

Coefficient de corrélation linéaire

1. Calculez le coefficient de corrélation r pour les six paires de mesures suivantes :

x	1	2	3	5	6	7
y	5	6	6	8	8	9

Effectuez ces calculs à la main. Si votre calculatrice dispose d'une fonction prédéfinie pour calculer r , assurez-vous que vous savez l'utiliser en vérifiant votre résultat.

2. Démontrez l'égalité :

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - N \cdot \bar{x} \cdot \bar{y}$$

Prouvez alors que le coefficient de corrélation r peut se réécrire :

$$r = \frac{\sum x_i y_i - N \cdot \bar{x} \cdot \bar{y}}{\sqrt{(\sum x_i^2 - N\bar{x}^2)(\sum y_i^2 - N\bar{y}^2)}}$$

La plupart des calculatrices utilisent cette expression pour obtenir r , évitant ainsi la sauvegarde de toutes les données avant le calcul des moyennes et des écarts types.

3. Dans l'effet photo-électrique, l'énergie cinétique K des électrons éjectés du métal par la lumière varie linéairement avec la fréquence de la lumière incidente f :

$K = hf - \zeta$, où h et ζ sont des constantes.

Une étudiante vérifie cette loi en mesurant K pour n valeurs différentes de f puis en calculant le coefficient de corrélation r . Cinq mesures donnant $r = 0.7$ confirment-elles significativement la relation linéaire ? Et si $n = 20$, on avait $r = 0.5$?

Régression linéaire

4. Utilisez la méthode des moindres carrés pour trouver la droite $y = a x + b$ s'ajustant le mieux aux trois points (1,6), (5,3), (5,1). Tracez ces points et la droite associée. Votre calculatrice possède certainement une fonction intégrée permettant le calcul de a et de b . Profitez en pour apprendre à vous en servir en vérifiant votre résultat.

5. Droite passant par l'origine : Soit deux variables x et y satisfaisant la relation $y = ax$ dont le graphe est une droite passant par l'origine, comme par exemple la loi d'Ohm $U=RI$. Ayant N mesures (x_i, y_i) d'incertitudes négligeables en x et toutes égales en y , démontrez que la meilleure estimation de a s'écrit :

$$a = \frac{\sum xy}{\sum x^2}$$

6. Nous désirons déterminer si le taux de natalité peut être expliqué uniquement par le taux d'urbanisation. Il s'agit donc d'estimer le taux de natalité en fonction du taux d'urbanisation, à l'aide d'une droite de régression.

Pays	Taux de natalité	Taux d'urbanisation
Canada	16.2	55.0
Costa Rica	30.5	27.3
Cuba	16.9	33.3
États Unis	16.0	56.5
El Salvador	40.2	11.5
Guatemala	38.4	14.2
Haïti	41.3	13.9
Honduras	43.9	19.0
Jamaïque	28.3	33.1
Mexique	33.9	43.2
Nicaragua	44.2	28.5
Trinidad	24.6	6.8
Panama	28.0	37.7
Rep. Dominicaine	33.1	37.1

- Calculez les meilleurs estimateurs de a et de b pour $y = ax + b$
- Calculez R^2
- Quelle proportion de variation de la variable Y est expliquée par la relation linéaire ainsi déterminée ?
- Testez l'hypothèse $a=0$ sur la pente.
- Donnez l'intervalle de confiance de a à 95%
- Même chose que les questions d et e mais sur l'ordonnée à l'origine.

Autres courbes d'ajustement par moindres carrés

7. L'accélération d'un corps en chute libre peut se déduire de mesures successives de sa hauteur y_i à des temps t_i régulièrement espacés (évalués avec un stomboscope par exemple) après détermination du meilleur ajustement au polynôme attendu :

$$y = y_0 + v_0 t - \frac{1}{2} g t^2$$

Utilisez la méthode aux moindres carrés pour obtenir les meilleures estimations des trois coefficients de cette loi et par la même, la meilleure estimation de g, à partir des cinq mesures du tableau :

X (temps t)	-2	-1	0	1	2
Y (hauteur h)	131	113	89	51	7

Hauteur (en cm) en fonction du temps (en dixième de seconde) d'un corps en chute

8. Le rythme des désintégrations d'un échantillon radioactif décroît exponentiellement au cours du temps. Un étudiant suit cette décroissance exponentielle en enregistrant à l'aide d'un compteur le nombre de désintégrations durant le laps de temps de 15 s. Répétant cinq fois ses mesures à intervalle de 10 minutes, il obtient les résultats suivants :

X (temps t min)	10	20	30	40	50
Y (v désintégrations)	409	304	260	192	170

Nombre $v(t)$ de désintégrations dans des intervalles de 15 s, en fonction du temps total écoulé.

Si la loi de désintégration est exponentielle, le nombre $v(t)$ vérifie :

$$v(t) = v_0 e^{-t/\tau}$$

où τ est la vie moyenne (inconnue) de l'échantillon et v_0 une autre constante inconnue.

Que vaut la vie moyenne τ ? Combien de désintégrations aurait-il relevé en 15 secondes à $t=0$?