CONTROL ID: 3389450
PRESENTER: Hyoyoung Choo-Wosoba
PRESENTER (INSTITUTION ONLY): National Cancer Institute
TITLE: A Hidden Markov Modeling Approach for Identifying Tumor Subclones in Next-Generation Sequencing Studies
ABSTRACT BODY:
Abstract Body: Allele-specific copy number alteration (ASCNA) analysis is for identifying copy number abnormalities in tumor cells. Unlike normal cells, tumor cells are heterogeneous as a combination of dominant and minor subclones with distinct copy number profiles. Estimating the clonal proportion and identifying mainclone and subclone genotypes across the genome is important for understanding tumor progression. Several ASCNA tools have recently been developed, but they have been limited to the identification of subclone regions, and not the genotype of subclones. In this talk, I propose subHMM, a hidden Markov model-based approach that estimates both subclone region as well as region-specific subclone genotype and clonal proportion. I specify a hidden state variable representing the conglomeration of clonal genotype and subclone status. I propose a two-step algorithm for parameter estimation, where in the first step, a standard hidden Markov model with this conglomerated state variable is fit. Then, in the second step, region-specific estimates of the clonal proportions are obtained by maximizing region-specific pseudo-likelihoods. I apply subHMM to study renal cell carcinoma datasets in The Cancer Genome Atlas. In addition, I conduct simulation studies that show the good performance of the proposed approach.
AUTHORS/INSTITUTIONS: H. Choo-Wosoba, P. Albert, B. Zhu, National Cancer Institute, Rockville, Maryland, UNITED STATES|

CONTROL ID: 3391870
PRESENTER: Xinyuan Song
PRESENTER (INSTITUTION ONLY): Chinese University of Hong Kong
TITLE: Bayesian Adaptive Group Lasso with Semiparametric Hidden Markov Models
ABSTRACT BODY:

Abstract Body: This study develops a Bayesian adaptive group least absolute shrinkage and selection operator (lasso) method to conduct simultaneous model selection and estimation under semiparametric hidden Markov models (HMMs). We specify the conditional regression model and the transition probability model in the HMM into additive nonparametric functions of covariates. A basis expansion is adopted to approximate the nonparametric functions. We introduce multivariate conditional Laplace priors to impose adaptive penalties on regression coefficients and different groups of basis expansions under the Bayesian framework. An efficient Markov chain Monte Carlo algorithm is then proposed to identify the non-existent, constant, linear and nonlinear forms of covariate effects in both conditional and transition models. The empirical performance of the proposed methodology is evaluated via simulation studies. We apply the proposed model to analyse a real dataset that was collected from the Alzheimer's Disease Neuroimaging Initiative study. The analysis identifies important risk factors on cognitive decline and the transition from cognitive normal to Alzheimer's disease.

AUTHORS/INSTITUTIONS: X. Song, Statistics, Chinese University of Hong Kong, Shatin, HONG KONG|

CONTROL ID: 3402741
PRESENTER: Abhra Sarkar
PRESENTER (INSTITUTION ONLY): The University of Texas at Austin
TITLE: Bayesian Semiparametric Higher Order Hidden Markov Models
ABSTRACT BODY:

Abstract Body: We consider the problem of flexible modeling of higher order hidden Markov models when the number of latent states and the nature of the serial dependence, including the true order, are unknown. We propose flexible Bayesian methods based on tensor factorization techniques that can characterize any transition probability with a specified maximal order, allowing automated selection of the important lags and capturing higher order interactions among the lags. Theoretical results provide insights into identifiability of the emission distributions and consistency of the posterior. We design Markov chain Monte Carlo algorithms for posterior computation. In simulation experiments, the method vastly outperforms competitors not just in higher order settings, but, remarkably, also in first order cases. Practical utility is illustrated using real world applications.

AUTHORS/INSTITUTIONS: A. Sarkar, Statistics and Data Sciences, The University of Texas at Austin, Austin, Texas, UNITED STATES|