Travaux dirigés

Promotion 115

Table des matières

1 Test d'ajustement du χ^2

Exercice 1 Analyse de la répartition des variétés de maïs au Sénégal

Au Sénégal, pour faire face aux irrégularités des pluies, une ONG a distribué aux agriculteurs un mélange de trois variétés de semences de maïs, chacune ayant une tolérance différente à la sécheresse. L'ONG avait préparé des sacs contenant théoriquement :

- 50% de semences de variété A (cycle très court, très résistante au manque d'eau) ;
- 30% de semences de variété B (cycle moyen, moyennement résistante) ;
- 20% de semences de variété C (cycle long, intensive mais gourmande en eau).

Après la récolte, un conseiller agricole souhaite vérifier si la proportion des trois variétés observées dans les greniers des producteurs correspond toujours à la répartition théorique du sac initial, ou si les conditions du milieu ont favorisé le développement d'une variété plutôt qu'une autre.

Il tire au hasard et identifie 400 épis de maïs dans les stocks d'un village et obtient la répartition suivante :

- Variété A : 224 épis
- Variété B : 101 épis
- Variété C : 75 épis

1. Formulez l'hypothèse nulle H_0 et l'hypothèse alternative H_1.
2. Déterminez les effectifs théoriques attendus pour chaque variété sous l'hypothèse H_0 pour l'échantillon considéré.
3. Calculez la valeur de la statistique du χ^2 observée, noté χ^2_{obs}.
4. Au seuil de risque $\alpha = 5\%$, déterminer la valeur seuil lue dans la table statistique, notée χ^2_{seuil}.
5. Énoncez la règle de décision, puis déduisez-en si l'on doit accepter ou rejeter l'hypothèse nulle H_0.

Correction ▼ [stat-0045]

Exercice 2 Infestation des cultures de sorgho par les pucerons au Botswana

Dans la région du *Central District* au Botswana, la culture du sorgho est essentielle pour la sécurité alimentaire locale, mais elle est fréquemment attaquée par le puceron du sorgho (*Melanaphis sacchari*). Un agronome cherche à comprendre comment ces insectes colonisent les parcelles afin de guider les petits producteurs dans leurs traitements.

Selon la théorie écologique, si les pucerons arrivent de manière totalement aléatoire et indépendante d'une plante à l'autre, le nombre de colonies de pucerons observées par plant de sorgho doit suivre une *loi de Poisson* :

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

L'agronome examine un échantillon de 200 plants de sorgho choisis au hasard dans une parcelle. Les observations récoltées sont résumées dans le tableau suivant :

Nombre de colonies par plant	*Nombre de plants de sorgho*
0	85
1	70
2	32
3 ou plus	13
Total	200

À partir de ces données, l'agronome a calculé que le nombre moyen de colonies par plant sur cet échantillon est de $\lambda = 0,9$.

L'agronome souhaite tester, au seuil de risque $\alpha = 5\%$, si la distribution des pucerons observée sur le terrain s'ajuste bien à la loi de Poisson de paramètre $\lambda = 0,9$.

1. Formulez l'hypothèse nulle H_0 et l'hypothèse alternative H_1.
2. En utilisant la relation de la loi de Poisson (avec $\lambda = 0,9$), calculez la probabilité théorique associée à chaque catégorie, c'est à dire $P(X = 0,\ 1,\ 2,\ 3)$, où X désigne le nombre de colonies par plant.
3. En déduire les effectifs théoriques sous l'hypothèse H_0 pour l'échantillon constitué.
4. Calculez la valeur de la statistique du χ^2 observée, χ^2_{obs}.
5. Au seuil de risque $\alpha = 5\%$, déterminer la valeur seuil lue dans la table statistique, notée χ^2_{seuil}.
6. Énoncez la règle de décision, puis déduisez-en si l'on doit accepter ou rejeter l'hypothèse nulle H_0.

Correction ▼ [stat-0046]

2 Test d'indépendance du χ^2

Exercice 3 Influence de la fertilisation sur la qualité du cacao en Côte d'Ivoire

En Côte d'Ivoire, premier producteur mondial de cacao, une coopérative agricole souhaite savoir si le type d'engrais utilisé par les planteurs a une influence directe sur la qualité des fèves de cacao récoltées.

Les fèves de cacao sont classées en deux catégories de qualité : *Qualité Supérieure* (fèves bien fermentées et de gros calibre) et *Qualité Standard*.

L'ingénieur de la coopérative mène une enquête auprès de 500 producteurs de la région de Soubré. Il répartit les données selon le type d'engrais appliqué (Engrais Chimique, Engrais Organique/Compost, ou Aucun engrais) et la qualité dominante de la récolte. Il obtient le tableau de contingence (effectifs observés) suivant :

Type d'engrais	*Qualité Supérieure*	*Qualité Standard*	*Total*
Engrais Chimique	110	90	200
Engrais Organique	95	55	150
Aucun engrais	65	85	150
Total	270	230	500

L'objectif est d'utiliser le test du χ^2 d'indépendance au seuil de risque $\alpha = 5\%$ pour déterminer s'il existe un lien significatif entre le mode de fertilisation et la qualité du cacao.

1. Formulez l'hypothèse nulle H_0 et l'hypothèse alternative H_1 associées.
2. Calculez les effectifs théoriques attendus pour chaque case du tableau de contigence sous l'hypothèse H_0.
3. Calculez la valeur de la statistique du χ^2, notée χ^2_{obs}.
4. Au seuil de risque $\alpha = 5\%$, déterminer la valeur seuil lue dans la table statistique, notée χ^2_{seuil}.
5. Énoncez la règle de décision, puis déduisez-en si l'on doit accepter ou rejeter l'hypothèse nulle H_0.

Correction ▼ [stat-0047]

Exercice 4 Pratiques de protection des cultures et circuits de commercialisation des mangues au Vietnam

Dans la province du Delta du Mékong au Vietnam, les producteurs de mangues font face à des exigences de marché de plus en plus strictes concernant l'utilisation des pesticides. Un sociologue rural et un ingénieur agronome s'associent pour étudier si la méthode principale de protection des vergers choisie par un agriculteur est liée à son mode de commercialisation (son circuit de vente principal).

Les chercheurs interrogent un échantillon aléatoire de 400 producteurs de mangues et enregistrent :

- La méthode de protection principale utilisée : *Lutte chimique* (pesticides de synthèse) ou *Lutte agroécologique* (pièges à phéromones, filets de protection, biopesticides).
- Le *circuit de commercialisation* principal : *Marché local* (intermédiaires traditionnels, marchés de district) ou *Exportation / Supermarchés* (filières organisées à forte valeur ajoutée).

Après l'enquête, ils regroupent les données dans le tableau de contingence (effectifs observés) suivant :

Méthode de protection	*Marché local*	*Exportation / Supermarchés*	*Total*
Lutte chimique	190	60	250
Lutte agroécologique	70	80	150
Total	260	140	400

L'objectif est d'utiliser le test du χ^2 d'indépendance au seuil de risque $\alpha = 5\%$ pour déterminer si la méthode de protection des cultures et le circuit de vente sont indépendants, ou s'il existe une association significative entre ces deux caractères.

1. Formulez l'hypothèse nulle H_0 et l'hypothèse alternative H_1 associées.
2. Calculez les effectifs théoriques attendus pour chaque case du tableau de contigence sous l'hypothèse H_0.
3. Calculez la valeur de la statistique du χ^2, notée χ^2_{obs}.
4. Au seuil de risque $\alpha = 5\%$, déterminer la valeur seuil lue dans la table statistique, notée χ^2_{seuil}.
5. Énoncez la règle de décision, puis déduisez-en si l'on doit accepter ou rejeter l'hypothèse nulle H_0.

Correction ▼ [stat-0048]

3 Test d'homogénéité du χ^2

Exercice 5 Consommation de bois-énergie et dynamique de déforestation au Mali

Au Mali, la surexploitation du bois de feu pour la cuisine est l'une des causes principales de la déforestation. Pour adapter au mieux les campagnes de sensibilisation et proposer des alternatives énergétiques efficaces, une ONG souhaite savoir si les habitudes de consommation en énergie de cuisson des foyers sont identiques d'une zone géographique à l'autre.

Les enquêteurs ont sélectionné et interrogé *trois échantillons indépendants* de 200 foyers chacun, situés dans trois régions économiques distinctes : la région de Mopti (zone sahélienne sèche), la région de Sikasso (zone soudanienne plus verte) et la zone urbaine de Bamako.

Chaque foyer a été classé selon sa source d'énergie principale pour la cuisine : *Bois de feu*, *Charbon de bois*, ou *Gaz*. Les résultats sont regroupés dans le tableau suivant :

Source d'énergie principale	*Échantillon Mopti*	*Échantillon Sikasso*	*Échantillon Bamako*	*Total*
Bois de feu	140	110	30	280
Charbon de bois	50	70	110	230
Gaz	10	20	60	90
Total	200	200	200	600

L'objectif est d'utiliser le test du χ^2 d'homogénéité au seuil de risque $\alpha = 5\%$ pour déterminer si la structure de consommation en énergie de cuisson est identique (homogène) dans ces trois régions.

