Université d'Ottawa · University of Ottawa

Faculté des sciences
Mathématiques et de statistique

Faculty of Science
Mathematics and Statistics

Echantillon d'Examen Final : MAT 2379-2779 (Automne 2018)

Professeur : M. Alvo, G. Lamothe, M. Mountassir

Nom : ____________________ Numéro d'étudiant : ____________________

Il est interdit de se servir de téléphones cellulaires, de dispositifs électroniques non autorisés ou de notes de cours (à moins qu'il s'agisse d'un examen à livre ouvert). Les téléphones et les dispositifs doivent être fermés et rangés dans votre sac : vous ne pouvez pas les laisser dans vos poches ou sur votre personne. Sinon, on pourrait vous demander de quitter l'examen immédiatement et des allégations de fraude scolaire pourraient être déposées dont le résultat pourrait être un 0 (zéro) pour l'examen.

En apposant votre signature, vous reconnaissez vous être assuré de respecter l'énoncé ci-dessus.

____

Question	Réponse	Question	Réponse	Question	Réponse
1		10		19	
2		11		20	
3		12		21	
4		13		22	
5		14		23	
6		15		24	
7		16		25	
8		17			
9		18			

NB : À la fin de l'examen, ne remettez que cette page.

- **C'est un examen à livre fermé**
- **Une feuille de formules (p. 22) et des tableaux de probabilités (p. 19) sont fournis.**
- **Seule une calculatrice non-programmable et non-graphique est permise.**
- **Les feuilles brouillons sont fournies à la page 26.**
- **Il y a 25 questions à choix multiples.**
- **Donnez votre réponse à chaque question dans le tableau de la première page.**

Multiple Choice Questions

1. La durée moyenne de la gestation humaine est d'environ 40,5 semaines. On pense que le diabète de la mère peut influencer la durée de la gestation. Dans une étude constituée de 20 femmes diabétiques enceintes, il a été observé que la durée moyenne de gestation était de 38,8 semaines et que l'écart type de la durée était 5 semaines. Est-ce que les preuves que la durée moyenne de la gestation chez les femmes diabétiques n'est pas 40,5 semaines sont significatives ? Formuler un test d'hypothèses approprié à un niveau de signification de $\alpha = 0,05$. Déterminer la valeur P et donner vos conclusions.

   A) La valeur P est entre 0,01 et 0,025. Il y a des preuves significatives que la durée moyenne de la gestation chez les femmes diabétiques est différente de 40,5 semaines.
   B) La valeur P est entre 0,025 et 0,05. Il y a des preuves significatives que la durée moyenne de la gestation chez les femmes diabétiques est différente de 40,5 semaines.
   C) La valeur P est entre 0,05 et 0,10. Les preuves que la durée moyenne de la gestation chez les femmes diabétiques est différente de 40,5 semaines ne sont pas significatives.
   D) La valeur P est entre 0,10 et 0,20. Les preuves que la durée moyenne de la gestation chez les femmes diabétiques est différente de 40,5 semaines ne sont pas significatives.
   E) La valeur P est entre 0,20 et 0,40. Les preuves que la durée moyenne de la gestation chez les femmes diabétiques est différente de 40,5 semaines ne sont pas significatives.

2. Le vaccin Bacille Calmette-Guérin (BCG) pour la tuberculose est obligatoire pour les enfants d'âge scolaire dans de nombreux pays européens. Au Canada, avant la vaccination par le BCG, le patient est testé pour la tuberculose en utilisant un test cutané à la tuberculine, appelé le test de Mantoux. Les gens qui ont été vaccins par le BCG aura souvent un résultat positif au test de Mantoux, même s'ils n'ont pas la tuberculose. Par conséquent, le test de Mantoux n'est pas un outil très efficace pour détecter la tuberculose. Dans une étude récente, 12% des sujets avaient un résultat positif au test de Mantoux. Parmi ceux qui ont un résultat positif au test, seulement 10% avaient la tuberculose. D'autre part, 1% des patients avec un résultat négatif au test avaient la tuberculose. Dans cette étude, quel pourcentage des patients avaient la tuberculose ?

A) 1,10% B) 2,08% C) 0,88% D) 1,20% E) 13,03%

3. Une équipe de chercheurs ont étudié l'effet de la thérapie de la réminiscence pour les femmes âgées souffrant de dépression. Ils ont mesuré la dépression (sur une échelle de 1 à 10) chez 20 femmes de 60 ou plus résidant pendant 3 mois ou plus dans un établissement de soins de longue durée. Une plus grande cote de dépression est interprété comme une dépression plus sévère. La dépression fut mesuré aussi après l'intervention. Dans `R`, nous avons attribué les cotes pré-intervention et post-intervention aux variables `pre` et `post`, respectivement. Nous avons aussi calculé la différence entre la cote de dépression post-intervention et la cote de dépression pré-intervention, pour chaque sujet.

```
> d=post-pre
> summary(d)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -5.00   -3.00   -0.50   -1.35    0.00    1.00
> sd(d)
[1] 2.084403
```

Supposons que la différence entre les mesures post-intervention et pré-intervention est normalement distribuée. Utiliser les statistiques descriptives ci-haut pour pour calculer un intervalle de confiance à 95% pour la moyenne de la différence des cotes de dépression post-intervention et pré-intervention. Basé sur cet intervalle de confiance, est-ce que nous sommes confiant que la dépression est moins sévère en moyenne après l'intervention ?

A) I.C.=[0,37 ; 2,33]. Nous sommes confiant que la dépression est plus sévère après l'intervention en moyenne.
B) I.C.=[-2,33 ; 0,37]. Nous ne sommes pas confiant que la dépression est moins sévère après l'intervention en moyenne.
C) I.C.=[-2,53 ; -0,57]. Nous sommes confiant que la dépression est moins sévère après l'intervention en moyenne.
D) I.C.=[-2,33 ; -0,37]. Nous sommes confiant que la dépression est moins sévère après l'intervention en moyenne.
E) I.C.=[-2,33 ; -0,37]. Nous ne sommes pas confiant que la dépression est moins sévère après l'intervention en moyenne.

