
UE 4 : Évaluation des méthodes d'analyse appliquées aux sciences de la vie et de la santé

Enseignements dirigés



Université Henri Poincaré

Année universitaire 2011–2012

Sommaire

ED1	3
– Fonctions d’une variable réelle	
ED2	6
– Fonctions réelles à une ou plusieurs variables réelles	
ED3	8
– Fonctions réelles de plusieurs variables réelles	
ED4	11
– Intérêts des Statistiques et de la Statistique en Santé.	
– Mesures, généralités	
– Probabilités	
– Les Lois de Probabilités	
– Autres Lois de Probabilités	
ED5	15
– La Statistique Descriptive	
– Variable aléatoire, estimation ponctuelle et par intervalle	
ED6	18
– Échantillonnage	
– Les Tests Statistiques	
– Adaptation d’une distribution expérimentale	
– Conformité d’une distribution expérimentale à une distribution théorique	
– Association entre caractères qualitatifs	
– Comparaison entre plusieurs distributions observées	
ED7	22
– Comparaison de deux variances	
– Comparaison de deux moyennes	
ED8	25
– Comparaison de deux pourcentages	
– Régression ou corrélation	
– Épidémiologie : types d’enquêtes	
Annexes : tables statistiques et papiers gradués.....	31

1 ED 1 : Fonctions réelles d'une variable réelle

1.1 Exercices

Exercice 1

Calculer

$$\sin\left(\frac{17\pi}{3}\right), \quad \cos\left(-\frac{19\pi}{4}\right), \quad \tan\left(\frac{19\pi}{6}\right), \quad \sin\left(-\frac{9\pi}{2}\right), \quad \arccos\left(\frac{1}{2}\right).$$

Exercice 2

Résoudre les équations :

$$\begin{aligned} 2 \ln(x-1) &= 4, \\ \exp(x+3) \exp(x+1) &= e^{-1}, \\ e^{2x} - 1 &= 0, \\ e^x + 4 &= \frac{2}{e^x}, \\ \sin(3x) &= \frac{1}{2}, \\ \tan(4x) &= 2, \\ \tan(4x) &= 1. \end{aligned}$$

Exercice 3

Étudier les fonctions définies par les formules suivantes :

1. $f(x) = \sqrt{2x+3}$.
2. $g(x) = \ln(1 + \sin x)$.
3. $h(x) = \frac{2x-1}{x+1}$.
4. $k(x) = e^{\cos(x)}$.

1.2 QCM

1. On considère le nombre réel $u = \sin\left(\frac{19\pi}{3}\right)$.

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. $|u| \leq 1$
- B. $u = \sin\left(\frac{\pi}{3}\right)$
- C. $u = \frac{1}{2}$
- D. u est positif
- E. u est un nombre imaginaire pur

2. On considère l'équation $(\star) : 2 \sin x = 1$.

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. Le nombre 0 est solution de l'équation (\star)
- B. $\frac{\pi}{6}$ est une solution de l'équation (\star)
- C. $-\frac{\pi}{6}$ est une solution de l'équation (\star)
- D. Toutes les solutions de l'équation (\star) sont les nombres $\frac{\pi}{6} + 2k\pi$ où k est un entier relatif
- E. Les solutions de l'intervalle $[0, 2\pi]$ sont $\frac{\pi}{6}$ et $\frac{5\pi}{6}$

3. Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. $\ln(2a) = \ln a^2$ pour tout réel strictement positif a
- B. $\ln(2 \exp a) = \ln 2 + a$ pour tout réel a
- C. $\ln(2a) = \ln 2 + \ln a$ pour tout réel strictement positif a
- D. $\ln(|2a|) = \ln 2 + \ln |a|$ pour tout réel a
- E. $\ln(e^a) = a$ pour tout réel a

4. On considère l'équation $(\star) : \ln(x - 2) + \ln(x + 2) = \ln 5$.

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. Le réel 3 est solution de l'équation (\star)
- B. Le réel -3 est solution de l'équation (\star)
- C. Puisque $\ln(x - 2) + \ln(x + 2) = \ln((x - 2)(x + 2)) = \ln(x^2 - 4)$, alors l'équation (\star) est équivalente à $\ln(x^2 - 4) = \ln 5$ pour n'importe quel réel x .
- D. L'équation (\star) se réécrit $\ln(x^2 - 4) = \ln 5$ pour tout réel x qui vérifie les conditions

$$x + 2 > 0$$

$$x - 2 > 0$$

- E. Toute solution de l'équation $x^2 - 4 = 5$ est solution de l'équation (\star) .

5. On cherche à préciser les paramètres a et λ d'une fonction f définie par $f(x) = a \exp(\lambda x)$. On sait que $f(0) = 2$ et que la dérivée de la fonction $\ln(f)$ vaut 5 en 0.
Précisez pour chaque affirmation si elle est juste ou fausse.
- A. Pour tout réel x , $f(x) = 5e^{2x}$
 - B. Pour tout réel x , $f(x) = 2e^{5x}$
 - C. f n'est pas définie pour $x < 0$
 - D. La dérivée de f vérifie $f'(0) = 0$.
 - E. La dérivée de la fonction $\ln(f)$ est constante
6. On considère l'équation $e^{2x+1} = 3$.
Indiquez pour chaque affirmation si elle est juste ou fausse.
- A. Cette équation n'a pas de solution réelle.
 - B. $x = 1$ est l'unique solution de cette équation.
 - C. L'unique solution de cette équation s'écrit $x = \frac{\ln 3 - 1}{2}$
 - D. L'unique solution de cette équation s'écrit $x = \ln 3 - 2$
 - E. Cette équation n'a pas de sens pour les x tels que $2x + 1 < 0$.
7. On considère une fonction $\varphi(x)$ dont la dérivée vaut $\varphi'(x) = e^x - 1$
Indiquez pour chaque affirmation si elle est juste ou fausse.
- A. φ est croissante sur \mathbb{R}
 - B. φ admet un minimum en $x = 0$
 - C. φ admet un maximum en $x = 0$
 - D. $\varphi(x)$ est nulle pour tout x
 - E. On peut déterminer $\varphi(0)$

2 ED 2 : Fonctions réelles à une ou plusieurs variables réelles

2.1 Exercices

Exercice 1

Déterminer le développement limité en 0 à l'ordre indiqué des fonctions suivantes :

$$\begin{aligned}
 (a) \quad f(x) &= x(x-2)(x-3) && \text{(ordre 2),} \\
 (b) \quad f(x) &= e^x(1+2x) && \text{(ordre 2),} \\
 (c) \quad f(x) &= \frac{\ln(1+x)}{x} && \text{(ordre 3),} \\
 (d) \quad f(x) &= x\sqrt{1+x} && \text{(ordre 2),} \\
 (e) \quad f(x) &= \sqrt{2+x} && \text{(ordre 2),} \\
 (f) \quad f(x) &= \frac{\sin x}{x} && \text{(ordre 5).} \\
 (g) \quad f(x) &= \frac{e^x}{x} && \text{(ordre 5).} \\
 (h) \quad f(x) &= \frac{\sin x}{x} + \frac{e^x}{x} && \text{(ordre 5).}
 \end{aligned}$$

Exercice 2

Calculer les limites des fonctions suivantes lorsque $x \rightarrow 0$:

$$\begin{aligned}
 (a) \quad f(x) &= \frac{\sin x}{x}, \\
 (b) \quad f(x) &= \frac{\sin(x)-x}{x^k}, \\
 (c) \quad f(x) &= \frac{\sin(x)-x}{x^k} && \text{pour } k \geq 2, \\
 (d) \quad f(x) &= \exp(\cos(x)), \\
 (e) \quad f(x) &= \exp(\sin(x)), \\
 (f) \quad f(x) &= \frac{\sin(x)-x}{\cos(x)-x}.
 \end{aligned}$$

Exercice 3

Calculer les dérivées partielles d'ordre un et deux de la fonction f définie sur \mathbb{R}^2 par :

$$f(x,y) = x^2y^3 + 3x^3y - 2x^2y.$$

À partir de quel ordre les dérivées partielles de f sont-elles toutes nulles ?

2.2 QCM

- On considère une fonction f dont le développement limité en 0 est $f(x) = 1 - x + x^2 + o(x^2)$.
Indiquez pour chaque affirmation si elle est juste ou fausse.
 - $f(0) = 0$
 - $f(0) = 1$
 - $f'(0) = -1$
 - $f''(0) = 1$
 - On peut déduire de ce développement la valeur exacte de $f(1)$

2. On considère la fonction f définie sur $] - \infty, 1[$ par : $f(x) = \ln(1 - x)$.

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. Le développement limité de f en 0 à l'ordre zéro est : $-x + o(x)$
- B. Le développement limité de f en 0 à l'ordre un est : $-x + o(x)$
- C. Le développement limité de f en 0 à l'ordre deux est : $-x - x^2 + o(x^2)$
- D. Le développement limité de f en 0 à l'ordre deux est : $x - x^2 + o(x^2)$
- E. On peut calculer $f(2)$ grâce au développement limité

3. On considère la fonction dite logistique f définie sur $] - \infty, \infty[$ par : $f(x) = \frac{1}{1 + \exp(-x)}$.

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. f prend la valeur 1 en 0
- B. Pour tout réel x , $f(x) = \frac{\exp(x)}{1 + \exp(x)}$
- C. f ne prend que des valeurs positives
- D. f ne prend que des valeurs inférieures à 1
- E. Pour tout réel x de l'intervalle $]0, 1[$, $f\left(\ln\left(\frac{x}{1-x}\right)\right) = x$

4. On souhaite résoudre l'équation $f(x) = y$ avec $y \in \mathbb{R}$. On note $g(y)$ la fonction réciproque (si elle existe).

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. La fonction réciproque $g(y)$ est définie par

$$x \text{ unique solution de } f(x) = y \Leftrightarrow x = g(y)$$

- B. Le domaine de définition de $g(y)$ est l'ensemble des valeurs y telles que l'équation $f(x) = y$ admet une solution unique
- C. Lorsque l'équation $f(x) = y$ admet plusieurs solutions, on ne peut pas déterminer de fonction réciproque en ce point y
- D. On sait $x = \text{Arccos } y$
- E. On écrit $x = \frac{y}{f}$

5. On considère la fonction f définie sur $] - 1, \infty[$ par : $f(x) = \ln(1 + 2x)$, de dérivées première et seconde vérifiant : $f'(x) = \frac{2}{1 + 2x}$, $f''(x) = \frac{-4}{(1 + 2x)^2}$.

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. Le développement limité de f en 0 à l'ordre un est : $1 + 2x + o(x)$
- B. Le développement limité de f en 0 à l'ordre deux est : $1 + 2x - x^2 + o(x^2)$
- C. Le développement limité de f en 0 à l'ordre deux est : $2x - 2x^2 + o(x^2)$
- D. Le développement limité de f en 0 n'existe pas
- E. On a toujours :

$$\begin{aligned} f(x) = f(0) &+ f'(0)x \\ &+ \frac{f''(0)}{2}x^2 \\ &+ o(x^2). \end{aligned}$$

3 ED 3 : fonctions réelles de plusieurs variables réelles

3.1 Exercices

Exercice 1

Calculer les dérivées partielles d'ordres un et deux de la fonction f définie sur $\mathbb{R}^2 \setminus \{0,0\}$ par :

$$f(x,y) = \exp\left(\frac{x}{x^2 + y^2}\right).$$

Exercice 2

Trouver les points critiques des fonctions de deux variables suivantes et déterminer ceux qui sont des maxima ou des minima de la fonction.

1. $f(x,y) = xy$,
2. $f(x,y) = 4xy - x^2 + y^2$,
3. $f(x,y) = x^2 - (y - 1)^2$,
4. $f(x,y) = x^2 + (2y - 1)^2$,
5. $f(x,y) = x^2 + y^2 - x(y - 1)$,
6. $f(x,y) = (2x - 1)y + xy + 2x^2 + 3y^2$,
7. $f(x,y) = x \cos(y)$,
8. $f(x,y) = \exp(x^2 + y^2)$,
9. $f(x,y) = x y^2$.

