Abbes Laghrour University Khanchela*
Faculty of Natural and Life Sciences
Departement of Agronomy
L3 Agro-Food Technology and Quality Control

Corrigé-type

Q1- L'hypothèse de normalité est principalement requise pour **les tests paramétriques**, en particulier lorsque la taille de l'échantillon est faible ; Les tests de comparaison de moyennes (t de Student, ANOVA) ; Tests basés sur des modèles linéaires (regression linéaire) ; Les intervalles de confiance basés sur l'estimation paramétrique (test Z, ...). (1pt)

Les outils pour tester la normalité : Méthodes graphiques (Histogramme de fréquence, Q-Q plot, Boxplot) (1pt); tests statistiques de normalité (test de shapiro-wilk, Kolmogorov-Smirnov (KS), Test du Khi-deux (χ^2), Test d'Anderson-Darling) (0.5pt); Coefficient d'asymétrie et d'aplatissement (0.5pt).

Q2- Bien que la corrélation et la régression linéaire soient souvent utilisées conjointement, elles répondent à des objectifs statistiques distincts et reposent sur des hypothèses différentes.

Objectif principal

La corrélation mesure le degré et le sens de l'association entre deux variables quantitatives. Elle indique si les variables évoluent ensemble, sans établir de relation causale. (0.75pt)

La régression vise à modéliser et expliquer la relation entre une variable dépendante (Y) et une ou plusieurs variables explicatives (X), et permet surtout la prédiction. (0.75pt)

Nature de la relation étudiée

Pour la corrélation : Relation symétrique (Le coefficient de corrélation (r) est identique pour (X, Y) et (Y, X)). (0.75pt)

La regression : Relation asymétrique, distinction claire entre variable explicative (X) et variable réponse (Y) (La droite de régression de Y sur X est différente de celle de X sur Y). (0.75pt)

Les résultats fournis

Corrélation :

Un coefficient sans unité ($r \in [-1, +1]$) : Indique l'intensité et la direction de la relation linéaire. (0.75pt)

Regression :

La régression délivre une équation complète Y=a+bX (pente b, ordonnée à l'origine a), un R^2 ajusté pour la qualité du modèle, et des intervalles de confiance pour les prédictions. Le R^2 de la régression s'appuie souvent sur r de la corrélation simple ($R^2=r^2$), renforçant leur usage joint. (0.75pt)

Exercice

1. Formulation des hypothèses (1 pts)

Hypothèse nulle (H_0): Le type de défaut (ou l'absence de défaut) est indépendant du poste de contrôle. Autrement dit, la répartition des défauts est la même sur les trois postes.

Hypothèse alternative (H_1): Il existe une association entre le type de défaut et le poste de contrôle. La répartition des défauts n'est pas la même selon le poste.

2. Calcul des effectifs théoriques

Sous H_0, l'effectif théorique pour chaque case se calcule par: $E_{ij} = \frac{Total\ de\ la\ ligne\ i\ X\ Total\ de\ la\ colonne\ j}{Total\ général}$ (0.5 pt)

	Post A	Post B	Post C
étanchéité	23.33 (0.25 pt)	23.33 (0.25 pt)	23.33 (0.25 pt)
étiquetage	30 (0.25 pt)	30 (0.25 pt)	30 (0.25 pt)

remplissage	26.67 (0.25 pt)	26.67 (0.25 pt)	26.67 (0.25 pt)
Aucun défaut	20 (0.25 pt)	20 (0.25 pt)	20 (0.25 pt)

3. Calcul de la statistique du Khi-deux

$\chi^2_{Obs} = \sum \frac{(Oij-Eij)^2}{Eij}$ (0.5 pt)

<u>Etanchéité :</u>

Post A : $\frac{(15-23.33)^2}{23.33} = 0.1195$ (0.25 pt); **Post B** : $\frac{(30-23.33)^2}{23.33} = 1.907$ (0.25 pt) ; **Post C** : $\frac{(15-23.33)^2}{23.33} = 2.975$ (0.25 pt)

<u>Etiquetage :</u>

Post A : $\frac{(40-30)^2}{30} = 3.33$ (0.25 pt);**Post B** : $\frac{(20-30)^2}{30} = 3.33$ (0.25 pt) ; **Post C** : $\frac{(30-30)^2}{30} = 0$ (0.25 pt)

<u>Remplissage :</u>

Post A : $\frac{(15-26.67)^2}{26.67} = 5.107$ (0.25 pt);**Post B** : $\frac{(40-26.67)^2}{26.67} = 6.663$ (0.25 pt) ; **Post C** : $\frac{(25-26.67)^2}{26.67} = 0.1046$ (0.25 pt)

<u>Aucun défaut :</u>

Post A : $\frac{(20-20)^2}{20} = 0$ (0.25 pt);**Post B** : $\frac{(10-20)^2}{20} = 5$ (0.25 pt) ; **Post C** : $\frac{(30-20)^2}{20} = 5$ (0.25 pt)0.

χ^2_{Obs} = 0.1195 + 1.907 + 2.975 + 3.33 + 3.33 + 0+ 5.107 + 6.663 + 0.1046 + 0 + 5 + 5 = **33.54** (0.5 pt)

4- Décision statistique

Degrés de liberté (ddl) = (4-1) x (3-1) = 6 (0.5 pt)

Valeur critique pour α=0,05 et ddl = 6 : $\chi^2_{Crit} = 12.59$ (0.5 pt)

Comparaison :

$\chi^2_{Obs} = 33.54 > \chi^2_{Crit} = 12.59$

Conclusion du test : On rejette l'hypothèse nulle H_0 au seuil de 5 %. Il existe une association significative entre le type de défaut (ou l'absence de défaut) et le poste de contrôle. (0.5 pt)

5- (0.5 pt)

Les résultats indiquent que la répartition des défauts n'est pas homogène entre les postes de contrôle. Cela peut suggérer :

- Des différences dans la sensibilité des contrôles selon le poste (par exemple, le poste B détecte plus de défauts de remplissage, mais moins de défauts d'étiquetage).
- Des variations dans les conditions de production ou de contrôle selon les lignes ou équipes.