Ecole supérieure d'agro-développement international
istom

Examen de statistiques

Mercredi 3 juin 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 du rendement du manioc en Côte d'Ivoire

Le manioc est une culture vivrière majeure pour la sécurité alimentaire et l'économie locale en Côte d'Ivoire. Pour aider les petits producteurs à augmenter leur productivité, une coopérative agricole teste l'effet de trois pratiques de gestion du sol sur le rendement final (exprimé en *tonnes par hectare, t/ha*).

L'ingénieur agronome répartit équitablement un essai sur $N = 15$ parcelles homogènes, soit 5 parcelles par méthode testée :

- *Méthode A :* Pratique traditionnelle (témoin, sans apport).
- *Méthode B :* Paillage végétal (utilisation des résidus de culture pour maintenir l'humidité).
- *Méthode C :* Apport de compost local (fertilisant organique).

À la récolte, les rendements obtenus (en t/ha) sur chaque parcelle sont renseignés ci-dessous :

- *Méthode A :* 10 ; 12 ; 9 ; 11 ; 13
- *Méthode B :* 14 ; 15 ; 13 ; 17 ; 16
- *Méthode C :* 18 ; 19 ; 21 ; 17 ; 20

On souhaite déterminer, au seuil de risque $\alpha = 5\,\%$, si le choix de la méthode de gestion du sol a un impact significatif sur le rendement moyen du manioc.

1. Formuler l'hypothèse nulle (H_0) et l'hypothèse alternative (H_1) de ce test de comparaison de moyennes (ANOVA).
2. Calculer les caractéristiques de tendance centrale de l'essai :
   (a) Déterminer le rendement moyen pour chaque groupe ($\overline{Y}_A$, $\overline{Y}_B$, $\overline{Y}_C$).
   (b) Déterminer le rendement moyen global $\overline{Y}$ de l'ensemble de l'essai.
3. Calculer la variabilité de l'essai :

(a) Calculer la Somme des Carrés Écarts Inter-groupes (SCE_{inter}).
(b) Calculer la Somme des Carrés Écarts Intra-groupes (SCE_{intra}).

4. Déterminer les degrés de liberté (ddl) associés à la variabilité Inter-groupes (ddl_{inter}) et à la variabilité Intra-groupes (ddl_{intra}).

5. Déduire des questions précédentes les carrés moyens (CM_{inter} et CM_{intra}), puis calculer la valeur de la statistique de test observée F_{obs}.

6. Sachant que la valeur seuil lue dans la table de la loi de Fisher au risque $\alpha = 5\,\%$ est $F_{\text{seuil}} = 3,89$, comparer F_{obs} et F_{seuil}, puis interpréter le résultat d'un point de vue agronomique pour les producteurs de la coopérative.

Correction ▼ [stat-0057]

Exercice 2 Adoption d'une technique de stockage du niébé au Niger

Au Niger, le niébé (une variété de haricot sec) est une culture cruciale pour la sécurité alimentaire. Cependant, les pertes post-récolte dues aux insectes ravageurs sont massives. Pour y faire face, une ONG a introduit une nouvelle technique de stockage hermétique (les sacs PICS).

Un agronome souhaite savoir si le taux d'adoption de cette nouvelle technologie est *homogène* entre deux types d'exploitations : les exploitations gérées par des *groupements de femmes* et celles gérées par des *hommes*. Il sélectionne de manière indépendante deux échantillons de producteurs et enregistre s'ils ont adopté ou non les sacs hermétiques. Voici les données collectées (effectifs observés) :

- Sur un échantillon de *100 exploitations gérées par des femmes*, 40 ont adopté la technique et 60 ne l'ont pas adoptée.
- Sur un échantillon de *100 exploitations gérées par des hommes*, 60 ont adopté la technique et 40 ne l'ont pas adoptée.

