



Examen Régression Linéaire Master 1

Durée : 1 heure 30 minutes.

Calculatrice autorisée. Une page de notes manuscrites A4 autorisée. Tout autre document interdit.

Les questions ont une unique bonne réponse.

Des points négatifs pourront être affectés à des mauvaises réponses.

Questions de cours

Les questions de cours suivantes portent sur un modèle linéaire gaussien $\mathbf{Y} = \mathbb{X}\beta + \epsilon$ avec constante, où \mathbb{X} est une matrice de taille $n \times p$ de rang p , et où $\epsilon \sim \mathcal{N}_n(0, \sigma^2 \mathbf{I}_n)$. On note $\hat{\beta}$ l'estimateur du maximum de vraisemblance de β et $\hat{\sigma}^2$ l'estimateur sans biais de la variance des bruits σ^2 .

Question 1 Parmi les conditions suivantes, laquelle implique que les composantes de $\hat{\beta}$ sont décorrélées ?

- A Les colonnes de \mathbb{X} sont deux à deux orthogonales.
- B \mathbb{X} est de rang p .
- C Les lignes de \mathbb{X} sont linéairement indépendantes.
- D Les colonnes de \mathbb{X} sont de normes égales à 1.
- E Les colonnes de \mathbb{X} sont linéairement indépendantes.

Question 2 Chaque résidu studentisé r_i^* défini par $r_i^* = \frac{\hat{\epsilon}_i}{\hat{\sigma}\sqrt{1-h_{ii}}}$, où h_{ii} est le $i^{\text{ème}}$ élément diagonal de la matrice chapeau, ne suit pas exactement une loi de Student à $n-p$ degrés de liberté ; pourquoi ?

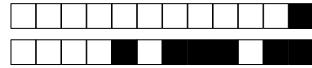
- A Car $\hat{\epsilon}_i$ ne suit pas exactement une loi normale.
- B Car $\hat{\epsilon}_i$ et $\hat{\sigma}$ ne sont pas indépendants.
- C Car $\hat{\epsilon}_i$ suit déjà une loi de Student à $n-p$ degrés de liberté.
- D Car cela dépend de la valeur de h_{ii} .
- E Car r_i^* suit une loi de Student à $n-p-1$ degrés de liberté.

Question 3 Quelle est l'expression de $\text{Cov}(\|\hat{\mathbf{Y}}\|^2, \hat{\sigma}^2)$?

- A $\sigma^2 \text{trace}(\mathbb{X}'\mathbb{X})$.
- B 0.
- C $(n-p)\sigma^2 \|\mathbb{X}\hat{\beta}\|^2$.
- D $\sigma^2 \|\mathbb{X}\beta\|^2$.
- E $\frac{(n-p)\|\mathbb{X}\beta\|^2}{\sigma^2}$.

Question 4 Quelle est la loi de \mathbf{Y} ?

- A $\mathcal{N}_n(\mathbb{X}\beta, \sigma(\mathbb{X}'\mathbb{X})^{-1})$.



- [B] $\mathcal{N}_n(\beta, \sigma^2 \mathbf{I}_n)$.
- [C] $\mathcal{T}(n - p)$.
- [D] $\mathcal{N}_n(\mathbb{X}\beta, \sigma^2 \mathbf{I}_n)$
- [E] $\chi^2(p)$.

Question 5 Quelle est la loi du vecteur des valeurs ajustées $\hat{\mathbf{Y}} = \mathbb{X}\hat{\beta}$?

- [A] $\mathcal{N}_n(\mathbb{X}\beta, \hat{\sigma}^2(\mathbb{X}\mathbb{X}')^{-1})$.
- [B] $\mathcal{N}_n(\mathbb{X}\beta, \sigma^2 \mathbb{X}(\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}')$.
- [C] $\mathcal{N}_n(\mathbb{X}\beta, \sigma^2(\mathbb{X}\mathbb{X}')^{-1})$.
- [D] $\mathcal{N}_n(\mathbb{X}\beta, \hat{\sigma}^2 \mathbb{X}(\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}')$.
- [E] $\mathcal{N}_n(\mathbb{X}\beta, \sigma^2 \mathbb{X}\mathbb{X}')$.

Question 6 Quelle est l'expression de la statistique de Fisher du test de l'hypothèse $(H_0) : \beta = m$ contre $(H_1) : \beta \neq m$.

- [A] $\frac{1}{p\hat{\sigma}^2} \|(\mathbb{X}'\mathbb{X})^{-1}(\hat{\beta} - m)\|^2$.
- [B] $\frac{1}{p\hat{\sigma}^2} \|\mathbb{X}'\mathbb{X}(\hat{\beta} - m)\|^2$.
- [C] $\frac{1}{p\hat{\sigma}^2} \|\mathbb{X}(\hat{\beta} - m)\mathbb{X}'\|^2$.
- [D] $\frac{1}{p\hat{\sigma}^2} \|\hat{\beta} - m\|^2$.
- [E] $\frac{1}{p\hat{\sigma}^2} \|\mathbb{X}(\hat{\beta} - m)\|^2$.

