Examen de statistiques

Mercredi 20 mai 2026

Promotion 115

Antoine Géré

Document(s) autorisé(s) : ☐ Oui ☒ Non   Calculatrice autorisée : ☒ Oui ☐ Non

Remarques :

- Les exercices sont indépendants.
- Il sera tenu compte de la propreté de votre copie, ainsi que de la clarté et de la qualité de la rédaction et du raisonnement.
- **Ne pas écrire avec un crayon papier**, sauf pour dessiner et/ou annoter des croquis, le cas échéant.
- Utiliser les **notations** indiquées dans le texte et **justifier toutes vos réponses**.
- Le sujet est à conserver par l'étudiant-e.

Exercice 1 Optimisation de la micro-dose d'engrais sur Sorgho au Sahel

Dans le cadre de la diffusion de la technologie de la « *micro-dose* », des ingénieurs agronomes d'un institut de recherche au Niger recommandent l'application ciblée d'engrais minéral (NPK) directement au moment du semis.

Pour la culture du sorgho, le réglage optimal préconisé pour l'épandeur est de délivrer une dose cible de $5{,}00$ g d'engrais par poquet[1]. Une dose plus faible limiterait le démarrage de la plante (effet starter insuffisant), tandis qu'une dose plus élevée risquerait de provoquer une toxicité saline pour la semence en cas de déficit hydrique précoce.

On admet que la quantité d'engrais X délivrée par l'épandeur suit une loi normale d'écart-type connu $\sigma = 0{,}30$ g, reflétant la variabilité de la granulométrie de l'engrais et l'irrégularité du geste de l'opérateur.

Un ingénieur agronome d'une coopérative souhaite vérifier si les épandeurs distribués aux groupements de producteurs sont toujours correctement calibrés avant le début de la campagne. Il réalise un contrôle sur un échantillon de 100 poquets et obtient une masse moyenne déposée de $5{,}07$ g.

1. Définissez l'hypothèse nulle H_0 et l'hypothèse alternative H_1 pour aider l'ingenieur agronome dans sa vérification.
2. Au seuil de risque de α de 3%, déterminez la règle de décision pour accepter ou rejeter H_0.
3. L'ingénieur peut-il affirmer que les épandeurs sont bien réglés au seuil choisi ?

Correction ▼   [stat-0042]

Exercice 2 Résistance de la Chenille Légionnaire d'Automne (*Spodoptera frugiperda*)

Au Sénégal, la culture du maïs peut être gravement menacée par la chenille légionnaire d'automne. Un programme de lutte utilise un bio-pesticide à base de *Bacillus thuringiensis* (Bt).

[1] Le poquet est le trou utilisé lors du semis pour y placer plusieurs graines

Les études de laboratoire ont établi que, pour une dose d'application standard, le taux de survie normal des larves (dû à une mauvaise exposition ou à la vigueur naturelle des individus) ne dépasse pas $12,5\%$. Si le taux de survie constaté sur le terrain est significativement supérieur à ce seuil, cela indique l'émergence d'une résistance au sein de la population locale, rendant le traitement inefficace à court terme.

Lors d'une mission de surveillance épidémiologique, un ingénieur agronome prélève des larves sur une parcelle traitée. Sur un échantillon de 800 larves, il constate que 134 ont survécu après l'application du produit. L'étude a pour but d'évaluer s'il y a un début de résistance des chenilles dans cette zone.

1. Formulez l'hypothèse nulle H_0 et l'hypothèse alternative H_1. Pourquoi un test unilatéral à droite est-il ici impératif pour l'analyse ?
2. Au seuil de risque α de 1%, déterminez la règle de décision pour accepter ou rejeter H_0.
3. L'ingénieur peut-il affirmer qu'il existe une résistance émergente dans cette zone de culture au seuil choisi ?