On souhaite tester, au seuil de risque $\alpha = 5\,\%$, si le comportement vis-à-vis de l'adoption de cette technologie est le même pour les deux groupes d'exploitants.

1. Construire le tableau de contingence à double entrée regroupant les effectifs observés (n_{ij}), en faisant apparaître les totaux en ligne et en colonne.
2. Formuler l'hypothèse nulle (H_0) et l'hypothèse alternative (H_1) de ce test du χ^2 d'homogénéité.
3. Calculer le tableau des effectifs théoriques ($n^\star_{ij}$) attendus sous l'hypothèse nulle.
4. Calculer la statistique globale χ^2_{obs}.
5. Déterminer le nombre de degrés de liberté (ddl) associé à ce test.
6. Sachant que la valeur seuil lue dans la table de la loi du χ^2 au seuil $\alpha = 5\,\%$ pour ce ddl est $\chi^2_{\text{seuil}} = 3,84$, comparer χ^2_{obs} et χ^2_{seuil}, et interpréter le résultat pour l'ONG.

Correction ▼ [stat-0058]

Exercice 3 Répartition des mangues « Francisque » selon les calibres en Haïti

En Haïti, la mangue « Francisque » est un produit d'exportation phare et une source de revenus vitale pour des milliers de petits producteurs ruraux (notamment dans le département de l'Artibonite). Pour être exportées vers les marchés internationaux, les mangues doivent respecter des normes de calibrage strictes basées sur leur poids.

Une coopérative d'exportation de Saint-Marc utilise un modèle théorique basé sur l'historique des campagnes précédentes. Selon ce modèle, une récolte saine et standard se répartit historiquement selon les proportions suivantes :

- *50 %* de mangues de *Calibre Moyen* (recherché pour les cartons standards),

- *40 %* de mangues de *Calibre Grand* (haute valeur commerciale),
- *10 %* de mangues de *Calibre Petit* (généralement réorientées vers le marché local).

À cause d'une perturbation climatique durant la floraison, l'agronome de la coopérative craint que la taille des fruits ait été modifiée. Il prélève un échantillon aléatoire de $N = 300$ mangues dans le centre de tri et observe les effectifs réels suivants :

- *135* mangues de Calibre Moyen,
- *132* mangues de Calibre Grand,
- *33* mangues de Calibre Petit.

On souhaite tester, au seuil de risque $\alpha = 5\,\%$, si la distribution des calibres observée cette année est toujours conforme au modèle théorique historique.

1. Formuler l'hypothèse nulle (H_0) et l'hypothèse alternative (H_1) de ce test du χ^2 d'ajustement.
2. À partir des proportions du modèle historique, calculer les effectifs théoriques (ou attendus), notés $n_i^\star$, pour chacune des trois catégories sur un échantillon total de 300 mangues.
3. Calculer la statistique χ^2_{obs}.
4. Déterminer le nombre de degrés de liberté (ddl) associé à ce test d'ajustement.
5. Prendre une décision statistique et conclure :
   (a) Sachant que la valeur seuil lue dans la table de la loi du χ^2 au seuil $\alpha = 5\,\%$ pour ce ddl est $\chi^2_{\text{seuil}} = 5{,}99$, comparer χ^2_{obs} et χ^2_{seuil} puis prendre la décision statistique associée.
   (b) Interpréter ce résultat d'un point de vue agronomique et commercial pour la coopérative d'exportation.

Correction ▼ [stat-0059]

Correction de l'exercice 1 ▲

1. Formulation des hypothèses du test

   Le test d'analyse de la variance (ANOVA) à un facteur permet de comparer simultanément les moyennes de plusieurs groupes indépendants afin de mesurer l'effet d'un facteur qualitatif sur une variable quantitative.

