

# Selección de ventana $h$ por convalidación cruzada ‘deje-uno-afuera’

Martin Acuña, Facundo Solar

2023-09-15

## Funciones utilizadas

$$\hat{f}_h^{(-i)}(x_i) = \frac{1}{(n-1)h} \sum_{j \neq i} K\left(\frac{x_i - x_j}{h}\right)$$

```
fh_i <- function(x, i, h){  
  n.x <- length(x)  
  suma <- 0  
  for(j in 1:n.x){  
    if (j!=i){  
      suma <- suma + dnorm((x[i]-x[j])/h)  
    }  
  }  
  return((1/((n.x-1)*h)) * suma)  
}
```

$$h_{CV}^* = \arg \max_{h \in H} \left( \frac{1}{n} \sum_{i=1}^n \log \hat{f}_h^{(-i)}(x_i) \right)$$

```
h.CV <- function(x,grilla.h){  
  n.x <- length(x)  
  n.h <- length(grilla.h)  
  maxi <- -9999  
  hmax <- NA  
  for(j in 1:n.h) {  
    h <- grilla.h[j]  
    suma <- 0  
    for (i in 1:n.x) {  
      suma <- suma + log(fh_i(x,i,h))  
    }  
    suma <- suma * (1/n.x)  
    if(suma>maxi){  
      maxi <- suma  
      hmax <- h  
    }  
  }  
  return(hmax)  
}
```

## La funcion principal

```
bw.loocv <- function(x, grilla.h=NA) {  
  if (is.na(grilla.h)[1]) {  
    grilla.h <- seq(bw.nrd0(x)/100, bw.nrd0(x)*100, length=100)  
  }  
  n.h <- length(grilla.h)  
  n.x <- length(x)  
  loglikes <- rep(NA, n.h)  
  r <- matrix(NA, ncol = n.x, nrow = n.h)  
  for(j in 1:n.h) {  
    h <- grilla.h[j]  
    for (i in 1:n.x) {  
      r[j,i] <- log(fh_i(x,i,h))  
    }  
    loglikes[j] <- mean(r[j,])  
  }  
  
  hmax <- h.CV(x, grilla.h)  
  
  return(list(h.opt=hmax, h.ret = grilla.h, loglikes=loglikes))  
}
```

## Generador de la muestra

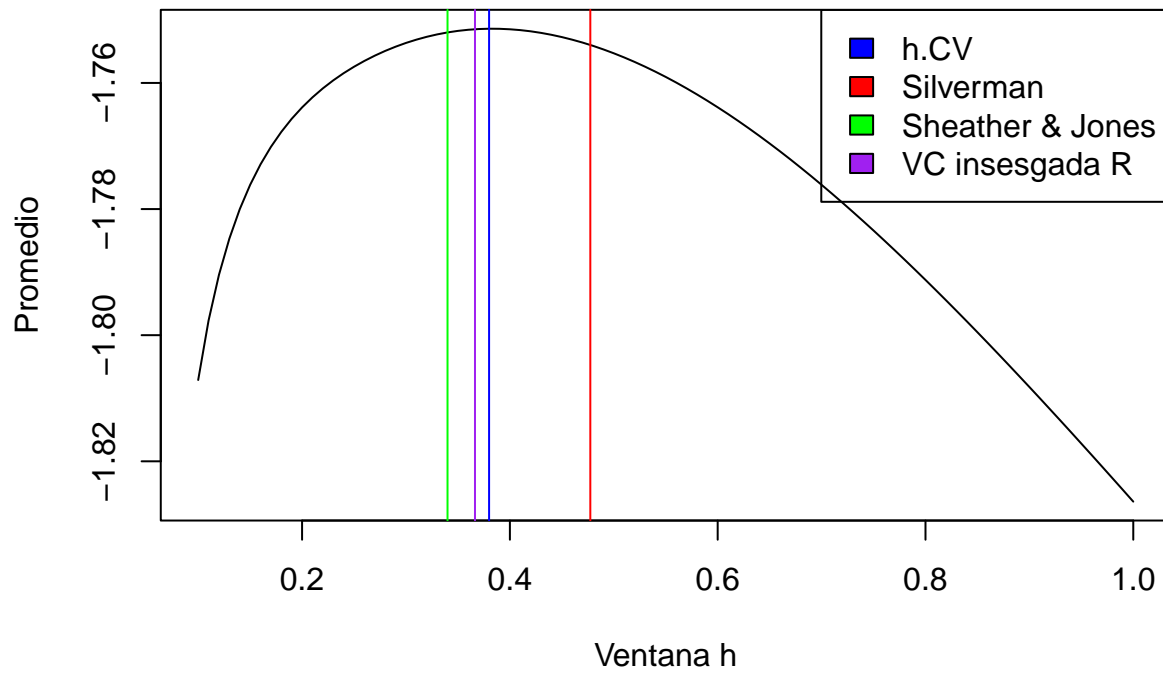
```
muestra <- function(seed, n=200) {  
  set.seed(1234)  
  binoms <- rbinom(n, size=1, p=0.75)  
  return (  
    binoms * rnorm(n, mean=0, sd=1)  
    + (1 - binoms) * rnorm(n, mean=3.25, sd=sqrt(0.5))  
  )  
}  
write.table(muestra(1234), "entrega1.txt", row.names=FALSE, col.names=FALSE)
```

## Datos utilizados

```
x <- read.table('entrega1.txt')$V1  
grilla.h <- seq(0.1,1,0.01)  
datos <- bw.loocv(x,grilla.h)
```

## Graficos

### Comparacion Log-verosimilitud



### Estimacion de la densidad

