Université d'Ottawa · University of Ottawa

Faculté des sciences
Mathématiques et de statistique

Faculty of Science
Mathematics and Statistics

Examen Final : MAT 2779 (Automne 2021)

Professeur : M'hammed Mountassir

Durée : 180 minutes+10 minutes pour soumettre la première page avec le tableau complété(en pdf). AUCUN RETARD N'EST ACCEPTÉ
19 Décembre 2021

Nom : ____________________ Numéro d'étudiant : ____________________

Examen à LIVRE OUVERT. Vous pouvez utiliser le logiciel R.

Question	Réponse	Question	Réponse	Question	Réponse
1		9		17	
2		10		18	
3		11		19	
4		12		20	
5		13		21	
6		14		22	
7		15		23	
8		16		24	

NB : À la fin de l'examen, remettez juste cette première page.

1. Dans une ville, 6 % de la population fume des cigares, 14 % de la population fume la cigarette, et 84 % de la population ne fume pas la cigarette et ne fume pas des cigares. Soit A l'événement qu'un résident choisi au hasard fume des cigares et soit B l'événement qu'un résident choisi au hasard fume la cigarette. Lequel des énoncés suivants est **faux** ?

   A) $P(A \cup B) = 0{,}16$ B) $P(A' \cup B') = 0{,}96$ C) $P(A \cap B) = P(A)P(B)$
   D) $P(A \cup B') = 0{,}90$ E) $P(A \cap B') < P(A' \cap B)$

2. On veut tester un nouveau type de pesticide sur trois différents types d'arbres fruitiers. 40% des arbres sont de type A, 25% sont de type B et 35% sont de type C. Ce pesticide est efficace à 90% sur les arbres de type A, à 85% sur ceux de type B et à 75% sur ceux de type C. On sélectionne au hasard un arbre et on note que le pesticide lui a été bénéfique. Quelle est la probabilité qu'il s'agisse d'un arbre de type B ?

   A) $0,2545$ B) $0,2321$ C) $0,2125$. D) $0,7455$. E) 0.3546.

3. Il est connu qu'en moyenne 12 patients arrivent à une clinique de santé toutes les heures. Pour vérifier l'exactitude de cette information, 22 heures sont surveillées à des jours différents. Pour chaque heure, le nombre de patients arrivés est déterminé. Ces 22 mesures sont utilisées pour produire un intervalle de confiance à 90% pour le nombre moyen d'arrivées en une heure en supposant que la population est normalement distribuée.
   Ci-dessous la sortie de R qui nous donne cet intervalle :

```
> t.test(x,conf.level = 0.90)$conf.int
[1] 12.14  15.28
```

   Déterminer l'écart type de l'échantillon.

   A) $4,28$ B) 5.28 C) 2.68 D) 3.54 E) 12.14

4. Deux maladies A et B sont répartis dans une certaine population. Il est connu que 10 % des personnes dans cette population va contracter la maladie A, 15 % vont contracter la maladie B et 3% de la population vont contracter les deux maladies. Nous choisissons une personne au hasard dans cette population. Quelle est la probabilité que cette personne va contracter les deux maladies étant donné que la personne a contracté au moins une de ces deux maladies ?

   A) 0.136 B) 0.552 C) 0.431 D) 0.452 E) 0.862

5. Le poids des paquets de pois-chiches d'une certaine marque est distribué normalement avec une moyenne $\mu = 500$ g et un écart type $\sigma = 119$ g. On sélectionne un échantillon aléatoire de taille n de ce type de paquets. Soit $\overline{X}$ dénote le poids moyen de cet échantillon. Déterminer la taille de l'échantillon n telle que $P(\overline{X} > 550) = 0,2$.

