

FEUILLE DE TRAVAUX DIRIGÉS N° 2

Décomposition de la variance - rapport de corrélation

Enseignant-Formateur : H. El-Otmany

A.U. : 2019-2020

Exercice n°1 Lors d'une étude menée sur 16 étudiants afin d'expliquer le score à un examen de mathématiques à partir du temps consacré à la préparation de cet examen. Pour chaque étudiant, on dispose :

- du temps de révision en heures (variable X),
- du score obtenu sur 800 points (variable Y).

X	5	11	9	10	5	15	8	2	7	1	13	5	6	4
Y	360	460	555	700	390	470	720	650	300	490	640	450	520	500

1. Représenter graphiquement le nuage des points.
2. Calculer et tracer la droite de régression linéaire sur le même graphique.
3. Calculer la Somme de Carrés des Écarts verticaux.
4. Calculer la variance $V(Y)$.
5. En notant \hat{Y} le vecteur des valeurs estimées par la droite de régression linéaire, calculer $V(\hat{Y})$.
6. On note $E = Y - \hat{Y}$, calculer $V(E)$. En déduire la somme des carrés des écarts verticaux.
7. Vérifier que $V(Y) = V(E) + V(\hat{Y})$.
8. Calculer le rapport de corrélation linéaire : $r_{Y/X}^2 = \frac{V(\hat{Y})}{V(Y)}$.
9. Quel est le pourcentage de variation expliquée par la régression linéaire ?
10. Calculer le coefficient de corrélation linéaire r . Interpréter le résultat.

Exercice n°2 On donne pour les six premiers mois de l'année 1982 les nombres d'offres d'emploi (concernant les emplois durables et à plein temps) et le nombre des demandes d'emploi (déposées par des personnes sans emploi, immédiatement disponibles et à la recherche d'un emploi durable et à plein temps). Les données sont exprimées en milliers d'individus.

Offres (X)	61	66.7	75.8	78.6	82.8	87.2
Demandes (Y)	2034	2003.8	1964.5	1928.2	1885.3	1867.1

1. Représenter le nuage de points. Le nuage de points vous semble-t-il aligné le long d'une droite ?
2. Trouver la droite de régression des demandes d'emploi en fonction des offres d'emploi et la tracer sur le graphique précédent.
3. Calculer le coefficient de corrélation entre X et Y . Commenter.
4. Calculer la Somme de Carrés des Écarts.
5. Reprendre les calculs en utilisant la calculatrice.

Exercice n°3 Dans un établissement universitaire, on souhaite étudier l'agrégation entre les notes d'admission d'étudiants à l'intégration d'une formation initiale et les notes obtenues durant cette formation. Les notes ci-dessous sont celles obtenues par 10 étudiants en mathématiques.

Notes d'admission Y	Notes de formation X
10	12
16	17
12	13
6	7
8	10
13	12
9	11
14	14
10	10
7	9

1. Représenter graphiquement le nuage des points.
2. Calculer et tracer la droite de régression linéaire sur le même graphique.
3. Est-ce que la moyenne (ou la somme) des écarts verticaux est nulle ?
4. Calculer la Somme de Carrés des Écarts verticaux.
5. Calculer la variance $V(Y)$.
6. En notant \hat{Y} le vecteur des valeurs estimées par la droite de régression linéaire, calculer $V(\hat{Y})$.
7. On note $E = V(Y) - V(\hat{Y})$, calculer $V(E)$. En déduire la somme des carrées des écarts verticaux.
8. Vérifier que $V(Y) = V(E) + V(\hat{Y})$.
9. Calculer le rapport de corrélation linéaire : $r_{Y/X}^2 = \frac{V(\hat{Y})}{V(Y)}$.
10. Quel est le pourcentage de variation expliquée par la régression linéaire ?
11. Calculer le coefficient de corrélation linéaire r . Interpréter le résultat.

Exercice n°4 On considère les mesures suivantes de deux variables quantitatives X et Y sur 10 individus équipondérés numérotés de 1 à 10.

X	2	2	4	4	4	6	6	6	8	8
Y	4	6	8	10	10	10	12	12	8	10

Le but de cette expérience est de déterminer une fonction permettant d'estimer Y à partir de X .

1. Représenter graphiquement le nuage des points.
2. Donner la droite de régression et tracer cette droite sur le nuage de points.
3. Calculer le coefficient de corrélation linéaire r_{XY} . Commenter le résultat.
4. En déduire le rapport de corrélation $r_{Y/X}^2$, ainsi la décomposition numérique de la variance de Y en variance expliquée par la régression linéaire et variance résiduelle.
5. En déduire que la somme des carrés des écarts verticaux.