UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO

Parte A.1

1) Si consideri il modello di regressione $y = \alpha + \beta x + u$, dove α e β sono parametri, x è un regressore e u è un errore di regressione. L'assunzione $E(u \mid x) = 0$ è violata quando:

   a. Esiste collinearità perfetta delle variabili
   b. Esiste collinearità imperfetta ma significativa delle variabili
   c. Viene omesso un secondo regressore rilevante
   d. Viene omessa una variabile esplicativa rilevante correlata con x

Risposta: d

2) Si consideri il modello $y = \alpha + \beta z + u$, dove α e β sono parametri, z è un regressore stocastico e u è un termine di errore. La stima OLS di tale modello è data da $y = \hat{\alpha} + \hat{\beta} z + \hat{u}$, dove $\hat{\alpha}$ e $\hat{\beta}$ sono le stime OLS di α e β. Quale delle seguenti condizioni è essenziale perché questo modello sia un modello di regressione?

   a. $E(\hat{u}\hat{z}) = 0$
   b. $E(\hat{u}z) = 0$
   c. $E(\hat{u}\hat{z}) = 0$
   d. $E(uz) = 0$

Risposta: d, poiché $E(u|z) = 0 => E(uz) = 0$

3) Se la popolazione è descritta da un modello di regressione con due variabili esplicative stocastiche positivamente correlate, omettere una delle due variabili comporta che

   a. Lo stimatore OLS del coefficiente della variabile rimanente diventa positivo.
   b. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all'ingiù (sottostimato) se, in popolazione, il coefficiente della variabile omessa è positivo.
   c. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all'insù (sovrastimato) se, in popolazione, il coefficiente della variabile omessa è positivo.

d. Lo stimatore OLS della variabile rimanente risulterà senz'altro distorto all'ingiù (sottostimato)
e. Lo stimatore OLS della variabile rimanente risulterà senz'altro distorto all'insù (sovrastimato)
f. La media del prodotto dei residui e delle variabili esplicative diventa non nulla

Risposta: c, poiché $corr(x_{1,i},\ x_{2,i}) > 0$ e $\beta_2 > 0$ e dunque $\beta_2 corr(x_{1,i},\ x_{2,i}) > 0$.

4) Se la popolazione è descritta da un modello di regressione con due variabili esplicative stocastiche negativamente correlate, omettere una delle due variabili comporta che

a. Lo stimatore OLS del coefficiente della variabile rimanente diventa positivo.
b. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all'ingiù (sottostimato) se, in popolazione, il coefficiente della variabile omessa è positivo.
c. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all'insù (sovrastimato) se, in popolazione, il coefficiente della variabile omessa è positivo.
d. Lo stimatore OLS della variabile rimanente risulterà senz'altro distorto all'ingiù (sottostimato)
e. Lo stimatore OLS della variabile rimanente risulterà senz'altro distorto all'insù (sovrastimato)
f. La media del prodotto dei residui e delle variabili esplicative diventa non nulla

Risposta: b

5) Se la popolazione è descritta da un modello di regressione con due variabili esplicative stocastiche positivamente correlate, omettere una delle due variabili comporta che

a. Lo stimatore OLS del coefficiente della variabile rimanente diventa positivo.
b. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all'ingiù (sottostimato) se, in popolazione, il coefficiente della variabile omessa è negativo.
c. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all'insù (sovrastimato) se, in popolazione, il coefficiente della variabile omessa è negativo.
d. Lo stimatore OLS della variabile rimanente risulterà senz'altro distorto all'ingiù (sottostimato)
e. Lo stimatore OLS della variabile rimanente risulterà senz'altro distorto all'insù (sovrastimato)
f. La media del prodotto dei residui e delle variabili esplicative diventa non nulla

Risposta: b, poiché $corr(x_{1,i},\ x_{2,i}) > 0$ e $\beta_2 < 0$ e dunque $\beta_2 corr(x_{1,i},\ x_{2,i}) < 0$.

6) Se la popolazione è descritta da un modello di regressione con due variabili esplicative stocastiche negativamente correlate, omettere una delle due variabili comporta che

a. Lo stimatore OLS del coefficiente della variabile rimanente diventa positivo.
b. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all'ingiù (sottostimato) se, in popolazione, il coefficiente della variabile omessa è negativo.
c. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all'insù (sovrastimato) se, in popolazione, il coefficiente della variabile omessa è negativo.
d. Lo stimatore OLS della variabile rimanente risulterà senz'altro distorto all'ingiù (sottostimato)
e. Lo stimatore OLS della variabile rimanente risulterà senz'altro distorto all'insù (sovrastimato)
f. La media del prodotto dei residui e delle variabili esplicative diventa non nulla

Risposta: c

7) Si consideri un insieme di N variabili aleatorie i.i.d., x_i, con valore atteso μ e deviazione standard finita σ, $i \in \{1,2,\dots,N\}$. Il teorema del limite centrale dice che:

a. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso $N\mu$ e deviazione standard $N\sigma$
b. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso μ e deviazione standard σ
c. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso μ e deviazione standard $\sigma/\sqrt{N}$
d. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso $\mu/\sqrt{N}$ e deviazione standard $\sigma/\sqrt{N}$
e. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso μ/N e deviazione standard $\sigma/\sqrt{N}$

Risposta: c

8) Si consideri un insieme di N variabili aleatorie i.i.d., x_i, con valore atteso μ e varianza finita σ, $i \in \{1,2,\dots,N\}$. Il teorema del limite centrale dice che:

a. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso $N\mu$ e deviazione standard $N\sigma$
b. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso μ e deviazione standard σ

c. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso μ e deviazione standard $\sigma/\sqrt{N}$
d. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso μ e deviazione standard $\sqrt{\sigma}/\sqrt{N}$
e. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso μ/N e deviazione standard $\sigma/\sqrt{N}$

Risposta: d

9) Si consideri un insieme di N variabili aleatorie i.i.d., x_i, con valore atteso μ e deviazione standard finita σ, $i \in \{1,2,\dots,N\}$. Il teorema del limite centrale dice che:

a. La media aritmetica delle variabili ha distribuzione normale con valore atteso $N\mu$ e deviazione standard $N\sigma$
b. La media aritmetica delle variabili ha distribuzione normale con valore atteso μ e deviazione standard σ
c. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso μ e deviazione standard σ
d. La media aritmetica delle variabili ha distribuzione normale con valore atteso μ e deviazione standard $\sigma/\sqrt{N}$
e. La media aritmetica delle variabili ha distribuzione asintoticamente normale con valore atteso μ e deviazione standard $\sigma/\sqrt{N}$

Risposta: e

10) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt.\,1{,}80$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0{,}35$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt.\,1{,}90$. Si può rigettare tale ipotesi in un test bilaterale al 5% di significatività?

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

Risposta: a. Spiegazione: La statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}80-1{,}90}{\frac{0{,}35}{\sqrt{100}}} = \frac{-0{,}10}{0{,}035} = -2{,}86.$

Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z > 2{,}86) = \Pr(Z < -2{,}86) = 0{,}002$. Quindi, il valore-p del test a due code è $\Pr(Z > 2{,}86) + \Pr(Z < -2{,}86) = 2 \times 0{,}002 = 0{,}004$. Poiché tale valore non è maggiore di 5%, possiamo rigettare l'ipotesi nulla a questo livello di significatività.

11) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt. 1{,}80$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0{,}35$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt. 1{,}90$. Si può rigettare tale ipotesi in un test bilaterale al 2,5% di significatività?

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

Risposta: a. Spiegazione: la statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}80-1{,}90}{\frac{0{,}35}{\sqrt{100}}} = \frac{-0{,}10}{0{,}035} = -2{,}86$.
Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z > 2{,}86) = \Pr(Z < -2{,}86) = 0{,}002$. Quindi, il valore-p del test a due code è $\Pr(Z > 2{,}86) + \Pr(Z < -2{,}86) = 2 \times 0{,}002 = 0{,}004 = 0{,}4\%$. Poiché tale valore non è maggiore di 2,5%, possiamo rigettare l'ipotesi nulla a questo livello di significatività.

12) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt. 1{,}80$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0{,}35$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt. 1{,}90$. Si può rigettare tale ipotesi in un test bilaterale al livello di significatività dell'1%?

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

Risposta: a. Spiegazione: la statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}80-1{,}90}{\frac{0{,}35}{\sqrt{100}}} = \frac{-0{,}10}{0{,}035} = -2{,}86$.
Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z > 2{,}86) = \Pr(Z < -2{,}86) = 0{,}002$. Quindi, il valore-p del test a due code è $\Pr(Z > 2{,}86) + \Pr(Z < -2{,}86) = 2 \times 0{,}002 = 0{,}004 = 0{,}4\%$. Poiché tale valore non è maggiore di 1%, possiamo rigettare l'ipotesi nulla a questo livello di significatività.

13) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt. 1{,}80$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0{,}35$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt. 1{,}90$. Si può rigettare tale ipotesi in un test bilaterale al livello di significatività dello 0,25%?

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

<u>Risposta</u>: b. Spiegazione: la statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}80-1{,}90}{\frac{0{,}35}{\sqrt{100}}} = \frac{-0{,}10}{0{,}035} = -2{,}86$.
Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z > 2{,}86) = \Pr(Z < -2{,}86) = 0{,}002$. Quindi, il valore-p del test a due code è $\Pr(Z > 2{,}86) + \Pr(Z < -2{,}86) = 2 \times 0{,}002 = 0{,}004 = 0{,}4\%$. Poiché tale valore è <u>maggiore</u> di 0,25%, <u>non possiamo rigettare</u> l'ipotesi nulla a questo livello di significatività.

14) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt.\,1{,}83$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0{,}35$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt.\,1{,}90$. Si può rigettare tale ipotesi in un test bilaterale al livello di significatività del 5%?

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

<u>Risposta</u>: a. Spiegazione: La statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}83-1{,}90}{\frac{0{,}35}{\sqrt{100}}} = \frac{-0{,}07}{0{,}035} = -2{,}00$.
Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z > 2{,}00) = \Pr(Z < -2{,}00) = 0{,}023$. Quindi, il valore-p del test a due code è $\Pr(Z > 2{,}00) + \Pr(Z < -2{,}00) = 2 \times 0{,}023 = 0{,}046 = 4{,}6\%$. Poiché tale valore <u>non è maggiore</u> di 5%, <u>possiamo</u> rigettare l'ipotesi nulla a questo livello di significatività.

15) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt.\,1{,}83$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0{,}35$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt.\,1{,}90$. Si può rigettare tale ipotesi in un test bilaterale al livello di significatività dell'1%?

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

Risposta: b. Spiegazione: La statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1,83-1,90}{\frac{0,35}{\sqrt{100}}} = \frac{-0,07}{0,035} = -2,00$.
Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z > 2,00) = \Pr(Z < -2,00) = 0,023$. Quindi, il valore-p del test a due code è $\Pr(Z > 2,00) + \Pr(Z < -2,00) = 2 \times 0,023 = 0,046 = 4,6\%$. Poiché tale valore è maggiore di 1%, non possiamo rigettare l'ipotesi nulla a questo livello di significatività.

16) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt. 1,83$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0,35$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt. 1,90$. Si può rigettare tale ipotesi in un test unilaterale al livello di significatività dell'1% contro l'ipotesi alternativa che l'altezza media sia minore di mt. 1,90? (ovvero, possiamo rigettare l'ipotesi nulla che l'altezza media sia maggiore di mt. 1,90?)

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

Risposta: b. Spiegazione: La statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1,83-1,90}{\frac{0,35}{\sqrt{100}}} = \frac{-0,07}{0,035} = -2,00$.
Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z < -2,00) = 0,023$. Quindi, il valore-p del test a una coda è $\Pr(Z < -2,00) = 0,023 = 2,3\%$. Poiché tale valore è maggiore di 1%, non possiamo rigettare l'ipotesi nulla a questo livello di significatività.

17) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt. 1,80$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0,35$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt. 1,90$. Si può rigettare tale ipotesi in un test unilaterale al livello di significatività dell'1% contro l'ipotesi alternativa che l'altezza media sia minore di mt. 1,90? (ovvero, possiamo rigettare l'ipotesi nulla che l'altezza media sia maggiore di mt. 1,90?)

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

Risposta: a. Spiegazione: La statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}80-1{,}90}{\frac{0{,}35}{\sqrt{100}}} = \frac{-0{,}10}{0{,}035} = -2{,}86$.
Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z < -2{,}86) = 0{,}002$. Quindi, il valore-p del test a una coda è $\Pr(Z < -2{,}86) = 0{,}002 = 0{,}2\%$. Poiché tale valore è non è maggiore di 1%, possiamo rigettare l'ipotesi nulla a questo livello di significatività.

18) In un ampio campione, che include 1000 studenti, la media delle altezze risulta essere pari a $mt. 1{,}83$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0{,}35$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt. 1{,}90$. Si può rigettare tale ipotesi in un test bilaterale al livello di significatività dell'1%?

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

Risposta: a. Spiegazione: La statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}83-1{,}90}{\frac{0{,}35}{\sqrt{1000}}} = \frac{-0{,}07}{0{,}011} = -6{,}32$.
Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z > 6{,}32) = \Pr(Z < -6{,}32) = 0{,}000$. Quindi, il valore-p del test a due code è $\Pr(Z > 6{,}32) + \Pr(Z < -6{,}32) = 2 \times 0{,}000 = 0{,}000 = 0{,}0\%$. Poiché tale valore non è maggiore di 1%, possiamo rigettare l'ipotesi nulla a questo livello di significatività.

19) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt. 1{,}83$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0{,}25$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt. 1{,}90$. Si può rigettare tale ipotesi in un test bilaterale al livello di significatività dell'1%?

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

Risposta: a. Spiegazione: La statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}83-1{,}90}{\frac{0{,}25}{\sqrt{100}}} = \frac{-0{,}07}{0{,}025} = -2{,}80$.
Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z > 2{,}80) = \Pr(Z < -2{,}80) = 0{,}003$. Quindi, il valore-p del test a due code è $\Pr(Z > 2{,}80) + \Pr(Z < -2{,}80) = 2 \times 0{,}003 = 0{,}006 = 0{,}6\%$. Poiché tale valore non è maggiore di 1%, possiamo rigettare l'ipotesi nulla a questo livello di significatività.

20) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt. 1{,}97$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0{,}35$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt. 1{,}90$. Si può rigettare tale ipotesi in un test bilaterale al livello di significatività del'2,5%?

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

Risposta: b. Spiegazione: La statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}97-1{,}90}{\frac{0{,}35}{\sqrt{100}}} = \frac{0{,}07}{0{,}035} = 2{,}00$.
Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z > 2{,}00) = \Pr(Z < -2{,}00) = 0{,}023$. Quindi, il valore-p del test a due code è $\Pr(Z > 2{,}00) + \Pr(Z < -2{,}00) = 2 \times 0{,}023 = 0{,}046 = 4{,}6\%$. Poiché tale valore è maggiore di 2,5%, non possiamo rigettare l'ipotesi nulla a questo livello di significatività.

21) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt. 1{,}97$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0{,}35$. Si sottoponga a verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt. 1{,}90$. Si può rigettare tale ipotesi in un test unilaterale al livello di significatività del'2,5% contro l'ipotesi alternativa che l'altezza media sia minore di mt. 1,90? (ovvero, possiamo rigettare l'ipotesi nulla che l'altezza media sia maggiore di mt. 1,90?)

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

Risposta: b. Spiegazione: La statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}97-1{,}90}{\frac{0{,}35}{\sqrt{100}}} = \frac{0{,}07}{0{,}035} = 2{,}00$.
Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z < 2{,}00) = 1 - \Pr(Z < -2{,}00) = 1 - 0{,}023 = 0{,}977 = 97{,}7\%$. Quindi, il valore-p del test a una coda è $\Pr(Z < 2{,}00) = 0{,}977 = 97{,}7\%$. Poiché tale valore è (molto!) maggiore di 2,5%, non possiamo rigettare l'ipotesi nulla a questo livello di significatività.

22) In un ampio campione, che include 100 studenti, la media delle altezze risulta essere pari a $mt. 1{,}97$. Si assuma che nella popolazione dalla quale sono tratti tali studenti l'altezza sia distribuita con una deviazione standard pari a $\sigma = 0{,}35$. Si sottoponga a

verifica l'ipotesi nulla che l'altezza media in popolazione degli studenti sia pari a $mt. 1{,}90$. Si può rigettare tale ipotesi in un test unilaterale al livello di significatività del'2,5% contro l'ipotesi alternativa che l'altezza media sia maggiore di mt. 1,90? (ovvero, possiamo rigettare l'ipotesi nulla che l'altezza media sia minore di mt. 1,90?)

a. Si
b. No
c. Non si può dire
d. Si, ma a condizione che si sia esogeneità debole dell'altezza.

<u>Risposta</u>: b. Spiegazione: La statistica per il test è $z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}97-1{,}90}{\frac{0{,}35}{\sqrt{100}}} = \frac{0{,}07}{0{,}035} = 2{,}00$.

Poiché il campione è grande, tale statistica è normalmente distribuita e dunque $\Pr(Z > 2{,}00) = \Pr(Z < -2{,}00) = 0{,}023 = 2{,}3\%$. Quindi, il valore-p del test a una coda è $\Pr(Z > 2{,}00) = 0{,}023 = 2{,}3\%$. Poiché tale valore è <u>minore di</u> 2,5%, <u>possiamo</u> rigettare l'ipotesi nulla a questo livello di significatività.

23) Si consideri il modello di regressione $y = \alpha + \beta x + u$, dove α e β sono parametri, x è un regressore non stocastico e u è un errore di regressione. L'assunzione $E(u \mid x) = 0$ è violata quando:

a. Mai
b. Esiste collinearità imperfetta ma significativa delle variabili
c. Viene omesso un secondo regressore rilevante
d. Viene omessa una variabile esplicativa rilevante

Risposta: a

24) Si consideri il modello $y = \alpha + \beta z + u$, dove α e β sono parametri, z è un regressore stocastico e u è un termine di errore. Si supponga che $E(u) = 0$ ed $E(uz) = 0$. Qual è la conseguenza?

a. La stima OLS di β è non distorta.
b. La stima OLS di β è consistente.
c. Esiste collinearità perfetta delle variabili
d. Esiste collinearità imperfetta ma significativa delle variabili
e. Il modello dato omette uno o più regressori

Risposta: b (quanto asserito in a non è garantito giacché richiede che u e z siano indipendenti, il che non è garantito dall'assenza di correlazione, b è corretta perché z è debolmente esogena e dunque per il CLT lo stimatore OLS è consistente…)

25) Si consideri il modello $y = \alpha + \beta z + u$, dove α e β sono parametri, z è un regressore stocastico e u è un termine di errore. Si supponga che $E(u) = 0$ ed $E(uz) = 0$. Qual è la conseguenza?

