

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ

ΔΙΟΙΚΗΣΗ ΕΠΙΧΕΙΡΗΣΕΩΝ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

«Βασικές Μέθοδοι Ανάλυσης Χρονολογικών
Σειρών»

Φοιτητές: Κολοβού Λαμπρινή ΑΜ:10958
Μπαστάνη Στέλλα ΑΜ:10832
Επιβλέπων Καθηγητής: Κουετάς Κων/νος

Περιεχόμενα

ΠΡΟΛΟΓΟΣ.....	2
ΠΕΡΙΛΗΨΗ.....	3
ΚΕΦΑΛΑΙΟ 1 ^ο : ΕΙΣΑΓΩΓΗ.....	4
ΚΕΦΑΛΑΙΟ 2 ^ο : ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ.....	11
2.1 Στοχαστική διαδικασία και χρονολογικές σειρές.....	12
2.2 Συνεχείς-Διακριτές χρονολογικές σειρές.....	12
2.3 Μέτρα χρονολογικών σειρών.....	14
2.4 Συνάρτηση αυτοσυσχετίσεως.....	20
ΚΕΦΑΛΑΙΟ 3 ^ο : ΣΤΑΣΙΜΕΣ ΣΤΟΧΑΣΤΙΚΕΣ ΔΙΑΔΙΚΑΣΙΕΣ.....	24
3.1 Αυτοπαλίνδρομη διαδικασία τάξης p , $AR(p)$	27
3.1.1 Αυτοπαλίνδρομη διαδικασία τάξης 1, $AR(1)$	28
3.1.2 Αυτοπαλίνδρομη διαδικασία τάξης 2, $AR(2)$	30
3.2 Η συνάρτηση μερικής αυτοσυσχέτισης.....	32
3.3 Διαδικασίες κινητού μέσου.....	35
3.3.1 Διαδικασία κινητού μέσου τάξης 1, $MA(1)$	36
3.3.2 Διαδικασία κινητού μέσου τάξης 2, $MA(2)$	36
3.3.3 Διαδικασία κινητού μέσου τάξης q , $MA(q)$	37
3.3.4 Σχέση μεταξύ $AR(p)$ και $MA(q)$ διαδικασιών.....	38
3.4 Μικτές Διαδικασίες ($ARMA(p,q)$).....	39
3.4.1 $ARMA(1,1)$ Διαδικασία.....	40
3.5 Το μοντέλο $ARIMA(p,d,q)$	41
3.6 Μεθολογία <i>Box-Jenkins</i>	41
ΚΕΦΑΛΑΙΟ 4 ^ο : ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ.....	44
4.1 Εισαγωγή στο πρόγραμμα PASW 18.....	45
4.2 Αναφορά-Στοιχεία για την ενεργειακή κατανάλωση στην Ευρώπη.....	46
4.3 Ανάλυση των δεδομένων της ενεργειακής κατανάλωσης της Γαλλίας.....	46
4.4 Ανάλυση των δεδομένων της ενεργειακής κατανάλωσης της Ελλάδας.....	60
4.5 Ανάλυση των δεδομένων της ενεργειακής κατανάλωσης της Αυστρίας.....	74
ΠΑΡΑΡΤΗΜΑ.....	90
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	156

Πρόλογος

Η παρούσα πτυχιακή εργασία με θέμα «Βασικές Μέθοδοι Ανάλυσης Χρονολογικών Σειρών», πραγματοποιήθηκε στο πλαίσιο της πτυχιακής εργασίας του τμήματος Επιχειρηματικού Σχεδιασμού και Πληροφοριακών Συστημάτων του Ανώτατου Τεχνολογικού Εκπαιδευτικού Ιδρύματος Πατρών. Η εργασία αυτή αναπτύσσεται σε τέσσερα κεφάλαια και ένα παράρτημα.

- Το **Κεφάλαιο 1**, περιλαμβάνει μια εισαγωγή στις χρονολογικές σειρές καθώς και μια αναφορά σε παραδείγματα χρήσης χρονολογικών σειρών στην καθημερινή ζωή.
- Στο **Κεφάλαιο 2**, αναπτύσσονται οι βασικές έννοιες χρονολογικών σειρών.
- Στο **Κεφάλαιο 3**, περιγράφονται οι στάσιμες στοχαστικές διαδικασίες.
- Στο **Κεφάλαιο 4**, γίνεται μια περιγραφή του λογισμικού PASW 18 που χρησιμοποιήθηκε στην παρούσα εργασία καθώς και η διαδικασία η οποία ακολουθήθηκε για την επεξεργασία των δεδομένων των ενεργειακών καταναλώσεων των ευρωπαϊκών χωρών.

Στο **Παράρτημα Α**, παρατίθενται συνοπτικά η διαδικασία και τα αποτελέσματα της επεξεργασίας των δεδομένων των ενεργειακών καταναλώσεων των Ευρωπαϊκών χωρών.

Στο σημείο αυτό αισθανόμαστε την ανάγκη να εκφράσουμε τις ειλικρινείς και θερμές ευχαριστίες μας σε όσους συνέβαλαν στην ολοκλήρωση αυτής της προσπάθειας. Πρώτα απ' όλα, στον επιβλέποντα καθηγητή μας κ. Κουνετά Κωνσταντίνο για τη συνεχή καθοδήγηση, τις ουσιώδεις συμβουλές, καθώς και την ενθάρρυνση που μας παρείχε όλο αυτό το διάστημα. Επίσης θα θέλαμε να ευχαριστήσουμε θερμά την κ. Μπουμπούλη Αθανασία ως πρώτη επιβλέπουσα. Τέλος, θέλουμε να ευχαριστήσουμε όλους εκείνους που μας βοήθησαν με την στήριξη τους τις οικογένεια μας, τους φίλους μας, τους συμφοιτητές μας. Σε αυτούς, που με την καθημερινή τους συμπαράσταση, την υπομονή τους και τη θετική τους σκέψη, συνέβαλαν στην εκπλήρωση του στόχου μας.

Πάτρα 2013

ΠΕΡΙΛΗΨΗ

Στην εργασία αυτή θα μελετηθούν και θα αναπτυχθούν οι βασικές μέθοδοι ανάλυσης των χρονολογικών σειρών. Χρονολογική σειρά είναι μια σειρά από παρατηρήσεις που λαμβάνονται σε ορισμένες χρονικές στιγμές ή περιόδους που ισαπέχουν ή όχι μεταξύ τους. Στην παρούσα εργασία αναπτύσσονται οι συνεχείς και διακριτές χρονολογικές σειρές καθώς επίσης και τα μέτρα των χρονολογικών σειρών.

Ασχολούμαστε με τις στάσιμες στοχαστικές διαδικασίες και τις μεθόδους ανάλυσής τους όπως είναι οι διαδικασίες κινητού μέσου τάξης $p,1,2$, η αυτοπαλίνδρομη διαδικασία τάξης $p,1,2$, η μεθοδολογία Box-Jenkins και τα μοντέλα ARMA και ARIMA.

Τέλος, αναλύουμε τη διαδικασία επεξεργασίας των δεδομένων των ενεργειακών καταναλώσεων Ευρωπαϊκών χωρών όπως είναι η Ελλάδα, η Αυστρία και η Γαλλία καθώς και τα αποτελέσματα στα οποία καταλήγουμε με τη χρήση του προγράμματος PASW 18 το οποίο είναι ένα πρόγραμμα που ενδείκνυται για τη στατιστική ανάλυση δεδομένων και προσφέρει στο χρήστη δυνατότητες για δημιουργία αναφορών, ανάλυση και μοντελοποίηση δεδομένων καθώς και για γραφική αναπαράστασή τους

Λέξεις κλειδιά: χρονολογικές σειρές, αυτοπαλίνδρομη διαδικασία, στοχαστικές διαδικασίες, μεθοδολογία Box-Jenkins, ARIMA, PASW 18

Key words: time series, autoregressive process, stochastic processes, Box-Jenkins methodology, ARIMA, PASW 18

Κεφάλαιο 1^ο

Εισαγωγή

Κεφάλαιο 1^ο : Εισαγωγή

Θεωρία και πράξη είναι δυο αντιπαραβαλλόμενες έννοιες που ταλάντευσαν τους επιστήμονες από τα αρχαία χρόνια. Στις θετικές επιστήμες, και όχι μόνο, αντικειμενικός στόχος είναι αυτές οι δυο έννοιες να έρθουν όσο το δυνατόν πιο κοντά. Η έλευση των υπολογιστικών συστημάτων από το 1964 και έπειτα βοήθησε πολύ στο να πραγματοποιηθούν μεγάλα άλματα προς αυτή την κατεύθυνση. Αυτό έγινε κυρίως με τη χρήση τεχνικών όπως η προσομοίωση και η πρόβλεψη. Συνδεδεικμένος κρίκος της θεωρίας και της πράξης είναι η δυνατότητα των μαθηματικών εξισώσεων να χρησιμοποιούν στοιχεία του 'παρελθόντος'. Με άλλα λόγια, η δυνατότητα κατανόησης του παρελθόντος να μας οδηγήσει στην πρόβλεψη του μέλλοντος.

Με την εξέλιξη της τεχνολογίας σε μεγάλο βαθμό βασικός στόχος κάθε ανθρώπινης δραστηριότητας είναι η διενέργεια προβλέψεων, δηλαδή η χρησιμοποίηση ενός εκτιμώμενου μοντέλου για την πρόβλεψη των μελλοντικών τιμών των μεγεθών. Η πρόβλεψη είναι απαραίτητη γιατί αποτελεί το βασικό εργαλείο για κάθε μελλοντική εξέλιξη και απόφαση. Παράλληλα οι προβλέψεις που γίνονται πρέπει να είναι έγκυρες και όσο πιο κοντά στο μέλλον που θα προκύψει. Η ποιότητα της πρόβλεψης εξαρτάται κυρίως από τον τρόπο που συλλέχθηκαν και αναλύθηκαν οι πληροφορίες. Έτσι, η ανάγκη για έγκυρες, άρα και πρακτικά χρήσιμες προβλέψεις οδήγησε στην ανάπτυξη πολλών τεχνικών και μεθόδων συλλογής και επεξεργασίας διαφόρων πρωτογενών πληροφοριών.

Τέτοιες μέθοδοι λοιπόν, προβλέψεων μπορεί να είναι ποιοτικές ή ποσοτικές. Με τις ποιοτικές μεθόδους ασχολούνται επιστήμες όπως η φιλοσοφία, θεολογία, οικονομία, πολιτικές επιστήμες κ.α., οι οποίες για την ανάλυση των δεδομένων χρησιμοποιούν την 'ανθρώπινη' λογική και την παρατήρηση. Με τον όρο 'ανθρώπινη' λογική εννοείται κάθε απόφαση που λαμβάνει κανείς χωρίς τη χρήση μαθηματικών εξισώσεων. Για παράδειγμα η εκτίμηση για την εξέλιξη μιας γλώσσας, της πολιτικής κατάστασης μιας χώρας, την πορεία της Ευρωπαϊκής Ένωσης και άλλα, είναι αποτέλεσμα λογικών διεργασιών σε διάφορα φόρουμ σκέψης. Σε όλες τις παραπάνω προβλέψεις δεν χρειάζεται καμία μαθηματική επεξεργασία. Αντίθετα, στις ποσοτικές μεθόδους η ανάλυση των δεδομένων γίνεται με τη βοήθεια μαθηματικών εκφράσεων υποστηριζόμενες από υπολογιστές. Μια από αυτές τις ποσοτικές μεθόδους, είναι η ανάλυση χρονολογικών σειρών.

Με τον όρο χρονολογικές σειρές εννοούμε μια σειρά από παρατηρήσεις που λαμβάνονται σε ορισμένες χρονικές στιγμές ή περιόδους που ισαπέχουν ή όχι μεταξύ τους. Έτσι, η τρέχουσα τιμή μιας μεταβλητής Y εκφράζεται ως συνάρτηση των προηγούμενων τιμών της, δηλαδή των τιμών της με χρονική υστέρηση.



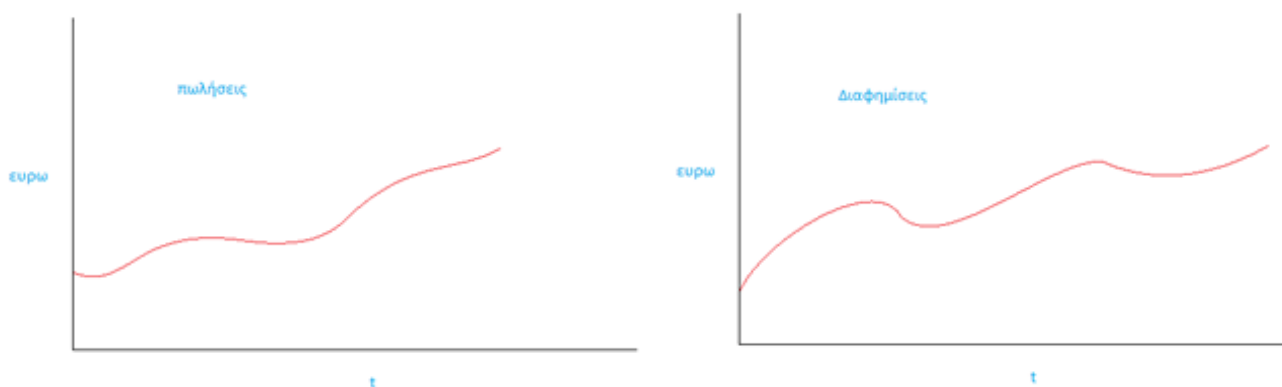
Σχήμα 1: Η αξία και ο όγκος μια μετοχής κατά την διάρκεια των ετών 2003 – 2008 [1]

Η ανάπτυξη και η χρησιμοποίηση τέτοιων μοντέλων υπήρξε ραγδαία τις τελευταίες τρεις δεκαετίες, ιδίως μετά τη δημοσίευση της εργασίας των Box&Jenkins. Οι προβλέψεις με χρονολογικές σειρές αποδείχτηκαν πολλές φορές ανώτερες άλλων παλαιότερων μεθόδων. Επίσης, προβλέψεις με χρονολογικές σειρές μπορούν να γίνουν σε ένα μεγάλο πεδίο επιστημών, από την οικονομετρία, την μηχανική, την φυσική, την ιατρική και πολλά άλλα. Βέβαια, στο σημείο αυτό είναι αναγκαίο να τονιστεί ότι οι χρονολογικές σειρές δεν χρησιμοποιούνται αποκλειστικά και μόνο για την εκτίμηση μελλοντικών τιμών, άλλα είναι ένα πολύ χρήσιμο εργαλείο για την ανάλυση και καλύτερη κατανόηση αυτού κάθε αυτού του φαινομένου. Οι χρονολογικές σειρές καταγράφουν την ιστορία των φαινομένων και όπως είναι γνωστό η ιστορία είναι αυτή που διδάσκει και βοηθά για την καλύτερη κατανόηση των γεγονότων. Η ιστορία είναι αυτή, που με την κατάλληλη ανάλυση μπορεί να βοηθήσει για την καλύτερη εκτίμηση του μέλλοντος.

Στόχος λοιπόν των χρονολογικών σειρών είναι, 'ενώνοντας' όλα τα γεγονότα που έχουν συλλεχθεί στο χρόνο, να δημιουργηθεί ένα χρονοδιάγραμμα, η μελέτη του οποίου θα δώσει μια γενική εικόνα της διαχρονικής εξέλιξης των φαινομένων ή χαρακτηριστικών. Με άλλα λόγια η ανάλυση χρονολογικών σειρών χρησιμοποιείται για να καθοριστούν μοντέλα που μετατρέπουν πληροφορίες από κανονικά χρονικά διαστήματα σε στατιστικά μέτρα. Οι κυριότερες μέθοδοι ανάλυσης χρονοσειρών είναι η μέθοδος της αυτοσυσχέτισης (auto correlation) όπου η χρονοσειρά αναπαριστάται με ένα δυναμικό μοντέλο (μοντέλο ARIMA, διαδικασία Box&Jenkins) όπου οι παρατηρήσεις θεωρούνται ως συναρτήσεις του παρελθόντος τους (και πιθανόν του παρελθόντος και άλλων μετρούμενων ή παρατηρούμενων μεταβλητών) $x_t = f(x_{t-1}, x_{t-2}, x_{t-3}, \dots)$ και η μέθοδος της φασματικής ανάλυσης (spectral analysis) όπου αναπαριστά τη χρονοσειρά με ένα κινητικό μοντέλο που οι παρατηρήσεις θεωρούνται συναρτήσεις του χρόνου $x_t = f(t)$. Στη παρούσα διπλωματική εργασία θα παρουσιαστούν οι μέθοδοι αυτοσυσχέτισης.

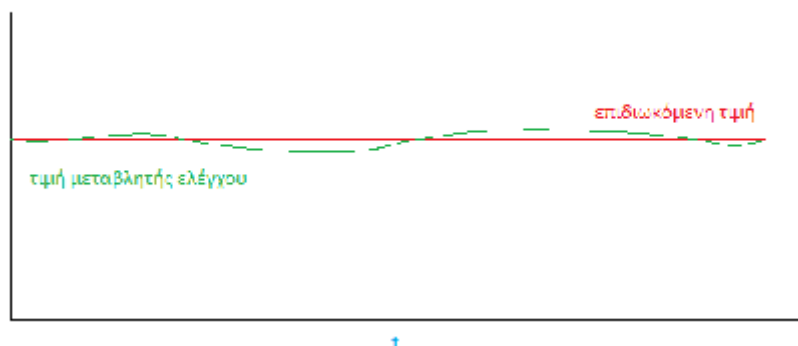
Συμπερασματικά, γίνεται κατανοητό από τα παραπάνω, ότι σκοποί της ανάλυσης χρονολογικών σειρών είναι η περιγραφή μιας χρονολογικής σειράς, η εξήγησή της, η πρόγνωση, ο σχεδιασμός, ο έλεγχος και τέλος η κατανόηση της διαδικασίας. Η περιγραφή μιας χρονολογικής σειράς μπορεί να γίνει με μια απλή παρατήρηση της γραφικής της παράστασης 'με το μάτι', όπου ανακαλύπτονται περιοδικότητες, τάσεις και άλλα. Επίσης, υπάρχουν περισσότερο σύνθετες και προχωρημένες περιγραφές που γίνονται με την βοήθεια των στοχαστικών μοντέλων. Η εξήγηση με τη σειρά της, είναι η συσχέτιση δυο ή περισσότερων χρονολογικών σειρών μεταξύ τους και η δυνατότητα η μια να εξηγεί την άλλη.

Στο παράδειγμα παρακάτω γίνεται η προσπάθεια να ερμηνευτεί κατά πόσο οι πωλήσεις συσχετίζονται με τις διαφημίσεις και αντίστροφα.



Σχήμα 2: Ο όγκος των διαφημίσεων και των πωλήσεων μιας εταιρίας σε € για ένα χρονικό διάστημα. Παρατηρούμε ότι όταν μειώνεται ο όγκος των πωλήσεων αυξάνεται ο όγκος των διαφημίσεων, για την καλύτερη προώθηση των προϊόντων.[1]

Επίσης, σημαντικός σκοπός στην ανάλυση των χρονολογικών σειρών είναι ο έλεγχος. Παραδείγματα στη σπουδαιότητα του ελέγχου μπορούν να παρουσιαστούν στις διαδικασίες παραγωγής και ποιότητας ενός προϊόντος.

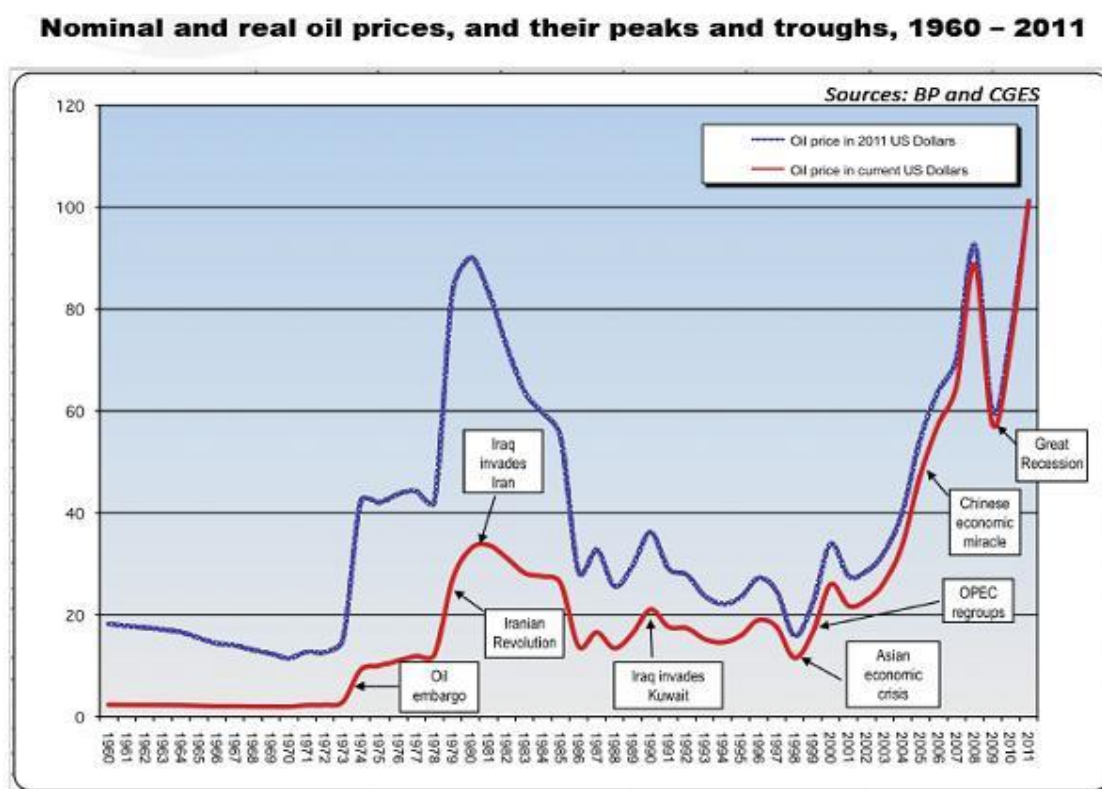


Σχήμα 3: Η τιμή της ποσότητας ζάχαρης και η προδιαγραφόμενη ποσότητα για την παρασκευή ενός γλυκίσματος σε ένα χρονικό διάστημα[1]

Στο παραπάνω σχήμα φαίνεται η ποσότητα ζάχαρης που χρησιμοποιείται για την παρασκευή ενός γλυκού σε μια βιομηχανία τροφίμων και πώς πραγματοποιείται ο έλεγχος με τη βοήθεια των χρονολογικών σειρών.

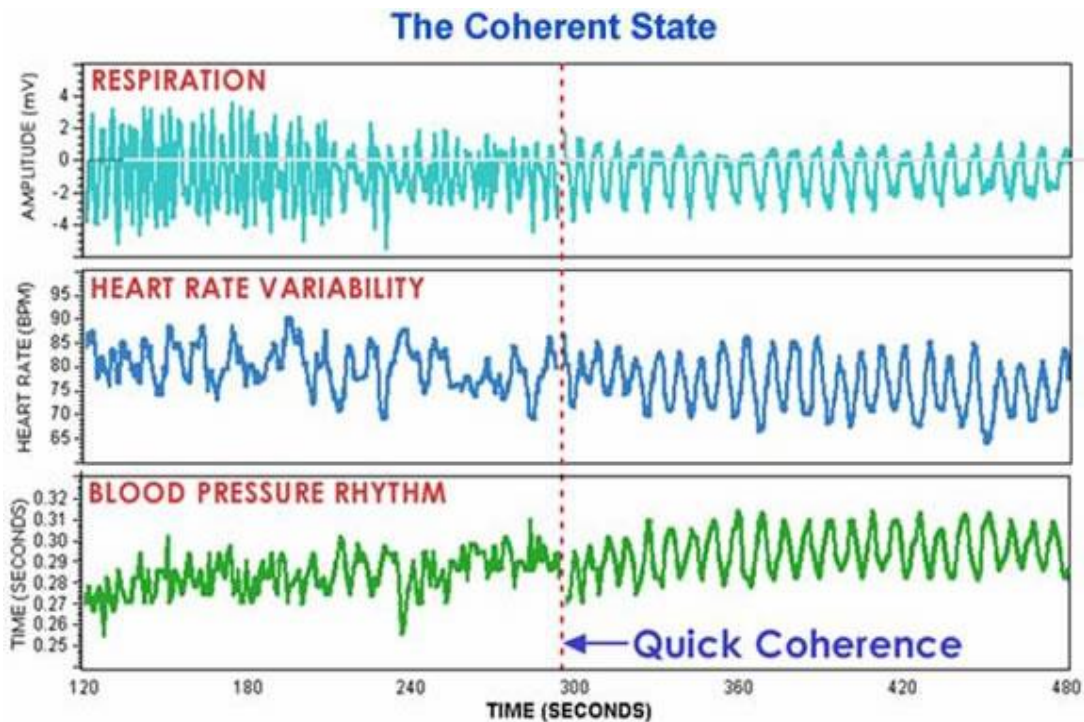
Τέλος, βασικός σκοπός της ανάλυσης χρονολογικών σειρών είναι η κατανόηση της διαδικασίας. Την κατανόηση παρέχει η στατιστική δίνοντας φορμαλιστικές περιγραφές των χρονοσειρών αυτών καθαυτών. Έτσι, μπορεί να παραχθεί μια χρονοσειρά από ένα άγνωστο σύστημα και συνεπώς να κατανοηθεί καλύτερα το ίδιο το σύστημα που έχει παράξει την χρονοσειρά αυτή. Για παράδειγμα μπορεί να γίνει κατανοητό εάν ένα σύστημα είναι περιοδικό και να βρεθεί η περίοδος του.

Παραδείγματα χρονολογικών σειρών μπορούν να παρουσιαστούν και στην Οικονομία. Τέτοια μπορεί να είναι η εξέλιξη της τιμής του πετρελαίου, η αξία μιας μετοχής, ο δείκτης συναλλαγών στο χρηματιστήριο και άλλα.



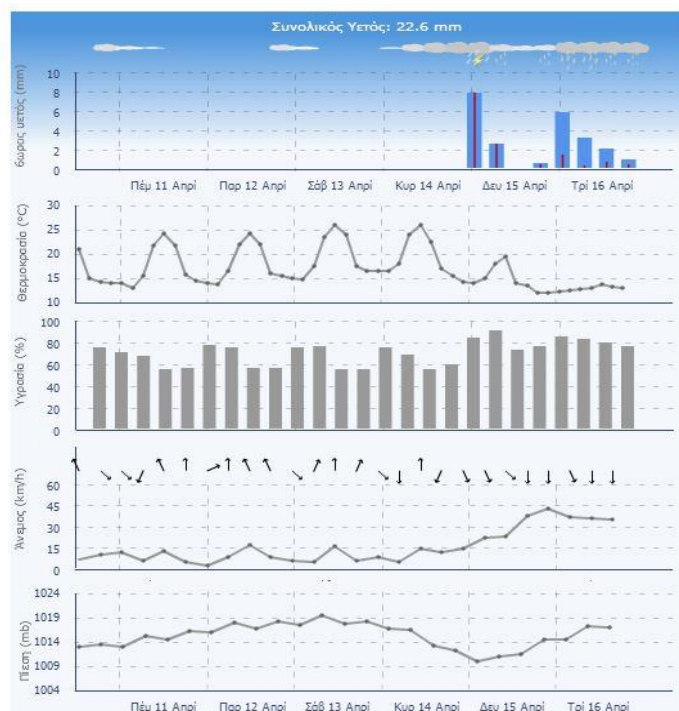
Σχήμα 4: Η τιμή του δείκτη αργού πετρελαίου απο το 1990-2011[1]

Ένα παράδειγμα χρονολογικών σειρών στο εμπόριο (marketing) επίσης είναι οι πωλήσεις ανά μήνα. Βασικός σκοπός στο εμπόριο εκτός της πρόβλεψης είναι ο σχεδιασμός, η σύγκριση και η ερμηνεία. Επίσης, παραδείγματα χρονολογικών σειρών έχουμε από την δημογραφία όπου μελετώνται οι εξελίξεις πληθυσμών. Από την ιατρική αναλύονται τα ηλεκτροεγκεφαλογραφήματα (ΗΕΓ, ΕΕΓ) και ηλεκτροκαρδιογραφήματα (ΗΚΓ, ΕCG)(Σχήμα 5). Εδώ αποσκοπούμε στην αναγνώριση ανωμαλιών, στην κατανόηση και στην εξήγηση της δυναμικής της καρδιάς και του εγκεφάλου και φυσικά στην πρόβλεψη διαταραχών.



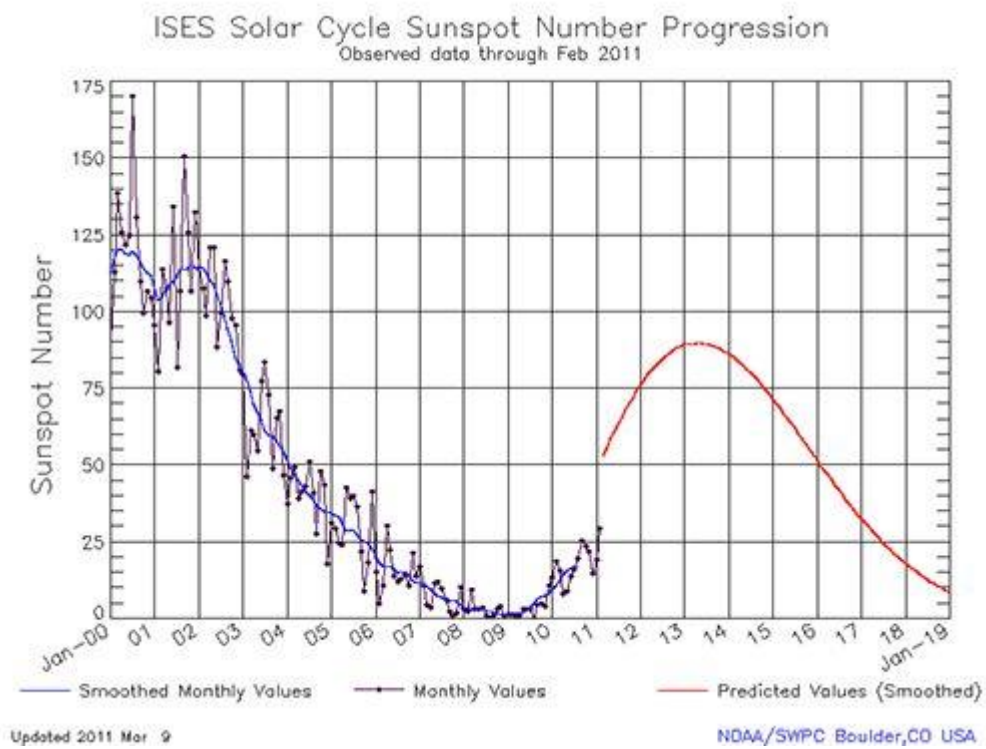
Σχήμα 5: Ηλεκτροεγκεφαλογράφημα. Ηλεκτροεγκεφαλικοί παλμοί για περίοδο 480 sec.[1]

Ακόμα, χρονολογικές σειρές συναντάμε στη φυσική, στη σεισμολογία, στη μετεωρολογία, στην αστροφυσική και αλλού, με αντικειμενικό στόχο την κατανόηση και την πρόβλεψη μετεωρολογικών φαινομένων ή άλλων φυσικών φαινομένων.



Σχήμα 6. Μετεώγραμμα για την Αθήνα[1]

Στην αστροφυσική με την ανάλυση χρονοσειρών μπορεί να περιγραφεί το μοντέλο και να γίνει σύγκριση με τα θεωρητικά μοντέλα που ήδη υπάρχουν (Σχήμα 7). Έτσι, με την κατανόηση των διαδικασιών που παράγουν τις χρονοσειρές καθοδηγούμαστε στη δημιουργία νέων μοντέλων. Τα νέα αυτά μοντέλα χρησιμοποιούνται μετά για προβλέψεις.



Σχήμα 7: Ακριβής και προβλεπόμενος αριθμός ηλιακών κηλίδων σαν συνάρτηση του χρόνου. Ο 24ος κύκλος ξεκινά τον Ιανουάριο του 2008.[1]

Κεφάλαιο 2^ο

Βασικές έννοιες

Κεφάλαιο 2ο : Βασικές Έννοιες

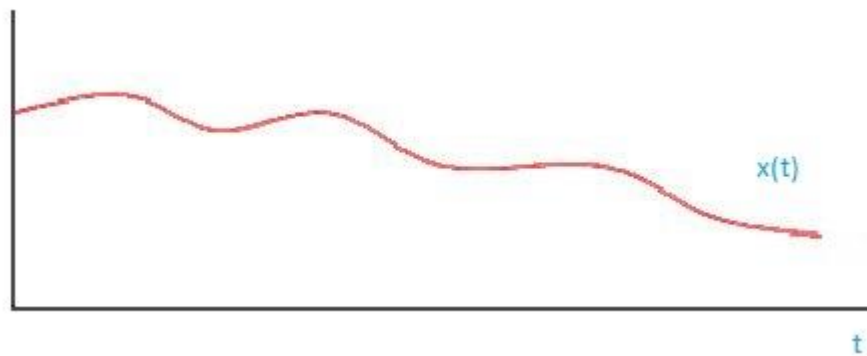
2.1 Στοχαστική Διαδικασία και Χρονολογικές Σειρές

Με τον όρο χρονολογική σειρά εννοούμε μια σειρά από παρατηρήσεις που λαμβάνονται σε ορισμένες χρονικές στιγμές ή περιόδους. Άρα χρονολογική σειρά είναι ένα δείγμα $Y_1, Y_2, Y_3, \dots, Y_T$ όπου ο δείκτης παριστάνει ισαπέχοντα ή μη, χρονικά σημεία ή διαστήματα, υποθέτοντας, ότι οι παρατηρήσεις $y_1, y_2, y_3, \dots, y_T$ είναι συγκεκριμένες τιμές ή συγκεκριμένες πραγματοποιήσεις των τυχαίων μεταβλητών $Y_1, Y_2, Y_3, \dots, Y_T$ και ότι επιπλέον οι τυχαίες μεταβλητές αυτές $Y_1, Y_2, Y_3, \dots, Y_T$ είναι μέρος μιας άπειρης σειράς τυχαίων μεταβλητών. Η άπειρη αυτή ακολουθία των τυχαίων μεταβλητών ονομάζεται στοχαστική ή τυχαία διαδικασία ή στοχαστική ανέλιξη και παριστάνεται ως $\{Y_t\}$.

Γενικά όπως και στην περίπτωση T τυχαίων μεταβλητών, μια στοχαστική διαδικασία μπορεί να περιγραφεί από μια συνάρτηση πιθανότητας $f(y_1, y_2, y_3, \dots, y_T)$. Εάν ήταν γνωστή η συνάρτηση πιθανότητας, τότε θα ήταν εύκολο να υπολογιστεί, για παράδειγμα, η πιθανότητα μιας συγκεκριμένης πραγματοποίησης ή η πιθανότητα μιας μελλοντικής τιμής. Ωστόσο η συνάρτηση πιθανότητας δεν είναι γνωστή, και η πλήρης εξειδίκευση της μορφής της είναι αδύνατη, σκοπός της ανάλυσης χρονολογικών σειρών είναι η διατύπωση μοντέλων που να μπορούν να περιγράψουν το μηχανισμό της στοχαστικής διαδικασίας από την οποία πρόεκυψε η συγκεκριμένη χρονολογική σειρά.

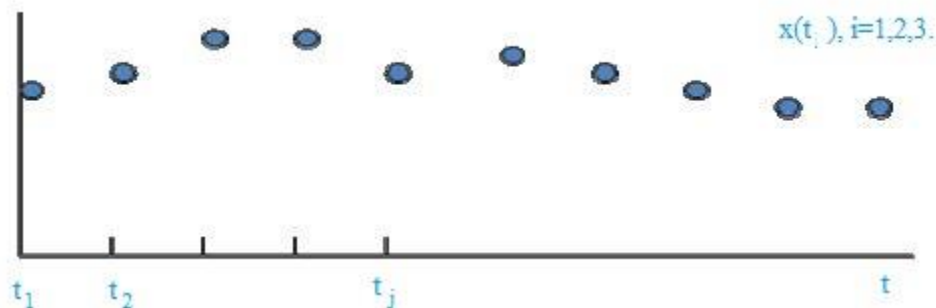
2.2 Συνεχείς - Διακριτές Χρονολογικές Σειρές.

Οι χρονολογικές σειρές διακρίνονται σε **συνεχείς** και σε **διακριτές χρονολογικές σειρές**. **Συνεχείς (continuous)** χρονολογικές σειρές είναι αυτές στις οποίες η τιμή του φαινομένου $X(t)$ παρατηρείται συνεχώς. Παράδειγμα συνεχών χρονολογικών σειρών είναι η συνεχής καταγραφή της θερμοκρασίας του αέρα ή η συνεχής παρακολούθηση των σεισμών.



Σχήμα 1. Συνεχής χρονολογική σειρά.[1]

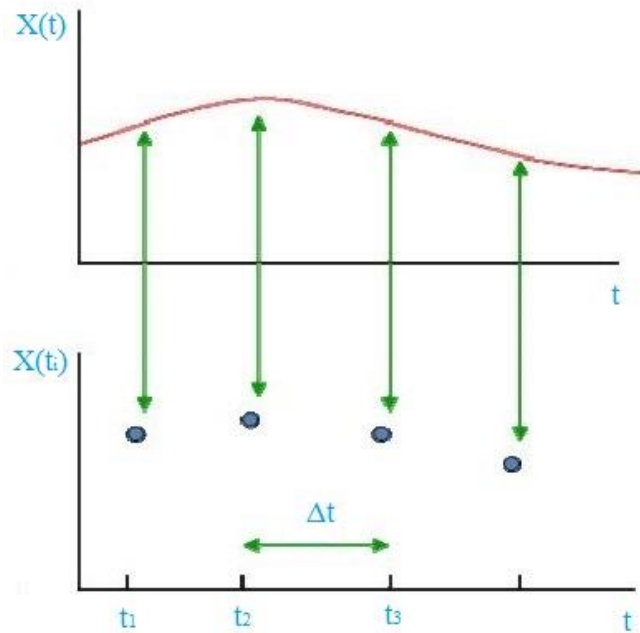
Διακριτές (discrete) χρονολογικές σειρές είναι αυτές όπου η τιμή του φαινομένου $X(t_i)$ καταγράφεται σε ορισμένα χρονικά διαστήματα $\Delta t, t_i = i\Delta t$.



Σχήμα 2. Διακριτή χρονολογική σειρά[1]

Παράδειγμα διακριτών χρονολογικών σειρών είναι η τιμή μιας μετοχής ανά ημέρα ή ο αριθμός των ηλιακών κηλίδων ανά έτος όπου υπάρχουν τιμές σε συγκεκριμένα χρονικά διαστήματα.

Οι διακριτές χρονολογικές σειρές είναι αυτές που μπορούν να ‘κατανοηθούν’ από έναν Η/Υ κατά συνέπεια αντικειμενικός στόχος είναι οι συνεχείς χρονοσειρές να μετατραπούν σε διακριτές. Η διαδικασία μετατροπής μιας συνεχούς χρονολογικής σειράς σε διακριτή ονομάζεται διακριτοποίηση ή δειγματοληψία (Sampling, read off, digitize). Πρόκειται ουσιαστικά για μια διαδικασία κατά την οποία διαβάζοντας μια συνεχή χρονολογική σειρά κρατάμε τιμές μόνο σε σημεία που απέχουν ορισμένη χρονική απόσταση Δt μεταξύ τους ή μετρώντας εξ αρχής μόνο σε διακριτές χρονικές στιγμές (Συνήθως αποφεύγεται η επιλογή μη σταθερού Δt καθώς δημιουργεί αρκετές δυσκολίες).



Σχήμα 3. Μετατροπή Συνεχούς χρονολογικής σειράς σε Διακριτή.[1]

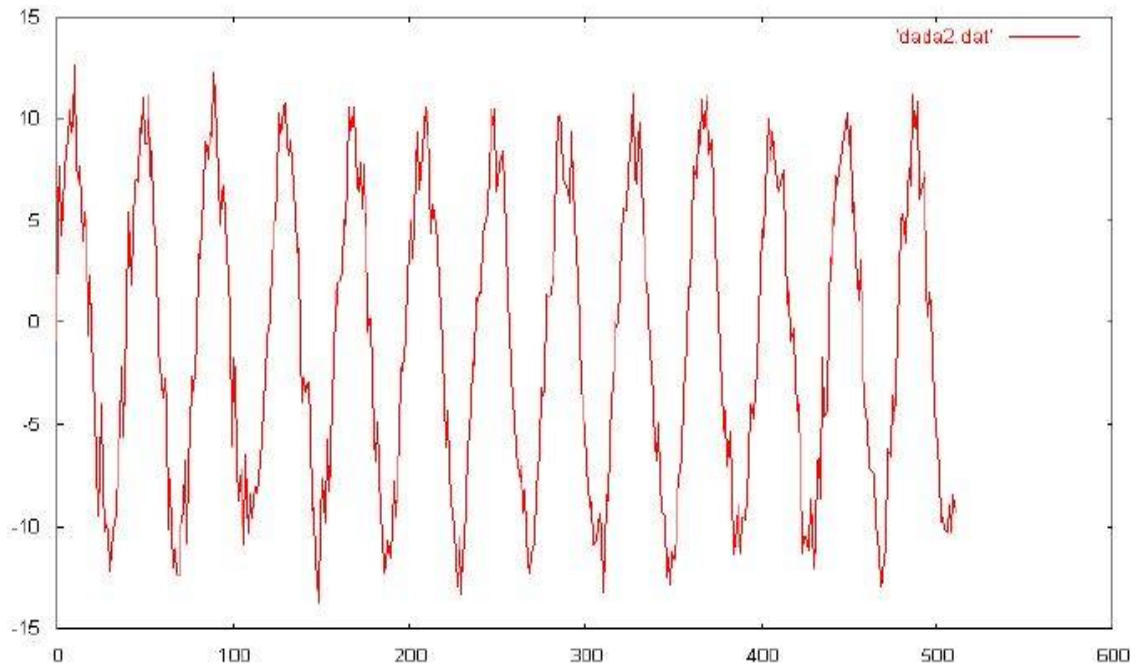
Ορισμένες φορές στην διαδικασία μετατροπής μιας συνεχούς χρονολογικής σειράς σε διακριτή είναι χρήσιμο να μην επιλέγεται απλά η τιμή της συνεχούς χρονολογικής σειράς στην χρονική στιγμή αυτή, αλλά το $X(t_i)$ να είναι το άθροισμα ή το ολοκλήρωμα για όλο το χρονικό διάστημα Δt . Για παράδειγμα, όταν το $X(t_i)$ είναι η βροχή σε τετραγωνικά χιλιοστά mm^2 ανά ημέρα, λαμβάνεται το άθροισμα όλων των βροχοπτώσεων για κάποιο συγκεκριμένο χρονικό διάστημα, οπότε σαν $X(t_i)$ λαμβάνεται $X(t_i) = \frac{1}{\Delta t} \int_{t_{i-1}}^{t_i} f(t) dt$ (2.1) όπου $f(t)$ η συνεχής βροχόπτωση και $t_i - t_{i-1} = \Delta t$

Συνήθως στην ανάλυση χρονολογικών σειρών οι διαδοχικές παρατηρήσεις δεν είναι ανεξάρτητες, άρα πρέπει να λαμβάνεται υπ' όψη η σειρά των παρατηρήσεων. Ακριβώς αυτή η εξάρτηση επιτρέπει την πρόγνωση του μέλλοντος με βάση το παρελθόν.

2.3 Μέτρα Χρονολογικών Σειρών

Μερικές ακόμη βασικές έννοιες των χρονολογικών σειρών είναι **η γραφική παράσταση, ο μέσος όρος και η διασπορά αυτής, ο κινητός μέσος όρος, η τάση και η στασιμότητα.**

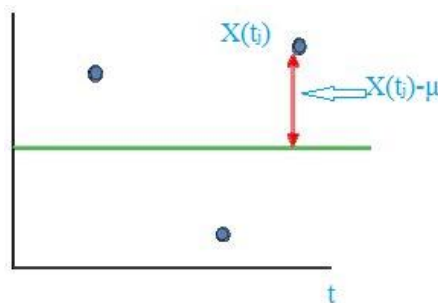
Με την έννοια **γραφική παράσταση** μιας χρονολογικής σειράς εννοούμε την καμπύλη που παράγεται σε ένα σύστημα ορθογώνιων αξόνων όπου ο οριζόντιος άξονας είναι ο χρόνος και ο κατακόρυφος οι μετρούμενες τιμές στο αντίστοιχο χρονικό διάστημα. Συμβολίζοντας με X_i τις N χρονικές στιγμές και με Y_i τις τιμές των αντίστοιχων παρατηρήσεων, δημιουργούμε N ζεύγη της μορφής (X_i, Y_i) τα οποία μπορούμε να παραστήσουμε σε ένα σύστημα αξόνων.



