UNIVERSITÉ PAUL SABATIER-IUT PONSAN — GEAP S2
NOM : ... — 2024-2025
PRENOM : ...

Éléments de correction du partiel du S2 (durée 1,5 h)
Mai 2025

<u>*Consignes :*</u> *La présentation, la clarté et la qualité de vos raisonnements tiendront une part importante dans la notation. Tout résulat non justifié ne sera pas pris en compte.*

<u>*Documents autorisés :*</u> *une feuille A4 de notes de cours, calculatrices.*

Exercice 1 Dans une situation de monopole sur la production d'un objet, une entreprise le conditionne et en fait la promotion. Une statistique a été établie pour étudier la liaison entre production et coût de publicité.

On notera q la quantité produite exprimée en centaines, y la part du coût de publicité en pourcentage. (On pourra garder 3 décimales significatives dans les calculs, sauf indication contraire en question 2a)

Quantité q_i (centaines)	1	10	20	50	100	150	200
Pourcentage y_i	4,20	3,70	3,30	2,30	1,20	0,65	0,35

Par exemple : Pour une production de 100 objets le coût de publicité est de $4,2\,\%$ du coût total.

1. Ci-joint en annexe le nuage de points (q_i ; y_i) qui sera rendu avec la copie.
   (a) Calculer le coefficient de corrélation $r(q, y)$

   On a $r(q,y) = \frac{cov(q,y)}{\sigma(q)\sigma(y)}$. On peut procéder par étape, en calculant des quantités intermédiaires :

   $$\left.\begin{array}{l} \bar{q} \approx 75.857 \\ \bar{y} \approx 2.243 \\ \sum y_i q_i \approx 509.7 \\ var(q) \approx 5031.551 \\ var(y) \approx 2.04 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} cov(q,y) \approx -97.322 \\ \sigma(q) \approx 70.933 \\ \sigma(y) \approx 1.428 \end{array}\right.$$

   Puis,

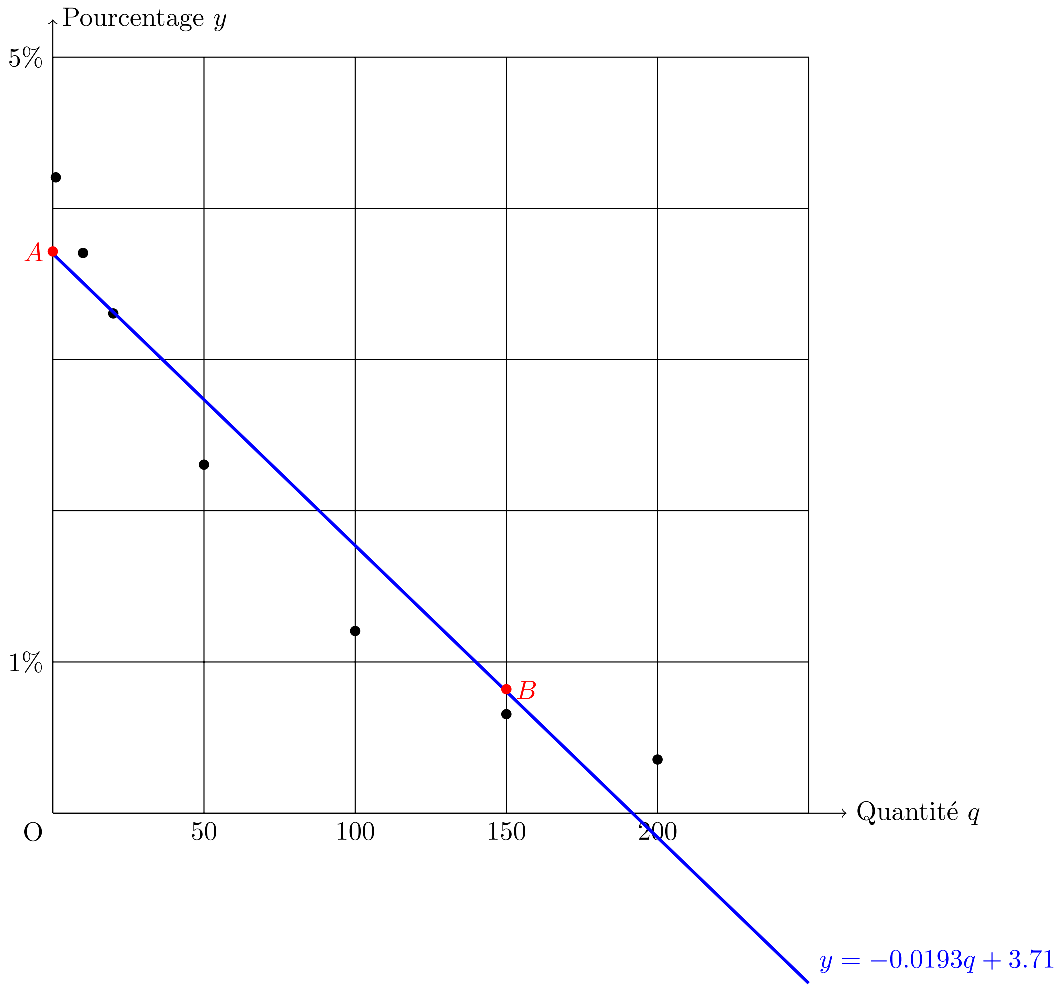
   $$\boxed{r(q,y) \approx -0.960}$$

   (b) Effectuer l'ajustement de y en fonction de q, puis tracer cette droite de régression sur la feuille donnée en annexe représentant le nuage de points. On expliquera le tracé en précisant par exemple 2 points utilisés.

