Modèle linéaire gaussien-Suite et fin

TD4-MAPI3

2016-2017

Exercice 1

On considère le modèle linéaire avec n observations et p variables explicatives. Pour les hypothèses suivantes, les mettre sous l'une des formes standards vues en cours, donner le type de loi de test, et les valeurs des paramètres.

---

1) L'hypothèse est $\beta_2^* = \beta_3^* = \beta_4^* = 0$ (avec $p > 4$)
2) L'hypothèse est $\beta_2^* = \beta_3^* = \beta_4^*$ (avec $p > 4$)
3) L'hypothèse est $\beta_{2i-1}^* = 2 * \beta_{2i}^*$ pour $i = 1, ..., p/2$ (avec p pair et $i = 1, ..., p/2$).
4) On a $n = 4, p = 3$,

$$X = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & -1 \end{pmatrix}.$$

L'hypothèse est $Y = X^{(0)}\beta^{(0)} + \epsilon$ avec

$$X^{(0)} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}.$$

Par exemple, pour le question 1), la forme standard est de tester $M^t\beta^* = 0$, avec 0_k le vecteur de taille $1 \times k$ dont toutes les composantes sont nulles,

$$M^t = \begin{pmatrix} 0 & 1 & 0 & 0 & 0_{p-4} \\ 0 & 0 & 1 & 0 & 0_{p-4} \\ 0 & 0 & 0 & 1 & 0_{p-4} \end{pmatrix}.$$

Ensuite, la statistique de test suit une loi de Fisher de paramètres 3 et $n - p$.

Exercice 2

On considère le modèle linéaire avec $p = 4$ et où tous les coefficients de la première colonne de X valent 1 (on a inclus un terme *intercept*). Un calcul préliminaire à donné

$$X^tX = \begin{pmatrix} 50 & 0 & 0 & 0 \\ 0 & 20 & 15 & 4 \\ 0 & 15 & 30 & 10 \\ 0 & 4 & 10 & 40 \end{pmatrix}, \quad X^tY = \begin{pmatrix} 100 \\ 50 \\ 40 \\ 80 \end{pmatrix}, \quad Y^tY = 640.$$

On admettra que

$$\begin{pmatrix} 20 & 15 & 4 \\ 15 & 30 & 10 \\ 4 & 10 & 40 \end{pmatrix}^{-1} = \frac{1}{13720} \begin{pmatrix} 1100 & -560 & 30 \\ -560 & 784 & -140 \\ 30 & -140 & 375 \end{pmatrix}.$$

1) Que vaut n?
2) Calculer $\hat{\beta}$ et $\hat{\sigma}$. (On pourra utiliser $||Y - X\hat{\beta}||^2 = ||Y||^2 - ||X\hat{\beta}||^2$)
3) Donner un intervalle de confiance à 95% pour β_2^*.
4) Tester, avec un niveau de test de 95%, l'hypothèse $\beta_2^* = \beta_3^* = \beta_4^* = 0$. On pourra utiliser que le quantile 95% de la loi de Fisher de paramètres 3 et 46 vaut environ 2.80.