Exercises solution

> **Solution** n°1

Exercice p. 3

- ***Calcul of statistics :***

1. ***Means of the Variables*** The mean of precipitation P is:
   $\bar{P} = \frac{12.04 + 17.18 + 11.83 + 6.23 + 16.99 + 3.87}{6} = \frac{68.14}{6} = 11.3567 = 11.36$
   The mean of maximum temperature $t_{\max}$ is:
   $\bar{t}_{\max} = \frac{23.7 + 15.5 + 13.1 + 13.5 + 21.1 + 20.3}{6} = \frac{107.2}{6} = 17.8667 = 17.87$
   The mean of minimum temperature $t_{\min}$ is:
   $\bar{t}_{\min} = \frac{5.9 + (-1.8) + 2.8 + (-2.4) + 7.2 + (-0.9)}{6} = \frac{10.8}{6} = 1.8.$
2. ***Variances***
   Variance of P (Precipitation)
   $\mathrm{Var}(P) = \frac{1}{5}\sum_{i=1}^{6}(P_i - \bar{P})^2$ $\mathrm{Var}(P) = \frac{1}{6}\left((12.04 - 11.36)^2 + (17.18 - 11.36)^2 + (11.83 - 11.36)^2 + (6.23 - 11.36)^2 + (16.99 - 11.36)^2 + (3.87 - 11.36)^2\right)$
   $\mathrm{Var}(P) = \frac{1}{5} \cdot 148.67 = 29.73\ \mathrm{cm}^2$. 148.66953 97.79333 84.86000
   Standard division of P6 (Precipitation):
   $\sigma_P = \sqrt{\mathrm{Var}(P)} = \sqrt{29.73} = 5.45.$
   Variance of $t_{\max}$ ***(Maximum Temperature)***
   $\mathrm{Var}(t_{\max}) = \frac{1}{5}\left((23.7 - 17.87)^2 + (15.5 - 17.87)^2 + (13.1 - 17.87)^2 + (13.5 - 17.87)^2 + (21.1 - 17.87)^2 + (20.3 - 17.87)^2\right)$
   $\mathrm{Var}(t_{\max}) = \frac{1}{5}(44.91 + 2.34 + 15.33 + 12.45 + 16.72 + 10.88)$
   $\mathrm{Var}(t_{\max}) = \frac{1}{5} \cdot 97.79 = 19.558\ °\mathrm{C}^2$
   Standard division of $t_{\max}$ ***(Maximum Temperature):***
   $\sigma_{t_{\max}} = \sqrt{\mathrm{Var}(t_{\max})} = \sqrt{17.10} = 4.42.$
   Variance of $t_{\min}$ ***(Minimum Temperature)***
   $\mathrm{Var}(t_{\min}) = \frac{1}{5}\left((5.9 - 1.80)^2 + (-1.8 - 1.80)^2 + (2.8 - 1.80)^2 + (-2.4 - 1.80)^2 + (7.2 - 1.80)^2 + (-0.9 - 1.80)^2\right).$
   $\mathrm{Var}(t_{\min}) = \frac{1}{5}(16.81 + 13.69 + 1.00 + 17.64 + 29.16 + 7.29)$
   $\mathrm{Var}(t_{\min}) = \frac{1}{5} \cdot 84.86 = 16.97\ °\mathrm{C}^2.$
   Standard division of $t_{\min}$ ***(Minimum Temperature)***
   $\sigma_{t_{\min}} = \sqrt{\mathrm{Var}(t_{\min})} = \sqrt{16.97} = 4.11.$

- ***Calcul of Dispersion Matrix:*** $\Sigma = \frac{1}{n-1}Z^tZ$ where $Z = X - I_n g^t$ is ***the centered data*** set:
  where g is the vector of column means and I_n identity matrix.
  We get :

The Centered Matrix:

Each element of the centered matrix is calculated as:

$$Z_{\text{centered}}(i,j) = X(i,j) - \bar{X}(j)$$

Thus, subtracting the column means from each element:

$$Z_{\text{centered}} = \begin{bmatrix} 12.04 - 11.36 & 23.7 - 17.87 & 5.9 - 1.8 \\ 17.18 - 11.36 & 15.5 - 17.87 & -1.8 - 1.8 \\ 11.83 - 11.36 & 13.1 - 17.87 & 2.8 - 1.8 \\ 6.23 - 11.36 & 13.5 - 17.87 & -2.4 - 1.8 \\ 16.99 - 11.36 & 21.1 - 17.87 & 7.2 - 1.8 \\ 3.87 - 11.36 & 20.3 - 17.87 & -0.9 - 1.8 \end{bmatrix}$$

