The Fifth International Workshop in Sequential Methodologies 2015

June 22 to June 24, 2015 at Columbia University in the City of New York

Organizing Committee:

Nitis Mukhopadhyay

Alexander Tartakovsky

Zhiliang Ying

Local Organizers:

Yunxiao Chen (Chair)

Xiaoou Li

Jingchen Liu

Zhezhen Jin

Rohit Kumar Patra

Bodhisattva Sen

Haolei Weng

Hongzhong Zhang

Sponsors:

Department of Statistics, Columbia University in the City of New York

Institute of Mathematical Statistics

Monday morning, June 22

[9:00 - 9:15]

Opening Remarks

Speaker: David Madigan, Executive Vice President
and Dean of the Faculty of Arts and Sciences, Columbia University

Monday morning, June 22

[9:15 - 10:15]

Plenary Session

Title: Optimal Detection in Multi-stream Data

Speaker: Hock Peng Chan, National University of Singapore

Abstract:

Consider multi-stream data and the problem of detecting sequentially that at an unknown time in the past, a small fraction of the data-streams has undergone a change in distribution. There have been a lot of detailed investigation on this, and influential works include Tartakovsky & Veeravalli (2008), Mei (2010) and Xie & Siegmund (2013). In particular Xie & Siegmund (2013) paid attention to designing tests that down-weight the contributions of data-streams that do not undergo change of distribution, so that the signal contributions are not drowned out by the noise. We consider score-based test statistics similar to that suggested in Xie & Siegmund (2013). However unlike in Xie & Siegmund (2013), our score is motivated by a limit of detectability argument rather than on supposed distribution changes. That is we first determine what strength (and fraction) of signals are not detectable, and then aim to do as well as possible on the boundary of the detectable region. We show that our tests achieve minimum detection delay under the usual average run length constraint, extending the optimality results of Moustakides (1986) to multi-stream data. We discuss various extensions of our score-based tests, for example the use of window-limited and geometric-sized windows, as suggested in Lai (1995), to reduce the computational burden. We consider the situations of non-normal data and signal strengths/frequencies that are not uniform across data-streams. We also consider how Mei's sum of CUSUM test, which can work well when change of distributions do not occur simultaneously across data-stream, can be modified via an appropriate score function that down-weights the contributions from data-streams that did not go through change in distributions.

Monday morning, June 22

[11:00 - 12:30]

Tests and Estimation

Organizer: Nitis Mukhopadhyay, University of Connecticut

Chair: Seyed Yaser Samadi, Southern Illinois University

Title: **Empirical Likelihood Based Detection Procedure for Change Point in Mean Residual Life Functions under Random Censorship**

Speaker: Wei Ning, Bowling Green State University

Abstract:

Mean residual life function is an important function in survival analysis which describes the expected remaining lifetime given survival time up to a certain age. In this talk, I will propose a nonparametric method based on empirical likelihood to test the change point of mean residual life functions of independent random variables. Monte Carlo Simulations under different lifetime settings are carried out to show the power performance. Two real data examples are analyzed to illustrate the testing procedure.

Title: **Confidence Interval Estimation Following SPRT in a Normal Distribution with Equal Mean and Variance**

Speaker: Debanjan Bhattacharjee, Utah Valley University

Abstract:

Confidence interval estimation following a sequential probability ratio test (SPRT) is an important and difficult problem (Siegmund 1978,1985). Considerable difficulties arise because following termination of SPRT, the customary estimator of an unknown parameter of interest, obtained from the randomly stopped sample (Anscombe 1952), is usually biased. As a result, the coverage probability of a naive confidence interval for the parameter based on the randomly stopped version of the customary estimator often falls below the target confidence-coefficient. In this article, we have addressed this problem for a normal distribution with equal mean and variance. Refer to Bhattacharjee and Mukhopadhyay (2011, 2012). This

problem is different from the confidence interval estimation for mean in a normal distribution where the mean and the variance are not explicitly related to each other (Siegmund 1978). We propose a methodology essentially based on the random central limit theorem due to Anscombe (1952) which is remarkably easier to implement than Siegmund's original methodology. We explore bias corrections and resulting more accurate variance estimation in order to come up with practically useful and easily implementable methodologies to deliver achieved coverage probability that is very close to the nominal goal.

Title: **Composite Sequential Monte Carlo Test for Post-Market Vaccine Safety Surveillance**

Speaker: Ivair Silva, Federal University of Ouro Preto

Abstract:

Group sequential hypothesis testing is now widely used to analyze prospective data. If Monte Carlo simulation is used to construct the signaling threshold, the challenge is how to manage the Type I error probability for each one of the multiple tests without losing control on the overall significance level. This paper introduces a valid method for a true management of the alpha spending at each one of a sequence of Monte Carlo tests. The method also enables the use of a sequential simulation strategy for each Monte Carlo test, which is useful for saving computational execution time. Thus, the proposed procedure allows for sequential Monte Carlo test in sequential analysis, and this is the reason why it is called composite sequential test. An upper bound for the potential power losses from the proposed method is deduced. The composite sequential design is illustrated through an application for post-market vaccine safety surveillance data.

Acknowledgement: This research was funded by the National Institute of General Medical Sciences, grant #RO1GM108999, by Conselho Nacional de Desenvolvimento Científico e Tecnológico(CNPq), Brazil, and by Fundação de Amparo á Pesquisa do Estado de Minas Gerais (FAPEMIG).

Monday morning, June 22

[11:00 - 12:30]

Sequential Designs in Drug Development

Organizer: Vladimir Dragalin, Johnson & Johnson

Chair: Vladimir Dragalin, Johnson & Johnson

Title: Dose Expansion Cohorts in Phase I Trials

Speaker: Alexia Iasonos, Memorial Sloan Kettering Cancer Center

Abstract:

A rapidly increasing number of Phase I dose-finding studies, and in particular those based on the standard 3+3 design, frequently prolong the study and include dose ex- pansion cohorts (DEC) with the goal to better characterize the toxicity profiles of experimental agents and to study disease specific cohorts. These trials consist of two phases: the usual dose escalation phase that aims to establish the maximum tolerated dose (MTD) and the dose expansion phase that accrues additional patients, often with different eligibility criteria, and where additional information is being collected. We propose methodology that allows monitoring of safety in the DEC by re-evaluating the MTD in light of additional information. Our working assumption is that, regardless of the design being used for dose escalation, during the DEC we are experimenting in the neighborhood of a target dose with an acceptable rate of toxicity. We refine our initial estimate of the MTD by continuing experimentation in the immediate vicinity of the initial estimate of the MTD and we stop accrual based on efficacy data using sequential probability ratio tests. The auxiliary information provided in this evaluation can include toxicity, pharmacokinetic, efficacy or other endpoints. We consider approaches specifically focused on the aims of DEC, that examine efficacy alone or simultaneously with safety and compare the proposed tests via simulations.

Title: Simple Benchmark for Planning and Evaluating Complex Dose Finding Designs

Speaker: Ken Cheung, Columbia University

Abstract:

While a general goal of early phase clinical studies is to identify an acceptable dose for further investigation, modern dose finding studies and designs are highly specific to individual clinical settings. In addition, as outcome-adaptive methods often involve complex algorithm, it is crucial to have diagnostic tools at the planning stage to evaluate the plausibility of a method's simulated performance and the adequacy of the algorithm. In this talk, I will

introduce a simple technique that provides an upper limit, or a benchmark, of accuracy for dose finding methods for a given design objective. The proposed benchmark is nonparametric optimal, and is demonstrated by examples to be a practical accuracy upper bound for model-based dose finding methods. We illustrate the implementation of the technique in the context of phase I trials that consider multiple toxicities and phase I/II trials where dosing decisions are based on both toxicity and efficacy, and apply the benchmark to several clinical examples considered in the literature. By comparing the operating characteristics of a dose finding method to that of the benchmark, we can form quick initial assessments of whether the method is adequately calibrated and evaluate its sensitivity to the dose-outcome relationships.

Title: **Estimation after Population Enrichment Design**

Speaker: Vladimir Dragalin, Johnson & Johnson

Abstract:

Recent breakthroughs in identifying clinically actionable biomarkers allow a paradigm shift in clinical trial design and a focus on patient subpopulations with potentially greater treatment effect. Instead of limiting the enrollment only to the enriched population, prospectively specified adaptive designs enable the data-driven selection of one or more pre-specified subpopulations in an interim analysis and the confirmatory proof of efficacy in the selected subset at the end of the trial. Strong control of the familywise Type I error rate is guaranteed by combining closure principle and p-value combination tests. However, much less is done on the estimation of the treatment effect at the end of such trials. In this presentation, some adjustments to the nave maximum likelihood estimators are proposed and their properties investigated.

Monday morning, June 22

[11:00 - 12:30]

Recent Advances in Sequential Change Detection

Organizers: Georgios Fellouris, University of Illinois at Urbana-Champaign
Olympia Hadjiliadis, City University of New York

Chair: Georgios Fellouris, University of Illinois at Urbana-Champaign

Title: **Quickest Detection and Isolation of Line Outages in Power Systems**

Speaker: Venugopal Veeravalli, University of Illinois at Urbana-Champaign

Abstract:

A method is proposed to detect and identify power system transmission line outages in near real-time. The method exploits the statistical properties of the small random fluctuations in electricity generation and demand that a power system is subject to as time evolves. To detect and identify transmission line outages, a linearized incremental small-signal power system model is used in conjunction with high-speed synchronized voltage phase angle measurements obtained from phasor measurement units. By monitoring the statistical properties of voltage phase angle time-series, line outages are detected and identified using techniques borrowed from the theory of quickest change detection. As illustrated through case studies, the proposed method is effective in detecting and identifying single- and double-line outages in an accurate and timely fashion.

Title: **High-dimensional Change-point Detection: Kernel, Sketching, and Multi-sensor**

Speaker: Xie Yao, Georgia Institute of Technology

Abstract:

Detecting change-points from high-dimensional streaming data is a fundamental problem that arises in many big-data applications such as video and speech processing, sensor networks, social networks, and genomic signal processing. Challenges herein include developing algorithms that can provide reliable results over larger classes of data distributions, that have low computational complexity and good statistical power, and that can exploit structures to detecting weak signals. I will present three aspects of our recent work that tackle these challenges: (1) developing kernel-based methods for sequential change-point detection; and (2) using sketching to reduce data dimensionality; and (3) utilizing signal sparsity to achieve quicker detection of abrupt or gradual changes. We provide theoretical performance bounds and demonstrate the performance of the algorithms using simulated and real data.

