ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΠΑΤΡΑΣ
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ: ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΕΠΟΧΙΚΑ ΜΟΝΤΕΛΑ ΧΡΟΝΟΣΕΙΡΩΝ.ΕΦΑΡΜΟΓΗ ΣΤΟ ΧΩΡΟ ΕΠΙΧΕΙΡΗΣΕΩΝ ΟΙΚΟΝΟΜΙΑΣ

ΣΠΟΥΔΑΣΤΗΣ: ΔΕΔΟΠΟΥΛΟΣ ΚΩΝΣΤΑΝΤΙΝΟΣ

ΕΙΣΗΓΗΤΡΙΑ: ΜΠΟΥΜΠΟΥΛΗ ΑΘΑΝΑΣΙΑ

ΠΑΤΡΑ, ΜΑΡΤΙΟΣ 2009

ΠΕΡΙΕΧΟΜΕΝΑ

<u>ΚΕΦΑΛΑΙΟ 3: ΜΑΘΗΜΑΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ ΣΤΑ ΔΙΑΦΟΡΑ ΜΟΝΤΕΛΑ ΧΡΟΝΟΣΕΙΡΩΝ</u>

<u>ΚΕΦΑΛΑΙΟ 4: ΕΠΟΧΙΚΟΤΗΤΑ ΚΑΙ ΟΙ ΕΦΑΡΜΟΓΕΣ ΤΗΣ</u>

Αφιέρωση,

Στους γονείς μου , Γεώργιο και Παρασκευή.

Ευχαριστίες,

Θα ήθελα να ευχαριστήσω πολύ όσους μου συμπαραστάθηκαν κατά τη διάρκεια της εργασίας αυτής, γιατί όντως υπήρξαν πολλές δύσκολες στιγμές.

Πιο συγκεκριμένα θέλω να ευχαριστήσω τις καθηγήτριές, μου κα Καρυώτη Βασιλική και κα Μπουμπούλη Αθανασία, για την καθοδήγησή τους κατά την εκπόνηση της εργασίας, καθώς και τη βιβλιοθήκη του Τει Πάτρας για την παροχή πληροφοριών που μου παρείχε.

ΚΕΦΑΛΑΙΟ 1 :   ΕΙΣΑΓΩΓΗ

Τα έθνη για να επιβιώσουν και να αναπτυχθούν πρέπει να προγραμματίζουν για το μέλλον. Οι σύγχρονες επιχειρήσεις πρέπει να καταρτίζουν προγράμματα για την παραγωγή, για τις επενδύσεις, για την έρευνα της αγοράς και την ανάπτυξη των πωλήσεών τους, για να αντιμετωπίζουν του ανταγωνισμό και να αυξήσουν τα κέρδη τους. Οι εκάστοτε κυβερνήσεις καταρτίζουν προγράμματα οικονομικής αναπτύξεως, με τα οποία προκαθορίζουν προσεγγιστικά, τη διαμόρφωση όλων των μεγεθών της οικονομίας , ώστε να γνωρίζουν κατά προσέγγιση τα έσοδά τους για να προγραμματίζουν τις δαπάνες τους.

Χρονοσειρά ονομάζεται το σύνολο από χρονολογικά ταξινομημένες παρατηρήσεις σε σχέση με μία μεταβλητή κατά τη διάρκεια διαδοχικών και ισόχρονων περιόδων, εφόσον σημειώσουμε πως η ανάλυση χρονοσειρών είναι μια στατιστική μέθοδος ανάλυσης όπου τα δεδομένα προκύπτουν από παρατηρήσεις που συλλέγονται κατά τη διάρκεια μιας σχετικά μακριάς χρονικής περιόδου. Οι μέθοδοι ανάλυσης των χρονοσειρών βρίσκουν εφαρμογή στην κλιματολογία, την οικονομία, την φυσική κ.λ.π.

Χαρακτηριστικό κάθε χρονοσειράς, είναι η άμεση εξάρτηση μεταξύ των διαδοχικών τιμών της, η οποία αποτελεί και το αντικείμενο μελέτης του κλάδου με απώτερο σκοπό, τη χρησιμοποίησή τους στη διενέργεια προβλέψεων. Πολλές εταιρίες βέβαια δεν ασχολούνται με την πρόβλεψη γιατί έχουν την πεποίθηση πως ούτε το μέλλον επιφυλάσσει σημαντικές αλλαγές, ούτε θα υπάρξει αρκετός χρόνος (στο μέλλον) ώστε να προλάβει να αντιδράσει στις αλλαγές των γεγονότων.

Κάθε ερευνητής ο οποίος ασχολείται με την ανάλυση των χρονοσειρών έχει σαν στόχο να γίνει ο προσδιορισμός της φύσης του φαινομένου που αντιπροσωπεύεται από την ακολουθία των επιχειρήσεων και στο τέλος να γίνει πρόβλεψη των μελλοντικών τιμών της χρονοσειράς. Αυτό θα επιτευχθεί με διάφορες μεθόδους οι οποίοι διαφέρουν ως προς την πολυπλοκότητα, την ταχύτητα και το κόστος υπολογισμού τους, τη διαθεσιμότητα των απαραίτητων δεδομένων και των άλλων παραγόντων κατά περίπτωση.

Με σκοπό την επίτευξη πρόβλεψης πρέπει να εξετασθούν παρελθοντικά δεδομένα, επομένως πρόβλεψη, αποτελεί η διαδικασία η οποία βάση της ανάλυσης , με τη βοήθεια κάποιων απαραίτητων στατιστικών μεθόδων, της πορείας ενός συγκεκριμένου μεγέθους, κατά το παρελθόν και το παρόν έχοντας ως σκοπό να μας δώσει τις τιμές αυτού στο τέλος. Ακολουθείται η παρακάτω διαδικασία:

ΙΣΤΟΡΙΚΑ ΔΕΔΟΜΕΝΑ + ΜΕΘΟΔΟΛΟΓΙΑ à ΚΑΤΑΣΚΕΥΗ ΥΠΟΔΕΙΓΜΑΤΟΣ à ΑΞΙΟΠΟΙΗΣΗ ΥΠΟΔΕΙΓΜΑΤΟΣ à ΠΡΟΒΛΕΨΕΙΣ

Οι προβλέψεις ανάλογα με τον χρονικό τους ορίζοντα χωρίζονται στις ακόλουθες κατηγορίες:

- **Βραχυπρόθεσμες**: από τους επόμενους μήνες εώς και τρία χρόνια
- **Μεσοπρόθεσμες**:μεταξύ τριών και πέντε χρόνων
- **Μακροπρόθεσμες**: μεταξύ 5 και 15 χρόνων
- Και τέλος **τεχνολογικές** προβλέψεις μεταξύ 15 και 30 χρόνια.

Οι μεσοπρόθεσμες προβλέψεις σχετίζονται με τον προγραμματισμό των εταιρειών, ενώ οι μακροπρόθεσμες με τον τεχνολογικό εξοπλισμό τους.

Σκοπός της συγκεκριμένης εργασίας είναι η μελέτη της έννοιας χρονοσειράς και πιο συγκεκριμένα των εποχικών μοντέλων χρονοσειρών και η εφαρμογή τους στο χώρο των επιχειρήσεων οικονομίας. Στο δεύτερο κεφάλαιο θα γίνει μια πρώτη επαφή με τις βασικές έννοιες των χρονοσειρών από περιγραφικής απόψεως, ενώ στο τρίτο θα γίνει προσπάθεια, για μία πιο μαθηματική προσέγγιση τη λεγόμενη ανάλυση στο πεδίο του χρόνου.

Παρακάτω, στο τέταρτο κεφάλαιο θα επεκταθούμε στην έννοια της εποχικότητας και των εποχικών δεικτών καθώς και στις προβλέψεις και τα χαρακτηριστικά τους. Η εποχικότητα ποικίλει από έτος σε έτος δείχνοντας ότι τα μοντέλα τα οποία βασίζονται σε μη προσαρμοσμένα δεδομένα είναι πολύ πιθανό να είναι περισσότερο ευέλικτα.

Θα αναφερθεί επίσης και η μεθοδολογία Box- Jenkins με τα πλεονεκτήματα, καθώς και τα μειονεκτήματά της.

Στο τελευταίο κεφάλαιο θα γίνει παρουσίαση και ανάλυση των αποτελεσμάτων μιας εφαρμογής, αναφερόμενης σε ένα παράδειγμα με δεδομένα τα οποία παρουσιάζουν εποχικότητα.

ΚΕΦΑΛΑΙΟ 2: ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΧΡΟΝΟΣΕΙΡΩΝ-ΠΕΡΙΓΡΑΦΙΚΕΣ ΜΕΘΟΔΟΙ

2.1 Εισαγωγή

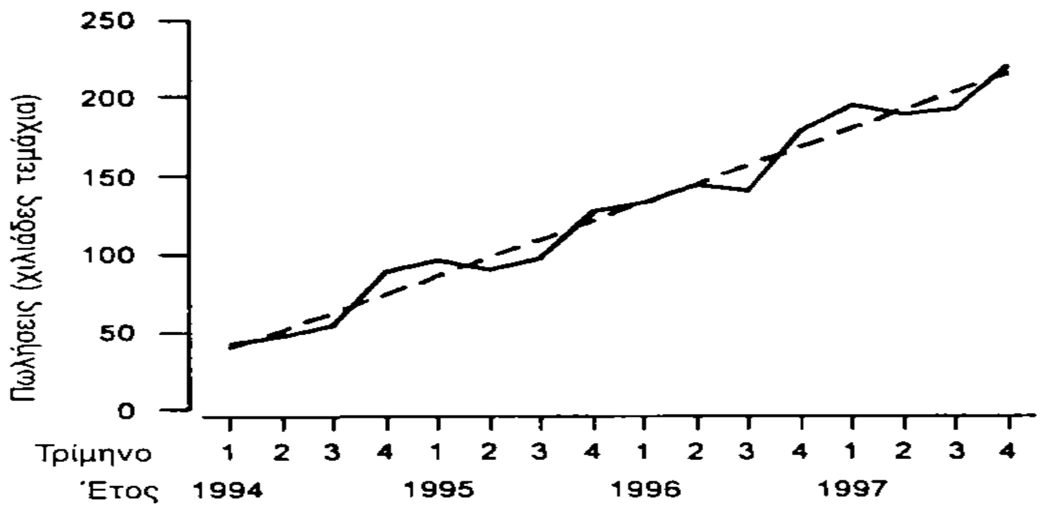
Στο παρόν κεφάλαιο θα ασχοληθούμε με την παρουσίαση από περιγραφικής απόψεως κάποιων βασικών μεθόδων ανάλυσης χρονοσειρών. Θα αναφερθούν οι τέσσερις συνιστώσες, που εμφανίζονται κατά τη μελέτη μιας χρονοσειράς, οι διάφορες τεχνικές με τις οποίες αυτές μελετώνται, καθώς και η έννοια της εποχικότητας που θα μας απασχολήσει και περισσότερο.

2.2 Ορισμός

Ένας ορισμός για τη χρονοσειρά θα μπορούσε να είναι ο εξής:
Χρονοσειρά ή χρονολογική σειρά είναι η σειρά των τιμών της μεταβλητής που παίρνει σε διαδοχικές χρονικές στιγμές ή περιόδους. Απαραίτητο στοιχείο είναι να γνωρίζουμε τη μονάδα μέτρησης (σε χρόνια, μήνες, εβδομάδες) η οποία συμβολίζεται με: Y_t ή X_t ή W_t ή Z_t (Καρυώτη Βασιλική, 2004). Μια μεταβλητή μπορεί να είναι συνεχής ως προς το χρόνο, δηλαδή να παίρνει τιμές σε κάθε χρονική στιγμή για παράδειγμα η θερμοκρασία, ή ασυνεχής όταν παίρνει τιμές μόνο σε ορισμένες χρονικές στιγμές για παράδειγμα το κλείσιμο τιμών χρηματιστηρίου.

2.3 Γραφική παράσταση χρονοσειράς

Όπως φαίνεται παρακάτω μια χρονοσειρά που περιλαμβάνει μια μεταβλητή Y παριστάνεται γραφικά εάν κατασκευάσουμε τη γραφική παράσταση του Y προς t. Έτσι και στο σχήμα έχουμε την γραφική παράσταση της χρονοσειράς που δείχνει τις τριμηνιαίες πωλήσεις των δίσκων CD κάποιας δισκογραφικής εταιρείας. Ο κάθετος άξονας απεικονίζει τις πωλήσεις ενώ ο οριζόντιος τα τρίμηνα 4 διαφορετικών ετών.



Σχήμα 2.1 Τριμηνιαίες πωλήσεις δίσκων CD κάποιας εταιρίας

Πηγή:Εκδόσεις Τζιόλα "Στατιστική 3η Έκδοση"

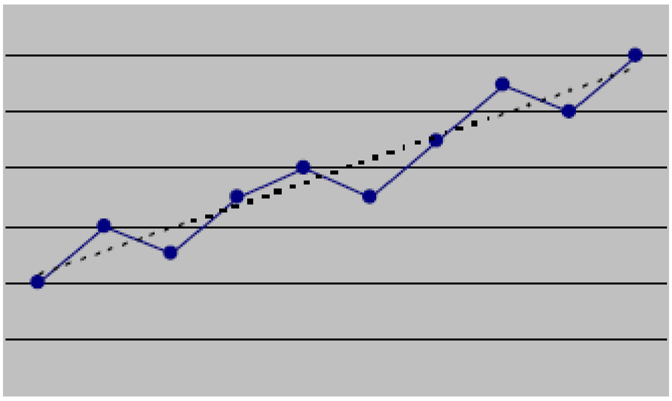
2.4 <u>Συνιστώσες Χρονοσειρών</u>

Οι τιμές των χρονολογικών σειρών που παρατηρούνται προέρχονται από το αποτέλεσμα της ταυτόχρονης επίδρασης τεσσάρων συνιστωσών οι οποίες είναι οι ακόλουθες: η τάση, ο κύκλος, η εποχικότητα και οι τυχαίες κυμάνσεις.

2.4.1 <u>Τάση</u>

Τάση ονομάζεται η μακροχρόνια και συνεχής κίνηση που ακολουθεί η χρονοσειρά. Αποτελεί την εξέλιξη της σειράς για μεγάλες χρονικές περιόδους, (πάνω από δέκα έτη), και είναι κατά μέσο όρο απαλλαγμένη από βραχυχρόνιες αυξομειώσεις. Γι'αυτό ονομάζεται και μακροχρόνια τάση. Η τάση μπορεί να εμφανιστεί με γραμμική ή καμπυλόγραμμη μορφή. Και μπορεί να είναι ανοδική ή καθοδική.

Η συνιστώσα αυτή αντιπροσωπεύει την επίδραση μακρόχρονων παραγόντων των οποίων το αποτέλεσμα τείνει να αλλάξει την πορεία του μεγέθους βαθμιαία. Τέτοιοι παράγοντες είναι δημογραφικά χαρακτηριστικά, γεωγραφική κατανομή του πληθυσμού, τεχνολογικά επιτεύγματα, οικονομικές εξελίξεις. Ακολουθεί ένα διάγραμμα που απεικονίζει μια μορφή ανοδικής τάσης.



Σχήμα 2.2 Πηγή:Μιχάλης Βαϊδάνης 2005

2.4.2 Κυκλική Συνιστώσα

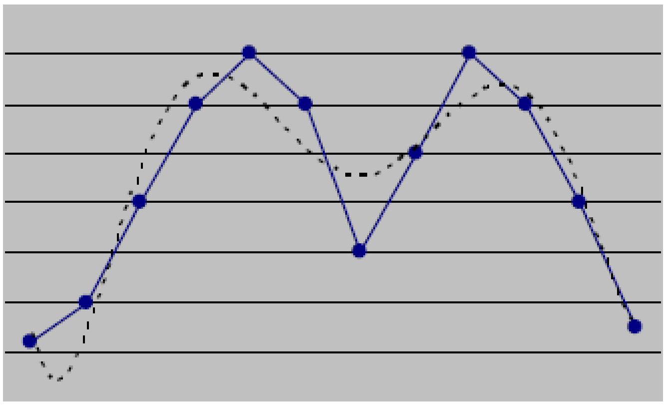
Εδώ πρόκειται για κυματοειδείς διακυμάνσεις γύρω από την τάση που καλύπτουν χρονικό διάστημα μεγαλύτερο του έτους και προέρχονται από μεταβολές της γενικής οικονομικής δραστηριότητας και μπορεί να έχουν είτε ανοδικές είτε καθοδικές φάσεις. Οι κυκλικές κινήσεις αποτελούν τη διαφορά μεταξύ της αναμενόμενης τιμής της χρονοσειράς που προέρχεται από την τάση και την πραγματική τιμή της χρονοσειράς.

Σε κάποιες δυτικές χώρες όπου οι οικονομίες τους είναι αρκετά ανεπτυγμένες, τα περισσότερα οικονομικά μεγέθη όπως επενδύσεις, τιμές, εισόδημα κ.λ.π. παρουσιάζουν λίγο ή πολύ έντονες κυμάνσεις. Το μεγαλύτερο πρόβλημα που αντιμετωπίζουμε στη μελέτη των οικονομικών κύκλων είναι ότι η διάρκειά τους δεν είναι σταθερή.

2.4.3 Εποχική συνιστώσα

Η εποχική συνιστώσα αποτελεί μια κυκλική κύμανση με περίοδο το έτος γιατί μέσα σ’αυτό εξαντλεί όλες τις ανοδικές και καθοδικές κινήσεις. Επειδή επαναλαμβάνεται ρυθμικά κάθε έτος είναι και περιοδική. (Μπόρα-Σέντα Ε., 1997). Η εποχική κύμανση που το όνομά της προέρχεται από το γεγονός ότι συνδέεται με τις εποχές, εμφανίζεται μόνο στις χρονολογικές σειρές με εποχικές παρατηρήσεις (τετραμηνιαία, τριμηνιαία κ.λπ.).

Τα διάφορα κοινωνικά φαινόμενα (εορτές, θερινές διακοπές κ.λ.π.) ο μεταβαλλόμενος αριθμός των εργάσιμων ημερών μεταξύ των μηνών του έτους, το διαφορετικό ωράριο των καταστημάτων, η κατανάλωση περισσότερων παγωτών κατά τους ζεστούς μήνες, είναι μερικές από τις αιτίες των περιοδικών κυμάνσεων που εμφανίζουν οι χρονολογικές σειρές με εποχικά δεδομένα.



Σχήμα 2.3 Πηγή:Μιχάλης Βαϊδάνης 2005

2.4.4 <u>Τυχαία Συνιστώσα</u>

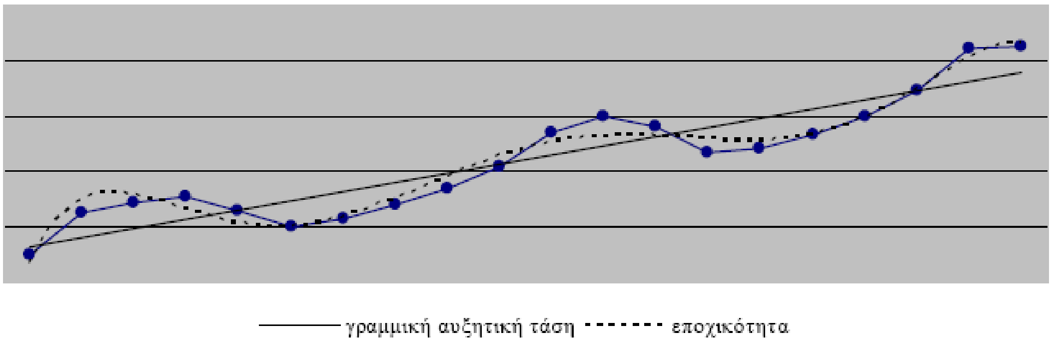
Οποιαδήποτε επίδραση στη διαμόρφωση της τιμής δεν οφείλεται σε κάποια από τις παραπάνω συνιστώσες θεωρείται τυχαία κύμανση.

Η τυχαία συνιστώσα εμφανίζεται ακανόνιστα με επιδράσεις άλλοτε αρνητικές και άλλοτε θετικές. Επίσης οι τυχαίες κυμάνσεις δεν μπορούν να προβλεφθούν. Παράδειγμα τέτοιων επιδράσεων αποτελούν οι εξαγγελίες κυβερνητικών μέτρων, ασυνήθιστες κλιματολογικές συνθήκες κ.α.

Πρέπει να αναφερθεί πως οι τρεις πρώτες συνιστώσες είναι συστηματικές, δηλαδή ακολουθούν συγκεκριμένο τρόπο επίδρασης και με κάποιες μεθόδους μπορούμε να τις απομονώσουμε και να τις μετρήσουμε. Αυτό όμως, δεν συμβαίνει και με την τέταρτη.

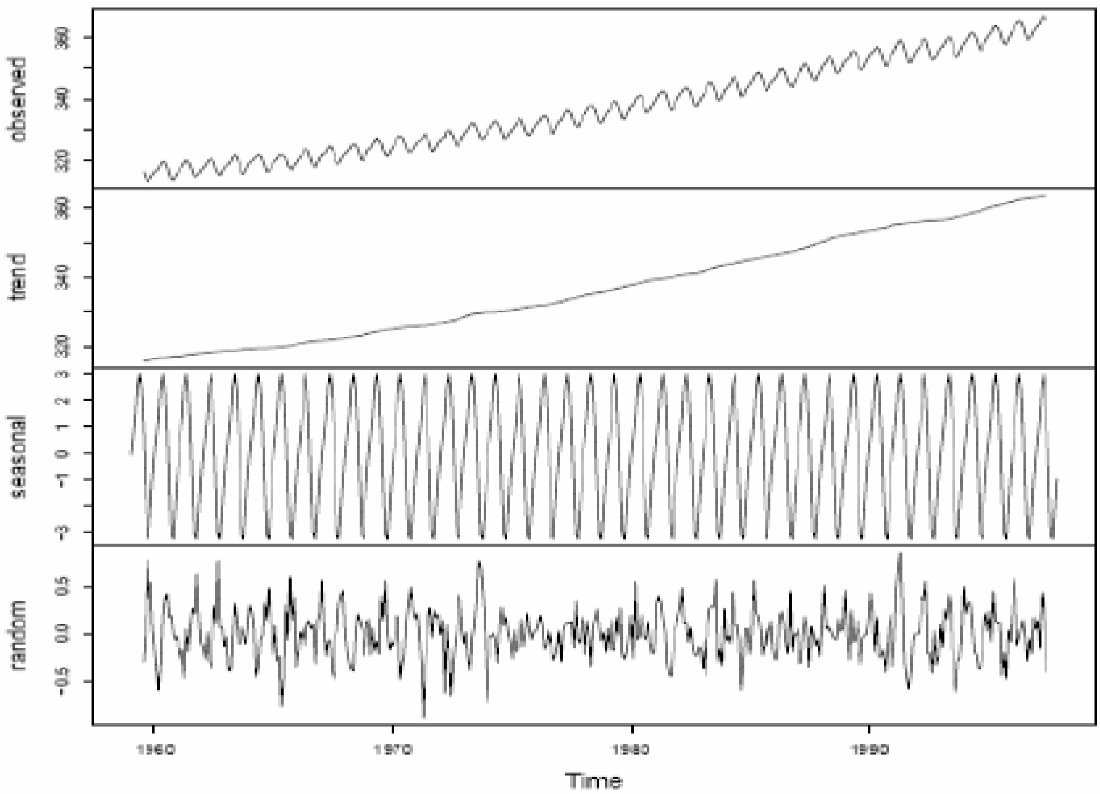
Αν εξετάσουμε μια οποιαδήποτε χρονοσειρά, διαπιστώνουμε ότι αποτελεί συνδυασμό ενός ή περισσότερων από τα παραπάνω στοιχεία.