Thus, the final centered matrix is:

$$Z_{\text{centered}} = \begin{bmatrix} 0.68 & 5.83 & 4.10 \\ 5.82 & -2.37 & -3.60 \\ 0.47 & -4.77 & 1.00 \\ -5.13 & -4.37 & -4.20 \\ 5.63 & 3.23 & 5.40 \\ -7.49 & 2.43 & -2.70 \end{bmatrix}$$

Computation of $\text{Zcentered}^T Z_{\text{centered}}$

Compute Z^T_{centered}

$$Z^T_{\text{centered}} = \begin{bmatrix} 0.68 & 5.82 & 0.47 & -5.13 & 5.63 & -7.49 \\ 5.83 & -2.37 & -4.77 & -4.37 & 3.23 & 2.43 \\ 4.10 & -3.60 & 1.00 & -4.20 & 5.40 & -2.70 \end{bmatrix}$$

The product is:

$$Z^T_{\text{centered}} Z_{\text{centered}} = \begin{bmatrix} 148.66 & 10.33 & 54.47 \\ 10.33 & 97.79 & 56.9 \\ 54.47 & 56.9 & 84.86 \end{bmatrix}$$

So the covariance matrix is given by:

- where : $\text{Var}(P) = 29.73, \quad \text{Var}(t_{\max}) = 19.56, \quad \text{Var}(t_{\min}) = 16.97$

$\text{Cov}(P, t_{\max}) = 2.06, \quad \text{Cov}(P, t_{\min}) = 10.89, \quad \text{Cov}(t_{\max}, t_{\min}) = 11.38$

- ***Calcul the correlation Matrix :*** $R = \dfrac{1}{n-1}(Z^*)^t Z^*$ where :

$Z^* = Z \times D(1/\sigma)$. ***The centered and reduced matrix for the given dataset is:***

$$Z^* = \begin{bmatrix} 0.13 & 1.32 & 1.00 \\ 1.07 & -0.54 & -0.87 \\ 0.09 & -1.08 & 0.24 \\ -0.94 & -0.99 & -1.02 \\ 1.03 & 0.73 & 1.31 \\ -1.37 & 0.55 & -0.66 \end{bmatrix}$$

The correlation matrix is then written as:

$$R = \frac{1}{n-1}(Z^*)^t Z^*) = \text{Corr} = \begin{bmatrix} 1.00 & 0.09 & 0.49 \\ 0.09 & 1.00 & 0.62 \\ 0.49 & 0.62 & 1.00 \end{bmatrix}$$

Comments:

$\text{Corr}(P, t_{max}) = 0.09$: This value is very close to zero, indicating a ***very weak*** or ***nonexistent relationship*** between precipitation and maximum temperature.

$Corr(P, t_{min}) = 0.49$: There is also a moderate positive correlation between precipitation (P) and minimum temperature (t_{min}), suggesting that higher precipitation is somewhat associated with an increase in minimum temperature.

$Corr(t_max, t_{min}) = 0.62$: This value indicates a ***moderate to strong positive*** correlation, meaning that warmer days tend to have warmer nights as well.

Remark: Correlation Coefficients can also be calculated as follows:

$$\mathrm{Corr}(P, t_{\max}) = \frac{\mathrm{Cov}(P, t_{\max})}{\sigma_P \cdot \sigma_{t_{\max}}} = \frac{10.57}{4.99 \cdot 4.14} = 0.51$$

$$\mathrm{Corr}(P, t_{\min}) = \frac{\mathrm{Cov}(P, t_{\min})}{\sigma_P \cdot \sigma_{t_{\min}}} = \frac{8.43}{4.99 \cdot 3.78} = 0.45.$$

$$\mathrm{Corr}(t_{\max}, t_{\min}) = \frac{\mathrm{Cov}(t_{\max}, t_{\min})}{\sigma_{t_{\max}} \cdot \sigma_{t_{\min}}} = \frac{12.04}{4.14 \cdot 3.78} = \frac{12.04}{15.65} = 0.77.$$