Contributed Session Abstracts

Tuesday, June 5

Contributed Session 1A
Tuesday, 130pm-3pm
Title: Change-Point Detection
Chair: Daniel R. Jeske, University of California – Riverside

Performance Evaluation of Certain Procedures for Reacting to a Change in Distribution when the Post-Change Parameter is Unknown

Sven Knoth, Helmut Schmidt University, Germany
Aleksey S. Polunchenko and Grigory Sokolov, University of Southern California

The simplest change-point detection problem assumes the observations are independent with baseline distribution known pre- and post-change. Regarding the change-point as unknown (but not random), we address a more general setting where the post-change distribution is only specified up to a parameter θ. The literature is rife with approaches to handle this case of (parametric) composite post-change hypotheses. By and large, the principal idea is usually one of the following: a) generalized likelihood ratio (GLR), b) likelihood ratio mixture, and c) adaptive, i.e., θ is estimated sequentially. Better compromise between optimality and practicality is offered by yet another idea – "approximating" the parameter space by a "sieve" of finitely many pre-selected reference points to then "wrap" each into a simple hypothesis and deal with the obtained discrete set of alternatives in any conventional multichart-type of way. See, e.g., Waldmann (EJOR, 1996) and Tartakovsky and Polunchenko (IEEE FUSION, 2008). It is this idea that is the centerpiece of this work. Specifically, we consider stopping with the "earliest" of the individual hypotheses (this is akin to the GLR argument). That is, a bundle of basis procedures is run in parallel. For the basis procedures, we focus on the CUSUM chart and the Shiryaev–Roberts rule. The goal is to develop a numerical framework to evaluate all major performance measures for any procedure of this type given any true value of the parameter. To undertake this task, we first derive exact renewal equations on the performance measures; the equations are multi-variate Fredholm integral equations of the second type and apply to a large class of procedures. We then supply a numerical method to solve these equations. We also provide a tight upper bound on the method's error. Lastly, we offer a case study to illustrate the procedures' design. Specifically, we consider a sequence of independent zero-mean unit-variance Gaussian random variables undergoing a surge in the mean. The surge is of unknown magnitude θ. We assess the procedures' performance via Pollak's (Ann. Stat., 1985) Supremum (conditional) Average Detection Delay (SADD) and via the Stationary ADD (STADD), each subject to a lower bound on the Average Run Length (ARL) to false alarm; the STADD measure is originally due to Shiryaev (Soviet Math. Dokl., 1961). The main conclusion is that even as few as 2 reference points is practically enough to detect the change efficiently: both the SADD and the STADD are at worst 25% less than those of the best procedures for known θ. This is a relatively small price to pay given the low computational cost of the parallel procedures.

Univariate and Data-Depth Based Multivariate Control Charts Using Trimmed Mean and Winsorized Standard Deviation

Kumaresh Dhara, Kushal Kr. Dey, Bikram Karmakar, and Sukalyan Sengupta, Indian Statistical Institute

Over the years, Shewhart's $\bar{X} - \mathrm{S}$ and $\bar{X} - \mathrm{R}$ control charts have remained the most popular tool for Statistical Process Control (SPC) in industries. A generalization of this chart for multivariate data has been proposed using the Hotelling's T^2 statistic. However, a major criticism with the Shewhart chart and its generalizations is its lack of robustness. Various robust alternatives based on repeated medians, MCD and depth scores have already been suggested. But most of them have high computational complexity, lack easy interpretability and rely heavily on the normality assumption.

In this article, we have proposed a computationally efficient and robust alternative to the Shewhart's $\bar{X} - \mathrm{S}$ chart, by using trimmed mean ($\bar{X}_{trim}$) and winsorized standard deviation (W) instead of the usual mean ($\bar{X}$) and standard deviation (S). We have also extended this notion to the multivariate scenario, where we have defined a multivariate trimmed mean and winsorized standard deviation using data depth. The desired properties like robustness, asymptotic distribution etc. for the control chart have been thoroughly studied. Also, our approach has also been extended to EWMA charts for both univariate and multivariate case. We have carried out extensive simulations with univariate and multivariate data, and have observed that our control chart clearly outperforms the Shewhart chart and the EWMA chart and their multivariate analogues, by some margin. We have also compared the performance of our chart with respect to various data depth measures due to Liu (1999), Oja (1983), Tukey (1971) and Sering (1992).

