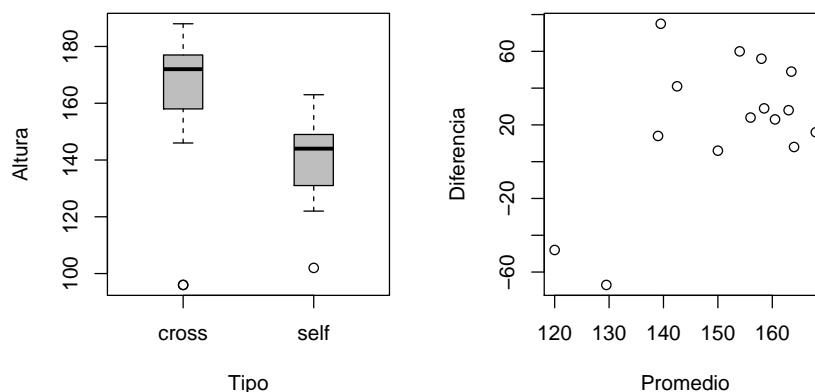


1. Se desarrolló un experimento para comparar la altura para dos grupos plantas de maíz sometidas a diferentes tratamientos de fertilización. Se obtuvo los siguientes resultados: Considere además

Tratamiento	n	min	Q_1	me	media	Q_3	max	s	γ_1	γ_2
Cross (x)	15	96	158	172	161.5	177	188	28.94	-1.50	0.90
Self (y)	15	102	131	144	140.6	149	163	16.41	-0.69	-0.07

los gráficos:



Suponga que: $\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = 0$ y sea

$$z_i = x_i - y_i, \quad i = 1, \dots, n.$$

Usando la información disponible:

- Realice un análisis descriptivo de la variable x : altura de las plantas para el tratamiento de fertilización Cross (escriba su respuesta en no más de 5 líneas).
- Obtenga la media (\bar{z}), varianza (s_z^2) y el coeficiente de variación CV_z .
- Considere la siguiente transformación

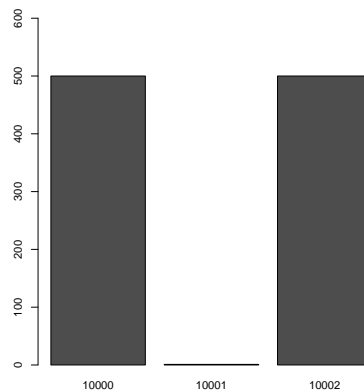
$$u_i = -1.76 y_i + 408.96$$

Obtenga la media (\bar{u}) y desviación estándar (s_u).

2. Considere las notas del Certamen 1 de MAT045 del año pasado en cada uno de los dos paralelos que se dictaron (P100 y P101):

Paralelo	n	min	P_{25}	mediana	media	P_{75}	max	s
P100	40	19	35	43.5	45.4	56.25	76	12.8
P101	13	21	34	39.0	41.8	43.00	78	17.8

- a) El profesor del paralelo P100 ha decidido sumar los puntos obtenidos en una tarea en clases al Certamen 1. Suponga que todos los alumnos obtuvieron 5 puntos en la tarea. Sea X : nota del Certamen 1 (incorporando puntos obtenidos en la tarea) del paralelo P100. Obtenga la media, desviación estándar y CV de X .
- b) El profesor del paralelo P101 ha decidido incorporar la asistencia a clases al Certamen 1, como un factor que pondera dicha nota. Suponga que todos los alumnos tienen un factor igual a 1.1. Sea Y : nota del Certamen 1 (ponderada por el factor) del P101. Calcule el CV de Y .
- c) Luego de las decisiones de los profesores, ¿Cuál paralelo es el que tuvo la mayor variabilidad en las notas del Certamen 1?
3. Suponga un conjunto de datos con 1001 observaciones, tal que 500 observaciones son 10000, 500 observaciones corresponden a 10002 y un dato es 10001, cuya representación gráfica es:



Calcule la media, mediana, desviación estándar, el coeficiente de asimetría b_1 y el coeficiente de curtosis b_2 .

4. Considere x_1, x_2, \dots, x_n un conjunto de n datos. Verifique que:

$$s^2 = \frac{1}{2n(n-1)} \sum_{i=1}^n \sum_{j=1}^n (x_i - x_j)^2.$$

5. Sea \bar{x}_n el promedio de un conjunto con n datos, digamos x_1, \dots, x_n . Entonces verifique que:

$$\bar{x}_n = \bar{x}_{n-1} + \frac{1}{n}(x_n - \bar{x}_{n-1}).$$

Adicionalmente suponga que se dispone de “pesos” $\omega_1, \dots, \omega_n$ asociados a cada uno de los datos. Obtenga una formula recursiva para el promedio ponderado

$$\bar{x}_n = \frac{\sum_{i=1}^n \omega_i x_i}{\sum_{i=1}^n \omega_i} = \bar{x}_{n-1} - \frac{\omega_n}{W_n} (x_n - \bar{x}_{n-1}), \quad W_n = \sum_{i=1}^n \omega_i$$

6. Los trabajadores de una empresa, cuya tarea es clasificar y envasar fruta, obtuvieron los siguientes salarios semanales (clasificados según sexo).

Ingreso (UM)	Mujeres	Hombres
65 – 75	10	0
75 – 85	15	0
85 – 95	60	5
95 – 105	15	10
105 – 115	10	50
115 – 125	0	25
125 – 135	0	10
Total	110	100

- a) Calcule la media y mediana para cada grupo, interprete resultados.
- b) Obtenga estadísticas de resumen para el total de trabajadores.
7. La planta de ropa Levi's de Albuquerque (New Mexico) llevó a cabo un estudio para comparar la calidad de las telas suministradas por dos proveedores A y B. Los datos corresponden a las medidas de desgaste debido a defectos en la tela. Los datos disponibles se refieren a 22 mediciones del desgaste por defectos para cada una de las fábricas.

Fábrica A		Fábrica B	
0.12	0.03	1.64	0.63
1.01	0.35	-0.60	0.90
-0.20	-0.08	-1.16	0.71
0.15	1.94	-0.13	0.43
-0.30	0.28	0.40	1.97
-0.07	1.30	1.70	0.30
0.32	4.27	0.38	0.76
0.27	0.14	0.43	7.02
-0.32	0.30	1.04	0.85
-0.17	0.24	0.42	0.60
0.24	0.13	0.85	0.29

- a) Obtenga el promedio, mediana, s^2 , coeficiente de variación CV y el coeficiente de sesgo de Galton, b_G para cada una de las fábricas.
- b) Sea x_i y y_i , la i -ésima medida de desgaste de la fábrica A y B, respectivamente. Obtenga, $z_i = x_i - y_i$, para $i = 1, \dots, 22$. Calcule el promedio, mediana, varianza y coeficiente de variación basado en los datos z_1, z_2, \dots, z_{22} .

Observación: Los datos están disponibles en el fichero `levis.csv`.