1. Tenemos el conjunto de datos:

$$\boldsymbol{x} = \{\underbrace{10\,000, 10\,000, \cdots, 10\,000}_{500 \text{ observaciones}}, 10\,001, \underbrace{10\,002, 10\,002, \cdots, 10\,002}_{500 \text{ observaciones}}\}.$$

De este modo, es evidente que

$$\mathsf{me}(\boldsymbol{x}) = 10\,001,$$

mientras que el promedio muestral es dado por:

$$\begin{aligned}
\overline{x} &= \frac{500 \cdot 10\,000 + 10\,001 + 500 \cdot 10\,002}{1001} = \frac{500 \cdot 10\,000 + 10\,000 + 1 + 500(10\,000 + 2)}{1001} \\
&= \frac{500 \cdot 10\,000 + 10\,000 + 500 \cdot 10\,000 + 1 + 500 \cdot 2}{1001} = \frac{10\,000 \cdot 1001 + 1001}{1001} \\
&= \frac{1001(10\,000 + 1)}{1001} = 10\,001.
\end{aligned}$$

Sea $u_i = x_i - \overline{x}$, para $i = 1, \dots, 1001$. Es decir, tenemos:

$$\boldsymbol{u} = \{\underbrace{-1, -1, \cdots, -1}_{500 \text{ obs}}, 0, \underbrace{1, 1, \cdots, 1}_{500 \text{ obs}}\}.$$

Podemos calcular la varianza muestral como:

$$s^2 = \frac{1}{1001 - 1} \sum_{i=1}^{1001} u_i^2 = \frac{1}{1000}\big(500(-1)^2 + 0 + 500(1)^2\big) = \frac{1000}{1000} = 1.$$

Como $s = 1$, sigue que $z_i = (x_i - \overline{x})/s = u_i$, para $i = 1, \dots, 1001$. Esto permite calcular

$$b_1 = \frac{1}{1001} \sum_{i=1}^{1001} \Big(\frac{x_i - \overline{x}}{s}\Big)^3 = \frac{1}{1001} \sum_{i=1}^{1001} z_i^3 = \frac{1}{1001}\big(500(-1)^3 + 0 + 500(1)^3\big) = 0.$$

Finalmente,

$$\begin{aligned}
b_2 &= \Big\{\frac{1}{1001} \sum_{i=1}^{1001} \Big(\frac{x_i - \overline{x}}{s}\Big)^4\Big\} - 3 = \frac{1}{1001} \sum_{i=1}^{1001} z_i^4 - 3 = \frac{1}{1001}\big(500(-1)^4 + 0 + 500(1)^4\big) - 3 \\
&= \frac{1000}{1001} - 3 = -2.001
\end{aligned}$$

2. Sabemos que $\widehat{\alpha} = \overline{y} - \widehat{\beta}\,\overline{x}$, de este modo

$$e_i = y_i - \widehat{\alpha} - \widehat{\beta}\,x_i = y_i - \overline{y} - \widehat{\beta}(x_i - \overline{x}).$$

En efecto,

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} \big(y_i - \overline{y} - \widehat{\beta}(x_i - \overline{x})\big) = \sum_{i=1}^{n} (y_i - \overline{y}) - \widehat{\beta} \sum_{i=1}^{n} (x_i - \overline{x}) = 0.$$

Por otro lado,

$$\begin{aligned}\sum_{i=1}^n e_i\,\widehat{y}_i &= \sum_{i=1}^n (y_i - \overline{y} - \widehat{\beta}(x_i - \overline{x}))(\overline{y} + \widehat{\beta}(x_i - \overline{x})) \\ &= \overline{y}\sum_{i=1}^n (y_i - \overline{y} - \widehat{\beta}(x_i - \overline{x})) + \widehat{\beta}\sum_{i=1}^n (y_i - \overline{y} - \widehat{\beta}(x_i - \overline{x}))(x_i - \overline{x}),\end{aligned}$$

el primer término es cero pues, $\sum_{i=1}^n e_i = 0$. Es decir, tenemos que:

$$\sum_{i=1}^n e_i\,\widehat{y}_i = \widehat{\beta}\Big[\sum_{i=1}^n (y_i - \overline{y})(x_i - \overline{x}) - \widehat{\beta}\sum_{i=1}^n (x_i - \overline{x})^2\Big].$$

Notando que $\widehat{\beta} = \mathsf{cov}(\boldsymbol{x}, \boldsymbol{y}) / \mathsf{var}(\boldsymbol{x})$, se verifica el resultado.

3.a. Considere $x_i = u_i + 5$, $i = 1, \ldots, 40$. De este modo,

$$\overline{x} = \overline{u} + 5 = 45.4 + 5 = 50.4, \qquad s_x = s_u = 12.8.$$

Mientras que

$$\mathsf{CV}_x = \frac{s_x}{\overline{x}} = \frac{12.8}{50.4} = 0.2540.$$

3.b. En este caso, $y_i = 1.1\, v_i$, $i = 1, \ldots, 13$. Así,

$$\overline{y} = 1.1\,\overline{v} = 1.1 \cdot 41.8 = 45.98, \qquad s_y = 1.1\, s_v = 1.1 \cdot 17.8 = 19.58.$$

Además

$$\mathsf{CV}_y = \frac{s_y}{\overline{y}} = \frac{17.8}{41.8} = 0.4258.$$

3.c. La mayor variabilidad se obtuvo en el Paralelo $\mathsf{P101}$.

