An Extreme Value Analysis of Advanced Age Mortality Data

Kathryn A. Robertson
Debbie J. Dupuis, PhD
Bruce L. Jones, PhD, FSA, FCIA

Abstract

Extreme value theory describes the behavior of random variables at extremely high or low levels. The application of extreme value theory to statistics allows us to fit models to data from the upper tail of a distribution. This paper presents a statistical analysis of advanced age mortality data, using extreme value models to quantify the upper tail of the distribution of human life spans.

Our analysis focuses on mortality data from two sources. Statistics Canada publishes the annual number of deaths in Canada, broken down by gender and age. We use the deaths data from 1949 to 1997 in our analysis. The Japanese Ministry of Health, Labour and Welfare also publishes detailed annual mortality data, including the 10 oldest reported ages at death in each year. We analyze the Japanese data over the period from 1980 to 2000.

Using the r-largest and peaks-over-threshold approaches to extreme value modeling, we fit generalized extreme value and generalized Pareto distributions to the life span data. Changes in distribution by birth cohort or over time are modeled through the use of covariates. We then evaluate the appropriateness of the fitted models, and discuss reasons for their shortcomings. Finally, we use our findings to address the existence of a finite upper bound on the life span distribution, and the behavior of the force of mortality at advanced ages.