2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Regression Models You Can See

Wei-Yin Loh, University of Wisconsin, Madison

There are numerous techniques for fitting regression models to data, ranging from classical multiple linear regression to highly sophisticated approaches such as spline-based, tree-based, rule-based, neural network, and ensemble methods. Although it is important in many applications that a regression model be interpretable, research in this area is mostly driven by prediction accuracy. It seems almost a fact that the more sophisticated an algorithm, the less interpretable its models become. In this talk, we discuss some basic problems that hinder model interpretation and propose that the most interpretable model is one that can be visualized graphically. The challenge is how to build such a model without unduly sacrificing prediction accuracy. We propose one solution and compare its prediction accuracy with other methods.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Multi-scale Analysis for A Class of Network Traffic

Mou-Hsiung (Harry) Chang, Mathematics Division, U. S. Army Research Office

This talk considers the buffer contents of multiple time scales on-off and infinite source Poisson high speed data networks. Under heavy tailed distribution assumption for either "on" or "off" session, it will be shown that the buffer content converges to a fractional Brownian motion under fast time scale, to a Levy process under slow time scale, and to a telecom process under intermediate time scale. The paper provides mathematical verifications for the long range dependence and self-similarity properties exhibited in empirical measurements of network traffic.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Heavy Traffic Methods in Wireless Systems: from Brownian Models Towards General Lévy Driven Models and Incorporating Time-Delays

Robert T. Buche and Jim X. Zhang, North Carolina State University

Ever-increasing capacity demands for wireless systems and the limited spectrum available for transmissions make accurate modeling and optimal control analysis of paramount importance. Heavy traffic methods for modeling and control of wireless queueing systems, which are well established for wireline models, will be first be outlined. Loosely speaking, the heavy traffic method assumes the system is operating at near capacity. Then through scaling time and space one can obtain a weak limit which well-approximates the queueing dynamics by a stochastic differential equation, typically driven by a Brownian motion. However, this Brownian system is obtained under essentially weak dependence and light tails assumptions. Including the strong dependence and heavy tailed case is important for modeling recent wireless applications (e.g. multimedia, gaming, etc.) where one expects a more general Lévy motion driving process — the fundamental issues in showing this will be discussed. Modeling time delay (e.g., in channel state and queue state information) is also important and leads to studying a stochastic differential equation with delay. We will discuss some progress and issues in the control analysis under this delay model.

Control of Multi-Node Mobile Communications Networks with Time Varying Channels via Stability Method

Harold Kushner, Brown University

Consider a communications network consisting of mobiles, each of which can be scheduled to serve as a receiver and/or transmitter. There are random external data processes arriving at some of the mobiles, each destined for some set of destinations. Each mobile can serve as a node in the possibly multi-hop path from source to destination. At each mobile the data is queued according to the source-destination pair until transmitted. Time is divided into small scheduling intervals. The capacity or quality of the connecting channels are randomly varying due to the motion of the mobiles and consequent scattering. At the beginning of the intervals, the channels are estimated via pilot signals and this information can be used for the scheduling decisions. The issues are the allocation of transmission power and/or time, bandwidth, and perhaps antennas, to the various queues at the various mobiles in a queue and channel-state dependent way to assure stability and good operation. Lost packets might or might not have to be retransmitted. The decisions are made at the beginning of the scheduling intervals. In a recent work, stochastic stability methods were used to develop scheduling policies for the simple system where there is a single transmitter communicating with many mobiles. The resulting controls were readily implementable and allowed a range of tradeoffs between current rates and queue lengths, under very weak conditions. Here the basic methods and results are extended to the network case. The choice of Liapunov function allows a choice of the effective performance criteria. All essential factors are incorporated into a “mean rate” function, so that the results cover many different systems. Because of the non-Markovian nature of the problem, we use the perturbed Stochastic Liapunov function method, which is designed for such problems. Extensions concerning acknowledgments, multicasting, non-unique routes, and others are given to illustrate the versatility of the method, and a useful method for getting the a priori routes is discussed.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

A Cross-Layer Design Approach to Opportunistic Resource Allocation in Mobile Wireless Networks

R. K. Prasanth and Joao B. D. Cabrera, Scientific Systems Company, Inc.
Cesar Santivanez and Ram Ramanathan, BBN Technologies

Consider a mobile wireless network with a fixed number of nodes. Some nodes in the network generate data destined for other nodes. The nodes have some shared resources such as frequency bands that can be used for communication and some private resources such as buffers where data can be queued. The use of shared resources is subject to time-varying interference constraints. We formulate the problem of transporting data to their destinations in multiple hops using the network resources as a model predictive control problem in which queue lengths at each node are available for feedback and the performance index is a measure of the total transport service performed by the network. The resource allocation results to illustrate the approach. The paper concludes with directions for future research.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Evaluating Maintenance Procedures Using Hypergeometric Sampling in Order to Reduce Resources Required for Logistics Demonstrations

John Sereno, US Army Evaluation Center

When an Army system is developed, procedures for system maintenance are developed as well, which must be evaluated. The typical Army system is sufficiently complex that many failure modes are possible, each potentially necessitating a unique procedure. The standard practice in a logistics demonstration is to verify correctness before an operation test and/or fielding by having maintainers perform each procedure according to the instructions. The resources required to check all of the maintenance procedures, in terms of people, funding and calendar time can be prohibitive. The application of hypergeometric sampling provides a statistical basis for successfully concluding the demonstration of the maintenance procedures, based on a sample, provided the results meet specific criteria. Sampling does introduce decision risk which in this case is the risk that one or more of the unevaluated procedures is incorrect; the amount of risk can be quantified and traded off against the benefits. As a result of collaboration between AEC and AMSAA, hypergeometric sampling of maintenance procedures has been used on a few systems already resulting in substantially shorter and less expensive logistics demonstrations.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

A MANPRINT Assessment of an Army’s Rapid Materiel Equipping Initiative for the Persistent Surveillance and Dissemination System Of Systems

Jock O. Grynovicki, Jean S. Breitenbach, and Teresa A. Branscome, US Army Research Laboratory
Steve Abdalla*, MSA Incorporated

