2015 Chapter Meeting

ICSA Midwest Chapter

October 25-26, 2015

AbbVie Campus at Mettawa, Illinois
This meeting is sponsored by

visit www.abbvie.com for more information.

AbbVie intern program (please see back of this brochure).
Meeting Organization Committees

**Executive Committee**

Chair, Lanju Zhang, AbbVie Inc  
Chapter Chair, Yang Bo, AbbVie Inc  
Program Chairs, Annie Qu, University of Illinois; Haoda Fu, Eli Lilly & Company  
Student Poster session, Lingsong Zhang, Purdue University  
Local Committee Chair, Suzana Bancila, AbbVie Inc  
Treasurer, Ying Zhang, Indiana University  
Publicity and fundraising Chairs, Ming-Dauh Wang, Eli Lilly and company

**Program Committee**

Chair: Annie Qu and Haoda Fu  
Short Course: Annie Qu  
Student Poster: Lingsong Zhang  
Organizers: Haoda Fu, Annie Qu, Yijie Zhou, Judith Xu, Yuan Ji, Peter Song

**Local Committee**

Chair, Suzana Bancila, AbbVie Inc  
Yihua Gu, AbbVie Inc  
Deli Wang, AbbVie Inc  
Ying Zhang, AbbVie Inc
Short Course

Oct 25 1:30pm - 5:00pm (Mexico City Room)

Title: Introduction to Machine Learning and Data Mining

Abstract: The course will start from the basic principles of machine learning and end with the frontiers in statistical learning and data mining. The lectures will cover a collection of methods that have been developed in recent years in statistical machine learning and data mining. Specifically, the topics will include: bias-variance trade-off, cross-validation, k-nearest neighbors, logistic regression, linear discriminant analysis and quadratic discriminant analysis, support vector machines and kernel methods, boosting and random forests.

Instructor: Professor Ji Zhu, University of Michigan

Bio: Ji Zhu obtained his B.Sc. in Physics from Peking University in 1996 and his Ph.D. in Statistics from Stanford University in 2003. His research interests include statistical learning, high-dimensional data and statistical network analysis. He is also interested in applications in medicine, computational biology, engineering, physics and business. Professor Zhu received a CAREER award from the NSF in 2008 and was elected a member of ISI in 2010 and a fellow of ASA in 2013.

Main Program

October 26, 2015

<table>
<thead>
<tr>
<th>Time</th>
<th>Session Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 – 10:00 am</td>
<td>Registration and Breakfast</td>
</tr>
<tr>
<td>10:00 – 10:15 am</td>
<td>Welcome</td>
</tr>
<tr>
<td>(Chicago Room)</td>
<td>Some tips on developing leadership skills within your organization</td>
</tr>
<tr>
<td></td>
<td>Walt Offen (AbbVie Inc)</td>
</tr>
<tr>
<td>10:15 –10:30 am</td>
<td>Keynote Address</td>
</tr>
<tr>
<td>(Chicago Room)</td>
<td>Opportunities and professional development for Statisticians</td>
</tr>
<tr>
<td></td>
<td>Wei Shen (Eli Lilly and Company)</td>
</tr>
<tr>
<td>10:30–10:45 am</td>
<td>Break/Networking</td>
</tr>
<tr>
<td>10:45 –12:00 pm</td>
<td>Parallel Session 1: Personalized medicine (Chicago Room)</td>
</tr>
<tr>
<td></td>
<td>Parallel Session 2: Challenges in high-dimensional data (Mexico City Room)</td>
</tr>
<tr>
<td>12:00 – 1:30 pm</td>
<td>Lunch &amp; Poster Session</td>
</tr>
<tr>
<td>1:30 pm – 2:45 pm</td>
<td>Parallel Session 3: Causal inference and observational studies (Chicago Room)</td>
</tr>
<tr>
<td></td>
<td>Parallel Session 4: Longitudinal and survival analysis (Mexico City Room)</td>
</tr>
<tr>
<td>2:45 –  3:00 pm</td>
<td>Break/Networking</td>
</tr>
<tr>
<td>3:00 pm – 4:15 pm</td>
<td>Parallel Session 5: Novel Statistical Methods and Considerations to Meet Modern Oncology Challenges (Chicago Room)</td>
</tr>
<tr>
<td></td>
<td>Parallel Session 6: Bayesian and bioinformatics</td>
</tr>
<tr>
<td>4:20 pm – 4:30 pm</td>
<td>Poster Awards and Closing Remarks</td>
</tr>
<tr>
<td></td>
<td>Lingsong Zhang (Perdue U), Lanju Zhang (Abbvie Inc) (Mexico City Room)</td>
</tr>
</tbody>
</table>
### Parallel Sessions October 26, 2015