a. La stima OLS di β è distorta.
b. La stima OLS di β non è distorta.
c. La stima OLS di β è consistente.
d. La stima OLS di β non è consistente.
e. Esiste collinearità perfetta delle variabili
f. Esiste collinearità imperfetta ma significativa delle variabili

Risposta: c (quanto asserito in a non è garantito giacché u e z potrebbero essere indipendenti, b è sbagliata perché la condizione necessaria è che u e z siano indipendenti che, per quanto non escludere, non si può dare per scontata date le assunzioni fatte…)

26) Se la popolazione è descritta da un modello di regressione lineare con due regressori stocastici, omettere una delle due variabili la quale sia correlata con l'errore comporta che

a. **La stima OLS del coefficiente della variabile inclusa è non distorta.**
b. **La stima OLS del coefficiente della variabile inclusa è distorta ma consistente.**
c. **La stima OLS del coefficiente della variabile inclusa è distorta ed inconsistente.**
d. **Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all’ingiù (sottostimato) se, in popolazione, il coefficiente della variabile omessa è positivo.**
e. **Lo stimatore OLS del coefficiente della variabile rimanente risulterà senz’altro distorto all’ingiù (sottostimato)**
f. **La media del prodotto dei residui e delle variabili esplicative diventa non nulla**

Risposta: c

27) Se la popolazione è descritta da un modello di regressione lineare con due regressori stocastici, omettere una delle due variabili comporta che

a. La stima OLS del coefficiente della variabile inclusa è non distorta.
b. La stima OLS del coefficiente della variabile inclusa è distorta ma consistente.
c. La stima OLS del coefficiente della variabile inclusa è distorta ed inconsistente.
d. La stima OLS del coefficiente della variabile inclusa è distorta ed inconsistente se il regressore omesso è correlato con l'errore.
e. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all’ingiù (sottostimato) se, in popolazione, il coefficiente della variabile omessa è positivo.
f. Lo stimatore OLS del coefficiente della variabile rimanente risulterà senz’altro distorto all’ingiù (sottostimato)

Risposta: d

28) Se la popolazione è descritta da un modello di regressione lineare con due regressori non stocastici, omettere una delle due variabili comporta che

a. La stima OLS del coefficiente della variabile inclusa è non distorta.
b. La stima OLS del coefficiente della variabile inclusa è distorta ma consistente.
c. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all’ingiù (sottostimato) se, in popolazione, il coefficiente della variabile omessa è positivo.
d. Lo stimatore OLS del coefficiente della variabile rimanente risulterà senz’altro distorto all’ingiù (sottostimato)
e. La media del prodotto dei residui e delle variabili esplicative diventa non nulla

Risposta: a

29) Se la popolazione è descritta da un modello di regressione lineare con due regressori stocastici mutualmente indipendenti, omettere una delle due variabili comporta che

a. La stima OLS del coefficiente della variabile inclusa rimane, in generale, non distorta.
b. La stima OLS del coefficiente della variabile inclusa è, in generale, distorta ma consistente.
c. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all'ingiù (sottostimato) se, in popolazione, il coefficiente della variabile omessa è positivo.
d. Lo stimatore OLS del coefficiente della variabile rimanente risulterà senz'altro distorto all'ingiù (sottostimato)
e. La media del prodotto dei residui e delle variabili esplicative diventa non nulla

Risposta: b

30) Se la popolazione è descritta da un modello di regressione lineare con due regressori stocastici mutualmente non correlati, omettere una delle due variabili comporta che

a. La stima OLS del coefficiente della variabile inclusa rimane, in generale, non distorta.
b. La stima OLS del coefficiente della variabile inclusa è, in generale, distorta ma consistente.
c. Lo stimatore OLS del coefficiente della variabile rimanente risulterà distorto all'ingiù (sottostimato) se, in popolazione, il coefficiente della variabile omessa è positivo.
d. Lo stimatore OLS del coefficiente della variabile rimanente risulterà senz'altro distorto all'ingiù (sottostimato)
e. La media del prodotto dei residui e delle variabili esplicative diventa non nulla

Risposta: b

31) Quale delle seguenti affermazioni è vera quando le variabili esplicative (ovvero, le X) sono stocastiche?

a. Lo stimatore OLS non è distorto
b. Lo stimatore OLS è consistente
c. Lo stimatore OLS è non distorto a condizione che le variabili esplicative e gli errori siano indipendentemente distribuiti e con varianza finita
d. Nessuna delle opzioni di cui sopra.

Risposta: c

32) Quale delle seguenti affermazioni è <u>in generale</u> vera quando le variabili esplicative (ovvero, le X) sono stocastiche?

a. Lo stimatore OLS è distorto
b. Lo stimatore OLS è non distorto
c. Lo stimatore OLS è consistente
d. Nessuna delle opzioni di cui sopra.

<u>Risposta</u>: a

33) Quale delle seguenti affermazioni è vera quando le variabili esplicative (ovvero, le X) sono stocastiche?

a. Lo stimatore OLS è distorto in campioni infiniti
b. Lo stimatore OLS è distorto in campioni sia finiti che infiniti
c. Non cambia nulla rispetto al caso in cui le X sono fisse in campionamenti ripetuti
d. Lo stimatore OLS è, <u>in generale</u>, distorto in campioni finiti ma è consistente in campioni infiniti a condizione che le variabili esplicative siano (debolmente) esogene

<u>Risposta</u>: d

34) Quale delle seguenti affermazioni è <u>in generale</u> vera quando le variabili esplicative (ovvero, le X) sono stocastiche?

a. Lo stimatore OLS non è distorto
b. Lo stimatore OLS è consistente
c. Lo stimatore OLS è distorto in campioni finiti ma è consistente in campioni infiniti a condizione che le variabili esplicative siano (debolmente) esogene
d. Nessuna delle opzioni di cui sopra.

<u>Risposta</u>: c

35) Quale delle seguenti affermazioni è vera quando le variabili esplicative (ovvero, le X) sono stocastiche?

a. Lo stimatore OLS non è distorto
b. Lo stimatore OLS è consistente
c. Lo stimatore OLS è distorto in campioni infiniti ma è consistente in campioni finiti a condizione che le variabili esplicative siano (debolmente) esogene
d. Nessuna delle opzioni di cui sopra.

<u>Risposta</u>: d

36) Nella regressione multipla, una conseguenza della multicollinearità imperfetta dei regressori è la seguente

a. nessuna, poiché molte variabili economiche in realtà sono imperfettamente correlate
b. lo stimatore OLS diventa distorto, anche se rimane consistente
c. lo stimatore OLS non è definito e non è calcolabile
d. lo stimatore OLS è definito e calcolabile ma aumentano gli errori standard

Risposta: d

37) Nella regressione multipla, una conseguenza della multicollinearità perfetta dei regressori è la seguente

a. nessuna, poiché molte variabili economiche in realtà sono imperfettamente correlate
b. lo stimatore OLS diventa distorto, anche se rimane consistente
c. lo stimatore OLS non è definito e non è calcolabile
d. lo stimatore OLS è definito e calcolabile ma aumentano gli errori standard

Risposta: c

38) Nella regressione multipla, occorre preoccuparsi della multicollinearità perfetta poiché

a. molte variabili economiche in realtà sono imperfettamente correlate
b. lo stimatore OLS diventa distorto, anche se rimane consistente
c. lo stimatore OLS non è definito e non è calcolabile
d. lo stimatore OLS è definito e calcolabile ma aumentano gli errori standard

Risposta: c

39) Che succede alle stime OLS se, in un modello di regressione per cui valga l'assunzione di esogeneità debole dei regressori, si includono variabili irrilevanti?

a. Possono aumentare gli errori standard delle stime dei coefficienti di quelle rilevanti (stime meno precise)
b. Lo stimatore OLS degli errori standard dei coefficienti del modello di regressione diventa incalcolabile
c. Lo stimatore OLS non è calcolabile
d. Lo stimatore OLS diventa distorto

Risposta: a

40) Che succede alle stime OLS se, in un modello di regressione per cui valga l'assunzione di esogeneità debole dei regressori, si includono variabili irrilevanti?

a. Possono aumentare gli errori standard delle stime dei coefficienti di quelle rilevanti (stime meno precise)
b. Lo stimatore OLS degli errori standard dei coefficienti del modello di regressione diventa incalcolabile
c. Lo stimatore OLS non è calcolabile
d. Lo stimatore OLS diventa non consistente

Risposta: a

41) Che succede alle stime OLS se, in un modello di regressione per cui valga l'assunzione di esogeneità debole dei regressori, si includono variabili irrilevanti?

a. Possono cambiare i livelli di significatività (valori-p) delle stime dei coefficienti delle variabili rilevanti per quanto le stime di tali coefficienti rimangano consistenti
b. Lo stimatore OLS degli errori standard dei coefficienti del modello di regressione diventa incalcolabile
c. Lo stimatore OLS non è calcolabile
d. Lo stimatore OLS diventa distorto

Risposta: a, per via della possibilità che la variabile aggiuntiva irrilevante sia correlata con quella rilevante e dunque gli errori standard risultino sovrastimati, con un conseguente effetto sui valori-p.

