

箱ヒゲ図

Shouhei TAKEUCHI

October 1, 2015

Contents

Rで箱ヒゲ図	2
このファイルの更新情報	2
データ作成	2
箱ヒゲ図の作図	2
各種パラメータについて	4
bxp() と boxplot() の違い	14
Tips	16

R で箱ヒゲ図

boxplot() 関数についてまとめたファイル。

綺麗な図を描くためには、boxplot() を使うよりも、bxp() を使った方が都合がいい。たとえば、箱、ヒゲ、ヒゲの先端、外れ値をそれぞれ個別に柔軟に対処できる。

このファイルの更新情報

このファイルは 2015-10-01 14:52:45 に更新されました。

- 更新日時の表示方法を変更した。
- github で公開してみた。
- データを付値するオブジェクトを x から dat に変更し、boxplot() 関数内のモデル式の書き方を変更した。

データ作成

適当なデータを作成しておく。その際、必ず外れ値が出るように、乱数はシードを固定しておく。

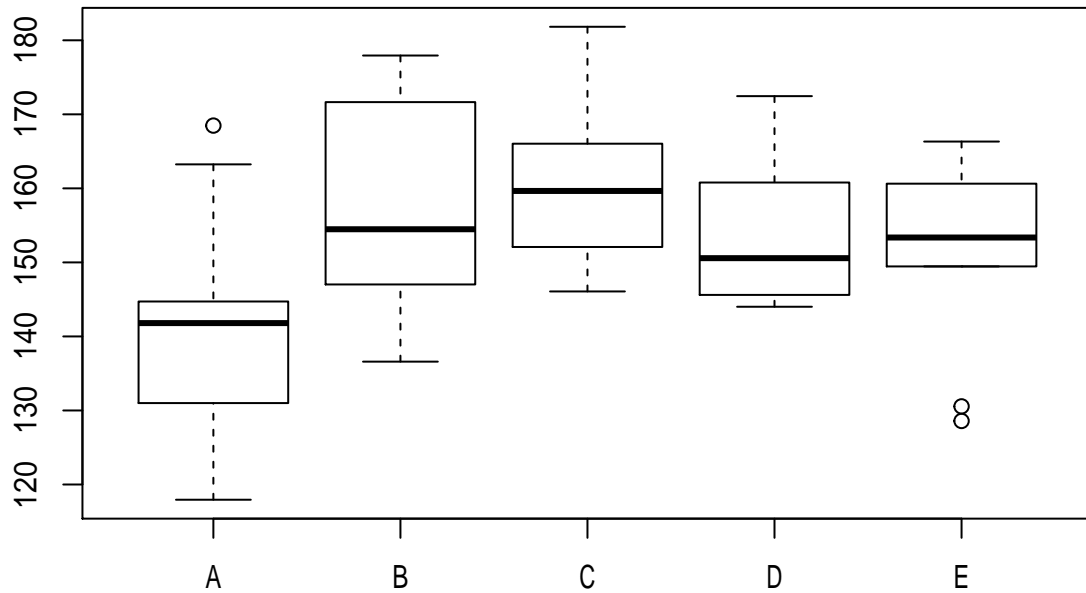
```
set.seed(14)
dat <- data.frame(
  height = 150 + 15 * rnorm(50) + sample(c(rep(0, 35), rep(30, 15)), 1),
  name = rep(LETTERS[1:5], 10)
)
```

箱ヒゲ図の作図

一番簡単なのは、boxplot() 関数を使うこと。ただし、適当な変数に付値しておくと、データのサマリーが見れるようになる。

```
# boxplot 関数は、bxp() 用のデータも作成する。
(bx.p <- boxplot(height ~ name, data = dat, main = "箱ヒゲ図"))
```

箱ヒゲ図



```
## $stats
##      [,1]      [,2]      [,3]      [,4]      [,5]
## [1,] 117.9454 136.5958 146.0785 144.0094 149.4579
## [2,] 130.9978 147.0213 152.0830 145.6078 149.4579
## [3,] 141.8023 154.4786 159.6523 150.5719 153.3621
## [4,] 144.7211 171.6504 166.0349 160.7900 160.6338
## [5,] 163.2420 177.9412 181.8250 172.4573 166.3213
##
## $n
## [1] 10 10 10 10 10
##
## $conf
##      [,1]      [,2]      [,3]      [,4]      [,5]
## [1,] 134.9456 142.1728 152.6814 142.9862 147.7782
## [2,] 148.6591 166.7843 166.6232 158.1576 158.9461
##
## $out
## [1] 168.4792 128.5835 130.5536
##
## $group
## [1] 1 5 5
```

