

## Exercice Supplémentaire (Quanti + Quali)

X : type de forêt	Y : hauteurs des arbres				Total
	de 21 à 23 m	de 23 à 25 m	25 à 27 m	27 à 29 m	
Type I	0	3	7	3	13
Type 2	1	6	6	1	14
Type 3	5	5	0	0	10
Total	6	14	13	4	37

Calculer les variances inter-classes, intra classe et la variance totale.

Donner le rapport de corrélation entre X et Y.

### Corrigé :

On a 2 variables.

X : type de forêt (de Type qualitatif)

Y : hauteur des arbres (quanti-continue).

On représente chaque classe de Y par leur centre.

$$\text{centre classe} = \frac{a+b}{2}$$

(1)

# ① Variance inter-classes :

$$\text{Var}^{(\text{inter})}(y) = \frac{1}{n} \sum_{l=1}^3 n_l \cdot (\bar{y}_l - \bar{y})^2$$

on a 3 médailles

$$\bar{y}_1 = \frac{1}{13} [(6 \times 22) + (3 \times 24) + (7 \times 26) + (3 \times 28)]$$

$\circlearrowleft$  eff type 1

$$= 26 \text{ m}$$

$\bar{y}_1, \bar{y}_2, \bar{y}_3$  sont les moyennes de Type 1, Type 2 et Type 3]

$$\bar{y}_2 = 25 \text{ m}$$

$$\bar{y}_3 = 23 \text{ m}$$

$$\bar{y} = \frac{1}{37} \cdot [(6 \times 22) + (14 \times 24) + (13 \times 26) + (4 \times 28)]$$

$\circlearrowleft$  total

$$= 24,8 \text{ m}$$

$$\text{Var}^{(\text{inter})}(y) = \frac{1}{37} \left[ 13(26 - 24,8)^2 + 14(25 - 24,8)^2 + 10(23 - 24,8)^2 \right]$$

$\text{Var}^{(\text{inter})}(y) = 1,47$

## ② Variance intra-classes

$$\text{Var}^{\text{intra}}(y) = \frac{1}{n} \sum_{l=1}^3 n_l \cdot S_{l,l}^2$$

$S_1^2 = \frac{1}{13} \left[ 0(\bar{y}_1 - \bar{y}_1)^2 + 3(\bar{y}_1 - 26)^2 + 7(26 - 26)^2 + 3(28 - 26)^2 \right] = 1,85$

$\bar{y}_1: \text{Moyenne de type 1}$

$S_2^2 = \frac{1}{14} \left[ 1(\bar{y}_2 - \bar{y}_2)^2 + 6(\bar{y}_2 - 25)^2 + 6(26 - 25)^2 + 1(28 - 25)^2 \right] = 2.14$

$S_3^2 = \frac{1}{10} \left[ 5(\bar{y}_3 - \bar{y}_3)^2 + 5(\bar{y}_3 - 23)^2 + 0(26 - 23)^2 + 0(28 - 23)^2 \right] = 0.5$

$$\text{Var}^{\text{intra}}(y) = \frac{1}{37} \left[ (13 \times 1,85) + (14 \times 2.14) + (10 \times 0.5) \right]$$

Moyenne des  
variancex

$$= 1.6$$

③

### ③ Variance totale: $V(y)$

$$\cdot V(y) = \frac{1}{n} \sum_{p=1}^n n_p (y_p - \bar{y})^2$$

Moyenne générale

$$= \frac{1}{37} \left[ 6(22 - 24,8)^2 + 14(24 - 24,8)^2 + 13(26 - 24,8)^2 + 4(28 - 24,8)^2 \right]$$

$$= 3,126 \text{ m}^2.$$

On ~  $V(y) = \text{Var}^{(\text{inter})}(y) + \text{Var}^{(\text{intra})}(y)$

$$= 1,437 + 1,6$$

$$= 3 \text{ m}^2$$

(Légère différence due au arrondissement des valeurs).

$$\frac{\text{Var}^{(\text{inter})}(y)}{\text{Var}^{\text{total}}(y)}$$

Rapport de corrélation:

$$= \frac{1,4}{3,126} \approx 0,45 = 45\%$$

On constate que le type de forêt explique à 45% la variabilité (ou variation) des hauteurs d'arbres.

④