Exercice 3

Le module de la résultante des forces exercées sur un barrage par l'eau retenue est $F = \rho g a h^2 / 2$, où ρ est la masse volumique de l'eau, a et h la longueur et la hauteur du barrage, et g l'accélération de la pesanteur.

Calculer l'incertitude relative, puis absolue, sur F pour les valeurs suivantes :

$$\begin{aligned}\rho &= 1,000 \pm 0,005 \text{ kg/m}^3 \\ g &= 9,809 \pm 0,001 \text{ m/s}^2 \\ a &= 35,20 \pm 0,01 \text{ m} \\ h &= 28,50 \pm 0,01 \text{ m}.\end{aligned}$$

3.2 QCM

1. On considère la fonction f définie sur \mathbb{R}^2 par : $f(x,y) = x^2 + y^2 + 4x - 2y$.

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. $\frac{\partial f}{\partial x}(x,y) = 2x$
- B. $\frac{\partial f}{\partial x}(x,y) = 2x + 4 - 2y$
- C. $\frac{\partial f}{\partial x}(x,y) = 2x + 4$
- D. $\frac{\partial^2 f}{\partial x \partial y}(x,y) = 2$
- E. $\frac{\partial^2 f}{\partial x^2}(x,y) = 2$

2. On considère la fonction f définie sur \mathbb{R}^2 par : $f(x,y) = e^{x^2+y^2}$, dont on peut calculer que $\frac{\partial f}{\partial x}(x,y) = 2x e^{x^2+y^2}$ et $\frac{\partial f}{\partial y}(x,y) = 2y e^{x^2+y^2}$.

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. Le point $(0,0)$ est l'unique point critique de f
- B. $(2x,2y)$ est un point critique de f pour toute valeur de x et y
- C. La fonction f n'admet aucun point critique
- D. Les points critiques sont les points où $s^2 - rt = 0$ avec $r = \frac{\partial^2 f}{\partial x^2}(x,y)$, $s = \frac{\partial^2 f}{\partial x \partial y}(x,y)$, et $t = \frac{\partial^2 f}{\partial y^2}(x,y)$
- E. Le point $(1,1)$ est un point critique de f

3. On considère la fonction f définie sur \mathbb{R}^2 par : $f(x,y) = x^2 + y^2 - xy$.

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. Tous les points de la droite $y = 2x$ sont des points critiques
- B. Tous les points de la droite $y = \frac{x}{4}$ sont des points critiques
- C. Le point $(0,0)$ est l'unique point critique
- D. Le point $(1,1)$ est l'unique point critique
- E. La fonction f n'admet aucun point critique

4. On note par f la densité de la loi normale standard bidimensionnelle ; cette fonction est définie sur

\mathbb{R}^2 par $f(x,y) = \frac{\exp\left(-\frac{x^2+y^2}{2}\right)}{2\pi}$, de dérivées partielles vérifiant $\frac{\partial f}{\partial x}(x,y) = -x f(x,y)$ et $\frac{\partial f}{\partial y}(x,y) = -y f(x,y)$.

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. La fonction f ne s'annule nulle part
- B. Le point $(0,0)$ est un point critique
- C. Le point $(1, -1)$ est un point critique
- D. Le point $(0,0)$ est l'unique point critique
- E. Tous les points des droites d'équations $y = 0$ et $x = 0$ sont des points critiques

5. On considère une fonction f définie sur \mathbb{R}^2 dont les dérivées partielles sont les suivantes : $\frac{\partial f}{\partial x}(x,y) = x(x-1)$ et $\frac{\partial f}{\partial y}(x,y) = y(y-2)$.

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. La fonction f n'admet aucun point critique
 - B. Le point $(0,0)$ est le seul point critique
 - C. La fonction f admet deux points critiques
 - D. La fonction f admet trois points critiques
 - E. La fonction f admet quatre points critiques
6. On considère une fonction f définie sur \mathbb{R}^2 dont les dérivées partielles à l'origine vérifient :
- $$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = \frac{\partial^2 f}{\partial y \partial x}(0,0) = 0 \text{ et}$$

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(0,0) &= 1 \\ \frac{\partial^2 f}{\partial y^2}(0,0) &= -1. \end{aligned}$$

Indiquez pour chaque affirmation si elle est juste ou fausse.

- A. La fonction f n'admet aucun point critique
- B. Le point $(0,0)$ est un point critique
- C. Le point $(0,0)$ est un maximum local
- D. Le point $(0,0)$ est un minimum local
- E. Le point $(0,0)$ est un point selle

4 ED 4 : Généralités, probabilités, lois de probabilité

4.1 Exercices

Exercice 1

Dans un service hospitalier, la charge de travail de l'année, des infirmières dépend de l'âge moyen des patients. Vous voulez étudier cette charge de travail dans un hôpital comptant 3 services. Dans le système d'information de l'hôpital vous pouvez disposer des éléments suivants :

- du nombre patients différents ayant fréquenté l'hôpital et de l'âge de chaque patient,
 - du nombre de séjours réalisés dans l'hôpital, répartis par service et pour chaque séjour de l'âge du patient calculé à partir de sa date de naissance et de sa date d'entrée à l'hôpital,
 - du nombre de pathologies différentes qui ont été accueillies, réparties par service et avec pour chacune d'elle l'âge moyen des patients.
- Quelle unité statistique et quels éléments (a, b ou c) allez-vous utiliser ? Justifiez votre réponse.
 - Faites-vous une étude exhaustive (recensement) ou un sondage ? Justifiez votre réponse.
 - Parmi les 3 services de l'hôpital, il y a un service de petite enfance qui accueille les enfants de 0 à 1 an et un service de gériatrie. Quelle précision allez-vous utiliser dans chacun de ces 2 services ?

Exercice 2

Une personne fait une tentative en ingérant une seule boîte de médicament. Elle avait à sa disposition 2 boîtes de médicament A, 1 boîte de B et 1 boîte de C. La probabilité de mourir si elle a ingéré A est de 0,1, de 0,2 si elle a ingéré B et de 0,3 si elle a ingéré C. On dispose d'un test biologique qui est positif dans 90% des cas quand on risque de mourir et négatif dans 85% des cas quand on ne risque pas de mourir.

- Quelle sont les probabilités que la personne ait ingéré A, B, C ?
- Quelle est la probabilité de mourir de cette personne ?
- Le test est positif, quelle est la probabilité de mourir ?
- Quelles sont les valeurs des rapports de vraisemblance du test ?

Exercice 3

Une variable aléatoire continue admet pour loi de densité de probabilité la fonction $f(x)$ telle que :

$$\begin{aligned} f(x) &= 0 & \forall x < 0 \\ f(x) &= 2 - ax & \text{pour } x \in [0, 2] \\ f(x) &= 0 & \forall x > 2 \end{aligned}$$

- Déterminer la valeur de a pour que $f(x)$ soit une fonction densité de probabilité.
- La loi ainsi obtenue est-elle une loi de densité de probabilité ?

Exercice 4

On lance 360 fois un dé à 6 faces, non pipé.

- A quelle distribution théorique obéit le nombre "6" obtenu ?
- Calculer la moyenne et l'écart-type de cette distribution.

Exercice 5

Si la probabilité pour qu'un individu ait une mauvaise réaction à l'injection d'un certain sérum est de 0,001, déterminer la probabilité pour que, sur 2 000 individus :

1. trois aient une mauvaise réaction,
2. plus de deux aient une mauvaise réaction.

Exercice 6

Une cantine sert des repas en très grand nombre. Chaque repas comporte une ration de viande. Les poids des rations de viande se répartissent suivant la loi normale, autour de la valeur moyenne $m = 117 g$ avec un écart-type de $15 g$.

1. Quelle est la probabilité pour que le poids d'une ration de viande soit compris entre $105 g$ et $132 g$?
2. Le 1^{er} mars, la cantine a servi 462 repas.
À combien peut-on évaluer le nombre de rations de viande dont le poids dépassait $135 g$?

4.2 QCM

1. Cochez si vrai pour chacune des propositions suivantes :
 - A. L'erreur de mesure est constituée uniquement de la précision ou incertitude de la mesure.
 - B. L'erreur systématique dans une mesure correspond à une erreur de calibration (tare d'une balance).
 - C. L'erreur totale de mesure est la somme de la précision, de la dispersion statistique et de l'erreur systématique.
 - D. L'utilisation d'instrument de mesure numérique à la place d'instrument analogique fait qu'il n'y a plus d'erreur de mesure.
 - E. Les différentes sources d'erreurs de mesure s'annulent toujours.
2. Cochez si vrai pour chacune des propositions suivantes :
 - A. En santé, la variabilité est due uniquement à l'erreur de mesure.
 - B. La variabilité de mesures répétées chez le même individu est la variabilité inter-individus.
 - C. La variabilité de mesures effectuées chez différents individus d'un même groupe est la variabilité inter-individus.
 - D. La variabilité totale est égale à :
(Variabilité inter-individus + Variabilité intra-individus) - Variabilité de la mesure.
 - E. Quand on mesure la tension artérielle plusieurs fois dans la journée, à plusieurs heures d'intervalle, chez le même individu, la variabilité observée est la somme de la variabilité intra-individu et de la variabilité de mesure.
3. Dans un article vous lisez : « L'objectif de cette étude, réalisée dans 2 centres est d'étudier les relations manifestations allergiques pulmonaires et profession dans la population générale. Dans un centre le diagnostic d'allergie est porté sur l'examen clinique, dans le second sur des examens biologique. Le recrutement a été effectué dans les services de pneumologie des 2 centres. Tous les patients hospitalisés pour manifestations allergiques pulmonaires de l'année 2010 ont été interrogés.»
Cochez si vrai pour chacune des propositions suivantes :
 - A. Il n'y a pas de biais dans cette étude.
 - B. Il y a un biais de caractérisation ou d'information dans cette étude.
 - C. Il y a un biais de sélection dans cette étude.
 - D. On ne doit pas craindre de facteur de confusion dans cette étude.
 - E. Pour l'année 2010, pour les 2 services c'est un recensement.

4. vous lisez : « L'étude des caractéristiques diagnostiques de la bandelette dans le diagnostic de l'infection du liquide d'ascite a été réalisée par une étude cas-témoin. La lecture de la bandelette donne comme résultat : Négatif, trace, +, ++, +++ . On a utilisé comme seuil de positivité du test diagnostique une lecture donnant un résultat supérieur ou égal à +. Le tableau ci-dessous donne les effectifs de chacune des situations : »

	Infection	Pas d'infection	Total
Test +	90	15	105
Test -	10	85	95
Total	100	100	200

Cocher si Vrai

- A. La lecture de la bandelette est une variable ordinale.
- B. Avec le tableau et les données fournies, on peut calculer les valeurs prédictives positive et négative du test
- C. La sensibilité du test vaut $90/100$. La spécificité du test vaut $90/105$.
- D. La prévalence de l'infection du liquide d'ascite dans la population est de $100/200 = 0,5 = 50\%$
- E. Quel que soit le seuil que l'on utilise les valeurs de sensibilité et de spécificité du test seront identiques.
5. Dans un article vous lisez « L'objectif de cette étude est de déterminer la survie des femmes atteintes d'un cancer du sein avec métastases suivies dans un centre de lutte contre le cancer. . . . Les résultats de l'étude de survie sont fournis dans le tableau ci-dessous. . . . »

Délai en années	Exposées	DCD	Tx instantané de décès	Tx instantané de survie	Tx cumulé de survie
1	200	0	0	1	1
...					
5	50	10	A	B	0,3
6	C	4	D	E	F

A, B, C, D, E sont des valeurs que vous devez calculer. Cochez les proposition exactes

- A. $A = 0,2$; $B = 0,8$; $C = 40$; $D = 0,1$; $E = 0,9$; $F = 0,27$
- B. $A = 0,2$; $B = 0,8$; $C = 40$; $D = 0,1$; $E = 0,9$; $F = 0,9$
- C. A, B, C, D, E ne peuvent pas être calculés
- D. Les résultats de cette étude peuvent être extrapolés (étendus) à toutes les femmes atteintes de cancer du sein.
- E. Par rapport à toutes les femmes atteintes d'un cancer du sein cette étude présente un biais de caractérisation (d'information)
6. En vous référant aux tables fournies en fin de fascicule, que pensez-vous de ces affirmations ?
- A. La table de la fonction de répartition de la loi normale est la table de $\pi(t)$
- B. Par complémentarité : $\pi(t) = 0,5 + \phi(t) \quad t > 0$
- C. Une loi de probabilité de la variable discontinue X est définie par les couples : (x_i, p_i)
- D. La table de l'écart réduit donne la probabilité α pour qu'une variable aléatoire normale centrée réduite soit extérieure à l'intervalle $[-t_\alpha; +t_\alpha[$
- E. La fonction de répartition d'une loi continue se définit par : $F(x_i) = P(X < x_i)$

7. Pour un examen, dix examinateurs ont préparé chacun deux sujets. On dispose donc de vingt sujets que l'on place dans 20 enveloppes identiques.

