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 SCIENCES ET TECHNOLOGIES DES ALIMENTS • CORRIGÉ DÉTAILLÉ

BTSA STA

Épreuve E4 : Traitement de données

Session 2026 • Code sujet : **26-BTS-STA-E4** • Durée : **180 minutes** • Coefficient : **3**

Sommaire

Cliquez sur une partie pour y accéder directement

1. **Modélisation statistique d'un paramètre de fabrication** (5 questions) — **Page 3**
2. **Contrôle de conformité et variable aléatoire** (5 questions) — **Page 6**
3. **Estimation et prise de décision** (4 questions) — **Page 8**
4. **Ce qu'il faut retenir & Questions fréquentes (FAQ)** (Synthèse & Q&A Examen) — **Page 10**

Sujet officiel (PDF) : https://btsasta.fr/assets/annales/btsa-sta/E4 - Mettre en œuvre un modèle mathématique et une solution informatique adaptés au traitement de données/2026/BTSA STA - E4 - Mettre en œuvre un modèle mathématique et une solution informatique adaptés au traitement de données - Session 2026 - 1-Sujet.pdf

1. Modélisation statistique d’un paramètre de fabrication

Une équipe de la Laiterie des Trois Vallons étudie le lien entre la température de maturation de cinq lots et leur indice de texture. Le DT 1 fournit les mesures et des sommes récapitulatives, tandis que le DR 1 est à compléter pour synthétiser les calculs.

Question 1.1. À partir du DT 1, calculer les moyennes x̄ et ȳ, puis déterminer Sxx et Sxy.

On divise chaque somme par l’effectif pour obtenir les moyennes : $\bar{x} = 100/5 = 20$ °C et $\bar{y} = 35,0/5 = 7,0$. Avec les formules du sujet, on calcule ensuite $S_{xx} = 2160 - 5 \times 20^2 = 160$, puis $S_{xy} = 722,8 - 5 \times 20 \times 7 = 22,8$. Ces quatre valeurs sont à reporter dans les premières lignes du DR 1 ; elles serviront au calcul de la droite de régression et du coefficient de corrélation.

Valeurs à inscrire dans le DR 1

Grandeur	Résultat
$\bar{x}$	20 °C
$\bar{y}$	7,0
S_{xx}	160
S_{xy}	22,8
Coefficient directeur a	
Ordonnée à l’origine b	
Coefficient de corrélation r	
Indice de texture estimé pour $x = 22$ °C	

✓ RÉPONSE À RETENIR

Valeurs à reporter pour l’ajustement affine : $\bar{x} = 20$ °C, $\bar{y} = 7,0$, $S_{xx} = 160$ et $S_{xy} = 22,8$.

Question 1.2. Déterminer une équation de la droite de régression de y en x par la méthode des moindres carrés. Donner les coefficients arrondis au millième et présenter les étapes de calcul.

1. Le coefficient directeur vaut $a = S_{xy}/S_{xx} = 22,8/160 = 0,1425$, soit $0,143$ au millième.
2. L’ordonnée à l’origine vaut $b = \bar{y} - a\bar{x} = 7 - 0,1425 \times 20 = 4,150$.
3. La droite de régression est donc $y = 0,1425x + 4,150$. Pour la présenter avec les coefficients demandés au millième, on écrit $y = 0,143x + 4,150$. Les estimations ultérieures se calculent avec les coefficients non arrondis.

Coefficients à inscrire dans le DR 1

Coefficient directeur a	a = 0,143
Ordonnée à l'origine b	b = 4,150
Coefficient de corrélation r	
Indice de texture estimé pour $x = 22\ °\text{C}$	

✓ RÉPONSE À RETENIR

Le modèle de régression demandé est $y = 0,143x + 4,150$, avec $a = 0,143$ et $b = 4,150$, coefficients arrondis au millième.

Question 1.3. Calculer le coefficient de corrélation linéaire r à l'aide des données du DT 1 et interpréter sa valeur dans le contexte de l'étude.

