

TD 04 - Statistiques bivariées

▷ **Exercice 1 :** Un constructeur automobile souhaite tester le freinage d'urgence de leur nouvelle voiture. On note v la vitesse initiale de la voiture (en km/h) et d la distance nécessaire pour que la voiture s'arrête. On relève les données suivantes :

| | | | | | | |
|----------|----|----|----|----|-----|-----|
| Vitesse | 30 | 50 | 70 | 90 | 110 | 130 |
| Distance | 7 | 16 | 31 | 52 | 78 | 123 |

1. Dessiner le nuage de point associée.
2. Un modèle de régression linéaire semble-t-il adéquat ?
3. Calculer le coefficient de régression linéaire entre d et v .
4. On souhaite à présent établir si la distance de freinage dépend du carré de la vitesse. On pose $t = v^2$. Déterminer la série statistique t .
5. Calculer le coefficient de régression linéaire entre d et t .

▷ **Exercice 2 :** Dans une entreprise, on observe l'ancienneté et les salaires de douze employés :

| | | | | | | |
|----------------|------|------|------|------|------|------|
| Ancienneté (X) | 7 | 15 | 15 | 16 | 5 | 12 |
| Salaire (Y) | 4100 | 6200 | 4400 | 7400 | 2900 | 5600 |
| Ancienneté (X) | 2 | 20 | 14 | 9 | 15 | 8 |
| Salaire (Y) | 2300 | 6500 | 6800 | 4100 | 5300 | 3500 |

1. Construire le nuage de points associées à cette série statistique.
2. Semble-t-il judicieux de définir un modèle de régression linéaire entre les deux données ?
3. Donner les lignes de code Python permettant de calculer $s_{X,Y}$, s_X , \bar{X} et \bar{Y} .
4. Le code précédent renvoie $s_{X,Y} = 6862,5$, $s_X^2 \simeq 25,58$, $\bar{X} = 11,5$ et $\bar{Y} = 4925$. Déterminer la droite de régression linéaire et la tracer sur le nuage de point.
5. Peut-on estimer le salaire d'un employé ayant 4 ans d'ancienneté ? 18 ans d'ancienneté ?

▷ **Exercice 3 :** L'évolution du chiffre d'affaire (en millions d'euros) d'une entreprise depuis sa création en 2002 est donnée dans le tableau suivant :

| | | | | | | | | |
|-----------------------|------|------|------|------|------|------|------|------|
| Année (X) | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 |
| Chiffre d'affaire (Y) | 0.7 | 1.6 | 2.0 | 2.4 | 2.5 | 2.8 | 3.0 | 3.0 |

On considère la série statistique double associée aux caractères X et Y ci-dessus.

Pour les tracés, on arrondira les valeurs numériques au centième près.

1. Dans un repère orthonormé, tracer le nuage de point associé.
2. (a) Calculer le coefficient de corrélation linéaire Y et X .
(b) Un ajustement linéaire est-il pertinent ?
3. (a) Calculer le coefficient de corrélation linéaire de X et $\exp(Y)$.
(b) En déduire un ajustement de Y en X de la forme :

$$Y = f(X)$$

où $f : x \mapsto \ln(ax + b)$ avec a et b des réels à déterminer.