2019

Exercice 1 :

On donnera une valeur approche de tous les résultats à 10^{-4} près.

Une étude cherche à estimer le pourcentage des femmes (entre 35 et 65 ans) ayant un cancer de sein.

Dans cette population de n femmes, on associe à chaque femme i ainsi étudiée, une variable aléatoire X_i. prenant la valeur 1 si la femme a un cancer (probabilité p) ; 0 dans le cas contraire (probabilité $1-p = q$).

1. Nombre de femme ayant un cancer de sein dans une population de n femme suit :

a. Une loi de Poison de paramètre $\lambda = p$.
b. Une loi normale $N(\mu = p, \sigma^2 = \frac{pq}{n})$.
c. Une loi normale $N(\mu = n.p, \sigma^2 = p.q)$.
d. Une loi binomiale $B(n.p, n.p.q)$.
e. Une loi binomiale $B(n,p)$.

2. Dans un échantillon de 100 femmes tiré de cette population, on a observé 70 cas de cancer de sein. Le pourcentage (théorique) des femmes (entre 35 et 65 ans) ayant un cancer de sein est estime par :

a. 7/1000
b. 70/1000
c. 70%
d. 0.21
e. La taille de l'échantillon ne permet pas d'estimer.

3. L'intervalle de confiance à 95% de proportion de femmes ayant un cancer de sein est :

a. [0.0451 - 0.0949]
b. [0.0541 - 0.0859]
c. [0.055 - 0.095]
d. [0.095 - 0.098]
e. Le nombre de femme ne permet pas d'estimer cet intervalle de confiance.

4. L'intervalle de confiance ainsi défini signifie que :

a. 95% des femmes de l'échantillon sont compris entre les bornes de l'intervalle.
b. 5% des femmes de l'échantillon sont compris entre les bornes de l'intervalle.
c. Le vrai pourcentage de la population cible de l'étude à 95 chances sur 100 de situer en dehors de l'intervalle.
d. Le vrai pourcentage de la population cible de l'étude à 95 chances sur 100 de situer dans l'intervalle.
e. Le vrai pourcentage de la population cible de l'étude à 5 chances sur 100 de situer dans l'intervalle.

5. Combien de femmes de cet intervalle d'âge faut-il étudier pour avoir une précision d'estimation (amplitude IC) = 1% au risque 5%

a. 958
b. 1245
c. 1763
d. 2154
e. 2504.

6. Le cancer du sein touche en effet 6% de femmes entre 35 et 65 ans. La fréquence du cancer de sein observée dans l'échantillon est-elle différente de la fréquence connue dans la population générale ? Pour cela on effectue :

a. Une comparaison de deux pourcentages observés.
b. Une comparaison de moyennes d'échantillons indépendants.
c. Une comparaison entre un pourcentage observé et un pourcentage théorique.
d. Une comparaison entre une moyenne observée et une moyenne théorique.
e. Une comparaison d'effectifs.

7. Le résultat du test (variable de décision) est :

a. $Z = 1.3333,$

b. $Z = 1.2346$

c. $x^2 = 19.80$

d. $t_{ddl} = 4.45$

e. Les conditions de validité du test ne sont pas remplies.

8. Que peut-on conclure (niveau de confiance 95%) :

a. Les résultats de l'échantillon sont conforme aux résultats de la population.
b. Les résultats de l'échantillon ne sont conforme aux résultats de la population.
c. L'échantillon est homogène.
d. L'échantillon est hétérogène.
e. Les conditions de validité du test ne sont pas remplies.

Exercice 2 :

On donnera une valeur approchée de tous les résultats à 10^{-4} prés.

Les résultats observés de l'évolution d'une certaine maladie à la suite de l'emploi de l'un ou l'autre des traitements A et B pour 1000 patients figurent dans le tableau ci-dessous :

Traitement Effet	Guérison	Amélioration	Etat stationnaire	Totale
A	280	210	110	600
B	220	90	90	400
Total	500	300	200	1000

9. Dans l'étude ; on test l'association entre :

a. Une variable qualitative binaire et une variable quantitative continue.
b. Deux variables qualitatives binaires.
c. Deux variables quantitatives continues.
d. Deux variables quantitatives discrètes.
e. Deux variables qualitatives.

10. Pour étudier l'existence d'une relation entre les traitements et leurs efficacités, on peut utiliser :

a. Test de conformité.
b. Test de corrélation.
c. Test d'ANOVA.
d. Test d'indépendance de x^2.
e. Test de Fisher d'égalité de 2 variances.

