The 13th Annual

SYMPOSIUM on STATISTICS in PSYCHIATRY

In Honor of

Thomas R. Ten Have, PhD, MPH

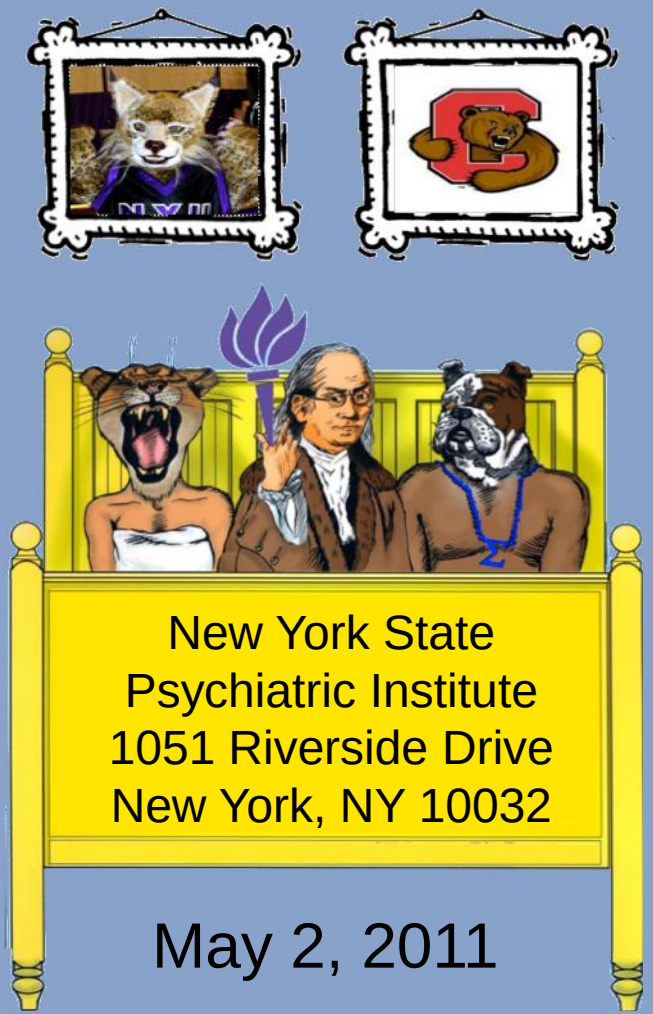
1051 Riverside Drive
New York, NY 10032
May 2, 2011

Hosted by the Division of Biostatistics in the
Department of Psychiatry at Columbia University
and New York State Psychiatric Institute

The 13th Annual

SYMPOSIUM on STATISTICS in PSYCHIATRY

Hosted by the Division of Biostatistics in the Department of Psychiatry at Columbia University and New York State Psychiatric Institute



Keynote Presentation
Xiao-Li Meng, Ph.D.

Harvard University

Mental Exercises for a Mental Health Study:

Is this a Simpson's Paradox?

ABSTRACT: Collaborative Psychiatric Epidemiology Surveys (CPES) offer a rich set of data for advancing our understanding of the complexity of mental disorders among the general population, especially minority groups. It also helps to advance our understanding of complex statistical conceptual and methodological issues, such as whether "equal" equals "fair" when assessing mental health service disparity and why dealing with Simpson's paradox is fundamentally dealing with bias-variance trade-off. This talk invites the audience---via an interactive "clicker" system---to enjoy a few mental exercises encountered in my involvement as a statistical consultant for analyzing CPES, demonstrating the need and benefit of critical thinking in statistical practice.

Lassell (Feihan) Lu, Ph.D.