4. Une étude a été menée pour estimer la sensibilité et la spécificité d'une nouvelle procédure pour identifier une maladie rénale chez les patients souffrant d'hypertension. Parmi 54 sujets hypertendus atteints de la maladie, la nouvelle procédure a identifié la maladie chez 45 sujets. Parmi 83 sujets hypertendus sans la maladie, la nouvelle procédure a identifié la maladie chez 24 sujets. Supposons que nous avons un patient hypertendu d'une population avec une prévalence de 8% pour cette maladie rénale. La nouvelle procédure identifie la maladie rénale chez ce patient. Quelle est la probabilité que le patient ait réellement cette maladie rénale ? Supposons que la sensibilité et la spécificité du test reste le même que dans l'étude mentionnée ci-haut.

   A) 0,8820 B) 0,2004 C) 0,3737 D) 0,5732 E) 0,0545

5. La pression intra-oculaire est la pression du fluide à l'intérieur de l'oeil. Le glaucome est une maladie oculaire qui se manifeste par une pression intraoculaire élevée. La répartition de la pression intraoculaire dans la population générale suit une loi normale de moyenne 16 mm Hg et d'écart type de 3 mm Hg. Une pression intra-oculaire normale est considérée comme étant comprise entre 12 mm Hg et 20 mm Hg (inclusivement). Laquelle des commandes suivantes de `R` donne la probabilité qu'une personne choisie au hasard ait une pression intra-oculaire normale ? (Une seule réponse est correcte.)

   A) `qnorm(20,16,3)-qnorm(12,16,3)`
   B) `pnorm(20,3,16)-pnorm(12,3,16)`
   C) `pnorm(20,16,3)-pnorm(12,16,3)`
   D) `pnorm(20,16,3)-pnorm(11,16,3)`
   E) `pnorm(20,16,9)-pnorm(12,16,9)`

6. Les autochtones du Canada ont un risque plus élevé de développer de nombreuses maladies chroniques par rapport au reste de la population. Dans une communauté autochtone particulière, 16% de la population ont la tuberculose, 20 % sont atteints de diabète et 8% ont les deux maladies. Quelle est la probabilité qu'une personne choisie au hasard dans cette communauté n'ait aucune de ces maladies ?

   A) 0,72 B) 0,28 C) 0,64 D) 0,85 E) 0,90

7. La maladie à virus Ebola, anciennement connu sous le nom de fièvre hémorragique à virus Ebola, est une maladie grave, souvent mortelle chez les humains. On pense que les chauves-souris de la famille des Pteropodidae sont des hôtes naturels du virus Ebola. Le virus est introduit dans la population humaine par contact étroit avec les fluides corporels d'animaux infectés. La période d'incubation (l'intervalle de temps de l'infection par le virus avant l'apparition des symptômes) est compris entre 2 à 21 jours. Les données ci-après donne la durée d'incubation (en jours) de 16 patients infectés par le virus Ebola :

$$4 \quad 5 \quad 6 \quad 6 \quad 7 \quad 8 \quad 9 \quad 9 \quad 11 \quad 12 \quad 13 \quad 15 \quad 15 \quad 17 \quad 20 \quad 21$$

Calculer la médiane ($\tilde{x}$), le premier quartile (q_1) et le troisième quartile (q_3) pour cet échantillon. Donner les valeurs des valeurs aberrantes (s'il y en a).

A) $\tilde{x} = 10;\ \ q_1 = 5;\ \ q_3 = 15\,;$ la valeur 21 est aberrante
B) $\tilde{x} = 10;\ \ q_1 = 6.25;\ \ q_3 = 15\,;$ il n'y a pas de valeurs aberrantes
C) $\tilde{x} = 11;\ \ q_1 = 6;\ \ q_3 = 17\,;$ la valeur 4 est aberrante
D) $\tilde{x} = 11;\ \ q_1 = 6,25;\ \ q_3 = 15\,;$ les valeurs 4 et 5 sont aberrantes
E) $\tilde{x} = 11;\ \ q_1 = 5,25;\ \ q_3 = 15,5\,;$ il n'y a pas de valeurs aberrantes

8. En biochimie et la pharmacologie, un récepteur est une molécule de protéine habituellement trouvée dans la surface de la membrane plasmique d'une cellule qui reçoit des signaux chimiques à partir de l'extérieur de la cellule. D'un échantillon de 9 cellules, on a observé une moyenne de 1203 fmoles récepteurs par milligramme de protéine membranaire, avec une erreur type estimée de la moyenne de 64 fmoles. (Une fmole est égale à 10^{-15} mole). Basé sur ces données, donner un intervalle de confiance à 95% pour le montant moyen (en fmoles) de récepteurs par milligramme trouvés dans la protéine de la membrane de ces cellules. On suppose que la quantité de récepteurs par milligramme de protéine membranaire est normalement distribuée.

A) $[1077,3; 1329,7]$ B) $[1153,8; 1252,2]$ C) $[0; 1322,8]$
D) $[1055,4; 1350,6]$ E) $[1098,1; 1308,9]$

9. La sortie de `R` suivante est un sommaire du poids (en livres) pour 50 bébés qui sont nés prématurés dans un grand hôpital.

```
> summary(x)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.432   4.607   6.158   6.809   8.212  18.700
> sd(x)
[1] 3.420575
```

On applique une transformation logarithmique :

```
> y=log(x)
```

Voici des statistiques descriptive pour le poids transformé :

```
> summary(y)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.3591  1.5276  1.8178  1.7990  2.1056  2.9285
> sd(y)
[1] 0.5064342
```

Calculer $a =$ la moyenne géométrique du poids (en livres) et $b =$ l'écart type géométrique du poids (en livres).

