

# ECONOMETRIA

## TERZO PROBLEM SET

### Esercizio 1

a) Partendo da un errore bianco  $u_t \sim iid N(0, \sigma_u^2)$ , costruite un processo  $\varepsilon_t$  che sia autocorrelato di ordine due.

b) Scrivete la formula della  $var(\varepsilon_t)$ .

c) Scrivete la  $cov(\varepsilon_t, \varepsilon_{t-1})$  e  $cov(\varepsilon_t, \varepsilon_{t-s})$ .

d) Scrivete la  $corr(\varepsilon_t, \varepsilon_{t-1})$ .

e) Scrivete la matrice varianza-covarianza di  $\varepsilon_t$  supponendo che  $T = 3$ .

### Esercizio 2

Considerate il processo:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

dove è un processo AR(1)

$$\begin{aligned} \varepsilon_t &= \rho \varepsilon_{t-1} + u_t \text{ con } u_t \sim iid N(0, \sigma_u^2). \\ \text{dove } \rho &= 1 \end{aligned}$$

a) Scrivete la  $var(\varepsilon_t)$ .

b) Scrivete la  $cov(\varepsilon_t, \varepsilon_{t-1})$ .

c) Scrivete la  $corr(\varepsilon_t, \varepsilon_{t-1})$ .

### Esercizio 3

Supponiamo di avere il seguente modello:

$$y_t = \beta_0 + \beta_1 C_t + \varepsilon_t \quad t = 1, \dots, T$$

dove  $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$  con  $u_t \sim iid N(0, \sigma_u^2)$ .

a) Proponete una trasformazione con  $t \geq 2$  in modo tale che l'errore non sia serialmente correlato. Come viene chiamata tale trasformazione?

b) Quale tipo di trasformazione considera anche la prima osservazione? Mostrate come modifichereste la trasformazione precedente.

### Esercizio 4

a) Spiegate a cosa serve il test di Durbin-Watson e scrivete la statistica del test. Nel caso di un campione di 50 osservazioni e con 2 regressori oltre alla costante, come potete commentare il valore del test di Durbin-Watson nei seguenti casi:

- DW=1.20
- DW=1.40
- DW=1.60
- DW=1.70

Usate la tabella qui sotto riportata coi valori critici di dL e dU per il test di Durbin-Watson. (Ipotizziamo come ipotesi alternativa la correlazione positiva).

Observations		X variables, excluding the intercept									
		1		2		3		4		5	
N	Prob.	D-L	D-U	D-L	D-U	D-L	D-U	D-L	D-U	D-L	D-U
15	0.05	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
	0.01	0.81	1.07	0.7	1.25	0.59	1.46	0.49	1.70	0.39	1.96
20	0.05	1.20	1.71	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
	0.01	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
25	0.05	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
	0.01	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
30	0.05	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
	0.01	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
40	0.05	1.44	1.54	1.39	1.60	1.34	1.66	1.39	1.72	1.23	1.79
	0.01	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
50	0.05	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
	0.01	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
60	0.05	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
	0.01	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
80	0.05	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
	0.01	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
100	0.05	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78
	0.01	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

b) Ipotizziamo un modello

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

dove

$$\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2} + \phi_3\varepsilon_{t-3} + u_t$$

Che tipo di processo è  $\varepsilon_t$ ? Che test possiamo usare per verificare il problema di autocorrelazione?

### Esercizio 5

a) Consideriamo la seguente regressione, dove le vendite mensili di giocattoli di un dato paese ( $y$ ) sono regredite sul consumo ( $c$ ).

Nel seguente output, sono presenti anche due test sui residui della regressione. Commentate.

```
. reg y c
```

Source	SS	df	MS	Number of obs =	524
Model	13.8769739	1	13.8769739	F( 1, 522) =	52.88
Residual	136.988471	522	.262430021	Prob > F =	0.0000
				R-squared =	0.0920
				Adj R-squared =	0.0902
Total	150.865445	523	.288461654	Root MSE =	.51228

y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
c	.4882883	.0671484	7.27	0.000	.356374 .6202027
_cons	.0040183	.022384	0.18	0.858	-.0399555 .0479921

```
. estat dwatson
```