An Adaptive Nonparametric Shewhart-Type Linear-Rank Control Chart

Gary R. Mercado and S. Chakraborti, The University of Alabama

We propose a new distribution-free Shewhart-type control chart based on a suitable linear rank statistic to detect possible shifts of a monitored process. From the efficiency and robustness standpoint, the choice of a chart plotting statistic should depend on the shape of the underlying process distribution, which may be symmetric or asymmetric and may have short, medium, or long tails. However, in most cases, practitioners have no clear idea about the underlying distribution. Thus, an adaptive procedure is useful which learns from a given set of data, leads to an efficient plotting statistic and hence to a "good" control chart. The linear rank statistics include many popular nonparametric tests such as the Wilcoxon rank-sum test and the Median test. In the first stage of the adaptive procedure, the unknown distribution is classified with respect to some measures like tailweight, skewness, or leptokurtosis. Then, in the second stage, an appropriate linear rank statistic is selected for that type of distribution. Finally, this statistic is used to construct a distribution-free Shewhart-type control chart. The proposed chart should be more attractive to the practitioner since it has a known and stable in-control performance and a numerical study reveals that it has better out-of-control performance than the Shewhart X-bar chart.

Partial Sliced Inverse Regression for Quality-Relevant Multivariate Statistical Process Monitoring

Yue Yu, University of Illinois, Chicago
Zhijie Sun, University of Southern California

The concept of the quality-relevant multivariate statistical process monitoring is to monitor the abnormal observations in the measurements. This paper introduces a popular dimension reduction method, sliced inverse regression (SIR), into this area. We provide an extension of SIR for the single-index model by adopting the idea from partial least squares (PLS). Our partial sliced inverse regression (PSIR) method has the merit of incorporating information from both predictors (X) and responses (Y), and it has capability of handling large, nonlinear, or "n<p" dataset. Two statistics with their corresponding distributions and control limits are given based on the X-space decomposition of PSIR for the purpose of fault detection in process monitoring. Simulations showed PSIR outperformed over PLS and SIR for both linear and nonlinear model.

Contributed Session 1B
Tuesday, 130pm-3pm
Title: Design of Experiments
Chair: Zhanpan Zhang, GE Global Research

An Application of Fractional Factorial Designs to Study Drug Combinations
Jessica Jaynes, Xianting Ding, Hongquan Xu, Weng Kee Wong, Chih-Ming Ho, UCLA

Herpes simplex virus type 1 (HSV-1) is known to cause diseases of various severities. There is increasing interest to find drug combinations to treat HSV-1 by reducing drug resistance and cytotoxicity. Drug combinations offer potentially higher efficacy and lower individual drug dosage. In this paper, we report a new application of fractional factorial designs to investigate a biological system with HSV-1 and six antiviral drugs, namely, Interferon-alpha, Interferon-beta, Interferon-gamma, Ribavirin, Acyclovir, and TNF-alpha. We show how the sequential use of two- and three-level fractional factorial designs can screen for important drugs and drug interactions, as well as determine potential optimal drug dosages through the use of contour plots. Our initial experiment using a two-level fractional factorial design suggests that there is model inadequacy and drug dosages should be reduced. A follow-up experiment using a blocked three-level fractional factorial design indicates that TNF-alpha has little effect and HSV-1 infection can be suppressed effectively by using a right combination of the other five antiviral drugs. These observations have practical implications in the understanding of antiviral drug mechanism that can result in better design of antiviral drug therapy.

How to Design Experiments when Categoric Mixture Components Go to Zero
Pat Whitcomb, Stat-Ease, Inc.

Mixture experiments are used when one wants to vary ingredients and the response depends on their relative proportion to the other ingredients. In some mixture experiments the formulator wants one (or more) of the ingredients to be present at alternate and mutually exclusive categoric levels. For example, an ingredient might be available from one of three vendors, or, perhaps, only one of two different materials can be used in a given formulation. The usual approach is to simply cross the mixture model with the categoric model. This works well so long as the proportion of the categoric ingredient does not go to zero. If any categoric components do go to zero, the crossed model contradicts itself by predicting different response values for the different levels of the categoric factor, even though it is completely absent from the blend! This presentation proposes a new form of mixture model that corrects this problem when the ingredient is zero and becomes equivalent to the crossed model when the categoric factor exceeds its zero level. A preservative blend (used to maximize shelf life of a food product) with a categoric factor whose proportion goes to zero is used to illustrate the method.