   A) $n = 15$ B) $n = 2$ C) $n = 6$ D) $n = 4$ E) $n = 10$

6. Une louve adulte donne naissance une fois par année. Soit X le nombre de louvtons par portée. Ci-dessous la fonction de masse des probabilités de X

Probabilités	X
0,10	2
0,18	3
0,53	4
0,12	5
0,07	6

   Déterminer l'écart type de X

   A) 0.8023 B) 0.9826 C) 0.8156 D) 0.8043 E) 0.8053

7. Pour étudier l'impact physiologique de la méditation, 16 individus ont été mesurés pour le niveau d'une substance biochimique, avant et après une heure de méditation. Voici un sommaire des données :

	Avant	Après
Moyenne de l'échantillon	$\bar{x}_1 = 88.3$	$\bar{x}_2 = 90.3$
Écart type de l'échantillon	$s_1 = 10.1$	$s_2 = 12.3$

L'écart type de l'échantillon pour les différences avant-après est $s_d = 5.6$. Tester l'hypothèse $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 < \mu_2$, où μ_1, μ_2 sont les niveaux moyens de cette substance chimique avant et après la méditation, respectivement. Supposez que la variable des différences est normalement distribuée.

(i) Quelle est la valeur de la statistique (t_0) du test ?

(ii) Déterminer la valeur P du test ?

A) (i) $t_0 = 1.18$; (ii) $0.05 < lavaleur - P < 0.1$

B) (i) $t_0 = 1.43$; (ii) $0.05 < lavaleur - P < 0.1$

C) (i) $t_0 = -1.43$; (ii) $0.05 < lavaleur - P < 0.1$

D) (i) $t_0 = -1.18$; (ii) $0.01 < lavaleur - P < 0.05$

E) (i) $t_0 = -2.85$; (ii) $0.005 < lavaleur - P < 0.01$

8. D'un côté on a traité 2560 patients avec un nouveau médicament contre la dépression et après une période de 3 mois on a trouvé que 480 parmi eux se sentent mieux. De l'autre côté on a donné un placébo à un groupe de 1870 patients qui sont dépressifs et après une période 3 mois on a trouvé que 316 parmi eux se sentent mieux. Peut-on affirmer que le nouveau médicament est plus efficace que le placebo (utiliser $\alpha = 5\%$) ? Donner la valeur-P de votre test. (Répondez par :OUI ou NON et donner la mesure de la valeur-P la plus proche).

A) OUI ; $0,113$ B) OUI ; $0,056$ C) NON ; $0,056$

D) NON ;$0,11$ E) Oui ; $0,03$

9. Soit X une variable aléatoire normalement distribuée avec une moyenne égale à 2 et une variance égale à 9. Quelle affirmation parmi les suivantes qui est **incorrecte** ?

   (A) $P(X < 1) = P(X > 3)$.
   (B) Si $P(X > a) = 0.025$, alors $a > 7$.
   (C) $P(X = 0) = 0$.
   (D) $\mathrm{E}(-2X + 2) = -2$.
   (E) $\mathrm{Var}(3X - 2) = 25$.

10. Un grossiste de fruits Canadien importe ses oranges du Maroc et d'Espagne et voudrait comparer les masses de ces oranges. En notant X, la masse (en grammes) des oranges du Maroc et par Y la masse (en grammes) des oranges de l'espagne. On va noter μ_X la moyenne de X et μ_Y la moyenne de Y. Pour un échantillon aléatoire de $n = 64$ oranges du Maroc, il a trouvé un intervalle de confiance de 95% pour μ_X égal à $[94.8975; 97.1025]$. Alors que pour un échantillon aléatoire de $m = 81$ oranges de l'Espagne, il a trouvé un intervalle de confiance de 95% pour μ_Y égal à $[84.216; 85.784]$. En supposant que les variables X et Y sont normalement distribuées avec des variances inconnues mais supposées égales, déterminer un intervalle de confiance de niveau 95% pour $\mu_X - \mu_Y$.