**Session 1: Personalized medicine**
Organizer: Haoda Fu (Eli Lilly and Company)

- Enrichment Design with patient population augmentation
  - Yijie Xhou (AbbVie Inc)
- **GUIDE the search of tailoring biomarkers**
  - Michael Man (Eli Lilly and Company)
- Searching for optimal treatment rule – a new look on personalized medicine and subgroup identification
  - Haoda Fu (Eli Lilly and Company)

**Session 2: Challenges in high-dimensional data**
Organizer: Annie Qu (University of Illinois at Urbana-Champaign)

- Regularized estimation of linear functional for high-dimensional time series
  - Xiaohui Chen (University of Illinois at Urbana-Champaign)
- Selection by partitioning the solution path
  - Peng Wang (University of Cincinnati)
- An augmented ADMM algorithm with application to high-dimensional statistical estimation problem
  - Yunzhang Zhu (Ohio State University)

**Session 3: Causal inference and observational studies**
Organizer: Judith Xu (Takeda)

- Indirect treatment comparisons
  - Shawn Yu (Takeda)
- Bayesian sensitivity analysis to assess the impact of unmeasured confounding in comparative observational study
  - Xiang Zhang (Eli Lilly and Company)
- Causal inference
  - Alan Fan (Astellas)

**Session 4: Longitudinal and survival analysis**
Organizer: Peter Song (University of Michigan)

- Conditional modeling of longitudinal data with terminal event
  - Shengchun Kong (Purdue University)
- Joint Frailty Models for Zero-Inflated Recurrent Events in the Presence of a Terminal Event
  - Lei Liu (Northwestern University)
- Mixtures of g-priors for hypothesis testing in ANOVA models with a diverging number of parameters
  - Min Wang (Technical University of Michigan)

**Session 5: Novel statistical methods and considerations to meet modern oncology challenges**
Organizer: Yijie Zhou (AbbVie Inc)

- Utilizing the relation between pCR and overall survival in trial design for neoadjuvant breast cancer therapies
  - Ming Zhu (AbbVie Inc)
- Multi-regional non-inferiority trials in oncology
  - Alan Rong (Astellas)
- Progression-free survival: how often you look matters
  - Jingyi Liu (Eli Lilly and Company)

**Session 6: Bayesian and bioinformatics**
Organizer: Yuan Ji (The University of Chicago)

- MetaXcan: a scalable gene-level association test
  - Hae Kyung Im (The University of Chicago)
- Zodiac: a comprehensive depiction of genetic interactions in cancer using TCGA data
  - Yuan Ji (The University of Chicago)
- A scalable algorithm for Bayesian variable selection (SAB) with application to miRNA-mRNA regulation in cancer
  - Feng Liang (University of Illinois at Urbana-Champaign)
Keynote Speakers

Walter W. Offen, PhD

Distinguished Research Fellow
Global Head of Statistical Innovation and Safety Statistics
Acting Global Head of Statistics
AbbVie Inc.

Walt is currently Distinguished Research Fellow and the Global Head of Statistical Innovation and Safety Statistics at AbbVie. He received his PhD in statistics from the University of Florida in 1980. Upon graduation he joined Eli Lilly and Company in Indianapolis, where he remained for 31½ years. He joined Abbott in April, 2012, and stayed with AbbVie when the pharmaceutical business separated from the rest of Abbott in January, 2013.

His interests include novel clinical trial design and analysis, Data Monitoring Committees, and multiplicity.

