Author Contributions Checklist Form

This form documents the artifacts associated with the article (i.e., the data and code supporting the computational findings) and describes how to reproduce the findings.

Part 1: Data

☐ This paper **does not** involve analysis of external data (i.e., no data are used or the only data are generated by the authors via simulation in their code).

☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Abstract

An earthquake data set for the period between 1885 and 1980 for a region close to Japan is used in Section 7.1; this data is taken from the cited paper from Ogata.
An Rocky Mountain Spotted Fever data set for California and Florida, raning from 1960 to 2011 is used; this data is also ued by the cited Schoenberg paper. The data is available from Project Tycho (open source).

Availability

☒ Data **are** publicly available
☐ Data **cannot be made** publicly available

If the data are publicly available, see the *Publicly available data* section. Otherwise, see the *Non-publicly available data* section, below.

Publicly available data

☐ Data are available online at:
☐ Data are available as part of the paper's supplementary material.
☒ Data are publicly available by request, following the process described here:

The data for the Ogata 1988 paper is provided in Table 1 of that paper; see pp. 14-15. Also see the file Ogata1988dataRAW.xlsx.
Simply make a free account on https://www.tycho.pitt.edu/accounts/login/. After this, the data for the Schoenberg paper can be downloaded.

☐ Data are or will be made available through some other mechanism, described here:

Non-publicly available data

Discussion of lack of publicly available data:

Description

File format(s)

☒ CSV or other plain text:
☐ Software-specific binary format (.Rda, Python pickle, etc.):
☐ Standardized binary format (e.g., netCDF, HDF5, etc.):
☒ Other (described here):

.xlsx

Data dictionary

☐ Provided by the authors in the following file(s):
☐ Data file(s) is (are) self-describiing (e.g., netCDF files)
☐ Available at the following URL:

Additional information (optional)

Data is described in Statistical models for earthquake occurrences and residual analysis for point processes, Ogata (1988) and A recursive point process model for infectious diseases, Schoenberg (2019). Apart from that, it is selfexplanatory, and instructions are given in TestingOgata.m and testingSchoenberg.m

Part 2: Code

Abstract

For simulations, we simulate sample paths, and using the simulated paths we perform Algorithm 1 from our paper. We do this for a large number of sample paths to find rejection rates and evaluate performance of our method under H0 and H1.

For the testing procedures, we import the data, preprocess it. For Ogata, we take their estimates, and compare tests based on random time change and our tests. For Schoenberg, we estimate the parameters and compare our tests to random time change based tests.

Description

Code format(s)

☐ Script files
- ☐ R ☐ Python ☒ Matlab
- ☐ Other:

☐ Package
- ☐ R ☐ Python ☒ MATLAB toolbox
- ☐ Other:

☐ Reproducible report
- ☐ R Markdown ☐ Jupyter notebook
- ☒ Other: txt file

☐ Shell script

☐ Other (described here):

Supporting software requirements

Version of primary software used

Matlab R2023a

Libraries and dependencies used by the code

cmtest.m, Cramer-von Mises test

Version 1.1.0.0 (17.5 KB) by Ahmed BenSaïda

Supporting system/hardware requirements (optional)

Parallelization used

☐ No parallel code used
☒ Multi-core parallelization on a single machine/node
Number of cores used: 8
☐ Multi-machine/multi-node parallelization
Number of nodes and cores used:

License

☒ MIT License (default)
☐ BSD
☐ GPL v3.0
☐ Creative Commons
☐ Other (described here):

Additional information (optional)

Part 3: Reproducibility workflow

Scope

The provided workflow reproduces:

☐ Any numbers provided in text in the paper

☒ The computational method(s) presented in the paper (i.e., code is provided that implements the method(s))

☒ All tables and figures in the paper

☒ Selected tables and figures in the paper, as explained and justified here:

Table1seeded.m can be run directly to reproduce the results from Table1 exactly. AAA_Replication_instructions.txt outlines how the remaining tables and figures can be reproduced; which files are needed and what needs to plugged in.

Workflow details

Location

The workflow is available:

☐ As part of the paper's supplementary material

☒ In this Git repository: https://github.com/JRB-UvA/GoFreproducibility

☐ Other:

Format(s)

☐ Single master code file

☐ Wrapper (shell) script(s)

☐ Self-contained R Markdown file, Jupyter notebook, or other literate programming approach

☒ Text file (e.g., a readme-style file) that documents workflow

☐ Makefile

☐ Other (more detail in 'Instructions' below)

Instructions

AAA_Replication_instructions.txt

Expected run-time

Approximate time needed to reproduce the analyses on a standard desktop machine:

☐ <1 minute

☐ 1-10 minutes

☐ 10-60 minutes

☐ 1-8 hours

☒ >8 hours

☐ Not feasible to run on a desktop machine, as described here:

On my pc with AMD Ryzes 7 5800H 8 core:
Table 1, 4 6 will take <30 minutes each, 2, 3 and 5 will take at most 3 days each. Results of Section 7 can be retrieved by less than 10 minutes computation time.

Additional documentation (optional)

Notes (optional)