

Corrélation et régression linéaire

Une **corrélacion** mesure une association (ou relation) entre 2 variables quantitatives (par exemple, X et Y). Le résultat est donné par un coefficient de corrélation nommé « r ».

Le coefficient de corrélation r :

- Aura toujours une valeur comprise entre -1 et $+1$: $-1 \leq r \leq 1$
- Un r positif (>0) indique une relation positive (croissante) entre les 2 variables étudiées :
Plus la variable X **augmente**, plus la variable Y **augmente**.
- Un r négatif (<0) indique une relation négative (décroissante) entre les 2 variables étudiées :
Plus la variable X **augmente**, plus la variable Y **diminue**.
- Un r proche de 0 ($\cong 0$) indique l'absence de relation entre les 2 variables étudiées.

Afin de voir s'il y a une association réelle entre les 2 variables étudiées X et Y, une p-valeur sera également donnée. Une p-valeur inférieure à 0,05 indique une association dite « statistiquement significative » c'est-à-dire que l'association entre les 2 variables peut être considérée comme réelle.

Une **régression linéaire** mesure également une association (ou relation) entre 2 variables quantitatives (par exemple, X et Y) mais les conditions d'application sont différentes. Le résultat est donné par une droite de régression :

$$\hat{Y} = a + b.x$$

où \hat{Y} est la moyenne de la variable Y en X,
 b est la pente,
et a est l'ordonnée à l'origine (=intercept) c'est-à-dire l'endroit où commence la droite.

Un coefficient b positif (>0) indique une relation positive (croissante) entre les 2 variables étudiées c'est-dire que plus la variable X **augmente**, plus la variable Y **augmente** ;

Un coefficient b négatif (<0) indique une relation négative (décroissante) entre les 2 variables étudiées c'est-à-dire que plus la variable X **augmente**, plus la variable Y **diminue**.

Un coefficient b proche de 0 ($\cong 0$) indique l'absence de relation entre les 2 variables étudiées.