

# PrivateTeacher

*Maîtriser les Sciences Exactes*

## STATISTIQUES

### Mesures d'association

### Série d'exercices

Julien RUPPEN

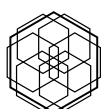
19 May, 2026

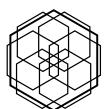
#### **Abstract**

Ce document aborde les mesures d'association à travers l'axe de la précision des variables disponibles. Selon le type de variables et les conditions respectées, on choisit la mesure la plus appropriée : covariance et corrélation de Pearson pour les variables quantitatives avec relation linéaire, corrélation de Spearman et tau de Kendall lorsque ces conditions ne sont pas remplies ou que les variables sont ordinales. Cette progression du plus précis au plus robuste permet à l'étudiant de comprendre non seulement comment calculer chaque mesure, mais surtout pourquoi et dans quel contexte la choisir.

# Contents

<b>1 Exercice 1 — Sommeil et attention</b>	<b>4</b>
1.1 Question de recherche . . . . .	4
1.2 Q1 — Choix de la mesure d'association . . . . .	4
1.3 Q2 — Graphique . . . . .	5
1.4 Q3 — Calcul de la covariance . . . . .	5
1.5 Q4 — Calcul du coefficient de corrélation de Pearson . . . . .	6
1.6 Q5 — Inférence . . . . .	7
1.7 Q6 — Interprétation . . . . .	7
<b>2 Exercice 2 — Sommeil et fatigue</b>	<b>8</b>
2.1 Question de recherche . . . . .	8
2.2 Q1 — Choix de la mesure d'association . . . . .	8
2.3 Q2 — Graphique . . . . .	9
2.4 Q3 — Calcul de la covariance . . . . .	9
2.5 Q4 — Calcul du coefficient de corrélation de Pearson . . . . .	10
2.6 Q5 — Inférence . . . . .	11
2.7 Q6 — Interprétation . . . . .	11
<b>3 Exercice 3 — Onset dépressif et sévérité actuelle</b>	<b>12</b>
3.1 Question de recherche . . . . .	12
3.2 Q1 — Choix de la mesure d'association . . . . .	12
3.3 Q2 — Graphique . . . . .	13
3.4 Q3 — Calcul de la corrélation de Spearman . . . . .	13
3.5 Q4 — Calcul du tau de Kendall . . . . .	14
3.6 Q5 — Inférence . . . . .	15
3.7 Q6 — Interprétation . . . . .	15
3.8 Q7 — Comparaison des deux mesures . . . . .	15
<b>4 Exercice 4 — Perception du risque et intention de vaccination</b>	<b>17</b>
4.1 Question de recherche . . . . .	17
4.2 Q1 — Choix de la mesure d'association . . . . .	17
4.3 Q2 — Graphique . . . . .	18
4.4 Q3 — Calcul de la corrélation de Spearman . . . . .	19
4.5 Q4 — Calcul du tau de Kendall . . . . .	19
4.6 Q5 — Inférence . . . . .	20
4.7 Q6 — Interprétation . . . . .	20
4.8 Q7 — Comparaison des deux mesures . . . . .	21





## 1. Exercice 1 — Sommeil et attention

Un chercheur souhaite savoir si le nombre d'heures de sommeil d'un étudiant est lié à sa performance cognitive le lendemain. Il recrute 6 étudiants en psychologie et mesure, pour chacun, le nombre d'heures de sommeil la nuit précédente ainsi que leur score à un test d'attention standardisé passé le matin.

Étudiant	Sommeil (h)	Score d'attention
1	7.5	84.5
2	6.5	61.2
3	6.1	72.5
4	7.7	75.6
5	6.6	72.9
6	5.5	66.8

### 1.1. Question de recherche

Existe-t-il un lien linéaire entre le nombre d'heures de sommeil et le score d'attention ?

### 1.2. Q1 — Choix de la mesure d'association

Étant donné la situation décrite, quelle mesure d'association est la plus adaptée ?

Trois cas de figure sont disponibles :

- **Cas 1** — Variables *quantitatives*, normalité et linéarité respectées.  
Mesure la plus adaptée : **Pearson** (via la covariance)
- **Cas 2** — Variables *quantitatives*, normalité ou linéarité **non** respectée.  
Mesure la plus adaptée : **Spearman** (Kendall possible pour comparaison)
- **Cas 3** — Variables *ordinales*.  
Mesure la plus adaptée : **Kendall** (Spearman possible pour comparaison)

Les Cas 1 et 2 ont les mêmes variables — c'est le respect des conditions qui les distingue. Les Cas 2 et 3 ne requièrent ni normalité ni linéarité — c'est le type de variable qui les distingue.

