

Estadística I.

Práctica 7: Estadística Descriptiva

1. Se tiene el siguiente conjunto de datos:

3 5 3 11 6
1 6 12 18 3

- (a) Hallar y graficar la función de distribución acumulada empírica.
- (b) Calcular las medidas de posición: media, mediana, media podada al 10%.
- (c) Calcular las medidas de dispersión: varianza muestral, desvío estándar muestral, IQR y MAD estandarizados.

2. Repetir el ejercicio 1 para el siguiente conjunto de datos:

2 7 5 2 5 8
4 9 7 6 6

3. Se tiene la siguiente muestra de “número de hijos” en 40 familias de una población:

0 0 0 0 0 0 0 0 0 0
0 0 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 2 2 2
2 2 2 2 3 3 3 3 4 4

- (a) Resumir los datos en una tabla (si no se le ocurre cómo, puede espiar el ejercicio 4).
 - (b) Repetir el ejercicio 1 para estos datos.
 - (c) Representar los datos en un histograma aproximado (es decir, dibujarlo “a mano”).
4. Repetir el ejercicio 1 para la siguiente muestra sobre “cantidad de televisores” en 120 hogares de una población.

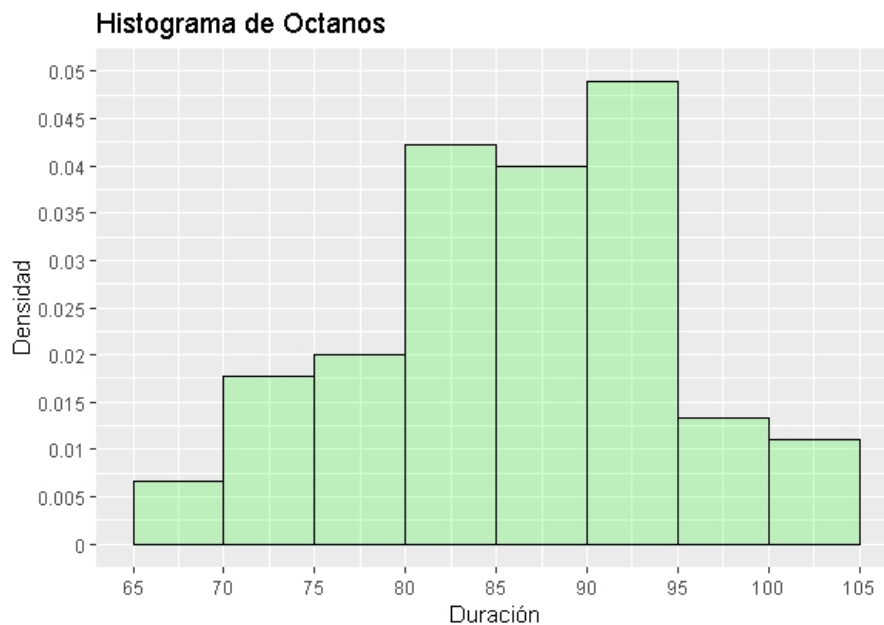
# Televisores	# Hogares
0	7
1	45
2	52
3	11
4	5

