# Ajustement

Avant la fin du XIX$^{e}$ siècle, on ne disposait d'aucun moyen fiable pour identifier une personne. Ce n'est qu'en 1882, grâce à Alphonse Bertillon, qu'ont été mises en place des fiches « d'anthropométrie », sur lesquelles on indiquait différentes mesures du corps : taille, envergure, longueur du buste, de la tête, de l'oreille droite... accompagnées des photos « profil/face » et des empreintes digitales, ces fiches ont été utilisées en France jusqu'en 1970.
En 1900, Charles Perrier, médecin à la prison de Nîmes, publie *Les criminels – Étude sur 859 condamnés*, d'où sont extraits les relevés suivants de mesures, en centimètre, effectuées sur dix hommes.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Taille $x_i$ | 154 | 158 | 162 | 163 | 164 | 165 | 167 | 168 | 170 | 174 |
| Envergure $y_i$ | 159 | 160 | 166 | 164 | 168 | 174 | 170 | 174 | 174 | 179 |

1. On considère un repère orthonormé du plan où les axes sont gradués à partir de la graduation 100 et où 1 cm représente 10 cm. Après avoir construit le nuage de points $(x_i; y_i)$, justifier pourquoi la taille et l'envergure sont deux grandeurs qu'on peut qualifier de « corrélées ».
2. On cherche à obtenir une relation affine entre taille et envergure, c'est- à-dire de la forme : envergure $= a \times \text{taille} + b$ .
   (a) En se basant sur 8 265 sujets, Bertillon écrit : « L'envergure l'emporte généralement sur la hauteur de la taille de 3 à 4 centimètres. » Tracer sur le graphique la droite $d$ d'équation $y = x + 4$.
   Que peut-on penser de l'affirmation de Bertillon ?
   (b) On considère maintenant la droite $\mathscr{D}$ d'équation $y = 1{,}06x - 5{,}8$.
   Laquelle des droites $d$ ou $\mathscr{D}$ semble « le mieux » ajuster le nuage de points ?
3. On souhaite maintenant comparer « numériquement » les ajustements du nuage de points par les droites $d$ ou $\mathscr{D}$.
   (a) Recopier et compléter le tableau suivant, donnant, pour chaque abscisse $x_i$, les carrés des écarts entre l'ordonnée $y_i$ et l'image $(x_i + 4)$ d'une part, et entre $y_i$ et l'image $(1{,}06x_i - 5{,}8)$ d'autre part, en arrondissant à $10^{-2}$.

| | | | | | |
|---|---|---|---|---|---|
| $x_i$ | 154 | 158 | ... | 174 | Total |
| $y_i$ | 159 | 160 | ... | 179 | |
| $(y_i - (x_i + 4))^2$ | 1 | 4 | ... | | |
| $(y_i - (1{,}06x_i - 5{,}8))^2$ | 2,43 | | ... | 0,13 | |

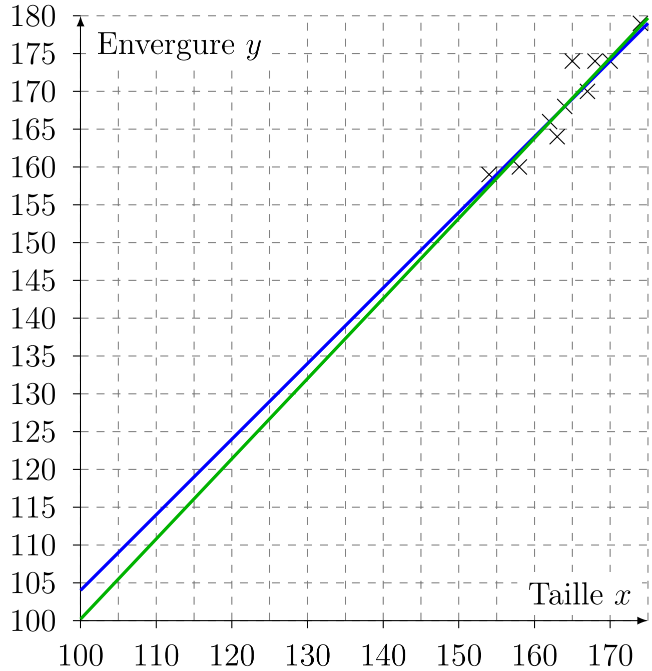
   (b) Quelle droite ajuste le mieux le nuage au sens « des moindres carrés » ?