1. Formulez l'hypothèse nulle H_0 et l'hypothèse alternative H_1 associées.
2. Calculez les effectifs théoriques attendus pour chaque case du tableau de contigence sous l'hypothèse H_0.
3. Calculez la valeur de la statistique du χ^2, notée χ^2_{obs}.
4. Au seuil de risque $\alpha = 5\%$, déterminer la valeur seuil lue dans la table statistique, notée χ^2_{seuil}.
5. Énoncez la règle de décision, puis déduisez-en si l'on doit accepter ou rejeter l'hypothèse nulle H_0.

Correction ▼ [stat-0049]

Exercice 6 Modernisation des systèmes d'irrigation face au stress hydrique dans les Andes (Pérou)

Dans la région de Cusco au Pérou, le changement climatique entraîne une fonte rapide des glaciers, réduisant la disponibilité de l'eau pour les cultures de pommes de terre et de maïs. Pour aider les communautés paysannes, le ministère de l'Agriculture a financé des projets de modernisation des infrastructures d'irrigation.

Un ingénieur agronome souhaite vérifier si l'adoption de ces nouvelles technologies s'est faite de la même manière (*de façon homogène*) dans les différentes vallées, ou si certaines communautés sont restées plus traditionnelles que d'autres.

Il sélectionne *trois communautés indépendantes* et interroge *150 agriculteurs dans chacune d'elles* (soit un échantillon total de 450 producteurs). Il classe chaque exploitation selon son système d'irrigation principal : *Irrigation traditionnelle par gravité* (canaux de terre en surface), *Aspersion mécanique*, ou *Goutte-à-goutte* (haute efficacité).

Voici les effectifs observés lors de son enquête :

Système d'irrigation principal	*Communauté A* (Haute vallée)	*Communauté B* (Moyenne vallée)	*Communauté C* (Basse vallée)	*Total*
Gravité (traditionnel)	90	65	40	195
Aspersion	45	55	65	165
Goutte-à-goutte	15	30	45	90
Total	150	150	150	450

L'objectif est d'utiliser le test du χ^2 d'homogénéité au seuil de risque $\alpha = 5\%$ pour déterminer si les trois communautés présentent la même structure d'équipement en systèmes d'irrigation.

1. Formulez l'hypothèse nulle H_0 et l'hypothèse alternative H_1 associées.
2. Calculez les effectifs théoriques attendus pour chaque case du tableau de contigence sous l'hypothèse H_0.
3. Calculez la valeur de la statistique du χ^2, notée χ^2_{obs}.
4. Au seuil de risque $\alpha = 5\%$, déterminer la valeur seuil lue dans la table statistique, notée χ^2_{seuil}.
5. Énoncez la règle de décision, puis déduisez-en si l'on doit accepter ou rejeter l'hypothèse nulle H_0.

Correction ▼ [stat-0050]

---

4 Analyse de la variance

Exercice 7 Effet du paillage biologique sur le rendement des tomates dans le Delta du Nil (Égypte)

Dans le delta du Nil en Égypte, les producteurs de tomates font face à une évaporation intense de l'eau des sols et à des problèmes de remontée de salinité. Pour limiter ces phénomènes, un agronome d'une station de recherche propose de tester l'impact de différents paillages végétaux déposés à la surface du sol.

Il met en place une expérimentation sur *12 parcelles identiques*, réparties en 3 groupes de 4 parcelles chacun (les 3 modalités du facteur « Paillage ») :

- *Groupe 1 : Paillage de paille de riz* (résidu très abondant dans le delta).
- *Groupe 2 : Paillage de tiges de roseaux broyées* (plante locale poussant au bord des canaux).
- *Groupe 3 : Témoin sur sol nu* (pratique traditionnelle).

À la récolte, le rendement en kilogrammes (kg) par parcelle est mesuré. Les résultats sont les suivants :

Paille de riz (A)	*Roseaux broyés* (B)	*Témoin sol nu* (C)
12	14	9
10	17	7
13	13	6
11	16	10

L'objectif est de vérifier, au seuil de risque $\alpha = 5\%$, si le type de paillage a une influence significative sur le rendement moyen des tomates.

1. Énoncez l'hypothèse nulle H_0 et l'hypothèse alternative H_1.
2. Calculez la moyenne de chaque groupe ($\overline{X}_A$, $\overline{X}_B$, $\overline{X}_C$), ainsi que la moyenne générale $\overline{X}$ de l'ensemble des 12 parcelles.
3. Calculez la statistique F observée, notée F_{obs}.
4. Au seuil de risque $\alpha = 5\%$, déterminer la valeur seuil lue dans la table statistique, notée F_{seuil}.
5. Énoncez la règle de décision, puis déduisez-en si l'on doit accepter ou rejeter l'hypothèse nulle H_0.

Correction ▼ [stat-0051]

Exercice 8 Effet du type d'engrais azoté sur le rendement du théier au Kenya

Dans la région de Kericho au Kenya, les petits producteurs de thé cherchent à maximiser la récolte de jeunes bourgeons et de feuilles tendres (la cueillette fine « deux feuilles et un bourgeon »). L'azote est l'élément nutritif principal indispensable pour stimuler la croissance foliaire. La coopérative locale souhaite tester trois formules de fertilisation différentes afin de conseiller au mieux ses membres.

Un agronome sélectionne *12 parcelles de théiers identiques* et applique trois traitements distincts (répartis en 3 groupes de 4 parcelles) :

- *Groupe 1 : Engrais chimique classique (Urée)*
- *Groupe 2 : Compost organique enrichi*
- *Groupe 3 : Fumier de ferme traditionnel*

Après plusieurs semaines de suivi, la masse de feuilles vertes récoltées (en kg) par parcelle est mesurée. Les résultats sont les suivants :

Urée (A)	*Compost enrichi* (B)	*Fumier de ferme* (C)
16	12	8
14	10	11
13	11	9
17	11	12

L'objectif est de vérifier, au seuil de risque $\alpha = 5\%$, si le type de fertilisant a une influence significative sur le rendement moyen du thé.

1. Énoncez l'hypothèse nulle H_0 et l'hypothèse alternative H_1.
2. Calculez la moyenne de chaque groupe ($\overline{X}_A$, $\overline{X}_B$, $\overline{X}_C$), ainsi que la moyenne générale $\overline{X}$ de l'ensemble des 12 parcelles.
3. Calculez la statistique F observée, notée F_{obs}.
4. Au seuil de risque $\alpha = 5\%$, déterminer la valeur seuil lue dans la table statistique, notée F_{seuil}.
5. Énoncez la règle de décision, puis déduisez-en si l'on doit accepter ou rejeter l'hypothèse nulle H_0.

Correction ▼ [stat-0052]

Exercice 9 Impact de l'agroforesterie (niveau d'ombrage) sur le rendement du café en Indonésie

Sur l'île de Sumatra en Indonésie, une coopérative de caféiculteurs de variété *Robusta* souhaite optimiser sa production tout en pratiquant une agriculture durable. Les agronomes locaux expliquent que maintenir des arbres d'ombrage au milieu des caféiers permet de réguler la température et de protéger les plants, mais qu'un ombrage trop dense pourrait freiner la photosynthèse et réduire la récolte.

Pour identifier le meilleur équilibre, un producteur teste trois niveaux d'ombrage différents sur *12 parcelles expérimentales de tailles identiques* (réparties en 3 groupes de 4 parcelles) :

- *Groupe 1 : Ombrage Fort* (densité d'arbres forestiers élevée).
- *Groupe 2 : Ombrage Modéré* (densité d'arbres intermédiaire).
- *Groupe 3 : Plein soleil* (système en monoculture traditionnelle, aucun arbre).

À la récolte, la quantité de café grain (en kg) produite par parcelle est pesée. Les résultats sont les suivants :

Ombrage Fort (A)	*Ombrage Modéré* (B)	*Plein soleil* (C)
10	15	11
8	18	9
11	14	8
7	17	12

L'objectif est de vérifier au seuil de risque $\alpha = 5\%$ si le niveau d'ombrage a une influence significative sur le rendement moyen des caféiers.

1. Énoncez l'hypothèse nulle H_0 et l'hypothèse alternative H_1.
2. Calculez la moyenne de chaque groupe ($\overline{X}_A$, $\overline{X}_B$, $\overline{X}_C$), ainsi que la moyenne générale $\overline{X}$ de l'ensemble des 12 parcelles.
3. Calculez la statistique F observée, notée F_{obs}.
4. Au seuil de risque $\alpha = 5\%$, déterminer la valeur seuil lue dans la table statistique, notée F_{seuil}.
5. Énoncez la règle de décision, puis déduisez-en si l'on doit accepter ou rejeter l'hypothèse nulle H_0.