11. Quelle est la valeur de la statistique calculée :

a. 12.0147
b. 14.125
c. 17.9166
d. 18.2365
e. 30.1824.

12. Quelle sont les degrés de liberté :

a. 0
b. 1
c. 2
d. 4
e. 6.

13. Quelle est la valeur de la statistique théorique : (niveau de confiance = 95%)

a. 2.71
b. 3.84
c. 4.61
d. 5.99
e. 12.59.

Exercice 3 :
On donnera une valeur approchée de tous les résultats à 10^{-3} près.
On veut étudier l'effet de deux médicaments sur le taux de lymphocytes d'animaux de laboratoires. On construit un plan factoriel dans lequel il y a trois groupes d'animaux d'effectifs 10 animaux par groupe. On garde un des groupes comme témoin et l'on administre les médicaments A et B aux deux autres groupes.

Groupe témoin	272, 193, 432, 259, 386, 349, 320, 247, 260, 478
Groupe traité par A	468, 383, 375, 398, 534, 451, 474, 278, 255, 528
Groupe traité par B	368, 290, 325, 298, 314, 350, 378, 321, 275, 401

Les données correspondent au modèle d'ANOVA : une variable de groupe, une variable continue dont on veut comparer les moyennes.
(Indications numériques $\sum_j x_{1j}$=3196, $\sum_j x_{2j}$=4094, $\sum_j x_{3j}$=3320 (somme de chaque ligne).

14. La taille globale des 3 échantillon est :
a. 10 b. 20 c. 30 d. 40 e. 60.

15. La moyenne globale est :
a. 256.667 b. 353.337 c. 415.325 d. 435,956 e. 563.754.

16. variabilité expliqué SSL est :
a. 47361.9 b. 51426.845 c. 54211.171 d. 62516.539 e. 65785.759.

17. Quelle sont les degrés de liberté :
a. (2,30) b. (2,27) c. (3,27) d. (3,30) e. (29,2).

18. Quelle est la valeur de la statistique calculée F sachant que SSR=176130 :
a. 3.63 b. 2.42 c. 4.003 d. 2.689 e. 6.84.

19. la valeur de la statistique théorique F : (niveau de confiance = 95%)
a. 2.96 b. 2.922 c. 3.354 d. 4.61 e. 12.59.

20. A quoi correspond le risque alpha :
a. A la probabilité de conclure à une différence significative
b. A la probabilité de conclure à tort à une absence de différence significative.
c. A la probabilité de ne pas conclure H1 alors que H1 est vraie.
d. A la probabilité d'accepter H0 alors que H0 est vraie.
e. A la probabilité de rejeter H0 alors que H0 est vraie.

NB :
- Pour une loi normale centrée réduite : $U_{0.975}$= 1.96, $U_{0.95}$ = 1.64, $U_{0.99}$=2.32
- Pour une loi de Student : t(7,0.975)= 2.365, t(7,0.95)= 1.985.
- Pour une loi du Khi-deux : $x^2_{0.95}(1)$= 3.84, $x^2_{0.95}(2)$= 5.99, $x^2_{0.95}(6)$= 12.59.
- Pour une loi du Fisher : $F_{0.95}(2,27)$= 3.354, $F_{0.95}(2,30)$= 3.316, $F_{0.95}(3,27)$= 2.96, $F_{0.95}(3,30)$= 2.922.

2020

Exercice 1 : **On donnera une valeur approchée de tous les résultats A 10^{-2} près.**

En vue de réaliser un programme de rééducation, des chercheurs ont soumis un questionnaire de neuropsychologue cognitive à 150 enfants dyslexiques tirés au sort. Le questionnaire comporte 20 questions et les chercheurs ont recueilli pour chaque enfant dyslexique le nombre x, de bonnes réponses. Les résultats ainsi récoltés sont tels que : $\sum x_i$= 1502 ; $\sum x_i^2$= 19486.

1. La population statistique étudiée est :

a. Les chercheurs.
b. Le programme de rééducation.
c. Les enfants dyslexiques.
d. Le questionnaire de neuropsychologie.
e. Le nombre de bonnes réponses.

2. La variable statistique X étudiée est :

a. Les chercheurs.
b. Le programme de rééducation.
c. Les enfants dyslexiques.
d. Le questionnaire de neuropsychologie.
e. Le nombre de bonnes réponses.

3. Les valeurs possibles de X sont :

a. Une seule valeur $\{20\}$
b. $\{0,1,\ldots,20\}$
c. $\{0,1,\ldots,50\}$
d. $\{juste, fausse\}$
e. a, b, c et d sont faux

4. La taille de l'échantillon est :

a. 20. b. 150. c. 1502. d. 3000. e. 19486.

