Seminar November 21, 2007

Robustifying the Hill estimator

Keith Knight

University of Toronto

Abstract

Suppose that $X_1, \cdots, X_n$ are i.i.d. random variables with $P(X_i > x) = x^{-\alpha}L(x)$ and define $X_{(1)} \geq X_{(2)} \geq \cdots \leq X_{(n)}$. The Hill estimator (Hill, 1975) of the tail index α is a pseudo-maximum likelihood estimator based on the exponential approximation of the normalized log-spacings $Y_j = j\ln(X_{(j)}/X_{(j+1)})$ for $j = 1, \cdots, k$. In practice, the Hill estimator can be extremely dependent on the choice of k and is inherently non-robust to large values Y_j, which bias the Hill estimator downward. In this paper, we introduce a simple robustification of the Hill estimator that has a bounded influence curve and is Fisher consistent under the exponential model. The estimator is straightforward to compute and can be tuned to have a specified asymptotic efficiency (with respect to the Hill estimator) between 0 and 1. We will also consider extensions of the exponential regression modifications of the Hill estimator (Feuerverger and Hall, 1999; Beirlant *et al.*, 1999) using smoothing.