TD III

Exercice 1. Montrer que la statistique de test F, introduite en cours pour tester la validité d'un modèle emboîté, peut s'écrire

$$F = \frac{n-p}{q}\frac{R^2 - R_0^2}{1-R^2}$$

où R^2 et R_0^2 sont les coefficients de détermination associés respectivement au modèle complet et au modèle emboîté.

Rappel : $R^2 = \dfrac{\|\hat{Y} - \mathbb{1}\bar{y}\|^2}{\|Y - \mathbb{1}\bar{y}\|^2}$ et $R_0^2 = \dfrac{\|\hat{Y}_0 - \mathbb{1}\bar{y}\|^2}{\|Y - \mathbb{1}\bar{y}\|^2}$

n : nbre de données.
p : dimension du modèle total : $rg(X)$
p_0 : dimension du sous-modèle
$q = (p - p_0)$

De plus, on a

$$F = \frac{n-p}{q}\frac{\|\hat{Y} - \hat{Y}_0\|^2}{\|Y - Y\|^2}$$

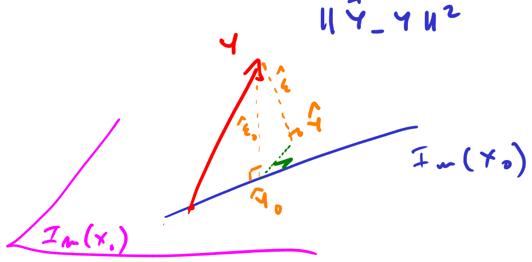
slide 29 CM5

or $R^2 = 1 - \dfrac{SCR}{SCT} = 1 - \dfrac{\|\hat{Y} - Y\|^2}{\|Y - \mathbb{1}\bar{y}\|^2}$ et $1 - R^2 = \dfrac{\|\hat{Y} - Y\|^2}{\|Y - \mathbb{1}\bar{y}\|^2}$

or
$$\frac{R^2 - R_0^2}{1-R^2} = \frac{\|\hat{Y} - \mathbb{1}\bar{y}\|^2 - \|\hat{Y}_0 - \mathbb{1}\bar{y}\|^2}{\cancel{\|Y - \mathbb{1}\bar{y}\|^2}} \times \frac{\cancel{\|Y - \mathbb{1}\bar{y}\|^2}}{\|\hat{Y} - Y\|^2}$$
$$= \frac{\|\hat{Y} - \hat{Y}_0 + \hat{Y}_0 - \bar{y}\mathbb{1}\|^2 - \|\hat{Y}_0 - \mathbb{1}\bar{y}\|^2}{\|\hat{Y} - Y\|^2}$$
$$= \frac{\|\hat{Y} - \hat{Y}_0\|^2 + \cancel{\|\hat{Y}_0 - \bar{y}\mathbb{1}\|^2} - \cancel{\|\hat{Y}_0 - \mathbb{1}\bar{y}\|^2}}{\|\hat{Y} - Y\|^2}$$
$$= \frac{\|\hat{Y} - \hat{Y}_0\|^2}{\|\hat{Y} - Y\|^2}$$

car $\begin{cases} \hat{Y} - \hat{Y}_0 \in Im(X) \cap Im(X_0)^{\perp} \\ \hat{Y}_0 - \mathbb{1}\bar{y} \in Im(X_0) \cap \mathbb{1}^{\perp} \end{cases}$

qui donne $(\hat{Y} - \hat{Y}_0) \perp (\hat{Y}_0 - \mathbb{1}\bar{y})$



Exercice 2. Dans le modèle de régression linéaire, il arrive parfois que l'on souhaite imposer des contraintes linéaires à β, par exemple que sa première coordonnée soit égale à 1. Nous supposerons en général que nous imposons q contraintes linéairement indépendantes à β, ce qui s'écrit sous la forme : $R\beta = r$, où R est une matrice $q \times p$ de rang $q < p$ et r un vecteur de taille q. Montrer que l'estimateur des moindres carrés sous contraintes s'écrit:

$$\hat{\beta}_c = \hat{\beta} + \left(X^t X\right)^{-1} R^t \left[R\left(X^t X\right)^{-1} R^t\right]^{-1} (r - R\hat{\beta}).$$

Calculer $\mathbb{E}\left(\hat{\beta}_c\right)$ et $\mathbb{V}\left(\hat{\beta}_c\right)$

Idée : optimisation sous contrainte $\Rightarrow$ Lagrangien.

on cherche à minimiser la somme des erreurs quadratiques :

$$\begin{cases} S(\beta) = (Y - X\beta)^t (Y - X\beta) \\ \text{s.c.} \quad R\beta - r = 0 \end{cases}$$