On calcule d'abord $S_{yy} = \sum y_i^2 - n\bar{y}^2 = 248,30 - 5 \times 7^2 = 3,30$, puis $r = S_{xy}/\sqrt{S_{xx}S_{yy}} = 22,8/\sqrt{160 \times 3,30} \approx 0,992$. Interprétation dans le contexte :

- La corrélation est fortement positive, car r est proche de 1.
- Pour ces cinq lots, l'indice de texture tend à augmenter avec la température de maturation. Cette corrélation décrit les mesures observées et ne démontre pas, à elle seule, un lien de cause à effet.

Coefficient à inscrire dans le DR 1

Ordonnée à l'origine b	
Coefficient de corrélation r	r ≈ 0,992
Indice de texture estimé pour $x = 22\ °\text{C}$	

✓ RÉPONSE À RETENIR

$r \approx 0,992$: la relation observée entre température et texture est fortement positive et linéaire pour les cinq lots étudiés.

Question 1.4. À l'aide du modèle obtenu à la question 1.2, estimer l'indice de texture pour une température de maturation de 22 °C. Donner le résultat au centième.

On utilise la droite calculée à la question 1.2 en conservant le coefficient directeur non arrondi : $y = 0,1425 \times 22 + 4,15 = 7,285$. À la précision demandée, cela donne $7,29$ unités conventionnelles. La température de 22 °C étant comprise entre les températures mesurées, 12 °C et 28 °C, cette estimation est une interpolation. La valeur à inscrire dans la dernière ligne du tableau DR 1 est le résultat arrondi au centième.

Estimation à inscrire dans le DR 1

Coefficient de corrélation r	.....
Indice de texture estimé pour $x = 22\ °\mathrm{C}$	..... 7,29 .....

Document réponse associé aux questions 1.1 à 1.4

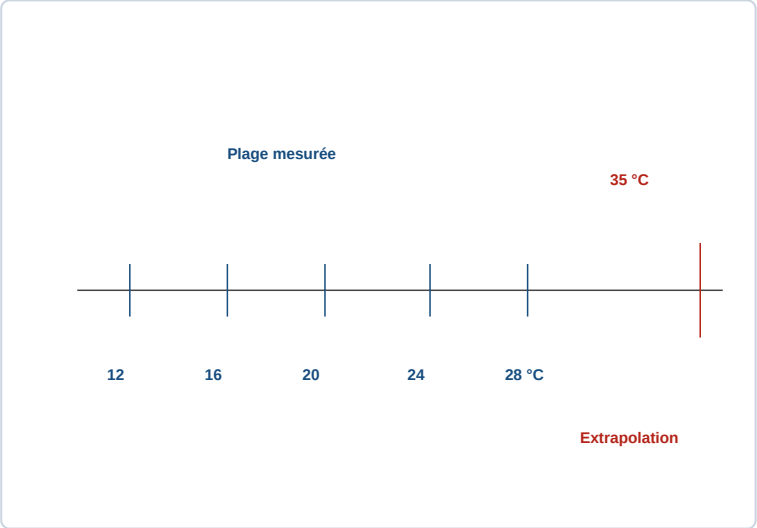
✓ RÉPONSE À RETENIR

Pour une maturation à 22 °C, l'indice de texture estimé est $7,29$ unités conventionnelles, au centième.

Question 1.5. Un responsable propose d'utiliser le modèle pour prévoir l'indice de texture à 35 °C. Indiquer si cette prévision est une interpolation ou une extrapolation, puis expliquer pourquoi elle doit être interprétée avec prudence.

La prévision à 35 °C est une extrapolation : la température dépasse l'intervalle des mesures, qui va de 12 °C à 28 °C. Le modèle a été ajusté sur cinq lots observés dans cette plage. Rien ne garantit que la relation reste linéaire au-delà de 28 °C ; le comportement du procédé pourrait changer. La prévision obtenue doit donc être considérée comme incertaine et ne remplace pas une mesure réalisée à cette température.

