

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΔΥΤΙΚΗΣ ΕΛΛΑΔΟΣ

ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ

ΤΜΗΜΑ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ



Τίτλος Εργασίας: "Ανάλυση και εφαρμογές της μεθοδολογίας BOX – JENKINS"

Πτυχιακή Εργασία των

Φωστηρόπουλος Ευάγγελος

Στογιάννος Νικόλαος

Επιβλέπων Καθηγητής

Παπαδόπουλος Δημήτριος

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

Περίληψη.....	5
Abstract.....	5
ΚΕΦΑΛΑΙΟ 1 ^ο : Εισαγωγή	6
ΚΕΦΑΛΑΙΟ 2 ^ο : Βασικές έννοιες Χρονολογικών Σειρών	8
2.1 Στοχαστική διαδικασία και Χρονολογικές Σειρές	8
2.1.1 Ορισμός Στοχαστικής διαδικασίας:.....	8
2.1.2 Ορισμός Χρονολογικών σειρών:.....	8
2.2 Διακριτές – Συνεχείς Χρονολογικές Σειρές.....	8
2.3 Μέτρα Χρονολογικών Σειρών	9
2.4 Συνιστώσες Χρονολογικών Σειρών.....	10
2.4.1 Μακροχρόνια τάση	11
2.4.2 Απομόνωση της τάσης.....	12
2.4.3 Εποχιακή Περιοδικότητα.....	12
2.4.4 Κυκλική Περιοδικότητα.....	13
2.4.5 Τυχαία Συνιστώσα	16
2.5 Στασιμότητα και Αυτοσυσχέτιση	16
2.6 Δειγματικοί συντελεστές αυτοσυσχέτισης.....	19
2.7 Διάγραμμα Αυτοσυσχέτισης	20
2.8 Έλεγχος Στασιμότητας	22
2.9 Μερική Αυτοσυσχέτιση.....	23
2.10 Τυχαία χρονολογική σειρά – Λευκός θόρυβος – iid Διαδικασία.....	24
2.10.1 Λευκός Θόρυβος.....	24
2.10.2 iid Διαδικασία.....	25
2.10.3 Υπόδειγμα τυχαίας διαδρομής λευκού θορύβου με σταθερά.....	25
ΚΕΦΑΛΑΙΟ 3 ^ο : Στάσιμα Υποδείγματα Χρονολογικών Σειρών.....	30
3.1 Αυτοπαλίνδρομο υποδείγματα – AR(p)	30
3.2 Αυτοπαλίνδρομο υπόδειγμα πρώτης τάξεως(Autoregressive of order 1)	31
3.3 Αυτοπαλίνδρομο υπόδειγμα δεύτερης τάξεως AR(2).....	35

3.4 Αυτοπαλίνδρομο υπόδειγμα $AR(p)$	41
3.5 Εξισώσεις Yule-Walker και συναρτήσεις αυτοσυσχέτιση.....	43
3.6 Υποδείγματα Κινητών Μέσων MA (Moving Average Models).....	44
3.7 Υπόδειγμα Κινητών Μέσων 1 ^{ης} Τάξης $-MA(1)$ (Moving Average of Order 1).....	45
3.8 Υπόδειγμα Κινητών Μέσων 2 ^{ης} Τάξης $-MA(2)$ (Moving Average of Order 2).....	47
3.9 Υποδείγματα Κινητών Μέσων Τάξης q $-MA(q)$	48
3.10 Αντιστρεψιμότητα	49
3.11 Σύγκριση AR και MA υποδειγμάτων	50
3.12 Μικτά Υποδείγματα $ARMA(p,q)$	50
3.13 Συμπεράσματα για τα υποδείγματα $ARMA(p,q)$	55
ΚΕΦΑΛΑΙΟ 4^ο: Μεθοδολογία Box-Jenkins-Υποδείγματα $ARIMA$	57
4.1 Μετατροπή μη στάσιμης διαδικασίας σε στάσιμη με Λήψη Διαφορών – Διαφοροποίηση (Differencing)	58
4.2 Εναλλακτικές Μορφές Υποδείγματος $ARIMA$	60
4.3 Παραδείγματα Υποδειγμάτων $ARIMA$	60
4.4 Εποχικά Υποδείγματα $SARIMA$	63
4.5 Μεθοδολογία Box – Jenkins.....	64
4.5.1 Ταυτοποίηση	64
4.5.2 Εκτίμηση	65
4.5.3 Διαγνωστικός Έλεγχος	66
4.5.3.1 Διαγνωστικός Έλεγχος – Υπερπροσαρμογή – Κριτήριο Akaike – Μπαϊεσιανό Κριτήριο.....	68
4.5.4 Προβλέψεις με AR , MA και $ARMA$ Υποδείγματα	69
4.5.4.1 Πρόβλεψη Διαδικασίας $MA(q)$	70
4.5.4.2 Πρόβλεψη Διαδικασίας $AR(p)$	71
4.5.4.3 Πρόβλεψη Διαδικασίας $ARMA(p, q)$	72
4.6 Αξιολόγηση προβλέψεων	74
ΚΕΦΑΛΑΙΟ 5: Η εφαρμογή της μεθοδολογίας των Box – Jenkins για τη μελέτη των γεννήσεων στην Ελλάδα.	77
ΒΙΒΛΙΟΓΡΑΦΙΑ	97

Περίληψη

Η μελέτη αυτή ασχολείται με τη μεθοδολογία Box – Jenkins, η οποία στα πλαίσια της ανάλυσης χρονολογικών σειρών συμβάλλει στην επιλογή του υποδείγματος που περιγράφει καλύτερα τη διαδικασία που παρήγαγε τα δεδομένα μας. Στα πρώτα κεφάλαια γίνεται μια εισαγωγή στις χρονολογικές σειρές, στη χρησιμότητά τους και τι μπορούν να περιγράψουν. Αναλύονται βασικές τους έννοιες όπως μέτρα, συνιστώσες, υποδείγματα και στασιμότητα, και αναλύονται τα χαρακτηριστικά τους. Έπειτα αναλύονται συγκεκριμένες κατηγορίες των χρονολογικών σειρών, όπως είναι οι στάσιμες και οι μη στάσιμες χρονοσειρές και στο τέλος παρουσιάζονται διάφορες μέθοδοι πρόβλεψης και γίνεται εφαρμογή των μεθόδων αυτών μέσω του στατιστικού προγράμματος PASW Statistics (ΣΤΑΤΙΣΤΙΚΟ ΠΡΟΓΡΑΜΜΑ).

Abstract

This study deals with Box – Jenkins methodology, which within the context of time series analysis, contribute to the assessment of which model describes better the process that generated our data. In the first chapters an introduction of time series analysis is presented, in order to understand their usage and what they can describe. Key concepts of time series are analyzed such as measures, models, stationary and various others analysis characteristics. In addition, specific categories of times series are analyzed, like stationary and non-stationary time series. In final chapter, various forecast methods are presented and then these methods are applied in practice through the computer statistics program PASW Statistics(ΣΤΑΤΙΣΤΙΚΟ ΠΡΟΓΡΑΜΜΑ).

Κεφάλαιο 1^ο: Εισαγωγή

Κύριο αντικείμενο της ανάλυσης χρονολογικών σειρών είναι η μελέτη παρελθοντικών δεδομένων ώστε να προβλέψει τις μελλοντικές τους τιμές. Παράλληλα, οι προβλέψεις πρέπει να είναι έγκυρες και ποιοτικές. Η ποιότητα και εγκυρότητα τους εξαρτάται κυρίως από τον τρόπο συλλογής και ανάλυσης των πληροφοριών. Η ανάγκη λοιπόν για έγκυρες προβλέψεις οδήγησε στη δημιουργία τεχνικών και μεθόδων συλλογής πληροφοριών. Οι μέθοδοι αυτοί διακρίνονται σε αντικειμενικές ή ποσοτικές (objective or quantitative) και υποκειμενικές ή ποιοτικές μεθόδους (subjective or qualitative). Εμάς θα μας απασχολήσουν οι ποσοτικές μέθοδοι, όπου η ανάλυση των δεδομένων γίνεται με τη βοήθεια στατιστικών μεθόδων. Οι προβλέψεις με χρονολογικές σειρές αποδείχτηκαν πολλές φορές ανώτερες από παλαιότερες μεθόδους. Προβλέψεις με χρονολογικές σειρές γίνονται σε ένα μεγάλο εύρος επιστημών όπως η ιατρική, η φυσική, η μετεωρολογία, η γεωφυσική, η οικονομετρία και η μηχανική. Είναι απαραίτητο να αναφερθεί πως οι χρονολογικές σειρές εκτός από πρόβλεψη μελλοντικών τιμών μεταβλητών, είναι πολύ χρήσιμες και για τη μελέτη και κατανόηση του φαινομένου στην πορεία του στο χρόνο. Μέσω της σωστής και καλής καταγραφής των δεδομένων οδηγούμαστε και στην καλύτερη εκτίμηση του μέλλοντος.

Στόχος της ανάλυσης χρονολογικών σειρών είναι με την ένωση δεδομένων που έχουν συλλεχτεί στο χρόνο, να δημιουργηθεί ένα χρονοδιάγραμμα, η μελέτη του οποίου θα μας δώσει μια γενική εικόνα της εξέλιξης του φαινομένου διαχρονικά. Η επιστήμη αυτή χρησιμοποιείται για να καθοριστούν μοντέλα που μετατρέπουν πληροφορίες από χρονικά διαστήματα σε στατιστικά μέτρα. Κύριες μέθοδοι των χρονοσειρών είναι η μέθοδος της φασματικής ανάλυσης (spectral analysis), όπου η χρονολογική σειρά αναπαρίσταται με ένα κινητικό μοντέλο που οι παρατηρήσεις του θεωρούνται συναρτήσεις του χρόνου, και η μέθοδος της αυτοσυσχέτισης, όπου η χρονολογική σειρά αναπαριστάται με ένα δυναμικό μοντέλο (μοντέλο ARIMA, διαδικασία Box-Jenkins) όπου οι παρατηρήσεις θεωρούνται ως συναρτήσεις του παρελθόντος τους.

Συμπερασματικά, μπορούμε να πούμε ότι σκοπός της ανάλυσης χρονολογικών σειρών είναι η περιγραφή, η εξήγηση, η πρόβλεψη, ο σχεδιασμός, ο έλεγχος και η κατανόηση μιας διαδικασίας

Με τον όρο χρονοσειρά ορίζεται ένα σύνολο παρατηρήσεων που συλλέγονται διαδοχικά στο χρόνο. Αν οι παρατηρήσεις σημειώνονται σε κάθε σημείο του χρόνου, πρόκειται για μια συνεχή (continuous) χρονολογική σειρά, όπως η θερμοκρασία. Αν οι μετρήσεις γίνονται σε προκαθορισμένα διαστήματα χρόνου, τότε γίνεται λόγος για διακριτή (discrete) χρονολογική σειρά, όπως ο μισθός ενός υπαλλήλου, η τιμή μιας μετοχής την ώρα που κλείνει το χρηματιστήριο. Η μελέτη μιας χρονολογικής σειράς προϋποθέτει τη συλλογή ενός μεγάλου αριθμού παρατηρήσεων.

Μερικά παραδείγματα χρονολογικών σειρών είναι οι μηνιαίες πωλήσεις μιας επιχείρησης, η τιμή ενός προϊόντος ανά τρίμηνο, ο αριθμός γεννήσεων κάθε έτος και η ημερήσια αξία μιας μετοχής.

Συμπεραίνοντας, σκοπός της ανάλυσης χρονολογικών σειρών όπως αναφέρθηκε παραπάνω, είναι η μελέτη, η πρόβλεψη, ο έλεγχος και η κατανόηση μιας διαδικασίας και ο σκοπός της μεθοδολογίας Box – Jenkins είναι η εύρεση του κατάλληλου ARIMA υποδείγματος που περιγράφει τη διαδικασία αυτή.

ΚΕΦΑΛΑΙΟ 2^ο: Βασικές έννοιες Χρονολογικών Σειρών

2.1 Στοχαστική διαδικασία και Χρονολογικές Σειρές

2.1.1 Ορισμός Στοχαστικής διαδικασίας:

Μια χρονική σειρά είναι ένα δείγμα με ισαπέχοντα χρονικά διαστήματα (έτη, τρίμηνα, μήνες κ.ο.κ). Αν οι παρατηρήσεις είναι συγκεκριμένες τιμές των τυχαίων μεταβλητών X_1, X_2, \dots, X_n και οι τυχαίες αυτές μεταβλητές είναι υποσύνολο μιας άπειρης σειράς τυχαίων μεταβλητών (ακολουθία τυχαίων μεταβλητών), τότε λέμε ότι η άπειρη αυτή ακολουθία των τυχαίων μεταβλητών ονομάζεται στοχαστική διαδικασία (stochastic process) και παριστάνεται ως X_t^1 .

2.1.2 Ορισμός Χρονολογικών σειρών:

Με τον όρο χρονολογική σειρά εννοούμε μια σειρά συνεχών δεδομένων με κύρια χαρακτηριστικά την καθορισμένη διάταξη των παρατηρήσεων διαχρονικά, και την καθορισμένη εξάρτηση μεταξύ των διαδοχικών παρατηρήσεων της σειράς. Η καθορισμένη αυτή διάταξη και εξάρτηση εγγυάται τη δημιουργία αξιόπιστων προβλέψεων. Η διάταξη μιας παρατήρησης προσδιορίζεται από ένα δείκτη t . Έτσι με τον όρο Z_t ορίζουμε την t -στή παρατήρηση της σειράς. Η προηγούμενη παρατήρηση ορίζεται ως Z_{t-1} και η επόμενη ως Z_{t+1}^2 .

Οι χρονολογικές σειρές διακρίνονται σε συνεχείς χρονολογικές σειρές και σε διακριτές.

2.2 Διακριτές – Συνεχείς Χρονολογικές Σειρές

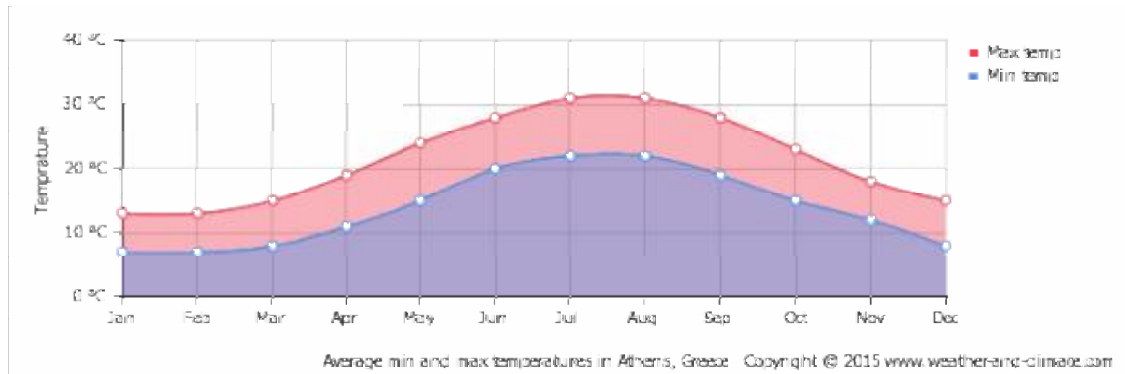
Συνεχείς (Continuous) χρονολογικές σειρές είναι αυτές όπου η τιμή του φαινομένου παρατηρείται σε οποιοδήποτε σημείο του χρόνου.

Διακριτές (Discrete) χρονολογικές σειρές είναι αυτές όπου η τιμή του φαινομένου καταγράφεται σε ορισμένα χρονικά διαστήματα. Ένα παράδειγμα μιας συνεχούς

¹ Θαλασσινός Ε.(1991, Κ.1 σ.17)

² Θαλασσινός Ε.(1991, Κ.1 σ.15)

χρονολογικής σειράς είναι η θερμοκρασία ενός τόπου ενώ αν τη θερμοκρασία τη μετράμε μια συγκεκριμένη ώρα κάθε μέρα πρόκειται για διακριτή σειρά.



Διάγραμμα 2.1 - Η θερμοκρασία στην Αθήνα για το 2015.

Διακριτές χρονολογικές σειρές μπορούν να δημιουργηθούν με διάφορους τρόπους. Από μια συνεχή χρονολογική σειρά, για παράδειγμα είναι δυνατό να δημιουργήσουμε μια διακριτή χρονολογική σειρά παίρνοντας παρατηρήσεις σε ισαπέχοντα χρονικά διαστήματα. Παράδειγμα αυτού του τύπου είναι η ημερήσια θερμοκρασία ενός τόπου στις 12:00 π.μ.

2.3 Μέτρα Χρονολογικών Σειρών

Σημαντικές έννοιες επίσης είναι η γραφική παράσταση που είναι η απεικόνιση της χρονολογικής σειράς, ο μέσος όρος, η διασπορά-διακύμανση και η τυπική απόκλιση της γραφικής παράστασης, ο κινητός μέσος όρος, η τάση και η στασιμότητα.³

Γραφική παράσταση της χρονολογικής σειράς εννοούμε την καμπύλη που δημιουργείται σε ένα σύστημα ορθογώνιων αξόνων όπου ο οριζόντιος άξονας συμβολίζει το χρόνο και ο κατακόρυφος τις παρατηρήσεις μας στο συγκεκριμένο χρονικό διάστημα. Η μελέτη της γραφικής αυτής παράστασης είναι χρήσιμη για τον προσδιορισμό βασικών χαρακτηριστικών της σειράς όπως η ύπαρξη τάσης, εποχικότητας και άλλων συνιστωσών που θα αναφερθούν παρακάτω αναλυτικότερα. Δημιουργούμε Ν ζεύγη που απεικονίζονται υπό τη μορφή (X_t, Y_t) όπου μπορούμε να παραστήσουμε σε αυτό το σύστημα αξόνων, συμβολίζοντας με X_t τις Ν χρονικές στιγμές και με Y_t τις τιμές των αντίστοιχων παρατηρήσεων.

³ Δημελή Σ.(2013, Κ.2 σ.27)

Μέσος όρος μ μιας χρονολογικής σειράς είναι η μέση τιμή όλων των τιμών της.

$$\mu = \frac{1}{N} \sum_{t=1}^N X(t_1) \quad (2.1)$$

Η διακύμανση ή διασπορά μιας τυχαίας μεταβλητής X συμβολίζεται συνήθως με $\text{Var}[x]$ και δηλώνει πόσο συγκεντρωμένες γύρω από τη μέση τιμή είναι οι τιμές της τυχαίας μεταβλητής.

$$\sigma^2 = \frac{1}{N-1} \sum_{t=1}^N (X(t_1) - \mu)^2 \quad (2.2)$$

Η θετική τετραγωνική ρίζα της διακύμανσης ονομάζεται τυπική απόκλιση και συμβολίζεται με σ .

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{N-1} \sum_{t=1}^N (X(t_1) - \mu)^2} \quad (2.3)$$

2.4 Συνιστώσες Χρονολογικών Σειρών

Η ανάλυση χρονολογικών σειρών αποσκοπεί στο διαχωρισμό αυτής της χρονολογικής σειράς στις επιμέρους συνιστώσες⁴ που περιλαμβάνουν:

- ∅ Τη μακροχρόνια τάση (long-run trend)
- ∅ Την εποχιακή περιοδικότητα (seasonal variation)
- ∅ Την κυκλική περιοδικότητα (cyclical variation)
- ∅ Την τυχαία περιοδικότητα (residual variation)

Ο διαχωρισμός της χρονολογικής σειράς στις επιμέρους συνιστώσες μπορεί να επιτευχθεί με δύο μη στοχαστικά υποδείγματα:

- ∅ Το πολλαπλασιαστικό υπόδειγμα (Multiplicative model), το οποίο εκφράζεται ως εξής:

$$Y_t = T_t C_t S_t R_t \quad (2.4)$$

- ∅ Το προσθετικό υπόδειγμα (Additive model) το οποίο εκφράζεται ως εξής:

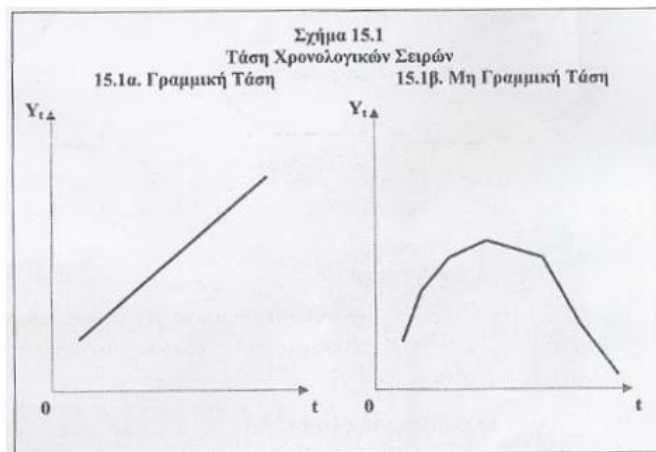
$$Y_t = T_t + C_t + S_t + R_t \quad (2.5)$$

⁴ Κουνετάς Κ.(2012, Κ.1 σ.12)

όπου T = μακροχρόνια τάση, C = κύκλική συνιστώσα, S = εποχιακή συνιστώσα, R = τυχαία συνιστώσα

2.4.1 Μακροχρόνια τάση

Η τάση μιας χρονολογικής σειράς αναφέρεται σε μια σχετικά σταθερή συμπεριφορά της χρονολογικής σειράς, η διάρκεια της οποίας είναι μεγαλύτερη του ενός έτους. Η τάση περιγράφει την καθαρή επιρροή μακροχρόνιων παραγόντων της χρονοσειράς από κυκλικές, εποχικές και τυχαίες επιδράσεις. Για παράδειγμα οι καταναλωτικές δαπάνες ενός καταναλωτή για ένα συγκεκριμένο αγαθό μπορεί να επηρεαστούν μακροχρόνια από παράγοντες όπως το εισόδημα, εποχικότητα του προϊόντος, καταναλωτικές συνήθειες. Όμως έχει αποδειχθεί εμπειρικά πως το εισόδημα είναι ο κυριότερος παράγοντας που επηρεάζει τις καταναλωτικές δαπάνες που μεταβάλλονται από περίοδο σε περίοδο με σταθερή αναλογία. Η καμπύλη που αναπαριστά την τάση μιας χρονικής σειράς δεν είναι απαραίτητα γραμμική. Για να μετρήσουμε την τάση ένας απλός τρόπος είναι το υπόδειγμα παλινδρόμησης το οποίο έχει εξαρτημένη μεταβλητή τις τιμές της χρονοσειράς και ανεξάρτητη μεταβλητή τη χρονική περίοδο.



Διάγραμμα 2.2 - Τάση Χρονολογικών Σειρών

Εμάς θα μας απασχολήσει η εξίσωση της γραμμικής τάσης (linear trend) η οποία έχει τη μορφή :

$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t \quad (2.6)$$

όπου $Y_t =$ η τάση της χρονολογικής σειράς για την περίοδο $t = (1, 2, \dots, N)$ και β_1 μετράει τη μέση μεταβολή της χρονοσειράς ανά μονάδα χρόνου.

2.4.2 Απομόνωση της τάσης

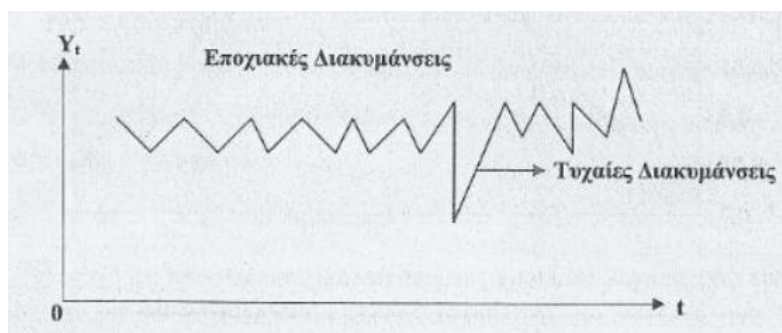
Για να εντοπιστούν οι υπόλοιπες συνιστώσες μιας χρονολογικής σειράς θα πρέπει να απομακρυνθεί η επίδραση της τάσης από τα στοιχεία του δείγματος. Στο πολλαπλασιαστικό υπόδειγμα η απομόνωση της τάσης επιτυγχάνεται διαιρώντας τις τιμές της χρονολογικής σειράς με τις αντίστοιχες της μακροχρόνιας τάσης. Προκύπτει ο εξής τύπος :

$$\left\{ \frac{Y_t}{T_t} \right\} = \left\{ \frac{T_t C_t S_t R_t}{T_t} \right\} = \{ C_t S_t R_t \} = Y^* \quad (2.7)$$

Όπου Y^* είναι οι τιμές της χρονολογικής σειράς απαλλαγμένης της τάσης.

2.4.3 Εποχιακή Περιοδικότητα

Η εποχιακή περιοδικότητα θεωρείται η πηγή των εμπορικών και οικονομικών χρονολογικών σειρών. Η βασική διαφορά που τη διακρίνει από την κυκλική συνιστώσα είναι η διάρκεια καθώς μπορεί να εμφανιστεί με διάρκεια μιας εβδομάδας, ενός μήνα, ή και ενός τριμήνου σε αντίθεση με την κυκλική περιοδικότητα που παρατηρείται σε χρονικά διαστήματα μεγαλύτερα του ενός έτους. Η εποχιακή συνιστώσα μετράται με τους εποχιακούς δείκτες (seasonal indices), οι οποίοι αναλύονται παρακάτω.



Διάγραμμα 2.3 - Χρονολογική σειρά που παρουσιάζει εποχιακές και τυχαίες διακυμάνσεις.

Οι εποχιακοί δείκτες μιας χρονολογικής σειράς δημιουργούνται με τον εξής τρόπο:

1^ο Στάδιο

Το πρώτο στάδιο αποτελείται από την εξομάλυνση της χρονοσειράς, η οποία επιτυγχάνεται με τον υπολογισμό των κινητών μέσων (moving averages), χαρακτηριστικό των οποίων είναι ότι στο πολλαπλασιαστικό υπόδειγμα απομονώνουν τις τυχαίες και τις εποχιακές επιδράσεις από τη χρονολογική σειρά.

2^ο Στάδιο

Στο δεύτερο στάδιο υπολογίζουμε το λόγο των τιμών της χρονολογικής σειράς και των κινητών μέσων.

$$\left\{ \frac{Y_t}{(KM)_t} \right\} = \left\{ \frac{T_t C_t S_t R_t}{T_t C_t} \right\} = \{S_t R_t\} \quad (2.8)$$

Όπου $(KM)_t = T_t C_t =$ κινητός μέσος. Ο κινητός μέσος μας δίνει την κυκλική συνιστώσα και τη συνιστώσα της τάσης.

3^ο Στάδιο

Στο τρίτο στάδιο για κάθε χρονική στιγμή υπολογίζουμε τους αριθμητικούς μέσους των λόγων $S_t R_t$. Η διαδικασία αυτή απομακρύνει τη χρονολογική σειρά από τυχαίες αποκλίσεις και αποτελεί ένα μέτρο εποχιακής συνιστώσας.

2.4.4 Κυκλική Περιοδικότητα

Η κυκλική περιοδικότητα υπολογίζει τις αποκλίσεις των τιμών γύρω από τη μακροχρόνια τάση της και έχει διάρκεια μεγαλύτερη του ενός έτους. Η απομόνωση και ο υπολογισμός της κυκλικής συνιστώσας είναι αρκετά περίπλοκη διαδικασία με εξαίρεση τους εμπορικούς κύκλους. Για να εντοπίσουμε και να υπολογίσουμε την κυκλική συνιστώσα, μπορούμε να χρησιμοποιήσουμε τη μέθοδο του ποσοστού της τάσης (percentage of trend). Το υπόδειγμα της χρονοσειράς στη γραμμική του μορφή υπολογίζεται με την μέθοδο ελάχιστων τετραγώνων (OLS).

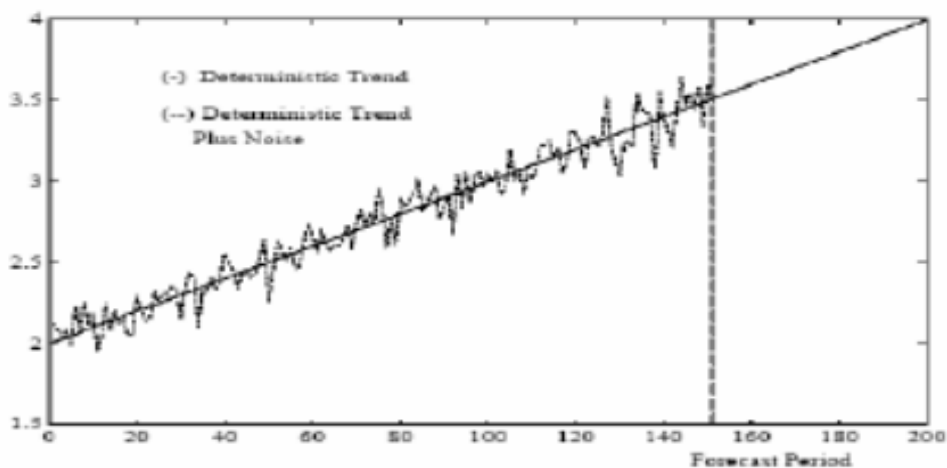
1^ο Στάδιο

Χρησιμοποιώντας τη γραμμή τάσης, υπολογίζουμε για κάθε χρονική περίοδο τις αντίστοιχες τιμές της.

2^ο Στάδιο

Υπολογίζουμε το λόγο των τιμών της χρονολογικής σειράς και των αντίστοιχων τιμών της τάσης. Στην περίπτωση που βρεθούν τυχαίες αποκλίσεις αυτής της διαίρεσης γύρω από τη μονάδα, αυτό οφείλεται στις τυχαίες και στις εποχιακές επιδράσεις στη χρονολογική σειρά.

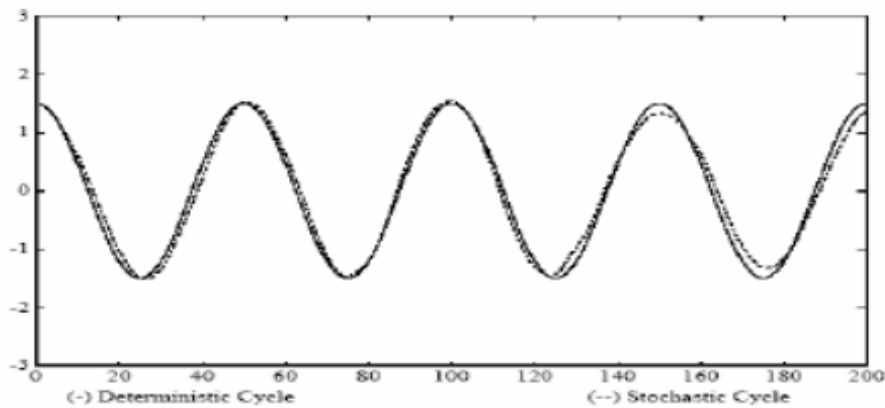
Η μακροχρόνια τάση συγκριτικά με τις υπόλοιπες συνιστώσες των χρονολογικών σειρών, είναι από τις περιπτώσεις που αντιμετωπίζονται πιο εύκολα γιατί η τάση εκφράζει τη χρονολογική σειρά για μια εκτεταμένη περίοδο. Στις χρονολογικές σειρές με μακροχρόνια τάση είναι εύκολη η κατανόηση της ιστορίας της μεταβλητής με αποτέλεσμα την πρόβλεψη μελλοντικών τιμών. Για να προσδιοριστεί η μακροχρόνια τάση, υπάρχουν ορισμένες μέθοδοι όπως η μέθοδος των δύο μέσων σημείων, η μέθοδος των κινητών μέσων, η μέθοδος της ευθείας ελάχιστων τετραγώνων, η μέθοδος της καμπύλης ελάχιστων τετραγώνων.



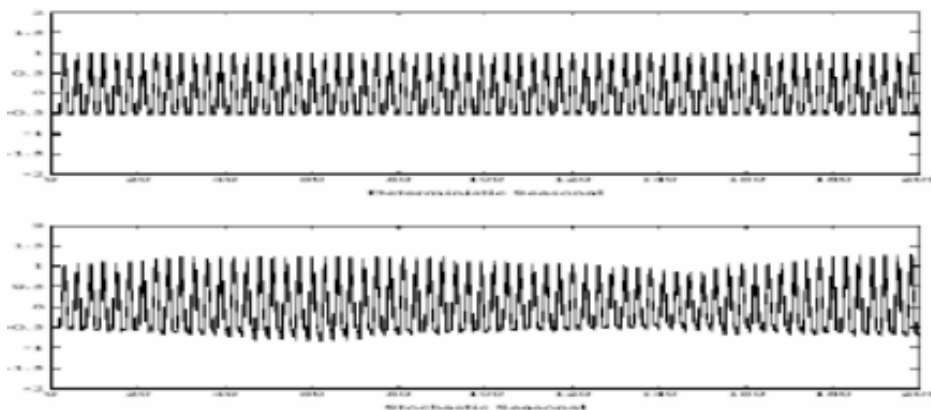
Διάγραμμα 2.4 – Μακροχρόνια Τάση

Η κυκλική κύμανση που παρουσιάζουν περιπτώσεις χρονολογικών σειρών έγκειται στο γεγονός ότι παρατηρούνται αυξομειώσεις της τιμής της μεταβλητής γύρω από γραμμή τάσης σε μια μακροχρόνια περίοδο. Στην πράξη τα σημεία της χρονολογικής σειράς για μια χρονική σειρά παρατηρήσεων βρίσκονται κάτω από την γραμμή τάσης και στην συνέχεια για άλλη χρονική σειρά τιμών πάνω από την γραμμή τάσης. Ο χρόνος για να έχουμε μια κυκλική αυξομείωση δεν είναι σταθερός.