Dans quel cas te trouves-tu ? Justifie ta réponse.

---



---



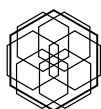
---



---



---



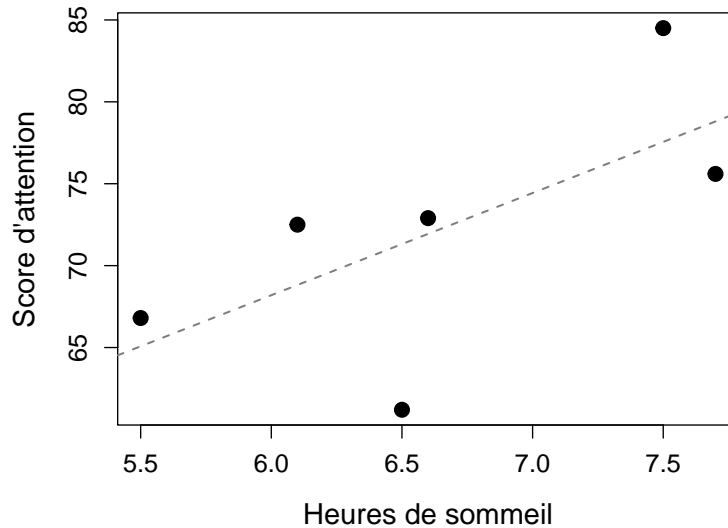
**1.3. Q2 — Graphique**

Figure 1: Diagramme de dispersion : heures de sommeil et score d'attention.

Décris ce que tu observes sur ce graphique. La relation entre les deux variables te semble-t-elle linéaire ? Positive ou négative ? Forte ou faible ?

---

---

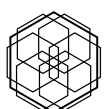
---

---

---

**1.4. Q3 — Calcul de la covariance**

Complète le tableau suivant, puis calcule la covariance  $S_{XY}$ .



1.5 Q4 — Calcul du coefficient de corrélation de Pearson **EXERCICE** SOMMEIL ET ATTENTION

$i$	$x_i$	$y_i$	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$
1	7.5	84.5			
2	6.5	61.2			
3	6.1	72.5			
4	7.7	75.6			
5	6.6	72.9			
6	5.5	66.8			
$\Sigma$					

$$\bar{x} = \frac{\sum x_i}{n} = \qquad \bar{y} = \frac{\sum y_i}{n} =$$

$$S_{XY} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n - 1} =$$

---



---



---

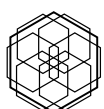
**1.5. Q4 — Calcul du coefficient de corrélation de Pearson**

Complète le tableau suivant pour obtenir les écarts-types de chaque variable, puis calcule  $r_{XY}$ .

$i$	$x_i$	$y_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$y_i - \bar{y}$	$(y_i - \bar{y})^2$
1	7.5	84.5				
2	6.5	61.2				
3	6.1	72.5				
4	7.7	75.6				
5	6.6	72.9				
6	5.5	66.8				
$\Sigma$						

$$s_X = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}} =$$

$$s_Y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n - 1}} =$$



$$r_{XY} = \frac{S_{XY}}{s_X \cdot s_Y} =$$

---

---

---

### 1.6. Q5 — Inférence

La statistique de test et la p-valeur associées au coefficient de Pearson sont :

$$t(4) = 1.743, \quad p = 0.156$$

Que concluons-nous sur la significativité de cette corrélation ? Le résultat te surprend-il compte tenu de la valeur de  $r$  ?

---

---

---

---

---

### 1.7. Q6 — Interprétation

Que signifie la valeur de  $r$  que tu viens de calculer ? Que dit-elle sur le lien entre le sommeil et l'attention ? La direction du lien est-elle cohérente avec ce que tu observes sur le graphique ?

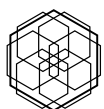
---

---

---

---

---



## 2. Exercice 2 — Sommeil et fatigue

Un chercheur souhaite savoir si le nombre d'heures de sommeil d'un étudiant est lié à son niveau de fatigue subjective le lendemain matin. Il recrute 6 étudiants en psychologie et mesure, pour chacun, le nombre d'heures de sommeil la nuit précédente ainsi que leur fatigue déclarée sur une échelle de 0 (pas du tout fatigué) à 100 (épuisé).

