Statistical Consulting in the Era of Data Science

Lillian Lin

Retired Statistical Consultant
E-mail: ls.lin.mt@gmail.com

Abstract: Statistics emerged from the examination of issues confronted by genetics, betting, and agriculture and is now an essential tool for nearly every scientific field. Yet, there has never been a focused interdisciplinary examination of how to construct a framework for statistical collaboration that satisfies the intellectual values of experts from fields of application as well as the statistics profession. Rather, these experts are satisfied to learn statistical recipes (e.g., how to use software for analyses typically deemed acceptable in their peer reviewed journals) rather than seeking out a statistician who can take responsibility for a study design and data analyses tailored to their specific investigation. Statisticians, in turn, have retreated to inventing ever more specialized methods (an approach supported by our peer reviewed journals) which results in ceding control of the definition of appropriate statistical collaboration to non-statisticians. In recent years, this problem has been compounded by the emergence of the field of data science. Statisticians are often pushed to perform technical tasks that stray far beyond our expertise and fail to receive proper credit for intellectual contributions. Drawing on decades of experience serving as a statistician collaborating with scientists and, more recently, in creating an academic statistical consulting service, I will propose and examine issues that every statistical consulting service, whether an individual consultant or an organization, must address in order to preserve intellectual identity for statisticians and our scientific colleagues and promote effective quantitative research.