S. ARIMA, S. POLETTINI, G. PASCULLI, L. GESUALDO, F. PESCE, D.A. PROCACCINI (2023) A Bayesian nonparametric approach to correct for underreporting in count data, submitted to Biostatistics

We provide the nimble code together with R functions to implement Bayesian inferences for underreported counts data.

Files contained in this repository can be used to obtain the results contained in Section 6 of the paper. Real data cannot be released for disclosure limitation but we provide synthetic data mimicking the CKD data. However, we provide synthetic data that are stored in the file CKD-PseudoData.rds.

The file "Proposed-Model-CKD.txt" is the model specification for the proposed methodology to be implemented through nimble.

R functions for importing the data and preparing all lists necessary for estimating the model using the nimble package have been provided in the file "Implementation.R". This file also contains convergence diagnostic plot and indices as well as some commands for producing plots and maps.

The file geoinfo.rds contains latitude, longitude and the shape file of the provinces of the Apulia region. This information are useful for drawing the maps (Figure 6 in the paper).

The file "R-Code-SimulationStudy.txt" reproduces the simulation scenarios in Section 5 of the paper: the nimble codes of the proposed model, the model in the De Oliveira et al. (2022) and the Poisson model are respective coded in the files "Proposed-Model.txt", "Lopes-Model.txt" and "Poisson-Model.txt". One could obtain the results of the three simulation scenarios by changing the value of epsilon, as stressed in the commented R code.