Étudiant	Sommeil (h)	Fatigue (0–100)
1	8.5	40.9
2	6.9	64.8
3	6.8	45.5
4	8.0	37.1
5	5.1	72.7
6	6.8	51.0

### 2.1. Question de recherche

Existe-t-il un lien linéaire entre le nombre d'heures de sommeil et le niveau de fatigue subjective ?

### 2.2. Q1 — Choix de la mesure d'association

Étant donné la situation décrite, quelle mesure d'association est la plus adaptée ?

Trois cas de figure sont disponibles :

- **Cas 1** — Variables *quantitatives*, normalité et linéarité respectées.  
Mesure la plus adaptée : **Pearson** (via la covariance)
- **Cas 2** — Variables *quantitatives*, normalité ou linéarité **non** respectée.  
Mesure la plus adaptée : **Spearman** (Kendall possible pour comparaison)
- **Cas 3** — Variables *ordinales*.  
Mesure la plus adaptée : **Kendall** (Spearman possible pour comparaison)

Les Cas 1 et 2 ont les mêmes variables — c'est le respect des conditions qui les distingue. Les Cas 2 et 3 ne requièrent ni normalité ni linéarité — c'est le type de variable qui les distingue.

Dans quel cas te trouves-tu ? Justifie ta réponse.

---



---



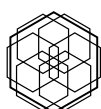
---



---



---



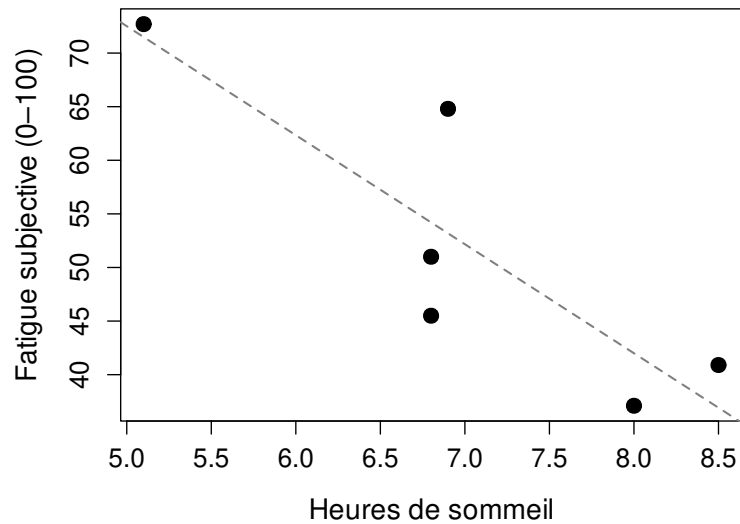
**2.3. Q2 — Graphique**

Figure 2: Diagramme de dispersion : heures de sommeil et fatigue subjective.

Décris ce que tu observes sur ce graphique. La relation entre les deux variables te semble-t-elle linéaire ? Positive ou négative ? Forte ou faible ?

---

---

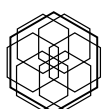
---

---

---

**2.4. Q3 — Calcul de la covariance**

Complète le tableau suivant, puis calcule la covariance  $S_{XY}$ .



2.5 Q4 — Calcul du coefficient de corrélation de Pearson — SOMMEIL ET FATIGUE

$i$	$x_i$	$y_i$	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$
1	8.5	40.9			
2	6.9	64.8			
3	6.8	45.5			
4	8.0	37.1			
5	5.1	72.7			
6	6.8	51.0			
$\Sigma$					

$$\bar{x} = \frac{\sum x_i}{n} = \quad \bar{y} = \frac{\sum y_i}{n} =$$

$$S_{XY} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n - 1} =$$

**2.5. Q4 — Calcul du coefficient de corrélation de Pearson**

Complète le tableau suivant pour obtenir les écarts-types de chaque variable, puis calcule  $r_{XY}$ .

$i$	$x_i$	$y_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$y_i - \bar{y}$	$(y_i - \bar{y})^2$
1	8.5	40.9				
2	6.9	64.8				
3	6.8	45.5				
4	8.0	37.1				
5	5.1	72.7				
6	6.8	51.0				
$\Sigma$						

$$s_X = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}} =$$

$$s_Y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n - 1}} =$$



$$r_{XY} = \frac{S_{XY}}{s_X \cdot s_Y} =$$

---

---

---

### 2.6. Q5 — Inférence

La statistique de test et la p-valeur associées au coefficient de Pearson sont :

$$t(4) = -3.311, \quad p = 0.03$$

Que concluons-nous sur la significativité de cette corrélation ? Compare ce résultat avec celui de l'exercice précédent.