A) $a = 905,96$ et $b = 30,59$ B) $a = 6,809$ et $b = 3,42$
C) $a = 6,04$ et $b = 1,66$ D) $a = 1,79$ et $b = 0,50$
E) $a = 0,59$ et $b = 1,36$

10. Les données ci-dessous donne les poids à la naissance (en onces) pour six livraisons à l'Hôpital Civic. En supposant que le poids à la naissance suit une loi normale, trouver un intervalle de confiance à 90 % pour le poids moyen (en onces) à la naissance.

97 117 140 78 99 148

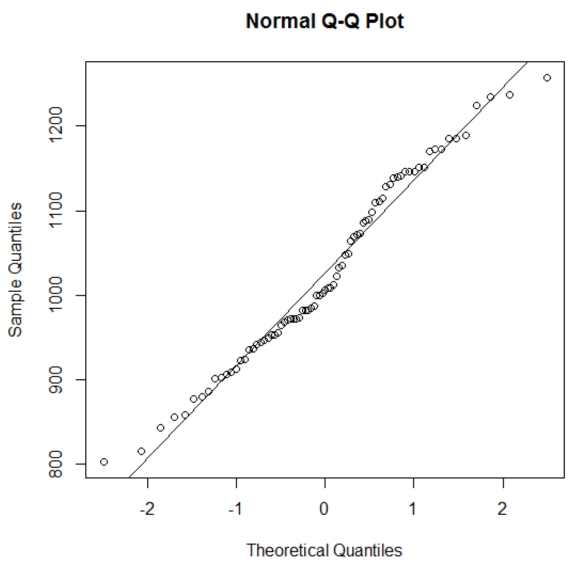
A) $[91,0; 135,4]$ B) $[84,8; 141,5]$ C) $[91,6; 134,8]$
D) $[95,0; 131,3]$ E) $[92,3; 133,6]$

11. Le Dryas récent (ou le *grand gel*) a été un événement de refroidissement brusque de l'hémisphère Nord qui s'est produit il y a environ 12 000 ans et peut être attribuable à une modification des courants de l'océan Atlantique (COA) qui auraient cessé de convoyer de l'eau réchauffée de l'équateur vers l'Europe. Le moyen le plus commun de ralentir les COA implique la réduction de la densité de l'eau de surface océanique par une augmentation du débit d'eau douce dans l'Atlantique Nord. Pour prévoir si un tel événement peut se reproduise, la densité de l'eau de mer près de la surface est étroitement surveillée. La sortie de `R` suivante donne un sommaire des 79 mesures de la densité de l'eau de l'océan Atlantique près de la surface (en kg$/m^3$), à une latitude de 45 degrés nord :

```
> summary(x)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  802.5   948.1  1006.0  1026.0  1122.0  1258.0
> var(x)
[1] 12013.36
> sd(x)
[1] 109.61
```

Le diagramme ci-bas est le diagramme quantile-quantile pour ces données, qu'on ait produit avec les commandes suivantes :

```
> qqnorm(x)
> abline(mean(x),sd(x))
```



Si l'équation de la droite sur ce diagramme est $y = a + bz$, où y est le quantile de l'échantillon et z est le quantile théorique, quelle sont les valeurs de a et b ?

A) $a = 1006$ et $b = 109,61$ B) $a = 109,61$ et $b = 1026$
C) $a = 1026$ et $b = 109,61$ D) $a = 1026$ et $b = 173,90$
E) $a = 1006$ et $b = 12013,36$

12. Voici les nombres d'attaques d'ours mortels en Amérique du Nord par décennie de 1900 à 1989.

$$2,\ 1,\ 4,\ 8,\ 6,\ 9,\ 9,\ 19,\ 20.$$

Calculer la moyenne et l'écart type du nombre d'attaques d'ours mortels en Amérique du Nord par décennie.

A) la moyenne est 8,667 et l'écart type est 5,6505
B) la moyenne est 8,0 et l'écart type est 19,0
C) la moyenne est 8,0 et l'écart type est 5,0
D) la moyenne est 8,667 et l'écart type est 46,0
E) La moyenne est 8,667 et l'écart type est 6,7823

13. Aux Etats-Unis, les groupes sanguins ont la répartition suivante :

$$41\%\ O,\quad 31\%\ A,\quad 22\%\ B\quad \text{et}\quad 6\%\ AB.$$

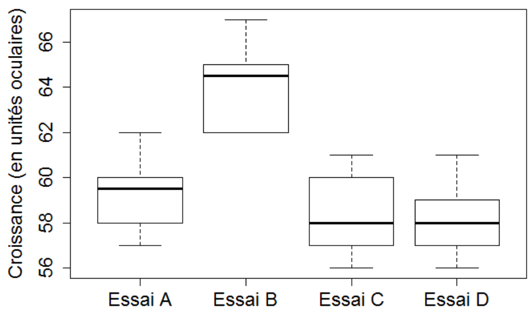
Il est connu que O est un donneur universel, A peut seulement donner à A et AB, B peut seulement donner à B et AB, et AB peut seulement donner à AB. Si un patient qui a besoin d'une transfusion de sang reçoit le sang d'un donneur choisi au hasard, et que les deux personnes sont indépendants les uns des autres, quelle est la probabilité que la transfusion soit réussie ?

A) 0,6607 B) 0,3393 C) 0,4101 D) 0,7314 E) 0,5899

14. 20 % des arbres dans une certaine forêt sont des érables. Dans cette forêt, 15% des érables sont des arbres matures, c'est-à-dire âgés entre 10 et 15 ans. Nous choisissons au hasard un arbre dans cette forêt. Quelle est la probabilité que c'est un érable âgé entre 10 et 15 ans ?

A) 0,03 B) 0,15 C) 0,20 D) 0,75 E) 0,175

15. Les boîtes à moustaches comparatifs ci-bas montrent les effets de différents sucres sur la croissance des sections de pois cultivés en culture de tissu, mesurées en unités oculaires. (Une unité oculaire est 0,114 cm.) À l'essai A, 2% de glucose a été ajouté à la culture. À l'essai B, 2% de saccharose a été ajouté à la culture. À l'essai C, 1% de glucose et 2% de fructose a été ajouté à la culture. Enfin, à l'essai D, 1% de fructose a été ajouté à la culture.



