MAP433 DM 1 EN 11

I/ Théorie :

1-1) ~~pour~~ Toute solution $\hat{u} \in \arg\min_{u \in \mathbb{R}^p} J_n(u)$

est une solution de l'équation : $\nabla J_n(u) = 0_{(p+1)\times 1}$

on obtient ainsi qu'elle est solution du système d'équations linéaires : $\frac{\partial J_n(u)}{\partial u_k} = 0$ pour $k \in \{0;p\}$

$$\Leftrightarrow \sum_{i=1}^{n} \left\{ Y_i - \sum_{l=0}^{p} u_l Z_i^l \right\} Z_i^k = 0 \text{ pour } k \in \{0,\dots,p\}$$

$$\Leftrightarrow \sum_{i=1}^{n} Y_i Z_i^k = \sum_{i=1}^{n} \left(\sum_{l=0}^{p} u_l Z_i^l \right) Z_i^k \quad k \in \{0,\dots,p\}$$

et donc en considérant les notations, le système d'équations peut s'écrire : $Z^T Y = Z^T Z \hat{u}$

et $\hat{u}$ est certainement solution de ce système.

2) • $Z^\# Z = (Z^T Z)^{-1} Z^T Z = I_p$.

• $Z Z^\# = Z (Z^T Z)^{-1} Z^T$. d'autre part soit $\hat{A}$ le projecteur orthogonale sur l'espace vectoriel engendré par les colonnes de la matrice Z.

Soit x un vecteur et bien sûr Hx son projeté sur notre espace ainsi il existe c un vecteur inconnu tq $Hx = Zc$. d'autre part x et sa projection Hx doivent vérifier la condition d'orthogonalité :

$\langle (x - Hx), Z_i \rangle = 0$ pour toutes les colonnes de Z.

Céci peut se traduit par $Z^T(x - Hx) = 0$

$$\Leftrightarrow Z^T(x - Zc) = 0 \Leftrightarrow Z^T x = Z^T Z c$$

ainsi comme $Z^T Z$ est inversible on a : $(Z^T Z)^{-1} Z^T x = c$

ainsi : $$Hx = Z(Z^T Z)^{-1} Z^T x$$ pour tout vecteur x et donc $H = Z(Z^T Z)^{-1} Z^T = Z Z^{\#}$

Ce qui fallait démontrer.

3) l'estimateur du moindre carré $\hat{\beta}$ ~~est solut~~ vérifie :

$$Z^T Z \hat{\beta} = Z^T Y \Rightarrow \hat{\beta} = (Z^T Z)^{-1} Z^T Y$$

et donc l'estimateur des moindres carrés est unique et a pour expression $\hat{\beta} = Z^{\#} Y$.

4) on a : $$\hat{\beta} = Z^{\#} Y = Z^{\#}(Z\beta + \sigma \mathcal{E}(\theta))$$
$$= \underbrace{Z^{\#} Z}_{I_p} \beta + \sigma Z^{\#} \mathcal{E}(\theta)$$
$$= \beta + \sigma \mathcal{E}(\theta)$$

ainsi $\hat{\beta} = \beta + \sigma Z^{\#} \varepsilon(\theta)$ et donc :

$$\mathbb{E}_\theta(\hat{\beta}) = \beta + \sigma Z^{\#} \underbrace{\mathbb{E}_\theta(\varepsilon(\theta))}_{0} = \beta$$

et donc l'estimateur des moindres carrés $\hat{\beta}$ est un estimateur sans biais de β.

