

Econométrie et Applications
1ère session

2013-14

La durée de l'examen est fixée à 2h00. Vous devez répondre aux questions dans l'ordre de l'énoncé. Chaque question nécessite une réponse concise et précise. Une courte explication est généralement nécessaire. Si possible, la notation adoptée au cours doit être utilisée. L'examen est évalué sur 20 points.

1. [5 points] Vrai ou faux. Une justification est toujours nécessaire.
 - (a) Le \bar{R}^2 (ou R^2 ajusté) est nécessairement compris entre 0 et 1.
 - (b) Sous les hypothèses de Gauss-Markov, l'estimateur des MCO est celui qui a la plus petite variance dans la classe des estimateurs non biaisés.
 - (c) Dans le modèle de régression linéaire simple, la somme du produit de la variable expliquée et du résidu est toujours égale à zéro: $\sum_{i=1}^N y_i \hat{u}_i = 0$.
 - (d) La distribution de Student converge vers la distribution normale lorsque le nombre de degrés de liberté tend vers l'infini.
 - (e) Dans le modèle de régression simple, la convergence des estimateurs des MCO peut être démontrée en utilisant la condition $\text{cov}(x, y) = 0$ à la place de $E(u|x) = 0$.
2. [4.5 points] Un chercheur utilise des données sur 27 individus pour étudier la relation entre une variable expliquée y_i et une variable explicative

x_i . Une analyse préliminaire donne les informations suivantes :

$$\begin{aligned} \sum_{i=1}^N y_i &= 405; & \sum_{i=1}^N x_i &= 270; \\ \sum_{i=1}^N (x_i - \bar{x})^2 &= 240; & \sum_{i=1}^N (y_i - \bar{y})^2 &= 300; \\ \sum_{i=1}^N (y_i - \bar{y})(x_i - \bar{x}) &= 120; & \sum_{i=1}^N \hat{u}_i^2 &= 54. \end{aligned}$$

avec $N = 27$. Utilisez ces informations afin de répondre à toutes les questions ci-dessous. Montrez tous vos calculs.

1. Calculez les estimateurs des MCO de la constante β_0 et de la pente β_1 .
 2. Calculez la valeur du R^2 , le coefficient de détermination pour les estimateurs des MCO, et interprétez le.
 3. Calculez l'estimateur de la variance pour l'estimateur du paramètre de la pente β_1 .
- 3. [4,5 points]** Considérez la relation suivante entre le logarithme du salaire et différentes variables explicatives :

$$\begin{aligned} \log(\text{sal}) &= 0,284 + 0,092 \cdot \text{educ} + 0,0041 \cdot \text{exper} - 0,0003 \cdot \text{exper}^2 \\ &\quad \begin{matrix} (0,104) & (0,007) & (0,0017) & (0,0007) \end{matrix} \\ &\quad + 0,022 \cdot \text{anc} \\ &\quad \quad \quad (0,003) \end{aligned}$$

$$N = 1526, R^2 = 0,316$$

où educ désigne le niveau d'éducation en années, exper l'expérience en années et anc l'ancienneté en années. On suppose que les hypothèses classiques sont satisfaites.

1. Testez l'hypothèse que $\beta_{\text{exper}^2} = 0$ (au seuil habituel de 5%) sous l'hypothèse alternative que $\beta_{\text{exper}^2} < 0$. Expliquez la démarche.
2. Testez l'hypothèse que $\beta_{\text{educ}} = 0,1$ (au seuil habituel de 5%) sous l'hypothèse alternative que $\beta_{\text{educ}} \neq 0,1$. Expliquez la démarche.
3. Supposez que l'on calcule la statistique de Fisher permettant de tester l'hypothèse $\beta_{\text{exper}} = \beta_{\text{exper}^2} = 0$. Celle-ci s'élève à 10,3. Quelle est la valeur du R^2 du modèle contraint correspondant?

4. [3 points] Considérez les deux modèles suivants (estimés sur des données individuelles) qui expliquent la durée du sommeil mesuré en heures (repos) en fonction de diverses variables :

$$\widehat{\text{repos}} = 3,638,25 - 0,148 \cdot \text{htrav} - 11,13 \cdot \text{educ} + 2,20 \cdot \text{age} \quad (1)$$

(112,28)
(0,017)
(5,88)
(1,32)

$$n = 45, \quad R^2 = 0.143$$

$$\widehat{\text{repos}} = 3,586,38 - 0,151 \cdot \text{htrav} \quad (2)$$

(38,91)
(0,017)

$$n = 45, \quad R^2 = 0.103$$

où htrav désigne le nombre d'heures de travail hebdomadaire, educ le niveau d'éducation en années et age l'âge en années. On suppose que les hypothèses classiques sont satisfaites.

