

## DEVOIR EN TEMPS LIBRE N° 8

Deux sujets sont proposés : le sujet 1 (exercices 1 et 2) et le sujet 2 (exercice 3). Un seul sujet sur les deux est traité. Les élèves n'ayant pas eu la moyenne au devoir sur le produit scalaire traiteront le sujet 1.

### EXERCICE I

Le plan est muni d'un repère orthonormé  $(O; \vec{i}, \vec{j})$ .

Cet exercice a pour objet d'établir une relation exprimant la distance entre un point et une droite en fonction du premier membre d'une équation cartésienne de la droite et des coordonnées du point.

1. Cette première question constitue une restitution organisée de compétence.

**Pré-requis :**

- deux vecteurs de coordonnées  $\begin{pmatrix} x \\ y \end{pmatrix}$  et  $\begin{pmatrix} x' \\ y' \end{pmatrix}$  ont pour produit scalaire :  $xx' + yy'$  ;
- deux vecteurs sont orthogonaux si, et seulement si, leur produit scalaire est nul ;
- un point, M, du plan est un point de la droite de vecteur normal,  $\vec{n}$ , passant par A, si, et seulement si :  $\overrightarrow{AM} \perp \vec{n}$ .

Soit D la droite de vecteur normal,  $\vec{n} \begin{pmatrix} a \\ b \end{pmatrix}$  (avec  $(a, b) \neq (0, 0)$ ), passant par A  $(x_0, y_0)$ .

Démontrer que D a une équation de la forme :  $ax + by + c = 0$  ; on exprimera c en fonction des coordonnées de  $\vec{n}$  et de A.

2. Soit M(x, y) un point du plan, D une droite de vecteur normal,  $\vec{n} \begin{pmatrix} a \\ b \end{pmatrix}$  (avec  $(a, b) \neq (0, 0)$ ), passant par A  $(x_0, y_0)$  et D' la droite passant par A et dirigée par  $\vec{n}$ . On désigne par H et H' les projetés orthogonaux respectifs du point M sur les droites D et D'.

a. Déterminer la nature du quadrilatère AHMH'.

b. On désigne par  $d(M, D)$  la distance<sup>1</sup> du point M à la droite D. Démontrer que :  $d(M, D) = MH$ .

c. Démontrer que :  $d(M, D) = \frac{|\overrightarrow{AM} \cdot \vec{n}|}{\|\vec{n}\|}$ .

d. Démontrer que :  $d(M, D) = \frac{|ax + by + c|}{\sqrt{a^2 + b^2}}$ .

3. **Application :** On considère les points A(1, 1), B(4, 7) et C(2, 6).

a. Déterminer une équation cartésienne de la droite (AB).

b. Calculer la distance entre la droite (AB) et le point C.

c. Déterminer l'aire du triangle ABC.

### EXERCICE II

Au XIX<sup>e</sup> siècle un explorateur découvre une nouvelle contrée et entreprend d'en dresser une carte et quelques relevés topographiques. Les figures proposées dans cet exercice ne sont pas à l'échelle. Les résultats des calculs de distances seront arrondis, s'il y a lieu, au décimètre.

1. Lors d'une étape, l'explorateur se trouve au bord d'une rivière infranchissable dont les berges sont parallèles. L'explorateur décide de mesurer la largeur du lit de la rivière. Il repère deux points A et B sur la berge qui est de son côté, et un troisième, C, sur la berge opposée. Il mesure la distance AB, puis en utilisant un théodolite (voir [2]) il mesure les angles  $\widehat{BAC}$  et  $\widehat{ABC}$  en utilisant le système sexagésimal (voir [1]). Il obtient

<sup>1</sup>La distance d'un point M à une droite D est la distance entre M et le point de D le plus proche de M.