   - H_0 *(Hypothèse nulle) :* $\mu_A = \mu_B = \mu_C$. Les rendements moyens théoriques sont identiques pour toutes les méthodes de gestion du sol. Les pratiques testées n'ont aucun impact significatif sur la productivité du manioc.
   - H_1 *(Hypothèse alternative) :* Au moins deux moyennes sont différentes. Le choix de la méthode de gestion du sol a un effet significatif sur le rendement moyen du manioc.

2. Caractéristiques de tendance centrale de l'essai

   Chaque groupe (ou échantillon conditionnel) possède un effectif égal : $n_A = n_B = n_C = 5$. L'effectif total de l'essai est $N = 15$.

   (a) *Rendement moyen pour chaque groupe :*

   - *Méthode A :* $\bar{Y}_A = \dfrac{10+12+9+11+13}{5} = \dfrac{55}{5} = \mathbf{11}$ t/ha
   - *Méthode B :* $\bar{Y}_B = \dfrac{14+15+13+17+16}{5} = \dfrac{75}{5} = \mathbf{15}$ t/ha
   - *Méthode C :* $\bar{Y}_C = \dfrac{18+19+21+17+20}{5} = \dfrac{95}{5} = \mathbf{19}$ t/ha

(b) *Rendement moyen global ($\bar{Y}$) :* Puisque l'essai est parfaitement équilibré ($n_j = 5$), la moyenne globale s'obtient indifféremment par la somme des 15 parcelles ou par la moyenne des moyennes de groupe :

$$\bar{Y} = \frac{\bar{Y}_A + \bar{Y}_B + \bar{Y}_C}{3} = \frac{11+15+19}{3} = \frac{45}{3} = \mathbf{15}\text{ t/ha}$$

3. Calcul des indicateurs de variabilité (SCE)

(a) *Somme des Carrés Écarts Inter-groupes (SCE_{inter}) :* Elle quantifie la variabilité expliquée par le facteur (l'effet des différentes méthodes de gestion).

$$SCE_{\text{inter}} = \sum_{j=1}^{3} n_j(\bar{Y}_j - \bar{Y})^2$$

$$SCE_{\text{inter}} = 5\times(11-15)^2 + 5\times(15-15)^2 + 5\times(19-15)^2$$
$$SCE_{\text{inter}} = 5\times(-4)^2 + 5\times(0)^2 + 5\times(4)^2 = 80+0+80 = \mathbf{160}$$

(b) *Somme des Carrés Écarts Intra-groupes (SCE_{intra}) :* Elle mesure la variabilité résiduelle, c'est-à-dire les fluctuations aléatoires à l'intérieur de chaque traitement (le « bruit »).

$$SCE_{\text{intra}} = \sum_j \sum_i (Y_{ij} - \bar{Y}_j)^2$$

Calculons la somme des carrés des écarts à la moyenne pour chaque groupe :

- Groupe A ($\bar{Y}_A = 11$) : $(10-11)^2 + (12-11)^2 + (9-11)^2 + (11-11)^2 + (13-11)^2 = 1+1+4+0+4 = 10$
- Groupe B ($\bar{Y}_B = 15$) : $(14-15)^2 + (15-15)^2 + (13-15)^2 + (17-15)^2 + (16-15)^2 = 1+0+4+4+1 = 10$
- Groupe C ($\bar{Y}_C = 19$) : $(18-19)^2 + (19-19)^2 + (21-19)^2 + (17-19)^2 + (20-19)^2 = 1+0+4+4+1 = 10$

$$SCE_{\text{intra}} = 10+10+10 = \mathbf{30}$$

4. Détermination des degrés de liberté (ddl)

Soit $k = 3$ le nombre de modalités du facteur et $N = 15$ l'effectif total.

- $ddl_{\text{inter}} = k - 1 = 3 - 1 = \mathbf{2}$
- $ddl_{\text{intra}} = N - k = 15 - 3 = \mathbf{12}$

Vérification : Le ddl total vaut $N - 1 = 14$, et on a bien $ddl_{\text{inter}} + ddl_{\text{intra}} = 2 + 12 = 14$.