Correction ▼ [stat-0043]

Exercice 3 Évaluation de deux stratégies de fertilisation azotée du Maïs au Sud-Bénin

Un ingénieur agronome cherche à évaluer l'impact d'un nouveau protocole de *fractionnement de l'urée* sur la culture du maïs. La pratique traditionnelle consiste à apporter la totalité de l'engrais au semis. La nouvelle méthode propose un fractionnement (1/3 au semis et 2/3 à la floraison) afin de limiter le lessivage de l'azote dû aux fortes précipitations sur la zone.

L'ingénieur dispose de données de rendement (exprimées en quintaux par hectare) récoltées sur deux zones expérimentales distinctes lors de la dernière campagne :

- *Zone A (Pratique traditionnelle) :* Le contrôle de 100 parcelles a donné un rendement moyen de 58 q/ha avec un écart-type de 3 q/ha.
- *Zone B (Fractionnement de l'urée) :* Le contrôle de 150 parcelles élémentaires a donné un rendement moyen de 56 q/ha avec un écart-type de 5 q/ha.

L'ingénieur souhaite déterminer si la différence de rendement observée entre ces deux zones est statistiquement significative au seuil de risque α de 4%.

1. Formulez l'hypothèse nulle H_0 et l'hypothèse alternative H_1.
2. Au seuil de risque α de 4%, déterminez la règle de décision pour accepter ou rejeter H_0.
3. L'ingénieur peut-il considérer que la stratégie de fertilisation a un impact réel sur le rendement moyen du maïs au seuil choisi ?

Correction ▼ [stat-0044]

Correction de l'exercice 1 ▲

1. Définition des hypothèses H_0 et H_1 :

   L'ingénieur cherche à vérifier la qualité du calibrage. Un déréglage est problématique qu'il soit en excès (toxicité saline) ou en défaut (effet starter insuffisant). Il s'agit donc d'un test bilatéral.

   Soit μ la moyenne réelle de la population des épandeurs distribués.

   - *Hypothèse nulle (H_0) :* $\mu = \mu_0 = 5,00$ g
     Interprétation agronomique : Les épandeurs sont correctement réglés et distribuent la dose cible préconisée.
   - *Hypothèse alternative (H_1) :* $\mu \neq 5,00$ g
     Interprétation agronomique : Les épandeurs sont mal réglés (ils sous-dosent ou sur-dosent de manière systématique).

2. Détermination de la règle de décision au seuil de risque $\alpha = 3\%$

   Sous l'hypothèse nulle H_0, la variable aléatoire $\overline{X}$ (qui modélise la moyenne d'un échantillon de taille 100) suit une loi normale de moyenne μ_0 et d'écart-type $\frac{\sigma}{\sqrt{n}}$:

   $$\overline{X} \sim \mathcal{N}\left(\mu_0\,;\,\frac{\sigma}{\sqrt{n}}\right) \implies \overline{X} \sim \mathcal{N}\left(5,00\,;\,\frac{0,30}{\sqrt{100}}\right) \implies \overline{X} \sim \mathcal{N}(5,00\,;\,0,03)$$

   On définit la variable de test centrée réduite Z, qui suit la loi standard $\mathcal{N}(0,1)$ sous H_0 :

   $$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{\overline{X} - 5,00}{0,03}$$

   Le test étant bilatéral au seuil $\alpha = 3\%$, la zone de rejet se divise en deux régions symétriques de probabilité $\frac{\alpha}{2} = 1,5\% = 0,015$ chacune, situées aux deux extrémités de la courbe. La zone d'acceptation centrale correspond à une probabilité de $1 - \alpha = 0,97$.