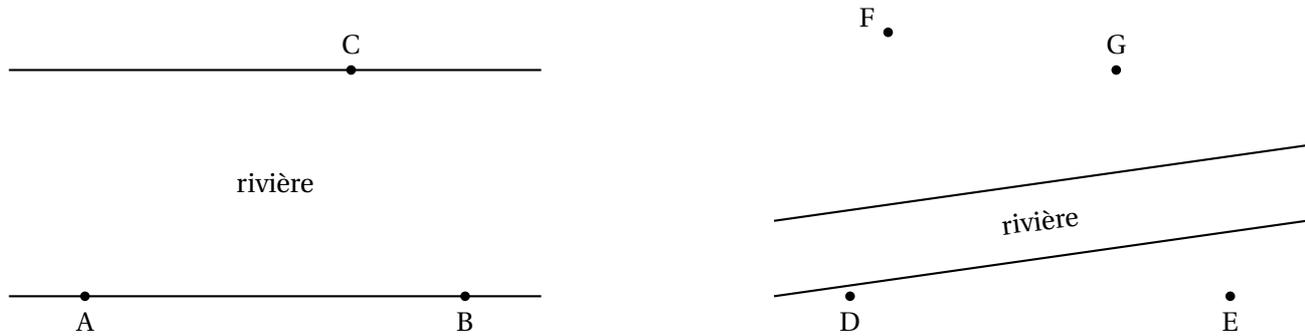


FIG. 1 – Problèmes de triangulation.

les résultats suivants :  $AB = 123 \text{ m}$  ;  $\widehat{BAC} = 31^\circ 53' 27''$  ;  $\widehat{ABC} = 59^\circ 29' 23''$ .

a. Exprimer la largeur du lit de la rivière en fonction des paramètres usuels du triangle  $ABC$  ( $a, b, c, \hat{A}, \hat{B}, \hat{C}$ ).

b. Déterminer, la largeur du lit de la rivière.

2. Lors d'une autre étape, alors qu'il est au point  $D$ , l'explorateur repère deux points inaccessibles,  $F$  et  $G$ , et un point accessible,  $E$ . On supposera que les points  $D, E, F$  et  $G$  sont dans un même plan horizontal. Il mesure les angles  $\widehat{FDE}$  et  $\widehat{GDE}$ , il obtient respectivement :  $78^\circ 11' 16''$  et  $45^\circ 40' 2''$ . L'explorateur se rend ensuite au point  $E$ . La distance  $DE$  mesure  $3450 \text{ m}$ .  $\widehat{FED}$  et  $\widehat{GED}$ , mesurent respectivement :  $14^\circ 15'$  et  $12^\circ 1' 5''$ .

a. Déterminer les distances  $DF$  et  $DG$ .

b. Déterminer la distance  $FG$ .

### EXERCICE III

Soit  $x$  et  $y$  deux séries statistiques issues de l'étude de deux caractères quantitatifs portant sur une même population de  $n$  individus ( $n \geq 3$ ). Pour tout élément,  $i$ , de  $[[1; n]]$ , on considère le point  $M_i$  de coordonnées  $(x_i, y_i)$ . On obtient ainsi un nuage de points.

