

Première Année Master M.A.E.F. 2018 – 2019

Econométrie II

Examen final, mai 2019

Examen de 3h00. Tout document ou calculatrice est interdit.

1. **(Sur 22 points)** Pour n et p deux entiers tels que $n \geq p$, on observe $Y = {}^t(Y_i)_{1 \leq i \leq n}$ défini par:

$$Y_i = \theta_1 X_i^{(1)} + \dots + \theta_p X_i^{(p)} + \varepsilon_i \quad \text{pour tout } i = 1, \dots, n, \quad (1)$$

avec une famille connue de réels $(X_i^{(j)})_{1 \leq i \leq n, 1 \leq j \leq p}$, telle que $X = \begin{pmatrix} X_1^{(1)} & X_1^{(2)} & \dots & X_1^{(p)} \\ \vdots & \vdots & & \vdots \\ X_n^{(1)} & X_n^{(2)} & \dots & X_n^{(p)} \end{pmatrix}$ soit une

matrice de rang p , $\theta = {}^t(\theta_j)_{1 \leq j \leq p}$ un vecteur de nombres réels inconnus et $\varepsilon = {}^t(\varepsilon_i)_{1 \leq i \leq n}$, où les (ε_i) sont des v.a.i.i.d. gaussiennes centrées non observées, avec $\text{var}(\varepsilon_1) = \sigma^2 > 0$, inconnue.

Notations: Pour $m \in \mathbf{N}^*$, I_m est la matrice identité de taille m . Pour M une matrice réelle quelconque, tM est la transposée de M . Pour u un vecteur colonne quelconque dans \mathbf{R}^m , $\|u\|^2 = {}^tu u$. Enfin, pour $i \in \{1, \dots, n\}$, on notera $x_i = {}^t(X_i^{(1)}, \dots, X_i^{(p)})$ et $H = (h_{ij})_{1 \leq i, j \leq n} = X ({}^tX X)^{-1} X$. On supposera que $|h_{ii}| < 1$ pour tout $i \in \{1, \dots, n\}$.

- (a) On note $\hat{\theta}$ l'estimateur de θ par moindres carrés ordinaires et $\hat{\sigma}^2$ l'estimateur non biaisé par moindres carrés ordinaires de σ^2 . Rappeler les expressions et les lois de $\hat{\theta}$ et $\hat{\sigma}^2$ **(1 pt)**.
- (b) On note $\hat{Y}_i = {}^t x_i \hat{\theta}$ et $\hat{\varepsilon}_i = Y_i - \hat{Y}_i$. Déterminer la loi de $\hat{\varepsilon}_i$ **(1 pt)**. En déduire la loi asymptotique de $\frac{\hat{\varepsilon}_i}{\hat{\sigma} \sqrt{1 - h_{ii}}}$ **(1 pt)**.
- (c) On va maintenant définir des résidus studentisés de manière à pouvoir travailler pour $n \geq p + 1$ quelconque. Pour $i \in \{1, \dots, n\}$ fixé, on notera $X_{(-i)}$ la matrice X privée de la ligne i (donc une matrice de taille $(n - 1) \times p$ que l'on supposera de rang p) et $Y_{(-i)}$ le vecteur Y privé de Y_i (donc de taille $n - 1$).
Montrer que ${}^tX X = {}^tX_{(-i)} X_{(-i)} + x_i {}^t x_i$ **(1.5 pts)**, puis que ${}^tX Y = {}^tX_{(-i)} Y_{(-i)} + x_i Y_i$ **(0.5 pts)**?
- (d) Montrer que $({}^tX X - x_i {}^t x_i)^{-1} = ({}^tX X)^{-1} + \frac{({}^tX X)^{-1} x_i {}^t x_i ({}^tX X)^{-1}}{1 - x_i ({}^tX X)^{-1} x_i}$ **(2 pts)**.
- (e) Montrer que $h_{ii} = x_i ({}^tX X)^{-1} x_i$ **(1 pt)** et $({}^tX_{(-i)} X_{(-i)})^{-1} = ({}^tX X)^{-1} + \frac{({}^tX X)^{-1} x_i {}^t x_i ({}^tX X)^{-1}}{1 - h_{ii}}$ **(0.5 pts)**.
- (f) Soit $\hat{Y}_i^{(-i)}$ la prédiction de Y_i à l'aide de l'estimateur par moindres carrés ordinaires du modèle initial privé de l'observation i . Déduire de ce qui précède que $\hat{Y}_i^{(-i)} = \frac{1}{1 - h_{ii}} \hat{Y}_i - \frac{h_{ii}}{1 - h_{ii}} Y_i$ **(2.5 pts)** et $\hat{\varepsilon}_i = (1 - h_{ii})(Y_i - \hat{Y}_i^{(-i)})$ **(1 pt)**.
- (g) Soit $\hat{\sigma}_{(-i)}^2$ l'estimation par moindres carrés ordinaires non biaisé de σ^2 du modèle initial privé de l'observation i . Montrer que $\hat{\sigma}_{(-i)}^2$ est indépendante de $\hat{\varepsilon}_i$ **(3 pts)**.
- (h) On définit le résidu studentisé par $\tilde{\varepsilon}_i = \frac{\hat{\varepsilon}_i}{\hat{\sigma}_{(-i)} \sqrt{1 - h_{ii}}}$. Déterminer la loi de $\tilde{\varepsilon}_i$ **(1.5 pts)**. A quoi peuvent servir concrètement les $\tilde{\varepsilon}_i$ **(0.5 pts)**?