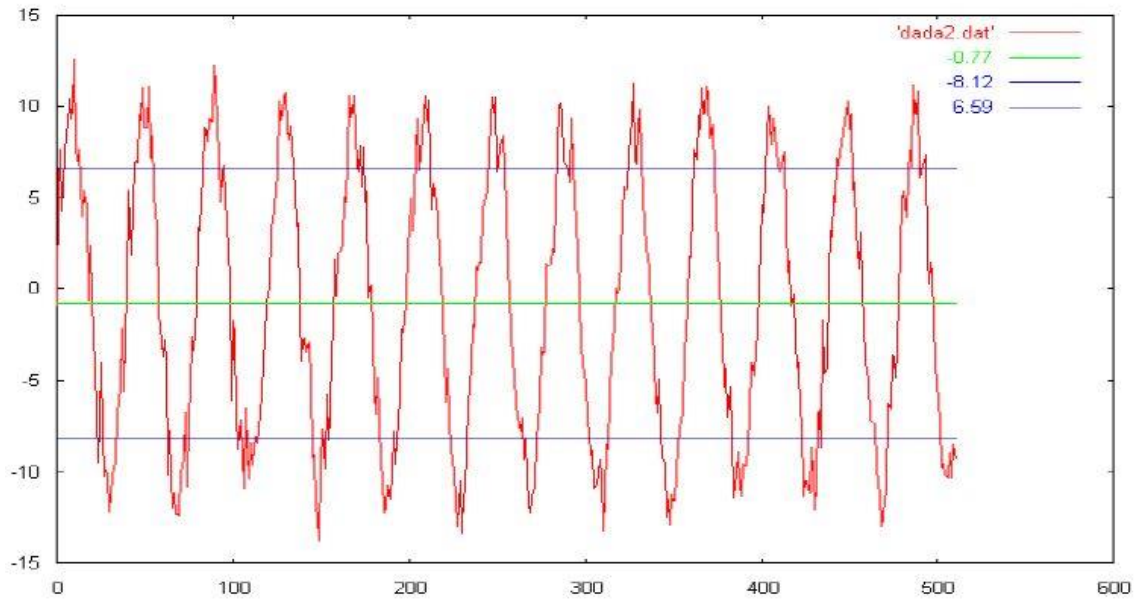
Σχήμα 4. Γραφική παράσταση μιας τυχαίας χρονολογικής σειράς η οποία παρουσιάζει περιοδικότητα.[1]

Με την απλή παρατήρηση μια γραφικής παράστασης μπορούν εύκολα να προκύψουν μερικά γρήγορα συμπεράσματα. Στο παράδειγμα του **σχήματος 4** βλέπουμε μια χρονολογική σειρά που δείχνει να είναι περιοδική με θόρυβο. Έστω η χρονολογική σειρά $X(t_i), i = 1, 2, 3, \dots, N$. Εύκολα βγάζει κανείς το συμπέρασμα ότι η σειρά των $X(t_i)$ δεν παίζει ρόλο, αυτό που μας ενδιαφέρει είναι ο **μέσος όρος μ** μιας χρονολογικής σειράς που είναι η μέση τιμή όλων των τιμών της, $\mu = \frac{1}{N} \sum_{i=1}^N X(t_i)$ (2.2)

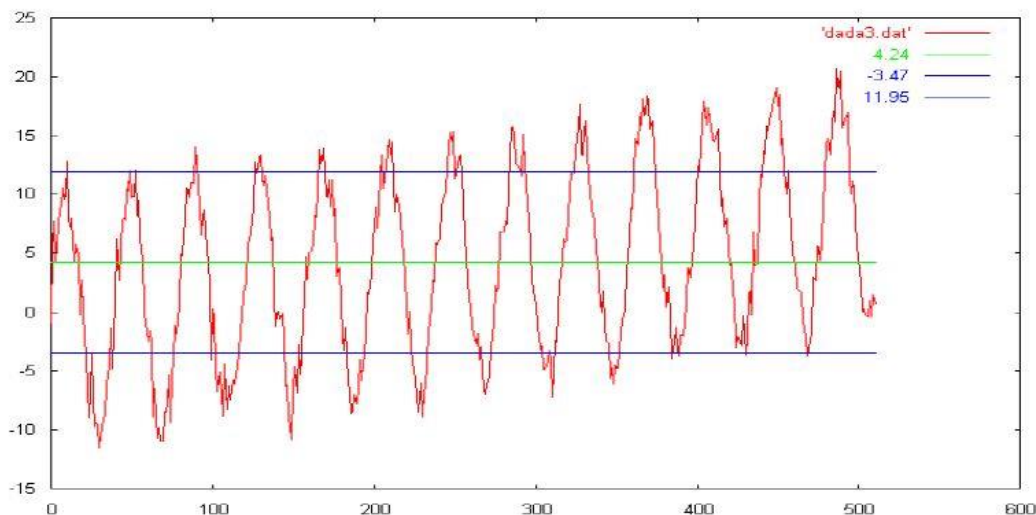
Διασπορά σ^2 μιας χρονολογικής σειράς είναι ο μέσος όρος των τετραγώνων των αποκλίσεων από την μέση τιμή, $\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (X(t_i) - \mu)^2$ (2.3)



Τυπική απόκλιση σ μιας χρονολογικής σειράς είναι η τετραγωνική ρίζα της διασποράς. Η τυπική απόκλιση είναι περισσότερο διαισθητική έννοια από ότι η διασπορά, $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X(t_i) - \mu)^2}$ (2.4)



Σχήμα 5. Γραφική παράσταση τυχαίας χρονολογικής σειράς, ο μέσος όρος και η τυπική απόκλιση αυτής.[1]

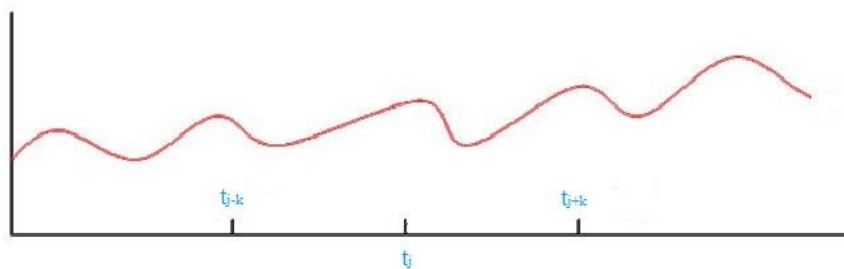


Σχήμα 6. Γραφική παράσταση χρονολογικής σειράς η οποία παρουσιάζει τάση.[1]

Μεταξύ των διαστημάτων $\mu - \sigma$ και $\mu + \sigma$ βρίσκονται τα περισσότερα σημεία της χρονολογικής σειράς και μέσα από το διάστημα αυτό προκύπτει και η διακύμανση των τιμών της. Υπάρχουν όμως χρονολογικές σειρές όπου τα μέτρα μ και $\mu \pm \sigma$ δεν δίνουν καλή περιγραφή της πραγματικής μέσης τιμής.

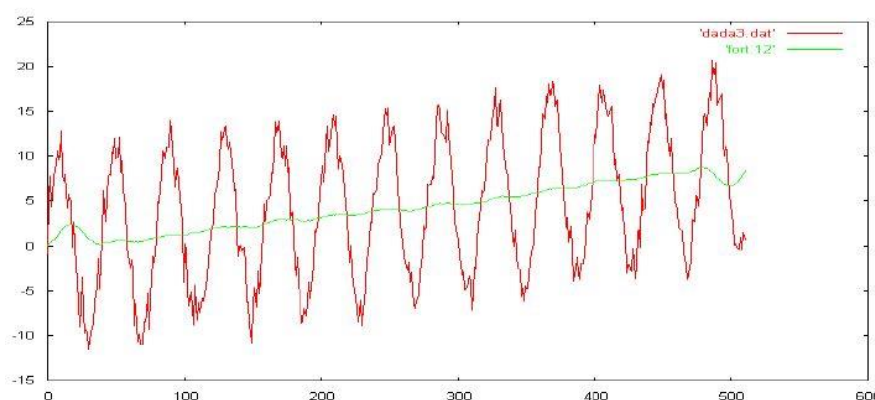
Αυτό συμβαίνει λόγω του ότι υπάρχει μια τάση και για αυτό η πραγματική μέση τιμή μ αλλάζει με τον χρόνο, δηλαδή η μέση τιμή είναι συνάρτηση του χρόνου $\mu = \mu(t)$. Έτσι, προκύπτει η ανάγκη να κατασκευαστεί ένα μοντέλο για την τάση. Ο πρώτος τρόπος για την κατασκευή αυτού του μοντέλου είναι να δεχτούμε αυθαίρετα ότι η τάση ακολουθεί κάποια γνωστή κατανομή. Στο **σχήμα 6**, σαν μια πρώτη προσέγγιση μπορούμε να δεχθούμε ότι η μέση τιμή μεταβάλλεται γραμμικά ως προς τον χρόνο $\mu(t) = a + bt$.

Συστηματοποιώντας την διαδικασία κατασκευής μοντέλου για την τάση κατασκευάστηκε ο **κινητός μέσος όρος (running mean, moving average)**.



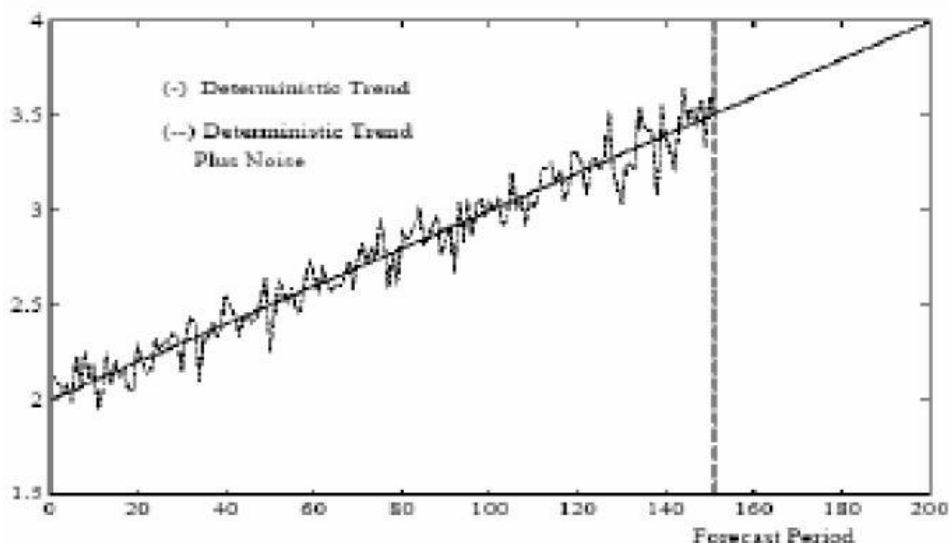
Έτσι, σαν κινητό μέσο όρο ορίζουμε $\mu(t_i, K) := \frac{1}{2K+1} \sum_{j=i-K}^{i+K} X(t_j)$ (2.5)

όπου, K είναι το μήκος του παραθύρου. Στο **σχήμα 7** φαίνεται ο κινητός μέσος όρος όπως αυτός ορίστηκε παραπάνω.



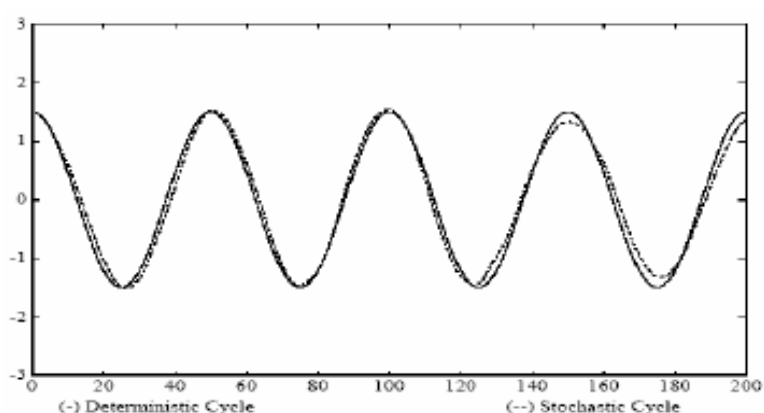
Σχήμα 7. Ο κινητός μέσος όρος για τυχαία χρονολογική σειρά.[1]

Στις χρονολογικές σειρές δεν υπάρχουν μόνο περιπτώσεις της παραπάνω συμπεριφοράς τάσης. Εκτός, λοιπόν, από την **μακροχρόνια τάση (Secular Trend)(σχήμα 8)**, στις χρονολογικές σειρές συναντάμε είτε κυκλική κύμανση αυτών, είτε περιοδική – εποχιακή μεταβολή, είτε ακανόνιστη μεταβολή. Συνήθως, στην πράξη συναντάμε περισσότερα από ένα χαρακτηριστικά. Στη μακροχρόνια τάση η τιμή της μεταβλητής τείνει να αυξηθεί ή να ελαττωθεί για μεγάλο χρονικό διάστημα.

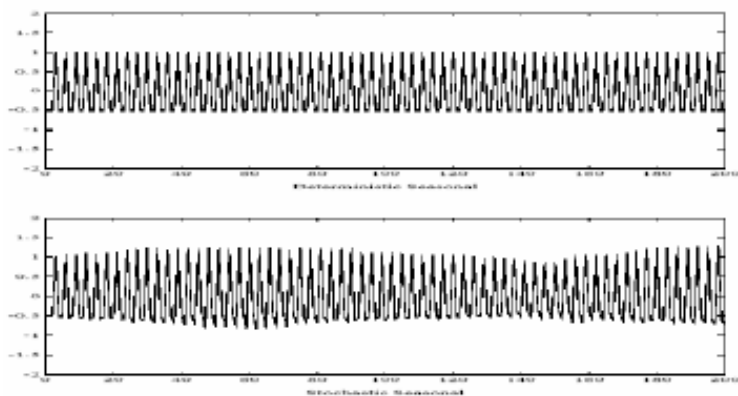


Σχήμα 8. Χρονολογική σειρά που παρουσιάζει μακροχρόνια τάση.[1]

Η μακροχρόνια τάση είναι από τις περισσότερο εύκολα αντιμετωπίσιμες περιπτώσεις γιατί η τάση εκφράζει τη χρονολογική σειρά για μια εκτεταμένη περίοδο. Οι χρονολογικές σειρές που έχουν μελετηθεί περισσότερο είναι αυτές με μακροχρόνια τάση γιατί η κατανόηση της ιστορίας της μεταβλητής είναι εύκολη με συνέπεια την πρόβλεψη μελλοντικών τιμών. Ορισμένες μέθοδοι προσδιορισμού της μακροχρόνιας τάσης είναι: 1)η μέθοδος των δυο μέσων σημείων, 2)η μέθοδος των κινητών μέσων, 3)η μέθοδος της ευθείας ελαχίστων τετραγώνων και 4) η μέθοδος της καμπύλης ελαχίστων τετραγώνων. Η κυκλική κύμανση που παρουσιάζουν περιπτώσεις χρονολογικών σειρών έγκειται στο γεγονός ότι παρατηρούνται αυξομειώσεις της τιμής της μεταβλητής γύρω από γραμμική τάση σε μια μακροχρόνια περίοδο. Στην πράξη τα σημεία της χρονολογικής σειράς για μια χρονική σειρά παρατηρήσεων βρίσκονται κάτω από τη γραμμική τάση και στη συνέχεια για άλλη χρονική σειρά τιμών πάνω από τη γραμμική τάση. Ο χρόνος μιας κυκλικής αυξομείωσης δεν είναι σταθερός.



Σχήμα 9. Χρονολογική σειρά που παρουσιάζει κυκλική κύμανση.[2]



Σχήμα 10. Χρονολογική σειρά που παρουσιάζει περιοδικότητα.[2]

Στην εφαρμογή οι κυκλικές αυξομειώσεις είναι οι πλέον δύσκολες να αντιμετωπιστούν γιατί η κυκλική κίνηση δεν ακολουθεί κανένα κανονικό μοντέλο αλλά έχει απρόβλεπτη κίνηση. Αντίθετα, από τις χρονολογικές σειρές που παρουσιάζουν κυκλικές μεταβολές (σχήμα 9) και είναι δύσκολα αντιμετωπίσιμες, οι χρονολογικές σειρές που παρουσιάζουν περιοδικές μεταβολές (σχήμα 10) είναι πολύ χρήσιμες γιατί ακολουθούν ένα κανονικό μοντέλο και έτσι μπορούν να δώσουν αξιόπιστες προβλέψεις για το μέλλον.

Τέλος, υπάρχουν και οι χρονολογικές σειρές που παρουσιάζουν ακανόνιστες μεταβολές που άλλοτε είναι μικρές, άλλοτε μεγάλες, θετικές ή αρνητικές χωρίς καμία κανονικότητα. Οι μεταβολές αυτές διακρίνονται σε συμπτωματικές και τυχαίες.

Η **στασιμότητα** είναι πολύ σημαντική έννοια καθώς είναι απαραίτητη προϋπόθεση για τα περισσότερα εργαλεία της ανάλυση χρονολογικών σειρών. Η χρονολογική σειρά μπορεί να θεωρηθεί ως στοχαστική διαδικασία πεπερασμένου πλήθους παρατηρήσεων, δηλαδή, μια πραγματοποίηση μιας διαδικασίας X_1, X_2, \dots, X_T . Στο εξής με τον όρο χρονολογική σειρά θα εννοούμε και το σύνολο των παρατηρήσεων της τυχαίας μεταβλητής.

Για να ορίσουμε ένα μοντέλο χρονολογικών σειρών, για ένα σύνολο παρατηρήσεων $\{x_t\}$ απαιτείται ο καθορισμός όλων των κοινών κατανομών μιας ακολουθίας X_t , πραγματοποίηση των οποίων πρέπει να είναι οι παρατηρήσεις $\{x_t\}$. Η αναζήτηση λοιπόν ενός μοντέλου, συνεπάγεται τον ορισμό όλων των κοινών κατανομών των (X_1, X_2, \dots, X_n) , $n=1, 2, \dots$ δηλαδή των πιθανοτήτων $P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_T \leq x_n)$ με $-\infty \leq x_1 \leq x_2 \leq \dots \leq x_n \leq \infty, n=1, 2, \dots$ μια διαδικασία η οποία είναι εξαιρετικά δύσκολη. Στη θέση της ορίζονται και χρησιμοποιούνται ροπές μικρότερης τάξης. Βασική και απαραίτητη υπόθεση ώστε να απλοποιηθούν σημαντικά προβλήματα που προκύπτουν από την μελέτη των ροπών είναι η υπόθεση της στασιμότητας.

Μια χρονολογική σειρά λέγεται **στάσιμη** εάν δεν υπάρχει συστηματική αλλαγή του μέσου όρου και της διασποράς της στο χρόνο. Με άλλα λόγια, εάν μια χρονολογική σειρά παρουσιάζει τάση τότε αυτή δεν θα είναι στάσιμη.

Μια στοχαστική διαδικασία είναι **αυστηρώς ή πλήρως στάσιμη (strictly – strongly – completely stationary)** όταν οι ιδιότητες της δεν επηρεάζονται από μια αλλαγή στην αρχή μετρήσεως του χρόνου.

Συνήθως, οι διαδικασίες αυτές λέγονται και n – τάξης στάσιμες. Αυτό σημαίνει ότι $F(y_t, y_{t+1}, \dots, y_{t+T}) = F(y_{t+s}, y_{t+1+s}, \dots, y_{t+T+s})$ για οποιοδήποτε s , δηλαδή η από κοινού συνάρτηση κατανομής με αρχή το χρονικό σημείο t είναι η ίδια με την από κοινού συνάρτηση κατανομής με αρχή το χρονικό σημείο $t+s$. Το s παριστάνει μια αυθαίρετη μετακίνηση κατά μήκος του άξονα του χρόνου είτε προς τα εμπρός είτε προς τα πίσω. Ο αυστηρός, όμως, ορισμός της στασιμότητας αναφέρεται σε όλες τις ιδιότητες μιας στοχαστικής διαδικασίας. Πράγμα αρκετά δύσκολο. Αδυνατίζοντας τον ορισμό για την στασιμότητα, παράλληλα όμως μετατρέποντάς τον σε περισσότερο εύηλο, ορίζεται η n -τάξεως ασθενώς στάσιμη διαδικασία εάν υπάρχουν οι ροπές n -τάξεως και είναι ανεξάρτητες του t . Δηλαδή, ισχύουν τα εξής

$$E(Y_t) = \mu \quad (2.6), \text{ ανεξάρτητη του } t$$

$$Var(Y_t) = \sigma^2 \quad (2.7), \text{ ανεξάρτητη του } t$$

$$Cov(Y_t, Y_{t+s}) = Cov(Y_{t+m}, Y_{t+m+s}) = \gamma_s \quad (2.8), \text{ ανεξάρτητη του } t$$

Όταν οι παραπάνω συνθήκες περιορίζονται και αναφέρονται μόνο στις πρώτες και δευτέρες ροπές μιλάμε για 2^{ης} τάξης ασθενώς στάσιμες διαδικασίες ή απλώς ασθενώς στάσιμες διαδικασίες ή κατά συνδιακύμανση στάσιμες. Από τη σχέση (2.8) είναι προφανές ότι $Cov(Y_t, Y_{t+s}) = Cov(Y_{t+s}, Y_t)$ δηλαδή $\gamma_s = \gamma_{-s}$. Επειδή Y_t και Y_{t+s} είναι παρατηρήσεις της ίδιας μεταβλητής που απέχουν χρονικά μεταξύ τους κατά s , η συνδιακύμανση $Cov(Y_t, Y_{t+s})$ αναφέρεται και ως αυτοσυνδιακύμανση. Είναι προφανές ότι για $s=0$, $\gamma_0 = Var(Y_t) = \sigma^2$. Τέλος, εάν μια χρονολογική σειρά είναι αυστηρά στάσιμη τότε θα είναι και ασθενώς στάσιμη 2^{ης} τάξης.

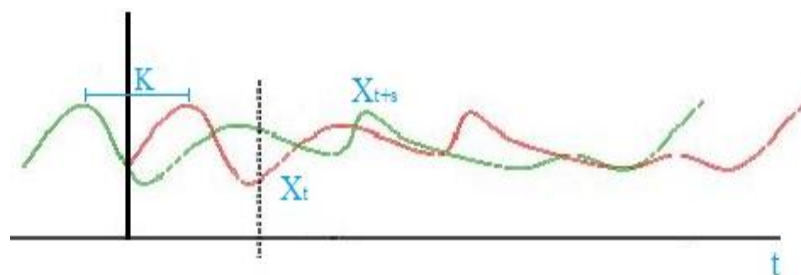
2.4 Συνάρτηση Αυτοσυσχετίσεως

Από τη θεωρία των πιθανοτήτων είναι γνωστό ότι ο λόγος της συνδιακύμανσης προς το γινόμενο των τετραγωνικών ριζών των διακυμάνσεων δυο μεταβλητών είναι ο συντελεστής συσχέτισής τους. Επίσης, ο συντελεστής συσχέτισης μας δίνει ένα μέτρο για το βαθμό της σχέσης μεταξύ δυο μεταβλητών.

Για τον συντελεστή συσχέτισης ισχύει ότι $-1 \leq \rho \leq 1$ και

- εάν $\rho = -1$ ή $\rho = 1$ τότε έχουμε την μέγιστη δυνατή συσχέτιση,
- εάν $\rho > 0$ υπάρχει θετική συσχέτιση, η όποια είναι τόσο πιο ισχυρή όσο πιο κοντά ο συντελεστής συσχέτισης είναι στο 1,
- εάν $\rho < 0$ υπάρχει αρνητική συσχέτιση, η όποια είναι τόσο πιο ισχυρή όσο πιο κοντά ο συντελεστής συσχέτισης είναι στο -1 ,
- εάν $\rho = 0$ δεν υπάρχει καμία συσχέτιση μεταξύ των δυο μεταβλητών.

Στην περίπτωση των χρονολογικών σειρών ο συντελεστής συσχέτισεως ανάμεσα στην Y_t και στην Y_{t+s} ονομάζεται συντελεστής αυτοσυσχετίσεως.



Σχήμα 11. Τα ζεύγη της χρονολογικής σειράς μετατοπισμένης κατά s . [1]

Σαν συνάρτηση αυτοσυσχέτισης (ACF) ονομάζεται η συνάρτηση

$$\rho_s(X) = \frac{\gamma_s(x)}{\gamma_0(x)} = \frac{\gamma_s}{\gamma_0} = \rho_s \quad (2.9) \quad \rho_s = \frac{\sum_{i=1}^{N-s} (X(t_i) - \mu_X)(X(t_{i+s}) - \mu_X)}{\sum_{i=1}^N (X(t_i) - \mu_X)^2} \quad (2.10)$$

όπου μ_X ο μέσος της χρονολογικής σειράς.

Δηλαδή, η σχέση που υπάρχει ανάμεσα στο συντελεστή αυτοσυσχετίσεως ρ_s και στο s . Το s μπορεί να πάρει τις τιμές από 0 μέχρι $N-1$. Για την συνάρτηση αυτοσυσχετίσεως υπάρχει το πρόβλημα ότι για μεγάλα s έχουμε μόνο λίγους όρους και

έτσι η αυτοσυσχέτιση έχει μεγάλο στατιστικό σφάλμα. Στην πράξη παίρνουμε υπ' όψη τα ρ_s μόνο μέχρι περίπου $N/4$ ή το πολύ $N/2$.

Η συνάρτηση $\gamma_s(X) = Cov(Y_t, Y_{t+s}) = Cov(Y_{t-s}, Y_t) = \gamma_s$ (2.11) ονομάζεται συνάρτηση αυτοδιασποράς (ACVF) μιας στάσιμης χρονολογικής σειράς.

Για τις παραπάνω συναρτήσεις ισχύουν τα εξής $\gamma_0 = \sigma_z^2, |\gamma_s| \leq \gamma_0, \gamma_s = \gamma_{-s}$ (2.12) και $\rho_0 = 1, \text{ και } |\rho_s| \leq 1, \rho_s = \rho_{-s}$ (2.13).

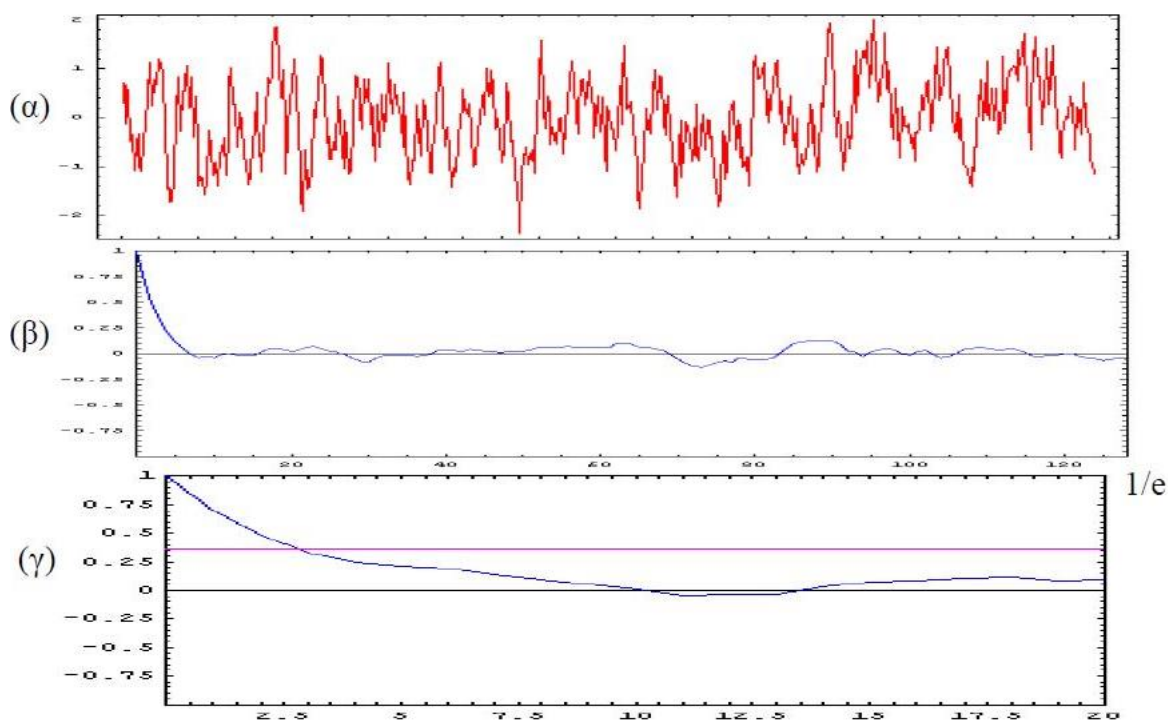
Με τον τρόπο αυτό ορίζονται, ο πίνακας αυτοδιασπορών $\Gamma_n = \begin{pmatrix} \gamma_0 & \dots & \gamma_{n-1} \\ \vdots & \ddots & \vdots \\ \gamma_{n-1} & \dots & \gamma_0 \end{pmatrix}$ και ο

πίνακας αυτοσυσχετίσεων $P_n = \begin{pmatrix} 1 & \dots & \rho_{n-1} \\ \vdots & \ddots & \vdots \\ \rho_{n-1} & \dots & 1 \end{pmatrix}$. Ισχύει ότι οι πίνακες είναι θετικά

ορισμένοι για στάσιμες χρονολογικές σειρές.

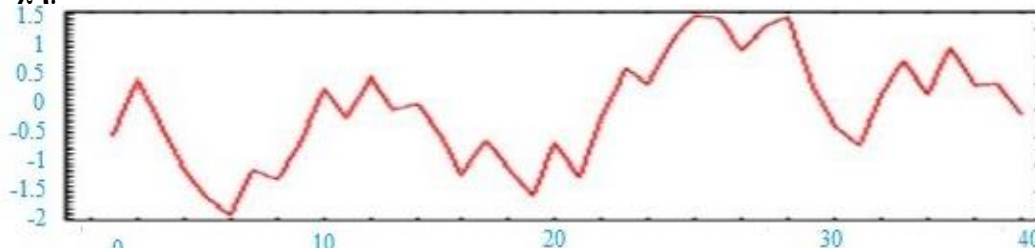
Η σημασία της συνάρτησης αυτοσυσχέτισης είναι πολύ μεγάλη, πρώτον γιατί δίνει το μέτρο της συσχέτισης των παρατηρήσεων – μετρήσεων οι οποίες απέχουν κατά το χρονικό διάστημα s και δεύτερον γιατί δείχνει τόσο το βαθμό (ένταση) όσο και το μήκος ή τη χρονική διάρκεια της μνήμης της στοχαστικής διαδικασίας.

Δηλαδή, εκφράζει κατά πόσο οι μετρήσεις με χρονική απόσταση s έχουν σχέση μεταξύ τους. Με άλλα λόγια, η συνάρτηση αυτοσυσχετίσεως μας δείχνει κατά πόσο το παρόν θυμάται το παρελθόν, και κατά πόσο το μέλλον θα επηρεαστεί από το παρόν.



Σχήμα 12. (α) Αρχική χρονολογική σειρά. Σαν θόρυβος με δομές. (β) Συνάρτηση αυτοσυσχετίσεως μέχρι $N/4$. (γ) Συνάρτηση αυτοσυσχετίσεως μέχρι $S=20$. Η συνάρτηση αυτοσυσχετίσεως τείνει στο μηδέν, παρόλα αυτά για τις πρώτες χρονικές στιγμές η χρονολογική βσειρά έχει μνήμη. Με αλλά λόγια υπάρχει χαρακτηριστικός χρόνος, δηλαδή, ένα χρονικό διάστημα για το οποίο η χρονολογική σειρά θυμάται το παρελθόν της. Χαρακτηριστικός χρόνος: Ορισμός 1: $t_c = 10.5$ Ορισμός 2: $t_c = 11$ Ορισμός 3: $t_c = 2.5$.[1]

Για το χαρακτηριστικό χρόνο υπάρχουν τρεις βασικοί τρόποι για τον ορισμό του. Ένας πρώτος ορισμός είναι η τιμή όπου, η συνάρτηση αυτοσυσχέτισης μηδενίζεται για πρώτη φορά. Δηλαδή υπολογίζεται το μικρότερο t_c για το οποίο ισχύει $\rho_{t_c} = 0$. Ένας δεύτερος ορισμός είναι ο χρόνος όπου η συνάρτηση αυτοσυσχετίσεως έχει το πρώτο ελάχιστο. Τέλος, τρίτος τρόπος για να οριστεί ο χαρακτηριστικός χρόνος, είναι η τιμή όπου η συνάρτηση αυτοσυσχέτισης πέφτει κάτω από $\frac{1}{e}$. Μια καλή επιλογή από τους παραπάνω ορισμούς είναι ο τρίτος, καθώς η συνάρτηση αυτοσυσχετίσεως πέφτει εκθετικά. Βέβαια, η επιλογή κάποιου από τους παραπάνω ορισμούς εξαρτάται από την εφαρμογή. Συχνά μπορούμε να αναγνωρίσουμε τον χαρακτηριστικό χρόνο και στην αρχική χρονολογική σειρά, όπως φαίνεται στο **σχήμα 13**.



Σχήμα 13. Η αρχική χρονολογική σειρά μέχρι την χρονική στιγμή **40**. Φαίνονται οι μικρές δομές $t_c = 11$ και στην συνεχεία οι ‘ταλαντώσεις’ $t_c = 2.5$.[1]

Καθώς οι τιμές του μέσου μ , της διακύμανσης σ^2 , της αυτοδιακύμανσης γ_s , και του συντελεστή αυτοσυσχέτισης ρ_s είναι άγνωστοι, είναι απαραίτητο στην πράξη η χρήση εκτιμητών, χρησιμοποιώντας τις αντίστοιχες ροπές του δείγματος.

Έτσι, για την μέση τιμή της χρονολογικής σειράς, ένας αμερόληπτος εκτιμητής της μέσης τιμής μ είναι ο δειγματικός μέσος $\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t$ (2.14)

Για την εκτίμηση της διακύμανσης σ^2 χρησιμοποιείται η δειγματική συνάρτηση $s^2 = \frac{1}{n} \sum_{t=1}^n (x_t - \bar{x})^2$ (2.15)

Η εκτίμηση της σχέσης (2.15) δεν είναι αμερόληπτη. Έτσι, πολλές φορές σαν εκτίμηση της διακύμανσης χρησιμοποιείται η δειγματική συνάρτηση $s^2 = \frac{1}{n-1} \sum_{t=1}^n (x_t - \bar{x})^2$ (2.16)

Η δειγματική συνάρτηση αυτοδιασποράς $\hat{\gamma} = c_s = \frac{1}{n} \sum_{t=1}^{n-s} (x_{t+s} - \bar{x})(x_t - \bar{x}), -n < s < n$ (2.17).

Αποδεικνύεται ότι $E\hat{\gamma}_s \cong \gamma_s - \frac{s}{n}\gamma_s - \left(\frac{n-s}{n}\right)Var(\bar{x})$ (2.18). Με βάση τα παραπάνω ότι $\hat{\gamma}_s$ δεν είναι αμερόληπτος εκτιμητής της συνάρτησης αυτοδιασποράς. Η δειγματική συνάρτηση αυτοδιασποράς μπορεί επίσης να οριστεί και από τη σχέση $\hat{\gamma} = c_s = \frac{1}{n-s} \sum_{t=1}^{n-s} (x_{t+s} - \bar{x})(x_t - \bar{x}), -n \leq s \leq n$ (2.19)

Στην περίπτωση αυτή $E\hat{\gamma}_s \cong \gamma_s - Var(\bar{x})$ (2.20). Ο δεύτερος εκτιμητής $\hat{\gamma}_s$ επίσης δεν είναι αμερόληπτος, αλλά ο $\hat{\gamma}_s$ έχει μεγαλύτερη αμεροληψία από τον $\hat{\gamma}_s$ ιδίως για s μεγάλο σε σχέση με το n . Τέλος, για την εκτίμηση της συνάρτησης αυτοσυσχετίσεως χρησιμοποιείται η δειγματική συνάρτηση αυτοσυσχετίσεως $\hat{\rho}_s = r_s = \frac{c_s}{c_0}$

(2.21). Το γράφημα (s, r_s) ονομάζεται διάγραμμα αυτοσυσχετίσεως (correlogram).

Έχει αποδειχτεί ότι για μεγάλο πλήθος παρατηρήσεων η κατανομή των δειγματικών αυτοσυσχετίσεων είναι κανονική και μάλιστα $r_s \sim N(\rho_s, Var(r_s))$ (2.22)

όπου $Var(r_s) \cong \frac{1}{n} \sum_{m=-\infty}^{\infty} (\rho_m^2 + \rho_{m+s}\rho_{m-s} + 2\rho_s^2\rho_m^2 - 4\rho_s\rho_m\rho_{m-s})$ (2.23) ή όταν το n είναι

πολύ μεγάλο $Var(r_s) \cong \frac{1}{n} \sum_{m=-\infty}^{\infty} \rho_m^2$ (2.24).

Κεφάλαιο 3^ο

Στάσιμες
Στοχαστικές
Διαδικασίες

Κεφάλαιο 3ο : Στάσιμες Στοχαστικές Διαδικασίες (Αυτοπαλίνδρομη Διαδικασία)

Στις παρακάτω παραγράφους θα αναφερθούν οι συνήθεις στάσιμες στοχαστικές διαδικασίες που χρησιμοποιούνται στην ανάλυση των χρονολογικών σειρών. Αυτές είναι η αυτοπαλίνδρομη διαδικασία (AR) , η διαδικασία κινητού μέσου (MA) και η μεικτή διαδικασία (ARMA).

Στα παρακάτω γίνεται η παραδοχή ότι η μέση τιμή της διαδικασίας είναι μηδέν, επειδή η συνάρτηση αυτοσυσχέτισης δεν επηρεάζεται εάν η σειρά έχει μέση τιμή διάφορη του μηδενός.

Επίσης, χρησιμοποιείται ο συμβολισμός $\{X_t\}$ για την διαδικασία και $\{e_t\}$ για την ακολουθία των τυχαίων σφαλμάτων, η οποία πάντα θα θεωρείται λευκός θόρυβος και είναι ασυσχέτιστη με την $\{X_t\}$.

Όπως έχει αναφερθεί, στο παρόν κεφάλαιο θα γίνει μελέτη των στάσιμων διαδικασιών. Σύμφωνα λοιπόν με το θεώρημα διαχωρισμού του Wold (*Wold's Decomposition Theorem*) κάθε στάσιμη στοχαστική διαδικασία μπορεί να εκφραστεί ως γραμμικός συνδυασμός μιας ακολουθίας ασυσχέτιστων τυχαίων μεταβλητών. Ένας τέτοιος γραμμικός συνδυασμός είναι επίσης γνωστός ως γραμμικό φίλτρο.

Έστω $\{X_t\}$ μια όχι κατά ανάγκη αυστηρώς στάσιμη στοχαστική διαδικασία, με μέσο μ . Το γραμμικό φίλτρο μπορεί να διατυπωθεί ως εξής:

$$X_t - \mu = e_t + \Psi_1 e_{t-1} + \Psi_2 e_{t-2} + \dots$$

Θέτοντας $\Psi_0 = 1$ η παραπάνω σχέση γράφεται:

$$X_t - \mu = \sum_{i=0}^{\infty} \Psi_i e_{t-i}$$

Όπως έχει ήδη αναφερθεί η ακολουθία $\{e_t\}$ για $t = 0, \pm 1, \pm 2, \dots$ είναι μια ακολουθία τυχαίων μεταβλητών για την οποία για κάθε t έχει γίνει η υπόθεση ότι:

$$E(e_t) = 0, \text{Var}(e_t) = \sigma^2, \text{Cov}(e_t, e_{t-k}) = 0 \text{ για κάθε } k \neq 0$$

Δηλαδή $e_t \sim N(0, \sigma_e^2)$ και πρόκειται για μια διαδικασία λευκού θορύβου.

Ο αριθμός των συντελεστών Ψ_t που είναι γνωστοί ως συντελεστές σταθμίσεως, μπορεί να είναι άπειρος ή πεπερασμένος. Εάν είναι άπειρος, γίνεται η υπόθεση ότι το άθροισμα τους συγκλίνει απολύτως, ότι δηλαδή $\sum_{t=0}^{\infty} |\Psi_t| < \infty$

Από την $X_t - \mu = e_t + \Psi_1 e_{t-1} + \Psi_2 e_{t-2} + \dots$ προκύπτει ότι για την συνάρτηση αυτόδιασποράς ισχύουν τα παρακάτω:

$$\gamma_0 = \text{Var}(X_t) = E(X_t - \mu)^2 = E(e_t + \Psi_1 e_{t-1} + \Psi_2 e_{t-2} + \dots)^2 = E(e_t^2) + \Psi_1^2 E(e_{t-1})^2 + \Psi_2^2 E(e_{t-2})^2 + \dots$$

καθώς ισχύει ότι: $E(e_t e_{t-k}) = 0, k \neq 0$.

$$\text{Επομένως } \gamma_0 = \sigma^2 + \Psi_1^2 \sigma^2 + \Psi_2^2 \sigma^2 + \dots = \sigma^2 \sum_{t=0}^{\infty} \Psi_t^2$$

$$\text{Και } \gamma_s = E(X_t - \mu)(X_{t-k} - \mu) = E(e_t + \Psi_1 e_{t-1} + \Psi_2 e_{t-2} + \dots + \Psi_k e_{t-k} + \Psi_{k+1} e_{t-k-1} + \dots) * \\ (e_{t-k} + \Psi_1 e_{t-k-1} + \Psi_2 e_{t-k-2} + \dots + \Psi_k e_{t-k-k} + \Psi_{k+1} e_{t-k-k-1} + \dots) = E(\Psi_k e_{t-k}^2) + E(\Psi_1 \Psi_k e_{t-k-1}^2) + \dots$$

Γιατί όλοι οι άλλοι όροι είναι μηδέν διότι $\text{Cov}(e_t, e_{t-k}) = 0, k \neq 0$

$$\text{Επομένως } \gamma_k = \sigma^2 \sum_{t=0}^{\infty} \Psi_t \Psi_{t+k}$$

$$\text{Τέλος από τις προηγούμενες σχέσεις προκύπτει ότι: } \rho_k = \frac{\sum_{t=0}^{\infty} \Psi_t \Psi_{t+k}}{\sum_{t=0}^{\infty} \Psi_t^2}$$

Γίνεται φανερό ότι εάν το άθροισμα των συντελεστών σταθμίσεως δεν συγκλίνει, η διακύμανση τείνει στο άπειρο και ο συντελεστής αυτοσυσχετίσεως στο μηδέν. Εάν η σειρά είναι στάσιμη, τότε η διακύμανση είναι πεπερασμένη.

Η εξίσωση $X_t - \mu = e_t + \Psi_1 e_{t-1} + \Psi_2 e_{t-2} + \dots$ αποτελεί μια γενική μορφή από την οποία, με διάφορες υποθέσεις σχετικά με τους συντελεστές σταθμίσεως, προκύπτουν διάφορες στοχαστικές διαδικασίες χρονολογικών σειρών.

Για παράδειγμα, έστω $\Psi_t = \varphi^t$ τότε από την $X_t - \mu = e_t + \Psi_1 e_{t-1} + \Psi_2 e_{t-2} + \dots$ προκύπτει: $X_t - \mu = e_t + \varphi e_{t-1} + \varphi^2 e_{t-2} + \dots = e_t + \varphi(e_{t-1} + \varphi e_{t-2} + \varphi^2 e_{t-3}) = e_t + \varphi(X_t - \mu)$ και $X_t = \mu(1 - \varphi) + \varphi X_{t-1} + e_t$

Η διαδικασία αυτή είναι γνωστή ως αυτοπαλίνδρομη διαδικασία πρώτης τάξης $AR(1)$.

Εάν υποθέσουμε ότι $\Psi_1 = \theta$ και $\Psi_t = 0$ για $t \geq 2$ τότε από την $X_t - \mu = e_t + \Psi_1 e_{t-1} + \Psi_2 e_{t-2} + \dots$ προκύπτει η διαδικασία: $X_t = \mu + e_t + \theta e_{t-1}$. Η διαδικασία αυτή είναι γνωστή ως διαδικασία κινητού μέσου πρώτης τάξης $MA(1)$. Στις επόμενες παραγράφους θα εξετάσουμε ξεχωριστά κάθε κατηγορία.

3.1 Αυτοπαλίνδρομη διαδικασία τάξης p , $AR(p)$

Η έκφραση μιας αυτοπαλίνδρομης διαδικασίας τάξης p είναι:

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + e_t \rightarrow (1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p) X_t = e_t \rightarrow \varphi(B) X_t = e_t$$

Όπου το $\varphi(B)$ πολυώνυμο $\varphi(B) = 1 - \sum_{j=1}^p \varphi_j B^j$

και B ο τελεστής της προς τα πίσω μετατόπισης: $B^j X_t = X_{t-j}$

Η ικανή και αναγκαία συνθήκη για να είναι στάσιμη η αυτοπαλίνδρομη διαδικασία, είναι η ρίζες της εξίσωσης $\varphi(\omega) = 0$ να βρίσκονται έξω από τον μοναδιαίο κύκλο ή ισοδύναμα εάν $\lambda_1, \lambda_2, \dots, \lambda_p$ οι ρίζες της εξίσωσης $\omega^p - \varphi_1 \omega^{p-1} - \dots - \varphi_p = 0$ να βρίσκονται μέσα στο μοναδιαίο κύκλο $|\lambda_j| < 1, j = 1, 2, \dots, p$

Για να είναι η σειρά στάσιμη, θα πρέπει η αυτοδιακύμανση γ_0 να είναι σταθερός θετικός αριθμός. Για να συμβεί αυτό, θα πρέπει κάθε όρος του αθροίσματος $\sigma^2 \sum_{t=0}^{\infty} \Psi_t^2$ να είναι θετικός αριθμός. Αυτό όμως μπορεί να συμβεί μόνο αν οι ρίζες της εξίσωσης $\omega^p - \varphi_1 \omega^{p-1} - \dots - \varphi_p = 0$ βρίσκονται μέσα στον μοναδιαίο κύκλο.

Ο μέσος όρος των X_t δίνεται από την σχέση: $\mu = \frac{\varphi}{1 - \varphi_1 - \varphi_2 - \dots - \varphi_p}$

Στην περίπτωση μας υποθέτουμε για χάρη απλότητας ότι $\varphi = 0$ οπότε προκύπτει $\mu = 0$.

Για την $AR(p)$ ισχύουν τα εξής:

Η συνάρτηση αυτοδιασποράς $ACVF$ της στάσιμης διαδικασίας είναι:

$$\gamma_k = \varphi_1 \gamma_{k-1} + \dots + \varphi_p \gamma_{k-p}, k > 0$$

ενώ η συνάρτηση αυτοσυσχετίσεως ACF δίνεται από την σχέση:

$$\rho_k = \varphi_1 \rho_{k-1} + \dots + \varphi_p \rho_{k-p}, k > 0$$

Επίσης διασπορά της αυτοπαλίνδρομης διαδικασίας είναι:

$$\gamma_0 = \sigma_x^2 = \sigma_e^2 + \varphi_1 \gamma_1 + \varphi_2 \gamma_2 + \dots + \varphi_p \gamma_p = \sigma_e^2 + \sum_{j=1}^p \varphi_j \gamma_j \rightarrow \gamma_0 = \frac{\sigma_e^2}{1 - \sum_{j=1}^p \varphi_j \rho_j}$$

Από την σχέση $\gamma_k = \varphi_1 \gamma_{k-1} + \dots + \varphi_p \gamma_{k-p}$ για $k=1, 2, \dots, p$ προκύπτουν οι παρακάτω p , Yule-Walker εξισώσεις:

$$\begin{aligned} \rho_1 &= \varphi_1 + \varphi_2 \rho_1 + \varphi_3 \rho_2 + \dots + \varphi_p \rho_{p-1} \\ \rho_2 &= \rho_1 \varphi_1 + \varphi_2 + \varphi_3 \rho_1 + \dots + \varphi_p \rho_{p-2} \\ \rho_3 &= \rho_2 \varphi_1 + \varphi_2 \rho_1 + \varphi_3 + \dots + \varphi_p \rho_{p-3} \\ &\vdots \\ \rho_p &= \rho_{p-1} \varphi_1 + \varphi_2 \rho_{p-2} + \varphi_3 \rho_{p-3} + \dots + \varphi_p \end{aligned}$$

Οι παραπάνω εξισώσεις συνιστούν ένα σύστημα p εξισώσεων, από την λύση του οποίου προκύπτουν οι τιμές για τις αυτοσυσχετίσεις, εάν είναι γνωστές οι τιμές των συντελεστών $\varphi_1, \varphi_2, \dots, \varphi_p$ οι οποίοι ονομάζονται συντελεστές αυτοπαλινδρόμησης.

Με την μορφή πινάκων το παραπάνω σύστημα γράφεται ως εξής: $R = \Pi \Phi$

$$\text{Όπου } R = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix}, \quad \Phi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix} \quad \text{και} \quad \Pi = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \dots & 1 \end{bmatrix}$$

Εάν οι αυτοσυσχετίσεις είναι γνωστές τότε οι συντελεστές αυτοπαλινδρόμησης δίνονται από την σχέση: $\Phi = \Pi^{-1}R$

3.1.1 Αυτοπαλίνδρομη διαδικασία τάξης 1, AR(1)

Η πιο απλή μορφή των αυτοπαλίνδρομων διαδικασιών είναι οι αυτοπαλίνδρομες διαδικασίες πρώτης τάξης.

