

Dossier TESTS2 - Analyse

1/ Construire un tableau permettant de reconstituer le protocole de base

	DEXT X	RAPI Y
i01	0	0
i02	1	4
i03	1	2
i04	1	1
(...)	(...)	(...)
(...)	(...)	(...)
i28	4	4
i29	4	3
i30	5	5

2) Indiquer comment retrouver la moyenne de X (2.57) à partir de la distribution bivariée.

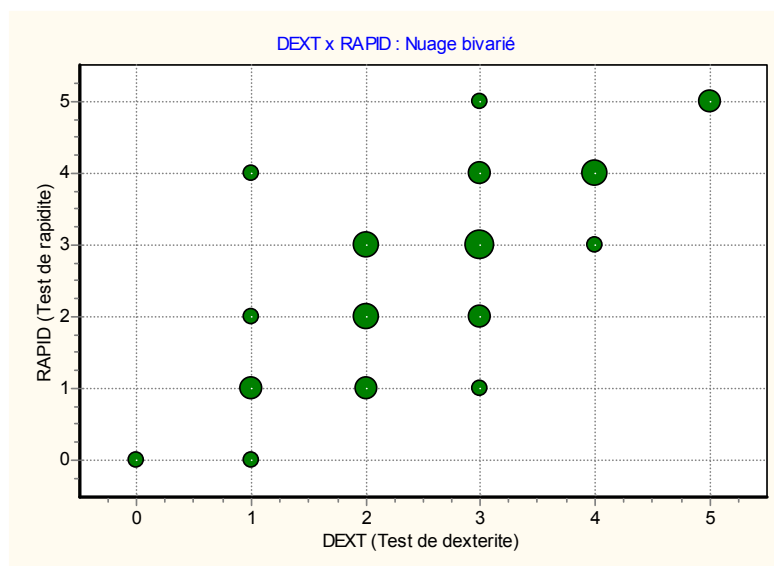
On calcule la distribution (univariée) de la variable X à partir de la distribution bivariée en calculant les totaux par colonne.

		0	1	2	3	4	5	Total
Scores au test de rapidité (variable Y)	5				1		2	
	4		1		2	3		
	3			3	4	1		
	2		1	3	2			
	1		2	2	1			
	0	1	1					
Total		1	5	8	10	4	2	30

$$Moy(X) = [(1 \times 0) + (5 \times 1) + (8 \times 2) + (10 \times 3) + (4 \times 4) + (2 \times 5)] / 30 = \dots$$

3) Représenter le nuage de points associé au tableau ci-dessus sur un graphique en coordonnées cartésiennes.

On représente le nuage pondéré par les effectifs. C'est le plus pertinent ici puisqu'on n'a pas les données individuelles.



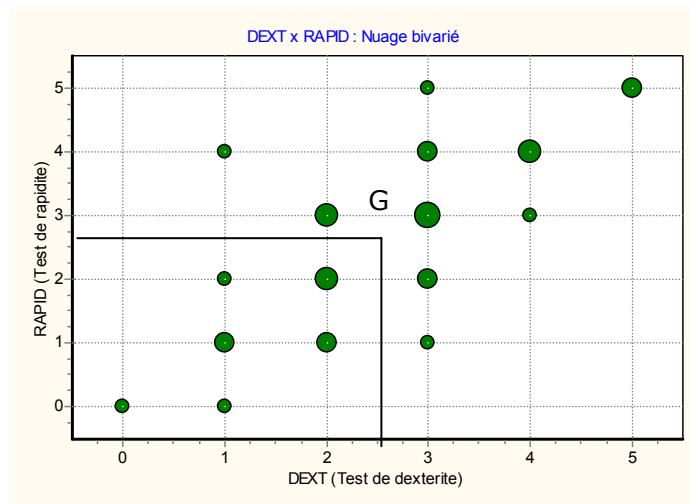
4) Quel type de relation suggère ce nuage ? Justifier votre réponse.

Relation

- de type linéaire (nuage de forme allongée)
- positive (nuage ascendant vers la droite)

5) Indiquer les coordonnées du point moyen du nuage et placer un repère sur le graphique.

