

Los datos que se muestran a continuación se hallan en el archivo buffalo.txt y corresponden a la mediciones de cantidad de nieve caída (en pulgadas) en Buffalo en los inviernos de 1910/1911 a 1972/1973.

126.4	82.4	78.1	51.1	90.9	76.2	104.5	87.4	110.5	25.0	69.3	53.5
39.8	63.6	46.7	72.9	79.6	83.6	80.7	60.3	79.0	74.4	49.6	54.7
71.8	49.1	103.9	51.6	82.4	83.6	77.8	79.3	89.6	85.5	58.0	120.7
110.5	65.4	39.9	40.1	88.7	71.4	83.0	55.9	89.9	84.8	105.2	113.7
124.7	114.5	115.6	102.4	101.4	89.8	71.5	70.9	98.3	55.5	66.1	78.4
120.5	97.0	110.0									

1. Realice un histograma para estos datos utilizando los parámetros por default. Repetir eligiendo como puntos de corte las siguientes secuencias: i) de 20 a 140 con paso 10 y ii) de 20 a 140 con paso 5. Comparar los tres histogramas obtenidos. ¿Tiene algún efecto el refinamiento de los bins?
2. Realice un histograma para estas observaciones utilizando puntos de corte (10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130). Repita corriendo el punto de inicio de los bins en 2 unidades 2 veces consecutivas. Compare los tres histogramas obtenidos. ¿Tiene algún efecto la elección del punto inicial en este ejemplo?
3. Implemente una función **densidad.est.parzen** que tenga por argumento un conjunto de datos $\mathbf{x} = (x_1, \dots, x_n)$, una ventana h y un punto x y devuelva $\hat{f}_h(x)$, el valor de la estimación de la densidad f en el punto x , utilizando el núcleo uniforme (también llamado rectangular).
4. Con la función **densidad.est.parzen** implementada, estime la densidad f en el intervalo (25,126.4) (mínimo y máximo de las observaciones) sobre una grilla de 200 puntos equiespaciados para $h = 10$. Grafique el estimador $\hat{f}_h(x)$ obtenido.
5. Estime la función de densidad f de la variable *pulgadas de nieve caída en invierno* a partir de los datos de Buffalo a través de la función **densidad.est.parzen** implementada en el ítem anterior usando $h = 10$. Realice un histograma para los datos de Buffalo y superponga la densidad estimada en los datos mediante la función **densidad.est.parzen** utilizando $h = 10, 20$ y 30 . Observe como varía la rugosidad de los estimadores de f obtenidos.

6. La función de R **density** computa un estimador de la densidad a partir de un conjunto de datos $\mathbf{x} = (x_1, \dots, x_n)$ y la evalúa en un conjunto de puntos intermedios. Mediante la función de R **density** estime la función de densidad f de la variable *pulgadas de nieve caída en invierno* a partir de los datos de Buffalo y utilizando el núcleo normal, el rectangular y el de Epanechnikov con ventana $h = 5$. Realice un gráfico en el que superpone las tres estimaciones de f y compare los resultados.
7. Para los datos de Buffalo halle la ventana óptima h_S por la regla de Silverman para el núcleo normal, calcule la estimación de f basada en el núcleo normal con dicha ventana y grafique la estimación de la densidad obtenida.
8. Se puede realizar la estimación de la densidad f a partir de un conjunto de datos a través de la función **density** en un punto deseado x (f.hat.x) de la siguiente forma:

```
f.hat.x<- density(datos, from=x, to=x, n=1)
```

También mediante la función **approxfun** de R se puede aproximar la densidad estimada por la función de R **density** en el punto deseado x a partir de un conjunto de datos de la siguiente manera:

```
df <- approxfun(density(datos))
f.hat.x<-df(x)
```

Calcule la estimación de la densidad usando **density** a partir de los datos de Buffalo usando la ventana de Silverman y con esta estimación obtenga estimaciones de la densidad en los valores observados de nieve caída en Buffalo. Grafique la densidad estimada y superponga en el mismo plot los puntos correspondientes a las observaciones y el valor estimado de la densidad correspondiente en rojo.

9. Llamemos $\hat{f}_h^{(-i)}$ a la densidad estimada sin utilizar al punto x_i y usando la ventana h . Calcule las estimaciones $\hat{f}_h^{(-i)}$ para $i = 17, 20, 51$ de la densidad f usando la función **density** a partir de los datos de Buffalo usando el núcleo de Epanechnikov y la ventana $h = 5$. Grafique en un mismo plot la estimación de f basada en todos los datos usando el núcleo de Epanechnikov y la ventana $h = 5$ y las estimaciones $\hat{f}_h^{(-i)}$ para $i = 17, 20, 51$. Compare las estimaciones obtenidas.
10. Halle la ventana de convalidación cruzada por máxima verosimilitud para la estimación de la densidad basada en el núcleo de Epanechnikov utilizando los datos de Buffalo mediante la función de R **density**. Utilice una grilla para la ventana h que varíe entre 5 y 25 con paso 0.25. Realice un gráfico de h versus la pseudo-log-verosimilitud calculada para cada h .

11. Grafique en un mismo plot las estimaciones de la densidad obtenidas con el núcleo de Epanechnikov usando la ventana óptima obtenida en el ítem anterior y usando la ventana h_S hallada para el núcleo normal. Compare los resultados.