Deux candidats se présentent : chacun choisit au hasard et simultanément deux sujets ; de plus, les sujets choisis par le premier candidat ne seront plus disponibles pour le deuxième.

Soit X la variable aléatoire égale au nombre de candidats qui ont choisi chacun deux sujets provenant d'un même examinateur.

La variable aléatoire X prend donc les valeurs 0, 1 ou 2.

La loi de probabilité de la variable aléatoire X est donnée ci-dessous :

k	0	1	2
$P(X = k)$	$\frac{290}{323}$	$\frac{32}{323}$	$\frac{1}{323}$

Quelle est la valeur de l'espérance mathématique de la variable aléatoire X ?

- A. $\frac{34}{323}$
 B. $\frac{15}{323}$
 C. $\frac{33}{323}$
 D. $\frac{2}{19}$
 E. $\frac{17}{323}$
8. Si 20% des pièces produites par une machine sont défectueuses, quelle est la probabilité, sur quatre pièces choisies au hasard, de n'en avoir aucune défectueuse ?
- A. 0,4096
 B. 0,3128
 C. 0,8192
 D. 0,6075
 E. 0,1808
9. On suppose que 2% des êtres humains sont gauchers. Quelle est la probabilité pour que, parmi 100 personnes, 3 ou plus soient gauchères ?
- A. 6,6%
 B. 67,7%
 C. 32,3%
 D. 57,7%
 E. 43,3%
10. Une machine met un médicament en sachets dont les masses se répartissent suivant la loi normale autour de la valeur moyenne de 1,20 g avec un écart-type de 0,06 g.
 Quelle est la probabilité pour que la masse d'un sachet soit comprise entre 1,14 g et 1,29 g ?
- A. 51,6%
 B. 59,7%
 C. 70,9%
 D. 77,5%
 E. 68,4%

5 ED 5 : Mesures et variables, statistiques descriptives, variables aléatoires, estimation

5.1 Exercices

Exercice 1

Exercice 1 : Dans une étude, on a tirés au sort des dossiers médicaux et l'on a noté le nombre d'anomalies constatées dans chaque dossier. On a obtenu la série suivante :

1 ; 4 ; 1 ; 2 ; 5 ; 3 ; 1 ; 2 ; 6 ; 1 ; 2 ; 3 ; 2 ; 3 ; 2 ; 4 ; 1 ; 2 ; 3 ; 2

1. Quel est le nombre de dossiers analysés ?
2. De quel type est la variable nombre d'anomalies ?
3. Mettre cette série sous forme d'une distribution non groupée.
4. Quels diagrammes pouvez-vous réaliser ?
5. Que vaut le premier quartile ?
6. Que vaut la moyenne ?

Exercice 2

Des sujets atteints de la maladie S reçoivent tous le même traitement M. Deux mois après le début du traitement M, chaque sujet est examiné. La constatation ou non d'une rémission est alors recherchée. Sur 904 sujets, 251 sont en rémission soit 27,77%.

1. Quelles sont les bornes de l'intervalle de confiance à 95% de ce résultat ? Sont-elles valides ?
2. En fait, parmi les 904 sujets se trouvent 4 sous-groupes qui diffèrent par la performance de leurs profils métaboliques. Les bornes indiquées sont les bornes de l'intervalle de confiance à 95% pour le pourcentage correspondant. Complétez les cases manquantes :

Groupe	effectif	en rémission	% de rémission	Borne inf	Borne sup
Groupe 1	236	75	31,78%		37,72%
Groupe 2	189	52		21,15%	33,88%
Groupe 3	260		16,15%	11,68%	20,63%
Groupe 4		82	37,44%	31,03%	

3. Dans la littérature, un article de référence donne 26,40% de rémission et nous considérerons cette valeur comme étant celle de la proportion Π (paramètre). Quels commentaires pouvez-vous faire à propos du groupe de 904 sujets et pour chacun des 4 sous-groupes ?

Exercice 3

La poursuite oculaire a été étudiée chez 64 sujets et donne les résultats suivants (en unités arbitraires ua). On considérera que la variable suit une loi normale.

Moyenne d'échantillon : 66,77 ua

Variance d'échantillon : 240,34 ua^2

1. Quel est l'intervalle de confiance à 95% de la moyenne de la poursuite oculaire ?
2. Quelle est alors la précision obtenue pour la moyenne ?

On souhaite réaliser une étude concernant la poursuite oculaire et on prendra cette fois pour variance, une variance de la littérature $\sigma^2 = 230ua^2$ considérée ici comme un paramètre.

3. Combien de sujets faudrait-il recruter pour obtenir une précision de $\pm 1ua$ sur la moyenne de la poursuite oculaire avec un risque $\alpha = 5\%$?

4. Combien de sujets faudrait-il recruter pour obtenir une précision de $\pm 3,83ua$ sur la moyenne de la poursuite oculaire avec un risque $\alpha = 5\%$?
5. Pourquoi ne trouve-t-on pas en réponse à la question précédente (4) une valeur plus proche de 64 sujets ?

5.2 QCM

1. Dans une étude on a les variables suivantes :
 - Sexe (codé 1 si masculin, 2 si féminin)
 - Age en années
 - Poids en kg
 - Douleurs notée 0 si absente, 1 si faible, 2 si moyenne, 3 si forte
 Cochez les propositions vraies
 - A. On peut calculer les effectifs et les fréquences des différentes valeurs de toutes les variables
 - B. On peut calculer la médiane de l'âge, du poids et des douleurs.
 - C. On peut calculer la moyenne de l'âge, du poids et des douleurs.
 - D. On peut calculer le moment d'ordre 2 de toutes les variables.
2. La moyenne du périmètre thoracique de 10 hommes vaut 90cm, celle de 20 femmes 96cm. Que vaut la moyenne du périmètre thoracique des 30 individus réunis ?
 - A. 93cm
 - B. 92cm
 - C. 94cm
 - D. 96cm
 - E. 95cm

Questions 3 à 4, énoncé commun.

Une étude a été réalisée de 1958 à 1964 afin d'étudier la mortalité (exprimée par un Taux de mortalité pour 100 000 habitants que nous appellerons plus simplement Taux de mortalité) en fonction de la Concentration en Calcium de l'eau de boisson consommée (exprimée en ppm).

Pour chacune des 61 villes d'Angleterre ayant participé à l'étude, vous disposez du Taux de mortalité et de la mesure de la Concentration en Calcium. Vous considérerez que vous disposez de toutes ces valeurs (il n'y a pas de donnée manquante) et que chaque variable suit une loi normale.

ville	Taux de mortalité	Concentration en Calcium (ppm)
Coventry	1307	78
Oxford	1175	107
...

D'après Hills and al : Statistical Methods

Vous disposez (dans le tableau ci-dessous) pour la Concentration en Calcium, de la moyenne, de l'écart type, de la valeur la plus faible (min) et de la valeur la plus élevée (max)

	m	s	min	max
Concentration en Calcium (ppm)	47,18	38,09	5	138

Les données fournies par les tables pour $\alpha = 0,05$ sont les suivantes :

$$\begin{aligned} \varepsilon_\alpha &= 1,96 \\ t_\alpha &= 1,96 \text{ pour } 61 \text{ ddl} & t_\alpha &= 1,96 \text{ pour } 59 \text{ ddl} \\ t_\alpha &= 1,96 \text{ pour } 60 \text{ ddl} & t_\alpha &= 1,96 \text{ pour } 58 \text{ ddl} \end{aligned}$$

3. Avec l'ensemble des données disponibles, vous pouvez obtenir
 - A. l'estimation par intervalle de confiance de la moyenne m de la Concentration en Calcium
 - B. l'estimation par intervalle de confiance de la moyenne μ de la Concentration en Calcium
 - C. l'estimation de la variance de la Concentration en Calcium de la population parente
 - D. la variance de la Concentration en Calcium de la population parente
 - E. la variance d'échantillon de la Concentration en Calcium
4. L'estimation par intervalle de confiance à 95% de la moyenne de la Concentration en Calcium
 - A. utilise la formule : $47,18 \pm 1,96 \times \sqrt{\frac{38,09}{61}}$
 - B. utilise la formule : $47,18 \pm 1,96 \times \frac{38,09}{\sqrt{60}}$
 - C. utilise la formule : $47,18 \pm 1,96 \times \frac{38,09}{\sqrt{61}}$
 - D. utilise la formule : $47,18 \pm 1,96 \times \sqrt{\frac{38,09}{60}}$
 - E. utilise une autre formule.
5. Estimation par intervalle et estimation ponctuelle ont toutes deux :
 - A. pour but d'estimer la valeur d'un paramètre
 - B. pour but d'estimer la valeur d'une statistique
 - C. pour but d'estimer la variance d'un paramètre
 - D. a prendre en compte le risque de première espèce dans leurs calculs respectifs
 - E. a prendre en compte le risque de deuxième espèce dans leurs calculs respectifs
6. Dans le cas général, un estimateur D_n du paramètre δ de la variable aléatoire X est sans biais
 - A. dès que l'effectif n est supérieur ou égal à 20
 - B. si $E(D_n) = \delta$
 - C. si $E(D_n) = 0$
 - D. si $E(D_n - \delta)^2 = 1$
 - E. si $E(D_n - \delta) = 1$
7. Dans le cas général, un estimateur D_n du paramètre δ de la variable aléatoire X est convergent
 - A. dès que l'effectif n est supérieur ou égal à 20
 - B. si $E(D_n)$ tend vers 1 quand n tend vers l'infini
 - C. si $E(D_n)$ tend vers 0 quand n tend vers l'infini
 - D. si $E(D_n - \delta)^2$ tend vers 0 quand n tend vers l'infini
 - E. si $E(D_n - \delta)$ tend vers δ quand n tend vers l'infini

6 ED 6 : Échantillonnage, principe des tests statistiques, adaptation aux lois, χ^2

6.1 Exercices

Exercice 1

On étudie un échantillon radioactif en déterminant le nombre de désintégrations qui se produisent pendant une durée de 30 secondes. Cette expérience est répétée 100 fois ce qui donne 100 valeurs de K qui se répartissent de la façon suivante :

k_i	0	1	2	3	4	5	6	7	8	9
n_i	1	4	11	14	18	18	16	11	6	1

1. Calculer la moyenne, la variance et l'écart-type de cette distribution.
2. Par quelle loi de probabilité peut-on ajuster la distribution observée ? Justifier votre réponse.
3. Quel test statistique utiliserez-vous pour vérifier le bien-fondé de votre choix ? Quelle est votre conclusion, au risque de 5 % ?

Exercice 2

Le tableau suivant donne, pour des enfants, la correspondance entre la taille des amygdales et la présence ou l'absence de streptocoques hémolytiques au prélèvement pharyngien.

		Porteurs de germes		Totaux
		Oui	Non	
Amygdales	non hypertrophiées	19	496	515
	hypertrophiées	30	570	600
	très hypertrophiées	23	262	285
Totaux		72	1328	1400

Y a-t-il association ou indépendance entre les caractères (seuil de risque de 5%) ?