The Army Research Laboratory Human Research and Engineering Directorate supported the Army Test and Evaluation Command during Rapid Equipment Fielding of the Persistent Surveillance Dissemination System of Systems (PSDS2) capability demonstration. Analysis of the MANPRINT data from the demonstration helped validate the tested architecture, which provides the capability to integrate data-feed from imaging sensors to assist in producing actionable intelligence, as well as to enhance situational awareness. The PSDS2 system demonstrated the capability to assist the soldier to detect, recognize, and infer the threat level of persons or object. This capability currently is not available in theater. Manpower, personnel, training, human factors, cognitive readiness, situational awareness, information management and workload were key areas that were addressed during the evaluation. Both qualitative and quantitative data supported the assessment. Data included demographics questionnaire, surveys, operator interviews, independent observer observations, and a baseline battery of cognitive questionnaires to obtain a general understanding of the knowledge, skills and ability of the operators, how they make decisions, and how they assess situations under uncertain conditions. This paper will present the approach, methodology, and issues associated in conducting a Rapid Materiel Equipping Initiative MANPRINT assessment. Key MANPRINT findings from the assessment will be briefly summarized.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

A Stochastic Usage Model for Testing Multi-Stage Systems

Donald P. Gaver and Patricia A. Jacobs, Naval Postgraduate School; Kevin D. Glazebrook, Lancaster University; Ernest A. Seglie, Office of the Director, Operational Test and Evaluation, OSD

A multi-stage system is represented by a network of nodes that represent subsystems (e.g. software algorithms or hardware elements such as mid-course missile propulsion). Usage of the system, visits to or usage of the various nodes, is represented as a time-dependent Poisson process, possibly favoring early nodes at the beginning of a system run or mission; later nodes tend to be accessed towards run end. Any visit to a node is assumed capable of activating any fault or design defect (DD) resident, of which there is a random, e.g. independently and node-specific Poisson, number at each node. Any DDs that activate during a test-mission are candidates for removal (by re-design) at the end of the test mission; successful removal occurs independently with node-specific probability. The model allows explicit calculation of the generating function of the number of DDs remaining in each node of the system (the Poisson usage model makes these independent) after a fixed number of test-missions. Numerical methods allow explicit numerical calculation of a fielded system success (no remaining DD failure) in field operation.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Toward Measuring Performance of Unmanned Ground Vehicles

Barry Bodt and Rick Camden, U.S. Army Research Laboratory

Unmanned Ground Vehicles (UGVs) will be an important part of the U.S. Army Future Combat Systems (FCS). Intended functions for UGVs include reconnaissance, surveillance, and target acquisition (RSTA), logistics support, and assault. The Robotics Program Office of the Army Research Laboratory manages the Robotics Collaborative Technology Alliance to pursue the three areas of advanced perception, intelligent control, and human-machine interface that are essential for the maturation of UGVs. Measuring improvements in these areas is challenging, especially when couched in terms of operations. In past work, we designed and conducted an experiment to stress autonomous mobility and evaluate workload in relevant environments. The formal three-site experiment allowed the Army to claim unparalleled autonomous mobility under a variety of experimental conditions. As we move forward, tactical and operational considerations become increasingly more pressing, and performance in light of those considerations must be measured. This talk provides an overview of four early forays toward measuring performance of UGVs and associated technologies. We will introduce the operational consideration in each case, define the measurement approach, and provide examples of results.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Wavelet Density Estimation for Streaming Data

LCDR Kyle Caudle, U. S. Naval Academy

There are many instances of situations where data are streaming. Point of Sales data, Internet Traffic Data, Military C4 communications are simple examples. It is often desired to estimate probability densities and other functions for such data. We propose an approach based on wavelets and adapted to streaming data. We examine asymptotic properties for recursive formulations and also propose an exponential smoothing formulation. We also provide some empirical evidence based on actual data.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Disaggregating Times Series Data

Shirley Bleasdale Joubert, Tom Burr*, and James C. Scovel, Los Alamos National Laboratory

This paper describes our experiences with disaggregating time series data. Suppose data is gathered every two seconds but there is a need to analyze the data at one-second intervals. Under certain assumptions, there are several reasonable disaggregation methods as well as several performance measures to judge their performance. We present results for both simulated and real data for two methods using several performance criteria and indicate when an existing disaggregation method can fail. Based on "Disaggregating Time Series Data," LAMS 13292.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Graphical Displays and Methods for Analyzing Internet Traffic Data for Potential Cyberattacks

Karen Kafadar, Univ of Colorado-Denver; David J. Marchette, Naval Surface Warfare Center; Edward J. Wegman, George Mason University

The analysis of massive, high-volume data sets stresses usual statistical software systems and requires new ways of drawing inferences beyond the conventional paradigm (optimal estimation of parameters from a hypothesized distribution), since the entire data set often cannot be read into the software system. Internet traffic data raise additional challenges: nearly continuous streams of observations from multiple computer systems that interact and exchange information in nondeterministic ways. These features invite cyber attacks, which can be introduced and spread rapidly, and which thus require methods that can detect very rapidly potential departures from “typical” behavior. This talk discusses data on Internet “sessions;”" we show that whole segments of observations are related to single activities, leading to challenges of (1) defining the characteristics for summarizing Internet data into “activity sessions” (2) displaying the summarized data, and (3) designing displays to identify potentially dangerous activities. We describe components of Internet traffic, propose some methods of visualizing them, and illustrate these methods on data collected at a university network. Some open issues in analyzing high-volume data in general are mentioned.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Data Normalization and Combination in Biological Studies: Application to Schizophrenia and Bipolar Disorder Study

Yuanzhang Li, David Cowan, David Niebuhr, Amy Millikan, Walter Reed Army Institute of Research
Robert Yolken, Johns Hopkins University

Large scale enzyme immunoassays are utilized widely in biological studies as they allow for high throughput of specimens. For example, serum and cerebral spinal fluid specimens were tested for antibodies to various viral and parasitic agents among patients with schizophrenia and bipolar disorder [Yolken, 1995]. When used as a screening tool, the output from large scale enzyme immunoassays can have substantial systematic error, and therefore statistical methods to control this error must often be used. Several statistical estimation approaches, including data normalization, meta-analysis and data combination are proposed to perform the analysis of large scale enzyme immunoassays data. We used the observed optical densities from our schizophrenia and bipolar disorder studies and simulated data to compare the different normalization and fusion methods. Data were generated from large scale enzyme immunoassays using normal procedures, with output provided as the log of the optical density. All case-control groups were assayed on the same plates.