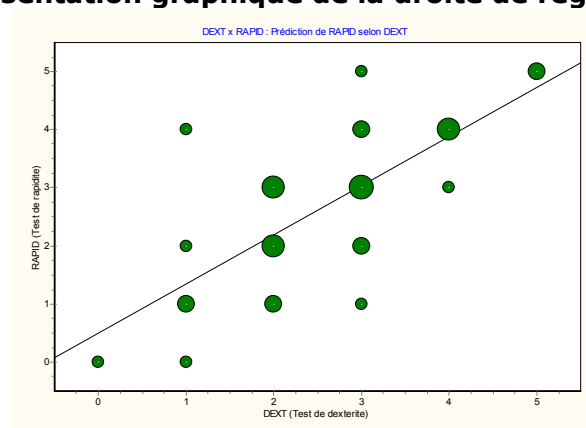
Le point moyen a pour coordonnées (2.57, 2.67), les moyennes de X et de Y..

**6/ Indiquer comment retrouver la valeur du coefficient de corrélation linéaire de Bravais-Pearson (RBP_{obs}) à partir de la covariance.**

$$RBP = r = \frac{Cov(X, Y)}{Ety(X) \cdot Ety(Y)} = \frac{1.22}{1.20 \times 1.40} = +0.73$$

7/ Élaborer une conclusion descriptive.

Pour ce groupe de 30 élèves, on observe une relation entre le score au test de dextérité (X) et le score au test de rapidité (Y). La liaison est positive : plus la note au test de dextérité est élevée et plus la note au test de rapidité l'est également. Cette liaison ($r = +.73$) peut être qualifiée de forte ($|r| > .40$).

8) L'équation de la droite de régression de Y en X est : $Y_{pred} = 0.85 X + 0.50$.**Représentation graphique de la droite de régression**

a/ Indiquer comment calculer les deux coefficients a et b

$$a = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = r \times \frac{\text{Ety}(Y)}{\text{Ety}(X)} \qquad a = \frac{1.22}{1.45} = 0.73 \times \frac{1.40}{1.20} = 0.85$$

$$b = \bar{y} - a \bar{x} \qquad b = 2.67 - 0.85 \times 2.57 = 0.50$$

Note : les résultats présentés ici ont été obtenus avec des calculs précis, avec "toutes" les décimales sur les statistiques utilisées dans les formules et non sur les valeurs arrondies présentées ici.

b/ Commenter les deux coefficients de régression

$$a = 0.85$$

En moyenne, quand le score en Dextérité (X) augmente de 1 point, celui de Rapidité (Y) tend à augmenter de 0.85 pts.

$$b = 0.50$$

La régression prédit un score de 0.50 en Rapidité (Y) à un élève qui aurait un score nul en Dextérité (X)

c/ Discuter de la qualité de cet ajustement

Pour évaluer la qualité de l'ajustement linéaire, on calcule le coefficient de détermination (R^2).

$$\text{On trouve : } R^2 = (r)^2 = 0.53 \text{ (0.5284)} = 53\%$$

On peut constater que la qualité d'ajustement est très bonne ($R^2 = 53\% > 16\%$).

9) Peut-on conclure à l'existence d'une liaison linéaire entre ces deux variables chez l'ensemble des élèves de CM2 de cet arrondissement parisien ?

$$t_{\text{obs}} = 5.60$$

$$ddl = 28$$

La table indique pour $ddl = 28$ et $t = 3.674$, $p/2 = 0.05\%$ et $p = 0.1\%$

Résultat du test : Le test est significatif au seuil 0.1%
(Note : le seuil exact pour $t=5.60$ est $p/2 = 0.00027\%$ et $p = 0.00054\%$)

a/ Si oui, que peut-on dire sur la force de cette "liaison parente" ?

Au vu du résultat du test *T de Student* on peut simplement dire que la liaison parente n'est pas nulle. On ne peut rien dire sur la force de cette liaison parente.

Par contre l'intervalle de confiance (**IC(5%) = [+0.50; +0.86]**) indique que +0.40 (et toutes les valeurs inférieures à +0.40) fait partie des valeurs de *RBPpar* pouvant être rejetées. On peut donc en conclure que la corrélation parente, *RBPpar*, est au moins supérieure à 0.40, donc importante / forte (au seuil 5%).

b/ Rédiger une conclusion inférentielle

Il semble qu'il existe, chez l'ensemble des enfants (dans la population parente) une liaison positive entre le score au test de dextérité et le score au test de rapidité ($t = 5.60$, $ddl = 28$, $p < 0.1\%$). Il semble que cette liaison est forte (supérieure à +0.40) (**IC(5%) = [+0.50; +0.86]**).

Note : La conclusion inférentielle commence par le rappel des aspects descriptifs (cf. conclusion descriptive) non rappelés ici.