5. Calcul des Carrés Moyens et de la statistique de test F_{obs}

Les Carrés Moyens s'obtiennent en rapportant chaque somme des carrés à son degré de liberté associé.

- *Carré Moyen Inter-groupes :* $CM_{\text{inter}} = \dfrac{SCE_{\text{inter}}}{ddl_{\text{inter}}} = \dfrac{160}{2} = \mathbf{80}$
- *Carré Moyen Intra-groupes :* $CM_{\text{intra}} = \dfrac{SCE_{\text{intra}}}{ddl_{\text{intra}}} = \dfrac{30}{12} = \mathbf{2{,}5}$
- *Calcul du F de Fisher observé (F_{obs}) :*

$$F_{\text{obs}} = \frac{CM_{\text{inter}}}{CM_{\text{intra}}} = \frac{80}{2{,}5} = \mathbf{32}$$

Ces résultats peuvent être avantageusement résumés dans le tableau classique de l'ANOVA ci-dessous :

Source de variation	*Somme des Carrés (SCE)*	*Degrés de liberté (ddl)*	*Carrés Moyens (CM)*	F_{obs}
Inter-groupes (Facteur)	160	2	80	*32*
Intra-groupes (Résidu)	30	12	2,5	
Total	*190*	*14*		

6. Règle de décision statistique et conclusion agronomique

   - *Comparaison statistique :* On compare la valeur observée à la valeur critique théorique fournie par l'énoncé :

   $$F_{\text{obs}} = 32 \quad \text{et} \quad F_{\text{seuil}} = 3{,}89$$

   On constate que $F_{\text{obs}} > F_{\text{seuil}}$ de manière extrêmement nette (32 est très largement supérieur à 3,89). La statistique calculée se situe ainsi profondément dans la zone de rejet de l'hypothèse nulle. Par conséquent, on *rejette fermement l'hypothèse nulle* H_0 au profit de l'hypothèse alternative H_1.

   - *Conclusion agronomique :* Au seuil de risque de 5 %, le choix de la méthode de gestion du sol exerce un impact hautement significatif sur le rendement moyen du manioc en Côte d'Ivoire. Les variations observées entre les trois modalités ne peuvent pas être imputées aux fluctuations aléatoires du hasard.

Correction de l'exercice 2 ▲

1. Tableau de contingence des effectifs observés (n_{ij})

   En croisant les deux variables qualitatives (*Type d'exploitant* et *Statut d'adoption*) et en calculant les totaux marginaux (lignes et colonnes) ainsi que l'effectif général N, on obtient le tableau de contingence suivant :

Type d'exploitant	**Adopté**	**Non adopté**	**Total (Lignes)**
Femmes	40	60	$n_{1\bullet} = \mathbf{100}$
Hommes	60	40	$n_{2\bullet} = \mathbf{100}$
Total (Colonnes)	$n_{\bullet 1} = \mathbf{100}$	$n_{\bullet 2} = \mathbf{100}$	$N = \mathbf{200}$

2. Formulation des hypothèses du test

   Le test du χ^2 d'homogénéité permet de vérifier si la distribution d'une variable qualitative est la même au sein de deux ou plusieurs sous-populations représentées par des échantillons indépendants.

   - H_0 **(Hypothèse nulle) :** Les deux échantillons sont homogènes. La proportion d'exploitants ayant adopté les sacs hermétiques est rigoureusement identique chez les femmes et chez les hommes. Le genre du gestionnaire de l'exploitation n'influence pas le comportement d'adoption.
   - H_1 **(Hypothèse alternative) :** Les deux échantillons ne sont pas homogènes. Le taux d'adoption des sacs de stockage dépend de manière significative du type d'exploitant (femmes ou hommes).