# Correction

∽

1. Voici ci-contre le nuage de points avec les unités demandées.

   On peut considérer que les deux grandeurs sont corrélées car les points semblent se situer autour d'une droite.



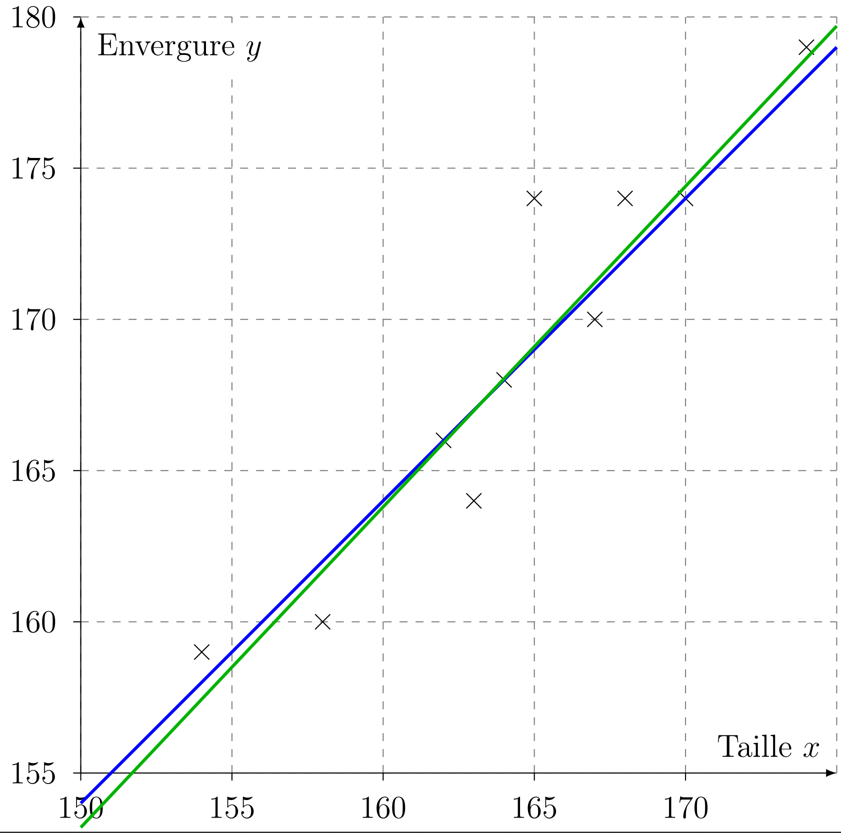

2. (a) La droite d'équation $y = x+4$ est tracée en bleu.

   Elle semble en effet passer assez près des points du nuage, donc Bertillon semble avoir raison.

   (b) La droite d'équation $y = 1{,}06x - 5{,}8$ est tracée en vert.

   Il est difficile de dire laquelle semble la mieux ajustée, mais la pente de la verte semble meilleure.

On peut agrandir la vue sur le nuage de points :




| $x_i$ | 154 | 158 | 162 | 163 | 164 | 165 | 167 | 168 | 170 | 174 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $y_i$ | 159 | 160 | 166 | 164 | 168 | 174 | 170 | 174 | 174 | 179 | |
| $(y_i - (x_i+4))^2$ | 1 | 4 | 0 | 9 | 0 | 25 | 1 | 4 | 0 | 1 | 45 |
| $(y_i - (1{,}06x_i - 5{,}8))^2$ | 2,43 | 2,82 | 0,01 | 8,88 | 0,00 | 24,01 | 1,49 | 2,96 | 0,16 | 0,13 | 42,89 |

3. (a) Voir le tableau ci-dessus (que l'on peut facilement obtenir à l'aide d'un tableur).

   (b) C'est pour la seconde droite que la somme des carrés des différences est la plus petite.

   Donc c'est la droite verte, d'équation $y = 1{,}06x - 5{,}8$ qui ajuste le mieux le nuage de point au sens des « moindres carrés ».