---

---

---

---

---

### 2.7. Q6 — Interprétation

Que signifie la valeur de  $r$  que tu viens de calculer ? Que dit-elle sur le lien entre le sommeil et la fatigue ? Que signifie le signe négatif ?

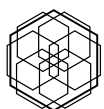
---

---

---

---

---



### 3. Exercice 3 — Onset dépressif et sévérité actuelle

Un clinicien souhaite savoir si l'âge auquel un patient a présenté ses premiers symptômes dépressifs est lié à la sévérité actuelle de son trouble. Il dispose des données de 5 patients, mesurées à l'aide de l'échelle de Hamilton (HAM-D, de 0 à 52).

Patient	Âge onset (ans)	HAM-D actuel
1	13	26
2	15	18
3	22	21
4	30	14
5	38	11

#### 3.1. Question de recherche

Existe-t-il un lien monotone entre l'âge des premiers symptômes dépressifs et la sévérité actuelle du trouble ?

#### 3.2. Q1 — Choix de la mesure d'association

Étant donné la situation décrite, quelle mesure d'association est la plus adaptée ?

Trois cas de figure sont disponibles :

- **Cas 1** — Variables *quantitatives*, normalité et linéarité respectées.  
Mesure la plus adaptée : **Pearson** (via la covariance)
- **Cas 2** — Variables *quantitatives*, normalité ou linéarité **non** respectée.  
Mesure la plus adaptée : **Spearman** (Kendall possible pour comparaison)
- **Cas 3** — Variables *ordinales*.  
Mesure la plus adaptée : **Kendall** (Spearman possible pour comparaison)

Les Cas 1 et 2 ont les mêmes variables — c'est le respect des conditions qui les distingue. Les Cas 2 et 3 ne requièrent ni normalité ni linéarité — c'est le type de variable qui les distingue.

Dans quel cas te trouves-tu ? Justifie ta réponse.

---



---



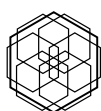
---



---



---



### 3.3. Q2 — Graphique

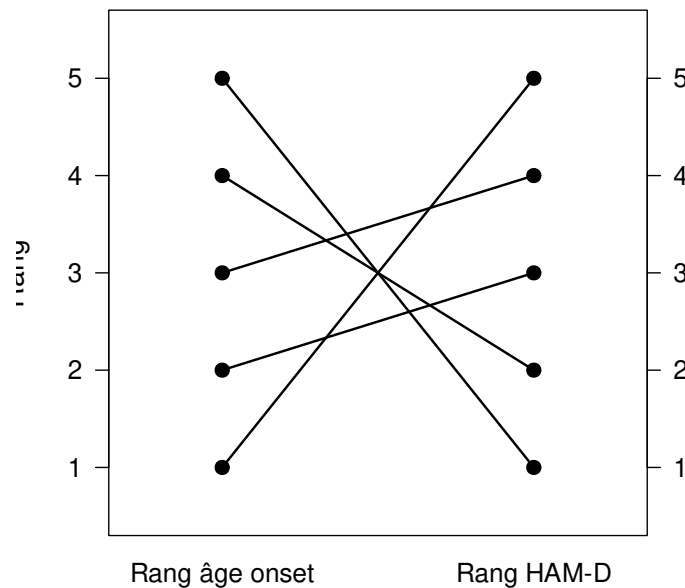


Figure 3: Graphique des rangs : âge d'onset et sévérité HAM-D. Chaque ligne représente un patient. Un croisement entre deux lignes correspond à une paire discordante.

Décris ce que tu observes. Les lignes sont-elles majoritairement parallèles ou se croisent-elles souvent ? Que signifient les croisements ?

---



---



---



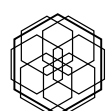
---



---

### 3.4. Q3 — Calcul de la corrélation de Spearman

Complète le tableau suivant en calculant les rangs, les différences et leur carré.