5) soit $\theta \in \Theta$. $\mathrm{Var}_\theta(\hat{\beta}) = \mathrm{Var}_\theta(Z^{\#} Y) = Z^{\#} \underbrace{\mathrm{Var}_\theta(Y)}_{\sigma^2 I_n} Z^{\#^T}$

ainsi $\cancel{\mathrm{Var}_\theta(Z^{\#} Y)} = \mathrm{Var}_\theta(\hat{\beta}) = \sigma^2 (Z^T Z)^{-1} Z^T Z \left[(Z^T Z)^{-1}\right]^T$

$$= \sigma^2 (Z^t Z)^{-1} (Z^t Z) (Z^T Z)^{-1}$$

$$= \sigma^2 (Z^T Z)^{-1}$$

$$\Rightarrow \mathrm{Var}_\theta(\hat{\beta}) = \sigma^2 (Z^T Z)^{-1}$$

6) $\theta \in \Theta$ $\hat{E}_\theta(\tilde{\beta}) = \beta \Leftrightarrow E_\theta(BY) = \beta \Leftrightarrow B E_\theta(Y) = \beta$

$\Leftrightarrow BZ\beta = \beta \Leftrightarrow (BZ - I_p)\beta = 0$ et

donc l'estimateur $\tilde{\beta}$ est sans biais ssi $BZ = I_p$.

7) soit $\theta \in \Theta$. $E_\theta\left((\tilde{\beta} - \beta)(\hat{\beta} - \beta)^T\right) = E_\theta\left[(BY - \beta)(\sigma Z^{\#} \varepsilon(\theta))^T\right]$

or $BY - \beta = B\left[Z\beta + \sigma \varepsilon(\theta)\right] - \beta = \underbrace{BZ}_{I_p}\beta + \sigma B\varepsilon(\theta) - \beta$

$$= \sigma B \varepsilon(\theta).$$

alors $\mathrm{Cov}_\theta(\tilde{\beta}, \hat{\beta}) = E_\theta\left(\sigma^2 B \varepsilon(\theta) \varepsilon(\theta)^T Z^{\#^T}\right)$

et donc $Cov_\theta(\hat{\beta}, \tilde{\beta}) = \sigma^2 B \underbrace{var_\sigma(\mathcal{E}(\sigma))}_{I_n} (Z^{\#})^T = \sigma^2 B (Z^{\#})^T$

ainsi $$Cov_\theta(\tilde{\beta}, \hat{\beta}) = \sigma^2 B (Z^{\#})^T = \sigma^2 B (Z^T Z)^{-1}$$

8) on a sous P_θ :

$$var_\theta(\tilde{\beta}) = \cancel{var_\theta(\tilde{\beta} - \hat{\beta}) + var_\theta(} \; var_\theta(\tilde{\beta} - \hat{\beta} + \hat{\beta})$$

$$= var_\theta(\tilde{\beta} - \hat{\beta}) + var_\theta(\hat{\beta}) + 2cov_\theta(\tilde{\beta} - \hat{\beta}, \hat{\beta})$$

avec $Cov_\theta(\tilde{\beta} - \hat{\beta}, \hat{\beta}) = \underbrace{Cov(\tilde{\beta}, \hat{\beta})}_{var_\theta(\hat{\beta}) = \sigma^2 (Z^T Z)^{-1}} - var_\theta(-\hat{\beta})$

et donc $Cov_\theta(\tilde{\beta} - \hat{\beta}, \hat{\beta}) = 0.$

ainsi $var_\theta(\tilde{\beta}) = var_\theta(\hat{\beta}) + var_\theta(\tilde{\beta} - \hat{\beta})$

$\iff var_\theta(\tilde{\beta}) - var_\theta(\hat{\beta}) = var_\theta(\tilde{\beta} - \hat{\beta})$

or comme $var_\theta(\tilde{\beta} - \hat{\beta})$ est une matrice symetrique positive alors : $\forall x \in \mathbb{R}^p$ $\quad {}^t x \, var_\theta(\tilde{\beta} - \hat{\beta}) x \geq 0$

$\iff {}^t x \, var_\theta(\tilde{\beta}) x \geq {}^t x \, var_\theta(\hat{\beta}) x$. Ceci : $\forall \theta$

et donc : $\forall \theta \in \Theta$ $var_\theta(\tilde{\beta}) \geq var_\theta(\hat{\beta})$.