Walt has been very active in professional activities. He has chaired a number of workshops and conferences, including the 7th Annual FDA/DIA Statistics Forum in April, 2013, and again in April, 2014. In the past few years he has served on panels at key statistical and scientific meetings on the topic of multiplicity of endpoints and enrichment strategies. Walt was inducted as a Fellow of the American Statistical Association in 2007.

Wei Shen, PhD

Senior Director
Global Statistical Sciences
Eli Lilly and Company

Dr. Wei Shen received his PhD in biostatistics from University of Minnesota in 1996. He joined Eli Lilly and Company in 1996 as a senior statistician. During his career at Lilly, Dr. Shen has provided extensive statistical support and leadership for drug development, registration and post-marketing.

Currently, Dr. Shen is Senior Director, Global Statistical Sciences, responsible for leading over 50 statisticians supporting clinical trials, advanced analytics, health outcomes research, and safety surveillance.

Dr. Shen has published over 40 manuscripts on statistical methodology and medical research.

Dr. Shen is President of the International Chinese Statistical Association (ICSA), and an associate editor for Journal of Biopharmaceutical Statistics.
Abstracts

Session 1: Personalized medicine
Organizer: Haoda Fu (Eli Lilly and Company)

**Enrichment Design with patient population augmentation**
*Lanju Zhang (AbbVie Inc)*

**Abstract:** Clinical trials can be enriched on subpopulations that may be more responsive to treatments to improve the chance of trial success. In 2012 FDA issued a draft guidance to facilitate enrichment design, where it pointed out the uncertainty on the subpopulation classification and on the treatment effect outside of the identified subpopulation. We consider a novel design strategy where the identified subpopulation (biomarker-positive) is augmented by some biomarker-negative patients. Specifically, after sufficiently powering biomarker-positive subpopulation we propose to enroll biomarker-negative patients, enough to assess the overall treatment benefit. We derive a weighted statistic for this assessment, correcting for the disproportionality of biomarker-positive and biomarker-negative subpopulations under enriched trial setting. Screening information is utilized for weight determination. This statistic is an unbiased estimate of the overall treatment effect as that in all-comer trials, and is the basis to power for the overall treatment effect. For analysis, testing will be first performed on biomarker-positive subpopulation; only if treatment benefit is established in this subpopulation will overall treatment effect be tested using the weighted statistic.

This design approach differs from typical enrichment design or stratified all-comer design in that the former enrolls only biomarker-positive patients and the latter enrolls a regular all-comer population. It also differs from adaptive enrichment by maintaining the trial design and analysis priority on biomarker-positive subpopulation. Therefore the proposed approach not only warrants a high probability of trial success on biomarker-positive subpopulation, but also efficiently assesses the overall treatment effect in the presence of an uncertain treatment benefit among biomarker-negative patients.

**GUIDE the search of tailoring biomarkers**
*Micahel Man (Eli Lilly and Company)*

**Abstract:** Identifying predictive biomarker for patient tailoring is becoming more and more important for drug development. Multi-marker approaches have become popular to complement single-marker analyses. Recently developed tree-based methods are particularly suitable for finding predictive biomarker in randomized trials with large number of candidate markers. A few examples will be provided to illustrate the GUIDE methodology. Challenges and future direction will also be discussed.

**Searching for optimal treatment rule – a new look on personalized medicine and subgroup identification**
*Haoda Fu (Eli Lilly and Company)*

**Abstract:** With new treatments and novel technology available, personalized medicine has become an important piece in the new era of medical product development. Traditional statistics methods for personalized medicine and subgroup identification primarily focus on single treatment or two arm randomized control trials. Motivated by the recent development of outcome weighted learning framework, we propose an alternative algorithm to search treatment assignments which connect with subgroup identification problems. Our method focuses on
applications from clinical trials to generate easy to interpret results. This framework is able to handle two or more than two treatments from both randomized control trials and observation studies. We implement our algorithm in C++, and connect it with R. The performance is evaluated by simulations, and we apply our method to a dataset from a diabetes study.

