

Exercice à Faire :

x_i	7.4	8.2	9.0	9.5	9.7	9.3	8.4	8.6
y_i	25	41	58	89	78	85	67	53

- Représenter le nuage de points (sur papier millimétré) avec une échelle convenable ;
- Déterminer l'équation de la droite de régression linéaire de y en x ;
- Représenter cette droite dans le repère précédent.

NB. Solutionner l'exercice par deux approches distinctes, l'une est purement analytique (à la main). L'autre approche est numérique, en utilisant le logiciel de traitement de données « R » ou « RStudio »

Exemple (Les étapes à suivre)

La distribution du nombre d'enfants (variable quantitative discrète) dans une population de 311 ménages, a donné lieu à la série statistique suivante :

x_i	10	10	10	10	10	50	50	50	50	50
y_i	13	18	16	15	20	86	90	88	88	92

I) Solution Analytique :

1 Nuage de points (sur papier millimétré) :

2 Droite de régression de y en x :

x_i	10	10	10	10	10	50	50	50	50	50	300
y_i	13	18	16	15	20	86	90	88	88	92	526
$x_i \cdot y_i$	130	180	160	150	200	4300	4500	4400	4400	4600	23020
x_i^2	100	100	100	100	100	2500	2500	2500	2500	2500	13000

On cherche une droite d'équation : $y = a \cdot x + b$

Avec : $b = \bar{y} - a \cdot \bar{x} = 52,6 - 1,81 \cdot 30 = -1,7$

$$\bar{x} = \frac{1}{N} \sum x_i = \frac{300}{10} = 30 ; \text{ avec le code R : mean(x)}$$

$$\bar{y} = \frac{1}{N} \sum y_i = \frac{526}{10} = 52,6 ; \text{ avec le code R : mean(y)}$$

$$\text{Et : } a = \frac{\text{cov}(x, y)}{v(x)} = \frac{724}{400} = 1,81$$

$$\text{cov}(x, y) = \frac{1}{N} \sum x_i \cdot y_i - \bar{x} \cdot \bar{y} = \frac{23020}{10} - 30 \cdot 52,6 = 724 ; \text{ avec le code R : cov}(x,y)$$

$$v(x) = \frac{1}{N} \sum x_i^2 - \bar{x}^2 = \frac{13000}{10} - 30^2 = 400 ; \text{ avec le code R : var}(x)$$

Donc l'équation de la droite de régression de y en x est la suivante : $y = 1,81 \cdot x - 1,70$

Remarque : Ces paramètres peuvent être obtenus en utilisant une calculatrice, un tableur excel, ou les fonctions du logiciel R (RStudio)

3 Tracer la droite de régression sur papier millimétré (avec le nuage de points)

II) Solution Numérique avec le logiciel d'analyse de données R, RStudio :

Commandes sous R

- Création du vecteur x : `x <- c(10,10,10,10,10,50,50,50,50,50)`
- Affichage du vecteur x : `x`
- Création du vecteur y : `y <- c(13,18,16,15,20,86,90,88,88,92)`
- Affichage du vecteur y : `y`
- Nuage de points : `plot(x,y)`
- Création du modèle linéaire : `modele <- lm(y ~ x)`
- Affichage du modèle linéaire : `modele`
- Traçage de la ligne de régression : `abline(modele)`