Correction ▼ [stat-0053]

Exercice 10 Effet de la fréquence de taille sur le rendement des cacaoyers au Guatemala

Dans la région d'Alta Verapaz au Guatemala, une association de petits producteurs de cacao de variété *Criollo* cherche à augmenter sa production de fèves de qualité pour le marché équitable. Les agronomes de l'association affirment que la taille régulière des branches élimine l'excès d'ombre et limite l'apparition de maladies cryptogamiques (champignons comme la moniliose) qui pourrissent les cabosses. Cependant, une taille trop fréquente demande beaucoup de main-d'œuvre et pourrait stresser les arbres.

Pour tester cela scientifiquement sur le terrain, un producteur applique trois rythmes de taille différents sur **12 parcelles homogènes** (réparties en 3 groupes de 4 parcelles) :

- *Groupe 1 : Taille Mensuelle* (entretien très intensif).
- *Groupe 2 : Taille Trimestrielle* (entretien modéré).
- *Groupe 3 : Témoin sans taille* (pratique traditionnelle).

À la fin de la saison, le poids de fèves de cacao séchées (en kg) par parcelle est enregistré. Les résultats sont les suivants :

Taille Mensuelle (A)	*Taille Trimestrielle* (B)	*Témoin sans taille* (C)
11	15	8
9	13	11
12	14	9
8	14	8

L'objectif est de vérifier, au seuil de risque $\alpha = 5\%$, si la fréquence de taille a une influence significative sur le rendement moyen en cacao.

1. Énoncez l'hypothèse nulle H_0 et l'hypothèse alternative H_1.
2. Calculez la moyenne de chaque groupe ($\overline{X}_A$, $\overline{X}_B$, $\overline{X}_C$), ainsi que la moyenne générale $\overline{X}$ de l'ensemble des 12 parcelles.
3. Calculez la statistique F observée, notée F_{obs}.
4. Au seuil de risque $\alpha = 5\%$, déterminer la valeur seuil lue dans la table statistique, notée F_{seuil}.
5. Énoncez la règle de décision, puis déduisez-en si l'on doit accepter ou rejeter l'hypothèse nulle H_0.

Correction ▼ [stat-0054]

Correction de l'exercice 1 ▲

1. **Formulation des hypothèses :**
   - **Hypothèse nulle (H_0) :** La répartition des variétés de maïs récoltées dans le village est **conforme** à la distribution théorique initiale du sac (50 %, 30 %, 20 %). Les conditions du milieu n'ont pas favorisé une variété plutôt qu'une autre.
   - **Hypothèse alternative (H_1) :** La répartition observée n'est **pas conforme** à la distribution théorique. Des facteurs exterieurs ont modifié de manière significative les proportions initiales.

2. **Détermination des effectifs théoriques (e_i) :**
   L'échantillon total comprend $N = 400$ épis de maïs. Sous l'hypothèse nulle H_0, les effectifs attendus pour chaque variété se calculent en multipliant la taille de l'échantillon par la proportion théorique associée ($n_i^\star = N \times p_i$) :
   - **Variété A :** $n_A^\star = 400 \times 0,50 = 200$ épis
   - **Variété B :** $n_B^\star = 400 \times 0,30 = 120$ épis
   - **Variété C :** $n_C^\star = 400 \times 0,20 = 80$ épis

   Condition de validité : Tous les effectifs théoriques sont strictement supérieurs à 5 . L'application du test du χ^2 d'ajustement est donc valide.

3. **Calcul de la statistique du χ^2 observée, notée χ^2_{obs} :**
   La relation est :
   $$\chi^2_{\text{obs}} = \sum_{i=1}^{k} \frac{(n_i - n_i^\star)^2}{n_i^\star}$$
   où k=3, car 3 variétés. Calculons la contribution de chaque variété (écart entre l'observé n_i et le théorique $n_i^\star$) :
   - **Variété A :** $\frac{(224-200)^2}{200} = \frac{24^2}{200} = \frac{576}{200} = 2,88$
   - **Variété B :** $\frac{(101-120)^2}{120} = \frac{(-19)^2}{120} = \frac{361}{120} \approx 3,0083$
   - **Variété C :** $\frac{(75-80)^2}{80} = \frac{(-5)^2}{80} = \frac{25}{80} = 0,3125$

   En faisant la somme de ces trois contributions, on obtient la statistique observée :
   $$\chi^2_{\text{obs}} = 2,88 + 3,0083 + 0,3125 = 6,2008$$

4. **Détermination de la valeur seuil, notée χ^2_{seuil} :** Le nombre de degrés de liberté (ddl) pour un test d'ajustement dépend du nombre de modalités k moins 1. Ici, nous avons $k = 3$ variétés, donc :
   $$\text{ddl} = k - 1 = 3 - 1 = 2$$
   En lisant dans la table statistique de la loi du χ^2 pour ddl $= 2$ au seuil de risque $\alpha = 5\%$, on trouve :
   $$\chi^2_{\text{seuil}} = 5,991$$

5. **Règle de décision et conclusion du test :**
   Si $\chi^2_{\text{obs}} > \chi^2_{\text{seuil}}$, **alors** on rejette l'hypothèse nulle H_0. **Sinon** si $\chi^2_{\text{obs}} \leq \chi^2_{\text{seuil}}$, on accepte H_0.
   Nous comparons nos valeurs.
   $$\chi^2_{\text{obs}} = 6,201 > \chi^2_{\text{seuil}} = 5,991$$
   La statistique calculée entre de justesse dans la zone de rejet. On rejette l'hypothèse nulle H_0 au profit de l'hypothèse alternative H_1.

   La structure des stocks de maïs récoltés dans le village ne correspond plus de manière significative à la composition théorique du sac distribué par l'ONG.

   L'analyse des écarts nous montre que la **variété A** (cycle court, très résistante) est surreprésentée dans les greniers (224 observés contre 200 attendus), tandis que la **variété B** est sous-représentée (101 observés contre 120 attendus).

Correction de l'exercice 2 ▲

1. **Formulation des hypothèses :**
   - **Hypothèse nulle (H_0) :** La distribution du nombre de colonies de pucerons par plant de sorgho suit une **loi de Poisson** de paramètre $\lambda = 0,9$. L'arrivée des pucerons se font de manière totalement aléatoire et indépendante sur les plants.
   - **Hypothèse alternative (H_1) :** La distribution observée **ne suit pas** une loi de Poisson de paramètre $\lambda = 0,9$. La colonisation n'est pas purement aléatoire (il peut y avoir des effets d'agrégation ou de répulsion entre les pucerons).

2. **Calcul des probabilités théoriques :** On utilise la formule donnée :

$$P(X = k) = e^{-0,9}\,\frac{0,9^k}{k!}$$

   - **Pour** $k = 0$ **:** $P(X = 0) = e^{-0,9}\dfrac{0,9^0}{0!} = 0,40657 \times \dfrac{1}{1} \approx 0,4066$
   - **Pour** $k = 1$ **:** $P(X = 1) = e^{-0,9}\dfrac{0,9^1}{1!} = 0,40657 \times \dfrac{0,9}{1} \approx 0,3659$
   - **Pour** $k = 2$ **:** $P(X = 2) = e^{-0,9}\dfrac{0,9^2}{2!} = 0,40657 \times \dfrac{0,81}{1 \times 2} \approx 0,1647$
   - **Pour** $k \geq 3$ **(« 3 ou plus ») :** Par complémentarité, la probabilité de cette dernière classe ouverte s'obtient en soustrayant les autres à la probabilité totale (qui vaut 1) :

$$P(X \geq 3) = 1 - (P(X = 0) + P(X = 1) + P(X = 2)) = 1 - (0,4066 + 0,3659 + 0,1647) = 0,0628$$

3. **Détermination des effectifs théoriques, noté $n_i^\star$:** L'échantillon comprend $N = 200$ plants de sorgho. Les effectifs attendus sous H_0 s'obtiennent par la relation $e_i = N \times P(X = k)$:
   - **Classe 0 :** $n_0^\star = 200 \times 0,4066 = 81,32$ plants
   - **Classe 1 :** $n_1^\star = 200 \times 0,3659 = 73,18$ plants
   - **Classe 2 :** $n_2^\star = 200 \times 0,1647 = 32,94$ plants
   - **Classe 3 ou plus :** $n_3^\star = 200 \times 0,0628 = 12,56$ plants

   Condition de validité : Tous les effectifs théoriques calculés sont bien supérieurs à 5. Les classes n'ont pas besoin d'être regroupées, le test du χ^2 est applicable.