New York University

Comparative Study of Variable Selection Methods in the Context of Developing Screening Instruments for Mental Disorders

ABSTRACT: The process of developing diagnostic or screening instruments for mental disorders involves item selection from, possibly, a very large pool of items from existing psychiatry questionnaires. A desirable screening instrument should consist of only a few important items and should accurately distinguish between cases and controls. Variable selection methods, such as t-test, classification trees, regularized linear models (e.g. Lasso, Elastic Net) and Random Forest can be used in developing such screening tools. Via a simulation study, the effect of various factors on the performance of those methods is investigated. Factors we consider are the correlations between the items; missing values for some items; the amount of information about diagnosis contained in the pool of items; the true relationship between the items and diagnosis; and the prevalence of the diagnosis in the sample. The results demonstrate an advantage of regularized linear models in variable selection, as well as the benefit of Random Forest in predicting future data. It is also shown that imputation of missing value via Random Forest followed by Lasso is an easy, yet efficient approach when there are missing data, and that t-test is a good pre-screening method. The variables selection methods are applied for the development of a short screening instrument for autism based on items from the currently used Autism Diagnostic Instrument-Revised version (ADI-R).

Knashawn H. Morales, Sc.D.

University of Pennsylvania

How to Know What You Don't Know:

Identifying explained and unexplained variation in multilevel models

ABSTRACT: Many studies now use statistical models designed to account for multilevel or clustered data. Answers to important questions related to explained and unexplained variation at the different levels of these models are rarely reported. We describe how to calculate how much of the total variance is accounted for by each level of clustering and what proportion of the variance at each level is accounted for by the included variables. This work is motivated by data from a multilevel study of characteristics associated with psychotropic medication use among Medicaid-enrolled children with autism. We use fixed effects models to calculate the total possible explainable variation at each level while accounting for all unmeasured factors related to outcome. Then, within each level, we calculate the proportion of that variance explained by included characteristics at each level. These results are compared to methods based on the random effects model which ignores the impact of unmeasured confounding.

Poster Session

Characterizing heterogeneity: an application of principal point classification to Autism data, *Jing Wang and Eva Petkova* from New York University

Meta-regression and the ecological fallacy in depression treatment studies, *Thaddeus Tarpey, Eva Petkova, Lei Huang and Liping Deng* from New York University

Lower-dimensional approximation for sparse functional data with an application to screening pubertal growth paths, *Wenfei Zhang* from New York University

Comparing alternative diagnostic criteria sets using Item Response Theory (IRT), *Melanie M. Wall, Deborah S. Hasin* from Columbia University

Theory of Self-learning Q-Matrix, *Jingchen Liu, Gongjun Xu, and Zhiliang Ying* from Columbia University

Prediction-Based Structured Variable Selection through the Receiver Operating Characteristic Curves, *Huaihou Chen, Yuanjia Wang, Naihua Duan, Runze Li, Roberto Lewis-Fernández* from Columbia University

The Poisson-Normal Model as an Alternative Approach for Analyzing Psychopharmacologic Challenge Study Data, *B. Pittman, J.H. Krystal, R. Gueorguieva* from Yale University

Yongtao Guan, Ph.D.

Yale University

On Measurement Error Problems with Predictors Derived from Stationary Stochastic Processes and Application to Cocaine Dependence Treatment Data

ABSTRACT: In a cocaine dependence treatment study, we use linear and nonlinear regression models to model post-treatment cocaine craving scores and first cocaine relapse time. A subset of the covariates are summary statistics derived from baseline daily cocaine use trajectories, such as baseline cocaine use frequency and average daily use amount. These summary statistics are subject to estimation error and can therefore cause biased estimators for the regression coefficients. Unlike classical measurement error problems, the error we encounter here is heteroscedastic with an unknown distribution, and there are no replicates for the error-prone variables or instrumental variables. We propose two robust methods to correct for the bias: a computationally efficient method-of-moment-based method for linear regression models and a subsampling extrapolation method that is generally applicable to both linear and nonlinear regression models.Simulations and an application to the cocaine dependence treatment data are used to illustrate the efficacy of the proposed methods. Asymptotic theory and variance estimation for the proposed subsampling extrapolation method and some additional simulation results are described in the online web supplementary material.