Exercice 3

On observe la survie de 100 malades atteints d'une maladie M pouvant prendre l'une ou l'autre des deux formes A et B.

L'une des formes est-elle plus grave que l'autre, compte tenu du tableau suivant (au coefficient de risque de 5 %) ?

	Vivants	Décédés	Totaux
Forme A	30	5	35
Forme B	35	30	65
Totaux	65	35	100

6.2 QCM

Pour chaque question, indiquez les propositions vraies

1. Pour mesurer un phénomène de santé
 - A. On l'estime dans la population à partir de l'échantillon
 - B. On le mesure dans un échantillon
 - C. Un échantillon permet d'obtenir un estimateur du taux
 - D. Une population donne une estimation d'un indicateur
2. Parmi les différents types de sondages
 - A. Un sondage empirique consiste à choisir des sujets de façon aléatoire
 - B. Un sondage aléatoire peut utiliser la méthode des quotas
 - C. Un sondage en grappe est un sondage au moins à deux degrés
 - D. Un sondage stratifié permet d'estimer un indicateur dans plusieurs strates distinctes indépendamment
 - E. Un sondage stratifié donne une estimation globale qui tient compte des caractéristiques des strates définies pour l'étude
3. La méthode des quotas
 - A. Est une méthode qui respecte le hasard
 - B. Consiste à échantillonner un nombre déterminé de sujets
 - C. Consiste à échantillonner des sujets dans des catégories définies
 - D. Est une méthode qui représente exactement la probabilité de répartition des catégories dans la population
4. la base de sondage
 - A. Représente l'ensemble de la population chez qui on étudie un phénomène de santé
 - B. Est la méthode choisie pour effectuer le sondage
 - C. Doit être vérifiée et fiable
 - D. Peut être, par exemple, les listes électorales
5. Un sondage où la présence de chaque individu dans l'échantillon est équiprobable est un sondage :
 - A. Empirique
 - B. Aléatoire
 - C. Stratifié
 - D. En grappe
6. Que pensez-vous de ces affirmations ?
 - A. α est appelé le seuil de sécurité (confiance)
 - B. Le nombre de degrés de liberté correspond au nombre de termes dépendants
 - C. Dans les tests statistiques, on se fixe généralement $\alpha = 0,05$. Il y correspond un seuil donné par la table
 - D. La loi de χ^2 tend vers une loi normale à partir de 30 degrés de liberté
 - E. Problèmes d'estimation, de conformité, d'homogénéité, de corrélation peuvent être examinés au cours d'une étude statistique

7. Que pensez-vous de ces affirmations ?
- A. Lors d'une étude statistique, les résultats sont accompagnés d'indications quantitatives fixant le degré de confiance qui peut leur être accordé
 - B. L'hypothèse d'un test statistique est imposée par construction du test
 - C. Lors de l'étude de plusieurs caractères, il est possible de chercher une liaison entre les valeurs des caractères
 - D. Un test paramétrique nécessite des conditions de validité
 - E. La loi de SNEDECOR à ν_1 et ν_2 degrés de liberté notée F_{ν_1, ν_2} vérifie $F_{\nu_1, \nu_2} = F_{\nu_2, \nu_1}$

8. Que pensez-vous de ces affirmations ?
- A. Les tests statistiques mettant en oeuvre deux caractères qualitatifs s'effectuent par un test non paramétrique
 - B. Un tableau de contingence peut permettre de présenter les résultats de deux caractères quantitatifs
 - C. Les fluctuations d'échantillonnage sont toujours présentes quelle que soit l'étude statistique non exhaustive effectuée
 - D. Deux caractères qualitatifs peuvent se présenter sous plusieurs modalités
 - E. La formule du χ^2 sans correction de YATES est légitime quelle que soit la taille de l'échantillon

9. Le taux de glycémie a été mesuré chez 80 enfants prématurés. On considère qu'un enfant est hypoglycémique si sa glycémie est égale ou inférieure à 30 cg/l et qu'il est normoglycémique si sa glycémie est supérieure à 30 cg/l.

On constate que sur 55 enfants dont la glycémie est normale, 24 sont des garçons, alors que sur 25 enfants dont la glycémie est abaissée, on dénombre 9 filles et 16 garçons.

On demande de déterminer, à l'aide d'un test approprié, si les filles seraient moins sujettes que les garçons à présenter une hypoglycémie, au coefficient de sécurité de 95%.

Que pensez-vous de ces affirmations ?

- A. Le problème à étudier est la comparaison de deux distributions
 - B. Le test choisi sera un χ^2 d'homogénéité
 - C. Le nombre de degrés de liberté est 1
 - D. On choisira la formule du χ^2 avec correction de YATES
 - E. L'effectif théorique du nombre de filles hypoglycémiques est $\frac{25}{2}$
10. On a relevé, pour 160 familles de 4 enfants, le nombre de garçons par famille. On obtient la distribution suivante :

Nombre de garçons	0	1	2	3	4
Nombre de familles	11	37	57	43	12

On fait l'hypothèse que la probabilité d'avoir un garçon est égale à 0,5. L'adéquation à une loi de probabilité est recherchée. Un test est choisi et donne comme valeur 1,1.

Tous les effectifs théoriques sont supérieurs à 5.

Au seuil de sécurité de 95%, on souhaite savoir si c'est vrai.

Que pensez-vous de ces affirmations ?

- A. Le nombre de degrés de liberté est 4
- B. Le test choisi est un test d'homogénéité
- C. La loi de probabilité choisie est la loi de Poisson
- D. Le test choisi est un test de conformité
- E. En utilisant la table appropriée, on peut dire que l'hypothèse émise est acceptée

11. La formule du nombre de degrés de liberté ν est, dans le cas général d'un test du χ^2 :

(les paramètres des lois sont estimés)

- A. d'indépendance : $(k - 1) \times (l - 1)$; k désignant le nombre de colonnes, l le nombre de lignes
 - B. de conformité à une loi binomiale : $k - 1$; k désignant le nombre de valeurs de la distribution
 - C. d'homogénéité : $k - 1 - r$; k désignant le nombre de possibilités et r le nombre de relations nécessaires
 - D. de conformité à une loi normale : $k - 3$; k désignant le nombre de valeurs de la distribution
 - E. de conformité à une loi de distribution : $k - r$; k désignant le nombre de valeurs et r le nombre de relations nécessaires
12. Lors d'un contrôle de la mise en suspension de pénicilline retard en flacons siliconés, le mode opératoire consiste à classer, selon des critères visuels, les flacons en trois catégories (mise en suspension bonne, mise en suspension correcte, mise en suspension mauvaise).

Pour déterminer la subjectivité des critères, on soumet à trois contrôleurs (C1, C2, C3), trois échantillons de 100 flacons tirés au hasard dans un même lot. Les résultats sont les suivants :

	Bonne	Correcte	Mauvaise	
C1	61	27	12	100
C2	56	29	15	100
C3	63	28	9	100

et conduisent à une valeur de test de 3,07.

Au coefficient de sécurité de 95%, on souhaite vérifier cette subjectivité.

Que pensez-vous de ces affirmations ?

- A. Bonne, Correcte et Mauvaise sont des modalités de la variable "Mise en suspension"
- B. Le nombre de degrés de liberté est 6
- C. L'effectif théorique du nombre de flacons examinés par le contrôleur C2 ayant "mauvaise mise en suspension" est : 12
- D. En utilisant la table appropriée, on peut affirmer que l'hypothèse émise est acceptée
- E. L'effectif théorique du nombre de flacons examinés par le contrôleur C3 ayant "bonne mise en suspension" est : 62

13. Que pensez-vous de ces affirmations ?

- A. La normalité d'une distribution peut se contrôler d'une manière graphique
- B. L'équation de l'écart centré réduit correspond à la formulation de la droite de Henry
- C. On reporte les fréquences cumulées d'une distribution selon l'axe vertical d'un papier gaussométrique ou gausso-logarithmique pour vérifier la normalité
- D. L'approximation de la moyenne ne peut s'obtenir par lecture sur le graphique gaussométrique
- E. Utiliser un papier millimétré ne permet pas de vérifier l'ajustement à une loi normale

7 ED 7 : Comparaison de moyennes et de variances, épidémiologie

7.1 Exercices

Exercice 1

Dans une étude sur l'infarctus du myocarde, on a comparé l'efficacité de deux traitements dans deux échantillons tirés au sort :

- le groupe A (32 sujets) a reçu le traitement classique de l'infarctus du myocarde.
- le groupe B (32 sujets) a reçu un traitement par cellules souches

Ces traitements sont censés améliorer le fonctionnement cardiaque en augmentant la Fraction d'Éjection Ventriculaire (FEV). La FEV a été évaluée chez chaque sujet par deux mesures : la première 1 jour après l'infarctus (notée FEV_1), et la seconde 6 mois plus tard (notée FEV_2). On a noté les résultats suivants :

groupe	FEV_1		FEV_2		$FEV_2 - FEV_1$	
	\bar{x}	(s)	\bar{x}	(s)	\bar{x}	(s)
A	51,3	(9,3)	52,0	(12,4)	0,7	(8,1)
B	50,0	(10,0)	56,7	(12,5)	6,7	(6,5)

\bar{x} =moyenne de l'échantillon, s =écart-type estimé

1. Y-a-t-il, au risque 5%, une différence de FEV entre les groupe A et B au jour 1 ? Argumentez les conditions d'applications du test utilisé
2. Peut-on conclure, au risque 5%, à une différence d'évolution de la FEV entre les groupes A et B ?
3. Peut-on conclure, au risque 5%, à une amélioration de la FEV dans le groupe A ?
4. Sachant que la fraction d'éjection ventriculaire d'un coeur sain vaut 60, que dire de la FEV à 6 mois dans le groupe B ?

7.2 QCM

Enoncé commun à tous les QCM.

On a mesuré le pourcentage X de polynucléaires éosinophiles dans la formule sanguine de deux groupes de sujets différents A et B. Les résultats sont les suivants :

- moyennes $\bar{x}_A = 10 \%$; $\bar{x}_B = 1 \%$
- effectifs : $n_A = 21$, $n_B = 41$
- Somme des carrés des écarts à la moyenne : $SCE_A = 0,2$, $SCE_B = 0,2$

1. Pour comparer X dans ces deux groupes, on pourrait utiliser :
 - A. un test de comparaison de moyennes sur échantillons indépendants
 - B. un test de comparaison de moyennes sur échantillons appariés
 - C. un test de comparaison de pourcentages sur échantillons indépendants
 - D. un test de comparaison de pourcentages sur échantillons appariés
 - E. un test de Mann-Whitney-Wilcoxon
2. Soit R le rapport des variances estimées de X_A et X_B . Quelles sont les propositions exactes ?
 - A. on dispose de suffisamment d'information pour utiliser un test paramétrique
 - B. il y a bien hétéroscédasticité, au risque de première espèce 0,10
 - C. Quelle que soit la distribution de X_A , \bar{X}_A suit approximativement une loi normale
 - D. Il y a une probabilité de 5% pour que R ou $1/R$ dépasse 2,287
 - E. Il y a une probabilité de 2,5% pour que R ou $1/R$ dépasse 1,994

On a compté, chez chaque sujet du groupe A, le nombre Y de polynucléaires éosinophiles avant puis après la prise d'un médicament. Les seules informations dont on dispose sont ces mesures :

avant (G/L)	0,30	0,30	0,25	0,70	0,10	0,40	0,25	0,10	0,30	0,40
après (G/L)	0,70	0,60	0,40	0,50	0,20	0,40	0,60	1,90	2,30	0,10

Vous choisissez le test le plus adapté pour comparer Y avant et après prise du médicament. Soit S la statistique associée à ce test.

