Workshop on Risk Management Approaches in Engineering Applications
Gainesville, Florida-USA, October 1-2, 2018

ABSTRACTS

Monday, October 1st

9:10-9:40am
Progressive Hedging in Nonconvex Stochastic Optimization
R. Tyrrell Rockafellar, University of Washington/Florida (rtr@uw.edu)

The progressive hedging algorithm minimizes an expected "cost" by iteratively decomposing into separate subproblems for each scenario. Up to now it has depended on convexity of the underlying "cost" function with respect to the decision variables and the constraints on them. However, a new advance makes it possible to obtain convergence to a locally optimal solution when the procedure is executed close enough to it and a kind of second-order local sufficiency condition is satisfied. This can moreover work not just for an expectation but also for minimizing a risk objective or buffered probability of exceedance.

9:40-10:10am
Engineering Design and Decision-Making
Johannes O. Royset, Naval Postgraduate School (joroyset@nps.edu)

Standard deviation, mean-squared error, and regression are integral parts of even the most rudimentary data analysis. Decision making based on utility theory, Markowitz risk, and failure probability is equally widely adopted. In this presentation, we review far-reaching extensions and modifications of these concepts that enable alternative approaches to risk mitigation, preference-driven data analysis, and decision making. We illustrate the framework by developing a risk-adaptive multi-fidelity approach to surrogate models in the context of design of a high-speed maritime vessel and earthquake engineering.

10:10-10:40am
Measuring Performance of Machine Learning Algorithms With Buffered Probability
Matt Norton, The Naval Postgraduate School (mnorton@nps.edu)

Popular performance metrics for machine learning (ML) algorithms, while often intuitive, can be difficult to handle when integrated into optimization problems and disregard important information about the quality of the prediction. These issues are increasingly relevant as performance requirements for ML algorithms, and the downstream decisions relying upon ML outputs, are more diverse than ever before. Integrating these requirements as constraints within ML algorithms would be natural, but is challenging given the nature of popular non-convex, discontinuous metrics. We discuss the use of Buffered Probability of Exceedance for formulating performance metrics that are intuitive, risk averse, and naturally implementable as constraints within ML algorithms. In particular, we apply these ideas to classification problems that require careful control over false alarm rates, a commonsense requirement in high-risk classification problems.

11:00-11:30am

A Primal-Dual Algorithm for Large-Scale Risk Minimization

Drew P. Kouri, Sandia National Laboratories (dpkouri@sandia.gov)

Many science and engineering applications necessitate the optimization of systems described by PDEs with uncertain inputs including noisy problem data and unknown boundary or initial conditions. One can formulate such problems as risk-averse optimization problems in Banach space, which upon discretization, become enormous risk-averse stochastic programs. For many popular risk models including the coherent risk measures, the resulting risk-averse objective function is non-smooth. This lack of differentiability complicates the numerical approximation of the objective function as well as the numerical solution of the optimization problem. To address these challenges, I present a general primal-dual algorithm for solving large-scale non-smooth risk-averse optimization problems. This algorithm is motivated by epigraphical regularization of risk measures. As a result, the algorithm solves a sequence of smooth optimization problems using Newton-type methods. I prove convergence of the algorithm even when the subproblem solves are performed inexactly and conclude with numerical examples demonstrating the efficiency of this method.

11:30am-12:00pm

A Locally Adapted Reduced Basis Method for Solving Risk-Averse PDE-Constrained Optimization Problems

Wilkins Aquino, Duke University (wilkins.aquino@duke.edu)
Drew Kouri, Sandia National Laboratories
Zilong Zou, Duke University

The numerical solution of large-scale risk-averse PDE-constrained optimization problems requires substantial computational effort due to the discretization in physical and stochastic dimensions. Managing the cost is essential to tackle such problems with high dimensional uncertainties. In this work, we combine an inexact trust-region (TR) algorithm from with a local, reduced basis (RB) approximation to efficiently solve risk-averse optimization problems with PDE constraints. The main contribution of this work is a numerical framework for systematically constructing surrogate models for the TR subproblem and objective function using local sample-based approximations. Under standard assumptions, the inexact TR algorithm is guaranteed to converge from any initial guess, provided that errors in the evaluation of the objective function and its gradient using our RB approach are adequately bounded. In this work, we provide conditions for which our RB approximations satisfy the necessary bounds and demonstrate the performance of our proposed approach through numerical examples. These examples demonstrate that we can efficiently solve risk-averse PDE-constrained optimization problems with significant computational savings when compared to Monte Carlo.