Οι κυκλικές αυξομειώσεις είναι δύσκολο να αντιμετωπιστούν γιατί η κυκλική κίνηση δεν ακολουθεί κανένα κανονικό μοντέλο αλλά κινείται απρόβλεπτα. Αντίθετα, οι χρονολογικές σειρές που παρουσιάζουν περιοδικές μεταβολές είναι πολύ χρήσιμες γιατί ακολουθούν κανονικό μοντέλο και έτσι μπορούν να δώσουν αξιόπιστες προβλέψεις για το μέλλον.



Διάγραμμα 2.5 - Χρονολογική σειρά που παρουσιάζει κυκλική κύμανση.



Διάγραμμα 2.6 - Χρονολογική σειρά που παρουσιάζει περιοδικότητα.

2.4.5 Τυχαία Συνιστώσα

Η τυχαία συνιστώσα μιας χρονολογικής σειράς δεν επηρεάζεται από τις άλλες συνιστώσες και δεν επαναλαμβάνεται σε τακτά χρονολογικά διαστήματα. Υπολογίζεται ως εξής :

$$R_t = \left\{ \frac{Y_t}{C_t T_t S_t} \right\} \quad (2.9)$$

2.5 Στασιμότητα και Αυτοσυσχέτιση

Η βασική ιδέα της στασιμότητας⁵ είναι ότι οι νόμοι της πιθανότητας που διαμορφώνουν μια στοχαστική διαδικασία δεν μεταβάλλονται με το χρόνο. Η στασιμότητα ορίζεται είτε με αυστηρή έννοια, είτε με ασθενέστερη που είναι και επικρατούσα. Μια στοχαστική διαδικασία χαρακτηρίζεται ως αυστηρά στάσιμη (strictly stationary) όταν οι στατιστικές της ιδιότητες δεν επηρεάζονται από μια μετακίνηση στο χρόνο, παραμένουν δηλαδή χρονικά αμετάβλητες. Ειδικότερα, μια στοχαστική διαδικασία $\{Y_t\}$ θεωρείται αυστηρά στάσιμη αν η από κοινού κατανομή πιθανότητας των $\{Y_t, Y_{t+1}, \dots, Y_{t+N-1}\}$ είναι ίδια με την από κοινού κατανομή του συνόλου $\{Y_{t+k}, Y_{t+k+1}, \dots, Y_{t+k+N-1}\}$ για οποιαδήποτε επιλογή του χρόνου t , του πλήθους N και των υστερήσεων/προηγέσεων k .

Μια χρονολογική σειρά θα είναι ασθενώς στάσιμη (weakly stationary) αν ο μέσος και η διακύμανση της δεν μεταβάλλονται με το χρόνο και η συνδιακύμανση μεταξύ των τιμών της σε δύο χρονικά σημεία εξαρτάται μόνο από την απόσταση ανάμεσα σε αυτά τα χρονικά σημεία, και όχι από τον ίδιο το χρόνο.

Αν η χρονολογική σειρά είναι στάσιμη, τότε ισχύει:

i) $E(Y_t) = \mu_y$, για όλα τα t (2.10)

ii) $\text{Var}(Y_t) = E[Y_t - E(Y_t)]^2 = \sigma_y^2$, για όλα τα t (2.11)

iii) $\text{Cov}(Y_t, Y_{t+k}) = \text{cov}(Y_{t+m}, Y_{t+m+k}) = \gamma_k$, για όλα τα t και $k, m \neq 0$ (2.12)

Όπου μ_y συμβολίζει το σταθερό μέσο της Y_t , το σ_y^2 συμβολίζει τη σταθερή διακύμανση της και το γ_k τη συνδιακύμανση μεταξύ δύο οποιονδήποτε τιμών της Y_t που

⁵ Δημελή Σ.(2013, Κ.2 σ.31)

απέχουν k χρονικές περιόδους, η οποία ονομάζεται αυτοσυνδιακύμανση (autocovariance) και ορίζεται ως :

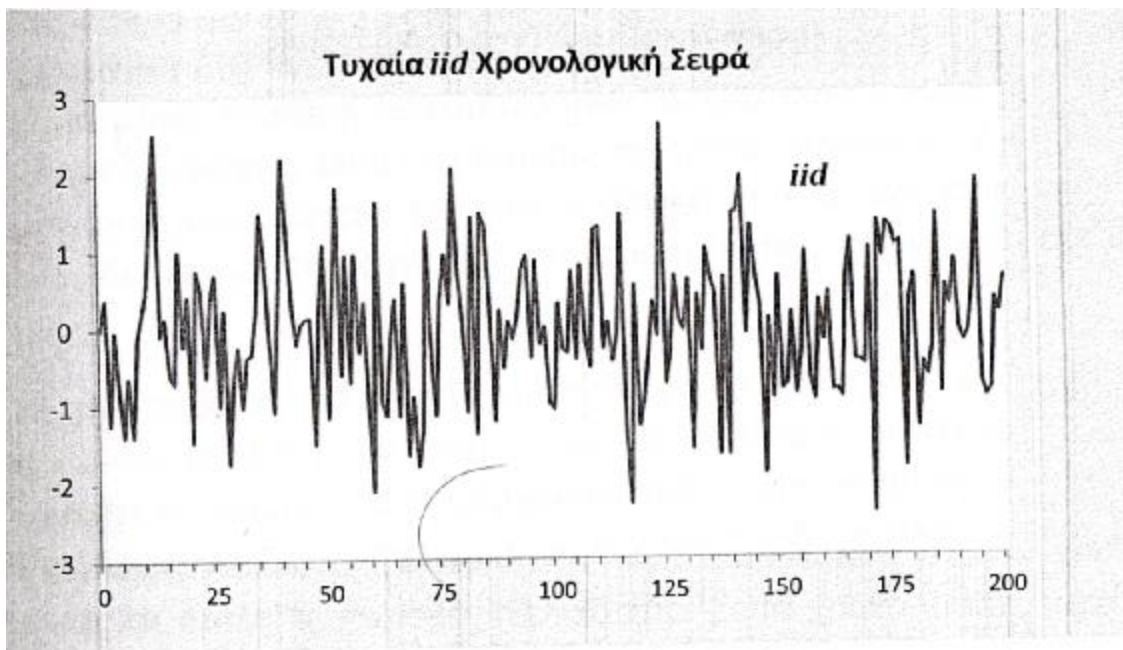
$$\gamma_k = cov(y_t, y_{t+k}) = E[y_t - E(y_t)][y_{t+k} - E(y_{t+k})] \quad (2.13)$$

Μελετώντας αυτές τις συνθήκες διαπιστώνεται ότι η πρώτη (i) συνθήκη υποδηλώνει το σταθερό μέσο και εξίσου η δεύτερη (ii) σταθερή διακύμανση για οποιαδήποτε χρονική στιγμή t . Η τρίτη (iii) δηλώνει ότι η συνδιακύμανση μεταξύ δύο οποιονδήποτε τιμών της Y_t που απέχουν k περιόδους είναι συνάρτηση μόνο του k , δηλαδή της χρονικής υστέρησης ή προήγησης των δύο αυτών τιμών. Είναι φανερό ότι η αυτοσυνδιακύμανση μηδενικής υστέρησης ($k=0$) είναι η διακύμανση, δηλαδή $\gamma_0 = cov(Y_t, Y_{t+k}) = var(Y_t) = \sigma_y^2$.

Έτσι, συνδυάζοντας τις δύο τελευταίες συνθήκες, οι συνθήκες στασιμότητας μπορούν να διατυπωθούν ως εξής.

- 1) Σταθερός μέσος σε όλα τα χρονικά σημεία t .
- 2) Συνδιακυμάνσεις γ_k ανεξάρτητες του χρόνου t , για οποιοδήποτε ακέραιο k .

Μία διαδικασία που ακολουθεί τις παραπάνω συνθήκες μπορεί να αναφερθεί και ως στάσιμη δεύτερης τάξεως ή στάσιμη ως προς τη συνδιακύμανση ή στάσιμη ως προς την ευρύτερη έννοια.



Διάγραμμα 2.7 - Τυχαία Χρονολογική Σειρά.

Στο παραπάνω διάγραμμα παρουσιάζεται η πορεία μιας τυχαίας χρονολογικής σειράς, όπου για την κατασκευή της χρησιμοποιήθηκε ένα δείγμα $N = 200$ παρατηρήσεων από έναν πληθυσμό τυχαίων αριθμών που ακολουθούν κανονική κατανομή. Η διαδικασία αυτή είναι εξ ορισμού στάσιμη διότι επαληθεύονται οι συνθήκες που αναφέρθηκαν παραπάνω.

Αν οι συνθήκες αυτές δεν ισχύουν τότε η σειρά δεν είναι στάσιμη.



Διάγραμμα 2.8 - Χρονολογική σειρά τυχαίας διαδρομής

Αντιθέτως, στο παραπάνω διάγραμμα παρουσιάζεται η στοχαστική διαδικασία της τυχαίας διαδρομής RW η οποία έχει τα χαρακτηριστικά μιας μη στάσιμης σειράς. Η σειρά απομακρύνεται διαχρονικά από τα αρχικά της επίπεδα, χωρίς καμία τάση επιστροφής σε ένα μακροχρόνιο επίπεδο. Η διακύμανσή της αυξάνεται και οι διαδοχικές τιμές της παρουσιάζουν έντονη αυτοσυσχέτιση. Οι ιδιότητες αυτές αναφέρονται, αποδεικνύονται και αναλύονται στις επόμενες σελίδες.

Για μια στάσιμη χρονολογική σειρά ισχύει :

$$E(Y_t) = E(Y_{t+k}) = \mu_y \quad (2.14)$$

$$\gamma_k = \text{cov}(Y_t, Y_{t+k}) = E[Y_t - \mu_y][Y_{t+k} - \mu_y] \quad (2.15)$$

$$\gamma_k = \gamma_{-k} \quad (2.16)$$

Πιο χρήσιμος είναι ο συντελεστής συσχέτισης μεταξύ δυο παρατηρήσεων Y_t και Y_{t+k} που απέχουν k χρονικές περιόδους και ονομάζεται συντελεστής αυτοσυσχέτισης τάξεως k (autocorrelation coefficient) :

$$\rho_k = \frac{cov(Y_t, Y_{t+k})}{\sqrt{var(Y_t)} \sqrt{var(Y_{t+k})}} \quad (2.17)$$

Στην περίπτωση όπου η χρονολογική σειρά είναι στάσιμη⁶, τότε η διακύμανση δε μεταβάλλεται σύμφωνα με το χρόνο και άρα $var(Y_t) = var(Y_{t+k}) = \sigma_y^2$, οπότε η (2.16) και με βάση τις (2.13)-(2.15) απλοποιείται ως ο λόγος:

$$\rho_k = \frac{cov(Y_t, Y_{t+k})}{var(Y_t)} = \frac{\gamma_k}{\gamma_0} \quad (2.18)$$

Ένα χαρακτηριστικό παράδειγμα στάσιμης χρονολογικής σειράς είναι το διάγραμμα 2.7, στο οποίο διάγραμμα η διαδικασία πληροί και τις τρεις συνθήκες στασιμότητας αφού έχει σταθερό μέσο, σταθερή διακύμανση και ανεξαρτησία μεταξύ των τιμών της. Η ανεξαρτησία συνεπάγεται πάντοτε με μηδενικές αυτοσυσχετίσεις, ενώ το αντίστροφο δεν ισχύει.

2.6 Δειγματικοί συντελεστές αυτοσυσχέτισης

Έχοντας ένα πεπερασμένο δείγμα παρατηρήσεων της χρονολογικής σειράς, μπορούν να υπολογιστούν οι δειγματικές τιμές⁷ των παραπάνω στατιστικών συντελεστών. Για παράδειγμα, ο δειγματικός μέσος (sample mean) μπορεί να υπολογισθεί ως η μέση των N παρατηρήσεων του δείγματος:

$$\bar{Y} = \frac{1}{N} \sum_{t=1}^N Y_t \quad (2.19)$$

Ο μέσος αυτός όμως, δεν εκφράζει το "συνολικό" μέσο $E(Y_t)$ της στοχαστικής διαδικασίας σε κάθε σημείο του χρόνου, αλλά μόνο ένα μέσο στο χρόνο. Για να συγκλίνει ο χρονικός μέσος στο $E(Y_t)$ θα πρέπει να ισχύει η υπόθεση της εργοδικότητας

⁶ Δημελή Σ.(2013, Κ.2 σ.35)

⁷ Δημελή Σ.(2013, Κ.2 σ.36)

(ergodicity). Μια στάσιμη χρονολογική σειρά Y_t είναι εργοδική στο μέσο εάν συγκλίνει στη "συνολική" αναμενόμενη τιμή :

$$\bar{Y} \rightarrow E(Y_t) \text{ καθώς } N \rightarrow \infty \quad (2.20)$$

Όπως αποδεικνύεται (Hamilton, 1994 και Henry 1995), μια σειρά θα είναι εργοδική στο μέσο, αν οι αυτοσυνδιακυμάνσεις γ_k τείνουν στο μηδέν αρκετά γρήγορα καθώς $k \rightarrow \infty$.

Παρόμοια ισχύουν για τη δειγματική διακύμανση (sample variance) και τη δειγματική αυτοσυνδιακύμανση (sample autocovariance).

$$\hat{\gamma}_0 = \hat{\sigma}_y^2 = \frac{1}{N} \sum_{t=1}^N (Y_t - \bar{Y})^2 \quad (2.21)$$

$$\hat{\gamma}_k = \frac{1}{N-k} \sum_{t=1}^{N-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y}) \quad (2.22)$$

Τα παραπάνω ισχύουν από την υπόθεση της εργοδιότητας ότι $\bar{\gamma}_k \rightarrow \gamma_k$ καθώς $N \rightarrow \infty$

Διαιρώντας τη δειγματική αυτοσυνδιακύμανση με τη συνδιακύμανση υπολογίζουμε το δειγματικό συντελεστή αυτοσυσχέτισης k -τάξεως:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{N-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^N (Y_t - \bar{Y})^2} \quad (2.23)$$

Όπου οι τιμές των ρ_k βρίσκονται μέσα στο όρια $-1 \leq \rho_k \leq 1$ και εξετάζουμε μόνο τις θετικές τιμές που παίρνει το k , επειδή ισχύει $\rho_k = \rho_{-k}$, δηλαδή η συνάρτηση αυτοσυσχέτισης είναι συμμετρική γύρω από το μηδέν. Η διαγραμματική παρουσίαση της συνάρτησης αυτής ονομάζεται διάγραμμα της συνάρτησης αυτοσυσχέτισης (correlogram).

2.7 Διάγραμμα Αυτοσυσχέτισης

Η διαγραμματική απεικόνιση⁸ των τιμών της $\hat{\rho}_k$ για $k = 1, 2, \dots$ μας δίνει το διάγραμμα της συνάρτησης αυτοσυσχέτισης και έχουμε μια οπτική εικόνα για το αν η χρονολογική σειρά είναι στάσιμη ή όχι. Για μια στάσιμη χρονολογική σειρά όπως

⁸ Δημελή Σ. (2013, Κ.2 σ.39)

παρατηρείται και παρακάτω οι συντελεστές αυτοσυσχέτισης φθίνουν γρήγορα προς το μηδέν καθώς μεγαλώνει ο αριθμός των υστερήσεων k ενώ αντιθέτως, δεν συμβαίνει το ίδιο στις μη στάσιμες χρονολογικές σειρές.



Σχήμα 2.1 – Αυτοσυσχετίσεις στάσιμης χρονολογικής σειράς

Στο πρώτο διάγραμμα παρατηρούμε πως η σειρά είναι στάσιμη ενώ στο δεύτερο η σειρά είναι μη στάσιμη λόγω της αργής μείωσης των διαδοχικών συντελεστών αυτοσυσχέτισης. Οι μη στάσιμες χρονολογικές σειρές μετατρέπονται σε στάσιμες παίρνοντας πρώτες ή δεύτερες διαφορές ή με τη χρήση λογαρίθμων.



Σχήμα 2.2 – Αυτοσυσχετίσεις μη στάσιμης χρονολογικής σειράς

2.8 Έλεγχος Στασιμότητας

i. Έλεγχος του Barlett (Barlett's Test)

Ο έλεγχος του Barlett⁹, βασίζεται στην υπόθεση ότι η χρονολογική σειρά είναι στάσιμη και οι συντελεστές αυτοσυσχέτισης του δείγματος ακολουθούν προσεγγιστικά την κανονική κατανομή με μηδενικό μέσο και διακύμανση $\frac{1}{N}$, όπου N = μέγεθος του δείγματος. Επομένως για τον έλεγχο της μηδενικής υπόθεσης έχουμε $H_0 : \rho_k = 0$ εναλλακτικά $H_1 : \rho_k \neq 0$ τα οποία συγκρίνουμε με την τιμή του στατικού t :

$$t = \frac{\hat{\rho}_k}{\sqrt{\frac{1}{N}}} = \sqrt{N} \hat{\rho}_k. \quad (2.24)$$

Ο έλεγχος Barlett ισχύει για μεγάλα δείγματα και εξετάζει μεμονωμένα κάθε συντελεστή αυτοσυσχέτισης. Για να ελέγξουμε την υπόθεση ότι από κοινού ένας αριθμός συντελεστών διαφέρει ή όχι από το μηδέν εφαρμόζουμε το στατικό κριτήριο Q των Box-Pierce.

ii. Q Στατιστική (Box-Pierce Test)

Ο έλεγχος στασιμότητας μιας χρονολογικής σειράς με τη στατιστική Box-Pierce χρησιμοποιείται για τον έλεγχο της συνδυαστικής υπόθεσης ότι όλοι οι συντελεστές αυτοσυσχέτισης είναι μηδέν και ορίζεται ως εξής :

$$Q = N \sum_{k=1}^m \hat{\rho}_k^2 \quad (2.25)$$

Όπου N = μέγεθος του δείγματος και m = μήκος χρονικής υστέρησης. Η στατιστική Q ακολουθεί την X^2 κατανομή με m βαθμούς ελευθερίας. Αν $Q > X^2(a,m)$ η χρονολογική σειρά δεν είναι στάσιμη και αντιστρόφως.

iii. Στατιστική των Box-Ljung

Η στατιστική των Box-Ljung αν και ακολουθεί την $X^2(a,m)$ κατανομή δίνει καλύτερα αποτελέσματα από την Q στατιστική όταν εφαρμόζεται σε μικρά δείγματα και ορίζεται ως εξής :

$$LB = N(N + 2) \sum_{k=1}^m \left(\frac{\hat{\rho}_k}{N-k} \right) \rightarrow X_m^2 \quad (2.26)$$

Αν $LB > X^2(a,m)$ τότε η υπόθεση της στασιμότητας της χρονολογικής σειράς απορρίπτεται.

⁹ Δημελή Σ.(2013, Κ.2 σ.40)

2.9 Μερική Αυτοσυσχέτιση

Μια άλλη συνάρτηση που θα μελετήσουμε και χρησιμοποιείται στην εξέταση των χαρακτηριστικών μιας χρονολογικής σειράς είναι η συνάρτηση των συντελεστών μερικής αυτοσυσχέτισης¹⁰. Με το συντελεστή μερικής συσχέτισης $Y_{xy}; z_1, \dots, z_n$ μετριέται η συσχέτιση μεταξύ των μεταβλητών x και y όταν αφαιρεθεί η επίδραση που ασκούν οι z_1, \dots, z_n πάνω σε αυτές. Στις χρονοσειρές ο συντελεστής μερικής αυτοσυσχέτισης μεταξύ y_t και y_{t+k} ορίζεται ως ο συντελεστής συσχέτισης μεταξύ τους όταν έχουν ληφθεί υπόψη οι συσχετίσεις όλων των ενδιάμεσων τιμών $y_{t+1}, y_{t+2}, \dots, y_{t+k-1}$.

Τα υποδείγματα αυτοπαλινδρόμησης μπορούν να εξηγήσουν καλύτερα την έννοια της μερικής αυτοσυσχέτισης. Έστω η παλινδρόμηση της y_t πάνω στις y_{t-1} και y_{t-2} :

$$y_t = \phi_{21}y_{t-1} + \phi_{22}y_{t-2} + \varepsilon_t \quad (2.27)$$

Ο πρώτος υποδείκτης του ϕ δηλώνει τη χρονική υστέρηση της μεταβλητής και ο δεύτερος υποδείκτης του ϕ δηλώνει τη μέγιστη τάξη παλινδρόμησης. Ο συντελεστής της y_{t-2} μετρά το συντελεστή αυτοσυσχέτισης δεύτερης τάξης $-\phi_{22}$ καθώς αυτός δείχνει τη συσχέτιση μεταξύ των y_t και y_{t-2} όταν έχει συμπεριληφθεί στην παλινδρόμηση η ενδιάμεση y_{t-1} . Γενικότερα ο συντελεστής μερικής αυτοσυσχέτισης p -οστής τάξεως συμβολίζεται με το γράμμα ϕ_{pp} και είναι ο συντελεστής του y_{t-p} στην παλινδρόμηση:

$$y_t = \phi_{1p}y_{t-1} + \phi_{2p}y_{t-2} + \phi_{pp}y_{t-p} + \varepsilon_t \quad (2.28)$$

εδώ ο συντελεστής πρώτης τάξης ϕ_{11} ταυτίζεται με τον απλό συντελεστή αυτοσυσχέτισης ρ_1 . Στην πράξη εκτιμούμε τέτοια υποδείγματα ξεκινώντας με μια χρονική υστέρηση της y_t και προσθέτοντας κάθε φορά, από μια υστέρηση διαδοχικά. Οι εκτιμήσεις ελάχιστων τετραγώνων των συντελεστών ϕ_{pp} σε κάθε τέτοιο υπόδειγμα για $p=1,2,3,\dots$ μας δίνει τη σειρά συντελεστών μερικής αυτοσυσχέτισης $\hat{\phi}_{11}, \hat{\phi}_{22}, \hat{\phi}_{33}, \dots$. Έτσι για παράδειγμα, οι τρεις πρώτοι συντελεστές προκύπτουν από τις εκτιμήσεις των υποδειγμάτων:

$$\begin{aligned} y_t &= \phi_{11}y_{t-1} + \varepsilon_t && \rightarrow \hat{\phi}_{11} \\ y_t &= \phi_{21}y_{t-1} + \phi_{22}y_{t-2} + \varepsilon_t && \rightarrow \hat{\phi}_{22} \\ y_t &= \phi_{31}y_{t-1} + \phi_{33}y_{t-3} + \varepsilon_t && \rightarrow \hat{\phi}_{33} \end{aligned} \quad (2.29)$$

¹⁰ Θαλασσινός Ε.(1991, Κ.1 σ.24)

Οι τιμές των ϕ_{kk} για τις διάφορες τιμές του $k = 1, 2, \dots$ αποτελούν τη συνάρτηση μερικής αυτοσυσχέτισης (partial autocorrelation function), η οποία χρησιμεύει στην ερμηνεία της μορφής της στοχαστικής διαδικασίας που δημιουργήσε τη δεδομένη χρονολογική σειρά.

2.10 Τυχαία Χρονολογική Σειρά – Λευκός θόρυβος – iid Διαδικασία

2.10.1 Λευκός Θόρυβος (white noise)

Η απλούστερη μορφή χρονολογικής σειράς είναι η στοχαστική διαδικασία λευκού θορύβου (white noise)¹¹. Μια χρονοσειρά ονομάζεται λευκός θόρυβος αν δεν έχει κανένα σαφές σχήμα ή πρότυπο. Αν μια τέτοια σειρά απεικονιστεί με ε_t , τότε είναι λευκός θόρυβος στην περίπτωση που έχει σταθερό μέσο, σταθερή διακύμανση και δεν αυτοσυσχετίζονται οι τιμές της. Πιο συγκεκριμένα :

$$E(\varepsilon_t) = 0, \text{ για όλα τα } t$$

$$\gamma_0 = E(\varepsilon_t)^2 = \sigma_\varepsilon^2, \text{ για όλα τα } t \quad (2.30)$$

$$\gamma_k = E(\varepsilon_t \varepsilon_{t-k}) = 0, \text{ για όλα τα } t \text{ και } k \neq 0$$

Η διαδικασία του λευκού θορύβου είναι εξ ορισμού στάσιμη σύμφωνα με τις συνθήκες (2.30) που αναφέρθηκαν παραπάνω και μπορεί να χαρακτηριστεί ως λευκός θόρυβος ή ως μια τυχαία διαταραχή και οι τιμές δεν επηρεάζονται από παρελθοντικές τιμές, δεν επηρεάζουν όμως ούτε τις μελλοντικές της τιμές. Το βασικό στοιχείο που χαρακτηρίζει το λευκό θόρυβο, είναι ότι οι αυτοσυνδιακυμάνσεις του και οι συντελεστές αυτοσυσχέτισης είναι μηδενικοί. Σε πολλές περιπτώσεις λευκών θορύβων, όπως είναι λόγω χάρη οι τυχεροί αριθμοί της κλήρωσης του ΚΙΝΟ ή του ΤΖΟΚΕΡ που καταγράφονται ημερησίως ή μηνιαία, οι προηγούμενες τιμές της σειράς δεν συντελούν καθόλου στην πρόβλεψη των μελλοντικών αριθμών που θα πάρει η σειρά.

¹¹ Δημελή Σ.(2013, Κ.2 σ.47)

2.10.2 iid Διαδικασία

Μια τυχαία διαδικασία¹², η οποία σχετίζεται στενά με τη διαδικασία του λευκού θορύβου είναι η διαδικασία ανεξάρτητων και ισόνομα κατανομημένων τυχαίων μεταβλητών ή αλλιώς iid διαδικασία. Μια ακολουθία τυχαίων μεταβλητών e_t θα πρέπει να έχει τις ακόλουθες ιδιότητες ώστε να χαρακτηριστεί ως iid.

$E(e_t) = \mu$, σταθερός (όχι απαραίτητα μηδέν)

$$\gamma_0 = E(e)^2 = \sigma_e^2, \text{ σταθερή για όλα τα } t \quad (2.31)$$

e_t ανεξάρτητα από τα e_s , για όλα τα t και $s, t \neq s$

Η πρώτη συνθήκη επιτρέπει στο μέσο να διαφοροποιείται από το μηδέν, η δεύτερη συνθήκη είναι ίδια με αυτή του λευκού θορύβου και η τρίτη συνθήκη της ανεξαρτησίας μεταξύ των τιμών της iid διαδικασίας συνεπάγεται πάντα με μηδενικές αυτοσυσχετίσεις.

2.10.3 Υπόδειγμα Τυχαίας Διαδρομής - Υπόδειγμα Τυχαίας Διαδρομής με Σταθερά

Στο υπόδειγμα τυχαίας διαδρομής (random walk model) ή τυχαίου περιπάτου¹³, κάθε τιμή της χρονοσειράς Y_t προκύπτει από την αμέσως προηγούμενη Y_{t-1} με την προσθήκη ενός τυχαίου σφάλματος. Επομένως έχουμε τη σχέση :

$$Y_t = Y_{t-1} + \varepsilon_t \quad (2.32)$$

Όπου ε_t είναι λευκός θόρυβος μηδενικού μέσου. Στην παραπάνω εξίσωση παρατηρείται πως οι διαδοχικές μεταβολές των τιμών της Y_t είναι ανεξάρτητες τυχαίες μεταβλητές.

$$\Delta Y_t = Y_t - Y_{t-1} = \varepsilon_t \quad (2.33)$$

Για να δούμε πως θα εξελιχθεί διαχρονικά αυτή η σειρά θέτουμε $t = 1$ και αντικαθιστούμε διαδοχικά :

¹² Δημελή Σ.(2013, Κ.2 σ.33)

¹³ Δημελή Σ.(2013, Κ.2 σ.48-53)

$$Y_1 = Y_0 + \varepsilon_1$$

$$Y_2 = Y_1 + \varepsilon_2 = Y_0 + \varepsilon_1 + \varepsilon_2 \quad (2.34)$$

$$Y_t = Y_0 + \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t = Y_0 + \sum_{i=1}^t \varepsilon_i$$

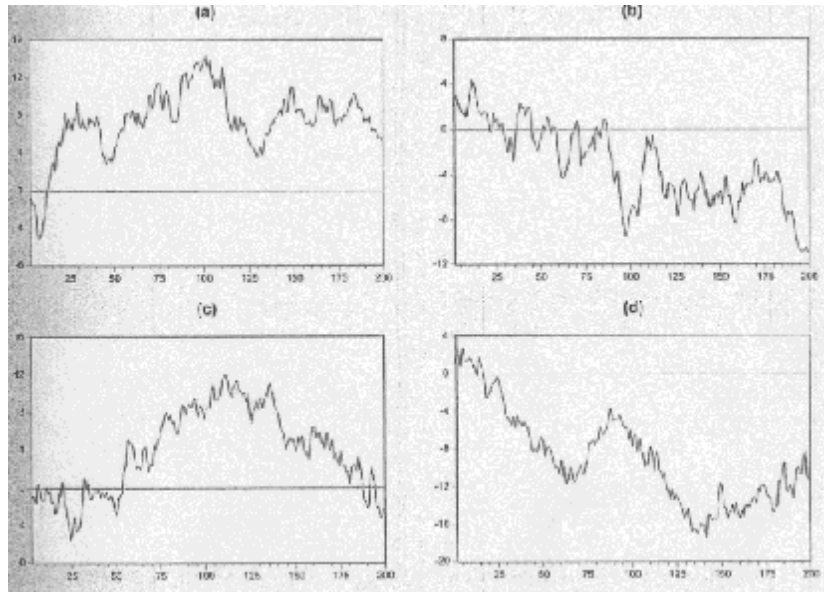
Η χρονοσειρά Y_t είναι το άθροισμα των τυχαίων μεταβλητών από την έναρξη της διαδικασίας έως τη χρονική στιγμή t . Η παραπάνω εξίσωση είναι και η γενική λύση της εξίσωσης διαφορών πρώτης τάξης που αντιπροσωπεύει η τυχαία διαδρομή. Γίνεται έλεγχος στασιμότητας σύμφωνα με τις συνθήκες στασιμότητας που αναφέρθηκαν στο κεφάλαιο 2.4. Υπολογίζουμε τον μέσο $E(Y_t)$ και τη διακύμανση $\text{var}(Y_t)$ της χρονολογικής σειράς :

$$\begin{aligned} E(Y_t) &= E(Y_0 + \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t) \\ &= Y_0 + E(\varepsilon_1) + E(\varepsilon_2) + \dots + E(\varepsilon_t) \quad (2.35) \\ &= Y_0, \quad \text{για όλα τα } t \end{aligned}$$

$$\begin{aligned} \gamma_0 = \text{var}(Y_t) &= E [Y_t - E(Y_t)]^2 \\ &= E(\varepsilon_1)^2 + E(\varepsilon_2)^2 + \dots + E(\varepsilon_t)^2 = t\sigma_\varepsilon^2 \quad (2.36) \end{aligned}$$

Τα ε_t παρουσιάζουν μηδενικό μέσο $E(\varepsilon_t) = 0$, σταθερή διακύμανση $E(\varepsilon_t)^2 = \sigma_\varepsilon^2$ και μηδενικές συνδιακυμάνσεις $E(\varepsilon_t \varepsilon_s) = 0$ για κάθε $t \neq s$. Παρατηρείται πως η τυχαία διαδρομή έχει ασταθείς διακυμάνσεις, παρόλο που έχει σταθερό μέσο $E(Y_t) = Y_0$, πράγμα που παραβιάζει τη δεύτερη συνθήκη στασιμότητας.

Η περίπτωση της τυχαίας διαδρομής είναι ένα κατατοπιστικό παράδειγμα για την κατανόηση της συμπεριφοράς μιας μη στάσιμης χρονολογικής σειράς. Η ισχυρή αυτοσυσχέτιση που παρουσιάζει ακόμα και σε μακρινές αποστάσεις οφείλεται στην ιδιότητα της εμμονής (persistence) ή αλλιώς στην «ισχυρή μνήμη» που παρουσιάζει η σειρά στο να «θυμάται» τις τυχαίες μεταβολές. Στην περίπτωση που κάποιο από τα ε_t μεταβληθεί ξαφνικά θετικά ή αρνητικά (positive or negative shock), τότε αυτή η μεταβλητή θα ενσωματωθεί για πάντα στην τιμή της σειράς Y_t χωρίς να εξασθενήσει, γι' αυτό και το υπόδειγμα της τυχαίας διαδρομής χαρακτηρίζεται ως υπόδειγμα με «άπειρη» μνήμη.