Ακολουθεί ένα διάγραμμα όπου φαίνεται μια χρονοσειρά με εμφανή την ύπαρξη αυξητικής τάσης και, ταυτόχρονα, εποχικότητας.



Σχήμα 2.4 Μιχάλης Βαϊδάνης 2005

Παρακάτω απεικονίζονται χωριστά οι τέσσερις συνιστώσες στις οποίες έγινε αναφορά νωρίτερα και φαίνεται καθαρά η διαγραμματική τους μορφή.



Σχήμα 2.5 ***ΕΜΠ ΜΜ ΒΔΕΕ 2006* | *Τζιραλής Γεώργιος*, gtzi@central.ntua.gr**

Στη συνέχεια θα ασχοληθούμε με το προσθετικό καθώς και το πολλαπλασιαστικό μοντέλο που αποτελούν τους δύο τρόπους με τους οποίους συνδέονται μεταξύ τους οι συνιστώσες των χρονοσειρών.

2.5.1 Προσθετικό Μοντέλο

Στο συγκεκριμένο μοντέλο η $\mathbf{Y_t}$ (η τιμή της μεταβλητής Y τη χρονική περίοδο t) προκύπτει από τον ακόλουθο τύπο:

$\Upsilon_t = T_t + C_t + S_t + I_t$, όπου:

Υ_t = η τιμή της σειράς Y στη χρονική περίοδο t

T_t = η τιμή της τάσης

C_t = η επίδραση του κύκλου

S_t = η επίδραση της εποχικής συνιστώσας

I_t = η επίδραση της τυχαίας συνιστώσας

Βασικό στοιχείο του προσθετικού μοντέλου είναι ότι και οι τέσσερις συνιστώσες έχουν την ίδια μονάδα μέτρησης, δηλαδή τη μονάδα μέτρησης της μεταβλητής Y.

2.5.2 Πολλαπλασιαστικό μοντέλο

Οι τέσσερις συνιστώσες της χρονοσειράς μπορούν να συνδεθούν και με τον πολλαπλασιασμό τους για κάθε χρονική περίοδο. Πιο συγκεκριμένα στην περίοδο t η παρατήρηση $\mathbf{Y_t}$ προκύπτει σύμφωνα με τον ακόλουθο τύπο:

$\Upsilon_t = T_t \cdot C_t \cdot S_t \cdot I_t$, ο οποίος ονομάζεται πολλαπλασιαστικό μοντέλο.

Εδώ σε αντίθεση με το προσθετικό μοντέλο, μόνο η συνιστώσα της τάσης (T) εκφράζεται στην ίδια μονάδα μέτρησης με την μεταβλητή Y. Οι άλλες τρεις συνιστώσες είναι καθαροί αριθμοί χωρίς μονάδα μέτρησης.

Για τους μελετητές των χρονοσειρών απαραίτητη ενέργεια αποτελεί η εκτίμηση των τιμών των συνιστωσών για όλες τις περιόδους που καλύπτει η χρονολογική σειρά. Η τάση ο κύκλος και εποχικότητα (συστηματικές συνιστώσες), ακολουθούν κάποιο συγκεκριμένο υπόδειγμα. Έτσι, με το στατιστικό προσδιορισμό του, θα μπορέσουμε να προβλέψουμε την εξέλιξη της σειράς στο μέλλον, με την υπόθεση ότι οι παραπάνω τρεις συνιστώσες να συνεχίσουν να έχουν και μελλοντικά την ίδια συμπεριφορά.

2.6 Εκτίμηση της τάσης

Η εκτίμηση της τάσης μιας χρονολογικής σειράς μπορεί να γίνει είτε με την εκτίμηση ενός υποδείγματος που περιγράφει τη μεταβολή της τάσης είτε με τη μέθοδο των κινητών μέσων όρων που απλώς εξομαλύνει τη χρονοσειρά και την απαλλάσσει από βραχυχρόνιες αυξομειώσεις.

Λέγοντας εκτίμηση ενός μοντέλου εννοούμε τη μαθηματική έκφραση της τάσης με τη βοήθεια μιας εξίσωσης όπου η ανεξάρτητη μεταβλητή είναι ο χρόνος (t) και η εξαρτημένη, η χρονολογική σειρά (Y).

Τα μοντέλα που κυρίως χρησιμοποιούνται είναι το γραμμικό και το εκθετικό, για τα οποία ισχύει:

$$\hat{Y}_t = a + bt \ (1) \quad \text{και} \quad Y_t = (a\,b^{X_t}) \ (2)$$

Τα δύο μοντέλα μπορούν να εκτιμηθούν με τη μέθοδο των ελαχίστων τετραγώνων με τη μόνη διαφορά ότι, επειδή το εκθετικό μοντέλο είναι μη γραμμικό ως προς τις παραμέτρους, πρέπει πρώτα να λογαριθμήσουμε τη σχέση (2) και στη συνέχεια να εφαρμόσουμε τη μέθοδο των ελαχίστων τετραγώνων.

2.6.1 Γραμμικό Μοντέλο

Οι μελετητές εδώ χρησιμοποιούν τη μέθοδο των ελαχίστων τετραγώνων για την εκτίμηση της τάσης. Ισχύει η παρακάτω σχέση:

$$\hat{Y}_t = \hat{a} + \hat{b}X_t \quad \text{όπου}$$

$\hat{Y}$ είναι η προβλεφθείσα τιμή της τάσης για μία χρονική περίοδο t

$\hat{a}$ είναι ο εκτιμητής της τιμής της τάσης όταν X= 0

$\hat{b}$ είναι ο εκτιμητής της κλίσης της εξίσωσης και

X_t είναι η παρατήρηση για κάθε περίοδο t.

2.6.2 Εκθετικό Μοντέλο

Εάν παρατηρηθεί γραμμική τάση τότε αυτή προσαρμόζεται καλύτερα με μια εκθετική τάση. Ισχύει ο τύπος :

$Y_t = (a\,b^{X_t})$. Εάν εφαρμόσουμε τη μέθοδο των ελαχίστων τετραγώνων η παραπάνω εξίσωση λογαριθμησμένη γίνεται: $\log \hat{Y}t = \log a + Xt \log b$ και υπολογίζουμε:

$$\log a = \frac{\Sigma \log \Upsilon}{n}$$

$$\log b = \frac{\Sigma(X \log \Upsilon)}{\Sigma X^2}.$$

2.6.3 Μέθοδος των κινητών μέσων όρων

Για την απεικόνιση της τάσης όμως επικρατέστερη μέθοδος φαίνεται να είναι αυτή τω κινητών μέσων όρων. Και αυτό γιατί , αφού η τάση είναι η μακροχρόνια κεντρική κίνηση της σειράς απαλλαγμένη από τις βραχυχρόνιες αυξομειώσεις έχουμε τη δυνατότητα εξομάλυνσης της αρχικής σειράς με τον υπολογισμό του μέσου αριθμητικού k διαδοχικών όρων. Αν λοιπόν το μήκος του κινητού μέσου είναι (κ=3), τότε οι όροι του κινητού μέσου n τιμών σειράς Y είναι:

$(Y_1+Y_2+Y_3)/3$, $(Y_2+Y_3+Y_4)/3$, …., $(Y_{n-2}+Y_{n-1}+Y_n)/3$. Στην περίπτωση που το k είναι μικρό η σειρά των κινητών μέσων ακολουθεί περισσότερο τις κινήσεις της αρχικής χρονοσειράς, ενώ αν το k έχει μεγάλες τιμές έχουμε μεγαλύτερο εξομαλυντικό αποτέλεσμα και η σειρά των κινητών μέσων είναι πιο ομαλή.

Ωστόσο και αυτή η μέθοδος εμφανίζει μειονεκτήματα. Καταρχήν δεν μπορεί να χρησιμοποιηθεί σε προβλέψεις αφού δεν χρησιμοποιεί κάποιο συγκεκριμένο υπόδειγμα, λόγω του οποίου ο κινητός μέσος χρησιμοποιείται μόνο για την εξομάλυνση των χρονοσειρών. Επίσης η επιλογή του μήκους του είναι υποκειμενική και γίνεται βάση της εμπειρίας του εκτιμητή. Το βασικότερο μειονέκτημα της μεθόδου όμως είναι πως χάνονται κάποιες τιμές στην αρχή και στο τέλος της χρονοσειράς.

2.7 Εκτίμηση του κύκλου

Αφού εκτιμήσουμε τις τιμές της τάσης το επόμενο βήμα είναι να απομονώσουμε σταδιακά τις συνιστώσες της χρονολογικής σειράς. Ο τρόπος που θα γίνει αυτό εξαρτάται από το υπόδειγμα που θα επιλέξει ο αναλυτής. Αν λοιπόν το υπόδειγμα είναι προσθετικό , τότε η αφαίρεση των τιμών της τάσης από τις τιμές της αρχικής σειράς θα δώσουν μια νέα σειρά τιμών που θα περιλαμβάνουν μόνο τον κύκλο και τις τυχαίες κυμάνσεις. Επομένως από τη σχέση $Y_t = T_t + C_t + I_t$ προκύπτει;

$$Y_t - T_t = C_t + I_t \ .$$

Αντίστοιχα στο πολλαπλασιαστικό υπόδειγμα διαιρούμε τις τιμές της αρχικής σειράς Y_t με τις τιμές της τάσης Y_t. Δηλαδή

$$Y_t = T_t \cdot C_t \cdot I_t \quad \text{ή} \quad \frac{Y_t}{T_t} = C_t \cdot I_t$$

2.8 Εποχικότητα

Μια χρονολογική σειρά ({Xt }, t= 0, ± 1 , ± 2..) η οποία εκφράζεται σε μηνιαία η τριμηνιαία στοιχεία, είναι δυνατό να παρουσιάζει εποχικές κυμάνσεις, δηλαδή κυμάνσεις με περίοδο ίση προς ένα έτος. Το φαινόμενο αυτό ονομάζεται εποχικότητα και η βασική αιτία εμφάνισής του είναι οι διάφορες κλιματολογικές συνθήκες, που διαφοροποιούνται κατά τη διάρκεια ενός έτους και που επαναλαμβάνονται με παρόμοιο τρόπο κάθε έτος. Επίσης και κάποιες χρονοσειρές που αφορούν στοιχεία παραγωγής πωλήσεων, τουριστικής δραστηριότητας κ.λ.π εμφανίζουν το φαινόμενο της εποχικότητας.

Η γνώση των εποχικών προτύπων είναι πολύ σημαντικός παράγοντας για τον επιχειρηματικό σχεδιασμό μιας εταιρείας, ενός καταστήματος κ.λ.π. Για παράδειγμα οι ιδιοκτήτες καταστημάτων παιδικών παιχνιδιών θα πρέπει να έχουν μεγάλα αποθέματα, τις ημέρες των Χριστουγέννων αφού οι πωλήσεις θα είναι ιδιαίτερα αυξημένες. Εκτός από τη δημιουργία του ετήσιου επιχειρηματικού προγράμματος, η γνώση των εποχικών κινήσεων μας δίνει τη δυνατότητα να αναλύουμε τις επιδόσεις του παρελθόντος να αλλάζουμε το διαχρονικό πρότυπο παραγωγής, ακόμη και να μεταβάλλουμε το ίδιο το εποχικό πρότυπο.

2.8.1 Εκτίμηση της Εποχικότητας

Η εποχικότητα περιέχεται σε χρονολογικές σειρές που οι τιμές τους αναφέρονται σε περιόδους μικρότερες του έτους (μηνιαία, τριμηνιαία κ.λ.π. δεδομένα). Η εκτίμηση των δεικτών της εποχικότητας γίνεται ως εξής: Θεωρούμε μια τριμηνιαία σειρά Υ. Σύμφωνα με το προσθετικό υπόδειγμα έχουμε:

$$Y_t = T_t + C_t + S_t + I_t \Leftrightarrow Y_t - T_t - C_t = S_t + I_t$$

ενώ για το πολλαπλασιαστικό υπόδειγμα:

$$Y_t = T_t \cdot C_t \cdot S_t \cdot I_t \Rightarrow \frac{Y_t}{T_t \cdot C_t} = S_t \cdot I_t$$

Σύμφωνα λοιπόν με τις παραπάνω σχέσεις, για να εκτιμήσουμε τους δείκτες της εποχικότητας πρώτα αφαιρούμε την επίδραση της τάσης και του κύκλου είτε με αφαίρεση είτε με διαίρεση, ανάλογα με το υπόδειγμα που ακολουθούμε. Από τον όρο ($S_t + I_t$) ή ($S_t \cdot I_t$) εκτιμούμε την κατά μέσο όρο επίδραση κάθε εποχής. Αυτές οι επιδράσεις είναι οι δείκτες της εποχικότητας. Στο επόμενο κεφάλαιο θα αναφερθεί ο τρόπος εκτίμησης των εποχικών δεικτών .

Είδαμε λοιπόν πως η εποχικότητα μετριέται από έναν δείκτη γνωστό ως εποχικό δείκτη, ο οποίος αντανακλά την τάση που έχουν οι τιμές της χρονοσειράς να είναι υψηλότερες ή χαμηλότερες από τις αντίστοιχες μεταβολές μιας άλλης υποπεριόδου του ίδιου μήκους κατά της διάρκεια της περιόδου παρατήρησης.
(Jeffrey Jarret, 2002)

Αναλυτικότερα οι εποχικοί δείκτες προσδιορίζονται με τη μέθοδο του κινητού μέσου .Βασικός στόχος είναι η εξάλειψη των άλλων συνιστωσών με σκοπό να μείνουν μόνο οι εποχικές διακυμάνσεις. Υπάρχει βέβαια η παραδοχή πως το εποχικό πρότυπο είναι απόλυτα ακριβές κατά τα διαδοχικά έτη, όμως υπάρχει μέσα σε λογικά όρια ο κινητός μέσος όρος μέσω του οποίου, ελαχιστοποιείται η επίδραση σε ένα τουλάχιστο βαθμό.

2.9 Σφάλματα στις προβλέψεις

Όλες οι προβλέψεις ενέχουν κάποιο επίπεδο αβεβαιότητας εξαιτίας της οποίας είναι αναγκαίο να κάνουμε υποθέσεις ή κρίσεις γύρω από τις σχετικές κινήσεις στο μέλλον και, επομένως, πρέπει να αναμένεται κάποιο σφάλμα στη πρόβλεψη. Έτσι λοιπόν η τυχαία συνιστώσα πρέπει να περιλαμβάνεται σε όλες τις προβλέψεις, αφού επηρεάζει σημαντικά την προβλεπτική ικανότητα. Αν λοιπόν το άρρυθμο στοιχείο όπως αλλιώς ονομάζεται αποτελεί ένα μεγάλο ποσοστό της συμπεριφοράς των δεδομένων τότε στην ουσία οι τιμές της μεταβλητής είναι σχεδόν αδύνατο να προβλεφθούν. Αντίθετα, αν είναι μικρό και είναι δυνατός ο προσδιορισμός των αποτελεσμάτων της τάσης και των εποχικών και κυκλικών σχημάτων, τότε θα είναι δυνατή μια ικανοποιητικά ακριβής πρόβλεψη (Ι.Πανάρετου & Ε. Ξεκαλάκη 1998).

Σφάλμα πρόβλεψης λοιπόν είναι η διαφορά μεταξύ πρόβλεψης και πραγματικής τιμής για μια δεδομένη περίοδο:

$$e_t = X_t - F_t \text{ όπου}$$

e_t το σφάλμα για την περίοδο t,

X_t η πραγματική τιμή για την περίοδο t,

F_t η πρόβλεψη για την περίοδο t.

2.9.1 MAD (mean absolute deviation) και MSE (mean squared error)

Για να έχουμε μια πληρέστερη εικόνα του σφάλματος σε βάθος χρόνου χρησιμοποιούμε τα μεγέθη της Μέσης Απόλυτης Απόκλισης (mean absolute deviation, MAD) και Μέσου Τετραγωνισμένου Σφάλματος (mean squared error, MSE). Όσο μικρότερη είναι η τιμή των μεγεθών αυτών, τόσο μεγαλύτερη η ακρίβεια .Για το MAD ισχύει:

$$MAD = \frac{\Sigma|e|}{n}$$ ενώ για το MSE:

$$MSE = \frac{\Sigma(e_t)^2}{n} = \frac{\Sigma(X_t - F_t)^2}{n}$$

Με τα μεγέθη αυτά μπορούμε να ελέγξουμε την ακρίβεια των μεθόδων πρόβλεψης και να αξιολογήσουμε τις μεθόδους εξομάλυνσης και να επιλέξουμε τη καταλληλότερη.

2.10 Μέθοδοι εξομάλυνσης

Ως μεθόδους εξομάλυνσης ονομάζουμε ένα σύνολο μεθόδων πρόβλεψης που χρησιμοποιείται έτσι ώστε να έχουμε ικανοποιητικές προβλέψεις βραχυπρόθεσμα αφού κάποιες φορές υπάρχουν δυσκολίες στην ανάπτυξη και την εφαρμογή κάποιων μεθόδων πρόβλεψης. Οι μέθοδοι εξομάλυνσης μιας χρονοσειράς αποσκοπούν στην απομάκρυνση της τυχαίας συνιστώσας και υποβαθμίζουν τόσο την τάση, όσο και την εποχική και την κυκλική συνιστώσα (Ανδρικόπουλος Ανδρέας, 1998) και βασίζονται στη διαδικασία καθορισμού του μέσου όρου παρατηρήσεων μιας χρονοσειράς που διαμορφώθηκαν στο παρελθόν. Οι επικρατέστερες μέθοδοι εξομάλυνσης είναι:

- Η Μέθοδος Κινητών Μέσων (Moving Average Method)
- Η Μέθοδος Εκθετικής Εξομάλυνσης (Exponential Smoothing method)

Πρώτα όμως θα πρέπει να αναφερθούμε στη μέθοδο του αριθμητικού μέσου ο οποίος είναι ένας εκτιμητής που ελαχιστοποιεί το λάθος εκτίμησης της τιμής του μέσου του πληθυσμού. Αν θέλουμε να χρησιμοποιήσουμε τη μέθοδο αυτή θα αυτό πρέπει να γίνει με την προϋπόθεση, ότι τα δεδομένα κατανέμονται τυχαία και ακολουθούν σταθερή εξέλιξη ως προς το χρόνο.

2.10.1 Απλός Μέσος Όρος

Όταν υποπτευόμαστε ότι η μεταβλητή που θέλουμε να προβλέψουμε παρουσιάζει επίπεδο μοτίβο, τότε η μέθοδος αυτή μπορεί να μας δώσει μια καλή εκτίμηση της μεταβλητής.
Η πρόβλεψη γίνεται με τον υπολογισμό της μέσης τιμής των δεδομένων:

$$F_{y+1} = \frac{\Sigma yi}{n} \quad (1)$$

Παράδειγμα 1
Ένα λατομείο θέλει να προβλέψει τις πωλήσεις του για το επόμενο δίμηνο έχοντας στη διάθεσή του τις πραγματοποιηθείσες (δηλ, τις πραγματικές) πωλήσεις για τα προηγούμενα πέντε δίμηνα.

Δίμηνο		Πραγματικές πωλήσεις (σε χιλιάδες τόνους)	Πρόβλεψη για επόμενο δίμηνο (σε χιλιάδες τόνους)
1°	Γενάρης-Φλεβάρης 2004	51	
2°	Μάρτης-Απρίλης 2004	53	
3°	Μάης-Ιούνιος 2004	48	
4°	Ιούλιος-Αύγουστος 2004	52	
5°	Σεπτέμβρης-Οκτώβρης 2004	50	
6°	Νοέμβρης-Δεκέμβρης 2004	;	50,8

(Πηγή: Μιχάλης Βαϊδάνης 2005)

Η λύση, με βάση τη σχέση 1, έχει ως εξής:

$F_6^{ου} = (A_1^{ου}+A_2^{ου}+A_3^{ου}+A_4^{ου}+A_5^{ου})/5 = (51+53+48+52+50)/5 = 254/5$

$F_6^{ου}$= 50,8 χιλιάδες τόνοι

Όταν παρέλθει το 6° δίμηνο και μάθουμε τις πραγματικές πωλήσεις γι' αυτό (έστω ότι ήταν 49 χιλ. τόνοι), επαναλαμβάνουμε την ίδια διαδικασία για την πρόβλεψη του 7^{ου} δίμηνου αυτή τη φορά με n=6:

$F_7^{ου} = (A_1^{ου}+A_2^{ου}+A_3^{ου}+A_4^{ου}+A_5^{ου}+A_6^{ου})/6 = (51+53+48+52+50+49)/6 = 303/6$

$F_7^{ου}$= 50,5 χιλιάδες τόνοι

2.10.2 Μέθοδος Κινητών Μέσων

Η πρόβλεψη χρονοσειρών με τη μέθοδο των κινητών μέσων εμπεριέχει τον υπολογισμό του μέσου όρου του δείγματος των παρατηρήσεων, καθώς και τη χρησιμοποίησή αυτού του μέσου σαν πρόβλεψη για την επόμενη περίοδο. Στην αρχή της διαδικασίας πρόβλεψης προσδιορίζεται ο αριθμός των παρατηρήσεων του δείγματος που συμπεριλαμβάνονται στον υπολογισμό του μέσου. Ο όρος κινητός μέσος χρησιμοποιείται, λόγω του υπολογισμού ενός νέου μέσου όρου από την εισαγωγή κάθε νέας παρατήρησης στο δείγμα και απόρριψης της παλαιότερης.

Επομένως κάθε πρόβλεψη χρησιμοποιεί τον ίδιο αριθμό παρατηρήσεων δείγματος από τη χρονοσειρά περιλαμβάνοντας μόνο την πιο πρόσφατη παρατήρηση. Η διαδικασία εξομάλυνσης των χρονοσειρών με την παραπάνω μέθοδο περιλαμβάνει τα εξής στάδια:

1. Υπολογίζουμε τα κινητά σύνολα, με τη χρήση των παρατηρήσεων της χρονοσειράς και βασιζόμενοι στην περίοδο την οποία επιθυμούμε τον υπολογισμό των κινητών μέσων.
2. Διαιρώντας τα κινητά σύνολα με τον αριθμό των παρατηρήσεων, σε κάθε υποσύνολο, το πηλίκο που προκύπτει μας δίνει τους κινητούς μέσους. Όταν έχουμε χρονοσειρές με περιττό αριθμό παρατηρήσεων δεν υπολογίζουμε κινητούς μέσους για τις πρώτες και τις τελευταίες (κ-1)/2 χρονικές περιόδους.
   κ = ο αριθμός των παρατηρήσεων για κάθε κινητό σύνολο

3. Υπολογίζουμε τη διαφορά ανάμεσα στις τιμές της χρονοσειράς και των κινητών μέσων. Μέσω αυτής βρίσκουμε τις αποκλίσεις της χρονοσειράς γύρω από τους κινητούς μέσους.