In order to handle the systematic measurement error and inconsistency among the variety of optical densities, we first used different scale transformations: normal score, mini max, baseline (negative and positive standard specimens) and median, etc. to reduce the systematic errors across the plates; both the with plate variance and between plate variance were adjusted by meta-analysis. Then the second data normalizations were used to reduce the inconsistency among the different optical densities. Finally, different data fusion methods, such as maximum, summation, etc. were used to combine the information from different optical densities. We found that the combined score from multiple optimal densities by selecting suitable scale normalization and a suitable data fusion technique might improve the prediction of case or control status and reduce the measurement bias compared to using individual optical densities. In addition, the simulation results shows that the median scale transformation is more robust and the normal scale transformation and mini-max transformation is more sensitive to outliers.

The approach presented in this study can also be employed in the analysis of other biological or biometric data.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

An Event Based Parametric Model for the Evaluation of Fit Factors in Full Face Respirator Masks

David Charles Bedard, US Army Evaluation Center

NIOSH testing for CBRN certified respirators has concentrated on observing overall cumulative behaviors in a simulated CBRN environment as a basis for determining expected laboratory respirator protection levels (LRPL) without looking into the events which drive the performance. A more analytical approach, applicable to face dam respirators and other protective gear with similar characteristics, is to model the main events driving the cumulative behavior, leading to the parameterization of the respirator behavior allowing better prediction of the LRPL. One method to do this with existing test facilities and instrumentation and with minimal modification of data storage procedures is presented.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Noise/Signal Separation
Andrew Cseko, Institute for Defense Analyses

Target detection systems contend with noise, a purely random process, and signal, a physical non-random process. Computer algorithms designed to discriminate between signal and noise have typically relied on 'ad-hoc' rules specific to circumstances and/or sensor types. We develop a simple and computationally efficient algorithm for processing classes of sensor data with characteristics comparable to over-the-horizon radar (OTHR) data, discriminating signal from noise, and simultaneously estimating the parameters of the underlying noise distribution. Our focus is on the OTHR-specific situation for which signal and noise have no or only little overlap, the distribution of the noise process is known to be Gaussian, lognormal or Rayleigh (but the parameters are to be estimated), and the functional representation of any signal is completely unknown and cannot be plausibly assumed. Order statistics and minimum least squares methods are exploited to develop Mathematica-based algorithms for estimating the unknown crossover point and noise parameters. Statistical accuracies are documented via Monte Carlo studies.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Reliability Reloaded

Sallie Keller-McNulty, Rice University

In this age of exponential growth in science and technology, the capability to evaluate the performance, reliability, and safety of complex systems presents new challenges. Today's methodology must respond to the ever increasing demands for such evaluations to provide key information for decision and policy makers at all levels of government and industry, problems ranging from national security to space exploration. Scientific progress in integrated reliability assessment requires the development of processes, methods, and tools that combine diverse information types (e.g., experiments, computer simulations, expert knowledge) from diverse sources (e.g., scientists, engineers, business developers, technology integrators, decision-makers) to assess quantitative performance metrics that can aid decision-making under uncertainty. These are highly interdisciplinary problems. The principle role of the statistician is to bring statistical sciences thinking and application to these problems. By the nature of our training, statisticians frequently assume the role of scientific integrator, hence are well poised to lead the development of integrated reliability assessments. However, this puts the statistician closer to policy pressures and politics. This talk will focus on the growing challenges facing statistical sciences in the domain of integrated reliability assessment and how we, as statisticians, must separate the scientific method from the politics of the scientific process to develop assessment methodology that will facilitate the decision making processes.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

X-Testing using Binary Data with Applications to Reliability Demonstration Plans

David Mease, University of California and Vijay Nair, University of Michigan

Reliability demonstration plans are frequently used in industry to formally verify that the reliability of a product exceeds a specified value with a certain degree of confidence. When the value of the reliability to be demonstrated is close to one, traditional reliability demonstration plans are problematic since they require extremely large sample sizes and have low power. One solution to this problem is to induce failure in the testing process by testing products under conditions in which they are more likely to fail. Under these conditions, it is sufficient to demonstrate a lower value of reliability which is then mapped back to the required reliability under standard conditions. Methods of inducing failure can include testing products for a longer period of time, testing at higher stress conditions, or testing weaker products through biased sampling. We give a general framework to describe this type of extreme testing, or "X-testing". Specifically, we consider the effect of X-testing on sample size and power of reliability demonstration plans based on binary data. This depends on the X-transform which is defined as the mapping of the reliability under standard conditions to those of the X-test. We study properties of various X-transforms with respect to zero failure plans, fixed sample size plans and fixed power plans and derive conditions under which X-transforms lead to inadmissible or universally efficient X-tests.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Systems Reliability and Experiment Planning

Alyson Wilson, Los Alamos National Laboratory, Statistical Sciences Group

Assessing systems reliability often requires combining information from many sources with many different characteristics. We will consider two recent models for combining both binary and lifetime data that occurs at different levels of the system. Once the reliability model is developed and fit, the next question is often "What tests should I do next?" The answer, of course, depends on what questions need to be answered, what is currently known, what information sources are available, the costs of collecting information, and the overall budget. We provide several examples to illustrate the issues.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Integrating System and Component Tests for Reliability Assessment
George Lopez, Marine Corps Programs Division, Fallbrook, CA and Todd Graves, Los Alamos National Laboratory

Consider a complex system that may undergo different types of destructive testing, either a full system functional pass/fail test, or a quality components test. In this paper we propose a new technique for integrating both sets of test data for a more accurate estimation of system reliability. Our Bayesian model allows the system reliability to depend on any number of covariates such as age, temperature, manufacturer, etc. We use Markov Chain Monte Carlo (MCMC) simulation for obtaining estimates of our model parameters, and illustrate this technique on a subset of the larger (proprietary) system.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Decision Theory Based Classification of High-Dimensional Vectors Based on Small Samples

David Bradshaw and Marianna Pensky, University of Central Florida

In this paper, an entirely new procedure for the classification of high-dimensional vectors on the basis of a few training samples is described. The proposed method is based on the Bayesian paradigm and provides posterior probabilities that a new vector belongs to each of the classes, therefore it adapts naturally to any number of classes. The classification technique is based on a small vector which can be viewed as a regression of the new observation onto the space spanned by the training samples which is similar to Support Vector Machine classification paradigm. This is achieved by employing matrix-variate distributions in classification.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Self-consistency in the Context of Bayesian Inference