1. Testez l'hypothèse que $\beta_{\text{educ}} = \beta_{\text{age}} = 0$ contre l'hypothèse alternative $\beta_{\text{educ}} \neq 0$ et/ou $\beta_{\text{age}} \neq 0$ en utilisant un seuil de signification de 5%.
2. Considérez une personne de 20 ans, ayant fait 12 années d'étude et ne travaillant pas. Quelles seront les heures de repos prédites pour cet individu par les deux modèles? Quelle est, selon vous, la plus précise de ces deux prédictions?

5. [3 points] Considérez le modèle de régression simple (avec une seule variable explicative) et supposez que le terme d'erreur est hétéroscédastique.

1. Qu'est-ce que cela signifie? Quelles sont les implications de l'hétéroscédasticité sur le biais des estimateurs des MCO et la validité de l'inférence (la validité des procédures de tests)?
2. Supposez que la variance du terme d'erreur puisse s'écrire:

$$\text{var}(u|x) = \sigma^2 \times h(x)$$

où $h(x)$ est une fonction positive, et considérez le cas où la fonction $h(x)$ est inconnue. Décrivez les différentes étapes permettant d'estimer ce modèle par les Moindres Carrés Pondérés (MCP) ou Moindres Carrés Généralisés "Faisables". Quels sont les avantages des MCP sur les MCO? Quels sont leurs inconvénients?

Critical Values of the F Distribution
($\alpha = .05$)

df within	df between										
	1	2	3	4	5	6	7	8	12	24	∞
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.68	4.53	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.79	2.61	2.41
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.69	2.51	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.20	2.01	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.18	1.98	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.16	1.96	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.12	1.91	1.66
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.00	1.79	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	1.92	1.70	1.39
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	1.88	1.65	1.33
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.85	1.63	1.28
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.83	1.61	1.26
∞	3.84	3.00	2.61	2.37	2.22	2.10	2.01	1.94	1.75	1.52	1.00

Critical Values of the t Distribution

df	2-tailed testing			1-tailed testing		
	••			••		
	0.1	0.05	0.01	0.1	0.05	0.01
5	2.015	2.571	4.032	1.476	2.015	3.365
6	1.943	2.447	3.707	1.440	1.943	3.143
7	1.895	2.365	3.499	1.415	1.895	2.998
8	1.860	2.306	3.355	1.397	1.860	2.896
9	1.833	2.262	3.250	1.383	1.833	2.821
10	1.812	2.228	3.169	1.372	1.812	2.764
11	1.796	2.201	3.106	1.363	1.796	2.718
12	1.782	2.179	3.055	1.356	1.782	2.681
13	1.771	2.160	3.012	1.350	1.771	2.650
14	1.761	2.145	2.977	1.345	1.761	2.624
15	1.753	2.131	2.947	1.341	1.753	2.602
16	1.746	2.120	2.921	1.337	1.746	2.583
17	1.740	2.110	2.898	1.333	1.740	2.567
18	1.734	2.101	2.878	1.330	1.734	2.552
19	1.729	2.093	2.861	1.328	1.729	2.539
20	1.725	2.086	2.845	1.325	1.725	2.528
21	1.721	2.080	2.831	1.323	1.721	2.518
22	1.717	2.074	2.819	1.321	1.717	2.508
23	1.714	2.069	2.807	1.319	1.714	2.500
24	1.711	2.064	2.797	1.318	1.711	2.492
25	1.708	2.060	2.787	1.316	1.708	2.485
26	1.706	2.056	2.779	1.315	1.706	2.479
27	1.703	2.052	2.771	1.314	1.703	2.473
28	1.701	2.048	2.763	1.313	1.701	2.467
29	1.699	2.045	2.756	1.311	1.699	2.462
30	1.697	2.042	2.750	1.310	1.697	2.457
40	1.684	2.021	2.704	1.303	1.684	2.423
50	1.676	2.009	2.678	1.299	1.676	2.403
60	1.671	2.000	2.660	1.296	1.671	2.390
80	1.664	1.990	2.639	1.292	1.664	2.374
100	1.660	1.984	2.626	1.290	1.660	2.364
120	1.658	1.980	2.617	1.289	1.658	2.358
••	1.645	1.960	2.576	1.282	1.645	2.327