



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ
ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

ΜΟΝΤΕΛΑ ΑΝΑΛΥΣΗΣ ΕΠΟΧΙΚΩΝ ΧΡΟΝΟΣΕΙΡΩΝ ΣΤΗΝ ΚΙΝΗΣΗ ΦΑΡΜΑΚΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Λεωνίδας Πετρίδης

Επιβλέπων : Γεώργιος Κοκολάκης
Καθηγητής ΕΜΠ

Αθήνα, Μάρτιος 2013



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ
ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

ΜΟΝΤΕΛΑ ΑΝΑΛΥΣΗΣ ΕΠΟΧΙΚΩΝ ΧΡΟΝΟΣΕΙΡΩΝ ΣΤΗΝ ΚΙΝΗΣΗ ΦΑΡΜΑΚΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Λεωνίδας Πετρίδης

Επιβλέπων : Γεώργιος Κοκολάκης
Καθηγητής ΕΜΠ

Εγκρίθηκε από την τριμελή επιτροπή τη 8^η Μαρτίου 2013.

Γεώργιος Κοκολάκης
Καθηγητής ΕΜΠ

Ιωάννης Σπηλιώτης
Αναπληρωτής Καθηγητής ΕΜΠ

Δημήτριος Φουσκάκης
Επίκουρος Καθηγητής ΕΜΠ

Αθήνα, Μάρτιος 2013

ΠΕΡΙΕΧΟΜΕΝΑ

1	Εισαγωγή.....	1
2	Προβλέψεις	2
2.1	Χαρακτηριστικά Προβλέψεων	3
2.2	Αξιολόγηση των μεθόδων πρόβλεψης.....	6
3	Αποθέματα	9
3.1	Έλεγχος αποθεμάτων με γνωστή (καθορισμένη) ζήτηση.....	11
3.2	Έλεγχος αποθεμάτων με άγνωστη (στοχαστική) ζήτηση.....	13
3.3	Συστήματα πολλαπλών προϊόντων – Ανάλυση ABC.....	15
3.4	Σύνδεση προβλέψεων με τον έλεγχο αποθεμάτων	16
4	Χρονοσειρές.....	17
4.1	Μερικά απλά μοντέλα χρονοσειρών.....	19
4.2	Στασιμότητα.....	20
4.3	Τάση και εποχικότητα.....	21
4.4	Εκτίμηση της τάσης.....	23
4.5	Εκτίμηση της τάσης και της εποχικότητας	25
4.6	Αυτοπαλινδρομική Διαδικασία Κινητού Μέσου ARMA(p,q)	26
4.7	Κριτήρια επιλογής μοντέλου	27
4.8	Έλεγχος των υπολοίπων	28
4.9	Επιλογή μοντέλου	30
4.10	Προβλέψεις για στάσιμες χρονοσειρές.....	31
5	Εφαρμογή.....	34
5.1	Zantac.....	34
5.2	Mesulid	41
5.3	Glucophage	48
5.4	Εμβόλια Ιών.....	54
6	Συμπεράσματα	60
7	Παράρτημα.....	61
	Βιβλιογραφία.....	63

1 ΕΙΣΑΓΩΓΗ

Σε πολλούς κλάδους της οικονομικής και κοινωνικής καθημερινότητας εμφανίζονται μετρήσιμα μεγέθη που προσφέρουν πολύτιμη βοήθεια στην εξήγηση των φαινομένων προκαλούν. Αυτά τα μεγέθη καλούμαστε να τα ανακαλύψουμε με την παρατήρηση πείραμα και την συστηματική ανάλυση, ώστε να αποκτήσουμε γνώση και να εξαγαθούμε συμπεράσματα. Πολλές φορές είναι χρήσιμο να μελετήσουμε τη διαχρονική εξέλιξη ενός φαινομένου και, βασιζόμενοι σε αυτή τη μελέτη, να αποκτήσουμε μια εικόνα για την πιθανή εξέλιξή του στο μέλλον. Για αυτόν το σκοπό έχει αναπτυχθεί πλήθος μεθόδων, οι οποίες, ανάλογα με την περίπτωση, είναι δυνατό να μας προσφέρουν την κατάλληλη βοήθεια για να πετύχουμε το στόχο μας.

Η χρήση των μεθόδων αυτών στο χώρο της οικονομίας και των επιχειρήσεων είναι εξαιρετικά χρήσιμη.