5. L'estimation ponctuelle sans biais de la moyenne (µ) de la variable dans la population est :

a. 15,02 b. 10,8 c. 10,01 d. 75,01 e. 79,05.

6. L'estimation ponctuelle sans biais de la variance de la variable dans la population est :

a. 28,3 b. 29,17 c. 29,7 d. 29,9 e. 30,59.

7. La moyenne de la variable dans la population à 95% de chance de situer dans l'intervalle :

a. [9.13 – 10.89]
b. [9.27 – 10.75]
c. [10.01 – 12.29]
d. [3.77 – 16.25]
e. [8.87 – 11.15]

8. Le probabilité que la moyenne de la variable dans la population se situe dans l'intervalle [9.27 – 10.75] est :

a. 0,05 b. 0,1 c. 0,9 d. 0,95 e. 0,99

9. La marge d'erreur dans l'estimation de la moyenne de la variable dans la population au niveau 99% est :

a. 0,99 b. 0,95 c. 0,88 d. 6,24 e. 1,16

Exercice 2 : On donnera une valeur approchée de tous les résultats à 10^{-4} près.

Soit p_1 la probabilité de guérison d'une maladie grâce à un traitement T1. Un groupe de 50 malades est soumis à ce traitement et 28 guérissent.

A. Peut-on dire que la probabilité p_1 de guérison par le traitement T1 est égale à 50% ou bien est supérieure à 50% ?

10. Pour cela on effectue :

a. Une comparaison de deux pourcentages observés.
b. Une comparaison de deux moyennes d'échantillons indépendants.
c. Une comparaison entre un pourcentage observé et un pourcentage théorique.
d. Une comparaison on entre une moyenne observée et une moyenne théorique.
e. Une comparaison d'effectifs.

11. Le résultat du test (variable de décision) est :

a. Z = 0,8571 b. Z = 0,8487 c. $x^2 = 3{,}84$ d. t = 4,45 e. F = 2.6

12. Que peut-on conclure (niveau de confiance = 95%)

a. L'échantillon est homogène.
b. L'échantillon est hétérogène.
c. Les résultats de l'échantillon ne sont conforme aux résultats théoriques.
d. Les résultats de l'échantillon sont conforme aux résultats théoriques.
e. Les conditions de validité du test ne sont pas remplies.

B. On s'intéresse maintenant à un nouveau traitement T2 permettant de soigner cette maladie. Sur 60 malades soumis à ce nouveau traitement. 34 guérissent. On se demande s'il y a une différence significative entre T1 et T2 quant à leur efficacité

13. Pour cela on effectue :

a. Une comparaison de deux pourcentages observés.
b. Une comparaison de deux moyennes d'échantillons indépendants.
c. Une comparaison entre un pourcentage observé et un pourcentage théorique.
d. Une comparaison on entre une moyenne observée et une moyenne théorique.
e. Une comparaison de deux variances.

14. Parmi conditions de validité du test, il y a :

a. Les variances des 2 populations sont égales.
b. Les distributions de la variable dans les populations d'où sont tires les échantillons donnent souvent une loi de Poisson.
c. $n * p_{th} \geq 5$ et $n * (1 - p_{th}) \geq 5$.
d. $n_1 * p_2 \geq 5$ et $n_1 * (1 - p_1) \geq 5, n_2 * p_2 \geq 5$ et $n_2 * (1 - p_2) \geq 5$.
e. Utilisable si petits effectifs.

15. Que peut-on conclure (niveau de confiance= 95)

a. Il y a une relation entre les deux traitements.
b. Les résultats de l'échantillon ne sont pas conforme aux résultats de la population.
c. Les résultats donnent une différence significative.
d. Les résultats donnent une différence non significative.
e. Les conditions de validité du test ne sont pas remplies.

NB :

1. Z suit une loi normale centrée réduite $N(0,1)$. Alors $P(Z \leq 1.28) = 0.9$, $P(Z \leq 1.64) = 0.95$, $P(Z \leq 1.96) = 0{,}975$, $P(Z \leq 2.33) = 0.99$, $P(Z \leq 2.58) = 0.995$.
2. Pour une loi de Student à 20 degrés de liberté : le fractile d'ordre 0.975 vaut 2.086, et le fractile d'ordre 0.95 vaut 1.725.
3. Pour une loi de Student à 19 degrés de liberté : le fractile d'ordre 0.975 vaut 2.093, et le fractile d'ordre 0.95 vaut 1.729.
4. Pour une loi du Khi-deux à un degré de liberté : le fractile d'ordre 0.95 vaut 3.84, le fractile d'ordre 0.90 vaut 2.71.
5. Pour une loi du Khi-deux à deux degrés de liberté : le fractile d'ordre 0.95 Vaut 5.99, le fractile d'ordre 0.90 vaut 4.61.