9) on pose $\hat{\varepsilon} = Y - \hat{Y} = Y - Z\hat{\beta} = Y - Z(Z^T Z)^{-1} Z^T Y$

$$= (I_n - H) Y .$$

en nous servant du modèle $Y = Z\beta + \sigma \varepsilon$

on a donc $\hat{\varepsilon} = \sigma (I_n - H) \varepsilon$. ainsi $E_\theta(\hat{\varepsilon}) = \sigma (I_n - H) E_\theta(\varepsilon) = 0$

et $\operatorname{var}(\hat{\varepsilon}) = \sigma^2 (I_n - H) \operatorname{var}(\varepsilon) (I_n - H)^T$

$$= \sigma^2 (I_n - H)(I_n - H)^T = \sigma^2 (I_n - H) .$$

un estimateur naturelle et sans biais de la variance résiduelle, est donnée par $\frac{1}{n} \sum_{i=1}^{n} (\hat{\varepsilon}_i - 0)^2 = \frac{1}{n} \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \frac{1}{n} \|\hat{\varepsilon}\|^2$

or comme $\|\hat{\varepsilon}\|^2$ est un scalaire, nous écrivons que ce scalaire est égale à sa trace, puis en nous servant de la propriété de la trace, nous obtenons:

$$E(\|\hat{\varepsilon}\|^2) = E(tr(\hat{\varepsilon}^t \hat{\varepsilon})) = E(Tr(\hat{\varepsilon}\hat{\varepsilon}^t)) = tr(E[\hat{\varepsilon}\hat{\varepsilon}^t])$$

$$= tr(\sigma^2 (I_n - H)) = \sigma^2 \underbrace{Tr(I_n - H)}_{n-p}$$

$$= \sigma^2 (n-p)$$

et donc $\hat{\sigma}^2 = \dfrac{\|\hat{\varepsilon}\|^2}{n-p} = (n-p)^{-1} SSE = \dfrac{1}{n-p} \|Y - Z\hat{\beta}\|^2$

est un estimateur sans biais de la variance σ^2.

10) on a : $\|Y\|^2 = \|\cancel{\delta} HY + (I-H)Y\|^2$, or H est un projection orthogonal alors $\langle HY, (I-H)Y \rangle = 0$ et donc

$$\|Y\|^2 = \underbrace{\|HY\|^2}_{RSS} + \underbrace{\|(I-H)Y\|^2}_{SS\bar{E}} = RSS + SS\bar{E}.$$

1-3)

11) La vraisemblance est donnée par :

$$\theta \longmapsto \prod_{k=1}^{n} \frac{1}{(2\pi)^{1/2}\sigma} e^{-\frac{1}{2\sigma^2}\left[Y_i - Z_i'^{T}\beta\right]^2}$$ et définie sur $\mathbb{R}^p \times \mathbb{R}_+^*$

maximiser la vraisemblance étant équivalent à maximiser son logarithme,

$$L_{n_m}(\theta = (\beta, \sigma^2)) \longmapsto C - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\|Y - Z\beta\|^2$$

est concave. ainsi maximiser la fonction $\beta \mapsto l_m(\theta)$ cette fonction revient à minimiser $\|Y - Z'\beta\|^2$ et donc comme on sait $\|Y - Z'\beta\|^2$ est minimisé pour $\hat{\beta} = \hat{\beta}$ l'estimateur du moindre carrés ainsi : $\forall \sigma \; \ln(\beta, \sigma^2) < \ln(\hat{\beta}, \sigma^2)$

$\forall \sigma \; \ln(\beta, \sigma^2) \leq \ln(\hat{\beta}, \sigma^2)$. considérons maintenant la fonction : $\sigma^2 \longmapsto C - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\|Y - Z'\hat{\beta}\|^2$

cette fonction atteint son maximum en

$$\hat{\sigma}^2 = \frac{1}{n}\|Y - Z'\hat{\beta}\|^2$$

Ainsi : $\forall (\beta, \sigma^2)$ $l_m(\beta, \sigma^2) \leq l_m(\hat{\beta}, \hat{\sigma}^2)$ et donc l'EMV de $\theta = (\beta, \sigma^2)$ est $(\hat{\beta}, \hat{\sigma}^2)$ avec $\hat{\beta}$ l'estimateur des moindres carrés $\hat{\beta} = (Z^T Z)^{-1} Z^T Y$

et $\hat{\sigma}^2 = \frac{1}{n} \|Y - Z \hat{\beta}\|^2$.