3. Calcul du tableau des effectifs théoriques ($n^\star_{ij}$)

   Les effectifs théoriques (ou attendus) correspondent à la distribution que l'on observerait si l'hypothèse nulle H_0 de parfaite homogénéité était vraie. La formule générale pour chaque cellule est :

   $$n^\star_{ij} = \frac{\text{Total Ligne } i \times \text{Total Colonne } j}{\text{Total Général } N}$$

   Puisque toutes les marges en ligne et en colonne valent précisément 100, le calcul est identique pour les quatre cases de notre tableau 2×2 :

   $$n^\star_{11} = n^\star_{12} = n^\star_{21} = n^\star_{22} = \frac{100 \times 100}{200} = \frac{10\,000}{200} = \mathbf{50}$$

   On dresse ainsi le tableau des effectifs théoriques attendus :

Type d'exploitant	**Adopté**	**Non adopté**	**Total**
Femmes	50	50	**100**
Hommes	50	50	**100**
Total	**100**	**100**	**200**

Tournez la page S.V.P.

4. Calcul de la statistique globale χ^2_{obs}

   La statistique du χ^2 synthétise l'écart global entre la réalité observée (n_{ij}) et le modèle d'homogénéité théorique ($n^\star_{ij}$). Elle s'obtient par la formule :

$$\chi^2_{\text{obs}} = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{(n_{ij} - n^\star_{ij})^2}{n^\star_{ij}}$$

   Calculons la contribution de chacune des quatre cases (on observe que l'écart au carré vaut systématiquement $(\pm 10)^2 = 100$) :

   - Case Femmes / Adopté : $\dfrac{(40-50)^2}{50} = \dfrac{100}{50} = \mathbf{2}$
   - Case Femmes / Non adopté : $\dfrac{(60-50)^2}{50} = \dfrac{100}{50} = \mathbf{2}$
   - Case Hommes / Adopté : $\dfrac{(60-50)^2}{50} = \dfrac{100}{50} = \mathbf{2}$
   - Case Hommes / Non adopté : $\dfrac{(40-50)^2}{50} = \dfrac{100}{50} = \mathbf{2}$

   En sommant ces contributions individuelles, on obtient la valeur globale de notre statistique :

$$\chi^2_{\text{obs}} = 2 + 2 + 2 + 2 = \mathbf{8}$$

5. Détermination du nombre de degrés de liberté (ddl)

   Pour un tableau de contingence comportant L lignes et C colonnes, le nombre de degrés de liberté est donné par la relation :

$$ddl = (L-1) \times (C-1)$$

   Dans le cadre de cette étude (tableau de dimension 2×2) :

$$ddl = (2-1) \times (2-1) = 1 \times 1 = \mathbf{1}$$

6. Décision statistique et interprétation pour l'ONG

   (a) **Comparaison et décision :** On confronte la valeur calculée sur notre échantillon à la valeur seuil critique fournie par l'énoncé :

$$\chi^2_{\text{obs}} = 8 \quad \text{et} \quad \chi^2_{\text{seuil}} = 3{,}84$$

   On constate que $\chi^2_{\text{obs}} > \chi^2_{\text{seuil}}$ (8 est strictement supérieur à 3,84). La statistique calculée se situe donc au-delà du seuil critique, dans la zone de rejet. Par conséquent, on **rejette l'hypothèse nulle** H_0 et on accepte l'hypothèse alternative H_1.

   (b) Au seuil de risque de 5 %, l'adoption de la nouvelle technique de stockage hermétique du niébé n'est pas homogène entre les exploitations gérées par des femmes et celles gérées par des hommes au Niger. Les hommes affichent un taux d'adoption de 60 %, significativement plus élevé que celui des groupements de femmes qui stagne à 40 %.

Correction de l'exercice 3 ▲

1. Formulation des hypothèses du test

   Le test du χ^2 d'ajustement permet de vérifier si une distribution de fréquences observée sur le terrain est compatible avec une loi de probabilité ou une distribution théorique connue historiquement.

