

NOM :
PRENOM :

Date : 28 Novembre au 2 Décembre 2011
Groupe :

Mathématiques pour la Biologie : Feuille-réponses du TD 9
Régression linéaire

Exercice 1. : On reprend l'exemple des 5 spécimens fossiles d'un animal disparu pour lesquels on possède les mesures de la longueur de leur fémur et de leur humérus.

1. Compléter le tableau suivant et en déduire les valeurs des variances et covariance :

i	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
1	38	41
2	56	61
3	59	70
4	64	72
5	75	84
Moyenne :

$Var(x) =$	$Var(y) =$	$cov(x, y) =$
------------	------------	---------------

2. Déterminer, par la méthode des moindres carrés ordinaires, l'équation de la droite de régression de y sur x .

$y =$

3. Cette droite passe-t-elle par G? Expliquez pourquoi.

4. Calculer le coefficient de corrélation linéaire. Commenter.

$\rho(x, y) =$

5. Calculer la longueur prédite par ce modèle pour l'humérus d'un spécimen dont le fémur mesurerait 50cm.

Exercice 2. :

Pour étudier les problèmes de malnutrition dans un pays pauvre, on a calculé le poids moyen par âge d'un échantillon de 2400 enfants répartis uniformément en 12 classes d'âge. On a obtenu les données suivantes :

classe d'âge	1	2	3	4	5	6	7	8	9	10	11	12
poids moyen	4,3	5,1	5,7	6,3	6,8	7,1	7,2	7,2	7,2	7,2	7,5	7,8
\hat{y}_i
ε_i

1. Un statisticien pressé a fait calculer par sa machine la droite des moindres carrés pour ces données et a trouvé la relation poids = 4,88 + 0,267 âge. S'est-il trompé? Expliquer.

2. A votre avis, quelle est la pertinence de son modèle?

3. Compléter le tableau en calculant les poids prédits par le modèle \hat{y}_i .

4. Calculer puis tracer les résidus en fonction de la classe d'âge¹.

¹Vous pouvez constater que deux résidus successifs sont beaucoup plus souvent du même signe que du signe opposé. Ceci n'est pas compatible avec le fait qu'ils soient supposés indépendants. On dit que les résidus sont *autocorrélés*. C'est une raison de rejeter le modèle et d'en rechercher un meilleur.