Acknowledgement: This is joint work with Yang Cao, Shuang Li, and Le Song at Georgia Institute of Technology.

Title: **Minimax Quickest Changepoint Detection for Dependent Data**

Speaker: Sergey Pergamenchtchikov, University of Rouen

Abstract:

In this paper we consider the quickest change point detection for dependent data. We consider two classes of sequential detection procedures introduced by Lai and Tartakovsky for which the probability or conditional probability of a false alarm within a period of some fixed length is less than the given small probability irrespective of when the period starts. Respectively, for each of these classes we consider two minimax risks: absolute delay risk and conditional absolute delay risk. Under some conditions for the observations we show that the Shiryaev - Roberts procedure is optimal with respect to both these risks. We check these conditions for the linear time series: autoregression and autoregressive GARCH.

Acknowledgement: The research is funded by the grant of the Government of Russian Federation No.14.A.12.31.0007, by the Russian Science Fondation (research project No. 14-49-00079) and by the Mendeleiev Fondation of Russian National Tomsk State University. This is joint work with Alexander Tartakovsky, University of Connecticut, USA

Title: **Robustness of the N-CUSUM Stopping Rule in a Wiener Disorder Problem**

Speaker: Neofytos Rodosthenous, Queen Mary, University of London

Abstract:

We study a Wiener disorder problem of detecting the minimum of N change-points in N observation channels coupled by correlated noises. It is assumed that the observations in each dimension can have different strengths and that the change-points may differ from channel to channel. The objective is the quickest detection of the minimum of the N change-points. We adopt a min-max approach and consider an extended Lordens criterion, which is minimized subject to a constraint on the mean time to the first false alarm. It is seen that, under partial information of the post-change drifts and a general nonsingular stochastic correlation structure in the noises, the minimum of N cumulative sums (CUSUM) stopping rules is asymptotically optimal as the mean time to the first false alarm increases without bound.

Monday morning, June 22

[11:00 - 12:30]

Inference Problems

Organizer: Nitis Mukhopadhyay, University of Connecticut

Chair: Xiaoou Li, Columbia University

Title: Issues in Inference with Response-Adaptive Designs

Speaker: Nancy Flournoy, University of Missouri

Abstract:

Following most response-adaptive designs, maximum likelihood estimates are no longer normally distributed. This includes k-treatment studies with early stopping rules and adaptive treatment allocation probabilities in a dose-response context. Exceptions are MLEs after random walk allocation rules. The problem, some solutions, and some new directions for research will be discussed.

Title: Sequential Inference for the Risk Ratio and a Measure of Reduction

Speaker: Hokwon Cho, University of Nevada

Abstract:

We propose sequential methods for obtaining approximate confidence intervals and optimal sample sizes for the ratio of two independent binomial variates, often called the risk ratio (RR), and the measure of reduction (MOR). The procedure is developed based on a modified maximum likelihood estimator (MLE) for the ratio and we study its desirable properties - unbiasedness, variances and normality. First-order asymptotic expansions are obtained to investigate large-sample properties of the proposed procedure. Monte Carlo experiment is carried out for various scenarios of samples for examining the finite sample behavior. Also, through illustrations, we compare the performance of the proposed methods, Wald-based confidence intervals with the likelihood-ratio confidence intervals in light of invariance, length and sample sizes.

Title: Sequential Nonparametric Testing Using the Martingale Law of the Iterated Logarithm

Speaker: Aaditya Ramdas, Carnegie Mellon University

Abstract:

We consider the problem of nonparametric two-sample mean testing, where one has access to i.i.d. samples from two multivariate distributions and wishes to test whether they have the same mean. We propose a new *sequential* test for this problem in data-rich, memory-constrained situations, that is novel in several ways: (a) it takes linear time and constant space to compute on the fly and has robust high-dimensional statistical performance; (b) it has comparable statistical performance to a batch/offline version of the test with the same memory/computational constraints, notably a matching power guarantee, at the same false positive rate; (c) it has a distinct computational advantage over a batch test, because it accesses only as many samples as are required, i.e. its stopping time is adaptive to the (unknown) difficulty of the problem; (d) its derivation and analysis are rigorously finite-sample, using a uniform empirical Bernstein version of the martingale law of the iterated logarithm (LIL), which may be of independent interest. We demonstrate how to extend this idea to nonparametric homogeneity and independence testing, and believe that our ideas are even more broadly applicable.

Acknowledgement: This is joint work with Akshay Balsubramani.

Monday afternoon, June 22

[14:00 - 15:30]

Optimal Stopping with Multiple Exercising or Partial Information in Financial Engineering

Organizer: Hongzhong Zhang, Columbia University

Chair: Hongzhong Zhang, Columbia University

Title: **Optimal Double Stopping & Optimal Switching under Mean Reversion**

Speaker: Tim Leung, Columbia University

Abstract:

We analyze an optimal double stopping problem to determine the optimal times to buy and subsequently sell an asset whose price is driven by an exponential OU or CIR process. In addition, we analyze a related optimal switching problem that involves an infinite sequence of trades, and identify the conditions under which the double stopping and switching problems admit the same optimal entry and/or exit timing strategies. Among our results, we find that the investor generally enters when the price is low, but may also find it optimal to wait if the price is sufficiently close to zero, leading to a disconnected continuation region.

Title: **Optimal Mean Reversion Trading with Transaction Costs and Stop-loss Exit**

Speaker: Xin Li, Columbia University

Abstract:

Motivated by the industry practice of pairs trading, we study the problem of trading a mean-reverting price spread. An optimal double stopping problem is formulated to analyze the timing to start and subsequently liquidate the position subject to transaction costs and a stop-loss constraint. Modeling the price spread by an Ornstein-Uhlenbeck process, we apply a probabilistic methodology and rigorously derive the optimal price intervals for market entry and exit. Using our analytic solutions, we examine the impact of transaction costs and the stop-loss constraint.

Title: Insiders' Hedging in a Stochastic Volatility Model

Speaker: Kiseop Lee, University of Louisville

Abstract:

We study a market where there are traders with different levels of information. Insiders observe exclusive, non-public information that affects the volatility of the price process, and the information levels are different even among insiders. By the nature of information, some information processes are continuous, while others are discrete. We study the local risk minimization hedging strategies of the insiders under stochastic volatility models, and compare them with an honest trader's strategy. A numerical example is provided.

Title: Optimal Multiple Stopping with Negative Discount Rate and Random Refraction Times under Levy Models

Speaker: Hongzhong Zhang, Columbia University

Abstract:

We study an optimal multiple stopping problem driven by exponential Levy processes. Our model allows for a negative effective discount rate, which arises in a number of financial applications such as stock loans and real options. Moreover, successive exercise opportunities are separated by i.i.d. random refraction times. Under a wide class of Levy models, we rigorously show that the optimal strategy to exercise successive call options is uniquely characterized by a sequence of up-crossing hitting times.

Monday afternoon, June 22

[14:00 - 15:30]

Recent Advances in Sequential Trial Design, Analyses and Evaluation

Organizer: Yuanjia Wang, Columbia University

Chair: Yuanjia Wang, Columbia University

Title: Sequential Evaluation of Youden's Index of Diagnostic Tests with Binary Outcomes.

Speaker: Aiyi Liu, NIH/NICHD

Abstract:

TBA

Title: Sequential Multiple Assignment Randomization Trials with EnRichment (SMARTer) Design

Speaker: Donglin Zeng, University of North Carolina at Chapel Hill

Abstract:

Sequential multiple assignment randomization trial (SMART) is a powerful design to study Dynamic Treatment Regimes (DTRs) and allows causal comparisons of DTRs. Practical challenges that may occur in the implementation of SMARTs include: 1) high dropout rates require a large sample size in the initial stage to ensure sufficient power for comparing DTRs; 2) it is time-consuming and challenging to manage and monitor a sequential multi-stage trial with a large number of participants. To handle these issues, we propose a SMART with EnRichment (SMARTer) design, which can potentially improve the design efficiency, shorten the recruitment period, and reduce the trial duration to make SMART more practical with limited time and resource. Specifically, at each subsequent stage of a SMART, we enrich the study sample with new patients who have received previous stages' treatments in a naturalistic fashion without randomization, and only randomize them among the current stage treatment options. One extreme case of the (SMARTer) is to synthesize separate independent single-stage randomized trials with patients who have received previous stage treatments. We show data from (SMARTer) allows for unbiased estimation of DTRs as SMART does under certain assumptions. Furthermore, we show analytically that the efficiency gain of the new design over SMART can be significant especially when the dropout rate is high. Lastly, extensive simulation studies are performed to demonstrate performance of (SMARTer) design, and sample size estimation in a scenario informed by real data from a SMART study is presented.

Title: An Enhanced Approach to Analyzing the Sequential Parallel Comparison Design/Doubly Randomized Delayed Start (SPCD/DRDS) Design

Speaker: Pilar Lim, Jansen Research & Development

Abstract:

High placebo response rates have been a major impediment in neuroscience clinical trials and the effects have been difficult to overcome in developing new compounds. It has led to high failure rates in mood disorder trials, among others, even in trials with known effective drugs. Current approaches to the analysis of the SPCD/DRDS design do not assess the consistency of treatment effect across the sequence of periods and do not provide an appropriate estimate of the treatment effect. To address this issue, this talk will discuss insights into assessing consistency and providing an assessment of the treatment effect in the SPCD/DRDS design. Results from a recently completed mood disorder clinical trial will be presented.

