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 GF

Traitement de données, toutes options

Session 2022 • Code sujet : **2022-BTS103-NOR-ME-RE-MA** • Durée : **180 minutes**

Sommaire

Cliquez sur une partie pour y accéder directement

1. **Modélisation de l’évolution du nombre de bornes de recharge** (10 questions) — **Page 3**
2. **Estimation de la proportion de bornes défectueuses** (4 questions) — **Page 7**
3. **Durée de recharge et loi normale** (8 questions) — **Page 9**
4. **Test du Khi2 d’indépendance entre fournisseur et zone géographique** (4 questions) — **Page 12**
5. **Ce qu’il faut retenir & Questions fréquentes (FAQ)** (Synthèse & Q&A Examen) — **Page 14**

Sujet officiel (PDF) : https://btsagf.fr/assets/annales/btsa-gf/E3 - Communiquer dans des situations et des contextes variés/2022/BTSA GF - E3 - Communiquer dans des situations et des contextes variés - Session 2022 - 1-Sujet.pdf

1. Modélisation de l'évolution du nombre de bornes de recharge

Zoé étudie l'évolution, de juin 2017 à juin 2020, du nombre de bornes publiques de recharge pour voitures électriques en France. Elle compare un ajustement affine à un modèle quadratique obtenu par changement de variable, puis utilise ce modèle pour estimer le nombre de bornes en juin 2022.

Question 1

Question 1.a. À partir des sept relevés, déterminer le coefficient de corrélation linéaire r1 entre le rang X et le nombre Y de bornes, exprimé en milliers.

Les moyennes sont $\bar{x} = 4$ et $\bar{y} = 16,514$. Les sommes des produits centrés et des carrés centrés valent respectivement $82,2$, 28 et $245,909$. Ainsi, $r_1 = \frac{82,2}{\sqrt{28 \times 245,909}}$. Le coefficient est positif et très proche de 1 : le rang et le nombre de bornes présentent une forte liaison linéaire croissante.

> **✓ RÉPONSE À RETENIR**
>
> $r_1 \approx 0,991$

Question 1.b. En tenant compte du nuage de points et de la valeur de r1, dire si le choix d'un ajustement affine entre X et Y est pertinent et justifier la réponse.

À première vue, l'ajustement affine est pertinent : le nuage est croissant et les points suivent globalement une direction rectiligne. Le coefficient $r_1 \approx 0,991$, très proche de 1, confirme une forte liaison linéaire positive. Cependant, ce coefficient ne suffit pas à valider le modèle : l'examen des résidus permet de vérifier si la droite rend compte de la tendance sans erreur systématique.

> **✓ RÉPONSE À RETENIR**
>
> L'ajustement affine paraît pertinent à première vue : nuage croissant et $r_1 \approx 0,991$.

Question 1.c. Déterminer, par la méthode des moindres carrés, une équation de la droite d'ajustement de Y en X.

La pente de la droite des moindres carrés vaut le quotient de la somme des produits centrés par la somme des carrés centrés des rangs : $82,2/28 \approx 2,936$. La droite passe par le point moyen $(4 ; 16,514)$, donc son ordonnée à l'origine vaut $16,514 - 2,936 \times 4 \approx 4,771$. Comme Y est exprimé en milliers, l'équation donne le nombre de bornes en milliers.

> **✓ RÉPONSE À RETENIR**
>
> $y \approx 2,936x + 4,771$, avec y en milliers.

Question 1.d. À partir du nuage des résidus de l'ajustement affine, expliquer pourquoi Zoé décide d'effectuer un autre ajustement.

Le nuage des résidus montre une structure courbe, et non une répartition aléatoire autour de zéro : les résidus sont positifs au début et à la fin, tandis qu'ils sont principalement négatifs au milieu. La droite surestime donc certaines valeurs et en sous-estime d'autres de façon systématique. Un modèle non affine est à rechercher pour mieux représenter l'évolution observée.

> ✓ **RÉPONSE À RETENIR**
>
> La forme courbe des résidus révèle que l'ajustement affine décrit mal la tendance.