Διαγράμματα 2.9 - Στοχαστικές διαδικασίες τυχαίας διαδρομής

Εάν στο υπόδειγμα τυχαίας διαδρομής προστεθεί ένας σταθερός όρος β τότε αυτό το υπόδειγμα είναι γνωστό ως υπόδειγμα τυχαίας διαδρομής με σταθερά. Η παρουσία σταθεράς προκαλεί μια σταθερή μετατόπιση και εκφράζει την τάση που δείχνουν κάποιες σειρές να έχουν είτε ανοδική ($\beta > 0$) είτε καθοδική ($\beta < 0$) πορεία.

$$Y_t = \beta + Y_{t-1} + \varepsilon_t \quad (2.37)$$

$$\Delta Y_t = Y_t - Y_{t-1} = \beta + \varepsilon_t \quad (2.38)$$

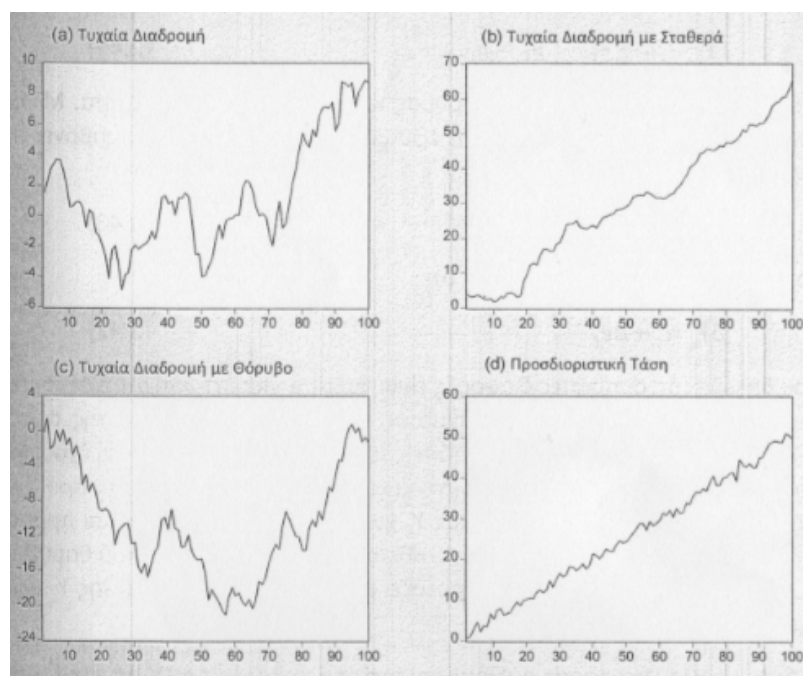
Οι πρώτες διαφορές είναι στάσιμες όπως παρατηρείται και στην περίπτωση της τυχαίας διαδρομής χωρίς σταθερά. Σε αντίθεση όμως με την τυχαία διαδρομή χωρίς σταθερά οι πρώτες διαφορές είναι εν μέρει σταθερές, δηλαδή ίσες με β , και εν μέρει στοχαστικές, δηλαδή ίσες με το λευκό θόρυβο ε_t . Η γενική λύση του υποδείγματος αυτού προκύπτει με τη μέθοδο διαδοχικών αντικαταστάσεων που εφαρμόσαμε και στην (2.34). Με δεδομένη την αρχική συνθήκη Y_0 προκύπτει:

$$Y_t = Y_0 + \beta t + \sum_{i=1}^t \varepsilon_i \quad (2.39)$$

όπου εμφανίζεται η προσδιοριστική τάση t . Αυτό δείχνει ότι η σειρά Y_t είναι μη στάσιμη και στη διακύμανση αλλά και στο μέσο¹⁴.

$$E(Y_t) = Y_0 + \beta t \quad (2.40)$$

Άρα ο μέσος της στοχαστικής διαδικασίας Y_t μεταβάλλεται γραμμικά με το χρόνο.

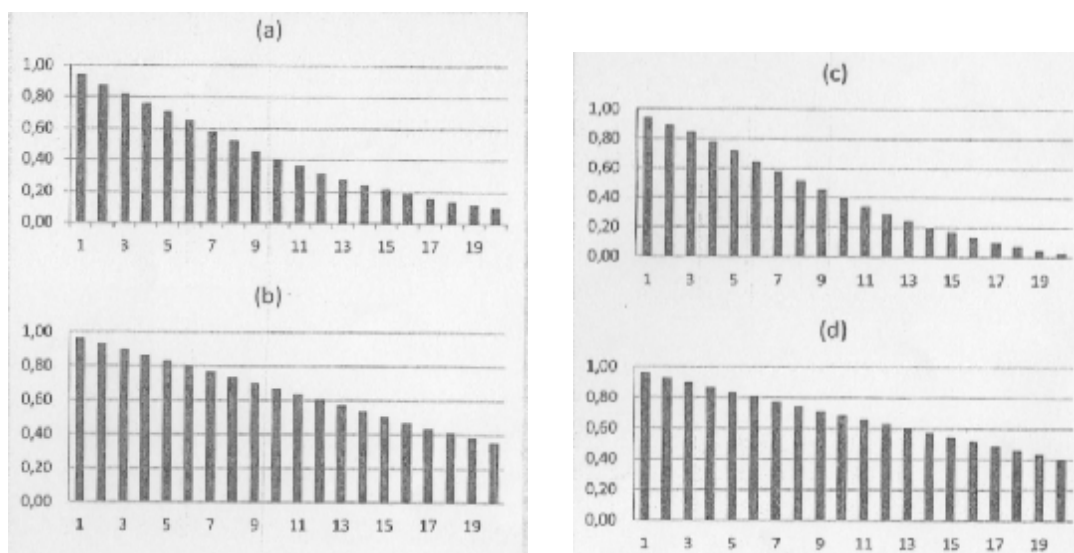


Διαγράμματα 2.10 - Υποδείγματα Στοχαστικής και Προσδιοριστικής Τάσης

Παραπάνω παρουσιάζονται προσομοιωμένα υποδείγματα τυχαίας διαδρομής με αρχική τιμή $Y_t = 0$ που δημιουργήθηκε με δείγμα $N = 100$ μιας ακολουθίας τυχαίων αριθμών από κανονικό πληθυσμό $\varepsilon_t \sim N(0,1)$. Το υπόδειγμα (b) προέκυψε από το ίδιο δείγμα προσθέτοντας στο υπόδειγμα σταθερά $\beta = 0,5$. Η επίδραση της προσδιοριστικής τάσης ($Y_0 + \beta t$) φαίνεται σημαντικά στην περίπτωση (d) που τη διαφοροποιεί από τη στοχαστική μορφή τάσης στην περίπτωση (a). Παρατηρείται πως η προσδιοριστική τάση υπερτερεί της στοχαστικής τάσης με αποτέλεσμα την καθαρά ανοδική πορεία της χρονολογικής

¹⁴ Δημελή Σ.(2013, Κ.2 σ.53)

σειράς όπως φαίνεται στην περίπτωση (b). Για τη δημιουργία της τελευταίας περίπτωσης προσομοιώθηκε η σειρά $Y_t = Y_0 + \beta t + \varepsilon_t$. Επίσης σημαντικό χαρακτηριστικό είναι πως η Y_t δεν επιστρέφει στην προσδιοριστική της τάση ($Y_0 + \beta t$) λόγω της συσσώρευσης των τυχαίων διαταραχών καθώς οι πρώτες διαφορές ($\Delta Y_t = \beta + \varepsilon_t$) κινούνται γύρω από το β . Στο υπόδειγμα τυχαίας διαδρομής χωρίς σταθερά η εμμονή που παρατηρείται στις τυχαίες διαταραχές οδηγεί τις τιμές σε μια πιο ασταθή συμπεριφορά προς τη μια ή την άλλη κατεύθυνση για μεγάλο διάστημα. Προσθέτοντας όμως μια σταθερά στο υπόδειγμα τυχαίας διαδρομής τότε η σειρά περιπλανιέται κατά την ποσότητα β προς τα πάνω για $\beta > 0$ ή προς τα κάτω για $\beta < 0$.



Σχήματα 2.3 –2.4 Συναρτήσεις αυτοσυσχέτισης των προσομοιωμένων υποδειγμάτων

Σε όλες τις περιπτώσεις παρατηρείται πως οι συντελεστές αυτοσυσχέτισης φθίνουν με αργό ρυθμό. Στο υπόδειγμα (b) τυχαίας διαδρομής με σταθερά, οι συντελεστές φθίνουν με αργό ρυθμό συγκριτικά με τους συντελεστές του υποδείγματος (a) χωρίς σταθερά. Στο υπόδειγμα (c) που αφορά τυχαία διαδρομή με θόρυβο οι συντελεστές φθίνουν γρηγορότερα σε σχέση με τους συντελεστές αυτοσυσχέτισης των υποδειγμάτων τυχαίας διαδρομής με ή χωρίς σταθερά. Η ύπαρξη δηλαδή του iid θορύβου μειώνει τις αυτοσυσχετίσεις συγκριτικά με της τυχαία διαδρομή. Τα υποδείγματα που παρουσιάστηκαν και οι ιδιότητες τους εξηγούνται αναλυτικά από τον Enders(2010, σ.188).

ΚΕΦΑΛΑΙΟ 3^ο: Στάσιμα Υποδείγματα Χρονολογικών Σειρών

3.1 Αυτοπαλίνδρομα υποδείγματα – AR(p)

Η γενική μορφή διατύπωσης ενός αυτοπαλίνδρομου υποδείματος¹⁵ AR(p) διατυπώνεται ως :

$$Y_t = \delta + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon_t \quad (3.1)$$

Το δ δηλώνει τη σταθερά της εξίσωσης, οι παράμετροι $\alpha_1, \alpha_2, \dots, \alpha_p$ είναι οι συντελεστές του αυτοπαλίνδρομου (autoregressive, AR) και το ε_t ονομάζεται λευκός θόρυβος. Η ονομασία αυτοπαλίνδρομο υποδηλώνει ότι η σχέση αυτή είναι ένα υπόδειγμα παλινδρόμησης, στο οποίο η εξαρτημένη μεταβλητή Y_t παλινδρομείται στις προηγούμενες τιμές της ίδιας μεταβλητής Y_t . Το p δείχνει την τάξη του αυτοπαλίνδρομου υποδείματος και αναφέρεται στο μήκος της υστέρησης ενώ τα $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$, είναι οι τιμές της χρονολογικής σειράς με υστέρηση.

Η χρησιμότητα του αυτοπαλίνδρομου υποδείματος AR, του υποδείματος κινητών μέσων MA και γενικότερα των ARMA υποδειγμάτων, είναι μεγάλη, διότι έχει αποδειχθεί πως ερμηνεύουν σε ικανοποιητικό βαθμό μεγάλο αριθμό πραγματικών δεδομένων στον τομέα των οικονομικών κυρίως επιστημών.

Στην ανάλυση χρονολογικών σειρών χρησιμοποιείται αρκετά ο **τελεστής υστέρησης (lag operator)**, ο οποίος διευκολύνει στην επίλυση αλγεβρικών πράξεων και συμβολίζεται με το γράμμα L ¹⁶.

$$L^j Y_t = Y_{t-j}, \quad j = 0, 1, 2 \dots \quad (3.2)$$

Χρησιμοποιώντας τον τελεστή υστέρησης, μπορούμε να παραστήσουμε το αυτοπαλίνδρομο υπόδειγμα AR(p) πιο συνοπτικά :

$$A(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p \quad (3.3)$$

¹⁵ Δημελή Σ.(2013, Κ.3 σ.83)

¹⁶ Δημελή Σ.(2013, Κ.3 σ.64)

3.2 Αυτοπαλίνδρομο υπόδειγμα πρώτης τάξεως(Autoregressive of order 1)

Αν στο γενικό υπόδειγμα αυτοπαλινδρόμησης θεωρήσουμε μια υστέρηση της χρονοσειράς Y_t και μόνο το τρέχον σφάλμα ε_t , τότε έχουμε το αυτοπαλίνδρομο υπόδειγμα πρώτης τάξης ή αλλιώς $AR(1)$ ¹⁷:

$$Y_t = \delta + \alpha Y_{t-1} + \varepsilon_t \quad (3.4)$$

Το αυτοπαλίνδρομο υπόδειγμα πρώτης τάξης και το γενικό υπόδειγμα $AR(p)$ αναλύονται απλούστερα αν τα επιλύσουμε με την επαναληπτική μέθοδο των διαδοχικών αντικαταστάσεων. Θεωρώντας ως αρχική τιμή το Y_0 , η ακολουθία των τιμών Y_1, Y_2, \dots, Y_t μπορεί να ληφθεί αντικαθιστώντας διαδοχικά στην (3.4) τις προηγούμενες τιμές της Y_t :

$$Y_1 = \delta + \alpha Y_0 + \varepsilon_1$$

$$Y_2 = \delta + \alpha(\delta + \alpha Y_0 + \varepsilon_1) + \varepsilon_2 = \delta(1 + \alpha) + \alpha^2 Y_0 + \alpha \varepsilon_1 + \varepsilon_2$$

...

$$Y_t = \delta \sum_{j=0}^{t-1} \alpha^j + \alpha^t Y_0 + \sum_{j=0}^{t-1} \alpha^j \varepsilon_{t-j} \quad (3.5)$$

Η (3.5) αποτελεί την αναδρομική μορφή (recursive form) του υποδείγματος, και τη χρησιμοποιούμε ώστε να μελετήσουμε της συνθήκες στασιμότητας της Y_t . Για τη στασιμότητα του μέσου λαμβάνουμε κατά μέλη την αναμενόμενη τιμή της (3.5):

$$E(Y_t) = \delta \sum_{j=0}^{t-1} \alpha^j + \alpha^t E(Y_0) + \sum_{j=0}^{t-1} \alpha^j E(\varepsilon_{t-j}) \quad (3.6)$$

Όπου τα δ και α είναι σταθερά. Ο τελευταίος όρος σε αυτή τη σχέση μηδενίζεται, καθώς $E(\varepsilon_t) = 0$ για όλα τα t και θεωρώντας δεδομένη την αρχική τιμή Y_0 , η εξίσωση του μέσου (3.6) απλοποιείται:

¹⁷ George E. P. Box, Gwilym M. Jenkins(2015. Κ.3, σ.54)

$$E(Y_t) = \delta \sum_{j=0}^{t-1} \alpha^j + \alpha^t Y_0 \quad (3.7)$$

$$= \delta \sum_{j=0}^{\infty} \alpha^j - \sum_{j=t}^{\infty} \alpha^j + \alpha^t Y_0$$

Η τελευταία είναι συνάρτηση του χρόνου και σημαίνει πως η Y_t δεν είναι στάσιμη. Στην (3.7) τα δυο αθροίσματα τείνουν στο άπειρο, οπότε για να συγκλίνουν θα πρέπει να ικανοποιείται η συνθήκη $|\alpha| < 1$. Στην περίπτωση που η συνθήκη αυτή ισχύει τότε τα αθροίσματα αποτελούν φθίνουσα γεωμετρική πρόοδο απείρου τάξεως με λόγο το α . Το όριο του πρώτου αθροίσματος είναι $1/(1-\alpha)$, ενώ του δεύτερου αθροίσματος είναι $\alpha^t/(1-\alpha)$. Με αλγεβρικές πράξεις προκύπτει πως ο μέσος ισούται με:

$$E(Y_t) = \delta \left(\frac{1}{1-\alpha} - \frac{\alpha^t}{1-\alpha} \right) + \alpha^t Y_0$$

$$= \delta / (1 - \alpha) \quad (3.8)$$

$$\text{Αφού } \alpha^t \rightarrow 0 \text{ καθώς το } t \rightarrow \infty \text{ για } |\alpha| < 1 \quad (3.9)$$

Η συνθήκη $|\alpha| < 1$ αποτελεί τη συνθήκη στασιμότητας του αυτοπαλίνδρομου υποδείγματος πρώτης τάξεως AR(1).

Επίσης από την (3.8) ισχύει ότι :

$$\delta = \mu(1 - \alpha), \mu = E(Y_t) \quad (3.10)$$

και αντικαθιστώντας το δ στην (3.4) έχουμε τη μορφή του υποδείγματος σε αποκλίσεις από το μέσο :

$$Y_t - \mu = \alpha(Y_{t-1} - \mu) + \varepsilon_t \quad \text{ή} \quad y_t = \alpha y_{t-1} + \varepsilon_t \quad (3.11)$$

Υψώνοντας και τα δύο μέλη της (3.11) στο τετράγωνο θα έχουμε :

$$E(y_t)^2 = E[(\alpha y_{t-1}) + (\varepsilon_t)]^2 = \alpha^2 E y_{t-1}^2 + 2\alpha E y_{t-1} \varepsilon_t + \varepsilon_t^2 \quad (3.12)$$

Το y_{t-1} εξαρτάται από το ε_{t-1} οπότε $E y_{t-1} \varepsilon_t = 0$ και επειδή η χρονολογική σειρά είναι στάσιμη το $E y_t^2 = E y_{t-1}^2 = \text{var}(y_t)$

Επομένως από την (3.12) προκύπτει :

$$\text{var}(y_t) = \alpha_1^2 \text{var}(Y_t) + \sigma^2 \quad \text{ή} \quad \gamma_0 = \text{var}(y_t) = \sigma_\varepsilon^2 \frac{1}{1-\alpha^2} \quad (3.13)$$

Πολλαπλασιάζοντας και τα δύο μέλη της (3.11) με y_{t-k} έχουμε:

$$y_t y_{t-k} = \alpha y_{t-1} y_{t-k} + \varepsilon_t y_{t-k}$$

$$E(y_t y_{t-k}) = \alpha E(y_{t-1} y_{t-k}) + E(\varepsilon_t y_{t-k}) \quad (3.14)$$

$\gamma_k = \alpha \gamma_{k-1}$ για κάθε $k > 1$ διότι $E(\varepsilon_t y_{t-k}) = 0$ και $E(y_{t-1} y_{t-k}) = \gamma_{k-1}$

$$\text{Οπότε γενικά θα ισχύει ότι} \quad \gamma_k = \alpha_k \gamma_0 \quad (3.15)$$

καθώς και $\gamma_1 = \alpha \gamma_0$, $\gamma_2 = \alpha^2 \gamma_0$, $\gamma_3 = \alpha^3 \gamma_0$.

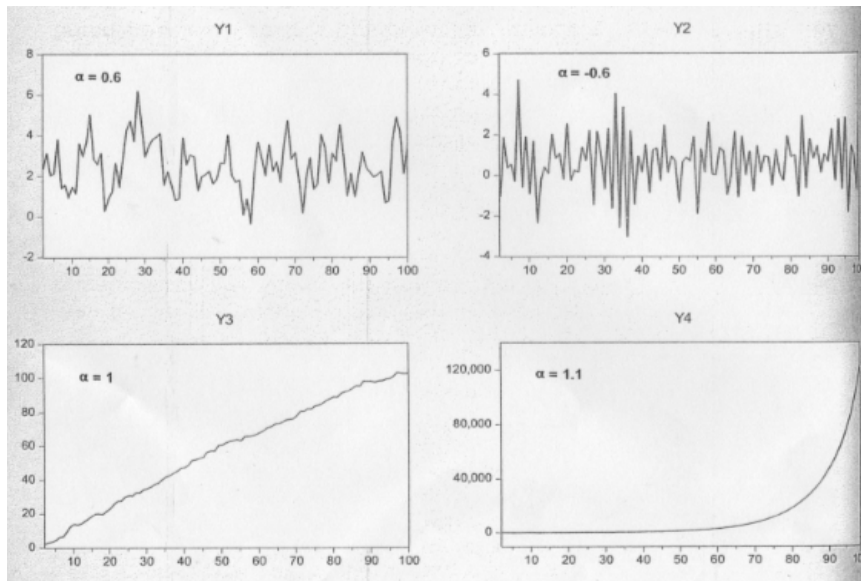
Διαιρώντας την (3.15) με την (3.13) προκύπτει η θεωρητική συνάρτηση αυτοσυσχέτισης:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \alpha_k, \quad k = 1, 2, 3, \dots, \rho_0 = 1 \quad (3.16)$$

Θα πρέπει το $|\alpha| < 1$ για να είναι στάσιμη η χρονολογική σειρά αφού η διακύμανση γ_0 δεν γίνεται να πάρει αρνητικές τιμές.

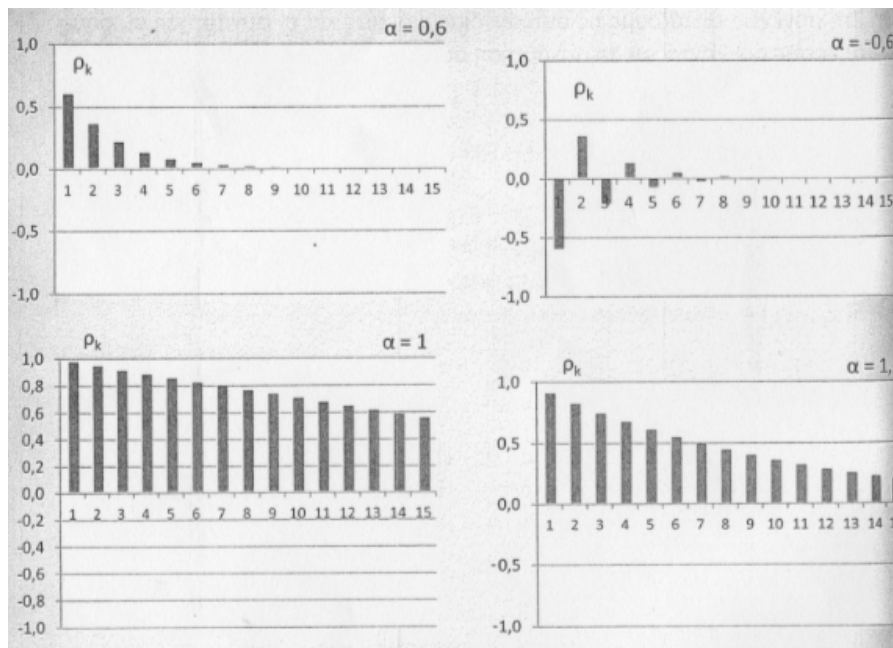
Για $\alpha > 0$ η συνάρτηση αυτοσυσχέτισης ξεκινώντας από τη μονάδα φθίνει γεωμετρικά και τείνει προς το 0 καθώς το k αυξάνει.

Το ίδιο ισχύει και στην περίπτωση που $\alpha < 0$ καθώς η συνάρτηση θα φθίνει γεωμετρικά προς το μηδέν αλλά με αρνητικό πρόσημο.



Σχήμα 3.1 – Αυτοπαλίνδρομα υποδείγματα πρώτης τάξεως AR(1)

Οι σειρές Y1 και Y2 ικανοποιούν τη συνθήκη σκοπιμότητας, δηλαδή έχουν σταθερό μέσο και σταθερή διακύμανση. Αντιθέτως στη σειρά Y3 παρατηρείται έντονη στοχαστική τάση, ενώ η σειρά Y4 έχει αποκλίνουσα πορεία.



Σχήμα 3.2 – Συναρτήσεις αυτοσυσχέτισης αυτοπαλίνδρομων υποδειγμάτων AR(1)

Οι συντελεστές αυτοσυσχέτισης φθίνουν με διαφορετικό ρυθμό σε αυτά τα υποδείγματα. Στη σειρά Y1, η οποία είναι στάσιμη σειρά, οι συντελεστές ρ_k φθίνουν με γρήγορο

ρυθμό και μονοτονικά και είναι AR(1) με θετικό αυτοπαλινδρομικό συντελεστή. Με γρήγορο ρυθμό και μονοτονικά φθίνουν επίσης και οι συντελεστές στην σειρά Y2 η οποία όμως έχει αρνητικό αυτοπαλινδρομικό συντελεστή. Στις σειρές Y3 και Y4 που είναι μη στάσιμες και ακολουθούν τυχαία διαδρομή και αποκλίνουσα πορεία, οι συντελεστές ρ_k φθίνουν με πολύ αργό ρυθμό.

3.3 Αυτοπαλίνδρομο υπόδειγμα δεύτερης τάξεως AR(2)

Τα αυτοπαλίνδρομα υποδείγματα δεύτερης τάξεως AR(2)¹⁸ (Autoregressive of order 2) στη γενική τους μορφή μπορούν να γραφτούν ως εξής :

$$Y_t = \delta + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \varepsilon_t \quad (3.17)$$

Το τυχαίο σφάλμα ε_t είναι λευκός θόρυβος και είναι ανεξάρτητο το υστερήσεων Y_{t-1} , Y_{t-2} , Y_{t-3}, \dots . Το αυτοπαλίνδρομο υπόδειγμα AR(2) σε αποκλίσεις από το μέσο γράφεται ως εξής :

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t \quad (3.18)$$

όπου $y_t = Y_t - \mu$ και $E(y_t) = 0$. Το υπόδειγμα AR(2) σε όρους του τελεστή υστέρησης L, μπορεί να γραφτεί ως :

$$(1 - \alpha_1 L - \alpha_2 L^2)y_t = \varepsilon_t \quad (3.19) \quad \text{ή} \quad A(L)y_t = \varepsilon_t \quad (3.20)$$

Το A(L) είναι ένα δευτερόνυμο πολυώνυμο, όπου οι ρίζες του είναι οι τιμές που λύνουν τη χαρακτηριστική εξίσωση (characteristic equation) του πολυωνύμου, στην ορολογία των εξισώσεων διαφορών.¹⁹

$$A(L) = (1 - \alpha_1 L - \alpha_2 L^2) = 0 \quad (3.21) \quad \text{και ως γινόμενο παραγόντων,}$$

¹⁸ George E. P. Box, Gwilym M. Jenkins(2015. Κ.3, σ.54)

¹⁹ Enders W.(2010, Κεφ.1) αναλυτική παρουσίαση της θεωρίας των εξισώσεων διαφορών που είναι τα υποδείγματα χρονολογικών σειρών χωρίς το στοχαστικό όρο.

$$A(L) = (1 - \lambda_1 L)(1 - \lambda_2 L) = 0 \quad (3.22)$$

Τα λ_1^{-1} και λ_2^{-1} είναι οι ρίζες του πολυώνυμου $A(L)$. Αντικαθιστώντας στην (3.21) το L με $1/X$ και πολλαπλασιάζοντας τα μέλη με X^2 προκύπτει η εξίσωση δευτεροβάθμιας μορφής :

$$X^2 - \alpha_1 X - \alpha_2 = 0 \quad (3.23)$$

Τα λ_1 και λ_2 μπορούν να θεωρηθούν ως οι ρίζες της παραπάνω χαρακτηριστικής εξίσωσης (3.23), και ονομαζόμενες και χαρακτηριστικές ρίζες συνδέονται με τις παραμέτρους α_1 και α_2 με τον εξής τρόπο :

$$\lambda_1 + \lambda_2 = \alpha_1 \quad \text{και} \quad \lambda_1 \lambda_2 = -\alpha_2 \quad (3.24)$$

και υπολογίζονται από τον τύπο $\lambda_1, \lambda_2 = (\alpha_1 \pm \sqrt{\Delta}) / 2$, $\Delta = \alpha_1^2 + 4\alpha_2$ (3.25)

Ένας εναλλακτικός τρόπος να γραφτεί το AR(2) είναι : $y_t = A(L)^{-1} \varepsilon_t$ (3.26) και αν αντικαταστήσουμε την (3.22) στην (3.26) και αναπτύσσοντας τα πολυώνυμα σε επιμέρους κλάσματα έχουμε:

$$y_t = \frac{1}{(1 - \lambda_1 L)(1 - \lambda_2 L)} \varepsilon_t = \frac{c}{1 - \lambda_1 L} \varepsilon_t + \frac{h}{1 - \lambda_2 L} \varepsilon_t \quad (3.26)$$

όπου $c = -\lambda_1 / (\lambda_2 - \lambda_1)$ και $h = \lambda_2 / (\lambda_2 - \lambda_1)$. Για να είναι συγκλίνουσα η χρονολογική σειρά θα πρέπει να συγκλίνουν οι όροι στο δεξί μέρος της εξίσωσης και για να συμβαίνει αυτό πρέπει να ισχύει:

$$|\lambda_1| < 1 \quad \text{και} \quad |\lambda_2| < 1 \quad (3.27)$$

οι οποίες συνθήκες είναι ισοδύναμες με τις συνθήκες $|1/\lambda_1| > 1$ και $|1/\lambda_2| > 1$ (3.28), οι οποίες είναι οι ρίζες της χαρακτηριστικής εξίσωσης (3.21).

Για να είναι στάσιμο το αυτοπαλίνδρομο υπόδειγμα δεύτερης τάξης AR(2) θα πρέπει οι ρίζες λ_1 και λ_2 της χαρακτηριστικής εξίσωσης να είναι μικρότερες της μονάδας σε απόλυτες τιμές ή αλλιώς να βρίσκονται εντός του μοναδιαίου κύκλου. Ισοδύναμα, θα

λέμε ότι το AR(2) είναι στάσιμο εάν οι ρίζες του πολυώνυμου $A(L)$ είναι μεγαλύτερες της μονάδας σε απόλυτες τιμές ή αλλιώς βρίσκονται εκτός του μοναδιαίου κύκλου. Αναλόγως με την τιμή της διακρίνουσας Δ στην (3.25), οι χαρακτηριστικές ρίζες διακρίνονται σε πραγματικές ($\Delta \geq 0$) ή μιγαδικές ($\Delta < 0$). Εφόσον ικανοποιούνται οι παραπάνω συνθήκες στασιμότητας τότε το AR(2) θα έχει σταθερό μέσο ίσο με :

$$\mu = \delta / (1 - \alpha_1 - \alpha_2) \quad (3.29)$$

Επίσης, η στασιμότητα συνεπάγεται σταθερή διακύμανση $\gamma_0 = E(y_t)^2 = E(y_{t-1})^2 = E(y_{t-2})^2$ για όλα τα t , επομένως:

$$\gamma_0 = E(\alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t)^2 = (\alpha_1^2 + \alpha_2^2) \gamma_0 + 2\alpha_1 \alpha_2 \gamma_1 + \sigma_\varepsilon^2 \quad (3.30)$$

η οποία είναι συνάρτηση της αυτοσυνδιακύμανσης πρώτης τάξης γ_1 .

Για να γίνει ο υπολογισμός της διακύμανσης, ορίζουμε τις εξισώσεις των αυτοσυνδιακυμάνσεων, οι οποίες υπολογίζονται πολλαπλασιάζοντας κατά μέλη το υπόδειγμα 3.18 με y_{t-1} , y_{t-2} ή γενικά με y_{t-k} :

$$\begin{aligned} \gamma_k &= E(y_{t-k} y_t) \\ &= E[y_{t-k} (\alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t)] \\ &= \alpha_1 \gamma_{k-1} + \alpha_2 \gamma_{k-2}, \quad \text{για } k = 1, 2, 3, \dots \end{aligned} \quad (3.31)$$

Για $k = 1$ η 3.31 γίνεται $\gamma_1 = \alpha_1 \gamma_0 + \alpha_2 \gamma_1$, την οποία αν τη λύσουμε μαζί με την 3.30 παίρνουμε την εξίσωση της διακύμανσης ως συνάρτηση των παραμέτρων του AR(2) και της διακύμανσης σ_ε^2 του λευκού θορύβου²⁰:

$$\gamma_0 = \frac{(1 - \alpha_2) \sigma_\varepsilon^2}{(1 - \alpha_2)(1 - \alpha_1^2 - \alpha_2^2) - 2\alpha_1 \alpha_2} = \frac{(1 - \alpha_2) \sigma_\varepsilon^2}{(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)} \quad (3.32)$$

²⁰ Δημελή Σ.(2013, Κ.3 σ.75)

Η διακύμανση είναι πάντα θετικός αριθμός, άρα κάθε όρος σε παρένθεση θα πρέπει να είναι θετικός και για να γίνει αυτό θα πρέπει να πληρούνται οι εξής συνθήκες από τις παραμέτρους α_1 και α_2 :

$$\alpha_1 + \alpha_2 < 1, \alpha_2 - \alpha_1 < 1 \text{ και } |\alpha_2| < 1 \quad (3.33)$$

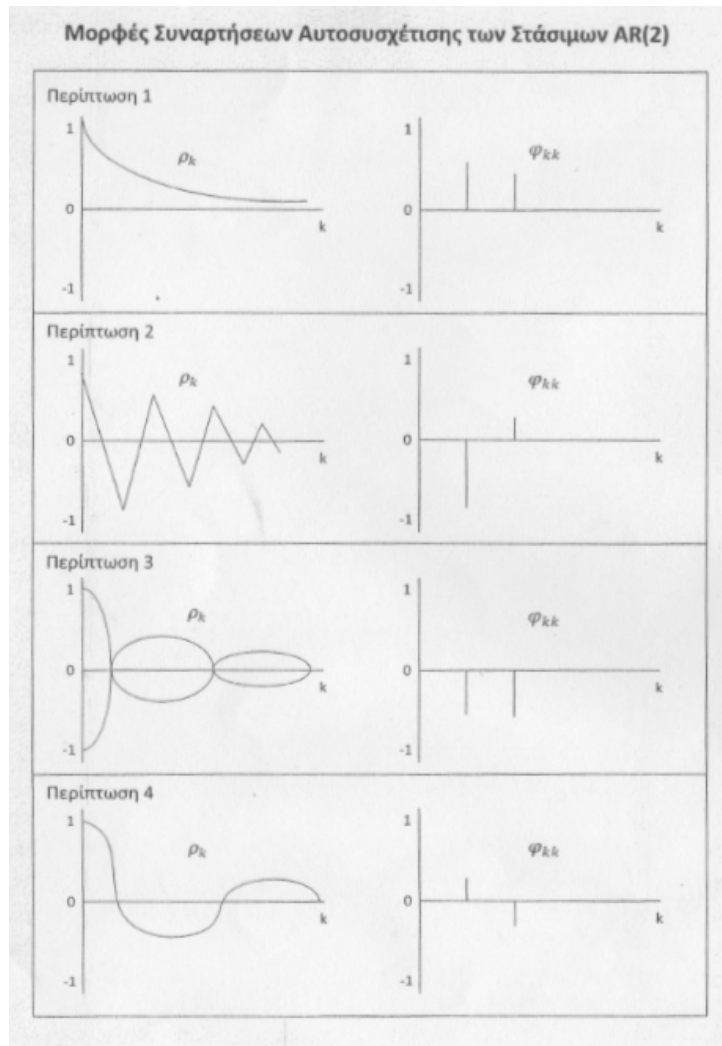
Διαιρώντας κατά μέλη την (3.31) με γ_0 παίρνουμε τη συνάρτηση αυτοσυσχέτισης $\rho_k = \gamma_k / \gamma_0$, που είναι και αυτή συνάρτηση των παραμέτρων α_1 και α_2 :

$$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2} \quad \text{για } k = 1, 2, 3 \quad (3.34)$$

Για $k = 1, 2$ λαμβάνουμε τις δυο πρώτες αυτοσυσχετίσεις ρ_1 και ρ_2 :

$$\begin{aligned} \rho_1 &= \alpha_1 + \alpha_2 \rho_1 \\ \rho_2 &= \alpha_1 \rho_1 + \alpha_2 \end{aligned} \quad (3.35)$$

Η λύση του συστήματος των δυο εξισώσεων, οι οποίες αποτελούν τις εξισώσεις Yule-Walker, εκφράζει τα ρ_1 και ρ_2 σε όρους των συντελεστών αυτοσυσχέτισης α_1 και α_2 . Οι συντελεστές αυτοσυσχέτισης ρ_k μεγαλύτερης τάξης ($k > 2$) υπολογίζονται από την (3.34) αντικαθιστώντας σε αυτή τις τιμές των ρ_1 και ρ_2 από την (3.35). Στο παρακάτω διάγραμμα παρουσιάζονται οι πιθανές μορφές των θεωρητικών συναρτήσεων αυτοσυσχέτισης των AR(2) υποδειγμάτων.