1:30-2:00pm

Optimal Bundling Strategies for Complements and Substitutes with Heavy-Tailed Valuations

Artem Prokhorov, The University of Sydney, Australia and St. Petersburg State University, Russia (artem.prokhorov@sydney.edu.au)
Rustam Ibragimov, Imperial College Business School, England and Innopolis University, Russia
Johan Walden, The University of California at Berkeley

We develop a general framework for modelling the optimal bundling strategies of a multiproduct monopolist providing goods that have extreme and dependent valuations by consumers. We show that the optimal bundling strategies crucially depend on the degree of heavy-tailedness of consumers' valuations and preferences for the goods, their dependence structure, and the degrees of complementarity and substitutability among the goods provided. For substitutes with sufficiently high degree of substitutability as compared to the degree of heavy-tailedness of consumers' valuations, the seller's optimal strategy is to provide goods separately. On the other hand, provision of the goods as a single bundle is optimal for the seller in the case of complements and substitutes with relatively low degree of substitutability. The conclusions hold regardless of the marginal costs of producing the goods. They further hold for a wide class of dependent valuations with possibly heavy-tailed shocks affecting all preferences and valuations of the consumers. We discuss how these results may help to explain several bundling strategies commonly observed in real-world markets.

2:00-2:30pm

Societal Cyber-Physical System for Sustainable and Resilient Building Design

Khalid M. Mosalam, University of California, Berkeley (mosalam@berkeley.edu)

In this talk, a framework of lifecycle analyses under uncertainty is presented for sustainable and resilient building design. Lifecycle economic and environmental metrics, including initial stage of construction as well as post-hazard repairs, are considered. As developed in recent research, the lifecycle economic analysis includes a complete cost model. The lifecycle multicriteria optimal design under uncertainty is determined through the Generalized Expected Utility (GEU), which allows, if needed, to model the risk aversion of the decision makers and their perception toward extreme events. This is a broadly general framework, which includes the maximization of the expected utility, the cumulative prospect theory and the minimization of the expected cost, as particular decision criteria. The GEU is a conditional expected utility, defined through a suitable superquantile of the utilities. The distributions of the utilities are obtained from the First Order Reliability Method (FORM) which, through the design point, gives also the most critical realizations of the consequences for different degrees of risk aversion. The quantification of the uncertainties is developed through the Performance-based Engineering; here the statistical characterization of the uncertain quantities are obtained through the Kernel Density Maximum Entropy Method (KDMEM) which provides the least biased and most honest distribution given the available information. The decision-making process is dynamic, in the sense that the optimal decision changes accordingly when new information is available. Such dynamic behavior is effectively represented using the Bayesian analysis, here modelled by

combining a decision support tool with the Bayesian Network (BN). The computational platform, in conjunction with an array of sensors and Internet of Things (IoT), may give rise to a semi-autonomous human-centered Cyber Physical System (CPS). Thus, it can be effectively used for the lifecycle holistic optimal design of a smart building. The framework is applied to a hypothetical building office, located in California. The analyses show that design by resilience means design by sustainability, since the more resilient design provides less environmental impact along the lifecycle and is also more advantageous from an economic point of view.

