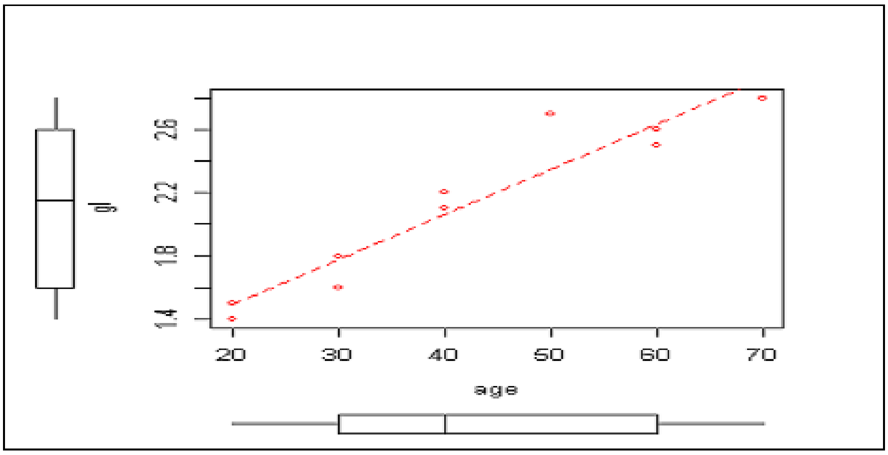
Régressions Linéaire simples

But : ´établir un lien entre une variable dépendante Y et une variable indépendante X pour pouvoir ensuite faire des prévisions sur Y lorsque X est mesurée.

Exemple de régression linéaire simple

Sur un échantillon de 10 sujets d’âges différents, on a recueilli l'âge et la concentration sanguine du cholestérol (en g/L) de 10 individus :

age (x_i)	30	60	40	20	50	30	40	20	70	60
gl (y_i)	1.6	2.5	2.2	1.4	2.7	1.8	2.1	1.5	2.8	2.6



Modèle de régression linéaire simple

$$\boldsymbol{y_i = ax_i + b}$$

Estimation de paramètres

1) **la méthode des moindres carrés**
   Paramètres $a$et b

$$a = \frac{COV\,(x,y)}{v(x)} = \frac{\sum_i (y_i - \bar{y})\,(x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

$$b = \bar{y} - a\bar{x}$$

Coefficient de corrélation

L’étudier la corrélation entre deux ou plusieurs variables aléatoire : est ´étudier l’intensité de la liaison qui peut être existée entre ces variables.
Une mesure de cette corrélation dans le cadre linéaire est obtenue par le calcul du coefficient appelé coefficient de corrélation

$$\boldsymbol{r = \frac{cov(x,y)}{\sigma_x \sigma_y}}$$

$$Cov(x,y) = \frac{1}{n}\sum(x_i - \bar{X})(y_i - \bar{Y})$$

$$\sigma_x^2 = \frac{1}{n}\sum(x_i - \bar{X})^2$$

$$\sigma_y^2 = \frac{1}{n}\sum(y_i - \bar{Y})^2$$

$Cov(x,y)$ **:** Covariance X, Y
σ_x^2 **:** Variance de X
σ_y^2 **:** Variance de Y

Le coefficient de corrélation est compris entre -1 et +1.
r = +1 corrélation positive parfaite
r = -1 corrélation négative parfaite
r = 0   absence totale de corrélation

Coefficient de détermination

Coefficient de détermination : Exprime la part de variabilité de Y expliquée par le modèle.

$R^2 \longrightarrow 1$, le modèle est excellent

$R^2 \longrightarrow 0$, le modèle ne sert à rien

$$R^2 = r^2$$