Reliability Experiments with Random Blocks and Subsampling
Jennifer Kensler, Virginia Tech

Reliability experiments provide important information regarding the life of a product, including how various factors affect product life. Current analyses of reliability data usually assume a completely randomized design. However, reliability experiments frequently contain subsampling which is a restriction on randomization. A typical experiment involves applying treatments to test stands, with several items placed on each test stand. In addition, raw materials used in experiments are often produced in batches. In some cases one batch may not be large enough to provide materials for the entire experiment and more than one batch must be used. These batches represent an additional restriction on randomization. This paper proposes a method for analyzing a reliability experiment with random blocks and subsampling.

Optimal blocking and Semifoldover Plans for Two-Level Factorial Designs
Po Yang, Depaul University

Semifoldover is a procedure for selecting follow-up experimental runs. The combined design obtained by joining the runs in the follow-up design to those of the initial design is called a semifolodver design. In this talk, we consider the problem of choosing optimal semifoldover designs when both blocking and semifoldover techniques are used. We show that a semifoldover blocked design can de-alias as many two-factor interactions as the corresponding foldover blocked design when the initial design is a resolution IV design. The optimal blocking and semifoldover plans in terms of the clear effect criterion are tabulated for 16 and 32 runs.

Contributed Session 1C
Tuesday, 130pm-3pm
Title: Control Charts
Chair: Ron Fricker, Naval Postgraduate School

Control Charts for Simultaneous Monitoring of Unknown Means and Variances of Normally Distributed Processes

Amanda Kaye McCracken and Subha Chakraborti, The University of Alabama
Amitava Mukherjee, Aalto University, Finland

Since their invention in the 1920s, Shewhart-type control charts have been popular tools for use in monitoring processes in fields as varied as manufacturing and healthcare. Most of these charts are designed to monitor a single process parameter, but recently, a number of charts for jointly monitoring the mean and variance of normally distributed processes have been developed. These joint monitoring charts are particularly relevant for processes in which special causes may result in a simultaneous shift in the mean and the variance. Among the available schemes for jointly monitoring the mean and the variance, the vast majority are designed for situations in which the true parameters are known rather than estimated from data. This paper presents a pair of Shewhart-type control charts for simultaneous monitoring of the mean and variance of a normally distributed process when both of these parameters are unknown and are estimated from data. The charts are adaptations of two existing procedures for the known parameter case.

Process Monitoring Strategies for Evaluating Process Performance

Joanne Wendelberger, Brian Weaver, Larry Ticknor, and David Collins, Los Alamos National Laboratory

Statistical Process Monitoring provides an approach for characterizing process behavior and understanding when unusual events may be present. In practice, there are many issues that must be addressed in order to establish that a process is stable and under control. An important part of establishing a process monitoring protocol is to understand the types of aberrant behavior that might occur and to develop appropriate procedures accordingly. Using ongoing statistical process monitoring can potentially result in significant savings and improved performance over static approaches based on occasional sampling or inventories.

Interactive SPC – A Textile Quality Case Study
Scott Wise, SAS Institute

Did you know that you can use Statistical Process Control (SPC) charts for data exploration, in addition to the traditional measurement and quality control purposes? Software developments have recently enabled the ability to use SPC as an interactive way to explore your data and create breakthrough learnings. This case will demonstrate with a real business case from Textile Quality how interactive SPC can be utilized to uncover true root causes that help drive quality savings.

1. Motivation or Background – Software advances now allows the practitioner to get more out of SPC by using it as an interactive tool.
2. Description of work done – Case study shows first-hand the real value of Interactive SPC in a challenging work environment (textile quality).
3. Significance – New point of view on how to use the traditional SPC tool.

How Changes Status Quo: This talk and case show the use of Statistical Process control (SPC) in a new way as a valuable interactive tool, rather than as only a tool for basic control and measurement.

Optimization and Efficiency Analysis of the EWMA Procedure for Detecting Changes in the Exponential Distribution
Aleksey S. Polunchenko, Grigory Sokolov, and Alexander G. Tartakovsky, University of Southern California

Along with CUSUM the Exponentially Weighted Moving Average (EWMA) chart is one of the most popular changepoint detection procedures, in particular in the Statistical Process Control community. One of the most important issues is to optimize EWMA by choosing a threshold and a smoothing parameter to guarantee the given false alarm rate (FAR) on one hand and simultaneously minimize the average delay to detection (ADD). In this work, we optimize the EWMA chart optimally designing the smoothing factor and perform its detailed analysis for the exponential model, i.e., assuming that the observations are independent exponentially distributed whose common mean may unexpectedly change from one known value to another. We examine the performance of the one-sided EWMA chart with respect to Pollak's maximal ADD (minimax setting) and Shiryaev's stationary ADD (in the multi-cyclic context).