Title: Hybrid Clinical Trials for Assessing Efficacy and Pragmatic Aims

Speaker: Christine Mauro, Columbia University

Abstract:

Complicated grief is a psychiatric disorder that affects many patients for long durations, with substantial impact on quality of life, functioning, and productivity. A psychotherapy, complicated grief treatment (CGT) has been shown to be efficacious in two previous randomized clinical trials. However there is an important need to assess efficacy for medication (citalopram), either as a stand-alone treatment, or in conjunction with CGT. Of primary interest is comparing: Aim 1. Citalopram vs. Placebo (without CGT, efficacy aim) and Aim 2. Citalopram + CGT vs. Placebo + CGT (efficacy aim). Of secondary interest is to assess benefits for CGT under naturalistic conditions, especially taking into consideration medication status, that is, Aim 3. Citalopram + CGT vs. Citalopram + no CGT (pragmatic aim) and Aim 4. CGT vs. no CGT (no Citalopram, pragmatic aim). Finally, there is also interest in understanding the interaction between Citalopram vs. Placebo vs. CGT vs. no CGT. In order to efficiently address all of these aims, a 2x2 factorial randomized trial to assess both efficacy and pragmatic aims was designed, implemented and is now completed. In this talk, we present analyses methods for examining the primary aims and new subgroup analysis methods to examine qualitative interaction.

Monday afternoon, June 22

[14:00 - 15:30]

Sequential Methods in Educational Testing

Organizer: Georgios Fellouris, University of Illinois Urbana-Champaign

Chair: Yunxiao Chen, Columbia University

Title: On Optimal Sequential Designs for Computerized Adaptive Testing

Speaker: Xiaoou Li, Columbia University

Abstract:

Computerized adaptive tests are tests whose items are tailored to examinees individual ability level. By selecting tests items appropriately according to the examinees previous responses, the classifications of the examinees can be made once sufficient information is cumulated. We investigate properties of optimal adaptive sequential designs for mastery tests that have minimal Bayes risk and provide numerical approximation.

Title: Sequential Design for Computerized Adaptive Testing that Allows for Response Revision

Speaker: Shiyu Wang, University of Illinois Urbana-Champaign

Abstract:

In computerized adaptive testing (CAT), items (questions) are selected sequentially, in the sense that the next item depends on the already observed responses. However, CAT has been heavily criticized for not allowing test-takers to review and revise their answers. In this talk, we propose a novel, flexible CAT design in which the examinee is allowed to revise the response to a previous item at any time during the test, as long as this is an item with $m > 2$ categories that has been revised at most $m - 2$ times. Assuming that both first attempts and revisions are governed by the (polytomous) nominal response model, we estimate adaptively the examinees ability using a partial likelihood estimator and we show that the resulting estimator is always strongly consistent as the number of administered items goes to infinity. Moreover, we show that such estimator is asymptotically normal when the items are selected to maximize the Fisher information at the current ability estimate and the number of revisions is much smaller than the number of items. Finally, we present a simulation study that is based on a discrete item pool and in which we compare the estimator of the proposed design with that of a regular CAT design. This is joint work with Georgios Fellouris (UIUC) and Hua-Hua Chang (UIUC).

Title: Sequential Change-Point Detection Procedures for Efficient Detection of Learning in Restricted Latent Class Models

Speaker: Jeff Douglas, University of Illinois Urbana-Champaign

Abstract:

TBA

Title: Some Sequential Procedures for Adaptive Mastery Testing

Speaker: Jay Bartroff, University of Southern California

Abstract:

Adaptive mastery tests are educational or psychological evaluations to assess whether an examinee's latent trait (e.g., ability) exceeds a given threshold. The fact that many of these tests are implemented via computer or handheld device affords wide adaptation and thus makes this fertile ground for applying sequential statistical methods. I will discuss two such applications: (1) sequential generalized likelihood ratios for improving performance of hypothesis based tests, and (2) stochastically curtailed confidence interval based tests. The latter involves computing the conditional probability that the latent trait's lower confidence bound will not cross the threshold before the end of the exam, and the performance of this method is improved by a slight extension of a result of Chang & Ying (Ann. Stat., 2009) for the conditional distribution of the sequentially designed modified MLE. Various parts of this work are joint with Matthew Finkelman, T. L. Lai, Haskell Sie, and Nathan Thompson.

Monday afternoon, June 22

[14:00 - 15:30]

Celebration of Alexander Tartakovsky's 60th Birthday

Session Description:

Organized by the former students and postdocs of Alexander Tartakovsky, this informal session brings together a diverse audience of colleagues and long-term collaborators of Dr. Tartakovsky to celebrate his 60th birthday by honoring his diverse and deep contributions to the theory and applications of change-point detection and sequential methodologies. Session participants will acknowledge Dr. Tartakovsky's work and give brief remarks on the occasion of his 60th anniversary.

Organizers: Georgios Fellouris, University of Illinois at Urbana-Champaign
Aleksey S. Polunchenko, State University of New York at Binghamton
Vasanthan Raghavan, Qualcomm Flarion Technologies
Grigory Sokolov, State University of New York at Binghamton

Chair: Aleksey S. Polunchenko, State University of New York at Binghamton

Speakers: Georgios Fellouris, University of Illinois at Urbana-Champaign
Olympia Hadjiliadis, City University of New York
Nitis Mukhopadhyay, University of Connecticut at Storrs
George V. Moustakides, University of Patras
Aleksey S. Polunchenko, State University of New York at Binghamton
Venugopal V. Veeravalli, University of Illinois at Urbana-Champaign

Monday afternoon, June 22

[15:50 - 17:20]

Tribute to Alexander Novikov: Sequential Tests and Estimation

Organizers: George Moustakides, University of Patras
Alexander Tartakovsky

Chair: Vlad Dragalin, Johnson & Johnson

Title: **Sequential Tests of Composite Hypotheses and Quickest Change Detection: Overview of Novikov's Contributions and Recent Results**

Speaker: Alexander Tartakovsky

Abstract:

This talk will consist of three parts. In the first part I will review Professor Novikov's general contributions to Probability, Statistics, and Financial Mathematics. In the second part I will consider his contributions to Sequential Analysis and Optimal Stopping in detail, with special emphasis on Multihypothesis Sequential Tests of Composite Hypotheses and, if time permits, some aspects of quickest changepoint detection. Finally, in the third part, possible extensions of sequential testing composite hypotheses for very general non-i.i.d. models will

be addressed. In particular I will show that a multihypothesis sequential generalized likelihood ratio test and a multihypothesis version of the Robbins-Siegmund adaptive sequential test are asymptotically optimal minimizing the expected sample size or, even more generally, higher moments of the stopping time distribution under very general conditions related to the strong law of large numbers for the log-likelihood ratio processes.

Title: Sequential Estimation for Stochastic Differential and Difference Equations

Speaker: Sergey Pergamenchtchikov, University of Rouen

Abstract:

We consider the sequential estimation problem for stochastic nonlinear differential equations observed in the discrete times. We construct a sequential kernel estimator based on the Novikov idea proposed for the parametric sequential estimation of the diffusion processes. In this paper we develop the Novikov approach for the nonparametric sequential estimation and we propose some truncated version of the sequential kernel estimation procedure. We obtain a non asymptotic upper bound for the pointwise absolute error risk of proposed procedure. The optimal convergence rate and a sharp constant in the bounds are found for the asymptotic minimax risk. As a consequence, we show that the proposed sequential procedure is efficient.

Acknowledgement:
The research is funded by the grant of the Government of Russian Federation No.14.A.12.31.0007, by the Russian Science Fondation (research project No. 14-49-00079) and by the Mendeleiev Fondation of Russian National Tomsk State University. This is joint work with Leonid Galtchouk, Tomsk State University, Tomsk, Russia

Title: Optimum Sequential Detection of Transitory Changes

Speaker: George Moustakides, University of Patras

Abstract:

We are interested in the sequential detection of a change in the statistics of a random process. Specifically we consider changes that are not abrupt but gradual. In other words the change exhibits a transitory phase before reaching its steady state form. Adopting the classical worst-case conditional detection delay proposed by Lorden as our performance measure and

constraining the average false alarm period, we derive the optimum sequential test that optimizes the proposed criterion. The resulting optimum rule resembles the well known CUSUM rule with the corresponding test-statistic-update being a function of the pdfs of the transitory and the steady state phases.

Tuesday morning, June 23

[9:00 - 10:30]

Abraham Wald Prize Ceremony and SQA Editor's Special Invited Presentation

Organizer: Nitis Mukhopadhyay, University of Connecticut

Chair: Nitis Mukhopadhyay, University of Connecticut

Title: Abraham Wald Prize in Sequential Analysis Ceremony

Speakers: Nitis Mukhopadhyay, University of Connecticut
T. K. S. Solanky, University of New Orleans

Title: SQA Editor's Special Invited Paper Presentation: A Review of the Efficient Score Vector in Sequential Monitoring

Speaker: Edit Gombay, University of Alberta

Abstract:

The use of the efficient score statistic in sequential monitoring procedures is reviewed and analysed. Various models are considered that arise in applications. Several of them were the subjects of separate publications, but with the new methodology they are easily calculated examples. The efficient score vector has the same optimality property as the generalized likelihood ratio, but it has a more simple structure, and this is especially important when the data have complicated structure. Furthermore, with the efficient score vector it is possible to detect which component of the parameter vector is different from the hypothetical or the historical value.

Tuesday morning, June 23

[10:45 - 12:15]

Celebration of George Moustakides's 60th Birthday

Organizers: Georgios Fellouris, University of Illinois at Urbana-Champaign
Olympia Hadjiliadis, City University of New York

Chair: Alexander Tartakovsky

Speakers: Georgios Fellouris, University of Illinois at Urbana-Champaign
Olympia Hadjiliadis, City University of New York
Igor Nikiforov, University Technology of Troyes
Alexander Tartakovsky
Venugopal Veeravalli, University of Illinois at Urbana-Champaign

Tuesday morning, June 23

[10:45 - 12:15]

Count Data Methodologies

Organizer: Nitis Mukhopadhyay, University of Connecticut

Chair: Michael Baron, American University and University of Texas at Dallas

Title: Two-Stage and Sequential Procedures to Estimate with Prescribed Precision the Parameter N of Binomial Distribution, B(N,p), When p Is Known

Speaker: Shelemyahu Zacks, Binghamton University

Abstract:

Two-Stage and purely sequential procedures are presented for estimating N with fixed-width confidence intervals, or with prescribed proportional closeness intervals, when p is known. The exact distributions of stopping times are derived with the exact formulas of the coverage probabilities. Numerical examples compare the exact analytical computations with corresponding simulations.