Σε αυτήν την εργασία θα δούμε τρόπους και μεθόδους πρόβλεψης της ζήτησης χρησιμοποιώντας ιστορικά στοιχεία πωλήσεων, προκειμένου να διασφαλίσουμε τη διαθεσιμότητα των προϊόντων προς πώληση. Η εφαρμογή θα γίνει στο χώρο των φαρμάκων, με στοιχεία πωλήσεων από φαρμακαποθήκες.

Είναι προφανές ότι η επιχείρηση πρέπει να αναζητήσει τον τρόπο με τον οποίο θα ελαχιστοποιήσει τα αποθέματα στο βαθμό που διασφαλίζει και τη διαθεσιμότητά τους σε κάποιο αποδεκτό επίπεδο εξυπηρέτησης (service level), αποφεύγοντας τόσο την τήρηση υπερβολικού αποθέματος (overstocking), όσο και την έλλειψη αποθέματος (stock-out).

Το πρόβλημα αυτό έχει πολλές παραμέτρους που εξαρτώνται από τη φύση των προϊόντων, τον τρόπο διάθεσής τους, την πηγή τους, τη συχνότητα ανανέωσης των αποθεμάτων, κ.α. Για παράδειγμα, μία επιχείρηση εμπορεύεται ένα προϊόν το οποίο αγοράζει από τον προμηθευτή, τοποθετώντας μία παραγγελία αγοράς, το παραλαμβάνει μετά από κάποιο χρονικό διάστημα, το αποθηκεύει και ύστερα το διαθέτει στους πελάτες της. Αν η επιχείρηση γνωρίζει εκ των προτέρων την ποσότητα του προϊόντος που θα διαθέσει στους πελάτες της, τότε θα παραγγείλει αυτήν την ποσότητα από τον προμηθευτή. Αυτό όμως δεν ισχύει πάντα. Συγκεκριμένα, στις περισσότερες επιχειρήσεις αυτή η ποσότητα, η ζήτηση δηλαδή, είναι άγνωστη. Συνεπώς πρέπει να κάνει κάποιου είδους πρόβλεψη για τη ζήτηση, ώστε να είναι σε θέση να την καλύψει με τον αποδοτικότερο τρόπο.

Είναι φανερό ότι ο έλεγχος αποθεμάτων είναι στενά συνδεδεμένος με τις προβλέψεις για τη ζήτηση. Παρακάτω παραθέτουμε μια εισαγωγή στις Προβλέψεις και στη συνέχεια κάποια βασικά στοιχεία από τη Θεωρία Ελέγχου Αποθεμάτων.

Στη βιβλιογραφία αναφέρονται πολλά διαφορετικά συστήματα προβλέψεων για τη ζήτηση, με εφαρμογή στον έλεγχο αποθεμάτων. Κάποια από αυτά είναι αρκετά απλά στη χρήση τους, αλλά μερικές φορές αποδεικνύονται ανεπαρκή, ειδικότερα σε περιπτώσεις στις οποίες η ακρίβεια στην πρόβλεψη είναι ζωτικής σημασίας. Για το λόγο αυτό έχουν αναπτυχθεί πολυπλοκότερα συστήματα προβλέψεων που απαιτούν γνώσεις Στατιστικής

και Ανάλυσης Χρονοσειρών. Θα αναλύσουμε στη συνέχεια της εργασίας μία από τις μεθόδους αυτές και θα δούμε εφαρμογές της στην πρόβλεψη της ζήτησης για προϊόντα, χρησιμοποιώντας πραγματικά στοιχεία από την αγορά του φαρμάκου. Η μέθοδος πρόβλεψης με χρήση μοντέλων ARMA είναι μια ισχυρή και ακριβής διαδικασία που, κάτω από τις κατάλληλες συνθήκες μπορεί να δώσει αξιόπιστες προβλέψεις για χρήση στον έλεγχο των αποθεμάτων μιας επιχείρησης. Για την περιγραφή τέτοιων μεθόδων παραθέτουμε επίσης μια εισαγωγή στην Ανάλυση Χρονοσειρών και στα μοντέλα ARMA(p,q) καθώς και μεθόδους αντιμετώπισης της εποχικότητας.