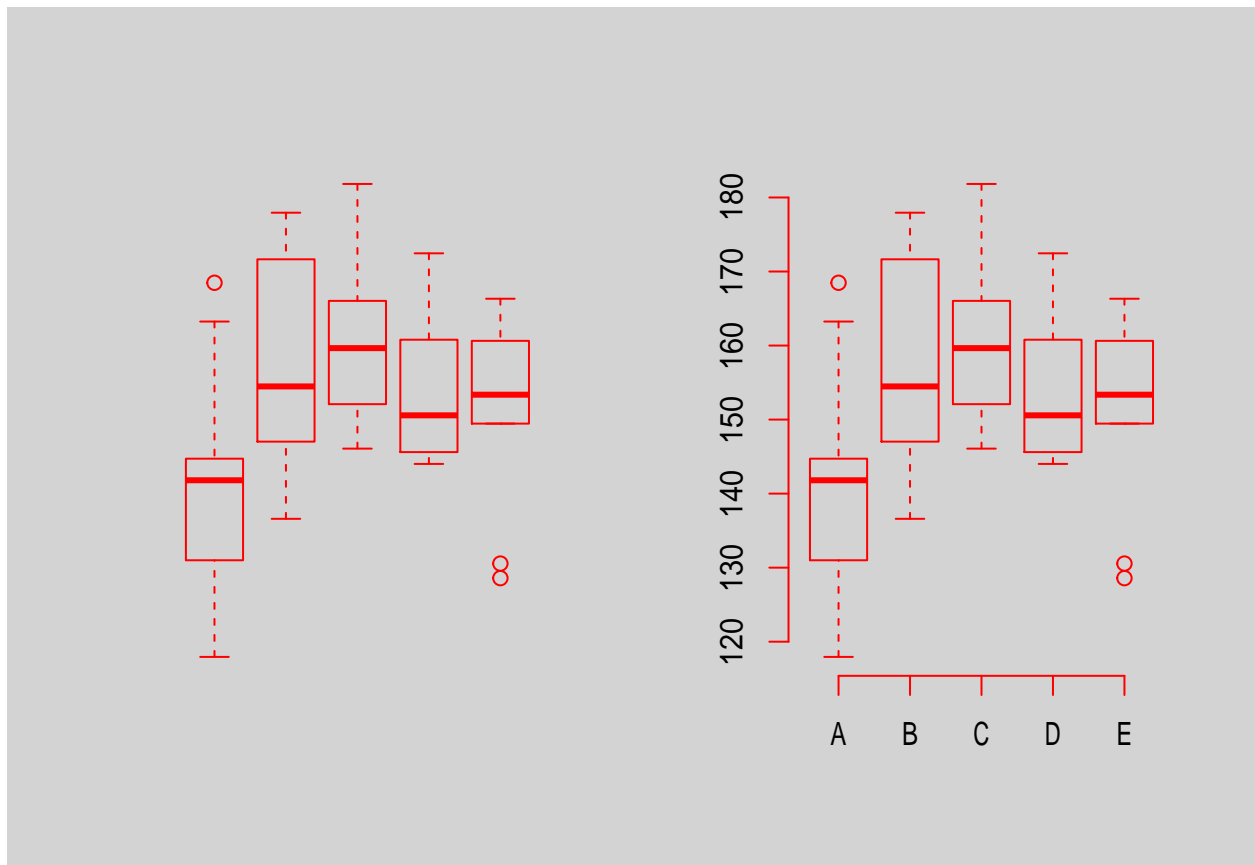
```
##
## $names
## [1] "A" "B" "C" "D" "E"
```

各種パラメータについて

基本的に、`bxp()` 関数を用いて指定する方が柔軟に対応できる。ただし、`ggplot2` のように背景をうまく変更する方法は見つけれなかった。

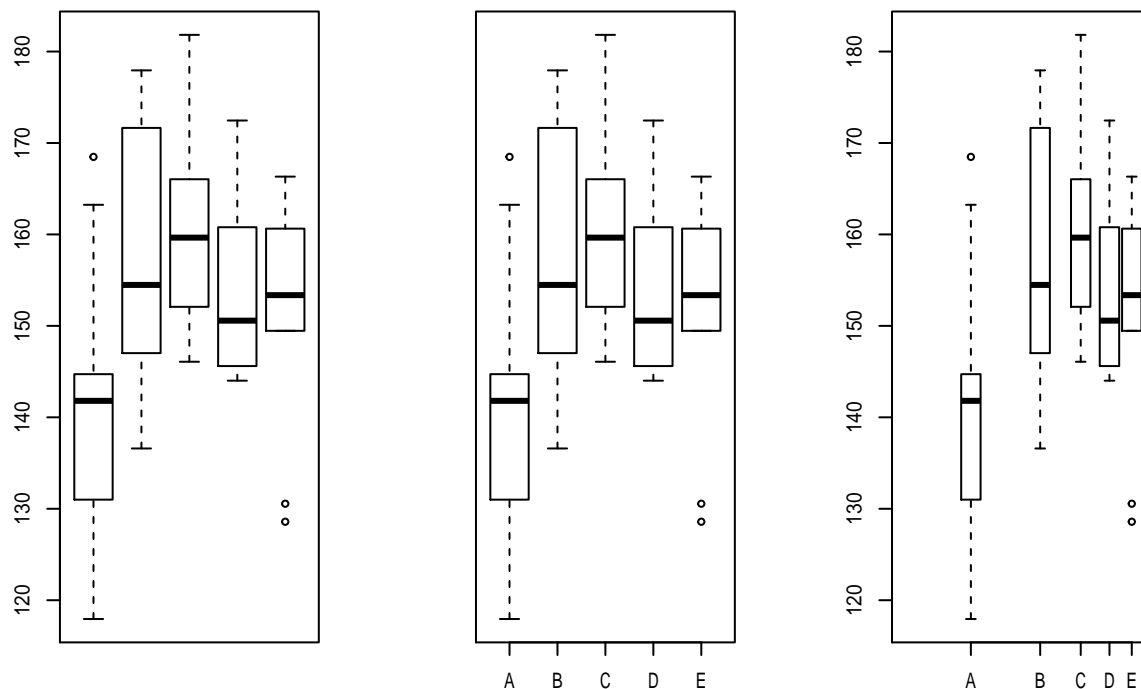
外枠・目盛りについて

```
# par(fg = "color") で指定すると、外枠も色が付く
# bxp(bx.p, border = "color") で指定すると、外枠はそのまま
# 外枠
bxp(bx.p, axes = FALSE)
bxp(bx.p, frame = FALSE) # スケールは残す。
```



```
# x 軸、y 軸の目盛り
bxp(bx.p, xaxt = "n")
```

```
bxp(bx.p, yaxt = "n")
# x軸をログスケールに変更。boxplot() では、箱のサイズは調整されない。
bxp(bx.p, log = "x")
```

