On Jiang's asymptotic distribution of the largest entry of a sample correlation matrix

Deli Li
(joint with Professors Yongcheng Qi and Andrew Rosalsky)

Department of Mathematical Sciences
Lakehead University, Canada
dli@lakeheadu.ca

Abstract: Let \( \{X, X_{k,i}; i \geq 1, k \geq 1\} \) be a double array of nondegenerate i.i.d. random variables and let \( \{p_n; n \geq 1\} \) be a sequence of positive integers such that \( n/p_n \) is bounded away from 0 and \( \infty \). This work is devoted to the solution to an open problem posed in Li, Liu, and Rosalsky (2010) on the asymptotic distribution of the largest entry \( L_n = \max_{1 \leq i < j \leq p_n} \left| \hat{\rho}_{i,j}^{(n)} \right| \) of the sample correlation matrix \( \Gamma_n = \left( \hat{\rho}_{i,j}^{(n)} \right)_{1 \leq i,j \leq p_n} \) where \( \hat{\rho}_{i,j}^{(n)} \) denotes the Pearson correlation coefficient between \( (X_{1,i}, \cdots, X_{n,i})' \) and \( (X_{1,j}, \cdots, X_{n,j})' \). We show under the assumption \( \mathbb{E}X^2 < \infty \) that the following three statements are equivalent:

1. \( \lim_{n \to \infty} n^2 \int_{(n \log n)^{1/4}}^{\infty} \left( \left( \frac{n \log n}{x} \right)^{-1} - \left( \frac{\sqrt{n \log n}}{x} \right)^{-1} \right) dF(x) = 0, \)

2. \( \left( \frac{n \log n}{\log n} \right)^{1/2} L_n \overset{p}{\to} 2, \)

3. \( \lim_{n \to \infty} \mathbb{P} \left( nL_n^2 - a_n \leq t \right) = \exp \left\{ -\frac{1}{\sqrt{8\pi}} e^{-t^2/2} \right\}, \quad -\infty < t < \infty \)

where \( F(x) = \mathbb{P}(|X| \leq x), x \geq 0 \) and \( a_n = 4 \log p_n - \log \log p_n, n \geq 2 \). To establish this result, we present six interesting new lemmas which may be beneficial to the further study of the sample correlation matrix.