2021

Exercice 1 : On donner une valeur approchée de tous les résultats à 10^{-5} prés.

Le pourcentage d'anomalies chromosomique dans les naissances d'une population donnée, était de 1% il y 10 ans.

1. Le nombre des naissances avec des anomalies chromosomiques parmi 100 naissances (il y a 10 ans) suit :

(A) une loi de poisson de paramètre $\lambda = 0.1$

(B) une loi normale $N(1\ ;\ 0.00099)$

(C) une loi normale $N(1\ ;\ 0.99)$

(D) une loi binomiale $B\ (10\ ;\ 0.01)$

(D) une loi binomiale $B\ (100\ ;\ 0.01)$

2. La probabilité d'avoir aucune naissance avec des anomalies chromosomiques dans cette population d'étude est :

(A) < 0.01 (b) = 0.01 (C) = 0.36603 (D) = 0.90438 (E) > 0.95

En 2021, on effectue un dépistage systématique (obtention des caryotypes à partir de prélèvement de sang) sur 600 naissance tirées au sort dans la population actuelle On observe 7 caryotypes anormaux

3. L'estimation ponctuelle sans biais du pourcentage (théorique) d'anomalies chromosomique dans les naissances de la population actuelle est :

(A) < 0.01 (b) = 0.014 (C) = 0.07 (D) = 0.7 (E) > 0.8

4. L'intervalle de confiance a 95% de proportion d'anomalies chromosomiques dans les naissances de la population actuelle est :

(A) (0.00371 - 0.02429) (B) (0.00839 - 001961) (C) (0,05971 - 0,08029) (D) (0.68971 - 0.71029) (E) Le nombre des nouveaux nées ne permet pas d'estimer cet intervalle de confiance

5. Au seuil de confiance de 95%, pour que la proportion d'anomalies chromosomiques dans la population actuelle appartient à l'intervalle (0.01 - 0.018) il faut observer combien de naissances :

(A) 61 (b) 3451 (C) 16275 (D) > 100000

(E) c'est imposable, car la proportion d'anomalies chromosomiques dans la population actuelle n'appartient pas à cet intervalle de confiance

6. Le pourcentage d'anomalies chromosomique est-il significativement différent qu'il y a 10 ans. Pour cela on effectue :

(A) une comparaison de deux pourcentages observes

(B) une comparaison de deux moyennes d'échantillons indépendantes

(C) une comparaison entre un pourcentage observe et un pourcentage théorique

(D) une comparaison entre une moyenne observe et une moyenne théorique

(E) une comparaison de deux intervalles de confiance.

7. Parmi les conditions de validité du test, il y a :

(A) Les variances des 2 populations sont égales

(B) La loi de la variable dans la population d 'ou sont tirés les échantillons suit loi du Poisson

(C) $n * p_{th} \geq 5$ et $n *(1 - p_{th}) \geq 5$

(D) $n * f \geq 5$ et $n *(1 - f) \geq 5$

(E) Utilisable si petits effectifs

8. Le résultat du test (variable de décision) est :

(A) 0.08988 (B) 0.38873 (C) 0.38873 (D) 0.80808 (E) 3.84

9. La valeurs de la statistique théorique (niveau de confiance = 95%) est :

(A) 0.38873 (B) -1.96 (C) 1.64 (D) 2.58 (E) 3.84

10. Que peut-on conclure (niveau de confiance = 95%) :

(A) Le pourcentage d'anomalies chromosomique est significativement différent qu'il y a 10 ans

(B) Le pourcentage d'anomalies chromosomique est non significativement différent qu'il y a 10 ans

(C) Les résultats de l'échantillon ne sont conforme aux résultats de la population

(D) L'échantillon est hétérogène

(E) Les conditions de validité du test ne sont pas remplies

Exercice 2 : On donnera une valeur approchée de tous les résultats à 10^{-2} près.