Acknowledgement: This is a joint work with Prof. Shyamal K. De.

Title: Bounded Length Confidence Intervals for the Binomial Proportion with Applications in Health and Ecological Sciences

Speaker: Swarnali Banerjee, Old Dominion University

Abstract:

Importance of the proportion parameter is very well known in statistics. For instance, estimation of (i) the probability of presence of infestation, (ii) the chance of getting a disease, and (iii) the chance of a relapse are very important in entomology and health studies. They frequently involve binary data modelled by a Bernoulli (p) distribution where p is an unknown parameter; $0 < p < 1$. In this paper, we elaborate on the existing procedures of confidence interval estimation of the proportion parameter, p, and illustrate how they may fail to estimate p efficiently. These arguments are backed with simulation results and analysis of real datasets. We introduce a new bounded length confidence interval for p using both purely sequential and two-stage methodologies. We have proved that the confidence interval will be of bounded length (< 1). This confidence interval not only overcomes the shortcomings cited earlier for the existing procedures, but also enjoys asymptotic properties of consistency and first-order efficiency. The performances of the new procedures are also summarized with simulations. The new procedures thus estimate probability of infestation more efficiently and would lead to better decisions in terms of control of error and balancing the budget. The real data sets analyzed for this paper involved estimating chances of relapse in bone marrow transplant patients (size: small), estimating chances of getting diabetes for Pima Indians (size: moderate) and estimating probability of infestation in a potato beetle data (size: large).

Title: A Robust Sequential Fixed-Width Confidence Interval for Count Data based on Bhattacharyya-Hellinger Distance Estimator

Speaker: Seyed Yaser Samadi, Southern Illinois University

Abstract:

For count data having an unknown mass function g_0, we use the minimum Bhattacharyya-Hellinger distance (MBHD) estimator and a stopping rule to construct a sequential fixed-width confidence interval for a functional $T(g_0) = \theta_0$, where f_{θ_0} is the best-fitting parametric model achieving the MBHD between g_0 and any member of a parametric class of mass functions. We establish the asymptotic *consistency* and *efficiency* properties of the sequential confidence interval and the expected sample size, respectively, as the half-width $d \to 0$. When

the count data comes from a gross-error contamination model $g_{\alpha,L} = (1-\alpha)f_\theta + \alpha\delta_{\{L\}}$ for $\alpha \in (0,1)$, where a parametric model f_θ is mixed with a point mass $\delta_{\{L\}}$ located at a value L, we reparametrize $L = L_d$ such that $L_d \to \infty$ as $d \to 0$, and theoretically show that the expected sample size is affected by α, while the coverage probability of the sequential confidence interval depends on the rate of $[T(g_{\alpha,L_d}) - \theta]/d$, as $d \to 0$. Our reparametrization fully exploits the MBHD estimator's inherent ability to progressively ignore increasing values of L, providing an asymptotically consistent sequential fixed-width interval estimator of θ. When f_θ is Poisson(θ), simulations are conducted to corroborate our theoretical results and to contrast the performance of the MBHD with that of the MLE of θ. A real data on *Death Notice*, modeled as a negative binomial, is analyzed to compare and contrast the performance of our sequential MBHD procedures with those of the MLE.

Tuesday morning, June 23

[10:45 - 12:15]

Sequential Decision Problems

Chair: Michael N. Katehakis, Rutgers University

Title: On the Short Term and the Asymptotic Behavior of Minimal-Exploration Sequential Allocation Policies

Speaker: Michael N. Katehakis, Rutgers University

Abstract:

Consider the problem of sampling sequentially from a finite number of $N \geq 2$ populations or 'bandits', where each population i is specified by a sequence of random variables $\{X_k^i\}_{k\geq 1}$, X_k^i representing the reward received the k^{th} time population i is sampled. For each i, the $\{X_k^i\}_{k\geq 1}$ are taken to be i.i.d. random variables with finite mean. For any slowly increasing function g, subject to mild regularity constraints, we construct two policies (the g-Forcing, and the g-Inflated Sample Mean) that achieve a measure of regret of order $O(g(n))$ almost surely as $n \to \infty$. Additionally, asymptotic probability one bounds on the remainder term are established. In our constructions, the function g effectively controls the 'exploration' of the classical 'exploration/exploitation' tradeoff.

When additional parametric assumptions can be made, one can construct policies that are asymptotically optimal in the sense of achieving the lower bound on the logarithmic rate of increase of the regret of Burnetas and Katehakis (1996). We present such asymptotically

optimal policies for the cases in which $\{X_k^i\}$ are: a) Normal with unknown means μ_i and unknown variances σ_i^2, and b) Uniform with unknown supports $[a_i, b_i]$.

Acknowledgement: This work has been supported by the National Science Foundation under grant CMMI-14-50743. This is joint work with Wesley Cowan, Department of Mathematics, Rutgers University

Title: Towards Minimax-Optimal Sequential Tests

Speaker: Michael Fauss, TU Darmstadt

Abstract:

The problem of minimax-optimal sequential hypothesis testing is considered under corridor-type uncertainties in the underlying probability distributions. It is shown that the minimax-optimal sequential test is determined by a triplet of least favorable distributions. This is in contrast to the fixed sample size case, where a pair of least favorable distributions is sufficient. Moreover, the least favorable triplet is shown to minimize a two-dimensional f-divergence that is in turn determined by the Lagrangian cost function of the test that is optimal with respect to this triplet. An implicit characterization of the least favorable distributions is given, as well as an iterative procedure for designing a minimax-optimal test by alternately finding an optimal test for a simple pair of hypotheses and finding the distributions that minimize the f-divergence induced by this test.

Title: Finite Horizon Markov Decision Problems and a Central Limit Theorem for Total Reward

Speaker: Alessandro Arlotto, Duke University

Abstract:

We prove a central limit theorem for a class of additive processes that arise naturally in the theory of finite horizon Markov decision problems. The main theorem generalizes a classic result of Dobrushin (1956) for temporally non-homogeneous Markov chains, and the principal innovation is that here the summands are permitted to depend on both the current state and a bounded number of future states of the chain. We show through several examples that this added flexibility gives one a direct path to asymptotic normality of the optimal total reward of finite horizon Markov decision problems. The same examples also explain why such results are not easily obtained by alternative Markovian techniques such as enrichment of the state space.

Title: **Minimax Optimality in Robust Detection of a Disorder Time in Poisson Rate**

Speaker: Nicole El Karoui, University Pierre et Marie Curie

Abstract:

We consider the minimax quickest detection problem of an unobservable time of change in the rate of an inhomogeneous Poisson process. We seek a stopping rule that minimizes the robust Lorden (1971) criterion, formulated in terms of the number of events until detection, both for the worst-case delay and the false alarm constraint. In the Wiener case, such a problem has been solved using the so-called cumulative sums (cusum) strategy by Shiryaev (1963, 2009), or Moustakides (2004) among others. In our setting, we derive the exact optimality of the cusum stopping rule by using finite variation calculus and elementary martingale properties to characterize the performance functions of the cusum stopping rule in terms of scale functions. These are solutions of some delayed differential equations that we solve elementarily. The case of detecting a decrease in the intensity is easy to study because the performance functions are continuous. In the case of an increase where the performance functions are not continuous, martingale properties require using a discontinuous local time. Nevertheless, from an identity relating the scale functions, the optimality of the cusum rule still holds. Finally, some numerical illustration are provided.

Tuesday morning, June 23

[10:45 - 12:15]

Sequential Methods for High Dimensional or Spatial Data

Chair: T.N. Sriram, University of Georgia

Title: **Detecting Changes in Spatial-temporal Image Data**

Speaker: Annabel Prause

Abstract:

We study detection methods for three-dimensional signals under dependent noise, i.e. for signals in the space-time domain. We model these signals either as $f\colon \mathbb{R}^3 \to \mathbb{R}$ or as $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^{N_y \times N_x}$ for $N_x, N_y \in \mathbb{N}$. Examples for such signals are multifaceted. They include geographic and climatic data as well as image data, that are observed over a fixed time horizon.

Title: Fixed-Size Confidence Regions in High-Dimensional Sparse Linear Regression Models

Speaker: Ching-Kang Ing, Academia Sinica

Abstract:

There is an extensive literature on fixed-size confidence regions for the regression parameters in a linear model with p regressors, attaining a prescribed coverage probability when p is fixed and the size d approaches 0. Motivated by recent developments in regression modeling in response to applications for which p is considerably larger than the sample size, we develop herein a more versatile sequential methodology for fixed-size confidence regions that can handle the case $p = p(d) \to \infty$ as $d \to 0$.

Title: Online Detection of Changes in Sparse Linear Regression Models

Speaker: Bingwen Zhang, Worcester Polytechnic Institute

Abstract:

Motivated by applications in statistical inference on complex dynamic data, we consider online detection of changes in sparse linear regression models. In the problem considered, the observer sequentially receives pairs of regressands and regressors, whose linear relationship changes at some unknown time. Our goal is to detect the presence of such a change with a minimal detection delay subject to certain false alarm constraints. In this paper, we put the constraint on the linear regression model that the coefficients are sparse. We present an efficient online algorithm to detect changes in the sparse linear coefficients. The proposed online algorithm exploits the sparsity structure of the linear regression models. Theoretical performance analysis of the proposed algorithm is provided. Numerical examples are provided to validate the analytical results.

Tuesday afternoon, June 23

[13:45 - 15:15]

Inference and Multinomial Problems

Organizer: Nitis Mukhopadhyay, University of Connecticut

Chair: T. K. S. Solanky, University of New Orleans

Title: **On an Algebraic Inequality Useful in Sequential Selection Procedures**

Speaker: Bruce Levin, Columbia University

Abstract:

We discuss an interesting inequality which implies the lower bound formula for the probability of correct selection and other selection-related events of interest in the Levin-Robbins-Leu family of sequential binomial subset selection procedures. We outline a strategy for the proof of the key inequality which provides an entirely complete and rigorous proof of the inequality for as many as seven competing populations.

