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, toutes options

Session 2014 • Code sujet : **2014-BTS103-RPL-ME-RE** • Durée : **3 heures**

Sommaire

Cliquez sur une partie pour y accéder directement

1. **Régression linéaire et évolution du parasitisme dans une pinède** (5 questions) **Page 3**

2. **Estimation d'une proportion et modélisation binomiale des oliviers commercialisables** (7 questions) **Page 5**

3. **Test du khi-deux sur l'influence de la nature du terrain** (1 question) **Page 8**

4. **Loi de la moyenne d'échantillon et calcul de probabilité** (2 questions) **Page 9**

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/2014/BTSA STA - E3 - Communiquer dans une langue étrangère - Session 2014 - 1-Sujet.pdf

1. Régression linéaire et évolution du parasitisme dans une pinède

Une population de pins sylvestres est parasitée par des chenilles processionnaires, contre lesquelles des prédateurs ont été introduits. Les nombres de pins parasités de 2001 à 2010 servent à étudier la tendance et à estimer quand les parasites pourraient disparaître.

Question 1.a. À l'aide de la calculatrice, déterminer le coefficient de corrélation linéaire entre le rang de l'année X et le nombre de pins parasités Y, puis commenter le résultat.

Le coefficient de corrélation des dix couples vaut environ $r = -0,9971$. Très proche de -1, il indique une liaison linéaire très forte et décroissante : le nombre de pins parasités baisse nettement quand le rang de l'année augmente.

> ✓ **RÉPONSE À RETENIR**
>
> $r \approx -0,9971$: corrélation linéaire très forte et négative entre le rang de l'année et le nombre de pins parasités.

Question 1.b. Déterminer, par la méthode des moindres carrés, une équation de la droite d'ajustement de Y en X. Arrondir les coefficients au centième.

Pour la droite des moindres carrés $y = ax + b$, la pente vaut $a = -1275,5/82,5 \approx -15,46$. Avec $\bar{x} = 5,5$ et $\bar{y} = 98,9$, on obtient $b = \bar{y} - a\bar{x} \approx 183,93$. Les deux coefficients sont arrondis au centième.

> ✓ **RÉPONSE À RETENIR**
>
> Droite d'ajustement de Y en X : $\hat{y} = -15,46x + 183,93$.

Question 2.a. En utilisant l'équation de la droite de régression précédente, recopier et compléter les deux dernières lignes du tableau pour les dix valeurs de x_i : les estimations $\hat{y}_i$ et les résidus e_i.

On utilise la droite arrondie $\hat{y}_i = 183,93 - 15,46x_i$, puis la définition $e_i = y_i - \hat{y}_i$. Les valeurs ci-dessous, arrondies au centième, complètent les dix colonnes du tableau.

Rang x_i	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Estimation $\hat{y}_i$	168,47	153,01	137,55	122,09	106,63	91,17	75,71	60,25	44,79	29,33
Résidu e_i	2,53	1,99	3,45	−2,09	−6,63	−3,17	−0,71	−1,25	0,21	5,67

Tableau complété : estimations et résidus

x_i	1	2	3	4	5	6	7	8	9	10
y_i	171	155	141	120	100	88	75	59	45	35
$\hat{y}_i$	168,47	153,01	137,55	122,09	106,63	91,17	75,71	60,25	44,79	29,33
e_i	2,53	1,99	3,45	−2,09	−6,63	−3,17	−0,71	−1,25	0,21	5,67

✓ RÉPONSE À RETENIR

Estimations : 168,47 ; 153,01 ; 137,55 ; 122,09 ; 106,63 ; 91,17 ; 75,71 ; 60,25 ; 44,79 ; 29,33.
Résidus : 2,53 ; 1,99 ; 3,45 ; −2,09 ; −6,63 ; −3,17 ; −0,71 ; −1,25 ; 0,21 ; 5,67.

Question 2.b. Rappeler la propriété de la moyenne des résidus, puis la vérifier sur l'exemple précédent.