On mesure par une note de qualité Q l'efficacité de différentes thérapies pour lutter contre la dépression nerveuse. On trouve les valeurs suivantes :

Y_1 : CO	3,5,4,5,6,7
Y_2 : PS	6,2,3,5,7,7
Y_3 : TE	7,7,6,3,0,1

CO : thérapie cognitive, PS : thérapie psychanalytique, TE: témoin (la thérapie consiste à discuter avec un acteur qui ne connait pas la psychologie). On donne $\overline{Y}_1 = 5$, $\overline{Y}_2 = 5$, $\overline{Y}_3 = 4$

$\sum\sum(Y_{ij}) = 84$ $\quad \sum\sum(Y_{ij})^2 = 462$

11. Dans cette étude, il y a combien de variables :

(A) 1 (B) 2 (C) 3 (D) 4 (E) 5

12. Le variables sont :

(A) tous quantitatives, (B) tous qualitatives, (C) 2 quantitatives et 1 qualitative,

(D) 1 quantitative et 1 qualitative, (E) 1 quantitatives et 2 qualitatives

13. Dans cette étude, il y a combien d'échantillon :

(A) 3 (B) 4 (C) 6 (D) 17 (E) 18

14. L'effectif total est :

(A) 6 (B) 12 (C) 17 (D) 15 (E) 16

15. La moyenne globale est :

(A) 5 (B) 4.94 (C) 4.67 (D) 14 (E) toutes les propositions précédente sont fausses.

16. On se demande s'il y a une différence significative entre les différentes thérapies pour lutter contre la dépression nerveuse. Pour cela on effective :

(A) une comparaison de plusieurs variances d'échantillon indépendantes.

(B) une comparaison de plusieurs moyennes d'échantillon indépendantes.

(C) une comparaison entre plusieurs pourcentages observés.

(D) une comparaison entre deux moyennes observées et une moyenne théorique.

(E) une comparaison entre deux pourcentage observées et un pourcentage théorique.

17. Parmi les conditions de validité du test, il y a :

(A) Les Y_{ij} sont des réalisations de la v.a. Y_{ij} – N(m_i . σ^2)

(B) L'écart-type (théorique) n'est pas le même dans toutes les populations.

(C) Il existe i tel que n_i < 30.

(D) Toutes les variables sont indépendantes les un des autres.

(E) Toutes les propositions précédentes sont fausses.

18. Quelle sont les degrés de liberté :

(A) 16 (B) 18 (C) (3;15) (D) (2;15) (E) (1.16)

19. Quelle est la valeur de la statistique calculée :

(A) 66 (B) 311 (C) 0.03 (D) 0.03 (E) 35.45

20. Supposant que les conditions de test sont vérifiées, et que la statistique théorique égale à 19.43, que peu conclure :

(A) Il y a une relation entre les variables étudiées

(B) Les différentes thérapies ont les mêmes effets pour lutter contre la dépression nerveuse

(C) Les échantillons sont homogènes

(D) Les résultats ne donnent pas une différence significative

(E) Chacune des 3 thérapies est différentes des autres

2022

Exercice 1 : On donnera une valeur approchée de tous les résultats à 10^{-3} près.
On considère le poids (X) des bébés en fonction de sexe (Y) et on obtient pour 5 filles (X_F) : 2.8 ; 3 ; 3 ; 3.2 ; 3.3 et pour 4 garçons (X_G) : 3 ; 3.2 ; 3.4 ; 3.6.
On donne les informations suivantes : $\sum X_F = 15.3$; $\sum X_G = 3.06$; $\sum X_F^2 = 46.97$; $\sum X_G^2 = 43.76$

1. **Dans cette étude, il y a combien de variables :**
a. 1.
b. 2.
c. 3.
d. 4.
e. 5.

2. **Pour comparer graphiquement deux sous population (le groupe des filles avec le groupe des garçons), on utilise :**
a. Nuage des points.
b. La courbe cumulatif.
c. Boite à moustache.
d. Diagramme en bétons.
e. Diagramme en secteur.

3. **On souhaite étudier l'influence du sexe sur le poids des bébés. Pour cela on effectue :**
a. Un test de conformité.
b. Un test d'homogénéité.
c. Le test d'ANOVA.
d. Un test d'ajustement.
e. Un test de corrélation.

4. **La statistique de test utiliser est :**
a. La statistique de Student 7.
b. La statistique de Z.
c. La statistique de Khi-deux.
d. La statistique de Fisher.
e. Toutes les propositions précédentes sont fausses.

5. **Quelle est la valeur de la statistique théorique (niveau de confiance = 95%)**
a. 1.96
b. 5.59
c. 3.84
d. 2.365
e. 1.895

6. **Quelle est la valeur de la statistique calculée :**
a. 2.759
b. 1.763
c. 1.6
d. 1.519
e. 1.41

7. **Supposant que les conditions de test sont vérifiées, que peut-on concluant :**
a. Les deux populations sont hétérogènes.
b. Le modèle est pertinent (globalement significatif).
c. Les résultats de l'échantillon ne sont pas conforme aux résultats de la population.
d. Il y a une relation entre les variables étudiées
e. Les variables sont indépendantes.