Durbin-Watson d-statistic( 2, 524) = 1.702273

```
. estat bgodfrey
```

Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df	Prob > chi2
1	14.561	1	0.0001

H0: no serial correlation

b) Nel seguente output, il modello è stato trasformato con l'uso della stima di Prais-Winsten. Le stime sono molto diverse rispetto a prima? Cosa ci indica la statistica di Durbin-Watson?

```
. prais y c
```

```
Iteration 0: rho = 0.0000
Iteration 1: rho = 0.1488
Iteration 2: rho = 0.1803
Iteration 3: rho = 0.1874
Iteration 4: rho = 0.1891
Iteration 5: rho = 0.1894
Iteration 6: rho = 0.1895
Iteration 7: rho = 0.1895
Iteration 8: rho = 0.1895
Iteration 9: rho = 0.1895
```

```
Prais-Winsten AR(1) regression -- iterated estimates
```

Source	SS	df	MS	Number of obs =	524
Model	6.56420242	1	6.56420242	F( 1, 522) =	25.73
Residual	133.146932	522	.25507075	Prob > F =	0.0000
				R-squared =	0.0470
				Adj R-squared =	0.0452
Total	139.711134	523	.2671341	Root MSE =	.50505

y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
c	.3495857	.068912	5.07	0.000	.2142067 .4849647
_cons	.0049985	.0272145	0.18	0.854	-.0484649 .0584619
rho	.1895324				

```
Durbin-Watson statistic (original) 1.702273
```

```
Durbin-Watson statistic (transformed) 2.007414
```

### Esercizio 6

Consideriamo il modello in cui regrediamo la variabile  $y$  (che rappresenta il salario di un individuo) sulla variabile  $x_1$  (anni di studio) e  $x_2^*$  (abilità individuale):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2^* + u$$

dove la variabile abilità non è osservabile e quindi è misurata con una proxy tale che:

$$x_2^* = \theta_0 + \theta_2 x_2 + v_2$$

a) Cosa rappresentano  $\theta_2$  e  $v_2$ ?

b) Scrivete l'equazione che stimereste in questo caso. (Suggerimento: sostituite la seconda equazione nella prima).

c) Quali sono le condizioni che garantiscono la correttezza degli stimatori OLS del modello al punto (b)? Spiegate.

d) Assumete adesso che:

$$x_2^* = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + v_2$$

Spiegate se gli stimatori OLS sono distorti in questo caso e calcolate la distorsione.

e) Dimostrate che  $Cov(x_2, \varepsilon)$  è diversa da zero, dove  $\varepsilon$  è l'errore del modello stimato al punto (b).

### Esercizio 7

Nel caso in cui :

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

dove  $E(\varepsilon) = 0$ ;  $cov(x_1, \varepsilon) = E(x_1 \varepsilon) \neq 0$

a) Quali sono le condizioni che una variabile strumentale  $z$  deve soddisfare?

b) Dimostrate che lo stimatore IV (di variabili strumentali) è consistente.

c) Cosa accade se  $cov(z, \varepsilon) \neq 0$ . Che problema si verifica?

### Esercizio 8

a) Scrivete lo stimatore IV multivariato nel caso di esatta identificazione e di sovraidentificazione. Spiegate perché è importante la condizione d'ordine.

b) Dimostrate che nel caso  $L=K$

$$\hat{\beta}_{IV} = \hat{\beta}_{2SLS}$$

c) Dimostrate la distorsione di  $\hat{\beta}_{2SLS}, \hat{\beta}_{IV}$

### Esercizio 9

Consideriamo il seguente modello stimato in STATA:

$$lw = \beta_0 + \beta_1 s + \beta_2 tenure + \beta_3 expr + \beta_4 iq + \varepsilon$$

dove  $lw$ = logaritmo dei salari;  $s$ =anni di scuola completati dall'individuo;  $tenure$ =anni di tenure;  $expr$ = anni di esperienza;  $iq$ =indicatore del quoziente di intelligenza.

