Chapitre 6.
Statistiques à 2 variables (2/2)

Yann Barsamian

École Européenne de Bruxelles 1

Année scolaire 2021–2022

Plan du chapitre

- Un nouvel ajustement linéaire ($y = ax + b$) : la droite de Mayer
- Ajustement exponentiel ($y = c \times d^x$)
- Ajustement logarithmique ($y = c\ln(x) + d$)

Rappels

Lors du chapitre 4, on s'est intéressés à des nuages de points qui avaient la forme d'une droite. On a vu la méthode des moindre carrés qui donne l'équation d'une droite d'ajustement.

Cet ajustement a de bonnes propriétés mais un problème subsiste : il n'est pas faisable "à la main", en tout cas pas à notre niveau : on doit se servir de la calculatrice.

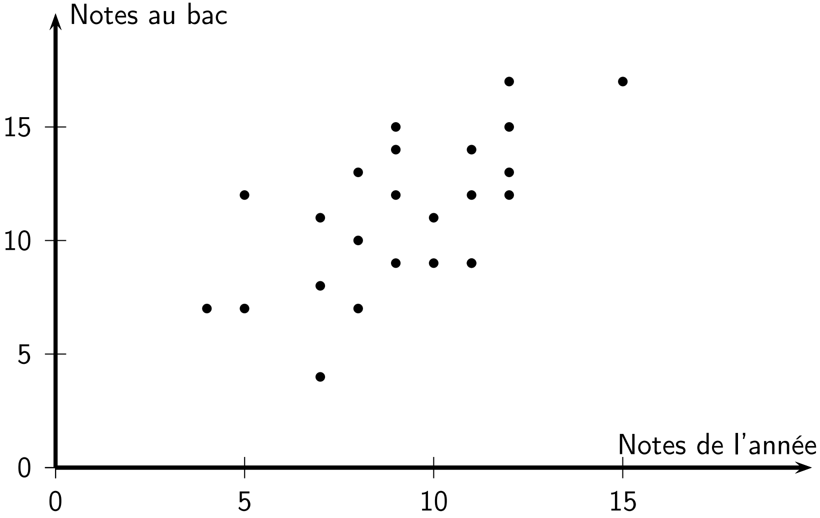
Dans ce chapitre, on va étudier un nouvel ajustement linéaire : la droite de Mayer. Cet ajustement est très simple à faire à la main !

Reprenons un exemple du chapitre précédent : pour une classe de Terminale en mathématiques, on regarde leur moyenne x au long de l'année et leur note y au baccalauréat (sur 20). Voici des notes fictives pour une classe de 24 élèves. Ainsi, chaque couple (x_i, y_i) (pour i de 1 à 24) représente les notes d'un élève de la classe.

x_i	8	9	7	15	12	12	10	8	11	11	7	8
y_i	7	9	4	17	13	15	9	13	14	9	11	10

x_i	11	11	12	12	7	9	9	5	9	5	10	4
y_i	9	12	17	12	8	15	12	7	14	12	11	7

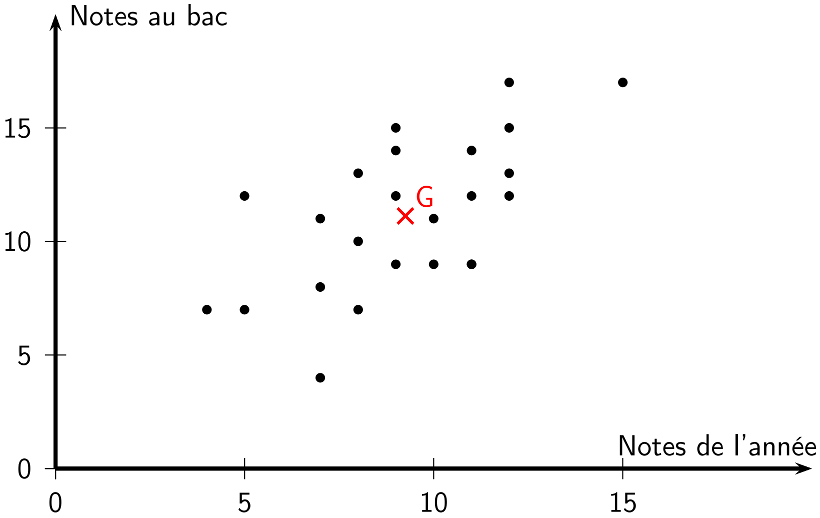
Le nuage de points $(x_i; y_i)$ était le suivant :