Από την σχέση $X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + e_t$ για $p=1$ προκύπτει :

$$X_t = \varphi X_{t-1} + e_t$$

Με διαδοχικές αντικαταστάσεις η παραπάνω σχέση γίνεται:

$$X_t = e_t + \varphi e_{t-1} + \varphi^2 e_{t-2} + \dots = \sum_{j=0}^{\infty} \varphi^j e_{t-j}$$

Έτσι για την $AR(1)$ προκύπτει ότι:

$$E(X_t) = 0, \text{Var}(X_t) = \sigma_e^2(1 + \varphi^2 + \varphi^4 + \dots) = \sigma_e^2 \sum_{j=0}^{\infty} \varphi^{2j}$$

Για να είναι στάσιμη η σειρά πρέπει $|\varphi| < 1$. Έτσι για $|\varphi| < 1$ η διασπορά της διαδικασίας $\{X_t\}$ δίνεται από την σχέση: $\sigma_x^2 = \gamma_0 = \text{Var}(X_t) = \frac{\sigma_e^2}{1 - \varphi^2}$

Απόδειξη της σχέσης: $\sigma_x^2 = \gamma_0 = \text{Var}(X_t) = \frac{\sigma_e^2}{1 - \varphi^2}$

Υψώνουμε στο τετράγωνο και τις δυο πλευρές της σχέσης $X_t = \varphi X_{t-1} + e_t$ και παίρνουμε τις μαθηματικές ελπίδες οπότε έχουμε: $E(x_t^2) = \varphi^2 E(x_{t-1}^2) + 2\varphi E(x_{t-1}e_t) + E(e_t^2)$
Αλλά $E(x_{t-1}e_t) = 0$ αφού το x_{t-1} εξαρτάται μόνο από το e_{t-1} το οποίο είναι λευκός θόρυβος. Επίσης $E(x_t^2) = E(x_{t-1}^2) = \text{Var}(x_t)$ αφού η σειρά είναι στάσιμη.

Επομένως $\text{Var}(x_t) = \varphi^2 \text{Var}(x_t) + \sigma_e^2$ ή $\gamma_0 = \text{Var}(X_t) = \frac{\sigma_e^2}{1 - \varphi^2}$ που είναι η σχέση που είδαμε παραπάνω.

Επίσης, η συνάρτηση αυτοδιασποράς δίνεται από τη σχέση:

$$\gamma_k = \text{Cov}(X_{t-k}, X_t) = \dots = \varphi^k \sigma_e^2 \sum_{j=0}^{\infty} \varphi^{2j} = \frac{\varphi^k \sigma_e^2}{1 - \varphi^2} = \varphi^k \sigma_x^2 = \varphi^k \gamma_0$$

Απόδειξη της παραπάνω σχέσης:

Πολλαπλασιάζουμε τις δυο πλευρές της $X_t = e_t + \varphi e_{t-1} + \varphi^2 e_{t-2} + \dots = \sum_{j=0}^{\infty} \varphi^j e_{t-j}$

επί x_{t-k} και παίρνοντας τις μαθηματικές ελπίδες έχουμε:

$$E(x_t x_{t-k}) = \varphi E(x_{t-1} x_{t-k}) + E(x_{t-k} e_t) \text{ ή } \gamma_k = \varphi \gamma_{k-1}, k > 0$$

γιατί $E(x_{t-k} e_t) = 0$ και $E(x_{t-1} x_{t-k}) = \gamma_{k-1}$

Από την παραπάνω σχέση προκύπτει ότι:

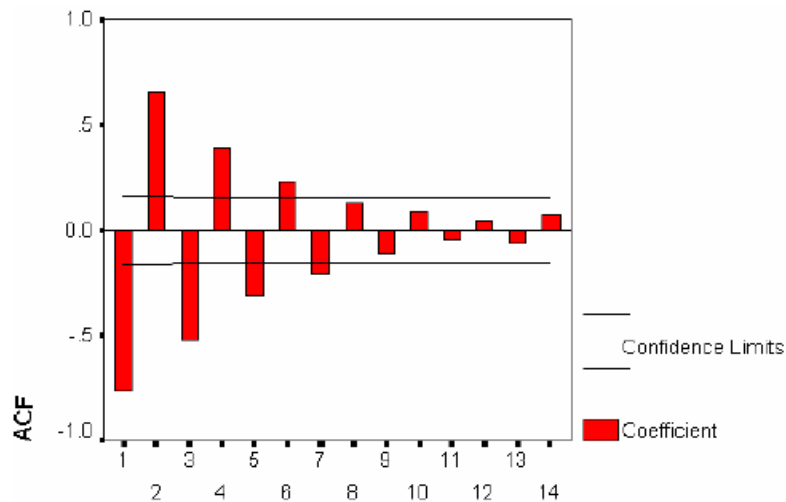
$$\begin{aligned} \gamma_1 &= \varphi \gamma_0 \\ \gamma_2 &= \varphi \gamma_1 = \varphi^2 \gamma_0 \\ \gamma_3 &= \varphi \gamma_2 = \varphi^3 \gamma_0 \\ &\vdots \\ \gamma_k &= \varphi^k \gamma_0 \end{aligned}$$

που είναι η σχέση που είδαμε παραπάνω.

Το παραπάνω αποτέλεσμα ισχύει εάν $|\varphi| < 1$ συνθήκη η οποία ισχύει όταν η $AR(1)$ διαδικασία είναι στάσιμη. Όταν όμως $0 < \varphi < 1$ τότε η συνάρτηση αυτοσυσχέτισης,

αρχίζοντας από την μονάδα αφού $\varphi_0 = 1$, ελαττώνεται εκθετικά και τείνει προς το μηδέν καθώς το k αυξάνει. Για $-1 < \varphi < 0$ πάλι η συνάρτηση αυτοσυσχετίσεως ACF ελαττώνεται και τείνει προς το μηδέν αλλά με εναλλασσόμενο πρόσημο.

Τα παραπάνω είναι απόρροια του ότι η συνάρτηση αυτοσυσχέτισης από τις σχέσεις που είδαμε παραπάνω γίνεται $\rho_k = \varphi^k, k = 0, 1, \dots$ και $\rho_k = \varphi^{|k|}, k = 0, \pm 1, \dots$



Σχήμα 1. Το γράφημα των ACF από μια $AR(1)$ διαδικασία [3]

3.1.2 Αυτοπαλίνδρομη διαδικασία τάξης 2, $AR(2)$

Οι αυτοπαλίνδρομες διαδικασίες δευτέρας τάξεως στην γενική τους μορφή μπορούν να γράφουν από την σχέση $X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + e_t$ για $p = 2$ ως εξής:

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + e_t$$

Όπως είναι ήδη γνωστό για να είναι στάσιμη η αυτοπαλίνδρομη διαδικασία δευτέρας τάξεως $AR(2)$ θα πρέπει οι ρίζες της εξίσωσης $\omega^2 - \varphi_1 \omega - \varphi_2 = 0$ να βρίσκονται μέσα

στον μοναδιαίο κύκλο. Οι ρίζες είναι : $\lambda_{1,2} = \frac{-\varphi_1 \pm \sqrt{\varphi_1^2 - 4\varphi_2}}{2}$.

Έτσι, η συνθήκη στασιμότητας απαιτεί $\left| \frac{-\varphi_1 \pm \sqrt{\varphi_1^2 - 4\varphi_2}}{2} \right| < 1$

συνθήκη που οδηγεί στις εξής σχέσεις μεταξύ των παραμέτρων της διαδικασίας :

$$\varphi_1 + \varphi_2 < 1 \text{ και } \varphi_2 - \varphi_1 < 1 \text{ και } -1 < \varphi_2 < 1$$

Ακόμα για την $AR(2)$ ισχύουν τα εξής:

$$\mu = \frac{\varphi_0}{1 - \varphi_1 - \varphi_2}$$

Για την συνάρτηση αυτοδιασποράς ισχύουν:

$$\gamma_0 = \text{Var}(X_t) = \varphi_1\gamma_1 + \varphi_2\gamma_2 + \sigma^2 \quad \text{ή} \quad \gamma_0 = \frac{(1-\varphi_2)}{(1-\varphi_2)(1-\varphi_1-\varphi_2)(1+\varphi_1-\varphi_2)} \sigma^2$$

Για να είναι η σειρά στάσιμη, θα πρέπει η αυτοδιακύμανση γ_0 να είναι σταθερός θετικός αριθμός. Για να συμβεί αυτό, θα πρέπει κάθε όρος στις παρενθέσεις στη παραπάνω σχέση να είναι θετικός αριθμός. Αυτό όμως μπορεί να συμβεί μόνο αν οι λύσεις της εξίσωσης να βρίσκονται μέσα στον μοναδιαίο κύκλο.

Επίσης $\gamma_k = \text{Cov}(X_t, X_{t-k}) = \varphi_1\gamma_{k-1} + \varphi_2\gamma_{k-2}, k > 0$.

Για την συνάρτηση αυτοσυσχέτισης προκύπτει ότι:

$$\rho_k = \varphi_1\rho_{k-1} + \varphi_2\rho_{k-2}, k > 0$$

Από την παραπάνω σχέση για $k=0,1,2,\dots$ βλέπουμε ότι,

$$\rho_1 = \varphi_1 + \varphi_2\rho_1$$

$$\rho_2 = \varphi_1\rho_1 + \varphi_2$$

Αφού $\rho_0 = 1$ και $\rho_{-1} = \rho_1$

Απόδειξη της σχέσης $\mu = \frac{\varphi_0}{1-\varphi_1-\varphi_2}$

Παίρνουμε τις μαθηματικές ελπίδες στη σχέση $X_t = \varphi_1X_{t-1} + \varphi_2X_{t-2} + e_t$ έχοντας υποθέσει ότι $\varphi_0 \neq 0$ οπότε: $E(X_t) = \varphi_0 + \varphi_1E(X_{t-1}) + \varphi_2E(X_{t-2}) + E(e_t)$ ή

$$\mu = \varphi_0 + \varphi_1\mu + \varphi_2\mu \quad \text{και} \quad \mu = \frac{\varphi_0}{1+\varphi_1+\varphi_2}$$

που είναι η σχέση που θέλουμε να αποδείξουμε

Απόδειξη της σχέσης $\gamma_0 = \text{Var}(X_t) = \varphi_1\gamma_1 + \varphi_2\gamma_2 + \sigma^2$

Πολλαπλασιάζουμε και τις δυο πλευρές της σχέσης $X_t = \varphi_1X_{t-1} + \varphi_2X_{t-2} + e_t$ επί x_t οπότε έχουμε: $X_t^2 = \varphi_1x_t x_{t-1} + \varphi_2x_t x_{t-2} + x_t e_t$

Παίρνουμε τις μαθηματικές ελπίδες:

$$E(x^2) = \text{Var}(X_t) = \varphi_1E(x_t x_{t-1}) + \varphi_2E(x_t x_{t-2}) + E(x_t e_t) \quad \text{ή} \quad \gamma_0 = \varphi_1\gamma_1 + \varphi_2\gamma_2 + E(x_t e_t)$$

Αλλά $E(x_t e_t) = E(\varphi_1x_{t-1} + \varphi_2x_{t-2} + e_t)e_t = \sigma^2$ και επομένως $\gamma_0 = \varphi_1\gamma_1 + \varphi_2\gamma_2 + \sigma^2$ που είναι η σχέση που θέλαμε να αποδείξουμε.

Απόδειξη της σχέσης $\gamma_k = \text{Cov}(X_t, X_{t-k}) = \varphi_1\gamma_{k-1} + \varphi_2\gamma_{k-2}, k > 0$

Πολλαπλασιάζουμε και τις δυο πλευρές της σχέσης $X_t = \varphi_1X_{t-1} + \varphi_2X_{t-2} + e_t$ επί x_{t-k} οπότε: $x_t x_{t-k} = \varphi_1x_{t-1}x_{t-k} + \varphi_2x_{t-2}x_{t-k} + x_{t-k}e_t$

Παίρνουμε τις μαθηματικές ελπίδες

$$E(x_t x_{t-k}) = \text{Cov}(X_t, X_{t-k}) = \varphi_1E(x_{t-1}x_{t-k}) + \varphi_2E(x_{t-2}x_{t-k}) + E(x_{t-k}e_t)$$

Αλλά $E(x_{t-k}e_t) = 0$ και εξ' ορισμού $E(x_{t-1}x_{t-k}) = \gamma_{k-1}$ και $E(x_{t-2}x_{t-k}) = \gamma_{k-2}$ οπότε

$\gamma_k = Cov(X_t, X_{t-k}) = \varphi_1 \gamma_{k-1} + \varphi_2 \gamma_{k-2}$ που είναι η σχέση που θέλουμε να αποδείξουμε.

Απόδειξη της σχέσης $\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2}, k > 0$

Διαιρούμε τη σχέση $\gamma_k = \varphi_1 \gamma_{k-1} + \varphi_2 \gamma_{k-2}$ με γ_0 και έχουμε

$$\gamma_k = \varphi_1 \gamma_{k-1} + \varphi_2 \gamma_{k-2} \Rightarrow \frac{\gamma_k}{\gamma_0} = \varphi_1 \frac{\gamma_{k-1}}{\gamma_0} + \varphi_2 \frac{\gamma_{k-2}}{\gamma_0} \Rightarrow \rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2}$$

που είναι και η σχέση που θέλουμε να αποδείξουμε.

Απόδειξη της σχέσης $\gamma_0 = \frac{(1-\varphi_2)}{(1-\varphi_2)(1-\varphi_1-\varphi_2)(1+\varphi_1-\varphi_2)} \sigma^2$

Η παραπάνω σχέση προκύπτει από την $\gamma_0 = Var(X_t) = \varphi_1 \gamma_1 + \varphi_2 \gamma_2 + \sigma^2$ εάν αντικαταστήσουμε τα γ_1, γ_2 όπως αυτά δίνονται από την σχέση

$$\gamma_k = Cov(X_t, X_{t-k}) = \varphi_1 \gamma_{k-1} + \varphi_2 \gamma_{k-2}, k > 0 \quad \begin{matrix} \rho_1 = \varphi_1 + \varphi_2 \rho_1 \\ \rho_2 = \varphi_1 \rho_1 + \varphi_2 \end{matrix}$$

είναι γνωστές ως εξισώσεις *Yule-Walker*, από τις οποίες προκύπτουν οι τιμές για τις δυο αυτοσυσχετίσεις, όταν βέβαια ξέρουμε τις τιμές των συντελεστών φ_1, φ_2 .

$$\text{Συγκεκριμένα προκύπτει } \rho_1 = \frac{\varphi_1}{1-\varphi_2} \text{ και } \rho_2 = \varphi_2 + \frac{\varphi_1^2}{1-\varphi_2}$$

Η συνάρτηση αυτοσυσχετίσεως μιας $AR(2)$ διαδικασίας θα τείνει προς το μηδέν καθώς αυξάνεται η υστέρηση k .

3.2 Η Συνάρτηση Μερικής Αυτοσυσχετίσεως

Για μια δοθείσα χρονολογική σειρά τα κύρια ερωτήματα που πρέπει να απαντηθούν είναι ποία στοχαστική διαδικασία θα δώσει ένα ικανοποιητικό μοντέλο, ποία είναι η τάξη του και ποιές είναι οι παράμετροί του. Όλες οι αυτοπαλίνδρομες διαδικασίες έχουν συναρτήσεις αυτοσυσχετίσεως, οι όποιες βαίνουν φθίνουσες καθώς αυξάνει το μήκος της υστέρησης k , με συνέπεια να είναι πολλές φορές δύσκολο να καθοριστεί η τάξη της διαδικασίας. Ως ένα πρόσθετο κριτήριο για τον σκοπό αυτό εισάγεται μια καινούργια συνάρτηση απαραίτητη στην ανάλυση των χρονολογικών σειρών, η συνάρτηση μερικής αυτοσυσχετίσεως.

Έστω, ότι η διαδικασία είναι $AR(p)$ και στην γενική περίπτωση έστω η μέση τιμή της, όποτε η έκφρασή της είναι: $X_t - \mu = \varphi_1 (X_{t-1} - \mu) + \dots + \varphi_p (X_{t-p} - \mu) + e_t$

$$\text{όπου } e_t \sim N(0, \sigma_e^2)$$

Η χρονολογική σειρά είναι μια πραγματοποίηση αυτής της διαδικασίας, άρα:

$$x_t - \mu = \varphi_1 (x_{t-1} - \mu) + \dots + \varphi_p (x_{t-p} - \mu) + e_t \text{ με } t = p+1, \dots, n$$

Οπότε γίνεται γνωστό το μοντέλο που θα προσαρμοστεί, όταν θα έχουν εκτιμηθεί οι παράμετροι $\varphi_1, \varphi_2, \dots, \varphi_p$, η μέση τιμή μ , η διασπορά του λευκού θορύβου σ_e^2 και η τάξη p του μοντέλου.

Ξεκινώντας την ανάλυση των δεδομένων, η τάξη του καλύτερα προσαρμοζόμενου μοντέλου είναι άγνωστη. Το πρόβλημα της εκτίμησης της τάξης του μοντέλου, είναι ανάλογο με το πρόβλημα στην παλινδρόμηση εάν μια ανεξάρτητη μεταβλητή πρέπει να συμπεριληφθεί ή όχι στο μοντέλο. Όπλο προς αυτή την κατεύθυνση είναι η συνάρτηση μερικής αυτοσυσχετίσεως (**Partial Autocorrelation Function PACF**).

Η μερική αυτοσυσχέτιση ανάμεσα στην X_t και στην X_{t-s} αναφέρεται στην συσχέτιση ανάμεσα στην X_t και στην X_{t-s} όταν έχουν αφαιρεθεί οι γραμμικές επιδράσεις των ενδιάμεσων μεταβλητών $X_{t-1}, X_{t-2}, \dots, X_{t-(k-1)}$.

Έστω X_t στάσιμη διαδικασία, χωρίς περιορισμό της γενικότητας και $E(X_t) = 0$. Έστω ένα μοντέλο παλινδρόμησης με εξαρτημένη μεταβλητή την X_{t+k} και ανεξάρτητες τις $X_{t+k-1}, X_{t+k-2}, \dots, X_t$ δηλαδή, $X_{t+k} = \varphi_{k1} X_{t+k-1} + \varphi_{k2} X_{t+k-2} + \dots + \varphi_{kk} X_t + e_{t+k}$ αφού φ_{kj} η j -παράμετρος της παλινδρόμησης και e_{t+k} το σφάλμα για το οποίο ισχύουν όλες οι γνωστές προϋποθέσεις. Εύκολα γίνεται αντιληπτό ότι ο συντελεστής φ_{kk} είναι ο συντελεστής αυτοσυσχετίσεως ανάμεσα στην X_t και στην X_{t-k} .

Πολλαπλασιάζοντας με X_{t+k-j} και τις δυο πλευρές τις $X_{t+k} = \varphi_{k1} X_{t+k-1} + \varphi_{k2} X_{t+k-2} + \dots + \varphi_{kk} X_t + e_{t+k}$ και παίρνοντας τις μέσες τιμές έχουμε $\gamma_j = \varphi_{k1} \gamma_{j-1} + \varphi_{k2} \gamma_{j-2} + \dots + \varphi_{kk} \gamma_{j-k}$ και $\rho_j = \varphi_{k1} \rho_{j-1} + \varphi_{k2} \rho_{j-2} + \dots + \varphi_{kk} \rho_{j-k}$.

Για $j = 1, 2, \dots, k$ ορίζεται το παρακάτω σύστημα εξισώσεων:

$$\begin{aligned} \rho_1 &= \varphi_{k1} \rho_0 + \varphi_{k2} \rho_{-1} + \dots + \varphi_{kk} \rho_{1-k} \\ \rho_2 &= \varphi_{k1} \rho_1 + \varphi_{k2} \rho_0 + \dots + \varphi_{kk} \rho_{2-k} \\ &\vdots \\ \rho_k &= \varphi_{k1} \rho_{k-1} + \varphi_{k2} \rho_{k-2} + \dots + \varphi_{kk} \rho_0 \end{aligned} \quad \text{ή } \Pi \Phi_k = R \quad \text{ή } \Phi_k = \Pi^{-1} R$$

Η παράμετρος φ_{kk} , που υπολογίζεται από το παραπάνω σύστημα με τον κανόνα Cramer, θεωρούμενη σαν συνάρτηση της υστέρησης k , λέγεται **Συνάρτηση Μερικής Αυτοσυσχετίσεως (PACF)**. Έτσι, για

$$k = 1 \text{ προκύπτει } \varphi_{11} = \rho_1,$$

$$k = 2 \text{ προκύπτει } \varphi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

$$\begin{array}{l} \varphi_{11} = \rho_1 \\ \varphi_{22} \neq 0 \\ \vdots \\ \varphi_{pp} \neq 0 \\ \varphi_{kk} = 0, k > p \end{array}$$

Τέλος για $AR(p)$ ισχύει:

3.3 Διαδικασίες Κινητού Μέσου

Οι διαδικασίες κινητού μέσου είναι χρήσιμες για περιγραφή φαινομένων στα οποία τα γεγονότα παράγουν ένα άμεσο αποτέλεσμα, η επίδραση του οποίου όμως δεν σταματά εκεί αλλά διαρκεί, αν και το ίδιο το γεγονός παύει να υφίσταται. Συνήθως επηρεάζει λιγότερο και για μικρό χρονικό διάστημα τις επόμενες χρονικές στιγμές. Για παράδειγμα μια απεργία επηρεάζει την οικονομία όχι μόνο την στιγμή που πραγματοποιείται αλλά και τους επόμενους μήνες αν και έχει λήξει.

Στην γενική του μορφή του μια διαδικασία κινητού μέσου q τάξεως ή $MA(q)$ γράφεται ως εξής:

$$X_t = e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \Rightarrow X_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) e_t \Rightarrow X_t = \theta(B) e_t$$

Όπου $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$

Η τάξη q αναφέρεται στο μήκος της υστερήσεως της μεταβλητής, για την οποία, όπως και πριν, γίνεται η υπόθεση ότι είναι λευκός θόρυβος. Ο όρος κινητός μέσος αναφέρεται στο γεγονός ότι η X_t εμφανίζεται ως ένα σταθμισμένο άθροισμα, όπως ο μέσος, των τιμών της e_t .

Προτού ξεκινήσει η μελέτη των διαδικασιών κινητών μέσων (MA) θα ορίσουμε την αντιστρέψιμη διαδικασία κινητού μέσου. Μια (MA) διαδικασία λέγεται αντιστρέψιμη όταν μπορεί να αντιστραφεί η έκφρασή της, δηλαδή να εκφραστούν τα e_t σαν συνάρτηση των X_t, X_{t-1}, \dots με μοναδικό τρόπο. Η ικανή και αναγκαία συνθήκη για να ισχύει η παραπάνω ιδιότητα είναι οι ρίζες του πολυωνύμου $\theta(B) = 0$ να βρίσκονται έξω από τον μοναδιαίο κύκλο. Με άλλα λόγια, μια διαδικασία (MA) είναι αντιστρέψιμη αν μπορεί να διατυπωθεί ως μια αυτοπαλίνδρομη διαδικασία με απείρους όρους.

3.3.1 Διαδικασία Κινητού Μέσου τάξης 1, MA(1)

Πρώτα θα εξεταστεί η περίπτωση μιας διαδικασίας κινητού μέσου τάξης 1 (MA(1) διαδικασία)

Για $q=1$ η σχέση ($X_t = e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$) γίνεται $X_t = e_t - \theta_1 e_{t-1}$

Αποδεικνύεται εύκολα ότι για την διαδικασία ισχύουν:

$$E(X_t) = 0 \quad , \quad \gamma_0 = \text{Var}(X_t) = (1 + \theta^2) \sigma_e^2 \quad , \quad \gamma_k = \text{Cov}(X_t, X_{t-k}) = \begin{cases} -\theta \sigma_e^2, & k = 1 \\ 0, & k > 1 \end{cases} \quad ,$$

$$\gamma_k = \text{Cov}(X_t, X_{t-k}) = \begin{cases} -\theta \sigma_e^2, & k = 1 \\ 0, & k > 1 \end{cases}$$

Τέλος, η συνάρτηση μερικής αυτοσυσχέτισης γίνεται $\varphi_{kk} = \frac{-\theta(1-\theta^2)}{1-\theta^{2(k+1)}}$ για $k \geq 1$

δηλαδή η PACF μιας διαδικασίας MA(1) φθίνει εκθετικά με εναλλασσόμενο πρόσημο εάν αρχίζοντας από θετικό ή με αρνητικό πρόσημο εάν $\theta > 0$ ή $\theta < 0$ αντίστοιχα ενώ μόνο η αυτοσυσχέτιση τάξης 1 είναι σημαντική.

Όπως εύκολα γίνεται φανερό από τις παραπάνω σχέσεις όλες οι αυτοσυνδιακυμάνσεις και συνεπώς οι αυτοσυσχέτισεις είναι μηδέν εκτός από την πρώτη. Αυτό σημαίνει ότι η 'μνήμη' της διαδικασίας δεν υπερβαίνει την μια περίοδο. Δηλαδή, μια οποιαδήποτε παρατήρηση της X_t , έστω X_T συσχετίζεται με την προηγούμενη X_{T-1} ή την επόμενη X_{T+1} , αλλά δεν συσχετίζεται με καμία άλλη.

Τέλος, μια MA(1) διαδικασία για να είναι αντιστρέψιμη θα πρέπει $|\theta| < 1$.

3.3.2 Διαδικασία Κινητού Μέσου Τάξης 2, MA(2)

Για $q=2$ η σχέση $X_t = e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$ γίνεται $X_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$

και αποτελεί την έκφραση για μια διαδικασία κινητού μέσου τάξης 2, MA(2).

Όπως είναι γνωστό μια κινητού μέσου διαδικασία είναι σταθερή και αντιστρέψιμη όταν οι ρίζες του πολυωνύμου $\theta(B) = 0$ βρίσκονται έξω από τον μοναδιαίο κύκλο.

Εύκολα, λοιπόν, αποδεικνύεται ότι για την διαδικασία 2, MA(2) ισχύουν οι σχέσεις:

$$\theta_1 + \theta_2 < 1 \quad , \quad \theta_2 - \theta_1 < 1 \quad , \quad -1 < \theta_2 < 1$$

Επίσης, ισχύουν οι σχέσεις $E(X_t) = 0$ και $\gamma_0 = \text{Var}(X_t) = (1 + \theta_1^2 + \theta_2^2) \sigma_e^2$

Για τις συναρτήσεις αυτοδιασποράς (ACVF) και αυτοσυσχέτισης ισχύουν

$$\gamma_k = \text{Cov}(X_t, X_{t-k}) = \begin{cases} (1 + \theta_1^2 + \theta_2^2)\sigma_e^2, & k = 0 \\ (-\theta_1 + \theta_1\theta_2)\sigma_e^2, & k = 1 \\ -\theta_2\sigma_e^2, & k = 2 \\ 0, & k > 2 \end{cases} \quad \text{και} \quad \rho_k = \begin{cases} \frac{-\theta_1(1-\theta_2)}{1 + \theta_1^2 + \theta_2^2}, & k = 1 \\ \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}, & k = 2 \\ 0, & k > 2 \end{cases}$$

ενώ η συνάρτηση μερικής αυτοσυσχέτισης PACF $\phi_{kk} = \begin{cases} \rho_1 = \frac{-\theta_1(1-\theta_2)}{1 + \theta_1^2 + \theta_2^2}, & k = 1 \\ \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}, & k = 2 \\ \frac{\rho_1^3 - \rho_1\rho_2(2 - \rho_2)}{1 - \rho_2^2 - 2\rho_1^2(1 - \rho_2)}, & k = 3 \end{cases}$

Σχέσεις οι οποίες προκύπτουν από το γεγονός ότι για $\rho_k = 0$ για $k > 2$

3.3.3 Διαδικασία Κινητού Μέσου τάξης q, MA(q)

Όπως έχει ήδη αναφερθεί η έκφραση $X_t = e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$ δίνει την διαδικασία κινητού μέσου τάξης q και συμβολίζεται με MA(q).

Μια διαδικασία κινητού μέσου είναι πάντα στάσιμη και αντιστρέψιμη όταν οι ρίζες του πολυωνύμου $\theta(B) = 0$ βρίσκονται έξω από το μοναδιαίο κύκλο. Εύκολα αποδεικνύεται ότι για την διαδικασία MA(q) ισχύουν οι σχέσεις: $E(X_t) = 0$ και

$$\gamma_0 = \text{Var}(X_t) = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)\sigma_e^2 = \sigma_e^2 \sum_{s=0}^q \theta_s^2 \quad \text{όπου} \quad \theta_0 = 1$$

Η ACVF της MA(q) είναι

$$\gamma_k = \text{Cov}(X_t, X_{t-k}) = \begin{cases} \sigma_e^2(-\theta_k + \theta_1\theta_{k+1} + \theta_{q-k}\theta_q), & k = 1, 2, \dots, q \\ 0, & k > q \end{cases} \quad \text{και} \quad \eta \quad \text{ACF} \quad \text{είναι}$$

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1\theta_{k+1} + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2}, & k = 1, 2, \dots, q \\ 0, & k > q \end{cases}$$

Τέλος οι μερικές αυτοσυσχετίσεις ϕ_{kk} μπορούν να εκφραστούν ως συναρτήσεις των αυτοσυσχέσεων ρ_s με τον ίδιο ακριβώς τρόπο όπως και για τις AR διαδικασίες.

Ο υπολογισμός γίνεται με βάση τις σχέσεις: $\Pi\Phi_k = R$ ή $\Phi_k = \Pi^{-1}R$ δηλαδή με την

$$\text{μέθοδο Cramer : } \varphi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}}$$

Σημειώνεται ότι, γενικά ενώ η συνάρτηση αυτοσυσχέτισης μιας AR(p) διαδικασίας μπορεί να εκτείνεται στο άπειρο, η συνάρτηση αυτοσυσχέτισης μιας AM(q) διαδικασίας τερματίζεται (μηδενίζεται) μετά από υστερήσεις. Η ‘μνήμη’ της, με άλλα λόγια, εξαντλείται σε περιόδους. Αντίθετα, η συνάρτηση μερικής αυτοσυσχέτισης μιας διαδικασίας τερματίζεται μετά από q υστερήσεις, ενώ η συνάρτηση μερικής αυτοσυσχέτισης μιας MA(q) διαδικασίας εκτείνεται στο άπειρο.

3.3.4 Σχέση μεταξύ MA(q) και AR(p) διαδικασιών.

Με αφορμή την προηγούμενη σημείωση και την παρατήρηση περί αντιστρεψιμότητας μιας διαδικασίας που θα μελετηθεί η σχέση μεταξύ MA(q) και AR(p) διαδικασιών.

Μια AR(p) διαδικασία μπορεί να γραφτεί ως εξής:

$$\Phi(B)X_t = e_t \rightarrow X_t = \frac{1}{\Phi(B)}e_t = \Psi(B)X_t \quad \text{με} \quad \Psi(B) = 1 + \psi_2 B^2 + \dots \quad \text{τέτοιο} \quad \text{ώστε}$$

$$\Phi(B)\Psi(B) = 1$$

Τα Ψ_j βάρη μπορούν να υπολογιστούν εξισώνοντας τους συντελεστές ιδίων δυνάμεων του B_j . Μάλιστα αποδεικνύεται ότι για μοντέλο AR(2), η ανάδρομη σχέση που συνδέει τα ψ_j με τα φ_j είναι $\psi_j = \psi_{j-1}\varphi_1 + \psi_{j-2}\varphi_2, j \geq 1, \psi_0 = 1$.

Δηλαδή, με άλλα λόγια, μια AR διαδικασία πεπερασμένης τάξης ισοδυναμεί με μια MA διαδικασία απείρου τάξης.

Μια MA(q) διαδικασία αντίστοιχα μπορεί να γραφτεί:

$$X_t = \theta(B)e_t \rightarrow \Pi(B)X_t = \frac{1}{\theta(B)}X_t = e_t \quad \text{με} \quad \Pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots \quad \text{τέτοιο} \quad \text{ώστε}$$

$$\Pi(B)\theta(B) = 1$$

Τα π_j μπορούν να υπολογιστούν όπως και προηγουμένως εξισώνοντας τους συντελεστές ιδίων δυνάμεων του B^j και έτσι προκύπτει το συμπέρασμα πως μια MA διαδικασία ισοδυναμεί με μια AR απείρου τάξης.

3.4 Μεικτές διαδικασίες ARMA(p,q)

Η έκφραση:

$$X_t - \varphi_1 X_{t-1} - \varphi_2 X_{t-2} - \dots - \varphi_p X_{t-p} = e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \rightarrow \Phi(B)X_t = \theta(B)e_t \rightarrow X_t = \frac{\theta(B)}{\Phi(B)}e_t$$

είναι συνδυασμός μιας $AR(p)$ και μιας $MA(q)$ διαδικασίας. Για αυτό αποκαλείται μεικτή αυτοπαλίνδρομη - κινητού μέσου διαδικασία τάξης (p, q) ή $ARMA(p, q)$. Είναι προφανές ότι η καθαρά αυτοπαλίνδρομη μορφή ή μια καθαρή μορφή κινητού μέσου μπορούν να θεωρηθούν ως ειδικές περιπτώσεις μιας $ARMA$ διαδικασίας. Δηλαδή

$$AR(p) = ARMA(p, 0) \text{ και } MA(q) = ARMA(0, q).$$

Αποδεικνύεται ότι ACF η της $ARMA(p, q)$ διαδικασίας είναι:

$$\rho_k = \begin{cases} \varphi_1 \rho_{k-1} + \dots + \varphi_p \rho_{k-p}, & k \geq q+1 \\ \varphi_1 \rho_{k-1} + \dots + \varphi_p \rho_{k-p} + \frac{1}{\gamma_0} E(X_{t-k} e_{t-q}), & k \leq q \end{cases}$$

Γενικά, η συνάρτηση αυτοσυσχετίσεως μιας $ARMA(p, q)$ διαδικασίας θα συμπεριφέρεται όπως αυτή μιας $AR(p)$ διαδικασία, ενώ η συνάρτηση μερικής αυτοσυσχετίσεως θα συμπεριφέρεται όπως αυτή μιας $MA(q)$ διαδικασίας για $k > q - p$. Συνοπτικά όλα τα παραπάνω και για τις σχέσεις των συναρτήσεων αυτοσυσχετίσεως και μερικής αυτοσυσχέτισης μεταξύ των στάσιμων χρονολογικών σειρών φαίνεται στον παρακάτω πίνακα.

Διαδικασία	Συνάρτηση Αυτοσυσχετίσεως	Συνάρτηση Μερικής Αυτοσυσχετίσεως
Λευκός Θόρυβος	Μηδέν	Μηδέν
$AR(p)$	Φθίνει εκθετικά ή φθίνει ακολουθώντας ημιτονοειδή συμπεριφορά	Μηδενίζεται μετά από p υστερήσεις
$MA(q)$	Μηδενίζεται μετά από p υστερήσεις	Φθίνει εκθετικά
$ARMA(p, q)$	Φθίνει εκθετικά	Φθίνει εκθετικά

Έτσι, όπως γίνεται φανερό, για τον προσδιορισμό μοντέλου χρησιμοποιούνται και οι μερικές δειγματικές αυτοσυσχετίσεις, με την βοήθεια των οποίων ορίζεται το είδος και η τάξη του προτεινόμενου μοντέλου. Βέβαια πάντα πρέπει να γίνεται η προσπάθεια για κατάληξη σε μοντέλα με μικρή τάξη.

Τέλος για $MA(q)$ διαδικασίες η δειγματική συνάρτηση αυτοσυσχέτισης σταματά να είναι σημαντική από την υστέρηση $k > q$. Ενώ για $AR(p)$ διαδικασίες η δειγματική συνάρτηση αυτοσυσχέτισης ελαττώνεται ομαλά, ενώ η δειγματική συνάρτηση μερικής αυτοσυσχέτισης σταματά να είναι σημαντική από την υστέρηση $k > q$. Έτσι, όπως γίνεται αντιληπτό, οι δειγματικές αυτοσυσχετίσεις και οι μερικές αυτοσυσχετίσεις μαρτυρούν την ταυτότητα των αντιστοιχών πληθυσμιακών.

Ξαναγυρίζοντας στις μεικτές διαδικασίες, μια $ARMA(p, q)$ διαδικασία είναι στάσιμη όταν οι ρίζες του AR πολυωνύμου είναι έξω από τον μοναδιαίο κύκλο και αντιστρέψιμη όταν οι ρίζες του MA πολυωνύμου είναι έξω από τον μοναδιαίο κύκλο. Μπορεί, όπως ήδη έχει δείχτει μια $ARMA$ διαδικασία να γραφεί ως μια καθαρά MA διαδικασία ή μια AR διαδικασία. Η διαφορά είναι ότι χρειάζεται μια υψηλής τάξεως MA διαδικασία από ότι με μια AR διαδικασία, δηλαδή απαιτεί περισσότερους συντελεστές. Έτσι, γίνεται φανερή, η 'οικονομία' που επιτυγχάνεται με την χρήση των μεικτών διαδικασιών. Για παράδειγμα, μια $ARMA(1,1)$ διαδικασία μπορεί να γραφεί ως μια $MA(\infty)$. Επίσης, στην συνάρτηση αυτοσυσχέτισης υπεισέρχεται ο συντελεστής από την $MA(1)$ διαδικασία αλλά μόνο για την αυτοσυσχέτιση πρώτης τάξης (ρ_1). Οι υπόλοιπες αυτοσυσχετίσεις εξαρτώνται μόνο από το αυτοπαλίνδρομο μέρος.

3.4.1 ARMA(1,1) διαδικασία.

Η απλούστερη μορφή μιας $ARMA$ διαδικασίας είναι η $ARMA(1,1)$ διαδικασία. Η σχέση: $X_t - \varphi_1 X_{t-1} - \varphi_2 X_{t-2} - \dots - \varphi_p X_{t-p} = e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$ για $p=1, q=1$ γίνεται: $X_t - \varphi X_{t-1} = e_t - \theta e_{t-1}$ για $-1 < \varphi < 1$

Επίσης εύκολα αποδεικνύετε ότι:

$$\gamma_0 = \text{Var}(X_t) = \varphi\gamma_1 + \sigma_e^2 - \theta(\varphi - \theta)\sigma_e^2 \text{ και } \gamma_k = \begin{cases} \varphi\gamma_0 - \theta\sigma_e^2, & k=1 \\ \varphi\gamma_{k-1}, & k \geq 2 \end{cases}$$

Για την συνάρτηση αυτοσυσχέτισης ισχύει:

$$\rho_k = \begin{cases} \frac{(1 - \varphi\theta)(\varphi - \theta)}{1 - \theta^2 - 2\varphi\theta}, & k=1 \\ \varphi\rho_{k-1}, & k \geq 2 \end{cases}$$

Όπως γίνεται φανερό η συνάρτηση αυτοσυσχέτισης ελαττώνεται εκθετικά καθώς το k αυξάνεται. Η μείωση αρχίζει όμως από το ρ_1 και από την μονάδα $\rho_0 = 1$ όπως στην περίπτωση $AR(1)$ της διαδικασίας. Η συνάρτηση μερικής αυτοσυσχέτισης συμπεριφέρεται όπως στην περίπτωση της $MA(1)$ διαδικασίας, δηλαδή φθίνει εκθετικά. Οι

μερικές αυτοσυσχετίσεις προσδιορίζονται όπως στην περίπτωση μιας διαδικασίας $AR(p)$. Για παράδειγμα, οι δυο πρώτες μερικές αυτοσυσχετίσεις με βάση τις γνώστες σχέσεις είναι οι εξής:

$$\varphi_{11} = \rho_{11} \text{ και } \varphi_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{\varphi\rho_1 - \rho_1^2}{1 - \rho_1^2}$$

3.5 Το μοντέλο $ARIMA(p,d,q)$

Σε ένα τυχαίο περίπατο οι διαφορές πρώτης τάξης οδηγούν σε διαδικασία λευκού θορύβου, δηλαδή σε στάσιμη διαδικασία. Στην γενική περίπτωση όμως, μια στάσιμη στοχαστική διαδικασία που προκύπτει παίρνοντας διαφορές κάποιας τάξης δεν είναι λευκός θόρυβος αλλά $ARMA(p,q)$ στάσιμη διαδικασία.

Οι *Box-Jenkins* προτείνουν για μια μη στάσιμη χρονολογική σειρά την χρήση διαφορών πρώτης, δεύτερης ή d τάξεως για την επίτευξη στασιμότητας. Στην νέα σειρά Y_t προσαρμόζεται ένα μοντέλο $ARMA(p,q)$.

Γενικά ένα $ARMA(p,q)$ μοντέλο που εφαρμόζεται σε μια ολοκληρωμένη σειρά d τάξεως ονομάζεται αυτοπαλίνδρομο ολοκληρωμένο μοντέλο κινητού μέσου τάξεως (p,q,d) (Autoregressive Integrated Moving Average) και συμβολίζεται με $ARIMA(p,q,d)$. Με άλλα λόγια, $ARIMA(p,q,d)$ διαδικασία, είναι μια διαδικασία η οποία ‘διαφορίζεται’ (κατά αναλογία με το ‘ολοκληρωμένη’) d φορές και παράγει $ARMA(p,q)$ διαδικασία.

Η έκφραση της ολοκληρωμένης διαδικασίας κινητού μέσου $ARIMA(p,q,d)$ με την βοήθεια των πολυώνυμων είναι:

$$\Phi(B)\nabla^d X_t = \Theta(B)e_t \text{ ή } \Phi(B)\nabla^d X_t = \Theta(B)e_t \text{ όπου } Y_t = \nabla^d X_t$$

Το πολυώνυμο $\Phi(B)(1-B)^d$ έχει μια ρίζα ίση με την μονάδα, τάξης d , και όλες τις άλλες έξω από τον μοναδιαίο κύκλο.

Η ανάπτυξη και η κατασκευή μοντέλων $ARIMA$ ως εργαλεία πρόβλεψης είναι γνωστή ως μεθοδολογία $ARIMA$ ή τεχνικές *Box-Jenkins*.

3.6 Μεθοδολογία *Box-Jenkins*

Η προσέγγιση *Box-Jenkins* στην ανάλυση χρονολογικών σειρών, δηλαδή η προσαρμογή $ARIMA$ μοντέλων σε κάποια δεδομένα, απαιτεί τα παρακάτω βήματα:

Πρώτο βήμα είναι να κατασκευαστούν τα γραφήματα της σειράς, των αυτόσυσχετίσεων και των μερικών αυτοσυσχετίσεων. Με το βήμα αυτό δίνεται η δυνατότητα να διαπιστωθούν ορισμένα κρίσιμα στοιχεία για την εξεύρεση ενός $ARIMA$ μοντέλου. Εάν όλες οι αυτοσυσχετίσεις είναι ασήμαντες, ενδεχομένως ένα μη γραμμικό μοντέλο να προσαρμόζεται. Καλό είναι στις περιπτώσεις αυτές να εφαρμόζεται και δεύτερος έλεγχος ανεξαρτησίας, έλεγχος που είναι γνωστός σαν έλεγχος μοναδιαίας ρίζας.

Επίσης, εάν η σειρά δεν είναι στάσιμη, εφαρμόζοντας κάποιον μετασχηματισμό ή παίρνοντας διαφορές κάποιας τάξης, μετατρέπεται σε στάσιμη. Δηλαδή λαμβάνουμε νέα δεδομένα που αντιστοιχούν σε στάσιμη χρονολογική σειρά. Με τα νέα αυτά δεδομένα προσδιορίζεται η τάξη του *ARIMA* μοντέλου.

Στο δεύτερο βήμα εφαρμόζεται κάποιο εκ των γνωστών μοντέλων, αυτοπαλίνδρομο (*AR*) , κινητού μέσου (*MA*) ή μεικτό (*ARMA*) στις νέες παρατηρήσεις. Δηλαδή ακολουθεί η εκτίμηση των p παραμέτρων $\varphi_1, \varphi_2, \dots, \varphi_p$ της αυτοπαλίνδρομης (*AR*) διαδικασίας και των q παραμέτρων $\theta_1, \theta_2, \dots, \theta_q$ της διαδικασίας κινητού μέσου (*MA*) .

Στο επόμενο βήμα γίνεται έλεγχος του μοντέλου για την προσαρμοστικότητα του. Δηλαδή, πόσο καλά ‘ταιριάζει’ το εκτιμώμενο μοντέλο με τα δεδομένα, αφού πιθανώς κάποιο άλλο μοντέλο *ARIMA* να προσαρμόζεται καλύτερα. Ο έλεγχος αυτός, αφορά πέραν των γνωστών στατιστικών ελέγχων για την σημαντικότητα των συντελεστών, την συμπεριφορά των καταλοίπων και την τάξη του μοντέλου.

Εάν το εκτιμώμενο *ARMA* μεικτό μοντέλο ταιριάζει με τα δεδομένα, τότε τα κατάλοιπα θα πρέπει να συμπεριφέρονται ως διαδικασία λευκού θορύβου. Δηλαδή τα κατάλοιπα δεν θα πρέπει να αυτοσυσχετίζονται. Ο έλεγχος των καταλοίπων γίνεται με την στατιστική συνάρτηση Q των *Box–Peirce* (*Box–Peirce Q Statistic*) με την οποία ελέγχεται η σημαντικότητα από κοινού ενός αριθμού συντελεστών αυτοσυσχετίσεων, έστω m .Δηλαδή ελέγχετε η μηδενική υπόθεση:

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_m = 0$$

$$H_1 : \text{διαφορετικά}$$

Η στατιστική Q ορίζεται ως εξής: $Q_{BP} = n \sum_{k=1}^m \hat{\rho}_k^2$ όπου $\hat{\rho}_k$ είναι οι δειγματικές αυτοσυσχετίσεις των καταλοίπων και ο αριθμός των παρατηρήσεων (καταλοίπων). Συνήθως, ο αριθμός των αυτοσυσχετίσεων των καταλοίπων που χρησιμοποιούνται για τον υπολογισμό της στατιστικής Q_{BP} ισούται με την τετραγωνική ρίζα του αριθμού των παρατηρήσεων $m = \sqrt{n}$. Η παραπάνω στατιστική Q_{BP} ακολουθεί προσεγγιστικά την χ^2 κατανομή με $m-p-q$ βαθμούς ελευθερίας. Για δεδομένο επίπεδο σημαντικότητας α , η μηδενική υπόθεση ότι όλοι οι συντελεστές αυτοσυσχετίσεως είναι μηδέν απορρίπτεται, εάν η τιμή της Q_{BP} υπερβαίνει την κρίσιμη τιμή της χ^2 , δηλαδή εάν $Q_{BP} > \chi_a^2$. Μια τροποποιημένη μορφή της στατιστικής των *Box–Peirce* είναι η στατιστική που πρότειναν οι *Ljung – Box* και η

$$\text{οποία ορίζεται ως εξής: } Q_{LB} = n(n+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{n-k}$$

Η παραπάνω στατιστική ακολουθεί προσεγγιστικά την κατανομή χ^2 με $m-p-q$ βαθμούς ελευθερίας. Η στατιστική Q_{LB} θεωρείται καταλληλότερη για μικρά δείγματα. Για μεγάλα δείγματα, και οι δυο στατιστικές έχουν ανάλογη συμπεριφορά.

Τέταρτο βήμα αποτελεί ο έλεγχος της τάξης του μοντέλου. Ο έλεγχος αυτός γίνεται συγκρίνοντας το μοντέλο με ένα άλλο μοντέλο μεγαλύτερης τάξης. Δηλαδή, το

εκτιμώμενο μοντέλο $ARMA(p, q)$ συγκρίνεται με τα μοντέλα $ARMA(p, q+1)$ και $ARMA(p+1, q)$. Εάν το μοντέλο που εκτιμήθηκε περιγράφει την διαδικασία που παρήγαγε τα δεδομένα ικανοποιητικά, οι επιπλέον συντελεστές στα μμεγαλύτερα τάξεως μοντέλα δεν θα πρέπει να είναι στατιστικά διαφορετικοί από το μηδέν.

