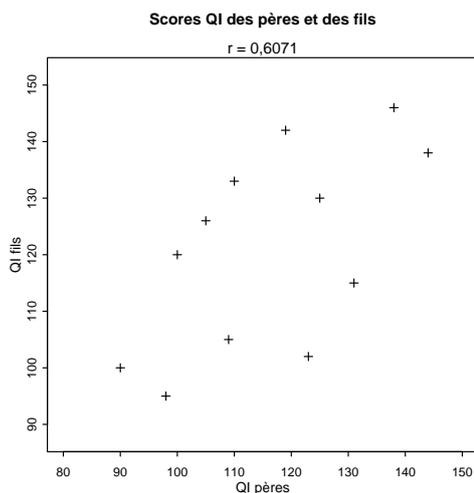


**Document Chapitre 3. Etude de la liaison entre deux variables quantitatives**

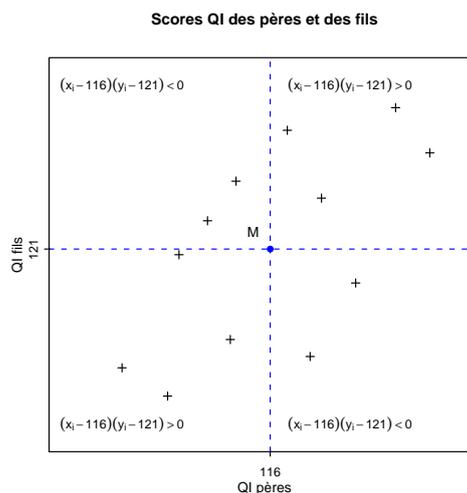
*Nuages de points et corrélations*

- *Exemple 1* Sur un échantillon de 12 couples (père, fils), on a relevé les QI des pères (variable  $X$ ) et les QI des fils (variable  $Y$ ) :

| couple (père, fils) n° i | 1   | 2   | 3   | 4   | 5  | 6   | 7   | 8   | 9   | 10  | 11  | 12  |
|--------------------------|-----|-----|-----|-----|----|-----|-----|-----|-----|-----|-----|-----|
| QI du père $x_i$         | 123 | 144 | 105 | 110 | 98 | 138 | 131 | 90  | 119 | 109 | 125 | 100 |
| QI du fils $y_i$         | 102 | 138 | 126 | 133 | 95 | 146 | 115 | 100 | 142 | 105 | 130 | 120 |