2 ΠΡΟΒΛΕΨΕΙΣ

Είναι γεγονός ότι στην παγκόσμια οικονομία ότι εναλλάσσονται οικονομικοί κύκλοι – περίοδοι ανάπτυξης όπου οι δείκτες ανεργίας πέφτουν και επικρατεί γενικά κλίμα ευφορίας και ευμάρειας, ακολουθούμενες από περιόδους ύφεσης όπου ανεβαίνει η ανεργία και επικρατεί κλίμα ανέχειας και αβεβαιότητας.

Η διαχείριση των αποθεμάτων παίζει σημαντικό ρόλο στους κύκλους αυτούς. Κατά καιρούς πολλοί οικονομολόγοι προσπάθησαν να κατασκευάσουν μοντέλα που να περιγράφουν επαρκώς αυτούς τους οικονομικούς κύκλους, αλλά μέχρι στιγμής δεν έχει κατασκευαστεί ένα μοντέλο για όλες ανεξαιρέτως τις περιπτώσεις. Ένα μοντέλο δηλαδή που να εφαρμόζεται κάθε στιγμή και κάτω από όλες τις συνθήκες.

Παρόλα αυτά, είναι εμφανές ότι ενώ κάθε οικονομικός κύκλος είναι διαφορετικός από τους προηγούμενους, ειδικά σε ότι αφορά τη χρονική στιγμή, διάρκεια και έντασή του, οι κύκλοι συνεχίζονται και μπορεί κανείς να παρατηρήσει ορισμένους κοινούς παράγοντες.

Συχνά παρατηρείται μία χρονική υστέρηση μεταξύ ενός γεγονότος και της στιγμής όπου αυτό το γεγονός γίνεται αντιληπτό. Αυτή η υστέρηση είναι ο κύριος λόγος για τον οποίο η πρόβλεψη και ο προγραμματισμός είναι χρήσιμοι. Αν η υστέρηση είναι πολύ μικρή ή κοντά στο μηδέν, τότε δεν υπάρχει λόγος για προγραμματισμό. Αν η υστέρηση είναι αρκετά μεγάλη και, ταυτόχρονα το αποτέλεσμα του τελικού γεγονότος εξαρτάται από προβλέψιμους παράγοντες, τότε ο προγραμματισμός μπορεί να διαδραματίσει σημαντικό ρόλο.

Υπάρχουν πολλές απόψεις ως προς τη χρησιμότητα των προβλέψεων. Κάποιος μπορεί να αμφιβάλλει για την εγκυρότητα που μπορεί να έχει μια πρόβλεψη μέσα σε ένα μεταβλητό και αβέβαιο περιβάλλον, και κατά συνέπεια να ισχυριστεί ότι μία πρόβλεψη δε θα ήταν χρήσιμη. Ωστόσο θα πρέπει να αναγνωριστεί πως έχουν γίνει τεράστια βήματα προς τα εμπρός στον τομέα των προβλέψεων τους τελευταίους αιώνες. Υπάρχει πληθώρα φαινομένων των οποίων τα αποτελέσματα μπορούν πλέον να προβλεφθούν με ακρίβεια, κάτι που ίσως φάνταζε αδύνατο μερικούς αιώνες, ή και δεκαετίες πριν. Π.χ. ο καιρός, οι τροχιές των πλανητών, η ταχύτητα ενός αντικειμένου κ.α..

Μία σημαντική πτυχή των προβλέψεων που πρέπει να έχει κανείς υπόψη είναι ο διαχωρισμός μεταξύ ανεξέλεγκτων εξωτερικών γεγονότων, όπως η παγκόσμια οικονομία,

οι κυβερνήσεις, πελάτες, ανταγωνιστές κ.α., και ελεγχόμενα εσωτερικά γεγονότα όπως η διαφήμιση, η απόφαση για παραγωγή ή για αγορά, κ.α.. Η επιτυχία μιας επιχείρησης εξαρτάται και από τους δύο τύπους γεγονότων, αλλά οι προβλέψεις αφορούν άμεσα το πρώτο, ενώ η λήψη αποφάσεων αφορά το δεύτερο. Ο προγραμματισμός είναι ο σύνδεσμος μεταξύ των δύο.