On se rappelle que le point moyen du nuage, c'est le point $G(\overline{x}, \overline{y})$.

$$x_G = \frac{8+9+7+15+12+12+10+8+11+11+7+8+11+11+12+12+7+9+9+5+9+5+10+4}{24}$$

$$y_G = \frac{7+9+4+17+13+15+9+13+14+9+11+10+9+12+17+12+8+15+12+7+14+12+11+7}{24}$$



Pour la méthode Mayer, on commence par séparer le nuage de points en deux parties, en ordonnant les points par x croissant.

x_i	4	5	5	7	7	7	8	8	8	9	9	9
y_i	7	7	12	4	8	11	7	10	13	9	12	14

x_i	9	10	10	11	11	11	11	12	12	12	12	15
y_i	15	9	11	9	9	12	14	12	13	15	17	17

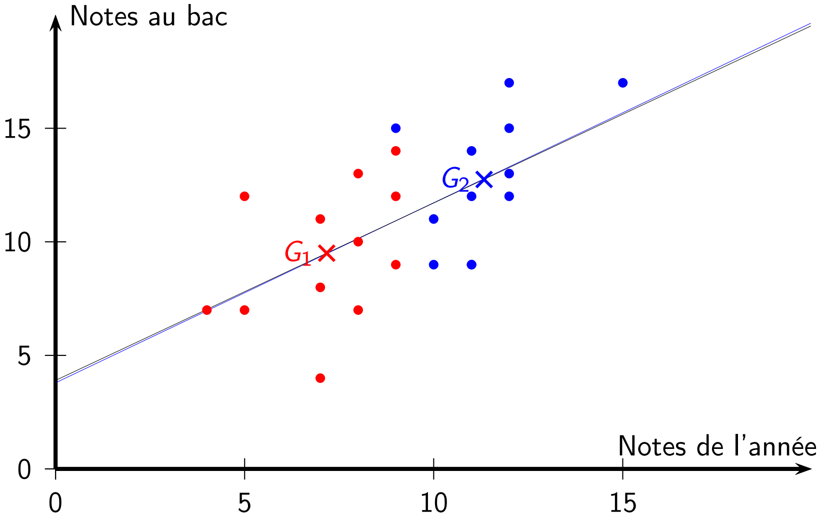
Pour chacun des deux sous-nuages, on calcule le point moyen :

$$\left\{\begin{array}{l} x_{G_1} = \frac{4+5+5+7+7+7+8+8+8+9+9+9}{12} \approx 7,17 \\ y_{G_1} = \frac{7+7+12+4+8+11+7+10+13+9+12+14}{12} = 9,5 \end{array}\right. \Rightarrow G_1(7,17;9,5)$$

$$\left\{\begin{array}{l} x_{G_2} = \frac{9+10+10+11+11+11+11+12+12+12+12+15}{12} \approx 11,33 \\ y_{G_2} = \frac{15+9+11+9+9+12+14+12+13+15+17+17}{12} = 12,75 \end{array}\right. \Rightarrow G_2(11,33;12,75)$$

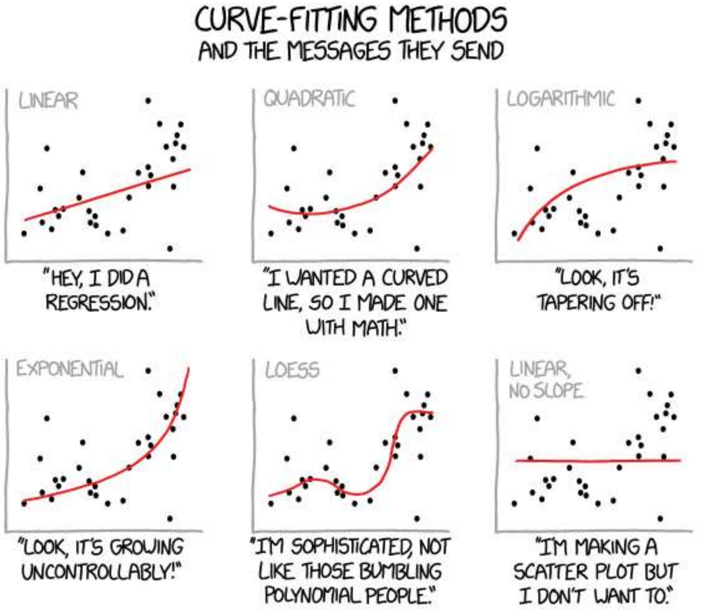
La droite de Mayer est la droite $(G_1 G_2)$! Voir la vidéo : `https://www.youtube.com/watch?v=ESHY4QPgriw`.