Title: **Detection of Changes in a Multinomial Process**

Speaker: Marlo Brown, Niagara University

Abstract:

We look at a multinomial distribution where the probabilities of landing in each category change at some unknown integer. We assume that the probability structure both before and after the change is known, and the problem is to find the probability that the probability structure has changed. For a loss function consisting of the cost of late detection and a penalty for early stopping, we develop, using dynamic programming, the one and two steps look ahead Bayesian stopping rules. We provide some numerical results to illustrate the effectiveness of the detection procedures. We show that the two step ahead procedure is a slight improvement over the one step ahead procedure. However the two procedures are very consistent in their stopping times.

Title: **Worst-Case Properties of the Slippage Configuration in Multinomial Selection**

Speakers: Craig A. Tovey, Georgia Institute of Technology and Timur Tankayev, Georgia Institute of Technology

Abstract:

The slippage configuration is known to be the worst-case probability vector for many multinomial selection procedures, in the sense that it has the minimum probability of correct selection (PCS) over all vectors outside the indifference zone. This worst-case property helps guide the design of selection procedures, the goal being to maximize the worst-case PCS subject to a cost constraint. Instead, we consider minimizing the worst-case expected cost, an idea due to Anton Kleywegt. We give sufficient conditions under which the slippage configuration has the highest expected cost over all probability vectors outside the indifference zone. We then show that these conditions hold for all potentially optimal procedures, in the case of two alternatives,

Our results combine with two other recent ones to rigorously solve the two-alternative multinomial selection problem, first proposed by Bechhofer, Elmaghraby and Morse in 1959, in the following sense: for a given indifference zone, upper bound on the maximum number of trials, and lower bound on the PCS , the linear program of Tollefson et al. (2014) generates a procedure that has the minimum worst-case expected number of trials among all procedures that satisfy the given bounds. The worst case is taken over all vectors outside the indifference zone. We also discuss a generalization to three or more alternatives.

Tuesday afternoon, June 23

[13:45 - 15:15]

Sequential Data Science

Organizer: Zhezhen Jin, Columbia University

Chair: Nancy Flournoy, University of Missouri

Title: A Joint Sequential Model for the Instrument Memory and the Person Memory

Speaker: Mounir Mesbah, University Pierre et Marie Curie

Abstract:

Specific longitudinal methods allowing for the analysis of latent variables have not yet been much developed despite the growing use of self-reported questionnaires in clinical trials aimed at measuring and evaluating many different latent variables such as QoL in cancer trials, dementia in Alzheimer's trials, etc. The benefit of combining sequential analysis and IRT modelling using Rasch models for binary items has already been studied in the context of clinical trials and seems very promising [1]. In this work, we will present a new longitudinal model where the memory of the instrument and the memory of the person are taken into account jointly. We present the model, its graphical properties in the context of a real longitudinal Health Related Quality of Life study.

References :

[1] Sebille, V., Hardouin, J.B.et Mesbah, M. (2007). Sequential analysis of latent variables using mixed-effect latent variable models: impact of non-informative and informative missing data. Statistics in Medicine. 26:48894904.

[2] Mesbah, M. (2015) Analysis of a Complex Longitudinal Health Related Quality of Life Data by a Mixed Logistic Model. To appear in "Applied Statistics in Biomedicine and Clinical Trials Design"- Chen Z_Chapter 19

Title: A Sequential Split-Conquer-Combine Approach for Analysis of Big Spatial Data

Speaker: Xie Minge, Rutgers, the State University of New Jersey

Abstract:

The task of analyzing massive spatial data is extremely challenging. In this talk, we propose a sequential split-conquer-combine (SCC) approach for analysis of dependent big data and illustrate it using a Gaussian random field model, along with a theoretical support. This sequential SCC approach can substantially reduce computing time and computer memory requirements. We also show that the sequential SCC approach is oracle in the sense that the result obtained using the approach is asymptotically equivalent to the one obtained from performing the analysis on the entire data in a super-super computer. The methodology is illustrated numerically using both simulation and a real data example of a computer experiment on modeling room temperatures. (Joint work with Chengrui Li and Ying Hung)

Title: Adaptive Cut Point Selection for Biomarkers in Censored Survival Data

Speaker: Zhezhen Jin, Columbia University

Abstract:

In biomedical research and practice, continuous biomarkers are often used for diagnosis and prognosis, with cut points being established to monitor treatment effect on survival or time to an event. We will study non-parametric procedure for the selection of time-dependent cut points with censored survival data. Numerical studies will be presented along with real applications.

Tuesday afternoon, June 23

[13:45 - 15:15]

Monitoring and Tracking

Organizer: Nitis Mukhopadhyay, University of Connecticut

Chair: Shelemyahu Zacks, Binghamton University

Title: Repeated Significance Tests Under Budget Constraints

Speaker: Vladimir Pozdnyakov, University of Connecticut

Abstract:

Typically a sequential testing procedure is stopped when a process associated with the test crosses a certain boundary. But sometimes it is practical to introduce an additional stopping rule that may force us to make a decision before the test statistic trajectory crosses the test boundary. Two examples are presented. The first example is a sequential RST for observations that come from a heavy tail distribution. The target sample size is random, and the test is designed in such a way that the target sample size adapts itself to an unknown non-linearly increasing total variation of a process linked to the test. In the second example, stopping is triggered by technological restrictions that are imposed on the total cost of monitoring. When an initial sample size is large both problems can be addressed with help of limit theorems. However, when the sample size is small, it is a more delicate question.

Title: Group Sequential Monitoring of Response-Adaptive Randomised Clinical Trials with Censored Survival Data

Speaker: Steve Coad, Queen Mary, University of London

Abstract:

Suppose that two treatments are being compared in a clinical trial with inherent right-censoring. Then a response-adaptive randomisation rule can be used to reduce the number of patients on the inferior treatment. Although most of the existing work deals with response-adaptive randomisation in the context of a fixed trial size, it is often more efficient to conduct a trial group sequentially. A group-sequential response-adaptive procedure is considered in which an error spending function is used to obtain the appropriate stopping boundaries at the different interim analyses. The finite-sample performance of two response-adaptive randomisation rules is compared with complete randomisation by simulation.

Acknowledgements:
This is joint work with Hsiao-Yin Liu at Queen Mary, University of London, who is supported by a Scholarship from the Ministry of Education in Taiwan.

Title: Recursive Tracking Algorithm for a Predictable Time-Varying Parameter of a Time Series

Speaker: Eduard Belitser, VU University Amsterdam

Abstract:

We propose a recursive algorithm for tracking a multi-dimensional time-varying parameter of a time series, which is also allowed to be a predictable process with respect to the underlying time series. The algorithm is driven by a gain function. For an arbitrary time series model and a gain function satisfying some conditions, we derive a general uniform non-asymptotic accuracy bound for the tracking algorithm in terms of chosen step size for the algorithm and the oscillations of the parameter of interest. We outline how appropriate gain functions can be constructed and give several examples of different variability settings for the parameter process for which our general result can be applied, leading to different convergence rates in different asymptotic regimes. The proposed approach covers many frameworks and models where stochastic approximation algorithms comprise the main inference tool for the data analysis. We consider a couple of concrete examples of tracking algorithms for the following (in general, time varying) quantities: the mean function of a conditionally Gaussian process (we also relate this to the famous Kalman filter), the signal in the "signal+noise" model, a (time-varying) root in the multi-dimensional Robbins-Monro setting, time-varying parameters in the ARCH(1) and AR(1) models, the intensity function of a Poisson process.

Acknowledgement:
This is joint work with P. Serra.

Tuesday afternoon, June 23

[13:45 - 15:15]

Sequential/Quickest Change Detection: Theory and Applications

Organizers:	Venugopal Veeravalli, University of Illinois at Urbana-Champaign George Moustakides, University of Patras
Chair:	Venugopal Veeravalli, University of Illinois at Urbana-Champaign

Title: Almost Optimal Sequential Detection in Multiple Data Streams

Speaker: Georgios Fellouris, University of Illinois at Urbana-Champaign

Abstract:

In this talk, we consider the problem of sequentially testing a simple null hypothesis against a finite set of alternatives. We establish various asymptotic optimality properties of sequential tests that are based on mixture likelihood ratio statistics and weighted generalized likelihood ratio statistics. In particular, based on high-order asymptotic expansions for their operating characteristics, we show that these tests can be designed to minimize within an asymptotically negligible term a weighted expected sample size under the alternative hypothesis. As an application, we consider the problem of signal detection in multiple streams, where observations are sequentially acquired from different sources and signal is present in an unknown subset of these streams. This is joint work with A. Tartakovsky.

Title: Sequential Joint Detection and Estimation: Optimum Tests and Applications

Speaker: Yasin Yilmaz, University of Michigan, Ann Arbor

Abstract:

We deal with the statistical inference problems in which one needs to detect and estimate simultaneously using as small number of samples as possible. Conventional methods treat the detection and estimation subproblems separately, ignoring the intrinsic coupling between them. However, a joint detection and estimation problem should be solved to maximize the overall performance. We address the sample size concern through a sequential and Bayesian setup. Specifically, we seek the optimum triplet of stopping time, detector, and estimator(s)

that minimizes the number of samples subject to a constraint on the combined detection and estimation cost. A general framework for optimum sequential joint detection and estimation is developed. The optimum detector and estimator(s) are strongly coupled with each other, proving that the separate treatment is strictly sub optimum. The theoretical results derived for a quite general model are then applied to several problems with linear quadratic Gaussian (LQG) models, including dynamic spectrum access in cognitive radio, and state estimation in smart grid with topological uncertainty. Numerical results corroborate the superior overall detection and estimation performance of the proposed schemes over the conventional methods that handle the subproblems separately.

Title: **Non-parametric Quickest Change Detection for Large Scale Random Matrices**

Speaker: Taposh Banerjee, University of Michigan, Ann Arbor

Abstract:

The problem of quickest detection of a change in the distribution of a $n \times p$ random matrix based on a sequence of observations having a single unknown change point is considered. The forms of the pre- and post-change distributions of the rows of the matrices are assumed to belong to the family of elliptically contoured densities with sparse dispersion matrices but are otherwise unknown. We propose a non-parametric stopping rule that is based on a novel summary statistic related to k-nearest neighbor correlation between columns of each observed random matrix. In the large scale regime of $p \to \infty$ and n fixed we show that, among all functions of the proposed summary statistic, the proposed stopping rule is asymptotically optimal under a minimax quickest change detection (QCD) model.