42) La distorsione dello stimatore OLS causata dalla presenza di errori di misurazione dei regressori può essere interpretata come un caso speciale di quale dei seguenti problemi?

a. Omissione di variabili rilevanti correlate con le variabili incluse
b. Omissione di variabili irrilevanti correlate con le variabili incluse
c. Selezione campionaria inversa
d. Non si può dire in generale

Risposta: a

43) Supponiamo che mi interessi stimare l'effetto di X su Y, ovvero β in $Y = \beta X + u$. Misuro però X imprecisamente come Z e dunque stimo $Y = \beta Z + e$ dove $Z = X + v$. In tale modello, l'errore sarà

a. $e = -\beta v + u$
b. $e = \beta v + u$
c. $e = -\beta u + v$
d. Non si può dire in generale

Risposta: a, poiché $Y = \beta Z + e = \beta Z + (-\beta v + u) = \beta Z - \beta v + u = \beta(Z - v) + u = \beta X + u$. In maniera del tutto equivalente, $Y = \beta Z + e = Y = \beta(X + v) + e = \beta X + \beta v + e = \beta X + (\beta v + e)$ e dunque $u = \beta v + e$ che implica $e = -\beta v + u$.

44) Il test basato sulla statistica F può essere utilizzato per sottoporre a verifica l'ipotesi nulla che

a. Tutti i coefficienti del modello di regressione siano pari a zero
b. Tutti i coefficienti del modello di regressione, con l'esclusione dell'intercetta, siano pari a zero
c. Tutti i coefficienti siano negativi
d. L'intercetta sia pari a zero

Risposta: b

45) Qual è la distribuzione asintotica di una statistica F con k e $n - k$ gradi di libertà?

a. La distribuzione è F con n gradi di libertà
b. La distribuzione è F con k gradi di libertà
c. La distribuzione è Chi-quadrato con n gradi di libertà
d. La distribuzione è Chi-quadrato con k gradi di libertà

Risposta: d

46) Si consideri il coefficiente di determinazione di un modello di regressione con k regressori (oltre alla costante) stimato usando un campione di n osservazioni, con n relativamente grande. Quale affermazione tra le seguenti è vera con riguardo alla statistica F del test della significatività congiunta dei k regressori?

a. La distribuzione è F con n gradi di libertà
b. La distribuzione è F con k gradi di libertà
c. La distribuzione è Chi-quadrato con n gradi di libertà
d. La distribuzione è Chi-quadrato con k gradi di libertà

Risposta: d

47) Si consideri il coefficiente di determinazione di un modello di regressione con k regressori stimato usando un campione di n osservazioni, con n relativamente grande. Quale affermazione tra le seguenti è vera con riguardo alla statistica F del test della significatività congiunta dei k regressori?

a. La distribuzione è F con n gradi di libertà
b. La distribuzione è F con k gradi di libertà

c. La distribuzione è Chi-quadrato con n gradi di libertà
d. La distribuzione è Chi-quadrato con k gradi di libertà
e. Nessuna delle risposte di cui sopra.

Risposta: e. La ragione è che ci sono solo $k-1$ regressori oltre la costante.

48) Si consideri il coefficiente di determinazione di un modello di regressione con k regressori (oltre alla costante) stimato usando un campione di n osservazioni, con n relativamente piccolo. Quale affermazione tra le seguenti è vera con riguardo alla statistica F del test della significatività congiunta dei k regressori?

a. La distribuzione è F con k ed $n-k$ gradi di libertà
b. La distribuzione è F con k e $k-n$ gradi di libertà
c. La distribuzione è Chi-quadrato con n gradi di libertà
d. La distribuzione è Chi-quadrato con k gradi di libertà

Risposta: a

49) Quale circostanza tra quelle elencate di seguito non inficia la consistenza dello stimatore OLS dei coefficienti di un modello di regressione?

a. Distorsione da variabile omessa
b. “Errors-in-variables”
c. Presenza di eteroschedasticità nel termine di errore
d. Causalità simultanea

Risposta: c

50) Quale circostanza tra quelle elencate di seguito inficia la consistenza dello stimatore OLS della varianza dei coefficienti di un modello di regressione?

a. Distorsione da variabile omessa
b. Presenza di eteroschedasticità nel termine di errore
c. “Errors-in-variables”
d. Causalità simultanea
e. Validità esterna

Risposta: b

51) Come si usano i cosiddetti criteri informativi (come il BIC o l’AIC) nella selezione del modello autoregressivo da utilizzarsi a fini previsivi?

a. Si sceglie il modello col numero di ritardi tale per cui il criterio informativo è più alto.

b. Si sceglie il modello col numero di ritardi tale per cui il criterio informativo è più basso.
c. Si sceglie il modello col numero di ritardi tale per cui il criterio informativo è più alto nel caso del BIC e più basso nel caso dell'AIC.
d. Si sceglie il modello col numero di ritardi tale per cui il criterio informativo è più basso nel caso del BIC e più alto nel caso dell'AIC.

Risposta: b

52) Cosa succede ai cosiddetti criteri informativi (come il BIC o l'AIC) quando vengono introdotte ulteriori variabili esplicative nel modello di regressione (o auto-regresione) considerato?

a. Il criterio aumenta
b. Il criterio diminuisce
c. Non si può dire
d. Il criterio aumenta solo se le variabili sono stocastiche
e. Il criterio diminuisce solo se le variabili sono stocastiche

Risposta: c

53) Cosa succede ai cosiddetti criteri informativi (come il BIC o l'AIC) quando vengono eliminate variabili esplicative da un modello di regressione (o auto-regresione)?

a. Il criterio aumenta
b. Il criterio diminuisce
c. Non si può dire
d. Il criterio aumenta se le variabili sono stocastiche
e. Il criterio diminuisce se le variabili sono stocastiche

Risposta: c

54) Cosa succede al coefficiente di determinazione di un modello di regressione quando vengono introdotte ulteriori variabili esplicative?

a. Il coefficiente aumenta
b. Il coefficiente diminuisce
c. Non si può dire
d. Il coefficiente aumenta se le variabili sono stocastiche
e. Il coefficiente diminuisce se le variabili sono stocastiche

Risposta: a

55) Cosa possiamo aspettarci che succeda al coefficiente di determinazione di un modello di regressione quando vengono eliminate variabili esplicative?

a. Il coefficiente aumenta
b. Il coefficiente diminuisce
c. Non si può dire
d. Il coefficiente aumenta se le variabili sono stocastiche
e. Il coefficiente diminuisce se le variabili sono stocastiche

Risposta: b

56) Cosa possiamo aspettarci che succeda ai cosiddetti criteri informativi (come il BIC o l'AIC) quando vengono introdotte ulteriori variabili esplicative rilevanti nel modello di regressione (o auto-regresione) considerato?

a. Il criterio aumenta
b. Il criterio diminuisce
c. Non si può dire
d. Il criterio aumenta solo se le variabili sono stocastiche
e. Il criterio diminuisce solo se le variabili sono stocastiche

Risposta: b

57) Cosa possiamo aspettarci che succeda ai cosiddetti criteri informativi (come il BIC o l'AIC) quando vengono rimosse variabili esplicative rilevanti del modello di regressione (o auto-regresione) considerato?

a. Il criterio aumenta
b. Il criterio diminuisce
c. Non si può dire
d. Il criterio aumenta solo se le variabili sono stocastiche
e. Il criterio diminuisce solo se le variabili sono stocastiche

Risposta: a

58) Con l'acronimo RMSFE si intende

a. Il "Random Mean Squared Forecast Error" di un modello previsivo
b. Il "Relative Mean Squared Forecast Error" di un modello previsivo al confronto di un'altro
c. Il "Ratio of Mean Squared Forecast Errors" di un modello previsivo al confronto di un'altro
d. Il "Root Mean Squared Forecast Error" di un modello previsivo

Risposta: d

59) Come si può stimare la incertezza che circonda le previsioni di un dato modello previsivo?

a. Si può stimare il RMSFE del modello.
b. Si può fare riferimento all'R^2 del modello stimato
c. Si può fare riferimento all'R^2 "pseudo out of sample" del modello stimato
d. Si può fare riferimento ad un criterio informativo

Risposta: a

60) Si supponga che occorra scegliere tra due modelli previsionali alternativi, ovvero il modello A e il modello B. Si supponga inoltre che la stima dell'RMSFE del modello A sia 5.2% e quella dell'RMSFE del modello B sia 3%. Quale modello risulta preferibile sulla base del criterio del RMSFE?

a. Nessuno poiché la differenza non è maggiore di 5%
b. Il modello A
c. Il modello B
d. Il modello A, a condizione che contenga rotture strutturali

Risposta: c

61) Previsioni "pseudo out of sample" possono essere utilizzate con finalità che escludano

a. Valutare la capacità previsiva effettiva del modello previsivo stimato
b. Stimare il RMSFE.
c. Fare inferenza sulla presenza di radici unitarie
d. Valutare la bontà relativa di due modelli previsivi alternativi

Risposta: c

62) Previsioni "pseudo out of sample" prodotte da modelli ADL possono essere utilizzate con finalità che escludano

a. Fare inferenza sulla presenza di relazioni di causalità simultanea tra le variabili del modello
b. Stimare il RMSFE.
c. Fare inferenza sulla presenza di radici unitarie
d. Valutare la bontà previsiva relativa di due modelli ADL alternativi

Risposta: a (poiché nei modelli ADL i regressori sono ritardati) e c (poiché per quello scopo dobbiamo usare ADF)

63) Previsioni “pseudo out of sample” prodotte da modelli ADL possono essere utilizzate con finalità che includono

a. Fare inferenza sulla presenza di relazioni di causalità simultanea tra le variabili del modello
b. Valutare la bontà previsiva relativa di due modelli ADL alternativi
c. Fare inferenza sulla presenza di radici unitarie
d. Valutare gli errori standard di un modello di regressione

Risposta: b (la a va esclusa poiché nei modelli ADL i regressori sono ritardati)

64) I modelli AR includono

a. Valori correnti e ritardi del termine di errore
b. Ritardi della variabile dipendente e di altre variabili previsive
c. Valori correnti e ritardi dei residui
d. Ritardi della variabile dipendente ma non di altre variabili previsive

Risposta: d

65) I modelli ADL includono, tra le variabili esplicative,

a. Valori correnti e ritardi del termine di errore
b. Ritardi della variabile dipendente e di altre variabili previsive
c. Valori correnti e ritardi dei residui
d. Ritardi della variabile dipendente ma non di altre variabili previsive

Risposta: b

66) I modelli ADL(2,3) includono, tra le variabili esplicative,

a. Tre ritardi della variabile dipendente e due ritardi di una variabile esogena
b. Due ritardi della variabile dipendente e tre ritardi di una variabile esogena
c. Tre ritardi della variabile dipendente e due ritardi del termine di errore
d. Due ritardi della variabile dipendente e tre ritardi del termine di errore

Risposta: b

67) Quale tra i seguenti modelli è un modello AR(1)?

a. Un modello con un ritardo di una variabile esogena
b. Un modello con un ritardo della variabile dipendente e un ritardo di una variabile esogena
c. Un modello ADL(1,0)

d. Un modello ADL(0,1)

Risposta: c, poiché il modello ADL(1,1) è $Y_t = \beta_0 + \beta_1 Y_{t-1} + \delta_1 X_{t-1} + u_t$ e quindi il modello AR(1) proposto ne è un caso speciale con $\delta_1 = 0$.