Lequel des énoncés suivants est correct. (Seulement un de ces énoncés est correct.)

A) Les croissances médianes sont semblables pour les essais C et D.
B) Les distances interquartiles des croissances sont semblables pour les essais A, C et D, mais pas pour B.
C) Il y a des valeurs aberrantes aux essais A, C, et D, mais il n'y a pas de valeur aberrante à l'essai B.
D) La distribution de la croissance pour l'essai B est approximativement symétrique.
E) La plus petite croissance fut à l'essai B.

16. Considérons un échantillon aléatoire de 50 femmes âgées de 20 à 29 ans. L'un des objectifs de l'étude est de décrire la répartition de l'indice de masse corporelle (IMC) pour ces femmes. Supposons que l'échantillon aléatoire est cueilli à partir d'une population de femmes ayant un IMC moyen de 26,8 et que l'écart type de la population est 7,42. Approximer la probabilité que l'IMC moyen de cet échantillon de 50 femmes sera plus grand que 29.

A) 0,0179 B) 0,0179 C) 0,6179 D) 0,3821 E) 0,0375

17. Une société pharmaceutique étudie un nouvel analgésique (médicament pour le soulagement de la douleur) sur un échantillon de six patients souffrant de migraine. Parmi ceux-ci, quatre patients ont rapporté que leurs migraines se sont atténuées après l'utilisation du médicament. Cependant, il est connu que 20 % des migraines vont s'atténuer de toute façon sans aucun traitement. Quelle est la probabilité que dans un échantillon de six patients souffrant de migraine, la migraine va s'atténuer sans aucun traitement pour exactement 4 des six patients ?

    A) 0,0016 B) 0,2534 C) 0,3523 D) 0,0154 E) 0,9992

18. La probabilité qu'un homme âgé de 40 à 55 va subir une crise cardiaque durant une période de 5 ans est 0,04. Un nouveau médicament est introduit dans l'espoir de réduire cette probabilité. Considérons une étude de 5 ans impliquant des hommes âgés de 40 à 55. Parmi les 2046 sujets donnés ce nouveau médicament, il y avait 56 sujets qui ont subi une crise cardiaque. Soit p soit la proportion d'hommes âgés de 40 à 55 utilisant ce nouveau médicament qui vont subir une crise cardiaque dans une période de 5 ans. Donner un intervalle de confiance à 95% pour p. Basé sur cet intervalle de confiance, sommes-nous confiant que le nouveau médicament est efficace pour réduire les risques d'une crise cardiaque ?

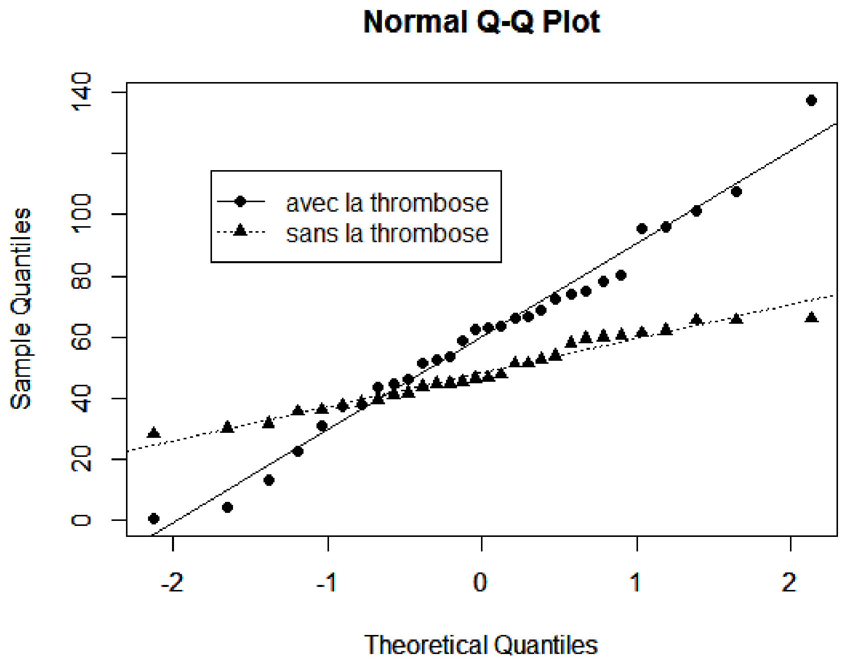
    A) I.C.=[0,020 ; 0,034]. Nous ne sommes pas confiant que le nouveau médicament est efficace pour réduire les risques d'une crise cardiaque.
    B) I.C.=[0,020 ; 0,034]. Nous sommes confiant que le nouveau médicament est efficace pour réduire les risques d'une crise cardiaque.
    D) I.C.=[0,014 ; 0,041]. Nous sommes confiant que le nouveau médicament est efficace pour réduire les risques d'une crise cardiaque.
    D) I.C.=[0,041 ; 0,052]. Nous sommes confiant que le nouveau médicament est efficace pour réduire les risques d'une crise cardiaque.
    E) I.C.=[0,020 ; 0,056]. Nous ne sommes pas confiant que le nouveau médicament est efficace pour réduire les risques d'une crise cardiaque.

19. Continuons avec la situation de la Question 18. Formuler l'hypothèse nulle et l'hypthèse alternative pour vérifier que le nouveau médicament est efficace pour réduire les risques d'une crise cardiaque. Ensuite, calculez la valeur P correspondante.

