2. Sea X el número de bolas blancas obtenidas al sacar dos de una urna con 10 bolas de las que 8 son blancas. Calcular:

   a) Función masa de probabilidad y función de distribución.
   b) Media, mediana y moda, dando la interpretación de cada una de estas medidas.
   c) Intervalo intercuartílico, especificando su interpretación.

Sea X = "nº de bolas blancas obtenidas al sacar 2 de una urna con 10 bolas de las que 8 son blancas".

a)

- Primero calculamos la función masa de probabilidad:

Los posibles valores de la variable aleatoria X son $\{0,1,2\}$

$$P(X=0) = \frac{\binom{2}{2}\binom{8}{0}}{\binom{10}{2}} = \frac{1}{45}$$

$$P(X=1) = \frac{\binom{2}{1}\binom{8}{1}}{\binom{10}{2}} = \frac{16}{45} \qquad \Rightarrow \qquad p_X(x) = P(X=x) = \frac{\binom{2}{2-x}\cdot\binom{8}{x}}{\binom{10}{2}}$$

$$P(X=2) = \frac{\binom{2}{0}\binom{8}{2}}{\binom{10}{2}} = \frac{28}{45}$$

- Calculamos ahora la función de distribución:

$$F_X(x) = P(X \leq x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{1}{45} & \text{si} \quad 0 \leq x < 1 \\ \frac{17}{45} & \text{si} \quad 1 \leq x < 2 \\ 1 & \text{si} \quad x \geq 2 \end{cases}$$

b)

- Calculamos la media de X:

La media de X viene dada por:

$$E[X] = \sum_{i=1}^{3} x_i \cdot P[X=x_i] = 0\cdot\frac{1}{45} + 1\cdot\frac{17}{45} + 2\cdot\frac{28}{45} = 1{,}6\bar{2} \simeq 2 \text{ bolas blancas se sacan.}$$

Por tanto, la media de X será la suma de cada uno de los posibles valores de X multiplicados por sus respectivas probabilidades.

• Calculamos la mediana de X:

La mediana viene dada por:

$$M_e \in \mathbb{R} : \quad P(X \leq M_e) \geq \frac{1}{2} \quad \wedge \quad P(X \geq M_e) \geq \frac{1}{2}$$

Es evidente que el valor que lo cumple es $x_i = 2 \Rightarrow M_e(X) = 2$

No nos aporta demasiada información, ya que nos dice que en una mitad saldrán 2 bolas blancas o menos y en la otra mitad exactamente 2 bolas.

• Calculamos la moda de X:

La moda de X es el valor que maximiza la función masa de probabilidad.

$M_o(X) = 2$

c)

Vamos a considerar el intervalo o recorrido intercuartílico como $R = Q_3 - Q_1$.

$Q_3 = 2$, $Q_1 = 1 \Rightarrow R = Q_3 - Q_1 = 2 - 1 = 1$

Es en este intervalo donde se encuentra el 50% central de los datos.