- (i) Il peut être aussi intéressant d'utiliser les $\hat{Y}_i^{(-i)}$ pour sélectionner le "meilleur" choix de variables de régression. Pour simplifier, on notera m le modèle associé à X et on définit le critère $\text{PRESS}(m) = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i^{(-i)})^2$. Expliquer en quelques mots pourquoi et comment ce critère peut permettre d'optimiser un choix de modèle (**0.5 pts**). Si on suppose que m est le vrai modèle et que m' est un sur-modèle, démontrer que $0 \leq h_{ii}^{(m)} \leq h_{ii}^{(m')}$ pour tout $i \in \{1, \dots, n\}$ (**2 pts**). En déduire que $\mathbb{E}(\text{PRESS}(m)) \leq \mathbb{E}(\text{PRESS}(m'))$ (**2 pts**). Est-ce ce que l'on espérait (**0.5 pts**)?

2. (Sur 10 points) Exercice de TP utilisant le logiciel R

- (a) On s'intéresse à la croissance de rats nourris à base de maïs OGM ou bien de maïs sans OGM (ce sont des données réelles...). On commence donc par étudier l'évolution du poids d'un rat immatriculé B38625. Soit la suite de commandes suivantes:

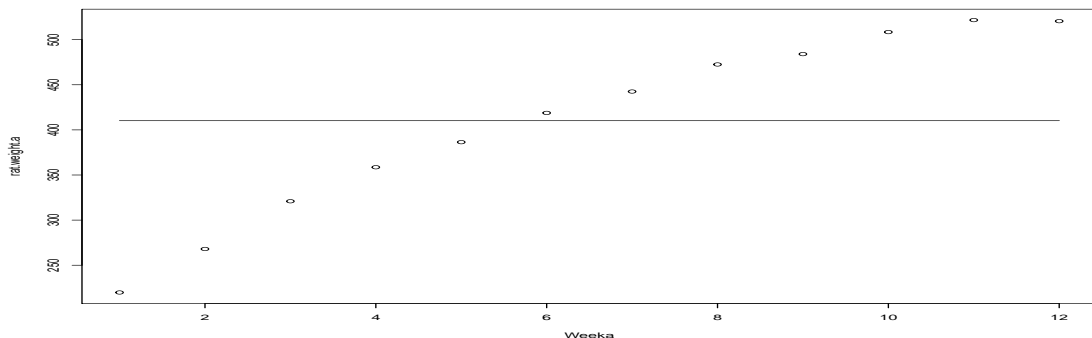
```
data.rat <- read.csv("ratWeight.csv")
rat.weight <- subset(data.rat, id=="B38625")
rat.weight.a= rat.weight$weight[1:12]
rat.weight.t= rat.weight$weight[13:14]
Week=c(1:14); Weeka=c(1:12); Weekt=c(13:14)
lm0=lm(rat.weight.a~1); summary(lm0)
plot(Weeka, rat.weight.a); lines(lm0$fitted.values)
```

