



Examen de Grado
Econometría
Agosto, 2017

Duración : 120 minutos
Fecha : 4 de agosto del 2017
Hora de comienzo : 8:30 horas

INSTRUCCIONES

Una vez leído en voz alta por el profesor en la sala, usted dispone de dos horas para responder este examen. No empiece a responder hasta que se le indique hacerlo.

Responda **exactamente 2** preguntas de las 3 preguntas propuestas. Si responde más de 2 preguntas solamente se considerarán las 2 peores respuestas.

Pregunta 1

(40 puntos) Considere el siguiente modelo de regresión múltiple:

$$y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + u_i,$$

donde x_1 es una variable endógena, x_2 es una variable exógena, y asuma que z_1 es una variable (exógena) excluida.

- ¿Qué requisitos debe cumplir z_1 para ser considerada una variable instrumental? Explíquelos, y luego refiérase a cómo probar la relevancia de este instrumento. ¿Por qué es tan importante que estas condiciones se satisfagan? **(7 puntos)**
- Imagine que z_1 es un instrumento válido para x_1 , y que x_2 es una variable inobservable de modo tal que se estima el modelo omitiendo a x_2 . ¿Es el estimador de variables instrumentales un estimador consistente? Justifique su respuesta, y entregue la intuición. **(7 puntos)**
- Imagine ahora que z_1 es un instrumento válido para x_1 y que cuenta con datos para x_2 . Considere que la forma reducida (o primera etapa) correcta para el estimador de variables instrumentales es $\hat{x}_1 = \hat{\pi}_0 + \hat{\pi}_1 z_1 + \hat{\pi}_2 x_2$, pero una vez estimado el econométrista construye los valores ajustados con esos coeficientes pero olvida de incluir x_2 , es decir construye $\hat{x}_1 = \hat{\pi}_0 + \hat{\pi}_1 z_1$. Demuestre que el estimador de MC2E de β_1 será consistente mientras que el de β_2 será inconsistente. Explique la intuición detrás de este resultado. **(14 puntos)**

Ayuda: recuerde que el estimador de MC2E de β_1 es igual a:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \hat{r}_{1,i} y_i}{\sum_{i=1}^n \hat{r}_{1,i}^2},$$

donde $\hat{r}_{1,i}$ son los residuos de la regresión de \hat{x}_1 sobre una constante y x_2 tal que $\hat{r}_{1,i} = \hat{x}_{1,i} - \hat{\alpha}_0 - \hat{\alpha}_1 x_{2,i}$.

- Finalmente, considere el caso en que tanto x_1 como x_2 son variables potencialmente endógenas, y que dispone de los siguientes instrumentos: z_1, z_2 y z_3 . Explique en detalle cómo implementar las pruebas de endogeneidad y de restricciones de sobre-identificación. Refiérase a las hipótesis nula y alternativa en cada caso, así como también a los estadísticos de contraste y su distribución. **(12 puntos)**

Pregunta 2 (40 puntos)

Suponga que Ud. trabaja en una empresa del sector financiero y está interesada en estimar el impacto de una campaña de tasas de interés reducidas en el gross dollar volume (GDV) de las tarjetas de crédito emitidas. Se define GDV como el flujo total de dinero o el valor de las transacciones globales que utilizan las tarjetas de crédito de la compañía en un período determinado. Esta métrica es útil para determinar el tamaño del mercado de la empresa en relación con sus competidores. Suponga que cuenta con variables de los tenedores de tarjetas como sus ingresos ($ingresos_i$) y una variable dummy (d_i) que indica si el tenedor de la tarjeta recibió la promoción o no. Por ahora Ud. desconoce los criterios de asignación de la promoción. Una colega le plantea estimar el siguiente modelo lineal

$$\log GDV_i = \beta_0 + \beta_1 ingresos_i + \beta_2 d_i + u_i$$

- a. Interprete económicamente los parámetros del modelo, específicamente interprete económicamente β_2 y refiérase a los signos esperados. ¿Bajo cuales supuestos respecto del término de error sus estimadores son consistentes? **(5 puntos)**
- b. Se tiene una muestra aleatoria simple de 1200 tenedores de tarjetas para un periodo posterior a la campaña. En el Cuadro 1 se presentan los resultados del modelo de regresión lineal planteado en el enunciado usando dicha muestra. Complete la información faltante e interprete los resultados. ¿Tiene algún efecto la campaña? **(6 puntos)**
- c. Una colega le plantea que la especificación sería incorrecta por cuanto la relación entre el log del GDV y los ingresos sería no lineal. Plantee un modelo alternativo que capture dicha no linealidad usando la información disponible y explique cómo seleccionaría entre el modelo planteado en el enunciado y el modelo propuesto. **(5 puntos)**
- d. Su asistente le entrega en el Cuadro 2 los resultados de un test estadístico acerca de la especificación del modelo estimado en (b). Interprete los resultados indicando la hipótesis nula y alternativa y si rechaza o no la hipótesis nula. **(6 puntos)**
- e. Un analista de riesgo le dice que la promoción de tasas rebajadas habría sido asignada a personas de alto ingreso. Discuta las implicancias de dicha asignación sobre la consistencia de los resultados del modelo estimado en b). **(6 puntos)**
- f. Ahora suponga que por un error de sistema, la asignación de la promoción fue mucho más frecuente en las tarjetas terminadas en los dígitos 9 y 0 y que los números de las tarjetas son asignados aleatoriamente. ¿Cómo podría usar esta información para estimar consistentemente los parámetros del modelo lineal planteado? Plantee una expresión para dicho estimador consistente y los supuestos necesarios. **(6 puntos)**