4. **Calcul de la statistique du χ^2 observée, notée χ^2_{obs} :** On applique la formule

$$\chi^2_{\text{obs}} = \sum_{i=1}^{k} \frac{(n_i - n_i^\star)^2}{n_i^\star}$$

   sur nos $k = 4$ catégories :
   - **Classe 0 :** $\dfrac{(85 - 81,32)^2}{81,32} = \dfrac{3,68^2}{81,32} = \dfrac{13,5424}{81,32} \approx 0,1665$
   - **Classe 1 :** $\dfrac{(70 - 73,18)^2}{73,18} = \dfrac{(-3,18)^2}{73,18} = \dfrac{10,1124}{73,18} \approx 0,1382$
   - **Classe 2 :** $\dfrac{(32 - 32,94)^2}{32,94} = \dfrac{(-0,94)^2}{32,94} = \dfrac{0,8836}{32,94} \approx 0,0268$
   - **Classe 3 ou plus :** $\dfrac{(13 - 12,56)^2}{12,56} = \dfrac{0,44^2}{12,56} = \dfrac{0,1936}{12,56} \approx 0,0154$

En faisant la somme de ces contributions :

$$\chi^2_{\text{obs}} = 0,1665 + 0,1382 + 0,0268 + 0,0154 = 0,3469$$

5. **Détermination de la valeur seuil, notée χ^2_{seuil} :** Pour un test d'ajustement où un paramètre de la loi théorique (ici λ) a été estimé à partir des données de l'échantillon, le nombre de degrés de liberté (ddl) se calcule par :

$$\text{ddl} = k - 1 - r$$

où $k = 4$ (le nombre de classes) et $r = 1$ (le nombre de paramètres estimés, ici λ).

$$\text{ddl} = 4 - 1 - 1 = 2$$

En lisant dans la table de la loi du χ^2 pour 2 ddl au seuil de risque $\alpha = 5\,\%$, on trouve :

$$\chi^2_{\text{seuil}} = 5,991$$

6. **Règle de décision et conclusion agronomique :**

Si $\chi^2_{\text{obs}} > \chi^2_{\text{seuil}}$, **alors** on rejette l'hypothèse nulle H_0. **Sinon** si $\chi^2_{\text{obs}} \leq \chi^2_{\text{seuil}}$, on accepte H_0.
Ici, nous constatons que :

$$\chi^2_{\text{obs}} = 0,3469 \leq \chi^2_{\text{seuil}} = 5,991$$

La statistique observée est très nettement inférieure à la valeur critique. **On accepte l'hypothèse nulle** H_0.

L'ajustement du nombre de colonies de pucerons à une loi de Poisson de paramètre $\lambda = 0,9$ est excellent. Les données de terrain confirment parfaitement l'hypothèse écologique de départ : les pucerons colonisent les plants de sorgho de manière totalement indépendante et aléatoire dans cette parcelle du Botswana.

Pour l'agronome, cela implique qu'il n'y a pas d'effet d'attraction locale ou de foyers d'infestation initiaux massifs. Pour guider les producteurs, les traitements localisés par taches ne sont pas pertinents : le risque est réparti uniformément dans le champ, et un suivi global de l'évolution moyenne par un échantillonnage régulier suffit pour décider du déclenchement d'une intervention.

Correction de l'exercice 3 ▲

1. **Formulation des hypothèses :**

   - **Hypothèse nulle (H_0) :** Le type d'engrais utilisé et la qualité des fèves de cacao récoltées sont **indépendants**. Le choix du mode de fertilisation n'a pas d'influence sur la qualité du cacao.
   - **Hypothèse alternative (H_1) :** Le type d'engrais et la qualité des fèves ne sont **pas indépendants** (ils sont dépendants). Il existe un lien statistique significatif entre le mode de fertilisation et la qualité de la récolte.

2. **Calcul des effectifs théoriques, noté $n^\star_{ij}$:**

   La relation générale pour calculer l'effectif théorique d'une case est :

$$n^\star_{ij} = \frac{\text{Total Ligne}_i \times \text{Total Colonne}_j}{\text{Total Global}} = \frac{n_{i\cdot} \times n_{\cdot j}}{N}$$

   Calculons la valeur attendue sous H_0 pour chacune des 6 cases du tableau de contingence (Total global $N = 500$) :

   - **Ligne « Engrais Chimique » (Total = 200) :**
     - Qualité Supérieure : $n^\star_{11} = \dfrac{200 \times 270}{500} = 108$
     - Qualité Standard : $n^\star_{12} = \dfrac{200 \times 230}{500} = 92$

- **Ligne « Engrais Organique » (Total = 150) :**
  - Qualité Supérieure : $n_{21}^{\star} = \dfrac{150 \times 270}{500} = 81$
  - Qualité Standard : $n_{22}^{\star} = \dfrac{150 \times 230}{500} = 69$
- **Ligne « Aucun engrais » (Total = 150) :**
  - Qualité Supérieure : $n_{31}^{\star} = \dfrac{150 \times 270}{500} = 81$
  - Qualité Standard : $n_{32}^{\star} = \dfrac{150 \times 230}{500} = 69$

Condition de validité : Tous les effectifs théoriques sont largement supérieurs à 5. Les conditions d'application du test du χ^2 sont donc respectées.

3. **Calcul de la statistique du χ^2 observée, noté χ^2_{obs} :**

   On applique la relation

$$\chi^2_{\text{obs}} = \sum_{ij} \frac{(n_{ij} - n_{ij}^{\star})^2}{n_{ij}^{\star}}$$

   pour les 6 cases du tableau :

   - **Engrais Chimique :**
     - Qualité Supérieure : $\dfrac{(110-108)^2}{108} = \dfrac{2^2}{108} = \dfrac{4}{108} \approx 0,0370$
     - Qualité Standard : $\dfrac{(90-92)^2}{92} = \dfrac{(-2)^2}{92} = \dfrac{4}{92} \approx 0,0435$
   - **Engrais Organique :**
     - Qualité Supérieure : $\dfrac{(95-81)^2}{81} = \dfrac{14^2}{81} = \dfrac{196}{81} \approx 2,4198$
     - Qualité Standard : $\dfrac{(55-69)^2}{69} = \dfrac{(-14)^2}{69} = \dfrac{196}{69} \approx 2,8406$
   - **Aucun engrais :**
     - Qualité Supérieure : $\dfrac{(65-81)^2}{81} = \dfrac{(-16)^2}{81} = \dfrac{256}{81} \approx 3,1605$
     - Qualité Standard : $\dfrac{(85-69)^2}{69} = \dfrac{16^2}{69} = \dfrac{256}{69} \approx 3,7101$

   En faisant la somme de ces 6 contributions :

$$\chi^2_{\text{obs}} = 0,0370 + 0,0435 + 2,4198 + 2,8406 + 3,1605 + 3,7101 = 12,2115$$

4. **Détermination de la valeur seuil, noté χ^2_{seuil} :**

   Soit L le nombre de lignes et C le nombre de colonnes du tableau de contingence (hors totaux). Nous avons $L = 3$ types d'engrais et $C = 2$ catégories de qualité.

$$\text{ddl} = (L-1) \times (C-1) = (3-1) \times (2-1) = 2 \times 1 = 2$$

   En lisant dans la table de la loi du χ^2 pour 2 degrés de liberté (ddl) au seuil de risque $\alpha = 5\%$, on obtient :

$$\chi^2_{\text{seuil}} = 5,991$$

5. **Règle de décision et conclusion agronomique :**

   Si $\chi^2_{\text{obs}} > \chi^2_{\text{seuil}}$, **alors** on rejette l'hypothèse nulle H_0. **Sinon** si $\chi^2_{\text{obs}} \leq \chi^2_{\text{seuil}}$, on accepte H_0.

   Nous comparons nos valeurs.

$$\chi^2_{\text{obs}} = 12,212 > \chi^2_{\text{seuil}} = 5,991$$

La statistique calculée est très nettement supérieure à la valeur limite de la table. **On rejette l'hypothèse nulle** H_0 et on accepte l'hypothèse alternative H_1.

Il existe un lien hautement significatif entre le type de fertilisation utilisé et la qualité marchande des fèves de cacao récoltées dans la région de Soubré. Le mode de fertilisation influence directement la qualité.

En comparant les effectifs observés aux effectifs théoriques, on constate que :

- L'absence totale d'engrais pénalise fortement la qualité (seulement 65 parcelles de Qualité Supérieure observées contre 81 attendues).
- L'utilisation d'Engrais Organique / Compost est le traitement qui favorise le plus la *Qualité Supérieure* (95 observés contre 81 attendues), surpassant légèrement l'engrais chimique dans cette catégorie de valeur.

Pour la coopérative ivoirienne, ces résultats justifient le déploiement d'un programme d'accompagnement technique incitant à la production et à l'application de compost organique.

Correction de l'exercice 4 ▲

1. **Formulation des hypothèses :**

   - **Hypothèse nulle (H_0) :** La méthode de protection des cultures et le circuit de commercialisation des producteurs de mangues sont **indépendants**. Le choix du mode de protection n'est pas lié à la filière de vente.
   - **Hypothèse alternative (H_1) :** La méthode de protection et le circuit de commercialisation ne sont **pas indépendants** (ils sont dépendants). Il existe une association significative entre les pratiques sanitaires des agriculteurs et leurs débouchés commerciaux.