Francisco J. Samaniego, University of California, Davis

The notion of self-consistency of inferential procedures has arisen in a number of statistical contexts. For a survival analysis setting, Efron (1967) defined self consistency in terms of a natural recursive relationship that nonparametric estimators might satisfy and showed that the Kaplan-Meier estimator was the unique self-consistent estimator of the survival function on the interval containing all deaths and censoring times. In the same context, Tsai and Crowley (1985) identified self-consistent estimates as the unique fixed points of nonparametric EM algorithms. In the present paper, the notion is defined for Bayesian estimation of a parameter θ of an exponential family of distributions. Specifically, we define a prior distribution (or Bayes estimator) as self consistent if the equation $E(\theta \mid X = E\,\theta) = E\,\theta$ is satisfied, where X is assumed to be a sufficient and unbiased estimator of θ (usually, the MVUE). This equation requires that the prior mean be a fixed point of the posterior mean function; it simply states that if your experimental outcome agrees with your prior opinion about the parameter, then the experiment should not change that opinion. Arguably, this condition is sufficiently natural and compelling that it might well be adjoined to the usual axioms on probability and utility that define coherent inference and decision making. Surprisingly, there are many prior distributions, including both “objective” and proper priors, which do not enjoy this property. Characterization results for families of self-consistent priors are discussed. The concept of conjugacy will be broadened quite substantially, and results are established specifying when Bayes estimators relative to priors in this broader family outperform classical procedures (in the sense of Samaniego and Reneau (1994)).

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

A Generalized Inference Approach to Estimating Hit Probabilities on a Rectangular Target

David W. Webb

The recent concepts of generalized p-values and generalized confidence intervals have been applied to many estimation problems (e.g., linear models, reliability analysis) where classical frequentist solutions are unavailable. In this presentation, we develop a method for obtaining a confidence intervals (or bound) for the probability of hitting a rectangular target given a set of impact locations whose underlying distribution is assumed to be bivariate normal. The mean interval widths and coverage probabilities for these intervals will be compared with other more traditional estimation methods.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

BANQUET ADDRESS
Capabilities Based Planning & Acquisition
Nancy L. Spruill, Director, Acquisition Resources & Analysis, OSD

Today, and in the future, the Department faces security challenges across a wider range of threats: Irregular, Disruptive, Traditional and Catastrophic. Capabilities based planning and acquisition seek to give decision makers the ability to apportion risk, and allocate resources, across, and within, these challenges. How is the Department doing this today, including in the Quadrennial Defense Review? What tools, data, and models are needed to allow us to better address this difficult problem?

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Spatial Network Design to Detect Regional Trends in Ground Level Ozone
Zhengyuan Zhu, University of North Carolina, Chapel Hill

One goal of the Clean Air Act Amendments of 1990 (CAAA) is to reduce ambient concentrations of atmospherically transported pollutants. Monitoring data from networks such as the Clean Air Act Status and Trends Monitoring Network (CASTNet) can be used to estimate regional trends of these pollutants to evaluate the effectiveness of the CAAA. This paper presents spatial network design methodology to optimize the network's ability to detect and quantify future regional trends in air pollution by adding or relocating monitoring sites. The 1997-2003 CASTNet ozone data is analyzed for trends using a two-stage approach similar to Holland et. al. (2000) to illustrate the design methodology. In the first stage, a site-specific trend is estimated after adjusting for the influence of meteorology and season. In the second stage, trend is assumed to vary over the eastern U.S. as a realization from a Gaussian random field, and Bayesian kriging methodology is used to estimate regional trends and uncertainties. A simulated annealing algorithm is then used to select the optimal locations for future monitoring stations under different practical considerations. We use the expected length of a Bayesian prediction interval as the design criterion to account for the uncertainty of estimating spatial covariance parameters underlying the estimates of regional trend. This design criterion is directly related to the network's capability to detect and quantify trends. Designs that minimize the Bayesian prediction interval are compared with the original design and designs that minimize the kriging variance. At the end we discuss how to determine the sample size of a network for a specific objective.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Putting Quantitative Risks in a Qualitative Context 0r: 95% Certainty in an Uncertain World

Barbara P. Billauer, Law and Science Education & Review Foundation for Law & Science Centers

In an age beset by upheaval, uncertainty, and novel risks, the legal and political communities are increasingly turning to statistics to provide the security, soundness of judgment, and reliability for decision-making once provided by experience, alone. Statisticians, both aware and comfortable with the limits of their discipline, are often unaware of the uses and misuses to which the data generated by their analytical tools are directed. In the legal community, the lack of a common conceptual conduit and indiscriminate use of terms meaning different things in law and science has resulted in legal doctrine replete with scientific gibberish and scientifically erroneous and invalid conclusions. In the policy-making realm, statistics are now used to model the unknowable and incalculable. Thus, parameters necessary to model individual, idiosyncratic, or ‘non-normal’ human responses are unavailable – requiring multiple assumptions for data generation – yet these assumptions are either outside the ken of the statistician or omitted by the researcher from the final conclusions or presentations. Perhaps, then, statistics, has been taken afar of its field by those without proper grounding? If so, what can the statistical community do to foster a more legitimate use of their métier and more informative use of their product?

This session explores two examples where statistically generated information was co-opted by the political world: predictions of health consequences following deployment of bioterrorist agents (e.g. smallpox and anthrax) and ensuing decisions of funding allocation, and projections of health consequences following natural disasters, such as Katrina. In addition, we will embark on a short discussion of failures of the legal community to grasp the basic underpinnings of the statistical method and a description of ‘the disconnect’ in default positions of proof in law and science.

Finally, suggestions to better use the statistical method and statistical data will be dealt with, including avenues for development of cross-communication between the scientific and legal/administrative/political communities to enable better understanding and use of the field and it products.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Deterrence Theory for Layered Defenses

Arthur Fries, Institute for Defense Analyses

This paper posits an empirically supported mathematical model of the psychology of deterrence to the analysis of layered defensive networks constructed to guard against potential acts of terrorism. In this setting, to successfully complete an assigned mission would-be terrorists and/or sympathizers would have to elude inspection, detection, apprehension, etc., at each of a number of prepared defensive obstacles. Traditional mathematical and statistical methods developed for designing and assessing the reliability of coherent collections of subsystems are extended to encompass the problem at hand. Throughout, comparisons are made to a nominal approach that considers only the chances of directly interdicting terrorists, i.e., without accounting for the substantial benefits attributable to deterrence effects.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Automated Metadata

