クレジット:
UTokyo Online Education 統計データ解析 I 2017 小池祐太

東京大学 THE UNIVERSITY OF TOKYO

CC BY NC ND

統計データ解析 (I) 第 5 回

小池祐太

2017 年 10 月 25 日

1. 基本的な描画 (続き)
   - ヒストグラム
   - 箱ひげ図
   - 棒グラフ
   - 円グラフ
   - 散布図行列
   - 3 次元のグラフ
   - プロット環境の設定

基本的な描画 (日本語を含む図の描画)

- OS によっては日本語を含む図を描画すると文字化けする場合がある
- その場合, 関数 `par()` のオプション `family` に適当なフォントファミリーを指定することで文字化けを回避できる場合がある
- 例えば, Mac OS のデフォルトの設定では日本語を含む図は文字化けしてしまうが, 以下のコマンドをコンソール上で実行することで文字化けを回避できる

```
par(family = "HiraginoSans-W4")
```

  - フォントファミリーとしてヒラギノ角ゴシック W4 を指定している (数字を変えると太さが変わる)

- 実行例 `plot-kion.r`

図の保存

- 作成したグラフは保存することができる
- RStudio の機能を使う場合, 右下ペインの「Plots」タブの「Export」をクリックすると, 形式やサイズを指定して保存できる (もしくはクリップボードにコピーもできる)
- コマンドで実行することも可能であるが, それについての詳細は `help(png)` や `help(dev.copy)` を参照すること

ヒストグラム

- **ヒストグラム**
  - データの値の範囲をいくつかの区間に分割し, 各区間に含まれるデータの個数を棒グラフにしたもの
  - 棒グラフの横幅が区間に対応し, 面積が区間に含まれるデータの個数に比例するようにグラフを作成する
  - データの分布の仕方 (どのあたりに値が集中しているか, どの程度値にばらつきがあるかなど) を可視化するのに有効
- ヒストグラムは関数 `hist()` で作成できる

ヒストグラム

- 基本書式

  `hist(x, breaks, freq)`

  - `x`: ヒストグラムを描画するベクトル
  - `breaks`: 区間の分割の仕方を指定．数字を指定するとデータ範囲をその数字に近い個数に等分割する．デフォルトの個数は Sturges の公式によって決定される．すなわち，データ数を n とすると，$\lceil \log_2 n + 1 \rceil$ である．[1] その他の指定方法もある (ヘルプ参照)
  - `freq`: `TRUE` 指定すると縦軸をデータ数にし，`FALSE` 指定すると縦軸をデータ数/全データ数とする．デフォルトは `TRUE` (`breaks` の指定によって変わる場合あり)
  - 他にも `plot` で指定できるオプションが利用可能

- 実行例 `hist3.r`

[1] $\lceil x \rceil$ は x 以下の最大の整数を表す．

ヒストグラム

演習 1

実行例 `hist3.r` で紹介した Weyl の一様分布定理において, 無理数 `a` を別の値に変更しても, `x` の分布の仕方は区間 $(0, 1)$ 上でほぼ均一となることを確かめよ (どの程度均一に近くなるかは `a` によって異なるため, いくつか確かめてみよ)

箱ひげ図

- **箱ひげ図**
  - データの中心, 散らばり具合および外れ値を考察するための図 (ヒストグラムの簡易版)
  - 複数のデータの分布の比較の際に有効
  - データの第 1 四分位点を下端, 第 3 四分位点を上端とする長方形 (箱) と, 第 1 四分位点, 第 3 四分位点からそれぞれ箱の長さの 1.5 倍以内にあるデータのうちの最小の値, 最大の値を下端, 上端とする直線 (ひげ) からなる
  - ひげの外側のデータは点で表示される
  - 中央値は太線で表示される
- 箱ひげ図は関数 `boxplot()` で描画できる

箱ひげ図

- ベクトル x に対する箱ひげ図は boxplot(x,...) で描画できる (... に関数 plot( ) と同様のオプションを指定可能)
- データフレーム x に対して, boxplot(x,...) は列ごとの箱ひげ図を描画
- データフレーム x において, 変数 A が「分類」を表す変数 (性別, 植物の種類など)[2] の場合, 別の変数 B に対して,

```
boxplot(B ~ A, data = x, ...)
```