3. Pour calculer s_{obs} , la réalisation de S , il faut :
 - A. calculer la somme des valeurs de Y avant et après
 - B. calculer la somme des rangs des valeurs de Y avant et après
 - C. calculer les différences avant/après des valeurs de Y
 - D. calculer la somme des différences avant/après des valeurs de Y
 - E. calculer les rangs des différences avant/après des valeurs de Y
4. On peut dire que :
 - A. il y a une différence significative du nombre de polynucléaires éosinophiles avant et après prise du médicament dans le groupe A
 - B. s_{obs} est supérieure à la valeur critique dans la table correspondante
 - C. $s_{obs} = 7$
 - D. $s_{obs} = 7,5$
 - E. il est nécessaire de connaître la distribution de Y pour réaliser le test

5. A propos des enquêtes descriptives
 - A. On estime l'incidence dans une étude longitudinale
 - B. On estime la prévalence dans une étude transversale
 - C. L'incidence compte l'ensemble des cas à un instant donné
 - D. La prévalence permet de décrire la charge en soin due à la maladie
6. les études d'incidence :
 - A. permettent de connaître la dynamique de la maladie
 - B. indiquent la charge en soin pour le système de santé à un moment donné
 - C. Consistent à compter l'ensemble des cas de la maladie
 - D. s'effectuent sur une période de temps
7. les études cas-témoins sont :
 - A. utiles pour connaître la performance diagnostique d'un examen complémentaire
 - B. utiles pour évaluer l'efficacité d'un traitement chirurgical
 - C. inadaptées pour identifier un facteur de risque de développer une maladie
 - D. inadaptées pour identifier un facteur pronostique
8. les études de cohorte :
 - A. permettent de sélectionner les sujets malades
 - B. identifient l'exposition dans le passé ou au début de l'étude
 - C. incluent des sujets indépendamment de l'évolution de leur maladie
 - D. comparent des sujets exposés et des sujets non exposés

8 ED 8 : Comparaison de pourcentages, régression, corrélation

8.1 Exercices

Exercice 1

Une enquête de satisfaction est réalisée sur un échantillon de 200 patients concernant la qualité des repas dans un service hospitalier. Le tableau suivant donne les résultats de l'indice global de satisfaction en fonction de la classe d'âge des patients.

	Satisfait	Insatisfait
< 60 ans	25 %	25 %
≥ 60 ans	20 %	30 %

Le pourcentage de patients insatisfaits est-il significativement différent chez les moins que chez les plus de 60 ans ?

Exercice 2

Une maladie touche 10% d'une population. On expérimente un traitement sur un échantillon de taille $n = 200$ et on trouve, à la suite du traitement, 5% des sujets touchés. Peut-on dire que le traitement est efficace avec un niveau de signification 0,05 ?

Exercice 3

Le dosage d'une substance sérique est réalisé en routine chez des enfants avec la méthode K1 (résultat en ng/mL). Une nouvelle méthode K2 (résultat en ng/mL) plus rapide a été mise au point et 45 dosages ont été réalisés avec chacune de ces deux méthodes. L'âge des sujets a également été recueilli. On supposera la normalité des distributions des variables.

L'objectif est d'étudier la relation entre les dosages réalisés par la nouvelle méthode K2 et les dosages de référence réalisés avec K1. On souhaiterait également savoir si il existe une relation entre l'âge et le résultat de ces dosages.

1. Quelle stratégie allez-vous utiliser pour répondre à chacune de ces problématiques ?

Les données sont les suivantes :

Dosage par K1 :

Moyenne d'échantillon $m_1 = 68,089 \text{ ng/mL}$ Variance d'échantillon $s_1^2 = 481,583 \text{ ng}^2/\text{mL}^2$

Dosage par K2 :

Moyenne d'échantillon $m_2 = 70,622 \text{ ng/mL}$ Variance d'échantillon $s_2^2 = 478,422 \text{ ng}^2/\text{mL}^2$

Age :

Moyenne d'échantillon $m_a = 8,61 \text{ ans}$ Variance d'échantillon $s_a^2 = 2,496 \text{ ans}^2$

Covariances :

$Cov(K1, K2) = 385,921 \text{ ng}^2/\text{mL}^2$ $Cov(K1, Age) = 0,199 \text{ ans} \times \text{ng/mL}$

$Cov(K2, Age) = 2,675 \text{ ans} \times \text{ng/mL}$

2. Quel est le résultat de l'analyse correspondant à la première problématique (relation K1 K2) ?
3. Quels sont les résultats de l'analyse correspondant à la seconde problématique (relation âge avec K1 K2) ?
4. On donne $r_{K1, K2} = 0,804$. De toutes les relations étudiées dans cet exercice, lesquelles sont statistiquement significatives ?
5. Aurait-on pu comparer les résultats entre K1 et K2 ?
6. Aurait-on pu comparer les résultats entre K1 et l'âge ou entre K2 et l'âge ?

7. Pourrait-on remplacer la méthode K1 par la méthode K2 ?
8. Pour une valeur de K1 égale à 60, quelle prédiction de la valeur de K2 peut-on faire ?
9. Pourquoi dans ce contexte, $r_{K1,K2} = 0,804$ est-il aussi proche de la valeur 0,801 ?

8.2 QCM

Questions 1 et 2, énoncé commun.

A la suite d'un même traitement on observe :

- 60% de bons résultats chez 500 jeunes patients
- 40% de bons résultats chez 500 patients âgés

On cherche à savoir si les effets du traitement sont les mêmes dans la population des jeunes patients que dans celle des patients âgés au risque de 5%.

1. Indiquez, pour chaque assertion, si elle est vraie ou fausse :
 - A. On peut réaliser un test de comparaison de 2 proportions observées sur échantillons indépendants
 - B. On peut réaliser un test de comparaison de 2 proportions observées sur échantillons appariés
 - C. On peut réaliser un test de comparaison d'une proportion observée à une proportion théorique
 - D. Pour pouvoir effectuer le test utilisant la loi normale, il faut que l'effectif $50 \times 0,5$ soient supérieur ou égal à 5
 - E. Pour pouvoir effectuer le test utilisant la loi normale, il faut que les effectifs $50 \times 0,6$ et $50 \times 0,4$ soient supérieurs ou égaux à 5
2. On effectue un test du χ^2 . Indiquez, pour chaque assertion, si elle est vraie ou fausse :
 - A. Hypothèse nulle H_0 : pas de différence significative entre les pourcentages de bons résultats dans les 2 échantillons
 - B. On trouve que la valeur prise par la statistique est $\chi^2 = 2$
 - C. On trouve que la valeur prise par la statistique est $\chi^2 = 4$
 - D. On compare le χ^2 trouvé à la valeur seuil de 3,84
 - E. On peut dire que les effets du traitement sont significativement différents dans la population des jeunes patients que dans celle des patients âgés.

On sait qu'en France le taux moyen de césariennes ces dernières années est de 20%. En 2010, à la maternité régionale, sur 900 naissances, 300 césariennes ont été pratiquées. On cherche à savoir si le taux de césarienne pratiqué à la maternité régionale en 2010 est significativement différent du taux national

3. Indiquez, pour chaque assertion, si elle est vraie ou fausse :
 - A. La variable aléatoire à laquelle on s'intéresse est une variable quantitative
 - B. On cherche à réaliser un test de comparaison de 2 proportions observées sur échantillons indépendants
 - C. On ne peut pas pratiquer le test car on ne connaît pas le nombre de naissance en France en 2010
 - D. On ne peut pratiquer le test car les conditions de validité de sont pas respectées
 - E. On réalise le test et on ne met pas en évidence de différence significative entre le taux de césarienne pratiqué à la maternité régionale en 2010 et le taux national

Questions 4 et 5 : énoncé commun.

Une étude a pour objectif d'évaluer, par IRM mammaire, la présence et la taille des reliquats tumoraux chez des patientes traitées pour un cancer du sein non inflammatoire par chimiothérapie néoadjuvante puis chirurgie. Cette étude a concerné 50 patientes ayant bénéficié d'une IRM post-chimiothérapique avant chirurgie. Les IRM ont été confrontées aux résultats anatomopathologiques postopératoires (analyse histologique de la pièce opératoire). La mesure de la taille du reliquat tumoral obtenue par l'examen histologique après chirurgie est considérée comme l'évaluation la plus précise que l'on puisse avoir. La confrontation des résultats IRM à l'analyse histologique a montré : 30 vrais positifs, 10 vrais négatifs, 10 faux positifs et aucun faux négatif. On cherche à savoir si le pourcentage de reliquats tumoraux identifiés par IRM est significativement différent de celui par analyse histologique, au risque 5%.

4. Indiquez, pour chaque assertion, si elle est vraie ou fausse :
 - A. La sensibilité de la méthode IRM à détecter un reliquat tumoral est de 75 %
 - B. On réalise un test de comparaison de 2 proportions observées sur échantillons indépendants
 - C. On réalise un test de comparaison de 2 proportions observées sur échantillons appariés
 - D. Les conditions d'application du test sont vérifiées
 - E. Les conditions d'application du test ne sont pas vérifiées mais on peut utiliser la correction de Yates
5. On trouve :
 - A. On obtient le résultat suivant avec le test $Z : |z| = 5$
 - B. On obtient le résultat suivant avec le test du $\chi^2 : \chi^2 = 10$
 - C. On obtient le résultat suivant avec le test du $\chi^2 : \chi^2 = 18,75$
 - D. On rejette l'hypothèse nulle au risque considéré
 - E. On ne montre pas de différence significative entre les pourcentages tumoraux identifiés par IRM et par analyse histologique

Questions 6 à 8 énoncé commun.

Une étude a été réalisée de 1958 à 1964 afin d'étudier la mortalité (exprimée par un Taux de mortalité pour 100 000 habitants que nous appellerons plus simplement Taux de mortalité) en fonction de la Concentration en Calcium de l'eau de boisson consommée (exprimée en ppm).

Pour chacune des 61 villes d'Angleterre ayant participé à l'étude, vous disposez du Taux de mortalité et de la mesure de la Concentration en Calcium. Vous considérerez que vous disposez de toutes ces valeurs (il n'y a pas de donnée manquante) et que chaque variable suit une loi normale.

Exemple de données :

ville	Taux de mortalité	Concentration en Calcium (ppm)
Coventry	1307	78
Oxford	1175	107
...

D'après Hills and al : Statistical Methods

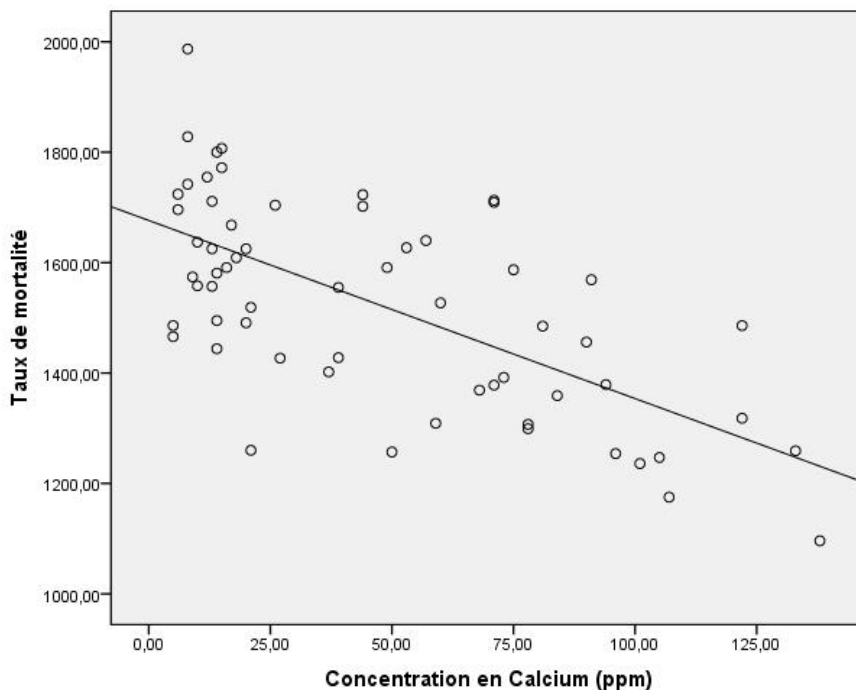
Vous disposez (dans le tableau ci-dessous) pour la Concentration en Calcium, de la moyenne, de l'écart type, de la valeur la plus faible (min) et de la valeur la plus élevée (max) :

	m	s	min	max
Concentration en Calcium (ppm)	47,18	38,09	5	138

Les données fournies par les tables pour $\alpha = 0,05$ sont les suivantes :

- $\varepsilon_\alpha = 1,96$
- $t_\alpha = 1,96$ pour 61 ddl
- $t_\alpha = 1,96$ pour 59 ddl
- $t_\alpha = 1,96$ pour 60 ddl
- $t_\alpha = 1,96$ pour 58 ddl

Le graphique ci-dessous vous donne la droite de régression du Taux de mortalité en fonction de la Concentration en Calcium ainsi que le nuage des 61 points correspondant.