   - H_0 **(Hypothèse nulle) :** La distribution des calibres de mangues observée cette année s'ajuste parfaitement au modèle historique de la coopérative. Les proportions réelles de la récolte respectent la loi fixe : 50 % de Moyen, 40 % de Grand et 10 % de Petit.

- H_1 **(Hypothèse alternative) :** La distribution observée cette année ne s'ajuste pas au modèle historique. La récolte présente une anomalie significative dans la répartition de ses calibres à cause des perturbations climatiques.

2. Calcul des effectifs théoriques ($n_i^\star$)

   Les effectifs théoriques (ou attendus) correspondent aux effectifs que l'on devrait observer si l'hypothèse nulle H_0 était strictement vraie. On les calcule en appliquant les proportions du modèle historique à l'effectif total de l'échantillon ($N = 300$).

   - **Calibre Moyen ($n_1^\star$) :** $300 \times 50\,\% = 300 \times 0{,}50 = \mathbf{150}$ mangues attendues.
   - **Calibre Grand ($n_2^\star$) :** $300 \times 40\,\% = 300 \times 0{,}40 = \mathbf{120}$ mangues attendues.
   - **Calibre Petit ($n_3^\star$) :** $300 \times 10\,\% = 300 \times 0{,}10 = \mathbf{30}$ mangues attendues.

   Vérification : $150 + 120 + 30 = 300$. On retrouve bien la totalité de notre échantillon initial.

3. Calcul de la statistique globale χ^2_{obs}

   La statistique du χ^2 mesure l'écart global existant entre la réalité du terrain (n_i) et le modèle théorique attendu ($n_i^\star$). Elle s'obtient par la formule :

   $$\chi^2_{\text{obs}} = \sum_{i=1}^{3} \frac{(n_i - n_i^\star)^2}{n_i^\star}$$

   Calculons la contribution individuelle de chaque calibre à la variabilité du test :

   - Calibre Moyen : $\dfrac{(135-150)^2}{150} = \dfrac{(-15)^2}{150} = \dfrac{225}{150} = \mathbf{1{,}5}$
   - Calibre Grand : $\dfrac{(132-120)^2}{120} = \dfrac{(12)^2}{120} = \dfrac{144}{120} = \mathbf{1{,}2}$
   - Calibre Petit : $\dfrac{(33-30)^2}{30} = \dfrac{(3)^2}{30} = \dfrac{9}{30} = \mathbf{0{,}3}$

   En faisant la somme de ces trois contributions, on obtient la valeur globale :

   $$\chi^2_{\text{obs}} = 1{,}5 + 1{,}2 + 0{,}3 = \mathbf{3{,}0}$$

4. Détermination du nombre de degrés de liberté (ddl)

   Dans un test d'ajustement à une loi théorique entièrement spécifiée (où aucun paramètre n'a été estimé à partir des données de l'échantillon), le nombre de degrés de liberté dépend uniquement du nombre de catégories ou modalités ($k = 3$) :

   $$ddl = k - 1 = 3 - 1 = \mathbf{2}$$

5. Règle de décision et conclusion

   (a) **Décision statistique :** On compare la statistique calculée sur notre échantillon à la valeur critique théorique fournie par la table :

   $$\chi^2_{\text{obs}} = 3{,}0 \quad \text{et} \quad \chi^2_{\text{seuil}} = 5{,}99$$

   On constate que $\chi^2_{\text{obs}} \leq \chi^2_{\text{seuil}}$ (3,0 est strictement inférieur à 5,99). La valeur calculée se situe dans la zone d'acceptation. On **ne peut donc pas rejeter l'hypothèse nulle** H_0 (on conserve H_0).

   (b) Au seuil de risque de 5 %, la différence constatée entre la répartition réelle des calibres et la loi historique n'est pas statistiquement significative. Elle s'explique uniquement par les fluctuations mineures du hasard de l'échantillonnage.