

Statistiques à deux variables

Exercice 1

Dépense moyenne par foyer et par semaine en alcool et en tabac (en livres), dans 10 régions de Grande Bretagne (Source : Department of Employment, 1981) :

Region	Alcool	Tabac
North	6.47	4.03
Yorkshire	6.13	3.76
Northeast	6.19	3.77
East Midlands	4.89	3.34
West Midlands	5.63	3.47
East Anglia	4.52	2.92
Southeast	5.89	3.20
Southwest	4.79	2.71
Wales	5.27	3.53
Scotland	6.08	4.51

Si x est la dépense en alcool et y la dépense en tabac, on trouve

$$\bar{x}^2 = 0.41 \quad \bar{y}^2 = 0.25 \quad \bar{x}\bar{y} = 0.25$$

Mettre ces points sur un graphique. Calculer la corrélation. Qu'en dire ?

Pour l'Irlande du Nord, on trouve $x = 4.02$ et $y = 4.56$. Qu'en dire ?

Exercice 2

Pour vérifier les relations d'allométrie entre insectes, on a retenu les deux variables

x = logarithme de la longueur de l'élytre

y = logarithme de la largeur de la tête.

Les mesures sur 50 insectes, notées (x_i, y_i) ont fourni les résultats suivants :

$$\begin{aligned} \sum_{i=1}^{50} x_i &= 155 & \sum_{i=1}^{50} y_i &= 125 & \sum_{i=1}^{50} x_i y_i &= 391,1 \\ \sum_{i=1}^{50} x_i^2 &= 482,5 & \sum_{i=1}^{50} y_i^2 &= 320,5 & \sum_{i=1}^{50} x_i^2 y_i^2 &= 3468,7 \end{aligned}$$

- Calculer
 - La moyenne et l'écart-type du caractère x sur l'échantillon observé
 - La moyenne et l'écart-type du caractère y sur l'échantillon observé
 - La covariance empirique et la corrélation empirique des variables x et y
 - L'équation de la droite de régression de y sur x obtenue par estimation sur ces données.
- En déduire la loi d'allométrie exprimant la largeur de la tête en fonction de la longueur de l'élytre ($\log(0,046) = -3,08$).

Exercice 3

On se propose d'étudier l'influence de la température sur la durée d'incubation des oeufs de grenouilles. On choisit 6 échantillons de 200 oeufs chacun. Le nombre x d'éclosion au 22-ème jour est le suivant

température t_i d'incubation en degrés Celsius	6	6,4	6,8	7,2	7,6	8
nombre x_i d'éclosions à la température t_i	131	144	157	170	190	189

- Dessiner le nuage des données et tracer "à l'oeil" une droite D qui a l'air de bien approcher ce nuage.
- Calculer le coefficient de corrélation observé et écrire l'équation de la droite de régression de x en t . Etudier la qualité de l'ajustement.
- Calculer le nombre d'éclosions prédit pour un échantillon de 200 oeufs au 22-ème jour pour une température de 7,5 degrés.

Exercice 4

On étudie un échantillon de taille $n = 100$ sur lequel ont été mesurés deux caractères x et y , on a observé les résultats suivants :

$$\begin{aligned} \sum_{i=1}^n x_i &= 800 & \sum_{i=1}^n y_i &= 1200 & \sum_{i=1}^n y_i^2 &= 16000 \\ \sum_{i=1}^n x_i^2 &= 7200 & \sum_{i=1}^n x_i y_i &= 10200 & & \end{aligned}$$

- Calculer l'équation de la droite de régression linéaire de la variable y sur la variable x .
- Calculer l'équation de la droite de régression linéaire de la variable x sur la variable y .

Exercice 5

On a mesuré les variables x et y sur 10 individus et obtenu les résultats suivants :

individu $n^o i$	1	2	3	4	5	6	7	8	9	10
x_i	18	20	19	16	19	16	19	21	15	17
y_i	43	110	70	17	91	29	80	134	15	34

On cherche une relation logarithmique entre x et y du type : $y = a \ln x + b$. Pour cela on posera $z = \ln x$.

- Représenter graphiquement le nuage de points (z_i, y_i) et calculer la droite de régression linéaire de y en z .
- Quelle valeur de y peut-on prédire pour un individu présentant la valeur $x = 22$.

Exercice 6

Les données suivantes représentent la taille de 161 enfants du village Égyptien de Kalama suivis entre 18 et 29 mois, moyennées par tranche d'âge¹ :

âge (mois)	18	19	20	21	22	23	24	25	26	27	28	29
taille (cm)	76.1	77	78.1	78.2	78.8	79.7	79.9	81.1	81.2	81.8	82.8	83.5

Tracer les points, la droite de régression (au jugé...) et en déduire graphiquement l'équation de la droite de régression (on cherchera deux points visiblement proches de la vraie droite de régression ce qui permettra de calculer son équation, à une légère erreur près).

¹Source : D.S. Moore & G.P. McCabe (1989), *Introduction to the Practice of Statistics*, p. 118.