Question 2

Question 2.a. Pour la variable U définie par U = (X + 2,5)², compléter le tableau de l'Annexe A avec les valeurs de ui associées aux rangs xi.

Rang x_i	**1**	**2**	**3**	**4**	**5**	**6**	**7**
u_i	12,25	20,25	30,25	42,25	56,25	72,25	90,25
y_i	9,1	10,2	12,8	15,9	18,9	22,3	26,4

Pour compléter la ligne, on applique à chaque rang la définition $u_i = (x_i + 2,5)^2$. Par exemple, pour le rang 3, $u_3 = (3 + 2,5)^2 = 30,25$. Les valeurs déjà inscrites dans l'annexe pour les rangs 1 et 2 sont conservées. La ligne des y_i reprend les données du tableau de l'énoncé.

Annexe A complétée

Rang x_i	1	2	3	4	5	6	7
u_i	12,25	20,25	30,25	42,25	56,25	72,25	90,25
y_i	9,1	10,2	12,8	15,9	18,9	22,3	26,4

> ✓ **RÉPONSE À RETENIR**
>
> **u_i, pour les rangs 1 à 7 : 12,25 ; 20,25 ; 30,25 ; 42,25 ; 56,25 ; 72,25 ; 90,25.**

Question 2.b. **Sachant que y = 0,226u + 6,072 est une équation de la droite d'ajustement de Y en U par la méthode des moindres carrés, déterminer le coefficient de corrélation linéaire r2 entre U et Y.**

Pour les sept couples $(u_i ; y_i)$, la somme des produits centrés vaut $1087,4$. Les sommes des carrés centrés valent 4816 pour U et $245,909$ pour Y. Le coefficient est donc $r_2 = 1087,4 / \sqrt{4816 \times 245,909}$. Il est positif, conformément à la pente de la droite, et extrêmement proche de 1 : la liaison linéaire entre U et Y est très forte.

> ✓ **RÉPONSE À RETENIR**
>
> $r_2 \approx 0,999$

Question 3. **En tenant compte des informations des questions 2 et 3, notamment du coefficient de corrélation et du nuage des résidus du deuxième ajustement, dire s'il est plus pertinent de choisir ce deuxième ajustement et justifier.**

Oui, les éléments conduisent à préférer le deuxième ajustement. Le coefficient $r_2 \approx 0,999$ indique une liaison linéaire presque parfaite entre U et Y. De plus, le nuage des résidus du deuxième modèle ne présente plus la forme courbe nette observée pour le premier ajustement : les points sont dispersés autour de zéro sans tendance systématique évidente.

> ✓ **RÉPONSE À RETENIR**
>
> Oui : $r_2 \approx 0,999$ et les résidus du deuxième ajustement ne montrent pas de tendance courbe nette.

Question 4

Question 4.a. **Justifier que y = 0,226x² + 1,13x + 7,485 est une équation de la courbe d'ajustement entre X et Y obtenue par le changement de variable de la question 2.**

On remplace u par $(x+2,5)^2$ dans $y = 0,226u + 6,072$. En développant, $y = 0,226(x^2 + 5x + 6,25) + 6,072 = 0,226x^2 + 1,13x + 7,4845$. À 10^{-3} près, la constante vaut $7,485$. On obtient donc bien l'équation de la courbe d'ajustement donnée.

> ✓ **RÉPONSE À RETENIR**
>
> $y = 0,226x^2 + 1,13x + 7,485$

Question 4.b. **En supposant que l'ajustement reste valable après juin 2020, calculer une estimation du nombre de bornes mises à disposition du public en juin 2022.**

Les relevés sont espacés de six mois. Juin 2020 correspond au rang 7 ; deux ans plus tard, juin 2022 correspond au rang 11. En évaluant le modèle à ce rang, on obtient $y = 0,226 \times 11^2 + 1,13 \times 11 + 7,485 = 47,261$. Comme y est en milliers, l'estimation est de 47 261 bornes.

✓ RÉPONSE À RETENIR

Estimation en juin 2022 : 47 261 bornes.