2:30-3:00pm
Sensitivity Analysis for Risk-Averse Optimization
Bart G Van Bloemen Waanders, Sandia National Laboratories (bartv@sandia.gov)
Joey Hart, Sandia National Laboratories

We present sensitivity analysis for risk-averse control problems constrained by partial differential equations (PDEs). The influence of uncertain parameters on the optimal solution is determined by computing the Fr?chet derivative of the solution of the PDE-constrained optimization problem. The most influential parameters are identified by computing the Singular Value Decomposition of this Fr?chet derivative. Our numerical algorithm approximates the leading singular pairs by solving a symmetric eigenvalue problem with a randomized algorithm, giving scalable parallel efficiency. We demonstrate our approach on a simple thermal control problem with a range of uncertain parameters. This motivates the use of sensitivity analysis to evaluate the most important ones so that the computational complexity of a risk-averse optimization problem (using CVaR) can be reduced. In addition, we demonstrate sensitivities on a deterministic thermal-fluid control problem

3:00-3:30pm
Stochastic Dominance of Lithuanian Pension Funds
Audrius Kabasinskas, Kaunas University of Technology, Lithuania
(audrius.kabasinskas@ktu.lt)

Pension funds are hot topic in media of Lithuania. Our analysis show that in the last decade conservative funds had better performance than most of stock funds. A paper has appeared in 2018. It is essential to look at this problem from point of view of stochastic dominance (FSD, SSD and TSD).

4:00-4:30pm
Continuity Properties of Minimax and Sequential Optimization under Uncertainty
Eugene A. Feinberg, Stony Brook University (eugene.feinberg@stonybrook.edu)

This talk describes continuity properties of minimax values and solutions to minimax problems based on the recently discovered generalization of Berge's maximum theorem to possibly noncompact decision sets. The talk also presents the theory for sequential decision making under uncertainty, also known as robust optimization of Markov decision processes, based on these continuity properties. This theory leads to more general results for robust optimization than the previously known ones.

The talk is based on joint papers with Pavlo O. Kasyanov and Michael Z. Zgurovsky from the National Technical University of Ukraine Kyiv Polytechnic Institute.

4:30-5:00pm

The Shortest Path Interdiction Problem With Arc Improvement Recourse: a Multiobjective Approach

Cole Smith, Clemson University (jcsmith@clemson.edu)

We consider the shortest path interdiction problem involving two agents, a leader and a follower, playing a Stackelberg game. The leader seeks to maximize the follower's minimum cost by interdicting certain arcs, thus increasing the travel time along the affected arcs. The follower may improve the network after the interdiction by lowering the cost of some arcs, subject to a budget restriction limiting the number of arc improvements. The leader and the follower are both aware of all problem data, with the exception that the leader is unaware of the follower's improvement budget. The effectiveness of an interdiction action is given by the length of a shortest path after arc costs are adjusted by both the interdiction and improvement actions. An element of risk that comes into play is the fact that the leader's interdiction actions are made without knowledge of the follower's improvement budget; as a result, an interdiction action that might be very effective with respect to an anticipated improvement budget value may be relatively ineffective against another possible value. Each possible interdiction action thus yields a vector of potential follower objectives, each corresponding to a different improvement budget value. We discuss these tradeoffs, implications with respect to risk models, and a multiobjective optimization model for this problem that generates the Pareto-optimal frontier.

5:00-5:30pm

Hospital Readmission Reduction Strategy Using Stochastic Programming

Michelle Alvarado, University of Florida (alvarado.m@ufl.edu)

In response to the Hospital Readmission Reduction Program (HRRP), hospitals are developing care strategies to reduce their 30-day readmission rates. The hospitals are evaluated for six conditions (e.g. heart failure, pneumonia) using a weighted, risk-adjusted measure and those hospitals whose readmission rates are worse than their peers, may receive up to a 3% penalty on reimbursement payments from the Centers for Medicare and Medicaid. Readmission rates are uncertain as is the effectiveness of the proposed care strategies. We develop a stochastic program with probabilistic constraints for the hospital's optimal care strategy in response to HRRP. The model minimizes the hospital's costs and utilizes probabilistic constraints to control uncertain readmission probabilities across all patients for each condition, with the goal of achieving targeted readmission rates with a given confidence level. We explore the trade-off between cost of care, reduce readmission rates, and confidence levels for the care strategies.