Η αναγνώριση του ρόλου των προβλέψεων και ο σωστός χειρισμός και αξιοποίησή τους στο περιβάλλον των στρατηγικών αποφάσεων είναι τόσο σημαντική όσο η επιλογή της ίδιας της μεθόδου πρόβλεψης. Υπάρχουν πολλές διαθέσιμες μέθοδοι προβλέψεων, που χρησιμοποιούνται στις σύγχρονες επιχειρήσεις. Εκτείνονται από τις απλές (*naive*), μεθόδους στις οποίες η πιο πρόσφατη παρατήρηση χρησιμοποιείται ως πρόβλεψη για την επόμενη χρονική περίοδο, μέχρι τις πολύπλοκες προσεγγίσεις που χρησιμοποιούν νευρωνικά δίκτυα και οικονομετρικά συστήματα ταυτόχρονων εξισώσεων.

Όπως αναφέρθηκε προηγουμένως, οι προβλέψεις παίζουν σημαντικό ρόλο στη διαδικασία λήψης αποφάσεων σε μια επιχείρηση. Η επιχείρηση θέτει στόχους, προσπαθεί να προβλέψει περιβαλλοντικούς παράγοντες και ύστερα επιλέγει να κάνει εκείνες τις ενέργειες που πιστεύει ότι θα έχουν ως αποτέλεσμα την επίτευξη των στόχων της. Η ανάγκη για ακριβείς προβλέψεις γίνεται μεγαλύτερη, όσο η επιχείρηση προσπαθεί να μειώσει την εξάρτησή της από τυχαία γεγονότα και όσο επιθυμεί να εισαγάγει επιστημονικές διαδικασίες στην αντιμετώπιση αυτής της τυχαιότητας.

Εφόσον κάθε τομέας μιας επιχείρησης συνδέεται άμεσα ή έμμεσα με όλους τους άλλους, μία άστοχη πρόβλεψη είναι σε θέση να επηρεάσει αρνητικά ολόκληρη την επιχείρηση. Η μεγάλη εξάρτηση μεταξύ των προβλέψεων που προέρχονται από όλους τους τομείς μιας επιχείρησης δεν πρέπει να παραλείπεται, αν θέλουμε το σύστημα προβλέψεων που εφαρμόζουμε να είναι αποτελεσματικό. Για παράδειγμα, ένα λάθος στην πρόβλεψη των πωλήσεων μπορεί να επηρεάσει τον ισολογισμό, τα λειτουργικά έξοδα, τη ρευστότητα, τις τιμές, τα αποθέματα, κ.ά..

2.1 Χαρακτηριστικά Προβλέψεων

Συχνά οι άνθρωποι που δεν είναι εξοικειωμένοι με ποσοτικές μεθόδους προβλέψεων πιστεύουν ότι το παρελθόν δεν είναι σε θέση να περιγράψει και να εξηγήσει το μέλλον με ακρίβεια επειδή τα πάντα αλλάζουν συνεχώς. Αφού όμως εξοικειωθούν με μεθόδους προβλέψεων, χειρισμό δεδομένων και τη στατιστική, είναι σε θέση να αναγνωρίσουν ότι ενώ τίποτα δεν παραμένει ακριβώς το ίδιο κατά το πέρασμα του χρόνου, κάποιες δομές που υπάρχουν στα ιστορικά δεδομένα, επαναλαμβάνονται κατά μία έννοια. Εφαρμόζοντας κάποιος την κατάλληλη μέθοδο μπορεί να αναγνωρίσει την εξάρτηση ή τη σχέση μεταξύ της μεταβλητής που πρόκειται να προβλεφθεί και του ίδιου του χρόνου (ή άλλων μεταβλητών), ώστε να καταλήξει σε μία ικανοποιητική πρόβλεψη.

Ιδιότητες προβλέψεων

- Οι προβλέψεις σχεδόν πάντοτε είναι ανακριβείς
- Μία καλή πρόβλεψη εμπεριέχει κάποιο μέτρο για το σφάλμα
- Η πρόβλεψη αθροιστικών στοιχείων είναι γενικά προτιμότερη από την πρόβλεψη μεμονωμένων μεγεθών
- Οι πιο μακροπρόθεσμες προβλέψεις τείνουν να είναι λιγότερο ακριβείς

Ποιοτικές (υποκειμενικές) μέθοδοι προβλέψεων

Αναφέρονται κυρίως σε μεθόδους που αντλούν πληροφορίες από άτομα ή ομάδες ατόμων. Οι πιο γνωστές από αυτές είναι:

- Υποκειμενικές προβλέψεις από πωλητές ή στελέχη της αγοράς
- Ερωτηματολόγια πελατών
- Απόψεις διευθυντικών στελεχών
- Η μέθοδος του μαντείου

Ποσοτικές (αντικειμενικές) μέθοδοι προβλέψεων

Υπάρχουν δύο διαφορετικές οικογένειες ποσοτικών μοντέλων πρόβλεψης: τα επεξηγηματικά μοντέλα και τα μοντέλα χρονοσειρών.

