Utilisation d'une calculatrice

Statistiques à deux variables et régression.

Dans le cadre des programmes qui demandent de connaître l'utilisation des calculatrices, je propose dans ce TP de montrer aux élèves ou aux étudiants de Bts d'apprendre à trouver les valeurs demandées, par simple lecture.

Énoncé

La pollution par le dioxyde de souffre et le dioxyde d'azote en région PACA a été donnée en 1998 par le tableau suivant.

Année t	***1990***	***1991***	***1992***	***1993***	***1994***	***1995***	***1996***	***1997***
dioxyde de souffre S	24	23	20	21	18	19	17	17
dioxyde d'azote N	50	49	50	48	48	47	42	39

Première partie : Évolution comparative de S et de N

1.Introduire les trois listes de valeurs dans votre calculatrice. Réaliser le nuage de points (S ;N)

2.Donner la valeur décimale arrondie à 10^{-3} près du coefficient de corrélation affine de la série double (S, N)

3.En acceptant 0,75 comme limite inférieure pour accepter un modèle affine, peut-on envisager un ajustement linéaire de S en N ?

4.Donner une équation de la droite de régression de N en S par la méthode des moindres carrés. (précision 10^{-3}). Recopier la valeur de $r^{2.}$ avec toutes les décimales affichées et mémorisez le coefficient directeur. (par exemple dans la mémoire A)

5.Faire de même pour la droite de régression de S en N. (mémorisez cette fois-ci dans la mémoire B)

6.Que constatez-vous en effectuant le produit des coefficients directeurs des droites trouvés aux deux question précédentes ?

Seconde partie : Évolution et prévision pour l'an 2000

Rappel de formules :

Covariance de x et y $\sigma_{xy} = \frac{1}{n}\sum_{1}^{n} x_i y_i - \overline{x}\,\overline{y}$

Coefficient de corrélation $r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$

1.Pour la série double (t, S) lire sur la calculatrice les valeurs nécessaires au calcul de *r* et effectuer ce calcul. Donner le coefficient de corrélation à 10^{-3} près.

2.Sans calcul, donner le coefficient de corrélation de la série double (t, N)

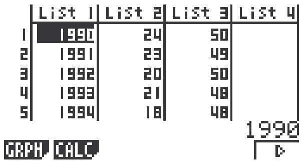
3.Laquelle de ces deux séries, justifie le plus un ajustement affine ? En déterminer une équation de la droite de régression en t par la méthode des moindres carrés.

4.Donner une estimation arrondie à l'unité de la valeur de la variable déterminée à la question précédente en l'an 2000.

L'objectif est de répondre à toutes les questions, sans introduire plusieurs fois les valeurs dans les listes de la calculatrice.
Le document est réalisé pour une calculatrice du type Casio Graph 35 ou 65, mais est adaptable sans difficulté à d'autres modèles (de marque TI également)

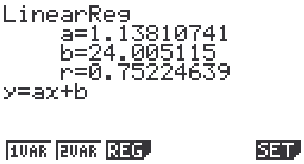
Première partie

Introduction des données.



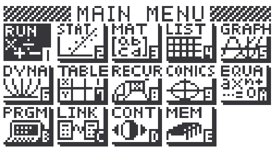
Pour la suite, nous décidons de placer les trois séries numériques dans les listes 1, 2 et 3 de la calculatrice.
Il est possible de réaliser le nuage de points sur la calculatrice, mais le tracé sur papier peut permettre d'autres utilisations.

