

STANFORD UNIVERSITY
DEPARTMENT OF STATISTICS
DEPARTMENTAL SEMINAR

4:15 p.m., Tuesday, November 25, 2003
Sequoia Hall Room 200
(Cookies at 3:45 in 1st Floor Lounge)

Noureddine El-Karoui
Stanford University

The Tracy-Widom law holds when $n, p, p/n \rightarrow \infty$, with application to PCA

Abstract:

Principal Component Analysis (PCA) is a tool used across the spectrum of scientific applications. In modern practice, it is often applied to $n \times p$ data matrices with n and p both large. Classical theory (Anderson 1963) fails to apply in this setting. Using random matrix theory, Johnstone (2000) recently shed light on some theoretical aspects of PCA in this setup. Specifically, when the entries of the $n \times p$ matrix X are iid $\mathcal{N}(0, 1)$ and $n/p \rightarrow \rho \in (0, \infty)$, he showed that $\lambda_{n,p}$, the largest eigenvalue of the empirical covariance matrix $X'X$, converges to the so-called Tracy-Widom distribution (after proper recentering and rescaling).

We will show that the result holds when $n, p \rightarrow \infty$ and $n/p \rightarrow 0$ or ∞ , in effect removing the need to worry about the limiting behavior of n/p . We will also present preliminary results for rates of convergence. Finally, we will illustrate how these and related theoretical insights might be used in practice.