   On cherche y sous la forme $y = aq + b$. Par les formules du cours et les quantités calculées à la question précédente, on obtient :
   $a = \frac{cov(q,y)}{var(q)} \approx -0.0193$ et $b = \bar{y} - a\bar{q} \approx 3.71$ L'équation de la droite de régression de y en q est :

$$\boxed{y = -0.0193q + 3.71}$$

Pour le tracé de la droite, on peut par exemple construire deux points A et B appartenant à la droite. Si $q = 0$, on a $y = 3.71$, donc le point $A(0; 3.71)$ est sur la droite d'ajustement. De même, on peut prendre $B(150; 0.815)$



(c) Quelle serait la part du coût de la publicité à prévoir pour une production de 25 000 objets ?

25 000 objets correspond à 250 centaines d'objets. Ainsi, on doit calculer $y(250)$. Numériquement, on obtient :

$$\boxed{y(250) = -1.125, \text{ résultat peu réaliste !}}$$

(d) Que pensez-vous de l'ajustement effectué dans cette question ?

La part du coût de publicité ne pouvant être négatif, on en déduit que :

$\boxed{\text{l'ajustement précédent n'est pas adapté à des quantités de production élevées.}}$

2. On considère un nouveau modèle en posant $z = \ln(100y)$.

   (a) Calculer le coefficient de corrélation $r(z, q)$. *Les valeurs de z_i seront données sous forme décimale arrondie au centième le plus proche.*

   On commence par fabriquer toutes les valeurs des z_i

Quantité q_i (centaines)	1	10	20	50	100	150	200
$z_i = \ln(100y_i)$	6,04	5,91	5,80	5,44	4,79	4,17	3,56

Pourcentage y_i	4,20	3,70	3,30	2,30	1,20	0,65	0,35

   On a $r(q, y) = \frac{cov(q,z)}{\sigma(q)\sigma(z)}$. On peut procéder par étape comme à la question 1a, en calculant des quantités intermédiaires manquantes :

   $$\left.\begin{array}{l} \bar{q} \approx 75.857 \\ \bar{z} \approx 5.101 \\ \sum z_i q_i \approx 2269.64 \\ var(q) \approx 5031.551 \\ var(y) \approx 0.783 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} cov(q,z) \approx -62.746 \\ \sigma(q) \approx 70.933 \\ \sigma(z) \approx 0.885 \end{array}\right.$$

   Puis, en arrondissant :

   $$\boxed{r(q, z) \approx 1 \text{ (valeur précise : } -0.99994)}$$

   (b) Déterminer une équation de la droite de régression de z en q de la forme $z = aq + b$ par la méthode des moindres carrés.

   On cherche z sous la forme $z = mq + p$. Par les formules du cours et les quantités calculées à la question précédente, on obtient :
   $m = \frac{cov(q,z)}{var(q)} \approx -0.012$ et $p = \bar{z} - m\bar{q} \approx 6.047$ L'équation de la droite de régression de z en q est :

   $$\boxed{z = -0.012q + 6.047}$$

   (c) Quelle serait la part du coût de la publicité à prévoir pour une production de 25 000 objets ?

   La relation précédente peut se re écrire $y = \frac{1}{100}e^{-0.012q+6.047}$ Soit, $y \approx 4.23e^{-0.012q}$. D'où, avec ce modèle :

   $$\boxed{y(250) \approx 0.187,}$$

   valeur plus réaliste qu'en question 1c.

Exercice 2 Questions courtes et indépendantes

1. Une entreprise en pleine expansion voit son chiffre d'affaires (CA) augmenté chaque année. Un étudiant de gea effectue son stage dans cette organisation et souhaite étudier la dépendance entre le CA et le temps. A

cet effet, il calcule le coefficient de corrélation et il trouve $-0,96$. Qu'en pensez vous ?

D'après l'énoncé, le CA est une fonction croissante du temps. Le coeff de corrélation ne peut donc être négatif.

L'étudiant s'est trompé !

2. Dans une chaîne de fabrication d'un article, deux paramètres u et z associés à la production, sont liés par une relation du type $u \approx \ln(\lambda z^3 + \beta)$. Quelles variables doit on considérer pour trouver les valeurs de α et β par un ajustement linéaire ?