2.10.3 Μέθοδος της Εκθετικής Εξομάλυνσης

Η εκθετική εξομάλυνση αποτελεί μια μέθοδο πρόβλεψης κατά την οποία οι προβλέψεις υπολογίζονται μετά την εξομάλυνση των δεδομένων, με στόχο να απομονωθούν τα πραγματικά πρότυπα από τις τυχαίες διακυμάνσεις. Βασική αρχή της μεθόδου είναι το ότι όσο πιο πρόσφατα είναι τα δεδομένα τόσο περισσότερες πληροφορίες περιέχουν. Με την εκθετική εξομάλυνση μπορούμε να διενεργήσουμε μια απλή και έγκαιρη πρόβλεψη η οποία επιτρέπει στους ερευνητές να διορθώσουν προγενέστερες ανακρίβειες στις προβλέψεις. Η Εκθετική Εξομάλυνση περιλαμβάνει πολλές μεθόδους οι οποίες χρησιμοποιούνται συνήθως για την πρόβλεψη μεγάλου αριθμού μεγεθών όπως στην περίπτωση προβλέψεων κέρδους, έλεγχο πωλήσεων κ.λ.π.

2.11 Συμπέρασμα

Στο κεφάλαιο αυτό λοιπόν παρατηρήσαμε πως η μελλοντική πορεία μιας μεταβλητής έχει άμεση εξάρτηση με τη συμπεριφορά της στο παρελθόν. Έτσι μπορεί να γίνει μια απόπειρα πρόβλεψης με κύριο μέλημα του εκάστοτε αναλυτή το όσο δυνατό μικρότερο σφάλμα, με τη βοήθεια της ανάλυσης των χρονοσειρών στις συνιστώσες της. Απαραίτητα στοιχεία στις έρευνες του κάθε αναλυτή είναι και ο προσδιορισμός των MAD και MSE. Καταλήγοντας, κάνοντας λόγο για το βασικό ρόλο στην προσπάθεια δημιουργίας ενός αξιόπιστου προτύπου για τη διενέργεια σωστών προβλέψεων έγινε αναφορά στις μεθόδους εξομάλυνσης. Οι πιο βασικές είναι του κινητού μέσου και της εκθετικής εξομάλυνσης.

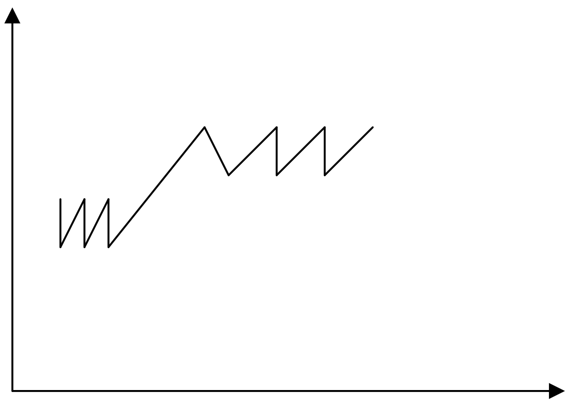
ΚΕΦΑΛΑΙΟ 3: ΜΑΘΗΜΑΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ ΣΤΑ ΔΙΑΦΟΡΑ ΜΟΝΤΕΛΑ ΑΝΑΛΥΣΗΣ ΧΡΟΝΟΣΕΙΡΩΝ.

3.1 Εισαγωγή

Παρακάτω θα παρουσιαστούν μερικά στοχαστικά μαθηματικά μοντέλα με τα οποία θα προσπαθήσουμε να περιγράψουμε τη διαχρονική εξέλιξη κάποιου φυσικού μεγέθους. Έτσι λοιπόν χρησιμοποιώντας διάφορα μοντέλα με τα οποία ασχολήθηκαν κατά καιρούς διάφοροι μελετητές, γνωστά και ως αλγεβρικά θα γίνει μια προσπάθεια εξεύρεσης του καταλληλότερου, για την ανάλυση κάθε μορφής χρονοσειράς.

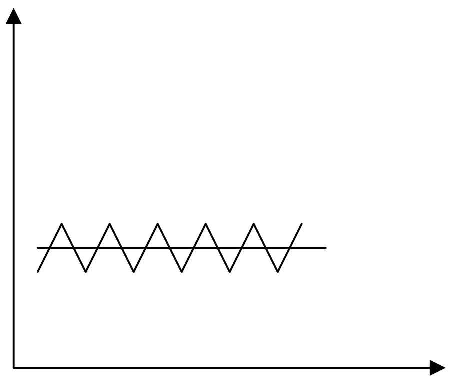
3.2 Στάσιμες και μη στάσιμες χρονοσειρές

Οι χρονοσειρές μπορεί να είναι στάσιμες ή μη στάσιμες. **Μη στάσιμη** θεωρείται μια χρονοσειρά όταν τα χαρακτηριστικά της μεταβάλλονται διαχρονικά και δεν μπορεί να περιγραφεί από κάποιο αλγεβρικό μοντέλο.



Σχήμα 3.1

Αντίθετα, μια χρονοσειρά ονομάζεται **στάσιμη** όταν τα χαρακτηριστικά της παραμένουν διαδοχικά σε ισορροπία γύρω από ένα σταθερό μέσο επίπεδο, ή πιο απλά όταν οι στατιστικές ιδιότητες της χρονοσειράς δεν αλλάζουν με το χρόνο. Δηλαδή τα δεδομένα κυμαίνονται από ένα σταθερό μέσο ανεξάρτητα του χρόνου και η διακύμανση παραμένει σταθερή.



Σχήμα 3.2

Αν θέλουμε να ελέγξουμε τη στασιμότητα ή μη μιας χρονοσειράς παρατηρούμε τη γραφική της παράσταση και:

- εάν δεν παρατηρείται αλλαγή της μέσης τιμής κατά μήκος του χρόνου τότε αυτή, είναι στάσιμη ως προς τη μέση τιμή
- εάν δεν παρατηρείται αλλαγή της διακύμανσης κατά μήκος του χρόνου τότε αυτή, είναι στάσιμη ως προς τη διακύμανση
- εάν παρατηρείται κάποια αλλαγή κατά μήκος του χρόνου τότε αυτή, δεν είναι στάσιμη. (Πτυχιακή:Βασικές μέθοδοι ανάλυσης χρονοσειρών, 2005)

3.3 Αυτοσυσχέτιση

Η συσχέτιση χαρακτηρίζει την ύπαρξη εξάρτησης μεταξύ των μεταβλητών. Αν και μπορεί να υπάρχει ένας μεγάλος αριθμός αιτών για τη ύπαρξη εξάρτησης, ο συντελεστής συσχέτισης δεν δίνει καμία πληροφορία σχετικά με αυτές τις αιτίες, και για το μόνο που μας πληροφορεί είναι, αν δύο ή περισσότερες μεταβλητές συμμεταβάλλονται θετικά ή αρνητικά.

Η διαγραμματική απεικόνιση των τιμών μιας χρονοσειράς μας δίνει το διάγραμμα της αυτοσυσχέτισης και βοηθά στο να διαπιστωθεί αν αυτή είναι στάσιμη ή όχι. Οι συντελεστές αυτοσυσχέτισης για μια στάσιμη σειρά φθίνουν σχετικά γρήγορα προς το μηδέν ενώ δεν συμβαίνει πάντα το ίδιο στις μη στάσιμες. Βέβαια τις μη στάσιμες μπορούν να μετατραπούν σε στάσιμες παίρνοντας π.χ πρώτες διαφορές, δεύτερες κ.λ.π. και ελέγχοντας τη συνάρτηση αυτοσυσχέτισης μέχρι να επιτευχθεί στασιμότητα. Γενικά με τη βοήθεια της συνάρτησης αυτοσυσχέτισης μετριέται ο βαθμός συσχέτισης ανάμεσα στις γειτονικές παρατηρήσεις των χρονοσειρών. Ορίζοντας την αυτοσυσχέτιση με υστέρηση k ισχύει:

$$r_k = \frac{\operatorname{cov}(y_t, y_{t-k})}{s_{y_t} \cdot s_{y_{t-k}}}$$

όπου,

cov η συνδιακύμανση μεταξύ y_t και y_{t-k} και s η τυπική απόκλιση της y_t και y_{t-k} σε δυο διαφορετικές χρονικές στιγμές.

3.4 <u>Μερική αυτοσυσχέτιση</u>

Μια άλλη συνάρτηση που χρησιμοποιείται στη μελέτη των χαρακτηριστικών μιας χρονολογικής σειράς είναι συνάρτηση μερικής αυτοσυσχέτισης. Γενικά οι συντελεστές μερικής αυτοσυσχέτισης μετρούν το βαθμό της σχέσης μεταξύ των Y_t και Y_{t-k}, όταν το αποτέλεσμα των άλλων χρονικών υστερήσεων στο Y_t, παραμένει σταθερό. Η συσχέτιση αυτών των μεταβλητών δεν επηρεάζεται από τις ενδιάμεσες μεταβλητές και ο συντελεστής τους συμβολίζεται με f_{kk} (k τάξεως).(Jeffrey Jarrett. 2002).

Οι μερικές αυτοσυσχετίσεις ορίζονται από την παρακάτω εξίσωση παλινδρόμησης:

$$y_t = f_{1k} y_{t-1} + f_{2k} y_{t-2} + \ldots + f_{kk} y_{t-k} + e_t$$

όπου e_t το σφάλμα παλινδρόμησης.

Γενικά, κάθε χρονοσειρά μας ενδιαφέρει να προσδιορίσουμε τα εάν:

- είναι τα δεδομένα τυχαία
- τα δεδομένα είναι σταθερά ή μη σταθερά, σε ποιο επίπεδο είναι μεταβλητά στο χρόνο
- τα δεδομένα είναι εποχικά, κι αν όντως είναι, ποια είναι η έκταση της εποχικότητας.

Σημείωση: Το διάγραμμα που αφορά τον συντελεστή αυτοσυσχέτισης λέγεται ACF ενώ αντίστοιχα του συντελεστή μερικής αυτοσυσχέτισης PACF.

3.5 Μετατροπή μη-στάσιμης χρονοσειράς σε στάσιμη

Ένα αλγεβρικό μοντέλο δεν μπορεί να περιγράψει μια μη-στάσιμη χρονοσειρά. Αρχικά θα πρέπει να γίνεται έλεγχος για το αν αυτή είναι στάσιμη. Με στόχο την ανάπτυξη του κατάλληλου μοντέλου θα πρέπει να εξαλειφθούν τα μη στάσιμα πρότυπα. Αυτό μπορεί να γίνει μέσω της μεθόδου της διαφόρισης .

Οι χρονοσειρές «διαφορών πρώτης τάξης» προκύπτουν από τις διαφορές των διαδοχικών παρατηρήσεων της αρχικής χρονοσειράς:

$$\nabla Y'_t = Y_t - Y_{t-1},$$

όπου το ∇ δείχνει τη διαφορά μεταξύ του Υ στην περίοδο t και στην προηγούμενη περίοδο $(t-1)$. Οι χρονοσειρές διαφορών πρώτης τάξης έχουν $n-1$ δεδομένα.

Οι χρονοσειρές «διαφορών δεύτερης τάξης» προκύπτουν από τις διαφορές των πρώτων διαφορών και έχουν $n-2$ δεδομένα :

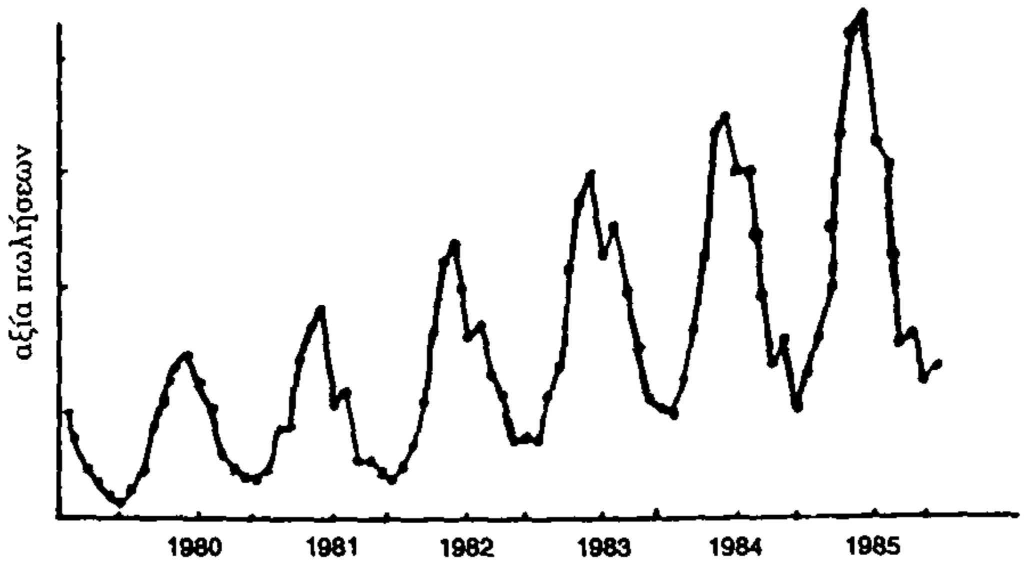
$$\nabla^2 Y''_t = Y'_t - Y'_{t-1} = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2},$$

όπου το ∇^2 δείχνει τη διαφορά μεταξύ του Y στην περίοδο $(t-1)$ και στην προηγούμενη περίοδο $(t-2)$.

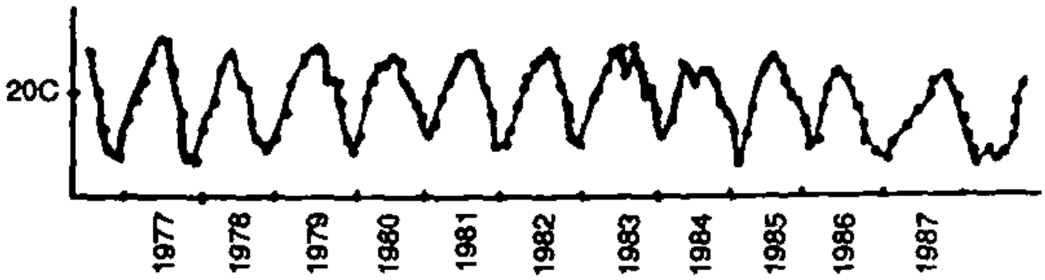
Με τον τρόπο αυτό μετατρέπεται μια χρονοσειρά από μη-στάσιμη σε στάσιμη διαφορίζοντάς την όσες φορές χρειάζεται. (Πτυχιακή: Βασικές μέθοδοι ανάλυσης χρονοσειρών, 2005)

3.6 Πηγές μεταβλητικότητας μιας χρονοσειράς

Μια σειρά παρατηρήσεων είναι κατά κανόνα ασταθής. Αυτές οι πηγές αστάθειας είναι συγχρόνως και πηγές μεταβλητικότητας μιας σειράς. Πολλές χρονοσειρές που αναφέρονται π.χ σε πωλήσεις, μετεωρολογικές παρατηρήσεις κ.λ.π. παρουσιάζουν μια μεταβλητικότητα η οποία είναι ετήσια στην περίοδο (βλέπε σχήματα 3.3, 3.4). Ο τύπος αυτός της μεταβλητικότητας που χαρακτηρίζεται ως εποχικότητα μπορεί να μετρηθεί ή ακόμη να απαλειφθεί δίνοντας νέα δεδομένα απαλλαγμένα από εποχικές κυμάνσεις. Είναι ευνόητο ότι η περίοδος δεν είναι ανάγκη να είναι δωδεκάμηνη. Πράγματι ένα παράδειγμα είναι η ημερήσια μεταβολή της θερμοκρασίας μιας “τυπικής” ημέρας, ο αριθμός των τηλεφωνικών κλήσεων π.χ σε δυο ώρες ένα τηλεφωνικό κέντρο κ.λ.π.



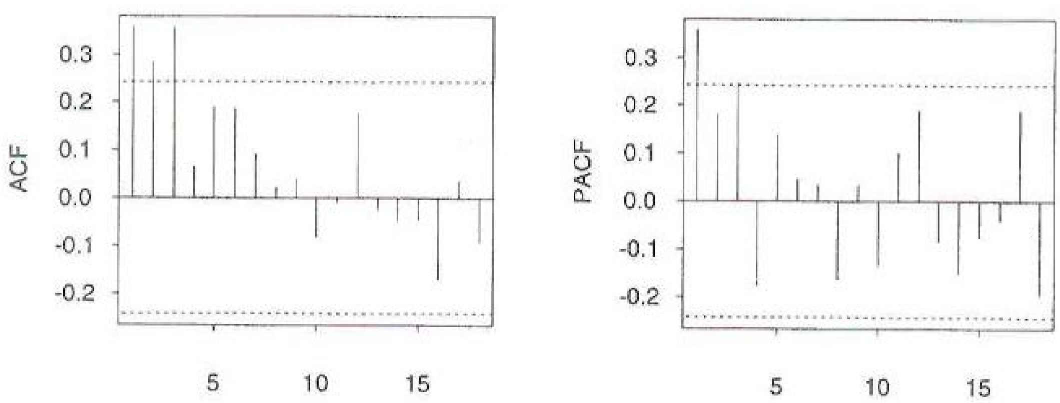
Σχήμα 3.3 Πωλήσεις μιας εταιρίας σε διαδοχικούς μήνες
(Πηγή:Όθωνας Παπαδήμας,1994)



Σχήμα 3.4 Μέση θερμοκρασία αέρος μιας πόλης σε διαδοχικούς μήνες (Πηγή: Όθωνας Παπαδήμας,1994)

3.7 <u>Αναγνώριση Εποχικότητας</u>

Όταν μια χρονοσειρά επαναλαμβάνεται κατά τη διάρκεια μιας περιόδου συνήθως ετήσιας θεωρείται ότι έχει εποχικότητα. Αν εξεταστούν οι συντελεστές αυτοσυσχέτισης, μπορεί να αναγνωριστεί ένα εποχικό μοντέλο. Έπειτα μετράμε περισσότερες από δυο ή τρεις χρονικές υστερήσεις ώστε να ορίσουμε εάν είναι στατιστικά διάφορες του μηδενός. Η ύπαρξη ενός προτύπου φανερώνεται από τις αυτοσυσχετίσεις που είναι σημαντικά διάφορες του μηδενός. Με σκοπό την εύρεση της εποχικότητας οφείλουμε να αναζητήσουμε σημαντικές αυτοσυσχετίσεις υψηλότερων τάξεων. Έτσι, για μηνιαία δεδομένα που παρουσιάζουν εποχικότητα, υψηλές αυτοσυσχετίσεις μπορούν να παρατηρηθούν για υστέρηση 12 αλλά και για καθυστερήσεις 24 και 36. Όταν τα δεδομένα είναι τριμηνιαία και παρουσιάζουν εποχικότητα, υψηλές αυτοσυσχετίσεις θα παρουσιαστούν για υστέρηση 4, 8, 12.



(ΠΗΓΗ:http://fsu.ece.ntua.gr/old/gr/texnikes_2004.html)

Παρατηρώντας τις παραπάνω γραφικές παραστάσεις, φαίνονται υψηλές αυτοσυσχετίσεις για υστερήσεις 1, 2 και 3 στην ACF και σημαντικές μερικές αυτοσυσχετίσεις για υστερήσεις 1 και 3 στην PACF. Επιπλέον φαίνεται ότι για καθυστέρηση 12, η αυτοσυσχέτιση είναι σχετικά υψηλή παρά το γεγονός ότι δεν είναι στατιστικά σημαντική, κάτι που ίσως οφείλεται στην ύπαρξη κάποιας εποχικότητας στη χρονοσειρά.

Στην περίπτωση που τα δεδομένα δεν ακολουθούν ένα πρότυπο π.χ τάση, που να δημιουργεί μη στάσιμα δεδομένα είναι τις περισσότερες φορές εύκολο να αναγνωριστεί η εποχικότητα. Βέβαια αφού υπάρχει πρότυπο, πρέπει πρώτα να γίνει μετατροπή των μη-στάσιμων δεδομένων σε στάσιμα και στην συνέχεια να εξετασθούν για την παρουσία εποχικότητας.

Όταν τα εποχιακά δεδομένα είναι μη στάσιμα, για να προκύψει η στάσιμη χρονοσειρά χρησιμοποιούνται οι **εποχιακές διαφορές**. Ως εποχιακή διαφορά ορίζεται η διαφορά μιας παρατήρησης και της αντίστοιχης παρατήρησης του προηγούμενου έτους. (Πτυχιακή: Βασικές μέθοδοι ανάλυσης χρονοσειρών, 2005)

Έτσι για μηνιαία δεδομένα που παρουσιάζουν εποχικότητα, η εποχική διαφορά ορίζεται :

$$Y_t' = Y_t - Y_{t-12} = Y_t - B^{12}Y_t = (1 - B^{12})Y_t$$

Γενικά, οι σειρές εποχικών διαφορών προκύπτουν από τις διαφορές μεταξύ δεδομένων που απέχουν μεταξύ τους s χρονικές περιόδους, όπου s η εποχικότητα. . Όμοια με τις διαφορές πρώτης τάξης, η διαφόριση μπορεί να επαναληφθεί ώστε να προκύψουν δεύτερης τάξης εποχιακές διαφορές, αν και κάτι τέτοιο δεν χρειάζεται συχνά.

3.8 Μοντέλα χρονοσειρών ARIMA

Υπάρχουν διάφορα μοντέλα πρόβλεψης από την αξιολόγηση των οποίων μπορεί κάποιος να επιλέξει το καταλληλότερο κατά την κρίση του με στόχο την πρόβλεψη μιας χρονοσειράς. Αυτά που χρησιμοποιούνται περισσότερο είναι τα μοντέλα ARIMA αφού φαίνεται πως τα αποτελέσματα τα οποία δίνουν στην πρόβλεψη των μελλοντικών τιμών ενός μεγέθους, να είναι ικανοποιητικά.

Τα μοντέλα πολλαπλής γραμμικής παλινδρόμησης περιγράφονται από την εξίσωση:

$MA:q$

όπου Y_t η εξαρτημένη μεταβλητή (υπό πρόβλεψη) και $X_1, X_2, \ldots X_p$ οι ανεξάρτητες μεταβλητές (επεξηγηματικές).

Τα μη εποχιακά ARIMA μοντέλα είναι διαδεδομένα ως ARIMA (p,d,q) όπου: $AR:p$ = τάξη του AR όρου, $I:d$ = η τάξη της διαφόρισης και $MA:q$ = η τάξη του MA όρου.

3.9 Αυτοπαλινδρομικά μοντέλα (AR)

Η εξίσωση $Y_t = b_0 + b_1 Y_{t-1} + b_2 Y_{t-2} + \ldots + b_p Y_{t-p} + e_t$ είναι μια εξίσωση πολλαπλής γραμμικής παλινδρόμησης, με τη διαφορά ότι οι ανεξάρτητες μεταβλητές είναι παλαιότερες τιμές της υπό πρόβλεψη μεταβλητής ή οι τιμές της μεταβλητής για χρονικές υστερήσεις 1,2, ... , *p*.

Το μοντέλο αυτό καλείται Αυτοπαλινδρομικό (Autoregression ή AR) τάξεως *p*. Τα συγκεκριμένα μοντέλα πρέπει να έχουν διαφορετικό χειρισμό από τα κλασσικά παλινδρομικά μοντέλα για τους εξής λόγους :

- Η βασική υπόθεση της ανεξαρτησίας των σφαλμάτων (υπολοίπων) μπορεί εύκολα να παραβιαστεί στα αυτοπαλινδρομικά μοντέλα, αφού οι επεξηγηματικές μεταβλητές έχουν συνήθως μια εξαρτημένη σχέση καθώς περιγράφουν την εξέλιξη του ίδιου μεγέθους.

