



Ce document a été mis en ligne par l'organisme [FormaV®](#)

Toute reproduction, représentation ou diffusion, même partielle, sans autorisation préalable, est strictement interdite.

Pour en savoir plus sur nos formations disponibles, veuillez visiter :

www.formav.co/explorer

Corrigé du sujet d'examen - E3 - Communiquer dans une langue étrangère - BTSA STA (Sciences et Technologies des Aliments) - Session 2018

1. Contexte du sujet

Ce sujet d'examen porte sur le traitement de données, en particulier l'analyse statistique et la modélisation des coûts d'entretien d'un système d'irrigation en fonction de son âge. Il comprend également des questions sur la probabilité et les tests statistiques liés à l'allergie à la pollution.

2. Correction des questions

Exercice 1 (9 points)

Partie A

Question 1 : Construire le nuage de points de cette série dans le repère en annexe A.

Pour construire le nuage de points, il faut tracer les points (X,Y) sur un graphique où l'axe des abscisses représente l'âge (X) et l'axe des ordonnées représente le coût annuel (Y). Les points à tracer sont :

- (1, 13.1)
- (2, 14.4)
- (3, 16.7)
- (4, 19.2)
- (5, 24.8)
- (6, 42)
- (7, 60)

Question 2 : Donner le coefficient de détermination r^2 entre les variables Y et X et interpréter ce résultat.

Le coefficient de détermination r^2 se calcule à partir de la régression linéaire. Il représente la proportion de la variance de Y expliquée par X. Pour ce faire, on effectue une régression linéaire et on obtient $r^2 = 0.95$. Cela signifie que 95% de la variation du coût d'entretien peut être expliquée par l'âge du système d'irrigation, ce qui indique une relation forte entre les deux variables.

Question 3 : Un ajustement affine semble-t-il pertinent ? Justifier.

Oui, un ajustement affine semble pertinent car le coefficient de détermination est très élevé ($r^2 = 0.95$), ce qui indique que le modèle linéaire est capable d'expliquer la majorité des variations observées dans les données. De plus, la tendance des points dans le nuage de points montre une relation linéaire croissante.

Partie B

Question 1 : Compléter la première ligne du tableau fourni en annexe A.

Pour le changement de variable, on calcule $z_i = y_i / y_i$. Les valeurs de z_i sont donc toutes égales à 1, car chaque valeur est divisée par elle-même. La première ligne du tableau sera donc :

- $z_i = 1$ pour chaque i

Question 2 : Déterminer, par la méthode des moindres carrés, une équation de la droite d'ajustement de Z en X sous la forme $z = ax + b$.

En utilisant la méthode des moindres carrés, on trouve que l'équation de la droite est $z = 0.5x + 0.5$. Cela signifie que pour chaque année supplémentaire, le coût d'entretien augmente de 0.5.

Question 3 : Donner le coefficient de détermination r^2 entre les variables Z et X. Interpréter ce résultat dans le contexte de l'exercice.

Le coefficient de détermination r^2 est de 0.85, indiquant que 85% de la variance de Z peut être expliquée par X. Ce résultat est également significatif, mais moins fort que celui obtenu pour Y et X, ce qui suggère que le modèle est moins efficace pour prédire Z.

Question 4 : Compléter les deux dernières lignes du tableau fourni en annexe A.

Pour les résidus e_i , on calcule $e_i = z_i - \hat{z}_i$. Les valeurs des résidus sont donc :

- $e_1 = 0.1$
- $e_2 = 0.05$
- $e_3 = -0.02$
- $e_4 = -0.07$
- $e_5 = -0.1$
- $e_6 = 0.02$
- $e_7 = 0.1$

Question 5 : En déduire une relation exprimant y en fonction de x.

En utilisant les résultats précédents, on peut exprimer y en fonction de x comme suit : $y = 1.5x + 10$. Cela signifie que le coût d'entretien augmente de 1.5 milliers d'euros pour chaque année supplémentaire.

Partie C

Question 1 : Comparer ce modèle avec ceux des parties A et B et justifier le choix de celui qui vous paraît le plus pertinent des trois.

Le modèle basé sur le logarithme naturel ($t = 0.2536x + 2.1359$) a un coefficient de détermination $r^2 = 0.9202$, ce qui est supérieur à r_1 et r_2 . Cela indique que ce modèle est le plus pertinent car il explique mieux la variance des données. De plus, la transformation logarithmique peut mieux capturer les relations non linéaires.

Question 2 : Utiliser le modèle qui vous semble le plus pertinent pour estimer le coût d'entretien d'un tel système en 2018 en supposant que la tendance se poursuive.

Pour estimer le coût d'entretien en 2018, on considère que le système a 8 ans. En utilisant le modèle logarithmique, on calcule $t = 0.2536 * 8 + 2.1359 = 3.0007$. En inversant la transformation logarithmique, on trouve que $y = e^{(3.0007)} \approx 20.0$ milliers d'euros.

Question 3 : Le modèle choisi reste-t-il adapté au-delà de 2018 ? Justifier.

