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 :
formav.co/explorer

BTSA STA

Traitement de données

Session 2019 • Code sujet : **2019-BTS103-RPL-ME-RE** • Durée : **180 minutes**

Sommaire

Cliquez sur une partie pour y accéder directement

1. Ajustement de l'effet des acariens prédateurs sur le dessèchement des rosiers (6 questions) **Page 3**

2. Modélisation du nombre d'abricotiers atteints de moniliose (4 questions) **Page 5**

3. Lien entre l'humidité du maïs et la qualité du foie gras (1 question) **Page 7**

4. Estimation de la matière grasse de la Brousse du Rove (4 questions) **Page 8**

5. Ce qu'il faut retenir & Questions fréquentes (FAQ) (Synthèse & Q&A Examen) **Page 10**

Sujet officiel (PDF) : https://btsasta.fr/assets/annales/btsa-sta/E3 - Communiquer dans une langue étrangère/2019/BTSA STA - E3 - Communiquer dans une langue étrangère - Session 2019 - 1-Sujet.pdf

1. Ajustement de l'effet des acariens prédateurs sur le dessèchement des rosiers

Un horticulteur en protection biologique intégrée teste, dans sept parcelles, l'effet du nombre de sacs d'acariens prédateurs sur le dessèchement des feuilles de rosiers. Le tableau et les graphiques permettent de comparer l'ajustement affine direct à un ajustement après transformation logarithmique.

Question 1. Expliquer pourquoi un ajustement affine entre les variables X, nombre de sacs, et Y, pourcentage de feuilles desséchées, n'est pas adapté.

Sur le graphique, les points ne suivent pas une droite : le pourcentage de feuilles desséchées chute fortement au début, puis la baisse ralentit. De plus, la droite représentée s'écarte nettement de certaines observations et prévoit un pourcentage négatif vers 6 sacs, ce qui est impossible. Une droite ne décrit donc pas correctement l'évolution du pourcentage en fonction du nombre de sacs.

✓ RÉPONSE À RETENIR

Le nuage est courbe et la droite prévoit un pourcentage négatif : l'ajustement affine direct de Y en X est inadapté.

Question 2. En posant $z_i = \ln(y_i)$ pour chaque parcelle, donner la valeur du coefficient de corrélation linéaire entre les variables X et Z.

Pour chaque parcelle, on transforme le pourcentage observé en calculant $z_i = \ln(y_i)$. Les sommes centrées obtenues sont $S_{xz} \approx -16{,}183$, $S_{xx} = 28$ et $S_{zz} \approx 9{,}572$. Le coefficient de corrélation vaut alors $r_{XZ} = S_{xz}/\sqrt{S_{xx}S_{zz}}$. Sa valeur proche de -1 indique une forte liaison linéaire décroissante après transformation logarithmique.

✓ RÉPONSE À RETENIR

Le coefficient est $r_{XZ} \approx -0{,}989$, ce qui indique une liaison linéaire négative très forte entre X et Z.

Question 3. Déterminer, par la méthode des moindres carrés, une équation de la droite d'ajustement de Z en X.

La pente de la droite des moindres carrés est $a = S_{xz}/S_{xx} = -16{,}182634/28 \approx -0{,}578$. Avec $\bar{x} = 3$ et $\bar{z} \approx 2{,}563713$, son ordonnée à l'origine est $b = \bar{z} - a\bar{x} \approx 2{,}563713 - (-0{,}577951 \times 3) = 4{,}298$. On arrondit les coefficients au millième.

✓ RÉPONSE À RETENIR

Une équation de la droite d'ajustement, avec coefficients arrondis au millième, est $\hat{z} = 4{,}298 - 0{,}578x$.

Question 4. À partir de l'équation de la droite d'ajustement obtenue à la question précédente, calculer le résidu $e_3 = z_3 - \hat{z}_3$.

1. Pour la parcelle 3, le tableau donne $x_3 = 2$ et $y_3 = 20$, donc $z_3 = \ln(20) \approx 2{,}995732$.
2. La droite obtenue à la question précédente donne $\hat{z}_3 = 4{,}297567 - 0{,}577951 \times 2 \approx 3{,}141665$.
3. Le résidu est la différence entre la valeur observée et la valeur estimée : $e_3 = z_3 - \hat{z}_3$.

> **✓ RÉPONSE À RETENIR**
>
> Pour la parcelle 3, le résidu observé moins estimé vaut $e_3 \approx -0{,}146$.