- Ο τρόπος προσδιορισμού του πλήθους των προηγούμενων τιμών της υπό πρόβλεψη μεταβλητής δεν είναι πάντοτε "ευθύς".
  (Πτυχιακή:Βασικές μέθοδοι ανάλυσης χρονοσειρών, 2005)

3.10 Μοντέλα κινητών μέσων όρων, (ΜΑ)

Τα υποδείγματα AR(p) δεν μπορούν να απομονώσουν ορισμένα πρότυπα δεδομένων, όταν το p είναι πολύ μικρό. Ένα άλλο λοιπόν μοντέλο, αυτό των κινητών μέσων (ΜΑ) ίσως απομονώσει το πρότυπο στις περιπτώσεις που τα AR(p) δεν μπορούν. Η γενική μορφή ενός υποδείγματος κινητών μέσων MA(q) έχει τη μορφή:

$$Y_t = e_t - q_1 e_{t-1} - q_2 e_{t-2} - \ldots - q_q e_{t-q}$$

(όπου τα θ είναι σταθεροί παράμετροι και et είναι λευκός θόρυβος) ή
με τη χρήση του τελεστή ολίσθησης B το μοντέλο μπορεί να γραφεί ως:

$$Y_t = \left(1 - q_1 \mathrm{B} - \ldots - q_q \mathrm{B}^q\right) e_t \,.$$

Η παρατήρηση Y_t εξαρτάται από τον όρο σφάλματος e_t και το προηγούμενο σφάλμα e_{t-q}. Τα σφάλματα e_t ακολουθούν την κανονική κατανομή με μέση τιμή 0 και διακύμανση 1. Μια διαδικασία κινητού μέσου ΜΑ είναι πάντα στάσιμη (Παπαδήμας 'Οθωνας, 1994).

Στο μοντέλο αυτό υποθέτουμε ότι η χρονολογική σειρά Y_t δημιουργείται ως ένας σταθμικός μέσος των τυχαίων σφαλμάτων των q προηγούμενων περιόδων και ονομάζεται μοντέλο κινητών μέσων όρων (moving average) τάξεως q συμβολιζόμενο ως MA(q). Σημειώνουμε ότι ο όρος « κινητός μέσος » δεν ανταποκρίνεται στην πραγματικότητα, γιατί οι σταθμίσεις δεν έχουν άθροισμα τη μονάδα.

3.11 Μοντέλο λευκού θορύβου

Το απλούστερο δυνατό σχήμα χρονολογικής σειράς είναι αυτό της τυχαίας μεταβλητής ή αλλιώς του ονομαζόμενου **λευκού θορύβου.** Μια σειρά είναι λευκός θόρυβος αν στην ουσία δεν έχει κανένα ευκρινές σχήμα ή πρότυπο. Αν συμβολίσουμε με e_t μια τέτοια σειρά τότε θα λέμε ότι είναι λευκός θόρυβος αν έχει σταθερό μέσο(συνήθως μηδέν), σταθερή διακύμανση και οι τιμές της δεν αυτοσυσχετίζονται.

$$E(e_t) = m$$
$$Var(e_t) = s^2_e$$
$$\operatorname{cov}(e_t, e_{t+k}) = 0 \text{, αν } k \neq 0$$

Μια τέτοια σειρά είναι πάντα στάσιμη και αποτελεί θεμελιώδες μοντέλο σε πολλές τεχνικές ανάλυσης χρονοσειρών. Παράδειγμα λευκού θορύβου αποτελούν οι τυχεροί αριθμοί του ΛΟΤΤΟ.

3.12 <u>Τυχαίος περίπατος</u>

Μια διαδικασία $Y_t,\ t = 1, 2, ..., \mathrm{T}$ καλείται τυχαίος περίπατος αν:

$$Y_t = Y_{t-1} + b + e_t \ ,$$

όπου e_t είναι λευκός θόρυβος.

Η παρατήρηση Y_t εξαρτάται μόνο από την προηγούμενη παρατήρηση Y_{t-1}. Δηλαδή κατά την χρονική περίοδο t η παρατήρηση Y_t θα παριστά την θέση ενός περιπατητή και η απόφαση για το που θα πάει την επόμενη περίοδο (μπροστά ή πίσω) δεν εξαρτάται από το που βρίσκεται αυτή τη στιγμή.

Οι τύποι που ισχύουν είναι:

Μέση τιμή $\quad \mathrm{E}(y_t) = \mathrm{E}(y_0)$

Διακύμανση $\quad Var(y_t) = ts_e^{\,2}$

Συνδιακύμανση $\quad C\mathrm{ov}(y_t, y_{t-k}) = (t-k)\,s^{\,2}_e$

Η συνδιακύμανση επειδή δεν είναι στάσιμη, αφού εξαρτάται από το χρόνο. Στο μοντέλο της τυχαίας διαδρομής μπορεί να υπάρχει και μια σταθερά β.

$$\text{Άρα } Y_t = \ b + e_t$$

όπου το β μας δείχνει την μετατόπιση της σταθεράς προς τα πάνω. Το μοντέλο για μη στάσιμα δεδομένα και παρουσιάζει μεγάλες περιόδους εμφανών τάσεων οι οποίες αλλάζουν κατεύθυνση με απρόβλεπτο τρόπο όπως για παράδειγμα οι τιμές των μετοχών.

3.13 Μικτά μοντέλα ARMA(p, q)

Αν συνδυαστούν μια AR(p) και μια MA (q) διαδικασία προκύπτει, ένα μικτό μοντέλο ARMA (p, q) το οποίο έχει την ακόλουθη μορφή:

$$Y_t - f_1 Y_{t-1} - \ldots - f_p Y_{t-p} = e_t - q_1 e_{t-1} - \ldots - q_q e_{t-q}$$

ή με τη χρήση του τελεστή B το μοντέλο μπορεί να γραφεί ως:

$$(1 - f_1 B - \ldots - f_p B^p) Y_t = \left(1 - q_1 \mathrm{B} - \ldots - q_q \mathrm{B}^q\right) e_t \text{ ή } \Phi(\mathrm{B}) Y_t = \Theta(\mathrm{B}) e_t .$$

όπου παρατήρηση Y_t εξαρτάται από την προηγούμενη τιμή Y_{t-p} και από το προηγούμενο σφάλμα e_{t-q}. Χρησιμοποιώντας τις δειγματικές και τις μερικές δειγματικές αυτοσυσχετίσεις, γίνεται ο προσδιορισμός του μοντέλου το οποίο είναι στάσιμο όταν οι ρίζες του πολυωνύμου είναι έξω από τον μοναδιαίο κύκλο και αντιστρέψιμο όταν οι ρίζες του πολυωνύμου είναι έξω από τον μοναδιαίο κύκλο (Παπαδημας Οθωνας, 1994).

3.14 Μοντέλα ARIMA(p, d, q)

Οι περισσότερες χρονοσειρές που συναντώνται σε κάποιες εφαρμογές της πρόβλεψης δεν είναι σταθερές. Ένα ARIMA (p,d,q) είναι ένα μοντέλο το οποίο αν διαφοριστεί d φορές παράγει ένα μοντέλο ARMA (p,d,q). Εάν η χρονοσειρά που εξετάζουμε δεν είναι στάσιμη τότε αντί για ένα ARMA μοντέλο εφαρμόζεται ένα ARIMA (p,d,q) μοντέλο. Η τάξη του δίνεται από τα τρία γράμματα αφού η τάξη του συντελεστή αυτοπαλινδρόμησης είναι p, της διαφόρισης που χρειάζεται για την επίτευξη σταθερότητας είναι d ενώ η τάξη του μέρους του κινητού μέσου είναι q. Η γενική του μορφή είναι:

$$f(\mathrm{B}) \nabla^d Y_t = q(\mathrm{B}) e_t .$$

Η πιο απλή περίπτωση είναι το ARIMA(1,1,1) με εξίσωση:

$$(1 - f_1 B)(1 - \mathrm{B}) Y_t = (1 - q_1 \mathrm{B}) e_t$$

.

Οι συναρτήσεις αυτοσυσχέτισης (ACF) και μερικής αυτοσυσχέτισης (PACF) των μοντέλων ARIMA (p,d,q) ακολουθούν ένα πλήθος προτύπων που καθιστά αδύνατη τη θέσπιση κανόνων για την αναγνώριση του καταλληλότερου μοντέλου. Το μοντέλο ARIMA(p, d, q) είναι το πλέον γενικό από όσα εξετάστηκαν και το περισσότερο χρησιμοποιημένο.

ΠΙΝΑΚΑΣ 3.1

Μοντέλο	**Συνθήκες Σταθερότητας**	**Συνθήκες Αντιστρεψιμότητας**	**Συντελεστές Αυτοσυσχέτισης**	**Συντελεστές Μερικής Αυτοσυσχέτισης**
AR(1)	$\|q_1\| < 1$	Καμία	Φθίνει εκθετικά προς τα κάτω	Διακόπτεται μετά από μία χρονική υστέρηση
AR(2)	$\|f_1\| < 1$ $f_1 + f_2 < 1$ $f_2 - f_1 < 1$	Καμία	Φθίνει εκθετικά προς τα κάτω	Διακόπτεται μετά από δύο χρονικές υστερήσεις
MA(1)	Καμία	$\|f_1\| < 1$	Διακόπτεται μετά από μία χρονική υστέρηση	Φθίνει εκθετικά προς τα κάτω
MA(2)	Καμία	$\|q_1\| < 1$ $q_1 + q_2 < 1$ $q_2 - q_1 < 1$	Διακόπτεται μετά από δύο χρονικές υστερήσεις	Φθίνει εκθετικά προς τα κάτω
ARMA(p,q)	$\|f_1\| < 1$	$\|q_1\| < 1$	Φθίνει εκθετικά προς τα κάτω	Φθίνει εκθετικά προς τα κάτω

(Πτυχιακή:Βασικές μέθοδοι ανάλυσης χρονοσειρών, 2005)

3.15 <u>Εποχικά Μοντέλα (SARIMA)</u>

Όπως έχει προαναφερθεί υπάρχουν ορισμένες χρονοσειορές οι οποίες σε ίσα χρονικά διαστήματα παρουσιάζουν την ίδια συμπεριφορά. Όπως τα μοντέλα ARIMA(p,d,q), έτσι και τα μοντέλα που εμφανίζουν εποχικότητα s παρουσιάζουν παρόμοιες ιδιότητες. Τα μοντέλα ARIMA που έχουν επεκταθεί για να χειρίζονται την εποχικότητα συμβολίζονται ως:

$$\mathrm{ARIMA(p,d,q)} \times \left(\mathrm{P,D,Q}\right)_s ,$$

όπου s η εποχικότητα, (p,d,q): μη εποχικοί όροι

και (P, D, Q) :εποχικοί όροι

Έστω ότι έχουμε ένα μοντέλο $\mathrm{SARIMA}\left(1,1,1\right)\left(1,1,1\right)_4$ το οποίο αν χρησιμοποιήσουμε τον τελεστή ολίσθησης Β παίρνει τη μορφή:

$$\left(1 - f_1 \mathrm{B}\right)\left(1 - \Phi_1 \mathrm{B}^4\right)\left(1 - \mathrm{B}\right)\left(1 - \mathrm{B}^4\right)\Upsilon_t = \left(1 - q_1 \mathrm{B}\right)\left(1 - \Theta_1 \mathrm{B}^4\right) e_t$$

(εποχικοί όροι)

Μέσω των αυτοσυσχετίσεων ACF και των μερικών αυτοσυσχετίσεων PACF αναγνωρίζονται οι εποχικοί όροι. Για ένα εποχικό AR μοντέλο, ARIMA(0,0,0)(1,0,0,)$_{12}$, η ACF θα φθίνει εκθετικά στις εποχικές υστερήσεις ενώ η PACF θα παρουσιάζει μόνο μια στατιστικά σημαντική τιμή για υστέρηση 12. Αντίστοιχα για ένα εποχιακό MA μοντέλο, ARIMA(0,0,0)(0,0,1)$_{12}$, η ACF θα παρουσιάζει μόνο μια στατιστικά σημαντική τιμή για υστέρηση 12 και η PACF θα φθίνει εκθετικά στις εποχικές υστερήσεις 12, 24, 36,…

3.16 <u>Συμπέρασμα</u>

Στο κεφάλαιο αυτό λοιπόν αναλύθηκαν κάποια βασικά μοντέλα χρονοσειρών στο πεδίο του χρόνου και είδαμε πως αυτά μπορούν να χρησιμοποιηθούν και να γίνουν πολύτιμα εργαλεία στα χέρια ειδικών αναλυτών οι οποίοι στοχεύουν σε υψηλά επίπεδα ακρίβειας .Επίσης έγινε αναφορά στα εποχικά μοντέλα χρονοσειρών γνωστά και ως SARIMA. Παρακάτω θα ασχοληθούμε εκτενέστερα με την εποχικότητα, τον τρόπο εντοπισμού των εποχικών δεικτών και τις εφαρμογές της στο χώρο των επιχειρήσεων οικονομίας.

ΚΕΦΑΛΑΙΟ 4: ΕΠΟΧΙΚΟΤΗΤΑ ΚΑΙ ΟΙ ΕΦΑΡΜΟΓΕΣ ΤΗΣ

4.1 Εισαγωγή

Πολλά δεδομένα οικονομικών χρονοσειρών που είναι διαθέσιμα για τις τράπεζες δεδομένων υπολογιστών καθώς και δημοσιεύσεις σε περιοδικά κ εφημερίδες, είναι εποχικά προσαρμοσμένα. Κάποιες φορές μπορεί να θεωρείται επιθυμητή η πρόβλεψη εποχικών χρονοσειρών που δεν είναι προσαρμοσμένες για εποχικότητα.

Η εποχικότητα ποικίλει από χρόνο σε χρόνο δείχνοντας ότι υποδείγματα που βασίζονται σε μη προσαρμοσμένα δεδομένα είναι πολύ πιθανό να είναι περισσότερο ευέλικτα ή χρήσιμα. Αν αναμένεται μεταβαλλόμενη εποχικότητα, είναι συνήθως προτιμότερο να λαμβάνεται αυτό υπόψη μέσω της ανάπτυξης ενός καταλλήλως εξειδικευμένου μοντέλου.

4.2 Προβλέψεις στις Χρονοσειρές

Ο σχεδιασμός είναι μια κεντρική δραστηριότητα της διοίκησης των επιχειρήσεων. Συχνά, η κατάλληλη απόφαση που πρέπει να πάρει μια επιχείρηση για οτιδήποτε την απασχολεί, εξαρτάται από μεταβλητές οι οποίες δεν βρίσκονται υπό τον έλεγχο της διοίκησης. Για παράδειγμα, ο αριθμός των εμπορικών αντιπροσώπων που χρειάζεται μια εταιρία για το επόμενο έτος , εξαρτάται από τη μελλοντική ζήτηση των καταναλωτών, για τα προϊόντα της. Θα ήταν λοιπόν επιθυμητό, να υπήρχε η δυνατότητα, οι ανεξέλεγκτες αυτές μεταβλητές να προβλεφθούν πριν ληφθούν αποφάσεις σχετικές με τον σχεδιασμό διαφόρων δραστηριοτήτων.

Οι μέθοδοι πρόβλεψης μπορούν να ταξινομηθούν σε δύο κατηγορίες:

- Ποιοτικές μέθοδοι πρόβλεψης και
- Ποσοτικές μέθοδοι πρόβλεψης

Οι ποιοτικές μέθοδοι πρόβλεψης είναι πολύ σημαντικές όταν δεν υπάρχουν διαθέσιμα ιστορικά στοιχεία, όπως για παράδειγμα στην περίπτωση που το *marketing* μιας εταιρίας θα ήθελε να προβλέψει τις πωλήσεις ενός νέου προϊόντος. Οι ποιοτικές μέθοδοι θεωρούνται πολύ υποκειμενικές.

Αντίθετα, οι ποσοτικές μέθοδοι πρόβλεψης χρησιμοποιούν τα ιστορικά δεδομένα αφού έχουν ως στόχο τη μελέτη των στοιχείων του παρελθόντος ώστε να γίνει βαθύτερη κατανόηση της δομής των δεδομένων και να παρθούν σωστές αποφάσεις για το μέλλον.

Η πρόβλεψη είναι μια από τις σημαντικότερες λειτουργίες μέσα σε μια επιχείρηση και εν γένει σε έναν οργανισμό, για τη λήψη κάθε κρίσιμης απόφασης: ο έλεγχος του κόστους, ο σχεδιασμός νέων προϊόντων, η πρόσληψη προσωπικού, ο όγκος της παραγωγής, το ύψος των αποθεμάτων, όλα καθορίζονται από την πρόβλεψη. Χωρίς αυτήν, κάθε απόφαση θα λαμβανόταν στην τύχη.

Η πρόβλεψη μπορεί να είναι βραχυπρόθεσμη, μεσοπρόθεσμη ή μακροπρόθεσμη ανάλογα με τον χρονικό ορίζοντα στον οποίο αναφέρεται. Π.χ. η πρόβλεψη των τιμών μιας μετοχής για το επόμενο τρίμηνο μπορεί να θεωρηθεί βραχυπρόθεσμη ενώ η πρόβλεψη της τιμής του δείκτη του Γενικού Χρηματιστηρίου Αθηνών, σε μια τριετία από τώρα μπορεί να θεωρηθεί ως μακροπρόθεσμη.

4.2.1 <u>Αρχές της Πρόβλεψης</u>

1) Καμία πρόβλεψη δεν είναι τέλεια: καθώς περιλαμβάνει το στοιχείο της αβεβαιότητας, η πρόβλεψη θα περιέχει κάποιο σφάλμα (δηλ, τη διαφορά μεταξύ της πρόβλεψης και της πραγματικότητας). Με βάση αυτό, στόχος της διαδικασίας πρόβλεψης είναι η ελαχιστοποίηση του σφάλματος για την όσο το δυνατόν ακριβέστερη προσέγγιση της πραγματικότητας.

2) Μια πρόβλεψη είναι περισσότερο ακριβής για ομάδες στοιχείων παρά για μεμονωμένα στοιχεία: π.χ. η πρόβλεψη της συνολικής ζήτησης για μια εταιρία παγωτών (ΕΒΓΑ, ΔΕΛΤΑ κτλ) για το επόμενο έτος θα είναι ακριβέστερη από την ζήτηση για ένα συγκεκριμένο παγωτό (π.χ. του STATUS).

3) Η πρόβλεψη είναι περισσότερο ακριβής όταν είναι βραχυπρόθεσμη παρά όταν είναι μακροπρόθεσμη: όσο κοντινότερος είναι ο χρονικός ορίζοντας της πρόγνωσης τόσο μικρότερος είναι ο βαθμός αβεβαιότητας και άρα τόσο μικρότερο το σφάλμα που θα περιέχει. Ένα κλασσικό παράδειγμα αφορά στην πρόβλεψη του καιρού: ένα μετεωρολογικό δελτίο για τις επόμενες δύο ή τρεις μέρες είναι πάρα πολύ πιθανό να είναι βγει αληθινό. Αντίθετα, η πρόγνωση για τον καιρό του επόμενου μήνα έχει μεγάλες πιθανότητες να αποδειχτεί λανθασμένη.

4.2.2 Πρόβλεψη και μέθοδοι εξομάλυνσης

Η περιγραφή της διαχρονικής εξέλιξης και η πρόβλεψη της μελλοντικής συμπεριφοράς μιας χρονολογικής σειράς αποτελεί ένα από τα πιο ενδιαφέροντα αντικείμενα μελέτης της οικονομίας, της επιχειρησιακής έρευνας, της δημογραφίας και πολλών άλλων επιστημών.

Όλες οι μέθοδοι ανάλυσης χρονολογικών σειρών αναζητούν υποδείγματα, που θα περιγράψουν κατά τον καλύτερο δυνατό τρόπο τα υπάρχοντα εμπειρικά δεδομένα. Τα υποδείγματα αυτά πρέπει να διαθέτουν τη μεγαλύτερη δυνατή απλότητα και συγχρόνως τον ελάχιστο δυνατό αριθμό παραμέτρων χωρίς όμως να βλάπτεται η ευελιξία τους.

Βέβαια κάθε αναλυτής που θέλει να κάνει μια πρόβλεψη αξιολογεί με τα δικά του κριτήρια αυτές τις διαδικασίες ανάλυσης και διάσπασης των χρονοσειρών. Άρα χρειάζεται καλή αξιολογική κρίση με σκοπό να χρησιμοποιηθούν αυτές οι μέθοδοι σε προβλήματα πρόβλεψης

Βασικό ρόλο ώστε να δημιουργήσουμε ένα αξιόπιστο πρότυπο με το οποίο θα προσπαθήσουμε να έχουμε και ακρίβεια στην πρόβλεψη, έχουν οι μέθοδοι εξομάλυνσης (Smoothed Methods), οι βασικότερες των οποίων είναι του κινητού μέσου και της εκθετικής εξομάλυνσης.

Σύμφωνα με μελέτες έχει αποδειχθεί πως οι μέθοδοι εκθετικής εξομάλυνσης έχουν ικανοποιητικά ποσοστά ακρίβειας σε σχέση με πιο πολύπλοκες μεθόδους πρόβλεψης. Και αυτό επειδή δεν επηρεάζονται από τις ιδιομορφίες των προτύπων των δεδομένων ή από περιστασιακά ακραίες τιμές που παρατηρούνται σε επιχειρησιακά δεδομένα.

Η επιλογή της μεθόδου ποικίλει ανάλογα με τον αναλυτή. Ακολουθούν κάποια βήματα για τη διενέργεια μιας σωστής πρόβλεψης:

1) Επιλέγουμε μια μέθοδο πρόβλεψης βασισμένη σε ήδη αποκτημένη γνώση γύρω από το παρατηρηθέν μοντέλο της χρονοσειράς.
2) Χρησιμοποιούμε τη μέθοδο πρόβλεψης για την απόκτηση εκτιμημένης τιμής των δεδομένων.
3) Υπολογίζουμε το σφάλμα της πρόβλεψης.
4) Παίρνουμε απόφαση για την καταλληλότητα του μοντέλου που βασίζεται στην μέτρηση του σφάλματος πρόβλεψης.

(Jeffrrey Jarrett, 2002)

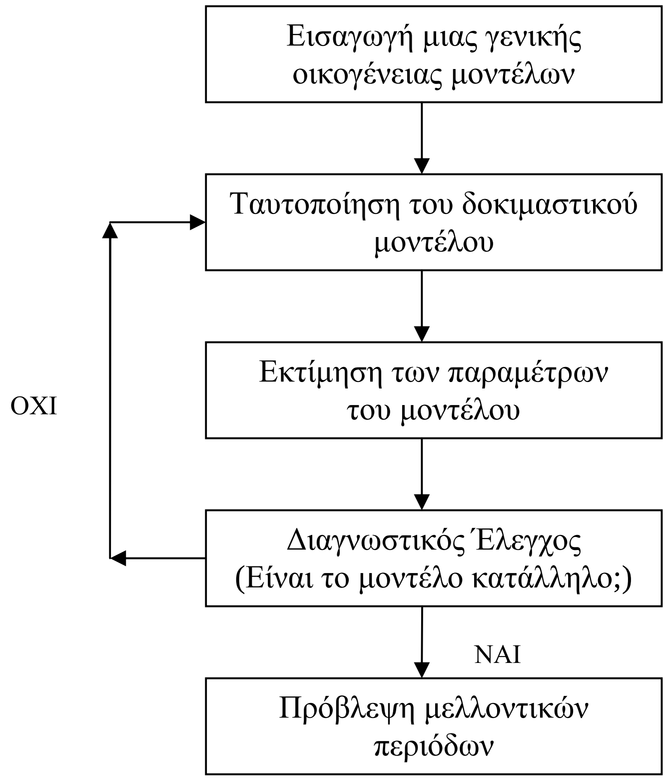
Τα τελευταία 25 χρόνια έχει αναπτυχθεί μια πολύ πιο προχωρημένη μεθοδολογία ανάλυσης χρονολογικών σειρών με πιο δημοφιλή την προσέγγιση που προτείνουν οι **Box - Jenkins** στο έργο τους: Time *series analysis. Forecasting and control. San Francisco 1970*, με την οποία θα ασχοληθούμε στη συνέχεια.