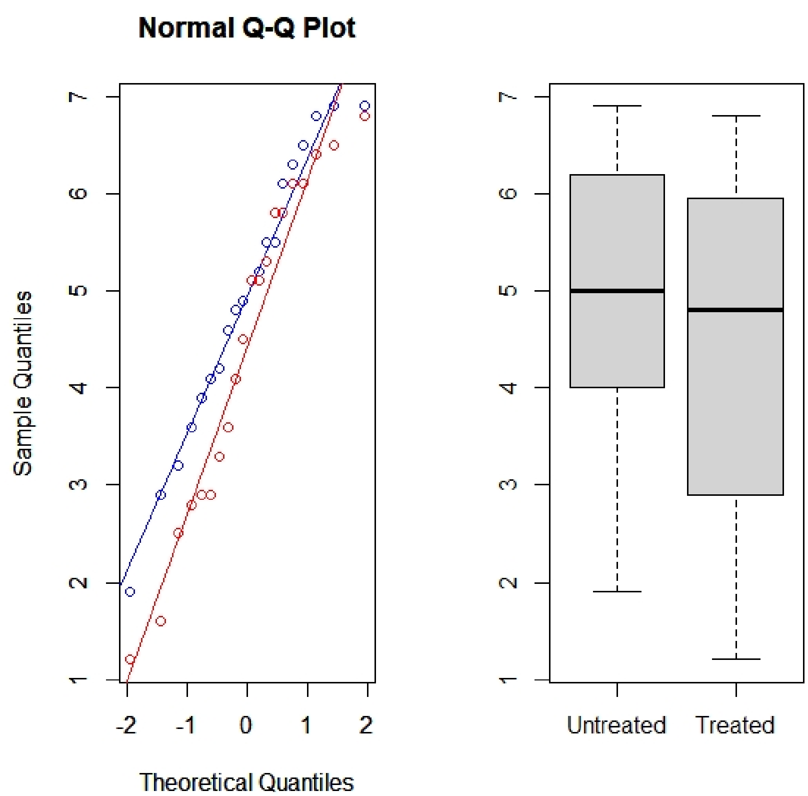
    (A)[10.681 ;11.318]
    (B)[8.756 ;10.657]
    (C)[10.564 ;12.114]
    (D)[9,693 ;12.307]
    (E)[8.563 ;9.657]

    INDICATION : Trouver d'abord l'écart type de chaque échantillon.

11. Pour étudier l'effet d'un certain pesticide sur la croissance des arbres fruitiers, un botaniste a mesuré la taille en mètres de deux types d'arbres après 5 ans. Un échantillon d'arbres non traités et un échantillons d'arbres traités et a obtenus les données suivantes (g_1 : les points rouges, représentent les arbres traités) :

```
g_1=c(3.2, 1.9, 3.6, 4.6, 4.1, 5.1, 5.5, 4.8, 2.9, 4.9,\\
 3.9, 6.9, 6.8, 6.1, 6.9, 6.5, 5.5, 5.2, 4.2, 6.3)
g_2=c(4.1, 2.5, 4.5, 3.6, 2.8, 5.1, 2.9, 5.1, 1.2, 1.6,\\
 6.4, 6.8, 6.5, 5.8, 2.9, 5.8, 3.3, 6.1, 5.3, 6.1)
```

Le botaniste a aussi fait quelques graphiques qui sont illustrés ci-après :



Parmi les affirmations suivantes, laquelle est correcte ?

(A) Les mesures des arbres non traités n'apparaissent pas normalement distribués, alors que celles des arbres traités le sont. .
(B) Les deux échantillons apparaissent normalement distribués avec des variances égales.
(C) Il y a des données aberrantes dans les arbres traités.
(D)Les deux échantillons apparaissent non normalement distribués avec des variances égales..
(E) Les deux échantillons apparaissent normalement distribués avec des variances non égales.

12. Le poids à la naissance des nouveaux-nés en Amérique du Nord est en moyenne de 120 onces. On veut tester l'hypothèse que les mères dont la condition socio-économique est défavorisée ont des bébés dont le poids moyen est inférieur à 120 onces. Soit μ le poids moyen des bébés des mères dont le statut socio-éconimique est défavorisé. Formuler les hypothèses qu'on doit tester et la signification de l'erreur type associée (Soit l'erreur de type I ou de type II) à l'affirmation correcte parmi les suivantes :