Le Lagrangien s'écrit:

$$L(\beta, \lambda) = S(\beta) - \lambda^t (R\beta - r) = (Y - X\beta)^t (Y - X\beta) - \lambda^t (R\beta - r)$$
$$= Y^t Y + \beta^t X^t X \beta - Y^t X \beta - \beta^t X^t Y - \lambda^t (R\beta - r)$$

on cherche un pt selle:

$$\begin{cases} \frac{\partial L}{\partial \beta}(\beta, \lambda) = 0 \\ \frac{\partial L}{\partial \lambda}(\beta, \lambda) = 0 \end{cases} \Leftrightarrow \begin{cases} 2X^t X\beta - 2X^t Y - R^t \lambda \overset{(*)}{=} 0 \quad \text{multipli } R(X^tX)^{-1} \\ R\beta - r = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} 2R\beta - 2R(X^tX)^{-1}X^tY - R(X^tX)^{-1}R^t\lambda = 0 \\ R\beta - r = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} 2r - 2R(X^tX)^{-1}X^tY = R(X^tX)^{-1}R^t\lambda \\ R\beta - r = 0 \end{cases}$$

et on a donc :

$$\lambda = 2\left[R(X^tX)^{-1}R^t\right]^{-1}\left[r - R(X^tX)^{-1}X^tY\right]$$

que l'on ré-injecte dans (*)

$$2(X^tX)\beta - 2X^tY - 2R^t\left[R(X^tX)^{-1}R\right]^{-1}\left[r - R(X^tX)^{-1}X^tY\right] = 0$$

Ce qui donne :

$$\beta = (X^tX)^{-1}X^tY + (X^tX)^{-1}R^t\left[R\,(X^tX)^{-1}R^t\right]^{-1}\left[r - R\,(X^tX)^{-1}X^tY\right]$$

$$= \hat{\beta} + (X^tX)^{-1}R^t\left[R\,(X^tX)^{-1}R^t\right]^{-1}\left[r - R\,\hat{\beta}\right]$$

Ainsi

$$E(\hat{\beta}_c) = E(\hat{\beta}) + (X^tX)^{-1}R^t\left[R\,(X^tX)^{-1}R^t\right]^{-1}\left[r - R\,E(\hat{\beta})\right]$$

$$= \beta + (X^tX)^{-1}R\left[R\,(X^tX)^{-1}R^t\right]^{-1}\left[r - R\beta\right]$$

et

$$Var(\hat{\beta}_c) = Var\left(\hat{\beta} + (X^tX)^{-1}R^t\left[R\,(X^tX)^{-1}R^t\right]^{-1}\left[r - R\,\hat{\beta}\right]\right)$$

$$= Var\left(\left[Id_p + \underbrace{(X^tX)^{-1}R^t\left[R\,(X^tX)^{-1}R^t\right]^{-1}R}_{=B}\right]\hat{\beta}\right)$$

$$= (Id_p - B)\; Var(\hat{\beta})\;(Id_p - B)^t$$

$$= (Id_p - B)\;\sigma^2(X^tX)^{-1}\;(Id_p - B)^t$$

$$= \sigma^2\left[(X^tX)^{-1} - \cancel{B(X^tX)^{-1}} - (X^tX)^{-1}B^t + \cancel{B(X^tX)^{-1}B^t}\right]$$

car $B\,(X^tX)^{-1}B^t = \left((X^tX)^{-1}R^t\left[R\,(X^tX)^{-1}R^t\right]^{-1}R\right)(X^tX)^{-1}\left(R^t\left[R(X^tX)^{-1}R^t\right]^{-1}R\,(X^tX)^{-1}\right)$

on a utilisé le fait de $(AB)^t = B^tA^t$ et que l'on inverse des matrices symétriques

$$= (X^tX)^{-1}R^t\left[R(X^tX)^{-1}R^t\right]^{-1}R\,(X^tX)^{-1} = B\,(X^tX)^{-1}$$

$$= \sigma^2\left[(X^tX)^{-1} - (X^tX)^{-1}R^t\left[R\,(X^tX)^{-1}R^t\right]^{-1}R\,(X^tX)^{-1}\right]$$

De plus l'estimateur $\hat{Y}_c = X\hat{\beta}_c$ a pour propriété :

$$E(\hat{Y}_c) = X\,E(\hat{\beta}_c) = X\beta + X(X^tX)^{-1}R^t\left[R\,(X^tX)^{-1}R^t\right]^{-1}\left[r - R\beta\right]$$

$$Var(\hat{Y}_c) = X\,Var(\hat{\beta}_c)\,X^t = \sigma^2\left[X\,(X^tX)^{-1}X^t - X(X^tX)^{-1}R^t\left[R\,(X^tX)^{-1}R^t\right]^{-1}(X^tX)\,X^t\right]$$

