Méthode 1

Représenter une série statistique double

Énoncé

Chaque année une association organise une tombola. Elle achète un certain nombre de lots puis vend des billets de tombola par carnets de 10. Le nombre de lots achetés et le nombre de carnets vendus lors des 9 dernières tombolas est donné dans le tableau ci-contre

Nombre de lots x_i	38	26	20	13	8	34	32	41
Nombre de carnets y_i	69	64	60	59	55	67	68	70

1. Représenter cette série sous la forme d'un nuage de points dans un repère orthonormé.
2. Calculer les coordonnées du point moyen de ce nuage et le placer dans le repère.
3. Quelle forme présente le nuage de points ?
4. Tracer « au jugé » la courbe associée.

Conseils & Méthodes

1 Étant donné les valeurs de la série y, il est plus pratique de commencer à graduer l'axe des ordonnées à partir de 50, afin de garder une échelle (1 carreau pour 2 unités) suffisamment précise.

2 Pour tracer la droite, on place sa règle sur le point moyen et on « ajuste » au mieux la « pente » de façon à passer, en moyenne, le plus proche possible de tous les points.

Solution

Méthode 2 Obtenir la droite de régression et le coefficient de corrélation linéaire à la calculatrice

Énoncé

x_i	38	26	20	13	8	34	32	41
y_i	69	64	60	59	55	67	68	70

On considère le nuage de points associé à la série statistique de la Méthode 1 **p.225 À l'aide de la calculatrice, déterminer l'équation de la droite d'ajustement du nuage par la méthode des moindres carrés et le coefficient de corrélation linéaire.**
(Les valeurs seront arrondies à 10^{-3}).

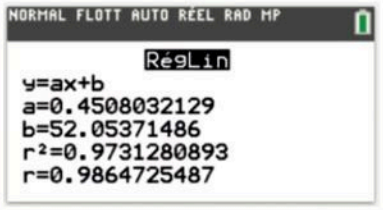
Conseils & Méthodes

Sur la calculatrice on peut aussi représenter le nuage de points. Cela permet de s'assurer qu'un ajustement affine a du sens.

Solution

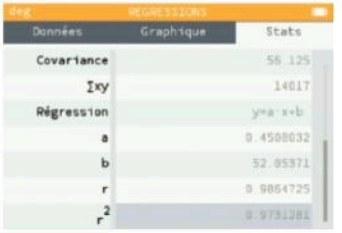
TI-83 Premium CE

Appuyer sur la touche stats de la calculatrice Sélectionner le menu **ÉDIT 1:Modifier** puis entrer les valeurs de la série statistique dans les colonnes L1 et L2.
Appuyer de nouveau sur la touche stats et sélectionner le menu **CALC 4 :RégLin(ax+b)**.
S'assurer qu'on a **Xliste : L1** et **Yliste : L2.** On obtient alors l'équation de la droite de régression de y en x, ainsi que le coefficient de corrélation linéaire r.



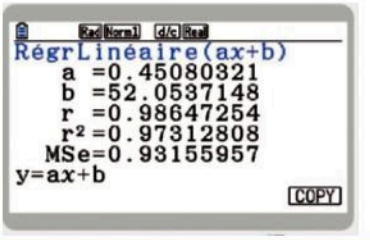
NUMWORKS

Appuyer sur la touche de la calculatrice. Sélectionner le menu Régressions puis entrer les valeurs de la série statistique dans les colonnes X1 et Y1. Appuyer sur la touche ↩. Puis sélectionner le menu **Stats** en vous déplaçant avec la flèche ▷. On obtient alors toutes les caractéristiques de la série statistique, et notamment, en se déplaçant vers le bas ▽, l'équation de la droite de régression, ainsi que le coefficient de corrélation linéaire r.



CASIO GRAPH 90+E

Appuyez sur la touche MENU de la calculatrice Sélectionner le menu Statistique puis EXE entrer les valeurs de la série statistique dans les colonnes.
List1 et **List2** et appuyer sur F2
Vérifier les réglages SET : **2varXliste : list1 et 2var Yliste : list2.**
Appuyer sur la touche REC puis sur F1 X puis sur F1 de nouveau : on obtient alors l'équation de la droite de régression de y en x, ainsi que le coefficient de corrélation linéaire r .



Méthode 3 Effectuer un ajustement exponentiel

Énoncé

On a relevé, pendant un an, sur différents parcours de même longueur, la vitesse moyenne x_i des véhicules et le nombre d'accidents mortels y_i au total sur l'année, pour i entier variant de 1 à 5.

Vitesse x_i (en km·h^{-1})	30	50	80	90	100
Nombre d'accidents y_i	1	6	41	66	102

1. Représenter ces données dans un repère orthogonal d'unités bien choisies.

2. On pose $y'_i = \ln(y_i)$ pour tout entier i de 1 à 5. Calculer les valeurs de y'_i arrondies au dixième.

3. Représenter le nuage de points $(x_i ; y'_i)$ dans un autre repère et vérifier que sa forme peut être ajustée par une droite.