a) Che problema può esserci nell'uso della variabile  $iq$ ? Come è possibile risolverlo? Osservando la tabella delle variabili disponibili, cercate una soluzione la problema.

```
obs:                758                Wages of Very Young Men, Zvi
                                                Griliches, J.Pol.Ec. 1976
vars:                27                31 Oct 2004 14:12
size:                68,978 (99.3% of memory free)
```

variable name	storage type	display format	value label	variable label
rns	float	%9.0g		residency in South
rns80	float	%9.0g		
mrt	float	%9.0g		marital status = 1 if married
mrt80	float	%9.0g		
smsa	float	%9.0g		reside metro area = 1 if urban
smsa80	float	%9.0g		
med	float	%9.0g		mother's education, years
iq	float	%9.0g		iq score
kww	float	%9.0g		score on knowledge in world of work test
year	float	%9.0g		
age	float	%9.0g		
age80	float	%9.0g		
s	float	%9.0g		completed years of schooling
s80	float	%9.0g		
expr	float	%9.0g		experience, years
expr80	float	%9.0g		
tenure	float	%9.0g		tenure, years
tenure80	float	%9.0g		
lw	float	%9.0g		log wage
lw80	float	%9.0g		
_Iyear_67	byte	%8.0g		year==67
_Iyear_68	byte	%8.0g		year==68
_Iyear_69	byte	%8.0g		year==69
_Iyear_70	byte	%8.0g		year==70
_Iyear_71	byte	%8.0g		year==71
_Iyear_73	byte	%8.0g		year==73
_est_iv	byte	%8.0g		esample() from estimates store

b) Commentate il seguente output di STATA

```
. ivreg lw s tenure expr (iq=kww med ), first
```

First-stage regressions

Source	SS	df	MS	Number of obs = 758		
Model	41951.1677	5	8390.23354	F( 5, 752) =	64.09	
Residual	98448.1581	752	130.915104	Prob > F =	0.0000	
				R-squared =	0.2988	
				Adj R-squared =	0.2941	
Total	140399.326	757	185.468066	Root MSE =	11.442	

  

	iq	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
s		2.480376	.2177002	11.39	0.000	2.053004	2.907749
tenure		.2372269	.2602616	0.91	0.362	-.2736987	.7481525
expr		-.4446049	.2109757	-2.11	0.035	-.8587763	-.0304335
kww		.3100834	.0637504	4.86	0.000	.1849335	.4352334
med		.4114398	.1626099	2.53	0.012	.0922165	.7306632
_cons		55.11403	3.050712	18.07	0.000	49.12511	61.10296

  

Instrumental variables (2SLS) regression

Source	SS	df	MS	Number of obs = 758		
Model	26.0174883	4	6.50437208	F( 4, 753) =	76.06	
Residual	113.268662	753	.150423189	Prob > F =	0.0000	
				R-squared =	0.1868	
				Adj R-squared =	0.1825	
Total	139.28615	757	.183997556	Root MSE =	.38784	

  

	lw	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
iq		.0178162	.0059301	3.00	0.003	.0061746	.0294578
s		.0520467	.019293	2.70	0.007	.0141722	.0899213
tenure		.0297485	.0090933	3.27	0.001	.0118973	.0475997
expr		.0442787	.0074261	5.96	0.000	.0297004	.058857
_cons		3.007415	.3826309	7.86	0.000	2.256265	3.758565

  

Instrumented: iq  
 Instruments: s tenure expr kww med

c) Commentate l'output del test di Sargan. A cosa serve tale test?

```
. overid
```

Tests of overidentifying restrictions:

Sargan N*R-sq test	0.236	Chi-sq(1)	P-value = 0.6271
Basmann test	0.234	Chi-sq(1)	P-value = 0.6284

d) Commentate l'output del test di Hausman. A cosa serve tale test?

. hausman iv .

	---- Coefficients ----			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	iv	.	Difference	S.E.
iq	.0178162	.0038838	.0139323	.0058275
s	.0520467	.0947161	-.0426694	.0180552
tenure	.0297485	.0362904	-.0065419	.0045474
expr	.0442787	.0390324	.0052463	.003694

b = consistent under Ho and Ha; obtained from ivreg  
B = inconsistent under Ha, efficient under Ho; obtained from regress

Test: Ho: difference in coefficients not systematic

chi2(4) = (b-B)'[(V\_b-V\_B)^(-1)](b-B)  
= 5.72  
Prob>chi2 = 0.2214