2. **Calcul des effectifs théoriques, noté $n^\star_{ij}$:** La formule pour calculer l'effectif théorique d'une case (attendu sous l'hypothèse d'indépendance) est :

$$n^\star_{ij} = \frac{\text{Total Ligne}_i \times \text{Total Colonne}_j}{\text{Total Global}} = \frac{n_{i\cdot} \times n_{\cdot j}}{N}$$

   Calculons ces valeurs pour les 4 cases intérieures du tableau (Total global $N = 400$) :

   - **Ligne « Lutte chimique » (Total = 250) :**
     - Marché local : $n^\star_{11} = \dfrac{250 \times 260}{400} = \dfrac{65000}{400} = 162,5$
     - Exportation / Supermarchés : $n^\star_{12} = \dfrac{250 \times 140}{400} = \dfrac{35000}{400} = 87,5$
   - **Ligne « Lutte agroécologique » (Total = 150) :**
     - Marché local : $n^\star_{21} = \dfrac{150 \times 260}{400} = \dfrac{39000}{400} = 97,5$
     - Exportation / Supermarchés : $n^\star_{22} = \dfrac{150 \times 140}{400} = \dfrac{21000}{400} = 52,5$

   Condition de validité : Tous les effectifs théoriques calculés sont strictement supérieurs à 5. Les conditions d'application du test du χ^2 sont parfaitement remplies.

3. **Calcul de la statistique du χ^2 observée, noté χ^2_{obs} :** On applique la formule $\chi^2_{\text{obs}} = \sum \frac{(n_{ij}-n^\star_{ij})^2}{n^\star_{ij}}$ pour chacune des 4 cases :

   - **Lutte chimique :**
     - Marché local : $\dfrac{(190 - 162,5)^2}{162,5} = \dfrac{27,5^2}{162,5} = \dfrac{756,25}{162,5} \approx 4,6538$
     - Exportation / Supermarchés : $\dfrac{(60 - 87,5)^2}{87,5} = \dfrac{(-27,5)^2}{87,5} = \dfrac{756,25}{87,5} \approx 8,6429$

- **Lutte agroécologique :**
  - Marché local : $\dfrac{(70-97,5)^2}{97,5} = \dfrac{(-27,5)^2}{97,5} = \dfrac{756,25}{97,5} \approx 7,7564$
  - Exportation / Supermarchés : $\dfrac{(80-52,5)^2}{52,5} = \dfrac{27,5^2}{52,5} = \dfrac{756,25}{52,5} \approx 14,4048$

En additionnant ces 4 contributions, on trouve la valeur globale de la statistique :

$$\chi^2_{\text{obs}} = 4,6538 + 8,6429 + 7,7564 + 14,4048 = 35,4579$$

4. **Détermination de la valeur seuil, noté χ^2_{seuil} :**
Le tableau possède $L = 2$ lignes et $C = 2$ colonnes (en excluant les marges de totaux). Le nombre de degrés de liberté (ddl) se détermine ainsi :

$$\text{ddl} = (L-1) \times (C-1) = (2-1) \times (2-1) = 1 \times 1 = 1$$

En consultant la table de la loi du χ^2 pour un seul degré de liberté (ddl $= 1$) au seuil de risque choisi de $\alpha = 5\%$, on trouve la valeur critique :

$$\chi^2_{\text{seuil}} = 3,841$$

5. **Règle de décision et conclusion sociologique et agronomique :**
Si $\chi^2_{\text{obs}} > \chi^2_{\text{seuil}}$, **alors** on rejette l'hypothèse nulle H_0. **Sinon** si $\chi^2_{\text{obs}} \leq \chi^2_{\text{seuil}}$, on accepte H_0.
Nous comparons nos valeurs.

$$\chi^2_{\text{obs}} = 35,458 > \chi^2_{\text{seuil}} = 3,841$$

La statistique observée est immensément supérieure à la valeur critique. **On rejette fermement l'hypothèse nulle** H_0 et on valide l'hypothèse alternative H_1.

Il existe une association hautement significative entre la méthode de protection des vergers de mangues et le circuit de commercialisation choisi par les agriculteurs dans le Delta du Mékong.

L'analyse des écarts entre les valeurs observées et théoriques met en évidence deux comportements marquants :

- Les producteurs insérés dans les filières d'Exportation / Supermarchés recourent beaucoup plus que prévu à la *lutte agroécologique* (80 observés contre 52,5 attendus).
- À l'inverse, les producteurs vendant sur le **marché local** restent fortement dépendants de la *lutte chimique* (190 observés contre 162,5 attendus).

Cette étude démontre que les structures de marché jouent un rôle de levier majeur dans la transition agroécologique au Vietnam.

Correction de l'exercice 5 ▲

1. **Formulation des hypothèses :**
   - **Hypothèse nulle (H_0) :** Les habitudes de consommation en énergie de cuisson sont **homogènes** (identiques) dans les trois régions étudiées (Mopti, Sikasso et Bamako). La répartition des sources d'énergie ne dépend pas de la zone géographique.
   - **Hypothèse alternative (H_1) :** Les structures de consommation ne sont **pas homogènes** (elles sont hétérogènes). Il existe des différences significatives de choix énergétiques entre ces trois zones économiques du Mali.

2. **Calcul des effectifs théoriques, noté $n^\star_{ij}$) :** Puisque les trois échantillons régionaux ont exactement la même taille (Total Colonne $= 200$), les effectifs théoriques sous l'hypothèse d'homogénéité seront identiques pour chaque région pour une même ligne. La formule appliquée est :

$$n^\star_{ij} = \frac{\text{Total Ligne}_i \times \text{Total Colonne}_j}{\text{Total Global}} = \frac{n_{i.} \times n_{.j}}{N}$$

Calculons les valeurs attendues pour les 9 cases intérieures du tableau :

- **Ligne « Bois de feu » (Total = 280) :**

$$n^\star_{11} = n^\star_{12} = n^\star_{13} = \frac{280}{3} \approx 93,33$$

- **Ligne « Charbon de bois » (Total = 230) :**

$$n^\star_{21} = n^\star_{22} = n^\star_{23} = \frac{230}{3} \approx 76,67$$

- **Ligne « Gaz » (Total = 90) :**

$$n^\star_{31} = n^\star_{32} = n^\star_{33} = \frac{90}{3} = 30,00$$

Condition de validité : Tous les effectifs théoriques calculés sont strictement supérieurs à 5. Les conditions d'application du test du χ^2 d'homogénéité sont validées.

3. **Calcul de la statistique du χ^2 observée, noté χ^2_{obs} :** On applique la formule

$$\chi^2_{\text{obs}} = \sum_{ij} \frac{(n_{ij} - n^\star_{ij})^2}{n^\star_{ij}}$$

pour les 9 cases :

- **Région de Mopti :**
  - Bois de feu : $\dfrac{(140-93,33)^2}{93,33} = \dfrac{46,67^2}{93,33} \approx 23,3346$
  - Charbon : $\dfrac{(50-76,67)^2}{76,67} = \dfrac{(-26,67)^2}{76,67} \approx 9,2775$
  - Gaz : $\dfrac{(10-30,00)^2}{30,00} = \dfrac{(-20)^2}{30} = \dfrac{400}{30} \approx 13,3333$
- **Région de Sikasso :**
  - Bois de feu : $\dfrac{(110-93,33)^2}{93,33} = \dfrac{16,67^2}{93,33} \approx 2,9775$
  - Charbon : $\dfrac{(70-76,67)^2}{76,67} = \dfrac{(-6,67)^2}{76,67} \approx 0,5802$
  - Gaz : $\dfrac{(20-30,00)^2}{30,00} = \dfrac{(-10)^2}{30} = \dfrac{100}{30} \approx 3,3333$
- **Zone de Bamako :**
  - Bois de feu : $\dfrac{(30-93,33)^2}{93,33} = \dfrac{(-63,33)^2}{93,33} \approx 42,9719$
  - Charbon : $\dfrac{(110-76,67)^2}{76,67} = \dfrac{33,33^2}{76,67} \approx 14,4910$
  - Gaz : $\dfrac{(60-30,00)^2}{30,00} = \dfrac{30^2}{30} = \dfrac{900}{30} = 30$

En additionnant ces 9 contributions, on obtient la valeur globale :

$\chi^2_{\text{obs}} = 23,3346 + 9,2775 + 13,3333 + 2,9775 + 0,5802 + 3,3333 + 42,9719 + 14,4910 + 30,0000 = 140,2993$

4. **Détermination de la valeur seuil, noté χ^2_{seuil} :**

5. Le tableau intérieur possède $L = 3$ lignes (sources d'énergie) et $C = 3$ colonnes (régions économiques). Le nombre de degrés de liberté (ddl) se détermine ainsi :

$$\text{ddl} = (L-1) \times (C-1) = (3-1) \times (3-1) = 2 \times 2 = 4$$

En lisant dans la table de distribution de la loi du χ^2 pour 4 degrés de liberté au seuil de risque de $\alpha = 5\%$, on trouve la valeur critique :

$$\chi^2_{\text{seuil}} = 9,488$$

6. **Règle de décision et conclusion environnementale :**
Si $\chi^2_{\text{obs}} > \chi^2_{\text{seuil}}$, **alors** on rejette l'hypothèse nulle H_0. **Sinon** si $\chi^2_{\text{obs}} \leq \chi^2_{\text{seuil}}$, on accepte H_0.
Nous comparons nos valeurs.
$$\chi^2_{\text{obs}} = 140,30 > \chi^2_{\text{seuil}} = 9,488$$
La statistique observée est extraordinairement supérieure à la valeur critique. **On rejette fermement l'hypothèse nulle** H_0 au profit de l'hypothèse alternative H_1.