12) sous P_θ $\theta \in \Theta$. $\hat{\beta} = (Z^T Z)^{-1} Z^T Y = (Z^T Z)^{-1} Z^T [Z\beta + \sigma G]$

$$= \beta + \underbrace{\sigma (Z^T Z)^{-1} Z^T G}_{\text{vecteur gaussien centré}}$$

$$\mathrm{Cov}\left(\sigma (Z^T Z)^{-1} Z^T G\right) = \sigma^2 (Z^T Z)^{-1} Z^T \underbrace{\mathrm{Cov}(G)}_{I_n} Z (Z^T Z)^{-1}$$

$$= \sigma^2 (Z^T Z)^{-1} Z^T Z (Z^T Z)^{-1}$$

$$= \sigma^2 (Z^T Z)^{-1}$$

ainsi $\hat{\beta}_n \sim \mathcal{N}\left(\beta, \sigma^2 (Z^T Z)^{-1}\right)$.

13) sous P_θ $\theta \in \Theta$. $\|Y - Z\hat{\beta}\|^2 = \|(I_n - H) Y\|^2$

or $Y = Z\hat{\beta} + \sigma G$ ainsi : $\|Y - Z\hat{\beta}\|^2 = \|(I_n - H)(Z\hat{\beta} + \sigma G)\|^2$

or $Z\hat{\beta}$ est la projection sur l'~~ensemble~~ espace engendré par les colonnes de Z ainsi $(I_n - H) Z \hat{\beta} = 0$ et donc $\|(I_n - H) Y\|^2 = \|(I_n - H) G\|^2$

et donc par le théorème de Cochran :

$\|(I_n-H)G\|^2 \sim \chi^2(\underbrace{rg(I_n-H)}))$ et donc

$= n - rg(H)$

$= n - p$

on a : $\hat{\sigma}^2_{EMV} = \frac{1}{n}\|Y - Z\hat{\beta}\|^2 \sim \frac{\sigma^2}{n}\chi^2_{n-p}$

$\Leftrightarrow \boxed{\hat{\sigma}^2_{EMV} \sim \frac{\sigma^2}{n}\chi^2_{n-p}}$

14) par le théorème de Cochran on sait que : $(I_n-H)G$ et HG sont indépendantes. or deux vecteurs gaussiens

nous avons vu que $\hat{\sigma}^2$ est une fonction deterministe de $(I_n-H)G$ $\left[\hat{\sigma}^2 = \frac{\sigma^2}{n}\|(I_n-H)G\|^2\right]$

et ~~$\hat{\beta}$ = H~~ $\hat{\beta}$ une fonction deterministe de HG $\left[\text{car } \hat{\beta} = \beta + \sigma \underbrace{Z^{-1}Z(Z^TZ)^{-1}Z^T}_{H}G\right]$

d'où l'indépendance de $\hat{\sigma}^2$ et de $\hat{\beta}$.

1-4/

15) ~~soit~~ sous P_θ. $x^T \hat{\beta} - x^T \beta = x^T (\hat{\beta} - \beta) = x^T \left(\cancel{\beta} + \sigma (Z^T Z)^{-1} Z^T G \right)$

$= \sigma x^T (Z^T Z)^{-1} Z^T G$.

ainsi $x^T \hat{\beta} - x^T \beta$ est une ~~vecteur~~ variable gaussienne car ~~toute combinaison linéaire des variables gaussiennes est gaussienne~~