6. **Au vu de ce graphique**, et si l'on admet que les conditions d'application des calculs de la régression linéaire et du coefficient de corrélation linéaire sont remplies, on peut dire :
- que la comparaison entre Taux de mortalité et Concentration en Calcium est significative
 - qu'il existe une relation significative entre Taux de mortalité et Concentration en Calcium
 - que la pente de cette droite est négative
 - que le coefficient de corrélation linéaire correspondant sera positif
 - que le coefficient de corrélation linéaire correspondant sera négatif
7. La valeur absolue du coefficient de corrélation linéaire calculé pour le Taux de mortalité et la Concentration en Calcium est égale à 0,655 (vous admettrez ce résultat). Le test de ρ à la valeur 0
- est donné par la formule $t = \frac{0,665}{\sqrt{\frac{1-0,665^2}{61-2}}}$
 - est donné par la formule $t = \frac{0,665}{\sqrt{\frac{1-0,665^2}{61-2}}}$
 - est donné par la formule $t = \frac{-0,665}{\sqrt{\frac{1-0,665^2}{61}}}$
 - est donné par la formule $t = \frac{0,665}{\sqrt{\frac{1-0,665^2}{61-1}}}$
 - est donné par la formule $t = \frac{-0,665}{\sqrt{\frac{1-0,665^2}{61-2}}}$
8. Si l'on admet que le test de ρ à la valeur 0 (dans l'étude de la corrélation linéaire entre le Taux de mortalité et la Concentration en Calcium) est significatif, on peut dire :
- que plus la Concentration en Calcium augmente, plus le Taux de mortalité augmente
 - que plus la Concentration en Calcium augmente, plus le Taux de mortalité diminue
 - que plus la Concentration en Calcium diminue, plus le Taux de mortalité diminue
 - qu'il existe une liaison significative entre les deux variables
 - que la comparaison entre les deux variables est significative

Questions 9 et 10 énoncé commun.

On pose la formulation suivante :

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

9. Cette formulation
- correspond au modèle de la régression linéaire de Y en X
 - correspond au modèle de la régression linéaire de X en Y
 - permet de tester l'égalité des pentes
 - donne comme résultat la valeur du coefficient de corrélation linéaire entre Y et X
 - suppose l'existence d'une relation linéaire entre X et Y

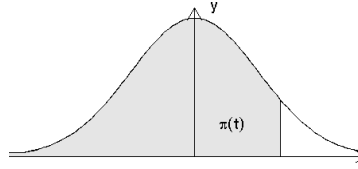
10. Dans cette formulation
- A. β_0 et β_1 correspondent aux paramètres du modèle
 - B. β_0 est le coefficient de corrélation linéaire
 - C. ε est l'ordonnée à l'origine
 - D. ε correspond à la composante aléatoire
 - E. X est la variable à expliquer
11. Le coefficient de corrélation linéaire
- A. est égal à la somme des pentes
 - B. a le même signe que la variance des x
 - C. est compris entre -1 et 1
 - D. est égal (en valeur absolue) à la moyenne géométrique des pentes
 - E. est égal à la moyenne des pentes

Tables statistiques¹ et papiers gradués

1. Générées avec R : R Development Core Team (2010). R : A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

Table de $\pi(t)$

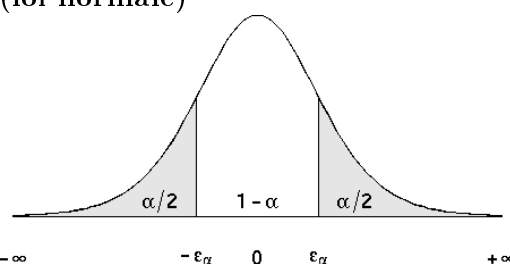
$$\pi(t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$



t	-0,09	-0,08	-0,07	-0,06	-0,05	-0,04	-0,03	-0,02	-0,01	-0,00
-3,0	0,001 0	0,001 0	0,001 1	0,001 1	0,001 1	0,001 2	0,001 2	0,001 3	0,001 3	0,001 3
-2,9	0,001 4	0,001 4	0,001 5	0,001 5	0,001 5	0,001 6	0,001 6	0,001 7	0,001 8	0,001 9
-2,8	0,001 9	0,002 0	0,002 1	0,002 1	0,002 2	0,002 3	0,002 3	0,002 4	0,002 5	0,002 6
-2,7	0,002 6	0,002 7	0,002 8	0,002 9	0,003 0	0,003 1	0,003 2	0,003 3	0,003 4	0,003 5
-2,6	0,003 6	0,003 7	0,003 8	0,003 9	0,004 0	0,004 1	0,004 3	0,004 4	0,004 5	0,004 7
-2,5	0,004 8	0,004 9	0,005 1	0,005 2	0,005 4	0,005 5	0,005 7	0,005 9	0,006 0	0,006 2
-2,4	0,006 4	0,006 6	0,006 8	0,006 9	0,007 1	0,007 3	0,007 5	0,007 8	0,008 0	0,008 2
-2,3	0,008 4	0,008 7	0,008 9	0,009 1	0,009 4	0,009 6	0,009 9	0,010 2	0,010 4	0,010 7
-2,2	0,011 0	0,011 3	0,011 6	0,011 9	0,012 2	0,012 5	0,012 9	0,013 2	0,013 6	0,013 9
-2,1	0,014 3	0,014 6	0,015 0	0,015 4	0,015 8	0,016 2	0,016 6	0,017 0	0,017 4	0,017 9
-2,0	0,018 3	0,018 8	0,019 2	0,019 7	0,020 2	0,020 7	0,021 2	0,021 7	0,022 2	0,022 8
-1,9	0,023 3	0,023 9	0,024 4	0,025 0	0,025 6	0,026 2	0,026 8	0,027 4	0,028 1	0,028 7
-1,8	0,029 4	0,030 1	0,030 7	0,031 4	0,032 2	0,032 9	0,033 6	0,034 4	0,035 1	0,035 9
-1,7	0,036 7	0,037 5	0,038 4	0,039 2	0,040 1	0,040 9	0,041 8	0,042 7	0,043 6	0,044 6
-1,6	0,045 5	0,046 5	0,047 5	0,048 5	0,049 5	0,050 5	0,051 6	0,052 6	0,053 7	0,054 8
-1,5	0,055 9	0,057 1	0,058 2	0,059 4	0,060 6	0,061 8	0,063 0	0,064 3	0,065 5	0,066 8
-1,4	0,068 1	0,069 4	0,070 8	0,072 1	0,073 5	0,074 9	0,076 4	0,077 8	0,079 3	0,080 8
-1,3	0,082 3	0,083 8	0,085 3	0,086 9	0,088 5	0,090 1	0,091 8	0,093 4	0,095 1	0,096 8
-1,2	0,098 5	0,100 3	0,102 0	0,103 8	0,105 6	0,107 5	0,109 3	0,111 2	0,113 1	0,115 1
-1,1	0,117 0	0,119 0	0,121 0	0,123 0	0,125 1	0,127 1	0,129 2	0,131 4	0,133 5	0,135 7
-1,0	0,137 9	0,140 1	0,142 3	0,144 6	0,146 9	0,149 2	0,151 5	0,153 9	0,156 2	0,158 7
-0,9	0,161 1	0,163 5	0,166 0	0,168 5	0,171 1	0,173 6	0,176 2	0,178 8	0,181 4	0,184 1
-0,8	0,186 7	0,189 4	0,192 2	0,194 9	0,197 7	0,200 5	0,203 3	0,206 1	0,209 0	0,211 9
-0,7	0,214 8	0,217 7	0,220 6	0,223 6	0,226 6	0,229 6	0,232 7	0,235 8	0,238 9	0,242 0
-0,6	0,245 1	0,248 3	0,251 4	0,254 6	0,257 8	0,261 1	0,264 3	0,267 6	0,270 9	0,274 3
-0,5	0,277 6	0,281 0	0,284 3	0,287 7	0,291 2	0,294 6	0,298 1	0,301 5	0,305 0	0,308 5
-0,4	0,312 1	0,315 6	0,319 2	0,322 8	0,326 4	0,330 0	0,333 6	0,337 2	0,340 9	0,344 6
-0,3	0,348 3	0,352 0	0,355 7	0,359 4	0,363 2	0,366 9	0,370 7	0,374 5	0,378 3	0,382 1
-0,2	0,385 9	0,389 7	0,393 6	0,397 4	0,401 3	0,405 2	0,409 0	0,412 9	0,416 8	0,420 7
-0,1	0,424 7	0,428 6	0,432 5	0,436 4	0,440 4	0,444 3	0,448 3	0,452 2	0,456 2	0,460 2
-0,0	0,464 1	0,468 1	0,472 1	0,476 1	0,480 1	0,484 0	0,488 0	0,492 0	0,496 0	0,500 0
t	+0,00	+0,01	+0,02	+0,03	+0,04	+0,05	+0,06	+0,07	+0,08	+0,09
+0,0	0,500 0	0,504 0	0,508 0	0,512 0	0,516 0	0,519 9	0,523 9	0,527 9	0,531 9	0,535 9
+0,1	0,539 8	0,543 8	0,547 8	0,551 7	0,555 7	0,559 6	0,563 6	0,567 5	0,571 4	0,575 3
+0,2	0,579 3	0,583 2	0,587 1	0,591 0	0,594 8	0,598 7	0,602 6	0,606 4	0,610 3	0,614 1
+0,3	0,617 9	0,621 7	0,625 5	0,629 3	0,633 1	0,636 8	0,640 6	0,644 3	0,648 0	0,651 7
+0,4	0,655 4	0,659 1	0,662 8	0,666 4	0,670 0	0,673 6	0,677 2	0,680 8	0,684 4	0,687 9
+0,5	0,691 5	0,695 0	0,698 5	0,701 9	0,705 4	0,708 8	0,712 3	0,715 7	0,719 0	0,722 4
+0,6	0,725 7	0,729 1	0,732 4	0,735 7	0,738 9	0,742 2	0,745 4	0,748 6	0,751 7	0,754 9
+0,7	0,758 0	0,761 1	0,764 2	0,767 3	0,770 4	0,773 4	0,776 4	0,779 4	0,782 3	0,785 2
+0,8	0,788 1	0,791 0	0,793 9	0,796 7	0,799 5	0,802 3	0,805 1	0,807 8	0,810 6	0,813 3
+0,9	0,815 9	0,818 6	0,821 2	0,823 8	0,826 4	0,828 9	0,831 5	0,834 0	0,836 5	0,838 9
+1,0	0,841 3	0,843 8	0,846 1	0,848 5	0,850 8	0,853 1	0,855 4	0,857 7	0,859 9	0,862 1
+1,1	0,864 3	0,866 5	0,868 6	0,870 8	0,872 9	0,874 9	0,877 0	0,879 0	0,881 0	0,883 0
+1,2	0,884 9	0,886 9	0,888 8	0,890 7	0,892 5	0,894 4	0,896 2	0,898 0	0,899 7	0,901 5
+1,3	0,903 2	0,904 9	0,906 6	0,908 2	0,909 9	0,911 5	0,913 1	0,914 7	0,916 2	0,917 7
+1,4	0,919 2	0,920 7	0,922 2	0,923 6	0,925 1	0,926 5	0,927 9	0,929 2	0,930 6	0,931 9
+1,5	0,933 2	0,934 5	0,935 7	0,937 0	0,938 2	0,939 4	0,940 6	0,941 8	0,942 9	0,944 1
+1,6	0,945 2	0,946 3	0,947 4	0,948 4	0,949 5	0,950 5	0,951 5	0,952 5	0,953 5	0,954 5
+1,7	0,955 4	0,956 4	0,957 3	0,958 2	0,959 1	0,959 9	0,960 8	0,961 6	0,962 5	0,963 3
+1,8	0,964 1	0,964 9	0,965 6	0,966 4	0,967 1	0,967 8	0,968 6	0,969 3	0,969 9	0,970 6
+1,9	0,971 3	0,971 9	0,972 6	0,973 2	0,973 8	0,974 4	0,975 0	0,975 6	0,976 1	0,976 7
+2,0	0,977 2	0,977 8	0,978 3	0,978 8	0,979 3	0,979 8	0,980 3	0,980 8	0,981 2	0,981 7
+2,1	0,982 1	0,982 6	0,983 0	0,983 4	0,983 8	0,984 2	0,984 6	0,985 0	0,985 4	0,985 7
+2,2	0,986 1	0,986 4	0,986 8	0,987 1	0,987 5	0,987 8	0,988 1	0,988 4	0,988 7	0,989 0
+2,3	0,989 3	0,989 6	0,989 8	0,990 1	0,990 4	0,990 6	0,990 9	0,991 1	0,991 3	0,991 6
+2,4	0,991 8	0,992 0	0,992 2	0,992 5	0,992 7	0,992 9	0,993 1	0,993 2	0,993 4	0,993 6
+2,5	0,993 8	0,994 0	0,994 1	0,994 3	0,994 5	0,994 6	0,994 8	0,994 9	0,995 1	0,995 2
+2,6	0,995 3	0,995 5	0,995 6	0,995 7	0,995 9	0,996 0	0,996 1	0,996 2	0,996 3	0,996 4
+2,7	0,996 5	0,996 6	0,996 7	0,996 8	0,996 9	0,997 0	0,997 1	0,997 2	0,997 3	0,997 4
+2,8	0,997 4	0,997 5	0,997 6	0,997 7	0,997 7	0,997 8	0,997 9	0,997 9	0,998 0	0,998 1
+2,9	0,998 1	0,998 2	0,998 2	0,998 3	0,998 4	0,998 4	0,998 5	0,998 5	0,998 6	0,998 6
+3,0	0,998 7	0,998 7	0,998 7	0,998 8	0,998 8	0,998 9	0,998 9	0,998 9	0,999 0	0,999 0

