STAT 516 LAB 1

1. Let (X_k) be a Markov chain with transition matrix

$$P = \begin{pmatrix} .1 & .2 & .3 & .3 & .1 \\ .5 & .1 & .1 & .1 & .2 \\ .1 & .5 & .1 & .1 & .2 \\ .1 & .2 & .5 & .1 & .1 \\ .2 & .2 & .1 & .4 & .1 \end{pmatrix}$$

How long do you need to observe the chain in order to estimate P well? You decide what you mean by "well".

Functions you may need include generateMarkovChain and count from the class software page.

2. Starting from state 1, find $\mathbf{E}^{1}(T_{5})$

(a) based on the true P in 1

(b) based on the estimated *P* in 1.

3. The expected value in question 2 is a certain function of

$$F_{15}(t; P) = \mathbf{P}^1 \{ T_5 \le t \}.$$

It is complicated to determine this distribution analytically, even when P is known. If P is unknown, it is necessary to apply simulation methods. The bootstrap (Efron, 1978, Ann. Stat.) can be modified to do this. Let \hat{P} be the estimated transition matrix. Generate a path from \hat{P} , and use that to estimate the transition matrix the usual way, yielding a matrix \hat{P}^* . The bootstrap idea is to say that the distribution of

$$n^{1/2}(F_{15}(t;\hat{P}) - F_{15}(t;P))$$

is well approximated by

$$n^{1/2}(F_{15}(t;\hat{P}^*) - F_{15}(t;\hat{P})).$$

In practice, one would use repeated samples from \hat{P} to get realizations of the hitting time, and then use the empirical distribution function of these hitting times to estimate $F_{15}(t; P)$ as well as whatever function of it you are interested in. The approximation argument above is the theoretical basis for this procedure. Compute a bootstrap estimate using the sample generated in part 1. Use a bootstrap sample of size 100. Compare the mean of the resulting estimate to the answes in 2. Can you use the bootstrap distribution to estimate the variance of the estimated mean?