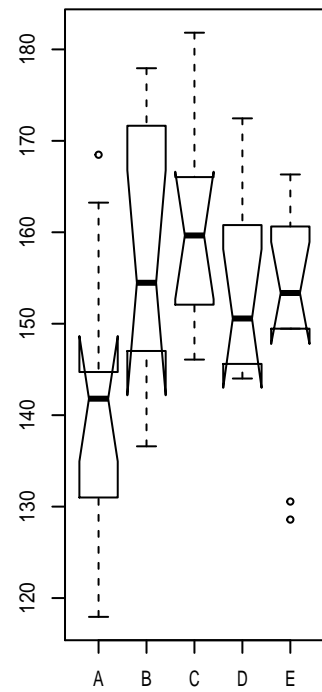
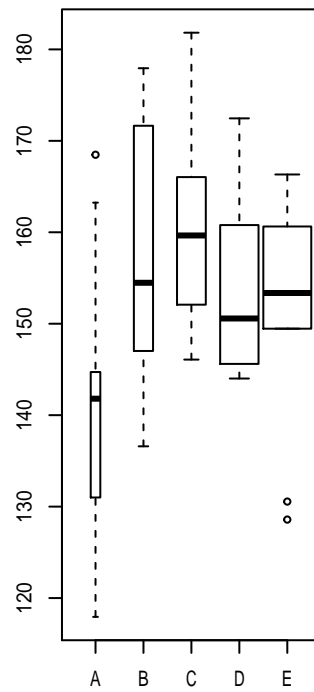
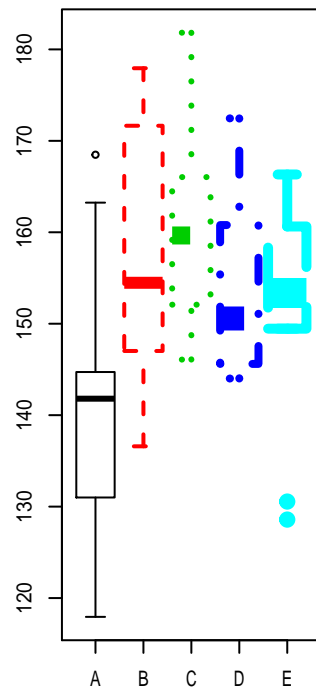


箱ヒゲ図について (全体)

```
# 線種・太さ・色・箱の大きさ
bxp(bx.p, lty = 1:5, lwd = 1:5, border = 1:5)
bxp(bx.p, boxwex = c(0.2, 0.4, 0.6, 0.8, 1)) # default: 0.8

# 箱の形 (中央値の 95% 信頼区間?)
bxp(bx.p, notch = TRUE)
```

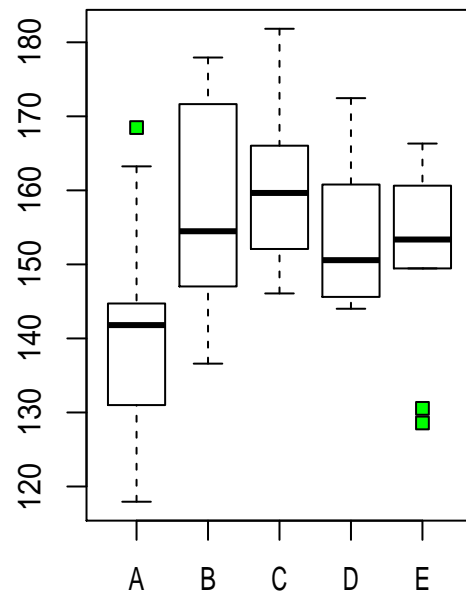
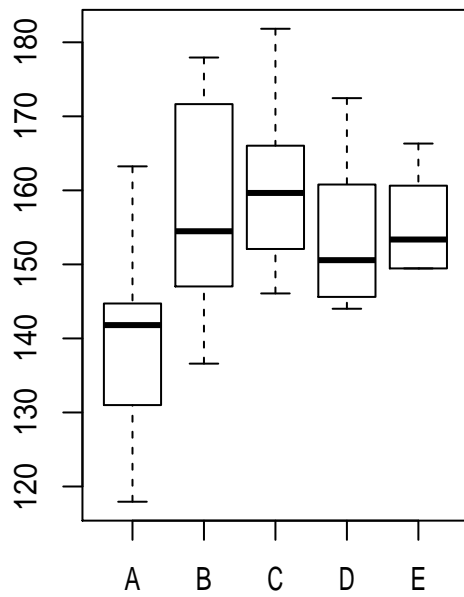
```
## Warning in bxp(bx.p, notch = TRUE): some notches went outside hinges
## (' box '): maybe set notch=FALSE
```