   On peut donc écrire

   $$P\,(5 - d \leq Z \leq 5 + d) = 0.97$$

   donc

   $$\begin{aligned}
   &P\left(\frac{5-d-5}{0,03} \leq Z \leq \frac{5+d-5}{0,03}\right) = 0.97\\
   &P\left(\frac{-d}{0,03} \leq Z \leq \frac{d}{0,03}\right) = 0.97\\
   &P\left(Z \leq \frac{d}{0,03}\right) - P\left(Z \leq \frac{-d}{0,03}\right) = 0.97\\
   &P\left(Z \leq \frac{d}{0,03}\right) - P\left(Z \geq \frac{d}{0,03}\right) = 0.97\\
   &P\left(Z \leq \frac{d}{0,03}\right) - \left(1 - P\left(Z \leq \frac{d}{0,03}\right)\right) = 0.97\\
   &2P\left(Z \leq \frac{d}{0,03}\right) - 1 = 0.97\\
   &P\left(Z \leq \frac{d}{0,03}\right) = 0.985
   \end{aligned}$$

   A l'aide de la table de la loi normale centrée réduite on trouve :

   $$\frac{d}{0,03} = 2,17 \quad \iff d = 0.0651$$

Règle de décision :

Si la moyenne observée est comprise entre $5 - d = 4.9349$ et $5 + d = 5,0651$, alors on accepte H_0 au seuil de 3%, autrement on rejette H_0.

3. On a mesuré une moyenne de $5,07\ g$, on rejette donc H_0 au profit de l'hypothèse alternative H_1.

   Au seuil de risque de 3%, l'ingénieur **ne peut pas affirmer que les épandeurs sont bien réglés**. L'excès de masse constaté ($+0,07$ g en moyenne) est statistiquement significatif et met en évidence un sur-dosage systématique des appareils distribués. Pour préserver l'économie des intrants des exploitants de la coopérative et écarter tout risque de toxicité saline pour le jeune sorgho au Niger, l'ingénieur doit imposer un **recalibrage complet** du matériel avant le début des semis.

Correction de l'exercice 2 ▲

Il s'agit d'un **test d'hypothèse sur une proportion (ou fréquence)** issu d'un grand échantillon ($N = 800$). Nous allons utiliser une méthode d'approximation par la **loi normale centrée réduite**.

1. Formulation des hypothèses et choix du sens du test

   - **Hypothèse nulle H_0 :** Le taux de survie théorique des larves sur le terrain reste conforme au taux normal de référence établi en laboratoire. Le bio-pesticide à base de *Bacillus thuringiensis* (Bt) fonctionne normalement, il n'y a pas d'émergence de résistance.

   $$H_0 : p = p_0 = 0,125 \quad (\text{soit } 12,5\,\%)$$

   - **Hypothèse alternative H_1 :** Le taux de survie constaté sur le terrain est supérieur au seuil normal toléré. Cela traduit une perte d'efficacité du traitement par sélection d'individus résistants.

   $$H_1 : p > 0,125$$

   Justification du test unilatéral à droite :

   L'objectif de l'ingénieur agronome est de mener une mission de surveillance épidémiologique pour détecter une menace spécifique : l'apparition d'une résistance.

   - Si le taux de survie des chenilles est inférieur ou égal à $12,5\,\%$, le traitement est considéré comme parfaitement efficace.
   - Seule une **augmentation significative** du taux de survie ($p > 0,125$) met en évidence l'anomalie recherchée. L'intérêt agronomique et la zone de rejet se situent donc exclusivement dans la queue supérieure (à droite) de la distribution gaussienne.

2. Détermination de la règle de décision au seuil de risque $\alpha = 1\%$

   Sous l'hypothèse nulle H_0, la variable aléatoire X (qui modélise le taux de survie des larves d'un échantillon de taille 800) suit une loi normale de propotion p_0 et d'écart-type $\sqrt{\dfrac{p_0(1-p_0)}{n}}$:

   $$X \sim \mathcal{N}\left(p_0\,;\sqrt{\frac{p_0(1-p_0)}{n}}\right) \implies X \sim \mathcal{N}\left(0,125\,;\sqrt{\frac{0,125 \times (1-0,125)}{800}}\right) \implies X \sim \mathcal{N}(0,125\,;0,01169)$$

   On définit la variable de test centrée réduite Z, qui suit la loi standard $\mathcal{N}(0,1)$ sous H_0 :

   $$Z = \frac{X - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{X - 0,125}{0,01169}$$

   Le test étant unilatéral à droite au seuil $\alpha = 0,01$, la zone de rejet de H_0 se trouve entièrement à l'extrême droite de la courbe de distribution.