4.a. Sea

X : mediciones de desgaste por defectos en la fábrica A
Y : mediciones de desgaste por defectos en la fábrica B

Tenemos

$$\sum_{i=1}^{22} x_i = 9.95, \qquad \sum_{i=1}^{22} y_i = 19.43,$$

de ahí que

$$\overline{x} = \frac{9.95}{22} = 0.4523, \qquad \overline{y} = \frac{19.43}{22} = 0.8832.$$

Además

$$\sum_{i=1}^{22} (x_i - \overline{x})^2 = 21.1348, \qquad \sum_{i=1}^{22} (y_i - \overline{y})^2 = 49.5031,$$

esto permite obtener

$$\mathsf{var}(\boldsymbol{x}) = \frac{21.1348}{22-1} = 1.0064, \qquad \mathsf{var}(\boldsymbol{y}) = \frac{49.5031}{22-1} = 2.3573.$$

Por tanto,

$$\begin{aligned}\mathsf{CV}_x &= \frac{s_x}{\overline{x}} = \frac{\sqrt{1.0064}}{0.4523} = \frac{1.0032}{0.4523} = 2.2181,\\ \mathsf{CV}_y &= \frac{s_y}{\overline{y}} = \frac{\sqrt{2.3573}}{0.8832} = \frac{1.5353}{0.8832} = 1.7384.\end{aligned}$$

Por otro lado, para los datos de la fábrica A:

$$Q_1(\boldsymbol{x}) = -0.0450, \qquad Q_2(\boldsymbol{x}) = 0.1950, \qquad Q_3(\boldsymbol{x}) = 0.3150.$$

De ahí que,

$$\begin{aligned}b_{\mathsf{G}}(\boldsymbol{x}) &= \frac{(Q_3(\boldsymbol{x}) - Q_2(\boldsymbol{x})) - (Q_2(\boldsymbol{x}) - Q_1(\boldsymbol{x}))}{Q_3(\boldsymbol{x}) - Q_1(\boldsymbol{x})}\\ &= \frac{(0.3150 - 0.1950) - (0.1950 + 0.0450)}{0.3150 + 0.0450} = \frac{0.1200 - 0.2400}{0.3600}\\ &= -\frac{0.1200}{0.3600} = -0.3333.\end{aligned}$$

Análogamente, para los datos de la fábrica B,

$$Q_1(\boldsymbol{y}) = 0.3850, \qquad Q_2(\boldsymbol{y}) = 0.6150, \qquad Q_3(\boldsymbol{y}) = 0.8875.$$

Esto lleva a

$$\begin{aligned}b_{\mathsf{G}}(\boldsymbol{y}) &= \frac{(0.8875 - 0.6150) - (0.6150 - 0.3850)}{0.8875 - 0.3850} = \frac{0.2725 - 0.2300}{0.5025}\\ &= \frac{0.0425}{0.5025} = 0.0846.\end{aligned}$$

4.b. Note que

$$\sum_{i=1}^{22} x_i y_i = 6.3194.$$

De este modo,

$$\begin{aligned}\mathsf{cov}(\boldsymbol{x}, \boldsymbol{y}) &= \frac{1}{22-1}\Big(\sum_{i=1}^{22} x_i y_i - n\overline{x}\overline{y}\Big) = \frac{1}{21}\big(6.3194 - 22 \cdot 0.4523 \cdot 0.8832\big)\\ &= -\frac{2.4683}{21} = -0.1175.\end{aligned}$$

Esto permite obtener,

$$r = \mathsf{cor}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\mathsf{cov}(\boldsymbol{x}, \boldsymbol{y})}{\sqrt{\mathsf{var}(\boldsymbol{x})\,\mathsf{var}(\boldsymbol{y})}} = \frac{-0.1175}{1.0032 \cdot 1.5353} = -0.0763.$$

4.c. Para los datos transformados es fácil notar que

$$\begin{aligned}\overline{z} &= \frac{1}{n}\sum_{i=1}^{n} z_i = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i) = \frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{n}\sum_{i=1}^{n} y_i = \overline{x} - \overline{y} \\ &= 0.4523 - 0.8832 = -0.4309\end{aligned}$$

y

$$\begin{aligned}\mathsf{var}(\boldsymbol{z}) &= \mathsf{var}(\boldsymbol{x}) + \mathsf{var}(\boldsymbol{y}) - 2\,\mathsf{cov}(\boldsymbol{x}, \boldsymbol{y}) = 1.0064 + 2.3573 - 2(-0.1175) \\ &= 3.5988.\end{aligned}$$

De este modo,

$$\mathsf{CV}_z = \frac{\sqrt{\mathsf{var}(\boldsymbol{z})}}{\overline{z}} = -\frac{1.8970}{0.4309} = -4.4024.$$

Finalmente, los datos ordenados $z_{(1)}, z_{(2)}, \dots, z_{(22)}$ son:

-6.88	-1.77	-1.69	-1.52	-1.36	-0.79	-0.70	-0.61
-0.60	-0.59	**-0.55**	**-0.55**	-0.36	-0.16	-0.16	-0.06
0.28	0.96	1.00	1.51	1.61	3.51		

De ahí que,

$$\mathsf{me}(\boldsymbol{z}) = \frac{1}{2}\big(-0.55 + (-0.55)\big) = -0.55.$$