**Session 2: Challenges in high-dimensional data**

*Organizer: Annie Qu (University of Illinois at Urbana-Champaign)*

<table>
<thead>
<tr>
<th>Regularized estimation of linear functional for high-dimensional time series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xiaohui Chen (University of Illinois at Urbana-Champaign)</td>
</tr>
</tbody>
</table>

**Abstract:** We consider regularized estimation of linear functionals of high-dimensional linear processes. Our framework covers the broad regime from i.i.d. samples to long-range dependent time series and from sub-Gaussian innovations to those with mild polynomial moments. We show that the regularization parameter and the rate of convergence depend on the degree of temporal dependence and the moment conditions in a subtle way. The regularized estimator is demonstrated on large-scale sparse Markowitz portfolio allocation, optimal linear estimation and prediction for time series, where the asymptotic optimality is established. New sharp concentration inequalities of quadratic forms under dependence are established and they can be applied to other high-dimensional estimation problems for the second-order structures of time series data, such as characterizing the spectral and Frobenius norm rates of convergence for the hard- and soft-thresholded covariance matrix estimates. In addition, asymptotic rate of the graphical Lasso estimate of precision matrices is derived for Gaussian processes. The effect of long range dependence and distributional properties of the linear processes is illustrated on the simulated data. Finally, these second-order estimation procedures are applied to classify the cognitive states for a real fMRI dataset.

<table>
<thead>
<tr>
<th>Selection by partitioning the solution path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peng Wang (University of Cincinnati)</td>
</tr>
</tbody>
</table>

**Abstract:** The performances of penalized likelihood approaches profoundly depend on the selection of the tuning parameter; however there has not been a common agreement on the criterion for choosing the tuning parameter. Moreover, penalized likelihood estimation based on a single value of the tuning parameter would suffer from several drawbacks. This project introduces a novel approach for feature selection based on the whole solution paths rather than choosing one single tuning parameter, which significantly improves the selection accuracy. Moreover, it allows for feature selection using ridge or other strictly convex penalties. The key idea is to classify the variables as relevant or irrelevant at each tuning parameter and then select all the variables which have been classified as relevant at least once. We establish the theoretical properties of the method, and illustrate the advantages of the proposed approach with simulation studies and a data example.

<table>
<thead>
<tr>
<th>An augmented ADMM algorithm with application to high-dimensional statistical estimation problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yunzhang Zhu (Ohio State University)</td>
</tr>
</tbody>
</table>

**Abstract:** In this talk, I will present a fast and stable algorithm for solving a class of linearly regularized statistical estimation problem. This type of problems arises in many statistical estimation procedures, such as high-dimensional linear regression with fused lasso regularization, convex clustering, and trend filtering, among others. We propose a so-called augmented alternating direction methods of multipliers (ADMM) algorithm to solve this class of problems. As compared to a standard ADMM algorithm, our proposal significantly reduces the amount of computation at each iteration while maintaining the same overall rate of convergence. We demonstrate the superior performance of the augmented ADMM algorithm on a generalized lasso problem. Finally, we discuss a possible extension and some interesting connections to two well-known algorithms in imaging.
## Session 3: Causal inference and observational studies

### Organizer: Judith Xu (Takeda)

#### Indirect treatment comparisons

**Shawn Yu (Takeda)**

**Abstract:** In the past 5 years, the number of publications related to indirect treatment comparisons has increased significantly. More and more Health Technology Agencies (HTA), payers and decision makers are using results from indirect treatment comparisons in assessing comparative effectiveness of medicines and devices. This presentation is aimed at statisticians new to this research area to facilitate their general understanding of indirect treatment comparisons, more from a statistical perspective. Basic concepts and underlying statistical principles will be introduced. Major types of analysis methods and quality control measures will be described. Some applications of the methods and a few case studies will be covered.