4.3 Η μεθοδολογία Box- Jenkins στην πρόβλεψη

Η μεθοδολογία των Box-Jenkins αποτελεί μια στατιστικά εξειδικευμένη προσέγγιση στην ανάλυση και την κατασκευή ενός υποδείγματος πρόβλεψης, με στόχο την όσο το δυνατόν καλύτερη αναπαράσταση μιας χρονοσειράς.

Οι Box- Jenkins το 1976 ενοποίησαν με επιτυχία τη μεθοδολογία που εκφράζει τις ανάγκες των γνώσεων για την κατανόηση και χρησιμοποίηση απλών μοντέλων ARMA με μια μεταβλητή και περιλαμβάνει κατά σειρά τις παρακάτω διαδικασίες:

- Ταυτοποίηση του υποδείγματος
- Εκτίμηση του υποδείγματος
- Διαγνωστικούς ελέγχους του υποδείγματος
- Πρόβλεψη μελλοντικών περιόδων , οι οποίες παρουσιάζονται στο σχήμα που ακολουθεί



Σχήμα 4.1 Πηγή: Πτυχιακή Εργασία "Βασικές Μέθοδοι Ανάλυσης Χρονοσειρών"2005

Σκοπός της φάσης της ταυτοποίησης είναι να γίνει η επιλογή ενός ιδιαίτερου μοντέλου ARMA από τη γενική ομάδα τςν ARIMA(p, d, q). Η επιλογή των κατάλληλων τιμών p και q προϋποθέτει την εξέταση των συντελεστών μερικής αυτοσυσχέτισης και αυτοσυσχέτισης αντίστοιχα που έχουν υπολογιστεί για τα δεδομένα.

Το δεύτερο στάδιο στην κατασκευή του υποδείγματος είναι η εκτίμηση ή το στάδιο προσαρμογής. Τα υποδείγματα ARIMA μπορούν να εκτιμηθούν με τη μέθοδο των ελαχίστων τετραγώνων. Οι Box-Jenkins (1976) προτείνουν μια ποικιλία διαδικασιών εκτίμησης των παραμέτρων στις οποίες ο αναγνώστης που ενδιαφέρεται μπορεί να προστρέξει για λεπτομέρειες.

Ακολούθως, πριν το μοντέλο χρησιμοποιηθεί για πρόβλεψη θα πρέπει να γίνει ο έλεγχος της επάρκειάς του. Αυτός ο διαγνωστικός έλεγχος γίνεται με την εξέταση των όρων του σφάλματος e_t, για να βεβαιωθούμε ότι είναι τυχαίοι. Αν οι όροι του σφάλματος είναι στατιστικά διάφοροι του μηδενός, το υπόδειγμα δεν θεωρείται επαρκές.

Μετά την ταυτοποίηση και τον έλεγχο καταλληλότητας ενός υποδείγματος, μπορούν να γίνουν προβλέψεις μιας ή περισσοτέρων περιόδων στο μέλλον. Βέβαια καθώς η περίοδος της πρόβλεψης γίνεται πιο μακρινή μέσα στο μέλλον, οι πιθανότητες σφάλματος πρόβλεψης μεγαλώνουν.

Η προσέγγιση αυτή έχει έναν αριθμό πλεονεκτημάτων έναντι των άλλων μεθόδων ανάλυσης χρονοσειρών όπως:

1) Η μέθοδος είναι λογική και στατιστικά ακριβής
2) Αποσπά ένα μεγάλο αριθμό πληροφοριών από τα ιστορικά δεδομένα της χρονοσειράς
3) Επιδρά στη ακρίβεια της πρόβλεψης, ενώ συγκρατεί τον αριθμό των παραμέτρων σε ένα ελάχιστο επίπεδο σε σύγκριση με άλλες παρόμοιες διαδικασίες κατασκευής υποδειγμάτων
4) Είναι απαλλαγμένη από ρεαλιστικές υποθέσεις όπως αυτές που βασανίζουν τα οικονομετρικά υποδείγματα

Τα μοντέλα Box- Jenkins για εξομάλυνση χρονοσειρών έχουν τις παρακάτω δύο επιθυμητές ιδιότητες:

- Αποκαλύπτουν κατά τρόπο ικανοποιητικό την υπάρχουσα νομοτέλεια η οποία πιθανόν υφίσταται μεταξύ των διαδοχικών παρατηρήσεων μιας μεταβλητής, και

- Δίνουν άριστες προβλέψεις για την μελλοντική εξέλιξη μιας μεταβλητής, έτσι ώστε να παρέχεται σε κάθε ενδιαφερόμενο η δυνατότητα να αξιοποιήσει κατάλληλα αυτές τις προβλέψεις ή αν χρειαστεί να κάνουν τους απαραίτητους χειρισμούς ή ακόμη να πάρουν τα κατάλληλα μέτρα στην περίπτωση που οι προβλέψεις δείχνουν δυσμενή εξέλιξη.

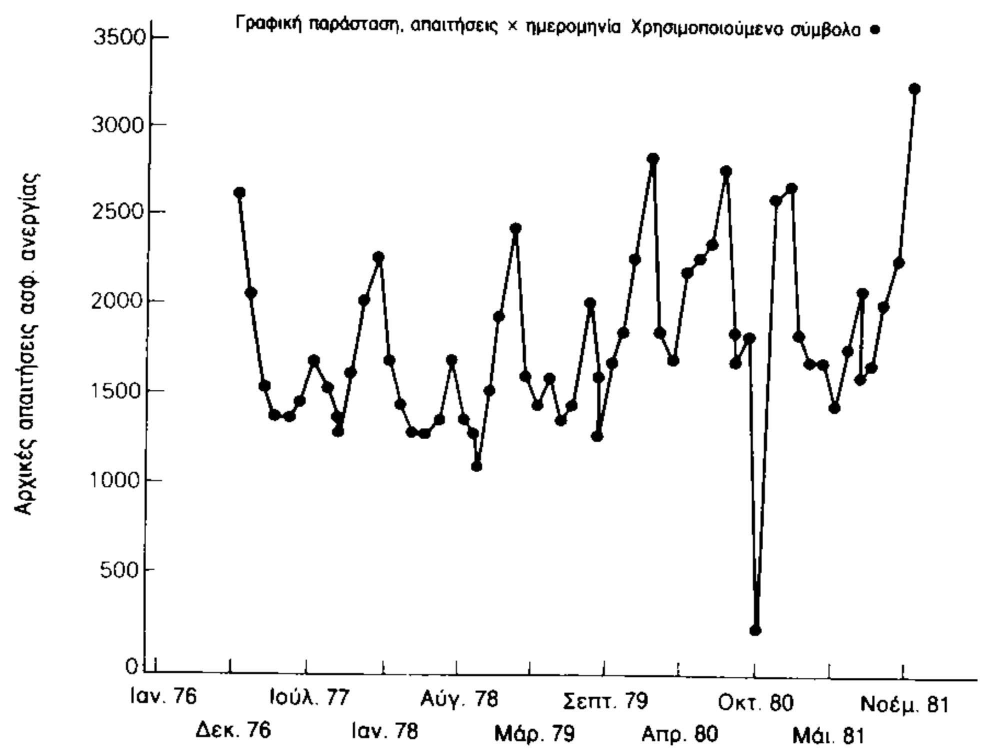
Βέβαια πρέπει να αναφέρουμε και κάποια μειονεκτήματα της μεθοδολογίας αυτής. Σύμφωνα με τους Pack και Downing (1983) η αιτία που μπορεί να αποτύχει η προσέγγιση των Box- Jenkins στο να δώσει καλύτερα και πιο ακριβή αποτελέσματα οφείλεται στους παρακάτω παράγοντες:

1) στο μέγεθος (μήκος της χρονοσειράς).
2) στο περιεχόμενο των πληροφοριών (αυτοσυσχέτιση) της χρονοσειράς.
3) στην ύπαρξη ακραίων τιμών ή διαρθρωτικών μεταβολών της χρονοσειράς.
4) στη χρησιμοποίηση δεδομένων που εκφράζουν μέσους όρους της χρονοσειράς.
5) στην επιλογή της έναρξης του χρόνου πρόβλεψης.

4.4 <u>Εποχικότητα και Μοντέλα ARIMA</u>

Τα εποχικά πρότυπα παρατηρούνται σε δεδομένα που έχουν ταξινομηθεί ανά τρίμηνο, μήνα ή εβδομάδα. Επίσης λαμβάνουν χώρα ή επανακυκλώνονται μέσα στην περίοδο ενός χρόνου, επαναλαμβάνονται από τη μια χρονιά στην άλλη και μπορούν εύκολα να προβλεφθούν. Ένα παράδειγμα εποχικής διακύμανσης φαίνεται στο παρακάτω σχήμα όπου εμφανίζονται τα μηνιαία δεδομένα αρχικών απαιτήσεων των πολιτειακών ασφαλιστικών προγραμμάτων ανεργίας. Οι απαιτήσεις παρουσιάζουν υψηλές τιμές τον Ιανουάριο και σχεδόν μέχρι το Μάιο πέφτουν, φθάνουν σε νέα κορυφή συνήθως κατά τον Ιούλιο, μετά ξαναπέφτουν και αρχίζουν την ανοδική τάση από τον Σεπτέμβριο έως τον Ιανουάριο της επόμενης χρονιάς.

Όπως παρατηρείται οι απαιτήσεις των επιδομάτων ανεργίας είναι πολύ υψηλές τον Ιανουάριο, αφού πολλοί άνθρωποι αποδεσμεύονται από την εποχική απασχόληση κατά την περίοδο των διακοπών του Δεκεμβρίου. Έως την περίοδο των ζεστότερων μηνών οι απαιτήσεις μειώνονται. Την εποχή αυτή επειδή τα σχολεία και τα κολέγια σταματούν τα μαθήματά τους η προσφορά εργασίας αυξάνεται, ενώ οι τιμές που αφορούν την ανεργία συνεχίζουν να είναι υψηλές ως τον Σεπτέμβριο όπου τα μαθήματα ξεκινούν πάλι.



Σχημα 4.2: Αρχικές απαιτήσεις επιδομάτων ανεργίας- πολιτειακά προγράμματα (εκατοντάδες) .(Πηγή:Jeffrrey Jarrett, 2002)

Υπάρχουν χρονοσειρές που παρουσιάζουν σε ίσα χρονικά διαστήματα την ίδια συμπεριφορά. Όπως τα μοντέλα ARIMA (p,d,q) έτσι και αυτά που εμφανίζουν εποχικότητα s παρουσιάζουν παρόμοιες ιδιότητες.

Τα μοντέλα ARIMA που έχουν επεκταθεί για να χειρίζονται την εποχικότητα συμβολίζονται ως :

$$\mathrm{ARIMA}(p,d,q)\times(P,D,Q)_s$$ **,όπου s η εποχικότητα.**

Για χρονοσειρές που περιέχουν μια εποχική συνιστώσα που επαναλαμβάνεται μετά από κάθε s παρατήρηση, μπορεί να χρησιμοποιηθεί ένα βοηθητικό υπόδειγμα στο μη- εποχικό υπόδειγμα ARIMA . Για μια εποχική χρονοσειρά με περίοδο s, το s = 12 αφορά μηνιαία δεδομένα και το s= 4 αφορά τριμηνιαία. Τα εποχικά υποδείγματα $(P,D,Q)_s$ μπορούν να αναπτυχθούν για εποχικά δεδομένα με τρόπο παρόμοιο της αρχικής διαδικασίας ARIMA.

Όπως έχει ήδη αναφερθεί μια χρονοσειρά λέμε ότι έχει εποχικότητα όταν παρουσιάζει ομοιότητες κάθε s διαστήματα ή ισοδύναμα , η ίδια συμπεριφορά επαναλαμβάνεται κάθε s παρατηρήσεις.

Πολλά φαινόμενα όπως θερμοκρασία, υγρασία , αφίξεις επιβατών, πωλήσεις προϊόντων κ.λ.π. δίνουν τέτοιες χρονικές σειρές. Μπορούμε να έχουμε και περισσότερες από μία εποχικότητες στην ίδια χρονική σειρά.

Αναπαριστάνοντας μια χρονοσειρά σε όρους ενός πολλαπλασιαστικού υποδείγματος, είναι συχνά δυνατό να μειώνεται ο αριθμός των παραμέτρων που πρόκειται να εκτιμηθούν. Επίσης βοηθά στην ερμηνεία της δομής του υποδείγματος.

Θεωρούμε ένα μηνιαίο εποχικό υπόδειγμα το οποίο έχει εποχική διαφόριση (σε μήνες) και επίσης, κανονικές και εποχικές παραμέτρους αυτοπαλινδρόμησης πρώτης τάξης.

Συνεπώς,

$(p, d, q) \times (P, D, Q)_s = (1, 0, 0) \mathrm{X} (1, 1, 0)_{12}$

Σε όρους του Y_t έχουμε:

$Y_t = \varphi Y_{t-1} + (1+\varphi_{12})Y_{t-13}\ \varphi_{12}\ Y_{t-24} + \varphi_1\varphi_{12}Y_{t-25} + e_t$

Η χρήση του τελεστή της προς τα πίσω μετατόπισης μας δίνει τη δυνατότητα να εκφράζουμε την τάξη και τον τύπο της διαδικασίας σε όρους του εκθέτη του B και του τύπου της παραμέτρου.

<u>Σημείωση 1:</u>

Εάν μελετάμε εποχικά δεδομένα, μια διαδικασία ARMA μπορεί να αποτελείται από δύο μέρη:

1.Το μη εποχικό μέρος (αναλογία)
2.Το εποχικό μέρος (αναλογία)

Για σκοπούς που εξυπηρετούν μια διαδικασία ARMA ,διαιρούμε την διαδικασία σε δύο τμήματα. Για την ταυτοποίηση του εποχικού προτύπου αγνοούμε την μη-εποχική διαδικασία και εξετάζουμε αν η εποχικότητα ορίζεται ή όχι από μια διαδικασία ARIMA, αφού συγκεντρώσουμε την προσοχή μας στους συντελεστές των εποχικών όρων.

Αν υποθέσουμε ότι το μη-εποχικό τμήμα είναι ένα ARIMA (1,0,1) και ότι η χρονοσειρά δίνει ένα τριμηνιαίο εποχικό πρότυπο, τότε το πλήρες εποχικό υπόδειγμα γράφεται:

$$\left(1-f_1\mathrm{B}\right)\left(1-f_4\mathrm{B}^4\right)\mathrm{Y}_t=\left(1-q_1\mathrm{B}\right)e_t$$

εάν η εποχικότητα βρίσκεται στο μέρος (αναλογία) AR,ή

$$\left(1-f_1\mathrm{B}\right)\ \mathrm{Y}_t\ =\left(1-q_1\mathrm{B}\right)\left(1-q_4\mathrm{B}^4\right)e_t$$

εάν η εποχικότητα βρίσκεται στο μέρος MA, όπου

$$\left(1-f_4\mathrm{B}^4\right)\mathrm{Y}_t=\mathrm{Y}_t\text{-}f_4\mathrm{B}^4\ \mathrm{Y}_t\quad \text{ομοίως}\quad \left(1-q_4\mathrm{B}^4\right)e_t\ =\ e_t q_4\mathrm{B}^4e_t=e_t-q_4e_{t-12}\,.$$

Επομένως τα παραπάνω υποδείγματα περικλείουν εποχικές παραμέτρους που χρησιμοποιούν τα Y_t ή E_t των προηγούμενων τεσσάρων τριμήνων ώστε η εποχικότητα να λαμβάνεται υπόψη.

<u>Σημείωση 2:</u>

Σε μια εποχική διαδικασία MA(2) με ένα επίπεδο διαφόρισης, έχουμε την εξής έκφραση:

$$\left(1-B\right)\left(1-B^4\right)Y_t=\left(1-q_1\mathrm{B}-q_2\mathrm{B}^2\right)e_t$$

(διαφόριση ενός επιπέδου)(εποχικό πρότυπο στο R)=(MA(2)μη εποχικό)

$$\left(1-B\right)\ Y_t=\left(1-q_1\mathrm{B}-q_2\mathrm{B}^2\right)\left(1-q_4\mathrm{B}^4\right)e_t$$

(διαφόριση ενός επιπέδου) = (MA(2)μη εποχικό)(εποχικές παράμετροι στο MA)

(Jeffrrey Jarrett, 2002)

4.5 <u>Χρονολογικές σειρές με Εποχικότητα</u>

Μελετώντας τις χρονικές σειρές με περιοδικότητα ας δούμε ένα παράδειγμα.

Παράδειγμα 1 :

Υποθέτουμε ότι έχουμε μέσες μηνιαίες θερμοκρασίες xt, t=1, 2, 3 για μερικά έτη. Οι χρονοσειρές αυτές παρουσιάζουν περιοδικότητα ανά 12 μήνες δηλ. s = 12.
Αν λοιπόν οι παρατηρήσεις μας είναι:

χ1, χ2, χ3, ….., χ12
χ13, χ14, χ15,……,χ24
χ25, χ26, χ27, …...,χ36

περιμένουμε, μέσα σε κάθε έτος και για όλα τα έτη δηλ. οι παρατηρήσεις χ1, χ2, χ3, ….., χ12 καθώς και οι χ13, χ14, χ15,……,χ24 δηλαδή οι Ιανουάριοι, οι Φεβρουάριοι κ.λ.π. όλων των ετών να περιέχουν επίσης την ίδια συμπεριφορά.

Το ότι όλοι οι Ιανουάριοι κ.λ.π. δηλαδή οι μήνες που απέχουν μεταξύ τους s = 12 χρονικά διαστήματα έχουν σχέση σημαίνει ότι ο συντελεστής αυτοσυσχέτισης Κ =12 , θα έχει μια υψηλή τιμή, σημαντικά μεγαλύτερη από την τιμή των γειτονικών αυτοσυσχετίσεων. Επίσης είναι πιθανό και οι αυτοσυσχετίσεις που αντιστοιχούν σε υστέρηση πολλαπλάσια του Κ = 12 να είναι επίσης σημαντικές.

Έστω τώρα ότι μελετώντας τις παρατηρήσεις που αντιστοιχούν στον ίδιο μήνα διαφορετικών ετών (π.χ. στους Ιανουάριους) βρίσκουμε ότι το μοντέλο που προσαρμόζεται είναι ένα ARIMA (P,D,Q) μοντέλο που αναλυτικά γράφεται:

$\Phi\rho(B^s)\nabla_S^D X_t = \Theta_Q (B^s)a_t$ (1) όπου s=12

Στο μοντέλο αυτό εμφανίζεται ο τελεστής B^s διότι οι παρατηρήσεις διαφέρουν κατά s=12 και για τον ίδιο λόγο αντί του ∇^D εμφανίζεται ο τελεστής ∇_S^D.

Είναι εύκολο να υποθέσουμε ότι αν χρησιμοποιήσουμε τις παρατηρήσεις που αντιστοιχούν σε άλλο μήνα π.χ στους Φεβρουάριους θα καταλήγαμε σε ένα μοντέλο παρόμοιο με το (1) και το ίδιο θα συνέβαινε με τον οποιοδήποτε μήνα. Μπορούμε ακόμη να υποθέσουμε ότι όλα αυτά τα μοντέλα του ίδιου τύπου, έχουν τις ίδιες τιμές παραμέτρων, δηλαδή τα ίδια $Φp(B^s)$ και $Θ_Q (B^s)$.

Το μοντέλο ARIMA $(p,d,q)(P,D,Q)_S$ είναι το γενικότερο μοντέλο των χρονικών σειρών. Η τάξη του καθορίζεται από επτά σταθερές p, d, q, P, D, Q, s, οι οποίες πρέπει να προσδιοριστούν και να είναι γνωστές προτού αρχίσει η εκτίμηση των παραμέτρων του μοντέλου. Στην περίπτωση του γενικού πολλαπλασιαστικού μοντέλου, οι παράμετροι είναι σε πλήθος p+q+P+Q. Για τον προσδιορισμό των σταθερών S, D, P, Q, d, p, q με τη σειρά που αναφέρονται, ακολουθούμε την παρακάτω διαδικασία:

1. Κάνουμε το γράφημα των αυτοσυσχετίσεων, οι οποίες όπως αναφέρθηκε θα παρουσιάζουν μια πτώση αργή ή γρήγορη, εκθετική ή κυματοειδή. Αν για κάποια υστέρηση K=S , ο αντίστοιχος συντελεστής r_s εμφανίζεται πολύ σημαντικός σε σχέση με τους γειτονικούς του, τότε θεωρούμε ότι το μοντέλο μας έχει εποχικότητα S. Υπολογίζουμε τότε τις αυτοσυσχετίσεις r_s, r_{2s} , r_{3s} … και στη συνέχεια βρίσκουμε την τιμή
   του D.
2. Στη συνέχεια από το γράφημα των αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων υστέρησης s, 2s, 3s προσδιορίζουμε τις τιμές των Q και P αντίστοιχα.
3. προσδιορίζουμε το εποχικό μοντέλο ARIMA (P, D, Q) στις παρατηρήσεις μας και εκτιμούμε τα υπόλοιπα $\hat{a}_t$.Θεωρώντας την σειρά των $\hat{a}_t$ ως χρονική σειρά, θα αναζητήσουμε κατάλληλο ARIMA (p, d, q) μοντέλο και στη συνέχεια εκτιμούμε τις παραμέτρους του συγκεκριμένου μοντέλου στο οποίο καταλήξαμε.

Όλα όσα προαναφέρθηκαν στον καθορισμό του μοντέλου είναι ενδεικτικά. Είναι πολύ δύσκολο να καθοριστεί επακριβώς το μοντέλο από τις αυτοσυσχετίσεις και τις μερικές αυτοσυσχετίσεις γιατί η εικόνα που φαίνεται στα γραφήματά τους δεν είναι τόσο σαφής. Είναι όμως ο πιο απλός τρόπος και δεν προϋποθέτει γνώση της ανάλυσης των χρονοσειρών σε βάθος.(Θεόδωρου Η. Αποστολόπουλου)

4.6 Προσδιορισμός των εποχικών μεταβολών- Ο Εποχικός δείκτης

Για να προσδιορίσουμε τον εποχικό δείκτη θα πρέπει να προσδιορίσουμε τον τρόπο με τον οποίο μεταβάλλονται τα δεδομένα της χρονοσειράς από μήνα σε μήνα κατά τη διάρκεια ενός τυπικού έτους. Ένα σύνολο αριθμών το οποίο δείχνει τις σχετικές τιμές μιας μεταβλητής κατά τη διάρκεια των μηνών ενός έτους καλείται όπως έχουμε προαναφέρει, εποχικός δείκτης (seasonal index) της μεταβλητής. Για παράδειγμα εάν γνωρίζουμε ότι οι πωλήσεις κατά τους μήνες Ιανουάριο, Φεβρουάριο, Μάρτιο κ.λ.π. είναι 50%, 120, 90%... της μέσης πώλησης, ανά μήνα , κατά τη διάρκεια ενός έτους τότε αυτοί λέγονται τιμές του εποχικού δείκτη(seasonal index numbers). Η μέση τιμή του εποχικού δείκτη για όλο το έτος, θα πρέπει να είναι 100%, συνεπώς το άθροισμα των εποχικών δεικτών των 12 μηνών του έτους θα πρέπει να είναι 1200%. Για τον υπολογισμό του εποχικού δείκτη υπάρχουν διάφορες μέθοδοι:

4.6.1 Η μέθοδος μέσων ποσοστών

Στη μέθοδο αυτή τα δεδομένα κάθε μήνα εκφράζονται ως ποσοστά επί του συνόλου του έτους. Αθροίζονται τα επί τοις εκατό ποσοστά για τους αντίστοιχους μήνες μιας σειράς ετών και προσδιορίζεται η μέση τιμή ή διάμεσος. Εάν χρησιμοποιηθεί η μέση τιμή καλό είναι να αποφεύγονται οι ακραίες τιμές που τυχόν εμφανίζονται. Τα 12 ποσοστά που προκύπτουν δίνουν τον εποχικό δείκτη. Εάν η μέση τους τιμή δεν είναι 100%(εάν δεν είναι δηλαδή το άθροισμά τους 1200%) θα πρέπει να κανονικοποιηθούν κάτι που επιτυγχάνεται με πολλαπλασιασμό τους με έναν κατάλληλο παράγοντα.