Question 7 Quelle est la loi du vecteur de paramètres β ?

- [A] $\mathcal{N}_p(0, \sigma^2 \mathbf{I}_p)$.
- [B] $\chi^2(p)$.
- [C] $\mathcal{N}_p(\beta, \sigma^2(\mathbb{X}'\mathbb{X})^{-1})$.
- [D] Aucune car β n'est pas aléatoire.
- [E] $\mathcal{T}(n - p)$.

Question 8 Pour la première observation, la valeur du résidu studentisé par validation croisée est $t_i^* = 4.09$. Que peut-on dire de cette observation ?

- [A] Elle n'a pas un effet levier.
- [B] Il s'agit d'une observation aberrante et ayant un effet levier.
- [C] Il s'agit d'une observation aberrante.
- [D] Elle a un effet levier.
- [E] Il ne s'agit pas d'une observation aberrante.

Etude d'un modèle gaussien



Soit le modèle linéaire gaussien $\mathbf{Y} = \mathbb{X}\beta + \epsilon$ où $\epsilon \sim \mathcal{N}_n(0, \sigma^2 \mathbf{I}_n)$ et $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)'$. On suppose que $\mathbb{X} = [\mathbf{1} \mathbf{u} \mathbf{v} \mathbf{w}]$, où $\mathbf{1} \in \mathbb{R}^n$ est le vecteur dont toutes les composantes sont égales à 1, et où \mathbf{u} , \mathbf{v} et \mathbf{w} sont trois vecteurs de \mathbb{R}^n . On suppose également que $\mathbb{X}'\mathbb{X} = \text{diag}(50, 1, 1, 1)$, que $\mathbb{X}'\mathbf{Y} = (100, 4, 3, 27)'$, et que $\|\mathbf{Y}\|^2 = 1032$.

On note $\hat{\beta}$ l'estimateur du maximum de vraisemblance de β , $\hat{\sigma}^2$ l'estimateur sans biais de la variance des bruits, $\hat{\mathbf{Y}} = \mathbb{X}\hat{\beta}$ le vecteur des valeurs ajustées, et R^2 le coefficient de détermination.

On considère également le sous-modèle $\mathbf{Y} = \tilde{\mathbb{X}}\theta + \tilde{\epsilon}$ où $\tilde{\mathbb{X}} = [\mathbf{1} \mathbf{u}]$ et l'on note $\hat{\theta}$ l'estimateur du maximum de vraisemblance de θ .

On donne les quantiles suivants :

Lois de Student : $t_{46}(0.975) = 2.01$.

Lois du χ^2 : $c_{46}(0.025) = 29.160$ et $c_{46}(0.975) = 66.616$.

Lois de Fisher : $f_{1,46}(0.95) = 4.05$, $f_{2,46}(0.95) = 3.20$ et $f_{3,46}(0.95) = 2.80$.

Question 9 On souhaite comparer les deux modèles au moyen d'un test de Fisher. On note $F(\mathbf{Y})^{\text{obs}}$ la valeur prise par la statistique de test. Quelle est l'assertion vraie ?

- [A] $F(\mathbf{Y})^{\text{obs}} = 17.1$ et au niveau 5% l'on accepte l'hypothèse selon laquelle les observations peuvent être issues du sous-modèle.
- [B] $F(\mathbf{Y})^{\text{obs}} = 217.6$ et au niveau 5% l'on rejette l'hypothèse selon laquelle les observations peuvent être issues du sous-modèle.
- [C] $F(\mathbf{Y})^{\text{obs}} = 4.1$ et au niveau 5% l'on accepte l'hypothèse selon laquelle les observations peuvent être issues du sous-modèle.
- [D] $F(\mathbf{Y})^{\text{obs}} = 172.7$ et au niveau 5% l'on rejette l'hypothèse selon laquelle les observations peuvent être issues du sous-modèle.
- [E] $F(\mathbf{Y})^{\text{obs}} = 163.7$ et au niveau 5% l'on rejette l'hypothèse selon laquelle les observations peuvent être issues du sous-modèle.

Question 10 Quelle est la valeur de SCT , la somme des carrés totale ?

- [A] 832.
- [B] 582.
- [C] 982.
- [D] 232.
- [E] 1019.5.

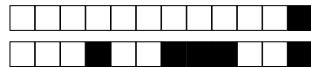
Question 11 Quelle est la valeur du coefficient de détermination R^2 ?

- [A] 0.2131.
- [B] 0.0519.
- [C] 0.7392.
- [D] 0.9358.
- [E] 0.9063.

Question 12 On souhaite tester $(H_0) : \beta_1 = 0$ contre $(H_1) : \beta_1 \neq 0$ au moyen d'un test de Student. On note $T(\mathbf{Y})^{\text{obs}}$ la valeur prise par la statistique de test. Quelle est l'assertion vraie ?