    A) $H_0 :\ p = 0,04$ contre $H_1 :\ p > 0,04$. La valeur P est 0,9982.
    B) $H_0 :\ p = 0,04$ contre $H_1 :\ p < 0,04$. La valeur P est 0,0036.
    C) $H_0 :\ p = 0,04$ contre $H_1 :\ p < 0,04$. La valeur P est 0,0018.
    D) $H_0 :\ p = 0,04$ contre $H_1 :\ p \neq 0,04$. La valeur P est 0,0036.
    E) $H_0 :\ p = 0,04$ contre $H_1 :\ p < 0,04$. La valeur P est 0,0154.

20. L'objectif d'une étude était de déterminer s'il y avait des concentrations différentes de l'anticardiolipine IgG (en mg/dl) chez les sujets avec et sans thrombose. Dans `R`, nous avons attribué la concentration d'IgG pour les 30 sujets avec la thrombose à la variable x. Alors que la concentration d'IgG pour les 30 sujets sans la thrombose a été attribuée à la variable y. Nous avons produit des statistiques descriptives pour les deux variables et des diagrammes quantile-quantile superimposés.

```
> summary(x)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.527  43.850  62.680  60.240  74.840 137.600
> summary(y)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  28.01   39.81   46.48   48.14   58.94   66.06
```



On veut utiliser la fonction `t.test()` pour vérifier que la concentration moyenne d'IgG pour les patients avec la thrombose est différente de la concentration moyenne d'IgG pour les patients sans la thrombose.
Lequel des énoncés suivant est correct ? (Seulement un des énoncés est correct.)

A) C'est évident que nous avons des échantillons de populations non-normales, alors on ne devrait pas utilisé la fonction `t.test`
B) C'est raisonnable de supposer que les populations sont normales et la bonne commande est `t.test(x,y)`
C) C'est raisonnable de supposer que les populations sont normales et la bonne commande est `t.test(x,y,var.equal=TRUE)`
D) C'est raisonnable de supposer que les populations sont normales et la bonne commande est `t.test(x,y,alternative="greater")`
E) C'est raisonnable de supposer que les populations sont normales et la bonne commande est `t.test(x,y,paired="TRUE")`

21. Un vaccin expérimental contre le cancer a été mis au point pour réduire la taille d'une tumeur. Nous tenons à tester l'hypothèse nulle que le vaccin n'a pas d'effet sur la taille d'une tumeur, contre l'hypothèse alternative que le vaccin est efficace pour réduire la taille d'une tumeur. Soit μ_D la moyenne de la différence entre la taille du tumeur avant le vaccin et la taille du tumeur 3 mois après la vaccination. Formuler des hypothèses appropriées et interpréter une erreur de première espèce et une erreur de deuxième espèce dans ce contexte. Choisir l'énoncé qui est correct dans la liste ci-bas. (Un seul énoncé est correct.)

A) $H_0 : \mu_D = 0$ contre $H_1 : \mu_D > 0$. On commet une erreur de deuxième espèce, si on conclue que le vaccin est efficace à réduire la taille de la tumeur, mais dans la réalité le vaccin n'a pas d'effet.
B) $H_0 : \mu_D = 0$ contre $H_1 : \mu_D < 0$. On commet une erreur de première espèce, si on conclue que le vaccin est efficace à réduire la taille de la tumeur, mais dans la réalité le vaccin n'a pas d'effet.
C) $H_0 : \mu_D = 0$ contre $H_1 : \mu_D > 0$. On commet une erreur de première espèce, si on conclue que le vaccin n'a pas d'effet sur la taille de la tumeur, mais dans la réalité le vaccin est efficace à réduire la taille du tumeur.
D) $H_0 : \mu_D = 0$ contre $H_1 : \mu_D \neq 0$. On commet une erreur de première espèce, si on conclue que le vaccin est efficace à réduire la taille du tumeur, mais dans la réalité le vaccin n'a pas d'effet.
E) $H_0 : \mu_D = 0$ contre $H_1 : \mu_D > 0$. On commet une erreur de deuxième espèce, si on conclue que le vaccin n'a pas d'effet sur la taille de la tumeur, mais dans la réalité le vaccin est efficace à réduire la taille du tumeur.

22. Une étude est menée pour étudier le lien entre le nombre X d'heures d'exercice par semaine et la pression artérielle systolique Y pour les hommes de 50 ans. Les données suivantes ont été obtenues sur les 10 sujets :

Nombre d'heures x_i	Pression artérielle systolique y_i	x_i^2	$x_i\,y_i$
4	120	16	480
10	110	100	1100
2	120	4	240
3	135	9	405
3	140	9	420
5	115	25	575
1	150	1	150
2	165	4	330
2	160	4	320
0	180	0	0

Pour ces données, nous avons :

$$\sum_{i=1}^{10} x_i = 32, \quad \sum_{i=1}^{10} y_i = 1395, \quad \sum_{i=1}^{10} x_i^2 = 172, \quad \sum_{i=1}^{10} x_i\,y_i = 4020$$

Donner la droite de régression estimée et estimer la pression artérielle systolique moyenne pour un individu qui fait 6 heures d'exercices par semaine.

A) La droite de régression estimée est $\hat{y} = 181,19 - (8,6)\,x$ et une estimation pour la pression artérielle systolique moyenne pour un individu qui fait 6 heures d'exercices par semaine est 130.
B) La droite de régression estimée est $\hat{y} = 167,31 - (7,12)\,x$ et une estimation pour la pression artérielle systolique moyenne pour un individu qui fait 6 heures d'exercices par semaine est 125.
C) La droite de régression estimée est $\hat{y} = 138,90 - (5,25)\,x$ et une estimation pour la pression artérielle systolique moyenne pour un individu qui fait 6 heures d'exercices par semaine est 107.
D) La droite de régression estimée est $\hat{y} = 145,76 - (3,42)\,x$ et une estimation pour la pression artérielle systolique moyenne pour un individu qui fait 6 heures d'exercices par semaine est 125.
E) La droite de régression estimée est $\hat{y} = 159,91 - (6,38)\,x$ et une estimation pour la pression artérielle systolique moyenne pour un individu qui fait 6 heures d'exercices par semaine est 122.

