

CONSTRUCCIÓN E INTERPRETACIÓN DEL DIAGRAMA DE CAJA Y BRAZOS

Los diagramas de caja y brazos son gráficas muy apropiadas para mostrar el comportamiento de los datos cuando interesa presentarlos estratificados por alguna variable cualitativa. Para la construcción del diagrama de caja y brazos es necesario realizar el cálculo de los cuartiles.

Para la construcción de un diagrama de caja se siguen los siguientes pasos:

1. Se construye una escala de referencia, ya sea horizontal o vertical.
2. Se calculan los cuartiles (Q_1 , Q_2 y Q_3) y el rango intercuartílico (RI).

$$RI = Q_3 - Q_1$$

3. Se calculan dos valores f_1 y f_3 que llamaremos barreras interiores, de la siguiente manera:

$$f_1 = Q_1 - 1.5RI \quad f_3 = Q_3 + 1.5RI$$

4. Se calculan los puntos a_1 y a_3 . Llamados valores adyacentes. El punto a_1 es el dato más cercano a f_1 (f_1 puede coincidir con un valor de los datos) sin ser el menor de esa barrera, El punto a_3 es el dato más cercano a f_3 (f_3 puede coincidir con un valor de los datos) pero no mayor que esa barrera.
5. Se localizan todos los puntos en la escala horizontal o vertical, según se haya elegido.
6. Se dibuja una caja con los extremos en el primer y tercer cuartil. Se dibuja la mediana con una línea interior en el lugar adecuado.
7. Los valores adyacentes se unen a la caja por medio de líneas, esto genera los brazos de las cajas.
8. Si existen datos que queden fuera de las barreras interiores, se dibujaran con círculos abiertos. Estos datos se conocen como datos atípicos.

Ejemplo 1:

Los siguientes datos corresponden a los tiempos de hospitalización, en días, después de una cirugía de cráneo.

8, 9, 9, 12, 13, 15, 15, 17, 23, 24, 21, 28, 33, 36, 37, 26, 38, 21, 45, 44, 78

Realizando los cálculos para los cuartiles se tiene:

$$Q_1=14, \quad Q_2=23, \quad Q_3= 36;$$

El rango intercuartílico es

$$RI = 36 - 14 = 22$$

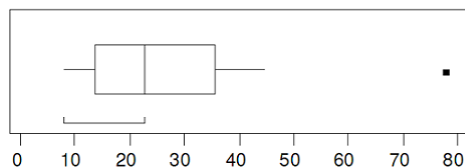
y las barreras interiores son:

$$f_1 = 14 - 1,5 (22) = -19,$$

$$f_3 = 36 + 1,5 (22) = 69$$

Solamente tenemos un dato que cae fuera de las barreras interiores, en el lado derecho, que es el 78. Se procede a construir el diagrama de caja, quedando de la siguiente manera:

Tiempos de hospitalización

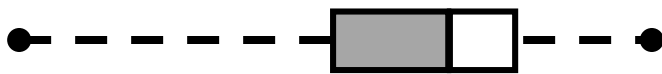


Del diagrama de caja podemos establecer análisis como los siguientes:

El 78 es un dato atípico, el 50% de los pacientes duraron hospitalizados 23 o más días, etc.

Tenga en cuenta para el análisis la información que proporcionan los cuartiles.

En el diagrama de caja se puede observar que los tiempos de hospitalización están sesgados (inclinados) a la derecha.



Distribución sesgada a la izquierda



Distribución sesgada a la derecha



Distribución rectangular