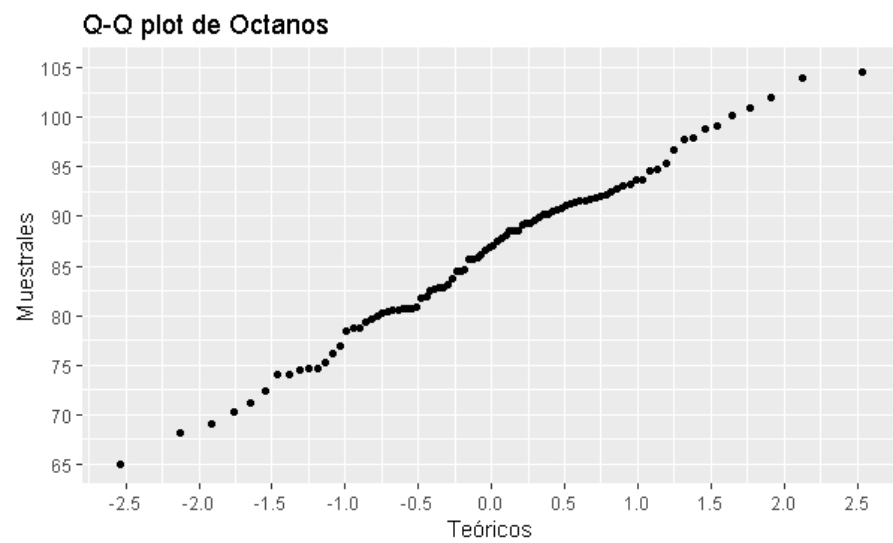
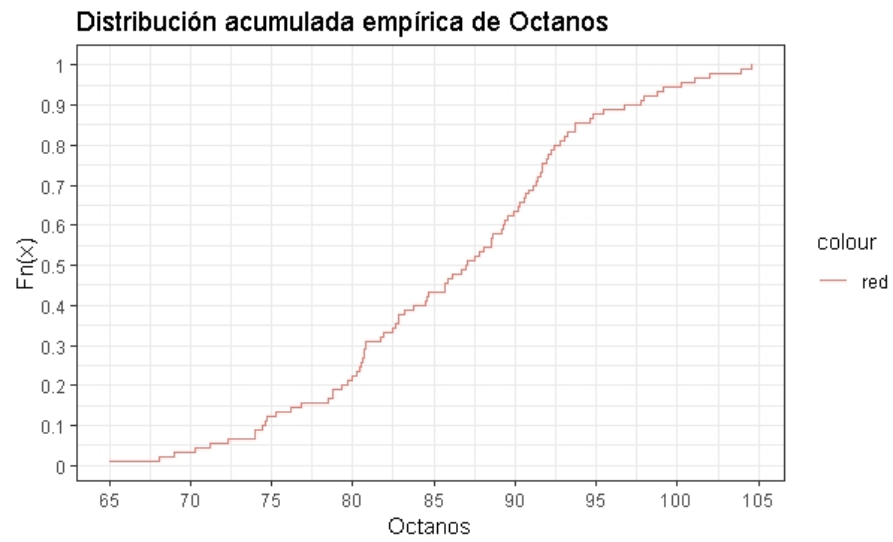
5. Calcular la mediana **poblacional** y los cuantiles 0.1, 0.4 y 0.8 de la **variable aleatoria** X , teniendo X distribución:

- (a) uniforme en $[0,1]$,
- (b) exponencial de parámetro 2,
- (c) normal con media 3 y varianza 9.