G est un vecteur gaussien et donc pour tout $x \in \mathbb{R}^p$ $t_x G$ est une va gaussienne. ici $\sigma x^T (Z^T Z)^{-1} Z^T \in \mathbb{R}^p$ et

donc $x^T \hat{\beta} - x^T \beta$ est une v. a gaussienne centré car $E_\theta(G) = 0$. d'autre part $\mathrm{Cov}_\theta(x^T \hat{\beta} - x^T \beta) = \mathrm{Cov}_\theta\left(\sigma x^T (Z^T Z)^{-1} Z^T G\right)$

$$\Rightarrow \mathrm{Cov}_\theta(x^T \hat{\beta} - x^T \beta) = \sigma^2 x^T (Z^T Z)^{-1} Z^T I_n Z (Z^T Z)^{-1} x$$

$$= \sigma^2 x^T (Z^T Z)^{-1} x$$

ainsi $$\frac{x^T \hat{\beta} - x^T \beta}{\sigma \sqrt{x^T (Z^T Z)^{-1} x}} \sim \mathcal{N}(0,1).$$

d'autre par $\frac{\hat{\sigma}^2}{\sigma^2} \sim \frac{\chi^2_{n-p}}{n-p}$ ainsi en somme on a

$$\frac{x^T \hat{\beta} - x^T \beta}{\frac{\hat{\sigma}}{\sigma} \sigma \sqrt{x^T (Z^T Z)^{-1} x}} = \frac{W}{\sqrt{V/_{n-p}}}$$ avec $W \sim \mathcal{N}(0,1)$ et $V \sim \chi^2(n-p)$

et donc $$\frac{x^T \hat{\beta} - x^T \beta}{\hat{\sigma} \sqrt{x^T (Z^T Z)^{-1} x}} \sim T(n-p).$$

16) soit $q^{n-p}_{\alpha/2}$ et $q^{n-p}_{1-\alpha/2}$ les quantils d'ordre $\alpha/2$ et $1-\alpha/2$ du loi de student à $(n-p)$ degrés $T(n-p)$. ainsi

$$P_\theta\left(q^{n-p}_{\alpha/2} \leq \frac{x^T\hat{\beta} - x^T\beta}{\hat{\delta}\sqrt{x^T(Z^TZ)^{-1}x}} \leq q^{n-p}_{1-\alpha/2} \right) = 1-\alpha$$

$$\Rightarrow P_\theta\left(\hat{\delta}\sqrt{x^T(Z^TZ)^{-1}x}\, q^{n-p}_{\alpha/2} \leq x^T\hat{\beta} - x^T\beta \leq \hat{\delta}\sqrt{x^T(Z^TZ)^{-1}x}\, q^{n-p}_{1-\alpha/2} \right) = 1-\alpha$$

$$\Rightarrow P_\theta\left(x^T\hat{\beta} - \hat{\delta}\sqrt{x^T(Z^TZ)^{-1}x}\, q^{n-p}_{1-\alpha/2} \leq x^T\beta \leq x^T\hat{\beta} - \hat{\delta}\sqrt{x^T(Z^TZ)^{-1}x}\, q^{n-p}_{\alpha/2} \right) = 1-\alpha$$

ainsi l'intervalle $I_\alpha(Z) = \left[x^T\hat{\beta} - \hat{\delta}\, q^{n-p}_{1-\alpha/2}\sqrt{x^T(Z^TZ)^{-1}x} \,;\, x^T\hat{\beta} - \hat{\delta}\, q^{n-p}_{\alpha/2}\sqrt{x^T(Z^TZ)^{-1}x} \right]$

est un intervalle de confiance de niveau $1-\alpha$

17) d'après la question précédente, on peut considérer le test : $\phi_0(Z) = \mathbb{1}_{\{ m_0 \notin I(Z) \}}$.

la taille de ce test est égale à α puisque.

