.

Ejercicio 1

En este ejercicio trabajaremos con la función de distribución empírica.

1. Fijar una semilla y generar $n = 25$ datos con distribución $\mathcal{U}(0, 1)$.
2. Graficar la función de distribución empírica asociada a los datos generados y superponer el gráfico de la función de distribución acumulada (teórica) de una $\mathcal{U}(0, 1)$. Comparar.
3. Repetir el ítem anterior con 3 semillas distintas y disponer los 4 gráficos en un mismo plot, inluyendo además la función de distribución acumulada. ¿Qué sugieren estos gráficos?
4. Repetir los ítems anteriores con $n = 1000$ y comparar con los gráficos anteriores.
5. Repetir los ítems previos, pero generando datos con distribución $\mathcal{E}(1)$.

   Comparar los resultados obtenidos con la Uniforme y la Exponencial: los gráficos de la función de distribución empírica para ambas distribuciones, ¿son similares o distintos? En base a eso, ¿parece razonable utilizar dicha función para identificar la distribución de la cual provienen los datos?
6. Representar en un mismo gráfico las funciones de densidad correspondientes a una $\mathcal{E}(1)$ y una $\mathcal{U}(0, 1)$ para $0 < x < 1$ y comparar.

Ejercicio 2

Seguimos trabajando con el conjunto de datos de Buffalo abordado en las "Actividades de Clase - Unidad 4"

Utilizar los comandos de `R`, `density` y `approxfun` para resolver los siguientes puntos:

1. Estimar la densidad en $x_0 = 80$ para `h` $= 10$ mediante el estimador de Parzen (es decir utilizando el núcleo uniforme) y verificar que los resultados coinciden con lo calculado en el ítem 11 de las "Actividades de Clase - Unidad 4".
   (*Aclaración: para que los resultados le den igual a los calculados a mano, hay que dividir al ancho de banda por* $\sqrt{3}$)
2. Graficar la densidad estimada.
3. Repetir los ítems anteriores con núcleo normal.

4. Estimar la función de densidad f de la variable pulgadas de nieve caída en invierno a partir de los datos de Buffalo utilizando el núcleo normal, el rectangular y el de Epanechnikov con ventana `h = 5`. Realizar un gráfico en el que se superpongan las tres estimaciones de f y comparar los resultados. En particular compare las suavidades de los 3 estimadores, ¿a qué cree que se debe lo que observa?
   (*Aclaración: ahora al estimar la densidad con el núcleo rectangular, usar* `bw = 5` *sin dividir por* $\sqrt{3}$)

5. Hallar la ventana óptima $\widehat{h}_{rot}$ por la regla de Silverman para el núcleo normal. Calcular la estimación de la densidad usando núcleo normal y $\widehat{h}_{rot}$, y obtener estimaciones de la densidad en los valores observados de la nieve caída en Buffalo.

   Graficar la densidad estimada y superponer los puntos correspondientes a las observaciones y el valor estimado de la densidad correspondiente en rojo.

6. Llamemos $\widehat{f}_h^{(-i)}$ a la densidad estimada sin utilizar la observación x_i y usando la ventana h. Calcular las estimaciones $\widehat{f}_h^{(-i)}$ para $i = 17, 20, 51$ de la densidad f usando la función `density` con núcleo de Epanechnikov y ventana `h=5`. Representar en un mismo gráfico la estimación de f basada en todos los datos usando el núcleo de Epanechnikov y ventana `h=5` y las estimaciones $\widehat{f}_h^{(-i)}$ para $i = 17, 20, 51$. Comparar los gráficos obtenidos.

7. Hallar la ventana óptima de Validación Cruzada $\widehat{h}_{cv}$ para la estimación de la densidad basada en el núcleo normal utilizando el comando `bw.ucv`.

8. Representar en un mismo gráfico las estimaciones de la densidad obtenidas con el núcleo de normal usando las ventanas $\widehat{h}_{rot}$ y $\widehat{h}_{cv}$ obtenidas en los ítems 5 y 7 respectivamente. Comparar los resultados.