Επεξηγηματικά μοντέλα

Τα επεξηγηματικά μοντέλα υποθέτουν ότι η μεταβλητή για την οποία θα γίνει η πρόβλεψη παρουσιάζει μια σχέση εξάρτησης με μία ή περισσότερες ανεξάρτητες μεταβλητές. Για παράδειγμα,

$$ΑΕΠ = f(\text{νομισματικές πολιτικές, φορολογία, πληθωρισμός, εισαγωγές, εξαγωγές, σφάλμα})$$

Αξίζει να σημειωθεί ότι η σχέση δεν είναι ακριβής. Πάντα θα υπάρχουν μεταβολές στο ΑΕΠ που δεν μπορούν να εξηγηθούν από τις μεταβλητές του μοντέλου και συνεπώς κάποιο τμήμα του ΑΕΠ μεταβάλλεται απρόβλεπτα. Ο όρος «σφάλματα» εμπεριέχει αυτές τις τυχαίες μεταβολές, πέρα από τις μεταβλητές του μοντέλου που επηρεάζουν την τιμή του ΑΕΠ.

Τα επεξηγηματικά μοντέλα μπορούν να εφαρμοστούν σε πλήθος συστημάτων όπως η εθνική οικονομία, η αγορά, το νοικοκυριό. Ο σκοπός του επεξηγηματικού μοντέλου είναι να καθοριστεί η σχέση μεταξύ ανεξάρτητων και εξαρτημένων μεταβλητών και στη συνέχεια να χρησιμοποιηθεί στην πρόβλεψη της μεταβλητής που μας ενδιαφέρει. Σύμφωνα με αυτήν την οικογένεια μοντέλων οι όποιες μεταβολές στις τιμές εισόδου θα επηρεάσουν την έξοδο του συστήματος κατά έναν προβλέψιμο τρόπο, εφόσον η σχέση παραμένει αναλλοίωτη.

Μοντέλα χρονοσειρών

Αντίθετα με τα επεξηγηματικά μοντέλα, τα μοντέλα πρόβλεψης με χρήση χρονοσειρών αντιμετωπίζουν το σύστημα ως ένα «μαύρο κουτί» και δεν προσπαθούν να ανακαλύψουν παράγοντες που επηρεάζουν τη συμπεριφορά του (μονοδιάστατα μοντέλα χρονοσειρών). Έτσι, η πρόβλεψη των μελλοντικών τιμών βασίζεται αποκλειστικά σε ιστορικές τιμές και

ιστορικά σφάλματα, αλλά όχι σε επεξηγηματικές μεταβλητές που, ενδεχομένως επηρεάζουν το σύστημα. Σκοπός τέτοιων μοντέλων πρόβλεψης είναι να αποκαλυφθούν υπάρχουσες δομές στα ιστορικά στοιχεία και η προβολή αυτής της δομής στο μέλλον.

Υπάρχουν δύο βασικοί λόγοι για τους οποίους είναι προτιμότερο να αντιμετωπίσουμε ένα σύστημα ως μαύρο κουτί. Αρχικά, ενδέχεται το σύστημα να είναι εξαιρετικά πολύπλοκο ώστε να μην μπορεί να εξηγηθεί ως προς τους παράγοντες που το επηρεάζουν και, κατά συνέπεια, να μην είναι δυνατό να ποσοτικοποιηθούν οι σχέσεις που υποτίθεται επηρεάζουν τη συμπεριφορά του. Δεύτερος λόγος είναι πως μπορεί να μη μας ενδιαφέρει το πώς ή γιατί συμβαίνει κάτι, αλλά μόνο το να προβλέψουμε ότι θα συμβεί.

