RESUME ET EXERCICES

Fomesoutra.com

ça soutra !

STATISTIQUES

SUJETS ET RESUME DE COURS

MATHS

TLE D

BY TEHUA

2025

FICHE DE COURS

Nuage de points

Le plan est muni d'un repère orthogonal.

L'ensemble des points M_{ij} de coordonnées (x_i ; y_j) est appelé nuage de points associé à la série.

Point moyen d'un nuage de points

On appelle point moyen d'un nuage de points représentant une série,

le point G de cordonnées (X_G ; Y_G) où $X_G=\bar{X}=\frac{\sum x_i}{N}$ et $Y_G=\bar{Y}=\frac{\sum y_j}{N}$.

Covariance

La covariance d'une série statistique à 2 caractères X et Y, de moyennes respectives $\bar{X}$ et $\bar{Y}$ et d'effectif total N est le nombre réel noté COV(X, Y) ou σ_{XY} tel que :

$$COV(X,Y)=\frac{\sum n_{ij}\ \ x_i-\bar{X}\ \ \ y_j-\bar{Y}}{Effectif\ total} \quad \text{ou} \quad COV(X,Y)=\frac{\sum n_{ij}x_i y_j}{Effectif\ total}-\bar{X}\bullet\bar{Y}$$

Droite de régression de Y en X

La droite (D) de régression de Y en fonction de X passe par le point $G\left(\bar{X},\bar{Y}\right)$, point moyen du nuage de point.

Son équation est : y = ax + b. $Avec\ \ a=\frac{COV(X,Y)}{V(X)}\ \ et\ \ b=\bar{Y}-a\bar{X}$.

Droite de régression de X en Y

La droite (D') de régression de X en fonction de Y passe par le point $G\left(\bar{X},\bar{Y}\right)$, point moyen du nuage de point.

Son équation est : $x=a'y+b'$. $Avec\ \ a'=\frac{COV(X,Y)}{V(Y)}\ \ et\ \ b'=\bar{X}-a\bar{Y}$.

Remarque $G\left(\bar{X},\bar{Y}\right)$ appartient aux 2 droites de régression.

Pour tracer ces droites, il suffit donc de déterminer un autre point vérifiant l'équation.

Corrélation linéaire

Deux variables statistiques X et Y sont dites en corrélation linéaire lorsque la courbe de régression de Y en X et la courbe de régression de X en Y sont des droites.

Coefficient de corrélation linéaire

On appelle **coefficient de corrélation linéaire** d'une série statistique **double de** caractère (X, Y), le nombre réel **r** défini par : $r = \dfrac{COV(X,Y)}{\sqrt{V(X) \bullet V(Y)}}$

Propriétés

* Si $0,87 \leq r \leq 1$ alors la **corrélation** linéaire entre les 2 variables est **forte.**
* Si $r = 1$ alors la **corrélation** est **parfaite.**

Estimation, Prévision

Lorsqu'il existe une bonne corrélation linéaire, il est possible de prévoir la valeur de y connaissant celle de x, en utilisant l'équation de la droite de régression de y en x ; et vice-versa.

METHODES PRATIQUES

M1 Comment calcule- t-on la variance de la variable X ?

Pour calculer la variance de la variable X, on peut procéder comme suit :

- Calculer $\overline{X}$ la moyenne de la variable X, puis calculer $\overline{X}^2$ le carré de la moyenne (1).
- Déterminer le carré de chaque modalité.
- Multiplier colonne par colonne le carré de chaque modalité par son effectif correspondant.
- Calculer la somme des produits obtenus ci-dessus.
- Diviser cette somme par l'effectif total : on obtient ainsi la moyenne des carrées (2)
- Soustraire de la moyenne des carrés (2) le carré de la moyenne (1): on obtient la variance.

NB : On procèdera de même pour calculer la variance de la variable Y.

M2 Comment calcule-t-on la covariance des variables X et Y ?

Pour calculer la covariance des variables X et Y, on peut procéder comme suit :

- Calculer $\overline{X}$ la moyenne de la variable X.
- Calculer $\overline{Y}$ la moyenne de la variable Y.
- Calculer le produit $\overline{X} \bullet \overline{Y}$ (1)
- Déterminer le carré de chaque modalité.
- Multiplier colonne par colonne chaque variable X à la variable Y correspondante.
- Calculer la somme des produits obtenus ci-dessus.
- Diviser cette somme par l'effectif total. (2)
- Calculer la différence (2) – (1) : on obtient la covariance des variables X et Y.

EXERCICE 1 BAC 2009 Session normale

L'entreprise Ivoirbois, spécialisée dans l'industrie du bois, envisage de faire des prévisions pour l'année 2007 du coût de production de feuilles de contre-plaqués en fonction du chiffre d'affaires.