Table de l'écart réduit (loi normale)

La table donne la probabilité α pour que l'écart réduit égale ou dépasse, en valeur absolue, une valeur donnée ε , c'est à dire la probabilité extérieure à l'intervalle $[-\varepsilon; \varepsilon]$



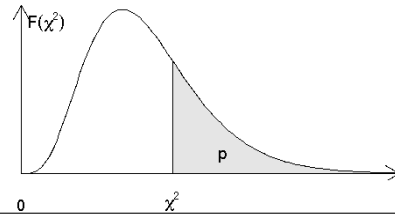
α	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	∞	2,575 829	2,326 348	2,170 090	2,053 749	1,959 964	1,880 794	1,811 911	1,750 686	1,695 398
0,1	1,644 854	1,598 193	1,554 774	1,514 102	1,475 791	1,439 531	1,405 072	1,372 204	1,340 755	1,310 579
0,2	1,281 552	1,253 565	1,226 528	1,200 359	1,174 987	1,150 349	1,126 391	1,103 063	1,080 319	1,058 122
0,3	1,036 433	1,015 222	0,994 458	0,974 114	0,954 165	0,934 589	0,915 365	0,896 473	0,877 896	0,859 617
0,4	0,841 621	0,823 894	0,806 421	0,789 192	0,772 193	0,755 415	0,738 847	0,722 479	0,706 303	0,690 309
0,5	0,674 490	0,658 838	0,643 345	0,628 006	0,612 813	0,597 760	0,582 842	0,568 051	0,553 385	0,538 836
0,6	0,524 401	0,510 073	0,495 850	0,481 727	0,467 699	0,453 762	0,439 913	0,426 148	0,412 463	0,398 855
0,7	0,385 320	0,371 856	0,358 459	0,345 126	0,331 853	0,318 639	0,305 481	0,292 375	0,279 319	0,266 311
0,8	0,253 347	0,240 426	0,227 545	0,214 702	0,201 893	0,189 118	0,176 374	0,163 658	0,150 969	0,138 304
0,9	0,125 661	0,113 039	0,100 434	0,087 845	0,075 270	0,062 707	0,050 154	0,037 608	0,025 069	0,012 533

Table pour les petites valeurs de la probabilité

α	ε
0,002	3,090 23
0,001	3,290 53
0,000 1	3,890 59
0,000 01	4,417 17
0,000 001	4,891 64
0,000 000 1	5,326 72
0,000 000 01	5,730 73

Table du χ^2 de Pearson

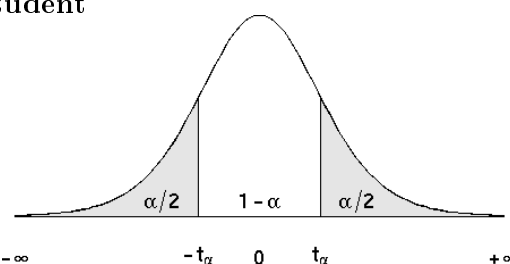
Valeurs de χ^2 ayant la probabilité p d'être dépassées



$\nu \backslash p$	0,99	0,975	0,95	0,9	0,1	0,05	0,025	0,01
1	0,000 16	0,000 98	0,003 93	0,015 8	2,71	3,84	5,02	6,63
2	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21
3	0,11	0,22	0,35	0,58	6,25	7,81	9,35	11,34
4	0,30	0,48	0,71	1,06	7,78	9,49	11,14	13,28
5	0,55	0,83	1,15	1,61	9,24	11,07	12,83	15,09
6	0,87	1,24	1,64	2,20	10,64	12,59	14,45	16,81
7	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48
8	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09
9	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67
10	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21
11	3,05	3,82	4,57	5,58	17,28	19,68	21,92	24,72
12	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22
13	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69
14	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14
15	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58
16	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00
17	6,41	7,56	8,67	10,09	24,77	27,59	30,19	33,41
18	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81
19	7,63	8,91	10,12	11,65	27,20	30,14	32,85	36,19
20	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57
21	8,90	10,28	11,59	13,24	29,62	32,67	35,48	38,93
22	9,54	10,98	12,34	14,04	30,81	33,92	36,78	40,29
23	10,20	11,69	13,09	14,85	32,01	35,17	38,08	41,64
24	10,86	12,40	13,85	15,66	33,20	36,42	39,36	42,98
25	11,52	13,12	14,61	16,47	34,38	37,65	40,65	44,31
26	12,20	13,84	15,38	17,29	35,56	38,89	41,92	45,64
27	12,88	14,57	16,15	18,11	36,74	40,11	43,19	46,96
28	13,56	15,31	16,93	18,94	37,92	41,34	44,46	48,28
29	14,26	16,05	17,71	19,77	39,09	42,56	45,72	49,59
30	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89

Table du t de Student

La table donne la probabilité α pour que t égale ou dépasse, en valeur absolue, une valeur donnée t_α , c'est à dire la probabilité extérieure à l'intervalle $[-t_\alpha; t_\alpha]$



d.d.l. \ α	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,599
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,768
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,254	0,387	0,526	0,678	0,846	1,043	1,292	1,664	1,990	2,374	2,639	3,416
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,675	0,842	1,036	1,282	1,645	1,960	2,327	2,576	3,291

Table des valeurs critiques pour le test de Mann-Whitney-Wilcoxon. (p=0.05, bilatéral)

La table donne les valeurs critiques u_{theo} telles que, pour deux échantillons A et B , $\min(u_A, u_B) \leq u_{theo}$ entraîne le rejet de H_0 au risque 0.05.

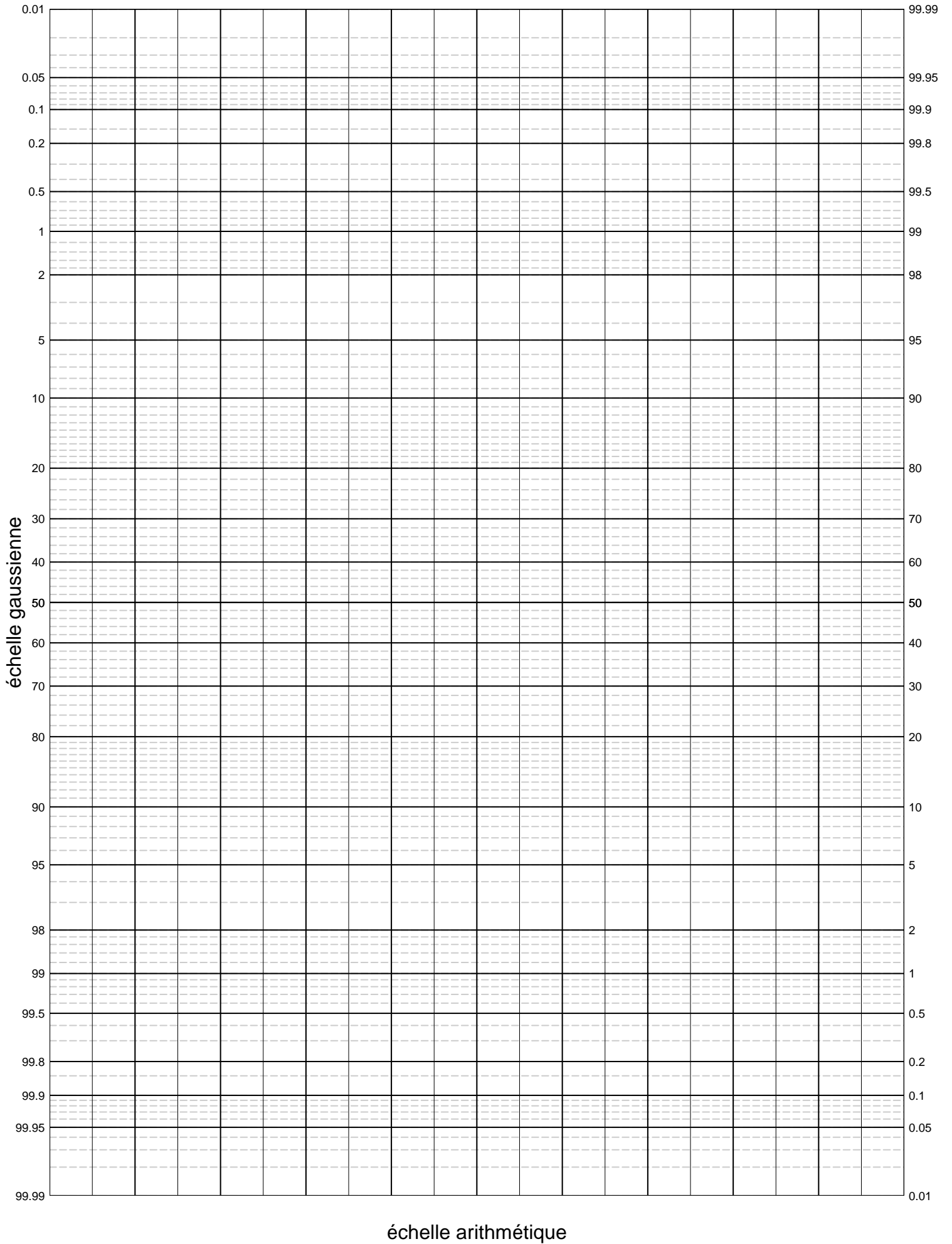
$n_A \backslash n_B$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	0	0	0	0	1	1	1	1	1	2	2	2	2
3	-	-	-	-	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	-	-	-	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
5	-	-	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	-	-	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	-	-	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	-	0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	-	0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10	-	0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	-	0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	-	1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	-	1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	-	1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	69	74	78	83
15	-	1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	-	1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	-	2	6	11	17	22	28	34	39	45	51	57	63	69	75	81	87	93	99	105
18	-	2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	-	2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	-	2	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