4.6.2 Η μέθοδος της ποσοστιαίας τάσης.

Εδώ τα δεδομένα για κάθε μήνα ως ποσοστά των μηνιαίων τιμών τάσης. Ο ζητούμενος εποχικός δείκτης προκύπτει από τον υπολογισμό της κατάλληλης μέσης τιμής, των ποσοστών που προκύπτουν για τους αντίστοιχους μήνες. Ομοίως με την προηγούμενη μέθοδο εάν τα ποσοστά δεν δίνουν άθροισμα 1200% πρέπει να κανονικοποιηθούν. Η διαίρεση κάθε μηνιαίας τιμής Y με την αντίστοιχη τιμή της τάσης T δίνει Y/T =CSI ,από την εξίσωση του πολλαπλασιαστικού μοντέλου και ο προσδιορισμός της μέσης τιμής των πηλίκων παράγει τους εποχικούς δείκτες. Αν αυτοί περιλαμβάνουν κυκλικές και ακανόνιστες μεταβολές τότε προκύπτει σημαντικό μειονέκτημα, ειδικά αν οι προαναφερθείσες μεταβολές είναι έντονες.

4.6.3 Η μέθοδος της ποσοστιαίας κυλιόμενης μέσης τιμής

Στη μέθοδο αυτή υπολογίζουμε την κυλιόμενη μέση τιμή για τους 12 μήνες. Αν το αποτέλεσμα που πάρουμε εμπίπτουν μεταξύ διαδοχικών μηνών (αντί για τα μέσα των μηνών, όπου αναφέρονται τα αρχικά δεδομένα), υπολογίζουμε την κυλιόμενη μέση τιμή 2 μηνών των παραπάνω κυλιόμενων μέσων τιμών των 12 μηνών. Το αποτέλεσμα συνήθως ονομάζεται κεντρωμένη κυλιόμενη τιμή 12 μηνών.

Στη συνέχεια εκφράζοντας τα αρχικά δεδομένα κάθε μήνα ως ποσοστά, επί της τιμής αυτής, που αντιστοιχεί στα αρχικά δεδομένα. Έτσι προσδιορίζοντας τη μέση τιμή των ποσοστών των αντίστοιχων μηνών προκύπτει ο επιθυμητός εποχικός δείκτης. Ισχύει το ίδιο σχετικά με την κανονικοποίηση.

Η μέθοδος αυτή απορρέει από την εξίσωση του πολλαπλασιαστικού μοντέλου. Η κυλιόμενη μέση τιμή 12 μηνών του Y επιτρέπει την εξάλειψη των εποχικών και ακανόνιστων κινήσεων S και I άρα ισούται με το γινόμενο TC. Στη συνέχεια διαιρώντας τα αρχικά δεδομένα Y με TC προκύπτει το γινόμενο SI. Η λήψη μέσων τιμών στους αντίστοιχους μήνες επιτρέπει την εξάλειψη της μη κανονικότητας και έτσι αυτό που υπολείπεται είναι ο ζητούμενος εποχικός δείκτης S.

Παράδειγμα 2:

Ο παρακάτω πίνακας 4.1 παρουσιάζει τον αριθμό των νέων κατοικιών (σε χιλιάδες) που δημιουργήθηκαν στις ΗΠΑ από τον Ιανουάριο του 1990 έως τον Δεκέμβριο του 1995.

Α) Απεικονίστε τα δεδομένα σε ένα διάγραμμα

Β) Προσδιορίστε τον εποχικό δείκτη με τη βοήθεια της μεθόδου των μέσων ποσοστών.

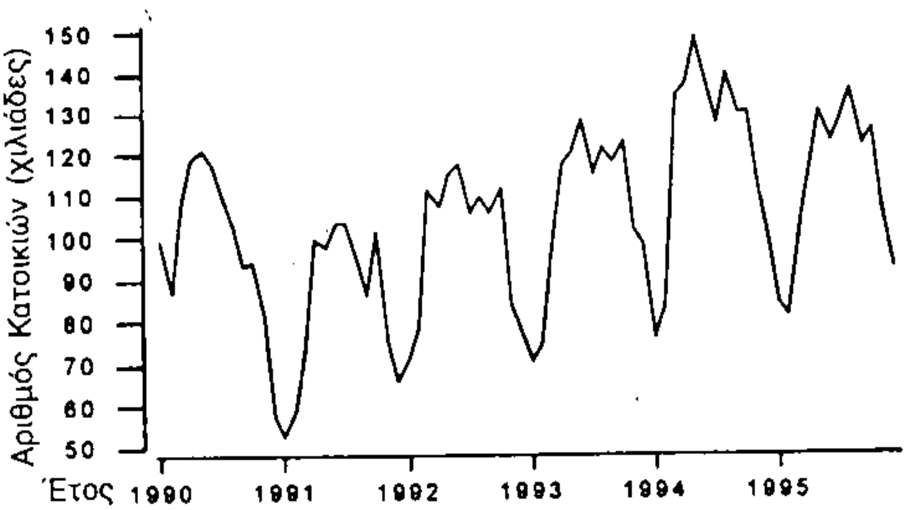
Πίνακας 4.1

Μήνας	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**
Ιαν	99.2	52.5	71.6	70.5	76.2	84.5
Φεβ	86.9	59.1	78.8	74.6	83.5	81.6
Μαρ	108.5	73.8	111.6	95.5	134.3	103.8
Απρ	119.0	99.7	107.6	117.8	137.6	116.9
Μαϊ	121.1	97.7	115.2	130.9	148.8	130.5
Ιουν	117.8	103.4	117.8	128.5	136.4	123.4
Ιουλ	111.2	103.5	106.2	115.3	127.8	129.1
Αυγ	102.8	94.7	109.9	121.8	139.8	135.8
Σεπτ	93.1	86.6	106.0	118.5	130.1	122.4
Οκτ	94.2	101.8	111.8	123.2	130.6	126.2
Νοε	81.4	75.6	84.5	102.3	113.4	107.2
Δεκ	57.4	65.6	78.6	98.7	98.5	92.8

Πηγή: Γραφείο απογραφών ΗΠΑ, Τρέχουσες Αναφορές Κατασκευών

Λύση:

Α) Η διαγραμματική απεικόνιση φαίνεται στο σχήμα 4.3



Σχήμα 4.3:Δημιουργία νέων κατοικιών στις ΗΠΑ από τον Ιανουάριο του 1990 έως τον Δεκέμβριο του 1995.

Ο πίνακας 4.2 αναγράφει τα αθροίσματα και τις μέσες τιμές των μηνών για κάθε έτος μεταξύ 1990 και 1995.Διαιρώντας τα δεδομένα για κάθε μήνα του 4.1 με την αντίστοιχη μέση μηνιαία τιμή κάθε έτους όπως φαίνεται στον 4.2 και εκφράζοντας το αποτέλεσμα ως ποσοστό, δημιουργούμε τις τιμές του πίνακα 4.3. για παράδειγμα, η τιμή για τον Ιανουάριο του 1990 προκύπτει διαιρώντας 99.2/99.4= 99.8%. Η τιμή για τον Ιανουάριο του 1991, 52.5/84.5= 62.1% κ.λ.π .

Η τελευταία στήλη του 4.3 παρουσιάζει για κάθε μήνα την μέση τιμή των ποσοστών. Αν το άθροισμα των ποσοστών της τελευτάς στήλης είναι 1199.8 πρέπει να πολλαπλασιαστούν με τον παράγοντα 1200/1199.8 ώστε να προκύψει συνολικό άθροισμα 1200. Ωστόσο αυτό δεν αλλάζει σημαντικά τα μηνιαία ποσοστά άρα οι αριθμοί αυτής της στήλης δίνουν τον ζητούμενο εποχικό δείκτη, ο οποίος δείχνει ότι κατά μέσο όρο, οι ελάχιστες κατοικίες ξεκινούν κατά τους μήνες Δεκέμβριο Ιανουάριο και Φεβρουάριο και ότι οι περισσότερες ξεκινούν να κτίζονται το Μάϊο και τον Ιούνιο.

Πίνακας 4.2

Έτος	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**
Σύνολο	1192.6	1014.0	1199.6	1287.6	1457.0	1354.2
Μέση τιμή	99.4	84.5	100.0	107.3	121.4	112.9

<u>Πίνακας 4.3</u>

Μήνας	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**Άθροισμα**	**Μ.Τιμή**
Ιαν	99.8	62.1	71.6	65.7	62.8	74.8	436.8	72.8
Φεβ	87.4	69.9	78.8	69.5	68.8	72.3	446+.7	74.5
Μαρ	109.25	87.3	111.6	89.0	110.6	91.9	599.7	99.9
Απρ	119.7	118.0	107.6	109.8	113.3	103.5	672.0	112.0
Μαϊ	121.8	115.6	115.2	112.7	122.6	115.6	703.5	117.2
Ιουν	118.5	122.4	117.8	119.8	112.4	109.3	700.1	116.7
Ιουλ	111.9	122.5	106.2	107.5	105.3	114.3	667.6	111.3
Αυγ	103.4	112.1	109.9	113.5	115.2	120.3	674.3	112.4
Σεπτ	93.7	102.5	106.0	110.4	107.2	108.4	628.2	104.7
Οκτ	94.8	120.5	111.8	114.8	107.6	111.8	661.2	110.2
Νοε	81.9	89.5	84.5	95.3	93.4	95.0	539.6	89.9
Δεκ	57.7	77.6	78.6	92.0	81.1	82.2	469.3	78.2

Παράδειγμα 3:

Α) Να προσδιοριστεί ο εποχικός δείκτης για το παραπάνω παράδειγμα χρησιμοποιώντας τη διάμεσο αντί για τον αριθμητικό μέσο.

<u>Λύση:</u>

Α) Αν τοποθετήσουμε σε αύξουσα σειρά τους αριθμούς της γραμμής του Ιανουαρίου στον πίνακα 4.3 είναι 62.1, 62.8, 65.7, 71.6, 74.8, και 99.8, οπότε η διάμεσος είναι:

(65.7+71.6):2 =68.7.

Ομοίως θα υπολογίσουμε και τις διάμεσους για τους άλλους μήνες και φαίνονται στη δεύτερη στήλη του πίνακα 4.4.Αυτές έχουν άθροισμα 1197.8 και τις διορθώνουμε πολλαπλασιάζοντας με τον παράγοντα 1200/1197.8. Έτσι προκύπτει η τρίτη στήλη του πίνακα. Στην πράξη όταν τα δεδομένα που προκύπτουν από τον αριθμητικό μέσο διαφέρουν από αυτά που προκύπτουν από τη διάμεσο, είναι καλύτερα να χρησιμοποιούμε τη διάμεσο, έτσι ώστε να μη λαμβάνονται υπόψη οι ακραίες τιμές των δεδομένων.

Πίνακας 4.4

Μήνας	**Διάμεσος**	**Εποχικός Δείκτης**
Ιαν.	68.7	68.8
Φεβ.	71.1	71.2
Μαρ.	100.5	100.7
Απρ.	111.6	111.8
Μαϊ.	115.6	115.8
Ιουν.	118.2	118.4
Ιουλ.	109.7	109.9
Αυγ.	112.8	113.0
Σεπτ.	106.6	106.8
Οκτ.	111.8	112.0
Νοε.	91.4	91.6
Δεκ.	79.9	80.0

4.7 Απαλοιφή της εποχικότητας

Οι δείκτες της εποχικότητας έχουν μεγάλη σημασία στην ανάλυση των χρονολογικών σειρών γιατί χρησιμοποιούνται συνήθως για την απαλοιφή της εποχικότητας, δηλαδή για την απαλλαγή των όρων μιας χρονολογικής σειράς από την επίδραση του εποχικού παράγοντα, ώστε να μπορέσουμε να τους συγκρίνουμε μεταξύ τους και στη συνέχεια να μπορέσουμε να μετρήσουμε τις κυκλικές κυμάνσεις. Επίσης όπως είπαμε, η μελέτη των εποχικών κυμάνσεων των χρονολογικών σειρών έχει ως αντικειμενικό σκοπό τη μέτρηση των εποχικών μεταβολών , ώστε να γίνεται ο κατάλληλος προγραμματισμός από τον ιδιωτικό και το δημόσιο τομέα της οικονομίας.

Για την απαλοιφή της εποχικότητας διαιρούμε τα αρχικά δεδομένα(μηνιαία ή τριμηνιαία) μιας χρονολογικής σειράς δια των αντίστοιχων δεικτών εποχικότητας και πολλαπλασιάζουμε τα πηλίκα με το 100. Τα αποτελέσματα που προκύπτουν από τις εποχικές κυμάνσεις όχι όμως και από τις άλλες συνιστώσες(τάση, κυκλική και τυχαία)

Παράδειγμα 4: Για τα μηνιαία δεδομένα του 1988 του παρακάτω παραδείγματος η απαλοιφή της εποχικότητας γίνεται φανερή από τον παρακάτω πίνακα:

Πίνακας 4.5

1988 ΜΗΝΑΣ Y_t	ΔΕΙΚΤΗΣ ΕΠΟΧ. S_t	100 Y_t	T_t=100 Y_t /S_t
ΙΑΝ 53	115.3	5300	45.96
ΦΕΒ 48	103.8	4800	46.24
ΜΑΡ 46	100.8	4600	45.63
ΑΠΡ 42	91.5	4200	45.90
ΜΙΑ 40	86.2	4000	46.40
ΙΟΥΝ 38	81.2	3800	46.79
ΙΟΥΛ 40	84.3	4000	47.44
ΑΥΓ 42	90.5	4200	46.40
ΣΕΠΤ 45	98.0	4500	45.92
ΟΚΤ 49	108.2	4900	45.28
ΝΟΕ 53	115.3	5300	45.88
ΔΕΚ 56	122.8	5600	45.60

Οι αριθμοί της τελευταίας στήλης είναι απαλλαγμένοι από τις εποχικές κυμάνσεις. Βέβαια η ανοδική τάση δεν είναι δυνατόν να φανεί από τα δεδομένα ενός έτους.

4.8 Συμπέρασμα

Σε αυτό το κεφάλαιο λοιπόν παρατηρήσαμε λοιπόν πως οι χρονοσειρές κάτω από την υπόθεση ότι με βάση την συμπεριφορά στο παρελθόν μπορεί να προβλεφθεί η μελλοντική πορεία της διακύμανσης της μεταβλητής που παρατηρούμε. Παρ'όλο που στην πρόβλεψη μελλοντικών καταστάσεων θα έχουμε πάντα σφάλμα , η ανάλυση και ο διαχωρισμός των χρονοσειρών στις συνιστώσες της έχει αξία στην διαδικασία της πρόβλεψης γιατί έτσι μπορεί να μειωθεί το σφάλμα αυτό.

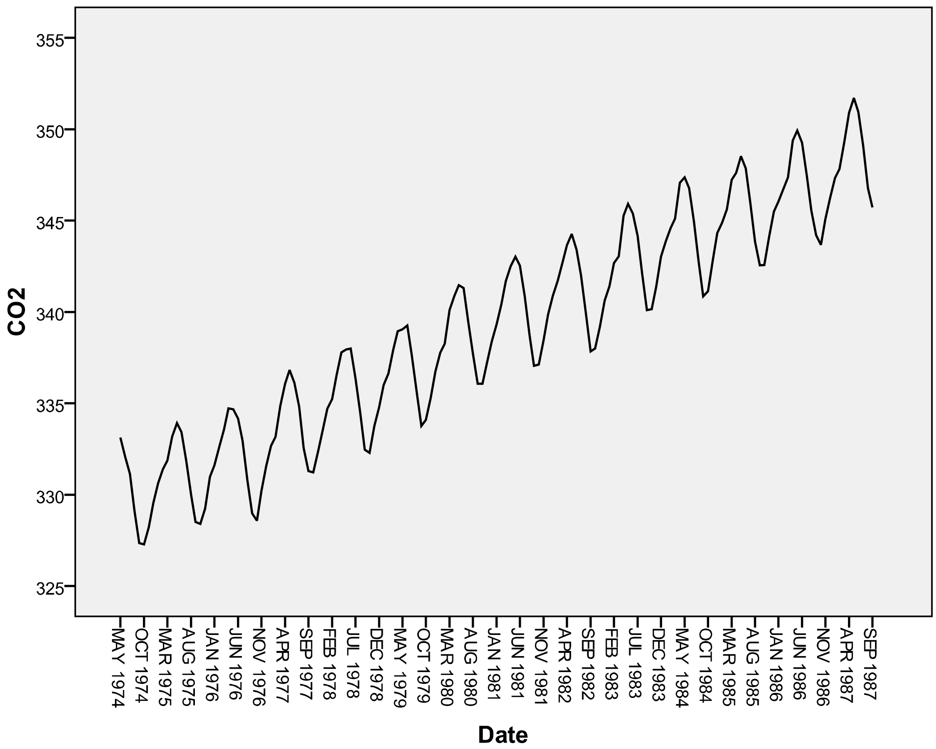
Επίσης, όσον αφορά τα πεδία εφαρμογής της μεθοδολογίας Box- Jenkins, σχετικά με την πρόβλεψη, είδαμε πως αναφέρονται είτε σε βραχυπρόθεσμο είτε σε μεσοπρόθεσμο ορίζοντα. Παράδειγμα αποτελεί η πρόβλεψη για τα κέρδη που μπορεί να επιτύχει μια επιχείρηση, ο έλεγχος των αποθεμάτων κ.α. Επίσης κάτι άλλο που πρέπει να σημειωθεί είναι το ότι η εφαρμογή της απαιτεί μεγάλη προσοχή και πως ο κάθε αναλυτής πρέπει να γνωρίζει πως δεν μπορεί να εφαρμόζεται σε όλες τις χρονοσειρές. Τα αποτελέσματά της βέβαια όπως είναι φανερό είναι πάρα πολύ χρήσιμα στις έρευνες των αναλυτών.

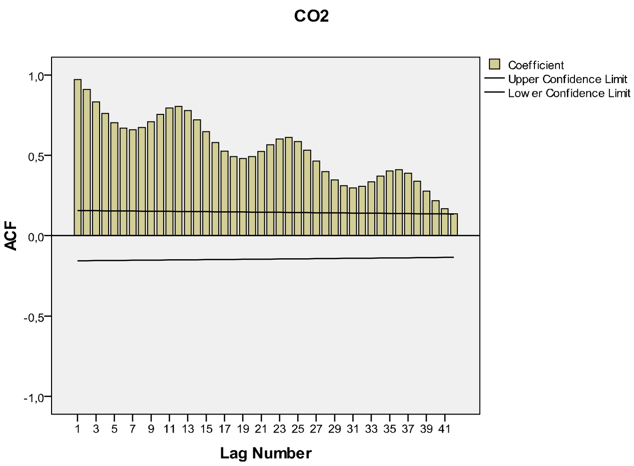
ΚΕΦΑΛΑΙΟ 5: ΕΦΑΡΜΟΓΗ

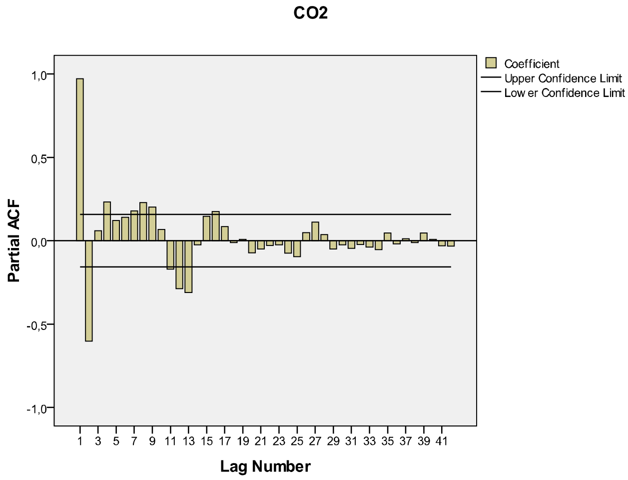
5.1 <u>Εισαγωγή</u>

Στα προηγούμενα κεφάλαια παρουσιάστηκαν μοντέλα ανάλυσης εποχικών χρονοσειρών. Στο κεφάλαιο αυτό θα γίνει ανάλυση ενός εποχικού μοντέλου η οποία βασίζεται σε πραγματικά δεδομένα. Η διαδικασία μοντελοποίησης έγινε με τη χρήση του στατιστικού προγράμματος SPSS 17.0.

Τα δεδομένα που χρησιμοποιούνται στο μοντέλο αφορούν μηνιαίες μέσες συγκεντρώσεις του CO2 στο παρατηρητήριο Mauna Loa από τον Μάϊο του 1974 έως τον Σεπτέμβριο του 1987. Αρχικά θα εξετάσουμε το εποχικό μέρος του μοντέλου και στη συνέχεια το μη εποχικό μέρος.