Température proposée hors plage mesurée



✓ RÉPONSE À RETENIR

À 35 °C, il s'agit d'une extrapolation hors de la plage observée $[12; 28]$ °C ; la prévision est donc incertaine.

2. Contrôle de conformité et variable aléatoire

Le DT 2 présente les contrôles de 200 lots fabriqués sur les lignes A et B, avec leurs effectifs conformes et non conformes. Une seconde analyse porte sur un prélèvement aléatoire de 20 lots, avec une probabilité de non-conformité constante de 0,08 et des résultats supposés indépendants.

Question 2.1. À partir du tableau du DT 2, déterminer les lois marginales des événements « ligne A ou ligne B » et « conforme ou non conforme », puis donner les probabilités correspondantes.

Événement	Effectif	Probabilité
Ligne A	160	0,800
Ligne B	40	0,200
Conforme C	174	0,870
Non conforme $\overline{C}$	26	0,130

Les lois marginales s'obtiennent à partir des totaux de lignes et de colonnes du tableau du DT 2. Pour chaque catégorie, on divise son effectif par les 200 lots contrôlés. Les probabilités des deux catégories de chaque variable se complètent à 1 : $0,800 + 0,200 = 1$ et $0,870 + 0,130 = 1$.

> ✓ **RÉPONSE À RETENIR**
>
> $P(A) = 0,800$, $P(B) = 0,200$, $P(C) = 0,870$ et $P(\overline{C}) = 0,130$.

Question 2.2. Calculer P(C | A) et P(C | Ā), puis interpréter les deux résultats dans le contexte du contrôle qualité.

Une probabilité conditionnelle se calcule en rapportant les lots conformes à l'effectif de la ligne concernée. Pour la ligne A, $P(C \mid A) = 144/160 = 0,900$. Pour la ligne B, complémentaire de A, $P(C \mid \overline{A}) = 30/40 = 0,750$. Dans l'échantillon contrôlé, la conformité est donc plus fréquente sur la ligne A, avec un écart de 0,150, soit 15 points de pourcentage. Ces proportions décrivent les lots observés.

> ✓ **RÉPONSE À RETENIR**
>
> $P(C \mid A) = 0,900$ et $P(C \mid \overline{A}) = 0,750$: la conformité observée est plus élevée sur la ligne A.

Question 2.3. Les événements A et C sont-ils indépendants ? Justifier la réponse par un calcul faisant intervenir les probabilités marginales et une probabilité conjointe.

Si A et C étaient indépendants, leur probabilité conjointe serait égale au produit de leurs probabilités marginales. Or, d'après le tableau, $P(A \cap C) = 144/200 = 0,720$, tandis que $P(A)P(C) = (160/200)(174/200) = 0,696$. Comme ces deux valeurs diffèrent, A et C ne sont pas indépendants dans les données observées. Cela indique une association entre la ligne de production et la conformité dans cet échantillon.

✓ RÉPONSE À RETENIR

A et C ne sont pas indépendants dans l'échantillon : $P(A \cap C) = 0,720$, différent de $P(A)P(C) = 0,696$.

Question 2.4. On note X le nombre de lots non conformes dans le prélèvement de 20 lots. Justifier que X suit une loi binomiale et préciser ses paramètres. Calculer P(X ≤ 1), arrondie au millième.

1. Le prélèvement comprend 20 essais indépendants. Pour chacun, la probabilité de non-conformité reste égale à 0,08. Le nombre X de lots non conformes suit donc la loi binomiale $\mathcal{B}(20; 0,08)$.
2. $P(X \leq 1) = P(X = 0) + P(X = 1) = 0,92^{20} + 20 \times 0,08 \times 0,92^{19} \approx 0,517$.

✓ RÉPONSE À RETENIR

$X \sim \mathcal{B}(20; 0,08)$ et $P(X \leq 1) \approx 0,517$, arrondi au millième.