3.5 Q4 — Calcul du tau de Kendall

$i$	$x_i$	$y_i$	$\text{rang}(x_i)$	$\text{rang}(y_i)$	$d_i$	$d_i^2$
1	13	26				
2	15	18				
3	22	21				
4	30	14				
5	38	11				
$\Sigma$						

$$r_{Sp} = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} =$$

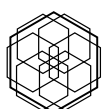
3.5. Q4 — Calcul du tau de Kendall

Pour chaque paire de patients  $(i, j)$  avec  $i < j$ , indique le sens de la différence sur X, le sens de la différence sur Y, puis si la paire est concordante (C) ou discordante (D).

Note : les patients sont numérotés par ordre croissant d'âge d'onset. Pour toutes les paires, le sens sur X est donc + (le patient  $j$  est toujours plus âgé que le patient  $i$ ).

Paire	Sens X	Sens Y	Résultat
(1, 2)			
(1, 3)			
(2, 3)			
(1, 4)			
(2, 4)			
(3, 4)			
(1, 5)			
(2, 5)			
(3, 5)			
(4, 5)			

$$\tau = \frac{C - D}{C + D} =$$



3.6 Q5 — Inférence **EXERCICE 3 — ONSET DÉPRESSIF ET SÉVÉRITÉ ACTUELLE**

---

---

---

**3.6. Q5 — Inférence**

La statistique de test et la p-valeur associées à la corrélation de Spearman sont :

$$t(3) = -3.576, \quad p = 0.037$$

Que concluons-nous sur la significativité de cette corrélation ?

---

---

---

---

---

**3.7. Q6 — Interprétation**

Que signifient les valeurs de  $r_{Sp}$  et de  $\tau$  calculées ? Que disent-elles sur le lien entre l'âge d'onset et la sévérité actuelle ? Les deux mesures convergent-elles ?

---

---

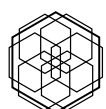
---

---

---

**3.8. Q7 — Comparaison des deux mesures**

Cette section a calculé à la fois la corrélation de Spearman et le tau de Kendall. Compare les deux mesures en répondant aux points suivants :



3.8 Q7 — Comparaison de deux mesures EXERCICE 3 — ONSET DÉPRESSIF ET SÉVÉRITÉ ACTUELLE

1. Les valeurs obtenues sont-elles proches ? Dans quel sens diffèrent-elles ?
2. Les deux mesures sont-elles significatives au seuil  $\alpha = .05$  ?
3. Sur quoi repose le calcul de chacune ? En quoi les bases de calcul sont-elles différentes ?
4. Laquelle des deux est la plus conservatrice ? Pourquoi est-ce attendu en général ?

---

---

---

---

---

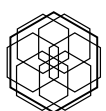
---

---

---

---

---



## 4. Exercice 4 — Perception du risque et intention de vaccination

Une chercheuse s'intéresse à la relation entre la perception du risque de contracter la grippe saisonnière et l'intention de se faire vacciner. Elle interroge 5 adultes à l'aide de deux échelles ordinales en cinq niveaux : la perception du risque (1 = très faible, 5 = très élevée) et l'intention de vaccination (1 = certainement pas, 5 = certainement oui).

Participant	Risque perçu (X)	Intention vaccination (Y)
1	1	1
2	2	3
3	3	2
4	4	4
5	5	5

### 4.1. Question de recherche

Une perception plus élevée du risque de contracter la grippe est-elle associée à une plus grande intention de se faire vacciner ?

### 4.2. Q1 — Choix de la mesure d'association

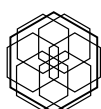
Étant donné la situation décrite, quelle mesure d'association est la plus adaptée ?

Trois cas de figure sont disponibles :

- **Cas 1** — Variables *quantitatives*, normalité et linéarité respectées.  
Mesure la plus adaptée : **Pearson** (via la covariance)
- **Cas 2** — Variables *quantitatives*, normalité ou linéarité **non** respectée.  
Mesure la plus adaptée : **Spearman** (Kendall possible pour comparaison)
- **Cas 3** — Variables *ordinales*.  
Mesure la plus adaptée : **Kendall** (Spearman possible pour comparaison)

Les Cas 1 et 2 ont les mêmes variables — c'est le respect des conditions qui les distingue. Les Cas 2 et 3 ne requièrent ni normalité ni linéarité — c'est le type de variable qui les distingue.

Dans quel cas te trouves-tu ? Justifie ta réponse.



