any x and any $n \geq 0$,

$$||P_x^n \; - \; \pi|| \; < \; a\, b^n$$

Recall also that the efficiency with which a chain can be used to estimate the expectation of some function of state, $h(x)$, assuming that it has converged, is determined by the "autocorrelation time", defined as

$$\tau \; = \; 1 \; + \; 2 \sum_{i=1}^{\infty} \rho_i$$

where ρ_i is the autocorrelation of $h(x)$ at lag i, which can be defined in terms of autocovariances as $\rho_i = \gamma_i / \gamma_0$ with $\gamma_i \; = \; E[(h(x_t) - \mu)(h(x_{t+i}) - \mu)]$ and $\mu = E[h(x)]$. Specifically, for large n, the variance of $\bar{h}_n = (h(x_t) + \cdots + h(x_{t+n-1}))/n$ is

$$\mathrm{Var}(h(x)) \,/\, (n/\tau) \; = \; \gamma_0 \,/\, (n/\tau) \; = \; \gamma_0 \tau / n$$

In this assignment, you should prove that (phrased informally) a chain that is uniformly ergodic, with a small value for b (and hence converges quickly), will produce good estimates of expectations of functions of state. Specifically, you should:

a) Prove that for any distributions P and Q for x, and any function $h(x)$, the total variation distance between the distribution of $h(x)$ under P and $h(x)$ under Q is no larger than the total variation distance of P and Q. That is,

$$||P_h - Q_h|| \; \leq \; ||P - Q||$$

Here, P_h is the distribution for $h(x)$ implied by the distribution P for x, and similarly for Q_h.

b) Prove that if $h(x)$ is some function for which $|h(x)| \leq C$ for all x, and P and Q are distributions for x, then

$$|E_P(h(x)) \; - \; E_Q(h(x))| \; \leq \; 2C||P - Q||$$

c) For a chain that is uniformly ergodic, with $||P_x^n \; - \; \pi|| \; < \; ab^n$, state and prove as good an upper bound as you can on the variance of $\bar{h}_n$ for some function of state $h(x)$ that satisfies $|h(x)| \leq C$ for all x. Hints: It's easier to work with $\gamma_0 \tau$ than with τ. Recall the theorem of probability that says $E[Z] = E[E[Z|Y]]$.