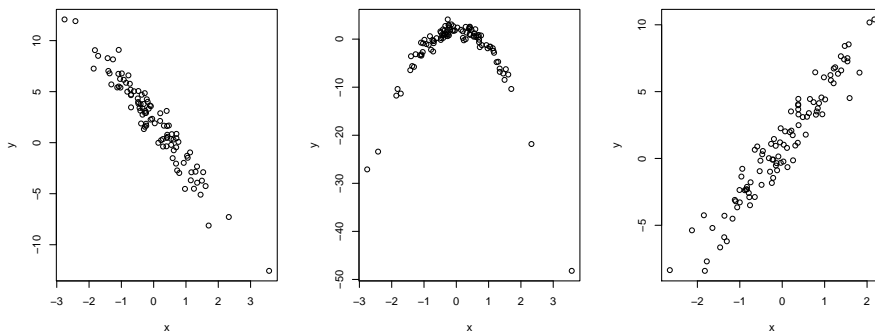


## TD n° 6 : Tests d'un ou plusieurs coefficients de corrélation

**Exercice 1.** On considère les trois nuages de points suivants, chacun étant associés aux données de 2 caractères  $X$  et  $Y$  :



Les trois coefficients de Pearson associés sont, dans un ordre aléatoire :  $r_1 = -0.1221756$ ,  $r_2 = -0.9644077$  et  $r_3 = 0.9602034$ . Associer à chaque nuage de points son coefficient de Pearson.

**Exercice 2.** Sur chacun des 11 spécimens fossiles d'un animal disparu, on dispose de la longueur de leur humérus et leur fémur. Les résultats, en centimètres, sont les suivants :

humérus	45	61	71	73	87	45	61	73	77	90	72
fémur	41	56	59	65	77	39	53	63	69	78	62

1. Tracer le nuage de points des données. Que peut-on en dire ?
2. Donner le coefficient de Pearson des données. Que peut-on en dire ?
3. Peut-on affirmer, avec un faible risque de se tromper, que la longueur de l'humérus et la longueur du fémur de l'animal sont liés ? Si oui, quel est le degré de significativité ?

**Exercice 3.** On relève les tailles en centimètres et les poids en kilogrammes de 15 personnes choisies au hasard. Les résultats sont les suivants :

taille	175	185	171	176	172	178	169	171	188	163	178	169	172	187	163
poids	71	85	69	75	71	81	65	68	86	61	78	65	66	92	55

1. Tracer le nuage de points des données. Que peut-on en dire ?
2. Donner le coefficient de Pearson des données. Que peut-on en dire ?
3. Peut-on prétendre, avec un faible risque de se tromper, que la taille et le poids d'une personne sont corrélées ? Si oui, quel est le degré de significativité ?

**Exercice 4.** Un sylviculteur veut tester une méthode pour mesurer de façon optique les hauteurs d'arbre. Pour cela, il prélève au hasard 8 arbres dont il mesure la hauteur, dans un premier temps, avec la méthode optique, puis au sol après abattage. L'unité est le mètre. Les résultats sont les suivants :

mesure optique	12.4	23.7	18.5	19.1	20.1	15.4	13.4	14.1
mesure au sol	9.2	22.3	15.4	17.3	18.4	12.2	11.7	12.6

Peut-on prétendre, avec un faible risque de se tromper, que les 2 méthodes sont corrélées ?

**Exercice 5.** Des candidats à un concours ont eut le choix de suivre 2 formations, notées A et B. Sur un échantillon de 30 candidats, on relève la formation choisie, la note sur 20 obtenue au contrôle final de la formation et la note sur 20 obtenue au concours. Les résultats sont les suivants :

note formation	12.5	13	8.5	9.5	6	11	15	12	10.5	13	8	9.5	13	8.5	15
note concours	13	14	9	9	10	11	16	12	8.5	14	8	10.5	13	11	18
formation	A	A	B	A	B	B	A	A	B	A	B	B	A	B	A

11	12.5	15	11	9.5	10.5	11.5	16	5.5	12	17	13.5	11	11.5	13
10	15	15.5	10	9	8.5	12.5	13	9.5	14	18	14	12	14.5	17
B	A	A	B	A	B	B	A	B	A	A	B	B	B	A

- Tracer un unique graphique affichant le nuage de points des données des 2 notes, avec les candidats ayant suivi la formation A avec des carrés verts et les candidats ayant suivi la formation B avec des triangles bleus, et une légende adaptée. Que peut-on en dire ?
- Donner le coefficient de Pearson des données associées aux 2 notes pour les candidats à la formation A, puis le coefficient de Pearson des données associées aux 2 notes pour les candidats à la formation B. Est-ce que les 2 valeurs obtenues vous semblent proches ?
- Peut-on prétendre, avec un faible risque de se tromper, que les 2 notes d'un candidat de la formation A et les 2 notes d'un candidat de la formation B sont corrélées différemment ? Si oui, quel est le degré de significativité ?

**Exercice 6.** Dans une grande surface, on s'intéresse à l'homogénéité des hommes et des femmes quant à la force du lien linéaire entre le (grand) nombre de produits achetés, caractère noté X, et la dépense en euros, caractère noté Y. Les résultats sont les suivants.

- Sur un échantillon de 54 hommes le coefficient de Pearson des données est de 0.35,
- Sur un échantillon de 31 femmes, le coefficient de Pearson des données est de 0.54.

Peut-on affirmer que les hommes et les femmes diffèrent significativement quant à la force du lien linéaire entre X et Y ?