#### Bayesian sensitivity analysis to assess the impact of unmeasured confounding in comparative observational study

**Xiang Zhang (Eli Lilly and Company)**

**Abstract:** The use of retrospective observational research as a tool for medical decision making, particularly with data from healthcare claims database and electronic medical records, has been growing in recent years. However, the use of such observational data for comparative effectiveness is challenged by selection bias and potential for unmeasured confounding. Unfortunately, the quantitative assessment of the potential influence of unmeasured confounders in observational data analysis is rare, despite the reliance of the validity of any cohort comparison on the "no unmeasured confounders" assumption. This presentation aims to introduce a new statistical method- Bayesian twin regression modeling- to address this challenge. Concepts, statistical principles, and simulation studies will be presented.

## Session 4: Longitudinal and survival analysis

### Organizer: Peter Song (University of Michigan)

#### Conditional modeling of longitudinal data with terminal event

**Shengchun Kong (Purdue University)**

**Abstract:** We consider a stochastic random effects model for longitudinal data with the occurrence of an informative terminal event that is subject to right censoring. Existing methods for analyzing such data include the joint modeling approach using latent frailty and the marginal estimating equation approach using inverse probability weighting, and in both cases the effect of terminal event on the response variable is inexplicit thus not easily interpretable. In contrary, we explicitly model the terminal event time as a covariate, which provides a straightforward interpretation while keeps the usual relationship of interest between the longitudinally measured response variable and covariates when time is far from the occurrence of the terminal event.
A two-stage semiparametric likelihood-based approach is proposed for estimating the regression parameters, where the conditional distribution of the right-censored terminal event time given other covariates is estimated prior to maximizing the likelihood function for the regression parameters. The method is illustrated by numerical simulations and the end-stage renal disease medical cost data. Desirable asymptotic properties are provided.

Joint Frailty Models for Zero-Inflated Recurrent Events in the Presence of a Terminal Event
Lei Liu (Northwestern University)

Abstract: Recurrent event data arise frequently in longitudinal medical studies. In many situations, there are a large portion of subjects without any recurrent events, manifesting the "zero-inflated" nature of the data. Some of the zero events may be "structural zeros" as patients are unsusceptible to recurrent events, while others are "random zeros" due to censoring before any recurrent events. On the other hand, there often exists a terminal event which may be correlated with the recurrent events. In this paper, we propose two joint frailty models for zero-inflated recurrent events in the presence of a terminal event, combining a logistic model for "structural zero" status (Yes/No) and a joint frailty proportional hazards model for recurrent and terminal event times. The models can be fitted conveniently in SAS Proc NLMIXED. We apply the methods to model recurrent opportunistic diseases in the presence of death in an AIDS study, and tumor recurrences and a terminal event in a sarcoma study.

Mixtures of g-priors for hypothesis testing in ANOVA models with a diverging number of parameters
Min Wang (Michigan Technological University)

Abstract: We consider Bayes approaches for the hypothesis testing problem in the multi-way analysis-of-variance (ANOVA) models. With the help of the singular value decomposition, we reparameterize the models with constraints for uniqueness into a standard linear regression model without constraints. We then derive Bayes factors based on mixtures of g-priors and study their corresponding consistency issues under the scenario in which the model dimension grows with the sample size. It is shown that two commonly used hyper-priors for g yield inconsistent Bayes factors due to the presence of a small inconsistency region around the null hypothesis. This observation motivates us to propose a new class of suitable hyper-priors to circumstance this undesirable asymptotic behavior. Simulation studies on two-way ANOVA models are conducted to compare the performance of the proposed priors with that of other mixtures of g-priors in the literature.

Session 5: Novel statistical methods and considerations to meet modern oncology challenges
Organizer: Yijie Zhou (AbbVie Inc) Jun Zhao (AbbVie Inc)

Utilizing the relation between pCR and overall survival in trial design for neoadjuvant breast cancer therapies
Ming Zhu (AbbVie Inc)

Abstract: The United States Food and Drug Administration (FDA) initiated the Accelerated Approval Program to allow faster approval of drugs for serious conditions that fill an unmet medical need. In 2012, FDA issued a draft Guidance for Industry on the accelerated approval of breast cancer therapies based on the surrogate endpoint “pathologic complete response” (pCR). Following the accelerated approval, clinical benefit of the drug shall still be demonstrated by subsequent confirmatory trials with endpoint like overall survival (OS). In this research, we investigate the correlation between pCR and OS and its limit properties, under both exponential and Weibull distribution of survival time. Considering a conventional design of two-arm comparison, extensive simulation studies are performed to further evaluate
the characteristics of the relationship between pCR and OS. Results of this research finding may provide some insight and guidance on designing clinical trials utilizing pCR as the surrogate endpoint.

