

Nonparametric estimation of unrestricted distributions and their jumps
Mynbaev, Martins-Filho and Henderson

Layman's abstract

Many disciplines apply statistical techniques to their data. However, the assumptions made by theoretical statisticians often do not hold true. For nonparametric estimation of distribution functions, most statistical theory requires two continuous derivatives (smoothness). In areas such as biomedicine, economics and finance, this assumption is often not valid. For example, in economics, when looking at wage distributions, it is common to find many individuals with hourly wages listed in integer amounts. This phenomenon not results in jumps in the distribution. In this paper, the authors relax smoothness restrictions on the distribution. By relying on Stieltjes integrals, they are able to construct estimates of the distribution and potential jumps in the distribution that are consistent and asymptotically normal. Perhaps more important is the fact that the authors are able to provide explicit estimates for convergence rates, including new inversion theorems for characteristic functions. The proposed estimators perform well both in simulations and empirical examples. The first empirical example looks at the distribution of land elevation data where there is a natural jump in the distribution due to the presence of lakes in the area under consideration. The second example looks for jumps (in the distribution of reported test statistics in top economic journals) that are artificial in nature due to the so-called “p-hacking” phenomenon in empirical research (the misuse of statistical analysis whereby researchers perform many statistical tests and only report those that are statistically significant). These methods should prove useful in practice as researchers will now be able to smoothly estimate these distributions with knowledge of convergence rates.