68) Quale tra i seguenti modelli è un modello AR(1)?

a. Un modello con un ritardo di una variabile esogena
b. Un modello con un ritardo della variabile dipendente e un ritardo di una variabile esogena
c. Un modello ARMA(1,0)
d. Un modello ARMA(0,1)

Risposta: c, poiché il modello ARMA(1,1) è $Y_t = \beta_0 + \beta_1 Y_{t-1} + \delta_1 u_{t-1} + u_t$ e quindi il modello AR(1) proposto ne è un caso speciale con $\delta_1 = 0$.

69) Quando si lavora con serie temporali, occorre cautela nell'includere regressori non ritardati perché

a. Potrebbero non essere disponibili i dati
b. Si potrebbe verificare un problema di causalità simultanea, che darebbe luogo ad inconsistenza dello stimatore OLS se non opportunamente trattato (ma il trattamento richiede conoscenze economiche ed econometriche notevoli)
c. Spesso non è importante in pratica
d. Si potrebbe verificare un problema di causalità inversa, che non è di agevole trattamento

Risposta: b e d

70) Quale delle circostanze seguenti è una conseguenza della presenza di un trend stocastico in una variabile:

a. Lo stimatore del coefficiente autoregressivo in modelli AR(1) è distorto verso lo zero
b. Si verifica eteroschedasticità
c. Statistiche di test *t* per sottoporre a verifica ipotesi concernenti i coefficienti del modello di cui al punto precedente avranno una distribuzione normale anche in campioni finiti
d. La variabile cresce all'infinito

Risposta: a

71) Quale delle seguenti scelte è appropriata quando una delle variabili oggetto di studio contenga una radice unitaria, prima di includerla in un modello di regressione (ma non di cointegrazione)?

a. Si differenzia la variabile, ovvero nel modello di regressione si includono le sue differenze prime
b. Si cerca una appropriata correzione degli errori standard
c. La regredisce la variabile su un trend deterministico per depurarla da questa componente
d. Semplicemente si ignora la variabile

Risposta: a

72) Quale delle seguenti scelte è appropriata quando una e solo una delle variabili oggetto di studio contenga un trend stocastico, prima di includerla in un modello di regressione?

a. Si differenzia la variabile, ovvero nel modello di regressione si includono le sue differenze prime
b. Si cerca una appropriata correzione degli errori standard
c. La regredisce la variabile su un trend deterministico per depurarla da questa componente
d. Semplicemente si ignora la variabile

Risposta: a

73) Quale delle seguenti scelte è appropriata quando due o più delle variabili oggetto di studio contengono un trend stocastico? [**da non fare se non si è trattato il tema della cointegrazione**]

a. Si differenziano le variabili prima di includerle in un modello di regressione
b. Si verifica se vi sia una relazione di cointegrazione tra le variabili prima di includerle in un modello a correzione dell'errore
c. Si regrediscono le variabili su un trend deterministico per depurarle dalla radice unitaria prima di includerle in un modello di regressione
d. Semplicemente si ignorano le variabili

Risposta: b

74) Quale delle seguenti scelte è appropriata quando la variabile dipendente di un modello di regressione oggetto di studio contenga un trend deterministico?

a. Si differenzia la variabile, ovvero nel modello di regressione si includono le sue differenze prime

b. Si cerca una appropriata correzione degli errori standard
c. Si include un trend deterministico tra i regressori
d. Si usano al posto di questa variabile i residui di una regressione della variabile stessa su un trend deterministico
e. Semplicemente si ignora la variabile

Risposta: c

75) Quale delle seguenti affermazioni è vera?

a. Le serie stazionarie contengono trend deterministici
b. Le serie stazionarie non contengono né trend né rotture strutturali
c. Le serie non stazionarie non contengono trend
d. Le serie stazionarie contengono rotture strutturali

Risposta: b

76) Quali delle seguenti affermazioni è vera riguardo al test di Dickey-Fuller?

a. l'ipotesi nulla sottoposta a verifica è che il processo stocastico oggetto di studio contenga una radice unitaria
b. l'ipotesi nulla sottoposta a verifica è che il processo stocastico oggetto di studio non contenga una radice unitaria
c. l'ipotesi nulla sottoposta a verifica è che il processo stocastico oggetto di studio sia stazionario
d. l'ipotesi alternativa è che il processo stocastico oggetto di studio contenga una radice unitaria

Risposta: a

77) Si consideri il processo stocastico AR(1) seguente

$$y_t = a + b\, y_{t-1} + e_t$$

Quale dei seguenti è il valore atteso di lungo periodo di y_t quando $b < 1$?

a. $\lim_{j \to +\infty} E(y_{t+j}) = \frac{1}{1-b}$

b. $\lim_{j \to +\infty} E(y_{t+j}) = y_t \frac{1}{1-b}$

c. $\lim_{j \to +\infty} E(y_{t+j}) = \frac{a}{1-b}$

d. $\lim_{j \to +\infty} E(y_{t+j}) = y_t \frac{a}{1-b}$

Risposta: c

78) Si consideri il processo stocastico AR(1) seguente

$$y_t = a + b\, y_{t-1} + e_t$$

Quale delle seguenti è una condizione necessaria perché il valore atteso di lungo periodo di y_t sia definito?

a. $b = 1$
b. $b < a$
c. $b > a$
d. $|b| < 1$

Risposta: d

79) Si consideri il processo stocastico AR(1) seguente

$$y_t = a + b\, y_{t-1} + e_t$$

Si assuma che $a = 1{,}0$, $b = 0{,}8$. Si assuma inoltre che l'osservazione relativa al tempo $t = 10$ sia pari a $y_{10} = 5{,}0$. Il valore atteso al tempo $t = 10$ del valore di y_{11} sarà

a. Maggiore di 5
b. Minore di 5
c. Uguale a 5
d. Non si può dire

Risposta: c, poiché $E_{t-1}(y_t) = a + by_{t-1} = 1 + 0{,}8 \times 5{,}0 = 5{,}0$. Infatti, il valore di $y_{10} = 5{,}0$ è pari alla media di lungo $\bar{y} = \frac{a}{1-b} = \frac{1}{1-0{,}8} = \frac{1}{0{,}2} = 5.0$ e dunque non vi è tendenza all'aumento o alla diminuzione. Ciò si può vedere ricordando che il processo si può riscrivere come

$$y_t = a + b\,\bar{y} + b\,(y_{t-1} - \bar{y}) + e_t$$

Dunque, abbiamo che $E\left(y_{t|t-1}\right) = a + b\,\bar{y} + b\,(y_{t-1} - \bar{y}) = 5{,}0$ perché $a + b\,\bar{y} = 1{,}0 + 0{,}8 \times 5{,}0 = 5{,}0$ e il termine di correzione dell'errore risulta essere $b\,(y_{t-1} - \bar{y}) = 0{,}8 \times (5 - 5) = 0$.

80) Si consideri il processo stocastico AR(1) seguente

$$y_t = 2.3 + 0.9\, y_{t-1} + e_t$$

Questo processo è almeno debolmente stazionario?

a. No
b. Si
c. Dipende dall'ampiezza della costante
d. Dipende dalla volatilità dell'errore.

Risposta: b (poiché il coefficiente del ritardo è minore di uno in valore assoluto).