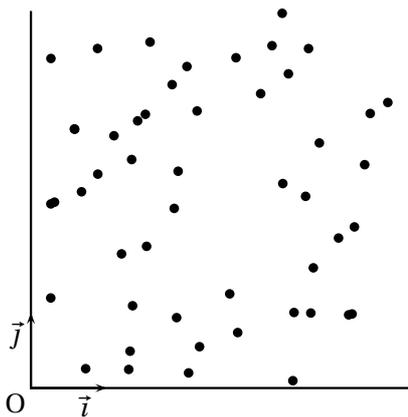


FIG. 2 –

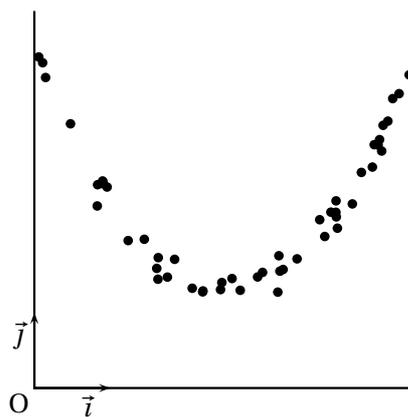


FIG. 3 –

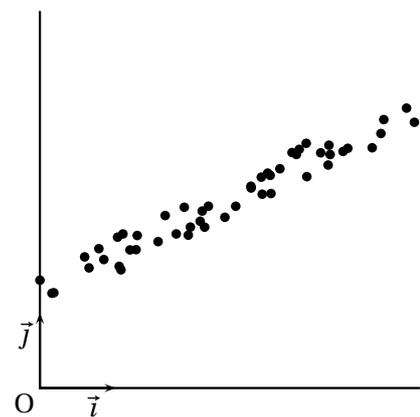


FIG. 4 –

Sur la figure 2, il est impossible de détecter de façon immédiate un lien entre  $x$  et  $y$ . Sur la figure 3, une tendance se dégage, les points du nuage semblent être disposés de part et d'autre d'une parabole invisible. Sur la figure 4, les points du nuage semblent être disposés de part et d'autre d'une droite invisible. C'est le dernier cas qui est l'objet de cet exercice. On se propose de déterminer le coefficient directeur et l'ordonnée à l'origine de la « droite invisible ». Cette droite est appelée droite de régression de  $y$  en  $x$ , la méthode de détermination de son équation réduite est appelée : « méthode des moindres carrés ».

Le plan est muni d'un repère orthonormé  $(O; \vec{i}, \vec{j})$ . On considère une droite  $(D)$  d'équation réduite :  $y = ax + b$ . À tout point  $M_i$  du nuage on associe son projeté,  $N_i$ , sur  $(D)$  parallèlement à l'axe des ordonnées. Pour trouver les valeurs de  $a$  et  $b$  pour lesquelles la droite  $(D)$  passe « au plus près » du nuage, on se propose

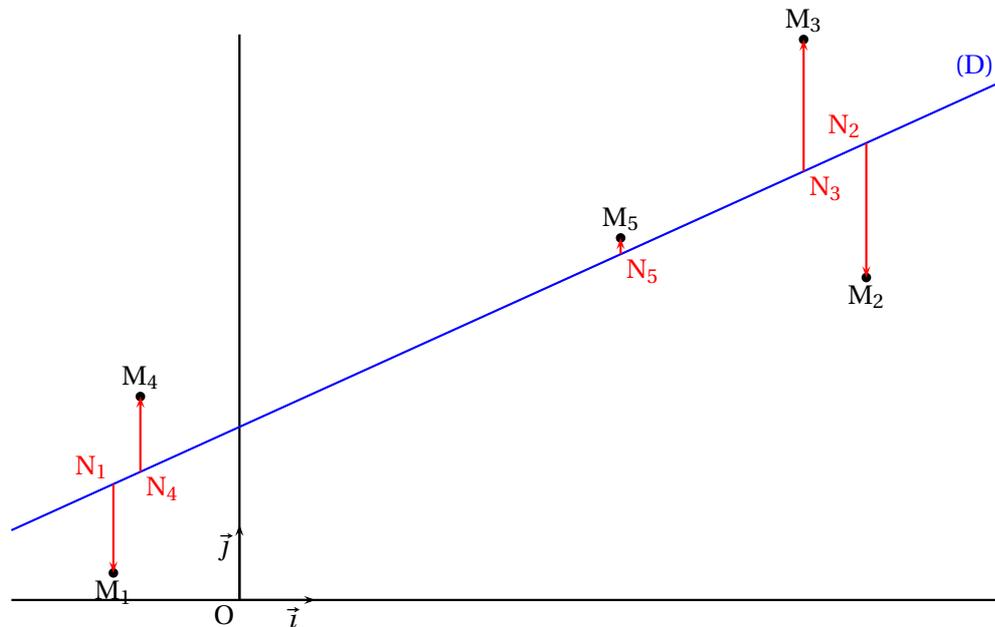


FIG. 5 – Droite de régression d'un nuage de points.

de déterminer le couple  $(a, b)$  qui réalise le minimum de la fonction  $S$  définie par :

$$S(a, b) = \sum_{i=1}^n N_i M_i^2.$$

Les caractéristiques de position et de dispersion utilisées sont les suivantes.