Question 2.5. Calculer l'espérance de X et interpréter ce résultat pour un grand nombre de prélèvements de 20 lots.

Pour une variable aléatoire suivant une loi binomiale de paramètres $n = 20$ et $p = 0,08$, l'espérance vaut $E(X) = np$. Ainsi, $E(X) = 20 \times 0,08 = 1,6$. Sur un grand nombre de prélèvements de 20 lots réalisés dans les conditions décrites, le nombre moyen de lots non conformes par prélèvement sera proche de 1,6.

✓ RÉPONSE À RETENIR

$E(X) = 1,6$: en moyenne, 1,6 lot non conforme par prélèvement de 20 lots.

3. Estimation et prise de décision

La laiterie souhaite estimer la proportion de lots non conformes et la moyenne d'un paramètre de fabrication. Le DT 2 fournit les résultats des contrôles, les données d'un échantillon de 36 mesures et les formules d'intervalles de confiance approximatifs à 95 %.

Question 3.1. Dans l'étude de conformité du DT 2, calculer la fréquence observée de lots non conformes parmi les 200 lots contrôlés.

Le tableau du DT 2 indique 26 lots non conformes parmi les 200 lots contrôlés. La fréquence observée est le quotient de cet effectif par l'effectif total : $f = \frac{26}{200} = 0{,}13$. En pourcentage, cela représente $0{,}13 \times 100 = 13\,\%$. Cette fréquence décrit les lots de l'échantillon contrôlé et sert d'estimation de la proportion de non-conformité dans la production.

> ✓ **RÉPONSE À RETENIR**
>
> Fréquence observée : $0{,}13$, soit $13\,\%$ de lots non conformes.

Question 3.2. Déterminer un intervalle de confiance approximatif à 95 % de la proportion de lots non conformes à partir des résultats du DT 2. Donner les bornes au millième et expliciter la formule utilisée.

Avec la fréquence observée $f = 0{,}13$ et l'effectif $n = 200$, on applique la formule du DT 2 : $\left[f - 1{,}96\sqrt{f(1-f)/n}\ ;\ f + 1{,}96\sqrt{f(1-f)/n}\right]$.

1. La marge d'erreur vaut $1{,}96\sqrt{\frac{0{,}13(1-0{,}13)}{200}} \approx 0{,}04661$.
2. Les bornes sont $0{,}13 - 0{,}04661 \approx 0{,}083$ et $0{,}13 + 0{,}04661 \approx 0{,}177$.

L'intervalle estime la proportion de lots non conformes dans la production.

> ✓ **RÉPONSE À RETENIR**
>
> Intervalle de confiance approximatif à 95 % : $[0{,}083\ ;\ 0{,}177]$.

Question 3.3. Dans l'étude du paramètre de fabrication du DT 2, déterminer un intervalle de confiance approximatif à 95 % de la moyenne du paramètre mesuré. Donner les bornes au centième.

Le DT 2 donne $\bar{x} = 4{,}82$, $s = 0{,}30$ et $n = 36$. Selon la formule fournie, la marge d'erreur vaut $1{,}96\frac{s}{\sqrt{n}} = 1{,}96\frac{0{,}30}{\sqrt{36}} = 0{,}098$. Les bornes sont donc $4{,}82 - 0{,}098 = 4{,}722$ et $4{,}82 + 0{,}098 = 4{,}918$. Arrondies au centième, elles donnent l'intervalle de confiance approximatif demandé.

> ✓ **RÉPONSE À RETENIR**
>
> Intervalle de confiance approximatif à 95 % : $[4{,}72\ ;\ 4{,}92]$.

Question 3.4. La spécification interne impose une moyenne du paramètre mesuré inférieure ou égale à 4,90. À partir de l'intervalle obtenu à la question 3.3, indiquer si les résultats permettent de conclure que cette exigence est respectée. Justifier la décision et préciser une limite de cette conclusion.

