JASA A&CS Reproducibility Initiative - Author Contributions Checklist Form

The purpose of the Author Contributions Checklist (ACC) Form is to document the code and data supporting a manuscript, and describe how to reproduce its main results.

1. Data

Abstract

We have included a synthetic data set, "Simulated_Data.csv, which was generated by the design detailed in Section 3 of the main paper with q=2 episodic components, p=1 regular component, n=1000 subjects, and m=3 replicates per subject. The synthetic dataset itself is also analyzed in the main paper. The data file comprises a data matrix Data=[Inds V1 V2 V3] with 1000 rows and 4 columns. The column contents are given below.

- i**nds** = an N(=n*m) element vector containing identifying labels for different subjects,
  V1 = the subject and sampling occasion specific recalls for the first episodic component, containing exact zeros.
- **V2** = the subject and sampling occasion specific recalls for the second episodic component, containing exact zeros.
- **V3** = the subject and sampling occasion specific recalls for the only regular component, containing all continuously measured values.

inds	V1	V2	V3
1	0.758299157	0.000000000	1.176854626
1	0.510284356	2.929304830	1.785666686
1	0.961566867	5.208036103	1.453759331
2	1.304722389	3.702188008	2.120903315
2	1.880341668	3.536992553	0.606317593
2	1.066852224	3.653615302	2.239304921
3	0.274565920	1.133716282	2.977757786
…	…	…	…

Availability

The real EAT dataset analyzed in the main paper can be accessed from the National Cancer Institute arranging a Material Transfer Agreement.

2. Codes

Abstract

The codes are written in R.
The codes comprise three main programs.

(a) **Bayes_Copula_Decon_MVT.R** - Implements the Bayesian multivariate copula-based deconvolution model developed in the main paper.

(b) **Bayes_Copula_Decon_Univariate_Episodic.R** - Implements a univariate deconvolution model for zero inflated data. Results produced by this method are used as initial values for the multivariate sampler. The method can also be independently used univariate deconvolution problems with zero-inflated data.

(c) **Bayes_Copula_Decon_Univariate_Regular.R** - Implements a univariate deconvolution model for continuously measured data. Results produced by this method are used as initial values for the multivariate sampler. The method can also be independently used for univariate deconvolution problems.

There are two additional files.

(d) **Bayes_Copula_Decon_Functs.R** - Contains utility functions used by the main programs.

(e) **Bayes_Copula_Decon_MVT_Run.R** - Calls the main function in the file Bayes_Copula_Decon_MVT.R to run the multivariate deconvolution model. Illustrated using the included synthetic dataset.

The implementation is highly automated - the main function implementing the multivariate copula deconvolution model takes in the data matrix (as described above) as an argument and a few additional parameters with default recommended values. Additional descriptions and instructions are included as detailed comments in the body of the codes.

Description
The codes are included in a zipped file. The file to run is **Bayes_Copula_Decon_MVT_Run.R**. The file calls functions that are also included in the zipped file. To run the codes, all files should be included the working directory.

The following R libraries are utilized and have to be pre-installed.
    foreach
    doParallel
    ks
    mvtnorm
    MCMCpack
    msm
    corpcor
    RColorBrewer

Reproducibility
The results for a simulation scenario described in Section 3 of the main paper can be reproduced using the codes and the included synthetic dataset. These include parts of Figures 5, 6 and 7 in the main paper and Figure S.2 in the supplementary material that summarized results produced by the copula deconvolution model.