Question 5. À l'aide des informations obtenues et du nuage des résidus, justifier la pertinence de l'ajustement affine entre les variables X et Z.

La corrélation calculée est très forte et négative : $r \approx -0{,}989$. Cela confirme que les points transformés sont proches d'un alignement. Le nuage des résidus complète ce constat :

- les résidus sont répartis de part et d'autre de zéro ;
- ils ne montrent pas de tendance régulière en fonction du nombre de sacs.

L'ajustement affine entre X et Z est donc pertinent.

> **✓ RÉPONSE À RETENIR**
>
> L'ajustement affine entre X et Z est pertinent : corrélation forte et résidus sans tendance systématique.

Question 6. À l'aide de l'ajustement choisi, estimer le nombre de sacs nécessaires dans la parcelle pour que le pourcentage de feuilles de rosiers desséchées soit inférieur à 1 % au bout d'un mois.

Le pourcentage estimé est inférieur à 1 lorsque son logarithme est inférieur à $\ln(1) = 0$. Il faut donc que $4{,}297567 - 0{,}577951x < 0$, soit $x > 4{,}297567/0{,}577951 \approx 7{,}436$. Le nombre de sacs étant entier, le minimum est 8. Cette estimation extrapole au-delà des 6 sacs observés ; avec 8 sacs, le modèle estime environ 0,722 % de feuilles desséchées.

> **✓ RÉPONSE À RETENIR**
>
> Le modèle estime qu'il faut au minimum 8 sacs par parcelle pour obtenir moins de 1 % de feuilles desséchées.

2. Modélisation du nombre d'abricotiers atteints de moniliose

Au printemps 2018, 11 % des abricotiers d'un secteur d'Occitanie sont atteints de moniliose. L'exercice étudie le nombre d'arbres malades dans un échantillon aléatoire simple de 100 abricotiers et les probabilités associées.

Question 1. Déterminer la loi de probabilité de la variable aléatoire X, qui compte le nombre d'abricotiers malades dans l'échantillon de 100 arbres, et justifier la réponse.

On assimile chaque sélection à une épreuve de Bernoulli : « malade » est le succès, de probabilité $0{,}11$, et « non malade » a pour probabilité $0{,}89$. Pour un échantillon de 100 arbres, le nombre de succès suit donc une loi binomiale de paramètres $n = 100$ et $p = 0{,}11$. Pour tout $k \in \{0, \dots, 100\}$, $P(X = k) = \binom{100}{k}(0{,}11)^k(0{,}89)^{100-k}$.

> ✓ **RÉPONSE À RETENIR**
>
> $X \sim \mathcal{B}(100\,;\,0{,}11)$.

Question 2. Déterminer la probabilité que dix abricotiers soient malades dans l'échantillon.

La probabilité d'obtenir exactement dix arbres malades est le terme de la loi binomiale correspondant à $k = 10$. Le coefficient $\binom{100}{10}$ compte les choix possibles des dix arbres malades ; chaque choix a une probabilité $(0{,}11)^{10}(0{,}89)^{90}$. Ainsi, $P(X = 10) = \binom{100}{10}(0{,}11)^{10}(0{,}89)^{90}$, soit environ $0{,}125$ à 10^{-3} près.

> ✓ **RÉPONSE À RETENIR**
>
> $P(X = 10) \approx 0{,}125$.

Question 3. Indiquer par quelle loi on peut approcher la loi de la variable aléatoire X et justifier cette approximation.

On peut approcher une loi binomiale par une loi normale lorsque ses deux paramètres de dispersion np et $n(1-p)$ sont suffisamment grands. Ici, ils valent respectivement $100 \times 0{,}11 = 11$ et $100 \times 0{,}89 = 89$, valeurs supérieures à 5. La loi normale d'approximation a pour moyenne $np = 11$ et pour écart type $\sqrt{np(1-p)} = \sqrt{9{,}79} \approx 3{,}129$.

> ✓ **RÉPONSE À RETENIR**
>
> Approximation par $\mathcal{N}(11\,;\,3{,}129)$, où 3,129 est l'écart type.

Question 4. En utilisant la méthode de votre choix, calculer la probabilité qu'il y ait au moins 10 abricotiers malades dans l'échantillon.