Αν λοιπόν ο σκοπός μας είναι να προβλέψουμε τις μελλοντικές τιμές του ΑΕΠ χωρίς να μας απασχολούν οι λόγοι για τους οποίους πήρε μία συγκεκριμένη τιμή, τότε μία προσέγγιση με χρήση χρονοσειρών θα ήταν καταλληλότερη. Επιπλέον, αν γνωρίζουμε ότι το ΑΕΠ δε μεταβάλλεται δραστικά από μήνα σε μήνα, ούτε από χρόνο σε χρόνο, τότε μπορούμε να υποθέσουμε ότι η αμέσως μελλοντική τιμή του θα εξαρτάται από το επίπεδο στο οποίο βρίσκεται τώρα, ή τις προηγούμενες χρονικές στιγμές. Σύμφωνα με την πιο πάνω παρατήρηση θα είναι:

$$ΑΕΠ_{t+1} = f(ΑΕΠ_t, ΑΕΠ_{t-1}, \dots, σφάλμα)$$

Η διαφορά μεταξύ των δύο μεθόδων εκτίμησης του ΑΕΠ είναι ότι στη δεύτερη σχέση οι εισοδοί είναι το ίδιο το ΑΕΠ σε προηγούμενες χρονικές στιγμές και όχι κάποιες ειδικές μεταβλητές που πρέπει να αναγνωριστούν και μετά να μετρηθούν. Σε κάθε περίπτωση όμως, η σχέση που κυβερνάει τις δύο εκφράσεις πρέπει να ανακαλυφθεί και να μετρηθεί.

Και οι δύο τύποι μοντέλων έχουν πλεονεκτήματα υπό κάποιες συνθήκες. Συνήθως τα μοντέλα χρονοσειρών αποτελούν την ευκολότερη λύση για να προβλέψουμε ένα μέγεθος, ενώ τα επεξηγηματικά μοντέλα εφαρμόζονται καλύτερα στη διαδικασία λήψης αποφάσεων ή στη διαμόρφωση πολιτικών. Όποτε είναι διαθέσιμα τα απαραίτητα στοιχεία μπορούμε να προβούμε σε υποθέσεις ως προς τη σημαντικότητα των σχέσεων μεταξύ των επεξηγηματικών μεταβλητών ή ως προς το χρόνο και να ελέγξουμε αυτές τις υποθέσεις. Όπως είδαμε και από το παράδειγμα του ΑΕΠ, συχνά μπορούμε να κάνουμε προβλέψεις είτε χρησιμοποιώντας επεξηγηματικές μεταβλητές είτε χρονοσειρές. Επίσης, υπάρχουν φορές που οι δύο μέθοδοι μπορούν να συνδυαστούν (πολυμεταβλητά μοντέλα χρονοσειρών).

Βασικά βήματα στη διαδικασία πρόβλεψης

Σε κάθε διαδικασία πρόβλεψης για την οποία είναι διαθέσιμα ποσοτικά δεδομένα, υπάρχουν πέντε βασικά βήματα.

- Καθορισμός του προβλήματος. Αυτό το βήμα είναι ίσως και το πιο δύσκολο, καθώς απαιτεί τη βαθιά κατανόηση των τρόπων με τους οποίους θα χρησιμοποιηθεί η πρόβλεψη, αλλά και του ποιος θα τη χρησιμοποιήσει.
- Συλλογή δεδομένων. Τα δεδομένα εκτός από ποσοτικά (τιμές, παρατηρήσεις), μπορεί να είναι και εμπειρικά.

