Statistique à deux variables quantitatives

Fiche tableaux : les formules à connaître
(Mathématiques complémentaires Terminale)

Série à une variable

Indicateur	Formule	Changement $y = ax + b$
Moyenne	$\overline{x} = \dfrac{1}{n}\sum x_i$	$\overline{y} = a\,\overline{x} + b$
Variance	$V(x) = \dfrac{1}{n}\sum (x_i - \overline{x})^2 = \dfrac{1}{n}\sum x_i^2 - \overline{x}^2$	$V(y) = a^2\,V(x)$
Écart type	$\sigma_x = \sqrt{V(x)}$	$\sigma_y = \lvert a\rvert\,\sigma_x$

Exemple : la série 1 ; 2 ; 6 a pour moyenne 3 et pour variance $\dfrac{4+1+9}{3} = \dfrac{14}{3}$.

Série double

Objet	Formule	À retenir
Point moyen	$G(\overline{x}\,;\overline{y})$	rarement un point du nuage
Covariance	$\mathrm{Cov}(x,y) = \dfrac{1}{n}\sum (x_i - \overline{x})(y_i - \overline{y})$	signe = sens du nuage
Droite des moindres carrés	$a = \dfrac{\mathrm{Cov}(x,y)}{V(x)}$, $b = \overline{y} - a\,\overline{x}$	passe par G
Résidu	$e_i = y_i - (ax_i + b)$	somme nulle pour cette droite
Coefficient de corrélation	$r = \dfrac{\mathrm{Cov}(x,y)}{\sigma_x\,\sigma_y}$	$-1 \leqslant r \leqslant 1$, $a = r\,\dfrac{\sigma_y}{\sigma_x}$

Exemple : $\overline{x} = 2$, $\overline{y} = 10$, $\mathrm{Cov}(x,y) = 6$, $V(x) = 3$ donnent $a = 2$ et $b = 10 - 2\times 2 = 6$.

Lire le coefficient de corrélation

Valeur de r	Lecture (repères usuels)
$r = 1$ **ou** $r = -1$	points parfaitement alignés
$\lvert r\rvert \geqslant 0{,}95$	alignement très net : ajustement affine justifié
$0{,}8 \leqslant \lvert r\rvert < 0{,}95$	tendance affine nette, dispersion visible
$\lvert r\rvert < 0{,}8$	ajustement affine discutable : regarder le nuage
r **proche de** 0	pas de lien affine (un lien courbé reste possible)

Le signe de r est celui de a : $r > 0$ pour un nuage croissant, $r < 0$ pour un nuage décroissant.

Changements de variable

Indice sur le nuage	Changement de variable	Ajustement affine obtenu	Retour à y
y **est multiplié par un nombre à peu près constant quand** x **augmente d'un même pas**	$z = \ln y$	$z = \alpha x + \beta$	$y = \mathrm{e}^{\beta}\,\mathrm{e}^{\alpha x}$
y **semble varier comme l'inverse de** x	$z = \dfrac{1}{x}$	$y = \alpha z + \beta$	$y = \dfrac{\alpha}{x} + \beta$
modèle puissance $y = k\,x^p$	$u = \ln x$, $v = \ln y$	$v = p\,u + \ln k$	$y = k\,x^p$

Exemple : $\ln y = 0{,}05x + 2$ donne $y = \mathrm{e}^{2}\,\mathrm{e}^{0{,}05x} \approx 7{,}39\,\mathrm{e}^{0{,}05x}$.

Prévoir avec un ajustement

Valeur de x	Nom	Fiabilité
dans $[x_{\min}\,;x_{\max}]$	interpolation	raisonnable
hors de $[x_{\min}\,;x_{\max}]$	extrapolation	« si la tendance se poursuit », fragile si l'on s'éloigne

Question réciproque : résoudre $ax + b = y_0$ ou $ax + b > y_0$ (sens inversé si $a < 0$).

Ce document est la propriété intellectuelle de Ben, professeur particulier (benprofparticulier.fr), protégée par le Code de la propriété intellectuelle (art. L. 111-1 et L. 122-4). Il est mis à disposition pour un usage personnel de révision. Toute reproduction, diffusion, mise en ligne ou revente, totale ou partielle, sans autorisation écrite est interdite et constitue une contrefaçon (art. L. 335-2).