Package 'COLP'

January 20, 2025

Type Package

Title Causal Discovery for Categorical Data with Label Permutation

Version 1.0.0

Date 2022-09-23

Description Discover causality for bivariate categorical data. This package aims to enable users to discover causality for bivariate observational categorical data. See Ni, Y. (2022) <arXiv:2209.08579> ``Bivariate Causal Discovery for Categorical Data via Classification with Optimal Label Permutation. Advances in Neural Information Processing Systems 35 (in press)".

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

Imports MASS, combinat, stats

URL https://github.com/nySTAT/COLP

BugReports https://github.com/nySTAT/COLP/issues

NeedsCompilation no

Author Yang Ni [aut, cre] (<https://orcid.org/0000-0003-0636-2363>)

Maintainer Yang Ni <yni@stat.tamu.edu>

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2022-09-29 08:40:12 UTC

Contents

CatPairs *Categorical Cause-Effect Pairs*

Description

Cause-effect pairs extracted from R packages MASS and datasets for which the pairwise causal relationships are clear from the context, and at least one of the variables in each pair is categorical. For non-categorical variable, we discretized it at 5 evenly spaced quantiles.The current version contains 33 categorical cause-effect pairs.

Usage

```
data(CatPairs)
```

Format

A list of length 2. The first element is a list of 33 cause-effect pairs as data frames with the first column being the cause and the second column being the effect. The second element is a list of sources of each pair.

COLP *Causal Discovery for Bivariate Cateogrical Data*

Description

Estimate a causal directed acyclic graph (DAG) for ordinal cateogrical data with greedy or exhaustive search.

Usage

```
COLP(y, x, algo = "E")
```

Arguments

y	factor, a potential effect variable
x	factor, a potential cause variable
algo	exhaustive search (algo="E") of category ordering or greedy search (algo="G")

Value

A list of length 3. cd = 1 if x causes y; cd = 0 otherwise. P is the optimal odering of the effect variable. epsilon is the difference in log-likelihood favoring x causes y.

Examples

```
fit = COLP(CatPairs[[1]][[1]]$Diffwt,CatPairs[[1]][[1]]$Treat,algo="E")
fit$cd
```

Index