Edward J. Wegman and Faleh Alshameri, George Mason University

Homeland security implies searching massive databases for information involving possible terrorists and the threats they are likely to bring. Many of these databases include free-form text such as intercepted emails and transcripts of phone calls. The implication is that these massive databases are sufficiently large that they cannot be thoroughly examined by humans. Generally metadata involve information about the format of the data, but not necessarily the actual content of the data. The concept of automated metadata is to use data mining tools to extract features from the data and to attach the features to the data as digital objects in the form of metadata. Thus an investigator could search for specific datasets having some desired features. We have employed this notion with a dataset involving 16,000 articles gathered from CNN and Reuters. This work is joint with Faleh Alshameri and is part of his Ph.D. dissertation work.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Inspection and Other Strategies for Homeland Security
James R. Thompson, Rice University

A survey of possible threats to the American homeland reveals America's high vulnerability to terrorist attack. With limited resources available, we need to concentrate on holes in the inspection system where the highest benefit is achieved. Al-Quaida has exhibited a pattern of attacks which can only be described as stylized and ritualistic. How can we use this fact to improve security of the nation? Beyond that, we need to ask the question as to what strategies beyond passive inspection are available. A key issue is the apparent absence of targets of retaliation against terrorist actions.

2005 "Eleventh Army Conference on Applied Statistics", sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

A New Look at the Regression Statistic R-Squared

Donald MacKenzie, Wyle Laboratories, Inc.

This investigation reveals surprising and useful relationships between the Ordinary Least Squares (OLS) regression statistic R^2 and basic parameters of underlying relationships governing dependent and independent variable behavior. Specifically, for the simple relationship Y = A+BX, with normally distributed X and Y variables and constant estimating errors, the expected value of R^2 can be determined from a simple equation. The equation gives the mean R^2 as a function of the standard error of the OLS regression (OSE) divided by the coefficient B: OSE/B. Thus, for any value of OSE/B, we can determine the expected value of R^2! Moreover, the expected value of the F-Stat parameter can also be determined from an equation in OSE/B.

These same type of relationships for OLS egressions also carry over to nonlinear relationships of the form $Y=AX^B$, where X and Y are lognormally distributed and the estimating errors are proportional to the value of the estimate. In this case, the R^2 statistic is a "generalized" version – it is the Pearson correlation between the estimates and the actual values of Y, denoted here by GR^2. Like the OLS case, the expected value of GR^2 can be determined from the ratio of the Standard Percent Error (SPE) and the exponent B: SPE/B. Although the log-transformed OLS method was used to bestfit the X-Y samples in this case, other methods, such as the Iteratively Re-Weighted Least Square (IRLS) method, yield similar results.

The basic approach to studying the R^2, GR2 and F-Stat behavior used Monte Carlo sampling to generate N X-Y pairs for specified values of A, B and OSE. An OLS regression was performed with the X-Y samples, after which the OSE, SPE, R^2, GR^2 and F-Stat values were calculated. This was repeated many times to obtain fairly accurate expected values for all of the statistics. The assumed values of B and OSE were varied systematically to generate curves of R^2 and F-Stat vs. both OSE and B for the Y=A+BX model, and curves of GR2 vs. SPE and B for the $Y=AX^B$ model. Since both R^2 and GR^2 decrease with increasing OSE, and increase with increasing B, it is natural to plot them against the ratio OSE/B. The result was that separate curves of R^2 vs. OSE and R^2 vs. B "collapsed" into a single curve of R^2 vs. Log(OSE/B). This curve was accurately fit with a "reversed" lognormal distribution: R^2 = 1.0 - LNF(OSE/B), where LNF() is the lognormal distribution function. The GR^2 curves did not completely collapse, but they did fall in a fairly narrow band around a reverse lognormal fit curve. The OLS F-Stat curves came very close to a single straight line when Log(F-Stat) was plotted against Log(OSE/B). These results can be used to assess the "reasonableness" of both OLS and nonlinear regressions with the assumed forms (Y=A+BX, $Y=AX^B$, respectively). If a regression R^2 (or GR^2) is close to the collapsed curve, the model has presumably yielded a reasonable result. If it is far away from the curve, then the model might be inappropriate for representing the underlying X-Y relationship.

2005 "Eleventh Army Conference on Applied Statistics", sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Binary and Ordinal Probit Models and the Method of Maximum Likelihood Estimation Applied to Small Sample Sizes

Douglas R. Sommerville, US Army Edgewood Chemical Biological Center

The U.S. Environmental Protection Agency has applied a categorical logistic regression approach to regressing ordered response categories on one or more factors and/or covariates describing toxicant exposure. This approach has also been used at the Edgewood Chemical Biological Center (ECBC) over the past several years in analyzing data from acute inhalation toxicity studies involving chemical warfare agents (both past and present). Many of the studies have used large sample sizes in the estimation of median effective dosages (four or more animals per exposure run, and 20 or more animals per exposure duration). When analyzing binary or ordinal response data from studies with large sample sizes, it is usually possible to simultaneously solve for both the median effective log(dosage) (μ) and the standard deviation (σ) of the effective log(dosages) via the maximum likelihood estimation (MLE) procedure with a normit-link function. However, recent ECBC toxicology work has been forced by experimental constraints to use smaller sample sizes (one animal per exposure run, five to eight animals per exposure duration or condition). For small sample sizes, established mathematical procedures exist for analyzing binary response data using MLE, typically involving the fixing of σ at some set value (based on historical knowledge of the system under study) while solving for μ. This method was expanded to permit the analysis of ordinal response data (using a normit link function), with σ being fixed and solving for multiple μ's corresponding to the endpoints of interest (severe effects and lethality). The method was then used successfully in acute inhalation/ocular, intravenous and subcutaneous toxicity studies involving the Gottingen minipig exposed to GB (sarin) and GF (cyclosarin). The mathematical derivation and algorithm of the MLE procedure for the ordinal normit model applied to small sample sizes and examples of its application to the ECBC minipig studies are presented.