Cuadro 1: Estimación MCO

```
. reg lgdv ingresos d
```

Source	SS	df	MS	Number of obs	=	1,200
				F(2, 1197)	=	80.09
Model	13.0279129	2	6.51395646	Prob > F	=	0.0000
Residual	97.3496192	1,197	.081328003	R-squared	=	
				Adj R-squared	=	0.1166
Total	110.377532	1,199	.092057992	Root MSE	=	.28518

lgdv	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ingresos		.005615		0.000	.0504285 .0724611
d	.1000559	.0164708		0.000	.0677411 .1323707
_cons	7.480301	.017989	415.83	0.000	7.445008 7.515594

Cuadro 2: Test

```
. estat ovtest
```

Ramsey RESET test using powers of the fitted values of lgdv

Ho:

F(3, 1194) = 1.29
 Prob > F = 0.2758

Cuadro 3: Valores Críticos para una distribución t-Student

n-k	90 %	95 %	97.50 %	99 %	99.50 %
1	3.078	6.314	12.71	31.82	63.66
2	1.886	2.92	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.44	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
.					
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
.					
65	1.295	1.669	1.997	2.385	2.654
78	1.292	1.664	1.991	2.375	2.641
∞	1.281	1.644	1.959	2.326	2.576

Pregunta 3

Usted quiere evaluar el premio por educación para lo cual cuenta con los *salarios en logaritmos* (y) de los individuos, sus *años de educación* (x) y la *cercanía* (z) a su colegio. Esta última es una variable dicotómica que toma valor 1 si el individuo vive cerca del colegio ó 0 en caso contrario, y que además sirve para separar la muestra en 2 grupos. Usted cuenta con la siguiente información para cada uno de los grupos:

Grupo	N	$\sum_{i=1}^N y_i$	$\sum_{i=1}^N x_i$	$\sum_{i=1}^N y_i^2$	$\sum_{i=1}^N x_i^2$	$\sum_{i=1}^N x_i y_i$	SCR modelo simple
Cerca del colegio	380	642	4920	1194	66248	8530	90.5
Lejos del colegio	146	212	1688	341	20792	2529	28.1
Todos	526	854	6608	1535	87040	11059	120.8

Considere los siguientes modelos y métodos de estimación:

M1: estimación MCO del modelo lineal sin controlar por cercanía al colegio.

M2: estimación MCO del modelo lineal controlando además por la cercanía del colegio.

M3: estimación VI del modelo lineal simple usando la cercanía como instrumento.

Ayuda 1: Note que la información relacionada con la variable "cercanía" aparece en la tabla en forma indirecta, por ejemplo: $\sum_{i=1}^N z_i = 380$, $\sum_{i=1}^N y_i z_i = 642$ y $\sum_{i=1}^N x_i y_i z_i = 8530$.

Ayuda 2: Para los siguientes cálculos recuerde que para cualquier par de variables v y w , se cumple que $\sum_{i=1}^N (v_i - \bar{v})(w_i - \bar{w}) = (\sum_{i=1}^N v_i w_i) - \frac{(\sum_{i=1}^N v_i)(\sum_{i=1}^N w_i)}{N}$.

Responda las siguientes preguntas:

- Calcule el premio por educación para la muestra completa usando M1 y comente sobre su significancia estadística y económica. **[6 puntos]**
- Realice un test de Chow para verificar si el premio por educación varía con la cercanía de los individuos a su colegio. **[8 puntos]**
- Discuta sobre la posible endogeneidad de la variable *educación* y proponga un test formal para verificar si esto es efectivo. Usted no debe realizar dicho test, pero sí debe ser detallado en su planteamiento en cuanto a variables utilizadas e hipótesis a verificar. **[6 puntos]**
- Estime la relación entre los años de educación y la cercanía al colegio y discuta si ésta es estadísticamente significativa. **[8 puntos]**
- Calcule el premio por educación para la muestra completa usando el modelo M3 y comente sobre su significancia estadística y económica. **[8 puntos]**
- Si debiera elegir 1 de los 3 modelos propuestos, ¿con cuál se quedaría? ¿Por qué? **[4 puntos]**