23. Une hypothèse importante dans la recherche de l'hypertension est que la réduction du sodium peut abaisser la tension artérielle. Puisque la réduction de sodium est difficile de maintenir pendant une longue période de temps, des conseils diététiques sont parfois utilisés pour atteindre cet objectif. Des niveaux de sodium urinaires ont été obtenus de huit individus inscrits dans un programme de réduction de sodium. Ils ont mesurés la teneur de sodium avant l'intervention et une semaine après le début de l'intervention. Voici les données :

```
> Semaine0=c(7.85,12.03,21.84,13.94,16.68,41.78,14.97,12.07)
> Semaine1=c(5.59,8.5,4.55,12.78,10.69,23.51,5.46,11.95)
```

Le chercheur A utilise la commande suivante :

```
> t.test(Semaine0,Semaine1,alternative="greater")
```

Voici la résultat :

```
data:  Semaine0 and Semaine1
t = 1.6814, df = 11.267, p-value = 0.06008
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.4779779        Inf
sample estimates:
mean of x mean of y
 17.64500  10.37875
```

Le chercheur B utilise la commande suivante :

```
> t.test(Semaine0,Semaine1,paired=TRUE,alternative="greater")
```

Voici le résultat :

```
data:  Semaine0 and Semaine1
t = 2.8835, df = 7, p-value = 0.01177
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 2.492072      Inf
sample estimates:
mean of the differences
                7.26625
```

Supposons que chacun de ces chercheurs ont vérifié la condition d'application de normalité avant d'utiliser la fonction `t.test`.

Lequel de ces énoncés est correct ? (Seulement un énoncé est correct.)

A) Le chercheur A a utilisé la commande correctement. À $\alpha = 0,05$, il y a des preuves significatives que la teneur de sodium ait diminué en moyenne après l'intervention.
B) Le chercheur B a utilisé la commande correctement. À $\alpha = 0,05$, il y a des preuves significatives que la teneur de sodium ait diminué en moyenne après l'intervention.
C) Le chercheur A a utilisé la commande correctement. À $\alpha = 0,05$, les preuves que la teneur de sodium ait diminué en moyenne après l'intervention ne sont pas significatives.
D) Le chercheur B a utilisé la commande correctement. À $\alpha = 0,05$, les preuves que la teneur de sodium ait diminué en moyenne après l'intervention ne sont pas significatives.
E) Ni l'un ni l'autre a utilisé la commande correctement. Une statistique t ne devrait pas être utilisé dans cette situation.

24. Le lien plante-eau joue un rôle important dans la physiologie des plantes. Nous considérons une expérience dans laquelle 16 jeunes plants de bouleau ont été inondées avec de l'eau pour une journée et 13 autres plants ont été utilisé comme un groupe témoin (c'est-à-dire des contrôles). À la fin de l'expérience, les racines de toutes les plantes ont été analysées pour le niveau de l'adénosine triphosphate (ATP), en tant que mesure pour le transfert d'énergie intracellulaire. Voici un sommaire des données :

	groupe inondé	groupe témoin
taille de l'échantillon	$n_1 = 16$	$n_2 = 13$
moyenne de l'échantillon	$\bar{x}_1 = 1,17$	$\bar{x}_2 = 1,91$
écart type de l'échantillon	$s_1 = 0,16$	$s_2 = 0,23$

Donner un intervalle de confiance à 90% pour $\mu_1 - \mu_2$, où μ_1 est l'ATP moyen pour les plants inondés et μ_2 est l'ATP moyen pour les plants non-inondés. Basé sur cet intervalle, peut-on conclure que les inondations causes une augmentation ou une diminution des niveaux moyens d'ATP ? (Supposons que le niveau d'ATP d'une plante inondée et le niveau d'ATP d'une plante non-inondée sont normalement distribués avec des variances égaux.)

A) [0,5673 ; 0,7614] ; l'inondation cause une augmentation du niveau moyen d'ATP.
B) [0,4532 ; 0,6719] ; l'inondation cause une augmentation du niveau moyen d'ATP.
C) [-0,6182 ; -0,4820] ; l'inondation cause une diminution du niveau moyen d'ATP.
D) [-0,8635 ; -0,6165] ; l'inondation cause une diminution du niveau moyen d'ATP.
E) [-0,0346 ; 0,3471] ; on ne peut pas conclure que l'inondation cause une diminution ou une augmentation du niveau moyen d'ATP.

25. Le niveau moyen de pression artérielle systolique d'une certaine population est approximativement égale à 125 mm Hg. Un sujet d'intérêt récent est que l'utilisation extensive de contraceptif oral peut provoquer une réduction de la pression artérielle systolique moyenne. Une étude est planifiée pour vérifier cette hypothèse. Les n femmes qui ont participé à cette étude ont utilisé des contraceptifs oraux pour une période de trois mois. À la fin de l'étude, leur pression artérielle systolique a été mesurée. Ces pressions artérielles systoliques furent attribuées à la variable `x` dans `R`. On affiche une sortie de `R` :

```
> mean(x)
[1] 120.4
> sd(x)
[1] 13.23
> t.test(x,mu=125,alternative="less")

        One Sample t-test

data:  x
t = -1.0998, df =?, p-value = 0.15
alternative hypothesis: true mean is less than 125
95 percent confidence interval:
    -Inf 128.067
sample estimates:
mean of x
    120.4
```

Déterminer le nombre n de participants pour cette étude.
(N.B. Dans la sortie ci-haut, on a remplacé le nombre de degrés de liberté par ?.)