Τέλος, θα αναφερθούν ορισμένα κριτήρια επιλογής του κατάλληλου μοντέλου. Είναι φανερό ότι αυξάνοντας την τάξη του μοντέλου, δηλαδή προσθέτοντας υστερήσεις για το αυτοπαλίνδρομο τμήμα ή και για το τμήμα του κινητού μέσου, θα μειώνεται το άθροισμα των τετραγώνων των καταλοίπων αλλά θα μειώνονται και οι βαθμοί ελευθερίας, αφού εκτιμούνται περισσότεροι παράμετροι. Δηλαδή, δεν υπάρχει μόνο 'κέρδος' από την προσθήκη μεταβλητών αλλά και 'κόστος'.

Στην ανάλυση χρονολογικών σειρών χρησιμοποιούνται ευρέως τα παρακάτω δυο κριτήρια για την επιλογή του καταλληλότερου μοντέλου. Το Κριτήριο Πληροφοριών *Akaike* (*Akaike Information Criterion AIC*) και το Μπαϊεσιανό Κριτήριο *Schwartz* (*Schwartz Bayesian Criterion SBC*).

Τα κριτήρια αυτά ορίζονται ως εξής:

$$AIC = \ln \frac{\sum \hat{u}^2}{n} + \frac{2s'}{n} \quad \text{και} \quad SBC = \ln \frac{\sum \hat{u}^2}{n} + \frac{s'}{n} \ln n$$

όπου $\sum \hat{u}^2$ το άθροισμα των τετραγώνων των καταλοίπων, n ο αριθμός των παρατηρήσεων και s' ο αριθμός των παραμέτρων που εκτιμώνται ($p+q+1$).

Με βάση τα παραπάνω κριτήρια, επιλέγεται το μοντέλο με την μικρότερη τιμή. Και τα δυο κριτήρια μπορούν να πάρουν και αρνητικές τιμές.

Το τελικό μοντέλο που έχει παραχθεί με την παραπάνω διαδικασία χρησιμοποιείται για την πρόβλεψη.

Κεφάλαιο 4^ο

Παράδειγμα εφαρμογής χρονολογικών σειρών

Κεφάλαιο 4ο : Παράδειγμα εφαρμογής χρονολογικών σειρών

4.1 Εισαγωγή στο πρόγραμμα PASW 18

Στην παρούσα πτυχιακή εργασία ως εφαρμογή των χρονολογικών σειρών, έγινε επεξεργασία δεδομένων ενεργειακών καταναλώσεων διαφόρων ευρωπαϊκών χωρών. Η επεξεργασία αυτών των δεδομένων έγινε μέσω του προγράμματος PASW 18 Superior Performance Software System (Σύστημα Λογισμικού Μέγιστης Παραγωγικότητας) που είναι το πιο διαδεδομένο πρόγραμμα για τη στατιστική ανάλυση δεδομένων.

Το PASW 18 είναι ένα στατιστικό πακέτο ανάλυσης δεδομένων, το οποίο προσφέρει στο χρήστη δυνατότητες για δημιουργία αναφορών, ανάλυση και μοντελοποίηση δεδομένων καθώς και για γραφική αναπαράστασή τους. Διαθέτει πολλές στατιστικές συναρτήσεις για ανάλυση δεδομένων μέσα από ένα εύχρηστο γραφικό περιβάλλον. Με τη βοήθεια του PASW 18 όλα τα στάδια της αναλυτικής διαδικασίας ολοκληρώνονται κάτω από ένα ενοποιημένο περιβάλλον εργασίας καλύπτοντας την ανάλυση από άκρο σε άκρο.



Εικόνα 1. PASW 18 [<http://en.wikipedia.org/wiki/SPSS>]

4.2 Αναφορά-Στοιχεία για την ενεργειακή κατανάλωση στην Ευρώπη

Στη σημερινή εποχή η κατανάλωση ενέργειας είναι ένα ζήτημα που απασχολεί ιδιαίτερα σε παγκόσμιο επίπεδο. Για το λόγο αυτό παρουσίασε ενδιαφέρον και θεωρήθηκε σκόπιμο να γίνει ανάλυση δεδομένων ενεργειακών καταναλώσεων σε διάφορες χώρες της Ευρώπης.

Η Ευρωπαϊκή Ένωση καταναλώνει όλο και περισσότερη ενέργεια και εισάγει όλο και περισσότερα ενεργειακά προϊόντα. Σήμερα η ενεργειακή κατανάλωση καλύπτεται κατά 41% από το πετρέλαιο, κατά 22% από το φυσικό αέριο, κατά 16% από τα στερεά καύσιμα, κατά 15% από την πυρηνική ενέργεια και κατά 6% από τις ανανεώσιμες πηγές ενέργειας.

Κατά καιρούς πολλές χώρες έχουν καταβάλλει προσπάθειες να περιορίσουν τις ενεργειακές τους καταναλώσεις και ιδιαίτερα την εξάρτησή τους από στερεά καύσιμα και πετρέλαιο. Ένα παράδειγμα είναι και αυτό της Γαλλίας. Με βάση το εθνικό Σύμφωνο Άνθρακα, το οποίο συνήφθη μεταξύ κοινωνικών εταίρων το 1995, η εξόρυξη άνθρακα μειώνεται προοδευτικά και θα σταματήσει οριστικά το 2005. Όλα τα ορυχεία έχουν συμπεριληφθεί σε ένα σχέδιο παύσης λειτουργίας και εισπράττουν, αποκλειστικά για την κάλυψη των απωλειών εκμετάλλευσης, ενισχύσεις για τη μείωση της δραστηριότητάς τους.

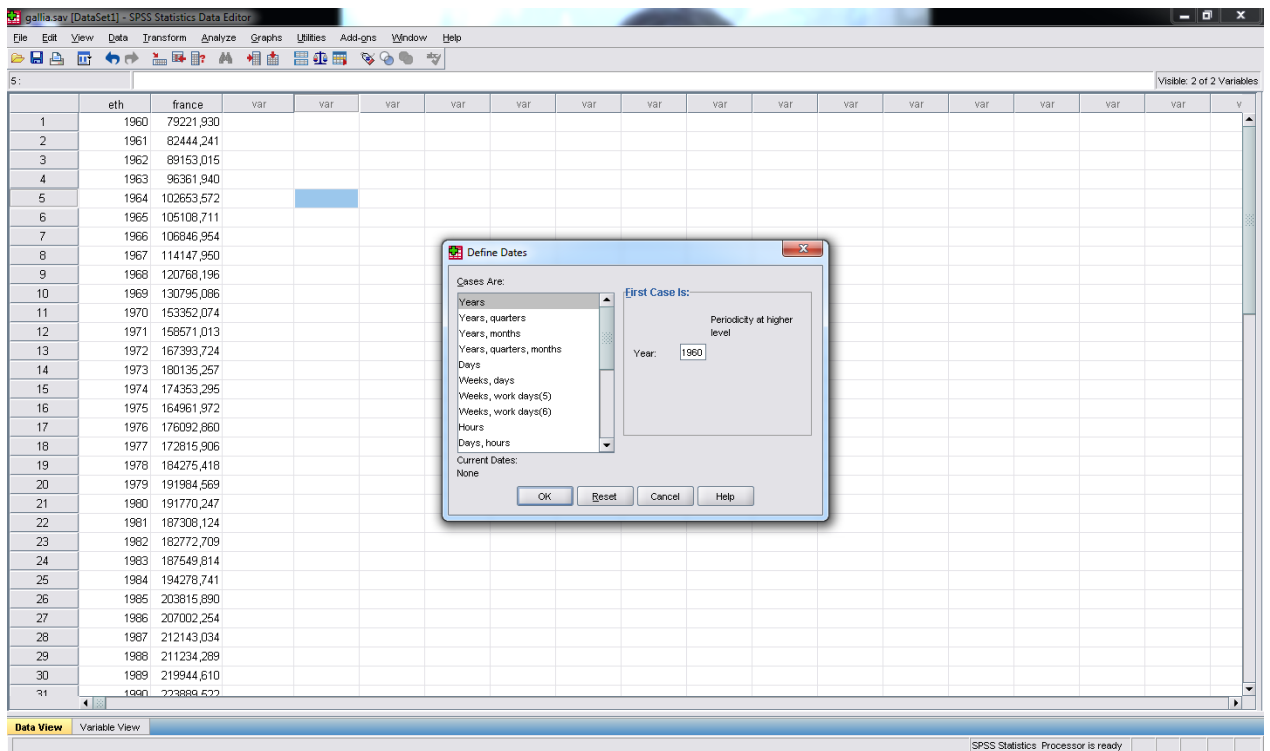
4.3 Ανάλυση των δεδομένων της ενεργειακής κατανάλωσης της Γαλλίας

Στο παράδειγμα που θα αναλύσουμε στην επόμενη ενότητα χρησιμοποιήθηκαν τα δεδομένα της ενεργειακής κατανάλωσης της Γαλλίας. Αρχικά εισάγονται στο PASW 18 οι τιμές της κατανάλωσης ενέργειας της Γαλλίας για τα έτη από το 1960 έως και το 2011.

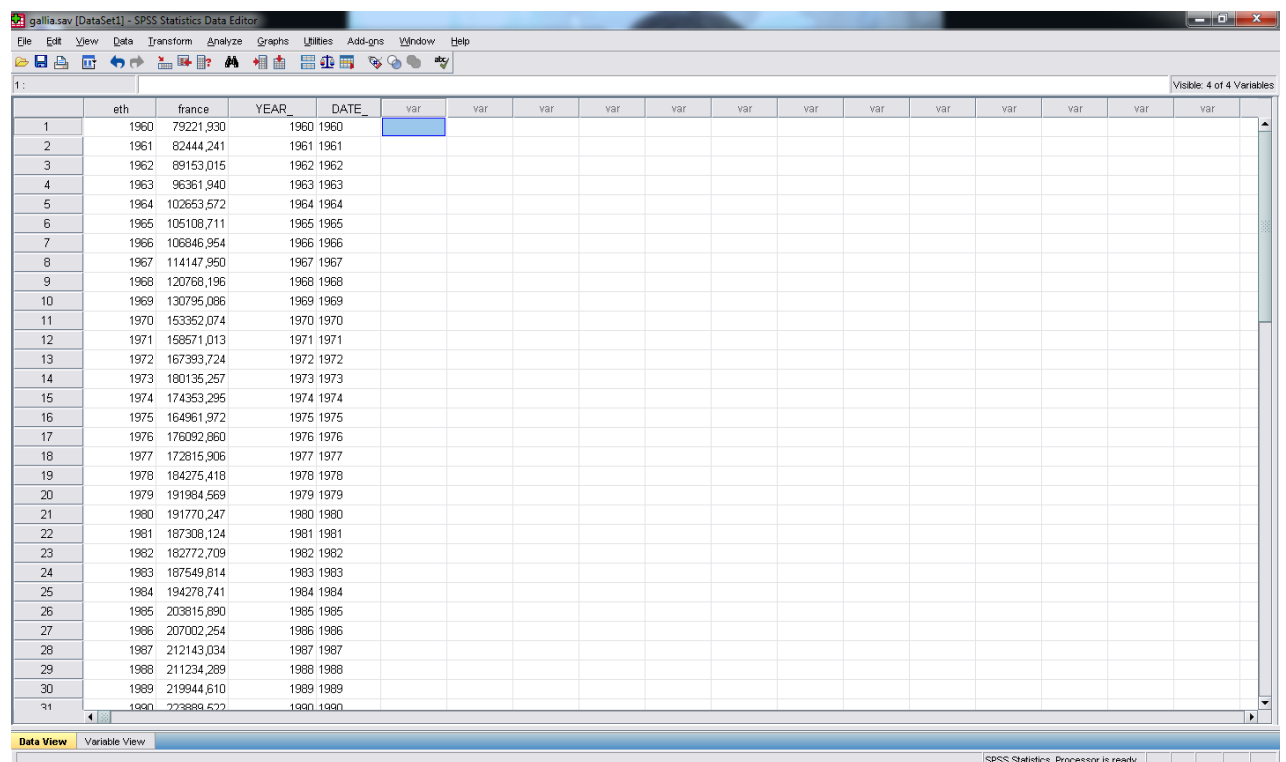
Το πρώτο βήμα είναι να ορίσουμε μέσα από το menu του προγράμματος τις χρονολογίες που μας ενδιαφέρουν προκειμένου να προβούμε στην περαιτέρω επεξεργασία των στοιχείων τους. Έτσι από το menu επιλέγω

Data→Define dates→Years

και στο παράθυρο years συμπληρώνουμε 1960 για να ορίσουμε την ημερομηνία έναρξης των αποτελεσμάτων.



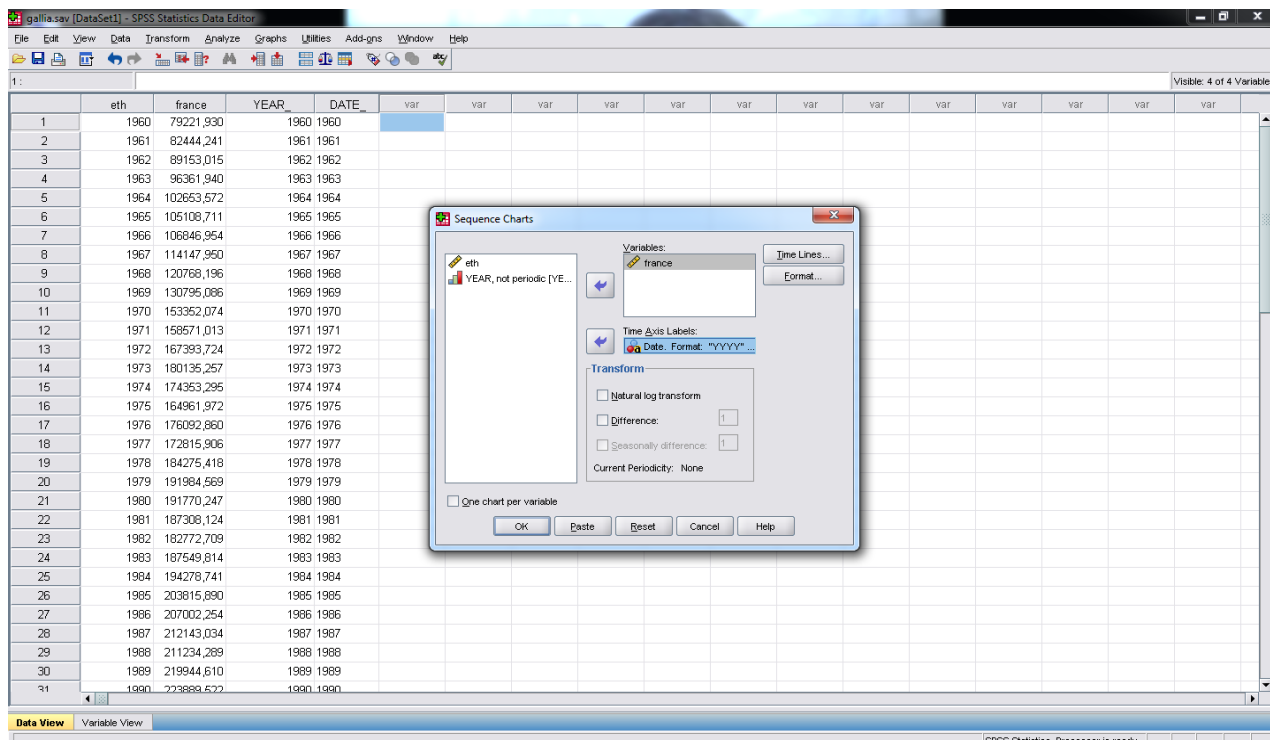
Σχήμα 1. Ορισμός ημερομηνίας



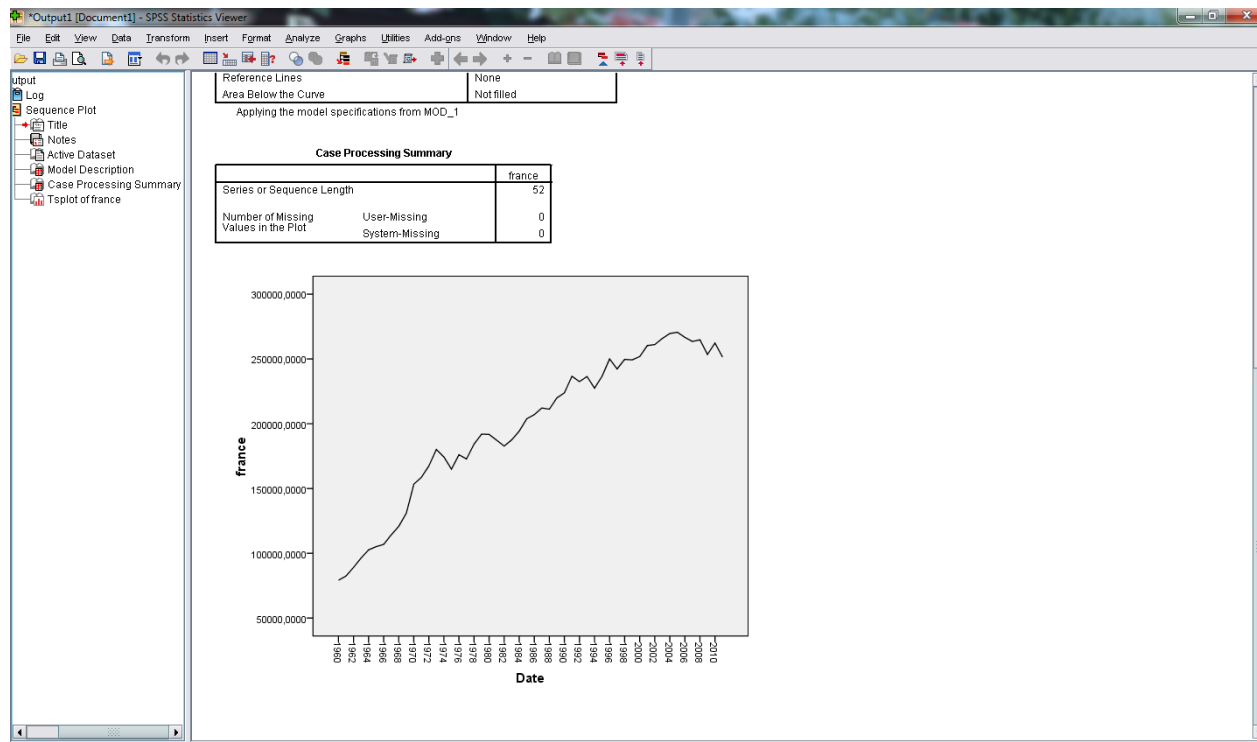
Σχήμα 2. Χρονολογία-Ημερομηνία

Στη συνέχεια θα κάνουμε έλεγχο της στασιμότητας των δεδομένων. Από το μενού επιλέγουμε

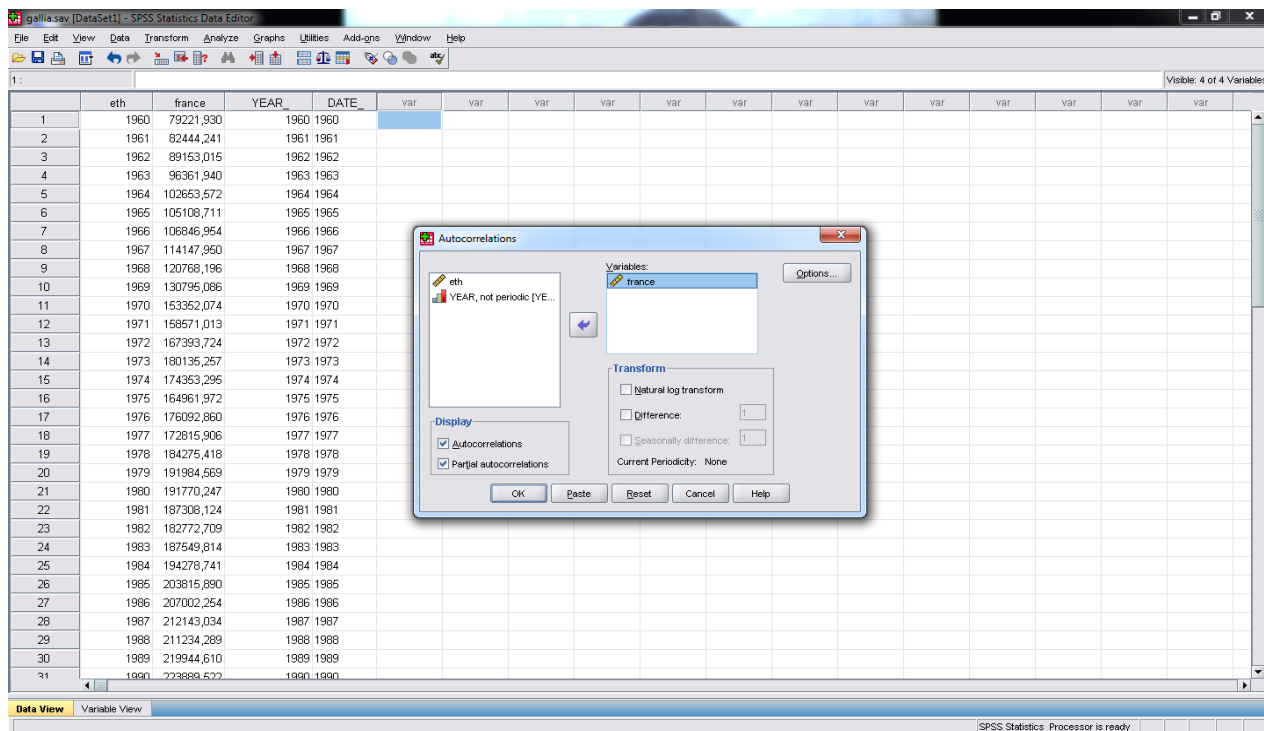
Analyze → Forecasting → Sequences Charts και στην επιλογή Variable βάζουμε τη μεταβλητή France στο time axis lab το date και πατάμε ok.



Σχήμα 3. Έλεγχος στασιμότητας



Σχήμα 4. Διαγράμμα Sequence



Σχήμα 5. Εξαγωγή Διαγράμματος ACF, PACF

Από την παραπάνω διαδικασία προκύπτει ο παρακάτω πίνακας αυτοσυσχετίσεων.

Autocorrelations

Series:france

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,941	,135	48,753	1	,000
2	,878	,133	92,033	2	,000
3	,817	,132	130,242	3	,000
4	,752	,131	163,357	4	,000
5	,692	,129	191,960	5	,000
6	,628	,128	216,030	6	,000
7	,558	,127	235,452	7	,000
8	,489	,125	250,698	8	,000
9	,422	,124	262,348	9	,000
10	,361	,122	271,062	10	,000
11	,314	,121	277,832	11	,000
12	,272	,119	283,045	12	,000
13	,234	,118	286,992	13	,000
14	,203	,116	290,038	14	,000
15	,168	,115	292,173	15	,000
16	,120	,113	293,293	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 1. Πίνακας αυτοσυσχετίσεων

Στον πίνακα αυτό εμφανίζονται οι παρακάτω όροι.

- ❖ Χρονολογικές υστερήσεις (lag)
- ❖ Τιμή του συντελεστή αυτοσυσχετίσης (Autocorrelation)
- ❖ Τυπικό σφάλμα (Std. error)
- ❖ Τιμή της στατιστικής συνάρτησης Box-Jung (Value)
- ❖ Βαθμοί ελευθερίας (df)
- ❖ Τιμή πιθανότητας στη στατιστική συνάρτηση (Sig)

Από τον πίνακα προκύπτει ότι για όλες τις χρονολογικές υστερήσεις υπάρχει αυτοσυσχετίση γιατί όλα τα p είναι μηδέν με αποτέλεσμα η χρονολογική σειρά να μην είναι στάσιμη με διαφορές. Έτσι θα κάνουμε δοκιμή με τη μέθοδο των πρώτων διαφορών. Από το μενού επιλέγουμε Analyze→Forecasting→Autocorrelation και επιλέγουμε Difference και πληκτρολογούμε 1. Έτσι παράγονται τα παρακάτω αποτελέσματα.

Autocorrelations

Series:france

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,052	,136	,148	1	,700
2	,160	,135	1,562	2	,458
3	,049	,133	1,694	3	,638
4	-,163	,132	3,232	4	,520
5	,047	,130	3,361	5	,645
6	,160	,129	4,896	6	,557
7	,065	,128	5,159	7	,641
8	-,022	,126	5,189	8	,737
9	,020	,125	5,214	9	,815
10	-,105	,123	5,937	10	,821
11	-,064	,122	6,216	11	,859
12	,066	,120	6,522	12	,888
13	-,047	,119	6,680	13	,918
14	,012	,117	6,691	14	,946
15	,062	,115	6,984	15	,958
16	-,122	,114	8,141	16	,945

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 2. Πίνακας αυτοσυσχετίσεων με τη μέθοδο των πρώτων διαφορών

Εφόσον οι τιμές δεν είναι μηδενικές κάνουμε έλεγχο στασιμότητας.

$H_0: \rho=0$

V_s

$H_1: \rho \neq 0$

Με την υπόθεση H_0 δεν υπάρχει αυτοσυσχέτιση, άρα είναι στάσιμη ενώ με την H_1 υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη. Ο έλεγχος αυτός πραγματοποιείται για όλα τα lags. Οπότε:

1ο lag:

$H_0: \rho_1=0$ $\rho_1=0,700 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_1 \neq 0$

2o lag:

$H_0: \rho_2=0$ $\rho_2=0,458 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_2 \neq 0$

3o lag:

$H_0: \rho_3=0$ $\rho_3=0,638 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_3 \neq 0$

4o lag:

$H_0: \rho_4=0$ $\rho_4=0,520 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_4 \neq 0$

.....

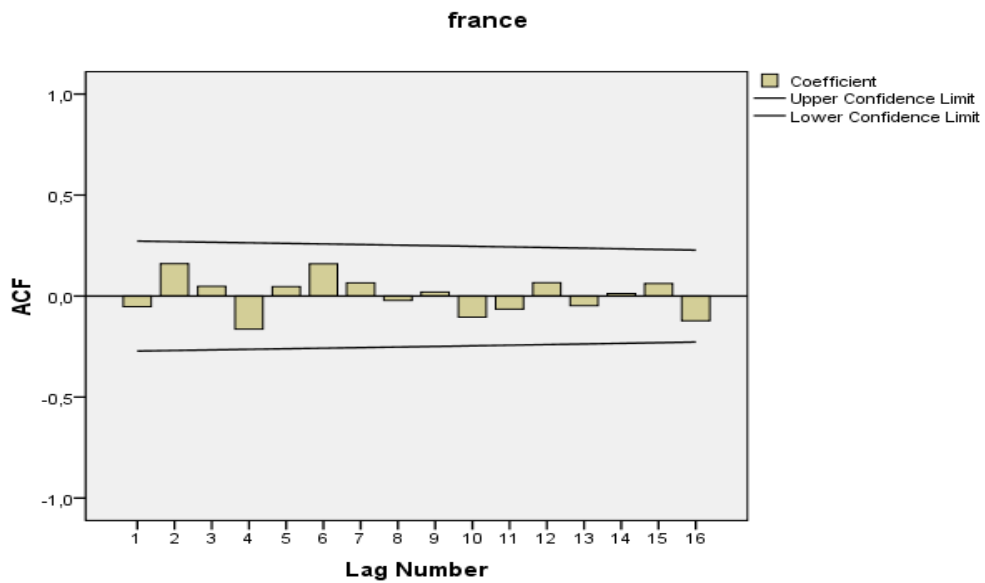
16o lag:

$H_0: \rho_{16}=0$ $\rho_{16}=0,945 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

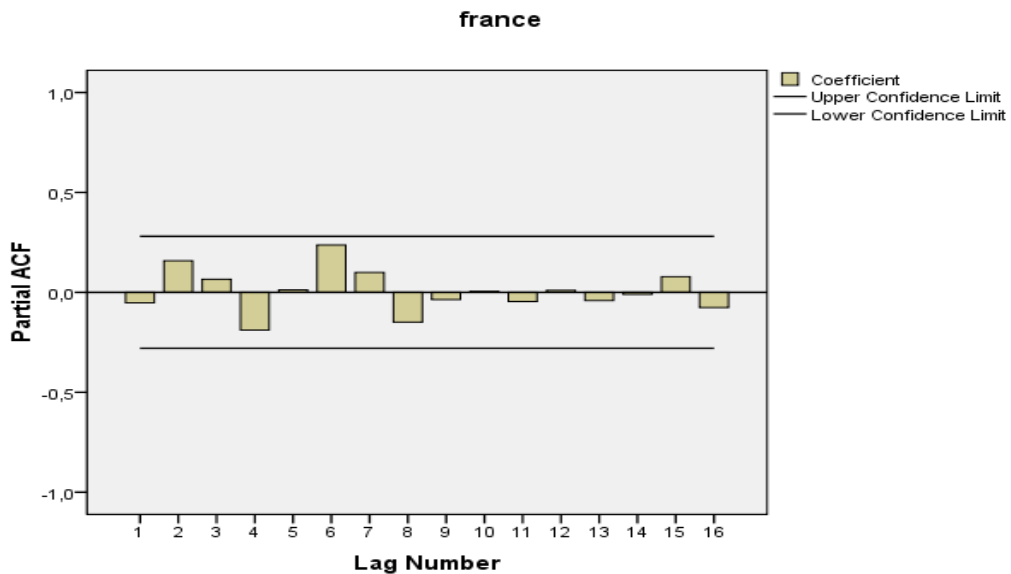
V_s

$H_1: \rho_{16} \neq 0$

Βάσει του ελέγχου κανονικότητας προκύπτει ότι η διαδικασία είναι στάσιμη με πρώτες διαφορές. Στη συνέχεια, από τα διαγράμματα ACF και PACF που εξάγονται από το πρόγραμμα εντοπίζονται τα q και p αντίστοιχα.



Σχήμα 6. Διάγραμμα ACF



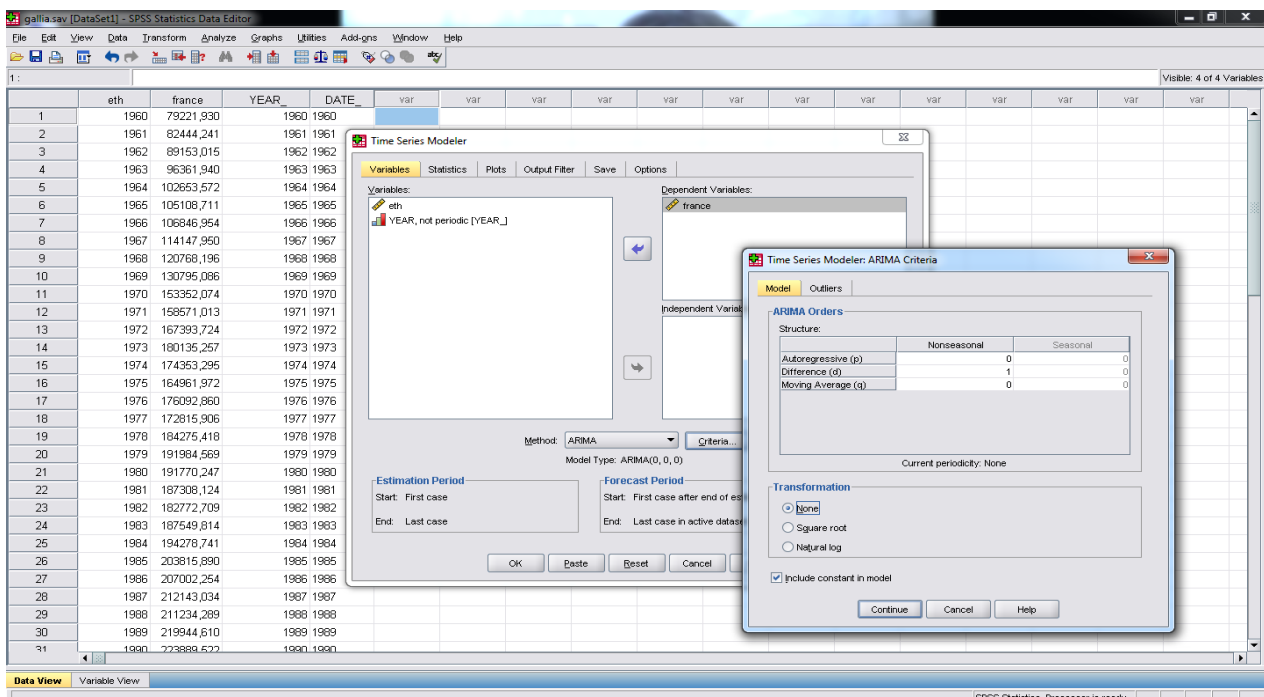
Σχήμα 7. Διάγραμμα PACF

Από τα διαγράμματα ACF και PACF προκύπτει $q=0$ και $p=0$. Το υπόδειγμα είναι της μορφής $ARIMA(p,d,q)$ - $ARIMA(0,1,0)$, ($d=1$ λόγω πρώτων διαφορών). Η θεωρητική μορφή του υποδείγματος είναι :

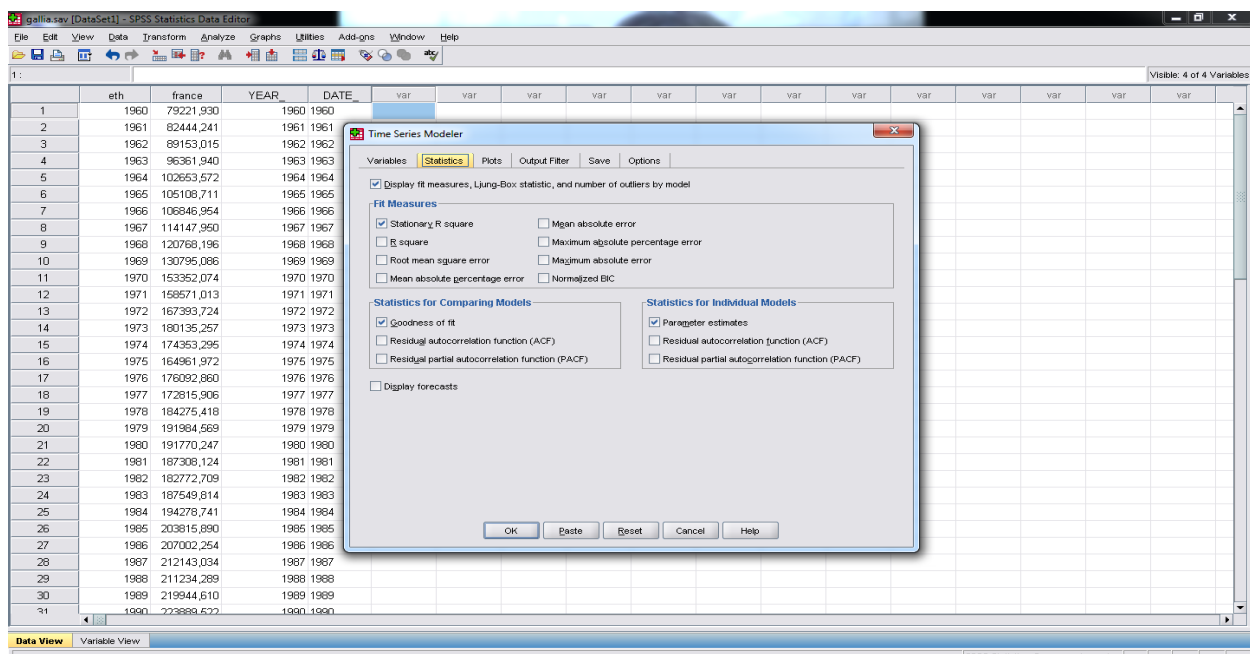
$$y_t = a_0 + a_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Έπειτα επιστρέφοντας στο PASW 18 ακολουθούμε τις παρακάτω διαδικασίες. Από το μενού επιλέγουμε Analyze → Forecasting → Create models. Στο Dependent Variable εισάγουμε τη μεταβλητή France και στο method επιλέγουμε ARIMA. Επιλέγοντας το criteria εμφανίζεται ο πίνακας ARIMA Criteria και στο nonseasonal εισάγονται

οι τιμές των p, d, q που έχουν εντοπιστεί δηλαδή 0,1 και 0 αντίστοιχα. Πατώντας το Continue και πηγαίνοντας στο Statistics επιλέγουμε Parameter Estimates και στη συνέχεια πατάμε ok.

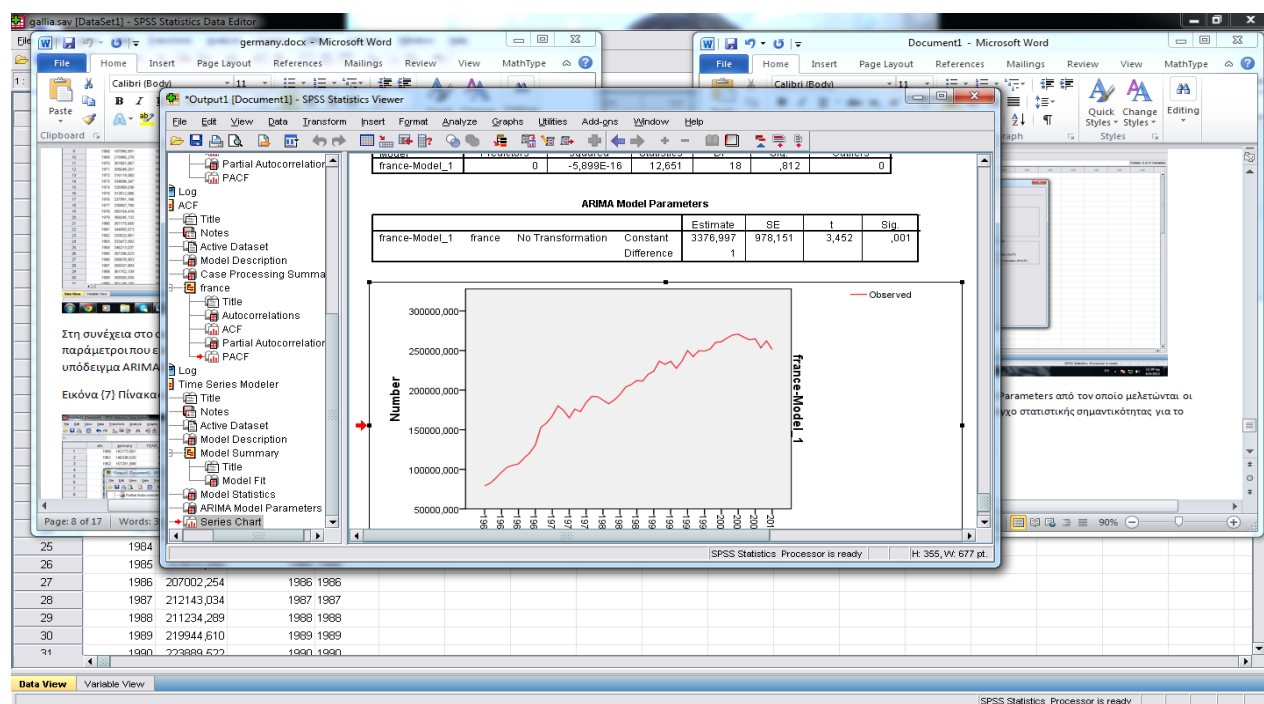


Σχήμα 8. Καθορισμός υποδείγματος ARIMA (0,1,0)



Σχήμα 9. Διαδικασία για έλεγχο στατιστικής σημαντικότητας

Στη συνέχεια εξάγεται ο πίνακας ARIMA Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί κάνοντας έλεγχο της στατιστικής σημαντικότητας για το υπόδειγμα ARIMA (0,1,0).



Σχήμα 10. Πίνακας ARIMA Models Parameters

Στη συνέχεια γίνεται έλεγχος στατιστικής σημαντικότητας p-value των εκτιμητών του υποδείγματος.

$$H_0: p=0$$

Vs

$$H_1: p \neq 0$$

Με την υπόθεση H_0 ο εκτιμητής μας είναι στατιστικά μη σημαντικός, ενώ με την H_1 ο εκτιμητής μας είναι στατιστικά σημαντικός.

$$1] H_0: p_0=0 \quad p_0=0,001 < \alpha=0,05 \quad \text{ο εκτιμητής είναι στατιστικά σημαντικός}$$

Vs

$$H_1: p_0 \neq 0$$

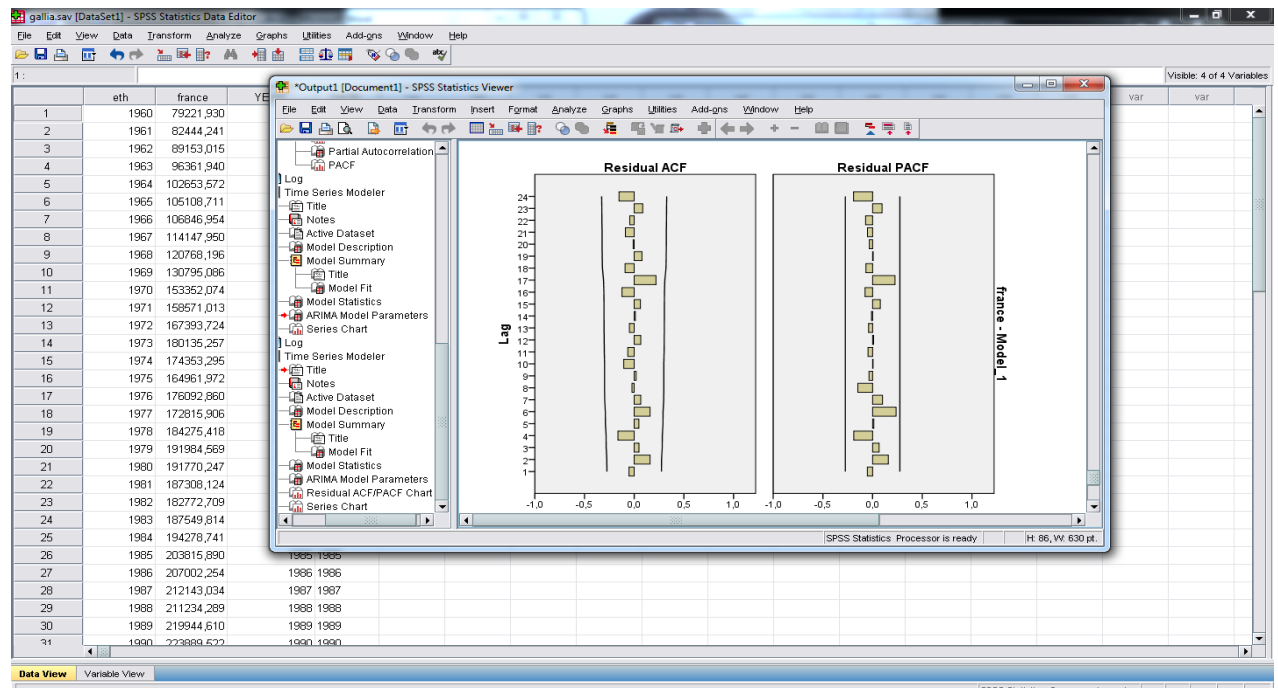
Αφού βρέθηκε ο εκτιμητής μας στατιστικά σημαντικός και γνωρίζοντας τις παραμέτρους του υποδείγματος ARIMA(0,1,0) θεωρείται ως το καταλληλότερο. Η μαθηματική μορφή του υποδείγματος θα είναι :

$$y_t = 3376,997 + E_t$$

Για να είναι το υπόδειγμα Y_t αποδεκτό, για να θεωρηθεί δηλαδή ότι περιγράφει την διαδικασία που παρήγαγε τα δεδομένα, θα πρέπει τα κατάλοιπα να είναι λευκός

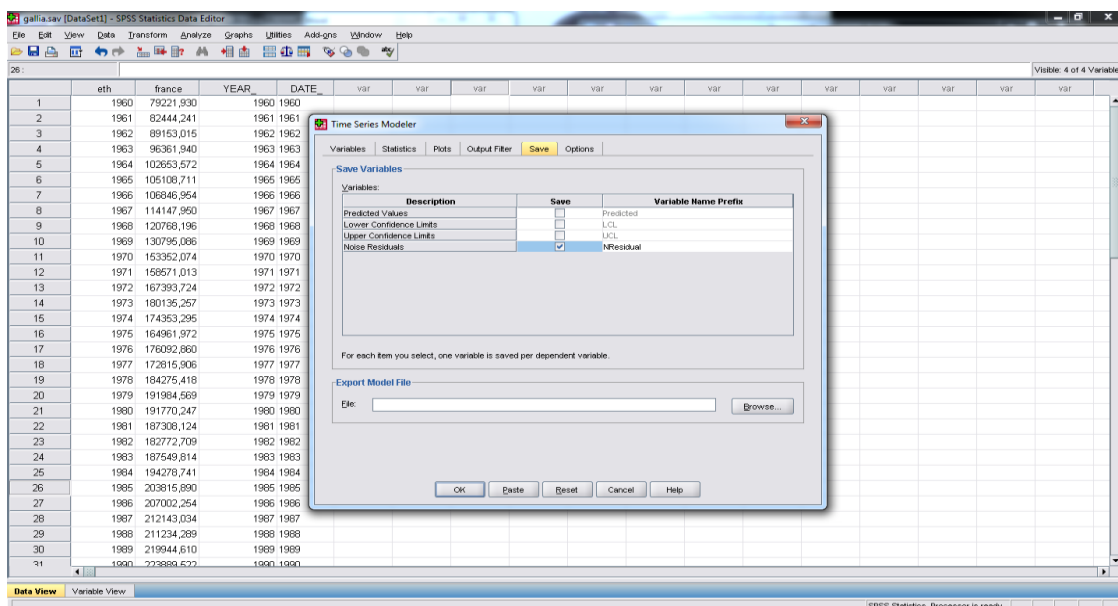
θόρυβος, δηλαδή δεν θα πρέπει να αυτοσυσχετίζονται. Για το λόγο αυτό θα γίνει έλεγχος στασιμότητας στα κατάλοιπα (residuals) επαληθεύοντας έτσι αν τα αποτελέσματα είναι αληθή για το ARIMA(0,1,0).

Έτσι στο PASW 18 επιλέγουμε Analyze→Forecasting→Create models→Plots και στη συνέχεια επιλέγουμε Residuals autocorrelations function (ACF) και Residuals partial autocorrelation function (PACF) και στη συνέχεια πατάμε ok.



Σχήμα 11. Εύρεση των καταλοίπων (residuals)

Παρατηρούμε ότι είναι και τα δύο μηδενικά οπότε έχουμε λευκό θόρυβο. Επιστρέφοντας και πάλι στο PASW 18 και στα δεδομένα από τις επιλογές Analyze→Forecasting→Create models και στην επιλογή Save επιλέγουμε το Noise Residuals και στη συνέχεια πατάμε ok.