La structure de consommation en énergie de cuisson n'est absolument pas homogène entre les trois régions du Mali.

Pour l'ONG, cette étude démontre qu'une campagne de sensibilisation uniforme à l'échelle nationale serait inefficace. Les stratégies doivent être différenciées.

Correction de l'exercice 6 ▲

1. **Formulation des hypothèses :**
   - **Hypothèse nulle (H_0) :** La structure d'équipement en systèmes d'irrigation est **homogène** (identique) au sein des trois communautés de la région de Cusco. Le choix ou l'accès à la technologie ne dépend pas de la vallée d'implantation.
   - **Hypothèse alternative (H_1) :** Les structures d'équipement ne sont **pas homogènes** (elles sont hétérogènes). Il existe des différences statistiquement significatives dans l'adoption des technologies d'irrigation entre les trois communautés.

2. **Calcul des effectifs théoriques, noté $n^\star_{ij}$:** Étant donné que les trois échantillons communautaires disposent rigoureusement du même effectif (Total Colonne = 150), les effectifs théoriques attendus sous H_0 seront identiques pour les trois communautés sur une même ligne. La formule générale appliquée est :
$$n^\star_{ij} = \frac{\text{Total Ligne}_i \times \text{Total Colonne}_j}{\text{Total Global}} = \frac{n_{i\cdot} \times n_{\cdot j}}{N}$$
Calculons les valeurs théoriques pour les 9 cases du tableau :
   - **Ligne « Gravité » (Total = 195) :**
$$n^\star_{11} = n^\star_{12} = n^\star_{13} = \frac{195}{3} = 65$$
   - **Ligne « Aspersion » (Total = 165) :**
$$n^\star_{21} = n^\star_{22} = n^\star_{23} = \frac{165}{3} = 55$$
   - **Ligne « Goutte-à-goutte » (Total = 90) :**
$$n^\star_{31} = n^\star_{32} = n^\star_{33} = \frac{90}{3} = 30$$

   Condition de validité : Tous les effectifs théoriques calculés sont strictement supérieurs à 5. Le test du χ^2 est parfaitement valide sans nécessiter de regroupement de classes.

3. **Calcul de la statistique du χ^2 observée, notée χ^2_{obs} :** On utilise la formule de calcul des écarts :
$$\chi^2_{\text{obs}} = \sum_{ij} \frac{(n_{ij} - n^\star_{ij})^2}{n^\star_{ij}}$$
pour chaque cellule :

- **Communauté A (Haute vallée) :**
  - Gravité : $\frac{(90-65)^2}{65} = \frac{25^2}{65} = \frac{625}{65} \approx 9,6154$
  - Aspersion : $\frac{(45-55)^2}{55} = \frac{(-10)^2}{55} = \frac{100}{55} \approx 1,8182$
  - Goutte-à-goutte : $\frac{(15-30)^2}{30} = \frac{(-15)^2}{30} = \frac{225}{30} = 7,5000$
- **Communauté B (Moyenne vallée) :**
  - Gravité : $\frac{(65-65)^2}{65} = \frac{0^2}{65} = 0$
  - Aspersion : $\frac{(55-55)^2}{55} = \frac{0^2}{55} = 0$
  - Goutte-à-goutte : $\frac{(30-30)^2}{30} = \frac{0^2}{30} = 0$
- **Communauté C (Basse vallée) :**
  - Gravité : $\frac{(40-65)^2}{65} = \frac{(-25)^2}{65} = \frac{625}{65} \approx 9,6154$
  - Aspersion : $\frac{(65-55)^2}{55} = \frac{10^2}{55} = \frac{100}{55} \approx 1,8182$
  - Goutte-à-goutte : $\frac{(45-30)^2}{30} = \frac{15^2}{30} = \frac{225}{30} = 7,5$

En faisant la somme cumulée de ces 9 contributions :

$$\chi^2_{\text{obs}} = 9,6154 + 1,8182 + 7,5000 + 0 + 0 + 0 + 9,6154 + 1,8182 + 7,5000 = 37,8672$$

4. **Détermination de la valeur seuil, notée χ^2_{seuil} :** Le tableau intérieur compte $L = 3$ lignes (types d'irrigation) et $C = 3$ colonnes (communautés indépendantes). Le nombre de degrés de liberté (ddl) se définit par :

$$\text{ddl} = (L-1) \times (C-1) = (3-1) \times (3-1) = 2 \times 2 = 4$$

En lisant dans la table de distribution de la loi du χ^2 pour 4 degrés de liberté au seuil de risque choisi $\alpha = 5\%$, on extrait :

$$\chi^2_{\text{seuil}} = 9,488$$

5. **Règle de décision et conclusion agronomique :**

Si $\chi^2_{\text{obs}} > \chi^2_{\text{seuil}}$, **alors** on rejette l'hypothèse nulle H_0. **Sinon** si $\chi^2_{\text{obs}} \leq \chi^2_{\text{seuil}}$, on accepte H_0.
Nous comparons nos valeurs.

$$\chi^2_{\text{obs}} = 37,87 > \chi^2_{\text{seuil}} = 9,488$$

La statistique calculée sur le terrain est considérablement supérieure à la valeur critique. **On rejette fermement l'hypothèse nulle** H_0 au profit de l'hypothèse alternative H_1.

L'équipement en technologies de modernisation de l'eau n'est absolument pas homogène entre les vallées de Cusco. L'analyse fine des écarts met en relief un gradient topographique très marqué dans l'adoption de l'irrigation.

Correction de l'exercice 7 ▲

1. **Formulation des hypothèses :**

- **Hypothèse nulle (H_0) :** $\mu_A = \mu_B = \mu_C$. Les rendements moyens de tomates sont identiques pour les trois modalités de couverture du sol. Le type de paillage n'a pas d'influence significative sur la production.
- **Hypothèse alternative (H_1) :** Au moins un des types de paillage donne un rendement moyen statistiquement différent des autres.

2. **Calcul des moyennes :** Chaque groupe comporte $n = 4$ parcelles, pour un total de $N = 12$ observations.
   - **Groupe A (Paille de riz) :** $\overline{X}_A = \frac{12+10+13+11}{4} = \frac{46}{4} = 11{,}5$ kg
   - **Groupe B (Roseaux broyés) :** $\overline{X}_B = \frac{14+17+13+16}{4} = \frac{60}{4} = 15{,}0$ kg
   - **Groupe C (Témoin sol nu) :** $\overline{X}_C = \frac{9+7+6+10}{4} = \frac{32}{4} = 8{,}0$ kg
   - **Moyenne générale ($\overline{X}$) :**
   $$\overline{X} = \frac{11{,}5+15{,}0+8{,}0}{3} = \frac{34{,}5}{3} = 11{,}5 \text{ kg}$$
3. **Calcul de la statistique F observée (F_{obs}) :** Pour obtenir F_{obs}, il est nécessaire de calculer la Somme des Carrés Inter-groupes (SCE_{Inter}) et la Somme des Carrés Intra-groupes (SCE_{Intra}).
   - **Somme des Carrés Inter-groupes (SCE_{Inter}) :**
   $$SCE_{\text{Inter}} = n \times \sum_i (\overline{X}_i - \overline{X})^2$$
   $$SCE_{\text{Inter}} = 4 \times \left[(11{,}5-11{,}5)^2 + (15{,}0-11{,}5)^2 + (8{,}0-11{,}5)^2\right]$$
   $$SCE_{\text{Inter}} = 4 \times \left[0^2 + (3{,}5)^2 + (-3{,}5)^2\right] = 4 \times [0 + 12{,}25 + 12{,}25]$$
   $$SCE_{\text{Inter}} = 4 \times 24{,}5 = \mathbf{98{,}0}$$
   - **Somme des Carrés Intra-groupes (SCE_{Intra}) :** On calcule la somme des carrés des écarts à la moyenne au sein de chaque groupe :
     - Groupe A : $(12-11{,}5)^2 + (10-11{,}5)^2 + (13-11{,}5)^2 + (11-11{,}5)^2 = 0{,}25 + 2{,}25 + 2{,}25 + 0{,}25 = 5{,}0$
     - Groupe B : $(14-15)^2 + (17-15)^2 + (13-15)^2 + (16-15)^2 = 1+4+4+1 = 10{,}0$
     - Groupe C : $(9-8)^2 + (7-8)^2 + (6-8)^2 + (10-8)^2 = 1+1+4+4 = 10{,}0$
   $$SCE_{\text{Intra}} = 5{,}0 + 10{,}0 + 10{,}0 = \mathbf{25{,}0}$$
   (Note : La Somme des Carrés Totale vaut ainsi $SCE_{Totale} = 98 + 25 = 123$).
   - **Degrés de liberté (ddl) et Carrés Moyens (CM) :** Soit $k = 3$ groupes et $N = 12$ observations.
     - $\text{ddl}_{\text{Inter}} = k - 1 = 3 - 1 = 2$
     - $\text{ddl}_{\text{Intra}} = N - k = 12 - 3 = 9$
     - $CM_{\text{Inter}} = \frac{SCE_{\text{Inter}}}{\text{ddl}_{\text{Inter}}} = \frac{98{,}0}{2} = 49{,}0$
     - $CM_{\text{Intra}} = \frac{SCE_{\text{Intra}}}{\text{ddl}_{\text{Intra}}} = \frac{25{,}0}{9} \approx 2{,}7778$
   - **Calcul du F observé :**
   $$F_{\text{obs}} = \frac{CM_{\text{Inter}}}{CM_{\text{Intra}}} = \frac{49{,}0}{\frac{25}{9}} = \frac{49 \times 9}{25} = \frac{441}{25} = \mathbf{17{,}64}$$
4. **Détermination de la valeur seuil (F_{seuil}) :** On cherche dans la table de la loi de Fisher-Snedecor la valeur critique pour un risque $\alpha = 5\%$ avec les degrés de liberté $\text{ddl}_1 = 2$ (numérateur) et $\text{ddl}_2 = 9$ (dénominateur). On trouve :
   $$F_{\text{seuil}} = \mathbf{4{,}26}$$
5. **Règle de décision et conclusion agronomique :**

