

## Μεταπτυχιακό Πρόγραμμα στη Στατιστική και Ε.Ε.

### Χρονοσειρές: Εργασία 1 Παράδοση έως Τρίτη 9/1/2024

Α. Στο αρχείο **cobb\_douglas\_data.txt** περιέχονται δεδομένα για το Ελληνικό ΑΕΠ ( $Q_t$ ), το Κεφάλαιο ( $K_t$ ) και την Εργασία ( $L_t$ ) για τα έτη 1961-2003 (με αυτή τη σειρά στις στήλες του αρχείου δεδομένων). Να χρησιμοποιηθούν αυτά τα δεδομένα για να εκτιμηθεί η συνάρτηση Cobb Douglas για την Ελληνική Οικονομία. Συγκεκριμένα, θεωρήστε το μοντέλο

$$Q_t = AK_t^{b_1} L_t^{b_2} e^{\varepsilon_t}$$

α) Να εκτιμήσετε τα  $b_1, b_2$  με τη μέθοδο ελαχίστων τετραγώνων και να υπολογίσετε 90% διαστήματα εμπιστοσύνης για αυτά, αφού κάνετε τις απαιτούμενες υποθέσεις για τους τυχαίους όρους  $\varepsilon_t$ . Ερμηνεύστε τα  $b_1, b_2$ . Τι συμπεράσματα βγάζετε από τα διαστήματα εμπιστοσύνης;

β) Να ελεγχθεί η στατιστική σημαντικότητα των παραμέτρων  $b_1, b_2$ , καθώς και η στατιστική σημαντικότητα του γραμμικού μοντέλου, σε επίπεδο σημαντικότητας  $\alpha=10\%$ .

γ) Να γίνουν οι παρακάτω έλεγχοι υποθέσεων σε επίπεδο στατιστικής σημαντικότητας  $\alpha=10\%$ :

$$(i) \begin{matrix} H_0 : b_1 = 1 \\ H_1 : b_1 < 1 \end{matrix}, \quad (ii) \begin{matrix} H_0 : b_2 = 1 \\ H_1 : b_2 > 1 \end{matrix}, \quad (iii) \begin{matrix} H_0 : b_1 + b_2 = 1 \\ H_1 : b_1 + b_2 \neq 1 \end{matrix}, \quad (iv) \begin{matrix} H_0 : b_1 + b_2 = 1 \\ H_1 : b_1 + b_2 > 1 \end{matrix}.$$

Για τους ελέγχους αυτούς να υπολογίσετε τα αντίστοιχα p-values. Τι συμπεράσματα βγάζετε;

δ) Να γίνει πρόβλεψη του ΑΕΠ για (i)  $K=5000000, L=3500000$  και (ii)  $K = \bar{K}, L = \bar{L}$ , όπου  $\bar{K}, \bar{L}$  οι δειγματικοί μέσοι για το Κεφάλαιο και την Εργασία. Να υπολογίσετε 90% διαστήματα εμπιστοσύνης για τις προβλέψεις.

ε) Θεωρήστε τα παρακάτω μοντέλα

$$M1: Q_t = AK_t^{b_1} L_t^{b_2} e^{\varepsilon_t}$$

$$M2: Q_t = AL_t^{b_2} e^{\varepsilon_t}$$

$$M3: Q_t = AK_t^{b_1} e^{\varepsilon_t}$$

$$M4: Q_t = Ae^{\varepsilon_t}$$

Να επιλεγεί το καταλληλότερο μοντέλο για την Ελληνική Οικονομία σύμφωνα με (i) το  $R_{adj}^2$ , (ii) το AIC, (iii) το BIC.

στ) Για τη συνάρτηση Cobb Douglas καθώς και για το υπομοντέλο που επιλέγεται (αν είναι διαφορετικό) να ελέγξετε αν ισχύουν οι υποθέσεις που κάνατε για τους τυχαίους όρους, με γραφικό έλεγχο καταλοίπων. Σχολιάστε.