Σχήμα 12. Διαδικασία για την εύρεση των καταλοίπων (residuals)

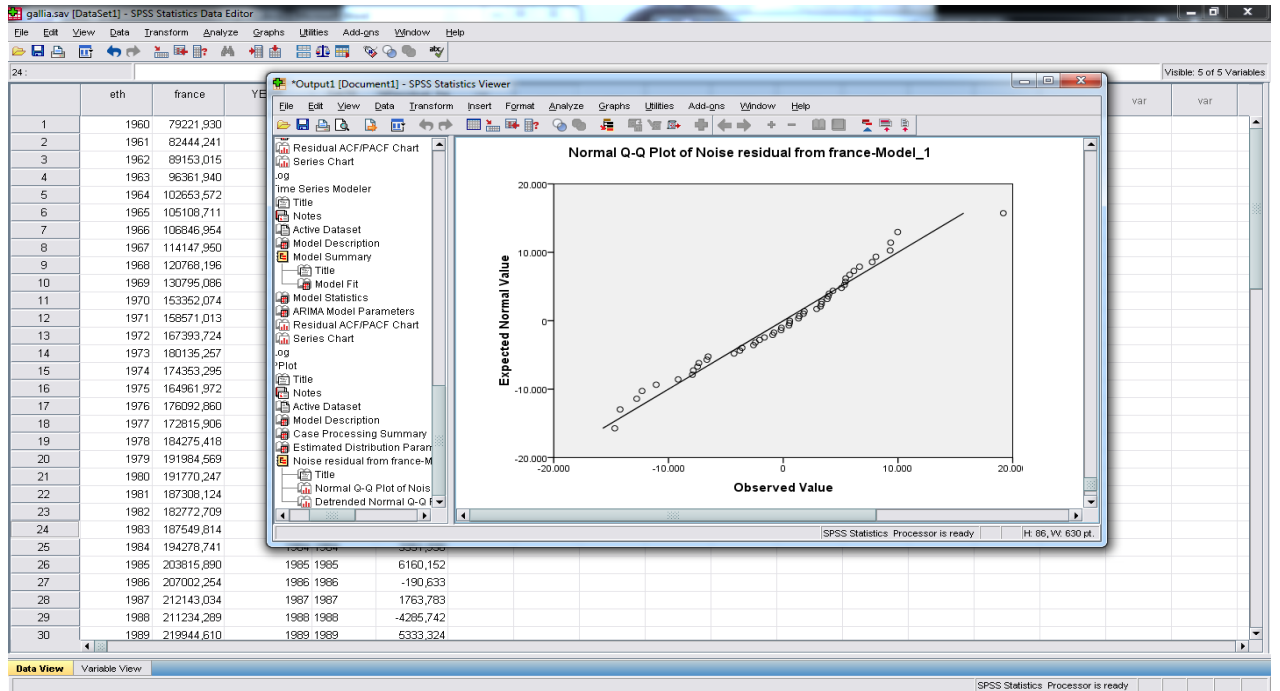
Πατώντας ok εμφανίζεται η μεταβλητή Noise Residuals France στο PASW 18 που δείχνει τα υπόλοιπα.

	eth	france	YEAR_	DATE_	NResidual_france_Model_1
1	1960	79221,930	1960	1960	.
2	1961	82444,241	1961	1961	-154,686
3	1962	89153,015	1962	1962	3331,777
4	1963	96361,940	1963	1963	3831,928
5	1964	102653,572	1964	1964	2914,635
6	1965	105108,711	1965	1965	-921,858
7	1966	106846,954	1966	1966	-1638,754
8	1967	114147,950	1967	1967	3923,999
9	1968	120768,196	1968	1968	3243,249
10	1969	130795,086	1969	1969	6649,893
11	1970	153352,074	1970	1970	19179,991
12	1971	158571,013	1971	1971	1841,942
13	1972	167393,724	1972	1972	5445,714
14	1973	180135,257	1973	1973	9364,536
15	1974	174353,295	1974	1974	-9159,959
16	1975	164961,972	1975	1975	-12768,320
17	1976	176092,860	1976	1976	7753,891
18	1977	172815,906	1977	1977	-6653,951
19	1978	184275,418	1978	1978	8082,515
20	1979	191984,569	1979	1979	4332,154
21	1980	191770,247	1980	1980	-3691,319
22	1981	187308,124	1981	1981	-7839,120
23	1982	182772,709	1982	1982	-7912,412
24	1983	187549,814	1983	1983	1400,108
25	1984	194278,741	1984	1984	3351,930
26	1985	203815,890	1985	1985	6160,152
27	1986	207002,254	1986	1986	-190,633
28	1987	212143,034	1987	1987	1763,783
29	1988	211234,289	1988	1988	-4265,742
30	1989	219944,610	1989	1989	5333,324

Σχήμα 13. Μεταβλητή noise residuals France

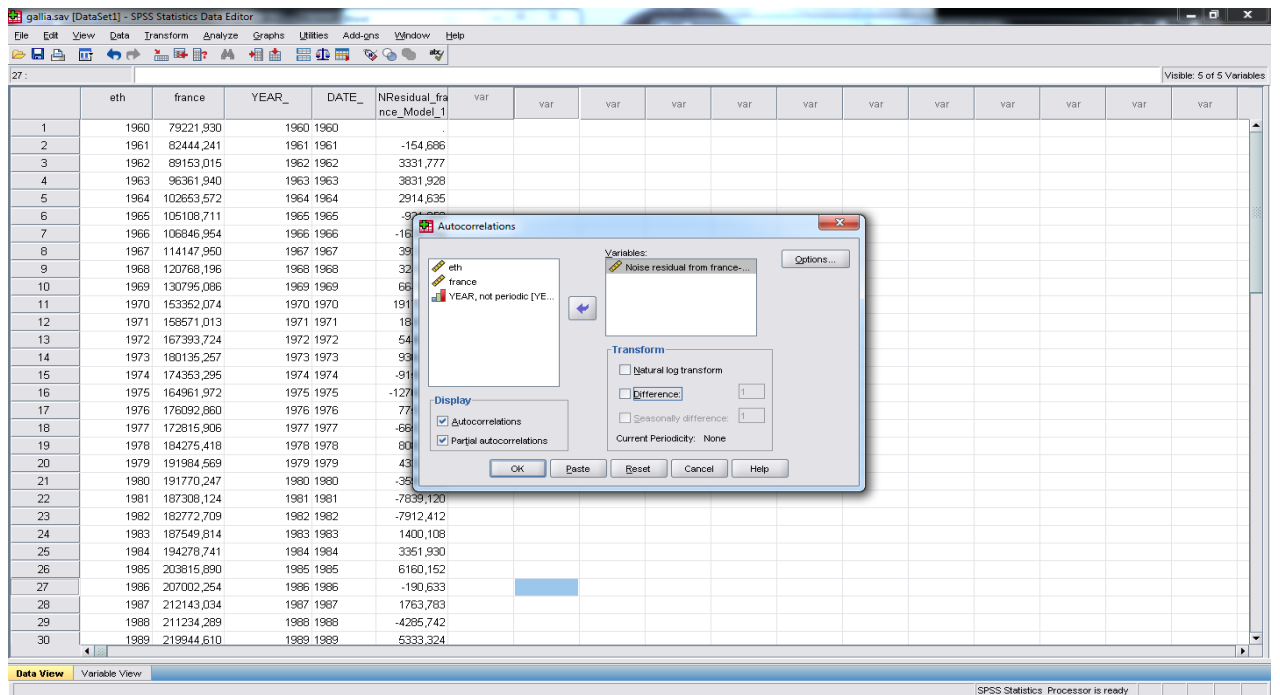
Από το μενού του PASW 18 επιλέγουμε Analyze → Descriptive Statistics → Q-Plots και στο Variable εισάγουμε τη μεταβλητή Noise Residual from france και πατάμε ok.

Σχήμα 14. Διαδικασία για έλεγχο στασιμότητας των Residuals



Σχήμα 15. Διάγραμμα κατανομής των Residuals

Στο διάγραμμα που εξάγεται παρατηρούμε ότι η κατανομή είναι κανονική αφού όλες οι παρατηρήσεις βρίσκονται κοντά στη γραμμή. Επιπλέον θα κάνουμε έλεγχο στασιμότητας και για τα Residuals. Από το μενού του PASW 18 επιλέγουμε Analyze → Forecasting → Autocorrelation και εισάγουμε τη μεταβλητή noise residuals from france και στη συνέχεια πατάμε ok.



Σχήμα 16. Έλεγχος στασιμότητας Residuals

Από το PASW 18 εξάγεται ο παρακάτω πίνακας.

Autocorrelations

Series:Noise residual from france-Model_1

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,052	,136	,148	1	,700
2	,160	,135	1,562	2	,458
3	,049	,133	1,694	3	,638
4	-,163	,132	3,232	4	,520
5	,047	,130	3,361	5	,645
6	,160	,129	4,896	6	,557
7	,065	,128	5,159	7	,641
8	-,022	,126	5,189	8	,737
9	,020	,125	5,214	9	,815
10	-,105	,123	5,937	10	,821
11	-,064	,122	6,216	11	,859
12	,066	,120	6,522	12	,888
13	-,047	,119	6,680	13	,918
14	,012	,117	6,691	14	,946
15	,062	,115	6,984	15	,958
16	-,122	,114	8,141	16	,945

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 3. Πίνακας αυτοσυσχετίσεων για τα residuals

Στη συνέχεια γίνεται έλεγχος στατιστικής σημαντικότητας p-value των εκτιμητών του υποδείγματος για όλα τα lags.

$H_0: \rho=0$

V_s

$H_1: \rho \neq 0$

Με την υπόθεση H_0 ο εκτιμητής μας είναι στατιστικά μη σημαντικός, ενώ με την H_1 ο εκτιμητής μας είναι στατιστικά σημαντικός.

1ο lag

$H_0: \rho_1=0$

στάσιμη)

V_s

$H_1: \rho_1 \neq 0$

2ο lag

$H_0: \rho_2=0$

στάσιμη)

V_s

$\rho_1=0,700 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα

$\rho_2=0,458 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα

$H_1: \rho_2 \neq 0$

3ο lag

$H_0: \rho_3 = 0$
(στάσιμη)

$\rho_3 = 0,638 > \alpha = 0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα

V_s

$H_1: \rho_3 \neq 0$

4ο lag

$H_0: \rho_4 = 0$
(στάσιμη)

$\rho_4 = 0,520 > \alpha = 0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα

V_s

$H_1: \rho_4 \neq 0$

.....

16ο lag

$H_0: \rho_{16} = 0$
(στάσιμη)

$\rho_{16} = 0,945 > \alpha = 0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα

V_s

$H_1: \rho_{16} \neq 0$

Άρα από τον έλεγχο καταλοίπων φαίνεται ότι η διαδικασία είναι στάσιμη καθώς για όλες τις χρονολογικές υστερήσεις τα ρ είναι στάσιμα. Έτσι το υπόδειγμα ARIMA(0,1,0) επιβεβαιώνεται ότι είναι το πλέον κατάλληλο.

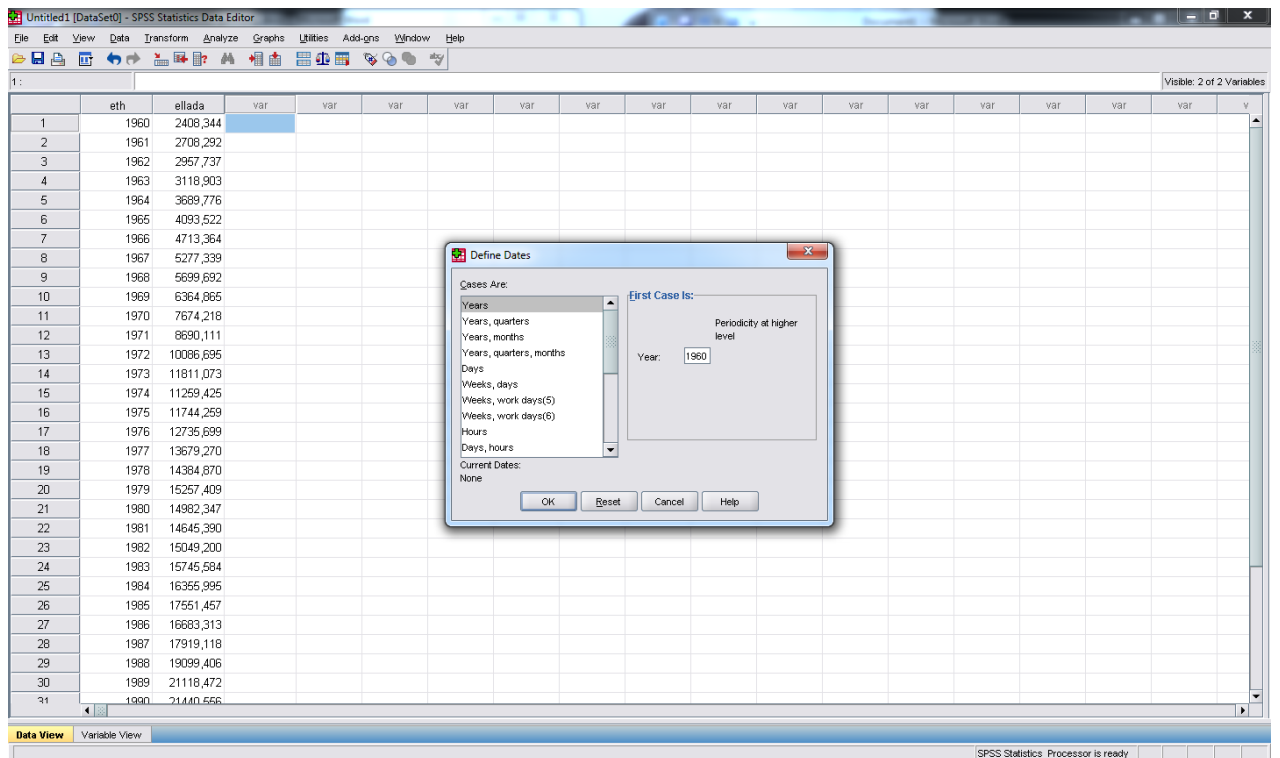
4.4 Ανάλυση των δεδομένων της ενεργειακής κατανάλωσης της Ελλάδας

Στο παράδειγμα που θα αναλύσουμε στην επόμενη ενότητα χρησιμοποιήθηκαν τα δεδομένα της ενεργειακής κατανάλωσης της Ελλάδας. Αρχικά εισάγονται στο PASW 18 οι τιμές της κατανάλωσης ενέργειας της Ελλάδας για τα έτη από το 1960 έως και το 2011.

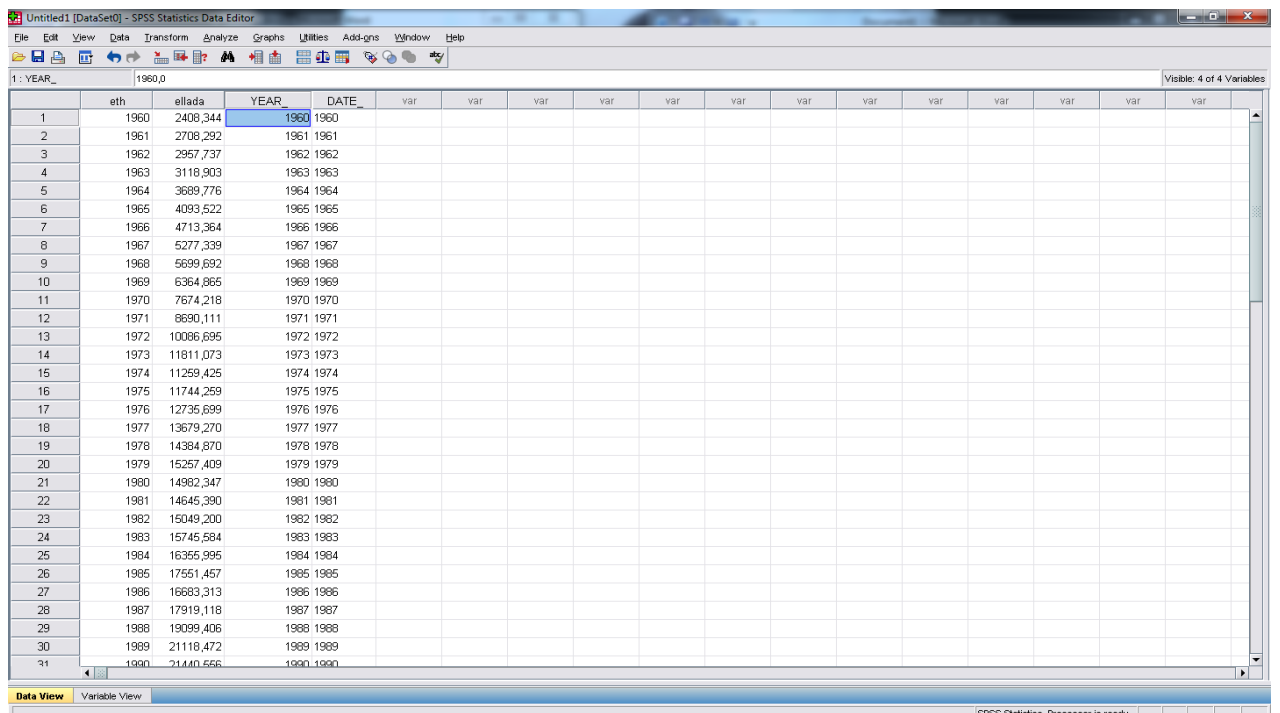
Το πρώτο βήμα είναι να ορίσουμε μέσα από το menu του προγράμματος τις χρονολογίες που μας ενδιαφέρουν προκειμένου να προβούμε στην περαιτέρω επεξεργασία των στοιχείων τους. Έτσι από το menu επιλέγω

Data → Define dates → Years

και στο παράθυρο years συμπληρώνουμε 1960 για να ορίσουμε την ημερομηνία έναρξης των αποτελεσμάτων.



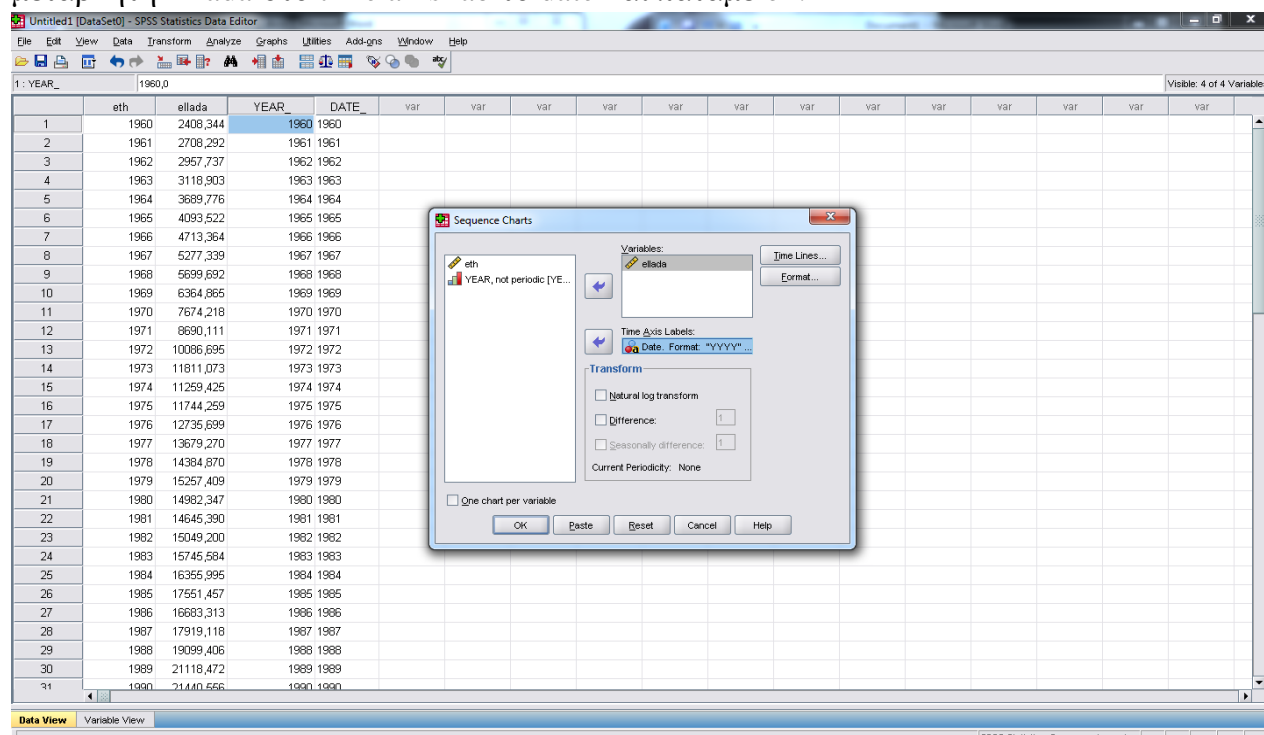
Σχήμα 17. Ορισμός ημερομηνίας



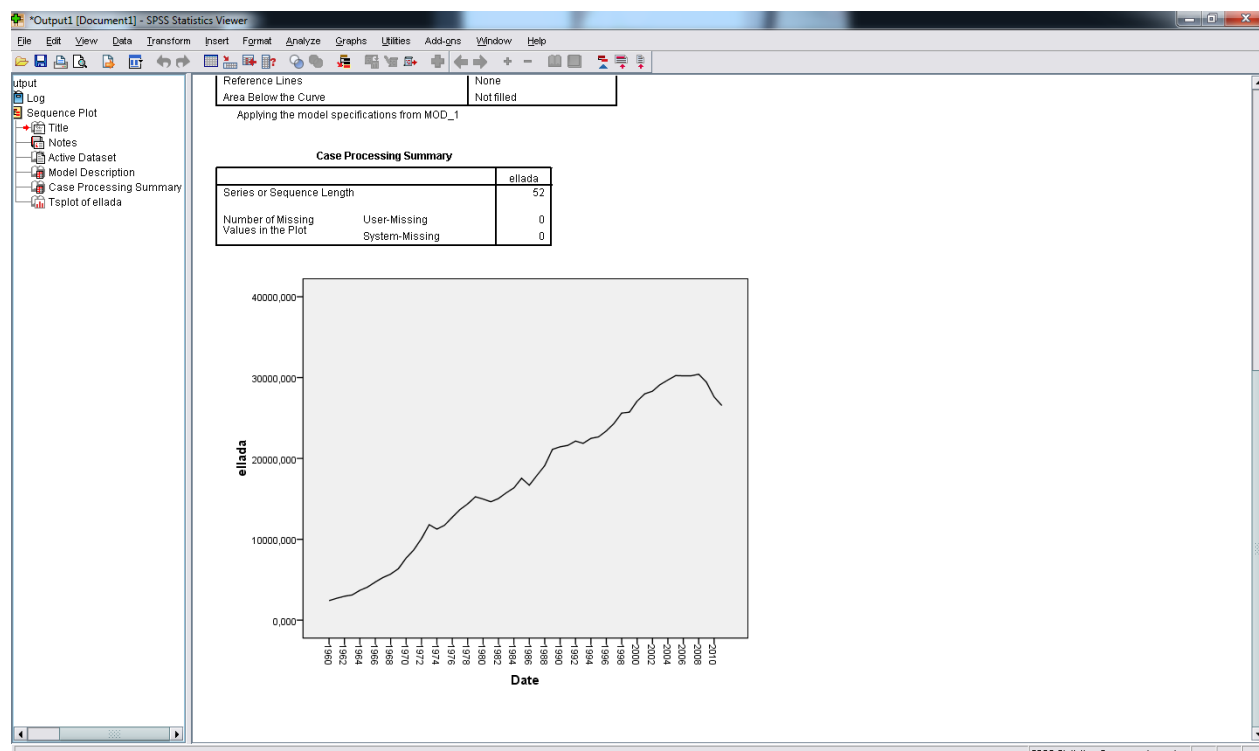
Σχήμα 18. Χρονολογία-Ημερομηνία

Στη συνέχεια θα κάνουμε έλεγχο της στασιμότητας των δεδομένων. Από το μενού επιλέγουμε

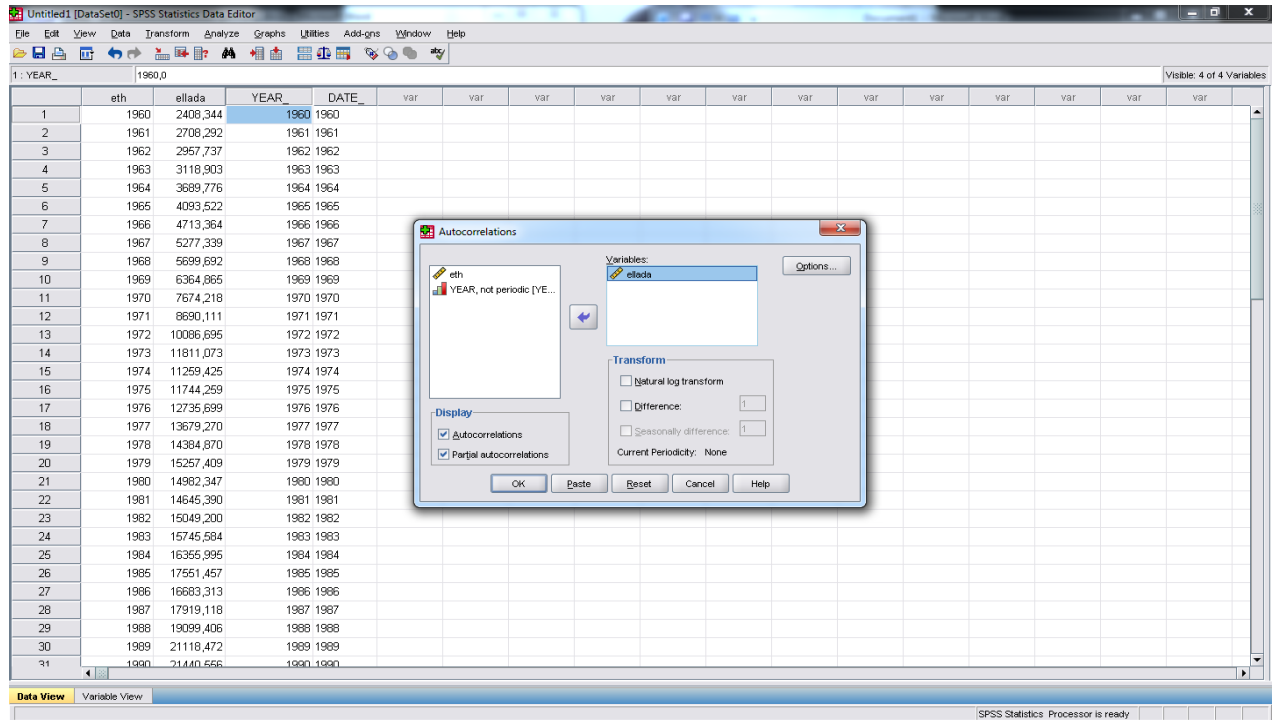
Analyze → Forecasting → Sequences Charts και στην επιλογή Variable βάζουμε τη μεταβλητή Ελλάδα στο time axis lab το date και πατάμε ok.



Σχήμα 19. Έλεγχος στασιμότητας



Σχήμα 20. Διάγραμμα Sequence



Σχήμα 21. Εξαγωγή Διαγράμματος ACF, PACF

Από την παραπάνω διαδικασία προκύπτει ο παρακάτω πίνακας αυτοσυσχετίσεων.

Autocorrelations

Series:ellada

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,958	,135	50,578	1	,000
2	,911	,133	97,145	2	,000
3	,855	,132	139,080	3	,000
4	,795	,131	176,071	4	,000
5	,734	,129	208,255	5	,000
6	,671	,128	235,719	6	,000
7	,606	,127	258,652	7	,000
8	,542	,125	277,406	8	,000
9	,479	,124	292,374	9	,000
10	,417	,122	304,017	10	,000
11	,359	,121	312,868	11	,000
12	,305	,119	319,390	12	,000
13	,255	,118	324,081	13	,000
14	,210	,116	327,340	14	,000
15	,164	,115	329,380	15	,000
16	,120	,113	330,494	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 4. Πίνακας αυτοσυσχετίσεων

Στον πίνακα αυτό εμφανίζονται οι παρακάτω όροι.

- ❖ Χρονολογικές υστερήσεις (lag)
- ❖ Τιμή του συντελεστή αυτοσυσχετίσης (Autocorrelation)
- ❖ Τυπικό σφάλμα (Std. error)
- ❖ Τιμή της στατιστικής συνάρτησης Box-Jung (Value)
- ❖ Βαθμοί ελευθερίας (df)
- ❖ Τιμή πιθανότητας στη στατιστική συνάρτηση (Sig)

Από τον πίνακα προκύπτει ότι για όλες τις χρονολογικές υστερήσεις υπάρχει αυτοσυσχετίση γιατί όλα τα p είναι μηδέν με αποτέλεσμα η χρονολογική σειρά να μην είναι στάσιμη με διαφορές. Έτσι θα κάνουμε δοκιμή με τη μέθοδο των πρώτων διαφορών. Από το μενού επιλέγουμε Analyze→Forecasting→Autocorrelation και επιλέγουμε Difference και πληκτρολογούμε 1. Έτσι παράγονται τα παρακάτω αποτελέσματα.

Autocorrelations

Series:ellada

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,345	,136	6,449	1	,011
2	,177	,135	8,184	2	,017
3	,038	,133	8,267	3	,041
4	,044	,132	8,378	4	,079
5	,025	,130	8,413	5	,135
6	,046	,129	8,539	6	,201
7	-,050	,128	8,692	7	,276
8	-,187	,126	10,886	8	,208
9	-,105	,125	11,591	9	,237
10	-,172	,123	13,533	10	,195
11	-,025	,122	13,575	11	,257
12	,082	,120	14,036	12	,298
13	-,027	,119	14,089	13	,368
14	-,093	,117	14,720	14	,398
15	,005	,115	14,721	15	,472
16	,128	,114	15,980	16	,454

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 5. Πίνακας αυτοσυσχετίσεων με τη μέθοδο των πρώτων διαφορών

Εφόσον οι τιμές δεν είναι μηδενικές κάνουμε έλεγχο στασιμότητας.

$H_0: \rho=0$

V_s

$H_1: \rho \neq 0$

Με την υπόθεση H_0 δεν υπάρχει αυτοσυσχέτιση, άρα είναι στάσιμη ενώ με την H_1 υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη. Ο έλεγχος αυτός πραγματοποιείται για όλα τα lags. Οπότε:

1ο lag:

$H_0: \rho_1=0$ $\rho_1=0,011 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

V_s

$H_1: \rho_1 \neq 0$

2o lag:

$H_0: \rho_2=0$ $\rho_2=0,017 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

V_s

$H_1: \rho_2 \neq 0$

3o lag:

$H_0: \rho_3=0$ $\rho_3=0,041 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

V_s

$H_1: \rho_3 \neq 0$

4o lag:

$H_0: \rho_4=0$ $\rho_4=0,079 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_4 \neq 0$

.....

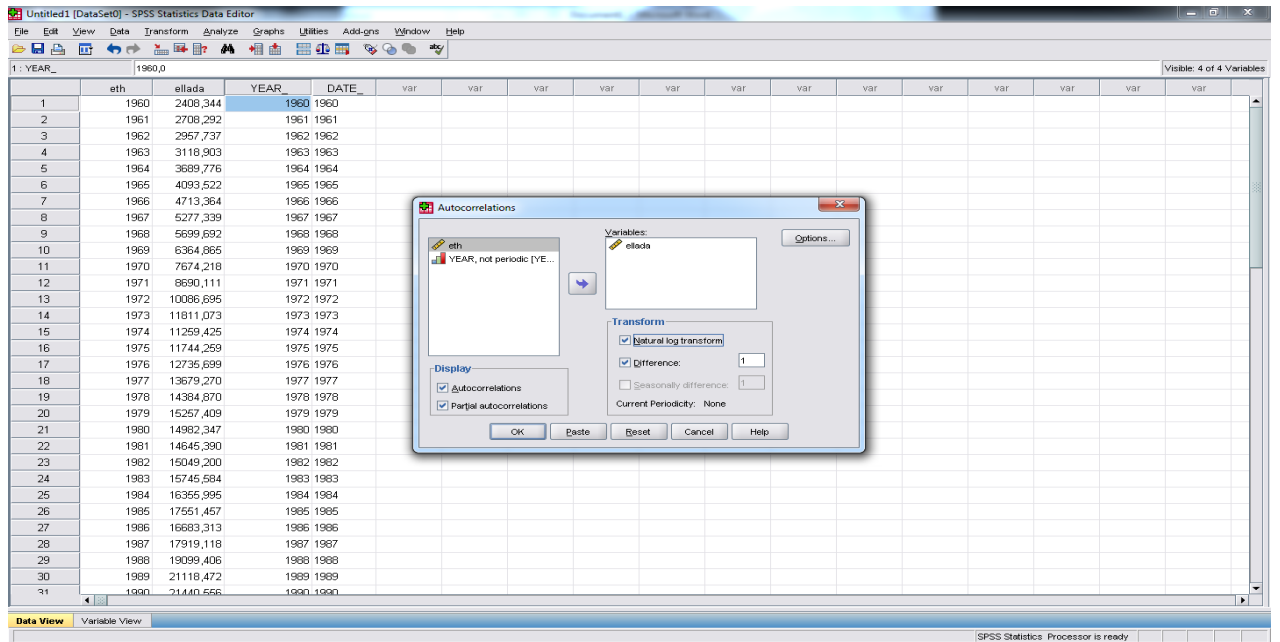
16o lag:

$H_0: \rho_{16}=0$ $\rho_{16}=0,368 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_{16} \neq 0$

Βάσει του ελέγχου κανονικότητας προκύπτει ότι η διαδικασία είναι δεν στάσιμη με πρώτες διαφορές. Οπότε παίρνουμε πρώτες διαφορές και λογάριθμο μαζί.



Σχήμα 22. Μέθοδος πρώτων διαφορών και λογάριθμος μαζί
Autocorrelations

Series:ellada

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,530	,136	15,189	1	,000
2	,433	,135	25,535	2	,000
3	,403	,133	34,658	3	,000
4	,311	,132	40,212	4	,000
5	,342	,130	47,091	5	,000
6	,420	,129	57,709	6	,000
7	,246	,128	61,442	7	,000
8	,113	,126	62,239	8	,000
9	,148	,125	63,643	9	,000
10	-,007	,123	63,646	10	,000
11	,088	,122	64,173	11	,000
12	,167	,120	66,109	12	,000
13	-,020	,119	66,136	13	,000
14	-,044	,117	66,276	14	,000
15	,011	,115	66,285	15	,000
16	,014	,114	66,300	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 6. Πίνακας αυτοσυσχετίσεων

Παρατηρούμε ότι τα lags είναι όλα μηδέν οπότε δεν είναι στάσιμη. Έτσι θα κάνουμε δοκιμή με τη μέθοδο των δεύτερων διαφορών.

The screenshot shows the SPSS Statistics Data Editor interface. The main window displays a data table with columns: eth, ellada, YEAR, DATE, and several empty columns labeled 'var'. The data rows range from 1960 to 1990. Overlaid on this is the 'Autocorrelations' dialog box. In the 'Variables' list, 'eth' and 'YEAR, not periodic [YE...]' are selected. The 'Transform' section has 'Difference' checked with a lag of 2, and 'Seasonally difference' is unchecked. The 'Display' section has 'Autocorrelations' and 'Partial autocorrelations' checked. The 'Current Periodicity' is set to 'None'. Buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help' are visible at the bottom of the dialog.

	eth	ellada	YEAR	DATE	var	var	var	var	var	var	var	var	var	var	var	var
1	1960	2408,344	1960	1960												
2	1961	2708,292	1961	1961												
3	1962	2957,737	1962	1962												
4	1963	3118,903	1963	1963												
5	1964	3689,776	1964	1964												
6	1965	4093,522	1965	1965												
7	1966	4713,364	1966	1966												
8	1967	5277,339	1967	1967												
9	1968	5699,692	1968	1968												
10	1969	6364,865	1969	1969												
11	1970	7674,218	1970	1970												
12	1971	8690,111	1971	1971												
13	1972	10086,695	1972	1972												
14	1973	11811,073	1973	1973												
15	1974	11259,425	1974	1974												
16	1975	11744,259	1975	1975												
17	1976	12735,699	1976	1976												
18	1977	13679,270	1977	1977												
19	1978	14304,670	1978	1978												
20	1979	15257,409	1979	1979												
21	1980	14982,347	1980	1980												
22	1981	14645,390	1981	1981												
23	1982	15049,200	1982	1982												
24	1983	15745,584	1983	1983												
25	1984	16355,395	1984	1984												
26	1985	17551,457	1985	1985												
27	1986	16683,313	1986	1986												
28	1987	17919,118	1987	1987												
29	1988	19099,406	1988	1988												
30	1989	21118,472	1989	1989												
31	1990	21440,566	1990	1990												

Σχήμα 23. Δεύτερες διαφορές

Autocorrelations

Series:ellada

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,440	,137	10,260	1	,001
2	,016	,136	10,274	2	,006
3	-,059	,134	10,464	3	,015
4	,009	,133	10,468	4	,033
5	-,034	,132	10,533	5	,061
6	,124	,130	11,437	6	,076
7	,034	,129	11,509	7	,118
8	-,167	,127	13,228	8	,104
9	,100	,126	13,867	9	,127
10	-,148	,124	15,295	10	,122
11	,057	,122	15,511	11	,160
12	,115	,121	16,419	12	,173
13	,012	,119	16,429	13	,227
14	-,146	,118	17,977	14	,208
15	-,029	,116	18,039	15	,261
16	,125	,114	19,238	16	,257

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 7. Πίνακας αυτοσυσχετίσεων

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας

$H_0: \rho=0$

V_s

$H_1: \rho \neq 0$

Με την υπόθεση H_0 δεν υπάρχει αυτοσυσχέτιση, άρα είναι στάσιμη ενώ με την H_1 υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη.

1ο lag:

$H_0: \rho_1=0$ $\rho_1=0,001 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

V_s

$H_1: \rho_1 \neq 0$

2ο lag:

$H_0: \rho_2=0$ $p_2=0,006 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

V_s

$H_1: \rho_2 \neq 0$

3ο lag:

$H_0: \rho_3=0$ $p_3=0,015 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

V_s

$H_1: \rho_3 \neq 0$

4ο lag:

$H_0: \rho_4=0$ $p_4=0,033 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

V_s

$H_1: \rho_4 \neq 0$

.....

16ο lag:

$H_0: \rho_{16}=0$ $p_{16}=0,257 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_{16} \neq 0$

Βάσει του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία δεν είναι στάσιμη με δεύτερες διαφορές. Οπότε παίρνουμε δεύτερες διαφορές και λογάριθμο μαζί.

SPSS Statistics Data Editor window showing a data table and the Autocorrelations dialog box.

	eth	ellada	YEAR_	DATE_	var	var	var	var	var	var	var	var	var	var	var	var	var	var
1	1960	2408,344	1960	1960														
2	1961	2708,292	1961	1961														
3	1962	2957,737	1962	1962														
4	1963	3118,903	1963	1963														
5	1964	3689,776	1964	1964														
6	1965	4093,522	1965	1965														
7	1966	4713,364	1966	1966														
8	1967	5277,339	1967	1967														
9	1968	5699,692	1968	1968														
10	1969	6364,865	1969	1969														
11	1970	7674,218	1970	1970														
12	1971	8690,111	1971	1971														
13	1972	10086,695	1972	1972														
14	1973	11811,073	1973	1973														
15	1974	11259,425	1974	1974														
16	1975	11744,259	1975	1975														
17	1976	12736,699	1976	1976														
18	1977	13679,270	1977	1977														
19	1978	14384,870	1978	1978														
20	1979	15257,409	1979	1979														
21	1980	14982,347	1980	1980														
22	1981	14645,390	1981	1981														
23	1982	15049,200	1982	1982														
24	1983	15745,584	1983	1983														
25	1984	16355,995	1984	1984														
26	1985	17551,457	1985	1985														
27	1986	16683,313	1986	1986														
28	1987	17919,118	1987	1987														
29	1988	19099,406	1988	1988														
30	1989	21118,472	1989	1989														
31	1990	21440,556	1990	1990														

The Autocorrelations dialog box is open, showing the following settings:

- Variables: ellada
- Transform:
 - Natural log transform
 - Difference: 2
 - Seasonally difference: 1
 - Current Periodicity: None
- Display:
 - Autocorrelations
 - Partial autocorrelations

Σχήμα 24. Δεύτερες διαφορές και λογάριθμος

Autocorrelations

Series:ellada

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,435	,137	10,040	1	,002
2	-,044	,136	10,145	2	,006
3	,036	,134	10,215	3	,017
4	-,118	,133	11,008	4	,026
5	-,076	,132	11,338	5	,045
6	,314	,130	17,150	6	,009
7	-,030	,129	17,205	7	,016
8	-,208	,127	19,889	8	,011
9	,168	,126	21,688	9	,010
10	-,248	,124	25,694	10	,004
11	,015	,122	25,709	11	,007
12	,271	,121	30,737	12	,002
13	-,062	,119	31,003	13	,003
14	-,142	,118	32,461	14	,003
15	,035	,116	32,550	15	,005
16	-,017	,114	32,572	16	,008

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 8. Πίνακας αυτοσυσχετίσεων

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας

$H_0: \rho=0$

V_s

$H_1: \rho \neq 0$

Με την υπόθεση H_0 δεν υπάρχει αυτοσυσχέτιση, άρα είναι στάσιμη ενώ με την H_1 υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη.

1ο lag:

$H_0: \rho_1=0$ $\rho_1=0,002 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

V_s

$H_1: \rho_1 \neq 0$

2o lag:

$H_0: \rho_2=0$ $p_2=0,006 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

V_s

$H_1: \rho_2 \neq 0$

3o lag:

$H_0: \rho_3=0$ $p_3=0,017 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

V_s

$H_1: \rho_3 \neq 0$

4o lag:

$H_0: \rho_4=0$ $p_4=0,026 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

V_s

$H_1: \rho_4 \neq 0$

.....

16o lag:

$H_0: \rho_{16}=0$ $p_{16}=0,008 < \alpha=0,05$ Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας

Παρατηρούμε ότι συνεχίζει να μην είναι στάσιμη ακόμη και με δεύτερες διαφορές και λογάριθμο μαζί.

Συμπέρασμα λοιπόν ότι η χρονοσειρά αυτή δεν είναι στάσιμη.

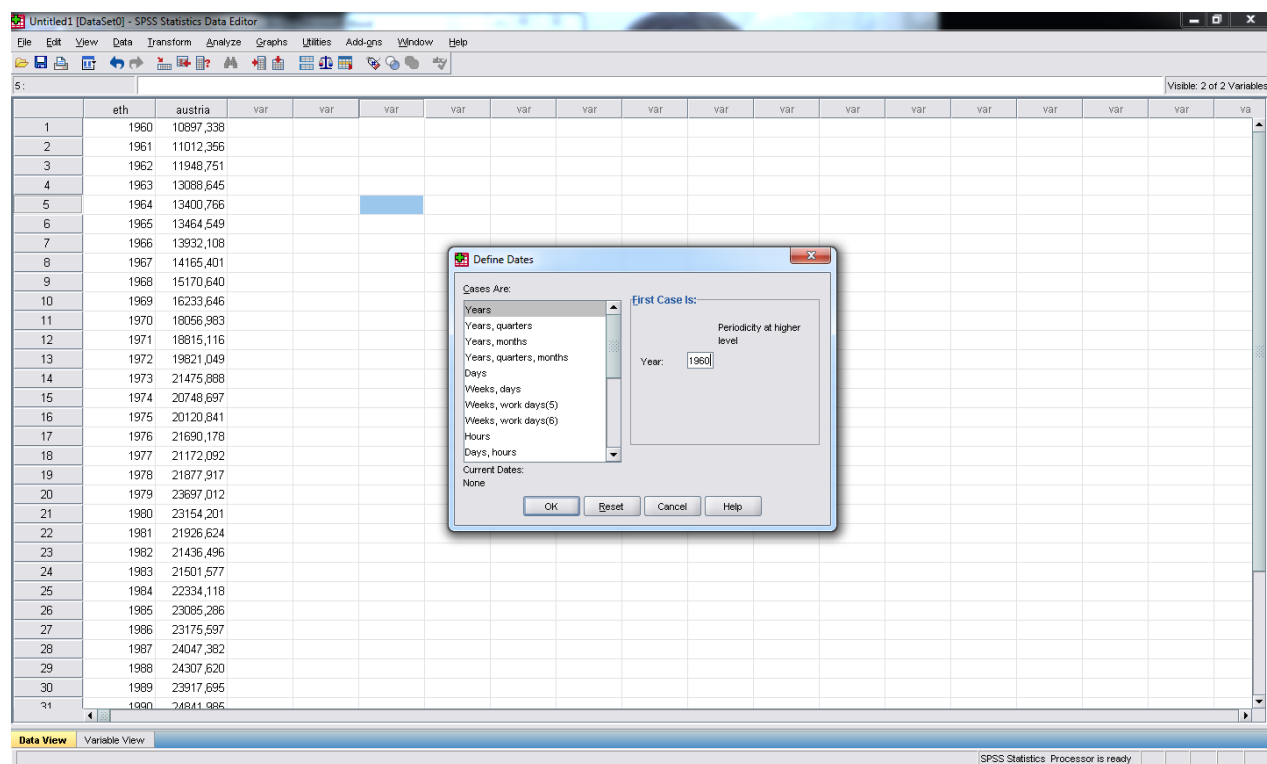
4.5 Ανάλυση των δεδομένων της ενεργειακής κατανάλωσης της Αυστρίας

Στο παράδειγμα που θα αναλύσουμε στην επόμενη ενότητα χρησιμοποιήθηκαν τα δεδομένα της ενεργειακής κατανάλωσης της Αυστρίας. Αρχικά εισάγονται στο PASW 18 οι τιμές της κατανάλωσης ενέργειας της Αυστρίας για τα έτη από το 1960 έως και το 2011.

Το πρώτο βήμα είναι να ορίσουμε μέσα από το menu του προγράμματος τις χρονολογίες που μας ενδιαφέρουν προκειμένου να προβούμε στην περαιτέρω επεξεργασία των στοιχείων τους. Έτσι από το menu επιλέγω

Data → Define dates → Years

και στο παράθυρο years συμπληρώνουμε 1960 για να ορίσουμε την ημερομηνία έναρξης των αποτελεσμάτων.