   Si $F_{\text{obs}} > F_{\text{seuil}}$, **alors** on rejette l'hypothèse nulle H_0. **Sinon** si $F_{\text{obs}} \leq F_{\text{seuil}}$, on accepte H_0.

   Nous comparons nos valeurs.
   $$F_{\text{obs}} = 17{,}64 > F_{\text{seuil}} = 4{,}26$$
   La statistique observée est largement supérieure à la valeur limite de la table. **On rejette l'hypothèse nulle** H_0 et on valide l'hypothèse alternative H_1.

   Le type de couverture du sol a un effet hautement significatif sur le rendement moyen des cultures de tomates dans le delta du Nil.

Correction de l'exercice 8 ▲

1. **Formulation des hypothèses :**
   - **Hypothèse nulle (H_0) :** $\mu_A = \mu_B = \mu_C$. Les rendements moyens en feuilles de thé sont identiques pour les trois types de fertilisants. L'engrais n'a pas d'effet sur la croissance foliaire.
   - **Hypothèse alternative (H_1) :** Au moins un des fertilisants donne un rendement moyen statistiquement différent des autres.

2. **Calcul des moyennes :** Chaque groupe comporte un effectif de $n = 4$ parcelles, pour un total de $N = 12$ observations.
   - **Groupe A (Urée) :** $\overline{X}_A = \dfrac{16+14+13+17}{4} = \dfrac{60}{4} = 15,0$ kg
   - **Groupe B (Compost enrichi) :** $\overline{X}_B = \dfrac{12+10+11+11}{4} = \dfrac{44}{4} = 11,0$ kg
   - **Groupe C (Fumier traditionnel) :** $\overline{X}_C = \dfrac{8+11+9+12}{4} = \dfrac{40}{4} = 10,0$ kg
   - **Moyenne générale ($\overline{X}$) :**
   $$\overline{X} = \frac{15,0+11,0+10,0}{3} = \frac{36,5}{3} = 12,0 \text{ kg}$$

3. **Calcul de la statistique F observée (F_{obs}) :** Pour obtenir F_{obs}, il faut décomposer la variabilité en calculant la Somme des Carrés Inter-groupes (SCE_{Inter}) et la Somme des Carrés Intra-groupes (SCE_{Intra}).
   - **Somme des Carrés Inter-groupes (SCE_{Inter}) :** Elle mesure la dispersion entre les moyennes des groupes et la moyenne générale.
   $$SCE_{\text{Inter}} = n \times \sum_i (\overline{X}_i - \overline{X})^2$$
   $$SCE_{\text{Inter}} = 4 \times \left[(15,0-12,0)^2 + (11,0-12,0)^2 + (10,0-12,0)^2\right]$$
   $$SCE_{\text{Inter}} = 4 \times \left[3^2 + (-1)^2 + (-2)^2\right] = 4 \times [9+1+4]$$
   $$SCE_{\text{Inter}} = 4 \times 14 = \mathbf{56,0}$$
   - **Somme des Carrés Intra-groupes (SCE_{Intra}) :** Elle mesure la dispersion résiduelle (les variations au sein de chaque groupe). Calculons la somme des carrés des écarts à la moyenne propre de chaque groupe :
     - Groupe A : $(16-15)^2 + (14-15)^2 + (13-15)^2 + (17-15)^2 = 1+1+4+4 = 10,0$
     - Groupe B : $(12-11)^2 + (10-11)^2 + (11-11)^2 + (11-11)^2 = 1+1+0+0 = 2,0$
     - Groupe C : $(8-10)^2 + (11-10)^2 + (9-10)^2 + (12-10)^2 = 4+1+1+4 = 10,0$
   $$SCE_{\text{Intra}} = 10,0 + 2,0 + 10,0 = \mathbf{22,0}$$
   (Note : La Somme des Carrés Totale vaut donc $SCE_{Totale} = 56 + 22 = 78$).
   - **Degrés de liberté (ddl) et Carrés Moyens (CM) :** Soit $k = 3$ groupes et $N = 12$ observations.
     - $\text{ddl}_{\text{Inter}} = k - 1 = 3 - 1 = 2$
     - $\text{ddl}_{\text{Intra}} = N - k = 12 - 3 = 9$
     - $CM_{\text{Inter}} = \dfrac{SCE_{\text{Inter}}}{\text{ddl}_{\text{Inter}}} = \dfrac{56,0}{2} = \mathbf{28,0}$
     - $CM_{\text{Intra}} = \dfrac{SCE_{\text{Intra}}}{\text{ddl}_{\text{Intra}}} = \dfrac{22,0}{9} \approx \mathbf{2,4444}$
   - **Calcul du F observé :**
   $$F_{\text{obs}} = \frac{CM_{\text{Inter}}}{CM_{\text{Intra}}} = \frac{28,0}{\frac{22}{9}} = \frac{28 \times 9}{22} = \frac{252}{22} \approx \mathbf{11,4545}$$

4. **Détermination de la valeur seuil (F_{seuil}) :**
   On consulte la table de Fisher-Snedecor pour un risque de première espèce $\alpha = 5\%$ avec les degrés de liberté $\text{ddl}_1 = 2$ (numérateur) et $\text{ddl}_2 = 9$ (dénominateur). On obtient la valeur critique :

$$F_{\text{seuil}} = \mathbf{4,26}$$

5. **Règle de décision et conclusion agronomique :**
   Si $F_{\text{obs}} > F_{\text{seuil}}$, **alors** on rejette l'hypothèse nulle H_0. **Sinon** si $F_{\text{obs}} \leq F_{\text{seuil}}$, on accepte H_0.
   Nous comparons nos valeurs.

$$F_{\text{obs}} = 11,45 > F_{\text{seuil}} = 4,26$$

   La statistique calculée tombe très nettement dans la zone de rejet. **On rejette l'hypothèse nulle** H_0 et on valide l'hypothèse alternative H_1.

   Le type de fertilisant azoté a une influence hautement significative sur le rendement en feuilles de thé.

Correction de l'exercice 9 ▲

1. **Formulation des hypothèses :**
   - **Hypothèse nulle (H_0) :** $\mu_A = \mu_B = \mu_C$. Les rendements moyens de café Robusta sont identiques pour les trois niveaux d'ombrage. La densité d'arbres d'ombrage n'a pas d'influence significative sur la production.
   - **Hypothèse alternative (H_1) :** Au moins un des niveaux d'ombrage donne un rendement moyen statistiquement différent des autres.