Σχήμα 3.3 – Μορφές συναρτήσεων αυτοσυσχέτισης των στάσιμων AR(2)

Η συνάρτηση αυτοσυσχέτισης ρ_k , $k=1,2,\dots$ του υποδείγματος AR(2) αποτελείται από τις εξισώσεις:

$$\rho_1 = \frac{\alpha_1}{1-\alpha_2} \quad (3.36) \quad , \quad \rho_2 = \frac{\alpha_2(1-\alpha_2) + \alpha_1^2}{1-\alpha_2} \quad (3.37)$$

και

$$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2} \quad \text{για } k = 3,4 \quad (3.38)$$

Έχουμε δηλαδή μια εξίσωση διαφορών δεύτερης τάξης με δεδομένες τις δυο πρώτες τιμές της. Επίσης, η εξίσωση (3.38) είναι ανάλογη με την εξίσωση διαφορών (3.18) της ίδιας στοχαστικής διαδικασίας Y_t χωρίς όμως το τυχαίο σφάλμα ε_t .

Τα ρ_k είναι συνάρτηση των παραμέτρων α_1 και α_2 , οι οποίες είναι όμοιες με τις παραμέτρους του υποδείγματος (3.18). Άρα για στάσιμη σειρά AR(2), η συνάρτηση αυτοσυσχέτισης θα ακολουθεί μια πορεία φθίνουσα είτε εκθετικά, είτε κυματιστά όπως φαίνεται και στο διάγραμμα. Αν οι ρίζες (3.24) είναι πραγματικές τότε τα ρ_k φθίνουν εκθετικά και μονοτονικά (διάγραμμα 3.3 – περίπτωση 1) ή με εναλλασόμενο πρόσημο (διάγραμμα 3.3 – περίπτωση 2). Αν οι ρίζες είναι μιγαδικές τότε τα ρ_k θα φθίνουν με ημιτονοειδή τρόπο (διάγραμμα 3.3- περίπτωση 3 και 4).

Αναφορικά με τους τους συντελεστές μερικής αυτοσυσχέτισης φ_{kk} του AR(2), μόνο οι δυο πρώτοι θα διαφέρουν από το μηδέν. Αυτοί υπολογίζονται βάσει των Yule – Walker εξισώσεων²¹:

$$\varphi_{11} = \rho_1 \quad \text{και} \quad \varphi_{22} = (\rho_2 - \rho_1^2) / (1 - \rho_1^2)$$

Στο διάγραμμα 3.3, όλοι οι συντελεστές φ_{kk} για $k > 2$ είναι μηδενικοί. Η τάξη μετά την οποία μηδενίζονται οι συντελεστές φ_{kk} προσδιορίζει και την τάξη ενός AR υποδείγματος.

Από τις τιμές των ρ_1 και ρ_2 (3.36),(3.37) και από τις συνθήκες στασιμότητας προκύπτει πως οι αποδεκτές τιμές των ρ_1 και ρ_2 για το στάσιμο υπόδειγμα AR(2) είναι :

$$|\rho_1| < 1, \quad |\rho_2| < 1, \quad \rho_1^2 < \frac{1}{2}(\rho_2 + 1) \quad (3.39)$$

Ο έλεγχος για την ιδιότητα της αντιστρεψιμότητας είναι εξαιρετικά χρήσιμος, αν δηλαδή μπορούμε να εκφράσουμε το υπόδειγμα AR(2) σε όρους τυχαίων μόνο σφαλμάτων. Αν πληρούνται οι συνθήκες στασιμότητας (3.27) τότε κάθε κλάσμα της (3.26) αποτελεί όριο γεωμετρικής πρόοδου άπειρης τάξης, οπότε μετά από τις απαραίτητες αλγεβρικές πράξεις το υπόδειγμα (3.26) παίρνει τη μορφή:

$$y_t = (1 + \psi_1 L + \psi L^2 + \psi L^3 + \dots) \varepsilon_t = \sum_{j=0}^{\infty} \psi_j L^j \varepsilon_t \quad \text{για} \quad \psi_0 = 1 \quad (3.40)$$

²¹ Δημελή Σ.(2013, Κ.3 σ.85)

όπου οι σταθμίσεις ψ_j είναι συναρτήσεις των λ_1 και λ_2 . Η μορφή (3.40) αποτελεί ένα υπόδειγμα κινητών μέσων άπειρης τάξης, $MA(\infty)$.

3.4 Αυτοπαλίνδρομο υπόδειγμα AR(p)

Η γενική μορφή ενός αυτοπαλίνδρομου υποδείγματος p τάξεως²², σε αποκλίσεις από το μέσο, έχει τη μορφή :

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \varepsilon_t \quad (3.41)$$

και σε όρους του τελεστή υστέρησης L :

$$(1 - \alpha_1 L - \dots - \alpha_p L^p) y_t = \varepsilon_t \quad (3.42) \quad \text{ή} \quad A(L) y_t = \varepsilon_t \quad (3.43)$$

όπου το y_t έχει μηδενικό μέσο. Για τις συνθήκες στασιμότητας του AR(p) υποδείγματος πράττουμε όπως και στην περίπτωση του AR(2) υποδείγματος. Αναλύουμε το πολυώνυμο σε παράγοντες:

$$A(L) = (1 - \lambda_1 L) \dots (1 - \lambda_p L) = \prod_{j=1}^p (1 - \lambda_j L) \quad (3.44)$$

όπου $\lambda_1^{-1}, \dots, \lambda_p^{-1}$ είναι οι ρίζες του πολυώνυμου $A(L)=0$. Το υπόδειγμα (3.42, 3.43) γράφεται ως:

$$y_t = A^{-1}(L) \varepsilon_t \quad (3.45)$$

Αντικαθιστώντας την (3.44) στην τελευταία λαμβάνουμε :

$$y_t = \frac{1}{(1 - \lambda_1 L) \dots (1 - \lambda_p L)} \varepsilon_t = \sum_{j=1}^p \left(\frac{\xi_j}{1 - \lambda_j L} \right) \varepsilon_t \quad (3.46)$$

²² Δημελή Σ.(2013, Κ.3 σ.84)

Αυτή η διαδικασία συγκλίνει μόνο αν $|\lambda_i| < 1$ για $i = 1, 2, \dots, p$. Αυτές οι συνθήκες ισοδυναμούν με τις συνθήκες $|\lambda_i^{-1}| > 1$, για $i = 1, 2, \dots, p$. Οι ρίζες δηλαδή της χαρακτηριστικής εξίσωσης του $AR(p)$ θα πρέπει να είναι εντός του μοναδιαίου κύκλου ή εναλλακτικά οι ρίζες του πολυωνύμου $A(L)$ στην (3.44) να είναι εκτός του μοναδιαίου κύκλου. Στην περίπτωση που ισχύουν οι συνθήκες στασιμότητας, η σειρά συγκλίνει αφού κάθε κλάσμα στο δεξί μέλος της (3.46) αποτελεί όριο γεωμετρικής προόδου με άπειρους όρους. Μετά την αντικατάσταση των σειρών αυτών προκύπτει ένα υπόδειγμα MA άπειρης τάξης:

$$y_t = (\sum_{j=1}^{\infty} \psi_j L^j) \varepsilon_t \quad (3.47)$$

όπου τα ψ_j , $j=1, 2, \dots$ συμβολίζουν τις σταθμίσεις του $MA(\infty)$. Άρα, τα στάσιμα υποδείγματα $AR(p)$ είναι πάντα αντιστρέψιμα.

Με βάση την ανάλυση που έγινε στα υποδείγματα $AR(1)$ και $AR(2)$, αποδεικνύεται ότι για μια στάσιμη $AR(p)$ σειρά, ο μέσος ισούται με :

$$\mu = E(Y_t) = \delta / (1 - \alpha_1 - \alpha_2 - \dots - \alpha_p) \quad (3.48)$$

Αν πολλαπλασιάσουμε την εξίσωση (3.41) με y_t κατά μέλη και παίρνοντας αναμενόμενες τιμές, προκύπτει η εξίσωση της διακύμανσης:

$$\gamma_0 = \alpha_1 \gamma_1 + \alpha_2 \gamma_2 + \dots + \alpha_p \gamma_p + \sigma_\varepsilon^2 \quad (3.49)$$

Αν θέσουμε $\gamma_0 = \rho_k \gamma_0$ για $k=1, 2, \dots$ τότε για τη διακύμανση έχουμε:

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \alpha_1 \rho_1 - \alpha_2 \rho_2 - \dots - \alpha_p \rho_p} \quad (3.50)$$

Αυτή η εξίσωση μας δίνει τη διακύμανση Y_t ως συνάρτηση των α_j ($j=1, 2, \dots, p$), της διακύμανσης σ_ε^2 του λευκού θορύβου και των συντελεστών αυτοσυσχέτισης ρ_k .

Αν πολλαπλασιάσουμε την εξίσωση (3.43) με y_{t-k} κατά μέλη και παίρνοντας αναμενόμενες τιμές, λαμβάνουμε τις εξισώσεις των αυτοσυνδιακυμάνσεων:

$$\gamma_k = \alpha_1 \gamma_{k-1} + \alpha_2 \gamma_{k-2} + \dots + \alpha_p \gamma_{k-p}, \quad k > 0 \quad (3.51)$$

αφού ισχύει ότι

$$E(y_{t-m} y_{t-k}) = \gamma_{k-m} \quad \text{και} \quad E(y_{t-k} \varepsilon_t) = \mathbf{0} \quad \text{για } k > 0 \quad (3.52)$$

Η (3.52) σε συνδυασμό με την παράσταση $MA(\infty)$ (3.47) για $t-k$ μας δίνει :

$$y_{t-k} = \left(\sum_{j=1}^{\infty} \psi_j L^j \right) \varepsilon_{t-k} = \psi(L) \varepsilon_{t-k} \quad \text{για } k > 0 \quad (3.53)$$

$$\text{Άρα} \quad E(y_{t-k} \varepsilon_t) = \psi(L) E(\varepsilon_{t-k} \varepsilon_t) = 0 \quad (3.54)$$

Κάθε y_{t-k} λοιπόν είναι συνάρτηση όλων των λαθών του παρελθόντος μέχρι και $t-k$, αλλά δεν σχετίζεται με το τρέχον ε_t . Διαιρώντας κατά μέλη την (3.51) με τη διακύμανση γ_0 προκύπτουν οι εξισώσεις των συντελεστών αυτοσυσχέτισης, οι οποίες είναι γνωστές ως εξισώσεις Yule-Walker:

$$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2} + \dots + \alpha_p \rho_{k-p} \quad \text{για } k > 0 \quad (3.55)$$

3.5 Εξισώσεις Yule-Walker και συναρτήσεις αυτοσυσχέτισης

Από τη σχέση (3.55) προκύπτουν οι συντελεστές αυτοσυσχέτισης ρ_k ενός $AR(p)$ υποδείγματος και εκφράζονται σε όρους των συντελεστών αυτοπαλινδρόμησης $\alpha_i (i=1,2,\dots,p)$. Γράφοντας τη σχέση αυτή για κάθε $k=1,2,\dots,p$ και δεδομένου ότι $\rho_0=1$ και $\rho_k = \rho_{-k}$ προκύπτει το σύστημα:

$$\rho_1 = \alpha_1 + \alpha_2 \rho_{k-1} + \dots + \alpha_p \rho_{p-1}$$

$$\rho_2 = \alpha_1 \rho_1 + \alpha_2 + \dots + \alpha_p \rho_{p-2}$$

...

$$\rho_k = \alpha_1 \rho_{p-1} + \alpha_2 \rho_{p-2} + \dots + \alpha_p \quad \text{ή αλλιώς} \quad \rho_p = P_p \varphi_p \quad (3.56)$$

όπου

$$\rho_p = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix}, \quad \varphi_p = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix}, \quad P_p = \begin{bmatrix} \mathbf{1} & \rho_1 & \rho_2 & \cdots & \rho_{p-1} \\ \rho_1 & \mathbf{1} & \rho_1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & \mathbf{1} \end{bmatrix}$$

Λύνοντας το σύστημα (3.56), το οποίο είναι χρήσιμο για την εκτίμηση του διανύσματος των παραμέτρων $a_i, i = 1, \dots, p$ ως προς φ_p προκύπτει το σύστημα:

$$\varphi_p = P_p^{-1} \rho_p \quad (3.56a)$$

Το σύστημα των γραμμικών εξισώσεων $\rho_p = P_p \varphi_p$ αποτελεί το σύστημα εξισώσεων των Yule-Walker²³. Επιλύοντας το σύστημα αυτό επιτυγχάνουμε την εκτίμηση των πρώτων p συντελεστών αυτοσυσχετίσης $\rho_k (k = 1, 2, \dots, p)$ ως συνάρτηση των παραμέτρων $a_i, i = 1, \dots, p$ του αυτοπαλίνδρομου AR(p). Η εκτίμηση των συντελεστών για $k > p$ προκύπτει από την (3.55) αφού αντικατασταθούν σε αυτή οι ήδη γνωστές ρ_k και $\alpha_k (k = 1, 2, \dots, p)$ από το σύστημα των Yule-Walker εξισώσεις (3.56).

Στην περίπτωση που οι θεωρητικές τιμές των αυτοσυσχετίσεων ρ_k δεν είναι γνωστές, αντικαθιστούμε τις δειγματικές τιμές ρ_k , όποτε μπορούμε να υπολογίσουμε τις παραμέτρους του AR(p) χωρίς να μελετήσουμε οικονομετρικά το υπόδειγμα.

Συμπεραίνοντας, σε μια διαδικασία AR(p) οι αυτοσυσχετίσεις ρ_k φθίνουν γεωμετρικά ή κυματιστά ανάλογα με τη μορφή των ριζών (πραγματικές ή μιδαγικές) του πολυωνύμου $A(L)$, ενώ οι συντελεστές μερικής αυτοσυσχετίσης είναι μηδέν μετά το συντελεστή φ_{pp} .

3.6 Υποδείγματα Κινητών Μέσων MA (Moving Average Models)

Τα υποδείγματα κινητών μέσων αρχικά αναπτύχθηκαν από τους Stutsky(1927) και Wold(1938) και είναι χρήσιμα για περιγραφή φαινομένων όπου τα γεγονότα παράγουν ένα άμεσο αποτέλεσμα, η επίδραση του οποίου δε σταματά εκεί ακόμα και αν το γεγονός παύει να υφίσταται. Συνήθως επηρεάζει λιγότερο και για μικρό χρονικό διάστημα τις επόμενες χρονικές στιγμές. Χαρακτηριστικά παραδείγματα είναι οι αργίες, οι απεργίες, το «μποϊκοτάζ», όπου επηρεάζουν την οικονομία όχι μόνο βραχυχρόνια αλλά και

²³ Δημελή Σ.(2013, Κ.3 σ.85)

μακροχρόνια. Στη γενική τους μορφή τα υποδείγματα κινητών μέσων στη γενική τους μορφή γράφονται ως²⁴:

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (3.57)$$

όπου μ είναι η σταθερά της εξίσωσης, τα θ είναι σταθερές παράμετροι και ε_t τα τυχαία σφάλματα που αποτελούν λευκό θόρυβο. Στο υπόδειγμα αυτό υποθέτουμε ότι η χρονολογική σειρά Y_t αποτελεί ένα σταθμισμένο μέσο των τυχαίων σφαλμάτων των q προηγούμενων περιόδων και ονομάζεται υπόδειγμα κινητών μέσων q τάξεως, συμβολιζόμενο ως MA(q).

3.7 Υπόδειγμα Κινητών Μέσων 1^{ης} Τάξης -MA(1) (Moving Average of Order 1)

Για $q = 1$ το υπόδειγμα κινητών μέσων MA²⁵ γράφεται ως εξής:

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (3.58)$$

όπου για απλούστευση $\theta = \theta_1$. Παίρνοντας την αναμενόμενη τιμή κατά μέλη προκύπτει ότι ο μέσος της Y_t ισούται με τη σταθερά μ της εξίσωσης:

$$E(Y_t) = \mu \quad (3.59)$$

αφού $E(\varepsilon_t) = E(\varepsilon_{t-1}) = \mathbf{0}$ λόγω των ιδιοτήτων του λευκού θορύβου.

Το MA(1) υπόδειγμα σε αποκλίσεις από το μέσο γράφεται ως:

$$y_t = Y_t - \mu = \varepsilon_t \theta_{t-1} \text{ , για το οποίο ισχύει } E(y_t) = \mathbf{0} \quad (3.60)$$

²⁴ George E. P. Box, Gwilym M. Jenkins (2015. K.3, σ.68)

²⁵ George E. P. Box, Gwilym M. Jenkins (2015. K.3, σ.70)

Όσον αφορά τη διακύμανση αυτή είναι :

$$\begin{aligned} \gamma_0 &= \text{var}(y_t) = E(Y_t - \mu)^2 \\ &= E(\varepsilon_t)^2 + \theta_1^2 E(\varepsilon_{t-1})^2 - 2\theta_1 E(\varepsilon_t \varepsilon_{t-1}) = (1 + \theta_1^2) \sigma_\varepsilon^2 \end{aligned} \quad (3.61)$$

Εδώ η διακύμανση της Y_t είναι πεπερασμένη και ανεξάρτητη του χρόνου t . Οπότε το MA(1) υπόδειγμα παρουσιάζει στασιμότητα μέσου και διακύμανσης.

Για την αυτοσυνδιακύμανση θα έχουμε :

$$\gamma_k = E(y_t y_{t-k}) = E[y_t (\varepsilon_{t-k} - \theta \varepsilon_{t-k-1})] = \mathbf{0}, k > 1 \quad (3.62)$$

αφού για $k > 1$ παρουσιάζονται μόνο συνδιακυμάνσεις μεταξύ των τιμών του λευκού θορύβου ε_t . Οπότε, η συνάρτηση αυτοσυσχέτισης για το υπόδειγμα MA(1) έχει τη μορφή:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} \frac{-\theta}{1+\theta^2} & \text{για } k=1 \\ \mathbf{0} & \text{για } k > 1 \end{cases} \quad (3.63)$$

Η συνάρτηση αυτοσυσχέτισης για το MA(1) υπόδειγμα γίνεται μηδέν μετά από μια υστέρηση, δηλαδή, η εξάρτηση των τιμών της y_t εξαντλείται στη μια περίοδο. Κάθε τιμή y_t σχετίζεται με την y_{t-1} ή y_{t+1} αλλά όχι με υστερήσεις μεγαλύτερες της μιας περιόδου.

Οι συντελεστές μερικής αυτοσυσχέτισης για το υπόδειγμα MA(1) υπολογίζονται σύμφωνα με τα συστήματα αυτοπαλινδρόμησης ή τις εξισώσεις των Yule και Walker :

$$\begin{aligned} \varphi_{11} &= \rho_1 = -\theta / (1 + \theta^2) \\ \varphi_{22} &= (\rho_2 - \rho_1^2) / (1 - \rho_1^2) = -\theta^2 / (1 + \theta^2 + \theta^4) \end{aligned} \quad (3.64)$$

Παρατηρούμε πως οι συντελεστές φ_{kk} , για $|\theta| < 1$, φθίνουν σχεδόν εκθετικά προς το μηδέν, καθώς το k μεγαλώνει.

Η θεωρητική συνάρτηση αυτοσυσχέτισης ρ_k στο υπόδειγμα MA(1) γίνεται μηδέν ή διακόπτεται μετά την πρώτη χρονική υστέρηση, ενώ η συνάρτηση μερικής αυτοσυσχέτισης φθίνει σταδιακά προς το μηδέν. Αντίθετα, σε ένα στάσιμο AR(1) υπόδειγμα η θεωρητική συνάρτηση ρ_k φθίνει βαθμιαία χωρίς να μηδενίζεται ποτέ, ενώ η συνάρτηση μερικής αυτοσυσχέτισης μηδενίζεται μετά την πρώτη χρονική υστέρηση.

3.8 Υπόδειγμα Κινητών Μέσων 2^{ης} Τάξης -MA(2) (Moving Average of Order 2)

Το υπόδειγμα MA(2)²⁶ θα γράφεται ως εξής:

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \quad (3.65)$$

και σε όρους του τελεστή υστέρησης L :

$$y_t = (1 - \theta_1 L - \theta_2 L^2) \varepsilon_t, \quad y_t = Y_t - \mu \quad (3.66)$$

Όπως και στο υπόδειγμα κινητών μέσω MA(1) αποδεικνύεται πως ο μέσος της σειράς είναι ίσος με τη σταθερά μ και ότι οι ροπές δεύτερης τάξης είναι πεπερασμένες και ανεξάρτητες του χρόνου:

$$\begin{aligned} \gamma_0 &= E(Y_t - \mu)^2 = (1 + \theta_1^2 + \theta_2^2) \sigma_\varepsilon^2 \\ \gamma_1 &= E(y_t y_{t-1}) = -\theta_1 (1 - \theta_2) \sigma_\varepsilon^2 \\ \gamma_2 &= -\theta_2 \sigma_\varepsilon^2 \\ \gamma_k &= 0, \quad \text{για } k > 2 \end{aligned} \quad (3.67)$$

Η θεωρητική συνάρτηση αυτοσυσχέτισης αυτοσυσχέτισης για το υπόδειγμα MA(2) θα είναι:

²⁶ George E. P. Box, Gwilym M. Jenkins (2015. K.3, σ.72)

$$\begin{aligned}
\rho_1 &= -\theta_1 (1 - \theta_2) / (1 + \theta_1^2 + \theta_2^2) \\
\rho_2 &= -\theta_2 / (1 + \theta_1^2 + \theta_2^2) \\
\rho_k &= \mathbf{0}, \quad \text{για } k > 2
\end{aligned} \tag{3.68}$$

Οπότε στο υπόδειγμα MA(2) μόνο οι δύο πρώτοι συντελεστές ρ_1 και ρ_2 δεν είναι μηδενικοί, δηλαδή, y_t επηρεάζεται από τις y_{t-1} και y_{t-2} , αλλά δεν επηρεάζεται από τιμές της y_t μεγαλύτερης υστέρησης.

Οι φ_{kk} συντελεστές μερικής αυτοσυσχέτισης στην περίπτωση του υποδείγματος MA(2) υπολογίζονται από το σύστημα εξισώσεων Yule και Walker αντικαθιστώντας σε αυτούς τις τιμές ρ_k . Εφόσον πληρούνται οι συνθήκες αντιστρεψιμότητας, οι τιμές των συντελεστών φ_{kk} θα φθίνουν προς το μηδέν είτε γεωμετρικά, είτε ημιτονοειδώς, αναλόγως το πρόσημο και το μέγεθος των παραμέτρων θ_1 και θ_2 . Συγκρίνοντας τα παραπάνω χαρακτηριστικά του υποδείγματος MA(2) με τα αντίστοιχα του υποδείγματος AR(2), παρατηρείται μια δυαδικότητα, κάτι που παρατηρείται και στα υποδείγματα πρώτης τάξης.

3.9 Υποδείγματα Κινητών Μέσων Τάξης q -MA(q)

Το γενικό υπόδειγμα MA q τάξεως²⁷ (3.57) έχει παρόμοια χαρακτηριστικά με τα υποδείγματα MA(1) και MA(2). Ακολουθώντας την ίδια διαδικασία, προκύπτει πως τα MA(q) ικανοποιούν τις συνθήκες στασιμότητας αφού:

$$E(Y_t) = \mu$$

$$\gamma_0 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_\varepsilon^2 \tag{3.69}$$

$$\gamma_k = \begin{cases} (-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q) \sigma_\varepsilon^2, & k = 1, \dots, q \\ \mathbf{0} & k > q \end{cases} \tag{3.70}$$

$$\rho_k = \begin{cases} \frac{\sum_{i=1}^{q-k} \theta_i \theta_{i+k} - \theta_k}{1 + \sum_{i=1}^q \theta_i^2} & k = 1, 2, \dots, q \\ \mathbf{0} & k > q \end{cases} \tag{3.71}$$

²⁷ George E. P. Box, Gwilym M. Jenkins (2015. K.3, σ.73)

Η συνάρτηση αυτοσυσχέτισης μηδενίζεται μετά την υστέρηση q τάξεως. Η συνάρτηση αυτή μπορεί να λάβει οποιαδήποτε μορφή για τις πρώτες q υστερήσεις. Οπότε μπορούμε να πούμε πως η τάξη ενός υποδείγματος MA ανταποκρίνεται στο μέγιστο k για το οποίο η αυτοσυσχέτιση ρ_k δεν είναι μηδενική. Όσον αφορά τους συντελεστές μερικής αυτοσυσχέτισης φ_{kk} που προκύπτουν από τις εξισώσεις Yule-Walker, στην περίπτωση των MA υποδειγμάτων δεν μηδενίζονται ποτέ, αλλά φθίνουν προς το μηδέν, καθώς το k αυξάνει, υπό την προϋπόθεση ότι η σειρά Y_t ικανοποιεί τις συνθήκες αντιστρεψιμότητας.

3.10 Αντιστρεψιμότητα

Η ιδιότητα της αντιστρεψιμότητας²⁸ αφορά τη μετατροπή ενός υποδείγματος MA(q) σε υπόδειγμα AR(∞) άπειρης τάξης. Αντίστοιχα ένα υπόδειγμα AR(p) είναι αντιστρέψιμο αν μπορεί να λάβει τη μορφή ενός υποδείγματος MA(∞). Τα υποδείγματα AR(p) είναι αντιστρέψιμα εφόσον είναι στάσιμα. Αντίθετα για τα υποδείγματα MA θα πρέπει να ικανοποιούν κάποιες προϋποθέσεις:

- Το MA(1) υπόδειγμα είναι αντιστρέψιμο αν και εφόσον $|\theta| < 1$.
- Στην περίπτωση του MA(2) υποδείγματος οι συνθήκες αντιστρεψιμότητας ικανοποιούνται όταν ισχύουν οι εξής συνθήκες για τις παραμέτρους θ_1 και θ_2 του υποδείγματος:

$$\begin{aligned}\theta_1 + \theta_2 &< 1 \\ \theta_2 - \theta_1 &< 1 \\ |\theta_2| &< 1\end{aligned}\tag{3.72}$$

- Για την αντιστρεψιμότητα ενός MA(q) υποδείγματος θα πρέπει οι ρίζες της χαρακτηριστικής εξίσωσης του $\Theta(L)$ να είναι σε απόλυτες τιμές μικρότερες της μονάδας ή αλλιώς οι ρίζες του πολυωνύμου $\Theta(L) = 0$ να είναι σε απόλυτες τιμές μεγαλύτερες της μονάδας.

²⁸ Δημελή Σ.(2013, Κ.3 σ.95)

3.11 Σύγκριση AR και MA υποδειγμάτων

Η μελέτη των υποδειγμάτων AR και MA αποδεικνύει μια δυαδικότητα (duality) που χαρακτηρίζει αυτές τις δυο μορφές χρονολογικών σειρών. Από τη μελέτη των χαρακτηριστικών ιδιοτήτων των AR και MA υποδειγμάτων διαπιστώνεται το εξής:

- i. Ένα στάσιμο αυτοπαλίνδρομο υπόδειγμα πεπερασμένης τάξης, έστω $AR(p)$, μπορεί να μετατραπεί σε ένα μη πεπερασμένο υπόδειγμα κινητών μέσων, $MA(\infty)$.
- ii. Ένα αντιστρέψιμο υπόδειγμα κινητών μέσων πεπερασμένης τάξης $MA(q)$ μπορεί να μετατραπεί σε ένα μη πεπερασμένο αυτοπαλίνδρομο υπόδειγμα $AR(\infty)$.
- iii. Το $AR(p)$ υπόδειγμα είναι στάσιμο όταν οι ρίζες του πολυωνύμου $A(L)$ είναι εκτός του μοναδιαίου κύκλου, αλλά είναι πάντα αντιστρέψιμο.
- iv. Το $MA(q)$ υπόδειγμα είναι πάντα στάσιμο, αλλά για να είναι αντιστρέψιμο πρέπει οι ρίζες του πολυωνύμου $\Theta(L)$ να είναι εκτός του μοναδιαίου κύκλου.
- v. Το πεπερασμένο AR υπόδειγμα έχει μια συνάρτηση αυτοσυσχέτισης που το μήκος της είναι άπειρο και φθίνει εκθετικά ή-και ημιτονοειδώς(κυμματιστά), ενώ η συνάρτηση μερικής αυτοσυσχέτισης μηδενίζεται πέρα από ένα συγκεκριμένο σημείο.
- vi. Το πεπερασμένο MA υπόδειγμα έχει μια συνάρτηση αυτοσυσχέτισης που μηδενίζεται πέρα από ένα συγκεκριμένο σημείο, αλλά επειδή είναι ισοδύναμο με ένα άπειρο AR υπόδειγμα, η συνάρτηση μερικής αυτοσυσχέτισης έχει άπειρο μήκος και φθίνει εκθετικά ή-και ημιτονοειδώς.

Τα παραπάνω επιβεβαιώνονται συγκρίνοντας τις συναρτήσεις αυτοσυσχέτισης και μερικής αυτοσυσχέτισης των $AR(2)$ και $MA(2)$ υποδειγμάτων²⁹.

3.12 Μικτά Υποδείγματα ARMA(p,q)

Τα μικτά υποδείγματα ARMA είναι ένας συνδυασμός από αυτοπαλίνδρομα υποδείγματα $AR(p)$ και διαδικασίες κινητού μέσου $MA(q)$. Ένα $ARMA(p,q)$ υπόδειγμα θα είναι στάσιμο αν οι ρίζες του πολυωνύμου $A(L)$ βρίσκονται εκτός του μοναδιαίου κύκλου, δηλαδή οι ρίζες να είναι σε απόλυτες τιμές μεγαλύτερες της μονάδας, και αντιστρέψιμο αν οι ρίζες του πολυωνύμου $\Theta(L)$ βρίσκονται εκτός του μοναδιαίου κύκλου. Στη γενική τους μορφή τα $ARMA(p,q)$ υποδείγματα γράφονται ως εξής:

$$A(L)Y_t = \delta + \theta(L)\varepsilon_t \quad (3.73)$$

²⁹ Δημελή Σ.(2013, Κ.3 σ.99)

Τα στάσιμα και αντιστρέψιμα ARMA υποδείγματα μπορούν να πάρουν τόσο τη μορφή της άπειρης τάξης κινητών μέσων MA(∞), όσο και την άπειρη αυτοπαλίνδρομη μορφή AR(∞). Πράγματι:

$$\begin{aligned}
 Y_t &= \delta A^{-1}(L) + A^{-1}(L)\theta(L)\varepsilon_t = \\
 &= \frac{\delta}{1-a_1L-a_2L^3-\dots-a_pL^p} + \frac{1-\theta_1-\theta_2L^1-\dots-\theta_qL^q}{1-a_1L-a_2L^2-\dots-a_pL^p}\varepsilon_t \\
 &= \frac{\delta}{1-a_1-\dots-a_p} + (\mathbf{1} + \psi_1L + \psi_2L^2+\dots)\varepsilon_t \\
 &= \frac{\delta}{1-a_1-\dots-a_p} + \psi(L)_t \tag{3.74}
 \end{aligned}$$

Όπου το το πολυώνυμο $\psi(L)$ προέκυψε από τη διαίρεση των πολυωνύμων $A(L)$ και $\Theta(L)$. Το ARMA έχει πάρει τη μορφή ενός ARMA(0, ∞) δηλαδή ενός MA(∞)³⁰.