Σχήμα 25. Ορισμός ημερομηνίας

	eth	austria	YEAR	DATE	var	var	var	var	var	var	var	var	var	var	var	var	var	var
1	1960	10897,338	1960	1960														
2	1961	11012,356	1961	1961														
3	1962	11948,751	1962	1962														
4	1963	13088,645	1963	1963														
5	1964	13400,766	1964	1964														
6	1965	13464,549	1965	1965														
7	1966	13932,108	1966	1966														
8	1967	14165,401	1967	1967														
9	1968	15170,640	1968	1968														
10	1969	16233,646	1969	1969														
11	1970	18056,983	1970	1970														
12	1971	18815,116	1971	1971														
13	1972	19821,049	1972	1972														
14	1973	21475,888	1973	1973														
15	1974	20748,697	1974	1974														
16	1975	20120,841	1975	1975														
17	1976	21690,178	1976	1976														
18	1977	21172,092	1977	1977														
19	1978	21877,917	1978	1978														
20	1979	23697,012	1979	1979														
21	1980	23154,201	1980	1980														
22	1981	21926,624	1981	1981														
23	1982	21436,496	1982	1982														
24	1983	21501,577	1983	1983														
25	1984	22334,118	1984	1984														
26	1985	23085,286	1985	1985														
27	1986	23175,597	1986	1986														
28	1987	24047,382	1987	1987														
29	1988	24307,620	1988	1988														
30	1989	23917,695	1989	1989														
31	1990	24841,985	1990	1990														

Σχήμα 26. Χρονολογία-Ημερομηνία

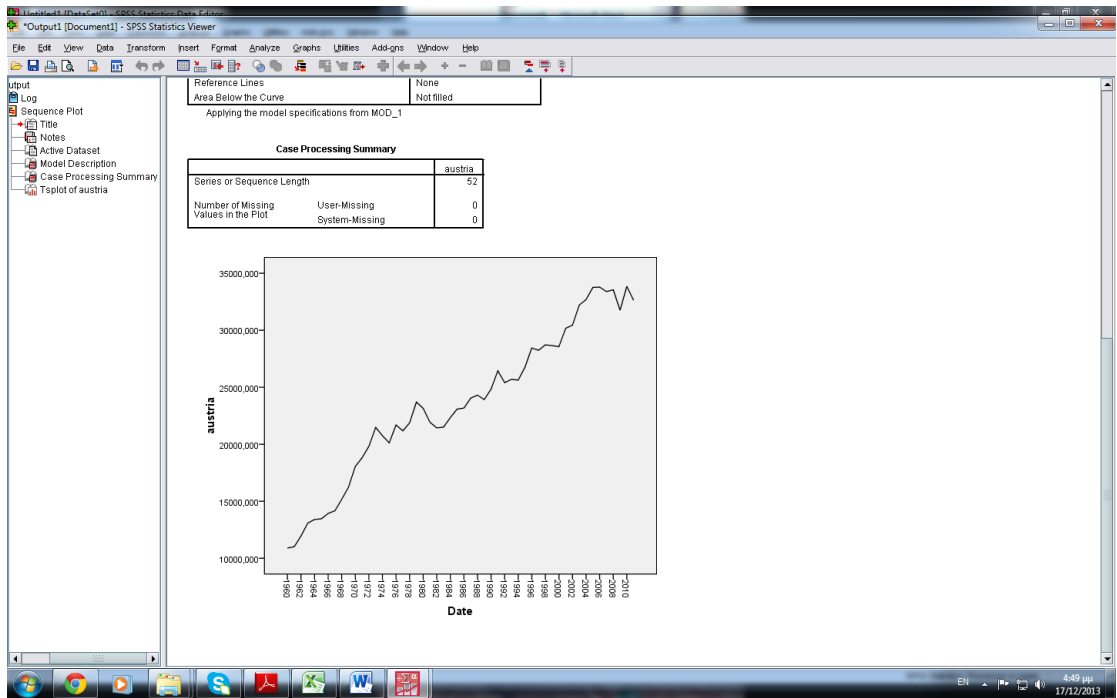
Στη συνέχεια θα κάνουμε έλεγχο της στασιμότητας των δεδομένων. Από το μενού επιλέγουμε

Analyze → Forecasting → Sequences Charts και στην επιλογή Variable βάζουμε τη μεταβλητή Austria στο time axis lab to date και πατάμε ok.

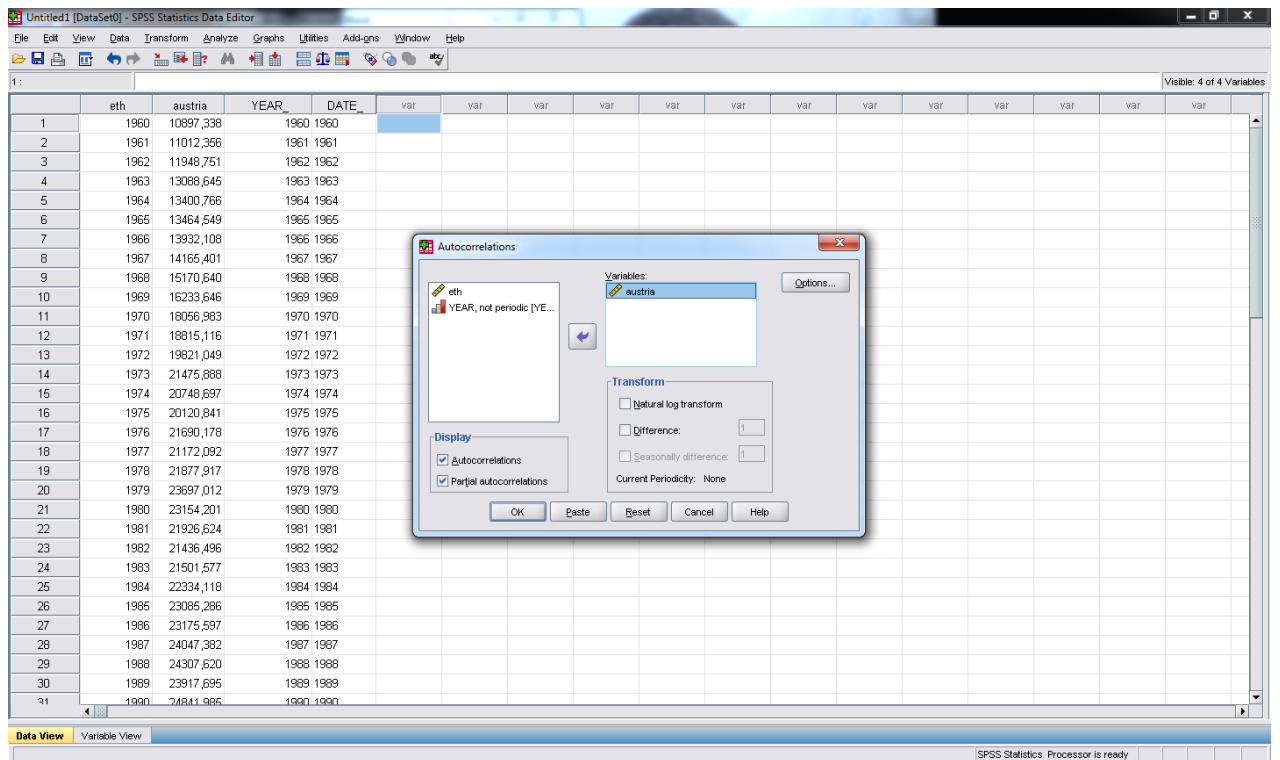
The 'Sequence Charts' dialog box is shown with the following settings:

- Variables:** austria
- Time Axis Labels:** YEAR, not periodic [YE...]
- Time Axis Labels:** Date, Format: "YYYY"
- Transform:**
 - Natural log transform
 - Difference
 - Seasonally difference
- Current Periodicity:** None
- One chart per variable

Σχήμα 27. Έλεγχος στασιμότητας



Σχήμα 28. Διαγράμμα Sequence



Σχήμα 29. Εξαγωγή Διαγράμματων ACF, PACF

Από την παραπάνω διαδικασία προκύπτει ο παρακάτω πίνακας αυτοσυσχετίσεων.

Autocorrelations

Series:austria

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,935	,135	48,162	1	,000
2	,864	,133	90,119	2	,000
3	,803	,132	127,044	3	,000
4	,735	,131	158,688	4	,000
5	,670	,129	185,539	5	,000
6	,599	,128	207,472	6	,000
7	,527	,127	224,785	7	,000
8	,453	,125	237,909	8	,000
9	,385	,124	247,583	9	,000
10	,326	,122	254,673	10	,000
11	,276	,121	259,908	11	,000
12	,236	,119	263,825	12	,000
13	,198	,118	266,637	13	,000
14	,170	,116	268,765	14	,000
15	,137	,115	270,179	15	,000
16	,096	,113	270,898	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 9. Πίνακας αυτοσυσχετίσεων

Στον πίνακα αυτό εμφανίζονται οι παρακάτω όροι.

- ❖ Χρονολογικές υστερήσεις (lag)
- ❖ Τιμή του συντελεστή αυτοσυσχετίσης (Autocorrelation)
- ❖ Τυπικό σφάλμα (Std. error)
- ❖ Τιμή της στατιστικής συνάρτησης Box-Jung (Value)
- ❖ Βαθμοί ελευθερίας (df)
- ❖ Τιμή πιθανότητας στη στατιστική συνάρτηση (Sig)

Από τον πίνακα προκύπτει ότι για όλες τις χρονολογικές υστερήσεις υπάρχει αυτοσυσχετίση γιατί όλα τα p είναι μηδέν με αποτέλεσμα η χρονολογική σειρά να μην είναι στάσιμη με διαφορές. Έτσι θα κάνουμε δοκιμή με τη μέθοδο των πρώτων διαφορών. Από το μενού επιλέγουμε Analyze→Forecasting→Autocorrelation και επιλέγουμε Difference και πληκτρολογούμε 1.

Έτσι παράγονται τα παρακάτω αποτελέσματα.

Autocorrelations

Series:austria

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,138	,136	1,023	1	,312
2	,025	,135	1,056	2	,590
3	,041	,133	1,151	3	,765
4	-,219	,132	3,911	4	,418
5	,026	,130	3,952	5	,556
6	,009	,129	3,957	6	,682
7	,119	,128	4,828	7	,681
8	-,055	,126	5,018	8	,756
9	,020	,125	5,043	9	,831
10	-,105	,123	5,776	10	,834
11	-,046	,122	5,921	11	,879
12	,051	,120	6,101	12	,911
13	-,226	,119	9,722	13	,716
14	,029	,117	9,784	14	,778
15	,013	,115	9,797	15	,832
16	-,087	,114	10,386	16	,846

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 10. Πίνακας αυτοσυσχετίσεων με τη μέθοδο των πρώτων διαφορών

Εφόσον οι τιμές δεν είναι μηδενικές κάνουμε έλεγχο στασιμότητας.

$H_0: \rho=0$

V_s

$H_1: \rho \neq 0$

Με την υπόθεση H_0 δεν υπάρχει αυτοσυσχέτιση, άρα είναι στάσιμη ενώ με την H_1 υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη. Ο έλεγχος αυτός πραγματοποιείται για όλα τα lags. Οπότε:

1ο lag:

$H_0: \rho_1=0$ $\rho_1=0,312 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_1 \neq 0$

2ο lag:

$H_0: \rho_2=0$ $\rho_2=0,590 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_2 \neq 0$

3ο lag:

$H_0: \rho_3=0$ $\rho_3=0,765 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_3 \neq 0$

4ο lag:

$H_0: \rho_4=0$ $\rho_4=0,418 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_4 \neq 0$

.....

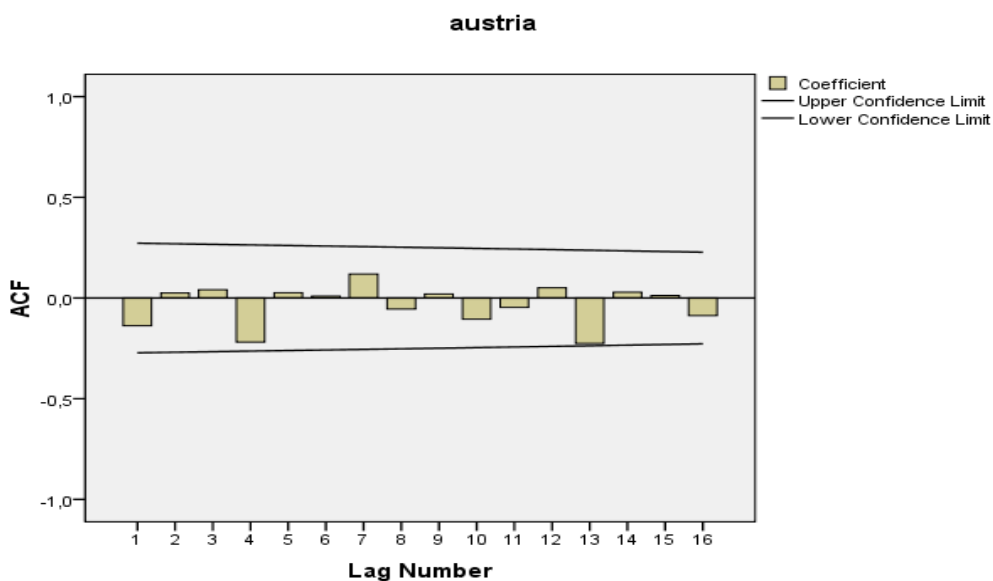
16ο lag:

$H_0: \rho_{16}=0$ $\rho_{16}=0,846 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

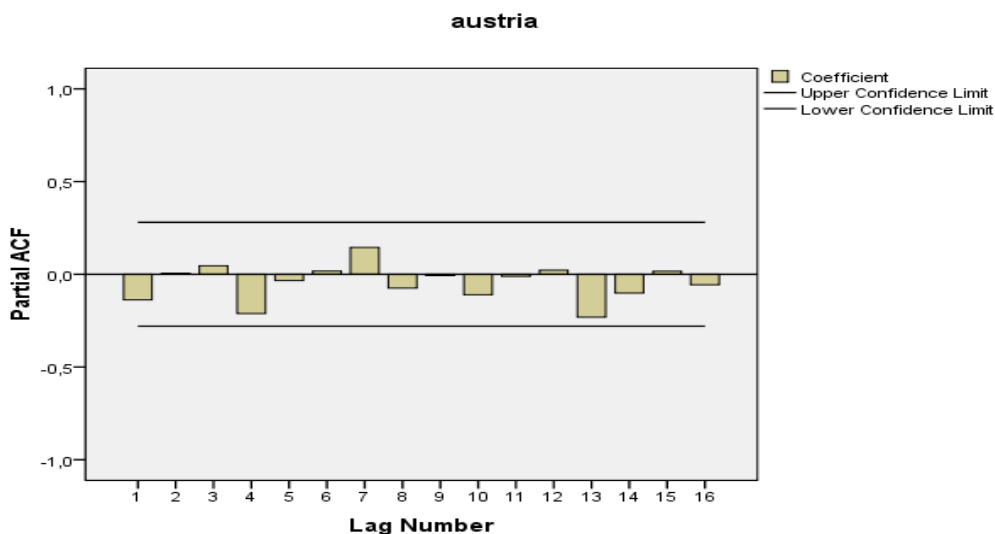
V_s

$H_1: \rho_{16} \neq 0$

Βάσει του ελέγχου κανονικότητας προκύπτει ότι η διαδικασία είναι στάσιμη με πρώτες διαφορές. Στη συνέχεια, από τα διαγράμματα ACF και PACF που εξάγονται από το πρόγραμμα εντοπίζονται τα q και p αντίστοιχα.



Σχήμα 30. Διάγραμμα ACF



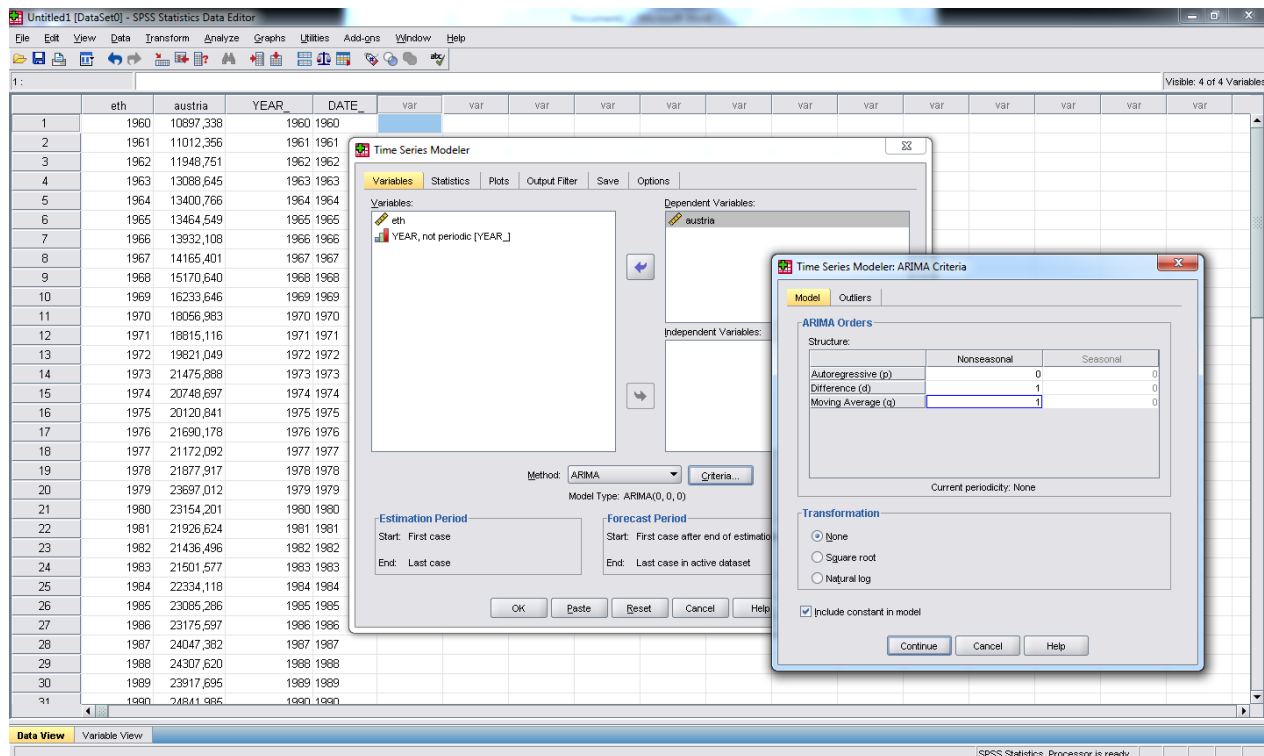
Σχήμα 31. Διάγραμμα PACF

Από τα διαγράμματα ACF και PACF προκύπτει $q=0$ και $p=0$. Το υπόδειγμα είναι της μορφής $ARIMA(p,d,q)$ - $ARIMA(0,1,0)$, ($d=1$ λόγω πρώτων διαφορών). Η θεωρητική μορφή του υποδείγματος είναι :

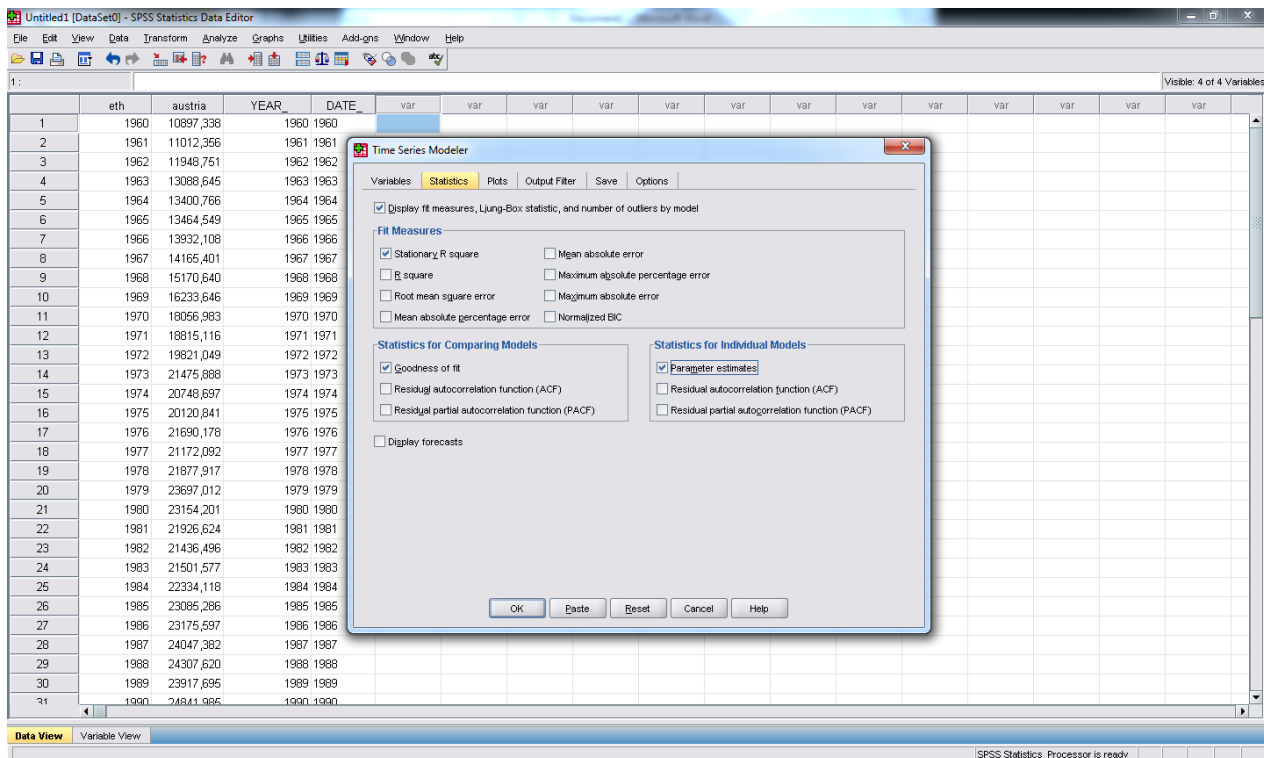
$$y_t = a_0 + a_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Έπειτα επιστρέφοντας στο PASW 18 ακολουθούμε τις παρακάτω διαδικασίες. Από το μενού επιλέγουμε Analyze → Forecasting → Create models. Στο Dependent Variable εισάγουμε τη μεταβλητή Austria και στο method επιλέγουμε ARIMA. Επιλέγοντας το criteria εμφανίζεται ο πίνακας ARIMA Criteria και στο nonseasonal εισάγονται οι τιμές των p,d,q που έχουν εντοπιστεί δηλαδή 0,1 και 1

αντίστοιχα. Πατώντας το Continue και πηγαίνοντας στο Statistics επιλέγουμε Parameter Estimates και στη συνέχεια πατάμε ok.

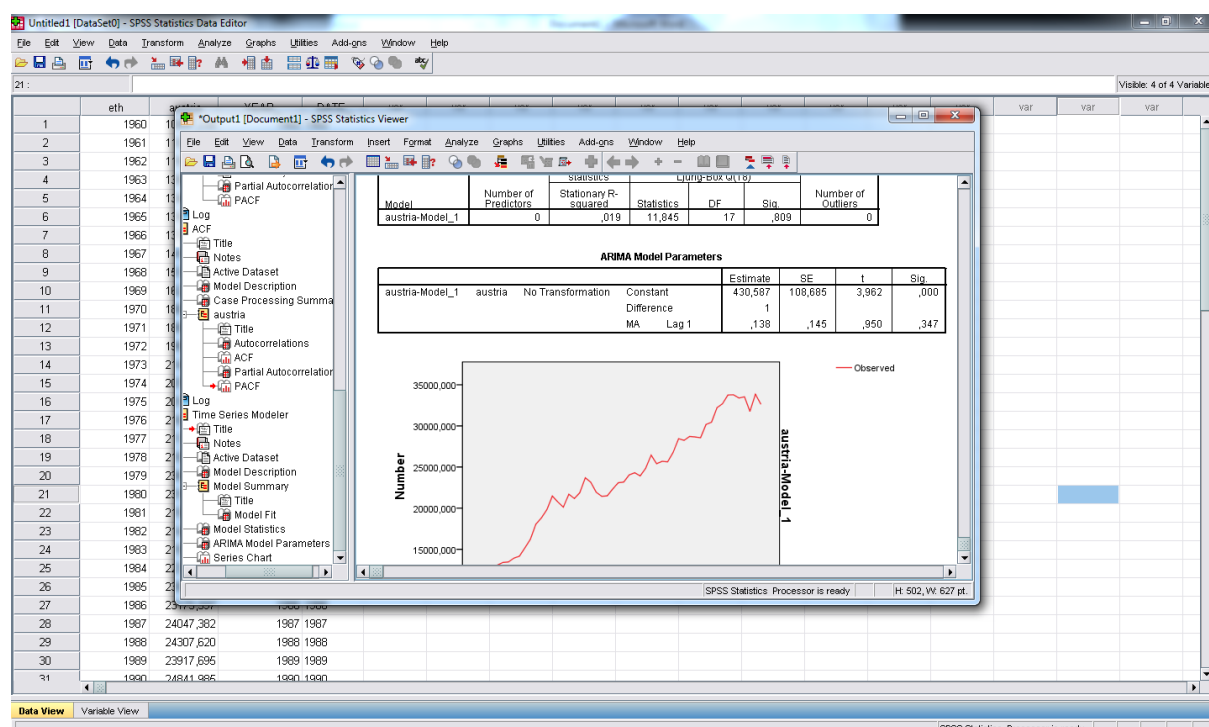


Σχήμα 32. Καθορισμός υποδείγματος ARIMA (0,1,1)



Σχήμα 33. Διαδικασία για έλεγχο στατιστικής σημαντικότητας

Στη συνέχεια εξάγεται ο πίνακας ARIMA Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί κάνοντας έλεγχο της στατιστικής σημαντικότητας για το υπόδειγμα ARIMA (0,1,1).



Σχήμα 34. Πίνακας ARIMA Models Parameters

Στη συνέχεια γίνεται έλεγχος στατιστικής σημαντικότητας p-value των εκτιμητών του υποδείματος.

1ο:

$H_0: \rho_0=0$ $\rho_0=0 < \alpha=0,05$ ο εκτιμητής είναι στατιστικά σημαντικός

V_s

$H_1: \rho_0 \neq 0$

2ο:

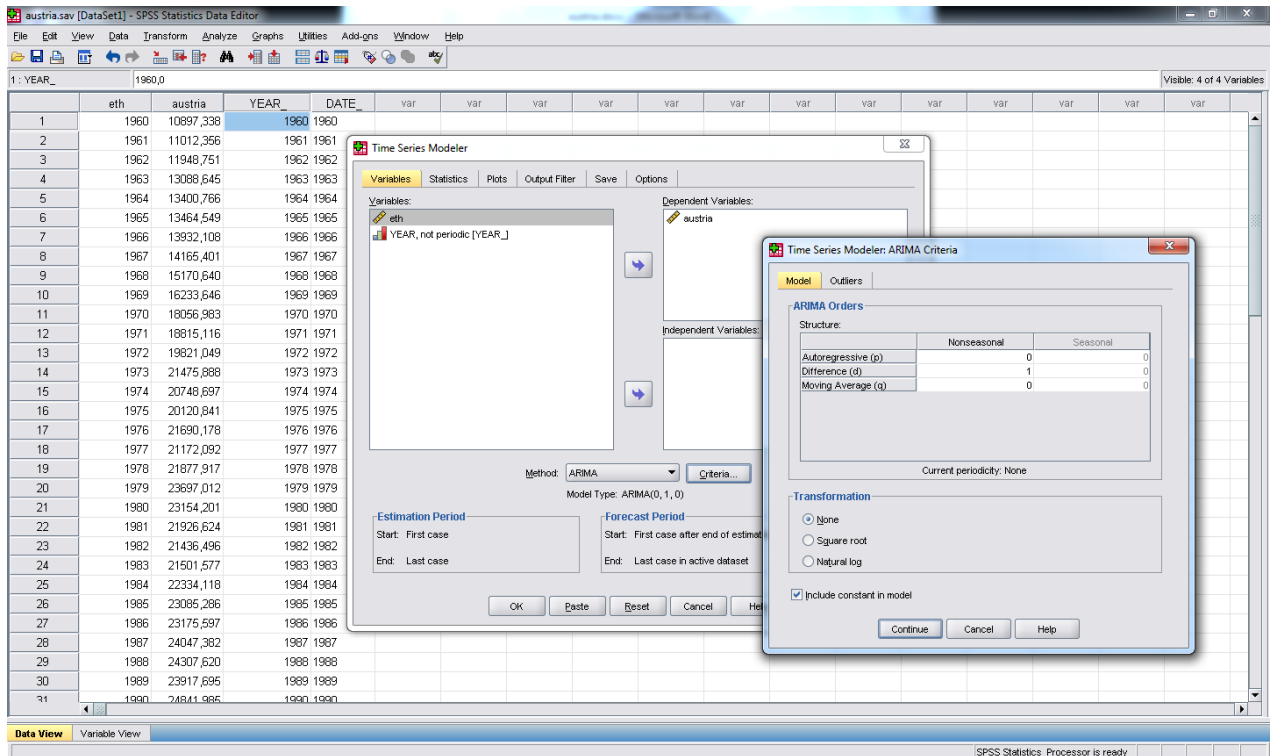
$H_0: \rho_1=0$ $\rho_1=0,347 > \alpha=0,05$ ο εκτιμητής είναι στατιστικά μη σημαντικός

V_s

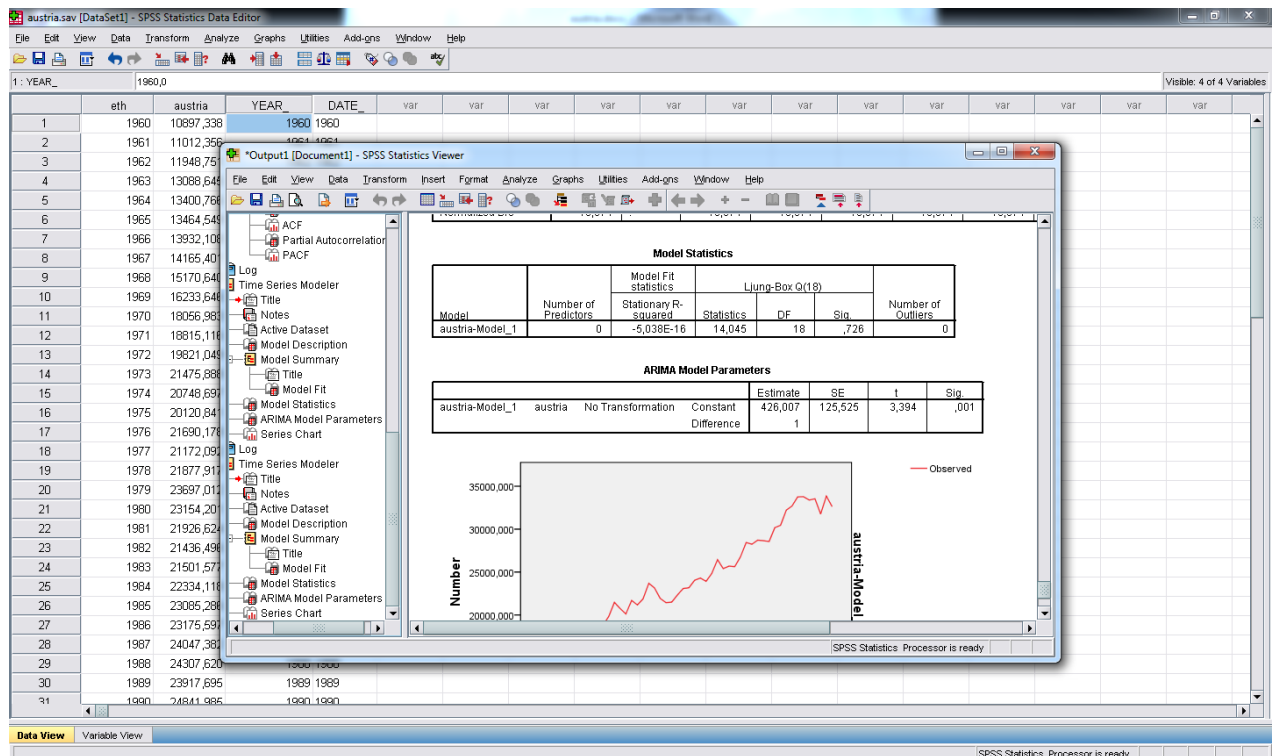
$H_1: \rho_1 \neq 0$

Αφού βρέθηκαν οι εκτιμητές μας στατιστικά μη σημαντικοί αφαιρώ τις στατιστικά μη σημαντικές παραμέτρους, μια μια την φορά, ξεκινώντας από αυτή με το μεγαλύτερο p-

value. Δηλαδή ξανα προσαρμόζω το μοντέλο βάζοντας μηδέν στην παράμετρο που πρέπει (p ή d ή q).



Σχήμα 35. Αναπροσαρμογή μοντέλου με $q=0$



Σχήμα 36. Πίνακας ARIMA model parameters

Στη συνέχεια γίνεται έλεγχος στατιστικής σημαντικότητας p-value των εκτιμητών του υποδείγματος για όλα τα lags.

$H_0: \rho=0$

V_s

$H_1: \rho \neq 0$

Με την υπόθεση H_0 ο εκτιμητής μας είναι στατιστικά μη σημαντικός, ενώ με την H_1 ο εκτιμητής μας είναι στατιστικά σημαντικός.

1ο lag:

$H_0: \rho_0=0$ $\rho_0=0,001 < \alpha=0,05$ ο εκτιμητής είναι στατιστικά σημαντικός

V_s

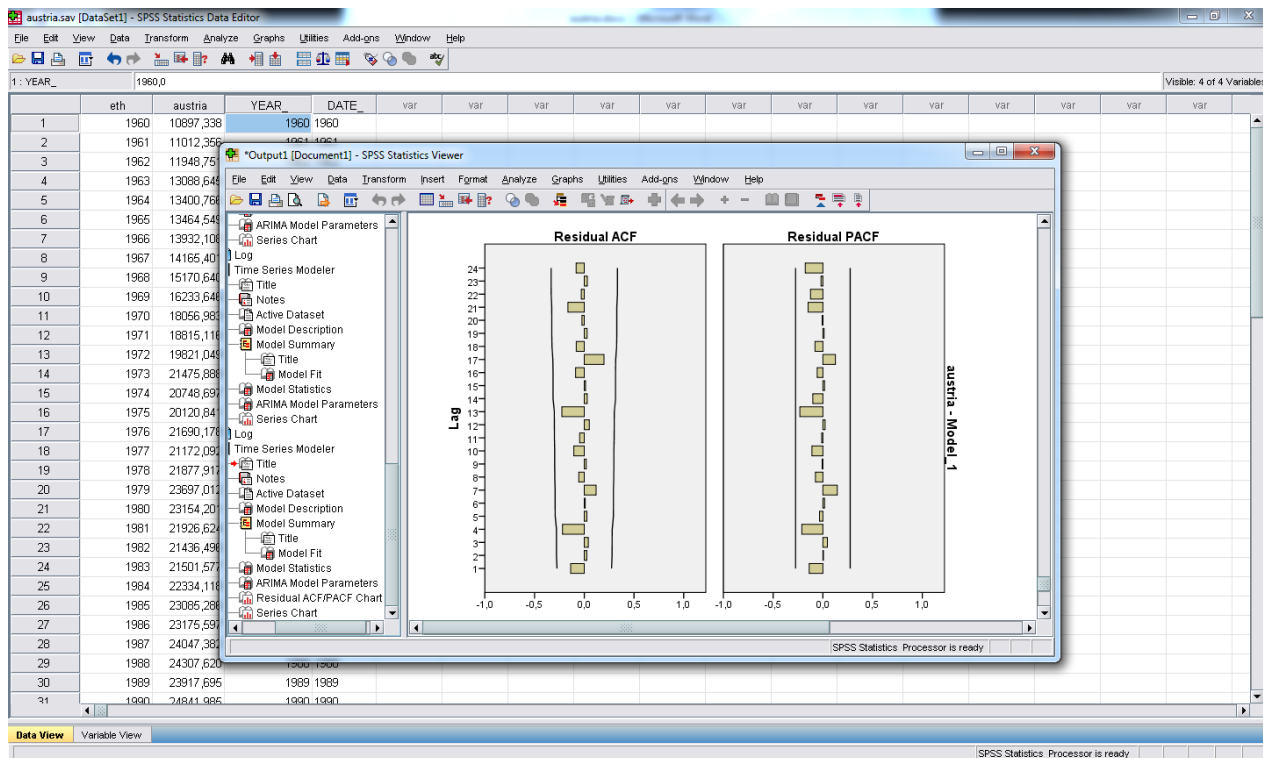
$H_1: \rho_0 \neq 0$

Αφού βρέθηκε ο εκτιμητής μας στατιστικά σημαντικός και γνωρίζοντας τις παραμέτρους του υποδείγματος ARIMA(0,1,0) θεωρείται ως το καταλληλότερο. Η μαθηματική μορφή του υποδείγματος θα είναι :

$$Y_t = 426,007 + E_t$$

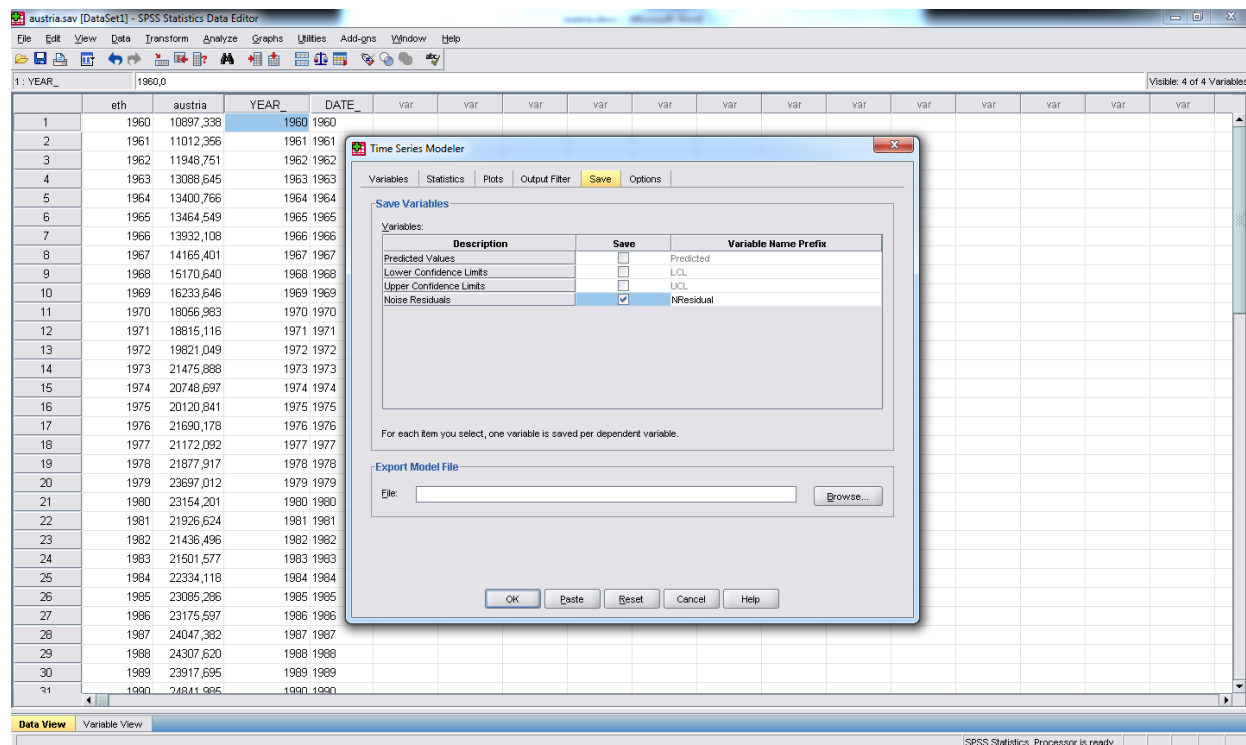
Για να είναι το υπόδειγμα Y_t αποδεκτό, για να θεωρηθεί δηλαδή ότι περιγράφει την διαδικασία που παρήγαγε τα δεδομένα, θα πρέπει τα κατάλοιπα να είναι λευκός θόρυβος, δηλαδή δεν θα πρέπει να αυτοσυσχετίζονται. Για το λόγο αυτό θα γίνει έλεγχος στασιμότητας στα κατάλοιπα (residuals) επαληθεύοντας έτσι αν τα αποτελέσματα είναι αληθή για το ARIMA(0,1,0).

Έτσι στο PASW 18 επιλέγουμε Analyze → Forecasting → Create models → Plots και στη συνέχεια επιλέγουμε Residuals autocorrelations function (ACF) και Residuals partial autocorrelation function (PACF) και στη συνέχεια πατάμε ok.



Σχήμα 37. Εύρεση των καταλοίπων (residuals)

Παρατηρούμε ότι είναι και τα δύο μηδενικά οπότε έχουμε λευκό θόρυβο. Επιστρέφοντας και πάλι στο PASW 18 και στα δεδομένα από τις επιλογές Analyze → Forecasting → Create models και στην επιλογή Save επιλέγουμε το Noise Residuals και στη συνέχεια πατάμε ok.



Σχήμα 38. Διαδικασία για την εύρεση των καταλοίπων (residuals)

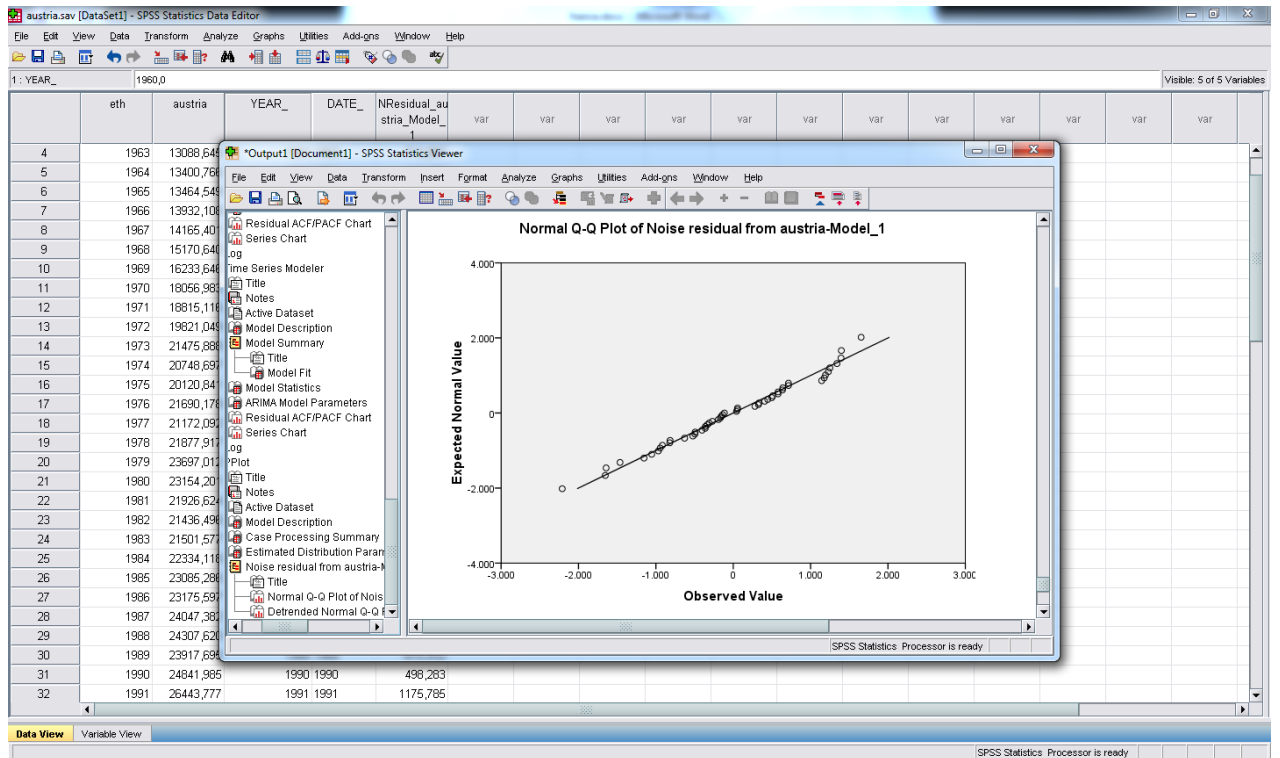
Πατώντας ok εμφανίζεται η μεταβλητή Noise Residuals Austria στο PASW 18 που δείχνει τα υπόλοιπα.

	eth	austria	YEAR_	DATE_	NResidual_austria_Model_1	var	var	var	var	var	var	var	var	var	var	var
4	1963	13088,645	1963	1963	713,887											
5	1964	13400,766	1964	1964	-113,886											
6	1965	13464,549	1965	1965	-362,224											
7	1966	13932,108	1966	1966	41,552											
8	1967	14165,401	1967	1967	-192,714											
9	1968	15170,640	1968	1968	579,232											
10	1969	16233,646	1969	1969	636,999											
11	1970	18056,983	1970	1970	1397,330											
12	1971	18815,116	1971	1971	332,126											
13	1972	19821,049	1972	1972	579,926											
14	1973	21475,888	1973	1973	1228,832											
15	1974	20748,697	1974	1974	-1153,198											
16	1975	20120,841	1975	1975	-1053,863											
17	1976	21690,178	1976	1976	1143,330											
18	1977	21172,092	1977	1977	-944,093											
19	1978	21877,917	1978	1978	279,818											
20	1979	23697,012	1979	1979	1393,088											
21	1980	23154,201	1980	1980	-968,818											
22	1981	21926,624	1981	1981	-1653,584											
23	1982	21436,496	1982	1982	-916,135											
24	1983	21501,577	1983	1983	-360,926											
25	1984	22334,118	1984	1984	406,534											
26	1985	23085,286	1985	1985	325,161											
27	1986	23175,597	1986	1986	-335,696											
28	1987	24047,382	1987	1987	445,778											
29	1988	24307,620	1988	1988	-165,769											
30	1989	23917,695	1989	1989	-815,932											
31	1990	24841,985	1990	1990	498,283											
32	1991	26443,777	1991	1991	1175,785											

Σχήμα 39. Μεταβλητή noise residuals France

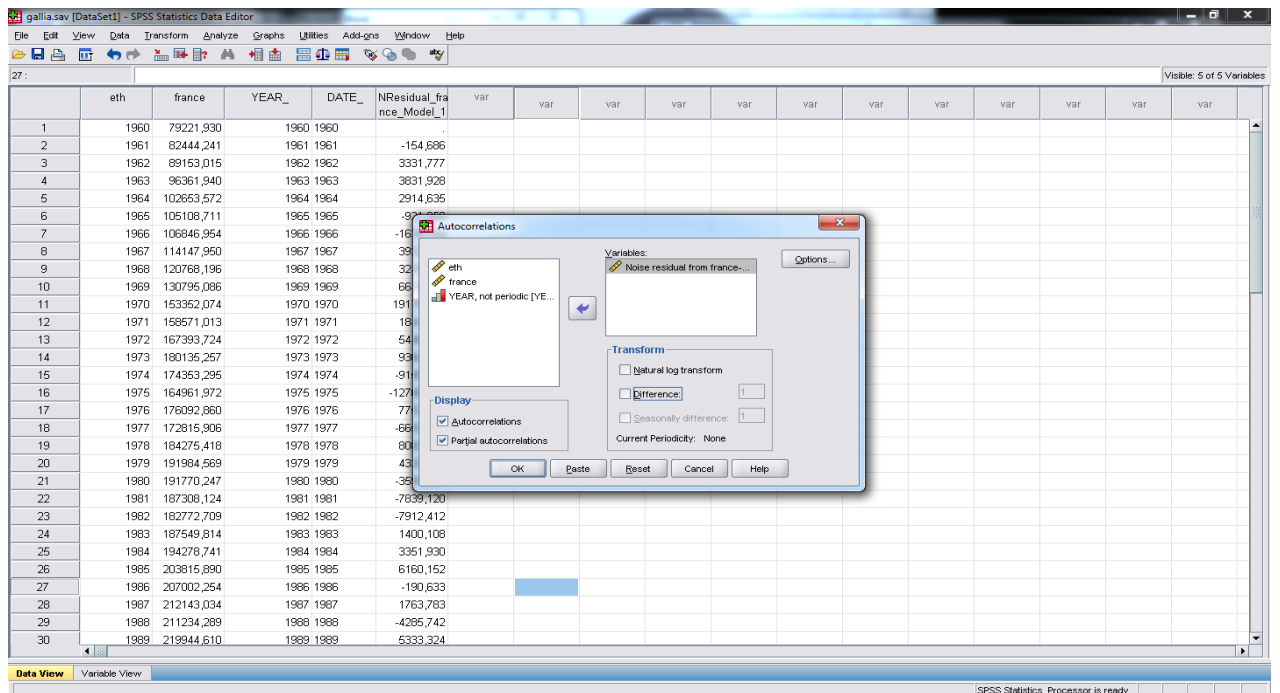
Από το μενού του PASW 18 επιλέγουμε Analyze→Descriptive Statistics→Q-Plots και στο Variable εισάγουμε τη μεταβλητή Noise Residual from austria και πατάμε ok.