A) $H_0 : \mu = 120$ contre $H_1 : \mu < 120$. Une erreur de Type II signifie qu'on a conclut que le poids moyen des bébés de ces mères est inférieur à 120 onces alors qu'en fait ce n'est pas vrai.
B) $H_0 : \mu = 120$ contre $H_1 : \mu < 120$. Une erreur de Type I signifie qu'on a conclut que le poids moyen des bébés de ces mères est inférieur à 120 onces alors qu'en fait ce n'est pas vrai.
C) $H_0 : \mu \geq 120$ contre $H_1 : \mu < 120$. Une erreur de Type I signifie qu'on a conclut que le poids moyen des bébés de ces mères est inférieur à 120 onces alors qu'en fait ce n'est pas vrai.
D) $H_0 : \mu \geq 120$ contre $H_1 : \mu < 120$. Une erreur de Type II signifie qu'on a conclut que le poids moyen des bébés de ces mères est égal à 120 onces alors qu'en fait il est inférieur à 120 onces.
E) $H_0 : \mu = 120$ contre $H_1 : \mu < 120$. Une erreur de Type I signifie qu'on a conclut que le poids moyen des bébés de ces mères est égal à 120 onces alors qu'en fait il est inférieur à 120 onces

13. Certains professeurs et étudiants (qui travaillent sur une nouvelle étude) sont intéressés à estimer la proportion p des oies infectés par un virus. Déterminer la taille n de l'échantillon pour pouvoir estimer cette proportion, telle que la marge d'erreur de l'estimation de p ne dépasse pas $\varepsilon = 0.03$ avec 95% de niveau de confiance.

    A) 1068 B) 1067 C) 510 D) 511 E) 208

14. Un sapin prend 10 ans pour atteindre une taille adulte. On veut estimer la taille moyenne μ d'un sapin de 4 ans. Supposons que la taille (en cm) d'un sapin de 4 ans est normalement distribuée. Un échantillon aléatoire de 20 sapins de 4 ans a donné une moyenne échantillonale de 25,25 cm et un écart type échantillonal de 4,5 cm. Cet échantillon a aussi donné un intervalle de confiance de largeur 2,673. On sélectionne un second échantillon aléatoire d'au moins 40 sapins qui a donné un écart type échantillonnal de 5,6 cm et on construit un intervalle de confiance pour μ en utilisant *le même niveau de confiance* que dans le premier intervalle. Ce nouveau intervalle de confiance est de largeur 2,186. Quelle est la taille du second échantillon ?

    *Indication :*Utiliser le premier intervalle pour d'abord trouver le niveau de confiance.

    (A) 52 (B) 62 (C) 43 (D) 58 (E) 66

15. Nous avons 10 échantillons de sang. Supposons que 2 des 10 échantillons sont contaminés. On cueille 2 échantillons sans-remise au hasard parmi ces 10 échantillons. Calculer la probabilité qu'exactement 1 des 2 échantillons sélectionnés sera contaminé.

    A) 0,1596 B) 0,4775 C) 0,2000 D) 0,2684 E) 0,3556

16. Soient deux échantillons aléatoires sélectionnés de tailles $n = m = 10$ de deux popullations normales avec des variances inconnues mais supposées égales. Soient $\bar{X}_1$ et $\bar{X}_2$ les moyennes des 2 échantillons. Calculer
$$P(\frac{\bar{X}_1 - \bar{X}_2}{S_p\sqrt{2/10}} > -2,101)$$

(A) 0,95 (B) 0,975 (C) 0,025 (D) 0,05 (E) 0,01

17. Les recherches cliniques tentent d'itentifier les risques d'exposition de la population dans le milieu du travail. Dans un groupe de 46 hommes âgés entre 25 et 36 ans, on a mesuré les changements dans la fonctionnalité de leur poumon. Une des variables d'intérêt était le volume de la force d'expiration (VFE), qui est le volume d'air expiré par seconde. À la fin d'une certaine période, on a trouvé que le VFE de cet échantillon a baissé avec une moyenne de 0,27 litre et un écart type échantillonnal de 0,52 litre. On sait que le VFE décroit de 0,1 litre chez les hommes de même tranche d'âge qui sont en bonne santé durant la même période du temps. Peut-on conclure que le milieu du travail de ce groupe joue un rôle significatif dans la décroissance du VFE supérieure à 0,1 litre ? Choisissez le test d'hypothèses approprié à faire. Donner la $valeur - P$ et votre conclusion. Prenez $\alpha = 0.02$.