Αντιστρόφως η (3.75) γράφεται ως εξής:³¹

$$\theta^{-1}(L)A(L)Y_t = \delta\theta^{-1}(L) + \varepsilon_t \tag{3.75}$$

Διαιρώντας το πολυώνυμο $A(L)$ με το $\Theta(L)$ προκύπτει το πολυώνυμο $\pi(L) = \theta^{-1}(L)A(L)$ με σταθμίσεις τα $\pi_j, j = 1, 2, \dots$ επομένως:

$$(1+\pi_1L + \pi_2L^2 + \dots)y_t = \delta\theta^{-1}(L) + \varepsilon_t \tag{3.76}$$

Οι σταθμίσεις ψ_j και π_j υπολογίζονται από τις εξισώσεις :

$$A(L)\psi(L) = \Theta(L) \text{ και } \Theta(L)\pi(L)=A(L) \tag{3.77}$$

³⁰ George E. P. Box, Gwilym M. Jenkins(2015. Κ.3, σ.75)

³¹ Tsay (2005, σ.62).

Παρακάτω, στον πίνακα 3.1 παρουσιάζονται συνοπτικά τα χαρακτηριστικά όλων των ARMA υποδειγμάτων, τα οποία μας βοηθούν στο να κατανοήσουμε τη μορφή του υποδείγματος που είναι πιθανότερο να δημιουργήσει το δείγμα παρατηρήσεων που μελετάμε.

Πίνακας 3.1 - Χαρακτηριστικά των ARMA(p,q) Υποδειγμάτων

Υπόδειγμα	Συνάρτηση Αυτοσυσχέτισης (ACF)	Συνάρτηση Μερικής Αυτοσυσχέτισης (PACF)
Λευκός θόρυβος	Όλες μηδέν: $\rho_k = 0$	Όλες μηδέν: $\varphi_{kk} = 0$
AR(1)	Φθίνουν προς το μηδέν γεωμετρικά (εκθετικά) από $\rho_0 = 1$ <ul style="list-style-type: none"> • ευθέως αν $\alpha > 0$ • εναλλάξ (-, +) αν $\alpha < 0$ 	Μηδέν μετά το $\varphi_{kk} = \rho_1$
AR(2)	Φθίνουν προς το μηδέν από ρ_2 <ul style="list-style-type: none"> • γεωμετρικά για πραγματικές ρίζες • κυματιστά (ημιτονοειδώς) για μιγαδικές ρίζες 	Μηδέν μετά το φ_{22}
AR(p)	Φθίνουν προς το μηδέν γεωμετρικά (ευθέως ή εναλλάξ) ή/και κυματιστά	Μηδέν μετά το φ_{pp}
MA(1)	Μηδέν μετά το ρ_1	Φθίνει σχεδόν γεωμετρικά από το φ_{11}
MA(2)	Μηδέν μετά το ρ_2	Φθίνει γεωμετρικά (ευθέως ή εναλλάξ) ή/και κυματιστά
MA(q)	Μηδέν μετά το ρ_q	Φθίνει γεωμετρικά (ευθέως ή εναλλάξ) ή/και κυματιστά
ARMA(1,1)	Φθίνει γεωμετρικά από το ρ_1 <ul style="list-style-type: none"> • ευθέως αν $\alpha > 0$ • εναλλάξ (-, +) αν $\alpha < 0$ 	Φθίνει γεωμετρικά από το φ_{11} <ul style="list-style-type: none"> • ευθέως αν $\alpha > 0$ • εναλλάξ (-, +) αν $\alpha < 0$
ARMA(p,q)	Φθίνει γεωμετρικά ή/και κυματιστά από το ρ_q	Φθίνει γεωμετρικά (ευθέως ή εναλλάξ) ή/και κυματιστά από το φ_{pp}

Στον πίνακα 3.2 συνοψίζονται οι συνθήκες στασιμότητας και αντιστρεψιμότητας των AR, MA και ARMA υποδειγμάτων.

Πίνακας 3.2 – Συνθήκες Στασιμότητας και Αντιστρεψιμότητας των ARMA Υποδειγμάτων

Υπόδειγμα	Εξίσωση	Στασιμότητα	Αντιστρεψιμότητα
AR(1)	$y_t = \alpha_1 y_{t-1} + \varepsilon_t$	$ \alpha_1 < 1$	Πάντα αντιστρέψιμο εφόσον στάσιμο
AR(2)	$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t$	$\alpha_1 + \alpha_2 < 1$ $\alpha_2 - \alpha_1 < 1$ $ \alpha_2 < 1$	Πάντα αντιστρέψιμο εφόσον στάσιμο
AR(p)	$A(L)y_t = \varepsilon_t$	Οι ρίζες του $A(L)$ να κείνται εκτός του μοναδιαίου κύκλου	Πάντα αντιστρέψιμο εφόσον στάσιμο
MA(1)	$y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$	Πάντα στάσιμο	$ \theta_1 < 1$
MA(2)	$y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$	Πάντα στάσιμο	$\theta_1 + \theta_2 < 1$ $\theta_2 - \theta_1 < 1$ $ \theta_2 < 1$
MA(q)	$y_t = \theta(L)\varepsilon_t$	Πάντα στάσιμο	Οι ρίζες του $\theta(L)$ να κείνται εκτός του μοναδιαίου κύκλου
ARMA(1,1)	$y_t = \alpha_1 y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$	$ \alpha_1 < 1$	$ \theta_1 < 1$
ARMA(p,q)	$A(L)y_t = \theta(L)\varepsilon_t$	Οι ρίζες του $A(L)$ να κείνται εκτός του μοναδιαίου κύκλου	Οι ρίζες του $\theta(L)$ να κείνται εκτός του μοναδιαίου κύκλου

Ο μέσος της ARMA(p,q) διαδικασίας προκύπτει παίρνοντας αναμενόμενες τιμές κατά μέλη στην (3.73):

$$\begin{aligned} E(Y_t) &= E[\delta A^{-1}(L)] + A^{-1}(L)\theta(L)E(\varepsilon_t) \\ &= \delta / (1 - \alpha_1 - \alpha_2 - \dots - \alpha_p) \end{aligned} \quad (3.78)$$

Αφού $a_j L^j = a_j$ και συμπίπτει με το μέσο (3.48) του AR(p) υποδείγματος. Οι συναρτήσεις αυτοσυσχέτισης των ARMA(p,q) υποδειγμάτων είναι αρκετά σύνθετες στις πρώτες q υστερήσεις, αφού είναι λύσεις εξισώσεων διαφόρων υψηλής τάξης. Επειδή το υπόδειγμα MA(q) έχει "μνήμη" μόνο q περιόδων, η συνάρτηση αυτοσυσχέτισης μετά από q περιόδους συμπεριφέρεται ως συνάρτηση αυτοσυσχέτισης ενός AR υποδείγματος. Παρομοίως, η συνάρτηση μερικής αυτοσυσχέτισης μετά από p υστερήσεις συμπεριφέρεται ως μια AR διαδικασία. Η συνάρτηση αυτοσυσχέτισης ανακύπτει με τον ίδιο τρόπο όπως και για τα AR υποδείγματα (βλ. 3.56). Γράφοντας την (3.73) σε αποκλίσεις από το μέσο προκύπτει:

$$y_t = a_1 y_{t-1} + \dots + a_p y_{t-p} - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (3.79)$$

Μετά από πολλαπλασιασμό κατά μέλη με y_{t-k} προκύπτει η εξίσωση διαφορών:

$$\begin{aligned} y_k &= a_1 y_{t-1} + \dots + a_p y_{t-p} - \theta_1 E(y_{t-k} \varepsilon_{t-1}) - \theta_2 E(y_{t-k} \varepsilon_{t-2}) - \dots \\ &\quad - \theta_q E(y_{t-k} \varepsilon_{t-q}) + E(y_{t-k} \varepsilon_t) \end{aligned} \quad (3.80)$$

Από τη μορφή MA(∞) του υποδείγματος $y_{t-k} = \psi(L)\varepsilon_{t-k}$ προκύπτει πως το y_{t-k} εξαρτάται από όλα τα προηγούμενα τυχαία σφάλματα μέχρι το χρόνο t-k. Επομένως:

$$\begin{aligned} E(y_{t-k} \varepsilon_t) &= \mathbf{0} && \text{για } k > 0, \\ E(y_{t-k} \varepsilon_t) &= \psi_{-k} \sigma_\varepsilon^2 && \text{για } k \leq 0 \end{aligned} \quad (3.81)$$

Οι αυτοσυνδιακυμάνσεις γ_k προκύπτουν από την εξίσωση :

$$\gamma_k = a_1\gamma_{k-1} + \dots + a_p\gamma_{k-p} - \sigma_\varepsilon^2(\theta_k\psi_0 + \theta_{k+1}\psi_1 + \dots + \theta_q\psi_{q-k}),$$

$$\theta_0 = -1 \tag{3.82}$$

και η συνάρτηση αυτοσυσχέτισης $\rho_k = \gamma_k/\gamma_0$ θα έχει τη μορφή:

$$\rho_k = a_1\rho_{k-1} + a_2\rho_{k-2} + \dots + a_p\rho_{k-p}, \quad k > q$$

$$A(L)\rho_k = 0, \quad k > q \tag{3.83}$$

3.13 Συμπεράσματα για τα υποδείγματα ARMA(p,q)

Σύμφωνα με τους πίνακες 3.1 και 3.2 όπου συνοψίζονται τα κυριότερα χαρακτηριστικά και οι συμπεριφορές των AR(p), MA(q) και ARMA(p,q) υποδειγμάτων καθώς και την ανάλυση που έγινε στα υποδείγματα αυτά προκύπτουν τα εξής συμπεράσματα.³²

Οι πρώτοι $k=q$ συντελεστές αυτοσυσχέτισης $\rho_1, \rho_2, \dots, \rho_q$ εξαρτώνται άμεσα και από τις παραμέτρους a του AR μέρους και από τις παραμέτρους θ του MA μέρους και αυτό προκύπτει από την (3.82).

Η συνάρτηση αυτοσυσχέτισης (3.82) είναι ίδια με αυτή του υποδείματος AR(p) για τιμές του k μεγαλύτερες από το q . Η ρ_k δηλαδή, μετά την q υστέρηση έχει συμπεριφορά μιας ξεκάθαρα AR διαδικασίας. Για αρχικές τιμές της εξίσωσης διαφορών (3.82) που είναι απαραίτητες για τον υπολογισμό των ρ_k , $k > q$, χρησιμοποιούνται οι p τιμές $\rho_q, \rho_{q-1}, \dots, \rho_{q-p+1}$ οι οποίες χαρακτηρίζουν πλήρως τους επόμενους συντελεστές αυτοσυσχέτισης.

Για τον προσδιορισμό της τάξης (p,q) του ARMA υποδείματος εξετάζεται:

- i. Αν $q-p < 0$, τότε η συνάρτηση αυτοσυσχέτισης $\rho_k, k=1,2,\dots$ θα φθίνει συνεχώς με εκθετικό τρόπο ή- και ημιτονοειδή τρόπο και οι τιμές της ρ_k θα εξαρτώνται από το είδος των ριζών του πολωνύμου $A(L)$ και από τις αρχικές τιμές της.

³² Βλέπε Δημελή Σ.(2013,Κ.3 σ.109-110)

- ii. Αν $q-p > 0$, τότε θα υπάρχουν $q-p+1$ αρχικές τιμές του $\rho(\rho_0, \rho_1, \rho_2, \dots, \rho_{q-p})$ που δεν ακολουθούν αυτό το σχήμα.

Η συνάρτηση μερικής αυτοσυσχέτισης φθίνει προς το άπειρο. Επειδή το υπόδειγμα μπορεί να λάβει τη μορφή (3.75) ενός αυτοπαλίνδρομου υποδείγματος άπειρης τάξης, οι συντελεστές αυτοσυσχέτισης φ_{kk} δεν διακόπτονται, αλλά συνεχίζουν μειούμενοι. Τα φ_{kk} για $k > p$ συμπεριφέρονται όπως και οι συντελεστές αυτοσυσχέτισης ενός MA υποδείγματος αφού η επίδραση του AR μέρους διακόπτεται στην p υστέρηση. Συμπεραίνεται λοιπόν ότι οι συντελεστές μερικής αυτοσυσχέτισης του ARMA(p, q) υποδείγματος θα φθίνουν μονοτονικά ή- και κυμματιστά από την p υστέρηση, συμπεριφερόμενη από εκείνη τη στιγμή και μετά ως μια MA διαδικασία.

Παρακάτω εξετάζονται τα μη στάσιμα στοχαστικά υποδείγματα. Σε αυτή την περίπτωση είναι απαραίτητος ο μετασχηματισμός της σειράς ώστε να διασφαλίζεται η στασιμότητα και εφαρμογή της μεθολογίας Box και Jenkins (1976) για την επιλογή του καταλληλότερου υποδείγματος πρόβλεψης μιας μη στάσιμης χρονολογικής σειράς.

Κεφάλαιο 4^ο: Μεθοδολογία Box-Jenkins-Υποδείγματα ARIMA

Η ανάλυση που έγινε στα υποδείγματα AR, MA και ARMA στα προηγούμενα κεφάλαια αναφέρεται σε στάσιμες διαδικασίες, που σημαίνει ότι ο μέσος και η διακύμανση δεν εξαρτώνται από τον χρόνο t , ενώ οι αυτοσυνδιακυμάνσεις εξαρτώνται μόνο από τη χρονική υστέρηση s . Σε αυτό το κεφάλαιο θα εξεταστούν οι μη στάσιμες διαδικασίες καθώς και μέθοδοι για μετατροπή αυτών των διαδικασιών σε στάσιμες.

Μη στάσιμες είναι συνήθως οι χρονολογικές σειρές που παρουσιάζουν τάση ή εποχικές διακυμάνσεις, όπως για παράδειγμα οι χρονολογικές σειρές του Ακαθάριστου Εγχώριου Προϊόντος, των Επενδύσεων, των Δημόσιων Δαπανών, και σειρές τυχαίων διαδρομών (random walk). Πολλές από αυτές τις μη στάσιμες διαδικασίες εμφανίζουν μεταβολές στο μέσο κατά διαστήματα, αλλά διατηρούν παρόλα αυτά μια ομοιογένεια, δηλαδή ο μέσος αλλάζει επίπεδο διαχρονικά, αλλά η χρονολογική σειρά συμπεριφέρεται ομοιόμορφα στα μεμονωμένα χρονικά διαστήματα.

Όπως αναφέρθηκε παραπάνω, υπάρχει τρόπος μετατροπής μιας μη στάσιμης χρονοσειράς σε στάσιμη και αυτό επιτυγχάνεται με τη μεθοδολογία Box-Jenkins, μετατρέποντας τις διαδικασίες σε στάσιμες με τη χρήση πρώτων, δεύτερων, κ.οκ. διαφορών. Όταν εξασφαλιστεί η στασιμότητα με τις d διαφορές τότε ακολουθεί η ανάλυση προσαρμογής του κατάλληλου ARMA(p,q) υποδείγματος στη μετασηματισμένη σειρά. Ένα υπόδειγμα που ακολουθεί στάσιμη διαδικασία ARMA(p,q) μετά από διαφορές d τάξεως αποτελεί μια νέα κατηγορία υποδειγμάτων, που ονομάζονται αυτοπαλίνδρομα ολοκληρωμένα υποδείγματα κινητών μέσων (Autoregressive Integrated Moving Average) τάξεως (p,d,q) και συμβολίζεται ως ARIMA(p,d,q).

4.1 Μετατροπή μη στάσιμης διαδικασίας σε στάσιμη με Λήψη Διαφορών – Διαφοροποίηση (Differencing)

Στην περίπτωση που η χρονοσειρά δεν είναι στάσιμη, το ARMA υπόδειγμα (3.73) δεν ικανοποιεί τις συνθήκες στασιμότητας. Οπότε, θεωρώντας τη μορφή (3.73) υποθέτοντας αυτή τη φορά ότι η Y_t έχει μια ρίζα ίση με τη μονάδα, τότε η μοναδιαία ρίζα μπορεί να παραγοντοποιηθεί αλγεβρικά:

$$\Phi(L)Y_t = A(L)(1-L)Y_t = \delta + \theta(L)\varepsilon_t \quad (4.1)$$

όπου το $A(L)$ συμβολίζει το «στάσιμο» πολυώνυμο, αυτό δηλαδή που έχει ρίζες εκτός του μοναδιαίου κύκλου και το $\Phi(L) = A(L)(1-L)$ που συμβολίζει το «μη στάσιμο» πολυώνυμο, αυτό δηλαδή που έχει ρίζες εντός του μοναδιαίου κύκλου.

Το υπόδειγμα (4.1) γράφεται και ως:

$$A(L)\Delta Y_t = \delta + \theta(L)\varepsilon_t \quad (4.2)$$

όπου το $\Delta Y_t = (1-L)Y_t = Y_t - Y_{t-1}$ εκφράζει τις πρώτες διαφορές της σειράς. Σε αυτή την περίπτωση λέγεται ότι η σειρά Y_t είναι ολοκληρωμένη πρώτης τάξης.

Αν μια χρονολογική σειρά μετατραπεί σε στάσιμη σειρά μετά από διαφορές d τάξεως³³, τότε ονομάζεται ολοκληρωμένη (integrated) d τάξεως και συμβολίζεται με $I(d)$. Το υπόδειγμα (4.1) αποτελεί ένα ολοκληρωμένο υπόδειγμα πρώτης τάξης ARMA(p,q), δηλαδή ένα ARIMA(p,1,q).

Γενικά, αν το πολυώνυμο $\Phi(L)$ έχει d ρίζες που ισούνται με τη μονάδα, ενώ οι υπόλοιπες ρίζες βρίσκονται εκτός του μοναδιαίου κύκλου, τότε έχουμε το αυτοπαλίνδρομο ολοκληρωμένο υπόδειγμα κινητών μέσων ARIMA(p,d,q):

$$\Phi(L)Y_t = A(L)(1-L)^d Y_t = \delta + \theta(L)\varepsilon_t \quad (4.3)$$

$$\Phi(L) = A(L)(1-L)^d \quad (4.4)$$

³³ Δημελή Σ.(2013, Κ.4 σ.118)

Επομένως το χαρακτηριστικό στοιχείο αυτού του μη στάσιμου ARIMA υποδείγματος της Y_t είναι ότι μπορεί να λάβει τη μορφή ενός στάσιμου ARMA υποδείγματος της στοχαστικής διαδικασίας $(1 - L)^d Y_t$, δηλαδή των διαφορών d τάξεως της Y_t . Συμβολίζοντας αυτές τις διαφορές με w_t προκύπτει:

$$w_t = \Delta^d Y_t = (1 - L)^d Y_t \quad (4.5)$$

και αντιστρέφοντας προκύπτει:

$$Y_t = S^d w_t, \quad \text{όπου } S = (1 - L)^{-1} = 1 + L + L^2 + \dots \quad (4.6)$$

Το S είναι το άθροισμα άπειρων όρων ή διαφορετικά το ολοκλήρωμα του τελεστή υστέρησης L και αυτό δείχνει ότι η μη στάσιμη χρονολογική σειρά Y_t μπορεί να προκύψει αθροίζοντας – ολοκληρώνοντας τη στάσιμη σειρά w_t κατά d φορές. Από τους Box και Jenkins προτείνεται οι σειρές να μετατρέπονται σε στάσιμες παίρνοντας διαφορές πρώτης (οι οποίες αφαιρούν τη γραμμική τάση), δεύτερης (οι οποίες αφαιρούν τη δευτεροβάθμια τάση) ή και μεγαλύτερης τάξης (γενικά οι διαφορές d τάξεως αφαιρούν τάση πολυωνυμικής μορφής τάξεως d).³⁴

Η διαφοροποίηση (differencing) μια χρονολογικής σειράς για να μετατραπεί σε στάσιμη προτείνεται όταν η σειρά εμφανίζει στοχαστική τάση ή/και εποχικότητα, όπου στην περίπτωση εποχικότητας οι διαφορές πρέπει να λαμβάνονται στις εποχικές περιόδους (ανά δωδεκάμηνο αν τα στοιχεία είναι μηνιαία κ.ο.κ.). Αφού η στασιμότητα εξασφαλιστεί με τις απαιτούμενες διαφορές τότε ακολουθεί η ανάλυση προσαρμογής του κατάλληλου ARMA(p,q) υποδείγματος στη τροποποιημένη σειρά.

Μια χρονολογική σειρά μπορεί να είναι στάσιμη στο μέσο, αλλά να είναι μη στάσιμη στη διακύμανση. Στην περίπτωση που η χρονολογική σειρά έχει αυξανόμενη διακύμανση χρησιμοποιούνται άλλοι μετασχηματισμοί για επίτευξη στασιμότητας, όπως ο μετασχηματισμός δύναμης (power transformation) που προτάθηκε από τους Box και Cox (1964). Η διαφοροποίηση και ο μετασχηματισμός δύναμης μπορούν να εφαρμοσθούν μόνο για θετικές τιμές της χρονολογικής σειράς. Στην περίπτωση αρνητικών ή μηδενικών τιμών, πρέπει πρώτα να προστεθεί σε όλες τις τιμές της χρονοσειράς ένας σταθερός αριθμός, αρκετά μεγάλος, ώστε να προκύψουν μόνο θετικοί αριθμοί και μετά να γίνει ο μετασχηματισμός.

³⁴ Δημελή Σ.(2013, Κ.9)

4.2 Εναλλακτικές Μορφές Υποδείγματος ARIMA

Το γενικό υπόδειγμα ARIMA(p, d, q) μπορεί να λάβει τρεις ισοδύναμες μορφές³⁵.

- **Μορφή Α:** Συνάρτηση ενός αριθμού προηγούμενων τιμών Y_{t-j} , $j=1, p+d$, του τρέχοντος τυχαίου σφάλματος ε_t και ενός αριθμού προηγούμενων σφαλμάτων ε_{t-j} , $j=1, \dots, q$. Η μορφή αυτή είναι αποτελεί μια εξίσωση διαφοράς (difference equation form).
- **Μορφή Β:** Συνάρτηση του τρέχοντος τυχαίου σφάλματος ε_t καθώς και όλων των προηγούμενων τυχαίων σφαλμάτων ε_{t-j} , $j = 1, 2, \dots$. Η μορφή αυτή είναι γνωστή ως τυχαία διαταραχή (random shock form).
- **Μορφή Γ:** Συνάρτηση όλων των προηγούμενων τιμών Y_{t-j} , $j=1, 2, \dots$ της χρονολογικής σειράς και του τρέχοντος τυχαίου σφάλματος ε_t . Η μορφή αυτή είναι γνωστή ως η αντίστροφη μορφή (inverted form).

4.3 Παραδείγματα Υποδειγμάτων ARIMA

Παρακάτω παρουσιάζονται μερικά υποδείγματα³⁶ ARIMA τα οποία συναντώνται πιο συχνά στις εφαρμογές των χρονοσειρών, όπως το ARIMA(1,1,1), το ARIMA(1,1,0) και μορφές υψηλότερης τάξης όπως ARIMA(0,2,2) και ARIMA(1,2,0). Συνήθως δεν παρατηρούνται περιπτώσεις που να έχουν τιμές για p,q ή d μεγαλύτερες του 2. Αντιθέτως, στις περισσότερες περιπτώσεις κάποια από αυτές τις παραμέτρους είναι μηδενική.

- **Υπόδειγμα ARIMA(1,1,1)**

Θεωρείται το υπόδειγμα (4.2), για $p=1$ και $q=1$ ισχύει:

$$\Phi(L)y_t = (1 - aL)(1 - L)y_t = (1 - \theta L)\varepsilon_t \quad (4.7)$$

Και

$$\Phi(L) = 1 - (1 + a)L + aL^2 \quad (4.8)$$

³⁵ Δημελή Σ.(2013, Κ.4 σ.120)

³⁶ Δημελή Σ.(2013, Κ.4 σ.122)

Άρα το υπόδειγμα ARIMA(1,1,1) γράφεται με τον εξής τρόπο:

$$Y_t = \delta + (1 + \alpha)Y_{t-1} - \alpha Y_{t-2} + \varepsilon_t - \theta \varepsilon_{t-1} \quad (4.9)$$

· **Υπόδειγμα ARIMA(0,1,1)**

Το υπόδειγμα ARIMA(0,1,1) λόγω απουσίας του αυτοπαλινδρομικού μέρους λέγεται και IMA(1,1) και έχει τη μορφή:

$$(1 - L)Y_t = \delta + \varepsilon_t - \theta \varepsilon_{t-1} \quad (4.10)$$

ή

$$Y_t = \delta + Y_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1} \quad (4.11)$$

Το μπορεί να εκφραστεί σε όρους του τωρινού και των προηγούμενων τυχαίων σφαλμάτων ε_t λευκού θορύβου μετά από διαδοχικές αντικαταστάσεις των υστερήσεων της Y_t στην (4.11). Η σειρά είναι μη στάσιμη και για αυτό το λόγο είναι πιο δύσκολο να λάβουμε τη μορφή MA του υποδείγματος (4.11). Οι μη στάσιμες σειρές δεν βρίσκονται σε ισορροπία και έτσι δεν μπορεί να υποτεθεί ότι εκτείνονται μέχρι το $-\infty$ ή ότι αρχίζουν από το $-\infty$.³⁷

Έτσι υποθέτουμε ότι οι χρονολογικές σειρές αυτές αρχίζουν από κάποιο χρόνο $t = -m$ και ότι αυτός προηγείται του χρόνου $t = 1$ που παρατηρούμε για πρώτη φορά στη χρονολογική σειρά. Οπότε μπορούμε να μετατρέψουμε την (4.11) σε μορφή κινητών μέσων και για απλούστευση θεωρούμε το μέσο της Y_t μηδενικό:

$$Y_t = \varepsilon_t + (1 - \theta)\varepsilon_{t-1} + (1 - \theta)\varepsilon_{t-2} + \dots + (1 - \theta)\varepsilon_{-m} - \theta\varepsilon_{-m-1} \quad (4.12)$$

Από την (4.12) υπολογίζονται η διακύμανση και οι αυτοσυσχετίσεις των τιμών της Y_t :

$$var(Y_t) = [1 + \theta^2 + (1 + \theta)^2(t + m)]\sigma_\varepsilon^2 \quad (4.13)$$

³⁷ Cryer and Chan, 2008 K.5 σ.96

$$\rho_k = \frac{1-\theta+\theta^2+(1-\theta)^2(t+m-k)}{\sqrt{\text{var}(Y_t)}\sqrt{\text{var}(Y_{t-k})}} \approx \sqrt{\frac{t+m-k}{t+m}} \quad (4.14)$$

Οι τιμές της Y_t (4.12) προκύπτουν από το άθροισμα ενός μεγάλου αριθμού ισοσταθμίσεων τυχαίων σφαλμάτων, διότι αντίθετα με τα ARMA υποδείγματα, οι σταθμίσεις των όρων του λευκού θορύβου δεν φθίνουν προς το μηδέν, αλλά σε αυτή την περίπτωση είναι ισοβαρείς. Η διακύμανση είναι συνάρτηση του χρόνου και αυξάνει καθώς απομακρυνόμαστε στο χρόνο, ενώ οι αυτοσυσχετίσεις είναι θετικές και πολύ ισχυρές. Για μεγάλες τιμές των m και k , οι συντελεστές αυτοσυσχετίσης τείνουν στη μονάδα.

· Υπόδειγμα ARIMA(0,2,2)

Το υπόδειγμα ARIMA(0,2,2) αν θεωρήσουμε ότι έχει μηδενικό μέσο, γράφεται ως εξής:

$$(1-L)^2 Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \quad (4.15)$$

Η εξίσωση διαφορών της Y_t προκύπτει αναδιατάσσοντας τους όρους στην (4.15):

$$Y_t = 2Y_{t-1} - Y_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \quad (4.16)$$

Το υπόδειγμα αυτό μπορεί να γραφτεί σε όρους του τρέχοντος και των προηγούμενων σφαλμάτων του λευκού θορύβου αν γίνει πάλι η υπόθεση ότι η χρονολογική σειρά αρχίζει από κάποιο χρόνο $t = -m$:

$$Y_t = \varepsilon_t + \sum_{j=1}^{t+m} \psi_j \varepsilon_{t-j} - [(t+m+1)\theta_1 + (t+m)\theta_2] \varepsilon_{-m-1} - (t+m+1)\theta_2 \varepsilon_{-m-2} \quad (4.17)$$

όπου $\psi_j = 1 + \theta_2 + (1 - \theta_1 - \theta_2)_j$ για $j = 1, 2, 3, \dots, t+m$

Και σε αυτή την περίπτωση οι σταθμίσεις δεν φθίνουν, αλλά αποτελούν γραμμική συνάρτηση του j . Η διακύμανση που υπολογίζεται από την (4.17) διαπιστώνεται πως αυξάνει γρήγορα με το χρόνο και οι αυτοσυσχετίσεις τους πλησιάζουν τη μονάδα για αποστάσεις k μετρίου μεγέθους.

· **Υπόδειγμα ARIMA(1,2,0)**

Το υπόδειγμα αυτό έχει τη μορφή :

$$(1 - aL)(1 - L)^2 y_t = \varepsilon_t \quad (4.18)$$

Το υπόδειγμα αυτό αποτελεί ένα στάσιμο AR(1) υπόδειγμα της μετασχηματισμένης σειράς των διαφορών δεύτερης τάξης:

$$(1 - L)^2 y_t = \Delta(y_t - y_{t-1}) = \Delta^2(y_t) \quad (4.19)$$

Και γράφοντας το υπόδειγμα (4.18) ως προς y_t :

$$y_t = (2 - \alpha)y_{t-1} - (1 + 2\alpha)y_{t-2} + \alpha y_{t-3} + \varepsilon_t \quad (4.20)$$

Που είναι μη στάσιμο ARMA(3,0) υπόδειγμα.

4.4 Εποχικά Υποδείγματα SARIMA

Στοιχεία μικρότερης διάρκειας του έτους όπως μηνιαία, τριμηνιαία κ.οκ. φανερώνουν εποχικότητα³⁸. Το εποχικό μέρος του υποδείγματος ARIMA έχει παρόμοια δομή με αυτή ενός μη εποχικού υποδείγματος και γράφεται ως SARIMA (Sp, Sd, Sq) x ARIMA(p, d, q) . Στο εποχικό μέρος γίνονται πολλαπλασιασμοί της χρονικής υστέρησης S (S ο αριθμός περιόδων για μια εποχή) με τους συντελεστές p που είναι ο αριθμός των αυτοπαλίνδρομων εποχικών όρων (SAR) με τον αριθμό εποχικών διαφορών d και με το εύρος των στοιχείων των εποχικών όρων του κινητού μέσου q (SMA). Το φαινόμενο της εποχικότητας αποτελεί μια κανονική κύμανση μέσα στο χρονολογικό έτος οδηγώντας σε υψηλή συσχέτιση ανάμεσα στις τιμές της σειράς που αντιστοιχούν στην ίδια περίοδο ανάμεσα στα διαφορετικά έτη. Ένας τρόπος αντιμετώπισης της εποχικότητας είναι η αφαίρεσή της από τη χρονολογική σειρά και έπειτα η εφαρμογή της μεθοδολογίας Box – Jenkins. Μια άλλη μέθοδος, η οποία εφαρμόζεται περισσότερο, αφορά την ενσωμάτωση του εποχικού πρότυπου των στοιχείων μας στο κανονικό υπόδειγμα ARIMA και την εφαρμογή της μεθοδολογίας Box – Jenkins. Αυτή η μέθοδος έχει ως αποτέλεσμα την εκτίμηση υποδειγμάτων με περισσότερους παραμέτρους.

³⁸ Κουνετάς Κ.(2012, σ.91)

4.5 Μεθοδολογία Box – Jenkins

Η προσέγγιση των Box και Jenkins(1976) στην ανάλυση χρονολογικών σειρών, είναι μια μέθοδος εύρεσης ενός στατιστικού υποδείγματος ARIMA που να παριστάνει ικανοποιητικά τη στοχαστική διαδικασία από την οποία προήλθαν τα δεδομένα, δηλαδή το δείγμα μας. Η μέθοδος αυτή εφαρμόζεται σε μη στάσιμες χρονοσειρές και απαρτίζεται από τέσσερα στάδια, την ταυτοποίηση (identification), την εκτίμηση (estimation), το διαγνωστικό έλεγχο (diagnostic checking) και την πρόβλεψη (forecasting). Με λίγα λόγια η μεθοδολογία Box- Jenkins είναι η ανάπτυξη και η κατασκευή υποδειγμάτων ARIMA, τα οποία είναι γραμμικά μοντέλα που μπορούν να περιγράψουν ικανοποιητικά τις συνιστώσες της χρονολογικής σειράς³⁹.