Voici une part des résultats obtenus:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	410.18	29.18	14.06	2.25e-08

Residual standard error: 101.1 on 11 degrees of freedom



Questions I.1: Ecrire formellement les différentes variables et le modèle. Que représente la valeur 410.18? A quoi servent les objets `rat.weight.a` et `rat.weight.t`? Que pensez vous des résultats obtenus par la régression? (**2pts**)

- (b) On tape ensuite les commandes (la commande `poly` avec l'option `raw=T` permet une régression polynomiale classique):

```
lm1=lm(rat.weight.a~poly(Weeka,degree=1,raw=T))
lm2=lm(rat.weight.a~poly(Weeka,degree=2,raw=T))
lm3=lm(rat.weight.a~poly(Weeka,degree=3,raw=T))
lm4=lm(rat.weight.a~poly(Weeka,degree=4,raw=T))
lm5=lm(rat.weight.a~poly(Weeka,degree=5,raw=T))
BIC(lm0,lm1,lm2,lm3,lm4,lm5)
```

Voici ce que l'on obtient à la fin des résultats:

	df	BIC
lm0	2	148.75934
lm1	3	115.38075
lm2	4	78.35984
lm3	5	79.89050
lm4	6	76.23188
lm5	7	76.34590

Questions I.2: Qu'a-t-on fait en tapant ces commandes? Que conclure quant aux résultats obtenus? (1pt)

(c) On tape ensuite les commandes:

```
summary(lm2)
par(mfrow = c(1, 2))
plot(lm2, which=c(1,2))
```

Voici ce que l'on obtient à la fin des résultats:

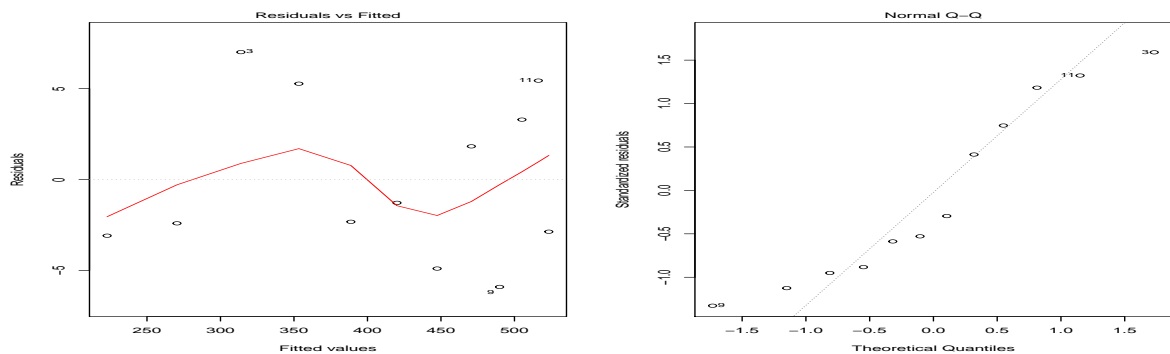
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	171.3341	4.9964	34.29	7.53e-11	***
poly(Weeka, degree = 2, raw = T)1	53.5810	1.7671	30.32	2.26e-10	***
poly(Weeka, degree = 2, raw = T)2	-2.0204	0.1323	-15.27	9.67e-08	***