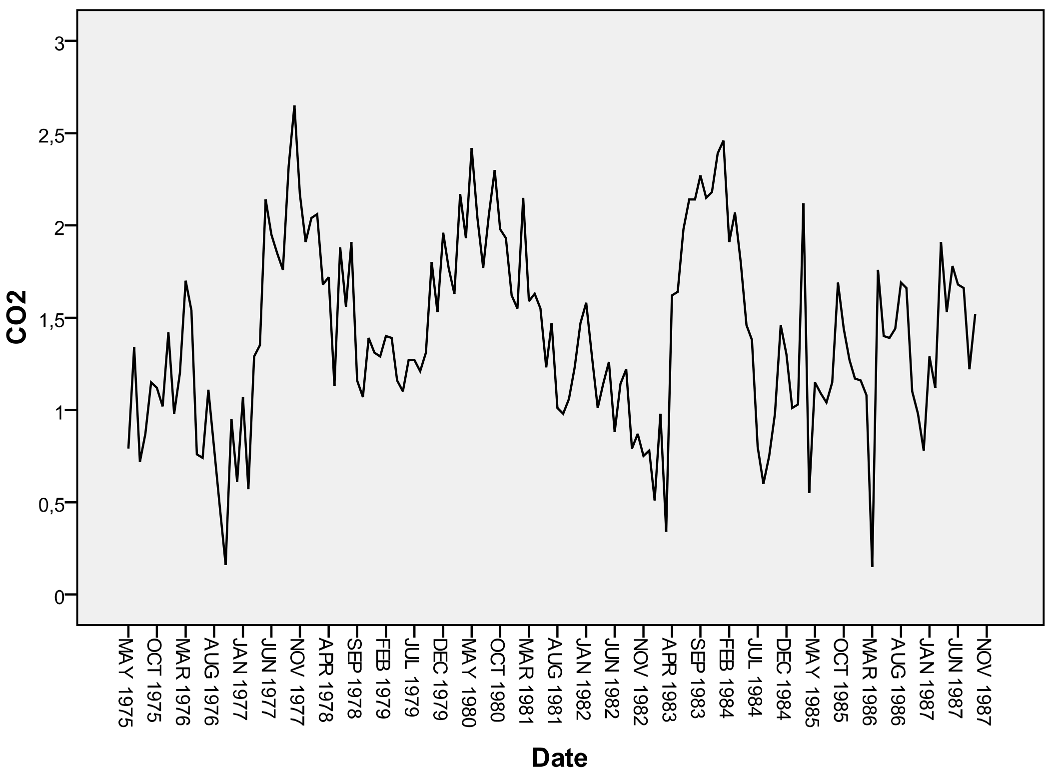




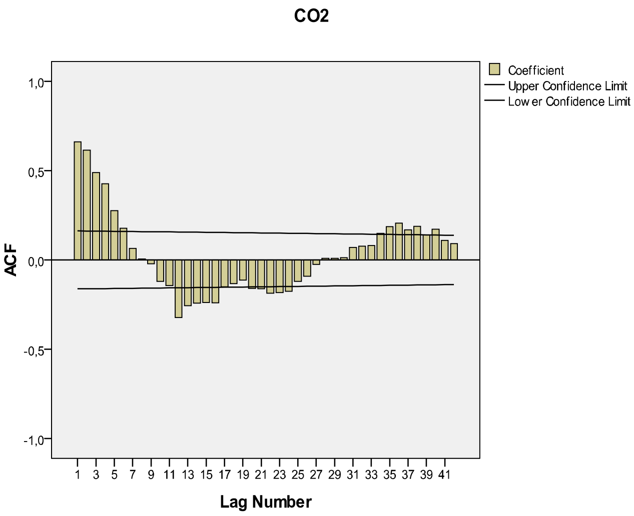
5.2 Ερμηνεία γραφικών παραστάσεων

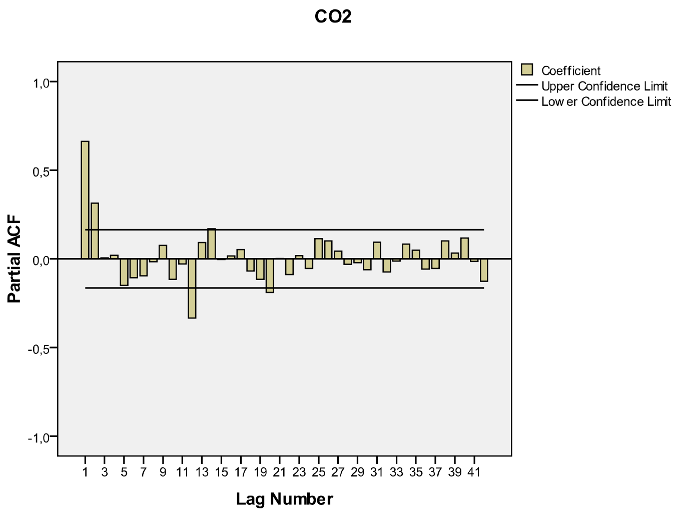
Παρατηρώντας την πρώτη γραφική παράσταση, «*Sequence*», είναι φανερό ότι η χρονοσειρά παρουσιάζει εποχικότητα αλλά φθίνει αργά, επομένως δεν παρουσιάζει στασιμότητα.

Στο διάγραμμα των αυτοσυσχετίσεων «*ACF*» βλέπουμε ότι οι παρατηρήσεις φθίνουν αργά και συγκεκριμένα παρουσιάζονται υψηλές αυτοσυσχετίσεις στις τάξεις 12, 24, 36 (αφού έχουν οριστεί οι τάξεις στις 42). Αντίστοιχα στο διάγραμμα των μερικών αυτοσυσχετίσεων «*PACF*» παρουσιάζεται μια υψηλή υστέρηση στην τάξη 12 και η χρονοσειρά φθίνει αργά, συνεπώς επιβεβαιώνεται η ύπαρξη μη στασιμότητας. Για το λόγο αυτό θα προχωρήσουμε σε εποχική διαφόριση. Τα αποτελέσματα φαίνονται στα διαγράμματα που ακολουθούν.



Transforms: seasonal difference(1, period 12)





Μετά την πρώτη εποχική διαφόριση παρατηρείται στην γραφική παράσταση «*Sequence*», ότι τα όρια της χρονοσειράς μαζεύτηκαν αρκετά. Δηλαδή από 355 που ήταν το ανώτατο όριο της πρώτης γραφικής παράστασης, κατέβηκε στην τάξη του 3 στη δεύτερη γραφική παράσταση. Φαίνεται πια η στασιμότητα της χρονοσειράς.

Επίσης στα διαγράμματα των αυτοσυσχετίσεων «*ACF*», όπως και των μερικών αυτοσυσχετίσεων «*PACF*» παρατηρείται ότι η χρονοσειρά φθίνει γρήγορα πια με μερικές υψηλές υστερήσεις, ιδιαίτερα στο διάγραμμα των αυτοσυσχετίσεων «*ACF*». Η χρονοσειρά είναι στάσιμη πια και θα εφαρμοστεί το εποχικό μοντέλο ARIMA(P,D,Q)s, αντικαθιστώντας P=1 γιατί στο διάγραμμα PACF η δωδέκατη υστέρηση βγαίνει έξω από τα όρια, το D=1 γιατί έχουμε διαφορίσει τη χρονοσειρά μια φορά και Q=3 γιατί στο διάγραμμα ACF οι υστερήσεις των τάξεων 12, 24 και 36 είναι λίγο έξω από τα όρια.

Αντικαθιστώντας τις παραπάνω τιμές το μοντέλο μας θα είναι το εξής:

$$ARIMA(1,1,3)_{12}$$

Τα αποτελέσματα εξετάζονται στους παρακάτω πίνακες:

Model Statistics

Model	Number of Predictors	Model Fit statistics: Stationary R-squared	Ljung-Box Q(18): Statistics	Ljung-Box Q(18): DF	Ljung-Box Q(18): Sig.	Number of Outliers
CO2-Model_1	0	,266	228,674	14	,000	0

ARIMA Model Parameters

					Estimate	SE	t	Sig.
CO2-Model_1	CO2	No Transformation	Constant		1,443	,011	128,908	,000
			AR, Seasonal	Lag 1	-,999	,152	-6,555	,000
			Seasonal Difference		1			
			MA, Seasonal	Lag 1	-,350	6,756	-,052	,959
				Lag 2	,989	8,275	,119	,905
				Lag 3	,355	2,173	,163	,871

Από τον πίνακα «ARIMA Model Parameters» έχουμε τα εξής:

- $\hat{\varphi}_1 = -0,999$ δηλαδή ο συντελεστής του AR_1 Seasonal, με p-value = 0. Συγκρίνοντας το 0 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μικρότερο, 0<0,05. Άρα το φ_1 είναι στατιστικά σημαντικός συντελεστής.
- $\hat{\theta}_1 = -0,350$ δηλαδή ο συντελεστής του MA_1 Seasonal, με p-value = 0,959. Συγκρίνοντας το 0,959 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μεγαλύτερο, 0,959>0,05. Άρα το θ_1 δεν είναι στατιστικά σημαντικός συντελεστής.
- $\hat{\theta}_2 = 0,989$ δηλαδή ο συντελεστής του MA_2 Seasonal, με p-value = 0,905. Συγκρίνοντας το 0,905 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μεγαλύτερο, 0,905>0,05. Άρα το θ_2 δεν είναι στατιστικά σημαντικός συντελεστής.
- $\hat{\theta}_3 = 0,355$ δηλαδή ο συντελεστής του MA_3 Seasonal, με p-value = 0,871. Συγκρίνοντας το 0,871 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μεγαλύτερο, 0,871>0,05. Άρα το θ_3 δεν είναι στατιστικά σημαντικός συντελεστής.
- $\hat{a} = 1,443$ δηλαδή η σταθερά με p-value = 0. Συγκρίνοντας το 0 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μικρότερο, 0<0,05. Άρα το α είναι στατιστικά σημαντικός συντελεστής.

Επαναλαμβάνεται η διαδικασία αφαιρώντας τον συντελεστή MA_1 Seasonal από το μοντέλο επειδή έχει τη μεγαλύτερη τιμή p-value. Αυτό θα γίνει θέτοντας στον συντελεστή Q την τιμή 2 από 3, που ήταν η προηγούμενη.

Επομένως τα αποτελέσματα θα έχουν ως εξής:

Model Statistics

Model	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Statistics	DF	Sig.	
CO2-Model_1	0	,265	239,379	15	,000	0

ARIMA Model Parameters

					Estimate	SE	t	Sig.
CO2-Model_1	CO2	No Transformation	Constant		1,446	,010	139,393	,000
			AR, Seasonal	Lag 1	-,354	,314	-1,129	,261
			Seasonal Difference		1			
			MA, Seasonal	Lag 1	,397	1,530	,259	,796
				Lag 2	,585	,826	,709	,480

Από τον πίνακα «ARIMA Model Parameters» έχουμε τα εξής:

- $\hat{\varphi}_1 = -0,354$ δηλαδή ο συντελεστής του AR_1 Seasonal, με p-value = 0,261. Συγκρίνοντας το 0,261 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μεγαλύτερο, 0,261>0,05. Άρα το φ_1 δεν είναι στατιστικά σημαντικός συντελεστής.
- $\hat{\theta}_1 = -0,397$ δηλαδή ο συντελεστής του MA_1 Seasonal, με p-value = 0,796. Συγκρίνοντας το 0,796 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μεγαλύτερο, 0,796>0,05. Άρα το θ_1 δεν είναι στατιστικά σημαντικός συντελεστής.
- $\hat{\theta}_2 = 0,585$ δηλαδή ο συντελεστής του MA_2 Seasonal, με p-value = 0,480. Συγκρίνοντας το 0,480 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μεγαλύτερο, 0,480>0,05. Άρα το θ_2 δεν είναι στατιστικά σημαντικός συντελεστής.
- α=1446 δηλαδή η σταθερά με p-value = 0. Συγκρίνοντας το 0 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μικρότερο, 0<0,05. Άρα το α είναι στατιστικά σημαντικός συντελεστής.

Επαναλαμβάνεται η διαδικασία αφαιρώντας τον συντελεστή MA_1 Seasonal από το μοντέλο επειδή έχει τη μεγαλύτερη τιμή p-value. Αυτό θα γίνει θέτοντας στον συντελεστή Q την τιμή 1 από 2.

Τα αποτελέσματα φαίνονται στους παρακάτω πίνακες:

Model Statistics

Model	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Statistics	DF	Sig.	
CO2-Model_1	0	,255	240,781	16	,000	0

ARIMA Model Parameters

					Estimate	SE	t	Sig.
CO2-Model_1	CO2	No Transformation	Constant		1,444	,011	132,958	,000
			AR, Seasonal	Lag 1	,231	,122	1,887	,061
			Seasonal Difference		1			
			MA, Seasonal	Lag 1	,989	1,586	,623	,534

Από τον πίνακα «ARIMA Model Parameters» τώρα έχουμε τα εξής:

- $\hat{\varphi}_1 = 0,231$ δηλαδή ο συντελεστής του AR_1 Seasonal, με p-value = 0,061. Συγκρίνοντας το 0,061 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μεγαλύτερο, 0,061>0,05. Άρα το φ_1 δεν είναι στατιστικά σημαντικός συντελεστής.
- $\hat{\theta}_1 = 0,989$ δηλαδή ο συντελεστής του MA_1 Seasonal, με p-value = 0,534. Συγκρίνοντας το 0,534 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μεγαλύτερο, 0,534>0,05. Άρα το θ_1 δεν είναι στατιστικά σημαντικός συντελεστής.
- α =1,444 δηλαδή η σταθερά με p-value = 0. Συγκρίνοντας το 0 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μικρότερο, 0<0,05. Άρα το α είναι στατιστικά σημαντικός συντελεστής.

Επαναλαμβάνεται η διαδικασία για τελευταία φορά αφαιρώντας τον συντελεστή MA_1 Seasonal από το μοντέλο επειδή έχει τη μεγαλύτερη τιμή p-value. Αυτό θα γίνει θέτοντας στον συντελεστή Q την τιμή 0 από 1.

Επομένως τα αποτελέσματα θα έχουν ως εξής:

Model Statistics

		Model Fit statistics	Ljung-Box Q(18)			
Model	Number of Predictors	Stationary R-squared	Statistics	DF	Sig.	Number of Outliers
CO2-Model_1	0	,108	245,963	17	,000	0

ARIMA Model Parameters

				Estimate	SE	t	Sig.
CO2-Model_1	CO2	No Transformation	Constant	1,413	,030	47,069	,000
			AR, Seasonal Lag 1	-,324	,079	-4,107	,000
			Seasonal Difference	1			

Από τον πίνακα «ARIMA Model Parameters» τώρα έχουμε τα εξής:

- $\hat{\phi}_1 = 0,324$ δηλαδή ο συντελεστής του AR_1 Seasonal, με p-value = 0. Συγκρίνοντας το 0 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μικρότερο, 0<0,05. Άρα το φ_1 είναι στατιστικά σημαντικός συντελεστής.
- α =1,413 δηλαδή η σταθερά με p-value = 0. Συγκρίνοντας το 0 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μικρότερο, 0<0,05. Άρα το α είναι στατιστικά σημαντικός συντελεστής.

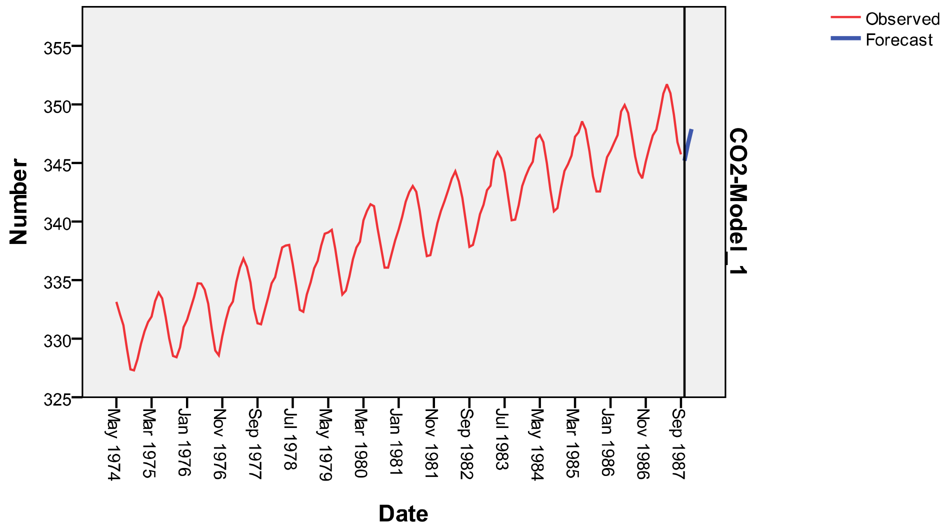
Ο συντελεστής MA_1 Seasonal δεν υπάρχει πια γιατί θέσαμε νωρίτερα την τιμή 0.

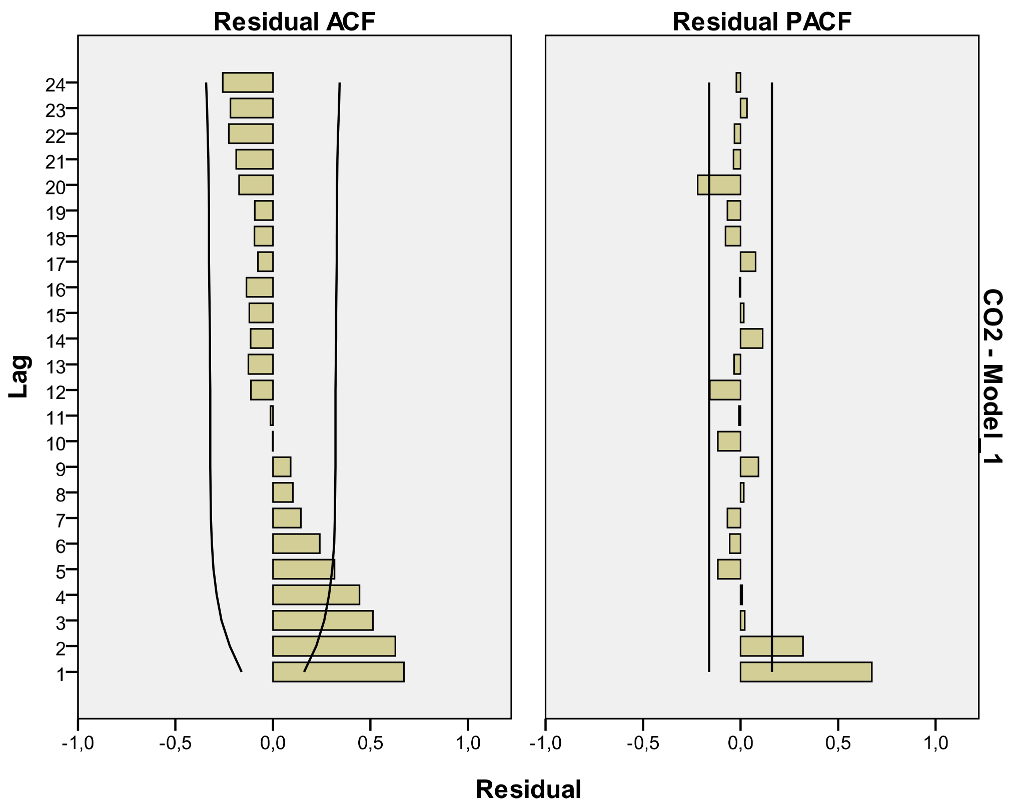
Το μοντέλο θα έχει την εξής μορφή:

$$ARIMA(1,1,0)_{12}$$

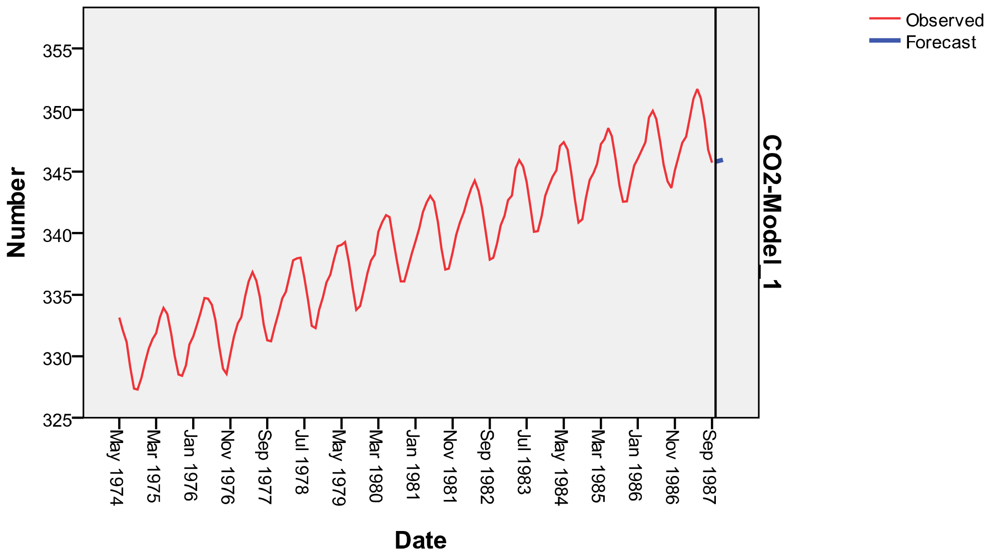
Δηλαδή από το μοντέλο $ARIMA(1,1,3)_{12}$, προχωρήσαμε στο $ARIMA(1,1,2)_{12}$, εν συνεχεία στο $ARIMA(1,1,2)_{12}$ και καταλήξαμε ότι το ιδανικό εποχικό μοντέλο είναι το $ARIMA(1,1,0)_{12}$.

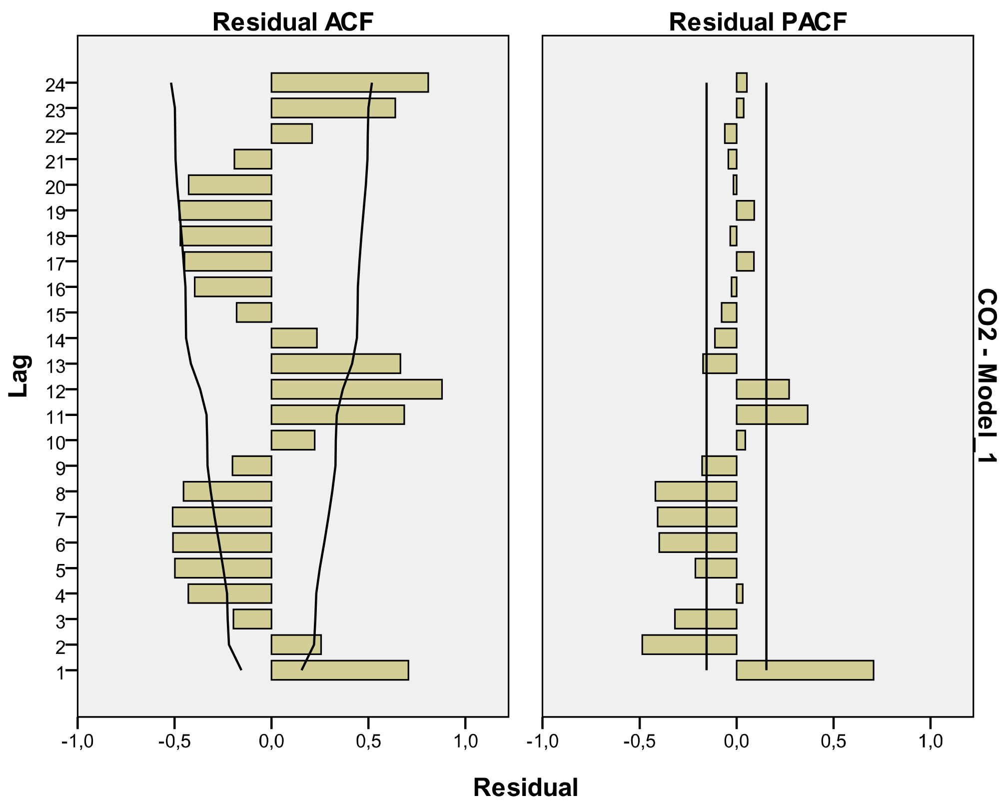
Στη συνέχεια εξετάζονται τα υπόλοιπα.