A) 12 B) 40 C) 10 D) 32 E) 25

Répartition de la loi $N(0,1)$: $\Phi(z) = P(Z \leq z)$

0,09	0,08	0,07	0,06	0,05	0,04	0,03	0,02	0,01	0,00	z
0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	−3,8
0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	−3,7
0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	−3,6
0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	−3,5
0,0002	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	−3,4
0,0003	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0005	0,0005	0,0005	−3,3
0,0005	0,0005	0,0005	0,0006	0,0006	0,0006	0,0006	0,0006	0,0007	0,0007	−3,2
0,0007	0,0007	0,0008	0,0008	0,0008	0,0008	0,0009	0,0009	0,0009	0,0010	−3,1
0,0010	0,0010	0,0011	0,0011	0,0011	0,0012	0,0012	0,0013	0,0013	0,0013	−3,0
0,0014	0,0014	0,0015	0,0015	0,0016	0,0016	0,0017	0,0018	0,0018	0,0019	−2,9
0,0019	0,0020	0,0021	0,0021	0,0022	0,0023	0,0023	0,0024	0,0025	0,0026	−2,8
0,0026	0,0027	0,0028	0,0029	0,0030	0,0031	0,0032	0,0033	0,0034	0,0035	−2,7
0,0036	0,0037	0,0038	0,0039	0,0040	0,0041	0,0043	0,0044	0,0045	0,0047	−2,6
0,0048	0,0049	0,0051	0,0052	0,0054	0,0055	0,0057	0,0059	0,0060	0,0062	−2,5
0,0064	0,0066	0,0068	0,0069	0,0071	0,0073	0,0075	0,0078	0,0080	0,0082	−2,4
0,0084	0,0087	0,0089	0,0091	0,0094	0,0096	0,0099	0,0102	0,0104	0,0107	−2,3
0,0110	0,0113	0,0116	0,0119	0,0122	0,0125	0,0129	0,0132	0,0136	0,0139	−2,2
0,0143	0,0146	0,0150	0,0154	0,0158	0,0162	0,0166	0,0170	0,0174	0,0179	−2,1
0,0183	0,0188	0,0192	0,0197	0,0202	0,0207	0,0212	0,0217	0,0222	0,0228	−2,0
0,0233	0,0239	0,0244	0,0250	0,0256	0,0262	0,0268	0,0274	0,0281	0,0287	−1,9
0,0294	0,0301	0,0307	0,0314	0,0322	0,0329	0,0336	0,0344	0,0351	0,0359	−1,8
0,0367	0,0375	0,0384	0,0392	0,0401	0,0409	0,0418	0,0427	0,0436	0,0446	−1,7
0,0455	0,0465	0,0475	0,0485	0,0495	0,0505	0,0516	0,0526	0,0537	0,0548	−1,6
0,0559	0,0571	0,0582	0,0594	0,0606	0,0618	0,0630	0,0643	0,0655	0,0668	−1,5
0,0681	0,0694	0,0708	0,0721	0,0735	0,0749	0,0764	0,0778	0,0793	0,0808	−1,4
0,0823	0,0838	0,0853	0,0869	0,0885	0,0901	0,0918	0,0934	0,0951	0,0968	−1,3
0,0985	0,1003	0,1020	0,1038	0,1056	0,1075	0,1093	0,1112	0,1131	0,1151	−1,2
0,1170	0,1190	0,1210	0,1230	0,1251	0,1271	0,1292	0,1314	0,1335	0,1357	−1,1
0,1379	0,1401	0,1423	0,1446	0,1469	0,1492	0,1515	0,1539	0,1562	0,1587	−1,0
0,1611	0,1635	0,1660	0,1685	0,1711	0,1736	0,1762	0,1788	0,1814	0,1841	−0,9
0,1867	0,1894	0,1922	0,1949	0,1977	0,2005	0,2033	0,2061	0,2090	0,2119	−0,8
0,2148	0,2177	0,2206	0,2236	0,2266	0,2296	0,2327	0,2358	0,2389	0,2420	−0,7
0,2451	0,2483	0,2514	0,2546	0,2578	0,2611	0,2643	0,2676	0,2709	0,2743	−0,6
0,2776	0,2810	0,2843	0,2877	0,2912	0,2946	0,2981	0,3015	0,3050	0,3085	−0,5
0,3121	0,3156	0,3192	0,3228	0,3264	0,3300	0,3336	0,3372	0,3409	0,3446	−0,4
0,3483	0,3520	0,3557	0,3594	0,3632	0,3669	0,3707	0,3745	0,3783	0,3821	−0,3
0,3859	0,3897	0,3936	0,3974	0,4013	0,4052	0,4090	0,4129	0,4168	0,4207	−0,2
0,4247	0,4286	0,4325	0,4364	0,4404	0,4443	0,4483	0,4522	0,4562	0,4602	−0,1
0,4641	0,4681	0,4721	0,4761	0,4801	0,4840	0,4880	0,4920	0,4960	0,5000	−0,0

Répartition de la loi $N(0,1)$: $\Phi(z) = P(Z \leq z)$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

Loi T avec ν degrés de liberté

	$F_T(t) = P(T \leq t)$						
	0,6	0,75	0,9	0,95	0,975	0,99	0,995
ν	$t_{0,40;\nu}$	$t_{0,25;\nu}$	$t_{0,10;\nu}$	$t_{0,05;\nu}$	$t_{0,025;\nu}$	$t_{0,01;\nu}$	$t_{0,005;\nu}$
1	0,325	1,000	3,078	6,314	12,706	31,821	63,657
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055
13	0,259	0,694	1,350	1,771	2,160	2,650	3,012
14	0,258	0,692	1,345	1,761	2,145	2,624	2,997
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947
16	0,258	0,690	1,337	1,746	2,120	2,583	2,921
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845
21	0,257	0,686	1,323	1,721	2,080	2,518	2,831
22	0,256	0,686	1,321	1,717	2,074	2,508	2,819
23	0,256	0,685	1,319	1,714	2,069	2,500	2,807
24	0,256	0,685	1,318	1,711	2,064	2,492	2,797
25	0,256	0,684	1,316	1,708	2,060	2,485	2,787
26	0,256	0,684	1,315	1,706	2,056	2,479	2,779
27	0,256	0,684	1,314	1,703	2,052	2,473	2,771
28	0,256	0,683	1,313	1,701	2,048	2,467	2,763
29	0,256	0,683	1,311	1,699	2,045	2,464	2,756
30	0,256	0,683	1,310	1,697	2,042	2,457	2,750
∞	0,253	0,674	1,282	1,645	1,96	2,326	2,576

