EDITORIAL: MEMORIAL ISSUE FOR CHARLES STEIN

BY RICHARD J. SAMWORTH¹ AND MING YUAN²

¹Statistical Laboratory, University of Cambridge, r.samworth@statslab.cam.ac.uk
²Department of Statistics, Columbia University, ming.yuan@columbia.edu

The Institute of Mathematical Statistics (IMS) Council approved a proposal from its Committee on Memorials to dedicate this issue of the Annals of Statistics to Charles M. Stein, who died in 2016 aged 96. This memorialisation is a reflection of Stein’s distinction as a mathematical statistician, whose work continues to have a profound impact on the discipline.

As co-editors, we have solicited four articles on some of the most remarkable and enduring of Stein’s contributions. Eaton and George (2021) describe the earliest of these, namely Stein’s work with Gil Hunt on understanding the relationship between invariant statistical procedures and minimaxity. It was already understood that many statistical problems possess invariance properties with respect to particular groups of transformations, and in such circumstances, it is natural to consider statistical procedures that respect this invariance. As Eaton and George explain, Hunt and Stein set out to understand when the minimax risk over all procedures matches the minimax risk over invariant ones. They cover not only the elegant solution as described by the Hunt–Stein theorem, but also the story of the lost manuscript and the way in which the ideas did finally come to be known.

Strawderman (2021) discusses Stein’s seminal result on the inadmissibility of the usual estimator of a multivariate normal mean in three or more dimensions (Stein (1956b)). Invariance played a significant role in Stein’s thinking here too, in that a spherically symmetric estimator that is admissible with respect to the set of all spherically symmetric estimators turns out also to be admissible among all estimators. Strawderman describes the way in which the 1956 paper became a launchpad for many further investigations into the Stein shrinkage phenomenon and (in)admissibility. Less directly, the idea of using shrinkage to reduce variance (at the expense of introducing some bias) lies at the heart of ridge regression and other penalised likelihood techniques that have become so popular in recent years.

Within the same volume of the famous Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Stein (1956a) laid the foundations for the theory of efficient semiparametric inference, as explained by van der Vaart and Wellner (2021). The authors describe Stein’s insight that when a parameter space is infinite-dimensional (non-parametric), there may be a one-dimensional least favourable subproblem that is as difficult as the original problem of estimating the underlying parameter. This facilitates an extension of the asymptotic theory of estimation in regular parametric models to one for smooth functionals in infinite-dimensional models. Among several examples and extensions, van der Vaart and Wellner present Stein’s classical example of estimating the population median in a location family generated by a symmetric density, showing that absence of knowledge of this symmetric density does not lead to an asymptotic deterioration in the estimation of this median.

Lastly, Chen (2021) outlines Stein’s method for normal approximation, beginning with his characterisation of the standard normal distribution in terms of the vanishing of the expectations of certain test functions. This turns out to spawn a powerful technique for bounding a discrepancy between the distribution of a given random variable and that of a standard normal random variable. Chen includes several anecdotes that give an insight into the way