L'intervalle obtenu est $[4{,}72\ ;\ 4{,}92]$, alors que la moyenne doit être inférieure ou égale à $4{,}90$. Sa borne supérieure dépasse cette limite : l'intervalle contient donc des valeurs compatibles avec l'exigence et d'autres qui ne le sont pas. Les résultats ne permettent pas de conclure que l'exigence est respectée. Cette conclusion reste limitée par l'incertitude d'échantillonnage et par le caractère approximatif de l'intervalle.

> **✓ RÉPONSE À RETENIR**
>
> On ne peut pas conclure au respect de l'exigence : la borne supérieure $4{,}92$ dépasse $4{,}90$.

Ce qu'il faut retenir de ce sujet (Session 2026)

Bilan de l'épreuve : Sujet de traitement de données de la session 2026, repéré 26-BTS-STA-E4, d'une durée de 180 minutes et de coefficient 3. Il porte sur les données de la Laiterie des Trois Vallons : températures de maturation, indices de texture, contrôles de conformité et mesures de fabrication. Il évalue l'ajustement affine, les probabilités, la loi binomiale et les intervalles de confiance. Le sujet précise qu'il reprend l'ancien référentiel M41, abrogé depuis 2025, et n'est pas une épreuve officielle du BTSA STA pour cette session.

Notions-clés, formules & règles à maîtriser

- Pour la régression de y en x, a = Sxy/Sxx et b = $\bar{y}$ - a$\bar{x}$; le modèle s'écrit y = ax + b.
- A et C sont indépendants si et seulement si P(A ∩ C) = P(A)P(C), ou si P(C | A) = P(C) lorsque P(A) > 0.
- Pour n essais indépendants de probabilité de succès p constante, le nombre de succès suit B(n,p) et son espérance vaut np.
- Les intervalles de confiance approximatifs à 95 % sont calculés selon les formules indiquées dans le DT 2.

Points de vigilance & attentes du jury

- Distinguer une interpolation, dans l'intervalle observé des températures, d'une extrapolation en dehors de cet intervalle.
- Pour une probabilité conditionnelle, prendre comme dénominateur l'effectif total du groupe conditionnant.
- Justifier la loi binomiale par le nombre fixe d'essais, leur indépendance et la constance de la probabilité.
- Comparer les bornes de l'intervalle à la spécification et préciser la limite de la conclusion tirée d'une estimation.

Questions fréquentes sur cette épreuve (Q&A Examen)

Q : Comment construire la droite de régression demandée dans la partie 1 ?

Calculez d'abord $\bar{x}$, $\bar{y}$, Sxx et Sxy à partir du DT 1. Utilisez ensuite a = Sxy/Sxx et b = $\bar{y}$ - a$\bar{x}$, puis écrivez y = ax + b. Respectez l'arrondi demandé et présentez les étapes.

Q : Comment choisir le dénominateur pour une probabilité conditionnelle du tableau ?

Pour P(C | A), divisez l'effectif conforme de la ligne A par le total de cette ligne. Pour P(C | $\bar{A}$), utilisez les effectifs de la ligne B, car $\bar{A}$ signifie que le lot provient de la ligne B.

Q : Comment gérer le temps et les points dans cette épreuve ?

L'épreuve dure 180 minutes et comporte trois parties indépendantes, notées respectivement sur 7, 7 et 6 points. Répartissez le temps selon ce barème, justifiez les calculs et complétez le document réponse DR 1.

Q : Quels critères de notation faut-il particulièrement respecter dans ce sujet ?

Les consignes demandent des calculs justifiés, des interprétations liées au contexte et les arrondis prescrits : probabilités au millième si nécessaire, autres résultats au centième sauf indication contraire. Utilisez les formules du DT 2.

Sujet officiel associé (PDF) : https://btsasta.fr/assets/annales/btsa-sta/E4 - Mettre en œuvre un modèle mathématique et une solution informatique adaptés au traitement de données/2026/BTSA STA - E4 - Mettre en œuvre un modèle mathématique et une solution informatique adaptés au traitement de données - Session 2026 - 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.