Question 4.c. Un cabinet spécialisé prévoit 50 000 bornes en juin 2022. Commenter cette affirmation au regard du modèle établi par Zoé.

Le modèle prévoit 47 261 bornes en juin 2022, contre 50 000 selon le cabinet. La prévision du cabinet dépasse celle du modèle de 2 739 bornes, soit environ 5,8 % de l'estimation de Zoé. Les deux prévisions sont du même ordre de grandeur : celle du cabinet est proche de la valeur du modèle, mais ne correspond pas à son estimation exacte.

✓ RÉPONSE À RETENIR

La prévision du cabinet dépasse celle du modèle de **2 739 bornes**, soit environ **5,8 %**.

2. Estimation de la proportion de bornes défectueuses

Un technicien de maintenance prélève un échantillon aléatoire simple et indépendant de 150 bornes publiques et y relève six équipements défectueux. L'exercice porte sur l'estimation de la proportion de défauts, puis sur le nombre de bornes défectueuses dans un échantillon lorsque la proportion est fixée à 0,04.

Question 1

Question 1.a. À partir des six bornes défectueuses parmi les 150 prélevées, déterminer une estimation ponctuelle de la proportion p de bornes défectueuses dans l'ensemble des bornes mises à disposition du public.

L'estimation ponctuelle de la proportion de bornes défectueuses est la fréquence observée dans l'échantillon. Il y a 6 bornes défectueuses sur 150 bornes prélevées, donc $\hat{p} = \frac{6}{150} = 0,04$. Cela correspond à une estimation de 4 % de bornes défectueuses dans l'ensemble du parc.

> ✓ **RÉPONSE À RETENIR**
>
> $\hat{p} = 0,04$, soit 4 %.

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

Au niveau de confiance 0,95, on utilise l'intervalle de confiance usuel $[f - \frac{1}{\sqrt{n}} \ ; \ f + \frac{1}{\sqrt{n}}]$, avec la fréquence observée $f = 0,04$ et l'effectif $n = 150$. Ainsi, les bornes sont $0,04 - \frac{1}{\sqrt{150}}$ et $0,04 + \frac{1}{\sqrt{150}}$. Après arrondi au millième, on obtient l'intervalle demandé.

> ✓ **RÉPONSE À RETENIR**
>
> $[-0,042 \ ; \ 0,122]$.

Question 2

Question 2.a. En supposant p = 0,04, justifier que la variable aléatoire X, qui compte les bornes défectueuses dans un échantillon de 150, suit une loi binomiale de paramètres n = 150 et p = 0,04.

Les 150 bornes sont prélevées de façon aléatoire et indépendante. Pour chacune, l'issue « défectueuse » a la même probabilité $p = 0,04$, et l'issue contraire a une probabilité $0,96$. La variable X compte le nombre de bornes défectueuses parmi ces 150 épreuves de Bernoulli indépendantes. Elle suit donc une loi binomiale de paramètres $n = 150$ et $p = 0,04$.

> ✓ **RÉPONSE À RETENIR**
>
> $X \sim \mathcal{B}(150 \ ; \ 0,04)$.

Question 2.b. Calculer la probabilité qu'un échantillon de 150 bornes contienne au moins huit bornes défectueuses lorsque p = 0,04.

« Au moins 8 » correspond à l'événement $X \geq 8$. On calcule sa probabilité par complément : $P(X \geq 8) = 1 - P(X \leq 7)$. Avec la loi $\mathcal{B}(150\ ;\ 0,04)$, la probabilité d'obtenir au plus 7 bornes défectueuses est environ $0,747$. On en déduit la probabilité recherchée, arrondie au millième.

> ✓ **RÉPONSE À RETENIR**
>
> $P(X \geq 8) \approx 0,253$.

3. Durée de recharge et loi normale

La durée de recharge d'une batterie vide avec une borne est modélisée par une variable aléatoire suivant une loi normale, dont la densité est représentée graphiquement. Une seconde partie étudie des probabilités et l'évolution de la durée après une augmentation de 10 % au bout de cinq ans.

