On the Evolution of Surnames

John Panaretos

School of Engineering, Division of Applied Mathematics, University of Patras, P.O. Box 1325, Patras, Greece

Summary

The problem of determining how family names evolve preoccupies both statistics and human biology. The determination of a proper and well justified probability model to describe the probability distribution of surnames has been confronted by many authors. In this paper two stochastic models giving rise to the Yule distribution are proposed to explain and fit some observed surname frequency distributions. The first is based on a contagion hypothesis in the sense that the more occurrences a surname has had the more likely it is to have a further occurrence. The second model is based on a weaker set of assumptions which also allows 'immigration' of new surnames. The distribution that arises from these models is then fitted to actual data and the fit is compared to that provided by the discrete Pareto distribution.

Key words: Discrete Pareto distribution; Surname distribution; Yule distribution.