- Προκαταρκτική ανάλυση. Προσπάθεια ανάλυσης των δεδομένων και άντληση πληροφοριών. Συνήθως είναι χρήσιμο να παρουσιάσουμε τα δεδομένα σε μορφή γραφήματος ώστε να γίνει μια πρώτη εκτίμηση για τη συμπεριφορά των παρατηρήσεων ως συνάρτηση του χρόνου. Στη συνέχεια μελετάμε βασικές στατιστικές ποσότητες, όπως μέση τιμή, διασπορά, μέγιστο, ελάχιστο, κτλ. Ακόμα, μπορούμε να εξαγάγουμε χρήσιμα συμπεράσματα για την ύπαρξη τάσης και εποχικότητας εκτελώντας την ανάλυση αποσύνθεσης της σειράς δεδομένων. Τέλος, ελέγχουμε τα δεδομένα για την ύπαρξη ακραίων τιμών και αν υπάρχουν προσπαθούμε να κατανοήσουμε τους λόγους που τις δημιούργησαν, ώστε να τις αντιμετωπίσουμε.
- Επιλογή μοντέλου και έλεγχοι εφαρμογής. Σε αυτό το βήμα δοκιμάζουμε διαφορετικά ποσοτικά μοντέλα στο σύστημα και ελέγχουμε την εφαρμογή τους. Έχουμε υπόψη μας ότι κάθε μοντέλο αποτελεί ένα σύνολο υποθέσεων για την προϋπάρχουσα δομή στα ιστορικά στοιχεία που διαθέτουμε. Σκοπός μας είναι να επιλέξουμε ένα μοντέλο το οποίο να εφαρμόζεται κατάλληλα στα δεδομένα αυτά ώστε να περιγράφει κατά κάποιον τρόπο το σύστημα που μελετάμε. Έχουμε στη διάθεσή μας πληθώρα στατιστικών μοντέλων που μπορούμε να δοκιμάσουμε, όπως, μοντέλα απλής παλινδρόμησης, εκθετικής εξομάλυνσης, μοντέλα ARIMA Box-Jenkins, ARMA, πολυμεταβλητά ARMA, και πολλά άλλα.
- Χρήση και αξιολόγηση του μοντέλου πρόβλεψης. Αφού επιλέξουμε το μοντέλο πρόβλεψης και τις παραμέτρους του, το χρησιμοποιούμε για να εκτιμήσουμε τις επιθυμητές προβλέψεις. Η εργασία όμως δεν τελειώνει εδώ. Μια πρόβλεψη είναι ολοκληρωμένη όταν αφού τη χρησιμοποιήσουμε για το σκοπό που την εκτιμήσαμε, την αξιολογήσουμε ως προς την εγκυρότητα και την ακρίβειά της. Για αυτόν το σκοπό, θα πρέπει να συγκρίνουμε την τιμή της πρόβλεψης με την πραγματική παρατήρηση όταν αυτή θα είναι διαθέσιμη (Makridakis, Wheelwright, & Hyndman, 1998).

2.2 Αξιολόγηση των μεθόδων πρόβλεψης

Το σφάλμα της πρόβλεψης για κάθε περίοδο e_t είναι η διαφορά μεταξύ της εκτίμησης F_t για την περίοδο t και της παρατήρησης X_t για την ίδια περίοδο ($e_t = F_t - X_t$). Υποθέτουμε ότι έχουμε $X_1, X_2, \dots, X_t, \dots$ παρατηρήσεις για τις περιόδους $1, 2, \dots, t, \dots$ αντίστοιχα. Αν $F_1, F_2, \dots, F_t, \dots$ είναι οι προβλέψεις για τις αντίστοιχες περιόδους, τότε ορίζουμε τα παρακάτω μεγέθη ως εξής:

Σφάλμα $e_t = F_t - X_t$ για προβλέψεις ενός βήματος,

ενώ για πολλαπλά βήματα τ $e_t = F_{t-\tau,t} - X_t$.

- ME (μέσο σφάλμα)

$$ME = \frac{1}{n} \sum_{i=1}^n e_i$$

- MAD (μέση τιμή των απόλυτων σφαλμάτων για n χρονικές περιόδους)

$$MAD = \frac{1}{n} \sum_{i=1}^n |e_i|$$

- MSE (μέσο άθροισμα τετραγωνικών σφαλμάτων για n χρονικές περιόδους)

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2$$

- MAPE (μέσο απόλυτο ποσοστιαίο σφάλμα για n χρονικές περιόδους)

$$MAPE = \left(\frac{1}{n} \sum_{i=1}^n \left| \frac{e_i}{\bar{X}_i} \right| \right) \times 100\%$$