外れ値について

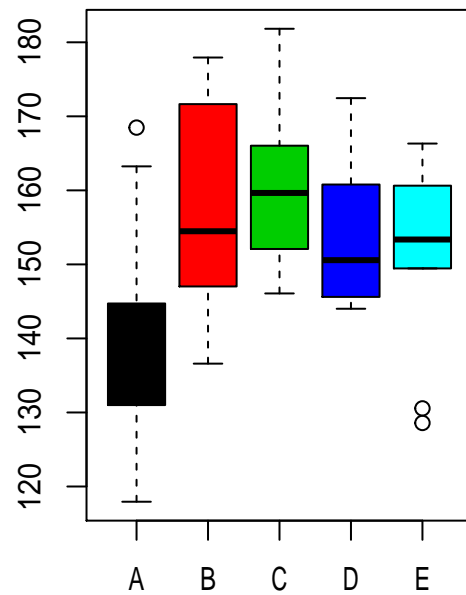
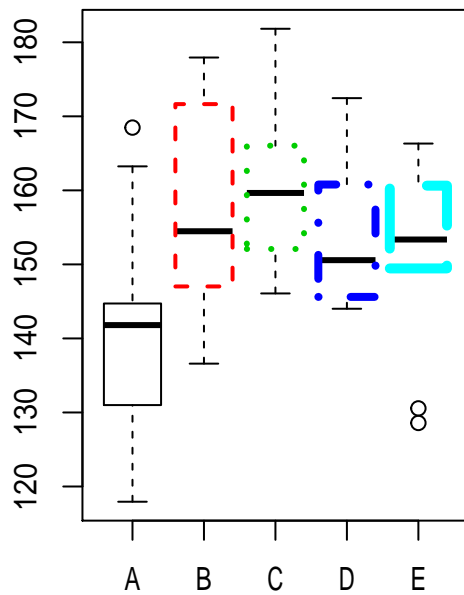
`bxp(bx.p, outline = FALSE)`

`bxp(bx.p, pch = 22, bg = "green")` # 外れ値に色を付ける (pch21 ~ 25 限定)



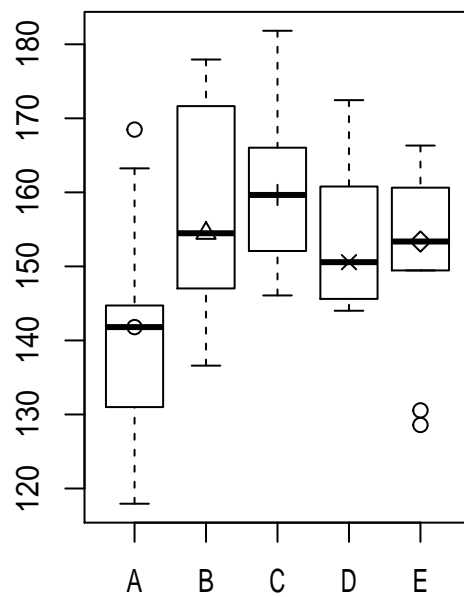
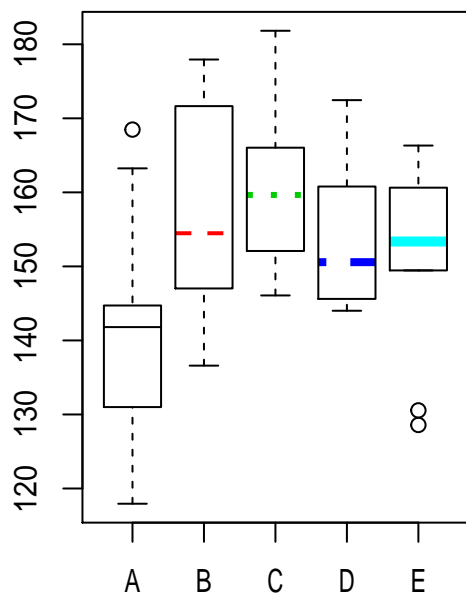
box (箱だけ) について

```
# 線種・太さ・色
bxp(bx.p, boxlty = 1:5, boxlwd = 1:5, boxcol = 1:5)
# 塗りつぶし
bxp(bx.p, boxfill = 1:5)
```