Avec l'approximation normale et la correction de continuité, $P(X \geq 10) \approx P(Y \geq 9{,}5)$, où Y a pour moyenne 11 et écart type $\sqrt{9{,}79}$. La valeur centrée réduite du seuil est $(9{,}5 - 11)/\sqrt{9{,}79} \approx -0{,}479$, soit $-0{,}48$ à la précision de la table. Par symétrie, la probabilité vaut environ $\Phi(0{,}48) = 1 - 0{,}3156$.

✓ **RÉPONSE À RETENIR**

$P(X \geq 10) \approx 0{,}684$.

3. Lien entre l'humidité du maïs et la qualité du foie gras

Une étude porte sur 60 foies gras issus de canards nourris avec du maïs-grain sec ou humide. Le tableau répartit les effectifs selon trois niveaux de qualité et les deux taux d'humidité afin d'évaluer s'il existe un lien entre ces variables.

Question. Au seuil de risque 0,05, peut-on considérer que la qualité du foie gras dépend du taux d'humidité du maïs ?

On teste l'indépendance entre la qualité du foie gras et le taux d'humidité du maïs. Sous cette hypothèse, l'effectif théorique d'une case vaut le total de sa ligne multiplié par le total de sa colonne, puis divisé par 60.

Qualité	Maïs sec	Maïs humide
Tout venant	7,25	7,75
Premier choix	7,25	7,75
Extra	14,5	15,5

Tous les effectifs théoriques étant supérieurs à 5, le test du Khi-2 est applicable. La statistique vaut $\chi^2 \approx 3{,}804$, avec $(3-1)(2-1) = 2$ degrés de liberté. Au seuil de 5 %, la valeur critique est 5,99.

✓ RÉPONSE À RETENIR

Comme 3,804 < 5,99, l'indépendance n'est pas rejetée. Au seuil de 0,05, on ne peut pas conclure que la qualité dépend de l'humidité du maïs.

4. Estimation de la matière grasse de la Brousse du Rove

Une fromagerie mesure la quantité de matière grasse de dix fromages Brousse du Rove de 100 grammes pour vérifier l'exigence minimale de 45 grammes. La quantité est supposée suivre une loi normale, et les mesures servent à estimer ses paramètres puis à examiner la conformité au cahier des charges.

Question 1. Déterminer la moyenne et l'écart-type des quantités de matière grasse mesurées sur cet échantillon de dix fromages.

La moyenne est la somme des dix mesures divisée par 10. Pour l'écart-type descriptif de l'échantillon, on divise la somme des carrés des écarts à cette moyenne par 10 :

$\bar{x} = 434/10 = 43{,}400$ g et $s = \sqrt{34{,}48/10} = 1{,}857$ g. La somme des carrés des écarts à la moyenne vaut 34,48. Ces résultats décrivent les dix fromages mesurés, avec l'écart-type calculé selon la convention descriptive.

> ✓ RÉPONSE À RETENIR
>
> Moyenne : 43,400 g ; écart-type descriptif : 1,857 g.

Question 2. Déduire des résultats de l'échantillon une estimation ponctuelle de la moyenne μ et de l'écart-type σ.

L'estimation ponctuelle de la moyenne de la population est la moyenne observée : $\hat{\mu} = 43{,}400$ g. Pour l'écart-type inconnu σ, on retient ici l'écart-type corrigé, qui tient compte de l'estimation de la moyenne sur le même échantillon :

$\hat{\sigma} = \sqrt{34{,}48/(10-1)} = 1{,}957$ g. Cette estimation corrigée est également celle utilisée pour construire l'intervalle de confiance de Student à la question suivante.

> ✓ RÉPONSE À RETENIR
>
> Estimations ponctuelles : $\hat{\mu} = 43{,}400$ g et $\hat{\sigma} = 1{,}957$ g.

Question 3. Déterminer une estimation par intervalle de confiance de la moyenne μ au niveau de confiance 0,95.

La variable est supposée normale et son écart-type est estimé à partir de l'échantillon. On utilise donc la loi de Student à $10 - 1 = 9$ degrés de liberté. Pour un intervalle bilatéral de niveau 0,95, la table donne $t_{0{,}975;9} = 2{,}26$.

La marge vaut $2{,}26 \times 1{,}957/\sqrt{10} \approx 1{,}399$ g. En centrant l'intervalle sur la moyenne observée de 43,400 g, on obtient les bornes arrondies au millième.

> ✓ RÉPONSE À RETENIR
>
> Intervalle de confiance à 95 % pour μ : [42,001 ; 44,799] g.