#### 4.3 EXERCICE Graphique PERCEPTION DU RISQUE ET INTENTION DE VACCINATION

---

---

---

---

---

#### 4.3. Q2 — Graphique

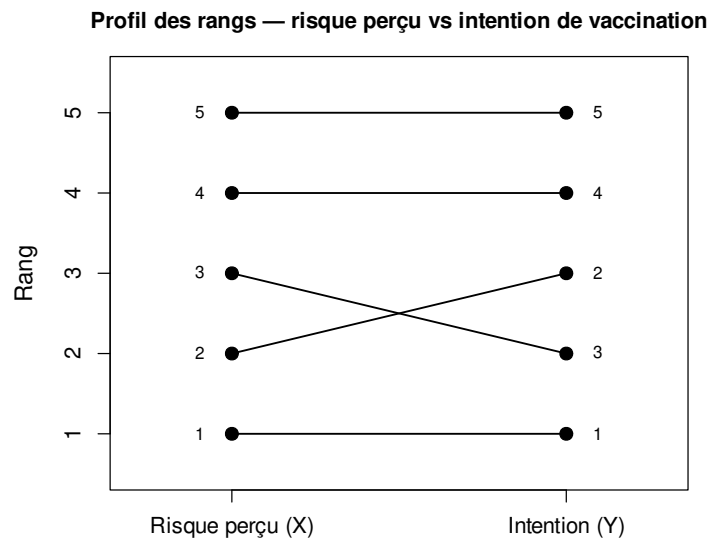


Figure 4: Profil des rangs : risque perçu et intention de vaccination. Chaque ligne représente un participant. Un croisement entre deux lignes correspond à une paire discordante.

Décris ce que tu observes. Les lignes sont-elles majoritairement parallèles ou se croisent-elles souvent ? Que signifient les croisements ?

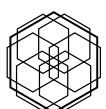
---

---

---

---

---



### 4.4. Q3 — Calcul de la corrélation de Spearman

Complète le tableau suivant en calculant les rangs, les différences et leur carré.

$i$	Risque ( $x_i$ )	Intention ( $y_i$ )	$\text{rang}(x_i)$	$\text{rang}(y_i)$	$d_i$	$d_i^2$
1	1	1				
2	2	3				
3	3	2				
4	4	4				
5	5	5				
$\Sigma$						

$$r_{Sp} = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} =$$

---



---



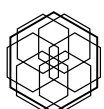
---

### 4.5. Q4 — Calcul du tau de Kendall

Pour chaque paire de participants  $(i, j)$  avec  $i < j$ , indique le sens de la différence sur X, le sens de la différence sur Y, puis si la paire est concordante (C) ou discordante (D).

Note : les participants sont numérotés par risque perçu croissant. Pour toutes les paires, le sens sur X est donc + (le participant  $j$  a toujours un risque perçu plus élevé que le participant  $i$ ).

Paire	Sens X	Sens Y	Résultat
(1, 2)			
(1, 3)			
(2, 3)			
(1, 4)			
(2, 4)			
(3, 4)			
(1, 5)			
(2, 5)			
(3, 5)			
(4, 5)			



#### 4.6 EXERCICE 4 — PERCEPTION DU RISQUE ET INTENTION DE VACCINATION

$$\tau = \frac{C - D}{C + D} =$$

---

---

---

#### 4.6. Q5 — Inférence

La statistique de test et la p-valeur associées au tau de Kendall sont :

$$z = 1.96, \quad p = 0.05$$

Que concluons-nous sur la significativité de cette corrélation ?

---

---

---

---

---

#### 4.7. Q6 — Interprétation

Que signifient les valeurs de  $\tau$  et de  $r_{Sp}$  calculées ? Que disent-elles sur le lien entre la perception du risque et l'intention de vaccination ?

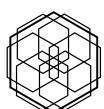
---

---

---

---

---



### 4.8. Q7 — Comparaison des deux mesures

Cette section a calculé à la fois le tau de Kendall et la corrélation de Spearman. Compare les deux mesures en répondant aux points suivants :

1. Les valeurs obtenues sont-elles proches ? Dans quel sens diffèrent-elles ?
2. Les deux mesures sont-elles significatives au seuil  $\alpha = .05$  ?
3. Sur quoi repose le calcul de chacune ? En quoi les bases de calcul sont-elles différentes ?
4. Laquelle des deux est la plus conservatrice ? La règle approximative  $\tau \approx \frac{2}{3} \times r_{Sp}$  est-elle vérifiée ici ?

---

---

---

---

---

---

---

---

---

---