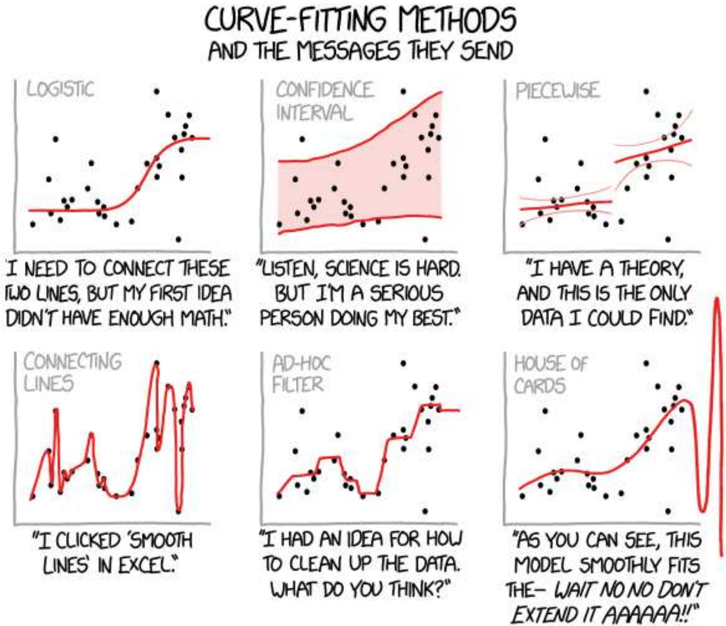
Sur le nuage, on peut facilement contrôler que G_1 et G_2 sont bien les points moyens de chacune des deux moitiés du nuage :



On a tracé (G_1G_2) en noir (et la droite des moindres carrés en bleu, quasiment identique). Comment calculer l'équation de (G_1G_2) ?

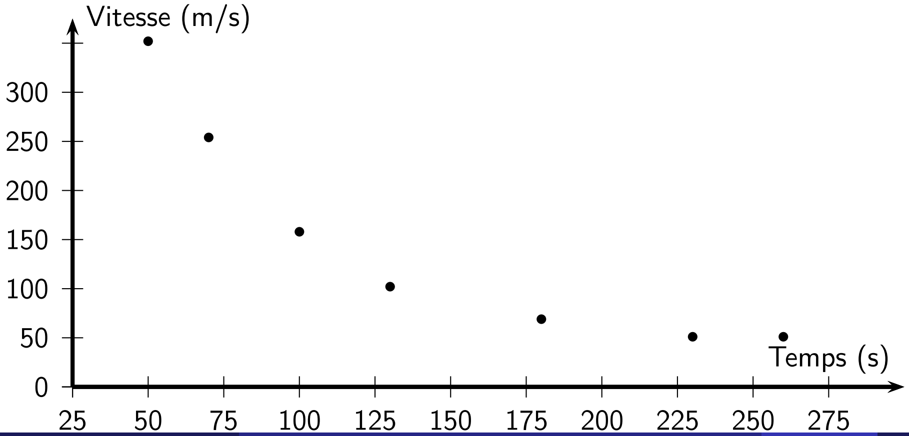
D'autres ajustements




CURVE-FITTING METHODS
AND THE MESSAGES THEY SEND
LOGISTIC
"I NEED TO CONNECT THESE TWO LINES, BUT MY FIRST IDEA DIDN'T HAVE ENOUGH MATH."
CONFIDENCE INTERVAL
"LISTEN, SCIENCE IS HARD. BUT I'M A SERIOUS PERSON DOING MY BEST."
PIECEWISE
"I HAVE A THEORY, AND THIS IS THE ONLY DATA I COULD FIND."
CONNECTING LINES
"I CLICKED 'SMOOTH LINES' IN EXCEL."
AD-HOC FILTER
"I HAD AN IDEA FOR HOW TO CLEAN UP THE DATA. WHAT DO YOU THINK?"
HOUSE OF CARDS
"AS YOU CAN SEE, THIS MODEL SMOOTHLY FITS THE– WAIT NO NO DON'T EXTEND IT AAAAAA!!"

Exemple du chapitre 4 : on a mesuré, pour un sauteur en parachute, à différents temps x depuis le saut, sa vitesse y.

x_i	50	70	100	130	180	230	260
y_i	352	254	158	102	69	51	51



III/ Ajustement logarithmique