Tuesday afternoon, June 23

[15:30 - 17:00]

Recent Developments in the Applications of Sequential Methodologies to Innovative Clinical Trial Designs, Queuing Network Control, and Fault Diagnosis in Multistage Manufacturing Processes

Organizer: Milan Shen, Stanford University

Chair: Milan Shen, Stanford University

Title: **Immuno-Oncology Trial Design with A Potentially Delayed Treatment Effect**

Speaker: Zheng Su, Deerfield Institute

Abstract:

In this talk we discuss a design and analysis scheme for time-sequential clinical trials to cope with potentially non-proportional hazards data in the proportional hazard model. A design scheme that combines the logrank test with the cumulative hazard difference is proposed to better evaluate the actual difference in the survival distributions between treatment and control groups. For immuno-oncology trials with a potentially delayed treatment effect we discuss a two-step design and analysis approach, which first determines whether a change point exists in the hazard ratio function and then applies the logrank test to either all the data or only the data beyond the change point if one is detected. Simulation results show that the proposed design preserves the type I error and has advantages over the logrank test under the non-proportionality settings.

Title: **Optimal Stopping and Singular Stochastic Control of Queueing Networks**

Speaker: Tiong-Wee Lim, National University of Singapore

Abstract:

Bounded variation follower problems (BVFPs) have been studied extensively. The equivalence between a BVFP and optimal stopping associated with a Dynkin game can be used to develop backward induction algorithms based solely on optimal stopping. In this talk, we extend this idea to develop a coupled algorithm to solve the problem of stochastic control of a criss-cross network, for which such equivalence does not hold.

Title: **Sequential Detection and Fault Diagnosis with Application in Multistage Manufacturing Processes**

Speaker: Milan Shen, Stanford University

Abstract:

The problem of detection and diagnosis of abrupt changes in a stochastic system on the basis of sequential observations has many applications, some of which are discussed in this thesis. In statistical process control (SPC), the past decade witnessed the emergence of

a new direction in quality control because of the availability of big data, making use of contemporaneous developments in the statistics literature on high-dimensional data analysis. It has been noticed that in multivariate and high-dimensional applications, only a sparse subset of quality characteristics or other variables of interest undergoes abnormal changes that lead to deviations from the state of statistical control. The past decade also witnessed major developments in surveillance over sensor networks, cyber security and information systems. After a brief review of these developments, we give a general theory for sequential fault detection in these stochastic models and also modify and extend it to the much less developed problem of fault diagnosis. This fault diagnosis (or change isolation problem) is to determine, upon detection of change in a system, which one in a set of possible changes has actually occurred. In this connection, we also develop a parallel theory of sequential multiple hypothesis testing. Motivated by applications to SPC of multistage manufacturing processes (e.g. semiconductor wafer fabrication), we also consider regression models for which variable selection is closely related to the fault diagnosis problem, and develop efficient fault detection-diagnosis procedure.

Tuesday afternoon, June 23

[15:30 - 17:00]

Frontiers in Quickest Change-Point Detection with Applications

Organizers:	Aleksey Polunchenko, State University of New York at Binghamton Grigory Sokolov, State University of New York at Binghamton
Chair:	Grigory Sokolov, State University of New York at Binghamton

Title: Some New Results on Using Quickest Detection in Computer Vision

Speaker: Olympia Hadjiliadis, City University of New York

Abstract:

TBA

Title: Multisensor Optimal Detection

Speaker: Grigory Sokolov, State University of New York at Binghamton

Abstract:

Consider the multisensor quickest detection problem, where the goal is to detect as soon as possible a disorder in the environment, which is reflected in the observations that the sensors collect.

In the special case when it is known in advance that the change will affect exactly one sensor, I will revisit the multichart CUSUM, according to which an alarm is raised the first time a local CUSUM statistic exceeds a user-specified threshold. In this context, I will propose a family of thresholds that makes the inflicted performance loss (relative to the optimal one that could be attained only if the identity of the affected subset was known) asymptotically bounded. In a more general case for the subset of affected sensors, I will show the second-order uniform asymptotic optimality of two families of detection rules: the generalized CUSUM and a mixture CUSUMfor any possible affected subset. This is joint work with Georgios Fellouris (Department of Statistics, University of Illinois at Urbana-Champaign).

Title: Change Points in Dependence Structures of Weak and Strong Dependent High–Dimensional Time Series

Speaker: Ansgar Steland, RWTH Aachen University

Abstract:

High-dimensional time series data arise in diverse areas. For high–dimensional vector time series of dimension $d = d_n$ depending on the sample size n, the case that d is large compared to n or is even larger than n is of particular interest. We are interest in large sample approximations to construct change-point statistics for such high–dimensional time series, in order to investigate the dependence structure for changes. By virtue of the construction of the procedure as a quadratic form related to the sample variance-covariance matrix, the proposed test can also be used to study changes in linear projections of high–dimensional data.

The question arises to which extent one may base such procedures on the sample variance-covariance matrix. Indeed, due to the lack of consistency of the sample variance-covariance matrix with respect to commonly used matrix norms such as the Frobenius norm, various regularized modifications have been proposed and studied, e.g. banding and tapering, [1], or shrinkage estimation, see [6], [7] and [3]. But in many cases one is mainly interested in bilinear forms of the sample variance-covariance matrix representing the covariance of linear projection statistics.

Within a high-dimensional time series model that allows for full covariance matrices, we propose novel large sample approximations for bilinear forms of the sample variance-

covariance matrix of a high–dimensional vector time series, in terms of strong approximations by Brownian motions, which extend approximations obtained by [4]. The results cover weakly as well as many long–range dependent linear processes and are valid for a large class of projection vectors that arise, naturally or by construction, in many statistical problems extensively studied for high–dimensional vector time series.

Among those key applications are the following: For sparse financial portfolio optimization, [2] proposed to construct regularized portfolio vectors. In order to conduct sparse principal component analysis, there are several recent proposals to construct ℓ_1-bounded components such as the SCotLASS (simplified component technique-lasso) approch of [5], the penalized matrix decomposition problem (PMD) studied by [10] or the lassoed principal components (LPC) method of [9]. Our results are also directly applicable to the problem of shrinkage estimation ([6], [7], [3]).

The large sample approximations finally allow us to propose a high-dimensional change-point analysis, in order to test for the presence of a change-point in the dependence structure.

Acknowledgement:
This is joint work with Rainer v. Sachs, Universite catholique Louvain, Belgium.

References

[1] Bickel, P.J. and Levina, E. (2008). Covariance regularization by thresholding. *Annals of Statistics*, 36(6): 2577–2604.

[2] Brodie, J., Daubechies, C., De Mol, C., Giannne, D. and Loris, I. (2009). Sparse and stable Markowitz portfolios. *Proceedings of the National Academy of Sciences of the U.S.A.* (PNAS), 106(30): 12267–12272.

[3] Boehm, H. and von Sachs, R. (2009). Shrinkage estimation in the frequency domain of multivariate time series. *Journal of Multivariate Analysis*, 100(5): 913–935.

[4] Kouritzin, M.A. (1995). Strong approximation for cross-covariances of linear variables with long-range dependence. *Stochastic Processes and Their Applications*, 60(2): 343-353.

[5] Jolliffe, L., Trendafilov, N. and Uddin, M. (2003). A modified principal component technique based on the lasso. *Journal of Computational and Graphical Statistics*, 12: 531–547.

[6] Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2): 365–411.

[7] Sancetta, A. (2008). Sample covariance shrinkage for high-dimensional dependent data. *Journal of Multivariate Analysis*, 99(5): 949–967.

[8] Steland, A. and von Sachs, R. (2015). Large sample approximations for variance-covariance matrices of high-dimensional time series, *revised version submitted.*

[9] Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: A retrospective. *Journal of the Royal Statistical Society, Series B*, 73(3): 273–282.

[10] Witten, D.M., Tibshirani, R. and Hastie, T. (2009). A penalized decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics* 10: 515–534.

Tuesday afternoon, June 23

[15:30 - 17:00]

Mathematical Methods and Complexity Estimates for Sequential Optimization

Organizers: Eugene A. Feinberg, Stony Brook University

Chair: Eugene A. Feinberg, Stony Brook University

Title: Uniform Fatou's Lemma and Its Application to Sequential Optimization with Incomplete Information

Speaker: Michael Z. Zgurovsky, National Technical University of Ukraine

Abstract:

Fatou's lemma is a classic fact in real analysis stating that the limit inferior of integrals of functions is greater than or equal to the integral of the inferior limit. In this talk we introduce a stronger inequality that holds uniformly for integrals on measurable subsets of a measurable space. The necessary and sufficient condition, under which this inequality holds for a sequence of finite measures converging in total variation, is provided. This statement is called the uniform Fatou's lemma, and it holds under the minor assumption that all the integrals are well-defined. The uniform Fatou's lemma improves the classic Fatou's lemma in the following directions: the uniform Fatou's lemma states a more precise inequality, it provides the necessary and sufficient condition, and it deals with variable measures. The uniform Fatou's lemma is useful for the analysis of Partially Observable Markov Decision Processes, and this application will be discussed. This talk is based on joint results with Eugene A. Feinberg and Pavlo O. Kasyanov.

Title: Berge's Maximum Theorem for Noncompact Action Sets and Its Applications

Speaker: Pavlo O. Kasyanov, National Technical University of Ukraine

Abstract:

For an upper semi-continuous set-valued mapping from one topological space to another and for a lower semi-continuous function defined on the product of these spaces, Berge's theorem states lower semi-continuity of the minimum of this function taken over the image sets. It assumes that the image sets are compact. For metrizable topological spaces, we provide the extensions of Berge's theorem and Berge maximum theorem to set-valued mappings with possible noncompact image sets and discuss their applications to Markov Decision Processes. This talk is based on joint results with Eugene A. Feinberg, Mark Voorneveld, and Nina V. Zadoianchuk.