**moyenne de  $x$  :**  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$

**moyenne de  $y$  :**  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$

**variance de  $x$  :**  $s_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \overline{x^2} - \bar{x}^2$  (d'après la formule de König).

**variance de  $y$  :**  $s_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 = \overline{y^2} - \bar{y}^2.$

**covariance de  $(x, y)$  :**  $s_{x,y} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = s_{y,x}.$

**écart type de  $x$  :**  $s_x = \sqrt{s_x^2}.$

**écart type de  $y$  :**  $s_y = \sqrt{s_y^2}.$

Nous utiliserons également l'isobarycentre des points du nuage, appelé *point moyen* du nuage; nous le désignerons par  $G$  et il a pour coordonnées  $(\bar{x}, \bar{y})$ . Les points du nuage sont alignés verticalement ou horizontalement si, et seulement si, l'une des séries  $x$  ou  $y$  est constante. Afin d'éliminer ce cas peu intéressant, nous supposons qu'aucune de ces deux séries statistiques n'est constante.

### Partie A – Droite de régression de $y$ en $x$

1. Démontrer la formule de König concernant la covariance :  $s_{x,y} = \overline{xy} - \bar{x} \cdot \bar{y}.$
2. **a.** Déterminer une expression de  $S(a, b)$  en fonction des nombres  $a, b$  et des nombres  $x_i, y_i$  (avec  $i \in \llbracket 1; n \rrbracket$ ).
- b.** Déterminer une expression de  $S(a, b)$  ayant la forme d'un polynôme d'indéterminée  $b$  où  $a$  et les nombres  $x_i, y_i$  (avec  $i \in \llbracket 1; n \rrbracket$ ) apparaîtront comme des paramètres.
- c.** Déterminer la valeur,  $b_0$ , pour laquelle le minimum de la fonction  $b \mapsto S(a, b)$  est atteint (on exprimera

le nombre  $b_0$  en fonction des nombres  $a$ ,  $\bar{x}$  et  $\bar{y}$ ).

- d. En déduire une interprétation géométrique concernant le point moyen.
3. a. Déterminer une expression de  $S(a, b_0)$  ayant la forme d'un polynôme d'indéterminée  $a$  où l'effectif de la population, les variances et la covariance apparaîtront comme des paramètres.  
b. Déterminer la valeur,  $a_0$ , pour laquelle le minimum de la fonction  $a \mapsto S(a, b_0)$  est atteint (on exprimera le nombre  $a_0$  en fonction des variances et de la covariance des séries  $x$  et  $y$ ).
4. a. Justifier que la fonction  $(a, b) \mapsto S(a, b)$  présente un minimum en  $(a_0, b_0)$ .  
b. Conclure.
5. Dans cette question on suppose que l'on dispose d'un graphique sur lequel apparaissent le nuage de points et la droite de régression. Comment lire sur la figure le signe de  $s_{x,y}$  ?

### Partie B – Coefficient de corrélation

Le coefficient de corrélation de la série  $(x, y)$  est le nombre  $R$  défini par :

$$R = \frac{s_{x,y}}{s_x \cdot s_y}.$$

1. Justifier que ce coefficient est effectivement défini.
2. Exprimer le minimum de la fonction  $(a, b) \mapsto S(a, b)$  en fonction du coefficient de corrélation, de l'effectif de la population et de la variance de  $y$ .
3. Démontrer que :  $|R| \leq 1$ .
4. Démontrer que les points du nuage sont alignés si, et seulement si,  $|R| = 1$ .

