```
##################################
##第九回「ノンパラメトリック密度推定法」
##################################


##################################
## 1991年にとられたナショナルフットボールリーグの給料データ
## QB: 正クォーターバックの給料 (1,000$) の密度を推定

x <- read.csv("qbackand.csv")
x <- x$QB

dens <- density(x) #デフォルトでGaussianカーネル, バンド幅はシルバーマンの方法
dens_cv <- density(x,bw="ucv") #不偏クロスバリデーション

plot(dens,lwd=2,col=3,xlab="1,000$",main="Quarter Back Salary") #推定した密度
lines(dens_cv,lwd=2,col=4)
rug(x)
legend(3000, 6e-4, legend = c("Silverman", "CV"), lty=c(1), col=c(3,4),title = "
bandwidth")

# カーネル関数による違い
plot(density(x, bw = "ucv"),     main = "7 different kernels")
for(i in 2:length(kernels))
   lines(density(x, bw = "ucv", kernel = kernels[i]),col = i)


##################################
## 世田谷区の中古マンション価格データ
## 価格 (price) と面積 (area) の分布を求める.
## 二次元データの密度関数
library(MASS)
library(KernSmooth)

x <- read.csv("setagaya_manshion.csv",header=T)
sman = data.frame(price=x[[1]],walk=x[[3]],area=x[[5]],str=ifelse(x[[6]]=="ＲＣ"
,1,0),kenpei=x[[7]],youseki=x[[8]],tikunen=x[[9]],kyuko=x[[10]]) #構造はダミー変

plot(sman$price,sman$area,xlab="price",ylab="area")

sman.kde2 <- kde2d(sman$price,sman$area,c(bandwidth.nrd(sman$price),bandwidth.nr
d(sman$area)),n=80)
# MASSライブラリ,2変量確率密度推定
# シルバーマンの方法でバンド幅を求める.
image(sman.kde2,xlab="price",ylab="area")
points(sman$price,sman$area)

sman2d_data <- matrix(c(sman$price,sman$area),ncol=2) #データを行列に変換
colnames(sman2d_data) <- c("price","area")

# KernSmoothパケージにあるbkde2Dも試してみる.
sman.bkde2 <- bkde2D(sman2d_data, bandwidth = apply(sman2d_data, 2,dpik)) #dpik
はヒューリスティクスによるハンド幅の選択
image(x = sman.bkde2$x1, y = sman.bkde2$x2, z = sman.bkde2$fhat,xlab = "price",
ylab = "area")
points(sman2d_data)
par(new=T)
contour(x = sman.bkde2$x1, y = sman.bkde2$x2, z = sman.bkde2$fhat,add=TRUE)

persp(x = sman.bkde2$x1,
  y = sman.bkde2$x2,
  z = sman.bkde2$fhat,
  xlab = "price",
  ylab = "area",
  zlab = "estimated density",
  theta = -25, axes = TRUE, box = TRUE)
```