5:30-6:00pm
Stochastic PDE-Constrained Optimization for Control of Multifunctional Composite Structures using Multi-Field Coupling Effects
Pavlo Krokhmal, University of Arizona (krokhmal@email.arizona.edu)

A two-stage stochastic partial differential equation (PDE)-constrained optimization methodology is developed for the active control of composite structures in the presence of uncertainties in mechanical loads. The methodology employs the effects of coupling between mechanical and electromagnetic fields in mechanically and electrically anisotropic materials. The PDE-constrained optimization procedure utilizes a first-order active-set algorithm with a conjugate gradient method, and automatic differentiation using hyper-dual arithmetic, as well as the adjoint method. The developed optimization methodology is applied to the problem of post-impact vibration control (via applied electromagnetic field) of an electrically conductive carbon fiber reinforced composite plate subjected to an uncertain impact load. The corresponding governing PDEs consist of a nonlinear coupled system of equations of motion and Maxwell's equations. A detailed computational analysis of the obtained solution is conducted, and a significant suppression of vibrations caused by the randomized impact load in all impact load scenarios is observed. The paper is based on joint work with Dmitry Chernikov and Olesya Zhupanska.

Tuesday, October 2nd

9:00-9:30am
Sensitivity analysis in applications with deviation, risk, regret, and error measures
Michael Zabarankin, Stevens Institute of Technology (mzabaran@stevens.edu)

The envelope formula is obtained for optimization problems with positively homogeneous convex functionals defined on a space of random variables. Those problems include linear regression with general error measures and optimal portfolio selection with the objective function being either a general deviation measure or a coherent risk measure subject to a constraint on the expected rate of return. The obtained results are believed to be novel even for Markowitz's mean-variance portfolio selection but are far more general and include explicit envelope relationships for the rates of return of portfolios that minimize lower semivariance, mean absolute deviation, deviation measures of L^p-type and semi-L^p type, and conditional value-at-risk. In each case, the envelope theorem yields explicit estimates for the absolute value of the difference between deviation/risk of optimal portfolios with the unperturbed and perturbed asset probability distributions in terms of a norm of the perturbation.

9:30-10:00am
An Optimal Control Framework for Efficient Training of Deep Neural Networks
Lars Ruthotto, Emory University (lruthotto@emory.edu)

One of the most promising areas in artificial intelligence is deep learning, a form of machine learning that uses neural networks containing many hidden layers. Recent success has led to breakthroughs in applications such as speech and image recognition. However, more theoretical insight is needed to create a rigorous scientific basis for designing and training deep neural networks, increasing their scalability, and providing insight into their reasoning.

In this talk, we present a new mathematical framework that simplifies designing, training, and analyzing deep neural networks. It is based on the interpretation of deep learning as a dynamic optimal control problem similar to path-planning problems. We will exemplify how this understanding helps design, analyze, and train deep neural networks. First, we will focus on ways to ensure the stability of the dynamics in both the continuous and discrete setting and on ways to exploit discretization to obtain adaptive neural networks. Second, we will present new multilevel and multiscale approaches, derived from he continuous formulation. Finally, we will discuss adaptive higher-order discretization methods and illustrate their impact on the optimization problem.

The talk is joint work with Eldad Haber and based in part on the manuscripts:
https://arxiv.org/abs/1705.03341
https://arxiv.org/abs/1703.02009
https://arxiv.org/abs/1709.03698
https://arxiv.org/abs/1804.04272

10:00-10:30am

Dynamic Maximization of the Sharpe Ratio and bPOE

Mikhail Zhitlukhin, Steklov Mathematical Institute, Moscow, Russia (mikhailzh@mi.ras.ru)

Consider the Black-Scholes model of a stock market, where two assets are traded: one is riskless with a constant rate of return, the other one is risky with its log-return process being a Brownian motion. For this model, I will speak about a problem of finding a trading strategy which has the maximal Sharpe ratio (the expected return divided by the standard deviation of return). The main difficulty in this problem is that it's not possible to directly apply the Bellman principle to find the optimal strategy. I will present a new approach, which allows to solve the problem easily, and can also be applied in some similar settings. The main idea is that maximization of the Sharpe ratio can be reduced to maximization of the bPOE in L2, for which the dual representation allows to greatly simplify the problem.