4.5.1 Ταυτοποίηση

Στο στάδιο της ταυτοποίησης του υποδείγματος πρέπει αρχικά να προσδιοριστεί η τάξη της μη στασιμότητας και η τάξη των AR ή/και πολυωνύμων. Για να το πετύχουμε αυτό συγκρίνουμε τη μορφή των δειγματικών συναρτήσεων αυτοσυσχέτισης και μερικής αυτοσυσχέτισης με τη μορφή θεωρητικών συναρτήσεων αυτοσυσχέτισης και μερικής αυτοσυσχέτισης που αντιστοιχούν σε διαδικασίες με άπειρο πλήθος ορών. Σε αυτό το στάδιο γίνεται η εξειδίκευση ενός ARIMA υποδείγματος με βάση τις πληροφορίες που λαμβάνονται από το δείγμα. Αυτό σημαίνει ότι καθορίζονται οι τιμές των d , p και q . Καθορίζεται δηλαδή, ο αριθμός d των διαφορών που απαιτούνται για να μετατραπεί η σειρά σε στάσιμη και στη συνέχεια καθορίζεται η τάξη p της αυτοπαλίνδρομης διαδικασίας και η τάξη q της διαδικασίας κινητού μέσου. Για να διαπιστωθεί η στασιμότητα ή όχι της σειράς, εξετάζεται η συμπεριφορά της δειγματικής συνάρτησης αυτοσυσχέτισης. Στην περίπτωση που οι αυτοσυσχετίσεις συγκλίνουν με γρήγορο ρυθμό προς το μηδέν, σημαίνει ότι η σειρά είναι μάλλον στάσιμη. Αντίθετα, αν οι αυτοσυσχετίσεις φθίνουν με αργό ρυθμό, είναι ένδειξη ότι η σειρά είναι μη στάσιμη και πρέπει να μετατραπεί σε στάσιμη. Η μετατροπή σε αυτή την περίπτωση γίνεται παίρνοντας τις πρώτες ή τις δεύτερες κ.οκ. διαφορές. Μετά τη μετατροπή προσδιορίζεται η τάξη του υποδείγματος ARIMA, δηλαδή προσδιορίζονται οι τιμές p και q .

³⁹ Θαλασσινός Ε.(1991, Κ.4 σ.95)

4.5.2 Εκτίμηση

Το δεύτερο στάδιο στη μεθοδολογία Box – Jenkins είναι αυτό της εκτίμησης⁴⁰ των αυτοπαλίνδρομων παραμέτρων $\alpha_1, \alpha_2, \dots, \alpha_p$ καθώς και των παραμέτρων $\theta_1, \theta_2, \dots, \theta_p$ των κινητών μέσων του ARIMA(p,d,q) υποδείγματος που ταυτοποιήθηκε στο προηγούμενο στάδιο. Η εκτίμηση ενός υποδείγματος ARIMA (p, d, q) προϋποθέτει την ύπαρξη ενός μεγάλου δείγματος παρατηρήσεων Y_1, Y_2, \dots, Y_N της χρονοσειράς Y_t . Αυτό το υπόδειγμα ισοδυναμεί με ένα στάσιμο ARMA(p, q) στις διαφορές d τάξεως του Y_t . Για λόγους απλούστευσης μπορεί να θεωρηθεί ότι η σειρά Y_t είναι στάσιμη αναλύοντας τις μεθόδους εκτίμησης των παραμέτρων των στάσιμων ARMA (p, q).

Η δυσκολία εκτίμησης των παραμέτρων ενός ARMA(p, q) υποδείγματος εξαρτάται από την τάξη των p και q. Αν πρόκειται για αυτοπαλίνδρομη σειρά, οι παράμετροι τις μπορούν να εκτιμηθούν με τη μέθοδο των ροπών (method of moments). Όσον αφορά τα υποδείγματα MA ή τις μικτές μορφές, εφαρμόζονται πιο σύνθετες στατιστικές μέθοδοι λόγω της παρουσίας λαθών προηγούμενων περιόδων τα οποία δεν είναι γνωστά. Οι πιο γνωστές μέθοδοι εκτίμησης στην κατηγορία των γραμμικών υποδειγμάτων παλινδρόμησης είναι η μέθοδος ελάχιστων τετραγώνων (Least Squares, LS) και η μέθοδος μέγιστης πιθανοφάνειας (Maximum Likelihood, ML).

Διαδικασία	Συνάρτηση Αυτοσυσχέτισης γ_k	Συνάρτηση Μερικής Αυτοσυσχέτισης α_k
AR(1)	Φθίνει προς το μηδέν. Στη θετική πλευρά εάν $\alpha_1 > 0$. Στην αρνητική πλευρά με εναλλαγή προσήμων εάν $\alpha_1 < 0$.	Μηδενίζεται μετά το α_1 – κορυφή στην πρώτη υστέρηση. Θετική κορυφή για $\alpha_1 > 0$. Αρνητική κορυφή για $\alpha_1 < 0$.
AR(2)	Φθίνει προς το μηδέν.	Μηδενίζεται μετά τη 2 ^η υστέρηση (α_2).

Πίνακας 4.1 Ισολογισμός Walker- Wilson

⁴⁰ Θαλασσινός Ε.(1991, Κ.6 σ.137)

Διαδικασία	Συνάρτηση Αυτοσυσχέτισης r_k	Συνάρτηση Μερικής Αυτοσυσχέτισης a_k
Αυτοπαλίνδρομη Διαδικασία p τάξεως	Φθίνει γεωμετρικά ή φθίνει ακολουθώντας ημιτονοειδή συμπεριφορά.	Μηδενίζεται μετά από p υστερήσεις.
Λευκός θόρυβος	Μηδέν	Μηδέν
MA(1)	Μηδενίζεται μετά την 1 ^η υστέρηση (r_1). Θετική κορυφή όταν $\theta_1 < 0$. Αρνητική κορυφή όταν $\theta_1 > 0$.	Φθίνει γεωμετρικά στην αρνητική πλευρά για $\theta_1 > 0$. Εναλλαγή προσήμων αρχίζοντας από τη θετική πλευρά $\theta_1 < 0$.

Πίνακας 4.2 – Ισολογισμός Walker - Wilson

4.5.3 Διαγνωστικός Έλεγχος

Στο στάδιο του διαγνωστικού ελέγχου⁴¹ προσπαθούμε να ελέγξουμε αν το υπόδειγμα που εκτιμήθηκε παραπάνω είναι το άριστο. Αν δηλαδή δεν υπάρχει άλλο υπόδειγμα διαφορετικής μορφής, βαθμού ή παραμέτρων που να υπερτερεί του αρχικού υποδείγματος που εκτιμήθηκε, σε ότι αφορά τη συμπεριφορά των σφαλμάτων ή την προβλεπτική του ικανότητα. Σε ένα αποτελεσματικό υπόδειγμα τα κατάλοιπα χαρακτηρίζονται καθαρά (λευκός θόρυβος) δηλαδή κατανέμονται ομοιόμορφα με μηδενικό μέσο και διακύμανση σ_ε^2 .

Η πιο αποτελεσματική μέθοδος είναι αυτή στην οποία γίνεται προσπάθεια, με τη μέθοδο της δοκιμής και αποδοχής/ απόρριψης, να εκτιμηθούν πιο πολύπλοκα υποδείγματα σε βάρος του αρχικού, μετά από την εφαρμογή του στα στοιχεία της σειράς. Το πρόβλημα που δημιουργείται σε αυτή την περίπτωση, είναι ο τρόπος με τον οποίο πρέπει να τροποποιηθεί το αρχικό υπόδειγμα. Είναι προτιμότερο στην περίπτωση αυτή να απορρίψουμε το αρχικό υπόδειγμα και να επανέλθουμε στο στάδιο της ταυτοποίησης παρά να προσπαθήσουμε να τροποποιήσουμε το αρχικό υπόδειγμα, διότι η προσθήκη νέων όρων στα σχήματα του υποδείγματος δεν είναι καλή λύση.

Μια άλλη μέθοδος διαγνωστικού ελέγχου είναι αυτή που βασίζεται στην ανάλυση καταλοίπων. Αυτή η μέθοδος εφαρμόζεται είτε στη συνάρτηση αυτοσυσχέτισης των καταλοίπων της νέας σειράς, δηλαδή της σειράς που προκύπτει από το υπόδειγμα που

⁴¹ Θαλασσινός Ε.(1991, Κ.7 σ.149)

εκτιμήθηκε, είτε στο αθροιστικό περιοδιάγραμμα (cumulative period gram) των καταλοίπων της σειράς. Η συνάρτηση αυτοσυσχέτισης συγκρίνεται με τη δειγματική συνάρτηση αυτοσυσχέτισης των καταλοίπων της αρχικής σειράς. Αν οι δυο αυτές συναρτήσεις παρουσιάζουν σημαντικές διαφορές τότε το υπόδειγμα πρέπει να επανεκτιμηθεί. Αν οι δυο συναρτήσεις δεν παρουσιάζουν σημαντικές διαφορές, τότε γίνεται ποσοτική ανάλυση των υπολοίπων της σειράς.

Τα κατάλοιπα πρέπει να συμπεριφέρονται σαν μια διαδικασία λευκού θορύβου. Δεν πρέπει να υπάρχει αυτοσυσχέτιση μεταξύ των καταλοίπων. Ο έλεγχος για τα κατάλοιπα γίνεται με τη στατιστική Q των Box – Pierce. Με τη στατιστική Q ελέγχεται από κοινού η σημαντικότητα ενός αριθμού συντελεστών αυτοσυσχέτισης, έστω m. Τότε η μηδενική υπόθεση θα είναι $H_0 : \rho_1 = \rho_2 = \dots = \rho_m$, όπου $\rho_i, i = 1, 2, \dots, m$ είναι οι συντελεστές αυτοσυσχέτισης των καταλοίπων. Η στατιστική Q των Box – Pierce εξ ορισμού γράφεται ως εξής:

$$Q_{BP} = T \sum_{s=1}^m \hat{\rho}_s^2 \quad (4.21)$$

όπου $\hat{\rho}_s^2$ είναι οι δειγματικές αυτοσυσχετίσεις των καταλοίπων και T ο αριθμός των παρατηρήσεων. Ο αριθμός των αυτοσυσχετίσεων ισούται με την τετραγωνική ρίζα του αριθμού των παρατηρήσεων $m = \sqrt{T}$. Η στατιστική Q των Box – Pierce ακολουθεί την κατανομή X^2 με m-p-q βαθμούς ελευθερίας. Η μηδενική υπόθεση θα απορρίπτεται αν η τιμή Q_{BP} είναι μεγαλύτερη από την κρίσιμη τιμή της κατανομής X_a^2 για δεδομένο επίπεδο α .

Οπότε ισχύει:

$$\text{Απόρριψη} : H_0, \text{ αν } Q_{BP} > X_a^2$$

$$\text{Αποδοχή} : H_0, \text{ αν } Q_{BP} < X_a^2$$

Το κριτήριο Q έχει τεθεί σε διάφορες τροποποιήσεις από την αρχική του διατύπωση και στις πρακτικές του εφαρμογές σήμερα χρησιμοποιείται κυρίως η τροποποίηση των Box – Ljung:

$$Q_{LB} = n(n + 2) \sum_{j=1}^k \frac{\hat{\rho}_j^2(\varepsilon)}{n-j} \quad (4.22)$$

Αυτή η στατιστική, όπως και η Q_{BP} , ακολουθεί προσεγγιστικά την κατανομή χ^2_α με $m-p-q$ βαθμούς ελευθερίας. Όπως και στη στατιστική Q_{BP} έτσι και εδώ η μηδενική υπόθεση θα απορρίπτεται αν $Q_{LB} > \chi^2_\alpha$ για δεδομένο επίπεδο σημαντικότητας α . Η στατιστική Q_{LB} χρησιμοποιείται συνήθως για μικρά δείγματα, ενώ για μεγάλα δείγματα και οι δύο στατιστικές χρησιμοποιούνται το ίδιο κατάλληλα.

4.5.3.1 Διαγνωστικός Έλεγχος – Υπερπροσαρμογή – Κριτήριο Akaike – Μπαϊεσιανό Κριτήριο

Στη διαδικασία της προσαρμογής⁴² γίνεται έλεγχος της καταλληλότητας του εκτιμημένου υποδείγματος συγκρίνοντάς το με ένα υπόδειγμα μεγαλύτερης τάξης. Το υπόδειγμα ARIMA(p, q) που εκτιμήθηκε, συγκρίνεται με τα υποδείγματα ARIMA (p+1, d, q) και ARIMA(p, d, q+1). Αν το υπόδειγμα που εκτιμήθηκε, περιγράφει τη διαδικασία από την οποία παράχθηκαν τα δεδομένα μας και στα μεγαλύτερα υποδείγματα οι επιπλέον συντελεστές είναι μηδέν, τότε το υπόδειγμα αυτό είναι το καταλληλότερο για τα δεδομένα μας. Αν οι συντελεστές δεν είναι μηδέν, τότε υπάρχει άλλο, καταλληλότερο υπόδειγμα για τα δεδομένα μας από το εκτιμημένο. Παρακάτω αναφέρονται δύο κριτήρια που βοηθούν στην επιλογή του κατάλληλου υποδείγματος, τα οποία χρησιμοποιούνται ευρέως στην ανάλυση χρονολογικών σειρών. Αυτά τα κριτήρια είναι, το κριτήριο πληροφοριών Akaike (Akaike Information Criterion, AIC) και το Μπαϊεσιανό κριτήριο Schwartz (Schwartz Bayesian Criterion, SBC) και ορίζονται ως εξής:

$$AIC = \ln(s^2) + \frac{2n}{T} \quad (4.23)$$

$$SBC = \ln(s^2) + n \ln(T) \quad (4.24)$$

Όπου s^2 είναι η εκτίμηση της διακύμανσης των καταλοίπων, n ο αριθμός εκτιμώμενων παραμέτρων υποδείγματος ($p + q + 1$) όπου η μονάδα αντιστοιχεί στην σταθερά, αν υπάρχει, και T είναι ο αριθμός παρατηρήσεων που χρησιμοποιούνται στην παλινδρόμηση.

Αν αυξηθεί η τάση του υποδείγματος προσθέτοντας υστερήσεις είτε για το αυτοπαλίνδρομο μέρος AR, είτε προσθέτοντας υστερήσεις για τους κινητούς μέσους MA, θα μειώνεται το άθροισμα των τετραγώνων των καταλοίπων και ταυτόχρονα μειώνονται οι βαθμοί ελευθερίας λόγω της εκτίμησης περισσότερων παραμέτρων. Με την προσθήκη μιας επιπλέον μεταβλητής στο υπόδειγμα μειώνεται το άθροισμα των

⁴² Σαρριανίδης Ν.(2015 σ.61)

τετραγώνων των καταλοίπων, οπότε μειώνεται και η διακύμανση σ^2 , ταυτόχρονα όμως αυξάνεται το n στους τύπους AIC και SBC. Επομένως αν η προστιθέμενη μεταβλητή δεν έχει ερμηνευτική ικανότητα, τότε οι τιμές των κριτηρίων AIC και SBC θα αυξηθούν. Δηλαδή, η επιλογή των υποδειγμάτων γίνεται με βάση την μικρότερη τιμή των κριτηρίων.

Επειδή $\ln(T) > 2$, το SBC κριτήριο επιβάλλει μεγαλύτερη ποινή στον επιπλέον αριθμό εκτιμώμενων παραμέτρων από το κριτήριο AIC. Για αυτό το λόγο το κριτήριο SBC οδηγεί πάντα στην επιλογή ενός υποδείγματος του οποίου ο αριθμός των παραμέτρων δεν είναι σε καμία περίπτωση μεγαλύτερος από αυτόν που επιλέχθηκε με το κριτήριο AIC.

Η εφαρμογή αυτών των κριτηρίων χρησιμεύει στην επιλογή του καταλληλότερου ARIMA υποδείγματος από πλευράς αριθμού υστερήσεων που θα πρέπει να περιληφθούν. Τα κριτήρια AIC και SBC εφαρμόζονται και σε άλλα υποδείγματα, όπως το υπόδειγμα κατανομημένων χρονικών υστερήσεων για την επιλογή του αριθμού των υστερήσεων των ανεξάρτητων μεταβλητών. Αν η σύγκριση γίνεται μεταξύ υποδειγμάτων με τον ίδιο αριθμό παραμέτρων, τότε τα AIC και SBC οδηγούν στην επιλογή του υποδείγματος με το μεγαλύτερο R^2 .

Τέλος, θα πρέπει να αναφερθεί ως διαγνωστικό στοιχείο ελέγχου, η προβλεπτική ικανότητα του μοντέλου σε τιμές εκτός δείγματος εκτίμησης. Στην πράξη αυτό σημαίνει τον διαχωρισμό του δείγματος σε δυο δείγματα. Με βάση το πρώτο και μεγαλύτερο (in sample) γίνεται η εκτίμηση του δείγματος, ενώ με βάση το δεύτερο (out of sample) γίνεται η αξιολόγηση βάσει στατιστικών στοιχείων.

4.5.4 Προβλέψεις με AR, MA και ARMA Υποδείγματα

Μετά την επιλογή του κατάλληλου υποδείγματος μορφής AR(p), MA(q), ARMA(p,q), το οποίο υπόδειγμα προσαρμόζεται καλύτερα από άλλα στα δεδομένα μιας χρονολογικής σειράς και μετά το στάδιο της εκτίμησης και του ελέγχου, το υπόδειγμα μπορεί να χρησιμοποιηθεί για τη διενέργεια προβλέψεων. Η πρόβλεψη των τιμών μιας χρονολογικής σειράς σε μελλοντικές περιόδους γίνεται βάσει το εκτιμημένο υπόδειγμα που προσαρμόστηκε στα δεδομένα. Έστω Y_1, Y_2, \dots, Y_T , τα δεδομένα μιας χρονολογικής σειράς, στην οποία προσαρμόστηκε ένα υπόδειγμα ARMA(p, q). Η πρόβλεψη την επόμενη περίοδο (T+1) θα είναι η υπό συνθήκη αναμενόμενη τιμή (Conditional Expectation) της σειράς αυτής. Δηλαδή, με βάση τις προηγούμενες τιμές της χρονοσειράς αυτής, η υπό συνθήκη αναμενόμενη τιμή αποτελεί την πιο πιθανή τιμή της

σειράς στο μέλλον σύμφωνα με όλες τις προηγούμενες τιμές της. Μια πρόβλεψη που γίνεται την περίοδο T , για την επόμενη περίοδο $T + 1$ συμβολίζεται με Y_{T+1} και γράφεται ως:

$$Y_{t+1} = E \frac{(Y_{T+1})}{(Y_T, \dots, Y_1)} = E_T(Y_{T+h}) \quad (4.25)$$

Εδώ ο υποδείκτης T στην αναμενόμενη τιμή υποδεικνύει ότι στηρίζεται στις πληροφορίες μέχρι και την περίοδο T .

Η πρόβλεψη για h περιόδους συμβολίζεται ως:

$$Y_{T+h} = E_T(Y_{T+h}) \quad (4.26)$$

Το λάθος πρόβλεψης εκφράζεται ως e_{T+h} . Πρόκειται για τη διαφορά μεταξύ της τιμής του Y_T την περίοδο $T + h$ και της αναμενόμενης τιμής που λάβαμε. Δηλαδή:

$$e_{T+h} = Y_{T+h} - Y_{T+h} \quad \text{όπου } h = 1, 2, 3, \dots \text{ μελλοντικές περιόδοι} \quad (4.27)$$

4.5.4.1 Πρόβλεψη Διαδικασίας MA(q)

Η "μνήμη" μιας διαδικασίας MA(q) έχει μήκος q και αυτό περιορίζει τις προοπτικές πρόβλεψης. Έστω το υπόδειγμα MA(1)⁴³:

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} \quad (4.28)$$

Υποθέτουμε ότι οι παράμετροι του υποδείγματος είναι σταθεροί και θα ισχύουν και για τις επόμενες χρονικές περιόδους $t+1, t+2, \dots$. Για την περίοδο $t+1$, η πρόβλεψη μπορεί να επιτευχθεί λαμβάνοντας την υπό συνθήκη αναμενόμενη τιμή του υποδείγματος με παράλληλη μετατόπιση των δεικτών κατά ένα.

$$\hat{y}_{t+1} = E_T(\mu + \theta_1 \varepsilon_t + \varepsilon_t/t) = \mu + \theta_1 \varepsilon_t \quad (4.29)$$

⁴³ Σαρριανίδης Ν.(2015 σ.69)

Η υπό συνθήκη αναμενόμενη τιμή έχει την έννοια πληροφοριών που είναι διαθέσιμες μέχρι τη χρονική στιγμή t . Η αναμενόμενη τιμή των σφαλμάτων μετά τη χρονική στιγμή t είναι μηδέν, αφού σύμφωνα με την υπόθεση των μοντέλων κινητού μέσου MA, το e_t είναι λευκός θόρυβος και συνεπώς:

$$E(\varepsilon_{t+1}/t) = 0 \quad (4.30)$$

Οι τιμές των σφαλμάτων μέχρι την περίοδο t παίρνουν συγκεκριμένες τιμές καθώς υπολογίζονται από την εκτίμηση του υποδείγματος.

Γενικά για προβλέψεις για $h > 1$ περιόδους ισχύει ότι:

$$\hat{Y}_{T+h} = \mu \quad (4.31)$$

$$V(\hat{\varepsilon}_{T+h}) = \sigma^2(1 + \theta_1^2) \quad (4.32)$$

Από τη σχέση (4.29) συμπεραίνεται ότι ένα υπόδειγμα κινητού μέσου MA(1) είναι κατάλληλο για προβλέψεις μόνο μια περίοδο μπροστά, αφού για $h > 1$ η πρόβλεψη είναι πάντα ο μέσος.

4.5.4.2 Πρόβλεψη Διαδικασίας AR(p)

Αντίθετα με τη διαδικασία κινητών μέσων MA, η διαδικασία AR έχει απεριόριστη μνήμη⁴⁴. Έστω το υπόδειγμα:

$$Y_t = \delta + aY_{t-1} + \varepsilon_t \quad (4.33)$$

Για $t = T + 1$ το υπόδειγμα (4.33) γίνεται :

$$Y_{t+1} = \delta + a_1Y_{t-1} + \varepsilon_t \quad (4.34)$$

⁴⁴ Σαρριανίδης Ν.(2015 σ.71)

Αν τα δ και α_1 είναι γνωστά, τότε σύμφωνα με τις πληροφορίες που έχουμε μέχρι την περίοδο T , μια πρόβλεψη για την περίοδο $T + 1$ θα είναι η υπό συνθήκη αναμενόμενη τιμή της Y_{t+1} . Θέτοντας ως \hat{Y}_{t+1} την πρόβλεψη, τότε:

$$\hat{Y}_{t+1} = E_T Y_{T+1} \quad (4.35)$$

Και γενικεύοντας προκύπτει:

$$\hat{Y}_{t+h} = E_T Y_{T+h} \quad (4.36)$$

που είναι μια πρόβλεψη για h περιόδους μπροστά βάσει των πληροφοριών που έχουμε μέχρι την περίοδο T . Η πρόβλεψη αυτή θεωρείται άριστη (optimal forecast) διότι ελαχιστοποιεί το μέσο του τετραγώνου του σφάλματος (Mean Square Error), δηλαδή ελαχιστοποιεί την ποσότητα $E_T (Y_{T+h} - \hat{Y}_{T+h})^2$. Για προβλέψεις για h περιόδους μπροστά ισχύει:

$$\hat{Y}_{T+h} = \delta + \alpha_1 Y_{T+h-1} \quad (4.37)$$

$$Var(\hat{\epsilon}_{T+h}) = \sigma_\epsilon^2 (1 + \alpha_1^2 + \alpha_1^4 + \dots + \alpha_1^{2(h-1)}) \quad (4.38)$$

Από τη σχέση (4.38) διαπιστώνεται πως καθώς αυξάνει η περίοδος πρόβλεψης, η διακύμανση του σφάλματος πρόβλεψης αυξάνει μη γραμμικά.

4.5.4.3 Πρόβλεψη Διαδικασίας ARMA(p, q)

Η πρόβλεψη για μια διαδικασία ARMA(p, q) μπορεί να παραχθεί με τη συνδυασμένη εφαρμογή των παραπάνω διαδικασιών που ισχύουν για τις προβλέψεις AR και MA αντίστοιχα⁴⁵.

Η πρόβλεψη μιας διαδικασίας ARIMA πραγματοποιείται με την πρόβλεψη της αντίστοιχης διαδικασίας ARMA, και έπειτα την αντικατάσταση των διαφορών $\Delta y_{t+1} = y_{t+1} - y_t$ για την εύρεση της προβλεπόμενης τιμής y_{t+1} .

⁴⁵ Δημελή Σ. (2013, Κ.6 σ.200)

- Έστω το υπόδειγμα ARMA(1, 1):

$$Y_t = \delta + \alpha_1 Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} \quad (4.39)$$

Για $t = T + 1$ το υπόδειγμα παίρνει τη μορφή:

$$\hat{Y}_{T+1} = \delta + \alpha_1 Y_T + \varepsilon_{T+1} + \theta_1 \varepsilon_T \quad (4.40)$$

Η άριστη πρόβλεψη για την επόμενη χρονική περίοδο είναι:

$$\hat{Y}_{T+1} = \delta + \alpha_1 Y_T + \theta_1 \varepsilon_T \quad (4.41)$$

Από τις παρακάτω σχέσεις προκύπτει το σφάλμα πρόβλεψης και η διακύμανση:

$$\hat{\varepsilon}_{T+1} = Y_{T+1} - \hat{Y}_{T+1} = \varepsilon_{T+1} \quad (4.42)$$

$$V(\hat{\varepsilon}_{T+1}) = V(\varepsilon_{T+1}) = \sigma^2 \quad (4.43)$$

- Έστω το υπόδειγμα ARIMA (1, 1, 1):

$$w_t = \delta + \alpha_1 w_{t-1} + \varepsilon_t, \text{ όπου } w_t = \Delta Y_t = Y_t - Y_{t-1} \quad (4.44)$$

Για να γίνει στάσιμη η μη στάσιμη σειρά Y_t παίρνουμε τις πρώτες διαφορές που είναι AR(1). Η άριστη πρόβλεψη για τη w_{T+1} είναι:

$$\hat{w}_{T+1} = \delta + \alpha_1 w_T \quad (4.45)$$

Στη συνέχεια γίνεται η πρόβλεψη της αρχικής σειράς Y , δηλαδή η πρόβλεψη για την Y_{T+1} είναι:

$$\begin{aligned} \hat{Y}_{T+1} &= Y_T + \hat{w}_{T+1} = Y_T + \delta + \alpha_1 w_T = \\ &= Y_T + \delta + \alpha_1 (Y_T - Y_{T-1}) = \delta + (\mathbf{1} + \alpha_1) Y_T - \alpha_1 Y_{T-1} \end{aligned} \quad (4.46)$$

Η πρόβλεψη για h περιόδους μπροστά θα είναι:

$$\hat{Y}_{T+h} = Y_T + \hat{w}_{T+1} + \hat{w}_{T+2} + \dots + \hat{w}_{T+h} \quad (4.47)$$

Στη μεθοδολογία των προβλέψεων με τη χρήση υποδειγμάτων ARMA(p, q), το τελευταίο στάδιο έχει να κάνει με την αξιοπιστία προβλέψεων. Κάνοντας μια πρόβλεψη, δεν μας ενδιαφέρει μόνο ένα σημείο πρόβλεψης \hat{Y}_{T+h} , αλλά και ένα μέτρο αξιοπιστίας της πρόβλεψης, δηλαδή το διάστημα εμπιστοσύνης της πρόβλεψης. Ένα διάστημα πρόβλεψης (1- α)% θα έχει τη μορφή $\hat{Y}_{T+h} \pm z_c$ (διακύμανση λάθους πρόβλεψης)^{1/2} όπου z_c είναι η $\alpha/2$ κριτική τιμή για την τυπική κανονική κατανομή $N(0, 1)$.

4.6 Αξιολόγηση προβλέψεων

Για να αξιολογηθεί η προβλεπτική ικανότητα ενός υποδείγματος θα πρέπει να γίνει σύγκριση των προβλέψεων με τα πραγματικά δεδομένα της χρονολογικής σειράς. Σε αυτή την περίπτωση ο λόγος γίνεται για τις εκ των υστέρων προβλέψεις, οι οποίες ονομάζονται "ex-post". Οι προβλέψεις "ex-post" είναι προβλέψεις που γίνονται μέσα στην περίοδο όπου έχουν ήδη πραγματοποιηθεί οι τιμές που θέλουμε να προβλέψουμε. Αντίθετα, όταν γίνεται σήμερα πρόβλεψη των μελλοντικών τιμών της μεταβλητής, τότε γίνεται λόγος για τις εκ των προτέρων προβλέψεις (ex-ante), δηλαδή προβλέψεως τιμών για περιόδους που δεν υπάρχουν ακόμα αληθινές τιμές. Οι προβλέψεις που γίνονται με ένα συγκεκριμένο υπόδειγμα μπορούν να αξιολογηθούν μόνο εκ των υστέρων(ex-post). Από ένα συνολικό δείγμα παρατηρήσεων, τις τελευταίες παρατηρήσεις δεν τις συμπεριλαμβάνουμε στην εκτίμηση ώστε να τις συγκρίνουμε με τις προβλέψεις που κάνουμε στο διάστημα αυτό. Για αυτό το σκοπό, είναι απαραίτητο να υπάρχει μεγάλος αριθμός παρατηρήσεων ώστε να υπάρχουν οι απαραίτητοι βαθμοί ελευθερίας για την εκτίμηση και τον έλεγχο του επιλεγμένου υποδείγματος. Υπάρχουν στατιστικά μέτρα που συντελούν στην αξιολόγηση των προβλέψεων. Τα μέτρα αυτά στηρίζονται στο μέγεθος του λάθους πρόβλεψης που κάνουμε. Για τη μέτρηση της ακρίβειας των "ex-post" προβλέψεων χρησιμοποιούνται κάποιοι τύποι, οι οποίοι παραθέτονται παρακάτω⁴⁶.

Μέσο Σφάλμα Τετραγώνου (Mean Squared Error, MSE)

$$MSE = \frac{1}{N} \sum_{t=1}^N e_t^2 \quad (4.48)$$

⁴⁶ Δημελή Σ. (2013, Κ.6 σ.215)

Τετραγωνική ρίζα Μέσου Σφάλματος Τετραγώνου (Root Mean Squared Error)

$$RMSE = \sqrt{\frac{\sum e_t^2}{N}} \quad (4.49)$$

Μέσο Απόλυτο Σφάλμα (Mean Absolute Error)

$$MAE = \frac{1}{N} \sum_{t=1}^N |e_t| \quad (4.50)$$

Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (Mean Absolute Percentage Error)

$$MAPE = \frac{100}{N} \sum \left| \frac{F_t - A_t}{A_t} \right| \quad (4.51)$$

Όπου A_t οι πραγματικές τιμές της χρονολογικής σειράς, F_t οι προβλεφθείσες τιμές και $e_t = F_t - A_t$ το λάθος πρόβλεψης.

Το μέσο Σφάλμα Τετραγώνου (MSE) και η Τετραγωνική Ρίζα του Μέσου Σφάλματος τετραγώνου (RMSE) δίνουν περισσότερη έμφαση στα μεγάλα λάθη διότι τετραγωνίζονται, σε αντίθεση με τα MAE και MAPE τα οποία υπολογίζουν μόνο τα απόλυτα σφάλματα. Όσο μικρότερες είναι οι τιμές των μεγεθών αυτών, τόσο καλύτερη είναι η προβλεπτική ικανότητα του υποδείγματος. Ένα μειονέκτημα των παραπάνω μεγεθών είναι πως όλα τους επηρεάζονται από τις μονάδες μέτρησης των μεταβλητών, επομένως χρειάζεται προσοχή κατά τη σύγκριση μεταξύ εναλλακτικών μετρήσεων ώστε η προβλεπόμενη μεταβλητή να εκφράζεται στις ίδιες μονάδες.

