Modern statistics is faced with a series of challenges as it addresses an expanding number of applications in machine learning, artificial intelligence, forecasting, cyber security, smart systems, social networks, and developing the internet of things. The challenges derive from the large-scale nature of these applications, presence of multi-attribute data of varying and unknown quality and relevance, nonstationarity of the arriving information, and multi-stage nature of the problems.

Optimization theories and algorithms have supported classical statistics well for decades as prominently illustrated by the deep understanding and wide application of regression models, likelihood maximization, and more generally M-estimators. These theories and algorithms, together with those of linear algebra and calculus of variations, have played an essential part in the development of rigorous and efficient statistical procedures that consistently provide users with reliable results. However, the expanding number of applications strain the existing mathematical foundations in several directions and has led to a proliferation of ad hoc and heuristic methods. This special issue contains 21 papers describing theoretical advances in several areas in the interface between statistics and optimization. In particular, the paper highlights the wider landscape presented by variational analysis, with its techniques for addressing nonsmoothness, nonconvexity, and multi-valuedness.

The issue starts with six papers on statistical inference. C.R. Doss describes likelihood ratio-based inference in concave regression. C. Huang and X. Huo enhance distributed averaging estimators by a single Newton step and achieve the same asymptotic properties as a centralized estimator. A. Shapiro gives statistical properties of estimators in minimum trace factor analysis, minimum rank matrix completion, and more generally for those obtained by semidefinite programming. M. Lamm and S. Lu give confidence intervals for solutions to stochastic variational inequalities and thereby establish foundations for statistical inference in numerous contexts. R. Bassett and J. Deride settle a long-standing ambiguity in the relationship between maximum a posteriori estimators and Bayes estimators. B. Grechuk and M. Zabarankin discuss measures
of error, stress the importance of asymmetry, and lay out far-reaching extensions of the well-known relationships between least-squares and density functions.

The next three papers show the possibilities that emerge when embracing nonsmoothness, nonconvexity, and multi-valuedness. A. Hantoute et al. show that despite inherent nonsmoothness, probability functions under Gaussian distributions nevertheless have tractable formulae for their subgradients. M. Nouiehed et al. establish that a large number of functions from statistical estimation, machine learning, and risk management can be expressed as the difference of two convex functions, which subsequently can be utilized in analysis and computations. A. Aswani lays out a new frontier of statistics with set-valued functions and their potential for addressing inverse problems.

Seven papers deal with the increasing need for solving large-scale optimization problems in statistical applications. P. Thompson and A. Jofré construct stochastic gradient-type algorithms with optimal complexity for smooth convex problems under relaxed assumptions. F. Roosta-Khorasani and M.W. Mahoney explore the use of Newton’s method, with gradients and Hessians being randomly sub-sampled, and establish finite sample size guarantees and local convergence results. M.-C. Yue et al. construct algorithms based on inexact sequential quadratic approximations for minimizing the sum-of-two-convex functions, one of which is possibly nonsmooth, and establish their convergence even under degeneracy with significant applications in regression and classification. A.Y. Aravkin et al. review and extend level-set methods in convex optimization and demonstrate their benefits in low-rank semidefinite optimization, sparse optimization, and gauge optimization. H. Attouch and J. Peypouquet study trajectories of a second-order differential equation with vanishing damping, governed by the Yosida regularization of a maximally monotone operator, which leads to novel regularized inertial proximal algorithms. P.L. Combettes and J.-Ch. Pesquet establish mean-square and linear convergence of the fundamental block-coordinate fixed-point iterations with random sweeping. R.T. Rockafellar and J. Sun enable the solution of large-scale stochastic variational inequalities by means of extensions of the progressive hedging algorithm.

The last five papers examine applications in risk management, classification, and graph clustering. S. Guo and H. Xu formulate distributionally robust models of uncertainty in the context of shortfall-risk minimization and examine the effect of approximations. M. Glanzer et al. examine acceptability prices of contingent claims and establish a large-deviation result that relates the bid and ask prices under stochastic ambiguity to the quality of the observed data. D.P. Kouri constructs risk quadrangles for a large class of spectral risk measures all with superquantiles as statistic. K. Fountoulakis et al. give variational formulations of local graph clustering problems and establish connections between numerical optimization and graph processing. M.D. Norton and S. Uryasev consider novel performance metrics based on buffered probabilities for classification problems.

The wide range of topics in the interface between statistics and variational analysis covered by this special issue reflects the interests and immense contributions of Distinguished Research Professor Roger J-B Wets, Department of Mathematics, University of California, Davis. Professor Wets celebrated his 80th birthday in 2017 and the special issue honors his leading role in stochastic optimization, probability theory, and
statistics. His strong law of large numbers (developed with H. Attouch, Z. Artstein, and L. Korf in the 90s) gives the most versatile consistency theory for M-estimators. It relies on his even more fundamental work on approximation theory for optimization and variational problems: Professor Wets coined the term epi-convergence in 1980, which is now accepted as the only “right” notion for approximating minimization problems; it ensures the convergence of optimal solutions and optimal values under the mildest possible assumptions. His contributions started in 1967 with a fundamental result about convergence and distances between convex cones as well as between their polars and include the first results on uniform approximations of sets and the convergence of measurable selections (both in 1981 with G. Salinetti), an Arzela-Ascoli Theorem for set-valued mappings (with A. Bagh in 1996), and a quantitative theory for epi-convergence (with Attouch in the 80s and early 90s).

Professor Wets developed the first algorithm directed specifically at two-stage stochastic optimization, the L-Shape Method, in 1969 with R. van Slyke. The method has been implemented in many different software packages and remains a standard approach to such problems. In his earliest work, Professor Wets recognized the special kind of induced constraints that emerge naturally in multi-stage decision problems and actually coined the term. Later, Wets introduced the term nonanticipativity constraints in multi-stage stochastic optimization problems to describe the constraints that enforce the necessity of making decisions based only on the information available at the time of the decision. This breakthrough led to two 1976-papers (with R.T. Rockafellar) that established duality theory for stochastic programs and fundamental insight about the price of information. Professor Wets developed the Progressive Hedging Algorithm for multi-stage stochastic programs in 1991 (with Rockafellar), which is now implemented in the widely used software Pyomo. Many of Professor Wets’ contributions are summarized and expanded in the seminal treatise entitled Variational Analysis (with Rockafellar in 1998).

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations