Validation d'une loi physique par regression linéaire

1 Régression linéaire

But :

1. on cherche à vérifier si des mesures expérimentales de **grandeurs liées** x et y sont compatibles avec une **relation linéaire** entre elles

valeurs mesurées de x	x_1	x_2	...	x_n
valeurs correspondantes de y	y_1	y_2	...	y_n

Relation linéaire $\boxed{y = ax + b}$ valable ?

2. et si c'est le cas, à estimer les valeurs de a et b ainsi que les incertitudes associées.

$$a = \cdots \pm \ldots \qquad b = \cdots \pm \ldots$$

1.1 Mise en œuvre de la régression

1. On mesure une série de $n \geq 5$ données [1] $x_1, x_2, \ldots, x_n$ et les valeurs correspondantes $y_1, y_2, \ldots, y_n$, accompagnées de leurs incertitudes respectives.

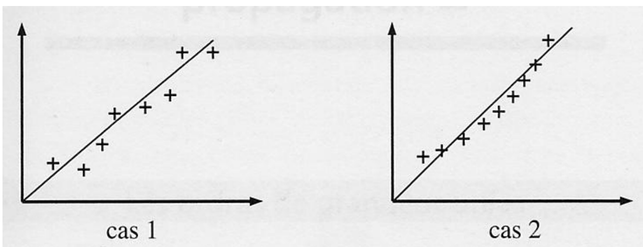
2. Deux méthodes
   - *méthode logicielle :* avec un tableur (`Latis Pro`, `Regressi`,...)
     - on rentre les valeurs dans deux colonnes du tableur
     - on visualise le graphe $y = f(x)$
     - on demande au logiciel d'utiliser un **modèle affine**. Il trace alors la droite moyenne en utilisant la **méthode des moindres carrées** [2].
   - *méthode manuelle :* sur papier millimétré et avec une règle
     - on place sur un graphe $y = f(x)$ (respecter les échelles) les points de mesure (x_i, y_i)
     - On trace **à la règle** la droite qui passe **le mieux** par tous les points (droite «moyenne»).

1.2 Validité du modèle linéaire

Faire une régression linéaire ne prouve nullement sa validité. Pour cela il faut **observé soigneusement la répartition des points de mesures** (x_i, y_i) **autour de la droite moyenne**.

- Si les points de mesures semble **répartis de manière aléatoire** autour de la droite moyenne, alors le modèle linéaire est validé (cas 1).

- Si les points de mesures semble **suivre une courbe non linéaire**, alors le modèle linéaire est invalidé (cas 2).



La seule façon valable de conclure à la validité d'une régression linéaire est une représentation graphique où l'on observe l'alignement des points avec la droite de régression.

1. un nombre plus faible de valeurs mesurées rend la régression peu fiable
2. L'ordinateur détermine la droite qui minimise la somme de carrés des écarts entre les points (x_i, y_i) et cette même droite.

1.2.1 Se méfier du coefficient de corrélation

Les calculatrices et certains logiciels fournissent un coefficient de corrélation r dont la valeur absolue est comprise entre 0 et 1. Plus sa valeur est proche de 1 (ou de -1), plus les écarts des points à la droite moyenne sont faibles. Cependant, il ne permet pas de dire si ces écarts sont *aléatoires* (loi linéaire validée) ou *systématiques* (loi linéaire invalidée), comme le montre l'exemple ci-dessous.

$$\begin{cases} a = -0{,}204 \pm 0{,}011 \\ b = \quad 2{,}110 \pm 0{,}067 \\ r = -0{,}989 \end{cases} \qquad \begin{cases} a = -0{,}175 \pm 0{,}013 \\ b = \quad 1{,}657 \pm 0{,}055 \\ r = -0{,}989 \end{cases}$$