$$P_{\beta^t x=0}\left(0 \notin I(Z)\right) = 1 - P_{\beta^t x=0}\left(0 \in I(Z)\right)$$

$$= 1 - P\left(q^{n-p}_{\alpha/2} \leq \frac{x^T\hat{\beta} - x^T\beta}{\hat{\sigma}\sqrt{x^T(Z^TZ)^{-1}x}} \leq q^{n-p}_{1-\alpha/2}\right)$$

$$= 1-(1-\alpha) = \alpha.$$

18) soit $0<\alpha<\alpha'<1$, or $q^{n-p}_{\alpha/2} < q^{n-p}_{\alpha'/2}$ et $q^{n-p}_{1-\alpha'/2} < q^{n-p}_{1-\alpha/2}$

alors $R_\alpha < R_{\alpha'}$. dans ce contexte : La p-valeur est

$$\bar{\alpha}(Z) = \inf\left\{\alpha \in [0,1] \quad 0 \notin I_\alpha(Z)\right\}.$$

$$= \inf\left\{\alpha \in [0,1] \quad P_{0,\beta^T x=0}\left(\frac{x^T\hat{\beta}}{\hat{\sigma}\sqrt{x^T(Z^TZ)^{-1}x}} \notin \left[q^{T(n-p)}_{\alpha/2}, q^{T(n-p)}_{1-\alpha/2}\right]\right.\right.$$

19) soit A une matrice de taille $q \times p$ de rang $q \leq p$

et donc ~~le~~ ~~rg~~ $Rg\left(A(Z^TZ)^{-1}A^T\right) = q$. d'autre

part $A(\hat{\beta}-\beta)$ est un vecteur gaussien de

moyenne nulle et de variance $A \underbrace{var_\theta(\hat{\beta})}_{\sigma^2(Z^TZ)^{-1}} A^T$

$= \sigma^2 A(Z^TZ)^{-1}A^T.$

et donc $\cancel{\frac{1}{\sigma^2}\{A(Z^TZ)A}$

$$\frac{1}{\sigma^2}\{A(\hat{\beta}-\beta)\}^T\left[A(Z^TZ)^{-1}A^T\right]^{-1}\{A(\hat{\beta}-\beta)\} \sim \chi^2_q$$

d'autre part $\frac{\hat{\sigma}^2}{\sigma^2} \sim \frac{\chi^2_{n-p}}{n-p}$ et donc

$$\frac{\sigma^2}{\hat{\sigma}^2}\frac{1}{\sigma^2}\{A(\hat{\beta}-\beta)\}^T\left[A(Z^TZ)^{-1}A^T\right]^{-1}\{A(\hat{\beta}-\beta)\} \sim \frac{\chi^2_q}{\frac{\chi^2_{n-p}}{n-p}}$$

donc en divisant par q on obtient que

$$\frac{1}{q\hat{\sigma}^2}\{A(\hat{\beta}-\beta)\}^T\left[A(Z^TZ)^{-1}A^T\right]^{-1}\{A(\hat{\beta}-\beta)\} \sim \frac{\chi^2_q/q}{\frac{\chi^2_{n-p}}{n-p}} \sim F(q,n-p)$$

20) on pose $A = \begin{array}{c} \overbrace{\hphantom{1\ 0\ 0 \cdots 0}}^{p \text{ Colonnes}} \\ \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \end{array}$ et

donc
$$A\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$
et donc d'aprés La question 10)) nous avons pour (β_1, β_2) La Région critique suivante :

$$RC_\alpha(\beta_1, \beta_2) = \left\{ \frac{1}{2\hat{\sigma}^2} \left[\hat{\beta}_1 - \beta_1, \hat{\beta}_2 - \beta_2\right] \left[A(Z^T Z)^{-1} A^T\right]^{-1} \begin{bmatrix} \hat{\beta}_1 - \beta_1 \\ \hat{\beta}_2 - \beta_2 \end{bmatrix} \leqslant q_{\alpha/2} \right\}$$

ou $q_{\alpha/2}$ est le quantile d'ordre $\alpha/2$ d'une loi de Fisher admettant $(q=2, n-p) = (2, n-p)$ dégrés de liberté