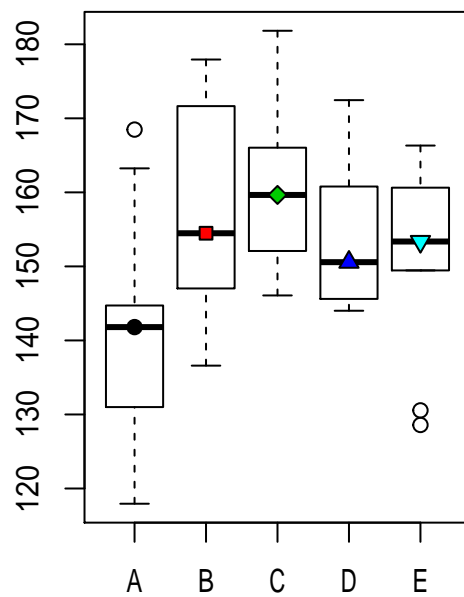
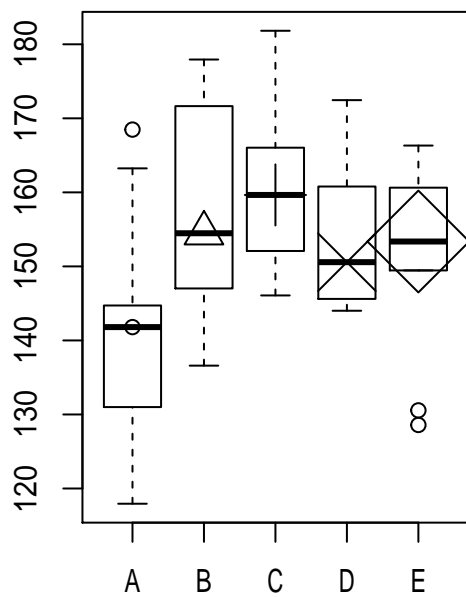


median (中央値のみ) について

```
# 線種・太さ・色
bxp(bx.p, medlty = 1:5, medlwd = 1:5, medcol = 1:5)
# マーク
bxp(bx.p, medpch = 1:5)
```

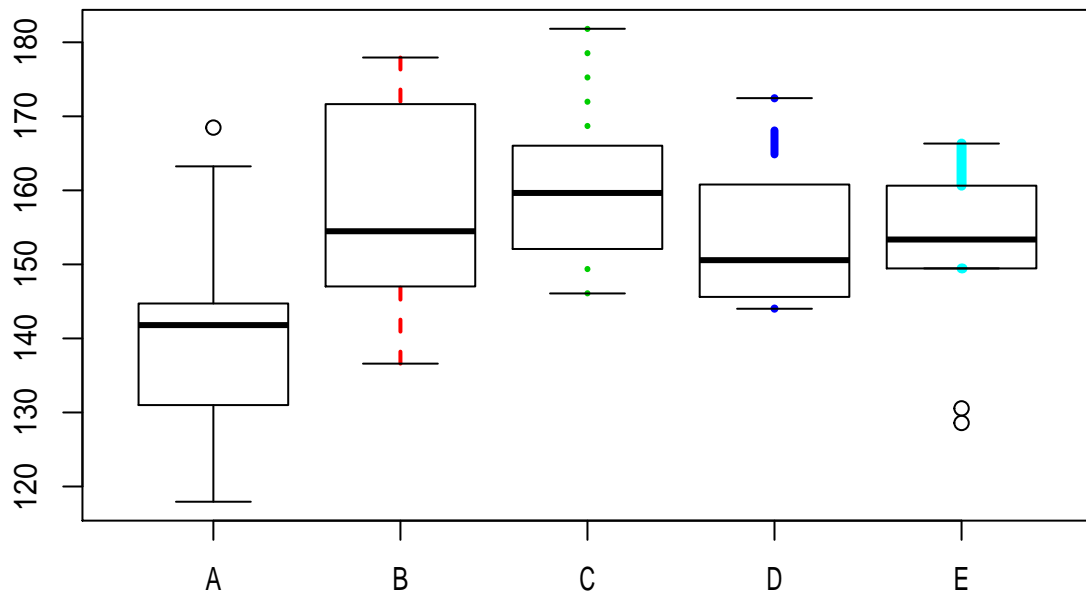
```
# medpch を指定したときの大きさ
bxp(bx.p, medpch = 1:5, medcex = 1:5)
# medpch で塗りつぶせる記号を選んだときの中身の色
bxp(bx.p, medpch = 21:25, medbg = 1:5)
```



whisker (ヒゲのみ) について

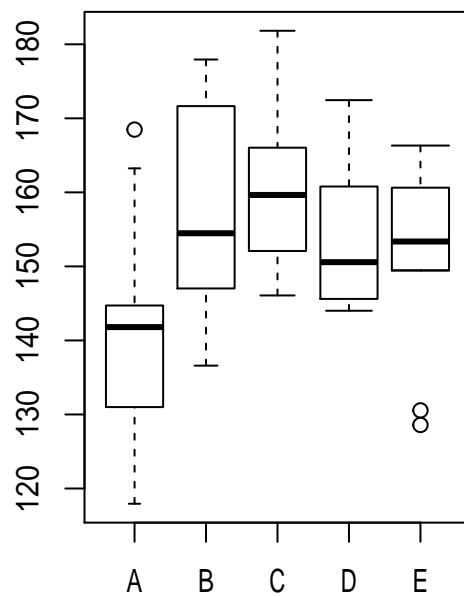
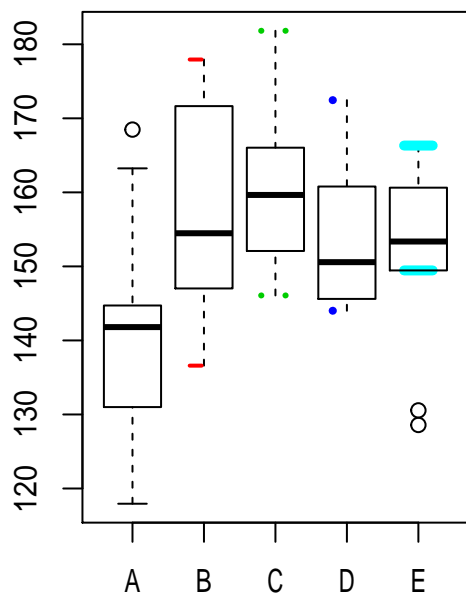
線種・太さ・色

```
bxp(bx.p, whisklty = 1:5, whisklwd = 1:5, whiskcol = 1:5)
```