Pour une régression des moindres carrés avec constante, la somme des résidus est nulle ; leur moyenne est donc nulle. Avec les dix résidus du tableau, leur somme vaut 0, donc $\bar{e} = 0/10 = 0$.

✓ RÉPONSE À RETENIR

La moyenne des dix résidus est nulle : $\bar{e} = 0$.

Question 3. Si la tendance observée se confirme, déterminer à partir de quelle année les parasites auront disparu de la pinède.

La droite atteint zéro lorsque $183,93 - 15,46x \leq 0$, soit $x \geq 183,93/15,46 \approx 11,89$. Le premier rang entier correspondant est 12. Puisque le rang 1 désigne 2001, le rang 12 désigne 2012.

✓ RÉPONSE À RETENIR

Selon la tendance linéaire, les parasites auraient disparu à partir de 2012.

2. Estimation d'une proportion et modélisation binomiale des oliviers commercialisables

Un pépiniériste replante de jeunes oliviers et cherche à estimer la proportion qui sera commercialisable cinq ans plus tard. Les questions portent sur un échantillon de 120 arbres, puis sur le nombre d'arbres commercialisables parmi 300, sous l'hypothèse p = 0,75.

Partie A

Question 1.a. En 2000, 90 des 120 jeunes oliviers replantés ont été commercialisables cinq ans plus tard. Donner une estimation ponctuelle de la proportion inconnue p.

L'estimation ponctuelle de la proportion inconnue est la fréquence observée dans l'échantillon. On compte 90 oliviers commercialisables parmi les 120 replantés. La fréquence correspondante fournit donc l'estimation de *p*.

✓ **RÉPONSE À RETENIR**

$\hat{p} = 90/120 = 0{,}75$, soit 75 %.

Question 1.b. On note F la proportion d'oliviers commercialisables cinq ans plus tard dans les échantillons aléatoires de taille 120. Donner une approximation de la loi de probabilité de F.

Caractéristique de F	Approximation
Moyenne	p
Écart type	$\sqrt{p(1-p)/120}$

Pour un effectif assez grand, la proportion échantillonnale est approximativement normale. Comme la valeur réelle de p est inconnue, on peut utiliser l'estimation observée $0{,}75$ pour obtenir une approximation numérique de l'écart type. La loi estimée est alors centrée en 0,75.

✓ **RÉPONSE À RETENIR**

$F \approx \mathcal{N}\left(p, \sqrt{p(1-p)/120}\right)$. Avec $p \approx 0{,}75$, l'écart type estimé vaut 0,0395.

Question 1.c. Déterminer une estimation de p par intervalle de confiance au niveau de confiance 0,95.

1. La fréquence observée est $90/120 = 0{,}75$. Pour le niveau de confiance 0,95, on utilise l'intervalle usuel $[f - 1/\sqrt{n}\,;\ f + 1/\sqrt{n}]$.
2. Avec $n = 120$, la marge vaut environ 0,0913. Les bornes sont donc environ 0,6587 et 0,8413, puis arrondies au centième.

✓ **RÉPONSE À RETENIR**

Intervalle de confiance au niveau 0,95 : [0,66 ; 0,84].

Question 2.a. **On admet p = 0,75 et le pépiniériste replante 300 jeunes oliviers en 2010. Justifier que le nombre X d'oliviers commercialisables cinq ans plus tard suit une loi binomiale et préciser ses paramètres.**

1. Chaque olivier correspond à une épreuve de Bernoulli : il est commercialisable ou non, avec une probabilité de succès constante $p = 0{,}75$.
2. Les 300 reprises étant indépendantes, X compte les succès de 300 épreuves indépendantes de même probabilité. C'est donc une variable binomiale de paramètres $n = 300$ et $p = 0{,}75$.

> ✓ **RÉPONSE À RETENIR**
>
> $X \sim \mathcal{B}(300\ ;\ 0{,}75)$.

Question 2.b. **Déterminer l'espérance mathématique et l'écart type de la variable aléatoire X.**

Pour une loi binomiale de paramètres n et p, l'espérance est np et l'écart type est $\sqrt{np(1-p)}$. En remplaçant par les paramètres de X, on obtient une moyenne de 225 oliviers commercialisables et un écart type de 7,5 oliviers.