Exercice 2 : On donnera une valeur approchée de tous les résultats à 10^{-3} près.
Un médicament est injecté en intraveineuse. Dans les heures qui suivent, la substance est éliminée par les reins. La quantité q_i (en mg) présente dans le sang à l'instant t_i (en h) a été mesurée par des prises de sang toutes les heures :

Temps t_i en h	0	2	4	6	8
Quantité q_i en mg	9.9	7.5	5.5	3.1	3

On donne les informations suivantes : $\sum t_i = 20$; $\sum t_i^2 = 120$; $\sum q_i = 29$; $\sum q_i^2 = 203.12$; $\sum t_i\, q_i = 79.6$

8. Le but principal de cette étude est de chercher s'il y a une :

a. Relation entre deux variables quantitatives.
b. Relation entre une variable quantitative et une variable qualitative.
c. Relation entre deux variables qualitative.
d. Relation entre plusieurs variables quantitatives.
e. Relation entre deux variable quantitatives et une variable qualitative.

9. Si l'équation de la droite de régression de q_i en fonction de t_i est donnée par : $q_i = -0.91\, t_i + 9.44$, alors la covariance estimée entre t_i et q_i ($s'_{t_i q}$) est égale à :

a. -0.974 b. 0.974 c. -1 d. -9.1 e. -7.28

10. La somme des erreurs est :

a. 0 b. -0.04 c. 0.04 d. -0.008 e. 0.2

11. Le MSE (la variance estimée des erreurs) est :

a. 0.008 b. 0.0016 c. 0 d. 0.013 e. 0.599

12. La qualité de l'ajustement de q_i en fonction de t_i est :

a. -1 b. 0 c. 0.948 d. 0.97 e. 1

13. Pour valider le modèle :

a. On calcule l'intervalle de confiance de la pente
b. On effectue un test de Fisher de signification globale
c. On dessiner le nuage des point
d. On vérifie que les erreurs suit une loi normale centrée
e. On effectue un test de comparaison de deux moyennes observées

14. Pour le test de corrélation (test de la pente), le (les) degré(s) de liberté est (sont) :

a. 5 b. 4 c. 3 d. (1,4) e. (1.3)

15. Pour $x_6 = 5$, l'intervalle de confiance de prédiction de Y_6 est égale à :

a. (4.027 - 5.753) b. (2.844 - 6.906) c. (2.274 - 7.726) d. (3.723 - 6.057) e. (2.164 - 7.616)

NB :

1. Z suit une loi normale centrée réduite $N(0,1)$, alors : $P(Z \leq 1.28) = 0.9$, $P(Z \leq 1.64) = 0.95$, $P(Z \leq 1.96) = 0.975$, $P(Z \leq 2.33) = 0.99$, $P(Z < 2.58) = 0.995$.
2. Pour une loi de Student à 3 degrés de liberté : le fractile d'ordre 0.975 vaut 3.182, et le fractile d'ordre 0.95 vaut 2.353.
3. Pour une loi de Student à 4 degrés de liberté : le fractile d'ordre 0.975 vaut 2.777, et le fractile d'ordre 0.95 vaut 2.132.
4. Pour une loi de Student à 5 degrés de liberté : le fractile d'ordre 0.975 vaut 2.571, et le fractile d'ordre 0.95 vaut 2.015.
5. Pour une loi de Student à 7 degrés de liberté : le fractile d'ordre 0.975 vaut 2.365, et le fractile d'ordre 0.95 vaut 1.895.
6. Pour une loi de Student à 8 degrés de liberté : le fractile d'ordre 0.975 vaut 2.306, et le fractile d'ordre 0.95 vaut 1.86.
7. Pour une loi du Khi-deux à un degré de liberté : le fractile d'ordre 0.95 vaut 3.84, le fractile d'ordre 0,90 vaut 2.71.
8. Pour une loi du Khi-deux à deux degrés de liberté : le fractile d'ordre 0.95 vaut 5.99, le fractile d'ordre 0.90 vaut 4.61.
9. Pour une loi du Khi-deux à 4 degrés de liberté : le fractile d'ordre 0.95 vaut 9.49, le fractile d'ordre 0.975 vaut 11.14.
10. Pour une loi du Fisher à (4, 5) degrés de liberté : le fractile d'ordre 0.95 vaut 5.19, le fractile d'ordre 0.975 vaut 7.39.
11. Pour une loi du Fisher à (1, 7) degrés de liberté : le fractile d'ordre 0.95 vaut 5.59, le fractile d'ordre 0.975 vaut 8.07.
12. Pour une loi du Fisher à (1, 4) degrés de liberté : le fractile d'ordre 0.95 vaut 7.71, le fractile d'ordre 0.975 vaut 12.5.
13. Pour une loi du Fisher à (1, 3) degrés de liberté : le fractile d'ordre 0.95 vaut 10.1, le fractile d'ordre 0.975 vaut 17.4.
14. Pour une loi du Fisher à (4, 3) degrés de liberté : le fractile d'ordre 0.95 vaut 6.59, le fractile d'ordre 0.975 vaut 15.1.
15. Pour une loi du Fisher à (3, 4) degrés de liberté : le fractile d'ordre 0.95 vaut 9.12, le fractile d'ordre 0.975 vaut 9.98.