NB : $z_\alpha = t_{\alpha,\infty}$

Feuille de formules

- Règle d'addition : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Probabilité conditionnelle de A sachant B :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Soient $E_1, E_2, \ldots, E_k$ une partition de l'ensemble d'échantillonage (c.-à-d. ils sont
mutuellement exclusifs et exhaustifs). On a la règle des probabilités totales :

$$P(A) = P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + \cdots + P(A|E_k)P(E_k).$$

- Formule de Bayes :

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

- Tests diagnostics :

$$\begin{aligned}
\text{taux des faux-positifs} &= P(\text{Test}+|\text{Vrai}-) \\
\text{taux des faux-négatifs} &= P(\text{Test}-|\text{Vrai}+) \\
\text{spécificité} &= P(\text{Test}-|\text{Vrai}-) \\
\text{sensibilité} &= P(\text{Test}+|\text{Vrai}+) \\
\text{valeur prédictive positive} &= P(\text{Vrai}+|\text{Test}+) \\
\text{valeur prédictive négative} &= P(\text{Vrai}-|\text{Test}-)
\end{aligned}$$

- Les événements A et B sont indépendants si $P(A \cap B) = P(A)P(B)$
- Valeur espérée d'une variable aléatoire discrète X :

$$\mu = E(X) = \sum_x x f(x), \quad \text{où} \quad f(x) = P(X = x)$$

- Variance d'une variable aléatoire discrète X :

$$\sigma^2 = \mathrm{Var}(X) = \sum_x (x-\mu)^2 f(x) = \left(\sum_x x^2 f(x)\right) - \mu^2, \quad \text{où} \quad f(x) = P(X = x)$$

- Fonction de répartition d'une variable aléatoire X : $F(x) = P(X \leq x)$

- Si X suit une loi binomiale de n épreuves et de probabilité de succès p, alors
$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$
- Centrer et réduire : Si X suit une loi normale de moyenne μ et de variance σ^2, alors
$$Z = \frac{X-\mu}{\sigma} \quad \text{suit une loi normale centrée et réduite}$$
- Moyenne de l'échantillon des valeurs $x_1, \ldots, x_n$: $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i$
- Variance de l'échantillon des valeurs $x_1, \ldots, x_n$:
$$s^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x})^2 = \frac{(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)^2/n}{n-1} = \frac{(\sum_{i=1}^n x_i^2) - n\,\overline{x}^2}{n-1}$$
- Pour calculer les centiles, on ordonne les données $x_1, x_2 \ldots, x_n$ en ordre croissant :
$$y_1 \le y_2 \le \ldots \le y_n$$
On calcul :
$$\text{rang du } k\text{ième centile} = (n+1)\,k\% = r + a/b,$$
où r est la partie entière et a/b est la partie fractionnelle.
Alors le $k$ième centile est
$$k\text{ième centile} = (1 - a/b) y_r + (a/b) y_{r+1}.$$
- Statistique pour les intervalles de confiance et des tests concernant une moyenne μ (population normale) :
$$T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \quad \text{suit une loi } T \text{ avec } \nu = n-1 \text{ d.l.}$$
- Quand n est grande, la statistique pour les intervalles de confiance et des tests concernant une moyenne μ :
$$Z = \frac{\bar{X}-\mu}{S/\sqrt{n}} \text{ suit une loi normale centrée et réduite}$$
- Afin d'être $(1-\alpha)\,100\%$ confiant que l'erreur de l'estimation de la moyenne soit inférieure à E, résoudre
$$n \ge \left(\frac{z_{\alpha/2}\,\sigma}{E}\right)^2$$

— Afin d'être $(1-\alpha)\,100\%$ confiant que l'erreur de l'estimation de la proportion soit inférieure à E, résoudre

$$n \geq \left(\frac{z_{\alpha/2}}{E}\right)^2 \cdot \frac{1}{4}$$

- Statistique pour les intervalles de confiance concernant une proportion p :

$$Z = \frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \text{ suit une loi normale centrée et réduite}$$

- Statistique pour le test $H_0 : p = p_0$ concernant une proportion p :

$$Z_0 = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} \text{ suit une loi normale centrée et réduite}$$

- Statistique pour des intervalles de confiance et des tests concernant μ_1 et μ_2 de deux populations normales indépendantes à variances **égaux** :

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S_p\sqrt{1/n_1 + 1/n_2}} \text{ suit une loi } T \text{ avec } \nu = n_1 + n_2 - 2 \text{ d.l.}$$

  où

$$S_p = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}.$$

- Statistique pour des intervalles de confiance et des tests concernant μ_1 et μ_2 de deux populations normales indépendantes à variances **inégaux** :

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \text{ suit une loi } T \text{ avec } \nu \text{ d.l.}$$

  où

$$\nu = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1-1) + (s_2^2/n_2)^2/(n_2-1)}$$

- Pour grandes n_1 et n_2, la statistique pour des intervalles de confiance et des tests concernant μ_1 et μ_2 de deux populations indépendantes :

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \text{ suit une loi normale centrée et réduite}$$

- Considérons les points (x_i, y_i) pour $i = 1, \ldots, n$. L'équation pour la droite de régression estimée est

$$\hat{y} = \hat{\alpha} + \hat{\beta}\, x,$$

où

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{(\sum_{i=1}^{n} x_i\, y_i) - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)/n}{(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)^2/n},$$

et

$$\hat{\alpha} = \overline{y} - \hat{\beta}\, \overline{x}.$$

Feuille Brouillon

Feuille Brouillon

Feuille Brouillon

Feuille Brouillon

Feuille Brouillon

Feuille Brouillon