Stockholm University

Mathematical Statistics
Stockholm University
Master Thesis **2018:2**
http://www.math.su.se

Characterizing Processing Times with Mixture Distributions

Simon Berggren*

March 2018

Abstract

In this text, we characterize the processing times of a large number of errands. Our goal with this characterization is to provide certain key values, such as a number of quantiles and the expected value. We propose a (finite) mixture distribution, more specifically a weighted mixture of a lognormal distribution and a truncated normal distribution. Goodness of fit of this mixture model is tested using the Kolmogorov- Smirnof statistic. We also develop a methodology, using maximum likelihood techniques and an iterative Expectation Maximization (EM) algorithm, for estimation of the parameters that define the mixture distribution. Occasionally, our processing times are only partially observed, i.e. some of the values are unknown at the time of estimation. To deal with this, an extension of the Stochastic EM (SEM) algorithm is adapted, and combined with the above mentioned estimation methodology for fully observed processing times. In this extended algorithm, repeatedly updated intermediate estimates, based on both the observable and randomly generated data, are weighted to produce the parameter estimates. Computer programs that are needed for parameter estimation, computation of the key values, evaluation of the models, etc. are implemented in C++ and R as a part of the work. The outlined methodology is assessed on simulated data, and finally applied to subsets of our real word data set.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: simon.berggren@gmail.com. Supervisor: Ola Hössjer.