2005 "Eleventh Army Conference on Applied Statistics", sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Does Translation Quality Predict Document Usefulness? Results from Modeling Task-Based Performance Using Automated Machine Translation Evaluation Metrics

Calandra R. Tate, US Army Research Laboratory and University of Maryland

A machine translation (MT) system is computer software designed to translate input texts automatically while preserving the meaning of the original text in the output. Translating texts from one language to another is quite a complex task because there may be a number of ways to express any given phrase. Even though researchers have developed many systems, which have advanced over the years to attain some success; in most cases, the outputs are far from perfect translations and rarely resemble coherent English. Still, developers, users, and stakeholders all want to be able to assess the performance of MT systems, whether for ranking several different systems or for evaluating different versions of the same system under development. The evaluation of MT is complicated and still an evolving field with no agreed-upon methods. Metrics on two levels have been explored: (i) intrinsic measures of performance which judge the system's translation based on similarity to some good human translation of the same text (Papineni et. al. 2002, Turian et. al 2003, Banerjee and Lavie 2005) and (ii) extrinsic measures of effectiveness such as, how well a task can be performed using an output of a translation engine (Taylor and White 1998, Doyon et al. 1999, Voss et. al 2004, Jones et. al 2005). Recent interests lie in identifying the relationship between the two approaches, namely if fast, reusable automated quality scores of translation output will correlate with the ability to do specific real world tasks with the output.

In a project performed by the Army Research Lab(ARL) and sponsored by the Center for Advanced Study for Language(CASL) at the University of Maryland, we constructed an experiment to evaluate MT system output quality based on subjects' performance on a series of tasks. This presentation summarizes the specification, fitting, and interpretation of generalized linear models describing the dependence on automated metrics and other document features of subject performance on the extraction of *Who, Where, When* information. Fifty-nine subjects participated in the Extraction task, each extracting people, places, and times from 18 machine translated documents across 3 Arabic-to-English MT systems. Subject responses were recorded and scored for correctness. We fitted several logistic models to these cross-classified data, using correct-subject-matching as response variables and incorporating document and machine effects as linear predictors for the log odds of correct matching. The best-fitting fixed effect models show that automated MT metrics are useful in distinguishing subject task performance on output from different MT engines. Goodness of fit statistics show that even when specified to include subject-specific random effects, these models do not quite achieve statistical adequacy. We conclude by discussing the consequences of such models, and in particular their effectiveness, and limitations as predictive tools.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Data Visualization with Lattice and Ggobie
Sam Buttrey, Naval Postgraduate School

The Army has approximately 30 Sikorksy UH-60 helicopters that have been equipped with IMD-HUMS, the Integrated Mechanical Diagnostics Helicopter Usage and Maintenance System (IMD-HUMS) developed by Goodrich Corporation. The system consists of a network of sensors that measure vibratons associated with components of the helicopter, such as gears, shafts and bearings. The data are acquired in high-frequency bursts, and are then reduced to several dozen “condition indicators” which change over time and in response to controllable inputs like torque and flight regime. The goal is to allow the efficient diagnosis and replacement of faulty components without relying on a time-based maintenance policy. The system already generates large amounts of data, and over the next few years about 1,500 aircraft will be equipped with the system. Therefore efficient and reliable measures of component condition are needed. In this talk we describe what the data look like, what steps we have taken thus far in the analysis, and directions for upcoming work.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

On the Accuracy of Measuring Rocket Submunition Impact Points Using Video Scoring
Lyn R. Whitaker and Robert A. Koyak, Naval Postgraduate School

In systems integration testing conducted by the U.S. Army at Yuma Proving Grounds, the impact locations of submunitions released from 2.75-inch rockets fired at a target from an air-to-surface weapon system are determined by technicians who visually inspect videotape of the test event for evidence of the impacts. This process, known as overhead scoring, is subject to positional errors (bias and variance), to non-detection of impacts, and to false detections. We use data from a series of systems integration tests conducted at Yuma Proving Ground by the Army in 2001 to statistically examine these errors. The data consist of impact locations measured three different ways: with two different overhead scoring techniques and with on-the-ground location of spent submunitions. Using a linear assignment algorithm, we estimate errors and matching probabilities. The accuracy of overhead scoring is found to depend on the technique used and/or the operator; and it is adversely affected by increasing the number of submunitions used in a test event.

Part of this research was conducted by ENS Derek Jennings and ENS Charles Larwood (USN) to fulfill the requirements of the Masters of Sciences degree in Operations Research conferred by the Naval Postgraduate School in June 2005.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Experimental Designs for Large-Scale Military Simulation Studies
Susan Sanchez, Tom Lucas and Paul Sanchez, Naval Postgraduate School

NPS researchers have produced several breakthrough techniques and technologies in the field of designed experiments within the past five years. Much of this work has been driven by the sheer size of DOD problems and the complexity of simulation models intended to address those problems. These "smart" experimental designs explore dozens or hundreds of factors, and reveal more and better insights about a simulation's behavior than do ad hoc methods or older, smaller designs. Incorporating experimental design into the model building process has also helped frame questions, provide structure to discussions of the application area, and even led to "accidental" V V & A. Beneficiaries of this work include not only the U.S. military services, but also many of our allies. Student theses and research have successfully dealt with diverse problem areas such as network-centric warfare, effective use of unmanned vehicles, future combat systems, peace support operations, logistics, and more. We provide an overview of our approach and the new mindset for simulation experiments, then highlight recent thesis work. We also provide links to software, spreadsheets, and other resources.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Application of the MARS/RSM Procedure to Determine Design Optima in Complex Structures

MAJ Scott T. Crino, US Army, University of Virginia; Donald E. Brown, University of Virginia

This research investigates the use of multivariate adaptive regression splines (MARS) in conjunction with a response surface methodology (RSM) for determining design optima in complex structures. MARS is a flexible regression technique that uses a modified recursive partitioning strategy to simplify high-dimension problems into smaller, yet highly-accurate models. The application of MARS/RSM improves on conventional RSM by addressing highly non-linear high-dimension problems that can be simplified into lower dimensions, yet maintains a low computational cost and better interpretability when compared to other popular metamodeling techniques like kriging or neural networks (NN). The ultimate goal of MARS/RSM is to converge to a simplified limit state function thereby reducing the number and cost of finite element model runs required to determine design optima. The MARS/RSM procedure is applied to a set of low-dimension test functions to determine its convergence and limiting properties. Favorable results indicate the need for further testing with high-dimension complex structures often associated with computationally expensive finite element modeling.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Statistical Analysis of Impurity Atoms in Semiconductor Crystals

Bernard Harris, University of Wisconsin and John M. Zavada,US Army Research Office