81) Si consideri il processo stocastico AR(1) seguente

$$y_t = a + b\, y_{t-1} + e_t$$

Quale delle seguenti è una condizione necessaria e sufficiente perché il valore atteso di lungo periodo di y_t sia pari a zero?

a. $a = 0$ e b finito
b. $a = 0$ e $|b| < 1$
c. $a = 0$ e $|b| = 1$
d. $a = 0$ e $b = \infty$

Risposta: b

82) Si consideri il processo stocastico AR(1) seguente

$$y_t = a + b\, y_{t-1} + e_t$$

Quale delle seguenti è una condizione necessaria e sufficiente perché il valore atteso di lungo periodo di y_t esista?

a. $a = 0$ e b finito
b. $a = 0$ e $|b| < 1$
c. $|b| < 1$
d. $a = 0$ e $b = \infty$

Risposta: c

83) Si consideri il seguente modello della variazione trimestrale del tasso di inflazione annualizzato, stimato col metodo OLS usando dati trimestrali annualizzati dal 1962:I al 2004:IV:

$$\widehat{\Delta Inf_t} = 0{,}017 - 0{,}238\ \Delta Inf_{t-1} - 0{,}200\ \Delta Inf_{t-2}$$

Si supponga inoltre che:

$Inf_{2004:II} = 0{,}6$

$Inf_{2004:III} = 1{,}6$

$Inf_{2004:IV} = 3{,}5$

Quale è la previsione di $\Delta Inf_{2005:I}$?

a. all’incirca -0,6
b. all’incirca -1,5
c. all’incirca +0,6
d. all’incirca +1,5

Risposta: a

Infatti, la previsione di $\Delta Inf_{2005:I}$ è:

$$\widehat{\Delta Inf}_{2005:I|2004:IV} = 0{,}017 - 0{,}238 \times 1{,}9 - 0{,}200 \times 1{,}0 = -0{,}64 \approx -0{,}6$$

const	0.017			
b1	-0.238			
b2	-0.2			
inf_{t-2}	0.6			
inf_{t-1}	1.6		delta_inf_·	1
inf_{t}	3.5		delta_inf_·	1.9
			delta_inf_·	-0.601

84) Con riferimento ai dati della domanda precedente, quale è la previsione di $Inf_{2005:I}$?

a. all'incirca 2,9%
b. all'incirca -2,9%
c. all'incirca 3,5%
d. all'incirca -3,5%

Risposta: a

Infatti:

$$\widehat{Inf}_{2005:I|2004:IV} = Inf_{2004:IV} + \widehat{\Delta Inf}_{2005:I|2004:IV} = 3{,}5 - 0{,}6 = 2{,}9\%$$

85) Si consideri il seguente modello della variazione trimestrale del tasso di inflazione annualizzato, stimato col metodo OLS usando dati trimestrali annualizzati dal 1962:I al 2004:IV:

$$\widehat{\Delta Inf}_t = 0{,}017 - 0{,}238\, \Delta Inf_{t-1}$$

Si supponga inoltre che:

$Inf_{2004:III} = 1{,}6$

$Inf_{2004:IV} = 3{,}5$

Quale è la previsione di $\Delta Inf_{2005:I}$?

a. all'incirca -0,4
b. all'incirca -1,5
c. all'incirca +0,4
d. all'incirca +1,5

Risposta: a

Infatti, la previsione di $\Delta Inf_{2005:I}$ è:

$$\widehat{\Delta Inf}_{2005:I|2004:IV} = 0{,}017 - 0{,}238 \times 1{,}9 = -0{,}44 \approx -0{,}4$$

86) Con riferimento ai dati della domanda precedente, quale è la previsione di $Inf_{2005:I}$?

a. all'incirca 3,1%
b. all'incirca -3,1%
c. all'incirca 3,5%
d. all'incirca -3,5%

Risposta: a

Infatti:

$$\widehat{Inf}_{2005:I|2004:IV} = Inf_{2004:IV} + \widehat{\Delta Inf}_{2005:I|2004:IV} = 3{,}5 - 0{,}4 = 3{,}1\%$$

Parte A.2

87) Si faccia riferimento alla Tabella A.1 in Appendice. Cosa si può desumere dal modello (2) riguardo all'effetto di Age sui guadagni orari medi (AHE)?

a. L'effetto è positivo ma non significativo ad alcun livello convenzionale di significatività
b. L'effetto è positivo e significativo al livello di significatività dell'1%
c. L'effetto è positivo e significativo al livello di significatività del 5% ma non al livello dell'1%
d. Non si può dire.

Risposta: b, poiché il coefficiente stimato risulta essere positivo e il valore-p associato è inferiore all'1%.

88) Si faccia riferimento alla Tabella A.1 in Appendice. Cosa si può desumere riguardo alle variabili esplicative considerate per quanto riguarda il modello 2 (in particolare, riguardo l'ipotesi nulla che nessuna delle variabili sia rilevante ai fini della spiegazione della variazione dei guadagni orari medi AHE nella popolazione da cui è estratto il campione)?

a. Si può rigettare al livello di significatività dell'1% l'ipotesi nulla che nessuna delle variabili sia rilevante ai fini della spiegazione della variazione dei guadagni orari medi (AHE) nella popolazione da cui è estratto il campione
b. Si può rigettare al livello di significatività del 5% ma non al livello dell'1% l'ipotesi nulla che nessuna delle variabili sia rilevante ai fini della spiegazione della variazione dei guadagni orari medi (AHE) nella popolazione da cui è estratto il campione
c. Non si può rigettare ad alcun livello di significatività convenzionale l'ipotesi nulla che nessuna delle variabili sia rilevante ai fini della spiegazione della variazione dei guadagni orari medi (AHE) nella popolazione da cui è estratto il campione.
d. Non si può dire.

Risposta: a (poiché la statistica F è altamente significativa almeno per il modello meno parsimonioso e, in questo caso, per ogni modello)

89) Si faccia riferimento alla Tabella A.1 in Appendice. Perché il coefficiente di determinazione R^2 del modello (8) non varia molto rispetto al coefficiente del modello (2)?

a. Le variabili addizionali degli altri modelli non sono molto rilevanti e, se lo sono, sono correlate con quelle del modello (2)
b. Le variabili addizionali degli altri modelli sono molto significative
c. Deve trattarsi di un errore di calcolo

d. Non si può dire.

Risposta: a

Tabella A.1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Variabile dipendente				
	AHE	**ln(AHE)**	**ln(AHE)**	**ln(AHE)**	**ln(AHE)**	**ln(AHE)**	**ln(AHE)**	**ln(AHE)**
Age	0,48	0,03		0,17	0,17	0,22	0,18	0,22
	(0,00)	(0,00)		(0,002)	(0,002)	(0,002)	(0,02)	(0,01)
ln(Age)			0,76					
			(0,00)					
Age^2				-0,002	-0,002	-0,003	-0,003	-0,003
				(0,008)	(0,008)	(0,007)	(0,04)	(0,02)
Female	-2,90	-0,17	-0,17	-0,16	-0,20	1,77	-0,20	2,10
	(0,00)	(0,00)	(0,00)	(0,00)	(0,00)	(0,27)	(0,00)	(0,20)
Female x Age						-0,12		-0,14
						(0,26)		(0,20)
Female x Age^2						0,002		0,002
						(0,30)		(0,24)
Bachelor	6,72	0,40	0,40	0,40	0,37	0,37	0,50	-0,10
	(0,00)	(0,00)	(0,00)	(0,00)	(0,00)	(0,00)	(0,75)	(0,95)
Bachelor x Age							-0,02	0,017
							(0,84)	(0,88)
Bachelor x Age^2							0,00	-0,00
							(0,75)	(0,98)
Female x Bachelor					0,07	0,07	0,08	0,07
					(0,007)	(0,01)	(0,01)	(0,01)
Intercetta	0,48	0,81	0,002	-0,31	-0,28	-1,14	-0,28	-0,99
F-statistic	338,29	341,44	342,37	258,20	208,31	149,74	150,05	117,77
	(0,00)	(0,00)	(0,00)	(0,00)	(0,00)	(0,00)	(0,00)	(0,00)
SER	7,72	0,45	0,45	0,45	0,45	0,45	0,45	0,45
R^2	0,184	0,1855	0,1859	0,1868	0,188	0,189	0,189	0,191
Adjusted R^2	0,183	0,1850	0,1854	0,1861	0,187	0,187	0,188	0,189
n	4500	4500	4500	4500	4500	4500	4500	4500

Note:. La tabella riporta otto diverse regressioni: la colonna (1) riporta la regressione della retribuzione oraria media (AHE) sull'età (Age), sul sesso (Female) e sull'istruzione (Bachelor); la colonna (2) indica la regressione di ln(AHE) su Age, Female e Bachelor; la colonna (3) evidenzia la regressione di ln(AHE) su ln(Age), Female e Bachelor; la colonna (4) mostra la regressione di ln(AHE) su Age, Age^2, Female e Bachelor; la colonna (5) riporta la regressione di ln(AHE) su Age, Age2, Female, Bachelor e Female x Bachelor; dalla colonna (6) si vede la regressione di ln(AHE) su Age, Age2, Female, Bachelor, Female x Bachelor, Female x Age e Female x Age^2; dalla colonna (7) si osserva la regressione di ln(AHE) su Age, Age2, Female, Bachelor, Female x Bachelor, Bachelor x Age e Bachelor x Age^2; infine la colonna (8) mostra la regressione su ln(AHE) di tutte le precedenti variabili. Per ciascun valore dei regressori viene riportato in parentesi il p-value associato. Il campione oggetto di studio contiene n = 4500 osservazioni. Si riportano inoltre: la statistica F (e relativo p-value) del test per la significatività congiunta dei regressori, la deviazione standard campionaria dei residui e due misure di bontà di adattamento del modello ai dati (R^2 e adjusted R squared). Le stime sono ottenute con un campione che comprende le prime 4500 osservazioni.