On peut donc écrire

$$P\,(Z \leq 0,125 + d) = 0,99$$

donc

$$P\left(Z \leq \frac{0,125 + d - 0,125}{0,01169}\right) = 0,99$$
$$P\left(Z \leq \frac{d}{0,01169}\right) = 0,99$$

A l'aide de la table de la loi normale centrée réduite on trouve :

$$\frac{d}{0,01169} = 2,33 \quad \Longleftrightarrow d = 0,02724$$

Règle de décision :

Si la fréquence observée est inferieur à $0,125 + d = 0,15224$, alors on accepte H_0 au seuil de 1%, autrement on rejette H_0.

3. Sur l'échantillon prélevé par l'ingénieur, le nombre de larves survivantes est de $x = 134$ pour un total de $N = 800$ observations, donc
$$p = \frac{134}{800} = 0,1675 \quad \text{(soit } 16,75\,\%)$$

Sous l'hypothèse H_0, l'échantillon doit être de taille suffisante pour justifier l'usage de la loi normale :

(a) $N = 800 \geqslant 30$ *(Validé)*

(b) $N \times p = 800 \times 0,125 = 100 \geqslant 5$ *(Validé)*

(c) $N \times (1 - p) = 800 \times 0,875 = 700 \geqslant 5$ *(Validé)*

Les conditions mathématiques sont parfaitement respectées.

Nous constatons sur notre échantillon que la fréquence observée ($16,75\,\%$) est superieur à $15,22\,\%$, donc on rejette H_0 au seuil de risque de $1\,\%$. Il existe une différence significative entre le taux théorique et la réalité du terrain. L'expert peut affirmer qu'il y a un **début de résistance émergente** de la chenille légionnaire d'automne au bio-pesticide Bt dans cette zone de culture au Sénégal.

Le taux de survie observé ($16,75\,\%$) ne peut plus être imputé au hasard. D'un point de vue décisionnel, il est urgent de préconiser une rotation des matières actives (stratégie de gestion des résistances) pour éviter un échec total des traitements à court terme.

Correction de l'exercice 3 ▲

Il s'agit d'un **test de comparaison de deux moyennes sur de grands échantillons indépendants**. Les tailles respectives des échantillons étant supérieures ou égales à 30 ($N_A = 100$ et $N_B = 150$), nous pouvons appliquer le test appris en cours basé sur la loi normale centrée réduite en substituant les écarts-types de la population par les écarts-types des échantillonaux (s_A et s_B).

1. Formulation des hypothèses H_0 et H_1

   Soit μ_A le rendement moyen théorique du maïs sous l'effet de la pratique traditionnelle (Zone A) et μ_B le rendement moyen théorique sous l'effet du fractionnement de l'urée (Zone B).

   - **Hypothèse nulle H_0 :** La stratégie de fertilisation (fractionnement) n'a aucun impact réel sur le rendement moyen du maïs. La différence observée entre les deux moyennes échantillonnales est purement fortuite et attribuable aux fluctuations aléatoires du hasard.

$$H_0 : m_A = m_B \quad \text{(soit } m_A - m_B = 0)$$

- **Hypothèse alternative H_1 :** Le protocole de fertilisation a un impact réel sur le rendement moyen du maïs. L'écart constaté est statistiquement significatif.

$$H_1 : m_A \neq m_B \quad (\text{soit } m_A - m_B \neq 0)$$

*Note : Le test est rigoureusement **bilatéral** car l'énoncé demande d'évaluer « l'impact » du protocole sans formuler d'hypothèse a priori sur la supériorité d'une méthode par rapport à l'autre.*

2. On commence par définir les variable aléatoire de décision. Appelons E_1 (resp. E_2) l'ensemble de toutes les parcelles de maïs cultivées selon le protocole de la Zone A (resp. de la Zone B).