2023

Exercice 1 :

1. Pour comparer graphiquement les caractéristiques statistiques d'une variable quantitative dans deux échantillons on utilise :

(A) Nuage des points.
(B) La courbe cumulatif.
(C) Boite à moustache.
(D) polygone des effectifs.
(E) Diagramme en secteur.

2. Pour prouver l'existence d'une relation entre deux variables quantitatives, on utilise :

(A) Nuage des points.
(B) La courbe cumulatif
(C) Boite à moustache
(D) Diagramme en bâtons
(E) Diagramme en secteur.

3. Pour vérifier si les moyennes de trois groupes (indépendant) ou plus sont différentes, on applique :

(A) Un test de conformité
(B) Un test d'homogénéité
(C) Un test d'ANOVA
(D) Un test de normalité
(E) Un test d'indépendance.

Exercice 2 : On donnera une valeur approchée de tous les résultats à 10^{-2} près.

Le tableau suivant donne les résultat du dosage de la ferritine chez les nouveau-nés de deux mois selon que la mère a reçu du fer ou un placebo pendant la grossesse.

Mère ayant reçu pendant la grossesse	Effectif	Moyenne (µg/L)	Variance s^2
Un placebo	25	130	4225
Du fer	24	190	9025

4. Le nombre de population concernée par cette étude est :

(A) 1
(B) 2
(C) 3.
(D) 4
(E) 5

5. L'effectif total concerné par l'étude est :

(A) 130
(B) 25
(C) 24.
(D) 47.
(E) 49.

6. Dans cette étude, il y a combien de variable(s) :

(A) 1
(B) 2
(C) 3.
(D) 4
(E) 5

7. Le but principal de cette étude est de chercher sil y a une :

(A) Relation entre deux variables quantitatives
(B) Relation entre deux variables qualitatives
(C) Relation entre plusieurs variables quantitatives
(D) Relation entre une variable quantitative et une variable qualitative.
(E) Relation entre deux variables quantitatives et une variable qualitative

8. La statistique de test utilisée est :

(A) La statistique de Student T.
(B) La statistique de Khi deux
(C) La statistique de Z.

(D) La statistique de Fisher
(E) Toutes les propositions précédentes sont fausses.

9. Quelle est la valeur de la statistique calculée :
(A) 1,41
(B) 2,52
(C) 2,57
(D) 2,59
(E) 30.

10. Supposant que les conditions de test sont vérifiées, que peut-on conclure :
(A) Les deux populations sont homogènes.
(B) Le modèle est pertinent (globalement significatif)
(C) Il y a une relation entre les variables étudiées
(D) Les résultats de l'échantillon ne sont pas conformes aux résultats de la population.
(E) Les variables sont indépendantes.

<u>Exercice 3</u> : On donnera une valeur approchée de tous les résultats à 10^{-3} près.
Sur un échantillon de 76 enfants, huit ont eu la rougeole.

11. L'estimation ponctuelle sans biais de la proportion (théorique) du rougeole dans la population des enfants est :
(A) < 0.01
(B) 0.105
(C) 8
(D) 0.095
(E) > 0.76.

12. L'intervalle de confiance à 95% de la proportion du rougeole dans la population actuelle est :
(A) [0.036 - 0.174]
(B) [0.026 - 0.1641]
(C) [0 - 0,325]
(D) [0.032 - 0.158]
(E) Le nombre des enfants ne permet pas d'estimer cet intervalle de confiance.