> ✓ **RÉPONSE À RETENIR**
>
> $E(X) = 225$ oliviers et $\sigma(X) = 7{,}5$ oliviers.

Question 2.c. **Indiquer par quelle loi on peut approcher la loi de X.**

On peut approcher une loi binomiale par une loi normale lorsque les nombres moyens de succès et d'échecs sont suffisamment grands. Ici, ils valent respectivement 225 et 75. L'approximation normale est donc adaptée ; elle conserve la moyenne et l'écart type de la loi binomiale.

- Moyenne : $300 \times 0{,}75 = 225$.
- Écart type : $\sqrt{300 \times 0{,}75 \times 0{,}25} = 7{,}5$.

> ✓ **RÉPONSE À RETENIR**
>
> X est approximativement normale de moyenne 225 et d'écart type 7,5.

Question 2.d. **En utilisant cette approximation, déterminer la probabilité que de 210 à 240 oliviers soient commercialisables cinq ans plus tard, puis la probabilité que plus de 80 % des oliviers soient commercialisables.**

Avec l'approximation normale de moyenne 225 et d'écart type 7,5, les bornes 210 et 240 correspondent aux valeurs centrées réduites −2 et 2. La table donne $\Phi(2) = 0{,}9772$ et, par symétrie, $\Phi(-2) = 0{,}0228$. L'événement B correspond à plus de 240 arbres, car 80 % de 300 vaut 240.

- $P(A) \approx \Phi(2) - \Phi(-2)$.

- $P(B) \approx 1 - \Phi(2)$.

> **✓ RÉPONSE À RETENIR**
>
> $P(A) \approx 0{,}9544$; $P(B) \approx 0{,}0228$.

3. Test du khi-deux sur l'influence de la nature du terrain

Un lot de 100 jeunes oliviers est réparti en deux groupes de 50, plantés dans des sols différents : limon argilo-sableux et sable argileux. Le tableau des résultats obtenus en 2011 sert à examiner si la nature du terrain influence le potentiel de commercialisation.

Partie B

Question Test. À partir du tableau des résultats, répondre à la question de l'influence de la nature du terrain sur le potentiel de commercialisation des oliviers en effectuant un test du khi-deux. Le seuil de risque retenu est 0,05.

On teste l'indépendance entre le type de sol et la commercialisation. Sous l'hypothèse d'indépendance, chaque effectif attendu est le produit du total de sa ligne par le total de sa colonne, divisé par le total général.

Sol	Commercialisables attendus	Non commercialisables attendus
Limon argilo-sableux	37,5	12,5
Sable argileux	37,5	12,5

La statistique du test vaut $\chi^2 = 1,33$, avec 1 degré de liberté. Au seuil de 5 %, la valeur critique est 3,84. Comme 1,33 est inférieur à 3,84, on ne rejette pas l'hypothèse d'indépendance : la différence observée n'est pas significative.

> **✓ RÉPONSE À RETENIR**
>
> Au seuil de 5 %, aucune influence significative de la nature du terrain sur la commercialisation n'est mise en évidence.

4. Loi de la moyenne d'échantillon et calcul de probabilité

La hauteur des oliviers commercialisables, exprimée en centimètres, suit une loi normale de moyenne 120 et d'écart type 7. On étudie la hauteur moyenne d'échantillons aléatoires simples de 25 arbres à l'aide de la table de la loi normale centrée réduite.

Partie C

Question 1. Rappeler la loi suivie par la variable aléatoire représentant la hauteur moyenne d'un échantillon de 25 oliviers, en précisant ses paramètres.

La moyenne d'un échantillon de 25 hauteurs indépendantes suit une loi normale, puisque les hauteurs individuelles suivent une loi normale. Son espérance est celle de la population : 120 cm. Son écart type vaut $\frac{7}{\sqrt{25}} = 1,4$ cm, donc sa variance est $1,4^2 = 1,96$ cm². Ainsi, avec la notation où le second paramètre est la variance, $\overline{X} \sim \mathcal{N}(120\,;\,1,96)$.