Title: Computational Complexity Estimates for Policy and Value Iteration Algorithms for Total-Cost and Average-Cost Markov Decision Processes

Speaker: Jefferson Huang, Stony Brook University

Abstract:

As proved by Yinyu Ye (2011), both the simplex method with Dantzig's rule and Howard's policy iteration algorithm are strongly polynomial for discounted Markov decision processes (MDPs) with finite state and action sets. In this talk, we present a simple example showing that the value iteration algorithm may not be strongly polynomial. This talk also describes reductions, which take place under certain conditions, of total-cost and average-cost MDPs to discounted ones. These reductions are used to develop strongly polynomial policy iteration and linear programming algorithms for certain classes of total-cost and average-cost MDPs. This talk is based on joint papers with Eugene A. Feinberg.

Tuesday afternoon, June 23

[15:30 - 17:00]

Adaptive Designs in Clinical Trials and Other Problems

Chair: Zhezhen Jin, Columbia University

Title: Efficient and Ethical Adaptive Clinical Trial Designs to Detect Treatment-covariate Interaction

Speaker: Seung Won Hyun, North Dakota State University

Abstract:

With the development of bioinformatics, numerous biomarkers have been found to be associated with diseases and some biomarkers have been demonstrated to account for the heterogeneity of patients responses to treatments. Personalized medicine has been proposed to tailor decisions, practices, or products to individual patient based on a variety of covariates such as biomarkers. It is desirable to offer efficient and ethical clinical trial designs for personalized medicine. Covariate adjusted response adaptive (CARA) randomization is a natural and innovative approach for such clinical trials, since it takes into account patients covariate profiles and it can achieve different objectives by sequentially updating the allocation probability based on collected information. In this paper, we focus on using a flexible and well performed CARA design to simultaneously target the efficient and ethical aims when logistic regression is used to detect the treatment-covariate interaction. The advantages of CARA designs are discussed and demonstrated through numerical studies.

Title: Adaptive Designs to Maximize Power in Clinical Trials with Multiple Treatments

Speaker: David Azriel, University of Pennsylvania

Abstract:

We consider a clinical trial with three competing treatments and study designs that allocate subjects sequentially in order to maximize the power of relevant tests. Two different criteria are considered: the first is to find the best treatment and the second is to order all three. The power converges to one in an exponential rate and we find the optimal allocation that maximizes this rate by large deviation theory. For the first criterion the optimal allocation has the plausible property that it assigns a small fraction of subjects to the inferior treatment. The optimal allocation depends heavily on the unknown parameters and, therefore, in order to implement it, a sequential adaptive scheme is considered. At each stage of the trial the parameters are estimated and the next subject is allocated according to the estimated

optimal allocation. We study the asymptotic properties of this design by large deviations theory and the small sample behavior by simulations. Our results demonstrate that, unlike the two-treatments case, adaptive design can provide significant improvement in power. Joint work with Paul Feigin.

Title: Efficient Importance Sampling for False Alarm Probability under K-Distributed Sea Clutter

Speaker: Huei-Wen Teng, National Central University Taiwan

Abstract:

We propose an adpative importance sampling algorithm for calculating small probabilities of false alarms (PFA) using the constant false alarm rate algorithm under the K-distribution. The importance sampling estimator is optimized by minimizing the variance of the estimator within a parametric family, and the optimal tilting parameter is characterized via a device called conjugate measure. Detection accuracy and efficiency have been analyzed by simulation. This is a joint work with Prof. Cheng-Der Fuh at National Central University.

Tuesday afternoon, June 23

[15:30 - 17:00]

Detection, Ranking-selection, and Other Problems in Sequential Methodology

Chair: Bodhisattva Sen, Columbia University

Title: A New Measure for Testing Independence

Speaker: Qingcong Yuan, University of Kentucky

Abstract:

We introduce a new measure for testing independence between two random vectors. Our measure differs from that of distance covariance, by using expected conditional difference of characteristic functions. We propose one empirical version by slicing on the responses. We show that this particular version under categorical response is equivalent to DISCO (Rizzo and Székely, 2010). This empirical measure is based on certain Euclidean distance. Its properties, asymptotics and applications in testing independence are discussed. Implementation and Monte Carlo results are also presented.

Title: On Restricted Subset Selection Rules

Speaker: Lifang Hsu, Le Moyne College

Abstract:

Restricted subset selection was first studied by Santner (1973, A Restricted Subset Selection Approach to Ranking and Selection Problems. Ph.D. Thesis (Mimeo. Ser. No. 318), Dept. of Statistics, Purdue University, West Lafayette, Indiana) and generalized by Hsu and Panchapakesan (2003, Selecting the normal population with largest mean: a restricted subset selection rule, In Balakrishnan, N., Kannan, N., Srinivasan, M. R., eds. Statistical Methods and Practice: Recent Advances, New Delhi: Narosa Publishing House, Chapter 10, pp. 145-161.) We will present recent advances in restricted subset selection and make comments on possible extensions to problems such as two-stage selection in terms of means when the variances are unknown, selection in terms of variance, and selection among cells in a multinomial distribution.

Title: Detecting a Change in Regression: Nonparametric CUSUM

Speaker: Victor Konev, Tomsk State University

Abstract:

In recent years much progress has been made in the analysis of asymptotic properties of Page's cumulative sum (CUSUM) procedure for detecting abrupt changes in several stochastic models. Lorden's minimax optimization theory has been extended far beyond the setting of independent and identically distributed observations before and after a change time. The CUSUM algorithms usually require the knowledge of (conditional) density functions before and after the change-point. The talk considers the problem of optimal detection of abrupt changes in a stochastic regression, with possibly dependent values, and unknown distribution of errors. The key idea of our approach is to retain the CUSUM rule, playing a dominating role in many discrete-time models, but replace the unknown likelihood ratio statistics in it by a special system of statistics computed from observations, appropriate for on-line implementation. When constructing such a system one should keep in mind that the optimality theory of the CUSUM detecting procedures heavily rests on the fact that the likelihood

ratio statistics have either identical or close (in some sense) statistical properties before the change-point and after it as well. Therefore it seems reasonable to seek for the statistics which not only react to the disruption but possess similar statistical properties. To build such statistics we use the least squares method and apply a sequential sampling scheme, which measures time in terms of accumulated Kullback-Leibler (K-L) divergence. The proposed approach suggests also an alternative performance criterion in the analysis of the procedure by replacing the expected detection delay by the corresponding K-L divergence. We have shown that, under false alarm probability constraint, the nonparametric CUSUM rule is optimal in the sense that it ensures the logarithmic asymptotic for the detection delay. The results of Monte-Carlo simulations are given.

Wednesday morning, June 24

[9:00 - 10:00]

Plenary Session

Title: **Sequential Detection/Isolation of Abrupt Changes with Some Applications**

Speaker: Igor Nikiforov, University of Technology of Troyes

Abstract:

The quickest change detection/isolation (multidecision) problem is of importance for a variety of applications. Efficient statistical decision tools are needed for detecting and isolating abrupt changes in the properties of stochastic signals and dynamical systems, ranging from on-line fault diagnosis in complex technical systems to detection/classification in radar, infrared, and sonar signal processing. The early on-line fault diagnosis (detection/isolation) in industrial processes helps in preventing these processes from more catastrophic failures. The problem of on-line target detection/classification is important when detecting and classifying a type of an abruptly appearing target.

The quickest multidecision detection/isolation problem is the generalization of the quickest changepoint detection problem to the case of M post-change hypotheses. It is necessary to detect the change in distribution as soon as possible and indicate which hypothesis is true after a change occurs. Both the rate of false alarms and the misidentification (misisolation) rate should be controlled by given levels. Several detection/isolation procedures that asymptotically minimize the tradeoff between the expected detection delay and the false alarm/misisolation rates in the worst-case scenario are discussed. These procedures are computationally simple and can be easily implemented on-line.

The theoretical results are illustrated by applications of the proposed detection/isolation procedures in diverse practical areas.

Acknowledgement:

This work was partially supported by the French National Research Agency (ANR) through ANR CSOSG Program (Project ANR-11-SECU-0005).

Wednesday morning, June 24

[10:30 - 12:00]

Markov Decision, Dynamic Network and Multiple Change-points in Credit Rating

Organizer: Haipeng Xing, Stony Brook University

Chair: Haipeng Xing, Stony Brook University

Title: Partially Observable Total-Cost Markov Decision Processes with Borel State Spaces and Their Applications

Speaker: Eugene A. Feinberg, Stony Brook University

Abstract:

This talk describes sufficient conditions for the existence of optimal policies for Partially Observable Markov Decision Processes (POMDPs) with Borel state, observation, and action sets and with the expected total costs. Action sets may not be compact and one-step cost functions may be unbounded. The introduced conditions are also sufficient for the validity of optimality equations, semi-continuity of value functions, and convergence of value iterations to optimal values. We also discuss applications of these results to filtration/identification problems and to inventory control. This talk is based on a joint paper with Pavlo O. Kasyanov and Michael Z. Zgurovsky.

Title: Latent Space Models for Dynamic Networks

Speaker: Yuguo Chen, University of Illinois at Urbana-Champaign

Abstract:

Dynamic networks are used in a variety of fields to represent the structure and evolution of the relationships between entities. We present a model which embeds longitudinal network data as trajectories in a latent Euclidean space. A Markov chain Monte Carlo algorithm is proposed to estimate the model parameters and latent positions of the nodes in the network. The model parameters provide insight into the structure of the network, and the visualization provided from the model gives insight into the network dynamics. We apply the latent space model to simulated data as well as real data sets to demonstrate its performance.

Title: A Recurrent Event Model with Multiple Change-points for Firms Rating Transitions under Market Structural Breaks

Speaker: Haipeng Xing, Stony Brook University

Abstract:

Various sudden shifts in financial market conditions over the past decades have demonstrated the significant impact of market structural breaks on firms' credit behavior. To characterize such effect quantitatively, we develop a continuous-time modulated Markov model for firms' credit rating transitions with the possibility of market structural breaks. The model takes a semi-parametric multiplicative regression form, in which the effects of firms' observable covariates and macroeconomic variables are represented parametrically and nonparametrically, respectively, and the frailty effects of unobserved firm-specific and marketwide variables are incorporated via the integration form of the model assumption. We further develop a mixtured-estimating-equation approach to make inference on the effect of market variations, baseline intensities of all firms' credit rating transitions, and rating transition intensities for each individual firm. We then use the developed model and inference procedure to analyze the monthly credit rating of U.S. firms from January 1986 to December 2012, and study the effect of market structural breaks on firms' credit rating transitions.