Exercice 3. Modèle de Cobb-Douglas Nous disposons pour n entreprises de la valeur du capital K_i, de l'emploi L_i et de la valeur ajoutée V_i. Nous supposons que la fonction de production de ces entreprises est du type Cobb-Douglas:

$$V_i = \lambda L_i^\alpha K_i^\gamma$$

1. Comment se ramène-t-on à un modèle de régression linéaire?
2. Pour $n = 1658$ entreprises, nous avons obtenu les estimateurs suivants :

$$\hat{\beta} = \begin{pmatrix} 3,136 \\ 0,738 \\ 0,282 \end{pmatrix}$$

avec $R^2 = 0,945$ et $SCR = 148,27$. Nous donnons aussi

$$\left(X^tX\right)^{-1} = \begin{pmatrix} 0,0288 & 0,0012 & -0,0034 \\ 0,0012 & 0,0016 & -0,0010 \\ -0,0034 & -0,0010 & 0,0009 \end{pmatrix} \text{ et } X^tX = \begin{pmatrix} 423 & 2231 & 4077 \\ 2231 & 13808 & 23769 \\ 4077 & 23769 & 42923 \end{pmatrix}$$

Calculer $\hat{\sigma}^2$ et une estimation de $\mathbb{V}(\hat{\beta})$.
3. Donner un intervalle de confiance au niveau 95% pour α. Idem pour γ.
4. Tester au niveau 5% l'hypothèse $H_0 : \gamma = 0$, contre $H_1 : \gamma > 0$.
5. Nous voulons tester l'hypothèse selon laquelle les rendements d'échelle sont constants (une fonction de production F est à rendement d'échelle constant si $\forall \theta \in \mathbb{R}^+, F(\theta L, \theta K) = \theta F(L, K)$). Quelles sont les contraintes verifiées par le modèle lorsque les rendements d'échelle sont constants? Tester au niveau 5% l'hypothèse H_0 : "les rendements sont constants", contre H_1 : "les rendements sont croissants".

1) on passe au log ! les paramètres sont λ, α, γ :

$$\ln(V_i) = \ln(\lambda) + \alpha \ln(L_i) + \gamma \ln(K_i)$$

$$y_i = \beta_0 + \beta_1 \ln(L_i) + \beta_2 \ln(K_i)$$

2) $$\hat{\sigma}^2 = \frac{SCR}{n-p} = \frac{148,27}{1658-3} = \frac{148,27}{1655} \approx 8,96 \cdot 10^{-2}$$

et $$\widehat{Var}(\hat{\beta}) = \hat{\sigma}^2 \left(X^tX\right)^{-1}$$

3) Intervalle de confiance pour α à 95 %

$$\hat{\alpha} \pm 1,96 \sqrt{\widehat{Var}(\hat{\alpha})} = 0,738 \pm 1,96 \sqrt{8,96 \cdot 10^{-2} \times 0,0016}$$
$$= [0,715 \,;\, 0,761]$$

Intervalle de confiance pour γ à 95 %

$$\hat{\gamma} \pm 1,96 \sqrt{\widehat{Var}(\hat{\gamma})} = 0,282 \pm 1,96 \sqrt{8,96 \cdot 10^{-2} \times 0,009}$$
$$= [0,264 \,;\, 0,300]$$

4) La statistique de test est

$$T = \frac{\hat{\gamma}}{\sqrt{\widehat{Var}(\hat{\gamma})}} = \frac{0,282}{\sqrt{8,96.10^{-2} \times 0,009}} \approx 31,4$$

c'est (très) largement supérieur au quantile de niveau 0,95 de la student $T_{1655} = 1,645$

on rejette (H_0). p-valeur ?

5) Condition de rendement d'échelle constant :

$$\forall \theta \in \mathbb{R}^+ \qquad F(\theta L, \theta K) = \theta F(L, K)$$

$$\iff \lambda (\theta L)^{\alpha} (\theta K)^{\gamma} = \theta \lambda L^{\alpha} K^{\gamma}$$

$$\iff \theta^{\alpha+\gamma} = \theta$$

$$\iff \alpha + \gamma = 1$$

on teste donc, au niveau 5%,

$$H_0 : \alpha + \gamma = 1 \qquad vs \qquad H_1 : \alpha + \gamma > 1$$

* Test global de Fisher

on utilise le résultat de l'exercice 2 : avec $R = (0 \ 1 \ 1)$ et $r = 1$

$$\hat{\beta}_c = \hat{\beta} + (X^t X)^{-1} R^t \left[R (X^t X)^{-1} R^t \right]^{-1} (r - R\hat{\beta})$$

avec • $r - R\hat{\beta} = 1 - \hat{\alpha} - \hat{\gamma} = 1 - 0,738 - 0,282 = -0,02$