Β. Στο αρχείο **consumption\_data\_cross\_sectional.txt** περιέχονται δεδομένα για την κατανάλωση (C) και το εισόδημα (Y) 326 ελληνικών νοικοκυριών για το έτος 2004 (στήλες

7 και 6 του αρχείου δεδομένων, αντίστοιχα). Για αυτές τις παρατηρήσεις υποθέτουμε το ακόλουθο απλό γραμμικό μοντέλο

$$C_i = b_0 + b_1 Y_i + \varepsilon_i.$$

α) Να εκτιμήσετε τα  $b_0, b_1$  με τη μέθοδο ελαχίστων τετραγώνων και να υπολογίσετε 95% διαστήματα εμπιστοσύνης για αυτά, αφού κάνετε τις απαιτούμενες υποθέσεις για τους τυχαίους όρους  $\varepsilon_i$ . Τι συμπεράσματα βγάζετε από τα διαστήματα εμπιστοσύνης;

β) Να γίνει ο έλεγχος ( $\alpha=5\%$ ) ετεροσκεδαστικότητας (i) Breusch-Pagan, θεωρώντας ότι η ετεροσκεδαστικότητα περιγράφεται από τη σχέση  $\sigma_i^2 = a_0 + a_1 Y_i$ ,  $i = 1, \dots, 326$ , (ii) White. Για τους ελέγχους αυτούς να υπολογίσετε τα αντίστοιχα p-values. Τι συμπεράσματα βγάζετε;

γ) Δεδομένης της ύπαρξης ισχυρής ετεροσκεδαστικότητας του τυχαίου όρου ως προς το εισόδημα  $Y_i$ , θεωρούμε ότι αυτή περιγράφεται από την ακόλουθη σχέση

$$\sigma_i^2 = a Y_i$$

Να εκτιμήσετε τα  $b_0, b_1$  με τη μέθοδο weighted least squares (WLS) και να υπολογίσετε 95% διαστήματα εμπιστοσύνης για αυτά. Συγκρίνετε τα αποτελέσματά σας με αυτά της μεθόδου least squares (LS - μ.ε.τ).

δ) Θεωρήστε ότι η ετεροσκεδαστικότητα περιγράφεται από τη σχέση

$$\sigma_i^2 = a_0 + a_1 Y_i$$

Να εκτιμήσετε τα  $b_0, b_1$  με τη μέθοδο generalized least squares (GLS) και να υπολογίσετε 95% διαστήματα εμπιστοσύνης για αυτά. Πόσο αλλάζουν τώρα τα αποτελέσματά σας; Τι συμπεράσματα βγάζετε;

ε) Στο αρχείο **consumption\_data\_cross\_sectional.txt**, στην μεταβλητή Res (στήλη 5) δίνεται πληροφορία για τον τόπο διαμονής των νοικοκυριών (0: αστική, 1: ημιαστική, 2: αγροτική). Να εκτιμήσετε (σημειικά με μ.ε.τ και με 95% διαστήματα εμπιστοσύνης) ένα γραμμικό μοντέλο για τη σχέση κατανάλωσης εισοδήματος που να λαμβάνει υπόψη την επίδραση του τόπου διαμονής των νοικοκυριών. Να ελέγξετε αν η επίδραση του τόπου διαμονής και η αλληλεπίδραση αυτού με το εισόδημα είναι στατιστικά σημαντικές και με βάση αυτό να επιλέξετε το καταλληλότερο γραμμικό μοντέλο για αυτά τα δεδομένα.

στ) Για το γραμμικό μοντέλο που θα επιλέξετε να γίνει ο έλεγχος ( $\alpha=5\%$ ) ετεροσκεδαστικότητας (i) Breusch-Pagan, θεωρώντας ότι η ετεροσκεδαστικότητα περιγράφεται από τη σχέση  $\sigma_i^2 = a_0 + a_1 Y_i$ ,  $i = 1, \dots, 326$ , (ii) White. Για τους ελέγχους αυτούς να υπολογίσετε τα αντίστοιχα p-values. Τι συμπεράσματα βγάζετε; Τέλος, θεωρώντας ότι η ετεροσκεδαστικότητα περιγράφεται από τη σχέση

$$\sigma_i^2 = a_0 + a_1 Y_i$$

να εκτιμήσετε τα  $b_0, b_1$  με τη μέθοδο generalized least squares (GLS) και να υπολογίσετε 95% διαστήματα εμπιστοσύνης για αυτά. Πόσο αλλάζουν τα αποτελέσματά σας σε σχέση με αυτά της μ.ε.τ; Τι συμπεράσματα βγάζετε;

Να παρουσιάσετε τα αποτελέσματα της εργασίας σας σε ένα κείμενο με τα σχόλια και τα συμπεράσματά σας και να παραθέσετε τον κώδικα R που θα χρησιμοποιήσετε σε παράρτημα.