Partie A

Question A.1.a. Par lecture graphique, déterminer et justifier l'espérance mathématique μ de la durée de recharge.

La densité d'une loi normale est symétrique autour de son maximum. Sur le graphique, ce maximum est à 5 heures : l'axe de symétrie donne l'espérance.

> ✓ **RÉPONSE À RETENIR**
>
> $\mu = 5$ heures.

Question A.1.b. Par lecture graphique, déterminer et justifier la probabilité que la durée de recharge de la batterie soit comprise entre 3 et 7 heures.

L'aire à droite de 7 heures vaut 0,025. Par symétrie autour de 5 heures, l'aire à gauche de 3 heures vaut aussi 0,025.

$$P(3 \leq Y \leq 7) = 1 - 0,025 - 0,025 = 0,950.$$

> ✓ **RÉPONSE À RETENIR**
>
> $P(3 \leq Y \leq 7) = 0,950$, soit 95 %.

Question A.2. À partir de la question précédente, déduire une valeur approchée de l'écart-type σ à l'unité et justifier le raisonnement.

1. L'intervalle $[3 ; 7]$, centré en 5, contient 95 % des valeurs.
2. Pour une loi normale, cet intervalle correspond à environ $[\mu - 1,96\sigma ; \mu + 1,96\sigma]$. Ainsi, $2/\sigma \approx 1,96$, donc $\sigma \approx 2/1,96 \approx 1,020$ heure.

> ✓ **RÉPONSE À RETENIR**
>
> À l'unité, $\sigma \approx 1$ heure.

Partie B

Question B.1. Dans cette partie, avec μ = 5 et σ = 1, calculer la probabilité que la durée de recharge de la batterie vide avec la borne soit inférieure à 6 h 30.

On centre et réduit le seuil : $u = (6,5 - 5)/1 = 1,5$. La table donne $P(U \leq 1,5) = 0,9332$. La loi normale étant continue, les probabilités avec une inégalité stricte ou large sont égales.

> ✓ **RÉPONSE À RETENIR**
>
> $P(Y < 6,5) = 0,933$, soit environ 93,3 %.

Question B.2.a. Au bout de cinq ans, la durée de recharge augmente de 10 % et Z = 1,1Y. Déterminer l'espérance mathématique de Z et interpréter le résultat dans le contexte de l'exercice.

Pour une constante positive, $E(1,1Y) = 1,1E(Y)$. Donc $E(Z) = 1,1 \times 5 = 5,5$ heures. Après cinq ans, la durée moyenne de recharge est de 5,5 heures, soit 0,5 heure, ou 30 minutes, de plus qu'au départ.

> ✓ **RÉPONSE À RETENIR**
>
> $E(Z) = 5,5$ heures : la durée moyenne augmente de 30 minutes.

Question B.2.b. Déterminer l'écart-type de la variable aléatoire Z = 1,1Y.

1. La variable Y a pour écart-type 1 heure.
2. La multiplication d'une variable aléatoire par 1,1 multiplie son écart-type par 1,1 : $\sigma_Z = 1,1 \times 1 = 1,1$ heure.

> ✓ **RÉPONSE À RETENIR**
>
> $\sigma_Z = 1,1$ heure.

Question B.2.c. À partir des caractéristiques de Z, en déduire sa loi de probabilité.

La multiplication d'une variable normale par une constante conserve une loi normale. Les paramètres obtenus sont l'espérance 5,5 heures et l'écart-type 1,1 heure. Ainsi, selon la notation $\mathcal{N}(\mu ; \sigma^2)$, sa variance vaut $1,1^2$.

> ✓ **RÉPONSE À RETENIR**
>
> $Z \sim \mathcal{N}(5,5 ; 1,1^2)$.

Question B.2.d. Calculer P(Z < 6,5) et interpréter le résultat dans le contexte de la durée de recharge après cinq ans.

Après cinq ans, Z suit une loi normale d'espérance 5,5 et d'écart-type 1,1. Le seuil réduit vaut $u = (6,5 - 5,5)/1,1 \approx 0,91$. La table donne $P(U \leq 0,91) = 0,8186$.