- [A] $T(\mathbf{Y})^{\text{obs}} = 1.85$ et l'on rejette (H_0) au niveau 5%.
- [B] $T(\mathbf{Y})^{\text{obs}} = 1.33$ et l'on accepte (H_0) au niveau 5%.
- [C] $T(\mathbf{Y})^{\text{obs}} = 0.87$ et l'on accepte (H_0) au niveau 5%.
- [D] $T(\mathbf{Y})^{\text{obs}} = 3.07$ et l'on rejette (H_0) au niveau 5%.
- [E] $T(\mathbf{Y})^{\text{obs}} = 0.90$ et l'on accepte (H_0) au niveau 5%.

Question 13 On souhaite tester la validité globale du modèle $\mathbf{Y} = \mathbb{X}\beta + \epsilon$ au moyen d'un test de Fisher. On note $F(\mathbf{Y})^{\text{obs}}$ la valeur prise par la statistique de test. Quelle est l'assertion vraie ?



- A $F(\mathbf{Y})^{\text{obs}} = 43.46$ et l'on considère le modèle gloablement valide au niveau 5%.
- B $F(\mathbf{Y})^{\text{obs}} = 4.15$ et l'on considère le modèle gloablement valide au niveau 5%.
- C $F(\mathbf{Y})^{\text{obs}} = 0.84$ et l'on ne considère pas le modèle globalement valide au niveau 5%.
- D $F(\mathbf{Y})^{\text{obs}} = 148.22$ et l'on considère le modèle globalement valide au niveau 5%.
- E $F(\mathbf{Y})^{\text{obs}} = 223.51$ et l'on considère le modèle globalement valide au niveau 5%.

Question 14 Parmi les intervalles ci-dessous, lequel est un intervalle de confiance de niveau 95% sur la variance des bruits σ^2 ?

- A [6.26; 14.30].
- B [13.63; 31.14].
- C [14.73; 33.64].
- D [3.23; 7.37].
- E [1.17; 2.67].

Question 15 On souhaite tester $(H_0) : \beta_3 = 8\beta_2$ au moyen d'un test de Fisher. On note $F(\mathbf{Y})^{\text{obs}}$ la valeur prise par la statistique de test. Quelle est l'assertion vraie ?

- A $F(\mathbf{Y})^{\text{obs}} = 3.15$ et l'on accepte (H_0) au niveau 5%.
- B $F(\mathbf{Y})^{\text{obs}} = 5.92$ et l'on rejette (H_0) au niveau 5%.
- C $F(\mathbf{Y})^{\text{obs}} = 8.25$ et l'on rejette (H_0) au niveau 5%.
- D $F(\mathbf{Y})^{\text{obs}} = 11.84$ et l'on rejette (H_0) au niveau 5%.
- E $F(\mathbf{Y})^{\text{obs}} = 12.37$ et l'on rejette (H_0) au niveau 5%.

Question 16 Quelle est la valeur de $\|\hat{\mathbf{Y}}\|^2$?

- A 954.
- B 817.
- C 615.
- D 124.
- E 51.

Question 17 Quelle est la valeur de $\hat{\sigma}^2$?

- A 9.0652.
- B 1.6957.
- C 4.6739.
- D 19.7391.
- E 21.3261.

Question 18 Quelle est la valeur de la somme des carrés résiduelles \widetilde{SCR} dans le sous-modèle ?

- A 816.
- B 318.
- C 747.
- D 945.
- E 447.

Question 19 Quelle est la valeur de la statistique du test de Student de nullité du coefficient β_0 ?

- A 10.86.
- B 2.69.
- C 1.54.
- D 21.54.
- E 5.37.

Question 20 Quelle est la valeur de \bar{Y} , la moyenne des composantes de \mathbf{Y} ?

- A 4.
- B 0.5.
- C 1.
- D 3.
- E 2.



Feuille de Réponses

0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9

← Codez votre numéro d'étudiant ci-contre, et inscrivez votre nom et prénom ci-dessous.

Nom et Prénom :

.....

Les réponses aux questions sont à donner exclusivement sur cette feuille. Les réponses données sur les feuilles précédentes ne seront pas prises en compte.

IMPORTANT: *Les cases contenant les réponses aux questions doivent être intégralement remplies.*

Question 1 : A B C D E

Question 2 : A B C D E

Question 3 : A B C D E

Question 4 : A B C D E

Question 5 : A B C D E

Question 6 : A B C D E

Question 7 : A B C D E

Question 8 : A B C D E

Question 9 : A B C D E

Question 10 : A B C D E

Question 11 : A B C D E

Question 12 : A B C D E

Question 13 : A B C D E

Question 14 : A B C D E

Question 15 : A B C D E

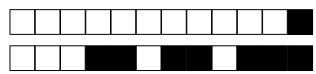
Question 16 : A B C D E

Question 17 : A B C D E

Question 18 : A B C D E

Question 19 : A B C D E

Question 20 : A B C D E



+1/6/55+