

Colorado-Wyoming Chapter

Hello everyone – Just a quick reminder that the CO/WY 2010 Spring meeting will be held at the end of this month. Details are included below. A job announcement is also included below.

REMINDER: SPRING CO/WY CHAPTER MEETING (April 23, 2010)

The spring ASA CO/WY meeting will be held on Friday, April 23, 2010 at the NCAR facility in Boulder, CO. There is no need to RSVP for this event. We will meet in the ML-239 Damon Room located at NCAR - 1850 Table Mesa Dr., Boulder, CO 80305. Directions - Take U.S. Highway 36 to Boulder Colorado • Exit at Table Mesa Drive • Head West on Table Mesa Drive • The Mesa Lab and Fleischmann Building are located at 1850 Table Mesa Drive where the road ends at the base of the Rocky Mountains

Event: CO/WY Spring Meeting 2010

Date: April 23, 2010

Location: NCAR, ML-239 Damon Room

Time: 9:00am – 4:00pm

Agenda

8am-9am: Refreshments

9am-11:30am Welcome and Presentations

11:30am-1:15pm: Lunch on your own or \$5 in the cafeteria

1:15pm-4pm: Presentations

Speakers

Myung-Hee Lee [mhlee.csu@gmail.com]

Title: Clustering High Dimensional, Low Sample Size Data using Maximal Data Piling: We present new hierarchical clustering method for high dimension, low sample size (HDLSS) data. The method utilizes the fact that each individual data vector accounts for exactly one dimension in the subspace generated by HDLSS data. The linkage that is used for measuring the distance between clusters is the orthogonal distance between a new subspaces generated by each cluster. The ideal implementation would be to consider all possible binary splits of data and choose the one that maximizes the distance in-between. Since this is not computationally feasible in general, however, we use singular value decomposition for its approximation. We provide theoretical justification of the method by studying high dimensional asymptotics. Also we obtain the probability distribution of the distance measure under the null hypothesis of no split, which we use to propose a criterion for determining the number of clusters. Simulation and real data analyses with microarray data show competitive clustering performance of the proposed method. (This is a joint work with Jeongyoun Ahn and Youngju Yoon)

A.M. Santos [pintyesantos@gmail.com]

Robust estimation of mixtures of heavy-tailed distributions A. M. Santos, IBM and UC-Denver Karen Kafadar, Indiana University. We examine the robustness and efficiency of a variant of the EM algorithm for estimating parameters of mixtures of h-distributions. Specifically, we use the biweight in an EM-like algorithm to estimate location, scale, and proportion of each component of a mixture of heavy-tailed distributions. This family of long-tailed distributions, indexed by a tail-length parameter ($h=0$ corresponds to the Gaussian), provides a measure by which robustness and efficiency can be assessed. We compare our results to those using the conventional EM-algorithm (that assumes Gaussian distributions) and with Scott's L₂E (2001) approach. Finally, we describe the application that motivated this research and our plans for future work.

Jun Zhu [jun.zhu.e@gmail.com]: Variable Selection in Spatial Linear Regression

Bahr, Timothy [TIMOTHY.BAHR@UCDENVER.EDU]

SLAM: Gaussian Dynamic Linear Analysis of MeDIP-chip Data

This paper presents an efficient and powerful algorithm, Gaussian Dynamic Linear Analysis of Methylated- chip data (SLAM), for profiling DNA methylation patterns from samples enriched for methylated DNA through immunoprecipitation (MeDIP) and hybridized on genome tiling microarrays (MeDIP-chip). SLAM is distinguished from other existing algorithms by its application of MeDIP-chip specific normalization, dynamic linear smoothing and a Probit transformation in order to accurately profile the percentage of methylated DNA bases across the genome. SLAM is the first MeDIP-chip analysis method that can directly compare the DNA methylation levels of multiple tissue types, as well as being the first method appropriate for the analysis of tiling array data for other epigenetic marks, such as histone methylation or acetylation. We apply SLAM to several epigenetic profiling datasets. Results show that SLAM can provide more efficient, accurate and consistent methylation estimates than other existing algorithms, implying that SLAM has the greater potential for biological discovery. SLAM is available as a command-line application or in a graphical user interface, and is freely available for download at: <http://statistics.byu.edu/johnson/SLAM/>.

Amber Hackstadt [ahacksta@lamar.colostate.edu]

A Bayesian Approach to Fitting Mixed Models Using Shape Restricted Regression Splines, Amber Hackstadt, Mary Meyer, and Jennifer Hoeting: We propose a Bayesian approach to fit shape-restricted regression splines for mixed models. Linear mixed-effects models are often used to analyze repeated measure and longitudinal data but often the only thing that is known about the data is the shape (monotone increasing/decreasing or concave/convex) and smoothness of the regression functions. Shape-restricted splines are used to model regression functions in mixed-effects models.

Yuan Wang [ywang.csu@gmail.com]

Carbon Flux in the Mid-Continent Region Project: The net exchange of CO₂ between the terrestrial biosphere and atmosphere remains a key uncertainty in the carbon cycle. An important campaign has been launched in the Mid-Continent region of USA, and our main goal is to compare, diagnose and reconcile two different estimates of CO₂ exchange over the study region.

Bruce Bugbee [brucebugbee@gmail.com]

A popular topic in e commerce research is the observation and quantifying of bidder behavior in online auction systems. A functional methodology and cluster analysis was used to quantify bidder behaviors across two distinct item types--collectibles (1968 Camaros) and commodities (generic digital camera). Similar behaviors were observed with varying proportions of occurrence and levels of auction success."

JOB ANNOUNCEMENT

Job Posting: Instructor in the Business School at the University of Colorado Denver.

We're seeking a full-time instructor to teach undergraduate and graduate (master's level) statistics courses. More information about the position is available at

https://www.jobsatcu.com/applicants/jsp/shared/position/JobDetails_css.jsp?postingId=221403. The job posting number is 809565.