Σχήμα 40. Διαδικασία για έλεγχο στασιμότητας των Residuals



Σχήμα 41. Διάγραμμα κατανομής των Residuals

Στο διάγραμμα που εξάγεται παρατηρούμε ότι η κατανομή είναι κανονική αφού όλες οι παρατηρήσεις βρίσκονται κοντά στη γραμμή. Επιπλέον θα κάνουμε έλεγχο στασιμότητας και για τα Residuals. Από το μενού του PASW 18 επιλέγουμε Analyze → Forecasting → Autocorrelation και εισάγουμε τη μεταβλητή noise residuals from austria και στη συνέχεια πατάμε ok.



Σχήμα 42. Έλεγχος στασιμότητας Residuals

Από το PASW 18 εξάγεται ο παρακάτω πίνακας.

Autocorrelations

Series:Noise residual from austria-Model_1

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,138	,136	1,023	1	,312
2	,025	,135	1,056	2	,590
3	,041	,133	1,151	3	,765
4	-,219	,132	3,911	4	,418
5	,026	,130	3,952	5	,556
6	,009	,129	3,957	6	,682
7	,119	,128	4,828	7	,681
8	-,055	,126	5,018	8	,756
9	,020	,125	5,043	9	,831
10	-,105	,123	5,776	10	,834
11	-,046	,122	5,921	11	,879
12	,051	,120	6,101	12	,911
13	-,226	,119	9,722	13	,716
14	,029	,117	9,784	14	,778
15	,013	,115	9,797	15	,832
16	-,087	,114	10,386	16	,846

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 11. Πίνακας αυτοσυσχετίσεων για τα residuals

Στη συνέχεια γίνεται έλεγχος στατιστικής σημαντικότητας p-value των εκτιμητών του υποδείγματος για όλα τα lags.

$H_0: \rho=0$

V_s

$H_1: \rho \neq 0$

Με την υπόθεση H_0 ο εκτιμητής μας είναι στατιστικά μη σημαντικός, ενώ με την H_1 ο εκτιμητής μας είναι στατιστικά σημαντικός.

1ο lag

$H_0: \rho_1=0$

στάσιμη)

V_s

$H_1: \rho_1 \neq 0$

2]

$H_0: \rho_2=0$

στάσιμη)

V_s

$\rho_1=0,312 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα

$\rho_2=0,590 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα

H₁: $\rho_2 \neq 0$

3]

H₀: $\rho_3 = 0$
στάσιμη)

$p_3 = 0,765 > \alpha = 0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα

V_s

H₁: $\rho_3 \neq 0$

4]

H₀: $\rho_4 = 0$
στάσιμη)

$p_4 = 0,418 > \alpha = 0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα

V_s

H₁: $\rho_4 \neq 0$

5]

.....

16]

H₀: $\rho_{16} = 0$
στάσιμη)

$p_{16} = 0,846 > \alpha = 0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα

V_s

H₁: $\rho_{16} \neq 0$

V_s

H₁: $\rho_{16} \neq 0$

Άρα από τον έλεγχο καταλοίπων φαίνεται ότι η διαδικασία είναι στάσιμη καθώς για όλες τις χρονολογικές υστερήσεις τα ρ είναι στάσιμα. Έτσι το υπόδειγμα ARIMA(0,1,0) επιβεβαιώνεται ότι είναι το πλέον κατάλληλο.

Παράρτημα

Συνοπτικά αποτελέσματα χωρών Ευρωπαϊκής Ένωσης

ΒΕΛΓΙΟ

Από το μενού κάνουμε τις εξής επιλογές: Data->Define Dates-> Years και στην επιλογή Years γράφουμε 1960.

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή belgium και στην time axis lab την date και πατάμε ok.

Από τον πίνακα autocorrelations προκύπτει ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς διαφορές. Έτσι θα δοκιμασούμε με πρώτες διαφορές . Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας. Οπότε βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία είναι στάσιμη με πρώτες διαφορές. Στη συνέχεια με τα διαγράμματα ACF και PACF που φαίνονται στο output εντοπίζονται τα q και τα p αντίστοιχα.

Από τα διαγράμματα ACF και PACF προκύπτει q=0 και p=0. Το υπόδειγμα είναι της μορφής ARIMA(p,d,q)-ARIMA(0,1,0), (d=1 λόγω πρώτων διαφορών). Η θεωρητική μορφή του υποδείγματος είναι :

$$Y_t = \alpha_0 + \alpha_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Έπειτα πηγαίνοντας στο PASW 18 ακολουθούμε τις παρακάτω διαδικασίες. Από το μενού επιλέγουμε Analyze-> Forecasting-> create models. Στο Dependent Variable περνάμε την μεταβλητή Belgium και στο method επιλέγουμε ARIMA. Πατώντας το criteria εμφανίζεται ο πίνακας arima criteria και στο nonseasonal μπαίνουν οι τιμές των p, d, q που έχουν εντοπιστεί 0, 1, 0, αντίστοιχα. Πατώντας το continue και πηγαίνοντας στο statistics κλικάρουμε την επιλογή parameter estimates και στη συνέχεια ok.

Στη συνέχεια στο output βγαίνει ο πίνακας Arima Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί κάνοντας έλεγχο στατιστικής σημαντικότητας για το υπόδειγμα ARIMA(0,1,0).

Αφού βρέθηκε ο εκτιμητής μας στατιστικά σημαντικός και γνωρίζοντας τις παραμέτρους του υποδείγματος ARIMA(0,1,0) θεωρείται ως το καταλληλότερο. Η μαθηματική μορφή του υποδείγματος θα είναι :

$$Y_t = 644,878 + E_t$$

Για να είναι το υπόδειγμα Y_t αποδεκτό, για να θεωρηθεί δηλαδή ότι περιγράφει την διαδικασία που παρήγαγε τα δεδομένα, θα πρέπει τα κατάλοιπα να είναι λευκός θόρυβος, δηλαδή δεν θα πρέπει να αυτοσυσχετίζονται. Για το λόγο αυτό θα γίνει έλεγχος στασιμότητας στα κατάλοιπα (residuals) επαληθεύοντας έτσι αν τα αποτελέσματα είναι αληθή για το ARIMA(0,1,0).

Από το μενού του PASW 18 Analyze-> Forecasting-> Create Models-> Plots και κλικάρουμε τις επιλογές

Residuals autocorrelations function (ACF)

Residuals partial autocorrelation function (PACF) και στη συνέχεια ok

Εύρεση των Residuals από τους δυο πίνακες που προκύπτουν παρατηρώντας ότι είναι και τα δύο μηδενικά οπότε έχουμε λευκό θόρυβο.

Πηγαίνοντας πάλι στο PASW 18 και στα δεδομένα από τις επιλογές Analyze-> Forecasting-> Create Models και στην επιλογή Save κλικάρουμε το Noise Residuals και στη συνέχεια ok.

Από το μενού του PASW 18 επιλέγουμε Analyze-> Descriptive statistics-> Q-Qplots και στο Variable περνάμε την μεταβλητή Noise Residual from Belgium και πατάμε ok.

Από το διάγραμμα του output παρατηρούμε ότι ακολουθεί κανονική κατανομή, αφού όλες οι παρατηρήσεις είναι κοντά στην γραμμή.

Επιπλέον θα κάνουμε έλεγχο στασιμότητας και για τα Residuals. Από το μενού του PASW 18 επιλέγουμε Analyze-> Forecasting-> Autocorrelation και περνάμε την μεταβλητή noise residuals from Belgium και ok.

Άρα από τον έλεγχο καταλοίπων φαίνεται ότι η διαδικασία είναι στάσιμη καθώς για όλες τις χρονολογικές υστερήσεις τα p είναι στάσιμα. Έτσι το υπόδειγμα ARIMA(0,1,0) επιβεβαιώνεται ότι είναι το πλέον κατάλληλο.

Πίνακας{1} -> Πίνακας Autocorrelations στο output

Autocorrelations

Series:belgium

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,926	,135	47,252	1	,000
2	,843	,133	87,179	2	,000
3	,768	,132	120,991	3	,000
4	,700	,131	149,667	4	,000
5	,631	,129	173,474	5	,000
6	,559	,128	192,522	6	,000
7	,479	,127	206,826	7	,000
8	,396	,125	216,815	8	,000
9	,329	,124	223,875	9	,000
10	,274	,122	228,883	10	,000
11	,224	,121	232,321	11	,000
12	,173	,119	234,428	12	,000
13	,139	,118	235,825	13	,000
14	,121	,116	236,903	14	,000
15	,099	,115	237,648	15	,000
16	,068	,113	238,009	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν

Πίνακας{2} Autocorrelations

Autocorrelations

Series:belgium

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,019	,136	,019	1	,891
2	,119	,135	,802	2	,670
3	-,109	,133	1,470	3	,689
4	,007	,132	1,473	4	,831
5	,096	,130	2,011	5	,848
6	,023	,129	2,043	6	,916
7	,159	,128	3,590	7	,826
8	-,118	,126	4,465	8	,813
9	,062	,125	4,715	9	,858
10	-,040	,123	4,823	10	,903
11	-,095	,122	5,434	11	,908
12	-,104	,120	6,186	12	,906
13	-,206	,119	9,198	13	,758
14	-,037	,117	9,299	14	,811
15	-,144	,115	10,863	15	,762
16	,063	,114	11,167	16	,799

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

ΕΛΕΓΧΟΣ ΣΤΑΣΙΜΟΤΗΤΑΣ

1: $H_0: \rho_1=0$ $\rho_1=0,891 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_1 \neq 0$

2: $H_0: \rho_2=0$ $\rho_2=0,670 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_2 \neq 0$

3: $H_0: \rho_3=0$ $\rho_3=0,689 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_3 \neq 0$

4: $H_0: \rho_4=0$ $\rho_4=0,831 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_4 \neq 0$

.....

16: $H_0: \rho_{16} = 0$
στάσιμη)

V_s

$H_1: \rho_{16} \neq 0$

$\rho_{16} = 0,799 > \alpha = 0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα

Πίνακας {3} Autocorrelations για τα Residuals

Autocorrelations

Series:Noise residual from belgium-Model_1

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,019	,136	,019	1	,891
2	,119	,135	,802	2	,670
3	-,109	,133	1,470	3	,689
4	,007	,132	1,473	4	,831
5	,096	,130	2,011	5	,848
6	,023	,129	2,043	6	,916
7	,159	,128	3,590	7	,826
8	-,118	,126	4,465	8	,813
9	,062	,125	4,715	9	,858
10	-,040	,123	4,823	10	,903
11	-,095	,122	5,434	11	,908
12	-,104	,120	6,186	12	,906
13	-,206	,119	9,198	13	,758
14	-,037	,117	9,299	14	,811
15	-,144	,115	10,863	15	,762
16	,063	,114	11,167	16	,799

ΕΛΕΓΧΟΣ ΣΤΑΣΙΜΟΤΗΤΑΣ

1: $H_0: \rho_1=0$ $\rho_1=0,891 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_1 \neq 0$

2: $H_0: \rho_2=0$ $\rho_2=0,670 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_2 \neq 0$

3: $H_0: \rho_3=0$ $\rho_3=0,689 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_3 \neq 0$

4: $H_0: \rho_4=0$ $\rho_4=0,831 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_4 \neq 0$

.....

16: $H_0: \rho_{16}=0$ $\rho_{16}=0,799 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_{16} \neq 0$

ΦΙΝΛΑΝΔΙΑ

Από το μενού κάνουμε τις εξής επιλογές: Data-> Define Dates-> Years και στην επιλογή Years γράφουμε 1960

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή finland και στην time axis lab την date και πατάμε ok.

Από τον πίνακα autocorrelations προκύπτει ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα ρ είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς

διαφορές. Έτσι θα δοκιμασουμε με πρώτες διαφορές . Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας. Οπότε βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία είναι στάσιμη με πρώτες διαφορές. Στη συνέχεια με τα διαγράμματα ACF και PACF που φαίνονται στο output εντοπίζονται τα q και τα p αντίστοιχα.

Από τα διαγράμματα ACF και PACF προκύπτει q=1 και p=0. Το υπόδειγμα είναι της μορφής ARIMA(p,d,q)-ARIMA(0,1,1), (d=1 λόγω πρώτων διαφορών). Η θεωρητική μορφή του υποδείγματος είναι :

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Έπειτα πηγαίνοντας στο PASW 18 ακολουθούμε τις παρακάτω διαδικασίες. Από το μενού επιλέγουμε Analyze-> Forecasting-> create models. Στο Dependent Variable περνάμε την μεταβλητή Finland και στο method επιλέγουμε ARIMA. Πατώντας το criteria εμφανίζεται ο πίνακας arima criteria και στο nonseasonal μπαίνουν οι τιμές των p, d, q που έχουν εντοπιστεί 0, 1, 1, αντίστοιχα. Πατώντας το continue και πηγαίνοντας στο statistics κλικάρουμε την επιλογή parameter estimates και στη συνέχεια ok.

Στη συνέχεια στο output βγαίνει ο πίνακας Arima Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί κάνοντας έλεγχο στατιστικής σημαντικότητας για το υπόδειγμα ARIMA(0,1,1).

Αφού βρέθηκαν οι εκτιμητές μας στατιστικά σημαντικοί και γνωρίζοντας τις παραμέτρους του υποδείγματος ARIMA(0,1,1) θεωρείται ως το καταλληλότερο. Η μαθηματική μορφή του υποδείγματος θα είναι :

$$Y_t = 495,407 - 0,332 \varepsilon_{t-1}$$

Για να είναι το υπόδειγμα Y_t αποδεκτό, για να θεωρηθεί δηλαδή ότι περιγράφει την διαδικασία που παρήγαγε τα δεδομένα, θα πρέπει τα κατάλοιπα να είναι λευκός θόρυβος, δηλαδή δεν θα πρέπει να αυτοσυσχετίζονται. Για το λόγο αυτό θα γίνει έλεγχος στασιμότητας στα κατάλοιπα (residuals) επαληθεύοντας έτσι αν τα αποτελέσματα είναι αληθή για το ARIMA(0,1,1).

Από το μενού του PASW 18 Analyze-> Forecasting-> Create Models-> Plots και κλικάρουμε τις επιλογές

Residuals autocorrelations function (ACF)

Residuals partial autocorrelation function (PACF) και στη συνέχεια ok

Εύρεση των Residuals από τους δυο πίνακες που προκύπτουν παρατηρώντας ότι είναι και τα δύο μηδενικά οπότε έχουμε λευκό θόρυβο.

Πηγαίνοντας πάλι στο PASW 18 και στα δεδομένα από τις επιλογές Analyze-> Forecasting-> Create Models και στην επιλογή Save κλικάρουμε το Noise Residuals και στη συνέχεια ok.

Από το μενού του PASW 18 επιλέγουμε Analyze-> Descriptive statistics-> Q-Qplots και στο Variable περνάμε την μεταβλητή Noise Residual from Finland και πατάμε ok.

Από το διάγραμμα του output παρατηρούμε ότι ακολουθεί κανονική κατανομή, αφού όλες οι παρατηρήσεις είναι κοντά στην γραμμή.

Επιπλέον θα κάνουμε έλεγχο στασιμότητας και για τα Residuals. Από το μενού του PASW 18 επιλέγουμε Analyze-> Forecasting-> Autocorrelation και περνάμε την μεταβλητή noise residuals from Finland και ok.

Άρα από τον έλεγχο καταλοίπων φαίνεται ότι η διαδικασία είναι στάσιμη καθώς για όλες τις χρονολογικές υστερήσεις τα p είναι στάσιμα. Έτσι το υπόδειγμα ARIMA(0,1,1) επιβεβαιώνεται ότι είναι το πλέον κατάλληλο.

Πίνακας{1} -> Πίνακας Autocorrelations στο output

Autocorrelations

Series:finland

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,938	,135	48,439	1	,000
2	,874	,133	91,357	2	,000
3	,820	,132	129,904	3	,000
4	,760	,131	163,673	4	,000
5	,691	,129	192,204	5	,000
6	,629	,128	216,375	6	,000
7	,576	,127	237,102	7	,000
8	,509	,125	253,635	8	,000
9	,445	,124	266,591	9	,000
10	,390	,122	276,770	10	,000
11	,338	,121	284,607	11	,000
12	,291	,119	290,547	12	,000
13	,249	,118	294,998	13	,000
14	,209	,116	298,229	14	,000
15	,165	,115	300,296	15	,000
16	,122	,113	301,450	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν

Πίνακας{2} Autocorrelations

Autocorrelations

Series:finland

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,251	,136	3,396	1	,065
2	-,082	,135	3,762	2	,152
3	,002	,133	3,763	3	,288
4	,195	,132	5,939	4	,204
5	-,296	,130	11,078	5	,050
6	,000	,129	11,078	6	,086
7	,185	,128	13,191	7	,068
8	-,058	,126	13,403	8	,099
9	,020	,125	13,430	9	,144
10	,090	,123	13,962	10	,175
11	-,131	,122	15,116	11	,177
12	-,103	,120	15,851	12	,198
13	,071	,119	16,213	13	,238
14	,023	,117	16,252	14	,298
15	-,167	,115	18,353	15	,245
16	,108	,114	19,252	16	,256

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

ΕΛΕΓΧΟΣ ΣΤΑΣΙΜΟΤΗΤΑΣ

1:H₀: ρ₁=0 ρ₁=0,065>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: ρ₁≠0

2:H₀: ρ₂=0 ρ₂=0,152>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: ρ₂≠0

3:H₀: ρ₃=0 ρ₃=0,288>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: ρ₃≠0

4:H₀: ρ₄=0 ρ₄=0,204>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: $\rho_4 \neq 0$

.....

16: H₀: $\rho_{16} = 0$ $\rho_{16} = 0,256 > \alpha = 0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

Vs

H₁: $\rho_{16} \neq 0$

Πίνακας {3} Autocorrelations για τα
Residuals **Autocorrelations**

Series: Noise residual from finland-Model_1

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,013	,136	,009	1	,925
2	-,058	,135	,193	2	,908
3	-,009	,133	,197	3	,978
4	,089	,132	,652	4	,957
5	-,280	,130	5,244	5	,387
6	-,018	,129	5,264	6	,510
7	,150	,128	6,654	7	,466
8	,002	,126	6,654	8	,574
9	,053	,125	6,833	9	,654
10	,070	,123	7,153	10	,711
11	-,148	,122	8,635	11	,656
12	-,146	,120	10,105	12	,607
13	,020	,119	10,134	13	,683
14	-,006	,117	10,137	14	,752
15	-,110	,115	11,051	15	,749
16	,171	,114	13,316	16	,650

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

ΕΛΕΓΧΟΣ ΣΤΑΣΙΜΟΤΗΤΑΣ

1:H ₀ : p ₁ =0 V _s H ₁ : p ₁ ≠0	p ₁ =0,925>α=0,05 Αποδοχή H ₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
2:H ₀ : p ₂ =0 V _s H ₁ : p ₂ ≠0	p ₂ =0,908>α=0,05 Αποδοχή H ₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
3:H ₀ : p ₃ =0 V _s H ₁ : p ₃ ≠0	p ₃ =0,978>α=0,05 Αποδοχή H ₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
4:H ₀ : p ₄ =0 V _s H ₁ : p ₄ ≠0	p ₄ =0,957>α=0,05 Αποδοχή H ₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
.....	
16:H ₀ : p ₁₆ =0 V _s H ₁ : p ₁₆ ≠0	p ₁₆ =0,650>α=0,05 Αποδοχή H ₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

ΓΕΡΜΑΝΙΑ

Από το μενού κάνουμε τις εξής επιλογές: Data-> Define Dates-> Years και στην επιλογή Years γράφουμε 1960

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή germany και στην time axis lab την date και πατάμε ok.

Από τον παραπάνω πίνακα προκύπτει ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς διαφορές. Έτσι θα δοκιμασούμε με πρώτες διαφορές. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας. Οπότε βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία είναι στάσιμη με πρώτες διαφορές. Στη συνέχεια με τα διαγράμματα ACF και PACF που φαίνονται στο output εντοπίζονται τα q και τα p αντίστοιχα.

Από τα διαγράμματα ACF και PACF προκύπτει q=0 και p=1. Το υπόδειγμα είναι της μορφής ARIMA(p,d,q)-ARIMA(1,1,0), (d=1 λόγω πρώτων διαφορών). Η θεωρητική μορφή του υποδείγματος είναι :

$$Y_t = \alpha_0 + \alpha_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Έπειτα πηγαίνοντας στο PASW 18 ακολουθούμε τις παρακάτω διαδικασίες. Από το μενού επιλέγουμε Analyze-> Forecasting-> create models. Στο Dependent Variable περνάμε την μεταβλητή germany και στο method επιλέγουμε ARIMA. Πατώντας το criteria εμφανίζεται ο πίνακας arima criteria και στο nonseasonal μπαίνουν οι τιμές των p, d, q που έχουν εντοπιστεί 1, 1, 0, αντίστοιχα. Πατώντας το continue και πηγαίνοντας στο statistics κλικάρουμε την επιλογή parameter estimates και στη συνέχεια ok.

Στη συνέχεια στο output βγαίνει ο πίνακας Arima Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί κάνοντας έλεγχο στατιστικής σημαντικότητας για το υπόδειγμα ARIMA(1,1,0).

Αφού βρέθηκαν οι εκτιμητές μας στατιστικά μη σημαντικοί αφαιρώ τις στατιστικά μη σημαντικές παραμέτρους, μια μια την φορά, ξεκινώντας από αυτή με το μεγαλύτερο p-value. Δηλαδή ξαναπροσαρμόζω το μοντέλο βάζοντας μηδέν στην παράμετρο που πρέπει (p ή d ή q).

Αναπροσαρμογή μοντέλου με p=0. Έλεγχος στατιστικής σημαντικότητας p-value των εκτιμητών του υποδείγματος

Η μαθηματική μορφή του υποδείγματος θα είναι :

$$Y_t = 247961,595 - 0,984 Y_{t-1}$$

Για να είναι το υπόδειγμα Y_t αποδεκτό, για να θεωρηθεί δηλαδή ότι περιγράφει την διαδικασία που παρήγαγε τα δεδομένα, θα πρέπει τα κατάλοιπα να είναι λευκός θόρυβος, δηλαδή δεν θα πρέπει να αυτοσυσχετίζονται. Για το λόγο αυτό θα γίνει έλεγχος στασιμότητας στα κατάλοιπα (residuals) επαληθεύοντας έτσι αν τα αποτελέσματα είναι αληθή για το ARIMA(0,1,1).

Από το μενού του PASW 18 Analyze-> Forecasting-> Create Models-> Plots και κλικάρουμε τις επιλογές

Residuals autocorrelations function (ACF)

Residuals partial autocorrelation function (PACF) και στη συνέχεια ok

Εύρεση των Residuals από τους δυο πίνακες που προκύπτουν παρατηρώντας ότι είναι και τα δύο μηδενικά οπότε έχουμε λευκό θόρυβο.

Πηγαίνοντας πάλι στο PASW 18 και στα δεδομένα από τις επιλογές Analyze-> Forecasting-> Create Models και στην επιλογή Save κλικάρουμε το Noise Residuals και στη συνέχεια ok.

Από το μενού του PASW 18 επιλέγουμε Analyze-> Descriptive statistics-> Q-Qplots και στο Variable περνάμε την μεταβλητή Noise Residual from germany και πατάμε ok.

Από το διάγραμμα του output παρατηρούμε ότι ακολουθεί κανονική κατανομή, αφού όλες οι παρατηρήσεις είναι κοντά στην γραμμή.

Επιπλέον θα κάνουμε έλεγχο στασιμότητας και για τα Residuals. Από το μενού του PASW 18 επιλέγουμε Analyze-> Forecasting-> Autocorrelation και περνάμε την μεταβλητή noise residuals from germany και ok.

Άρα από τον έλεγχο καταλοίπων φαίνεται ότι η διαδικασία είναι στάσιμη καθώς για όλες τις χρονολογικές υστερήσεις τα p είναι στάσιμα. Έτσι το υπόδειγμα ARIMA(0,1,1) επιβεβαιώνεται ότι είναι το πλέον κατάλληλο.

Πίνακας{1} -> Πίνακας Autocorrelations στο output

Autocorrelations

Series:germany

Lag	Autocorrelatio n	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,914	,135	45,987	1	,000
2	,821	,133	83,839	2	,000
3	,725	,132	113,955	3	,000
4	,630	,131	137,198	4	,000
5	,541	,129	154,694	5	,000
6	,457	,128	167,440	6	,000
7	,361	,127	175,560	7	,000
8	,255	,125	179,725	8	,000
9	,149	,124	181,182	9	,000
10	,046	,122	181,325	10	,000
11	,006	,121	181,327	11	,000
12	-,029	,119	181,385	12	,000
13	-,052	,118	181,581	13	,000
14	-,060	,116	181,847	14	,000
15	-,077	,115	182,296	15	,000
16	-,109	,113	183,225	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν

Πίνακας{2} Autocorrelations

Autocorrelations

Series:germany

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,142	,136	1,092	1	,296
2	,171	,135	2,706	2	,258
3	,109	,133	3,381	3	,337
4	-,070	,132	3,660	4	,454
5	-,086	,130	4,092	5	,536
6	,220	,129	7,000	6	,321
7	,139	,128	8,183	7	,317
8	,173	,126	10,055	8	,261
9	,155	,125	11,593	9	,237
10	-,047	,123	11,737	10	,303
11	-,070	,122	12,069	11	,358
12	-,083	,120	12,542	12	,403
13	-,060	,119	12,796	13	,464
14	,073	,117	13,181	14	,512
15	,066	,115	13,508	15	,563
16	,039	,114	13,622	16	,627

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

ΕΛΕΓΧΟΣ ΣΤΑΣΙΟΤΗΤΑΣ

1: $H_0: \rho_1=0$ $\rho_1=0,296 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_S

$H_1: \rho_1 \neq 0$

2: $H_0: \rho_2=0$ $\rho_2=0,258 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_S

$H_1: \rho_2 \neq 0$

3: $H_0: \rho_3=0$ $\rho_3=0,337 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_S

$H_1: \rho_3 \neq 0$

4: $H_0: \rho_4=0$ $\rho_4=0,454 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
 V_s
 $H_1: \rho_4 \neq 0$

.....
 16: $H_0: \rho_{16}=0$ $\rho_{16}=0,627 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
 V_s
 $H_1: \rho_{16} \neq 0$

Πίνακας {3} Autocorrelations για τα Residuals

Autocorrelations

Series: Noise residual from germany-Model_1

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,019	,136	,020	1	,888
2	,138	,135	1,069	2	,586
3	,105	,133	1,693	3	,638
4	-,077	,132	2,031	4	,730
5	-,108	,130	2,717	5	,744
6	,222	,129	5,669	6	,461
7	,089	,128	6,153	7	,522
8	,137	,126	7,329	8	,502
9	,142	,125	8,632	9	,472
10	-,058	,123	8,853	10	,546
11	-,056	,122	9,067	11	,616
12	-,067	,120	9,380	12	,670
13	-,058	,119	9,621	13	,725
14	,076	,117	10,040	14	,759
15	,054	,115	10,262	15	,803
16	,019	,114	10,291	16	,851

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

ΕΛΕΓΧΟΣ ΣΤΑΣΙΟΤΗΤΑΣ

1:H ₀ : $p_1=0$ στάσιμη) V _s H ₁ : $p_1 \neq 0$	$p_1=0,888 > \alpha=0,05$ Αποδοχή H ₀ (δεν υπάρχει αυτοσυσχέτιση, άρα
2:H ₀ : $p_2=0$ στάσιμη) V _s H ₁ : $p_2 \neq 0$	$p_2=0,586 > \alpha=0,05$ Αποδοχή H ₀ (δεν υπάρχει αυτοσυσχέτιση, άρα
3:H ₀ : $p_3=0$ στάσιμη) V _s H ₁ : $p_3 \neq 0$	$p_3=0,638 > \alpha=0,05$ Αποδοχή H ₀ (δεν υπάρχει αυτοσυσχέτιση, άρα
4:H ₀ : $p_4=0$ στάσιμη) V _s H ₁ : $p_4 \neq 0$	$p_4=0,730 > \alpha=0,05$ Αποδοχή H ₀ (δεν υπάρχει αυτοσυσχέτιση, άρα
.....	
16:H ₀ : $p_{16}=0$ στάσιμη) V _s H ₁ : $p_{16} \neq 0$	$p_{16}=0,851 > \alpha=0,05$ Αποδοχή H ₀ (δεν υπάρχει αυτοσυσχέτιση, άρα

ΙΡΛΑΝΔΙΑ

Από το μενού κάνουμε τις εξής επιλογές: Data-> Define Dates-> Years και στην επιλογή Years γράφουμε 1960

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή ireland και στην time axis lab την date και πατάμε ok.

Από τον πίνακα autocorrelations προκύπτει ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς διαφορές. Έτσι θα δοκιμασουμε με πρώτες διαφορές . Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας. Οπότε βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία είναι στάσιμη με πρώτες διαφορές. Στη συνέχεια με τα διαγράμματα ACF και PACF που φαίνονται στο output εντοπίζονται τα q και τα p αντίστοιχα.

Από τα διαγράμματα ACF και PACF προκύπτει $q=0$ και $p=1$. Το υπόδειγμα είναι της μορφής ARIMA(p,d,q)-ARIMA(1,1,0), ($d=1$ λόγω πρώτων διαφορών). Η θεωρητική μορφή του υποδείγματος είναι :

$$Y_t = \alpha_0 + \alpha_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Έπειτα πηγαίνοντας στο PASW 18 ακολουθούμε τις παρακάτω διαδικασίες. Από το μενού επιλέγουμε Analyze-> Forecasting-> create models. Στο Dependent Variable περνάμε την μεταβλητή ireland και στο method επιλέγουμε ARIMA. Πατώντας το criteria εμφανίζεται ο πίνακας arima criteria και στο nonseasonal μπαίνουν οι τιμές των p, d, q που έχουν εντοπιστεί 1, 1, 0, αντίστοιχα. Πατώντας το continue και πηγαίνοντας στο statistics κλικάρουμε την επιλογή parameter estimates και στη συνέχεια ok.

Στη συνέχεια στο output βγαίνει ο πίνακας Arima Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί κάνοντας έλεγχο στατιστικής σημαντικότητας για το υπόδειγμα ARIMA(1,1,0).

Αφού βρέθηκαν οι εκτιμητές μας στατιστικά μη σημαντικοί, εκτός του σταθερου που είναι στατιστικά σημαντικός, αφαιρώ τις στατιστικά μη σημαντικές παραμέτρους, μια μια την φορά, ξεκινώντας από αυτή με το μεγαλύτερο p-value. Δηλαδή ξαναπροσαρμόζω το μοντέλο βάζοντας μηδέν στην παράμετρο που πρέπει (p ή d ή q).

Αναπροσαρμογή μοντέλου με p=0. Έλεγχος στατιστικής σημαντικότητας p-value των εκτιμητών του υποδείγματος. Αφού βρέθηκε ο εκτιμητής μας στατιστικά σημαντικός και γνωρίζοντας τις παραμέτρους του υποδείγματος ARIMA(0,1,0) θεωρείται ως το καταλληλότερο. Η μαθηματική μορφή του υποδείγματος θα είναι :

$$Y_t = 191,851 + E_t$$

Για να είναι το υπόδειγμα Y_t αποδεκτό, για να θεωρηθεί δηλαδή ότι περιγράφει την διαδικασία που παρήγαγε τα δεδομένα, θα πρέπει τα κατάλοιπα να είναι λευκός θόρυβος, δηλαδή δεν θα πρέπει να αυτοσυσχετίζονται. Για το λόγο αυτό θα γίνει έλεγχος στασιμότητας στα κατάλοιπα (residuals) επαληθεύοντας έτσι αν τα αποτελέσματα είναι αληθή για το ARIMA(0,1,0).

Από το μενού του PASW 18 Analyze-> Forecasting-> Create Models-> Plots και κλικάρουμε τις επιλογές

Residuals autocorrelations function (ACF)

Residuals partial autocorrelation function (PACF) και στη συνέχεια ok

Εύρεση των Residuals από τους δυο πίνακες που προκύπτουν παρατηρώντας ότι δεν είναι και τα δύο μηδενικά οπότε δεν έχουμε λευκό θόρυβο.

Πίνακας{1} -> Πίνακας Autocorrelations στο output

Autocorrelations

Series:ireland

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,956	,135	50,340	1	,000
2	,905	,133	96,316	2	,000
3	,849	,132	137,590	3	,000
4	,785	,131	173,663	4	,000
5	,721	,129	204,739	5	,000
6	,656	,128	230,990	6	,000
7	,591	,127	252,799	7	,000
8	,527	,125	270,535	8	,000
9	,463	,124	284,545	9	,000
10	,399	,122	295,202	10	,000
11	,336	,121	302,924	11	,000
12	,281	,119	308,451	12	,000
13	,227	,118	312,154	13	,000
14	,178	,116	314,503	14	,000
15	,136	,115	315,903	15	,000
16	,093	,113	316,571	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν

Πίνακας{2} Autocorrelations

Autocorrelations

Series:ireland

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,069	,136	,254	1	,614
2	,203	,135	2,517	2	,284
3	,159	,133	3,949	3	,267
4	-,166	,132	5,531	4	,237
5	,003	,130	5,532	5	,354
6	,081	,129	5,927	6	,431
7	-,093	,128	6,454	7	,488
8	,139	,126	7,668	8	,467
9	-,166	,125	9,432	9	,398
10	-,139	,123	10,712	10	,380
11	,020	,122	10,738	11	,465
12	-,176	,120	12,879	12	,378
13	-,196	,119	15,610	13	,271
14	-,147	,117	17,192	14	,246
15	-,187	,115	19,810	15	,179
16	-,059	,114	20,076	16	,217

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

ΕΛΕΓΧΟΣ ΣΤΑΣΙΟΤΗΤΑΣ

1:H₀: ρ₁=0 ρ₁=0,614>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: ρ₁≠0

2:H₀: ρ₂=0 ρ₂=0,284>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: ρ₂≠0

3:H₀: ρ₃=0 ρ₃=0,267>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: ρ₃≠0

4:H₀: ρ₄=0 ρ₄=0,237>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_4 \neq 0$

.....

16: $H_0: \rho_{16} = 0$ $\rho_{16} = 0,217 > \alpha = 0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_{16} \neq 0$

ΙΤΑΛΙΑ

Από το μενού κάνουμε τις εξής επιλογές: Data-> Define Dates-> Years και στην επιλογή Years γράφουμε 1960

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή italy και στην time axis lab την date και πατάμε ok.

Από τον πίνακα autocorrelations προκύπτει ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα ρ είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς διαφορές. Έτσι θα δοκιμασούμε με πρώτες διαφορές. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας. Οπότε βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία είναι στάσιμη με πρώτες διαφορές. Στη συνέχεια με τα διαγράμματα ACF και PACF που φαίνονται στο output εντοπίζονται τα q και τα p αντίστοιχα

Από τα διαγράμματα ACF και PACF προκύπτει $q=1$ και $p=1$. Το υπόδειγμα είναι της μορφής ARIMA(p,d,q)-ARIMA(1,1,1), ($d=1$ λόγω πρώτων διαφορών). Η θεωρητική μορφή του υποδείγματος είναι :

$$Y_t = \alpha_0 + \alpha_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Έπειτα πηγαίνοντας στο PASW 18 ακολουθούμε τις παρακάτω διαδικασίες. Από το μενού επιλέγουμε Analyze-> Forecasting-> create models. Στο Dependent Variable περνάμε την μεταβλητή italy και στο method επιλέγουμε ARIMA. Πατώντας το criteria εμφανίζεται ο πίνακας arima criteria και στο nonseasonal μπαίνουν οι τιμές των p, d, q που έχουν εντοπιστεί 1, 1, 1, αντίστοιχα. Πατώντας το continue και πηγαίνοντας στο statistics κλικάρουμε την επιλογή parameter estimates και στη συνέχεια ok.

Στη συνέχεια στο output βγαίνει ο πίνακας Arima Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί κάνοντας έλεγχο στατιστικής σημαντικότητας για το υπόδειγμα ARIMA(1,1,1).

Αφού βρέθηκε ο εκτιμητής μας στατιστικά σημαντικός εκτός από τον σταθερό και γνωρίζοντας τις παραμέτρους του υποδείγματος ARIMA(1,1,1) ξαναπροσαρμόζω το

μοντέλο και στα κριτήρια ξεκλικάρω το include constant in model και παταμε continue.

Ξανα προσαρμογή μοντέλου. Έλεγχος στατιστικής σημαντικότητας p-value των εκτιμητών του υποδείγματος.

Αφού βρέθηκε ο εκτιμητής μας στατιστικά σημαντικός και γνωρίζοντας τις παραμέτρους του υποδείγματος ARIMA(1,1,1) θεωρείται ως το καταλληλότερο. Η μαθηματική μορφή του υποδείγματος θα είναι :

$$Y_t = -0,967Y_{t-1} - 0,788E_{t-1} + E_t$$

Για να είναι το υπόδειγμα Y_t αποδεκτό, για να θεωρηθεί δηλαδή ότι περιγράφει την διαδικασία που παρήγαγε τα δεδομένα, θα πρέπει τα κατάλοιπα να είναι λευκός θόρυβος, δηλαδή δεν θα πρέπει να αυτοσυσχετίζονται. Για το λόγο αυτό θα γίνει έλεγχος στασιμότητας στα κατάλοιπα (residuals) επαληθεύοντας έτσι αν τα αποτελέσματα είναι αληθή για το ARIMA(1,1,1).

Από το μενού του PASW 18 Analyze-> Forecasting-> Create Models-> Plots και κλικάρουμε τις επιλογές

Residuals autocorrelations function (ACF)

Residuals partial autocorrelation function (PACF) και στη συνέχεια ok

Εύρεση των Residuals από τους δυο πίνακες που προκύπτουν παρατηρώντας ότι δεν είναι και τα δύο μηδενικά οπότε δεν έχουμε λευκό θόρυβο.

Πίνακας{1} -> Πίνακας Autocorrelations στο output

Autocorrelations

Series:italy

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,929	,135	47,538	1	,000
2	,857	,133	88,777	2	,000
3	,787	,132	124,251	3	,000
4	,712	,131	153,901	4	,000
5	,636	,129	178,074	5	,000
6	,562	,128	197,345	6	,000
7	,489	,127	212,265	7	,000
8	,421	,125	223,587	8	,000
9	,360	,124	232,066	9	,000
10	,309	,122	238,445	10	,000
11	,268	,121	243,356	11	,000
12	,220	,119	246,747	12	,000
13	,179	,118	249,047	13	,000
14	,147	,116	250,646	14	,000
15	,121	,115	251,763	15	,000
16	,087	,113	252,349	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν

Πίνακας {2} Autocorrelations

Autocorrelations

Series:italy

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,090	,136	,439	1	,507
2	,277	,135	4,671	2	,097
3	,225	,133	7,529	3	,057
4	,160	,132	8,994	4	,061
5	-,005	,130	8,996	5	,109
6	,114	,129	9,775	6	,134
7	,019	,128	9,797	7	,200
8	,048	,126	9,941	8	,269
9	-,021	,125	9,969	9	,353
10	,001	,123	9,969	10	,443
11	-,009	,122	9,975	11	,533
12	-,077	,120	10,383	12	,582
13	-,125	,119	11,498	13	,569
14	-,210	,117	14,706	14	,399
15	,085	,115	15,243	15	,434
16	-,062	,114	15,541	16	,485

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

ΕΛΕΓΧΟΣ ΣΤΑΣΙΜΟΤΗΤΑΣ

1: $H_0: \rho_1=0$ $\rho_1=0,507 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_S

$H_1: \rho_1 \neq 0$

2: $H_0: \rho_2=0$ $\rho_2=0,097 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_S

$H_1: \rho_2 \neq 0$

3: $H_0: \rho_3=0$ $\rho_3=0,057 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_S

$H_1: \rho_3 \neq 0$

4:H₀: $p_4=0$ $p_4=0,061>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα
στάσιμη)
V_s
H₁: $p_4\neq 0$

.....
16:H₀: $p_{16}=0$ $p_{16}=0,485>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα
στάσιμη)
V_s
H₁: $p_{16}\neq 0$

ΟΛΛΑΝΔΙΑ

Από το μενού κάνουμε τις εξής επιλογές: Data-> Define Dates-> Years και στην επιλογή Years γράφουμε 1960

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή netherland και στην time axis lab την date και πατάμε ok.

Από τον πίνακα autocorrelations προκύπτει ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς διαφορές. Έτσι θα δοκιμασούμε με πρώτες διαφορές . Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας. Οπότε βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία είναι στάσιμη με πρώτες διαφορές. Στη συνέχεια με τα διαγράμματα ACF και PACF που φαίνονται στο output εντοπίζονται τα q και τα p αντίστοιχα.

Από τα διαγράμματα ACF και PACF προκύπτει $q=0$ και $p=0$. Το υπόδειγμα είναι της μορφής ARIMA(p,d,q)-ARIMA(0,1,0), ($d=1$ λόγω πρώτων διαφορών). Η θεωρητική μορφή του υποδείγματος είναι :

$$Y_t = \alpha_0 + \alpha_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Έπειτα πηγαίνοντας στο PASW 18 ακολουθούμε τις παρακάτω διαδικασίες. Από το μενού επιλέγουμε Analyze-> Forecasting-> create models. Στο Dependent Variable περνάμε την μεταβλητή netherland και στο method επιλέγουμε ARIMA. Πατώντας το criteria εμφανίζεται ο πίνακας arima criteria και στο nonseasonal μπαίνουν οι τιμές των p, d, q που έχουν εντοπιστεί 0, 1, 0, αντίστοιχα. Πατώντας το continue και πηγαίνοντας στο statistics κλικάρουμε την επιλογή parameter estimates και στη συνέχεια ok.

Στη συνέχεια στο output βγαίνει ο πίνακας Arima Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί κάνοντας έλεγχο στατιστικής σημαντικότητας για το υπόδειγμα ARIMA(0,1,0).

Αφού βρέθηκε ο εκτιμητής μας στατιστικά σημαντικός και γνωρίζοντας τις παραμέτρους του υποδείγματος ARIMA(0,1,0) θεωρείται ως το καταλληλότερο. Η μαθηματική μορφή του υποδείγματος θα είναι :

$$Y_t = 1109,374 + E_t$$

Για να είναι το υπόδειγμα Y_t αποδεκτό, για να θεωρηθεί δηλαδή ότι περιγράφει την διαδικασία που παρήγαγε τα δεδομένα, θα πρέπει τα κατάλοιπα να είναι λευκός θόρυβος, δηλαδή δεν θα πρέπει να αυτοσυσχετίζονται. Για το λόγο αυτό θα γίνει έλεγχος στασιμότητας στα κατάλοιπα (residuals) επαληθεύοντας έτσι αν τα αποτελέσματα είναι αληθή για το ARIMA(0,1,0).

Από το μενού του PASW 18 Analyze-> Forecasting-> Create Models-> Plots και κλικάρουμε τις επιλογές

Residuals autocorrelations function (ACF)

Residuals partial autocorrelation function (PACF) και στη συνέχεια ok

Εύρεση των Residuals από τους δυο πίνακες που προκύπτουν παρατηρώντας ότι είναι και τα δύο μηδενικά οπότε έχουμε λευκό θόρυβο.

Πηγαίνοντας παλι στο PASW 18 και στα δεδομένα από τις επιλογές Analyze-> Forecasting-> Create Models και στην επιλογή Save κλικάρουμε το Noise Residuals και στη συνέχεια ok.

Από το μενού του PASW 18 επιλέγουμε Analyze-> Descriptive statistics-> Q-Qplots και στο Variable περνάμε την μεταβλητή Noise Residual from Netherland και πατάμε ok.

Από το διάγραμμα του output παρατηρούμε ότι ακολουθεί κανονική κατανομή, αφού όλες οι παρατηρήσεις είναι κοντά στην γραμμή.

Επιπλέον θα κάνουμε έλεγχο στασιμότητας και για τα Residuals. Από το μενού του PASW 18 επιλέγουμε Analyze-> Forecasting-> Autocorrelation και περνάμε την μεταβλητή noise residuals from Netherland και ok.

Έλεγχος στατιστικής σημαντικότητας p-value των εκτιμητών του υποδείγματος

Άρα από τον έλεγχο καταλοίπων φαίνεται ότι η διαδικασία είναι στάσιμη καθώς για όλες τις χρονολογικές υστερήσεις τα p είναι στάσιμα. Έτσι το υπόδειγμα ARIMA(0,1,0) επιβεβαιώνεται ότι είναι το πλέον κατάλληλο.

Πίνακας{1} -> Πίνακας Autocorrelations στο output

Autocorrelations

Series:Netherland

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,921	,135	46,754	1	,000
2	,833	,133	85,694	2	,000
3	,750	,132	117,946	3	,000
4	,665	,131	143,833	4	,000
5	,581	,129	163,975	5	,000
6	,499	,128	179,178	6	,000
7	,415	,127	189,918	7	,000
8	,327	,125	196,751	8	,000
9	,249	,124	200,796	9	,000
10	,186	,122	203,106	10	,000
11	,140	,121	204,444	11	,000
12	,100	,119	205,152	12	,000
13	,083	,118	205,646	13	,000
14	,076	,116	206,069	14	,000
15	,064	,115	206,376	15	,000
16	,042	,113	206,516	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν

Πίνακας {2} Autocorrelations

Autocorrelations

Series:Netherland

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,047	,136	,121	1	,728
2	,105	,135	,725	2	,696
3	,137	,133	1,782	3	,619
4	-,027	,132	1,823	4	,768
5	,068	,130	2,097	5	,836
6	,045	,129	2,219	6	,899
7	,209	,128	4,894	7	,673
8	-,042	,126	5,006	8	,757
9	-,136	,125	6,188	9	,721
10	-,149	,123	7,644	10	,664
11	-,075	,122	8,021	11	,711
12	-,051	,120	8,201	12	,769
13	-,130	,119	9,412	13	,741
14	,033	,117	9,489	14	,798
15	-,014	,115	9,505	15	,850
16	-,040	,114	9,627	16	,885

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

ΕΛΕΓΧΟΣ ΣΤΑΣΙΜΟΤΗΤΑΣ

1: $H_0: \rho_1=0$ $\rho_1=0,728 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_S

$H_1: \rho_1 \neq 0$

2: $H_0: \rho_2=0$ $\rho_2=0,696 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_S

$H_1: \rho_2 \neq 0$

3: $H_0: \rho_3=0$ $\rho_3=0,619 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_S

$H_1: \rho_3 \neq 0$

4:H₀: $\rho_4=0$ $\rho_4=0,768>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: $\rho_4\neq 0$

.....

16:H₀: $\rho_{16}=0$ $\rho_{16}=0,885>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: $\rho_{16}\neq 0$

Πίνακας {3} Autocorrelations για τα Residuals

Autocorrelations

Series:Noise residual from Netherland-Model_1

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,047	,136	,121	1	,728
2	,105	,135	,725	2	,696
3	,137	,133	1,782	3	,619
4	-,027	,132	1,823	4	,768
5	,068	,130	2,097	5	,836
6	,045	,129	2,219	6	,899
7	,209	,128	4,894	7	,673
8	-,042	,126	5,006	8	,757
9	-,136	,125	6,188	9	,721
10	-,149	,123	7,644	10	,664
11	-,075	,122	8,021	11	,711
12	-,051	,120	8,201	12	,769
13	-,130	,119	9,412	13	,741
14	,033	,117	9,489	14	,798
15	-,014	,115	9,505	15	,850
16	-,040	,114	9,627	16	,885

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

ΕΛΕΓΧΟΣ ΣΤΑΣΙΜΟΤΗΤΑΣ

1:H₀: $\rho_1=0$ $\rho_1=0,728>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: $\rho_1\neq 0$

2:H₀: $p_2=0$ $p_2=0,696>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα
στάσιμη)
V_s
H₁: $p_2\neq 0$

3:H₀: $p_3=0$ $p_3=0,619>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα
στάσιμη)
V_s
H₁: $p_3\neq 0$

4:H₀: $p_4=0$ $p_4=0,768>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα
στάσιμη)
V_s
H₁: $p_4\neq 0$

.....