A) $Valeur - P < 0.02$; On ne peut pas dire que le milieu de travail de ce groupe cause une décroissance significative de VFE.
B) $Valeur - P < 0.02$; On peut dire que le milieu de travail de ce groupe cause une décroissance significative de VFE
C) $Valeur - P > 0.02$; On ne peut pas dire que le milieu de travail de ce groupe cause une décroissance significative de VFE
D) $Valeur - P > 0.02$; On peut dire que le milieu de travail de ce groupe cause une décroissance significative de VFE
E) La taille de l'échantillon ne nous permet pas de faire un tel test d'hypothèses.

18. Le 4 Novembre 2020, Les État-Unis ont officiellement quitté l'accord de Paris sur les changements climatiques. Cependant, une majorité des Américains souhaitent que leur pays réintègre cet accord. Des données d'un sondage sur un échantillon aléatoire de 128 Américains a été utilisé pour tester les hypothèses suivantes : $H_0 : p = 0.7 \quad contre \quad H_1 : p > 0.7$ avec p=la proportion des Américains qui sont favorables à l'accord de Paris. Ce test a donné une $valeur - P = 0,035$. Combien d'américains de l'échantillon étaient en faveur que les États-Unis réintègrent l'accord de Paris ?

(A) 99 (B) 80 (C) 75 (D) 107 (E) 115

19. Une étude a été menée pour comparer les temps que prennent des étudiants pour compléter un devoir dans un cours d'introduction en biostatistique. Deux échantillons indépendants, un de 50 hommes et l'autre de 50 femmes à qui on a donné des exercices similaires. Les résultats sont résumés dans le tableau suivant :

Hommes	Femmes
$n_1 = 50$	$n_2 = 50$
$\bar{x}_1 = 42$	$\bar{x}_2 = 38$
$s_1^2 = 18$	$s_2^2 = 14$

Peut-on dire qu'il y a une différence significative entre les temps moyens pour compléter ce devoir entre le groupe des hommes et celui des femmes. Faites votre test avec un seuil de signification 5%. Quelle est l'affirmation correcte parmi les suivantes ? (Supposez que les variances des 2 groupes sont similaires).

(A) On ne rejette pas l'hypothèse nulle. Donc, il y a une différence significative entre les moyennes des 2 groupes.
(B) On rejette l'hypothèse nulle. Donc il n'y a pas de différence significative entre les moyennes des 2 groupes.
(C) On rejette l'hypothèse nulle. Donc il y a une différence signiificative entre les moyennes des 2 groupes.
(D) On ne rejette pas l'hypothèse nulle. Donc il n'y a pas de différence significative entre les moyennes des 2 groupes.
(E) Aucune de ces réponses.

20. Pour comparer les effets de la méditation et de l'exercice physique sur une substance biochimique présente chez l'être humain. On a mesuré cette substance sur 11 personnes choisies au hasard une heure après une méditation intense et on a obtenu les résultats suivants :

$$n = 11; \bar{x} = 88,33; s_X = 10,1$$

on a aussi mesuré cette substance sur 15 personnes choisies au hasard, une heure après un exercice soutenu et on a obtenu les résultats suivants :

$$m = 15; \bar{y} = 90,65; s_Y = 12,3$$

Y-a-t-il une différence significative entre les taux moyennes de cette substance lors de ces deux activités ? On suppose que ce taux est distribué normalement dans les deux groupes avec des variances inconnues mais supposées égales. Prenez $\alpha = 1\%$. Donner votre décision et la valeur de t_0 obtenue.