Ο Συντελεστής Ανισότητας του Theil

Υπάρχει και ένα άλλο μέγεθος αξιολόγησης προβλέψεων, το οποίο είναι ανεξάρτητο των μονάδων μέτρησης και ονομάζεται συντελεστής ανισότητας του Theil. Ο συντελεστής αυτός συμβολίζεται με το γράμμα U και ορίζεται ως εξής:

$$U = \frac{\sqrt{\frac{1}{N} \sum (F_t - A_t)^2}}{\sqrt{\frac{1}{N} \sum (A_t)^2}} = \frac{RMSE}{\sqrt{\frac{1}{N} \sum (A_t)^2}} \quad (4.52)$$

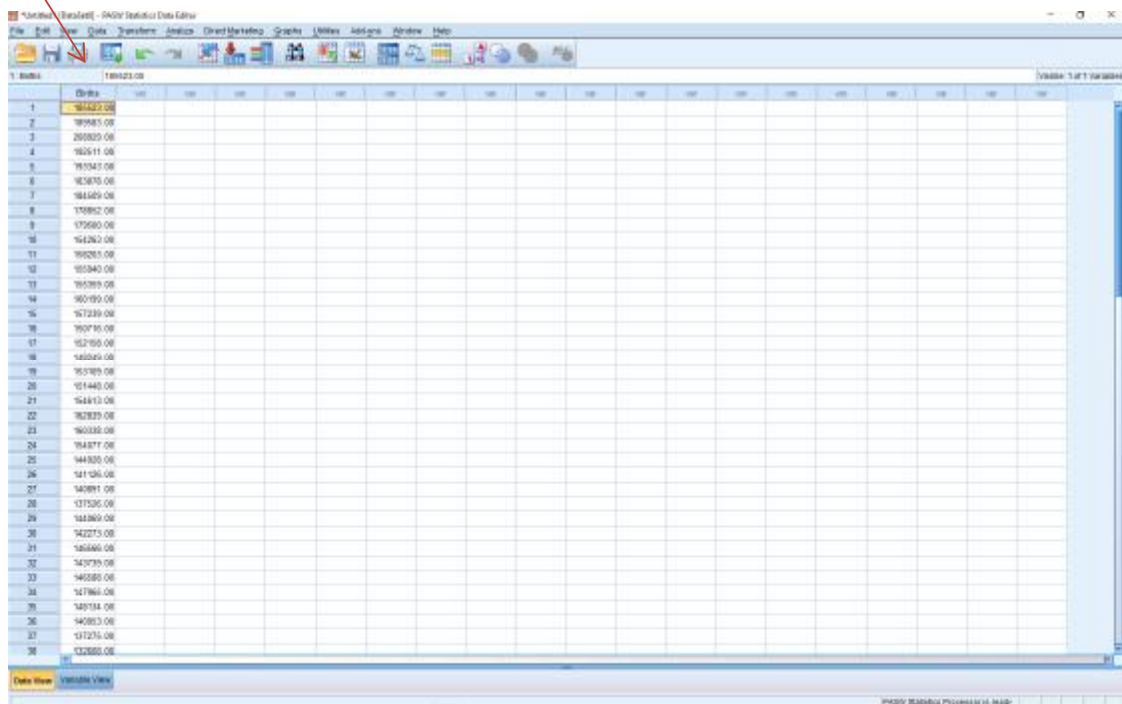
Σύμφωνα με τον Theil προτείνεται η χρήση ποσοστιαίων μεταβολών στη θέση των πραγματικών τιμών της χρονολογικής σειράς (A) και των προβλεφθεισών τιμών (F), αλλά ο τύπος χρησιμοποιείται και στις αρχικές τιμές. Όταν οι προβλέψεις είναι απόλυτα ακριβείς ο συντελεστής U παίρνει την τιμή μηδέν και όταν όλες οι προβλέψεις είναι μηδέν τότε ο συντελεστής U παίρνει την τιμή ένα. Αντιθέτως, αν ο συντελεστής U υπερβαίνει τη μονάδα τότε οι προβλέψεις είναι κακές. Δηλαδή, όσο πιο κοντά στο μηδέν είναι ο συντελεστής U, τόσο καλύτερες είναι οι προβλέψεις. Η σχέση του συντελεστή U αναλύεται σε τρεις συνιστώσες, το ποσοστό μεροληψίας (Bias Proportion, UM), το ποσοστό διακύμανσης (Variance Proportion, UV), το ποσοστό συνδιακύμανσης (Covariance Proportion) $UM + UV + UC = 1$. Χαρακτηριστικό των καλών προβλέψεων είναι πως θα έχουν μικρό ποσοστό μεροληψίας και διακύμανσης και μεγάλο ποσοστό συνδιακύμανσης.

ΚΕΦΑΛΑΙΟ 5: Η εφαρμογή της μεθοδολογίας των Box – Jenkins για τη μελέτη των γεννήσεων στην Ελλάδα.

Σε αυτό το κεφάλαιο αναλύεται η διεξαγωγή προβλέψεων με βάση το στατιστικό πρόγραμμα PASW Statistics. Για αυτή την ανάλυση έχουν επιλεγεί οι γεννήσεις στη χώρα μας.

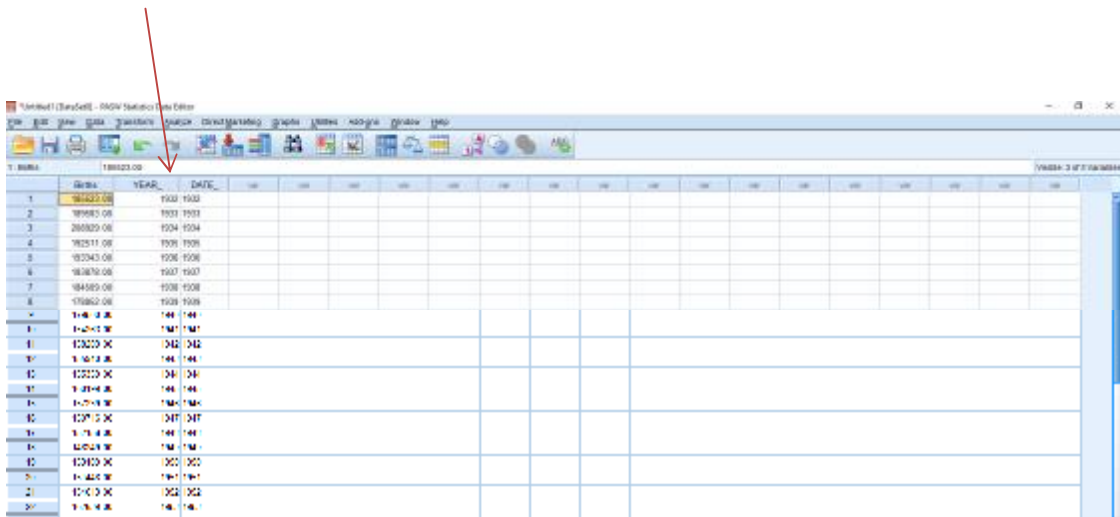
Στο παράδειγμα των γεννήσεων έχουν περαστεί τα δεδομένα στο στατιστικό πρόγραμμα για το διάστημα από 01/01/1932 έως 31/12/2001, δηλαδή 70 τιμές.

Αφού γίνει η εισαγωγή των δεδομένων στο στατιστικό πρόγραμμα το πρώτο βήμα είναι μέσω της επιλογής στο menu DATA-DEFINE DATES να ορίσουμε ότι η περιοδικότητα των δεδομένων είναι ανά έτος.



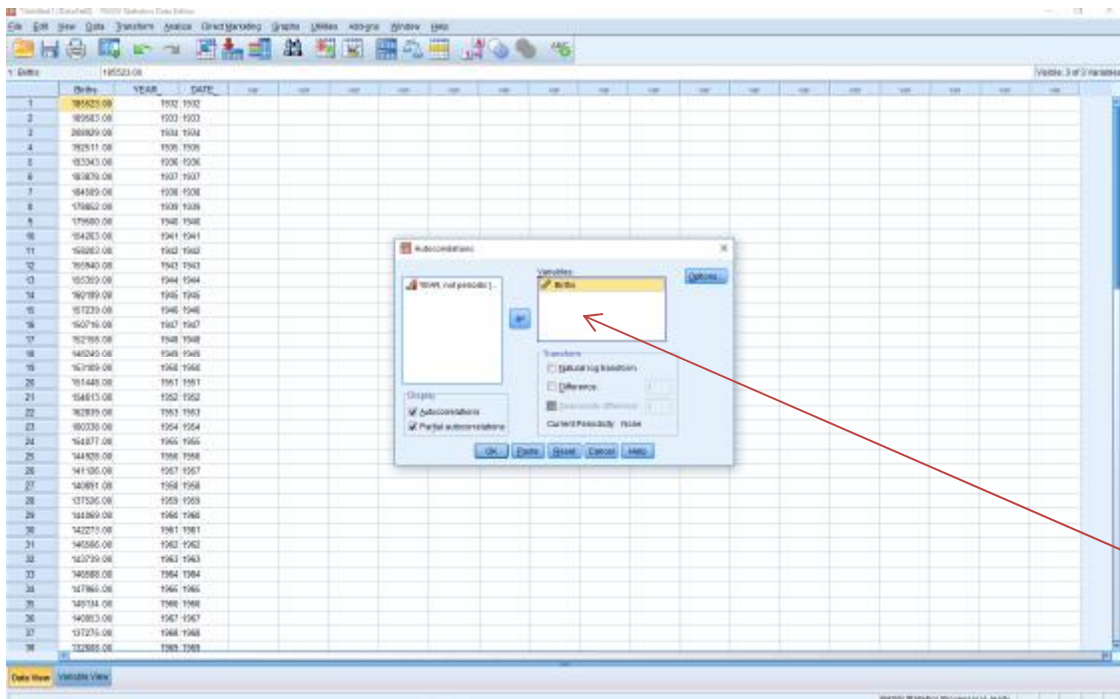
Όχι	Όχι
1	18142.00
2	18585.00
3	20020.00
4	18211.00
5	18542.00
6	18380.00
7	18120.00
8	17892.00
9	17580.00
10	16262.00
11	16025.00
12	15340.00
13	15385.00
14	16090.00
15	16229.00
16	16078.00
17	15250.00
18	14820.00
19	15100.00
20	15440.00
21	15410.00
22	16020.00
23	16030.00
24	16477.00
25	16430.00
26	16126.00
27	16091.00
28	13720.00
29	14860.00
30	14275.00
31	14560.00
32	14370.00
33	14020.00
34	12760.00
35	14914.00
36	14063.00
37	13276.00
38	13265.00

Εικόνα 1 (Εισαγωγή δεδομένων)



Εικόνα 2 (Μεταβλητές YEAR_, DATE_)

Από το menu του στατιστικό πρόγραμμα, και έπειτα από την επιλογή Analyze – Forecasting – Autocorrelations τοποθετείται η μεταβλητή Births στο Variables ώστε να γίνουν οι δοκιμές και να βρεθεί το καταλληλότερο υπόδειγμα κάνοντας τους ελέγχους στασιμότητας. Αρχικά δοκιμάζεται χωρίς διαφορές και επιλέγοντας το ok προκύπτουν τα εξής αποτελέσματα στο output.



Εικόνα 3 (Διαδικασία για έλεγχο στασιμότητας)

Births

Autocorrelations

Series: Births

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	.946	.117	65.341	1	.000
2	.890	.116	124.031	2	.000
3	.808	.115	173.168	3	.000
4	.747	.114	215.729	4	.000
5	.685	.114	252.100	5	.000
6	.640	.113	284.354	6	.000
7	.590	.112	312.162	7	.000
8	.555	.111	337.203	8	.000
9	.512	.110	358.890	9	.000
10	.493	.109	379.299	10	.000
11	.466	.108	397.884	11	.000
12	.441	.107	414.794	12	.000
13	.413	.106	429.881	13	.000
14	.384	.105	443.157	14	.000
15	.351	.104	454.475	15	.000
16	.324	.104	464.244	16	.000

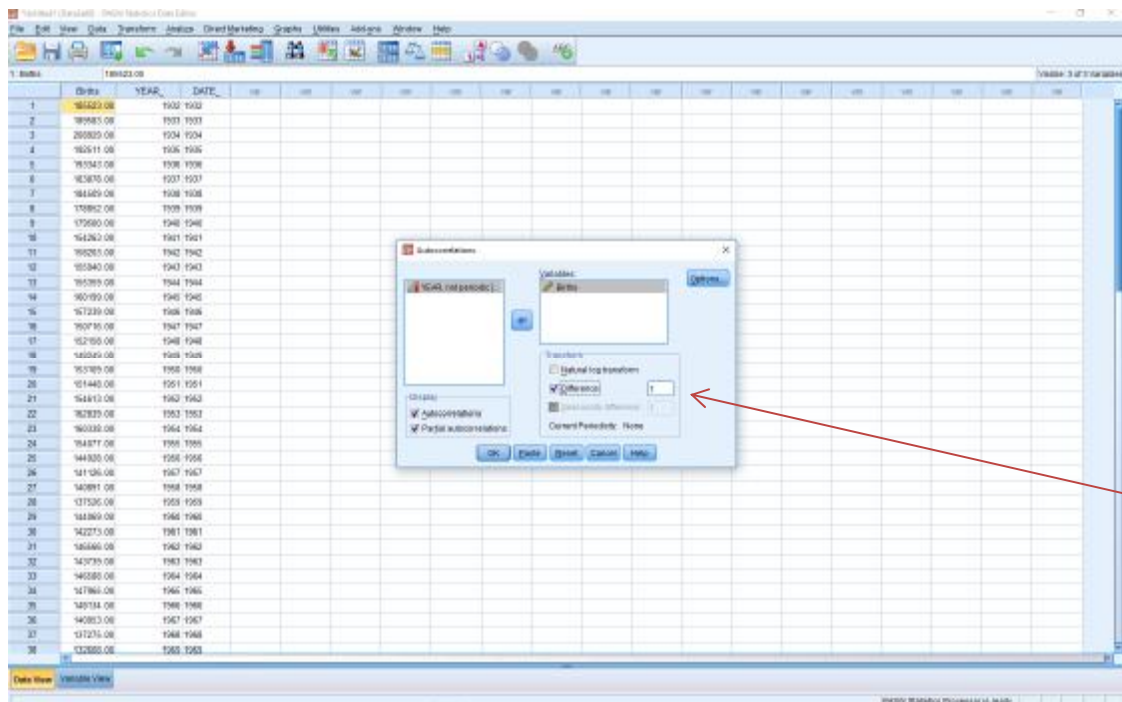
a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Πίνακας 1 (Autocorrelations)

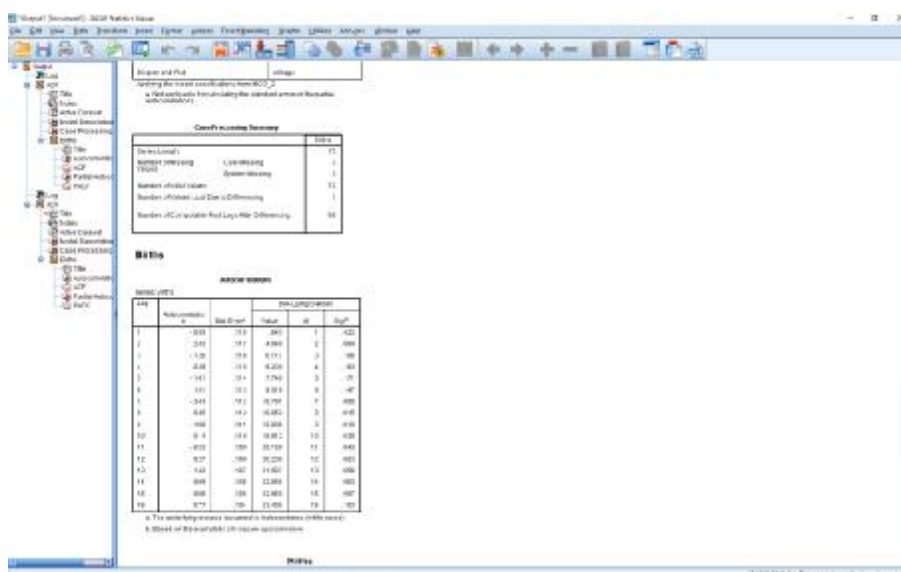
- Χρονολογικές υστερήσεις (Lag)
- Τιμή του συντελεστή αυτοσυσχέτισης (Autocorrelation)
- Τυπικό σφάλμα (Std Error)
- Τιμή της στατιστικής συνάρτησης Box-Ljung (Value)
- Βαθμοί ελευθερίας (df)
- Τιμή πιθανότητας της στατιστικής συνάρτησης (Sig)

Από τον πίνακα 1 είναι φανερό πως για όλα τα Lags υπάρχει αυτοσυσχέτιση γιατί όλα τα ρ είναι μηδέν, οπότε δεν είναι στάσιμη χωρίς διαφορές και θα δοκιμαστεί με πρώτες διαφορές.



Εικόνα 4 (Διαδικασία για έλεγχο στασιμότητας – πρώτες διαφορές)

Πατώντας ok εμφανίζονται τα εξής αποτελέσματα στον πίνακα Autocorrelations στο Output.



Εικόνα 5 (Αποτελέσματα στον πίνακα Autocorrelations στο Output)

Births

Autocorrelations

Series:Births

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-.095	.118	.645	1	.422
2	.243	.117	4.946	2	.084
3	-.126	.116	6.117	3	.106
4	.039	.115	6.230	4	.183
5	-.141	.114	7.749	5	.171
6	.151	.113	9.516	6	.147
7	-.343	.112	18.791	7	.009
8	.045	.112	18.952	8	.015
9	-.108	.111	19.896	9	.019
10	-.014	.110	19.912	10	.030
11	-.055	.109	20.168	11	.043
12	.027	.108	20.230	12	.063
13	-.142	.107	21.987	13	.056
14	.099	.106	22.855	14	.063
15	-.006	.105	22.858	15	.087
16	.077	.104	23.405	16	.103

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

b. Based on the asymptotic chi-square approximation.

Πίνακας 2 (Autocorrelations)

Έλεγχος στασιμότητας

$H_0: \rho = 0$

vs

$H_1: \rho \neq 0$

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

1) $H_0: \rho_1 = 0 \quad vs \quad H_1: \rho_1 \neq 0$

$\rho_1 = 0.422 > \alpha = 0.118$
στάσιμη

Αποδοχή H_0 - Δεν υπάρχει αυτοσυσχέτιση, άρα

2) $H_0 : \rho_2 = 0$ vs $H_1 : \rho_2 \neq 0$

$\rho_2 = 0.084 < \alpha = 0.117$ Απόρριψη H_0 - Υπάρχει αυτοσυσχέτιση

3) $H_0 : \rho_3 = 0$ vs $H_1 : \rho_3 \neq 0$

$\rho_3 = 0.106 < \alpha = 0.116$ Απόρριψη H_0 - Υπάρχει αυτοσυσχέτιση

4) $H_0 : \rho_4 = 0$ vs $H_1 : \rho_4 \neq 0$

$\rho_4 = 0.183 > \alpha = 0.115$ Αποδοχή H_0 - Δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη

5) $H_0 : \rho_5 = 0$ vs $H_1 : \rho_5 \neq 0$

$\rho_5 = 0.171 > \alpha = 0.114$ Αποδοχή H_0 - Δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη

6) $H_0 : \rho_6 = 0$ vs $H_1 : \rho_6 \neq 0$

$\rho_6 = 0.147 > \alpha = 0.113$ Αποδοχή H_0 - Δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη

7) $H_0 : \rho_7 = 0$ vs $H_1 : \rho_7 \neq 0$

$\rho_7 = 0.009 < \alpha = 0.112$ Απόρριψη H_0 - Υπάρχει αυτοσυσχέτιση

8) $H_0 : \rho_8 = 0$ vs $H_1 : \rho_8 \neq 0$

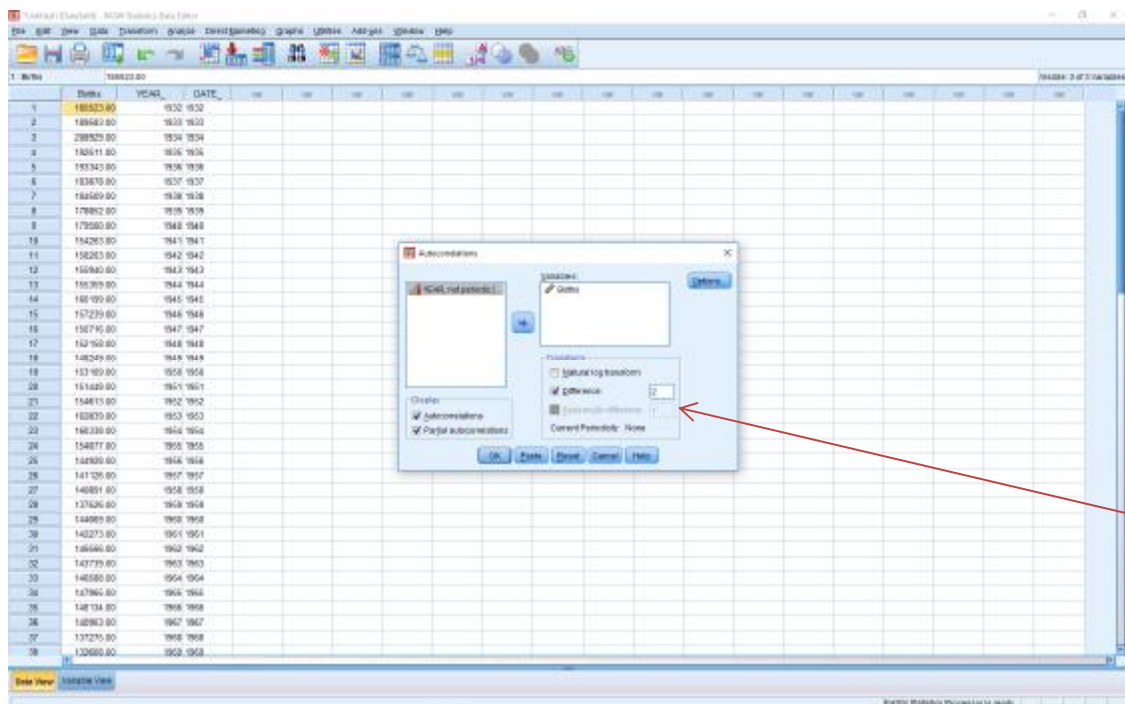
$\rho_8 = 0.015 < \alpha = 0.111$ Απόρριψη H_0 - Υπάρχει αυτοσυσχέτιση

....

$$16) H_0 : \rho_{16} = 0 \text{ vs } H_1 : \rho_{16} \neq 0$$

$$\rho_{16} = 0.103 < \alpha = 0.104 \quad \text{Απόρριψη } H_0 - \text{Υπάρχει αυτοσυσχέτιση}$$

Από τον έλεγχο υποθέσεων με πρώτες διαφορές υπάρχει αυτοσυσχέτιση στα περισσότερα από τα πρώτα 8 υποδείγματα. Επειδή τα $\rho_2, \rho_3, \rho_7, \rho_8$ εμφανίζουν αυτοσυσχέτιση, η διαδικασία θα δοκιμαστεί με δεύτερες διαφορές για να ελεγχθεί αν υπάρχουν καλύτερα αποτελέσματα.



Εικόνα 6 (Διαδικασία για έλεγχο στασιμότητας - δεύτερες διαφορές)

Births

Autocorrelations

Series:Births

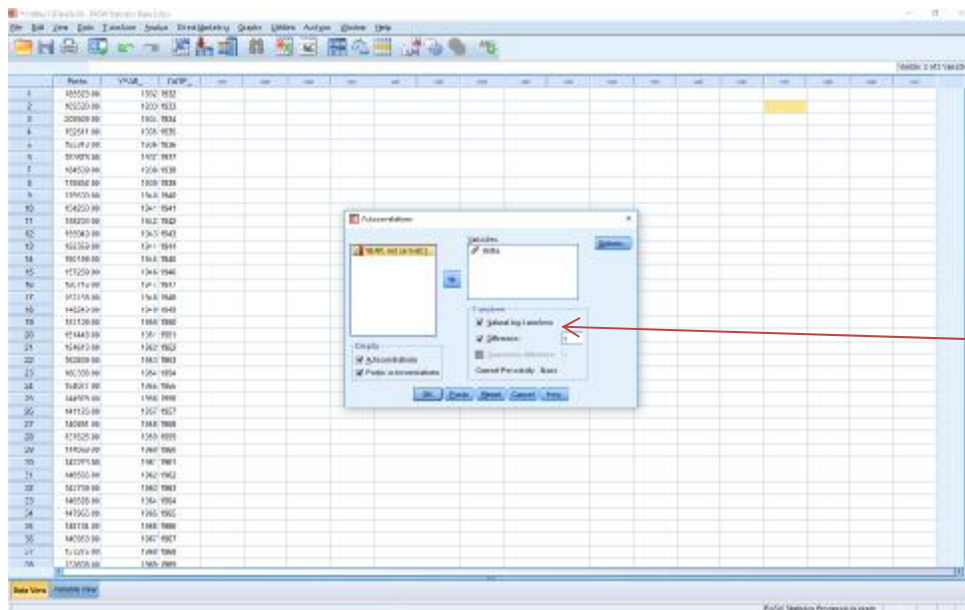
Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-.674	.119	32.309	1	.000
2	.364	.118	41.846	2	.000
3	-.263	.117	46.925	3	.000
4	.170	.116	49.063	4	.000
5	-.229	.115	53.009	5	.000
6	.367	.114	63.350	6	.000
7	-.413	.113	76.653	7	.000
8	.278	.112	82.764	8	.000
9	-.145	.111	84.467	9	.000
10	.069	.110	84.858	10	.000
11	-.058	.109	85.140	11	.000
12	.109	.108	86.156	12	.000
13	-.180	.107	88.949	13	.000
14	.163	.107	91.293	14	.000
15	-.096	.106	92.116	15	.000
16	.110	.105	93.221	16	.000

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

b. Based on the asymptotic chi-square approximation.

Πίνακας 3 (Autocorrelations)

Στον πίνακα 3 είναι φανερό ότι για όλα τα lags υπάρχει αυτοσυσχέτιση γιατί όλα τα ρ είναι μηδέν, οπότε η διαδικασία δεν είναι στάσιμη με δεύτερες διαφορές και δοκιμάζεται με πρώτες διαφορές και λογάριθμο. Πατώντας ok εμφανίζονται τα παρακάτω αποτελέσματα στον πίνακα Autocorrelations στο output.



Εικόνα 7 (Διαδικασία για έλεγχο στασιμότητας με πρώτες διαφορές και λογάριθμο)

Births

Autocorrelations

Series:Births

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	.033	.118	.077	1	.782
2	.300	.117	6.655	2	.036
3	-.069	.116	7.005	3	.072
4	.022	.115	7.042	4	.134
5	-.161	.114	9.026	5	.108
6	.089	.113	9.648	6	.140
7	-.329	.112	18.194	7	.011
8	.029	.112	18.260	8	.019
9	-.133	.111	19.701	9	.020
10	-.032	.110	19.783	10	.031
11	-.106	.109	20.739	11	.036
12	.034	.108	20.839	12	.053
13	-.172	.107	23.421	13	.037
14	.104	.106	24.387	14	.041
15	-.021	.105	24.426	15	.058
16	.077	.104	24.977	16	.070

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

b. Based on the asymptotic chi-square approximation.

Πίνακας 4 (Autocorrelations)

Έλεγχος Στασιμότητας

$$H_0 : \rho = 0$$

vs

$$H_1 : \rho \neq 0$$

Με την υπόθεση H_0 δεν υπάρχει αυτοσυσχέτιση, άρα η διαδικασία είναι στάσιμη ενώ με την υπόθεση H_1 υπάρχει αυτοσυσχέτιση και δεν προκύπτει στασιμότητα.

$$1) H_0 : \rho_1 = 0 \text{ vs } H_1 : \rho_1 \neq 0$$

$$\rho_1 = 0.782 > \alpha = 0.118 \quad \text{Αποδοχή } H_0 - \text{ Δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη}$$

$$2) H_0 : \rho_2 = 0 \text{ vs } H_1 : \rho_2 \neq 0$$

$$\rho_2 = 0.036 < \alpha = 0.117 \quad \text{Απόρριψη } H_0 - \text{ Υπάρχει αυτοσυσχέτιση}$$

3) $H_0 : \rho_3 = 0 \text{ vs } H_1 : \rho_3 \neq 0$

$\rho_3 = 0.072 < \alpha = 0.116$ Απόρριψη H_0 – Υπάρχει αυτοσυσχέτιση

4) $H_0 : \rho_4 = 0 \text{ vs } H_1 : \rho_4 \neq 0$

$\rho_4 = 0.134 > \alpha = 0.115$ Αποδοχή H_0 – Δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη

5) $H_0 : \rho_5 = 0 \text{ vs } H_1 : \rho_5 \neq 0$

$\rho_5 = 0.108 < \alpha = 0.114$ Απόρριψη H_0 – Υπάρχει αυτοσυσχέτιση

6) $H_0 : \rho_6 = 0 \text{ vs } H_1 : \rho_6 \neq 0$

$\rho_6 = 0.140 > \alpha = 0.113$ Αποδοχή H_0 – Δεν υπάρχει αυτοσυσχέτιση, άρα στάσιμη

7) $H_0 : \rho_7 = 0 \text{ vs } H_1 : \rho_7 \neq 0$

$\rho_7 = 0.011 < \alpha = 0.112$ Απόρριψη H_0 – Υπάρχει αυτοσυσχέτιση

8) $H_0 : \rho_8 = 0 \text{ vs } H_1 : \rho_8 \neq 0$

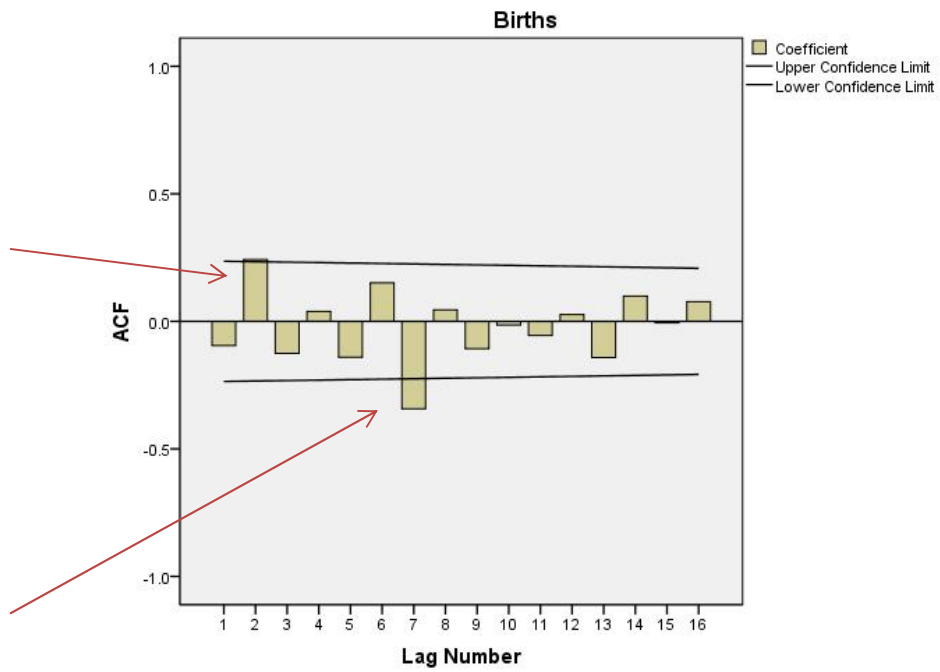
$\rho_8 = 0.019 < \alpha = 0.112$ Απόρριψη H_0 – Υπάρχει αυτοσυσχέτιση

....

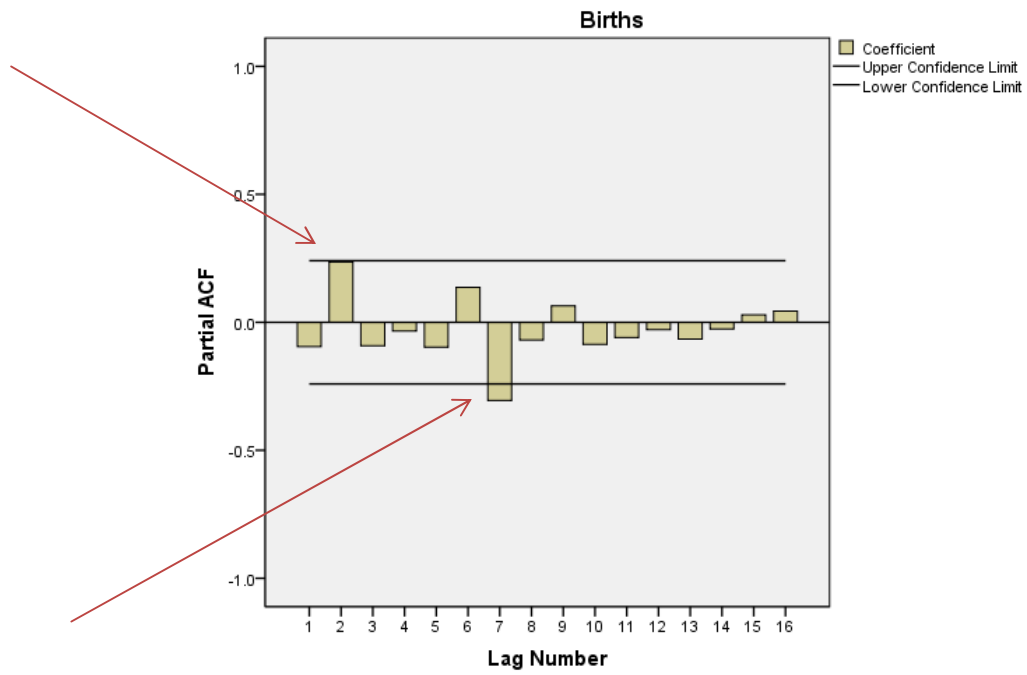
16) $H_0 : \rho_{16} = 0 \text{ vs } H_1 : \rho_{16} \neq 0$

$\rho_{16} = 0.070 < \alpha = 0.104$ Απόρριψη H_0 - Υπάρχει αυτοσυσχέτιση

Σύμφωνα με τον έλεγχο στασιμότητας η διαδικασία αυτή είναι στάσιμη με πρώτες διαφορές, καθώς με δεύτερες διαφορές και με πρώτες διαφορές και λογάριθμο δεν υπήρξαν καλύτερα αποτελέσματα. Ύστερα με τα αποτελέσματα που εμφανίζονται στο output των ACF και Partial ACF εντοπίζονται τα q και p αντίστοιχα.