16:H₀: $p_{16}=0$ $p_{16}=0,885>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα
στάσιμη)
V_s
H₁: $p_{16}\neq 0$

ΝΟΡΒΗΓΙΑ

Από το μενού κάνουμε τις εξής επιλογές: Data-> Define Dates-> Years και στην επιλογή Years γράφουμε 1960

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή norway και στην time axis lab την date και πατάμε ok.

Από τον πίνακα autocorrelations προκύπτει ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς διαφορές. Έτσι θα δοκιμασουμε με πρώτες διαφορές . Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Παρατηρούμε ότι πάλι για όλα τα lags υπάρχει αυτοσυσχέτιση. Οπότε θα παρουμε διαφορές και λογάριθμο μαζί. Από το μενού Analyze→ Forecasting→ Autocorrelations→ κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και κλικάρουμε και την επιλογή Natural log transform και ok.

Φαίνεται ότι η διαδικασία δεν είναι στάσιμη με πρώτες διαφορές και λογάριθμο μαζί αφού όλα σχεδόν τα p-value είναι μηδέν.. Οπότε παίρνουμε δεύτερες διαφορές. Από το μενού Analyze→ Forecasting→ Autocorrelations→ κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 2.

Βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία δεν είναι στάσιμη με δεύτερες διαφορές. Οπότε παίρνουμε δεύτερες διαφορές και λογάριθμο μαζί. . Από το μενού Analyze→ Forecasting→ Autocorrelations→ κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 2 και την επιλογή natural log transform.

Παρατηρούμε ότι συνεχίζει να μην είναι στάσιμη ακόμη και με δεύτερες διαφορές και λογάριθμο μαζί.

Συμπέρασμα λοιπόν ότι η χρονοσειρά αυτή δεν είναι στάσιμη.

Πίνακας { 1 } Πίνακας Autocorrelations στο output

Autocorrelations

Series:norway

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,930	,135	47,578	1	,000
2	,857	,133	88,783	2	,000
3	,793	,132	124,835	3	,000
4	,727	,131	155,724	4	,000
5	,670	,129	182,568	5	,000
6	,617	,128	205,790	6	,000
7	,566	,127	225,814	7	,000
8	,511	,125	242,507	8	,000
9	,459	,124	256,257	9	,000
10	,413	,122	267,682	10	,000
11	,362	,121	276,653	11	,000
12	,310	,119	283,400	12	,000
13	,258	,118	288,179	13	,000
14	,209	,116	291,394	14	,000
15	,162	,115	293,395	15	,000
16	,118	,113	294,489	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας {2} Autocorrelations

Autocorrelations

Series:norway

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,572	,136	17,704	1	,000
2	,353	,135	24,559	2	,000
3	-,258	,133	28,310	3	,000
4	-,044	,132	28,423	4	,000
5	-,001	,130	28,423	5	,000
6	-,068	,129	28,702	6	,000
7	,226	,128	31,850	7	,000
8	-,152	,126	33,309	8	,000
9	,111	,125	34,097	9	,000
10	-,042	,123	34,214	10	,000
11	,025	,122	34,257	11	,000
12	-,070	,120	34,596	12	,001
13	,068	,119	34,928	13	,001
14	-,057	,117	35,164	14	,001
15	,120	,115	36,246	15	,002
16	-,142	,114	37,811	16	,002

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον τα p-value είναι διάφορα του μηδενός κάνουμε έλεγχο στασιμότητας.

1:H₀: p₁=0 p₁=0,000<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: p₁≠0

2:H₀: p₂=0 p₂=0,000<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: p₂≠0

3:H₀: p₃=0 p₃=0,000<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: p₃≠0

4:H₀: p₄=0 p₄=0,000<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $\rho_4 \neq 0$

.....

16:H₀: $\rho_{16}=0$ $\rho_{16}=0,002 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $\rho_{16} \neq 0$

Πίνακας {3} Autocorrelations

Autocorrelations

Series:norway

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,218	,136	2,568	1	,109
2	,396	,135	11,220	2	,004
3	-,084	,133	11,621	3	,009
4	,038	,132	11,703	4	,020
5	,053	,130	11,868	5	,037
6	,075	,129	12,209	6	,057
7	,225	,128	15,305	7	,032
8	,033	,126	15,375	8	,052
9	,080	,125	15,790	9	,071
10	,016	,123	15,806	10	,105
11	,034	,122	15,882	11	,146
12	-,069	,120	16,208	12	,182
13	,099	,119	16,901	13	,204
14	,022	,117	16,937	14	,260
15	,111	,115	17,865	15	,270
16	-,127	,114	19,106	16	,263

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον τα p-value δεν είναι μηδενικά κάνουμε έλεγχο στασιμότητας.

1:H₀: $\rho_1=0$ $\rho_1=0,109 > \alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

Vs

H₁: $\rho_1 \neq 0$

2:H₀: $\rho_2=0$ $\rho_2=0,004 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

$H_1: \rho_2 \neq 0$
 3: $H_0: \rho_3 = 0$ $\rho_3 = 0,009 < \alpha = 0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)
 V_s
 $H_1: \rho_3 \neq 0$
 4: $H_0: \rho_4 = 0$ $\rho_4 = 0,020 < \alpha = 0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)
 V_s
 $H_1: \rho_4 \neq 0$

 16: $H_0: \rho_{16} = 0$ $\rho_{16} = 0,263 > \alpha = 0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα
 στάσιμη)
 V_s
 $H_1: \rho_{16} \neq 0$

Πίνακας {4} Autocorrelations

Autocorrelations

Series:norway

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,709	,137	26,645	1	,000
2	,404	,136	35,470	2	,000
3	-,203	,134	37,750	3	,000
4	,023	,133	37,781	4	,000
5	,035	,132	37,852	5	,000
6	-,123	,130	38,744	6	,000
7	,210	,129	41,401	7	,000
8	-,165	,127	43,092	8	,000
9	,063	,126	43,344	9	,000
10	-,025	,124	43,387	10	,000
11	,037	,122	43,478	11	,000
12	-,052	,121	43,661	12	,000
13	,085	,119	44,173	13	,000
14	-,096	,118	44,836	14	,000
15	,110	,116	45,731	15	,000
16	-,160	,114	47,696	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία δεν είναι στάσιμη με δεύτερες διαφορές. Οπότε παίρνουμε δεύτερες διαφορές και λογάριθμο μαζί.

Πίνακας {5} Autocorrelations

Autocorrelations

Series:norway

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,684	,137	24,791	1	,000
2	,377	,136	32,490	2	,000
3	-,196	,134	34,606	3	,000
4	,016	,133	34,620	4	,000
5	-,005	,132	34,622	5	,000
6	-,056	,130	34,809	6	,000
7	,131	,129	35,839	7	,000
8	-,054	,127	36,019	8	,000
9	-,028	,126	36,068	9	,000
10	,012	,124	36,077	10	,000
11	,035	,122	36,157	11	,000
12	-,090	,121	36,715	12	,000
13	,103	,119	37,467	13	,000
14	-,064	,118	37,766	14	,001
15	,099	,116	38,494	15	,001
16	-,169	,114	40,672	16	,001

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον τα p-value είναι διάφορα του μηδενός κάνουμε έλεγχο στασιμότητας.

1:H₀: $p_1=0$ $p_1=0,000 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $p_1 \neq 0$

2:H₀: $p_2=0$ $p_2=0,000 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $p_2 \neq 0$

3:H₀: $p_3=0$ $p_3=0,000 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

V_s
 $H_1: \rho_3 \neq 0$
 4: $H_0: \rho_4 = 0$ $\rho_4 = 0,000 < \alpha = 0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)
 V_s
 $H_1: \rho_4 \neq 0$

 16: $H_0: \rho_{16} = 0$ $\rho_{16} = 0,001 < \alpha = 0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)
 V_s
 $H_1: \rho_{16} \neq 0$

ΠΟΛΩΝΙΑ

Από το μενού κάνουμε τις εξής επιλογές: Data-> Define Dates-> Years και στην επιλογή Years γράφουμε 1960

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή poland και στην time axis lab την date και πατάμε ok.

Από τον πίνακα που προκύπτει φαίνεται ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα ρ είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς διαφορές. Έτσι θα δοκιμασουμε με πρώτες διαφορές . Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία δεν είναι στάσιμη με πρώτες διαφορές. Οπότε παίρνουμε πρώτες διαφορές και λογάριθμο μαζί. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και την επιλογή natural log transform και ok.

Βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία δεν είναι στάσιμη με πρώτες διαφορές και λογάριθμο μαζί. Οπότε παίρνουμε δεύτερες διαφορές. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 2

Βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία δεν είναι στάσιμη με δεύτερες διαφορές. Οπότε παίρνουμε δεύτερες διαφορές και λογάριθμο μαζί. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 2 και την επιλογή natural log transform και ok.

Παρατηρούμε ότι συνεχίζει να μην είναι στάσιμη ακόμη και με δεύτερες διαφορές και λογάριθμο μαζί.

Συμπέρασμα λοιπόν ότι η χρονοσειρά αυτή δεν είναι στάσιμη.

Πίνακας {1} **Autocorrelations**

Autocorrelations

Series: poland

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,926	,135	47,235	1	,000
2	,838	,133	86,643	2	,000
3	,745	,132	118,458	3	,000
4	,651	,131	143,219	4	,000
5	,555	,129	161,612	5	,000
6	,465	,128	174,827	6	,000
7	,378	,127	183,743	7	,000
8	,287	,125	188,997	8	,000
9	,190	,124	191,343	9	,000
10	,080	,122	191,774	10	,000
11	-,034	,121	191,855	11	,000
12	-,145	,119	193,334	12	,000
13	-,236	,118	197,360	13	,000
14	-,316	,116	204,735	14	,000
15	-,382	,115	215,827	15	,000
16	-,430	,113	230,261	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας {2} **Autocorrelations**

Autocorrelations

Series:poland

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,926	,135	47,235	1	,000
2	,838	,133	86,643	2	,000
3	,745	,132	118,458	3	,000
4	,651	,131	143,219	4	,000
5	,555	,129	161,612	5	,000
6	,465	,128	174,827	6	,000
7	,378	,127	183,743	7	,000
8	,287	,125	188,997	8	,000
9	,190	,124	191,343	9	,000
10	,080	,122	191,774	10	,000
11	-,034	,121	191,855	11	,000
12	-,145	,119	193,334	12	,000
13	-,236	,118	197,360	13	,000
14	-,316	,116	204,735	14	,000
15	-,382	,115	215,827	15	,000
16	-,430	,113	230,261	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον τα p-value είναι διάφορα του μηδενός κάνουμε έλεγχο στασιμότητας
 1: $H_0: \rho_1=0$ $\rho_1=0,022 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)

V_s

$H_1: \rho_1 \neq 0$

2: $H_0: \rho_2=0$ $\rho_2=0,030 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)

V_s

$H_1: \rho_2 \neq 0$

3: $H_0: \rho_3=0$ $\rho_3=0,070 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα
 στάσιμη)

V_s

$H_1: \rho_3 \neq 0$

4: $H_0: \rho_4=0$ $\rho_4=0,116 > \alpha=0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση,
 άρα στάσιμη)

Vs

H₁: $\rho_4 \neq 0$

.....

16:H₀: $\rho_{16}=0$ $\rho_{16}=0,004 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $\rho_{16} \neq 0$

Πίνακας {3} Autocorrelations

Autocorrelations

Series: poland

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,348	,136	6,546	1	,011
2	,261	,135	10,299	2	,006
3	,101	,133	10,872	3	,012
4	,176	,132	12,650	4	,013
5	,026	,130	12,691	5	,026
6	,096	,129	13,245	6	,039
7	,202	,128	15,750	7	,028
8	,260	,126	20,011	8	,010
9	,398	,125	30,214	9	,000
10	,260	,123	34,665	10	,000
11	,079	,122	35,084	11	,000
12	-,204	,120	37,961	12	,000
13	-,054	,119	38,168	13	,000
14	-,091	,117	38,774	14	,000
15	-,108	,115	39,650	15	,001
16	-,193	,114	42,516	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον τα p-value είναι διάφορα του μηδενός κάνουμε έλεγχο στασιμότητας

1:H₀: $\rho_1=0$ $\rho_1=0,011 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

$H_1: p_1 \neq 0$
 2: $H_0: p_2 = 0$ $p_2 = 0,006 < \alpha = 0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)
 V_S
 $H_1: p_2 \neq 0$
 3: $H_0: p_3 = 0$ $p_3 = 0,012 < \alpha = 0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)
 V_S
 $H_1: p_3 \neq 0$
 4: $H_0: p_4 = 0$ $p_4 = 0,013 < \alpha = 0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)
 V_S
 $H_1: p_4 \neq 0$

 16: $H_0: p_{16} = 0$ $p_{16} = 0,000 < \alpha = 0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)
 V_S
 $H_1: p_{16} \neq 0$

Πίνακας {4} **Autocorrelations**

Autocorrelations

Series: poland

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,408	,137	8,848	1	,003
2	,027	,136	8,886	2	,012
3	-,159	,134	10,287	3	,016
4	,134	,133	11,309	4	,023
5	-,157	,132	12,732	5	,026
6	-,010	,130	12,739	6	,047
7	,015	,129	12,752	7	,078
8	-,069	,127	13,042	8	,110
9	,252	,126	17,061	9	,048
10	,014	,124	17,073	10	,073
11	,064	,122	17,349	11	,098
12	-,325	,121	24,557	12	,017
13	,142	,119	25,984	13	,017
14	-,019	,118	26,010	14	,026
15	,042	,116	26,144	15	,037
16	-,206	,114	29,390	16	,021

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας

1:H₀: $p_1=0$ $p_1=0,003 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $p_1 \neq 0$

2:H₀: $p_2=0$ $p_2=0,012 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $p_2 \neq 0$

3:H₀: $p_3=0$ $p_3=0,016 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $p_3 \neq 0$

4:H₀: $p_4=0$ $p_4=0,023 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $\rho_4 \neq 0$

.....

16:H₀: $\rho_{16}=0$ $\rho_{16}=0,021 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $\rho_{16} \neq 0$

Πίνακας {5} Autocorrelations

Autocorrelations

Series: poland

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,438	,137	10,176	1	,001
2	,064	,136	10,397	2	,006
3	-,186	,134	12,316	3	,006
4	,173	,133	14,011	4	,007
5	-,168	,132	15,635	5	,008
6	-,029	,130	15,686	6	,016
7	,033	,129	15,753	7	,027
8	-,066	,127	16,024	8	,042
9	,218	,126	19,040	9	,025
10	,036	,124	19,126	10	,039
11	,078	,122	19,529	11	,052
12	-,336	,121	27,274	12	,007
13	,143	,119	28,712	13	,007
14	-,019	,118	28,738	14	,011
15	,055	,116	28,959	15	,016
16	-,235	,114	33,200	16	,007

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας

1:H₀: $\rho_1=0$ $\rho_1=0,001 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)

Vs

H₁: $\rho_1 \neq 0$

2:H₀: p₂=0 p₂=0,006<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)
V_s
H₁: p₂≠0

3:H₀: p₃=0 p₃=0,006<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)
V_s
H₁: p₃≠0

4:H₀: p₄=0 p₄=0,007<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)
V_s
H₁: p₄≠0

.....

16:H₀: p₁₆=0 p₁₆=0,007<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα δεν είναι στάσιμη)
V_s
H₁: p₁₆≠0

ΠΟΡΤΟΓΑΛΙΑ

Από το μενού κάνουμε τις εξής επιλογές: Data-> Define Dates-> Years και στην επιλογή Years γράφουμε 1960

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή Portugal και στην time axis lab την date και πατάμε ok.

Από τον πίνακα που προκύπτει φαίνεται ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς διαφορές. Έτσι θα δοκιμασουμε με πρώτες διαφορές . Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Από τον πίνακα που προκύπτει φαίνεται ότι για κάποια από τα lags υπάρχει αυτοσυσχέτιση, έχοντας ως αποτέλεσμα να μην είναι στάσιμη με πρώτες διαφορές. Έτσι θα δοκιμασουμε με πρώτες διαφορές και λογάριθμο μαζί. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και επιλέγουμε το natural log transform και ok.

Οπότε βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία είναι στάσιμη με πρώτες διαφορές και λογάριθμο μαζί. Στη συνέχεια με τα διαγράμματα ACF και PACF που φαίνονται στο output εντοπίζονται τα q και τα p αντίστοιχα.

Από τα διαγράμματα ACF και PACF προκύπτει q=1 και p=0. Το υπόδειγμα είναι της μορφής ARIMA(p,d,q)-ARIMA(0,1,1), (d=1 λόγω πρώτων διαφορών). Η θεωρητική μορφή του υποδείγματος είναι :

$$Y_t = \alpha_0 + \alpha_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Έπειτα πηγαίνοντας στο PASW 18 ακολουθούμε τις παρακάτω διαδικασίες. Από το μενού επιλέγουμε Analyze-> Forecasting-> create models. Στο Dependent Variable περνάμε την μεταβλητή portugal και στο method επιλέγουμε ARIMA. Πατώντας το criteria εμφανίζεται ο πίνακας arima criteria και στο nonseasonal μπαίνουν οι τιμές των p, d, q που έχουν εντοπιστεί 0, 1,1 αντίστοιχα. Πατώντας το continue και πηγαίνοντας στο statistics κλικάρουμε την επιλογή parameter estimates και στη συνέχεια ok.

Στη συνέχεια στο output βγαίνει ο πίνακας Arima Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί κάνοντας έλεγχο στατιστικής σημαντικότητας για το υπόδειγμα ARIMA(0,1,1).

Αφού βρέθηκε ο εκτιμητής μας στατιστικά μη σημαντικός εκτός από τον σταθερό, ξανά προσαρμόζω αφαιρώ τις στατιστικά μη σημαντικές παραμέτρους, μια μια την φορά, ξεκινώντας από αυτή με το μεγαλύτερο p-value. Δηλαδή ξανα προσαρμόζω το μοντέλο βάζοντας μηδέν στην παράμετρο που πρέπει (p ή d ή q). Στο output βγαίνει ο πίνακας Arima Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί.

Αφού βρέθηκε ο εκτιμητής μας στατιστικά σημαντικός και γνωρίζοντας τις παραμέτρους του υποδείγματος ARIMA(0,1,0) θεωρείται ως το καταλληλότερο. Η μαθηματική μορφή του υποδείγματος θα είναι :

$$Y_t = 395,339 + E_t$$

Για να είναι το υπόδειγμα Y_t αποδεκτό, για να θεωρηθεί δηλαδή ότι περιγράφει την διαδικασία που παρήγαγε τα δεδομένα, θα πρέπει τα κατάλοιπα να είναι λευκός θόρυβος, δηλαδή δεν θα πρέπει να αυτοσυσχετίζονται. Για το λόγο αυτό θα γίνει έλεγχος στασιμότητας στα κατάλοιπα (residuals) επαληθεύοντας έτσι αν τα αποτελέσματα είναι αληθή για το ARIMA(0,1,0).

Από το μενού του PASW 18 Analyze-> Forecasting-> Create Models-> Plots και κλικάρουμε τις επιλογές

Residuals autocorrelations function (ACF)

Residuals partial autocorrelation function (PACF) και στη συνέχεια ok

Παρατηρούμε ότι είναι και τα δύο μη μηδενικά οπότε δεν έχουμε λευκό θόρυβο.

Πίνακας { 1 } Autocorrelations

Autocorrelations

Series:portugal

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,964	,135	51,178	1	,000
2	,925	,133	99,244	2	,000
3	,881	,132	143,726	3	,000
4	,831	,131	184,164	4	,000
5	,777	,129	220,266	5	,000
6	,723	,128	252,169	6	,000
7	,660	,127	279,376	7	,000
8	,597	,125	302,139	8	,000
9	,534	,124	320,760	9	,000
10	,468	,122	335,390	10	,000
11	,403	,121	346,486	11	,000
12	,339	,119	354,547	12	,000
13	,274	,118	359,963	13	,000
14	,213	,116	363,317	14	,000
15	,158	,115	365,208	15	,000
16	,105	,113	366,069	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας {2} Autocorrelations

Autocorrelations

Series:portugal

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,059	,136	,186	1	,666
2	,109	,135	,844	2	,656
3	,385	,133	9,183	3	,027
4	-,029	,132	9,232	4	,056
5	,022	,130	9,260	5	,099
6	,077	,129	9,613	6	,142
7	,035	,128	9,689	7	,207
8	-,064	,126	9,947	8	,269
9	,058	,125	10,165	9	,337
10	,165	,123	11,951	10	,288
11	-,275	,122	17,053	11	,106
12	-,098	,120	17,726	12	,124
13	,031	,119	17,792	13	,166
14	-,310	,117	24,803	14	,037
15	-,006	,115	24,805	15	,053
16	-,006	,114	24,808	16	,073

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας

1:H₀: $p_1=0$ $p_1=0,666>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: $p_1\neq 0$

2:H₀: $p_2=0$ $p_2=0,656>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: $p_2\neq 0$

3:H₀: $p_3=0$ $p_3=0,027<\alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

V_s

H₁: $p_3\neq 0$

4:H₀: $p_4=0$ $p_4=0,056>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: $\rho_4 \neq 0$

.....

16:H₀: $\rho_{16}=0$ $\rho_{16}=0,073 > \alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

Vs

H₁: $\rho_{16} \neq 0$

Πίνακας {3} Autocorrelations

Autocorrelations

Series:portugal

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,091	,136	,452	1	,501
2	,122	,135	1,271	2	,530
3	,259	,133	5,060	3	,167
4	,121	,132	5,904	4	,206
5	-,057	,130	6,093	5	,297
6	,157	,129	7,572	6	,271
7	,061	,128	7,803	7	,350
8	,053	,126	7,982	8	,435
9	,154	,125	9,501	9	,392
10	,178	,123	11,588	10	,314
11	-,085	,122	12,075	11	,358
12	-,088	,120	12,608	12	,398
13	,113	,119	13,524	13	,408
14	-,262	,117	18,523	14	,184
15	,025	,115	18,570	15	,234
16	,025	,114	18,617	16	,289

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας

1:H₀: $\rho_1=0$ $\rho_1=0,501 > \alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

Vs

H₁: $\rho_1 \neq 0$

2:H₀: $p_2=0$ $p_2=0,530>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
V_s

H₁: $p_2\neq 0$

3:H₀: $p_3=0$ $p_3=0,167>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
V_s

H₁: $p_3\neq 0$

4:H₀: $p_4=0$ $p_4=0,206>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
V_s

H₁: $p_4\neq 0$

.....

16:H₀: $p_{16}=0$ $p_{16}=0,289>\alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
V_s

H₁: $p_{16}\neq 0$

Έλεγχο στατιστικής σημαντικότητας για το υπόδειγμα ARIMA(0,1,1)

H₀: $p_0=0$ $p_0=0,001<\alpha=0,05$ Απόρριψη της H₀ άρα ο εκτιμητής είναι στατιστικά σημαντικός

V_s

H₁: $p_0\neq 0$

H₀: $p_1=0$ $p_1=0,728<\alpha=0,05$ Αποδοχή της H₀ άρα ο εκτιμητής είναι στατιστικά μη σημαντικός

V_s

H₁: $p_1\neq 0$

Έλεγχο στατιστικής σημαντικότητας για το υπόδειγμα ARIMA(0,1,0)

H₀: $p_0=0$ $p_0=0,000<\alpha=0,05$ Απόρριψη της H₀ άρα ο εκτιμητής είναι στατιστικά σημαντικός

V_s

H₁: $p_0\neq 0$

ΙΣΠΑΝΙΑ

Από το μενού κάνουμε τις εξής επιλογές: Data-> Define Dates-> Years και στην επιλογή Years γράφουμε 1960

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή poland και στην time axis lab την date και πατάμε ok.

Από τον πίνακα που προκύπτει φαίνεται ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα p είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς διαφορές. Έτσι θα δοκιμασουμε με πρώτες διαφορές. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία δεν είναι στάσιμη με πρώτες διαφορές. Οπότε παίρνουμε πρώτες διαφορές και λογάριθμο μαζί. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και επιλέγουμε και το natural log transform.

Φαίνεται ότι η διαδικασία δεν είναι στάσιμη με πρώτες διαφορές και λογάριθμο μαζί αφού όλα σχεδόν τα p-value είναι μηδέν.. Οπότε παίρνουμε δεύτερες διαφορές. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 2.

Βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία δεν είναι στάσιμη με δεύτερες διαφορές. Οπότε παίρνουμε δεύτερες διαφορές και λογάριθμο μαζί. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 2 και επιλέγουμε και το natural log transform.

Παρατηρούμε ότι συνεχίζει να μην είναι στάσιμη ακόμη και με δεύτερες διαφορές και λογάριθμο μαζί.

Συμπέρασμα λοιπόν ότι η χρονοσειρά αυτή δεν είναι στάσιμη.

Πίνακας {1} Autocorrelations

Autocorrelations

Series:spain

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,956	,135	50,361	1	,000
2	,907	,133	96,575	2	,000
3	,855	,132	138,444	3	,000
4	,791	,131	175,049	4	,000
5	,723	,129	206,318	5	,000
6	,655	,128	232,500	6	,000
7	,586	,127	253,898	7	,000
8	,518	,125	271,044	8	,000
9	,456	,124	284,595	9	,000
10	,395	,122	295,032	10	,000
11	,337	,121	302,808	11	,000
12	,283	,119	308,414	12	,000
13	,232	,118	312,282	13	,000
14	,187	,116	314,861	14	,000
15	,147	,115	316,494	15	,000
16	,109	,113	317,426	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας {2} Autocorrelations

Autocorrelations

Series:spain

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,336	,136	6,109	1	,013
2	,228	,135	8,976	2	,011
3	,249	,133	12,481	3	,006
4	-,136	,132	13,544	4	,009
5	-,053	,130	13,709	5	,018
6	-,098	,129	14,288	6	,027
7	-,183	,128	16,356	7	,022
8	-,142	,126	17,622	8	,024
9	-,168	,125	19,431	9	,022
10	-,096	,123	20,042	10	,029
11	-,169	,122	21,975	11	,025
12	-,149	,120	23,521	12	,024
13	,048	,119	23,687	13	,034
14	-,080	,117	24,158	14	,044
15	,044	,115	24,300	15	,060
16	,211	,114	27,729	16	,034

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον τα p-value είναι διάφορα του μηδενός κάνουμε έλεγχο στασιμότητας.

1:H₀: p₁=0 p₁=0,013<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)
V_s

H₁: p₁≠0

2:H₀: p₂=0 p₂=0,011<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)
V_s

V_s

H₁: p₂≠0

3:H₀: p₃=0 p₃=0,006<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)
V_s

V_s

H₁: p₃≠0

4:H₀: p₄=0 p₄=0,009<α=0,05 Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)
V_s

V_s

H₁: $\rho_4 \neq 0$

.....

16:H₀: $\rho_{16}=0$ $\rho_{16}=0,034 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $\rho_{16} \neq 0$

Πίνακας {3} Autocorrelations

Autocorrelations

Series:spain

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,468	,136	11,844	1	,001
2	,490	,135	25,065	2	,000
3	,441	,133	35,992	3	,000
4	,182	,132	37,899	4	,000
5	,304	,130	43,342	5	,000
6	,179	,129	45,262	6	,000
7	,113	,128	46,050	7	,000
8	,042	,126	46,163	8	,000
9	,021	,125	46,190	9	,000
10	,034	,123	46,268	10	,000
11	-,041	,122	46,382	11	,000
12	-,036	,120	46,470	12	,000
13	-,034	,119	46,553	13	,000
14	-,047	,117	46,712	14	,000
15	-,034	,115	46,800	15	,000
16	,084	,114	47,339	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας {4} Autocorrelations

Autocorrelations

Series:spain

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,416	,137	9,173	1	,002
2	-,172	,136	10,771	2	,005
3	,356	,134	17,763	3	,000
4	-,323	,133	23,664	4	,000
5	,088	,132	24,113	5	,000
6	,049	,130	24,255	6	,000
7	-,079	,129	24,630	7	,001
8	,040	,127	24,730	8	,002
9	-,077	,126	25,104	9	,003
10	,110	,124	25,887	10	,004
11	-,060	,122	26,128	11	,006
12	-,132	,121	27,320	12	,007
13	,247	,119	31,624	13	,003
14	-,194	,118	34,330	14	,002
15	-,075	,116	34,743	15	,003
16	,378	,114	45,660	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον τα p-value είναι διάφορα του μηδενός κάνουμε έλεγχο στασιμότητας.

1:H₀: $p_1=0$ $p_1=0,002 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $p_1 \neq 0$ 2:H₀: $p_2=0$ $p_2=0,005 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $p_2 \neq 0$ 3:H₀: $p_3=0$ $p_3=0,000 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

H₁: $p_3 \neq 0$ 4:H₀: $p_4=0$ $p_4=0,000 < \alpha=0,05$ Απόρριψη H₀ (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

$H_1: \rho_4 \neq 0$

.....

16: $H_0: \rho_{16}=0$ $\rho_{16}=0,000 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

$H_1: \rho_{16} \neq 0$

Πίνακας {5} Autocorrelations

Autocorrelations

Series:spain

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,535	,137	15,195	1	,000
2	,030	,136	15,245	2	,000
3	,217	,134	17,855	3	,000
4	-,337	,133	24,283	4	,000
5	,221	,132	27,113	5	,000
6	-,051	,130	27,265	6	,000
7	,013	,129	27,276	7	,000
8	-,042	,127	27,383	8	,001
9	-,039	,126	27,478	9	,001
10	,076	,124	27,850	10	,002
11	-,056	,122	28,056	11	,003
12	-,014	,121	28,069	12	,005
13	,024	,119	28,110	13	,009
14	-,009	,118	28,115	14	,014
15	-,150	,116	29,785	15	,013
16	,309	,114	37,099	16	,002

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον τα p-value είναι διάφορα του μηδενός κάνουμε έλεγχο στασιμότητας.

1: $H_0: \rho_1=0$ $\rho_1=0,000 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

$H_1: \rho_1 \neq 0$

2: $H_0: \rho_2=0$ $\rho_2=0,000 < \alpha=0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη στάσιμη)

Vs

$H_1: \rho_2 \neq 0$
 3: $H_0: \rho_3 = 0$ $p_3 = 0,000 < \alpha = 0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)
 V_S
 $H_1: \rho_3 \neq 0$
 4: $H_0: \rho_4 = 0$ $p_4 = 0,000 < \alpha = 0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)
 V_S
 $H_1: \rho_4 \neq 0$

 16: $H_0: \rho_{16} = 0$ $p_{16} = 0,002 < \alpha = 0,05$ Απόρριψη H_0 (υπάρχει αυτοσυσχέτιση, άρα μη
 στάσιμη)
 V_S
 $H_1: \rho_{16} \neq 0$

ΣΟΥΗΔΙΑ

Από το μενού κάνουμε τις εξής επιλογές: Data-> Define Dates-> Years και στην επιλογή Years γράφουμε 1960

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή Norway και στην time axis lab την date και πατάμε ok.

Από τον πίνακα που προκύπτει φαίνεται ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα ρ είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς διαφορές. Έτσι θα δοκιμασούμε με πρώτες διαφορές. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Οπότε βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία είναι στάσιμη με πρώτες διαφορές. Στη συνέχεια με τα διαγράμματα ACF και PACF που φαίνονται στο output εντοπίζονται τα q και τα p αντίστοιχα.

Από τα διαγράμματα ACF και PACF προκύπτει $q=0$ και $p=0$. Το υπόδειγμα είναι της μορφής ARIMA(p,d,q)-ARIMA(0,1,0), ($d=1$ λόγω πρώτων διαφορών). Η θεωρητική μορφή του υποδείγματος είναι :

$$Y_t = \alpha_0 + \alpha_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Έπειτα πηγαίνοντας στο PASW 18 ακολουθούμε τις παρακάτω διαδικασίες. Από το μενού επιλέγουμε Analyze-> Forecasting-> create models. Στο Dependent Variable περνάμε την μεταβλητή sweden και στο method επιλέγουμε ARIMA. Πατώντας το criteria εμφανίζεται ο πίνακας arima criteria και στο nonseasonal μπαίνουν οι τιμές των p , d , q που έχουν εντοπιστεί 0, 1,0 αντίστοιχα. Πατώντας το continue και πηγαίνοντας στο statistics κλικάρουμε την επιλογή parameter estimates και στη συνέχεια ok.

Στη συνέχεια στο output βγαίνει ο πίνακας Arima Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί κάνοντας έλεγχο στατιστικής σημαντικότητας για το υπόδειγμα ARIMA(0,1,0).

$H_0: \rho_0=0$ $\rho_0=0,048 < \alpha=0,05$ Απόρριψη της H_0 άρα ο εκτιμητής είναι στατιστικά σημαντικός

V_s

$H_1: \rho_0 \neq 0$

Αφού βρέθηκε ο εκτιμητής μας στατιστικά σημαντικός και γνωρίζοντας τις παραμέτρους του υποδείγματος ARIMA(0,1,0) θεωρείται ως το καταλληλότερο. Η μαθηματική μορφή του υποδείγματος θα είναι :

$$Y_t = 572,562 + E_t$$

Για να είναι το υπόδειγμα Y_t αποδεκτό, για να θεωρηθεί δηλαδή ότι περιγράφει την διαδικασία που παρήγαγε τα δεδομένα, θα πρέπει τα κατάλοιπα να είναι λευκός θόρυβος, δηλαδή δεν θα πρέπει να αυτοσυσχετίζονται. Για το λόγο αυτό θα γίνει έλεγχος στασιμότητας στα κατάλοιπα (residuals) επαληθεύοντας έτσι αν τα αποτελέσματα είναι αληθή για το ARIMA(0,1,0).

Από το μενού του PASW 18 Analyze-> Forecasting-> Create Models-> Plots και κλικάρουμε τις επιλογές

Residuals autocorrelations function (ACF)

Residuals partial autocorrelation function (PACF) και στη συνέχεια ok

Παρατηρούμε ότι είναι και τα δύο μηδενικά οπότε έχουμε λευκό θόρυβο.

Πηγαίνοντας πάλι στο PASW 18 και στα δεδομένα από τις επιλογές Analyze-> Forecasting-> Create Models και στην επιλογή Save κλικάρουμε το Noise Residuals και στη συνέχεια ok.

Πατώντας ok εμφανίζεται η μεταβλητή Noise Residuals France στο PASW 18 που δείχνει τα υπόλοιπα.

Από το μενού του PASW 18 επιλέγουμε Analyze-> Descriptive statistics-> Q-Qplots και στο Variable περνάμε την μεταβλητή Noise Residual from sweden και πατάμε ok.

Από το διάγραμμα του output παρατηρούμε ότι ακολουθεί κανονική κατανομή, αφού όλες οι παρατηρήσεις είναι κοντά στην γραμμή.

Επιπλέον θα κάνουμε έλεγχο στασιμότητας και για τα Residuals. Από το μενού του PASW 18 επιλέγουμε Analyze-> Forecasting-> Autocorrelation και περνάμε την μεταβλητή noise residuals from sweden και ok.

Άρα από τον έλεγχο καταλοίπων φαίνεται ότι η διαδικασία είναι στάσιμη καθώς για όλες τις χρονολογικές υστερήσεις τα p είναι στάσιμα. Έτσι το υπόδειγμα ARIMA(0,1,0) επιβεβαιώνεται ότι είναι το πλέον κατάλληλο.

Πίνακας { 1 } Autocorrelations

Autocorrelations

Series:sweden

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,915	,135	46,087	1	,000
2	,835	,133	85,225	2	,000
3	,766	,132	118,867	3	,000
4	,686	,131	146,402	4	,000
5	,618	,129	169,210	5	,000
6	,559	,128	188,286	6	,000
7	,496	,127	203,660	7	,000
8	,431	,125	215,513	8	,000
9	,378	,124	224,857	9	,000
10	,325	,122	231,908	10	,000
11	,285	,121	237,482	11	,000
12	,241	,119	241,550	12	,000
13	,196	,118	244,303	13	,000
14	,158	,116	246,143	14	,000
15	,116	,115	247,159	15	,000
16	,079	,113	247,650	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας {2} Autocorrelations

Autocorrelations

Series:sweden

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,915	,135	46,087	1	,000
2	,835	,133	85,225	2	,000
3	,766	,132	118,867	3	,000
4	,686	,131	146,402	4	,000
5	,618	,129	169,210	5	,000
6	,559	,128	188,286	6	,000
7	,496	,127	203,660	7	,000
8	,431	,125	215,513	8	,000
9	,378	,124	224,857	9	,000
10	,325	,122	231,908	10	,000
11	,285	,121	237,482	11	,000
12	,241	,119	241,550	12	,000
13	,196	,118	244,303	13	,000
14	,158	,116	246,143	14	,000
15	,116	,115	247,159	15	,000
16	,079	,113	247,650	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον τα p-value είναι διάφορα του μηδενός κάνουμε έλεγχο στασιμότητας.

1:H₀: p₁=0 p₁=0,109>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: p₁≠0

2:H₀: p₂=0 p₂=0,274>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: p₂≠0

3:H₀: p₃=0 p₃=0,185>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: p₃≠0

4:H₀: p₄=0 p₄=0,160>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: $\rho_4 \neq 0$

.....
16:H₀: $\rho_{16}=0$ $\rho_{16}=0,472 > \alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: $\rho_{16} \neq 0$

Πίνακας {3} Autocorrelations

Autocorrelations

Series:Noise residual from sweden-Model_1

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,218	,136	2,575	1	,109
2	,017	,135	2,590	2	,274
3	,199	,133	4,828	3	,185
4	-,175	,132	6,581	4	,160
5	-,122	,130	7,454	5	,189
6	,158	,129	8,946	6	,177
7	-,037	,128	9,031	7	,250
8	,061	,126	9,266	8	,320
9	,200	,125	11,830	9	,223
10	,049	,123	11,989	10	,286
11	-,009	,122	11,995	11	,364
12	-,054	,120	12,197	12	,430
13	-,028	,119	12,252	13	,507
14	-,155	,117	14,008	14	,449
15	,012	,115	14,019	15	,524
16	,149	,114	15,727	16	,472

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Έλεγχος στατιστικής σημαντικότητας p-value των εκτιμητών του υποδείγματος

1:H₀: $\rho_1=0$ $\rho_1=0,109 > \alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: $\rho_1 \neq 0$

2:H₀: $\rho_2=0$ $\rho_2=0,274 > \alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: $\rho_2 \neq 0$

3:H₀: $\rho_3=0$ $\rho_3=0,185 > \alpha=0,05$ Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s
 $H_1: \rho_3 \neq 0$
 $4: H_0: \rho_4 = 0$ $\rho_4 = 0,160 > \alpha = 0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
 V_s
 $H_1: \rho_4 \neq 0$

 $16: H_0: \rho_{16} = 0$ $\rho_{16} = 0,472 > \alpha = 0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)
 V_s
 $H_1: \rho_{16} \neq 0$

ΗΝΩΜΕΝΟ ΒΑΣΙΛΕΙΟ

Από το μενού κάνουμε τις εξής επιλογές: Data-> Define Dates-> Years και στην επιλογή Years γράφουμε 1960.

Από το μενού επιλέγουμε Analyze-> Forecasting-> Sequences Charts και στην επιλογή Variable βάζουμε την μεταβλητή uk και στην time axis lab την date και πατάμε ok.

Από τον πίνακα που προκύπτει ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα ρ είναι μηδέν έχοντας ως αποτέλεσμα να μην είναι στάσιμη χωρίς διαφορές. Έτσι θα δοκιμασούμε με πρώτες διαφορές. Από τις επιλογές Analyze-> Forecasting-> Autocorrelation κλικάρουμε την επιλογή Difference και γράφουμε τον αριθμό 1 και ok.

Οπότε βάση του ελέγχου κανονικότητας φαίνεται ότι η διαδικασία είναι στάσιμη με πρώτες διαφορές. Στη συνέχεια με τα διαγράμματα ACF και PACF που φαίνονται στο output εντοπίζονται τα q και τα p αντίστοιχα.

Από τα διαγράμματα ACF και PACF προκύπτει $q=1$ και $p=1$. Το υπόδειγμα είναι της μορφής ARIMA(p,d,q)-ARIMA(1,1,1), ($d=1$ λόγω πρώτων διαφορών). Η θεωρητική μορφή του υποδείγματος είναι :

$$Y_t = \alpha_0 + \alpha_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Έπειτα πηγαίνοντας στο PASW 18 ακολουθούμε τις παρακάτω διαδικασίες. Από το μενού επιλέγουμε Analyze-> Forecasting-> create models. Στο Dependent Variable περνάμε την μεταβλητή uk και στο method επιλέγουμε ARIMA. Πατώντας το criteria εμφανίζεται ο πίνακας arima criteria και στο nonseasonal μπαίνουν οι τιμές των p, d, q που έχουν εντοπιστεί 1, 1, 1, αντίστοιχα. Πατώντας το continue και πηγαίνοντας στο statistics κλικάρουμε την επιλογή parameter estimates και στη συνέχεια ok.

Στη συνέχεια στο output βγαίνει ο πίνακας Arima Models Parameters από τον οποίο μελετώνται οι παράμετροι που είναι στατιστικά σημαντικοί κάνοντας έλεγχο στατιστικής σημαντικότητας για το υπόδειγμα ARIMA(1,1,1).

Αφού βρέθηκαν οι εκτιμητές μας στατιστικά μη σημαντικοί αφαιρώ τις στατιστικά μη σημαντικές παραμέτρους, μια μια την φορά, ξεκινώντας από αυτή με το μεγαλύτερο p-value. Δηλαδή ξαναπροσαρμόζω το μοντέλο βάζοντας μηδέν στην παράμετρο που πρέπει (p ή d ή q).

Αφού βρέθηκαν ξανά οι εκτιμητές μας στατιστικά μη σημαντικοί αφαιρώ τις στατιστικά μη σημαντικές παραμέτρους, μια μια την φορά, ξεκινώντας από αυτή με το μεγαλύτερο p-value. Δηλαδή ξαναπροσαρμόζω το μοντέλο βάζοντας μηδέν στην παράμετρο που πρέπει (p ή d ή q).

Αφού βρέθηκαν ξανά οι εκτιμητές μας στατιστικά μη σημαντικοί και συγκεκριμένα η σταθερά, ξαναπροσαρμόζω και στα κριτήρια ξεκλικάρω το include constant in model και continue.

Δεν υπάρχει κατί να εκτιμήσουμε.

Πίνακας{1} -> Πίνακας Autocorrelations στο output

Autocorrelations

Series:uk

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,866	,135	41,264	1	,000
2	,742	,133	72,218	2	,000
3	,607	,132	93,309	3	,000
4	,490	,131	107,364	4	,000
5	,389	,129	116,382	5	,000
6	,310	,128	122,248	6	,000
7	,214	,127	125,093	7	,000
8	,115	,125	125,942	8	,000
9	,024	,124	125,981	9	,000
10	-,056	,122	126,191	10	,000
11	-,096	,121	126,825	11	,000
12	-,094	,119	127,445	12	,000
13	-,084	,118	127,955	13	,000
14	-,027	,116	128,010	14	,000
15	-,014	,115	128,025	15	,000
16	-,016	,113	128,045	16	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας{2} Autocorrelations

Autocorrelations

Series:uk

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-,005	,136	,001	1	,970
2	,181	,135	1,814	2	,404
3	,086	,133	2,226	3	,527
4	-,066	,132	2,478	4	,649
5	,038	,130	2,562	5	,767
6	,149	,129	3,893	6	,691
7	,070	,128	4,196	7	,757
8	,032	,126	4,261	8	,833
9	,095	,125	4,837	9	,848
10	-,106	,123	5,578	10	,849
11	-,267	,122	10,397	11	,495
12	-,038	,120	10,496	12	,573
13	-,201	,119	13,362	13	,420
14	,141	,117	14,805	14	,392
15	-,096	,115	15,491	15	,417
16	-,092	,114	16,146	16	,443

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Εφόσον οι τιμές είναι μη μηδενικές κάνουμε έλεγχο στασιμότητας

1:H₀: p₁=0 p₁=0,970>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: p₁≠0

2:H₀: p₂=0 p₂=0,404>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: p₂≠0

3:H₀: p₃=0 p₃=0,527>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

H₁: p₃≠0

4:H₀: p₄=0 p₄=0,649>α=0,05 Αποδοχή H₀ (δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη)

V_s

$H_1: \rho_4 \neq 0$

.....

16: $H_0: \rho_{16} = 0$
στάσιμη)

V_s

$H_1: \rho_{16} \neq 0$

$p_{16} = 0,443 > \alpha = 0,05$ Αποδοχή H_0 (δεν υπάρχει αυτοσυσχέτιση, άρα

Βιβλιογραφία

Βιβλιογραφία

- [1] Καρβέλης Χαράλαμπος. (2008). «Ανάλυση χρονολογικών σειρών: ‘προβλέποντας το μέλλον, κατανοώντας το παρελθόν’». *Διπλωματική Εργασία στο Διατμηματικό Πρόγραμμα Μεταπτυχιακών σπουδών των τμημάτων Μαθηματικών και Μηχανικών Η/Υ & Πληροφορικής Πανεπιστημίου Πατρών: ‘Τα μαθηματικά των υπολογιστών και των αποφάσεων’*
- [2] Ζάρλα Αλεξάνδρα. «Ανάλυση χρονολογικών σειρών». Διπλωματική Εργασία *Διπλωματική Εργασία στο Διατμηματικό Πρόγραμμα Μεταπτυχιακών σπουδών των τμημάτων Μαθηματικών και Μηχανικών Η/Υ & Πληροφορικής Πανεπιστημίου Πατρών : ‘Τα μαθηματικά των υπολογιστών και των αποφάσεων’.*
- [3] Γεωργίου Παναγιώτης. (2008). «Οικονομετρική διερεύνηση της σχέσης συναλλαγών θεσμικών επενδυτών και χρηματιστηριακών αποδόσεων». *Διπλωματική Διατριβή στο Τμήμα Οικονομικών Επιστημών Πρόγραμμα Μεταπτυχιακών Σπουδών στη Χρηματοοικονομική Ανάλυση και Διαχείριση Πανεπιστημίου Πατρών*
- [4] Σαμάνη Μαρί-Άννα. (2012). «Μοναδιαίες Ρίζες και μέθοδοι συνολοκλήρωσης (Στη μακροοικονομία)». *Διατμηματικό πρόγραμμα μεταπτυχιακών σπουδών στην οικονομική επιστήμη Πανεπιστημίο Μακεδονίας*
- [5] Αντωνόπουλος Π. Γρηγόριος. (2008). «Ανάλυση μοντέλων χρονολογικών σειρών». *Διπλωματική Εργασία στο Διατμηματικό Πρόγραμμα Μεταπτυχιακών σπουδών των τμημάτων Μαθηματικών και Μηχανικών Η/Υ & Πληροφορικής Πανεπιστημίου Πατρών: ‘Τα μαθηματικά των υπολογιστών και των αποφάσεων’*
- [6] Pedeli Xanthi. (2011). “Modelling multivariate time series for count data”. *Thesis submitted to the Department of Statistics of the Athens University of Economics and Business in partial fulfillment of the requirements for the degree of PhD in Statistics.*