Implementation of the reparameterized inverse Gaussian distribution for inferential interpretability

Amanda van Dyk 18005889

WST795/STK795 Research Report

Submitted in partial fulfillment of the degree BSc(Hons) Mathematical Statistics

Supervisor(s): Prof A Bekker and Dr JT Ferreira

Department of Statistics, University of Pretoria

2021/10/29

1 A description of all the code used in the report and found in this folder:

This repository contains all the R Software code we used to obtain all the graphs and results in our research report.

The file called simulations study contains the code we used to simulate the contaminated reparameterized Inverse Gaussian (rIG) distributions and the EM algorithm we used to then estimate the parameters of the distribution.

The files labelled with SA or IT are all part of the code we used on the LIS [1] interface to obtain our analysis of the three income datasets. SA stands for South Africa and IT for Italy and the number after it indicates the year of the income data. The data was obtained from the Luxembourg Income Study [1].

All the files that end with "EDA" is the code we used to perform an exploratory data analysis of the income datasets. It includes code for calculating the mode and it shows how the weighted histograms were drawn.

All the files that contain "MWL contaminated" in the file name contains the code we used to estimate the the parameters of the contaminated rIG distribution that we fitted to the datasets using maximum weighted likelihood. It also includes the code to calculate the dAIC values.

All the files that contain "MWL rIG" in the file name contains the code we used to estimate the the parameters of the rIG distribution that we fitted to the datasets using maximum weighted likelihood. It also includes the code to calculate the dAIC values.

All the files that contain "fitted model" in the file name contains the code we used to obtain the graphs of the models we fitted to the datasets.

The two files that contain the word "graphs" in the file name contains the code we used to draw the graphs of the rIG, unit rIG and contaminated rIG density curves.

2 Bibliography

[1] Luxembourg Income Study (LIS) Database, http://www.lisdatacenter.org (multiple countries; 01/07/2021- 29/10/2021). Luxembourg: LIS.