   La relation qui lie u et z, se re écrit ainsi : $e^u \approx \lambda z^3 + \beta$. En posant

   $$\boxed{y = e^u \text{ et } x = z^3,}$$

   on a alors $y \approx \lambda x + \beta$, qui permet donc de se ramener à un cas linéaire.

3. Si on multiplie toutes les valeurs d'une série X par 10 et toutes les valeurs d'une série Y par 3, comment évolue le coefficient de corrélation r' de ces deux nouvelles séries ?

   Notons $X' = 10X$, $Y' = 3Y$ et n le nombre d'observations. On doit exprimer $r(X', Y') = \frac{cov(X',Y')}{\sigma(X')\sigma(Y')}$ en fonction de $r(X,Y) = \frac{cov(X,Y)}{\sigma(X)\sigma(Y)}$.

   Pour les écart-types, on a (cours S1) :

   $$\begin{cases} \sigma(X') & = \sigma(10X) = 10\sigma(X) \\ \sigma(Y') & = 3\sigma(Y). \end{cases}$$

   Pour les covariances, on a :

   $$\begin{aligned} cov(X', Y') &= \frac{1}{n}\sum_i x'_i y'_i - \bar{X}'\bar{Y}', \\ &= \frac{1}{n}\sum_i 10x_i \times 3y_i - 10\bar{X} \times 3\bar{Y}, \\ &= 30 \times cov(X, Y). \end{aligned}$$

   Ainsi,

   $r(X', Y') = \frac{cov(X',Y')}{\sigma(X')\sigma(Y')} = \frac{30 cov(X,Y)}{10\sigma(X) \times 3\sigma(Y)} = r(X, Y)$.

   Le coefficient de corrélation est inchangé !

4. Un négociant en vins a fait mener une étude visant à déterminer à quel prix maximal ses clients sont prêts à acheter une bouteille de vin. Les résultats sont regroupés dans le tableau suivant :

Prix maximal x_i en euros de la bouteille	5	10	15	20	25	30
Pourcentage y_i d'acheteurs potentiels	84	58	30	19	7	4

   Y a-t-il selon vous une dépendance entre ces quantités ? Justifiez votre réponse.

   On calcule le $r(x, y)$ et on trouve environ -0.959. Ce qui signifie qu'il y a

$\boxed{\text{une dépendance entre ces grandeurs.}}$

Calcul intermédiaires :

$$\begin{cases} \bar{x} = 17.5 \; et \; \bar{y} \approx 33.67 \\ var(x) \approx 72.92 \; et \; var(y) \approx 824.22 \\ \sum_i x_i y_i = 2125 \; et \; cov(x,y) = -235. \end{cases}$$

5. Soit la série $(u_i)_i$ définie par

Temps t	1	2	3	4	5	6
u_t	15	20	25	34	28	22

Calculer la moyenne mobile $M_2^{(2)}$ au temps $t = 2$ d'ordre 2, puis celle au temps $t = 4$ d'ordre 3, soit $M_4^{(3)}$.

On a

$$\begin{aligned} M_2^{(2)} &= \frac{1}{2}(\frac{u_1}{2} + u_2 + \frac{u_3}{2}), \\ &= \frac{\frac{15}{2} + 20 + \frac{25}{2}}{2}, \\ &= 20. \end{aligned}$$

Et de même, on a :

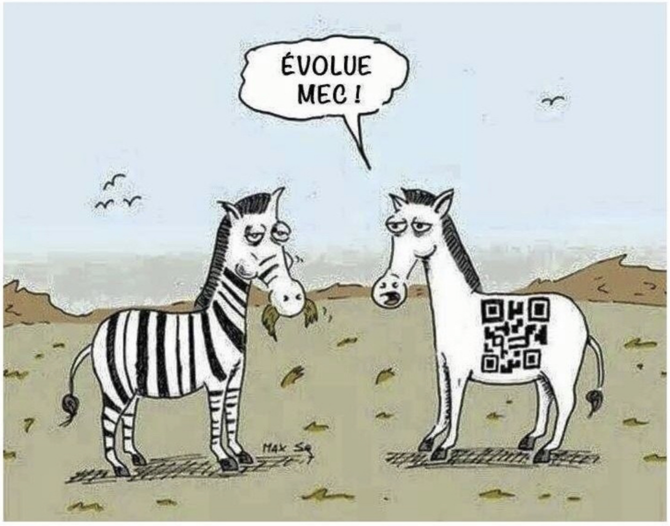
$$\begin{aligned} M_4^{(3)} &= \frac{1}{3}(u_3 + u_4 + u_5), \\ &= \frac{25 + 34 + 28}{3}, \\ &= 29. \end{aligned}$$

$$\boxed{M_2^{(2)} = 20 \; et \; M_4^{(3)} = 29.}$$

6. Quelles sont les 2 hypothèses sur la composante saisonnière dans les deux modèles des séries chronologiques ?

Les 2 hypothèses sont :

- la périodicité,
-et d'influence nulle sur une période.


ÉVOLUE MEC !