



UNIDAD TEMÁTICA 2 ESTADÍSTICA DESCRIPTIVA

ENUNCIADO

Las cuotas de ocupación de disco duro (en Mbytes) para diferentes usuarios de una cierta estación de trabajo son:

35 45 47 50 31 30 25 33 35 40 45 47 49 42 40 50 46 55 42 46

(1º) Indica si la muestra es homogénea, en el sentido que no existan puntuaciones atípicas. Siendo dicha muestra representativa de la población de internautas que acceden a los servicios de Internet de dicha estación de trabajo. (2º) Calcula el intervalo de confianza del 99 % para la desviación típica que presenta la cuota de ocupación de dicha estación de trabajo. (3º) Suponiendo que la distribución de las cuotas de ocupación es normal, ¿cuál es la probabilidad de que la cuota de ocupación de un internauta sea superior a 46 Mbytes o inferior a 34 Mbytes?

Resolución:

Que una serie estadística sea homogénea equivale a no encontrar observaciones atípicas (outliers) en dicha muestra.

En consecuencia, para responder a este ejercicio es necesario calcular los tres cuartiles (por ejemplo, construyendo la tabla de frecuencias correspondiente) previa ordenación de la serie estadística. La expresión que permite determinar cualquier cuantil es:

$$C_{p/s} = l_i + \frac{\frac{p}{s}n - F_{i-1}}{f_i} h$$

Para obtener el rango intercuartílico ($RIC = Q_3 - Q_1$) se han de determinar el tercer (75 %) y el primer (25 %) cuartil. En efecto:

$$Q_3: i = 15.5 \Rightarrow Q_3 = 47$$

$$Q_1: i = 5.5 \Rightarrow Q_1 = 35$$

Por consiguiente, $RIC = 12$, que incluye al 50 % de los datos de la serie estadística.

La determinación de las observaciones atípicas queda definida a partir de los valores frontera



$$\begin{aligned}Q_1 - 3RIC &= -1 \\Q_1 - 1.5RIC &= 17 \\Q_3 + 1.5RIC &= 65 \\Q_3 + 3RIC &= 83\end{aligned}$$

con lo que se definen las zonas de la figura



Se observa que la muestra es homogénea (no existen observaciones atípicas) ya que toda la serie estadística satisface que $Q_1 - 1.5RIC < x_i < Q_3 + 1.5RIC$, $\forall i$.

Para esta serie estadística se obtiene que: $\bar{x} = 41.65 Mb$ y $\hat{s}_x = 7.8959 Mb$ con

$$\sum_{i=1}^n x_i = 833 Mb; \quad \sum_{i=1}^n x_i^2 = 35879 Mb^2$$

Se deja al alumno su obtención, ya que serían necesarios para responder al resto de cuestiones del ejercicio, para cuyo efecto sería necesario disponer de la correspondiente tabla de frecuencias (como corresponden a temas posteriores se resolverán en su momento)