Table des valeurs critiques pour le test des rangs signés de Wilcoxon. (p=0.05, bilatéral)

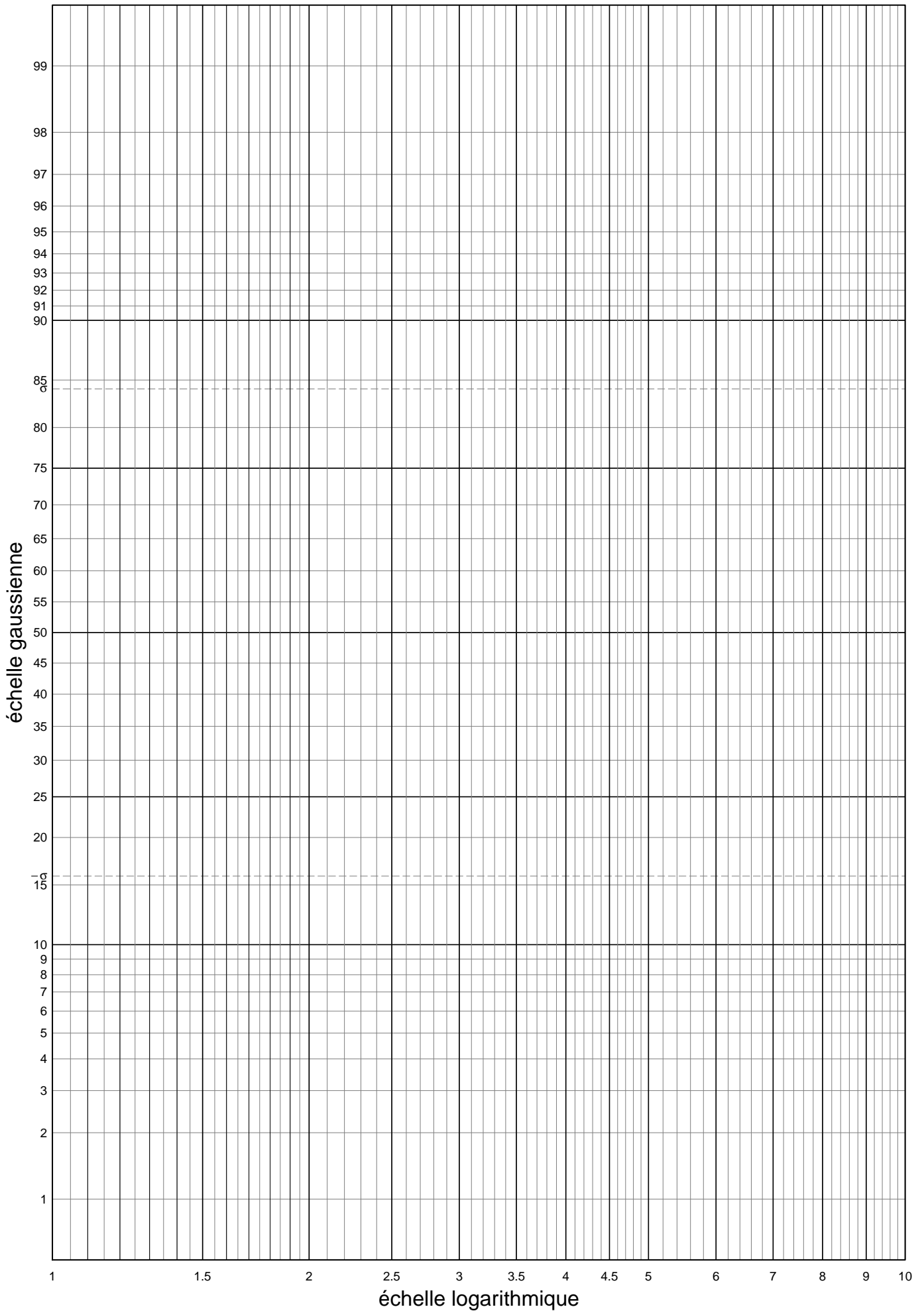
Soient p et n les sommes des rangs des différences positives et négatives. La table donne les valeurs critiques w_{theo} telles que $\min(p, n) \leq w_{theo}$ entraîne le rejet de H_0 .

Effectif	Seuil
6	0
7	2
8	3
9	5
10	8
11	10
12	13
13	17
14	21
15	25
16	29
17	34
18	40
19	46
20	52

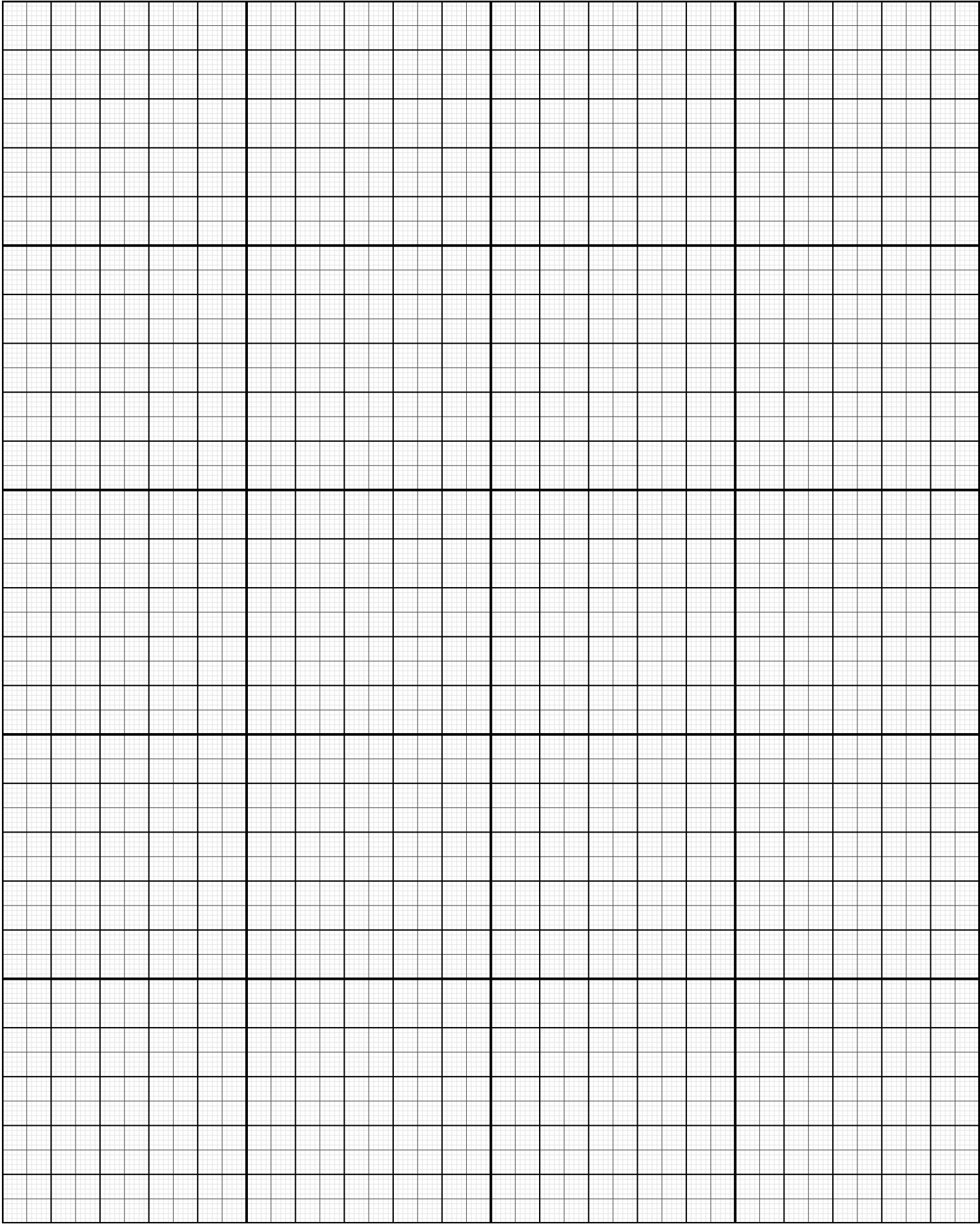
Papier gausso-arithmétique



Papier gaussio-logarithmique



Papier millimétré



Nom : _____

Prénom : _____

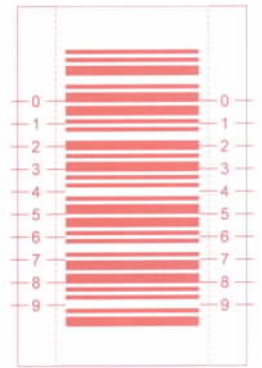
Date de naissance : _____

Examen : _____

Épreuve : _____

Date : _____

1



La manière de remplir ce questionnaire fait partie intégrante de l'épreuve . Les réponses aux QCM étant saisies par lecture optique, veuillez vous conformer strictement aux consignes suivantes :

- 1- Utilisez un stylo à bille noir ou un feutre noir.
- 2- Faites des marques franches dans les cases en regard des lettres choisies (A,B,C,D,E)
 - sur la ligne supérieure (V) pour les propositions jugées exactes ;
 - sur la ligne inférieure (F) pour les propositions jugées fausses.
- 3- En cas de changement de réponse, ne cherchez pas à effacer, mais utilisez la zone de reprise (R). Dans ce cas, répondez à nouveau à toutes les propositions.

Marquages corrects, si vous voulez répondre VRAI aux points A et C, et FAUX en B, D et E (après correction)

NE PAS FAIRE



Réservé



1

V A B C D E

F

R V A B C D E

F

11

V A B C D E

F

R V A B C D E

F

21

V A B C D E

F

R V A B C D E

F

31

V A B C D E

F

R V A B C D E

F

2

V A B C D E

F

R V A B C D E

F

12

V A B C D E

F

R V A B C D E

F

22

V A B C D E

F

R V A B C D E

F

32

V A B C D E

F

R V A B C D E

F

3

V A B C D E

F

R V A B C D E

F

13

V A B C D E

F

R V A B C D E

F

23

V A B C D E

F

R V A B C D E

F

33

V A B C D E

F

R V A B C D E

F

4

V A B C D E

F

R V A B C D E

F

14

V A B C D E

F

R V A B C D E

F

24

V A B C D E

F

R V A B C D E

F

34

V A B C D E

F

R V A B C D E

F

5

V A B C D E

F

R V A B C D E

F

15

V A B C D E

F

R V A B C D E

F

25

V A B C D E

F

R V A B C D E

F

35

V A B C D E

F

R V A B C D E

F

6

V A B C D E

F

R V A B C D E

F

16

V A B C D E

F

R V A B C D E

F

26

V A B C D E

F

R V A B C D E

F

36

V A B C D E

F

R V A B C D E

F

7

V A B C D E

F

R V A B C D E

F

17

V A B C D E

F

R V A B C D E

F

27

V A B C D E

F

R V A B C D E

F

37

V A B C D E

F

R V A B C D E

F

8

V A B C D E

F

R V A B C D E

F

18

V A B C D E

F

R V A B C D E

F

28

V A B C D E

F

R V A B C D E

F

38

V A B C D E

F

R V A B C D E

F

9

V A B C D E

F

R V A B C D E

F

19

V A B C D E

F

R V A B C D E

F

29

V A B C D E

F

R V A B C D E

F

39

V A B C D E

F

R V A B C D E

F

10

V A B C D E

F

R V A B C D E

F

20

V A B C D E

F

R V A B C D E

F

30

V A B C D E

F

R V A B C D E

F

40

V A B C D E

F

R V A B C D E

F

Réponses aux questions 1 à 40 - Faites des marques franches (tirets horizontaux nets)

Réponses aux questions 1 à 40 - Faites des marques franches (tirets horizontaux nets)

Réponses aux questions 1 à 40 - Faites des marques franches (tirets horizontaux nets)