は変数 B を変数 A で分類した場合の, 分類ごとの箱ひげ図を描画する

- 実行例 boxplot3.r

[2] **質的変数**と呼ばれる

箱ひげ図

演習 2

データセット `kikou2016.csv` において, 気温以外の変数についても月ごとの箱ひげ図を作成せよ (グラフタイトル・色などのグラフィックパラメーターは適当に調節せよ)

棒グラフ

- 棒グラフは関数 `barplot()` で作成できる
- 基本書式

```
barplot(x, width = 1, space = NULL, legend.text = NULL,
        beside = FALSE, args.legend = NULL)
```

  - `x`: 棒グラフを作成するベクトル/行列 (データフレームは不可)
  - `width`: 棒の幅を指定するパラメーター
  - `space`: 棒グラフ間・変数間のスペースを指定
  - `legend.text`: 凡例を指定
  - `beside`: 変数が複数あるときに, 棒グラフを縦に並べるか・横に並べるか
  - `args.legend`: 関数 `legend()` に渡す引数の指定
  - 他にも `plot` で指定できるオプションが利用可能

- 実行例 `barplot.r`

棒グラフ

演習 3

データセット `airquality` について, Month, Day 以外の変数の月ごとの平均の棒グラフを作成せよ

円グラフ

- 円グラフは関数 `pie()` で作成できる
- 基本書式

```
pie(x, clockwise = FALSE)
```

  - `x`: 円グラフを作成するベクトル
  - `clockwise`: 時計回りに書くか否か
  - 他にも `plot` で指定できるオプションが利用可能
- 実行例 `pie.r`

散布図行列

- データフレーム x に対して plot(x,...) もしくは pairs(x, ...) を実行すると, すべての列のペアに対する散布図を行列状に並べた図を作成する
  - 変数 A1, ..., Ak のみ考えたい場合, plot(~ A1 + ··· + Ak, data = x, ...) もしくは pairs(~ A1 + ··· + Ak, data = x, ...) を利用
- 実行例 pairs.r

3 次元のグラフ

- 3 次元のグラフを 2 次元に射影した俯瞰図は，関数 `persp()` を用いて描くことができる
- 基本書式

```
persp(x, y, z, theta = 0, phi = 15, expand = 1)
```

  - `x`, `y`, `z`: x, y, z 座標
    - ★ `z` は点 (`x[i]`, `y[j]`) に対応する値を (i, j) 成分とする行列で与える
  - `theta`, `phi`: グラフの角度を指定する極座標
  - `expand`: z 軸の拡大度を指定
  - 他にも `plot` で指定できるオプションが利用可能
- 実行例 `plot3d.r`

3 次元のグラフ

- 様々な種類の 3 次元のグラフを描画するために多くのパッケージが開発されている
- そのうちの 1 つパッケージ `scatterplot3d` には, 3 次元の散布図を書くための関数 `scatterplot3d()` が用意されている
- 基本書式

```
scatterplot3d(x, color, angle=40)
```

  - `x`: x, y, z 座標を指定するデータフレーム (関数 `persp()` のように直接指定することも可能)
  - `color`: 色を指定 (`col` ではない). デフォルトは黒
  - `angle`: x 軸と y 軸の間の角度
  - 他にも `plot` で指定できるオプションが利用可能
- 実行例 `plot3d.r`

プロット環境の設定

- プロットの際の線の種類や色, 点の形等のデフォルト値は関数 `par( )` で設定できる
- 設定可能なグラフィックスパラメーターは `help(par)` で確認できる
- 特に, 以下の例のように, 関数 `par( )` によってプロット環境の設定 (複数図の配置, 余白の設定) ができる
- 実行例 `par.r`

プロット環境の設定

- プロット環境は非常に細かく設定でき, またそれぞれの描画関数独自のパラメーターも存在するため, ここでは紹介しきれない. 必要に応じてヘルプファイル, またはインターネット上の情報を参照すること