4. Déterminer l'équation de la droite de régression de y' en x. (les coefficients seront arrondis à 10^{-3} près).

5. En déduire une expression de y en fonction de x.

Solution

Conseils & Méthodes

1 La forme du nuage obtenu invite à chercher un ajustement exponentiel.

2 Les valeurs y'_i étant beaucoup plus petites que les valeurs y_i, on sera la plupart du temps obligé de changer l'échelle de l'axe des ordonnées et donc d'utiliser un autre repère orthogonal.

3 On pourra éventuellement utiliser les propriétés de la fonction exponentielle pour écrire : $y = e^{0,066x} \times e^{-1,745}$. Puis, avec $e^{1,745} \approx 0,175$, on a : $y = 0,175e^{0,066x}$

Méthode 4

Effectuer des interpolations et des extrapolations

Énoncé

Pour l'achat d'une grosse quantité de ballons de handball, un fabricant propose un tarif dégressif selon la quantité d'articles commandés. Le tableau ci-contre présente un relevé des prix proposés.

Nombre de ballons x_i	100	500	1000	2000
Prix unitaire y_i (en €)	19,90	19	17,90	15,50

1. **Déterminer l'équation de la droite de régression de y en x et le coefficient de corrélation r.**
2. **D'après le coefficient r, l'ajustement affine de cette série statistique semble-t-il un bon choix ?**
3. **Déterminer le prix unitaire que devrait proposer le fabricant pour un achat de 1500 ballons.**
4. **Quelle quantité de ballons faudrait-il acheter pour obtenir un prix unitaire de 12 € ?**

Solution

Conseils & Méthodes

1 Plus r est proche de -1, plus sa valeur absolue est proche de 1 : c'est l'indicateur d'une très bonne corrélation linéaire entre les deux variables étudiées.

2 Les résultats d'interpolation et d'extrapolation ne sont que des valeurs indicatives qui découlent de beaucoup d'approximations et d'ajustements successifs, donc ils doivent toujours être donnés avec beaucoup de précautions (utiliser le subjonctif et de termes comme « environ », « probablement »…)

Méthode 5 Calculer par la méthode des moindres carrés

↪ Cours 2 p. 226

Énoncé

Le tableau ci-contre représente la masse x (en kg) et la taille y (en cm) d'un enfant, relevés par son médecin traitant lors des 6 dernières visites médicales.

Masse x_i (en kg)	4	5,4	10,2	11	12,6	19,8
Taille y_i (en cm)	53	61	72	78	94	113

1. Calculer la covariance de x et y puis la variance de x.
2. En déduire l'équation de la droite de régression de y en x.
3. Représenter le nuage de points $(x_i ; y_i)$ dans un repère puis tracer la droite de régression.

Solution

Conseils & Méthodes

1 $x_1 = 4$ et $\overline{x} = 10{,}5$ donc l'écart est égal à $4 - 10{,}5 = -6{,}5$.

2 Pour tracer la droite on fait un tableau de valeurs en choisissant deux valeurs de x relativement éloignées pour améliorer la précision du tracé et en calculant le y correspondant à l'aide de l'équation de la droite.

Méthode 6 D'autres ajustements

Énoncé

Le tableau ci-contre donne le nombre d'habitants y_i (avec i entier de 1 à 6) en Grèce, en millions d'habitants, tous les 4 ans depuis 1990 (année de rang $x_1=0$).

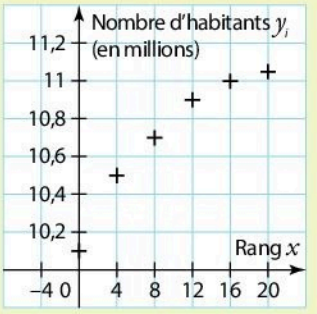
Années	1990	1994	1998	2002	2006	2010
Rang x_i	0	4	8	12	16	20
Nombre d'habitants y_i (en millions)	10,1	10,5	10,7	10,9	11	11,05

1. Au vu du nuage de points $(x_i ; y_i)$ ci-contre, un ajustement affine semble-t-il pertinent ?

2. On pose, pour tout entier i compris entre 1 et 6, $y'_i = e^{y_i}$. Calculer les valeurs y'_i.

3. Représenter le nuage de points $(x_i ; y'_i)$. Vérifier qu'il peut être ajusté par une droite.

4. Déterminer la droite de régression de y' en x.
En déduire une expression de y en fonction de x.



Solution

Conseils & Méthodes

1 Lorsque le nuage présente une forme incurvée non rectiligne il est très difficile de savoir s'il existe une courbe de corrélation entre les deux variables et surtout quelle courbe pourrait correspondre car plusieurs courbes de fonctions « usuelles » pourraient convenir. Le changement de variable proposé ici permet un assez bon ajustement du nuage avec une courbe « logarithme » mais qui n'est peut-être pas le meilleur.