Modern information technology is based on the electrical and optical properties of semiconductor crystals doped with various impurity atoms. The concentration and location of the impurity atoms determine the semiconductor properties and the possible device applications. Too low of an impurity concentration renders the semiconductor unusable. Too high a concentration often reduces the desirable effects. In this paper we present a statistical model for estimating the distribution of impurity atoms in an ordered lattice typical of a semiconductor crystal. This model assumes that the number of impurity atoms is much less than the number of possible sites in the lattice and that all ordering are equally probable. Asymptoticx probability distributions are obtained for several characteristics of the allocartion of impurity atoms.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

On the Construction of Restricted G-Optimal Central Composite Designs

Trevor A. Craney, Sikorsky

Many criteria exist for which experimental design matrices can be compared for efficiency to optimality. A G-optimal design matrix minimizes the maximum leverage point of the design. Most types of design matrices for first or second order models are limited from improvement with respect to G-optimality due to the fixed values associated with their construction. Central composite designs permit users to specify the axial values, thus allowing restricted optimization to be made of the design matrix. A method will be discussed for determining G-optimal central composite designs and the efficiencies of the associated design matrices. These efficiencies will be compared to efficiencies of central composite designs incorporating face-centered, spherical, and rotatable axial values for a varying number of factors and variable replication of center runs.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Statistical Analyses of Test Data on Mine Clearing

Jacqueline K. Telford, The Johns Hopkins University Applied Physics Laboratory

Naval Forces have a mission requirement to conduct rapid breaching of mine and obstacle fields to support amphibious landings on defended beaches by the US Marine Corps forces. The goal is to produce clear lanes through which the landing forces can move safely and rapidly. The objective of this project is to characterize the performance of the existing precision-guided bombs against mines and obstacles. This paper documents the statistical analyses that were performed on the data collected from tests conducted through June 2005. Statistical analyses using multiple methods were performed to identify strengths and weaknesses of the current test data. The statistical analysis methods used, ranging from graphical to quantitative, were:

1. Bubble Plot,
2. Qualitative Trend Analysis,
3. Contingency Table Analysis,
4. Logistic Regression Analysis, and
5. Loess Analysis.

Recommendations were provided on where to focus additional testing and modeling to get the largest payoff.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Reliability of Self-Perceived Health Changes in Pre- and Post-Deployment Health Assessment Questionnaires

Joey Zhou, Center for Health Promotion and Preventive Medicine

In pre- and post-deployment health assessment questionnaires (DD2795/2796), there are two ways to obtain self-perceived health changes during deployment. One way is the direct question in the post-deployment questionnaire: “Did your health change during this deployment? Got worse/otherwise” The other, the indirect way, calculates the difference in answering the same question in pre- and post-deployment questionnaires: “Would you say your health in general is: Excellent, Very Good, Good, Fair and Poor” It is found that there is substantial discrepancy in percentages of declining health changes obtained using the two ways. Taking a sample of about 600 soldiers deployed in Afghanistan as an example, self-reported declining health change (the direct way) is 21% while declining health change in self-rated health status (the indirect way) is 46%. Are the two measures of the self-perceived health change reliable? What factors determine the respondent inconsistency? In addition, the presentation proposes a more reliable measure to characterize self-perceived health changes by taking the respondent inconsistency into consideration.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Trends in Enlisted Military Applicants and Gains, 2000 - 2004

Timothy Powers, Weiwei Han, and Yuanzhang Li

The U.S. military depends on a fairly constant flow of new enlistees to satisfy its requirements for personnel who meet established standards of physical, psychological and academic fitness. Recent reports indicate, however, that the number of young people applying for all branches of enlisted military service has dropped considerably. A reduced applicant flow in the face of constant enlistment needs means that the services might have to be less selective in accepting applicants for service. The aims of this study are to document this reported drop in applications for enlisted service, and to assess the impact it has had on the numbers and qualifications of subsequent accessions.

Data on all applicants with no prior military service who underwent an accession medical examination during 2000–2004 at any of the 65 Military Entrance Processing Stations are used. Numbers of applications are first examined separately by service and component for each year of the period for any changes that extend beyond what might be expected by chance. Such changes are then further examined for demographic and other patterns. Categorical data analysis with trend testing and the multivariable log-linear models are used.

Reductions in applicants occurred across virtually all demographic groups, although not uniformly in magnitude. In particular, large reductions in female applicants were seen in each service other than the Marines, which has relatively very few female applicants in general. Considerable reductions in applicants from the youngest age group (age 17–20 years) were also seen in many service branch/component groups.

Data on first-time active duty accessions during 2000-2004 are examined in a similar manner. Demographic patterns are examined and compared to those seen in the applicant pool over the same time period. In addition, we assess trends in the general fitness to serve of new enlistees over this period.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Cluster-based Inference on Distribution of Spherical Data
Harry L. Hurd, Harry L. Hurd Assoc. and Adjunct Professor, UNC Chapel Hill

Motivated by clustering of N-dimensional microarray data using cosine or correlation distances, we develop inference methods to test for uniformity (complete spatial randomness) of the data at both global and local scales. The approach is based on analysis of the clustering dendrogram. Inference at the large scale is obtained from the top of the dendrogram and inference at the local scale is obtained from the bottom. The global analysis gives us information about the general "shape" of the underlying distribution whereas the local analysis identifies where the density of points is too large to be consistent with the general shape. We illustrate using microarray data.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Adaptation of an Alcohol Ecological Agent-Based to Homeland Security

Yasmin H. Said, George Mason University

This research is to instigate the prototype for developing a model framework for ethyl alcohol, which will provide an assessment for interventions, which are meant to minimize acute outcomes (intentional and unintentional injuries/death) that are influenced by the consumption of ethanol, without causing a financial or social burden, and imposing interventions that are ultimately ineffective (or even simply not cost effective). Our framework is ecological (individual agents and interactions are represented), stochastic (neither individual behavior nor consequences of interventions are certain) and very flexible. Constructing the framework raises deep issues in the domain science of alcohol, statistics, mathematics, and computer science. We have developed a time and space dependent stochastic digraph model of alcohol use and abuse. The model is intended as a social network model, which captures the dynamics of alcohol abuse and in particular the acute outcomes associated with alcohol abuse. The intent is to study potential interventions and investigate their effectiveness at reducing the overall prevalence of acute outcomes. Current interventions focus on one outcome at a time rather than simultaneously considering all outcomes. The work involves both sophisticated mathematics (stochastic digraphs) as well as intensive data collection. It is clear that a similar model structure of social networks can be applied to terrorists’ networks with the same ability to examine interventions in order to assess their effectiveness.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Estimating Enemy Using Capture-Recapture

Doug H. Frank, Indiana University of Pennsylvania

An example in a statistics text claims that during WWII the Germans numbered there tanks in sequential order. British intelligence used this knowledge to estimate the number of German tanks with the maximal order statistic. If tanks are randomly assigned serial numbers it is possible to use capture-recapture to estimate the total number of tanks. We review the procedure in two experimental models: direct and inverse. We show the probability behind estimators of the mean and variance of the estimator in both the with and without replacement experiments.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Music Plus One
Chris Raphael, Indiana University

I will discuss my ongoing work in creating a computer system that simulates a sensitively conducted orchestra in a non-improvisatory composition for soloist and orchestra.