Εικόνα 8 (Διάγραμμα αυτοσυσχετίσεων- Εύρεση τάξεως q)

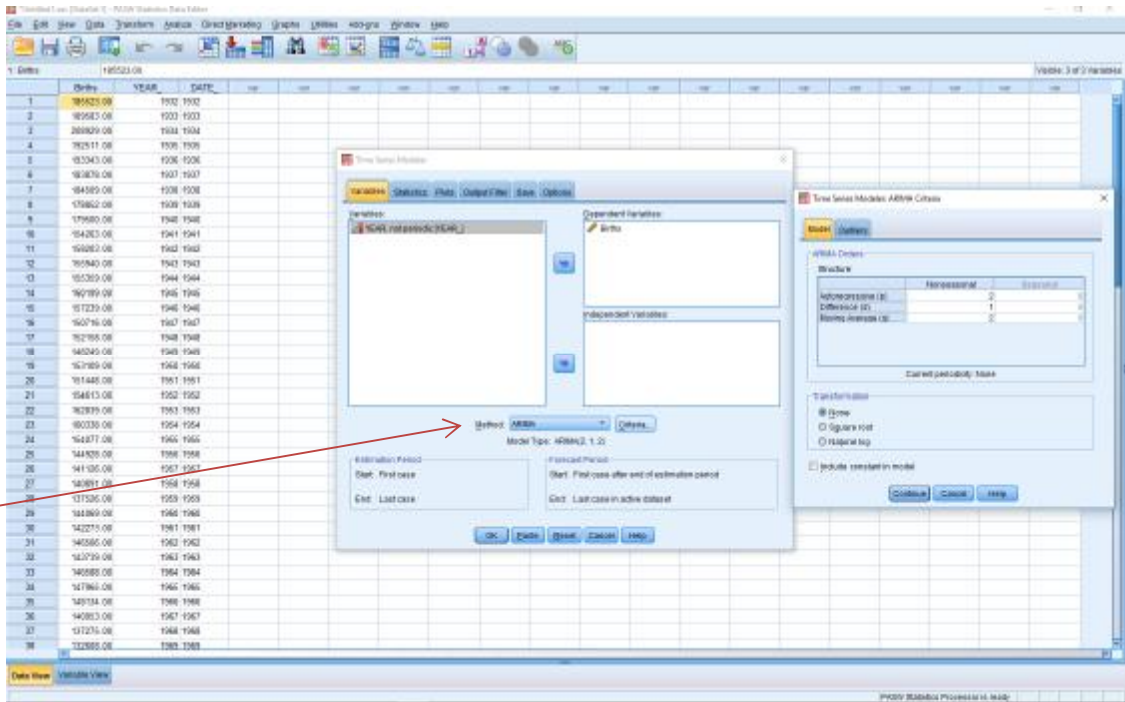


Εικόνα 9 (Διάγραμμα μερικών αυτοσυσχετίσεων- Εύρεση τάξεως p)

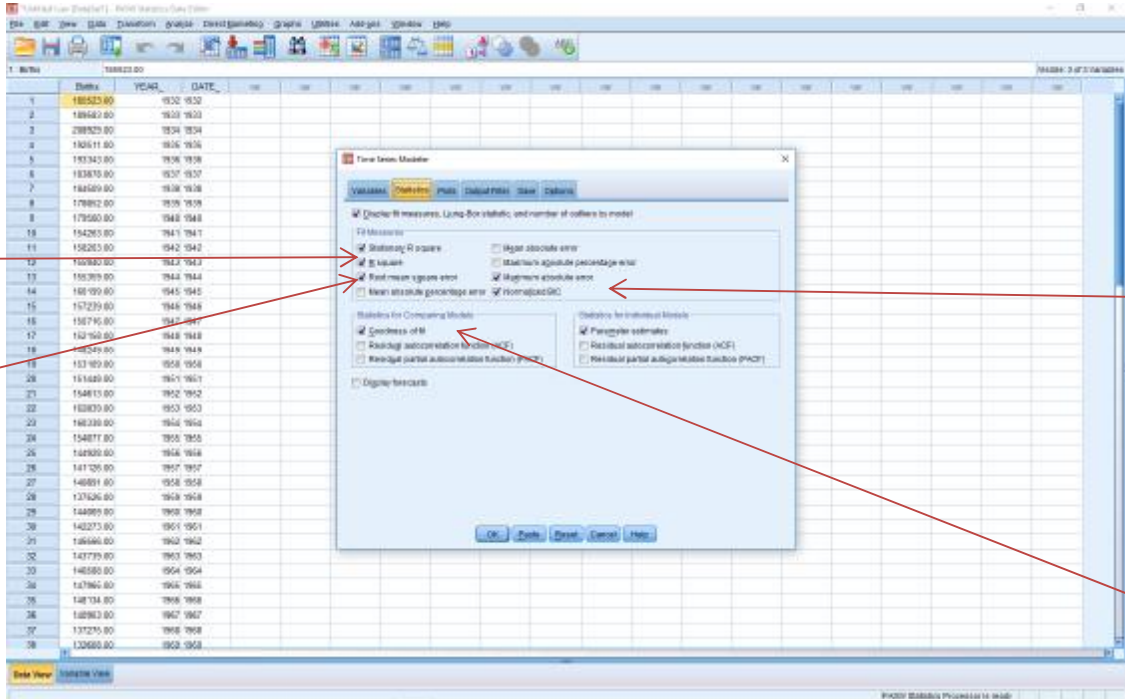
Οπότε $q=2$ και $p=2$. Το υπόδειγμα είναι της μορφής $ARIMA(p, d, q) - ARIMA(2, 1, 2)$ και η θεωρητική του μορφή είναι η εξής:

$$Y_t = \delta + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

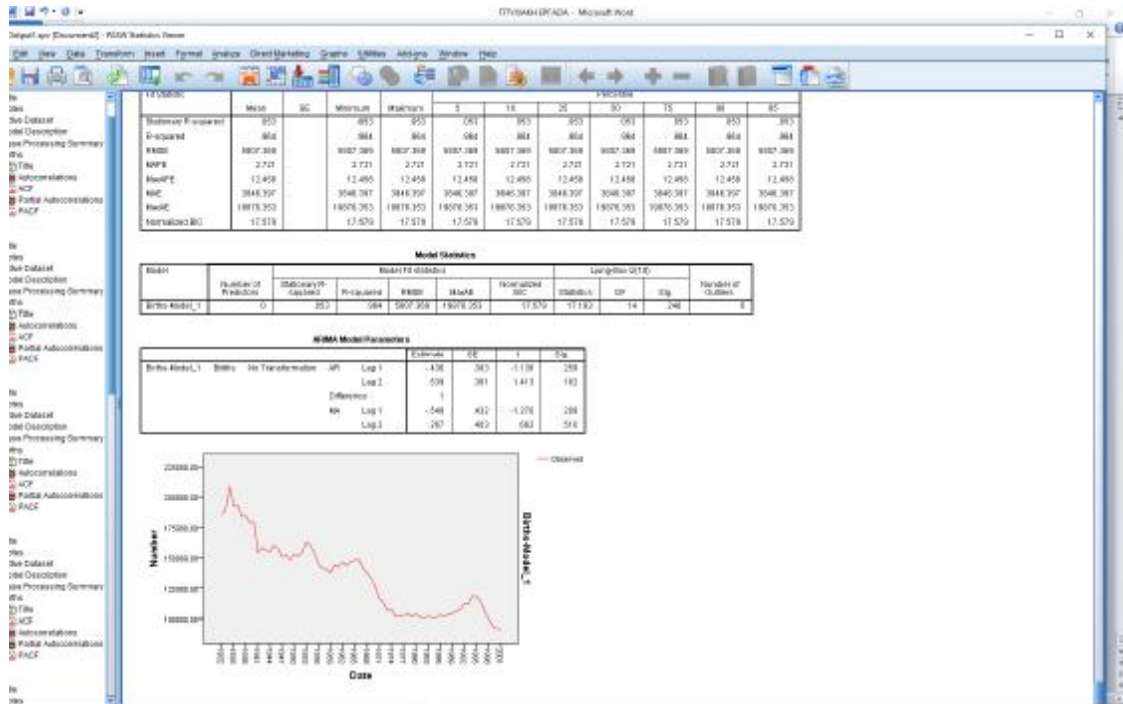
Για να εκτιμηθεί ένα $ARIMA(p, d, q)$ επιλέγουμε *Analyze-Forecasting-Create Models* και ανοίγει το παράθυρο του *Time Series Modeler*. Εδώ θέλουμε να εκτιμήσουμε ένα υπόδειγμα $ARIMA(2, 1, 2)$ για τη χρονοσειρά του παραδείγματός μας. Επιλέγουμε τη μεταβλητή *Births* που περιέχει τις τιμές της χρονοσειράς και τη μετακινούμε στο πλαίσιο *dependent Variable*. Στην κυλιόμενη μπάρα *Method* επιλέγουμε *ARIMA* και πατάμε το κουμπί *Criteria* για να ανοίξει το παράθυρο *Time Series Modeler: ARIMA Criteria*. Στο πλαίσιο του *ARIMA Orders Nonseasonal* πληκτρολογούμε $p = 2$, $d = 1$ και $q = 2$. Στις επιλογές *Transformation* υπάρχει η δυνατότητα εκτίμησης του υποδείγματος μετασχηματίζοντας τις τιμές της χρονοσειράς με λογαρίθμους ή με τετραγωνικές ρίζες. Στην περίπτωση που εξετάζεται αφήνουμε την επιλογή *None*. Η επιλογή *include constant in model* αναφέρεται στο μέσο όρο της χρονοσειράς. Ειδικότερα, εάν η χρονοσειρά έχει μη μηδενικό μέσο όρο κατά τη διαδικασία της εκτίμησης, η εκτίμηση του υποδείγματος γίνεται σε αποκλίσεις από τον μέσο όρο οπότε θα πρέπει να επιλέγεται το *include constant in model*. Στην χρονοσειρά που εκτιμάμε εφαρμόζουμε μετασχηματισμό σε πρώτες διαφορές, δηλαδή ο μέσος όρος της χρονοσειράς θα είναι μηδέν οπότε δεν πρέπει να επιλεγεί το *include constant in model*. Αφού εισαχθούν οι προηγούμενες επιλογές πατάμε *Continue* και επιστρέφουμε στο παράθυρο του *Time Series Modeler*.



Εικόνα 10 (Καθορισμός υποδείγματος ARIMA(2, 1, 2) στο στατιστικό πρόγραμμα



Εικόνα 11 (Διαδικασία για έλεγχο στατιστικής σημαντικότητας)
 Στο Output προκύπτει ο πίνακας ARIMA Model Parameters που χρειάζεται για τον έλεγχο στατιστικής σημαντικότητας στο υπόδειγμα ARIMA(2, 1, 2).



Εικόνα 12 (Πίνακας ARIMA Model Parameters – Output)

ARIMA Model Parameters

				Estimate	SE	t	Sig.	
Births-Model_1	Births	No Transformation	AR	Lag 1	-.436	.383	-1.138	.259
		Difference	1					
	MA	Lag 1		Lag 1	-.549	.432	-1.270	.209

Πίνακας 5 (ARIMA Model Parameters)

Έλεγχος στατιστικής σημαντικότητας p-value των εκτιμητών του υποδείγματος:

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

$$H_0 : \rho = 0 \text{ vs } H_1 : \rho \neq 0$$

$$1) H_0 : \rho_0 = 0 \text{ vs } H_1 : \rho_0 \neq 0$$

$\rho_0 = 0.259 < \alpha = 0.436$ - Απόρριψη H_0 , άρα ο εκτιμητής είναι στατιστικά

σημαντικός (Σ.Σ)

2) $H_0 : \rho_1 = 0$ vs $H_1 : \rho_1 \neq 0$

$\rho_1 = 162 < \alpha = 0.539$ - Απόρριψη H_0 , άρα ο εκτιμητής είναι στατιστικά

σημαντικός (Σ.Σ)

3) $H_0 : \rho_2 = 0$ vs $H_1 : \rho_2 \neq 0$

$\rho_2 = 0.209 < \alpha = 0.549$ - Απόρριψη H_0 , άρα ο εκτιμητής είναι στατιστικά

σημαντικός (Σ.Σ)

4) $H_0 : \rho_3 = 0$ vs $H_1 : \rho_3 \neq 0$

$\rho_3 = 0.510 > \alpha = 0.267$ - Αποδοχή H_0 , άρα ο εκτιμητής είναι στατιστικά μη σημαντικός (Σ.Μ.Σ.)

Αφού βρέθηκαν όλοι οι εκτιμητές στατιστικά σημαντικοί εκτός του σταθερού όρου ο οποίος δεν επηρεάζει το υπόδειγμα και γνωρίζοντας τις παραμέτρους του υποδείγματος ARIMA(2, 1, 2)

Θεωρείται ως το καλύτερο.

$$\alpha_1 = -0.436, \quad \alpha_2 = 0.539, \quad \theta_1 = -0.549, \quad \theta_2 = 0.267$$

Το δ είναι σταθερός όρος, τα α_1, α_2 εκτιμώμενοι παράμετροι αυτοπαλίνδρομης διαδικασίας AR και τα θ_1, θ_2 εκτιμώμενοι παράμετροι κινητού μέσου MA.

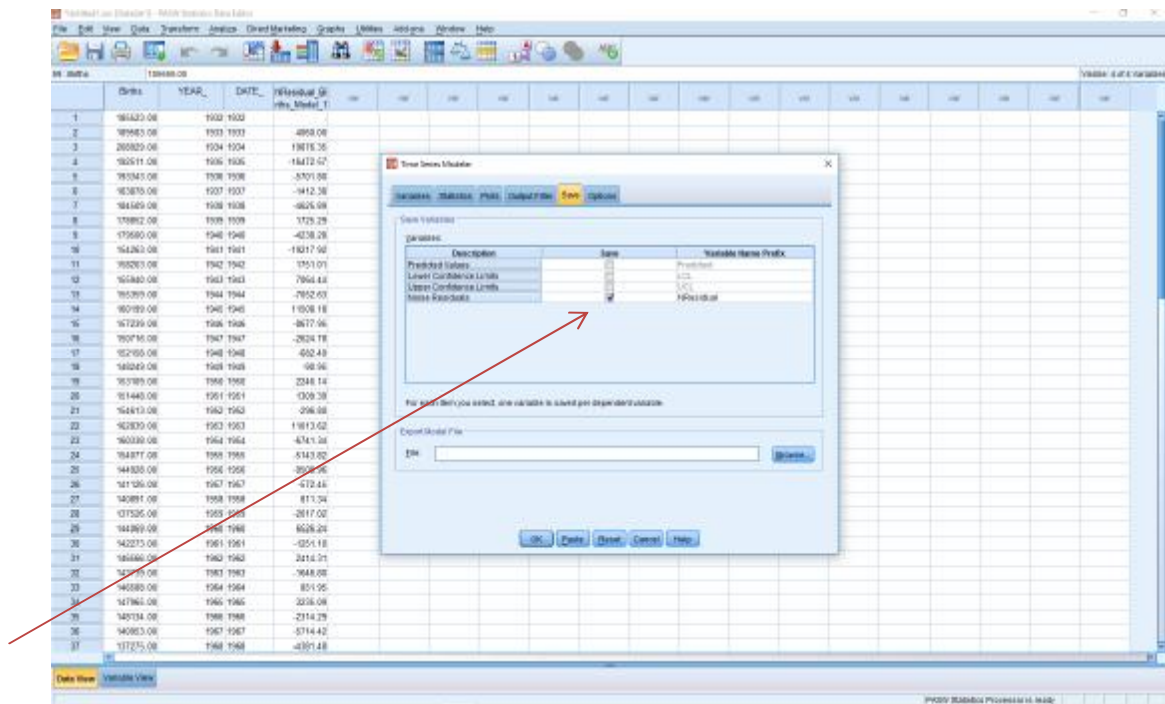
Το υπόδειγμα θα έχει την εξής μορφή:

$$Y_t = \delta - 0.436y_{t-1} + 0.539y_{t-2} - 0.549\varepsilon_{t-1} + 0.267\varepsilon_{t-2} + \varepsilon_t$$

Για να είναι το υπόδειγμα Y_t αποδεκτό, δηλαδή για να θεωρηθεί ότι το υπόδειγμα αυτό περιγράφει τη διαδικασία που παρήγαγε τα δεδομένα, θα πρέπει τα κατάλοιπα να είναι λευκός θόρυβος, οπότε δεν θα πρέπει να αυτοσυσχετίζονται. Για το λόγο αυτό θα γίνει

έλεγχος στασιμότητας στα κατάλοιπα (residuals) επαληθεύοντας έτσι την ορθότητα των αποτελεσμάτων για το ARIMA (2, 1, 2).

Επιστρέφοντας στο menu του ΣΤΑΤΙΣΤΙΚΟ ΠΡΟΓΡΑΜΜΑ πηγαίνουμε στο Analyze-Forecasting-Create Models και στο Time Series Modeler στο πεδίο Save επιλέγεται το Noise Residuals. Πατώντας ok εμφανίζεται η μεταβλητή Noise Residual Births στο ΣΤΑΤΙΣΤΙΚΟ ΠΡΟΓΡΑΜΜΑ που δείχνει τα κατάλοιπα(residuals).



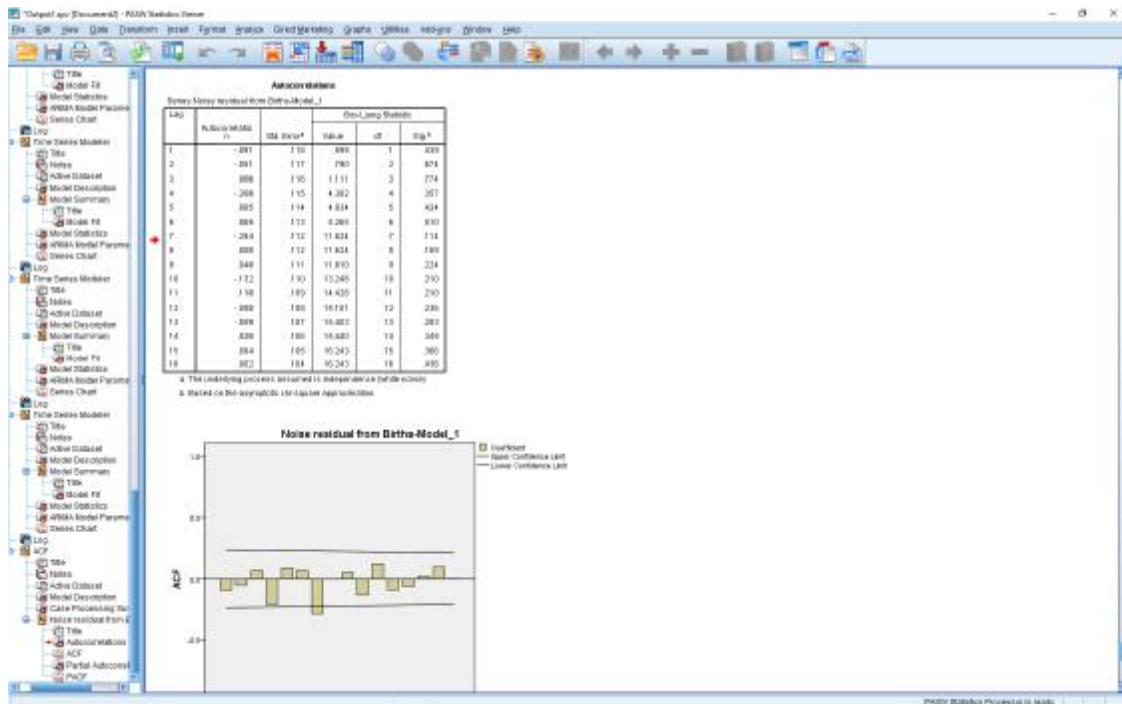
Εικόνα 13-14 (Διαδικασία για την εύρεση των Residuals και τιμές των Residuals)

	Date	YEAR	DATE	NResidual_9 rhw_MoM1_1
1	191420.00	1910	1020	
2	191485.00	1911	1003	-4998.00
3	191550.00	1912	1004	13818.30
4	191615.00	1913	1006	-18472.67
5	191680.00	1914	1008	-3701.00
6	191745.00	1915	1007	-1412.30
7	191810.00	1916	1008	-8626.00
8	191875.00	1917	1009	3729.29
9	191940.00	1918	1048	-4238.20
10	192005.00	1919	1011	-18217.90
11	192070.00	1920	1042	3551.01
12	192135.00	1921	1043	7604.44
13	192200.00	1922	1044	-7852.63
14	192265.00	1923	1045	11926.18
15	192330.00	1924	1046	-8677.90
16	192395.00	1925	1047	-2624.18
17	192460.00	1926	1048	-652.49
18	192525.00	1927	1049	-90.90
19	192590.00	1928	1048	2344.14
20	192655.00	1929	1051	1209.20
21	192720.00	1930	1050	-298.88
22	192785.00	1931	1053	11813.62
23	192850.00	1932	1054	4341.54
24	192915.00	1933	1055	5143.82
25	192980.00	1934	1056	3928.90
26	193045.00	1935	1057	-472.44
27	193110.00	1936	1058	811.24
28	193175.00	1937	1059	-2017.00
29	193240.00	1938	1060	8628.20
30	193305.00	1939	1061	-3251.18
31	193370.00	1940	1062	2114.21
32	193435.00	1941	1063	-3648.88
33	193500.00	1942	1064	821.95
34	193565.00	1943	1065	2238.08
35	193630.00	1944	1066	-2114.29
36	193695.00	1945	1067	5714.42
37	193760.00	1946	1068	-4291.48

Στο menu του στατιστικό πρόγραμμα πηγαίνοντας στο πεδίο Analyze-Forecasting-Autocorrelations βάζουμε τη μεταβλητή NResidual στο Variables και στη συνέχεια πατώντας ok προκύπτει ο πίνακας αυτοσυσχέτισης στο Output και έπειτα κάνουμε έλεγχο στασιμότητας.

	Date	YEAR	DATE	NResidual_9 rhw_MoM1_1
1	191420.00	1910	1020	
2	191485.00	1911	1003	-4998.00
3	191550.00	1912	1004	13818.30
4	191615.00	1913	1006	-18472.67
5	191680.00	1914	1008	-3701.00
6	191745.00	1915	1007	-1412.30
7	191810.00	1916	1008	-8626.00
8	191875.00	1917	1009	3729.29
9	191940.00	1918	1048	-4238.20
10	192005.00	1919	1011	-18217.90
11	192070.00	1920	1042	3551.01
12	192135.00	1921	1043	7604.44
13	192200.00	1922	1044	-7852.63
14	192265.00	1923	1045	11926.18
15	192330.00	1924	1046	-8677.90
16	192395.00	1925	1047	-2624.18
17	192460.00	1926	1048	-652.49
18	192525.00	1927	1049	-90.90
19	192590.00	1928	1048	2344.14
20	192655.00	1929	1051	1209.20
21	192720.00	1930	1050	-298.88
22	192785.00	1931	1053	11813.62
23	192850.00	1932	1054	4341.54
24	192915.00	1933	1055	5143.82
25	192980.00	1934	1056	3928.90
26	193045.00	1935	1057	-472.44
27	193110.00	1936	1058	811.24
28	193175.00	1937	1059	-2017.00
29	193240.00	1938	1060	8628.20
30	193305.00	1939	1061	-3251.18
31	193370.00	1940	1062	2114.21
32	193435.00	1941	1063	-3648.88
33	193500.00	1942	1064	821.95
34	193565.00	1943	1065	2238.08
35	193630.00	1944	1066	-2114.29
36	193695.00	1945	1067	5714.42
37	193760.00	1946	1068	-4291.48

Εικόνα 15 -16 (Διαδικασία για έλεγχο στασιμότητας των κατάλοιπων – Residuals και Πίνακας αυτοσυσχετίσεων στο Output)



Noise residual from Births-Model_1

Autocorrelations

Series: Noise residual from Births-Model_1

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-.091	.118	.599	1	.439
2	-.051	.117	.790	2	.674
3	.066	.116	1.111	3	.774
4	-.208	.115	4.382	4	.357
5	.085	.114	4.934	5	.424
6	.065	.113	5.265	6	.510
7	-.284	.112	11.624	7	.114
8	.000	.112	11.624	8	.169
9	.048	.111	11.810	9	.224
10	-.132	.110	13.248	10	.210
11	.118	.109	14.428	11	.210
12	-.088	.108	15.101	12	.236
13	-.059	.107	15.403	13	.283
14	.020	.106	15.440	14	.349
15	.094	.105	16.243	15	.366
16	.002	.104	16.243	16	.436

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

b. Based on the asymptotic chi-square approximation.

Πίνακας 6 (Autocorrelations)

Έλεγχος Στασιμότητας

$$H_0 : \rho = 0$$

vs

$$H_1 : \rho \neq 0$$

Με την υπόθεση H_0 δεν υπάρχει αυτοσυσχέτιση, άρα η διαδικασία των καταλοίπων είναι στάσιμη ενώ με την υπόθεση H_1 υπάρχει αυτοσυσχέτιση και δεν έχουμε στασιμότητα.

1) $H_0 : \rho_1 = 0$ vs $H_1 : \rho_1 \neq 0$

$$\rho_1 = 0.439 > \alpha = 0.118 \text{ Αποδοχή } H_0 - \text{ Δεν υπάρχει αυτοσυσχέτιση, άρα}$$

στάσιμη

2) $H_0 : \rho_2 = 0$ vs $H_1 : \rho_2 \neq 0$

$$\rho_2 = 0.674 > \alpha = 0.117 \text{ Αποδοχή } H_0 - \text{ Δεν υπάρχει αυτοσυσχέτιση, άρα}$$

στάσιμη

3) $H_0 : \rho_3 = 0$ vs $H_1 : \rho_3 \neq 0$

$$\rho_3 = 0.774 > \alpha = 0.116 \text{ Αποδοχή } H_0 - \text{ Δεν υπάρχει αυτοσυσχέτιση, άρα}$$

στάσιμη

4) $H_0 : \rho_4 = 0$ vs $H_1 : \rho_4 \neq 0$

$$\rho_4 = 0.357 > \alpha = 0.115 \text{ Αποδοχή } H_0 - \text{ Δεν υπάρχει αυτοσυσχέτιση, άρα}$$

στάσιμη

5) $H_0 : \rho_5 = 0$ vs $H_1 : \rho_5 \neq 0$

$$\rho_5 = 0.424 > \alpha = 0.114 \text{ Αποδοχή } H_0 - \text{ Δεν υπάρχει αυτοσυσχέτιση, άρα}$$

στάσιμη

6) $H_0 : \rho_6 = 0$ vs $H_1 : \rho_6 \neq 0$

$\rho_6 = 0.510 > \alpha = 0.113$ Αποδοχή H_0 - Δεν υπάρχει αυτοσυσχέτιση, άρα

στάσιμη

7) $H_0 : \rho_7 = 0$ vs $H_1 : \rho_7 \neq 0$

$\rho_7 = 0.114 > \alpha = 0.112$ Αποδοχή H_0 - Δεν υπάρχει αυτοσυσχέτιση, άρα

στάσιμη

8) $H_0 : \rho_8 = 0$ vs $H_1 : \rho_8 \neq 0$

$\rho_8 = 0.169 > \alpha = 0.111$ Αποδοχή H_0 - Δεν υπάρχει αυτοσυσχέτιση, άρα

στάσιμη

.....

16) $H_0 : \rho_{16} = 0$ vs $H_1 : \rho_{16} \neq 0$

$\rho_{16} = 0.436 > \alpha = 0.104$ Αποδοχή H_0 - Δεν υπάρχει αυτοσυσχέτιση, άρα

Στάσιμη

Ο έλεγχος καταλοίπων δείχνει ότι η διαδικασία είναι στάσιμη, καθώς για όλες τις χρονολογικές υστερήσεις τα ρ είναι στάσιμα. Έτσι το υπόδειγμα ARIMA(2, 1, 2) επιβεβαιώνεται ότι είναι το πλέον κατάλληλο.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Cryer, Jonathan D., Chan, Kung-Sik (2008) *Time Series Analysis, With Applications in R*, Springer-Verlag New York
- George E. P. Box, G. M. (2015) *Time Series Analysis: Forecasting and Control, 5th Edition*. WILEY
- Jenkins, G. M., & Box, G. E. (1976) *Time series analysis: Forecasting and Control 2nd Edition*. Oakland, California: Holden-Day.
- Ruey S. Tsay (2005) *Analysis of Financial Time Series, Second Edition*. John Wiley & Sons, Inc.
- Walter Enders, (2014) *Applied Econometric Time Series, 4th Edition*. WILEY
- Βενέτης, Ι. (2013). *Εισαγωγή στην Οικονομετρία*. Πάτρα: Gotsis Εκδόσεις.
- Δημελή, Σ. (2013). *Σύγχρονες Μέθοδοι Ανάλυσης Χρονολογικών Σειρών*. Αθήνα: ΕΚΔΟΣΕΙΣ ΟΠΑ ΑΕ.
- Θαλασσινός, Ε. (1991). *Ανάλυση Χρονολογικών Σειρών Μεθοδολογία Box Jenkins*. Αθήνα: ΕΚΔΟΣΕΙΣ ΣΤΑΜΟΥΛΗ.
- Παπαδημάς, Ο., & Κοΐλιας, Χ. (2002). *Εφαρμοσμένη Στατιστική*. Αθήνα: Εκδόσεις Νέων Τεχνολογιών.
- Ανδρουλάκης, Γ. (n.d.). *ΜΟΝΤΕΛΑ των Box & Jenkins*. Ανάκτηση από http://androulakis.bma.upatras.gr/mediawiki/index.php/%CE%9C%CE%BF%CE%BD%CF%84%CE%AD%CE%BB%CE%B1_%CF%84%CF%89%CE%BD_Box_%CE%BA%CE%B1%CE%B9_Jenkins#.CE.92.CE.B9.CE.B2.CE.BB.CE.B9.CE.BF.CE.B3.CF.81.CE.B1.CF.86.CE.AF.CE.B1
- Ανδρουλάκης, Γ. (n.d.). *ΔΙΑΜΕΡΙΣΜΟΣ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ ΣΕ ΣΥΝΙΣΤΩΣΕΣ*. Ανάκτηση από http://androulakis.bma.upatras.gr/mediawiki/index.php/%CE%94%CE%B9%CE%B1%CE%BC%CE%B5%CF%81%CE%B9%CF%83%CE%BC%CF%8C%CF%82_%CF%87%CF%81%CE%BF%CE%BD%CE%BF%CE%BB%CE%BF%CE%B3%CE%B9%CE%BA%CF%8E%CE%BD_%CF%83%CE%B5%CE%B9%CF%81%CF%8E%CE%BD_%CF%83%CE%B5_%CF%83%CF%8
- Ανδρουλάκης, Γ. (n.d.). *ΠΑΡΑΔΕΙΓΜΑΤΑ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ*. Ανάκτηση από http://androulakis.bma.upatras.gr/mediawiki/index.php/%CE%A0%CE%B1%CF%81%CE%B1%CE%B4%CE%B5%CE%AF%CE%B3%CE%BC%CE%B1%CF%84%CE%B1_%CF%87%CF%81%CE%BF%CE%BD%CE%BF%CE%BB%CE%BF%CE%B3%CE%B9%CE%BA%CF%8E%CE%BD_%CF%83%CE%B5%CE%B9%CF%81%CF%8E%CE%BD

- Ζάρλα, Α. (n.d.). *ΑΝΑΛΥΣΗ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ*. Ανάκτηση από
<http://nemertes.lis.upatras.gr/jspui/bitstream/10889/899/1/%CE%94%CE%B9%CF%80%CE%BB%CF%89%CE%BC%CE%B1%CF%84%CE%B9%CE%BA%CE%AE%CE%B5%CF%81%CE%B3%CE%B1%CF%83%CE%AF%CE%B1.pdf>
- Καρβέλης, Χ. (2008). *ΑΝΑΛΥΣΗ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ: ΠΡΟΒΛΕΠΟΝΤΑΣ ΤΟ ΜΕΛΛΟΝ, ΚΑΤΑΝΟΩΝΤΑΣ ΤΟ ΠΑΡΕΛΘΟΝ*. Ανάκτηση από
<http://nemertes.lis.upatras.gr/jspui/bitstream/10889/1401/1/%CE%9A%CE%B1%CF%81%CE%B2%CE%B5%CE%BB%CE%B7%CF%82%20%CE%A7%CE%B1%CF%81%CE%B1%CE%BB%CE%B1%CE%BC%CF%80%CE%BF%CF%82%20%CE%91%CE%9C.%20109%20%CE%91%CE%BD%CE%B1%CE%BB%CF%85%CF%83%CE%B7%20%CE%A7%CF%81%CE>
- Κουνέτας, Κ. (2012, Δεκέμβριος). *ΣΗΜΕΙΩΣΕΙΣ ΜΑΘΗΜΑΤΟΣ: ΤΕΧΝΙΚΕΣ ΠΡΟΒΛΕΨΕΩΝ & ΕΛΕΓΧΟΥ*. Ανάκτηση από
<http://eclass.teipat.gr/eclass/modules/document/file.php/728131/%CE%98%CE%95%CE%A9%CE%A1%CE%99%CE%91/%CE%A4%CE%B5%CF%87%CE%BD%CE%B9%CE%BA%CE%AD%CF%82%20%CE%A0%CF%81%CE%BF%CE%B2%CE%BB%CE%AD%CF%88%CE%B5%CF%89%CE%BD%20%26%20%CE%95%CE%BB%CE%AD%CE%B3%CF%87%CE>
- Σαρριανίδης, Ν. (n.d.). *ΑΝΑΛΥΣΗ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ*. Ανάκτηση από
<http://openclass.teiwm.gr/modules/document/file.php/BA-K117/4/%CE%91%CE%BD%CE%AC%CE%BB%CF%85%CF%83%CE%B7%20%CE%A7%CF%81%CE%BF%CE%BD%CE%BF%CE%BB%CE%BF%CE%B3%CE%B9%CE%BA%CF%8E%CE%BD%20%CE%A3%CE%B5%CE%B9%CF%81%CF%8E%CE%BD.pdf>