In the first part of the talk, I'll speak about a connection between the bPOE and the Sharpe ratio and introduce the notion of the monotone Sharpe ratio. In the second part, I'll present the solution of the Sharpe ratio maximization problem in two different settings.

11:00-11:30am

How to Supplement Safety Requirements to Prevent Major Technological Catastrophes?

Stan Uryasev, University of Florida (uryasev@ufl.edu)
Giorgi Pertaia, University of Florida

This paper discusses a new probabilistic characteristic called Buffered Probability of Exceedance (bPOE) for evaluation of tails of probabilistic distributions. bPOE equals a tail probability with known mean of the tail (i.e., it is a probability of the tail such that the mean of the tail equals some specified value). The objective of the paper is to show how to upgrade with bPOE safety requirements based on Probability of Exceedance (POE).

Let us explain definition of bPOE with a simple example. For instance, 4% of landfalling hurricanes in US have cumulative damage exceeding \$50 billion (i.e., POE = 0.04 for threshold=\$50 billion). It is estimated, that the average damage from the worst 10% of hurricanes is \$50 billion. In terms of bPOE, we say bPOE=0.1 for the threshold=\$50 billion. bPOE shows that the largest damages having magnitude around \$50 billion have frequency 10%. bPOE can be considered as an important supplement to POE.

The paper considers two application areas: 1) Materials strength regulations (A-basis, B-basis); 2) Ratings of Financial Companies (such as AAA, AA, ...). We demonstrate that these safety requirements can be efficiently managed/optimized with convex and linear programming algorithms. In particular, we discuss how to formulate and solve a Collateralized Debt Obligations (CDOs) structuring problem.

1:30pm - 2:00pm

Buffered Probability of Exceedance for Shortest Paths with Stochastic Arc Costs

Jeremy Jordan, Air Force Institute of Technology, (jeremy.jordan@afit.edu, jjordan810@gmail.com)

We consider the application of probability of exceedance (POE), value at risk (VAR), conditional value at risk (CVAR), and buffered probability of exceedance (bPOE) to shortest path problems with stochastic arc lengths. Problem formulations with bPOE and CVaR functions result in convex problems with binary variables, a significant improvement compared to formulations with POE and VAR functions. For larger shortest path networks, time to minimize bPOE becomes a limiting factor. We therefore explore several possibilities for heuristics, however none have proven successful thus far.

2:00 - 2:30pm

Applications of Buffered Probability of Exceedance to Structural Reliability Problems

Greg Zrazhevsky, Taras Shevchenko National University of Kyiv, Ukraine (zgrig@univ.kiev.ua)
Alex Golodnikov, Glushkov Institute of Cybernetics, Kyiv, Ukraine
Stan Uryasev, University of Florida

This paper presents results obtained during the first year of the project "Application of Buffered Probability of Exceedance (bPOE) to Structural Reliability Problems" supported by the European Office of Aerospace Research and Development. We demonstrate how bPOE concept can be applied to optimization of parameters of mechanical devices for excitation and formation of wave motion. We consider vibrations of hinged beam, which has defects, under the influence of a number of periodic concentrated forces as a simple model of such device. The problem is to determine an optimal number of forces and their characteristics (application points, amplitude and phase of oscillation), which provide the best approximation of a given shape and point-wise phase vibrations of the beam in a given frequency range with a given accuracy. Both deterministic and stochastic problem setups are considered. In deterministic case we assumed that there is a complete information about parameters of defects. In stochastic case, we assumed that only partial information (lower and upper bounds of parameters) is available. Presentation demonstrates mathematical methods for investigation of deterministic and stochastic versions of the problem.

