Université de Montpellier Année 2023-2024
HAX912X - Modèles linéaires généralisés

Feuille de Travaux Dirigés 1

Compléments sur la régression linéaire multiple

Exercice 1

On considère le modèle

$$y_i = bx_i + u_i, \quad \mathbb{E}[u_i] = 0, \quad \mathbb{E}[u_i^2] = \sigma^2, \quad \mathbb{E}[u_i u_j] = 0$$

où x_i est scalaire. Expliciter l'estimateur des moindres carrés $\widehat{b}$. Soit l'estimateur $\check{b} = \sum y_i / \sum x_i$. Comparer ces deux estimateurs en calculant leur biais et leur variance (On vérifiera que la propriété BLUE s'applique bien : la variance de $\widehat{b}$ est inférieure à celle de $\check{b}$). Sous quelle condition les variances sont-elles égales ?

Exercice 2

On fait une régression de y sur deux variables explicatives x et z, c-à-d $X = (\mathbf{1}, x, z)$; il y a en tout n individus. On a obtenu le résultat suivant :

$$X^T X = \begin{pmatrix} 5 & 3 & 0 \\ 3 & 3 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

1. Que vaut n ? Que vaut le coefficient de corrélation linéaire empirique entre x et z ? (Indication : penser à l'interprétation de chaque entrée de $X^T X$ en fonction des colonnes de X).

La régression linéaire fournit les résultats :

$$y = -1 + 3x + 4z + \widehat{u}, \quad RSS = 3.$$

2. Que vaut la moyenne empirique $\bar{y}$ (on utilisera la matrice $X^T X$) ?
3. Calculer $\|\widehat{y}\|^2$ et en déduire ESS, TSS et le coefficient de détermination R^2.

On s'intéresse au modèle privé du régresseur z :

$$y = X_0 \beta_0 + u_0, \quad X_0 = (\mathbf{1}_n, x).$$

4. Calculer numériquement $X_0^T y$ (commencer par calculer $X^T y$) ; en déduire $\widehat{\beta}_0$.
5. Calculer $\|\widehat{y}_0\|^2$. Démonter que $\|\widehat{u}_0\|^2 + \|\widehat{y}_0\|^2 = \|\widehat{u}\|^2 + \|\widehat{y}\|^2$.

Exercice 3

On considère le modèle de régression

$$y_i = ax_i + u_i, \quad i = 1, \dots N$$

avec : $E[u_i] = 0, \quad \mathbb{V}(u_i) = \sigma_i^2, \quad \mathbb{C}(u_i, u_j) = 0, \quad i \neq j.$
x_i et a sont scalaires.
Donner l'expression des estimateur MCO et MCG de a et comparer leur variance.

Exercice 4

On recueille J séries de mesures de modèle

$$y_{ij} = \mu + u_{ij}, \quad \mathrm{Var}(u_{ij}) = \sigma_i^2, \quad i = 1, \dots n, \quad j = 1, \dots J.$$

Les bruits ont donc des variances différentes dans l'espace des individus.
Donner l'expression de l'estimateur MCG de μ.