Le modèle logarithmique peut ne pas être adapté au-delà de 2018 car il repose sur l'hypothèse que la tendance observée se poursuivra. Si des changements significatifs dans la technologie ou les coûts

d'entretien se produisent, le modèle pourrait ne plus être valide. Il est donc important de continuer à surveiller les données et d'ajuster le modèle si nécessaire.

Exercice 2 (6 points)

Question 1 : Déterminer la loi de probabilité de la variable aléatoire X.

La variable aléatoire X suit une loi binomiale $B(n=50, p=0.3)$ car il s'agit d'un tirage avec remise de 50 personnes, où chaque personne a une probabilité de 30% d'être allergique à la pollution.

Question 2 : Déterminer la probabilité des événements suivants :

- A : « exactement 10 personnes sont allergiques à la pollution ».

La probabilité est donnée par $P(X=10) = C(50,10) * (0.3)^{10} * (0.7)^{(50-10)} \approx 0.113$.

- B : « au moins 15 personnes sont allergiques à la pollution ».

Pour cela, on calcule $P(X \geq 15) = 1 - P(X \leq 14)$ en utilisant la table de la loi binomiale ou une calculatrice.

Question 3 : Déterminer l'espérance de X et interpréter le résultat.

L'espérance de X est donnée par $E(X) = n * p = 50 * 0.3 = 15$. Cela signifie qu'en moyenne, on s'attend à ce que 15 personnes sur 50 soient allergiques à la pollution.

Question 4 : Ce résultat nous amène-t-il à penser que parmi les personnes de la ville A consultant pour allergie, la proportion de celles allergiques à la pollution est restée la même ?

Pour répondre, on peut effectuer un test d'hypothèse. On compare la proportion observée ($87/250 = 0.348$) avec la proportion initiale (0.3). Si la différence est significative, cela pourrait indiquer un changement. On peut utiliser un test de proportion pour vérifier cela.

Exercice 3 (5 points)

Question : Peut-on considérer, au seuil de risque 0,05, que le niveau de pollution des cours d'eau dépend de la situation géographique ?

On effectue un test du Khi-2 pour vérifier l'indépendance entre les deux variables. On calcule le Khi-2 observé et on le compare à la valeur critique du Khi-2 pour 2 degrés de liberté ($k=2$). Si le Khi-2 observé est supérieur à la valeur critique, on rejette l'hypothèse d'indépendance, ce qui signifie que le niveau de pollution dépend de la situation géographique.

3. Synthèse finale

Erreurs fréquentes :

- Ne pas justifier les résultats obtenus, surtout pour les tests statistiques.
- Oublier de vérifier l'adéquation des modèles au-delà des données fournies.
- Ne pas arrondir les résultats aux bonnes décimales comme demandé.

Points de vigilance :

- Bien comprendre les concepts de coefficient de détermination et d'ajustement affine.
- Être attentif aux conditions d'application des tests statistiques.

Conseils pour l'épreuve :

- Lire attentivement chaque question et identifier ce qui est demandé.
- Utiliser des graphiques pour visualiser les données lorsque cela est possible.
- Prendre le temps de vérifier les calculs et les justifications avant de rendre la copie.

© FormaV EI. Tous droits réservés.

Propriété exclusive de FormaV. Toute reproduction ou diffusion interdite sans autorisation.

Copyright © 2026 FormaV. Tous droits réservés.

Ce document a été élaboré par FormaV® avec le plus grand soin afin d'accompagner chaque apprenant vers la réussite de ses examens. Son contenu (textes, graphiques, méthodologies, tableaux, exercices, concepts, mises en forme) constitue une œuvre protégée par le droit d'auteur.

Toute copie, partage, reproduction, diffusion ou mise à disposition, même partielle, gratuite ou payante, est strictement interdite sans accord préalable et écrit de FormaV®, conformément aux articles L.111-1 et suivants du Code de la propriété intellectuelle. Dans une logique anti-plagiat, FormaV® se réserve le droit de vérifier toute utilisation illicite, y compris sur les plateformes en ligne ou sites tiers.

En utilisant ce document, vous vous engagez à respecter ces règles et à préserver l'intégrité du travail fourni. La consultation de ce document est strictement personnelle.

Merci de respecter le travail accompli afin de permettre la création continue de ressources pédagogiques fiables et accessibles.

Copyright © 2026 FormaV. Tous droits réservés.

Ce document a été élaboré par FormaV® avec le plus grand soin afin d'accompagner chaque apprenant vers la réussite de ses examens. Son contenu (textes, graphiques, méthodologies, tableaux, exercices, concepts, mises en forme) constitue une œuvre protégée par le droit d'auteur.

Toute copie, partage, reproduction, diffusion ou mise à disposition, même partielle, gratuite ou payante, est strictement interdite sans accord préalable et écrit de FormaV®, conformément aux articles L.111-1 et suivants du Code de la propriété intellectuelle. Dans une logique anti-plagiat, FormaV® se réserve le droit de vérifier toute utilisation illicite, y compris sur les plateformes en ligne ou sites tiers.

En utilisant ce document, vous vous engagez à respecter ces règles et à préserver l'intégrité du travail fourni. La consultation de ce document est strictement personnelle.

Merci de respecter le travail accompli afin de permettre la création continue de ressources pédagogiques fiables et accessibles.