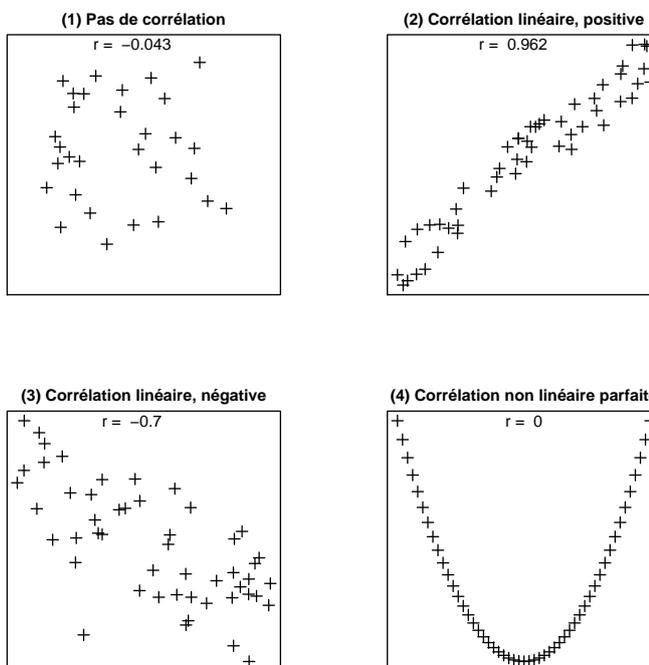


Graphique 1 :  
Nuage de points



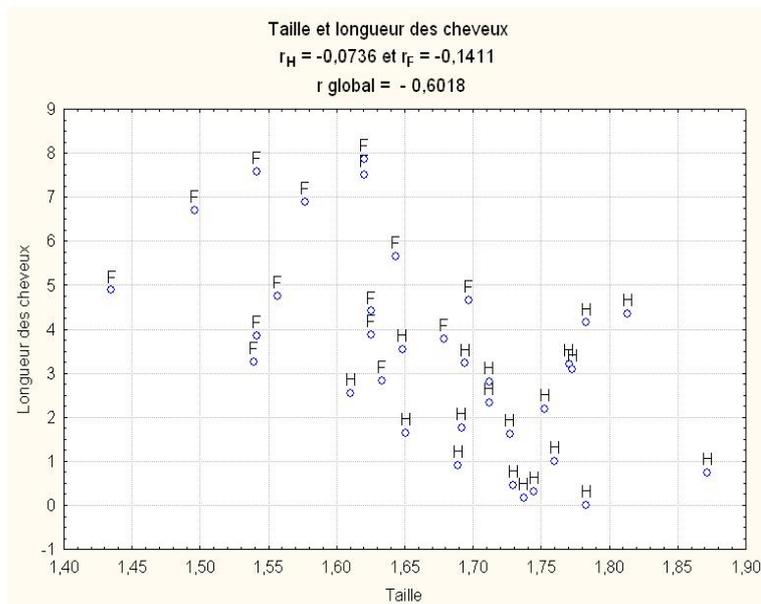
Graphique 3 :  
Dispersion des points autour du centre  $M$

- *Exemples-type*



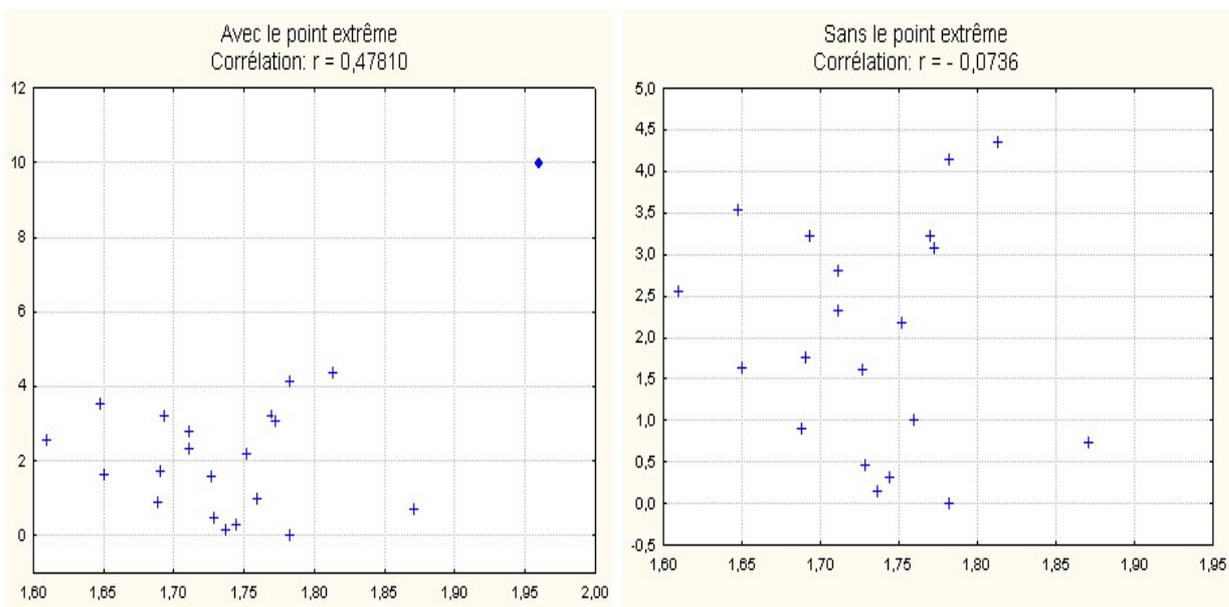
Graphique 2

- *Corrélation linéaire. Attention aux sous-groupes définis par une variable qualitative :*<sup>1</sup>



Graphique 4

- *Corrélation linéaire. Attention aux valeurs extrêmes :*



Graphique 5

1. Exemple emprunté à R. Rakotomalala.

## Coefficient de corrélation linéaire

Exemple 1. Résumés des données ; calcul de la covariance et du coefficient de corrélation linéaire observés :

$$\begin{aligned} \text{QI père : } \quad \sum x_i &= 1392 & \sum x_i^2 &= 164566 & \bar{x} &= 116 & s_x^* &= \sqrt{\frac{\sum x_i^2 - n\bar{x}^2}{n-1}} = 16,7712 \\ \text{QI fils : } \quad \sum y_i &= 1452 & \sum y_i^2 &= 179068 & \bar{y} &= 121 & s_y^* &= \sqrt{\frac{\sum y_i^2 - n\bar{y}^2}{n-1}} = 17,5188 \\ & \sum x_i y_i &= 170394 & & & & & \end{aligned}$$

$$\begin{aligned} \text{cov}^*(x, y) &= \frac{\sum x_i y_i - n\bar{x}\bar{y}}{n-1} = \frac{170394 - 12 \times 116 \times 121}{11} = 178,3636. \\ r(x, y) &= \frac{\text{cov}^*(x, y)}{s_x^* s_y^*} = \frac{178,3636}{16,7712 \times 17,5188} = 0,6071. \end{aligned}$$