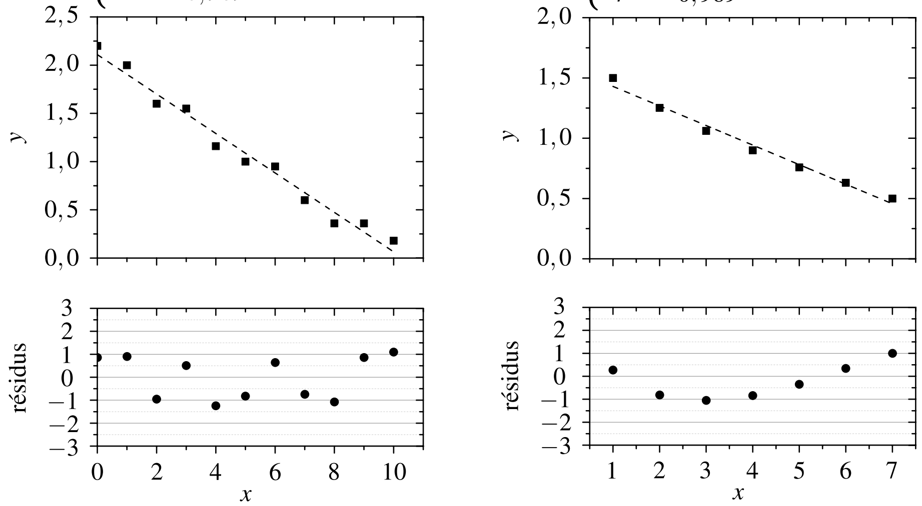


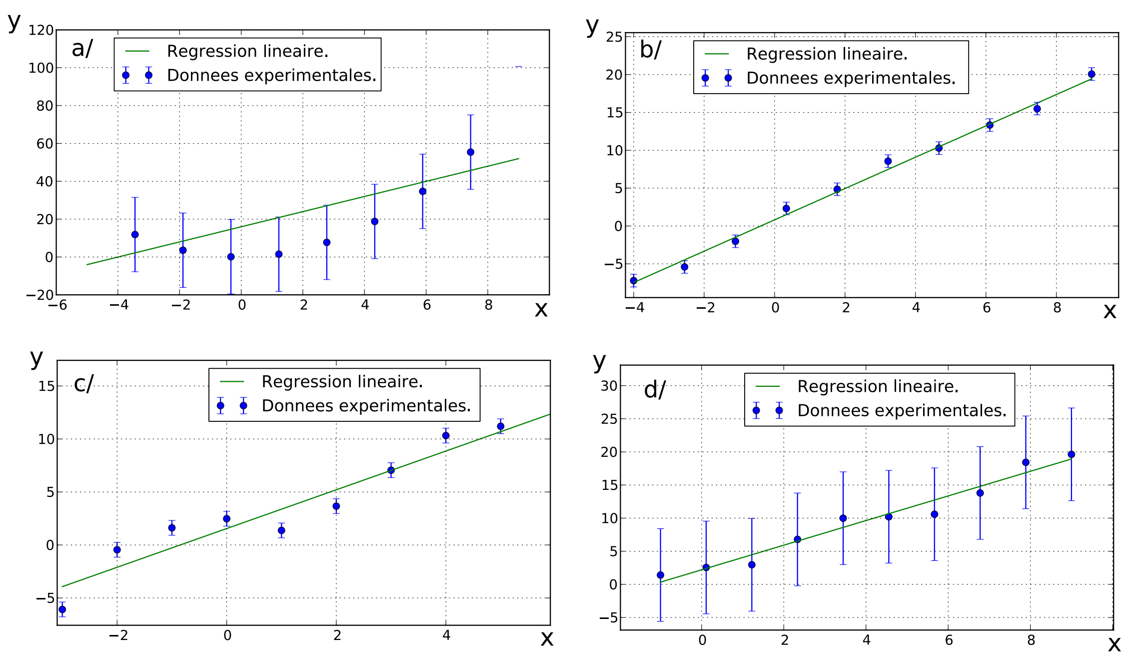
Figure 33.3 – Exemple de régression linéaire avec tracé des résidus normalisés :
à gauche : résidus aléatoires (points répartis aléatoirement autour de la droite) ;
à droite : résidus non aléatoires (points non répartis aléatoirement autour de la droite).

Résidus normalisés De nombreux logiciels de traitement de données proposent le tracé des résidus normalisés, qui représentent les écarts entre les points de mesure et la droite de régression divisés par l'écart quadratique moyen calculé sur l'ensemble des points. Si les résidus sont répartis aléatoirement de part et d'autre de 0, le modèle linéaire est validé.

1.2.2 Compatibilités des incertitudes avec un modèle linéaire

Si on dispose d'incertitudes types sur les mesures de y, il faut aussi vérifier que la droite de régression passe suffisamment près des points de mesures en tenant compte des barres d'incertitudes.

Exercice : conclure dans chaque cas



. . . .
. . . .
. . . .
. . . .
. . . .
. . . .
. . . .

1.3 Détermination des paramètres a et b

Méthode logicielle : les valeurs de a et de b, ainsi que leurs incertitudes sont fournies par le logiciel. Le fait que le logiciel renvoie des valeurs de a et b n'est certainement pas une garantie de succès : les calculs aboutissent toujours à quelque chose, même si la relation linéaire n'est pas du tout vérifiée.

Atttention !!!

Les logiciels (ou calculatrices) fournissent toujours des estimations des valeurs de a et b, que le modèle linéaire soit valable ou non !

Méthode manuelle («à la règle») : *pas d'incertitudes sur a et b dans ce cas.*

- b est l'ordonnée à l'origine de la droite moyenne.
- a est le coefficient directeur de la droite moyenne.

Plus le nombre n de valeurs mesurées (x_i, y_i) est important, plus les estimations de a et de b sont précises.

2 Se ramener à la vérification d'une loi linéaire

2.1 Méthode générale

Si la relation entre deux grandeur n'est pas linéaire, il faut se ramener à d'autres grandeurs facilement *calculables à partir des valeurs mesurées* et qui ont une relation linéaire entre elles.

2.2 Exemple

On mesure la puissance P dissipée par effet Joule dans une résistance, ainsi que l'intensité i la traversant.

valeurs mesurées de P (W)	$2,1$	$4,5$	$7,9$	$18,2$	$31,8$
valeurs mesurées de i (mA)	$1,0$	$1,5$	$2,0$	$3,0$	$4,0$

Incertitudes $\begin{cases} u(P) = 0,2 \text{ W} \\ u(i) = 0,1 \text{ mA} \end{cases}$

On veut vérifier expérimentalement la loi **non linéaire** $\boxed{P = Ri^2}$.

Étape 1 : identification des grandeurs *calculables* à partir de P et i vérifiant une loi *linéaire.*

. . . .
. . . .
. . . .

Étape 2 : Calcul des valeurs des nouvelles grandeurs à partir des valeurs mesurées de P et i, *sans oublier les incertitudes associées !*

. . . .
. . . .
. . . .

. . . .
. . . .
. . . .

Étape 3 : régression linéaire avec `Regressi` ou bien à la règle

. . . .
. . . .
. . . .