Assessing consistency in multi-regional non-inferiority oncology trials

Alan Rong (Astellas)

Abstract: Multi-regional Clinical Trials (MRCT) has become a norm in Oncology. However using MRCT to register new drugs in certain countries requires demonstrating that the regional treatment effect is consistent with the global effect. Most of the literature has focused on addressing this problem for superiority clinical trials, but methods for non-inferiority designs are needed as well. For these studies the objective is to show that the experimental drug is non-inferior to an existing standard treatment, which is a common trial setting in the Oncology field. The non-inferiority design adds complexity to the setting beginning with the definition of consistency. We utilize the fixed margin non-inferiority method and propose possible definitions for treatment consistency in Oncology trials. We develop methods to test for regional treatment consistency based on the proposed definitions and extend our methods to the design of future non-inferiority multi-regional clinical trials with a predefined country of interest. Our work is motivated by oncology clinical trials with participation from Asian countries including China and Japan.

Progression-free survival: how often you look matters

Jingyi Liu (Eli Lilly and Company)

Abstract: For oncology studies, OS remains to be the golden standard as the primary endpoint to evaluate the efficacy of an experimental treatment. However, Progression Free Survival (PFS) is increasingly used as the primary endpoint for phase III studies. Issues with analyses of PFS have been discussed, including role of central tumor image review, the assumption of right censoring, missing scan data etc.

Most recently, as the advance of the new oncology drug development, longer PFS duration is observed in patients in different tumor types, especially in a biomarker driven population. To reduce patients’ burden as well as potential image retaining cost, the tumor scan frequency is reduced during the conduct of the study, for example, every 6 month for the first 3 months then 12 weeks afterwards (LUX-Lung 6, 2014). The impact of the frequency of tumor assessments was explored by (Williams, 2002).

In this presentation, we will first lay out some mathematical framework to define Scan-interval Induced PFS (SI-PFS) and Flexible Scan-interval Induced PFS (FSI-PFS) and further explore the impact of changing of tumor scan frequency (switching point, frequency, assessment window) on operating statistics such as, power, size, median PFS and hazard ratio. Some mathematical properties and simulation results in different scenarios will be discussed. A real data example will also be provided. The talk will conclude with some practical recommendations and potential future work.

Session 6: Bayesian and bioinformatics

Organizer: Yuan Ji (The University of Chicago)

MetaXcan: a scalable gene-level association test

Hae Kyung Im (The University of Chicago)

Abstract: Genome wide association studies have been successful in discovering thousands of genetic variants robustly associated
with complex disease and traits but the mechanism underlying most of these discoveries are still not well understood. Given the growing consensus that regulation of gene expression levels may be mediating a large portion of the variability in phenotypes, we have proposed PrediXcan, which integrates reference transcriptome datasets with GWAS studies to prioritize genes that are likely to be in the causal pathway. PrediXcan takes genetic variation data and predicts expression levels based solely on genetic variation data. These predicted levels of expression are correlated with the phenotype of interest to assess the significance of the association. This is a gene based association test that leverages mechanism to improve the power to discover causal genes.

In this talk, I will describe MetaXcan, an extension of PrediXcan, that uses summary statistics data, such as generated by GWAS or meta analysis studies, instead of individual level data. MetaXcan, provides a fast and scalable approach to compute gene based association test without the need to assemble huge datasets and can be applied to many of the large scale meta analysis results currently available for the scientific community.