90) Si faccia riferimento alle Tabelle 14.a e 14.b in Appendice. Quale delle due riporta stime relative a un modello puramente autoregressivo?

a. La Tabella A
b. La Tabella B
c. Entrambe
d. Nessuna

Risposta: a

91) Si faccia riferimento alle Tabelle 14.a e 14.c in Appendice, facendovi riferimento come a Tabella A e Tabella B, rispettivamente. Quale delle due riporta stime relative a un modello puramente autoregressivo?

a. La Tabella A
b. La Tabella B
c. Entrambe
d. Nessuna

Risposta: c

92) Si faccia riferimento alle Tabelle 14.a e 14.b in Appendice. Quale delle due riporta stime relative a un modello autoregressivo a ritardi distribuiti con coefficienti non nulli delle variabili esogene?

a. La Tabella A
b. La Tabella B
c. Entrambe
d. Nessuna

Risposta: b

93) Si faccia riferimento alle Tabelle 14.a e 14.b in Appendice. Quale delle due riporta stime relative a un modello autoregressivo a ritardi distribuiti?

a. La Tabella A
b. La Tabella B
c. Entrambe
d. Nessuna

Risposta: c

94) Si faccia riferimento alla Tabella 14.a in Appendice. A quale livello di significatività minimo possiamo rifiutare l'ipotesi nulla che la capacità previsiva del modello AR(1) non sia superiore a quella di un modello che include solo la costante?

a. Al livello dell'1%
b. Al livello del 5%
c. Al livello del 10%
d. Al livello del 15%

Risposta: c

95) Si faccia riferimento alla Tabella 14.b in Appendice. A quale livello di significatività possiamo rifiutare l'ipotesi nulla che la capacità previsiva del modello (1) non sia superiore a quella di un modello che include solo la costante?

a. Al livello dell'1%
b. Al livello del 10%
c. Al livello del 15%
d. A nessuno dei livelli di cui sopra

Risposta: d

96) Si faccia riferimento alla Tabella 14.b in Appendice. A quale livello di significatività minimo possiamo rifiutare l'ipotesi nulla che la capacità previsiva del modello (3) non sia superiore a quella di un modello che include solo la costante?

a. Al livello del 5%
b. Al livello del 10%
c. Al livello del 15%
d. A nessuno dei livelli di cui sopra

Risposta: b

97) Si faccia riferimento al modello (3) nella Tabella 14.b in Appendice. A quale livello di significatività minimo possiamo rifiutare l'ipotesi nulla che il primo ritardo del rendimento in eccesso al tasso senza rischio non aiuti a prevedere la variabile dipendente?

a. Al livello del 5%
b. Al livello del 10%
c. Al livello del 15%
d. A nessuno dei livelli di cui sopra

Risposta: b

	(1)	(2)	(3)
	AR(1)	**AR(2)**	**AR(4)**
Exreturn(-1)	0,10	0,10	0,10
	(0,10)	(0,10)	(0,09)
Exreturn(-2)		-0,04	-0,03
		(0,46)	(0,56)
Exreturn(-3)			-0,10
			(0,06)
Exreturn(-4)			0,01
			(0,84)
Intercetta	0,55	0,58	0,63
Wald F-statistic	2,64	3,15	3,70
	(0,10)	(0,10)	(0,06)
R^2	0,010	0,012	0,023
Adjusted R squared	0,009	0,010	0,018
n	840	840	840

Tabella 14.a. I rendimenti in eccesso sono misurati in percentuale su base mensile. Le regressioni sono stimate per il periodo 1965:1 – 2004:12 (T=840 osservazioni). La tabella riporta tre diverse regressioni effettuate su campioni con n = 840 osservazioni: il primo campione nella colonna (AR1) riporta l'autoregressione di primo ordine del rendimento in eccesso atteso. Nella colonna (AR2) si riporta l'autoregressione di secondo ordine del rendimento in eccesso atteso; la colonna (AR4) riporta l'autoregressione di quarto ordine del rendimento in eccesso atteso. Per ciascun valore dei regressori viene riportato in parentesi il valore-p associato. Si riportano inoltre la statistica F di Wald (e relativo valore-p) del test per la non rilevanza congiunta dei regressori, la deviazione standard campionaria dei residui e due misure di bontà di adattamento del modello ai dati (R^2 e adjusted $\boldsymbol{R^2}$).

	(1)	(2)	(3)
	ADL(1,1)	**ADL(2,2)**	**ADL(1,1)**
Exreturn(-1)	0,009	0,105	0,106
	(0,53)	(0,40)	(0,08)
Exreturn(-2)		-0,098	
		(0,52)	
Δln(dividend Yield-1)	-0,01	0,0001	
	(0,92)	(0,99)	
Δln(dividend Yield-2)		-0,056	
		(0,67)	
ln(dividend Yield-1)			0,011
			(0,09)
Intercetta	0,56	0,60	4,33
Wald F-statistic	1,34	0,82	2,65
	(0,26)	(0,51)	(0,07)
SER	5,13	5,13	5,11
R^2	0,010	0,012	0,017
Adjusted R squared	0,0077	0,0075	0,014
n	840	840	840

Tabella 14.b. Nella prima colonna abbiamo effettuato una regressione della variabile rendimento in eccesso (del tasso di rendimento dell'attività priva di rischio) sul suo ritardo e sul ritardo della variabile $\boldsymbol{\Delta \ln}$ $(\textbf{dividend Yield})$; nella seconda colonna abbiamo considerato due ritardi di entrambe le variabili, mentre nella terza colonna abbiamo considerato solamente un ritardo della variabile rendimento in eccesso e il ritardo primo della variabile $\textbf{ln}$ $(\textbf{dividend Yield})$. Per ciascun valore dei regressori viene riportato in parentesi il p-value associato. Si riportano inoltre: la statistica F di Wald (con in parentesi il p-value associato) per la significatività congiunta dei regressori, la deviazione standard campionaria dei residui e due misure di bontà di adattamento del modello ai dati ($\boldsymbol{R^2}$ e adjusted $\boldsymbol{R^2}$).

	(1)	(2)	(3)
	ADL(1,1)	**ADL(2,2)**	**ADL(1,0)**
Exreturn(-1)	0,009	0,105	0,106
	(0,53)	(0,40)	(0,08)
Exreturn(-2)		-0,098	
		(0,52)	
Δln(dividend Yield-1)	-0,01	0,0001	
	(0,92)	(0,99)	
Δln(dividend Yield-2)		-0,056	
		(0,67)	
ln(dividend Yield-1)			
Intercetta	0,56	0,60	4,33
Wald F-statistic	1,34	0,82	2,65
	(0,26)	(0,51)	(0,07)
SER	5,13	5,13	5,11
R^2	0,010	0,012	0,017
Adjusted R squared	0,0077	0,0075	0,014
n	840	840	840

Tabella 14.c. Nella prima colonna abbiamo effettuato una regressione della variabile rendimento in eccesso (del tasso di rendimento dell'attività priva di rischio) sul suo ritardo e sul ritardo della variabile $\mathbf{\Delta ln}\ (\mathbf{dividend\ Yield})$; nella seconda colonna abbiamo considerato due ritardi di entrambe le variabili, mentre nella terza colonna abbiamo considerato solamente un ritardo della variabile rendimento in eccesso e il ritardo primo della variabile $\mathbf{ln}\ (\mathbf{dividend\ Yield})$. Per ciascun valore dei regressori viene riportato in parentesi il p-value associato. Si riportano inoltre: la statistica F di Wald per la significatività congiunta dei regressori, la deviazione standard campionaria dei residui e due misure di bontà di adattamento del modello ai dati ($\boldsymbol{R^2}$ e adjusted $\boldsymbol{R^2}$).

98) Si consideri il modello Probit seguente della relazione tra y_i e x_i:

$$y_i = \Phi(\alpha + \beta x_i) + u_i$$

Nel modello di cui sopra, α e β sono parametri, $\Phi()$ è la funzione di densità cumulata di una variabile aleatoria normale standardizzata e u_i è un errore di regressione. Si supponga che $\alpha = -8{,}2$ e $\beta = 3{,}2$. Quale è la probabilità che $y_i = 1$ se $x_i = 2{,}3$?

a. 0,5%
b. 10%
c. 20%
d. 80%
e. Nessuna delle risposte di cui sopra con un'approssimazione di un decimale

Risposta: c. Infatti:

Probit: y = F(a+bx) + u	
a	-8.2
b	3.2
x	2.3
z	-0.84
prob(y=1|x)	0.200454

99) Si consideri il modello Probit seguente della relazione tra y_i e x_i:

$$y_i = \Phi(\alpha + \beta x_i) + u_i$$

Nel modello di cui sopra, α e β sono parametri, $\Phi()$ è la funzione di densità cumulata di una variabile aleatoria standardizzata e u_i è un errore di regressione. Si supponga che

$$\alpha = -8{,}2 \text{ e } \beta = 3{,}2.$$

Di quanto aumenta la probabilità che $y_i = 1$ se $x_i = 2.3$ aumenta di un piccolo ammontare ε ?

a. L'aumento è di 3,2 × ε
b. L'aumento è di 89,71% × ε
c. L'aumento è di ε
d. L'aumento è di 20% × ε
e. Nessuna delle risposte di cui sopra con un'approssimazione di un decimale

Risposta: b. Infatti:

Probit: y = F(a+bx) + u	
a	-8.2
b	3.2
x	2.3
z	-0.84
prob(y=1|x)	0.200454
f(z)	28.03%
b * f(z)	89.71%

Parte B

(Quesiti a risposta aperta)

Quesito 1

Un economista dispone di dati ottenuti intervistando un campione di 3.763 soggetti ritenuto rappresentativo dei giovani residenti negli Stati Uniti. I dati sono sui seguenti caratteri dei soggetti:

- Salario orario
- Anni di scolarizzazione
- Anni di esperienza lavorativa

I soggetti intervistati avevano un'età compresa tra 14 e 21 anni nel 1979. L'economista usa questi dati per stimare quanto segue:

- Una regressione del logaritmo del salario orario su scolarizzazione, che chiameremo regressione (1)
- Una regressione del logaritmo del salario orario su scolarizzazione ed esperienza lavorativa, che chiameremo regressione (2)
- Una regressione del logaritmo del salario orario su scolarizzazione, esperienza lavorativa ed abilità cognitive (misurate tramite i punteggi in test di intelligenza), che chiameremo regressione (3)

I risultati di queste regressioni sono riportati nella tabella seguente. Nella prima riga della colonna, si riporta il numero del modello cui le stime riportate alle righe inferiori si riferiscono. Nelle righe successive, per ciascun modello, si mostrano le stime dei coefficienti delle variabili incluse, unitamente alle statistiche (mostrate tra parentesi) del test *t* della ipotesi nulla che il coefficiente della variabile cui ciascuna statistica si riferisce sia pari a zero.