> Environ 82 % des recharges après cinq ans durent moins de 6,5 heures.

✓ RÉPONSE À RETENIR

$P(Z < 6,5) \approx 0,819$, soit environ 81,9 %.

4. Test du Khi2 d’indépendance entre fournisseur et zone géographique

Une enquête répartit 2 400 bornes entre trois fournisseurs de maintenance et trois zones géographiques. Le QCM porte sur le test du Khi2 d’indépendance, son hypothèse nulle, ses degrés de liberté, le risque choisi et la conclusion à partir des valeurs observée et critique.

QCM

Question 1. Dans l’Annexe B, cocher l’énoncé correspondant à l’hypothèse nulle H0 pour le test visant à déterminer si les variables Fournisseur et Zone géographique sont indépendantes.

Dans un test du Khi2 d’indépendance, H_0 pose l’absence d’association entre les deux variables : la répartition des fournisseurs ne dépend pas de la zone géographique. Cela ne signifie pas que les effectifs observés dans chaque case soient égaux : les proportions peuvent varier dans l’échantillon. Parmi les quatre formulations proposées, cochez la première, qui énonce directement l’indépendance.

Annexe B : hypothèse nulle

✓	H_0 : « les variables Fournisseur et Zone géographique sont indépendantes »
	H_0 : « les variables Fournisseur et Zone géographique sont dépendantes »
	H_0 : « le nombre de bornes par zone est identique pour tous les fournisseurs »
	H_0 : « le nombre de bornes par zone n’est pas identique pour tous les fournisseurs »

✓ RÉPONSE À RETENIR

H_0 : les variables Fournisseur et Zone géographique sont indépendantes.

Question 2. Dans l’Annexe B, cocher le nombre de degrés de liberté de la loi du Khi2 suivie, sous H0, par la variable aléatoire de décision.

Le tableau de contingence croise 3 fournisseurs et 3 zones géographiques. Pour un test du Khi2 d’indépendance, le nombre de degrés de liberté est le produit du nombre de lignes moins 1 par le nombre de colonnes moins 1 : $(3-1)(3-1)=4$. Il faut donc cocher la réponse 4 dans l’Annexe B. Ce calcul détermine les degrés de liberté de la loi utilisée sous H_0.

Annexe B : degrés de liberté

degrés de liberté est : ✓

	2	✓	4		6		9

✓ RÉPONSE À RETENIR

4 degrés de liberté.

Question 3. Pour un risque α = 0,05, cocher l'énoncé qui décrit la probabilité associée à ce risque dans le cadre du test d'hypothèse.

Le risque
$alpha = 0,05$ correspond au risque de première espèce : rejeter l'hypothèse nulle alors qu'elle est vraie. Il ne s'agit pas de la probabilité d'accepter une hypothèse, ni de celle de rejeter H_1. Dans l'Annexe B, la proposition correspondante est la troisième.

- Événement : rejeter H_0 alors que H_0 est vraie.
- Probabilité : 0,05.

Annexe B : risque du test

	La probabilité d'accepter H_0 alors que H_1 est vraie est 0,05
✓	La probabilité d'accepter H_1 alors que H_1 est vraie est 0,05
	La probabilité de rejeter H_0 alors que H_0 est vraie est 0,05
	La probabilité de rejeter H_1 alors que H_0 est vraie est 0,05

✓ RÉPONSE À RETENIR

Rejeter H_0 alors qu'elle est vraie : probabilité 0,05.

Question 4. Sachant que χ² observé ≈ 94,96 et χ² théorique ≈ 9,49, cocher la conclusion du test au seuil de risque 0,05 concernant l'indépendance entre la répartition des fournisseurs et la zone géographique.

La règle du test consiste à rejeter H_0 lorsque la statistique observée dépasse la valeur critique. Ici, $94,96 > 9,49$: la statistique se situe dans la région de rejet au seuil de 0,05.