DISTRIBUCION T STUDENT

$P(t_{20} > 2.086) = 0.025$

	0.25	0.1	0.05	0.025	0.01	0.005	0.001
1	1.000	3.078	6.314	12.706	31.821	63.656	318.289
2	0.816	1.886	2.920	4.303	6.965	9.925	22.328
3	0.785	1.638	2.353	3.182	4.541	5.841	10.214
4	0.741	1.533	2.132	2.776	3.747	4.604	7.173
5	0.727	1.476	2.015	2.571	3.365	4.032	5.894
6	0.718	1.440	1.943	2.447	3.143	3.707	5.208
7	0.711	1.415	1.895	2.365	2.998	3.499	4.785
8	0.706	1.397	1.860	2.306	2.896	3.355	4.501
9	0.703	1.383	1.833	2.262	2.821	3.250	4.297
10	0.700	1.372	1.812	2.228	2.764	3.169	4.144
11	0.697	1.363	1.796	2.201	2.718	3.106	4.025
12	0.695	1.356	1.782	2.179	2.681	3.055	3.930
13	0.694	1.350	1.771	2.160	2.650	3.012	3.852
14	0.692	1.345	1.761	2.145	2.624	2.977	3.787
15	0.691	1.341	1.753	2.131	2.602	2.947	3.733
16	0.690	1.337	1.746	2.120	2.583	2.921	3.686
17	0.689	1.333	1.740	2.110	2.567	2.898	3.646
18	0.688	1.330	1.734	2.101	2.552	2.878	3.610
19	0.688	1.328	1.729	2.093	2.539	2.861	3.579
20	0.687	1.325	1.725	2.086	2.528	2.845	3.552
21	0.686	1.323	1.721	2.080	2.518	2.831	3.527
22	0.686	1.321	1.717	2.074	2.508	2.819	3.505
23	0.685	1.319	1.714	2.069	2.500	2.807	3.485
24	0.685	1.318	1.711	2.064	2.492	2.797	3.467
25	0.684	1.316	1.708	2.060	2.485	2.787	3.450
26	0.684	1.315	1.706	2.056	2.479	2.779	3.435
27	0.684	1.314	1.703	2.052	2.473	2.771	3.421
28	0.683	1.313	1.701	2.048	2.467	2.763	3.408
29	0.683	1.311	1.699	2.045	2.462	2.756	3.396
30	0.683	1.310	1.697	2.042	2.457	2.750	3.385
40	0.681	1.303	1.684	2.021	2.423	2.704	3.307
60	0.679	1.296	1.671	2.000	2.390	2.660	3.232
120	0.677	1.289	1.658	1.980	2.358	2.617	3.160
10000	0.674	1.282	1.645	1.960	2.326	2.576	3.090

Función de Distribución de la variable F, percentiles 95

	Grados de libertad del numerador																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	1000000
1	161.45	199.5	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.9	245.95	248.02	249.05	250.1	251.14	252.2	253.25	254.32
2	18.51	19	19.16	19.25	19.3	19.33	19.35	19.37	19.38	19.4	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.5
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.7	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6	5.96	5.91	5.86	5.8	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.5	4.46	4.43	4.4	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.1	4.06	4	3.94	3.87	3.84	3.81	3.77	3.74	3.7	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.3	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.5	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.9	2.86	2.83	2.79	2.75	2.71
10	4.96	4.1	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.7	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.2	3.09	3.01	2.95	2.9	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.4
12	4.75	3.89	3.49	3.26	3.11	3	2.91	2.85	2.8	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.3
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.6	2.53	2.46	2.42	2.38	2.34	2.3	2.25	2.21
14	4.6	3.74	3.34	3.11	2.96	2.85	2.76	2.7	2.65	2.6	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.9	2.79	2.71	2.64	2.59	2.54	2.48	2.4	2.33	2.29	2.25	2.2	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.2	2.96	2.81	2.7	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.1	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.9	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.1	2.87	2.71	2.6	2.51	2.45	2.39	2.35	2.28	2.2	2.12	2.08	2.04	1.99	1.95	1.9	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.1	2.05	2.01	1.96	1.92	1.87	1.81
22	4.3	3.44	3.05	2.82	2.66	2.55	2.46	2.4	2.34	2.3	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.8	2.64	2.53	2.44	2.37	2.32	2.27	2.2	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.4	3.01	2.78	2.62	2.51	2.42	2.36	2.3	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.6	2.49	2.4	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.88	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4	3.15	2.76	2.53	2.37	2.25	2.17	2.1	2.04	1.99	1.92	1.84	1.75	1.7	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.5	1.43	1.35	1.25
10000	3.84	3	2.6	2.37	2.21	2.1	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.01