Elle dispose à cet effet des statistiques résumées dans le tableau ci-dessous :

Années	2000	2001	2002	2003	2004	2005	2006
Chiffre d'affaires X (en millions de francs)	350	380	500	450	580	650	700
Coût de production Y (en millions de francs)	40	45	50	55	60	65	70

1 Représenter graphiquement le nuage de points associé à la série double (X, Y) dans le plan rapporté à un repère orthogonal (O, I, J).

On prendra 1 cm pour 50 millions de francs en abscisse et 1 cm pour 5 millions de francs en ordonnées.

2 a Calcule le chiffre d'affaires moyen $\overline{X}$.

b Calcule le coût moyen de production $\overline{Y}$.

3 a Vérifie qu'un arrondi à l'entier de la covariance cov (X, Y) de la série statistique est égal à **1193.**

b Justifie l'existence d'un ajustement linéaire entre X et Y.

4 a Détermine une équation de la droite (D) d'ajustement de Y en fonction de X par la méthode des moindres carrés.

b Construis (D) dans le repère (O, I, J).

5 Utilise l'ajustement précédent pour prévoir le coût de production de l'entreprise Ivoirbois de l'année 2007 si le chiffre d'affaires de l'année 2007 est de 800 millions de francs.

EXERCICE 2 BAC 2008 Session normale

Le tableau ci-dessous donne les notes sur 20 obtenues en mathématiques et en sciences physiques par huit candidats de la série D au baccalauréat 2005.

X_i est la note de mathématiques, Y_i la note en sciences physiques.

X_i	4	6	7	9	11	14	12	17
Y_i	3	4	6	8	10	12	9	14

1 Représente graphiquement le nuage de points associé à cette série statistique dans le plan muni d'un repère orthonormé (O, I, J). L'unité graphique est 1 cm.

2 Calcule les coordonnées du point moyen G du nuage puis place-le dans le repère.

3 a Vérifie que la covariance cov (X, Y) de la série statistique est égale à $\frac{57}{4}$

b Calcule le coefficient de corrélation linéaire entre X et Y.

4 Démontre qu'une équation de la droite (D) de régression de Y en fonction de X par la méthode des moindres carrés est : $Y = \frac{19}{22}X - \frac{17}{44}$

5 Sur la base de l'ajustement linéaire ainsi réalisé, calcule la note probable de mathématiques d'un candidat qui a obtenu 15 sur 20 en sciences physiques.

EXERCICE 3 Bac D 2000. Session de remplacement

Pour préparer la retraite de ses membres, une coopération a planté en 1991 des anacardiers qui sont rentrés en production trois ans plus tard. Le tableau statistique suivant donne l'évolution des productions depuis la première année de récolte.

Ordre X_i de l'année de production	1	2	3	4	5	6	7
Année de production			1996				
Quantité Y_i de production (en tonnes)	118	146	184	247	267	278	255

1. En quelle année cette coopérative a-t-elle obtenu 278 tonnes d'anacarde ?
2. Recopie et compléter le tableau statistique ci-dessus.
3. Représente dans le plan muni du repère orthogonal (O, I, J), le nuage de points de la série statistique double (X_i, Y_i).
On prendra : En abscisses 2 cm pour unité et en ordonnées 1 cm pour 20 tonnes.
4. Détermine l'arrondi d'ordre 2 du coefficient de corrélation linéaire de la distribution statistique (X_i, Y_i).
5. La corrélation entre les variables X et Y est-elle bonne ? Justifie.
6. Détermine par la méthode des moindres carrés, une équation de la droite de régression qui permet d'estimer l'année en fonction de la production.
7. En quelle année la coopérative produira-t-elle 350 tonnes ?

EXERCICE 4 Sénégal bac D 1999

L'étude du commerce extérieur d'un pays de 1990 à 1996 pour les importations et les exportations exprimés en milliards de francs donne le tableau suivant :

Importation X	2,8	3,2	3,8	4,4	6,4	5,7	7,4
Exportation Y	2	2,6	3,2	3,8	5	5,5	6,5

1. Calcule :
a. les moyennes $\overline{X}$ et $\overline{Y}$.
b. les variances V(X) et V(Y)
c. les écarts types σ(X) et σ(Y)
2. Calcule le coefficient de corrélation entre X et Y.
Existe t-il une corrélation entre les importations et les exportations.

EXERCICE 5 Sénégal bac D 2000

On donne la série statistique suivante à deux variables :

Xi	1,2	1,4	1,6	1,8	2
Yi	13	12	14	16	t

Par la méthode des moindres carrées, on a obtenu l'équation de la droite de régression de y en x, à savoir : **y = 9x + 0,6**

1. Calcule $\overline{X}$
2. Exprime $\overline{Y}$ en fonction de a
3. En utilisant 1. et 2., montre que t = 20.