• $R (X^t X)^{-1} = (0 \ 1 \ 1) (X^t X)^{-1} = (-0,0022 \quad 0,0006 \quad -0,0001)$

$$R (X^t X)^{-1} R^t = 0,0005 \qquad \text{et} \qquad \left[R (X^t X)^{-1} R^t \right]^{-1} = 2000$$

et

$$\hat{\beta}_c = \begin{pmatrix} 3,136 \\ 0,738 \\ 0,282 \end{pmatrix} + \begin{pmatrix} -0,0022 \\ 0,0006 \\ -0,0001 \end{pmatrix} \times 2000 \times (-0,02) = \begin{pmatrix} 3,224 \\ 0,714 \\ 0,286 \end{pmatrix}$$

La statistique de test est $F = \dfrac{n-p}{q} \dfrac{\|\hat{Y} - \hat{Y}_0\|^2}{\|Y - \hat{Y}\|^2}$

- $\|\hat{Y} - \hat{Y}_0\|^2 = \| 3,136 \, \mathbb{1} + 0,738 \, X_1 + 0,282 \, X_2$
$\qquad - 3,224 \, \mathbb{1} - 0,714 \, X_1 - 0,286 \, X_2 \|^2$

$$= \| -0,088 \, \mathbb{1} + 0,024 \, X_1 - 0,004 \, X_2 \|^2$$

$$= (-0,088)^2 n - 2 \times 0,088 \times 0,024 \sum_i x_{1,i} + 2 \times 0,088 \times 0,004 \sum_i x_{2,i}$$
$$+ (0,024)^2 \sum_i x_{1,i}^2 - 2 \times 0,024 \times 0,004 \sum_i x_{1,i} x_{2,i} + (0,004)^2 \sum_i x_{2,i}^2$$

$$= (-0,088)^2 \times 1658 - 2 \times 0,088 \times 0,024 \times 2230,8 + 2 \times 0,088 \times 0,004 \times 4076,9$$
$$+ (0,024)^2 \times 13807,7 - 2 \times 0,024 \times 0,004 \times 23769,2 + (0,004)^2 \times 42923,1$$

$$= 0,128 + 0,942 + 2,870 + 7,953 - 4,564 + 0,687$$

$$\approx 6,132$$

- $\|Y - \hat{Y}\|^2 = \|\hat{\varepsilon}\|^2 = 148,27$

et $\quad F = \dfrac{1655}{1} \times \dfrac{6,132}{148,27} = 68,45$

Cette statistique de test est à comparer au quantile de niveau 0,95 de $F_{q,n-p} = F_{1,1655}$ qui vaut 3,847 $\Rightarrow$ on rejette H_0.

* Test de Student :

Sous (H_0) nous savons que :

$$\frac{\hat{\gamma} + \hat{\alpha} - 1}{\hat{\sigma}_{\hat{\alpha}+\hat{\gamma}}} \sim \mathrm{Stud}_{n-3} = \mathrm{Stud}_{1655}$$

dont le quantile d'ordre 0,55 est 1645. Il suffit de calculer $\hat{\sigma}_{\hat{\alpha}+\hat{\gamma}}$. Comme on a :

$$\mathrm{Var}(\hat{\alpha} + \hat{\gamma}) = \mathrm{Var}(\hat{\alpha}) + \mathrm{Var}(\hat{\gamma}) + 2 \, \mathrm{Cov}(\hat{\alpha}, \hat{\gamma})$$

donc l'estimateur de $\hat{\sigma}_{\hat{\alpha}+\hat{\gamma}} = \sqrt{\widehat{Var}(\hat{\alpha}+\hat{\gamma})}$ où

$$\widehat{Var}(\hat{\alpha}+\hat{\gamma}) = \widehat{Var}(\hat{\alpha}) + \widehat{Var}(\hat{\gamma}) + 2\,Cov(\hat{\alpha},\hat{\gamma})$$

$$= (0{,}0016 + 0{,}0009 - 2 \times 0{,}001) \times \underbrace{8{,}96 \cdot 10^{-2}}_{\hat{\sigma}^2}$$

$$\simeq 4{,}5 \cdot 10^{-5}$$

on en déduit que la stat de test vaut :

$$T(u) = \frac{0{,}738 + 0{,}282 - 1}{\sqrt{4{,}5 \cdot 10^{-5}}} \simeq 2{,}98 > 1{,}645$$

on rejette (H_0)