13. Au seuil de confiance de 95% pour que la proportion du rougeole dans la population actuelle appartient a l'intervalle [0.11-0,14], il faut observer combien d'enfants :
(A) 50
(B) 100
(C) 1567.
(D) > 100000.
(E) C'est impossible, car la proportion du rougeole dans la population actuelle n'appartient pas à cet intervalle de confiance.

<u>Exercice 4</u> : On donnera une valeur approchée de tous les résultats à 10^{-3} près.
Une étude portant sur la mort de 2100 adultes a donné les résultats suivants :

Cause de la mort / Habitudes de fumeurs	Non fumeur	Fumeur modéré	Gros fumeur
Maladie respiratoire	55	120	162
Maladie du cœur	49	388	315
Autres	61	300	650

14. La (ou les) population(s) statistique étudiée(s) est (sont) :
(A) Les adultes mort
(B) les morts.
(C) les fumeurs et la cause de mort
(D) cause de mort
(E) les adultes.

15. La (les) variable(s) statistique(s) étudiée(s) est (sont) :
(A) Fumeur et adulte
(B) habitudes de fumeur et la cause de mort
(C) Le nombre de morts
(D) Maladie
(E) Fumeur.

16. Le nombre de mort par maladie de cœur est :

(A) 49.
(B) 315
(6) 752
(D) 1127
(E) 2100

17. Pour répondre à l'objectif de cette étude, on effectue :

(A) Un test de conformité
(B) Un test d'homogénéité,
(C) Un test d'ANOVA.
(D) Un test d'indépendance
(E) Un test de corrélation.

18, Parmi les conditions de validité du test, il y a :

(A) $n_i = f_i \geq 5$ et $n_i = (1 - f_i) \geq 5$,
(B) Utilisable si petits effectifs
(C) Les valeurs observées de(s) la (les) variables suit loi normale.
(D) L'écart-type (théorique) n'est pas le même dans toutes les populations.
(E) Toutes les variables sont indépendantes les uns des autres.

19. Quelle est (sont) le(s) degré(s) de liberté :

(A) (336, 1126)
(B) 1
(C) 3
(D) 4
(E) 2098

20. Supposant que les conditions de test sont vérifiées, et que la statistique de test calculée est égale à 132.82, au risque de 5%, que peut-on conclure :

(A) Il n'y a pas de relation entre les variables étudiées.
(B) Les résultats de l'échantillon ne sont pas conforme aux résultats théoriques.
(C) Les échantillons sont homogènes
(D) Les résultats ne donnent pas une différence significative.
(E) La cause de la mort est liée à l'habitude de fumeur.

NB :

1. Z **suit une loi normale centrée réduite** $N(0,1)$**, alors :** $P(Z \leq 1.28) = 0.9$, $P(Z \leq 1.64) = 0.95$, $P(Z \leq 1.96) = 0.975$, $P(Z \leq 2.33) = 0.99$, $P(Z \leq 2.58) = 0.995$.

2. Pour une loi de Student : $t_{23}(0.975) = 2.069$; $t_{24}(0.975) = 2.064$; $t_{25}(0.975) = 2.06$; $t_{47}(0.975) = 2.011$; $t_{49}(0.975) = 2.01$; $t_{24}(0.95) = 1.711$; $t_{25}(0.95) = 1.708$; $t_{47}(0.95) = 1.678$; $t_{48}(0.95) = 1.677$

3. Pour une loi du Khi-deux : $x^2_{0.05}(1) = 3.84$; $x^2_{0.025}(1) = 5.02$; $x^2_{0.05}(2) = 5.991$; $x^2_{0.1}(2) = 11.14$; $x^2_{0.05}(3) = 7.815$, $x^2_{0.05}(4) = 9.488$; $x^2_{0.05}(6) = 12.592$; $x^2_{0.05}(24) = 36.114$; $x^2_{0.05}(47) = 64.001$; $x^2_{0.05}(48) = 64.171$; $x^2_{0.05}(49) = 66.339$.

4. Pour une loi du Fisher à $F_{0.05}(4,5) = 5.19$; $F_{0.05}(1,70) = 5.59$; $F_{0.05}(2,22) = 3.44$; $F_{0.05}(2,24) = 3.4$; $F_{0.05}(2,46) = 3.2$; $F_{0.05}(2,49) = 3.19$; $F_{0.05}(3,24) = 3.01$; $F_{0.05}(3,46) = 2.98$; $F_{0.05}(24,25) = 1.96$; $F_{0.05}(25,24) = 1.98$; $F_{0.05}(23,24) = 2$; $F_{0.05}(24,23) = 2.01$; $F_{0.05}(336, 1126) = 1.43$