## Coefficient de corrélation de Spearman

1. Exemple 1. Calcul du coefficient de corrélation de Spearman observé :

Classements des données :

|                  |     |     |     |     |    |     |     |     |     |     |     |     |
|------------------|-----|-----|-----|-----|----|-----|-----|-----|-----|-----|-----|-----|
| QI du père $x_i$ | 123 | 144 | 105 | 110 | 98 | 138 | 131 | 90  | 119 | 109 | 125 | 100 |
| QI du fils $y_i$ | 102 | 138 | 126 | 133 | 95 | 146 | 115 | 100 | 142 | 105 | 130 | 120 |
| rang $x'_i$      | 8   | 12  | 4   | 6   | 2  | 11  | 10  | 1   | 7   | 5   | 9   | 3   |
| rang $y'_i$      | 3   | 10  | 7   | 9   | 1  | 12  | 5   | 2   | 11  | 4   | 8   | 6   |

Calcul direct du coefficient de corrélation de Spearman :

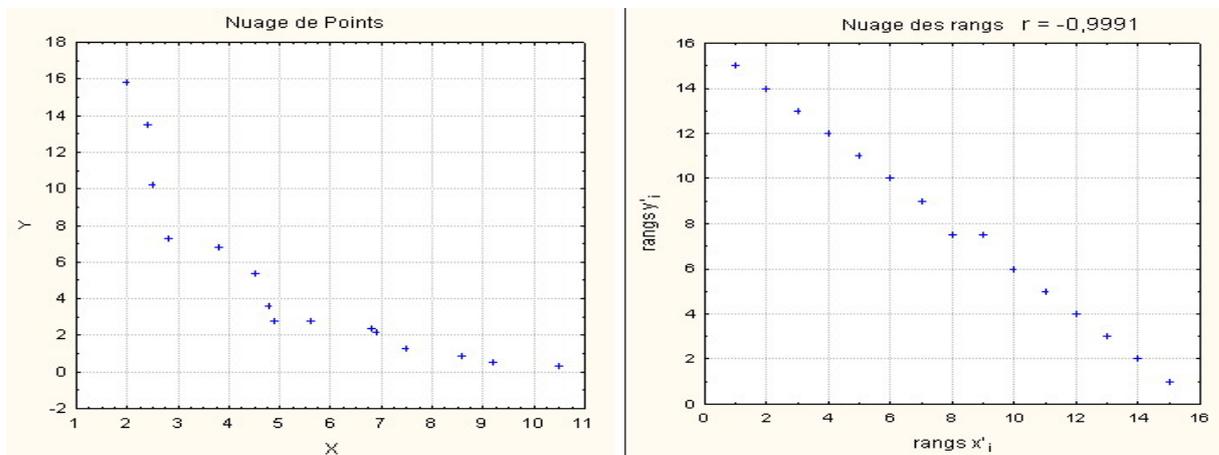
$$r_S(x, y) = \frac{\text{cov}^*(x', y')}{s_{x'}^* s_{y'}^*} = 0,643$$

Formule de calcul (sans ex-aequo) :

|                   |    |   |    |    |   |    |    |    |    |   |   |    |
|-------------------|----|---|----|----|---|----|----|----|----|---|---|----|
| $x'_i - y'_i$     | 5  | 2 | -3 | -3 | 1 | -1 | 5  | -1 | -4 | 1 | 1 | -3 |
| $(x'_i - y'_i)^2$ | 25 | 4 | 9  | 9  | 1 | 1  | 25 | 1  | 16 | 1 | 1 | 9  |

$$\sum (x'_i - y'_i)^2 = 102 \text{ et } r_S(x, y) = 1 - \frac{6 \times 102}{12(12^2 - 1)} = 0,643.$$

2. Exemple 2. Nuage des points  $(x_i, y_i)$  et nuage des rangs  $(x'_i, y'_i)$



Graphique 6. Deux valeurs  $y_i$  ex-aequo aux rangs 7 et 8 : on leur attribue le rang moyen 7,5.