> **✓ RÉPONSE À RETENIR**
>
> $\overline{X} \sim \mathcal{N}(120\,;\,1,96)$: moyenne 120 cm, écart type 1,4 cm.

Question 2. Déterminer la probabilité que la hauteur moyenne d'un échantillon de 25 oliviers soit inférieure à 116,5 cm.

L'écart type de $\overline{X}$ est $1,4$ cm. La valeur $116,5$ correspond donc à la valeur centrée réduite $u = \frac{116,5-120}{1,4} = -2,5$. Par symétrie de la loi normale centrée réduite, $P(\overline{X} < 116,5) = \Phi(-2,5) = 1 - \Phi(2,5)$. La table donne $\Phi(2,5) = 0,9938$, d'où la probabilité demandée.

> **✓ RÉPONSE À RETENIR**
>
> $P(\overline{X} < 116,5) = 0,0062$, soit environ 0,62 %.

Ce qu'il faut retenir de ce sujet (Session 2014)

Bilan de l'épreuve : Cette épreuve de traitement de données du BTSA, session de remplacement 2014, porte le code 2014-BTS103-RPL-ME-RE et dure 3 heures ; aucun coefficient n'est imprimé. Elle étudie la régression appliquée au parasitisme d'une pinède et la commercialisation de jeunes oliviers. Les compétences évaluées comprennent l'ajustement linéaire, les résidus, l'estimation, les lois binomiale et normale, ainsi que le test du khi-deux.

Notions-clés, formules & règles à maîtriser

- Le coefficient de corrélation linéaire renseigne sur le sens et l'intensité de la liaison entre deux variables quantitatives.
- En régression linéaire, le résidu est $e_i = y_i - \hat{y}_i$; la moyenne des résidus est nulle.
- Pour une loi binomiale B(n,p), l'espérance vaut np et l'écart type vaut √(np(1-p)).
- Si une population suit N(μ, σ), la moyenne d'un échantillon de taille n suit N(μ, σ/√n).

Points de vigilance & attentes du jury

- Vérifier que la droite calculée ajuste Y en X et arrondir ses coefficients au centième.
- Calculer les résidus avec le signe $e_i = y_i - \hat{y}_i$ et contrôler leur moyenne.
- Pour l'approximation d'une binomiale, identifier la loi approchante et standardiser correctement les bornes.
- Pour le test du khi-deux, formuler les hypothèses, déterminer les degrés de liberté et comparer à la valeur critique au seuil 0,05.

Questions fréquentes sur cette épreuve (Q&A Examen)

Q : Comment utiliser la droite d'ajustement pour estimer l'année de disparition des parasites ?

Poser Y égal à zéro dans l'équation de régression pour trouver le rang correspondant, puis convertir ce rang en année, sachant que le rang 1 correspond à 2001. La prévision suppose que la tendance se poursuit.

Q : Comment déterminer la loi de la moyenne d'un échantillon d'oliviers ?

Si les hauteurs individuelles suivent une loi normale de moyenne μ et d'écart type σ, la moyenne de n observations indépendantes suit une loi normale de moyenne μ et d'écart type σ/√n. Ici, n vaut 25.

Q : Comment répartir son temps dans cette épreuve ?

L'épreuve dure trois heures. Les points annoncés sont 5 pour l'exercice 1, puis 8, 5 et 2 pour les parties A, B et C. Adapter le temps à ces poids et garder une marge pour vérifier les calculs.

Q : Que faut-il présenter pour justifier la conclusion du test du khi-deux ?

Présenter les hypothèses, les effectifs théoriques, la statistique calculée, les degrés de liberté et la valeur critique au seuil 0,05. Conclure sur l'influence mise en évidence, sans confondre absence de preuve et preuve d'absence.

Sujet officiel associé (PDF) : https://btsasta.fr/assets/annales/btsa-sta/E3 - Communiquer dans une langue étrangère/2014/BTSA STA - E3 - Communiquer dans une langue étrangère - Session 2014 - 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.