2:30 - 3:00pm

Recursive Optimization of Mean-Semideviation Risk Measures with Variable Assessment

Dionysios Kalogerias, Princeton University (dkalogerias@princeton.edu)

We are concerned with data-driven optimization of a broad class of convex risk measures, termed meansemideviations, strictly generalizing the classical mean-upper-semideviation risk measure. Their construction relies on the concept of a risk regularizer, a one-dimensional nonlinear map with certain properties, essentially generalizing the positive part weighting function in the mean-upper-emideviation functional. Such risk measures are useful in cases where variable risk assessment relative to the values of the central deviation of the respective random cost function is desirable. First, we present a fundamental and intuitive constructive characterization result, demonstrating the generality of mean-semideviations, and providing an analytical device for constructing such risk measures with desirable characteristics. We then introduce and rigorously analyze the MESSAGEp algorithm, a compositional subgradient procedure for iteratively solving convex mean-semideviation problems to optimality. The MESSAGEp algorithm may be derived as an instance of the T-SCGD algorithm of Yang, Wang & Fang (2018), originally analyzed under a generic setting. By exploiting problem structure, we propose a substantially more flexible set of assumptions, under which we establish pathwise convergence of the MESSAGEp algorithm, in the same strong sense of the work of Yang et al. Via an explicit comparison of our assumptions against those of Yang et al., we also show that the new framework strictly generalizes that of Yang et al., allowing for much less restrictive structural requirements, mainly concerning the expansiveness of the cost function and the smoothness of the risk measure involved, including the case of mean-uppersemideviations. This work rigorously establishes the applicability of computationally efficient compositional stochastic optimization for a significantly broader spectrum of convex mean-semideviation risk-averse problems, strictly extending the current state of the art.

3:00 - 3:30pm

Sparsity-Exploiting Sample Average Approximation for Stochastic Programming and Applications to Designing Acoustic Tweezers

Hongcheng Liu, University of Florida (hliu@ise.ufl.edu)

As a common approach to stochastic programming, the sample average approximation (SAA) scheme yields a high sample complexity; it is stipulated that the number of samples should be polynomial in the number of problem dimensions in order to ensure proper optimization accuracy. In this research, we study a modified SAA in the scenarios where the global minimizer is sparse. By exploiting the sparsity through a regularization penalty referred to as the folded concave penalty (FCP), we show that the required number of samples can be significantly reduced: the required sample size is only poly-logarithmic in the number of dimensions. Numerical experiments indicate a shape improvement in the optimality quality even if the sample size is much smaller than the number of decision variables.

We further apply the modified SAA to the stochastic optimization problem of acoustic tweezer (AT) design for cell separation. The AT is a cutting-edge technique to conduct cell-/particle-level medical procedures through acoustic waves. The performance of the AT is very sensitive to the configurations that of the equipment and entails intrinsic stochasticity. Those configurations can potentially involve a large number of tuning parameters. To facilitate the AT design, we employ the modified SAA and evaluate its feasibility via AT simulations.

4:00 - 4:30pm

Asymptotic Bounds for Clustering Problems in Random Networks

Sergiy Butenko, Texas A&M University (butenko@tamu.edu)
Eugene Lykhovyd, Texas A&M University

Clustering is one of the most important tasks in network analysis. Determining the "natural" number of clusters in a network is a fundamental problem in unsupervised clustering. This problem can be approached by finding a maximum-cardinality subset of vertices such that each connected component of the corresponding induced subgraph forms a cluster of a given type (e.g., a clique or some clique relaxation). In this talk, we establish asymptotic bounds on the cardinality of such independent unions of clusters in uniform random graphs.

4:30 - 5:00pm

Decision Making Under Catastrophic Risk

Bogdan Grechuk, University of Leicester, England (bg83@leicester.ac.uk)

A nonstandard probabilistic setting for modeling of the risk of catastrophic events is presented. It allows random variables to take on infinitely large negative values with non-zero probability, which correspond to catastrophic consequences unmeasurable in monetary terms, e.g. loss of human lives. Thanks to this extension, the safety-first principle is proved to be consistent with traditional axioms on a preference relation, such as monotonicity, continuity, and risk aversion. Also, a robust preference relation is introduced, and an example of a monotone robust preference relation, sensitive to catastrophic events in the sense of Chichilnisky (2002), is provided. Probabilities of catastrophic events can be evaluated by distributionally robust method based on generalized Chebyshev's inequalities for various classes of distributions.

