

# R を用いた外国語教育データの整理・要約

川口 勇作（愛知学院大学）

ykawa@dpc.agu.ac.jp

外国語教育メディア学会 第 57 回全国研究大会ワークショップ @名古屋学院大学

2017/8/5

## 概要

外国語教育の研究や実務において、手元にあるデータを整理・要約することは、データの有り様をわかりやすく伝える上で必要不可欠です。そこで本ワークショップでは、フリーの統計解析環境である R を用いて、データの整理・要約の基本的な手法を習得することを目指します。まず R の導入を行い、その後、記述統計量（平均や標準偏差など）の算出、度数分布図の作成などについて、解説・実習を行います。

## R の基本知識・操作

- 命令を入力して Enter を押す
- #から始まる行はコメントアウトされる
- 使用する値は、矢印(<- / ->)やイコール(=)を用いて、変数に入れる  

```
> hako <- c(1, 2, 3, 4, 5)
```
- 行列から要素を取り出すには、**変数名[行数, 列数]** というように指定する
- データフレームから要素を取り出すには、**変数名\$列名** というように指定する

## クリップボードからデータ読み込み

1. まず、Excel 上で R に読み込みたいデータの範囲を選択し、コピー
2. そこから
  - Windows の場合：  

```
> read.table("clipboard", header= T)
```
  - Mac の場合：  

```
> read.table(pipe("pbpaste"), header = T)
```
  - 見出し行の有無の指定（指定がない場合は見出し行ありとして認識）  
ありの場合： header = T      なしの場合： header = F



- **psych** パッケージの **describe** 関数を用いて記述統計を一通り算出  
 ※引数はデータフレーム・行列でも可
  - 標本サイズ・平均・標準偏差・中央値・刈り込み平均・中央絶対偏差・最小値・最大値・データ範囲・歪度・尖度・標準誤差が算出される

```
> library(psych)
> describe(変数名)
```
- 度数分布表を作成する: **table(変数名)** ※引数はベクトル

```
> table(dat1[,1]) #身長 of 度数分布表
> table(dat1[,2]) #体重 of 度数分布表
```
- ヒストグラム (度数分布図)を作成する: **hist(変数名)** ※引数はベクトル

```
> hist(dat1[,1]) #身長 of ヒストグラム
> hist(dat1[,2]) #体重 of ヒストグラム
```
- 箱ひげ図 (box plot)を作成する/蜂群図 (bee swarm plot)を重ね描きする

```
> words <- read.csv("words.csv", header = T)
> boxplot(words)
> (
)
> library(beeswarm)
> beeswarm(words, add = T)
```

## 二変量データの要約

- 散布図を作成する

```
> toEIC <- read.csv("toEIC.csv", header = T)
> plot(toEIC$score, toEIC$words)
```
- 相関係数・相関行列を算出する
  - method では、"pearson", "spearman", "kendall" のいずれかを指定 (デフォルトは"pearson")

```
> cor(toEIC$score, toEIC$words, method = "相関係数の種類") # 相関係数
> cor(toEIC, method = "相関係数の種類") # 相関行列
```
- 相関係数・無相関検定・相関係数の信頼区間を算出する
  - method が"spearman", "kendall"の場合は、信頼区間は算出されない

```
> cor.test(toEIC$score, toEIC$words, method = "相関係数の種類")
```

## 困ったときのレファレンス

- seekR <http://seekr.jp/>
  - R 言語のための検索エンジン
- RjpWiki <http://www.okadajp.org/Rwiki/>
  - R に関する情報交換を目的とした Wiki
- R-Tips <http://cse.naro.affrc.go.jp/takezawa/r-tips/r.html>
  - R のコマンドがわかりやすくまとめられている
- R による統計処理 <http://aoki2.si.gunma-u.ac.jp/R/>
  - 使用方法の他にも、便利なスクリプトが公開されている
- r-wakalang (<https://r-wakalang.slack.com/>)
  - Slack 上で、R に関する情報交換を行っている。質問チャンネルもある
  - 解説記事: <http://qiita.com/uri/items/5583e91bb5301ed5a4ba>
- help 関数
  - 要インターネット環境
  - 引数には、パッケージ名や関数名などを入れる

## 連絡先

川口 勇作 (愛知学院大学)

Mail: [ykawa@dpc.agu.ac.jp](mailto:ykawa@dpc.agu.ac.jp)

サイト: <http://site.y-kawaguchi.com>

全国大会サイト上の WS 情報ページ: [http://bit.ly/LET2017WS\\_R](http://bit.ly/LET2017WS_R)

今回の資料等は、上記のサイト・ページにて公開予定です