Wednesday morning, June 24

[10:30 - 12:00]

Latest Advances in Change-Point Detection and Multiple Testing

Organizers: Shyamal K. De, National Institute of Science Education and Research
Aleksey Polunchenko, State University of New York at Binghamton

Chair: Vasanthan Raghavan, Qualcomm Flarion Technologies

Title: **Stepwise Methods of Multiple Testing Controlling Desired Error Rates in Sequential Experiments**

Speaker: Shyamal K. De, National Institute of Science Education and Research

Abstract:

The problem of multiple hypotheses testing often arises in sequential clinical trials where patients are collected sequentially (possibly in groups) to answer multiple questions about safety and efficacy of a new treatment. Typically safety and efficacy are assessed based on a number of endpoints/questions, e.g., intensity of pain, blood pressure, pulse rate, or other

health measurements. In such experiments, it is necessary to find a statistical answer to each posed question by testing each hypothesis while controlling some overall error probability of making wrong decisions. In this talk, we develop stopping rules and decision rules for sequential multiple testing such that desired error rates are controlled at pre-specified levels and the expected sample size is as low as possible. In case of simultaneous testing of a large number of hypotheses (e.g., in genetics), a few Type I and Type II errors can often be tolerated. For such applications, generalized family wise error rates GFWER-I and GFWER-II, defined as the probabilities of making at least k (¿ 1) Type I errors and at least m (¿ 1) Type II errors respectively, are controlled. Fixed-sample multiple test procedures cannot control GFWER-I and II simultaneously. Our proposed stepwise procedures control both GFWER-I and II at pre-specified levels. For large-scale multiple testing, False Discovery Proportion (FDP) and False Non-discovery Proportion (FNP) are also popular error rates. We propose a test procedure that controls the tail probabilities of both FDP and FNP at some prescribed levels. For testing simple versus simple hypothesis and for certain composite hypotheses, we provide a mathematical proof that the proposed method controls the probabilities of FDP and FNP being more than g1 and g2 respectively, where g1 and g2 are some fixed numbers between 0 and 1. Moreover, we show that the well-known FDR and FNR, defined as the expected values of FDP and FNP, can also be controlled simultaneously using the proposed procedure.

Title: On the Distribution of the Generalized Shiryaev–Roberts Diffusion

Speaker: Aleksey Polunchenko, State University of New York at Binghamton

Abstract:

We consider the quickest change-point detection problem where the aim is to detect the onset of a specified drift in "live"-monitored standard Brownian motion with the change-point assumed unknown (nonrandom). The topic of interest is the distribution of the Generalized Shryaev–Roberts (GSR) detection statistic set up to "sense" the presence of the drift. We derive closed-form formulae for the transition pdf of the time-homogeneous Markov diffusion generated by the GSR statistic (a) in the pre- as well as (b) in the post-change regimes. The two pdf formulae are found analytically, through direct solution of the respective Kolmogorov forward equations using the Fourier spectral method to achieve separation of the spacial and temporal variables. To conclude, we exploit the obtained formulae numerically and offer a brief study to characterize the behavior of the GSR statistic in the pre- and post-change regimes depending on three factors: (a) magnitude of the drift, (b) time, and (c) the GSR statistic's headstart.

Acknowledgement:
This is joint work with Dr. Grigory Sokolov (SUNY Binghamton). The effort of A.S. Polunchenko was supported, in part, by the Simons Foundation (`www.simonsfoundation.org`) through a Collaboration Grant in Mathematics (Award # 304574).

Title: **Comparison of CUSUM and Shiryaev-Roberts procedures for Autocorrelated data**

Speaker: Vasanthan Raghavan, Qualcomm Flarion Technologies

Abstract:

We consider the problem of quickest change-point detection where the observations form a first-order autoregressive (AR) process driven by temporally independent standard Gaussian noise with possible change in the drift and/or the correlation coefficient. The goal of this work is a comparative analysis of the CUSUM chart and the Shiryaev-Roberts procedure with respect to Pollak's Supremum Average Delay to Detection (SADD) constrained to a prespecified level of the Average Run Length (ARL) to false alarm. To the best of our knowledge, this is the first time the performance of the SR procedure is studied for autocorrelated data. Our studies show that both the CUSUM chart and the SR procedure are asymptotically second-order optimal, even though the CUSUM chart is found to be slightly better than the SR procedure, irrespective of the model parameters. Moreover, the existence of a worst-case post-change correlation parameter corresponding to the poorest detectability of the change for a given ARL to false alarm is established as well. Further case studies involving mismatched procedures are also presented for the setting where certain subsets of change process parameters are unknown. This is joint work with Aleksey Polunchenko.

Wednesday morning, June 24

[10:30 - 12:00]

Selection, Ranking, and Multiple Comparisons

Organizer: Nitis Mukhopadhyay, University of Connecticut

Chair: Nitis Mukhopadhyay, University of Connecticut

Title: **Statistical Partition Problem: Past, Present and Future**

Speaker: T. K. S. Solanky, University of New Orleans

Abstract:

The problem of partitioning a set of populations with respect to a control population is considered. Focusing on the origination of the problem in 1940s, some of commonly used formulations proposed for this problem are discussed. The issues related to the implementation of the proposed procedures are explored from the point of view of future needs. An example is included to illustrate the use of some selected formulations.

Title: Enhanced Stepwise Methods for Multiple Comparisons in Sequential Experiments

Speaker: Michael Baron, American University and University of Texas at Dallas

Abstract:

Recently introduced stepwise sequential methods for testing multiple hypotheses are more efficient than a simple Bonferroni adjustment of significance levels. They are able to control both Type I and Type II familywise error rates while requiring a generally smaller expected sample size than the corresponding Bonferroni schemes. However, the underlying Holm method for multiple comparisons treats all the tests equally, and it does not account for their possibly different level of difficulty, especially when different distributions are involved. Taking this into account, we derive sequential, truncated sequential, and non-sequential tests of multiple hypotheses with enhanced performance in terms of the expected sample size.

Title: A Stein-Type Two-Sample Procedure for Comparing Normal Means

Speaker: Elena Buzaianu, University of North Florida

Abstract:

In this paper, we propose a Stein-type two-sample procedure for comparing the means of $k > 1$ experimental normal populations among themselves and with the reference to the mean of a controlled normal population, when the variances of all $(k+1)$ populations are unequal and unknown. Our selection formulation follows closely to that by Bechhofer and Turnbull(1978), who considered the comparison of k normal means with a specific non-random standard value, when the variances are either known or unknown and equal. Instead of comparing the k experimental populations to a non-random standard value, our comparison is made with reference to a random controlled normal population. Moreover, we broaden their assumption of equal unknown variances to unequal unknown variances. The proposed procedure satisfies

two probability requirements: (1) the probability of selecting the control is at least pre-specified P_0^* when the largest experimental mean is significantly smaller than the mean of the control and (2) the probability of selecting the largest experimental mean is at least pre-specified P_1^* when the largest experimental mean is significantly largest than the second largest experimental mean and the mean of the control.

Acknowledgement:
This is joint work with Professor Pinyuen Chen (Syracuse University) .

Wednesday morning, June 24

[10:30 - 12:00]

Recent Advances in Sequential Change Detection

Organizers: Georgios Fellouris, University of Illinois at Urbana-Champaign
Olympia Hadjiliadis, City University of New York

Chair: Olympia Hadjiliadis, City University of New York

Title: Two Dimensional Hypothesis Testing and Sequential Decision Making in the Correlated Brownian Motion Model

Speaker: Michael Carlisle, City University of New York

Abstract:

We consider the problem of sequential decision making on the state of a two-sensor system with correlated noise. Each of the sensors is either receiving or not receiving a signal obstructed by noise, which gives rise to four possibilities: (noise, noise), (signal, noise), (noise, signal), (signal, signal). We set up the problem as a gambler's ruin problem in which we devise a decision rule that minimizes the length of continuous observation time required to make a decision about the state of the system subject to given error probabilities.

Title: Quickest Detection With Post-Change Distribution Uncertainty

Speaker: Heng Yang, City University of New York

Abstract:

We consider the problem of quickest detection of an abrupt change when there is uncertainty about the post-change distribution. In particular we examine this problem in the continuous-time Wiener model where the drift of observations changes from zero to a drift randomly chosen from a collection. We set up the problem as a stochastic optimization in which the objective is to minimize a measure of detection delay subject to a frequency of false alarm constraint. We consider a composite rule involving the CUSUM reaction period, that is the time between the last reset of the CUSUM statistic process and the CUSUM alarm, and show that by choosing parameters appropriately, such a composite rule can be asymptotically optimal of third order in detecting the change point as the average time to the first false alarm increases without bound. Because of the uncertainty, such a composite rule can also provide the information of the post-change distribution through the idea of its construction.

Title: Quickest Change Detection in Hidden Markov Models

Speaker: Cheng-Der Fuh, National Central University Taiwan

Abstract:

The quickest change detection problem is studied in sensor networks, where a set of sensors receive observations from a hidden Markov model (HMM) X and send sensor messages to a central processor, called the fusion center, which makes a final decision when observations are stopped. It is assumed that the parameter in the hidden Markov model for X changes from 0 to 1 at some unknown time. The problem is to determine the policies at the sensor and fusion center levels to jointly optimize the detection delay subject to the average run length (ARL) to false alarm constraint. In this article, a CUSUM-type fusion rule with stationary binary sensor messages is studied and a simple method for choosing the optimal local sensor thresholds is introduced. We also provide an efficient numerical computation method of the Kullback-Leibler information number for a two-state HMM.This is a joint work with Prof. Yajun Mei.

Wednesday morning, June 24

[12:15 - 12:30]

Conference Closing

Speaker: Alexander Tartakovsky