

Extreme value methods in informetrics/scientometrics

Goedele Dierckx

Joint work with:

J. Beirlant, W. Glänzel – Katholieke Universiteit Leuven
J.H.J. Einmahl – Tilburg University

Abstract

The last decade methods for quantifying the research output of individual researchers have become quite popular in research and university policy making. The Hirsch-index h (Hirsch, 2005) constitutes an interesting quality measure that has attracted a lot of attention recently and is now a standard measure available for instance on the Web of Science.

Let X_1, \dots, X_n be i.i.d. random variables with common distribution function F , denoting the numbers of citations of the n articles of a certain researcher. Then the hirsch index h is defined as

$$F(h) = 1 - \frac{h}{n}.$$

In recent papers (Glänzel, 2006, Egghe and Rousseau, 2006, and Burrell, 2007) the first theoretical properties of the hirsch index were derived. The underlying citation distribution was then assumed to be of Pareto type. In

extreme value theory these Pareto type distributions have been studied extensively. As such the techniques from this field can be used to propose new estimators and properties of the hirsch index.