### Partie C – Étude d'un cas

Dans un village, on se propose d'établir une éventuelle relation entre la taille (en cm) et la masse (en kg) des hommes adultes du village. Il arrive qu'un individu, sur conseil médical, désire ajusté sa masse à une masse idéale qui est fonction de sa taille. On cherchera donc à exprimer la masse en fonction de la taille. Une étude à été menée dont les résultats ont été regroupés dans le fichier `taillemasse.data` (voir [6]). Sur chaque ligne apparaît la taille puis la masse d'un individu. Deux lignes distinctes correspondent à deux individus distincts. Les réponses aux questions posées dans cette partie seront rendues dans un fichier WxMaxima intitulé `VotreNom_dm8.wxm` qui sera envoyer à [valere.bonnet@gmail.com](mailto:valere.bonnet@gmail.com). Toutefois il faudra également rendre, sur feuille et complété, le tableau 1 où les résultats seront arrondis, si besoin est, à trois chiffres après la virgule.

1. Entrer les données du fichier `taillemasse.data` dans une liste qui sera nommée `donnees`, chaque élément de cette liste sera une 2-liste (on pourra utiliser le module `numericalio`, voir [4] p 763-767 ou [5] p 12-13).
2. Extraire de la liste `donnees` les listes `taille` et `masse` (voir [4] p 445-450 ou [5] p 9-11).
3. Déterminer l'effectif de la population puis, en utilisant le module `descriptive`, déterminer les autres caractéristiques mentionnées dans le tableau 1 ( voir [4] p 539-558 ou [5] p 13-14), on ne demande pas la covariance.
4. Déterminer, sans utiliser le module `descriptive`, les caractéristiques mentionnées dans le tableau 1 (on pourra utiliser le produit scalaire, voir [4] p 285 ou [5] p 11).
5. Représenter sur une même figure le nuage de points et la droite de régression (on pourra utiliser `plot2d`, voir [4] p 102-110).
6. Tracer le diagramme en boîte de la taille (on utilisera : `boxplot`).
7. a. Calculer le coefficient de corrélation de la série (taille, masse).

$n$	
$\overline{\text{taille}}$	
$\overline{\text{masse}}$	
$s_{\text{taille}}$	
$s_{\text{masse}}$	
$s_{\text{taille,masse}}$	
min(taille)	
$D_1$ (taille)	
$Q_1$ (taille)	
med(taille)	
$Q_3$ (taille)	
$D_9$ (taille)	
max(taille)	
min(masse)	
max(masse)	

TAB. 1 – Caractéristiques de taillemasse.

b. On estime généralement que l'ajustement est valide dès que ce **coefficient** a une valeur absolue supérieure à :  $\frac{\sqrt{3}}{2}$ . L'ajustement obtenu est-il valide ?

8. Le fichier [6] a en fait été généré artificiellement en utilisant une formule. Retrouver le nom de cette formule (voir [7] et [8]).

## Références

- [1] Système sexagésimal - wikipedia  
[http://fr.wikipedia.org/wiki/Numération\\_sexagésimale](http://fr.wikipedia.org/wiki/Numération_sexagésimale)
- [2] Théodolite - Institut National de Recherche Pédagogique  
[http://www.inrp.fr/she/instruments/instr\\_mes\\_theodolite.htm](http://www.inrp.fr/she/instruments/instr_mes_theodolite.htm)
- [3] Régression linéaire - wikipedia  
[http://fr.wikipedia.org/wiki/Régression\\_linéaire](http://fr.wikipedia.org/wiki/Régression_linéaire)
- [4] Maxima Manual - Ver. 5.16.3  
<http://maxima.sourceforge.net/docs/manual/en/maxima.pdf>
- [5] Prise en main de Maxima  
<http://mathsaulyce.ee.info/memo/manmaxima.pdf>
- [6] taillemasse.data - fichier de données à exploiter  
<http://mathsaulyce.ee.info/memo/605/taillemasse.data>
- [7] La page « poids idéal » du site de DISTRIMED  
[http://www.distrimed.com/divers/page\\_poids.htm](http://www.distrimed.com/divers/page_poids.htm)
- [8] Un site traitant de diététique  
<http://www.poidsideal.info/>