1. On rejette l'hypothèse d'indépendance.
2. Les données indiquent une association entre fournisseur et zone géographique.

Il faut cocher la première proposition de l'Annexe B.

Annexe B : conclusion du test

Question 4 : Après calculs, on obtient $\chi_{observé} \approx 94,96$ or on prend la valeur critique : $\chi_{théorique} \approx 9,49$.

✓	Au seuil de risque $0,05$, la répartition des fournisseurs sur le territoire français n'est pas indépendante de la zone géographique
	Au seuil de risque $0,05$, la répartition des fournisseurs sur le territoire français est

✓ RÉPONSE À RETENIR

Au seuil de 0,05, la répartition des fournisseurs n'est pas indépendante de la zone géographique.

Ce qu'il faut retenir de ce sujet (Session 2022)

Bilan de l'épreuve : Épreuve de Traitement de données du BTSA, toutes options, session 2022, code 2022-BTS103-NOR-ME-RE-MA, d'une durée de 180 minutes, sans coefficient imprimé. Le sujet porte sur les bornes publiques de recharge électrique et évalue les ajustements de données, l'estimation d'une proportion, les lois binomiale et normale, ainsi qu'un test du Khi2 d'indépendance.

Notions-clés, formules & règles à maîtriser

- Le coefficient de corrélation mesure l'intensité d'un lien linéaire, mais l'analyse du nuage des résidus aide à repérer un modèle inadapté.
- Avec $U = (X + 2,5)^2$ et $Y = 0,226U + 6,072$, remplacer U par son expression en X pour obtenir la courbe quadratique.
- Pour n essais indépendants ayant chacun une probabilité p de défaut, le nombre de défauts suit une loi binomiale B(n,p).
- Si Y suit une loi normale $N(\mu,\sigma)$, alors aY suit $N(a\mu,|a|\sigma)$ pour tout réel a.

Points de vigilance & attentes du jury

- Conserver les unités : dans l'exercice 1, Y est exprimée en milliers de bornes et chaque rang correspond à un relevé effectué six mois plus tard.
- Ne pas juger un ajustement uniquement sur la corrélation : interpréter aussi la forme et la répartition des résidus.
- Pour une probabilité normale, centrer-réduire avec $(Y - \mu)/\sigma$ et utiliser la symétrie de la loi lorsque le seuil est négatif.
- Pour le test du Khi2, comparer la statistique observée à la valeur critique et formuler la conclusion en lien avec H0.

Questions fréquentes sur cette épreuve (Q&A Examen)

Q : Comment choisir entre un ajustement linéaire et un ajustement quadratique dans cet exercice ?

La corrélation renseigne sur la force d'un lien linéaire, mais ne suffit pas à valider le modèle. Examinez le nuage et les résidus : une structure systématique dans les résidus suggère que la droite décrit mal la tendance.

Q : Comment reconnaître et calculer la loi du nombre de bornes défectueuses ?

On utilise une loi binomiale lorsque le nombre d'essais est fixé, que les essais sont indépendants et qu'ils ont la même probabilité p de défaut. Ici, X compte les défauts parmi 150 bornes, avec n = 150 et p = 0,04.

Q : Comment gérer les annexes et le temps pendant cette épreuve de 180 minutes ?

Répartissez le temps en tenant compte des points et gardez une marge pour vérifier les calculs et les arrondis au millième. Complétez et numérotez les Annexes A et B, puis joignez-les à la copie comme demandé.

Q : Quel critère permet de conclure au test du Khi2 d'indépendance ?

Au seuil de risque choisi, comparez le Khi2 observé à la valeur critique. Si la statistique observée dépasse cette valeur, rejetez H0 d'indépendance ; sinon, les données ne permettent pas de rejeter H0 au seuil retenu.

Sujet officiel associé (PDF) : https://btsagf.fr/assets/annales/btsa-gf/E3 - Communiquer dans des situations et des contextes variés/2022/BTSA GF - E3 - Communiquer dans des situations et des contextes variés - Session 2022 - 1-Sujet.pdf
Ressources BTSA Gestion Forestière : 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.