Abstract: Regression remains an important framework in the era of big and complex data. In this talk, I will present some recent examples where we resort to the simple linear regression and its celebrated extensions in novel settings, including generalizing Kruskal-Wallis, Levene’s, and Pearson’s Chisq Hardy-Weinberg equilibrium tests. The Eureka moment came while reading Wu and Guan's (2015) comments on our generalized Kruskal-Wallis test (Acar and Sun 2013, Biometrics). Wu and Guan presented an alternative “rank linear regression model and derived the proposed GKW statistic as a score test statistic". Indeed, the regression framework eases the derivation and facilitates further extensions. More recently, we turned our attention to extending Levene's variance test for dependent sample with group uncertainty; this test is useful for latent Gene-Environment interaction analysis where data on E is unavailable. While a direct modification of the original test statistic is challenging, I will demonstrate how a two-stage regression framework eases the extension, leading to a generalized joint location-scale test (Soave and Sun 2017, Biometrics). Finally, I will discuss how to formulate the classical Pearson’s Chisq test of Hardy-Weinberg equilibrium as a linear regression, and how this formulation can be utilized to generalize the locally most powerful allelic association test (Zhang and Sun 2022, Biometrics), and develop methods for the complex X chromosome that is often omitted in the current genome-wide association studies and beyond. In summary, the crux of each work is reformulating the problem as a regression!

Bio: Dr. Lei Sun is a Professor in Statistics and Biostatistics at the University of Toronto. She studied mathematics at Fudan University and obtained her Ph.D. in statistics from the University of Chicago in 2001. Sun’s research area is in statistical genetics and genomics, and her research is funded by both the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Institutes of Health Research (CIHR). In 2017, Sun was honored to be the third female statistician to have received the prestigious CRM-SSC Prize in Statistics, awarded annually since 1999. In 2018, Sun was the recipient of the 2018 NSERC Discovery Accelerator Supplements Program Award, and in 2020, she served as the President of the Biostatistics Section of the Statistical Society of Canada.