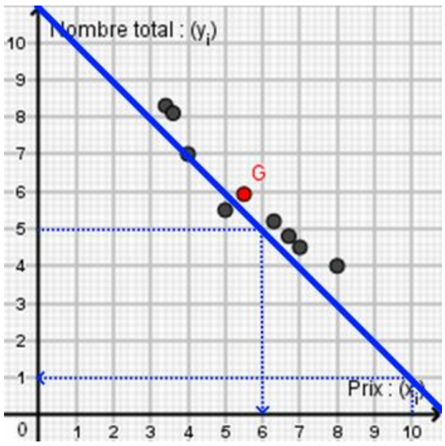
# Fiche P3.2 : Ajustement affine

## 1 – Droite d'ajustement

● Lorsque les points sont presque alignés, on peut approcher le nuage à l'aide d'une **droite d'ajustement**. On dit alors que l'on réalise un **ajustement affine** : On recherche une droite qui approxime au mieux tous les points du nuage, en passant au plus près de ceux-ci. La droite obtenue est appelée la droite de **régression.**

● Un ajustement permet de faire des estimations : Une **interpolation** lorsque la valeur à estimer se situe dans l'intervalle des valeurs connues et une **extrapolation** lorsqu'elle se situe à l'extérieur de cet intervalle.

Exemple 1 : On reprend la série sur le prix d'un paquet $(x_i)$ et le nombre total $(y_i)$ cigarettes vendues.

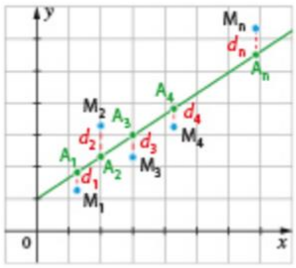



1) Un ajustement affine est possible car les points sont presque alignés.
2) Tracer « au jugé » un ajustement affine puis lire graphiquement l'équation de la droite de régression tracée : $y = 11 - x$
3) Utiliser cet ajustement affine pour estimer graphiquement :
   a. Le prix d'un paquet pour 50 milliards de cigarettes vendues :
   Interpolation avec $y = 5$. On trouve $x = 6$. Le paquet est vendu 6€.
   b. Le nombre de cigarettes vendus avec un paquet à 10€ :
   Extrapolation avec $x = 10$. On trouve $y = 1$ soit $10\ Milliards$.

## 2 – Droite des moindres carrés

Définition 1 : On appelle **droite des moindres carrés** l'ajustement affine obtenue en minimisant la somme des carrés des distances entre les points du nuage et la droite.

Explication : Parmi toutes les droites qui peuvent réaliser un ajustement affine du nuage de points $M_1, M_2, \ldots, M_n$, il en existe une telle que la somme des carrés des écarts $E = d_1^2 + d_2^2 + \cdots d_n{}^2$ soit minimale où $d_i = A_iM_i$ et $A_i$ est le point de la droite qui a la même abscisse que $M_i$. C'est en quelque sorte la droite qui passe « au plus près » de tous les points du nuage en même temps.




Propriété 1 : La **droite des moindres carrés** a pour équation $y = ax + b$ avec $a = \frac{cov(x;y)}{\sigma(x)}$ et $b = \bar{y} - a\bar{x}$

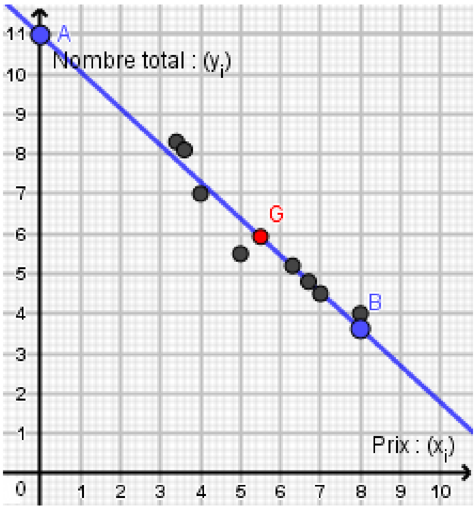
Remarque : Pour déterminer les coefficients $a$ et $b$ de l'équation, on utilise le tableur ou la calculatrice.