Από τα διαγράμματα παρατηρείται ότι τα υπόλοιπα του εποχικού μοντέλου δεν είναι κατανεμημένα κανονικά γιατί στα διαγράμματα των αυτοσυσχετίσεων «*ACF*», όπως και των μερικών αυτοσυσχετίσεων «*PACF*» οι πρώτες τιμές είναι εκτός των ορίων. Άρα θα διαφορίσουμε μια φορά τα υπόλοιπα. Τα αποτελέσματα έχουν ως εξής:


Observed
Forecast
Number
355
350
345
340
335
330
325
May 1974
Mar 1975
Jan 1976
Nov 1976
Sep 1977
Jul 1978
May 1979
Mar 1980
Jan 1981
Nov 1981
Sep 1982
Jul 1983
May 1984
Mar 1985
Jan 1986
Nov 1986
Sep 1987
Date
CO2-Model_1

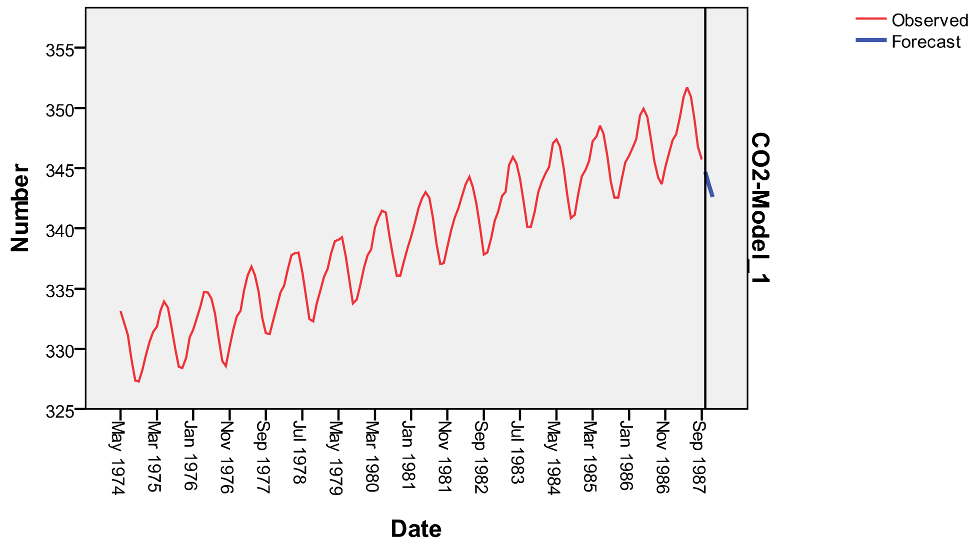

Residual ACF
Residual PACF
Lag
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
-1,0
-0,5
0,0
0,5
1,0
Residual
CO2 - Model_1

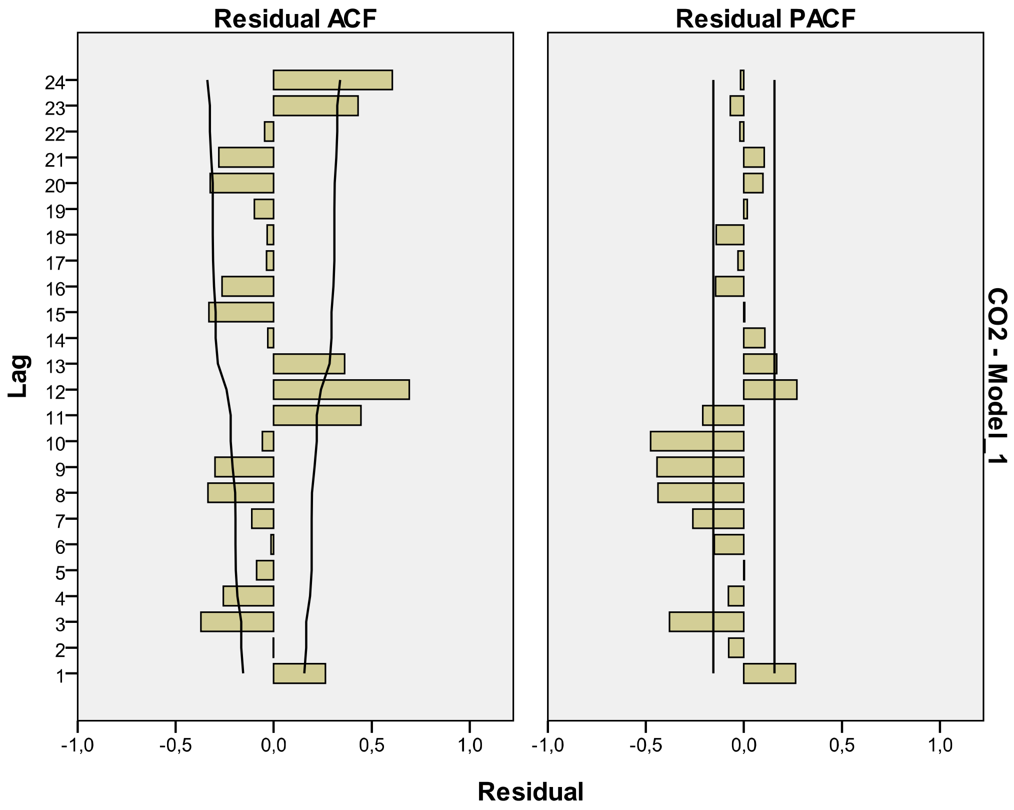
Model Statistics

Model	Number of Predictors	Model Fit statistics		Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Normalized BIC	Statistics	DF	Sig.	
CO2-Model_1	0	1,476E-16	,502	723,813	18	,000	0

Μετά την πρώτη διαφόριση των υπολοίπων παρατηρούμε στα διαγράμματα των αυτοσυσχετίσεων «*ACF*», όπως και των μερικών αυτοσυσχετίσεων «*PACF*» ότι οι τιμές μαζεύτηκαν αρκετά εντός ορίων. Όμως οι πρώτες δύο τιμές παραμένουν εκτός ορίων και το κριτήριο «Normalized BIC» έχει την τιμή 0,502, η οποία είναι αρκετά μεγάλη.

Επομένως θα διαφορίσουμε για δεύτερη φορά τα υπόλοιπα. Τα αποτελέσματα έχουν ως εξής:





Model Statistics

Model	Number of Predictors	Model Fit statistics		Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Normalized BIC	Statistics	DF	Sig.	
CO2-Model_1	0	,000	-,040	255,847	18	,000	0

Μετά την δεύτερη διαφόριση των υπολοίπων παρατηρούμε στα διαγράμματα των αυτοσυσχετίσεων «*ACF*», όπως και των μερικών αυτοσυσχετίσεων «*PACF*» ότι οι περισσότερς τιμές πια βρίσκονται εντός των ορίων. Όμως η πρώτη και η τρίτη τιμή παραμένουν εκτός ορίων ενώ το κριτήριο «Normalized BIC» έχει την τιμή -0,040, που σημαίνει ότι ελαττώθηκε πάρα πολύ σε σχέση με την προηγούμενη τιμή. Επομένως το μη εποχικό μοντέλο θα είναι της μορφής:

ARIMA(p,d,q).

Αντικαθιστώντας p=2 γιατί στο διάγραμμα PACF η δεύτερη και τρίτη τιμή είναι εκτός ορίων, d=2 γιατί έχουμε διαφορίσει τα υπόλοιπα δύο φορές και q=2 γιατί στο διάγραμμα ACF η δεύτερη και τρίτη τιμή είναι εκτός ορίων. Τα αποτελέσματα του μη εποχικού μοντέλου εξετάζονται στους παρακάτω πίνακες:

Model Statistics

Model	Number of Predictors	Model Fit statistics		Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Normalized BIC	Statistics	DF	Sig.	
CO2-Model_1	0	,349	-,317	181,577	14	,000	0

ARIMA Model Parameters

					Estimate	SE	t	Sig.
CO2-Model_1	CO2	No Transformation	Constant		,004	,070	,051	,959
			AR	Lag 1	,976	,020	48,768	,000
				Lag 2	-,983	,017	-57,437	,000
			Difference		2			
			MA	Lag 1	,852	,101	8,408	,000
				Lag 2	-,996	,220	-4,518	,000

Από τον πίνακα «ARIMA Model Parameters» έχουμε τα εξής:

- $\hat{\phi}_1 = 0{,}976$ δηλαδή ο συντελεστής του AR_1, με p-value = 0. Συγκρίνοντας το 0 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μικρότερο, 0<0,05. Άρα το φ_1 είναι στατιστικά σημαντικός συντελεστής.
- $\hat{\phi}_2 = -0{,}983$ δηλαδή ο συντελεστής του AR_2, με p-value = 0. Συγκρίνοντας το 0 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μικρότερο, 0<0,05. Άρα το φ_2 είναι στατιστικά σημαντικός συντελεστής.
- $\hat{\theta}_1 = 0{,}852$ δηλαδή ο συντελεστής του MA_1, με p-value = 0. Συγκρίνοντας το 0 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μικρότερο, 0<0,05. Άρα το θ_1 είναι στατιστικά σημαντικός συντελεστής.

- $\hat{\theta}_2 = -0,996$ δηλαδή ο συντελεστής του MA_2, με p-value = 0. Συγκρίνοντας το 0 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μικρότερο, 0<0,05. Άρα το θ_2 είναι στατιστικά σημαντικός συντελεστής.
- α=0,004 δηλαδή η σταθερά με p-value = 0. Συγκρίνοντας το 0,959 με το 0,05 επίπεδο σημαντικότητας βλέπουμε ότι είναι μεγαλύτερο, 0,959>0,05. Άρα το α δεν είναι στατιστικά σημαντικός συντελεστής.

Όλοι οι συντελεστές είναι στατιστικά σημαντικοί εκτός από την σταθερά α την οποία θα βγάλουμε από τα μοντέλο. Τα αποτελέσματα φαίνονται παρακάτω:

Model Statistics

Model	Number of Predictors	Model Fit statistics		Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Normalized BIC	Statistics	DF	Sig.	
CO2-Model_1	0	,349	-,355	181,067	14	,000	0

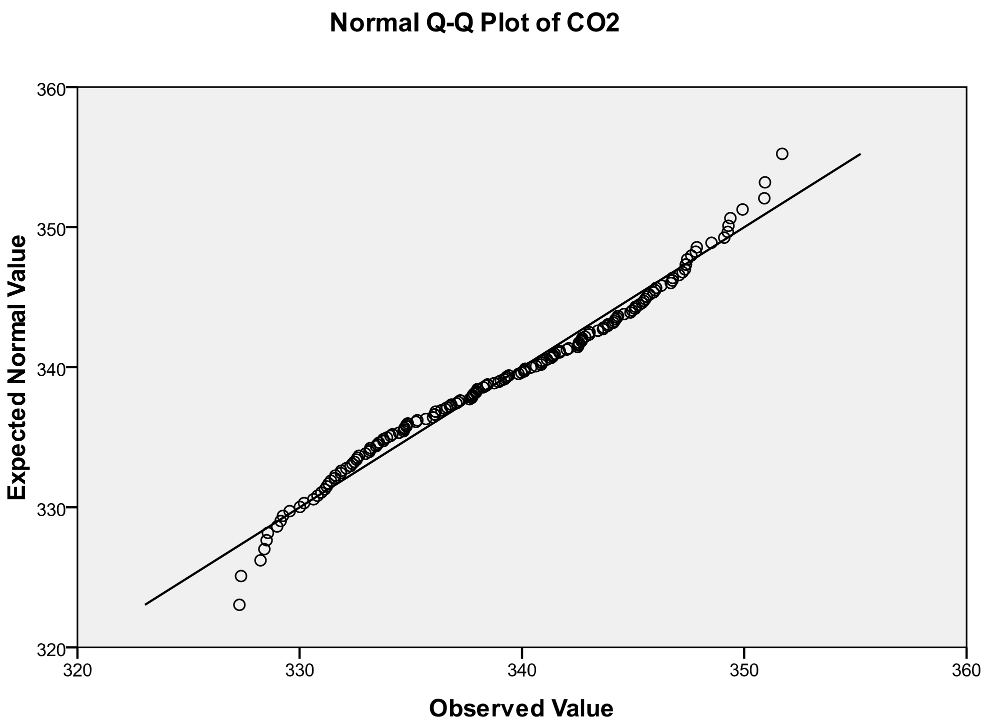
ARIMA Model Parameters

					Estimate	SE	t	Sig.
CO2-Model_1	CO2	No Transformation	AR	Lag 1	,976	,020	48,744	,000
				Lag 2	-,983	,017	-57,179	,000
			Difference		2			
			MA	Lag 1	,851	,070	12,120	,000
				Lag 2	-,993	,140	-7,103	,000

Όλοι οι συντελεστές είναι στατιστικά σημαντικοί πια γιατί τα p-value έχουν τιμή μηδέν. Επομένως το μοντέλο έχει την εξής μορφή:

$$ARIMA(2,2,2)x(1,1,0)_{12}$$

Εξετάζοντας το διάγραμμα Q-Q Plot παρατηρείται ότι τα υπόλοιπα είναι κατανεμημένα κανονικά γιατί τα περισσότερα από αυτά εφάπτονται στη διαγώνιο. Άρα τα δεδομένα μας ακολουθούν την κανονική κατανομή.



ΠΡΟΒΛΕΨΗ

Τα παρακάτω αποτελέσματα είναι οι τιμές τις πρόβλεψης για τους τελευταίους τρεις μήνες του έτους 1987.

YEAR	MONTH	DATE	FIT	LCL	UCL
1987	10	OCT 1987	346	344	347
1987	11	NOV 1987	346	342	349
1987	12	DEC 1987	345	339	351

Τον Οκτώβριο του 1987 η τιμή της πρόβλεψης είναι 346. Το κατώτερο όριο που ανήκει στο 95% διάστημα εμπιστοσύνης είναι 344 και το ανώτερο 347.

Τον Σεπτέμβριο του 1987 η τιμή της πρόβλεψης είναι 346. Το κατώτερο όριο που ανήκει στο 95% διάστημα εμπιστοσύνης είναι 342 και το ανώτερο 349.

Τον Δεκέμβριο του 1987 η τιμή της πρόβλεψης είναι 345. Το κατώτερο όριο που ανήκει στο 95% διάστημα εμπιστοσύνης είναι 339 και το ανώτερο 351.

Παρατηρούμε ότι οι τιμές της πρόβλεψης του μονοξειδίου του άνθρακα CO2 σε σχέση με τις πραγματικές τιμές των προηγούμενων μηνών θα παραμείνουν σχεδόν στα ίδια επίπεδα με μια μικρή αύξηση. Δηλαδή με το πέρασμα των χρόνων οι τιμές του μονοξειδίου του άνθρακα θα αυξάνονται.

ΚΕΦΑΛΑΙΟ 6: ΣΥΜΠΕΡΑΣΜΑ

Η συγκεκριμένη εργασία είχε σαν σκοπό μια πρώτη προσέγγιση στην έννοια των χρονοσειρών, των παραμέτρων τους και στον τρόπο εφαρμογής τους. Μέσω της ανάλυσης των χρονοσειρών, η οποία μπορεί να γίνει είτε στο πεδίο του χρόνου είτε στο πεδίο των συχνοτήτων, έγινε προσπάθεια να μελετηθεί η συμπεριφορά τους, χρησιμοποιώντας τις βασικές μεθόδους τους. Είναι αντιληπτό λοιπόν πως οι χρονοσειρές διαφέρουν ως προς την πολυπλοκότητα, τους τομείς εφαρμογής και την ακρίβεια των προβλέψεών τους.

Αρχικά παρουσιάστηκαν οι τέσσερις διαφορετικές συνιστώσες οι οποίες εξηγούν τη συμπεριφορά των τιμών χρονοσειρών,(τάση, κυκλική συνιστώσα, εποχική συνιστώσα και τυχαία συνιστώσα), καθώς κάποιες περιγραφικές μέθοδοι χρονοσειρών οι οποίες θεωρούνται απλές και δεν αναφέρονται σε στοχαστικά μοντέλα. Από αυτές άλλες (μέθοδοι εξομάλυνσης) χρησιμοποιούνται για βραχυπρόθεσμες προβλέψεις και άλλες (μέθοδοι ανάλυσης και διαχωρισμού των χρονοσειρών), για μακροπρόθεσμες.

Μελετώντας τις χρονοσειρές δημιουργούνται διάφορα ερωτήματα γύρω από αυτές:

- Είναι τυχαία τα δεδομένα;
- Είναι τα δεδομένα μη-σταθερά;
- Είναι τα δεδομένα εποχικά;

Οι απαντήσεις στα παραπάνω ερωτήματα μας δίνουν τη δυνατότητα καλύτερων προβλέψεων Κάτι επίσης, που δεν μπορεί να αποφευχθεί στην προσπάθεια για πρόβλεψη των μελλοντικών καταστάσεων είναι το σφάλμα, ωστόσο γίνονται μεγάλες προσπάθειες από τους αναλυτές για να επιτύχουν όσο το δυνατό μικρότερη τιμή για αυτό.

Οι μέθοδοι ανάλυσης στο πεδίο του χρόνου αναφέρονται στα λεγόμενα μοντέλα ARIMA τα οποία προτάθηκαν από τους Box και Jenkins με απώτερο σκοπό την εύρεση του καταλληλότερου μοντέλου στη διαδικασία της πρόβλεψης. Η προσέγγιση Box – Jenkins αποτελεί ένα πολύ χρήσιμο και ευέλικτο εργαλείο στην κατασκευή υποδειγμάτων ARIMA αφού αποτελεί μια πού καλή μέθοδο πρόβλεψης χρονοσειρών από τις τρέχουσες και τις ιστορικές τιμές τους.

Σε αντίθεση με τις περιγραφικές μεθόδους όπου μελετώνται η συμπεριφορά της χρονοσειράς κατά το παρελθόν, σ'αυτά τα μοντέλα, η πρόβλεψη γίνεται μέσω της εξέτασης του συνδυασμού, τόσο των παλαιότερων τιμών της χρονοσειράς, όσο και των εμφανισθέντων προτύπων που παρουσιάζονται.

Στη συνέχεια έγινε εκτενέστερη αναφορά στην έννοια της εποχικότητας, η οποία εντοπίζεται στις χρονοσειρές, και την προσπάθεια εξεύρεσης ή απαλοιφής της, με τη βοήθεια των εποχικών δεικτών μέσω κάποιων παραδειγμάτων. Επίσης αναλύθηκε περισσότερο η έννοια των προβλέψεων καθώς και η μεθοδολογία των Box και Jenkins.

Κάτι που παρατηρήθηκε επίσης είναι πως οι μέθοδοι των χρονολογικών σειρών είναι περισσότερο αποτελεσματικές όταν το περιβάλλον παραμένει σταθερό. Αν το περιβάλλον δεν παραμένει σταθερό, το μέλλον δεν θα μοιάζει με το παρελθόν. Επομένως, οι μέθοδοι προβολής χρονολογικών σειρών είναι περισσότερο αποτελεσματικές για βραχυπρόθεσμες προβλέψεις.

Καταλήγοντας πρέπει να σημειωθεί πως ανάλογα με την εμπειρία την ικανότητα και τις δυνατότητες κάθε αναλυτή, αυτός επιλέγει και το ποια μέθοδο θα επιλέξει με στόχο τα καλύτερα και τα πιο ικανοποιητικά αποτελέσματα στις έρευνές του.

<u>ΒΙΒΛΙΟΓΡΑΦΙΑ</u>

1. ΑΝΔΡΙΚΟΠΟΥΛΟΣ ΑΝΔΡΕΑΣ Α. «Οικονομετρία Θεωρία και Εμπειρικές Εφαρμογές 1998».
2. ΒΑΪΔΑΝΗΣ ΜΙΧΑΛΗΣ «Αρχές Διοίκησης και Οργάνωσης Παραγωγής 2005 ».
3. ΔΗΜΕΛΗ ΣΟΦΙΑ «Σύχρονες Μέθοδοι Ανάλυσης Χρονοσειρών », Εκδόσεις Κριτική 2002
4. ΕΚΔΟΣΕΙΣ ΤΖΙΟΛΑ «Στατιστική 3^{η} Έκδοση ».
5. ΚΑΡΥΩΤΗ ΒΑΣΙΛΙΚΗ «Σημειώσεις από το μάθημα Τεχνικές Προβλέψεων 2004 ».
6. ΚΙΝΤΗΣ Α.ΑΝΔΡΕΑΣ «Στατιστικές και Οικονομετρικές Μέθοδοι »Εκδόσεις Gutenberg.
7. Ι.ΠΑΝΑΡΕΤΟΥ & Ε. ΞΕΚΑΛΑΚΗ , 1998
8. ΜΠΟΡΑ- ΣΕΝΤΑ ΕΥΘΥΜΙΑ ΛΕΚΤΟΡΑΣ Α.Π.Θ , Χ ΜΩΥΣΙΑΔΗΣ ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ Α.Π.Θ «Εφαρμοσμένη στατιστική, Πολλαπλή Παλινδρόμηση, Ανάλυση Διασποράς »Εκδόσεις Gutenberg 1997.
9. ΞΕΝΑΚΗΣ Σ. ΑΝΔΡΕΑΣ « Ανάλυση Χρονολογικών Σειρών και Προβλέψεις-Ανάλυση Στο Πεδιο του Χρόνου 1998».
10. ΠΑΠΑΔΗΜΑΣ ΌΘΩΝΑΣ M .Sc Μαθηματικού Στατιστικού «Στατιστική ΙΙ», Μακεδονικές Εκδόσεις 1994.

11. Box, G.E.P. and Pierce, D.A. (1970). «Distribution of residual autocorrelations in ARIMA models» Oakland, California, Holden day
12. Jarrett Jeffrey «Μέθοδοι προβλέψεων Για Οικονομικές – Επιχειρηματικές Αποφάσεις», Εκδόσεις Gutenberg 2002.
13. Karioti Vassiliki Department of Mathematics «Master Time Series Analysis of Peak Expiratory Flow», September 1997.
14. Box, G.E.P. and Jenkins, G.M. (1976). «Time Series Analysis, Forecasting and Control» Oakland, California, Holden day.

ΠΑΡΑΡΤΗΜΑ

Δεδομένα συγκέντρωσης του CO2 στο παρατηρητήριο Mauna Loa από τον Μάϊο του 1974 έως τον Σεπτέμβριο του 1987.
(Ο πίνακας διαβάζεται κάθετα).

CO2						
333,13	334,72	336,54	338,26	340,9	343,02	344,11
332,09	334,68	337,79	340,1	341,7	343,87	345,49
331,13	334,17	337,95	340,88	342,7	344,59	346,04
329,14	332,96	338	341,47	343,65	345,11	346,7
327,36	330,8	336,37	341,31	344,28	347,07	347,38
327,29	328,98	334,47	339,41	343,42	347,38	349,38
328,23	328,57	332,46	337,74	342,02	346,78	349,93
329,55	330,2	332,29	336,07	339,97	344,96	349,26
330,62	331,58	333,76	336,07	337,84	342,71	347,44
331,4	322,67	334,8	337,22	338	340,86	345,55
331,87	333,17	336	338,38	339,2	341,13	344,21
333,18	334,86	336,63	339,32	340,63	342,84	343,67
333,92	336,07	337,93	340,41	341,41	344,32	345,09
333,43	336,82	338,95	341,69	342,68	344,88	346,27
331,85	336,12	339,05	342,51	343,04	345,62	347,33
330,01	334,81	339,27	343,02	345,27	347,23	347,82
328,51	332,56	337,64	342,54	345,92	347,62	349,29
328,41	331,3	335,68	340,88	345,4	348,53	350,91
329,25	331,22	333,77	338,75	344,16	347,87	351,71
330,97	332,37	334,09	337,05	342,11	346	350,94
331,6	333,49	335,29	337,13	340,11	343,86	349,1
332,6	334,71	336,76	338,45	340,15	342,55	346,77
333,57	335,23	337,77	339,85	341,38	342,57	345,73

ΠΗΓΗ: http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc4411.htm