My accompaniment system synthesizes a number of knowledge sources including the musical score, on-line analysis of the soloist's performance, and the musical interpretations demonstrated by both the soloist and orchestra in rehearsal. I present a probabilistic model -- a Bayesian Belief Network that represents these disparate knowledge sources in a coherent framework.

During live performance, my system "listens" to the soloist by using a hidden Markov model and conducts the orchestra through principled real-time decision-making engine that incorporates all currently available information for each decision. I will provide a live demonstration of my system on several examples.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Reliability and Survival in Financial Risk
Nozer D. Singpurwalla, George Washington University

In this talk we first create a platform for developing an interface between the mathematical theory of reliability and the mathematics of finance. This we are able to do because there exists an isomorphic relationship between the survival function of reliability and the asset pricing formula of fixed income investments. The isomorphism causes us to: Reinterpret the exponentiation formula of reliability from a more encompassing perspective. Characterize the asset pricing formula in non-parametric classes of functions and obtain its crossing properties. Import results from mathematical finance to reliability theory. Expand the scope of thinking in mathematical finance by invoking therein results from survival analysis vis a vis estimation and inference. Up until now, stochastic process theory has played a dominant role in mathematical finance. In this talk we show how other branches of statistics and probability also offer useful possibilities.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

A Nonparametric Reliability Growth Projection Based on Stein Estimation

Paul M. Ellner and J. Brian Hall, Army Materiel Systems Analysis Activity

We shall present methodology for statistically projecting the impact of delayed corrective actions on a complex system’s reliability. The projection is based on a Stein shrinkage estimator. The estimator is theoretical in the sense that it depends on the first and second population moments for the unknown failure mode initial rates of occurrence and the number of such potential failure modes. The theoretical projection is approximated by using closed form moment estimators to estimate the Stein shrinkage factor. The moment estimators do not require specification of a parent population for the initial mode rates of occurrence. Simulations conducted by the U.S. Army Materiel Systems Analysis Activity indicate that the accuracy of the reliability projection obtained from the approximated Stein shrinkage factor compare favorably to that of the reliability growth projection model adopted by the International Electrotechnical Commission (IEC). The new projection method also requires fewer assumptions and less failure data information than the IEC standard.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Reliability Sampling Methodology Using Simulation and Re-Sampling

John Nierwinski, Jr., Army Materiel Systems Analysis Activity

The U.S. Army is in the process of transforming its current legacy fleet into a future fleet that meets the future standards of the Army. This replacement strategy or transformation has evolved because the U.S. Army requires a more deployable and responsive force. The Army transformation is a three-part process consisting of development of the future fleet, maintenance & logistics of the interim fleet, and Recapitalizing (RECAP) of the legacy fleet. RECAP is a program that rebuilds and/or upgrades these legacy combat systems to a like-new condition in order to ensure and maintain a high level of operational readiness and lower operations & sustainment cost. The U.S. Army will utilize a military database to track and analyze a sample of the combat systems to measure whether the goals are satisfied for various phases of the Army Transformation (i.e. Fielding of RECAP combat systems and future fieldings of new combat systems). Other programs that require reliability sampling and performance tracking include fleet performance assessments for a given combat system.

2005 "Eleventh Army Conference on Applied Statistics", sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Replacing Variance for Assessing Portfolio Risk

MAJ Ernest Y. Wong, United States Military Academy

While constructing a simulation model to determine the optimal stock-bond mix, I discovered an, arguably, more judicious way of characterizing financial portfolio risk: rather than using variance or standard deviation as a way to assess the risk of the portfolio, I found it to be much more intuitive to compare each portfolio based on its probability of generating a negative return—or the chance of losing money. While consistent with the Central Limit Theorem taught in probability and statistics courses, my findings do question several conclusions from the Capital Asset Pricing Model (CAPM) taught in finance and economics classes. My analysis also shows how many commonly-held financial recommendations may be flawed. For instance, investment expert Chuck Carlson, author of the book *Eight Steps to Seven Figures,* suggests subtracting age from 110 to determine the percentage of stocks that should be held in a portfolio. Other rules of thumb recommend a portfolio in which the percentage of debt holdings equals one's age and percentage of equities equals 100 minus ones age. In this paper, I first describe how characterizing investment risk in terms of the chance losing money is more intuitive than using measures such as variance or standard deviation — vestiges of "steam engine era" statistics. I then explain how results from my simulation model challenge conclusions from the CAPM and contradict several financial rules of thumb. Finally, I suggest that leveraging today's simulation based capabilities enables us to account for and characterize risk in a much more discerning manner. By doing so, we are able to look at project risk in more intuitive ways, which may ultimately help to provide greater insights and lead to better decisions.

2005 “Eleventh Army Conference on Applied Statistics”, sponsored by Interface Foundation of North America with cooperation from U.S. Army Research Office

Evolving Structure in Multivariate Time Series with Application to Financial Systems
Katherine B. Ensor, Rice University and Ginger M. Davis, University of Virginia

Financial data lends itself to multivariate analysis due to its hierarchical structure (e.g. individual securities within sectors within markets). Many models exist for the joint analysis of several financial instruments such as securities due to the fact that they are not independent. These models often assume some type of constant behavior between the instruments over the time period of analysis. Instead of imposing this assumption, we are interested in understanding the dynamic covariance structure in our multivariate financial time series, which will provide us with an understanding of changing market conditions. In order to achieve this understanding, we first develop a multivariate model for the conditional covariance and then examine that estimate for changing structure using multivariate techniques. Specifically, we simultaneously model individual stock data that belong to one of three market sectors and examine the behavior of the market as a whole as well as the behavior of the sectors. Our aims are detecting and forecasting unusual changes in the system, such as market collapses and outliers, and understanding the issue of portfolio diversification in multivariate financial series from different industry sectors.