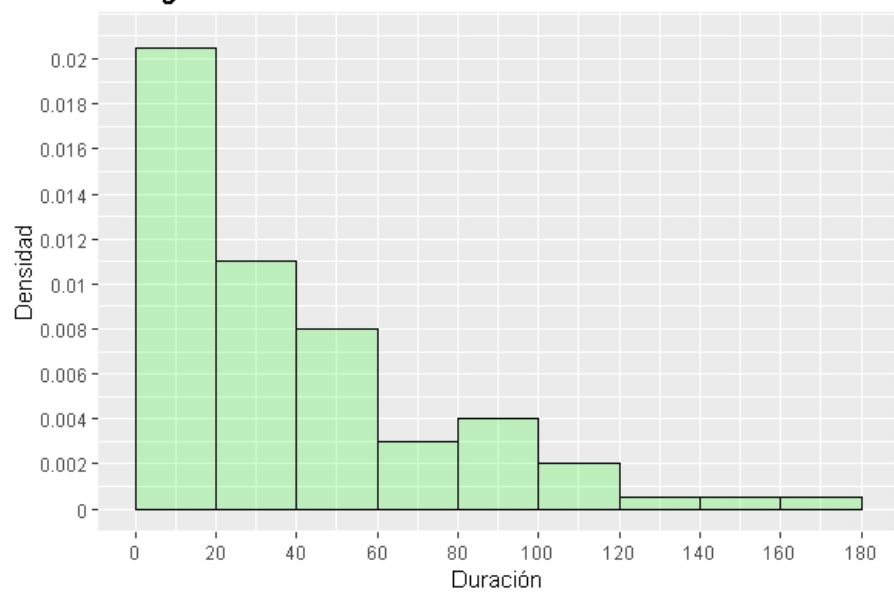
6. (optativo) Sea $X \sim \text{Exp}(1)$.
- Probar que $Y = \frac{1}{\lambda}X$ (con $\lambda > 0$) tiene distribución $\text{Exp}(\lambda)$.
 - ¿Qué resulta si graficamos los cuantiles de una $\text{Exp}(1)$ versus los cuantiles de una $\text{Exp}(\lambda)$?
7. Se tiene una muestra de 90 datos correspondientes a valores nominales de octanos en varios tipos de gasolina. Al final de la práctica se adjuntan los siguientes gráficos: histograma, q-q plot y función de distribución acumulada empírica.
- Mirando el q-q plot, ¿parecen los datos seguir una distribución normal? ¿Aproximadamente con qué parámetros?
 - Mirando el histograma, completar con valores aproximados:
 - La proporción de valores entre 75 y 80 es
 - La proporción de valores entre 80 y 90 es
 - El 40% de los datos es menor que
 - El 13% de los datos es mayor a
 - Repetir (b) mirando el gráfico de la función de distribución acumulada empírica. ¿Resultó más fácil o más difícil? Repetir la segunda pregunta de (b) mirando el q-q plot. ¿Resultó más fácil o más difícil?
8. Se tienen 101 datos que representan la duración (en minutos) de una cierta reacción química. Al final de la práctica se adjuntan los siguientes gráficos: histograma, q-q plot versus cuantiles normales, q-q plot versus cuantiles de $\text{Exp}(1)$ y función de distribución acumulada empírica.
- En base al q-q plot, ¿aparentan los datos seguir una distribución normal?
 - Mirando el histograma, ¿qué distribución parecería seguir la muestra?
 - Mirar el q-q plot versus cuantiles de una $\text{Exp}(1)$, ¿se ve una recta? ¿cuál parece ser aproximadamente el parámetro de la exponencial que representa a la muestra?
 - Mirando el histograma, completar con valores aproximados:
 - La proporción de valores mayores que 100 es
 - La proporción de valores entre la mediana y 100 es
 - Hay un 38% de los datos entre 20 y
 - Mirando el gráfico de la función de distribución acumulada empírica, completar:
 - El 27% de los datos es mayor a
9. Se tiene una muestra de 20 datos x_1, \dots, x_{20} cuya desviación estándar muestral es 7 y se tiene otra muestra de 30 datos y_1, \dots, y_{30} con desviación estándar muestral igual a 9. Se forma una muestra de 50 datos uniendo las dos anteriores $x_1, \dots, x_{20}, y_1, \dots, y_{30}$. Si se sabe que $\bar{x} = \bar{y}$, hallar la desviación estándar muestral de esta nueva muestra.

10. Sea x_1, \dots, x_n un conjunto de datos, cuya mediana es 4 y su MAD (sin estandarizar) es 5.
- Sean $y_i = 3x_i - 2$; $i = 1, \dots, n$. Hallar la mediana y el MAD (sin estandarizar) de y_1, \dots, y_n .
 - Repetir el ítem (a) pero considerando $y_i = -4x_i + 3$.
 - Hallar todos los valores de a y b para que los datos $y_i = ax_i + b$ tengan mediana 0 y MAD igual a 2.
11. Se realiza el q-q plot de una muestra ordenada x_1, \dots, x_n .
- Si la muestra es de 199 datos y el punto $(0.44, x_k)$ pertenece al gráfico, hallar k .
 - Si el punto $(2.06, a)$ pertenece al gráfico y la muestra proviene de una distribución $U[2, 9]$, hallar aproximadamente el valor de a .
 - Si el punto $(-1.21, a)$ pertenece al gráfico y la muestra proviene de una distribución $Exp(\frac{1}{3})$, hallar aproximadamente el valor de a .
 - Si el punto $(a, 1.816)$ pertenece al gráfico y la muestra proviene de una distribución $U[-1, 7]$, hallar aproximadamente el valor de a .





Histograma de Reacción Química



Distribución acumulada empírica de Reacción Química