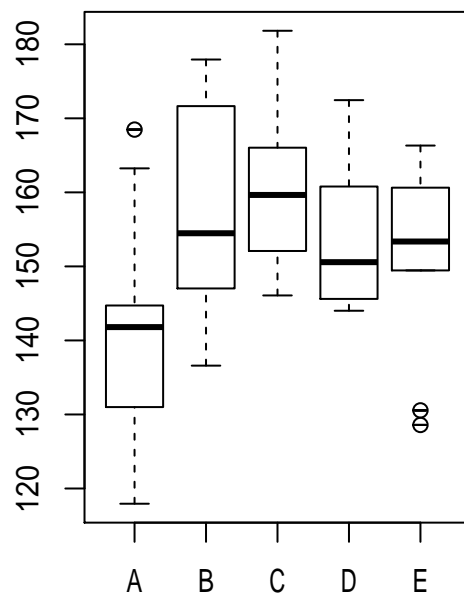
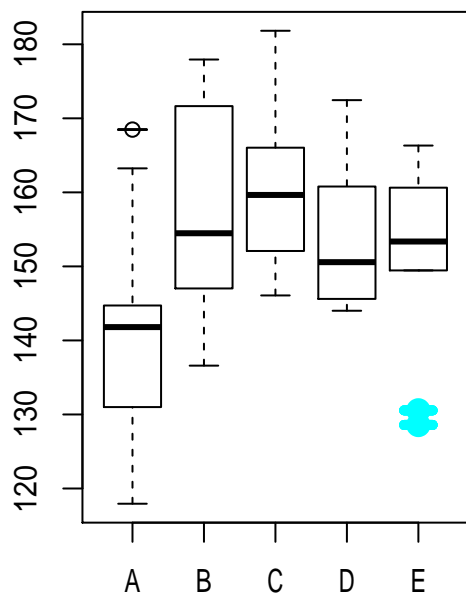
staple (ヒゲの天辺のみ) について

```
# 線種・太さ・色
bxp(bx.p, staplelty = 1:5, staplelwd = 1:5, staplecol = 1:5)
# ヒゲの長さ
bxp(bx.p, staplewex = c(0.2, 0.4, 0.6, 0.8, 1)) # default: 0.5
```

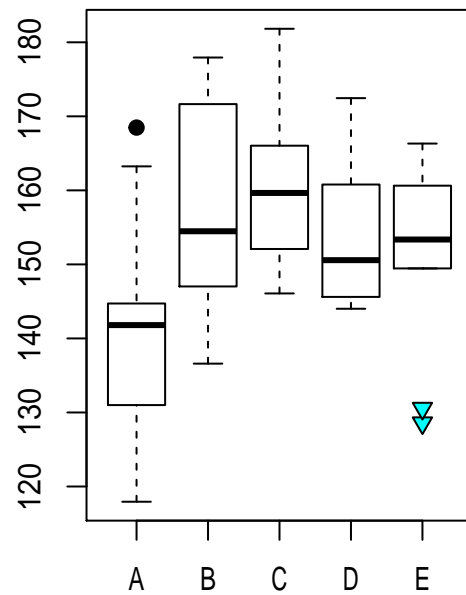
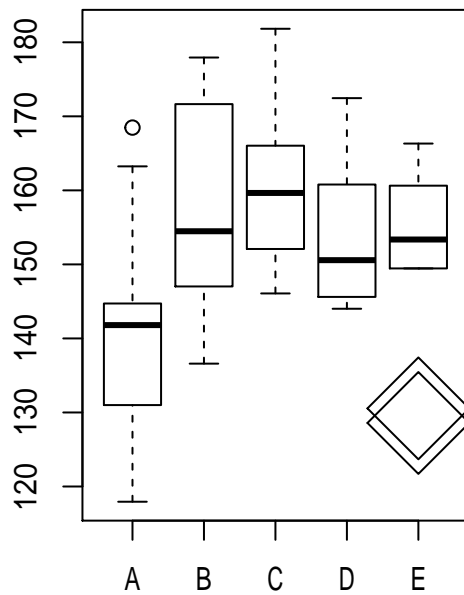


outlier について

```
# 線種・太さ・色
bxp(bx.p, outlty = 1:5, outlwd = 1:5, outcol = 1:5)
# 長さ
# default: 0.5 # 箱ごとの指定はできない。
bxp(bx.p, outwex = 0.2, outlty = 1:5)
```