Residual standard error: 4.834 on 9 degrees of freedom

Multiple R-squared: 0.9981, Adjusted R-squared: 0.9977

F-statistic: 2399 on 2 and 9 DF, p-value: 5.313e-13

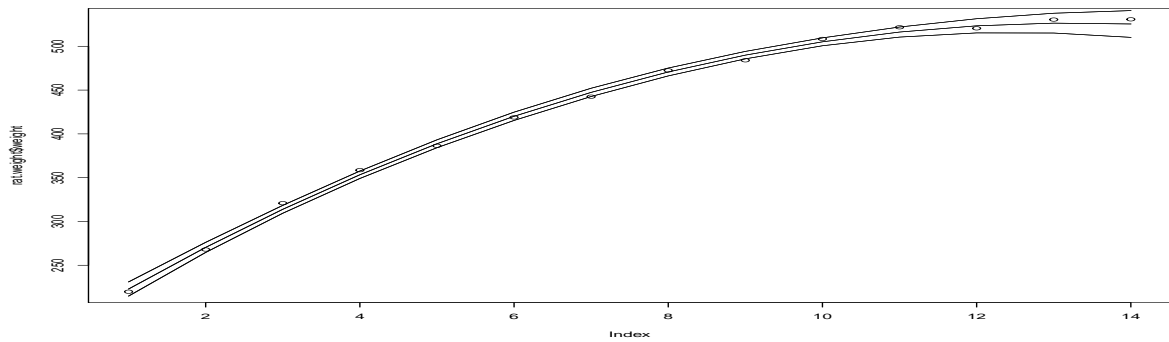


Questions I.3: Qu'a-t-on fait en tapant ces commandes? Ecrire précisément le modèle considéré. Que représente la valeur 30.32? Que pensez-vous des résultats obtenus? (1.5pts)

(d) On continue ainsi:

```
new=data.frame(Weeka=c(1:14))
poidspred=predict(lm2,newdata=new,interval="confidence",level=0.95)
lines(poidspred[1:14,1]); lines(poidspred[1:14,2]); lines(poidspred[1:14,3])
```

Et on obtient:



Questions I.4: Qu'a-t-on fait en tapant ces commandes? Quelles formules permettent d'obtenir les 3 courbes? Conclusion? (1.5pts)

- (e) Enfin, on en vient aux commandes suivantes, où l'on utilise des polynômes orthogonaux (on enlève l'option `raw=T`):

```
lm2.ortho <- lm(weight ~ poly(week, degree=2), data=rat.weight)
summary(lm2.ortho)
M2 <- model.matrix(lm2.ortho)
head(M2); t(M2)%*%M2
```

Et on obtient:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	427.386	1.199	356.36	< 2e-16 ***
poly(week, degree = 2)1	354.841	4.487	79.07	< 2e-16 ***
poly(week, degree = 2)2	-105.409	4.487	-23.49	9.46e-11 ***

Residual standard error: 4.487 on 11 degrees of freedom

Multiple R-squared: 0.9984, Adjusted R-squared: 0.9981

F-statistic: 3402 on 2 and 11 DF, p-value: 4.4e-16

	(Intercept)	poly(week, degree = 2)1	poly(week, degree = 2)2
1	1	-0.43094580	0.48181206
2	1	-0.36464645	0.25943726
3	1	-0.29834709	0.07412493
4	1	-0.23204774	-0.07412493
5	1	-0.16574839	-0.18531233
6	1	-0.09944903	-0.25943726
:	:	:	:

	(Intercept)	poly(week, degree = 2)1	poly(week, degree = 2)2
(Intercept)	1.400000e+01	-1.665335e-16	2.775558e-16
poly(week, degree = 2)1	-1.665335e-16	1.000000e+00	-2.775558e-17
poly(week, degree = 2)2	2.775558e-16	-2.775558e-17	1.000000e+00

Questions I.5: Quel produit scalaire est utilisé ici pour la notion d'orthogonalité? Quelle matrice est donnée par la commande `model.matrix`? Retrouver mathématiquement la valeur -0.43094580 . Que peut-on déduire de la deuxième matrice quant aux coefficients estimés? Quel est l'intérêt de travailler avec ces polynômes orthonormaux? (4pts)