Under the minimax setting EWMA's performance is compared against that of the Shiryaev– Roberts–*r* (SR–*r*) procedure (a tweak of the SR procedure that starts off at a specially designed point *r*), which has been recently introduced by Moustakides, Polunchenko and Tartakovsky (Statistica Sinica, 2011) and shown to be strictly minimax for certain scenarios and nearly minimax in the general case by Pulunchenko and Tartakovsky (Ann. Statist., 2010) and Tartakovsky, Pollak, and Polunchenko (Theory Prob. Appl., 2011). Similar prior work on EWMA's relative efficiency with respect to the average run length to detection was done by Novikov (Theory Prob. Appl., 1990) for the Gaussian model (assuming a change in the mean), who showed that in the limit, as the smoothing factor goes to 0 and the average run length to false alarm goes to infinity, EWMA is 22:5% worse than the CUSUM chart. Our study shows that EWMA is approximately 40% worse than the SR-*r* chart for small shifts (from 1 to 1.5) and roughly 20% – for larger shifts (from 1 to 3), with respect to the maximal ADD, which usually coincides with the average run length to detection.

For the multi-cyclic context, the obvious benchmark is the conventional SR procedure which minimizes the stationary ADD among all procedure with the given FAR expressed via the average run length between false alarms (see Shiryaev (Theory Prob. Appl., 1963) for Brownian motion and Pollak and Tartakovsky (Statist. Sinica 2009) for general discrete-time models). Srivastava and Wu (Ann. Statist., 1993) previously carried out the same comparison assuming continuous time (detection of a known shift in the drift of a Brownian motion). However, no prior analogous work in discrete time appears to have been done. Our conclusion is that EWMA and SR charts are quite competitive, which is what Srivastava and Wu (1993) also found.

Wednesday, June 6

Contributed Session 2A
Wednesday, 130pm-3pm
Title: Control Charts
Chair: Tatev Ambartsoumian, University of California – Riverside

Do I Need 5 or 10 Parts to Check if my Measurement Processes are Acceptable?

Victor Aguirre Torres, ITAM
Teresa López Alvarez, Kraft

Suppose that you are facing an audit of several measurement processes, and as usual resources are scarce and deadline is approaching fast. The usual set up for an R&R study is to have 10 parts in the study. You ask the question, can I get away with say 5 parts instead of 10? If the answer is yes then you could reduce in half the effort in making the measurements, this is particularly important if those measurements require an analytical procedure. In this work, using simulation, we evaluate this question from the point of view of assessing the chances that you have evidence that your measurement process is acceptable or not. Say to check that the %R&R is less than 30%. By focusing on the probability that the study gives evidence or not that the process is acceptable, we show that for several scenarios the end result of the R&R study is basically the same for 5 or 10 parts and hence that it is questionable to duplicate the measurement effort. Notice that a different conclusion would be obtained if the focus were on estimation of variance components.

Calculating Run Length Quantiles — Go Beyond the Famous and Simple ARL

Sven Knoth, Helmut Schmidt University, Germany

There is a vast and pretty mature literature on creating, evaluating and comparing control chart performance measures based on the expectation – recall the Average Run Length (ARL), the Conditional Expected Delay (CED), and further derivatives. There are a couple of reasons to analyze versions related to the median (MRL) or other quantiles. Besides their agreeable robustness they provide also a somewhat more natural way to sell a specific control chart design: *This chart triggers with a small probability, such as* $\alpha = 0.05$*, false alarms before, say, 100 observations (or whatever potential monitoring period one has in mind).* Only a small number of papers are published that deal with Run Length quantiles (the references fill one slide).

The talk will address the following topics:

- Extend the current collection of Run Length quantiles by transferring concepts from the “expectation” world. For instance, the Conditional Median Delay (CMD) as reasonable counterpart to the CED is created. Build the corresponding numerical routines.
- Provide the quantile calculation algorithms that are needed, e. g., for EWMA charts with varying limits.
- Implement all numerical recipes into available software – the R package spc.
- Compare several control charts in the light of the new measures.

