Biography of Pengfei Li

Born in 1979, Pengfei was raised in a small village in the Chinese province of Hubei. He studied Statistics at the Nankai University in China, where he obtained his Bachelor's and Master's degrees in Statistics in 2001 and 2004, respectively. Pengfei's master thesis considers optimal fractional factorial designs and their construction; it results in 7 publications in reputable statistical journals. In addition to finishing his master thesis, Pengfei worked as research assistant with Dr. Donggeng Wang at Hong Kong Baptist University for four months. During this period, they successfully applied uniform design to transportation study; their work leads to two publications in transportation journals.

Pengfei was admitted to the University of Waterloo for graduate studies in 2004, where he completed his Ph.D. in three years and four months, under the joint supervision of Professor Jiahua Chen (the 2014 SSC Gold Medalist) and Professor Paul Marriot. After postdoctoral studies with Professor Jiahua Chen at the University of British Columbia, Pengfei joined the Department of Mathematical and Statistical Sciences at the University of Alberta in 2008. The Department of Statistics and Actuarial Science at the University of Waterloo successfully recruited Pengfei back in 2012; he was promoted to the rank of Associate Professor in 2014 and the rank of Full Professor in 2019.

In 15 years after his Ph.D., Pengfei has established himself as an internationally renowned scholar with an outstanding research portfolio in the areas of finite mixture models, empirical likelihood, density ratio model, capture-recapture problems, and non-probability survey samples. Pengfei has around 70 journal publications and 2 book chapters. Of these, 8 appeared in The Canadian Journal of Statistics (CJS), and 18 appeared in Annals of Statistics (AOS), Biometrika (Bio), Biometrics, Journal of the American Statistical Association (JASA), and Journal of the Royal Statistical Society: Series B (JRSSB).

Among his original, significant, and impactful contributions, Pengfei's work in finite mixture models deserves special mention as detailed in all supporting letters for the nomination. Solving statistical problems under mixture models is excruciatingly difficult because of their lack of regularity properties. The early asymptotic results often stop at the proof-of-concept level, failing to develop concrete inference procedures. Pengfei's pioneering work on EM-tests (Li, Chen and Marriott, 2009, Bio; Chen and Li, 2009, AOS; Li and Chen, 2010, JASA; Chen, Li and Fu, 2012, JASA) has significantly pushed the boundary. To test the order of the finite mixture model, the new EM-tests examine how quickly rather than by how much the likelihood increases from the null to the alternative models. This has led to a reduced level of technical sophistication and produced many neat asymptotic results. More importantly, the theoretical work has led to efficient EM-tests that are readily implementable.

The EM-tests, like LASSO, require the specification of tuning parameter values. There are few examples to follow since the values of the tuning parameters do not affect the first-order asymptotics. Chen and Li (2011, CJS) instead invented the computer experiment approach for this purpose. The creativity and originality of this

contribution have been recognized in one supporting letter for the nomination, whose comments are quoted as ``These methods should also be of value in other areas where tuning parameters need to be specified.''

The semiparametric density ratio model (DRM) is a flexible platform for combining information from multiple sources, and it permits elegant inference solutions through the empirical likelihood. Pengfei and his collaborators advance the use of the DRM to several important areas of research and application. Li and Qin (2011, JASA) and Li, Liu and Qin (2017, JASA) employ the DRM to solve the unordered homologous chromosome-pair problem and form genetic semiparametric mixture models. They discovered that the asymptotic properties of the EL ratio test depend on the degenerateness of the Fisher information matrix. As highlighted in one supporting letter for the nomination, ``This is a unique observation that provides a fully satisfactory solution to the applied problem and a new area for theoretical research.'' Qin, Zhang, Li, Albanes, and Yu (2015, Bio) illustrated the potential of using the DRM to incorporate existing data from large cohort studies to improve improves the efficiency the analysis of a new study. In the same supporting letter mentioned above, it says ``The existence of such auxiliary information is common and this opens a potentially important area of research.''

Pengfei's path breaking Biometrika paper, Liu, Li and Qin (2017), revolutionizes the analysis of the capture-recapture data by utilizing auxiliary information through empirical likelihood. It leads to a much more satisfactory solution to the confidence interval for the population size. A referee urged them to broadcast their method widely to real-world users.

There is no sign that Pengfei is forming a comfort zone using these amazing achievements. His latest work on analyzing data from non-probability survey sampling breaks new ground (Chen, Li, and Wu, 2020, JASA). The paper develops a general framework, based on propensity scores, for incorporating non-probability survey samples into inference based on a properly designed survey sample. This is a novel approach in a very important area as non-probability survey samples are so easy to collect in the era of big data.

Pengfei's research excellence has been recognized by his NSERC funding and three Outstanding Performance Awards from the University of Waterloo. He also excels in statistical education: he has twice won teaching awards. He is an associate editor of The Canadian Journal of Statistics and Metrika. He has served on the organizing committees of eight international conferences. He is in charge of the scientific program of the SSC 2022 annual meeting.

Pengfei credits his success to his family, which has been fully supportive for his career, and to his inspirational and wonderful supervisor, mentors, collaborators, colleagues, and students. Pengfei and his wife, Weihong, have a son, Daniel, and a daughter, Katelyn.

The citation for the award reads:

"To Pengfei Li, for ground-breaking and pioneering research contributions to the EM-test for the order of finite mixture models; for original and creative methodological developments in the areas of the empirical likelihood, density ratio models, statistical genetics, non-probability sampling; for exceptional research productivity; and excellence in statistical education."