	(1)	**(2)**	**(3)**
Scolarizzazione	0,114 (30,16)	0,116 (28,99)	0,05 (14,53)
Esperienza lavorativa	-	0,038 (14,59)	0,037 (14,10)
Abilità cognitive	-	-	0,004 (4,79)
Costante	4,899 (74,59)	3,523 (65,02)	2,966 (32,69)
R^2	0,165	0,243	0,348
n (numero di soggetti)	3.763	3.763	3.763

Con riferimento a quanto sopra, si risponda alle seguenti domande:

a. Si spieghi come si possa effettuare, sulla base delle informazioni disponibili, un test della ipotesi nulla che i coefficienti di “Esperienza lavorativa” ed “Abilità cognitiva” siano congiuntamente significativi.
b. L’economista, contemplando i propri risultati, non riesce a spiegarsi come mai, nella regressione del modello (3), il coefficiente relativo agli anni di scolarizzazione sia così basso al confronto del coefficiente relativo agli anni di scolarizzazione nelle regressioni (1) e (2). In particolare, il nostro economista si chiede se il problema possa essere la dimensione del campione che, per quanto sembri grande, potrebbe sempre essere aumentata con uno sforzo di raccolta di dati maggiore. Lo studente offra la propria spiegazione del risultato empirico che ha sorpreso il nostro economista e ne valuti il ragionamento riguardo il possibile effetto dell’ampiezza del campione.
c. Un collega osserva che c’è notevole disparità tra i salari di soggetti diversi e, quindi, che l’ipotesi di omoschedasticità del termine di errore sia difficile da sostenere. Se il collega avesse ragione, quali problemi ciò comporterebbe per lo stimatore OLS e come si potrebbe ovviare?

Risposta:

(a) Si può effettuare un test F, calcolando la statistica F della differenza tra i coefficienti di determinazione dei modelli 3 e 1, che ha una distribuzione asintoticamente Chi-quadrato con 2 e 3.761 gradi di libertà. Comunque, anche un semplice test *t* sul coefficiente di “Esperienza lavorativa” o di “Abilità cognitiva” basta a rifiutare la ipotesi nulla.

(b) Chiaramente nei modelli (1) e (2) abbiamo un problema di variabile omessa rilevante, che porta ad una stima inconsistente del coefficiente degli anni di scolarizzazione. Quando questo problema viene corretto, ovvero nel modello (3), la stima OLS torna consistente. Per quanto riguarda l’ampiezza del campione, la inconsistenza non la si potrebbe rimediare nemmeno con un campione infinito, per quanto un numero maggiore di osservazioni sia sempre desiderabile poiché riduce gli errori standard delle stime.

(c) Il problema sarebbe la inconsistenza dello stimatore OLS degli errori standard. Si potrebbe ovviare utilizzando uno stimatore che apporti una correzione per tenere conto della eteroschedasticità, come per esempio lo stimatore degli errori standard proposto da White.

Quesito 2

Un economista dispone di dati ottenuti intervistando un campione di 3.763 soggetti ritenuto rappresentativo dei giovani residenti in una regione d'Italia. I dati sono sui seguenti caratteri dei soggetti:

- Salario orario (in Euro)
- Anni di scolarizzazione
- Anni di esperienza lavorativa

I soggetti intervistati avevano un'età compresa tra 14 e 21 anni nel 2009. L'economista usa questi dati per stimare quanto segue:

- Una regressione del logaritmo del salario orario su scolarizzazione, che chiameremo regressione (1)
- Una regressione del logaritmo del salario orario su scolarizzazione ed esperienza lavorativa, che chiameremo regressione (2)
- Una regressione del logaritmo del salario orario su scolarizzazione, esperienza lavorativa ed abilità cognitive (misurate tramite i punteggi in test di intelligenza), che chiameremo regressione (3)

I risultati di queste regressioni sono riportati nella tabella seguente. Nella prima riga della colonna, si riporta il numero del modello cui le stime riportate alle righe inferiori si riferiscono. Nelle righe successive, per ciascun modello, si mostrano le stime dei coefficienti delle variabili incluse, unitamente alle statistiche (mostrate tra parentesi) del test *t* della ipotesi nulla che il coefficiente della variabile cui ciascuna statistica si riferisce sia pari a zero.

	(1)	**(2)**	**(3)**
Scolarizzazione	0,114	0,116	0,05
	(29,16)	(28,99)	(14,53)
Esperienza lavorativa	-	0,038	0,037
		(2,59)	(1,10)
Abilità cognitive	-	-	0,004
			(1,79)
Costante	4,899	3,523	2,966
	(74,59)	(65,02)	(32,69)
R^2	0,165	0,243	0,288
n (numero di soggetti)	3.763	3.763	3.763

Con riferimento a quanto sopra, si risponda alle seguenti domande:

a. Si spieghi come si possa effettuare, sulla base delle informazioni disponibili, un test della ipotesi nulla che i coefficienti di "Esperienza lavorativa" ed "Abilità cognitiva" siano congiuntamente significativi.
b. L'economista, contemplando i propri risultati, non riesce a spiegarsi come mai, nella regressione del modello (3), il coefficiente relativo agli anni di scolarizzazione sia così basso al confronto del coefficiente relativo agli anni di scolarizzazione nelle regressioni (1) e (2). In particolare, il nostro economista si chiede se il problema possa essere la dimensione del campione che, per quanto sembri grande, potrebbe sempre essere aumentata con uno sforzo di raccolta di dati maggiore. Lo studente offra la propria spiegazione del risultato empirico che ha sorpreso il nostro economista e ne valuti il ragionamento riguardo il possibile effetto dell'ampiezza del campione.
c. Un collega osserva che c'è notevole disparità tra i salari di soggetti diversi e, quindi, che l'ipotesi di omoschedasticità del termine di errore sia difficile da sostenere. Se il collega avesse ragione, quali problemi ciò comporterebbe per lo stimatore OLS e come si potrebbe ovviare?

Risposta:

(a) Si può effettuare un test F, calcolando la statistica F della differenza tra i coefficienti di determinazione dei modelli 3 e 1, che ha una distribuzione asintoticamente Chi-quadrato con 2 gradi di libertà.

(b) Chiaramente nei modelli (1) e (2) abbiamo un problema di variabile omessa rilevante, che porta ad una stima inconsistente del coefficiente degli anni di scolarizzazione. Quando questo problema viene corretto, ovvero nel modello (3), la stima OLS torna consistente. Per quanto riguarda l'ampiezza del campione, la inconsistenza non la si potrebbe rimediare nemmeno con un campione infinito, per quanto un numero maggiore di osservazioni sia sempre desiderabile poiché riduce gli errori standard delle stime.

(c) Il problema sarebbe la inconsistenza dello stimatore OLS degli errori standard. Si potrebbe ovviare utilizzando uno stimatore che apporti una correzione per tenere conto della eteroschedasticità, come per esempio lo stimatore degli errori standard proposto da White.

Quesito 3

Un economista dispone di dati mensili dal 1987 al 2010 sui tassi di inflazione di Canada e Stati Uniti, denotati rispettivamente $\pi_{1,t}$ e $\pi_{2,t}$.

a. Si supponga che si sia effettuato un test di Dickey–Fuller con la finalità di fare inferenza sulla stazionarietà di $\pi_{1,t}$ e $\Delta\pi_{1,t}$ e che le relative statistiche-test assumano rispettivamente il valore di -1.26 e -3.10.
   i. Si spieghi quale sia la ipotesi nulla del test
   ii. Si illustrino le inferenze che sia possibile effettuare sulla base del risultato del test, tenendo presente i valori critici della distribuzione della statistica del test di Dickey–Fuller riportati di seguito:

Livello di significatività	Valore critico
1%	-2.569236
5%	-1.941408
10%	-1.616307

b. Il modello preso in considerazione per studiare la relazione tra i due tassi di inflazione è il seguente:

$$\pi_{1,t} = a + b_1\,\pi_{2,t} + u_t$$

dove i simboli hanno i significati ed interpretazione consueti. Si supponga che $\pi_{1,t}$ contenga una radice unitaria, ovvero che sia I(1), mentre $\pi_{2,t}$ sia I(0). Quali problemi ciò comporterebbe se volessimo stimare i parametri del modello di cui sopra tramite l'applicazione dello stimatore OLS?

c. Si supponga che il modello che descrive la relazione tra $\pi_{1,t}$ e $\pi_{2,t}$ sia quello descritto alla domanda precedente ma che sia $\pi_{1,t}$ che $\pi_{2,t}$ siano I(1). Quale problema ciò potrebbe comportare se volessimo stimare i parametri del modello tramite l'applicazione dello stimatore OLS?

Risposta (eventuale altro spazio per la risposta alle domande precedenti, anche se si raccomandano risposte succinte per quanto esaurienti):

Si veda discussione fatta in classe.