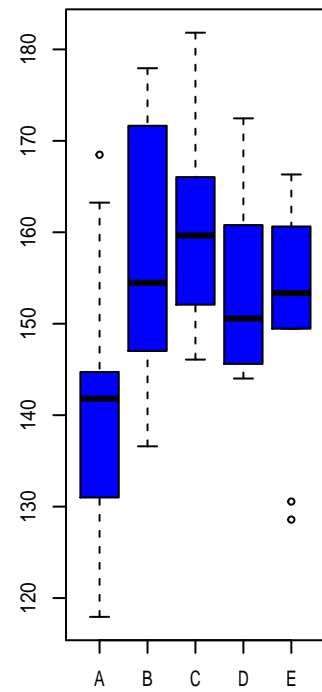
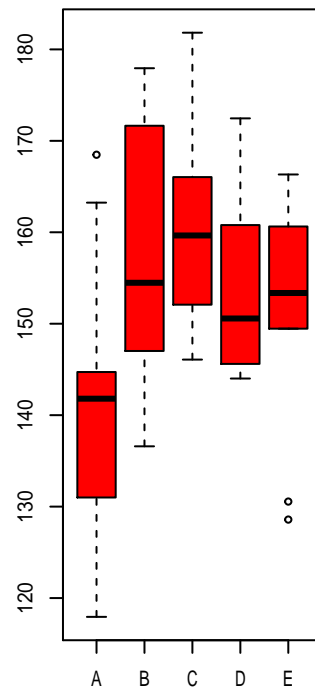
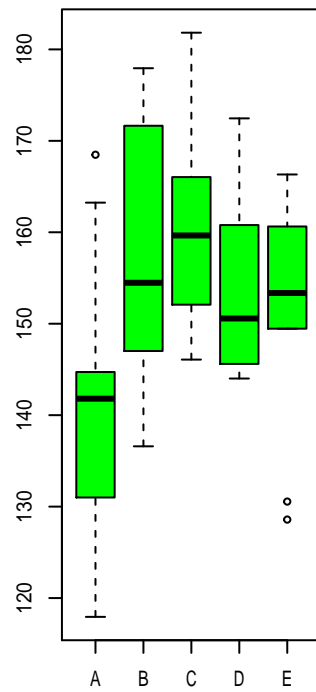


```
# 大きさ
bxp(bx.p, outpch = 1:5, outcex = 1:5)
# 色
# outpch で塗りつぶせる記号を選んだときの色
bxp(bx.p, outpch = 21:25, outbg = 1:5)
```



bxp() と boxplot() の違い

```
### 塗りつぶしの色について
bxp(bx.p, boxfill = "green")
boxplot(height ~ name, data = dat, col = "red")
boxplot(height ~ name, data = dat, boxfill = "blue")
```