Empirical Likelihood Confidence Regions for the Evaluation of Continuous-Scale Diagnostic Test in the Presence of Verification Bias

Binhuan Wang, Georgia State University

In a continuous-scale diagnostic test, when a cut-off level is given, the performance of the test in distinguishing diseased subjects from non-diseased subjects can be evaluated by its sensitivity and specificity. The joint inferences for sensitivity and specificity as well as cut-off level play an important role in the assessment of the diagnostic accuracy of the test. Most current studies on this topic focus on complete data cases. However, in some studies, only a portion of subjects given their screening test results ultimately has their true disease status verified. In addition, the verification may depend on the test result and the subject's observed characteristics. Directly applying full data methods to verified subjects results in biased estimates, known as verification bias. In this paper, based on a general framework that combines empirical likelihood and general estimation equations with nuisance parameters, we propose various bias-corrected joint empirical likelihood confidence regions for sensitivity and specificity with verification-biased data. Thorough simulation studies are conducted to compare the finite sample performance of the proposed confidence regions in terms of coverage probabilities, and some suggestions are provided accordingly. Finally, an example is provided to compare various proposed methods.

Empirical Likelihood Based Change Point Detection Method

Yijie Yue, University of Georgia

The change-point detection problem is an important topic in fields such as climate science, quality control and finance. We propose an empirical likelihood-based change-point detection approach based on the idea of sequential probability ratio test. We prove the optimality of the proposed stopping time without model assumptions for the simple null and alternative hypotheses; that is, the parameter is equal to one value in either hypothesis. We also discuss a situation of composite null hypothesis and simple alternative hypothesis, meaning the true value is from a range of values in the null hypothesis. Simulations and a real example application are given, followed by the conclusion and discussion.

Contributed Session 2B
Wednesday, 130pm-3pm
Title: Bayesian and Reliability Applications
Chair: J. Marcus Jobe, Miami University (Ohio)

Bayesian Method for Reliability Analysis

Ming Li , GE Global Research Center
William Q. Meeker, Iowa State University

Bayesian analysis has been part of statistical analysis from the very beginning when the foundations of modern statistics were established. Bayesian methods, however, are rarely used in the analysis of reliability data, mainly due to the lack of user friendly and efficient computation tools. With the development of freely available and efficient software package WinBUGs and OpenBUGs, there are more and more statisticians and engineers using Bayesian's idea to combine useful prior information and the field data. In this talk, we first introduce the basic ideas of the Bayesian method and WinBUGs/OpenBUGs software. Then we will show how to apply Bayesian methods to several reliability problems through WinBUGs/OpenBUGs. Some common mistakes and pitfalls for Bayesian application through WinBUGs/OpenBUGs are also discussed.

Estimation of Curvature: Large Data for Small Problems

James G. Wendelberger, Urban Science Applications, Inc. and The University of New Mexico

Measurement at the nano-scale has resulted in a huge number of measurements over very small regions. When determining the curvature of a silicon wafer one may utilize the large number of measurements to estimate curvature. How is this best done and how is that related to classical statistical analysis of the curvature estimation problem?

Bayesian Errors-in-Variables Calibration of a Nonlinear Metrology Tool

William Guthrie, National Institute of Standards and Technology

This talk will illustrate and assess the performance of the calibration of a hypothetical nonlinear metrology tool using Bayesian statistical methods when calibration standards with well-known values are not available. The Bayesian calibration allows for easy, yet statistically rigorous, incorporation the uncertainty from the measurement standards, the calibration measurements, and production measurement data to obtain the uncertainties associated with measurements made using the tool. Extensions to multivariate calibration will be discussed as time permits.

Bayesian Analysis to Determine Predictors' Relative Importance
Xiaoyin Wang, Towson University

While model selection is very well-known and studied in statistical research, there are not many studies on the relative importance analysis. Relative importance analysis is a very useful supplement to regression analysis. The purpose of determining predictor importance is not model selection but rather uncovering the individual contributions of the predictors relative to each other within a selected model. Past research has documented how indices commonly produced by multiple regression analyses fail to appropriately partition variance to the various predictors when they are correlated (Darlington 1968). In response, dominance analysis (Budescu 1993) has been developed that allow for more accurate variance partitioning among correlated predictors. The purpose of this article is to extend the current research practice by developing an original model the Bayesian framework to evaluate the relative importance of each predictor in a multiple regression model. In this paper, we present the state-of-the-art in Dominance Analysis (DA), and will then use our critic as a starting point to introduce a Bayesian approach, which we call Bayesian Dominance Hierarchies (BDH). The BDH will then be compared to the DA using simulated data sets in different contexts, followed by an empirical example in a business setting. Finally, the discussion will lead to reviewing potential future research avenues.