Directions

The New York State Psychiatric Institute is located at 1051Riverside Drive, between 165th and 169th Street. We also have locations at the Kolb Annex Building, located at 40 Haven Avenue and 168th Street. You may enter through the Kolb Annex to get to the Psychiatric Institute building.

Visitors are reminded to bring a valid ID to gain entrance to the building.

The symposium will be held in the Auditorium on the 1st floor.

No visitor parking available. There is a parking garage on Fort Washington Avenue and 165th Street (payment required).

INFO: Kelly Roska (212) 543-5589 or kr2343@columbia.edu

Keynote Presentation
Andrew C. Leon, Ph.D.

Weill Cornell Medical College

Evolution of Psychopharmacology Trial Design and Analysis:

Six Decades in the Making

ABSTRACT: The clinical trial methodology used to evaluate psychopharmacologic agents has evolved considerably over the past six decades. The first and most productive decade was characterized by case series, each with a small number of patients. These trials used non-standardized clinical observation as outcomes and seldom had a comparison group. The crossover design became widely used to examine acute psychiatric treatments in the 1950s and 1960s. Although this strategy provided comparison data, it introduced problems in study implementation and interpretation. In 1962, the FDA began to require "substantial evidence of effectiveness from adequate and well-controlled studies". Subsequent decades saw remarkable advances in clinical trial design, assessment, and statistical analyses. Standardized instruments were developed and double-blinding and placebo controls became the benchmark. Sample sizes increased and data analytic procedures were developed that could accommodate the problems of attrition. Randomized withdrawal designs examined maintenance therapies in the 1970s. Ethical principles for research became codified in the US at that time. A wave of regulatory approvals of novel antipsychotics, antidepressants and anticonvulsants came in the 1980s and 1990s, each based on data from randomized double-blind parallel group, placebo-controlled clinical trials. These trial designs often involved fixed-dose comparisons based, in part, on a greater appreciation that much of the benefit and harm in psychopharmacology was dose related. Despite this progress in design, discovery of new mechanisms of action and blockbuster interventions has slowed during the past decade. The evolution of design and analysis during the lifespan of psychopharmacology is examined here to understand this phenomenon.

Scientific Program

11:00-11:30	**Honoring of Thomas R. Ten Have, Ph.D, M.P.H.**
11:30-12:20	**KEYNOTE PRESENTATION**
	Xiao-Li Meng, Ph.D.
	Mental Exercises for a Mental Health Study: Is this a Simpson's Paradox?
12:20-12:30	**Discussion**
12:30- 1:30	**Lunch and Poster Session**
1:30- 2:05	**Lassell (Feihan) Lu, Ph.D.**
	Comparative Study of Variable Selection Methods in the Context of Developing Screening Instruments forMental Disorders
2:05- 2:15	**Discussion**
2:15- 2:50	**Knashawn H. Morales, Sc.D.**
	How to Know What You Don't Know: Identifying explained and unexplained variation in multilevel models
2:50- 3:00	**Discussion**
3:00- 3:15	**Refreshment Break**
3:15- 3:50	**Yongtao Guan, Ph.D.**
	On Measurement Error Problems with Predictors Derived from Stationary Stochastic Processes and Application to Cocaine Dependence Treatment Data
3:50- 4:00	**Discussion**
4:00- 4:50	**KEYNOTE PRESENTATION**
	Andrew C. Leon, Ph.D.
	Evolution of Psychopharmacology Trial Design and Analysis: Six Decades in the Making
4:50- 5:00	**Discussion**
5:00- 5:30	**Poster Session**
5:30- 6:00	**Reception at Coogan's**

Program Committee Members

Naihua Duan, Ph.D. – Columbia University

Andrew C. Leon, Ph.D. – Weill Cornell Medical College

Haiqun Lin, Ph.D. – Yale University

Eva Petkova, Ph.D. – New York University

Justine Shults, Ph.D. – University of Pennsylvania

Thomas R. Ten Have, Ph.D, M.P.H. – University of Pennsylvania