

R パッケージ catdap (Ver. 1.3.7-1)

1. はじめに

CATDAP (CATegorical Data Analysis Program)は、統計数理研究所で開発された最適な分割表（クロス表）の探索のためのプログラムである。CATDAP には、CATDAP-01 と CATDAP-02 とがある。CATDAP-01 は、カテゴリカルな（質的）データに対し二次元分割表の比較を行い各変数の間の関係の深さを検出するプログラムである。全ての変数がカテゴリカルであることが前提となる。CATDAP-02 は、ひとつの着目した項目（目的変数）を固定し他の項目（説明変数）の組合せで多次元分割表を作り、目的変数の分布の違いを最も適切に説明する説明変数組合せを探索する。このプログラムは量的な項目も適当に区分してカテゴリカルなデータに帰着することにより、変数が量的か質的にかかわらず適用できる。いずれも最適な説明変数の選択には、AIC (Akaike Information Criterion) が使われている。

R パッケージ catdap の関数 catdap1()と catdap2()は、それぞれ CATDAP-01 と CATDAP-02 を基にしている。CATDAP-02 では目的変数が量的な項目であるデータには適用できなかったが、catdap2() では機能を強化し

- ・ Base AIC (=説明変数なしモデルの AIC) の利用
- ・ 連続値目的変数に適用
- ・ 目的変数, 説明変数に欠測が含まれるデータに適用

が可能となった。BaseAIC の利用により、このプログラムであてはめられたモデルと、目的変数に正規分布を仮定する回帰モデル あるいは ロジスティック回帰モデルとの AIC の比較が可能である。

2. catdap の使い方

(1) R のインストール

CRAN (The Comprehensive R Archive Network) のホームページ (<https://cran.r-project.org/>) の Download and Install R の項目から利用 OS にあわせてインストーラーをダウンロードする。ダウンロードしたインストーラーを実行し、指示に従ってインストールを完了させる。

(2) catdap のインストールと読み込み

利用 OS に合わせて以下のいずれかをダウンロードする.

- ・ パッケージ・ソース catdap_1.3.7-1.tar.gz ,
- ・ Windows バイナリ (r-release) catdap_1.3.7-1.zip
- ・ MacOS バイナリ (r-release (arm64)) catdap_1.3.7-1.tgz

(補足)

catdap パッケージは 2016 年 CRAN に登録され, 2025/10/1 時点 (最新リリース R-4.5.1) では, catdap_1.3.7 がダウンロードファイルとしてリンクされていた. 本ページにリンクした catdap_1.3.7-1 は, -fimplicit-none オプションを付けてコンパイルしたときに出力された警告に関して修正したバージョンである.

R を起動し,

```
> install.packages("catdap_1.3.7-1.tar.gz", repos = NULL, type = "source")
```

または, Windows ならば

```
> install.packages("catdap_1.3.7-1.zip", repos=NULL)
```

を実行して catdap をインストール.

(RGui/パッケージ/Install packages(s) from local files... からインストールできる)

```
> library(catdap)
```

を実行して catdap を読み込む.

3. 参考文献

- [1] Y.Sakamoto and H.Akaike (1978). Analysis of Cross-Classified Data by AIC. Ann. Inst. Statist. Math., 30, pp.185-197.
- [2] K.Katsura and Y.Sakamoto (1980). A Categorical Data Analysis Program Package, Computer Science Monographs, No.14. The Institute of Statistical Mathematics, Tokyo.
- [3] Y.Sakamoto (1985). Categorical Data Analysis by AIC, Kluwer Academic publishers.
- [4] (株) NTT データ数理システム (2015). 情報量統計学的データ可視化ツール.
<http://hdl.handle.net/10787/3614>
- [5] 石黒 真木夫 (2016) CATDAP マニュアル. <http://hdl.handle.net/10787/3821>
- [6] 石黒 真木夫 (2016) 統計モデル可視化. <http://hdl.handle.net/10787/3823>
- [7] 石黒 真木夫 (2021) CATDAP 機能強化プラン.
<http://hdl.handle.net/10787/00034178>