4.9 Markov Chain Monte Carlo (MCMC) Methods

Note: The discussion of the following methods will be in the context of a discrete r.v., but the results hold true analogously in the continuous case.

- Suppose X is a (possibly vector-valued) discrete r.v. with pmf $P[X = x_j]$, $j \geq 1$.
- We wish to calculate the expected value of some specified function of X:

- Sometimes it is difficult to compute this analytically, so we can use Monte Carlo simulation to approximate θ.
- Generate a sequence of n iid r.v.'s $X_1, X_2, \ldots, X_n$ having pmf $P[X = x_j]$, $j \geq 1$.

- By the law of large numbers,

So if we make n large, then

<u>Problem:</u> It might be difficult to directly generate X_i's following the needed distribution.

- This is especially true if the X_i's are <u>random vectors</u> whose components are <u>dependent</u>.

<u>Another Problem:</u> Sometimes we don't know the whole pmf. We may know

where the b_j are known, but the constant C is unknown.

- Of course, since $\sum_j c b_j = 1$, then $c =$, but often is difficult or impossible to compute.

Markov Chain Monte Carlo (MCMC)

- In such cases, we may generate a sequence of <u>dependent</u> (rather than independent) r.v.'s which are the states of a <u>Markov chain</u> having stationary probabilities $P[X = x_j]$, $j \geq 1$.
- The LLN still holds in this case, and

 is still a consistent estimator of $\theta = E[h(X)]$.

The Metropolis-Hastings (M-H) Algorithm

- This is a famous MCMC method for generating r.v.'s following a certain distribution (which may be known only up to a normalizing constant).

- Let b_j $(j=1,2,\ldots)$ be positive numbers such that $B = \sum_{j=1}^{\infty} b_j < \infty$.
- We want to generate a Markov chain having stationary probabilities

- This probability distribution is called the <u>target</u> distribution.
- Let Q be the transition probability matrix of a specified irreducible Markov chain having state space $\{1,2,\ldots\}$.
- The elements of Q are $\{q(i,j)\}$, $i,j \in \{1,2,\ldots\}$.
- The Markov chain resulting from the M-H algorithm is $\{X_n\}$, defined as follows:
  * When $X_n = i$, generate a r.v. Y such that

* For $Y = j$, set

$X_{n+1} =$

Then $\{X_n\}$ is a Markov chain with transition probabilities

Note: The distribution $q(\cdot,\cdot)$ of the candidate r.v. Y is called the proposal distribution.

- The Markov chain is time-reversible with stationary probabilities $\pi(j)$ if

(the ________________ condition).

- This condition is satisfied if we take $\pi(j) = \frac{b_j}{B}$ and set

Why?