- Soit M_1 la variable aléatoire qui, à chaque échantillon de 100 parcelles issues de la population E_1, associe sa moyenne. Une estimation ponctuelle de la moyenne et de l'écart-type de la population de la Zone A est :

$$m_1 = 58$$

et

$$\sigma_1 = 3 \times \sqrt{\frac{100}{99}}$$

La taille des échantillons étant suffisamment grande, M_1 suit la loi $\mathcal{N}\left(m_1; \frac{\sigma_1}{\sqrt{100}}\right) = \mathcal{N}\left(58; \frac{3}{\sqrt{99}}\right)$.

- Soit M_2 la variable aléatoire qui, à chaque échantillon de 150 parcelles issues de la population E_2, associe sa moyenne. Une estimation ponctuelle de la moyenne et de l'écart-type de la population de la Zone B est :

$$m_2 = 56$$

et

$$\sigma_2 = 5 \times \sqrt{\frac{150}{149}}$$

La taille des échantillons étant suffisamment grande, M_2 suit la loi $\mathcal{N}\left(m_2; \frac{\sigma_2}{\sqrt{150}}\right) = \mathcal{N}\left(56; \frac{5}{\sqrt{149}}\right)$.

- La variable aléatoire $D = M_1 - M_2$ suit également une loi normale de paramètres :

$$E(D) = E(M_1) - E(M_2) = m_1 - m_2$$

$$V(D) = V(M_1) + V(M_2) = \frac{9}{99} + \frac{25}{149} = 0,2587$$

. D'où $\sigma_D = 0,51$. Donc D suit la loi $\mathcal{N}(m_1 - m_2; 0,51)$. D est la variable aléatoire de décision.

Supposons que l'hypothèse H_0 soit vraie, on a alors $m_1 - m_2 = 0$, alors D suit la loi normale $\mathcal{N}(0; 0,51)$. On cherche alors le réel d tel que

$$P(-d \leqslant D \leqslant d) = 0,96. \tag{1}$$

la variable aléatoire $T = \dfrac{D}{0,51}$ suit la loi normale centrée réduite $\mathcal{N}(0,1)$, on a alors :

$$\begin{aligned}
(1) \iff & P(-d < 0,51T < d) = 0,96 \\
\iff & P\left(-\frac{d}{0,51} \leqslant T \leqslant \frac{d}{0,51}\right) = 0,96 \\
\iff & 2\Pi\left(\frac{d}{0,51}\right) - 1 = 0,96 \\
\iff & \Pi\left(\frac{d}{0,51}\right) = 0,98
\end{aligned}$$

On trouve alors $\frac{d}{0,51} = 2,05$ soit $d = 1,0455 \approx 1,05$. Pour un seuil de risque de 4%, la zone critique est :

$$]-\infty; -1,05[\cup]1,05; +\infty[$$

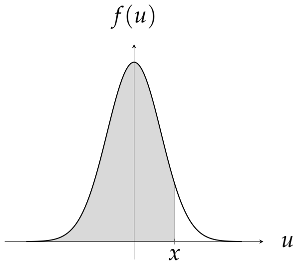
Règle de décision.

Si la différence des moyennes des deux échantillons est inférieure à $-1,05$ ou supérieure à $1,05$, alors l'hypothèse H_0, n'est pas validée.
Si la différence des moyennes des deux échantillons est comprise entre $-1,05$ et $1,05$, l'hypothèse H_0 est validée.

3. La différence des moyennes des deux échantillons est $58 - 56 = 2$.
D'après la règle de décision, on rejette H_0 et on décide que la stratégie de fertilisation a un impact réel sur le rendement moyen du maïs au seuil de risque de 4%.

Table de la Loi Normale centrée réduite

Probabilité de trouver une valeur inférieure à x.



$$F(x) = \int_{-\infty}^{x} f(u)\,du = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\,du$$

x	**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	0,9977	0,9978	0,9979	0,9979	0,9980
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