

# 統計学 2

## 5. ggplot2によるデータの可視化

矢内 勇生

2019年5月9日

高知工科大学 経済・マネジメント学群

# 今日の目標

- ggplot2の基本的な使い方を身につける

# データの可視化

- データ（分析結果）を**分かり易く**示す
  - データを要約する
  - 直感に訴える

# ggplot2 とは？

- Hadley Wickham が開発したRのパッケージ
- **g**rammar for **g**raphics
- 綺麗な図が、簡単に描ける
  - 文法 (grammar) を覚えれば、一貫した方法で作図ができる

# Hadley Wickham

- RStudio のChief Scientist
- 現在、Rの世界における最重要人物
- tidyverse を提唱
  - tidyr
  - ggplot2
  - dplyr
  - readr
  - etc.

# ggplot2による作図の基本

1. `ggplot()` 関数にデータを渡し、何を図にするか指定する
2. `geom_xxx()` で自分が作りたい図の層 (layer) を加える
3. `labs()` で軸ラベルを指定する
4. `plot()` または `print()` で図を表示する

# 1. ggplot()

- ggplot の第1引数は df (データフレーム)
  - aes (aesthetics) でデータフレーム内のどの変数を使うかを指定する
- 例：ddd という名前のデータフレーム内の、height という変数を横軸 (x軸) に、weight という変数を縦軸 (y軸) にした図を作る

```
ggplot(ddd, aes(x = height, y = weight))
```

## 2. geom\_xxx()

- geom (geometry) で始まる名前の層を加える
- xxx の部分は、作る図によって変える
  - ヒストグラム : geom\_histogram()
  - 散布図 : geom\_point()
- 使う geom によって aes() の中で指定すべきものが変わる



# 3. labs()

- labs で軸ラベルを指定する
- ラベルは引用符で囲む
- 横軸ラベル：x
- 縦軸ラベル：y
- 図のタイトル：title
- 例：

```
labs(x = “身長 (cm)”, y = “体重 (kg)”, title = “体重と身長散布図”)
```

## 4. plot() or print()

- ggplot で作った図は、オブジェクトとして保存し、`plot()` または `print()` で表示する

# 例：散布図

- データフレーム：ddd
- 横軸の変数：height
- 縦軸の変数：weight

```
p1 <- ggplot(ddd, aes(x = height, y = weight))
```

```
p1 <- p1 + geom_point()
```

```
p1 <- p1 + labs(x = “身長 (cm)”, y = “体重 (kg)”)
```

```
plot(p1)
```

# 体重と身長の散布図