(A) On rejette l'hypothèse nulle ; $t_0 = -0,51$
(B) On ne rejette pas l'hypothèse nulle ; $t_0 = -0,51$
(C) On rejette l'hypothèse nulle ; $t_0 = 1,43$
(D) On ne rejette rejette pas l'hypothèse nulle ; $t_0 = -1,43$
(E) On rejette l'hypothèse nulle ; $t_0 = -0,45$

21. Le poids à la naissance des nouveaux-nés en Amérique du Nord est en moyenne de 120 onces. On veut tester l'hypothèse que les mères dont la condition socio-économique est défavorisée ont des bébés dont le poids moyen est inférieur à 120 onces. Soit μ le poids moyen des bébés des mères dont le statut socio-éconimique est défavorisé. Formuler les hypothèses qu'on doit tester et la signification de l'erreur type associée (Soit l'erreur de type I ou de type II) à l'affirmation correcte parmi les suivantes :

A) $H_0 : \mu = 120$ contre $H_1 : \mu < 120$. Une erreur de Type II signifie qu'on a conclut que le poids moyen des bébés de ces mères est inférieur à 120 onces alors qu'en fait ce n'est pas vrai.
B) $H_0 : \mu = 120$ contre $H_1 : \mu < 120$. Une erreur de Type I signifie qu'on a conclut que le poids moyen des bébés de ces mères est inférieur à 120 onces alors qu'en fait ce n'est pas vrai.
C) $H_0 : \mu \geq 120$ contre $H_1 : \mu < 120$. Une erreur de Type I signifie qu'on a conclut que le poids moyen des bébés de ces mères est inférieur à 120 onces alors qu'en fait ce n'est pas vrai.
D) $H_0 : \mu \geq 120$ contre $H_1 : \mu < 120$. Une erreur de Type II signifie qu'on a conclut que le poids moyen des bébés de ces mères est égal à 120 onces alors qu'en fait il est inférieur à 120 onces.
E) $H_0 : \mu = 120$ contre $H_1 : \mu < 120$. Une erreur de Type I signifie qu'on a conclut que le poids moyen des bébés de ces mères est égal à 120 onces alors qu'en fait il est inférieur à 120 onces

22. Des recherches neurologiques ont montré que pour environ 80% de la population, les capacités linguistiques résident seulement dans le côté gauche du cerveau. Pour 10% de la population, c'est plutôt dans le côté droit, et les 10% restants la capacité est à deux-côtés. (Les deux derniers groupes sont principalement les gauchers.) Nous sélectionnons au hasard 25 personnes. Quelle est l'espérance et l'écart type du nombre de personnes pour qui les capacités linguistiques résident seulement dans le côté gauche du cerveau chez les individus sélectionnés ?

A) $\mu = 2.5\,;\ \sigma = 0.5$ B) $\mu = 20\,;\ \sigma = 4$ C) $\mu = 22.5\,;\ \sigma = 2.25$
D) $\mu = 20\,;\ \sigma = 2$ E) $\mu = 20\,;\ \sigma = 0.0064$

23. Un médecin de famille donne des consultations virtuelles de 10 minutes chacune en utilisant la plateforme Zoom. Le médecin consulte ses clients et prend le temps de répondre à leurs questions. La durée que prend un patient lors d'une consultation est normalement distribuée avec une moyenne de 9,5 minutes et un écart type de 2 minutes. Sur 6 consultations choisies au hasard, quelle est la probabilité qu'au moins une dure plus que 10 minutes ?

(A) 0.9958 (B) 0.9539 (C) 0.8041 (D) 0.7315 (E) 0.5490

24. Soient les variables X : la pression artérielle et Y : le niveau de calcium, on a mesuré ces deux variables sur un échantillon aléatoire de 38 personnes et on a trouvé la droite de régression linéaire suivante :

$$\hat{y} = -2,2 + 1,64475 * x$$

avec les écart types des échantillons $s_x = 0,45$ et $s_y = 1,667$. Déterminer la valeur du coefficient de détermination

$$R^2$$

entre ces deux variables. (Choisissez la bonne réponse parmi les choix suivants)

(A) $R^2 = 64,56\%$
(B) $R^2 = 59,71\%$
(C) $R^2 = 29,71\%$
(D) $R^2 = 19,71\%$
(E) $R^2 = 37,89\%$