Question 4. Peut-on conclure que la production respecte le cahier des charges imposant au minimum 45 grammes de matière grasse pour 100 grammes de fromage ?

Le cahier des charges fixe un minimum de 45 g de matière grasse pour 100 g de fromage. L'intervalle de confiance à 95 % de la moyenne est entièrement inférieur à ce seuil : sa borne supérieure est 44,799 g.

- La moyenne de la production est estimée inférieure au minimum requis au niveau de confiance retenu.
- On ne peut donc pas conclure que la production respecte le cahier des charges. Cet intervalle porte sur la moyenne, et non sur la teneur de chaque fromage pris séparément.

> **✓ RÉPONSE À RETENIR**
>
> Non : l'intervalle de confiance à 95 % de la moyenne est entièrement inférieur à 45 g.

Ce qu'il faut retenir de ce sujet (Session 2019)

Bilan de l'épreuve : Épreuve de traitement de données du BTSA, toutes options, session de remplacement 2019 en Métropole-Réunion, code 2019-BTS103-RPL-ME-RE, durée 180 minutes, sans coefficient imprimé. Quatre situations agricoles et agroalimentaires sont étudiées : protection des rosiers, moniliose de l'abricotier, qualité du foie gras et conformité de la Brousse du Rove. Les compétences portent sur l'ajustement, les probabilités, le test du khi-deux et l'estimation statistique.

Notions-clés, formules & règles à maîtriser

- La transformation z = ln(y) permet de tester si une relation exponentielle entre x et y peut être modélisée par une relation affine entre x et z.
- Le nombre de succès dans n essais indépendants de probabilité p suit une loi binomiale B(n, p).
- Le test du khi-deux d'indépendance compare les effectifs observés aux effectifs théoriques calculés sous l'hypothèse d'indépendance.
- Pour une moyenne normale avec écart-type inconnu, l'intervalle de confiance utilise la loi de Student à n - 1 degrés de liberté.

Points de vigilance & attentes du jury

- Observer le nuage de points avant de choisir un ajustement, puis examiner le nuage des résidus pour évaluer sa pertinence.
- Pour une approximation de loi, nommer la loi retenue et justifier que ses conditions d'application sont satisfaites.
- Dans un test du khi-deux, calculer les effectifs théoriques et utiliser les degrés de liberté adaptés au tableau.
- Pour l'intervalle de confiance, préciser la loi et les degrés de liberté, puis interpréter les résultats par rapport au seuil de 45 g.

Questions fréquentes sur cette épreuve (Q&A Examen)

Q : Comment vérifier si l'ajustement affine de X et Y convient ?

Observer le nuage de points : une tendance courbe indique qu'un ajustement affine direct convient mal. Après transformation logarithmique, vérifier que le nuage est approximativement linéaire et que les résidus ne présentent pas de structure marquée.

Q : Comment justifier l'approximation de la loi binomiale dans cet exercice ?

Ici, X suit une loi binomiale de paramètres 100 et 0,11. L'approximation normale est appropriée lorsque les effectifs théoriques np et n(1 - p) sont suffisamment grands ; la loi approchée a pour moyenne np et pour variance np(1 - p).

Q : Comment organiser le temps pour cette épreuve de 180 minutes ?

Le sujet comporte quatre exercices notés sur 20 points au total, de 4 à 6 points chacun. Répartir le temps selon le poids et la difficulté des exercices, puis garder quelques minutes pour vérifier les calculs, les arrondis au millième et les conclusions.

Q : Quelle loi utiliser pour l'intervalle de confiance sur la matière grasse ?

La quantité est supposée normale, l'écart-type de la production est inconnu et l'échantillon comprend dix fromages. L'intervalle de confiance de la moyenne s'appuie donc sur la loi de Student à neuf degrés de liberté.

Sujet officiel associé (PDF) : https://btsasta.fr/assets/annales/btsa-sta/E3 - Communiquer dans une langue étrangère/2019/BTSA STA - E3 - Communiquer dans une langue étrangère - Session 2019 - 1-Sujet.pdf
Ressources BTSA Sciences et Technologies des Aliments : Toutes les annales & corrigés • Fiches de révision complètes • Simulateur de notes d'examen

Copyright © 2027 FormaV®. Tous droits réservés.

FormaV® est une marque déposée auprès de l'INPI sous le numéro 5227897.

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.