This talk is based on joint work with M. Zabarankin, A. Molyboha, F. Faridafshin and A. Naess.

5:00 - 5:30pm

After-Tax Portfolio Optimization and the Observer Effect

Jaime Cuevas Dermody, Financial Engineering LLC (jaime@fe.net)
R. Tyrrell Rockafellar, University of Washington/Florida

Received models of stock-portfolio optimization are almost entirely before-tax. But some guidelines appear in practitioner publications and software literature for: allocation between asset classes defined by general tax characteristics; and harvesting losses. Here instead a model is developed in which investor-dependent after-tax values are incorporated in the optimization framework itself. This model accounts for multiple types of capital-gains and multiple types of ordinary income, each with multiple tax brackets, as well as, multiple tax-payment dates each year. Moreover, it handles complex rules for carrying losses forward and back between years, and for netting the various types of income and loss within years. The key innovation, that permits this new type of optimization, is the after-tax value of a portfolio, to a given investor (facing particular tax laws and situation) at a point in time. A new phenomena is also identified: the measurement of after-tax value changes that value, like the Observer Effect in particle physics. We also treat and measure the reduction in both mean and variance of return stemming from taxation.

5:30 - 6:00pm

Conservative Estimation of Tail Probabilities from Limited Sample Data

Charles Jekel, University of Florida and Sandia National Laboratories (cjekel@ufl.edu)
Vicente Romero, Sandia National Laboratories

Several uncertainty quantification (UQ) methods are compared for conservative but not overly conservative estimation of small tail probabilities involving responses that lay beyond specified thresholds in the tails of probability distributions. Sixteen very differently shaped distributions (or probability density functions, PDFs) and tail probability magnitudes ranging from 10-6 to 10-3 are considered in order for the study to be relevant to a wide range of risk analysis and quantification of margins and uncertainty (QMU) problems. The emphasis of the study is on sparse-data regimes ranging from N=2 to 20 samples, reflective of most experimental and some expensive computational situations. Relatively simple UQ methods tested for this regime involve statistical tolerance interval "Equivalent Normal" and related "Ensemble of Normals" and "Superdistribution" (SD) approaches, as described in [1] and [2]. The performance of these methods were improved for larger samples sizes through the use of a generalized Jackknife resampling technique, which determine a tail probability estimate by averaging estimates from smaller subsamples [3]. Several quantitative metrics for method conservatism and accuracy of tail probability estimation are used to assess and rank the methods' performance over many random trials for each test PDF and probability magnitude. The SD method generally had the best balance of conservatism and accuracy with N = 4 samples. For N = 5, the SD-Jackknife method generally performed best. A large-sample test problem with N=120 showed that several of the above methods performed better than a non-parametric Kernel Density Estimation or a Maximum Likelihood Estimation using the exact mathematical form of the bi-modal test PDF.

[1] Romero, V., Bonney, M., Schroeder, B. & Weirs, V. G., Evaluation of a Class of Simple and Effective Uncertainty Methods for Sparse Samples of Random Variables and Functions. Sandia National Laboratories report SAND2017-12349, Albuquerque, NM (United States), released November 2017.
[2] Romero, V., V.G. Weirs, "A Class of Simple and Effective UQ Methods for Sparse Replicate Data applied to Cantilever Beam End-to-End UQ Problem," paper AIAA-2018-1665, 20th AIAA Non-Deterministic Approaches Conference, AIAA SciTech 2018, Jan. 8-12, Kissimmee, FL.
[3] Jekel, CF, and Romero, VJ, Conservative Estimation of Tail Probabilities from Limited Sample Data—Expanded Investigation of a Class of Simple Uncertainty Methods for Sparse Data, Sandia National Laboratories report in preparation, Albuquerque, NM (United States).