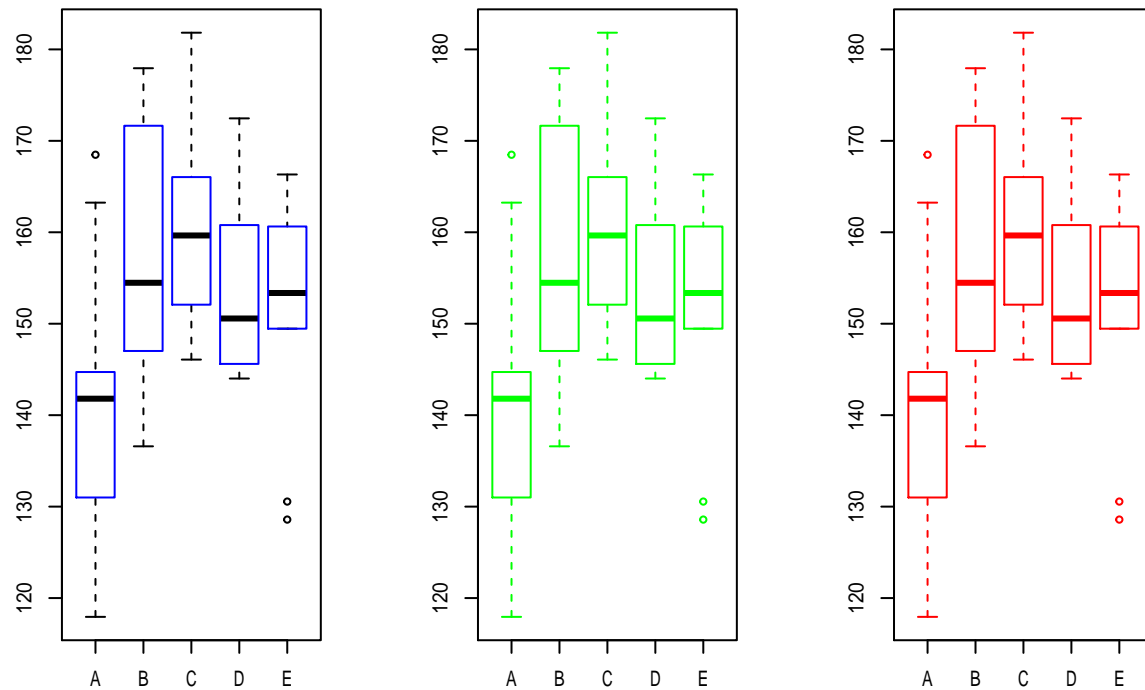


線の色について

`bxp(bx.p, boxcol = "blue")` # boxのみ

`bxp(bx.p, border = "green")` # 全体

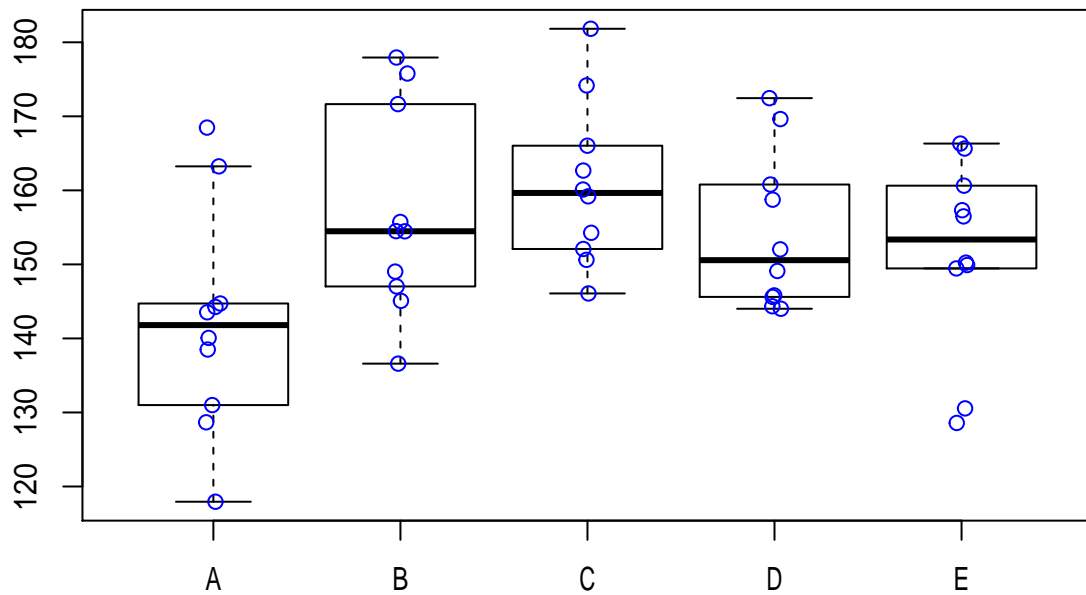
`boxplot(height ~ name, data = dat, border = "red")` # 全体



Tips

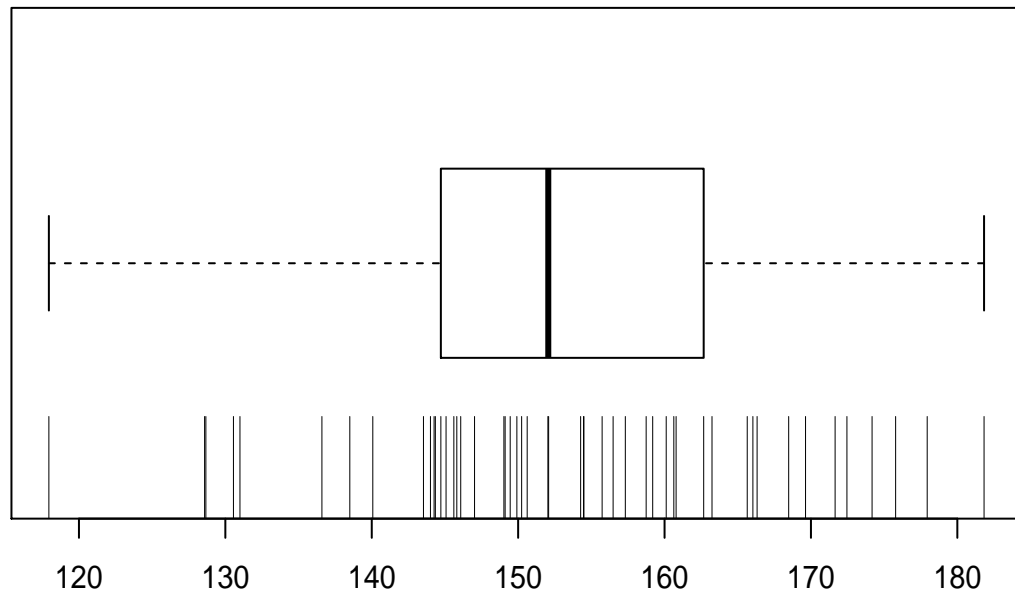
boxplot に jitter を重ねる。

```
# jitter を重ね描きする時は、外れ値を描かない方が良い。
bxp(bx.p, outline = FALSE)
# x 軸方向にずらすので、jitter() は横軸の変数にかける。
points(jitter(as.numeric(dat$name), 0.2), dat$height, col = "blue")
```

boxplot に rug を追加する

```
boxplot(dat$height, horizontal = TRUE)  
rug(dat$height, ticksize = 0.2)
```



grid の追加

grid は箱ヒゲ図に上に書きしてしまうので、grid を書いて、その上で bxp() 関数を使う。

```
bx.p <- boxplot(height ~ name, dat)
grid(NULL, NULL, col = "grey80", lty = 1, lwd = 1)
bxp(bx.p, add = TRUE)
```