2. **Calcul des moyennes :** Chaque groupe comporte $n = 4$ parcelles, pour un total de $N = 12$ observations.
   - **Groupe A (Ombrage Fort) :** $\overline{X}_A = \dfrac{10+8+11+7}{4} = \dfrac{36}{4} = \mathbf{9,0}$ kg
   - **Groupe B (Ombrage Modéré) :** $\overline{X}_B = \dfrac{15+18+14+17}{4} = \dfrac{64}{4} = \mathbf{16,0}$ kg
   - **Groupe C (Plein soleil) :** $\overline{X}_C = \dfrac{11+9+8+12}{4} = \dfrac{40}{4} = \mathbf{10,0}$ kg
   - **Moyenne générale ($\overline{X}$) :**

$$\overline{X} = \frac{9,0+16,0+10,0}{3} = \frac{35}{3} \approx \mathbf{11,67} \text{ kg}$$

3. **Calcul de la statistique F observée (F_{obs}) :** Décomposons la variabilité totale en calculant la Somme des Carrés Inter-groupes (SCE_{Inter}) et la Somme des Carrés Intra-groupes (SCE_{Intra}).
   - **Somme des Carrés Inter-groupes (SCE_{Inter}) :** Elle mesure l'effet du traitement (le niveau d'ombrage).

$$SCE_{\text{Inter}} = n \times \sum_i (\overline{X}_i - \overline{X})^2$$

$$SCE_{\text{Inter}} = 4 \times \left[\left(9 - \frac{35}{3}\right)^2 + \left(16 - \frac{35}{3}\right)^2 + \left(10 - \frac{35}{3}\right)^2\right]$$

$$SCE_{\text{Inter}} = 4 \times \left[\left(-\frac{8}{3}\right)^2 + \left(\frac{13}{3}\right)^2 + \left(-\frac{5}{3}\right)^2\right] = 4 \times \left[\frac{64}{9} + \frac{169}{9} + \frac{25}{9}\right]$$

$$SCE_{\text{Inter}} = 4 \times \frac{258}{9} = \frac{1032}{9} = \mathbf{114,67}$$

   - **Somme des Carrés Intra-groupes (SCE_{Intra}) :** Elle mesure la variabilité résiduelle (les fluctuations naturelles au sein des parcelles d'un même groupe) :

- Groupe A : $(10-9)^2+(8-9)^2+(11-9)^2+(7-9)^2=1+1+4+4=10,0$
- Groupe B : $(15-16)^2+(18-16)^2+(14-16)^2+(17-16)^2=1+4+4+1=10,0$
- Groupe C : $(11-10)^2+(9-10)^2+(8-10)^2+(12-10)^2=1+1+4+4=10,0$

$$SCE_{\text{Intra}} = 10,0+10,0+10,0 = \mathbf{30,0}$$

- **Degrés de liberté (ddl) et Carrés Moyens (CM) :** Avec $k=3$ groupes et $N=12$ observations :
  - $\text{ddl}_{\text{Inter}} = k-1 = 3-1 = 2$
  - $\text{ddl}_{\text{Intra}} = N-k = 12-3 = 9$
  - $CM_{\text{Inter}} = \dfrac{SCE_{\text{Inter}}}{\text{ddl}_{\text{Inter}}} = \dfrac{114,6667}{2} = \mathbf{57,3333}$
  - $CM_{\text{Intra}} = \dfrac{SCE_{\text{Intra}}}{\text{ddl}_{\text{Intra}}} = \dfrac{30,0}{9} \approx \mathbf{3,3333}$
- **Calcul du F observé :**

$$F_{\text{obs}} = \frac{CM_{\text{Inter}}}{CM_{\text{Intra}}} = \frac{\frac{1032}{18}}{\frac{30}{9}} = \frac{1032\times 9}{18\times 30} = \frac{1032}{60} = \mathbf{17,20}$$

4. **Détermination de la valeur seuil (F_{seuil}) :** En lisant la table statistique de Snedecor au seuil $\alpha = 5\%$ pour $\text{ddl}_1 = 2$ (numérateur) et $\text{ddl}_2 = 9$ (dénominateur), on relève la valeur critique suivante :

$$F_{\text{seuil}} = \mathbf{4,26}$$

5. **Règle de décision et conclusion agronomique :**

   Si $F_{\text{obs}} > F_{\text{seuil}}$, **alors** on rejette l'hypothèse nulle H_0. **Sinon** si $F_{\text{obs}} \leq F_{\text{seuil}}$, on accepte H_0.

   Nous comparons nos valeurs.

$$F_{\text{obs}} = 17,20 > F_{\text{seuil}} = 4,26$$

   La statistique calculée sur l'échantillon de Sumatra est largement supérieure à la valeur critique. **On rejette l'hypothèse nulle** H_0 et on adopte l'hypothèse alternative H_1.

   Le niveau d'ombrage a une influence statistique hautement significative sur le rendement moyen du café Robusta.

Correction de l'exercice 10 ▲

1. **Formulation des hypothèses :**
   - **Hypothèse nulle (H_0) :** $\mu_A = \mu_B = \mu_C$. Les rendements moyens de cacao sont identiques pour les trois rythmes de taille. La fréquence de taille n'a pas d'influence significative sur la production.
   - **Hypothèse alternative (H_1) :** Au moins un des rythmes de taille donne un rendement moyen statistiquement différent des autres.
2. **Calcul des moyennes :** Chaque groupe comporte un effectif de $n=4$ parcelles, pour un total de $N=12$ observations.
   - **Groupe A (Taille Mensuelle) :** $\overline{X}_A = \frac{11+9+12+8}{4} = \frac{40}{4} = \mathbf{10,0}$ kg
   - **Groupe B (Taille Trimestrielle) :** $\overline{X}_B = \frac{15+13+14+14}{4} = \frac{56}{4} = \mathbf{14,0}$ kg
   - **Groupe C (Témoin sans taille) :** $\overline{X}_C = \frac{8+11+9+8}{4} = \frac{36}{4} = \mathbf{9,0}$ kg
   - **Moyenne générale ($\overline{X}$) :**

$$\overline{X} = \frac{10,0+14,0+9,0}{3} = \frac{33}{3} = \mathbf{11,0} \text{ kg}$$

3. **Calcul de la statistique F observée (F_{obs}) :** Pour quantifier la variance, nous devons calculer la Somme des Carrés Inter-groupes (SCE_{Inter}) et la Somme des Carrés Intra-groupes (SCE_{Intra}).

- **Somme des Carrés Inter-groupes (SCE_{Inter}) :** Elle mesure la variabilité expliquée par le facteur (le rythme de taille).

$$SCE_{\text{Inter}} = n \times \sum_i (\overline{X}_i - \overline{X}_G)^2$$

$$SCE_{\text{Inter}} = 4 \times \left[(10,0 - 11,0)^2 + (14,0 - 11,0)^2 + (9,0 - 11,0)^2\right]$$

$$SCE_{\text{Inter}} = 4 \times \left[(-1)^2 + 3^2 + (-2)^2\right] = 4 \times [1 + 9 + 4]$$

$$SCE_{\text{Inter}} = 4 \times 14 = \mathbf{56,0}$$

- **Somme des Carrés Intra-groupes (SCE_{Intra}) :** Elle mesure la variabilité résiduelle ou l'erreur (les fluctuations au sein de chaque traitement). On calcule la somme des carrés des écarts à la moyenne de chaque groupe :
  - Groupe A : $(11-10)^2 + (9-10)^2 + (12-10)^2 + (8-10)^2 = 1+1+4+4 = 10,0$
  - Groupe B : $(15-14)^2 + (13-14)^2 + (14-14)^2 + (14-14)^2 = 1+1+0+0 = 2,0$
  - Groupe C : $(8-9)^2 + (11-9)^2 + (9-9)^2 + (8-9)^2 = 1+4+0+1 = 6,0$

$$SCE_{\text{Intra}} = 10,0 + 2,0 + 6,0 = \mathbf{18,0}$$

(Note : La Somme des Carrés Totale vaut ainsi $SCE_{Totale} = 56 + 18 = 74$).

- **Degrés de liberté (ddl) et Carrés Moyens (CM) :** Soit $k = 3$ groupes et $N = 12$ observations.
  - $\text{ddl}_{\text{Inter}} = k - 1 = 3 - 1 = 2$
  - $\text{ddl}_{\text{Intra}} = N - k = 12 - 3 = 9$
  - $CM_{\text{Inter}} = \dfrac{SCE_{\text{Inter}}}{\text{ddl}_{\text{Inter}}} = \dfrac{56,0}{2} = \mathbf{28,0}$
  - $CM_{\text{Intra}} = \dfrac{SCE_{\text{Intra}}}{\text{ddl}_{\text{Intra}}} = \dfrac{18,0}{9} = \mathbf{2,0}$
- **Calcul du F observé :**

$$F_{\text{obs}} = \frac{CM_{\text{Inter}}}{CM_{\text{Intra}}} = \frac{28,0}{2,0} = \mathbf{14,0}$$

4. **Détermination de la valeur seuil (F_{seuil}) :**

   On consulte la table statistique de la loi de Fisher-Snedecor pour un risque de première espèce $\alpha = 5\%$ avec les degrés de liberté $\text{ddl}_1 = 2$ (numérateur) et $\text{ddl}_2 = 9$ (dénominateur). On y lit la valeur critique suivante :

$$F_{\text{seuil}} = 4,26$$

5. **Règle de décision et conclusion agronomique :**

   Si $F_{\text{obs}} > F_{\text{seuil}}$, **alors** on rejette l'hypothèse nulle H_0. **Sinon** si $F_{\text{obs}} \leq F_{\text{seuil}}$, on accepte H_0.

   Nous comparons nos valeurs.

$$F_{\text{obs}} = 14,00 > F_{\text{seuil}} = 4,26$$

   La statistique calculée sur le terrain est largement supérieure à la valeur limite de la table. **On rejette l'hypothèse nulle** H_0 et on accepte l'hypothèse alternative H_1.

   Le rythme de taille des branches a un impact hautement significatif sur le rendement moyen en cacao dans cette région du Guatemala.