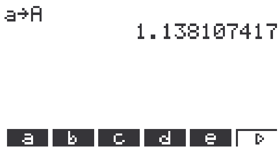
Mise en place des calculs pour la série (S ;N)



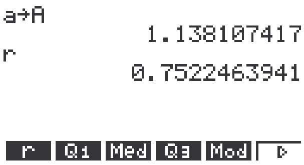
On obtient le coefficient de corrélation et le coefficient directeur
Au seuil indiqué l'ajustement linéaire est acceptable.
L'équation de la droite de régression de N en S est $y = 1{,}138x + 24{,}005$

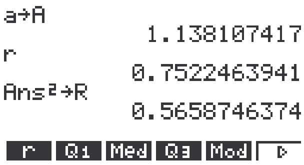
Mise en mémoire de a





La recherche de la valeur de a avec toutes ses décimales affichées se fait dans le mode RUN, par la touche VARS – STAT(F3) – GRPH(F3)

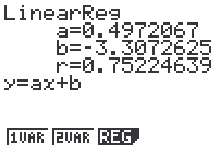


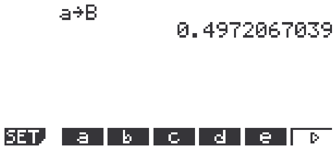


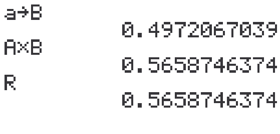
A l'aide de la flèche à droite (F6) on accède à r, on peut calculer son carré et mémoriser cette valeur dans R par exemple.

Mise en place des calculs pour la série (N ;S)

On procède de la même manière, que plus haut. Pour comparer le produit des coefficients directeurs au carré du coefficient de corrélation, on peut afficher les deux résultats sur le même écran.

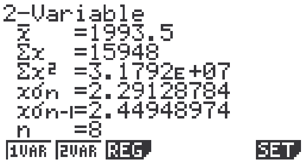






Seconde partie

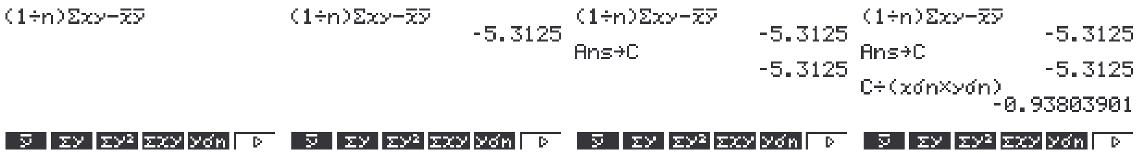
Objectif : Utiliser les mémoires de la calculatrice pour effectuer les opérations complètes à l'aide des formules rappelées. Les valeurs des sommes, des moyennes, des écart-types seront utilisées par l'appel des variables statistiques de la machine.



Etude de la série (t ;S)
On aura paramétré la calculatrice sur liste1, liste2 et effectué les calculs statistiques à deux variables

Calcul de la covariance

En mode RUN, on rappelle les variables statistiques. Dans le menu X (F1) on trouvera les valeurs concernant la première variable, et dans le menu Y (F2) celles concernant la seconde variable, et la somme des produits $x_i y_i$

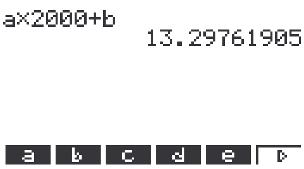


L'utilisation de la mémoire C au troisième écran n'est pas impératif. Mais il s'agit bien de la covariance des séries t et S. On pourra comparer le coefficient de corrélation r avec le résultat que donne la calculatrice en mode STAT – REG.

Comparaison des séries

Il suffira de paramétrer sur les listes 1 et 3 pour observer le coefficient de corrélation dont la valeur absolue est la plus proche de 1

Estimation pour l'an 2000



Après avoir affiché les valeurs obtenues par STAT – REG, les valeurs de a et de b de l'équation de la droite de régression sont en mémoire. VARS – STAT - GRPH
Pour faire les calculs de 2000a + b, on peut réutiliser les valeurs des variables

L'estimation sera donc de 13. Mais un ajustement affine sur une période trop longue semble injustifiable, puisque les valeurs peuvent tendre vers zéro, mais ne peuvent pas être négatives.