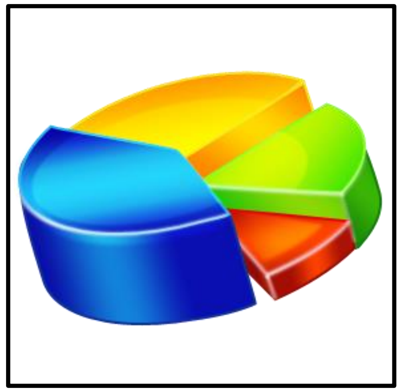
STATISTICS SEMINAR

Xiaojing Zhang

Master's Defense

Tuesday, May 3, 2016

Dickens Hall, Room 109, 10:30 am-12:00 pm



A Simulation Study of Confidence Intervals for the Transition Matrix of a Reversible Markov Chain

Let $\{X_i; i=0,1,2,...,\}$ be an irreducible, aperiodic, positive recurrent Markov chain with state a subset of space $\{0,1,...\}$, time-homogeneous transition matrix $\mathbf{P}$ = ($p_{ij} = P(X_1 = j \mid X_0 = i)$) and limiting distribution $\{\pi_i > 0\}$. Based on observing $\{X_i = x_i; i = 0,1,2,\ldots,t\}$, we study and compare two estimators of the transition probabilities $\{p_{ij}\}$:

(i) The maximum likelihood estimators $\{\hat{p}_{ij}\}$ of $\{p_{ij}\}$:

$$\hat{p}_{ij} = [N_{ij}(t)/N_i(t)]I(N_i(t)>0),$$

where

$$N_{ij}(t) = \sum_{l=0}^{t-1} I(X_l = i, X_{l+1} = j),\ N_i(t) = \sum_{l=0}^{t} I(X_l = i).$$

(ii) The symmetrized estimators $\{\hat{p}_{ij}^{(R)}\}$:

$$\hat{p}_{ij}^{(R)} = \left[(N_{ij}(t) + N_{ji}(t))/2N_i(t)\right] I(N_i(t) > 0).$$

It is well known that as $t \to \infty$, in distribution,

$$\sqrt{t}(\hat{p}_{ij} - p_{ij}) \to N(0, p_{ij}(1-p_{ij})/\pi_i). \qquad \text{(I)}$$

It was shown in Annis et. al. (2010) that *if the chain is reversible*, in distribution,

$$\sqrt{t}(\hat{p}_{ij}^{R} - p_{ij}) \to N(0, \sigma_{ij}^2(R)), \qquad \text{(II)}$$

with σ_{ij}^2(R) / $[p_{ij}(1-p_{ij})/\pi_i] \in [1/2, 1]$, implying that for a reversible chain $\hat{p}_{ij}^{(R)}$ is asymptotically as least as good as $\hat{p}_{ij}$ for reversible chains.

We designed and carried out a simulation study, using representative choices of n and $\mathbf{P}$, to compare the performance, in terms of coverage rate and mean width, of nominal 0.95 confidence intervals for the elements of the transition matrix $\mathbf{P}$ constructed using (I) and (II), where the limiting variances are replaced by appropriate sample estimates. When the chain is reversible, both intervals are asymptotically correct and the intervals based on II are asymptotically no wider than those based on I. However, our simulations indicate that for the finite sample sizes and models used here, the estimated coverage rates based on II appear to be considerably below their nominal values in some cases. Since, the coverage rates for the intervals based on (I) appear to approach their nominal levels as sample sizes increase, we recommend using them rather than the more complicated intervals based on II.

KEY WORDS: Stationary distribution; Reversibility; Transition probability estimation.