Tutoriel Casio : Dans le menu $STAT$: F3 ( REG ) puis F1 ( X ) puis F1 ( ax+b ) .

Tutoriel TI : Touche STAT, dans l'onglet $\boldsymbol{CALC}$ (touche ▶), sélectionner $\boldsymbol{RegLin(ax+b)}$ puis $\boldsymbol{Calculer}$.

Propriété 2 : La **droite des moindres carrés** passe par le point moyen du nuage $G(\bar{x};\bar{y})$

Exemple 2 : On reprend la série sur le prix d'un paquet $(x_i)$ et le nombre total $(y_i)$ cigarettes vendues.




1) Déterminer à la calculatrice une équation de la droite $(d)$ des moindres carrés, puis tracer cette droite sur la figure ci-contre.

   A la calculatrice on trouve : $a \approx -0.92$ et $b \approx 10.98$

   La droite des moindres carrés a pour équation $y = -0.92x + 10.98$

   Pour la tracer on calcule les coordonnées de 2 points de celle-ci :

   . $x = 0 : y = 10.98$ donc $A(0; 10.98)$

   . $x = 8 : y = -0.92 \times 8 + 10.98 = 3.62$ donc $B = (8; 3.62)$

2) Vérifier que le point moyen $G(5.5; 5.925)$ appartient à $(d)$

   $-0.92x_G + 10.98 = -0.92 \times 5.5 + 10.98 = 5.92 \approx y_G$

   Donc $G \in (d)$ car ses coordonnées vérifient l'équation de la droite.

3) Utiliser cet ajustement affine pour retrouver par le calcul les estimations précédentes :

   a. Le prix d'un paquet pour 50 milliards de cigarettes vendues : On sait que $y = 5$ et on cherche $x$.

   $-0.92x + 10.98 = 5 \Leftrightarrow -0.92x = -5.98 \Leftrightarrow x = \frac{5.98}{0.92} \approx 6.5$. On peut estimer à 6€ le prix du paquet.

   b. Le nombre de cigarettes vendus avec un paquet à 10€ : On sait que $x = 10$ et on cherche $y$.

   $y = -0.92 \times 10 + 10.98 = 1.78$. On peut estimer le nombre de cigarettes vendues à 17.8 Milliards.

4) Quel prix du paquet maximise la Recette totale réalisée par la vente de cigarettes ?

   $Recette = Prix\ d'une\ cigarette\ \times Nombre\ de\ ventes = \frac{x}{20} \times y$

   $R(x) = \frac{x}{20} \times (-0.92x + 10.98) = -\frac{0.92}{20}x^2 + \frac{10.98}{20}x = -0.0415x^2 + 0.599x$

   On dérive $R'(x) = -2 \times 0.0415x + 0.599 = -0.092x + 0.599$.

   On cherche quand la dérivée s'annule : $R'(x) = 0 \Leftrightarrow -0.092x + 0.599 = x = \frac{-0.599}{-0.092} = 6.5$.

   Le paquet doit être vendu 6.5€

## 3 – Coefficient de corrélation linéaire

Définition 1 : On considère une série statistique à deux variables $x = (x_1; \dots; x_n)$ et $y = (y_1; \dots; y_n)$.

On appelle **coefficient de corrélation linéaire** de la série, le nombre $r$ définie par $r = \frac{cov(x;y)}{\sigma(x)\sigma(y)}$

Propriété 1 : Le coefficient de corrélation linéaire est compris entre $-1$ et $1$ : $-1 \leq r \leq 1$

Remarque : Si $r$ est proche de $-1$ ou $1$, il y a une forte corrélation linéaire entre les deux variables et on pourra réaliser un ajustement affine. Si $r$ est proche de $0$, l'ajustement par une droite ne sera pas pertinent. Il se peut alors qu'un autre type de courbe puisse mieux ajuster le nuage (voir fiche suivante)

Calculatrice : Le coefficient de corrélation $r$ s'obtient à la calculatrice en même temps que les coefficients.

Exemple 2 : A la calculatrice on trouve $r \approx -0.96$ : L'ajustement affine est donc approprié.