<table>
<thead>
<tr>
<th>Zodiac: a comprehensive depiction of genetic interactions in cancer using TCGA data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yuan Ji (The University of Chicago)</td>
</tr>
<tr>
<td><strong>Abstract:</strong> The Cancer Genomes Atlas (TCGA) data are unique in that multimodal measurements across genomics features, such as copy number, DNA methylation, and gene expression, are obtained on matched tumor samples. The multimodality provides an unprecedented opportunity to investigate the interplay of these features. Graphical models are powerful tools for this task that address the interaction of any two features in the presence of others, while traditional correlation- or regression-based models cannot. We introduce Zodiac, an online resource consisting of a large database containing nearly 200 million interaction networks of multiple genomics features produced by applying novel Bayesian graphical models on TCGA data through massively parallel computation. Setting a new way of integrating TCGA data, Zodiac, publically available at <a href="http://www.compgenome.org/ZODIAC">www.compgenome.org/ZODIAC</a><a href="http://www.compgenome.org/ZODIAC">http://www.compgenome.org/ZODIAC</a>, is expected to facilitate the generation of new knowledge and hypotheses by the community.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A scalable algorithm for Bayesian variable selection (SAB) with application to miRNA-mRNA regulation in cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feng Liang (University of Illinois at Urbana-Champaign)</td>
</tr>
<tr>
<td><strong>Abstract:</strong> We propose a new computational framework for Bayesian variable selection. The key idea is to seek an approximation of the posterior distribution via an optimization procedure. The classical EM algorithm that gives the MAP estimate and the variational Bayes algorithm (VB) are special cases of this framework. A main feature of our algorithm (SAB) is its scalability. It is very fast in handling high dimensional data with large p and small n. We show that SAB achieves asymptotic consistency, albeit an approximation procedure. To address a critical biological problem, we apply SAB to a state-of-art cancer genomics data set with a goal to understand the complex regulatory relationship between miRNAs and mRNAs in cancer.</td>
</tr>
</tbody>
</table>
AbbVie 2016 Internship

Organization Name: AbbVie
Location: Lake County – North Chicago, IL
Number of Positions: Multiple
Type of Student: Graduate – 3rd - 4th year into PhD program
Deadline for Applying: January 31, 2016

Brief Description of Work:
The Data and Statistical Sciences (DSS) organization in AbbVie R&D anticipates having internship positions available for the summer of 2016. Internships will begin in May/June and are typically 10 – 12 weeks in duration, with housing provided to eligible applicants. During this time you will be provided with practical “hands-on” experience and given an opportunity to build your understanding of the pharmaceutical industry and of AbbVie. The successful candidate will be assigned specific projects within the DSS organization and work under the guidance of a senior statistician to perform statistical analyses and/or conduct statistical research in areas of interest to the organization (e.g., genomics, adaptive/novel design strategies in drug development, clinical trial simulation, benefit-risk assessment, etc.). The successful candidate will have the opportunity to participate in statistical seminars and workshops. At the conclusion of the internship, the intern is expected to give a presentation summarizing some aspect of their work while at AbbVie.

Other Relevant Information:

Basic Requirements:
- Currently enrolled in a graduate-level curriculum leading to a PhD in Statistics, Biostatistics, Bioinformatics or highly-related field
- Completed at least two full years of graduate study prior to the start of their internship
- GPA: 3.5/4.0 Be in good academic standing within their graduate program and overall at their university
- Track record of accomplishment
- Authorized to work in the U.S.
- Continue to be enrolled in graduate school for at least one semester following their internship.

Preferred Requirements:
- Authorized to work in the U.S.
- Minimum of one internship
- Strong working knowledge of SAS and/or R.
- Excellent problem-solving skills
- Exceptional interpersonal, communications, leadership and project management skills
- Proven track record of teamwork, adaptability, innovation, initiative and integrity

Recruitment Process:
We participate in several on-campus events (such as: information sessions, career fairs and conduct interviews while on campus). Interested candidates, please check the career services schedules at your school for the recruiting event schedule. AbbVie also posts our Science Intern positions on our career website, if your school isn’t listed, please visit www.abbviecareers.com.

Name and Address of Person to Contact:
DORTH KORST, Postdoctoral Program Manager
AbbVie
V32W, AP 51-1
1 North Waukegan Road
North Chicago, IL 60064
OFFICE +1 847-938-3599
FAX +1 847-937-4464
EMAIL dorth.korst@abbvie.com