Μέθοδοι πρόβλεψης στην πράξη

- *Η μέθοδος του απλού κινητού μέσου*
Εφαρμόζεται σε στάσιμες χρονοσειρές λαμβάνοντας τον αριθμητικό μέσο των n πιο πρόσφατων παρατηρήσεων και χρησιμοποιώντας αυτό το μέγεθος ως πρόβλεψη για την επόμενη χρονική περίοδο.
- *Η μέθοδος της εκθετικής εξομάλυνσης*
Επίσης για στάσιμες χρονοσειρές, βασίζεται σε έναν σταθμισμένο μέσο όρο των πιο πρόσφατων παρατηρήσεων και της προηγούμενης πρόβλεψης. Αν $0 < \alpha < 1$ είναι το βάρος ή συντελεστής βαρύτητας που εφαρμόζεται στην παρατήρηση τότε εφαρμόζεται ο συντελεστής $1 - \alpha$ στην πρόβλεψη.
- *Η μέθοδος της γραμμικής παλινδρόμησης*
Εφαρμόζεται σε χρονοσειρές με γραμμική ανοδική ή καθοδική τάση. Παλινδρόμηση μπορεί να εφαρμοστεί και σε δεδομένα που παρουσιάζουν διαφορετική συμπεριφορά, π.χ. εκθετική αύξηση.
- *Η μέθοδος της διπλής εκθετικής εξομάλυνσης του Holt,*
Χρησιμοποιεί ένα δεύτερο σύνολο εξισώσεων εκθετικής εξομάλυνσης για την εκτίμηση της κλίσης σε περίπτωση που υπάρχει αυξητική ή φθίνουσα τάση σε κάθε περίοδο.
- *Η μέθοδος των εποχικών συντελεστών*
Σε περίπτωση που έχουμε μία χρονοσειρά με έντονη εποχική συμπεριφορά, χρησιμοποιείται η μέθοδος των εποχικών συντελεστών. Αν υποθέσουμε ότι η χρονοσειρά παρουσιάζει περιοδική συμπεριφορά με σταθερή περίοδο π.χ. 12 μηνών και έχουμε στη διάθεσή μας αρκετές παρατηρήσεις π.χ. 6 ετών, λαμβάνουμε τους μέσους όρους για κάθε ίδιο μήνα του έτους, και

συγκρίνοντας τις τιμές αυτές με το μέσο για κάθε έτος, υπολογίζουμε τους μηνιαίους συντελεστές.

Έχουν γίνει δεκάδες δημοσκοπήσεις μεταξύ ανθρώπων που χρησιμοποιούν προβλέψεις στην εργασία τους. Αυτές οι δημοσκοπήσεις, μεταξύ άλλων, έχουν καταγράψει το βαθμό εξοικείωσης των ανθρώπων με τα διάφορα μοντέλα πρόβλεψης. Σύμφωνα με μία από τις πιο ολοκληρωμένες έρευνες που έγιναν πάνω στο θέμα (Mentzer & Cox, 1984) προκύπτουν διάφορα πολύ χρήσιμα συμπεράσματα σχετικά με τη διάδοση στατιστικών μεθόδων πρόβλεψης στις επιχειρήσεις, που φαίνονται στον Πίνακα 1.

Μέθοδος	Πολύ εξοικειωμένος	Λίγο εξοικειωμένος	Καθόλου εξοικειωμένος
<i>Υποκειμενικές</i>			
Ψηφοφορία διευθυντικών στελεχών	81	6	13
Ερωτηματολόγιο στελεχών πωλήσεων	79	5	16
Ερωτηματολόγιο πελατών	73	7	20
<i>Αντικειμενικές</i>			
Κινητός Μέσος	85	7	8
Γραμμική Προβολή	82	11	7
Εκθετική Εξομάλυνση	73	12	15
Παλινδρόμηση	72	8	20
Ανάλυση Γραμμής Τάσης	67	16	17
Simulation	55	22	23
Ανάλυση Κύκλου Ζωής	48	11	41
Κλασσική Αποσύνθεση	42	9	49
Box-Jenkins (ARIMA)	26	9	65

Πίνακας 1. Βαθμός εξοικείωσης με μεθόδους πρόβλεψης (ποσοστό ερωτηθέντων) Πηγή: (Mentzer & Cox, 1984)

Οι περισσότερες από τις παραπάνω μεθόδους χρησιμοποιούνται εκτενώς στις επιχειρήσεις με σκοπό την πρόβλεψη μεγεθών που έχουν σημασία για αυτές. Οι υποκειμενικές μέθοδοι με τις οποίες είναι περισσότερο εξοικειωμένοι οι χρήστες, δεν απαιτούν τη χρήση κάποιου στατιστικού μοντέλου και βασίζονται στην ερμηνεία του εκάστοτε χρήστη. Η χρησιμότητά τους έγκειται στο μακροχρόνιο προγραμματισμό και στη χάραξη μακροπρόθεσμης στρατηγικής. Οι απλές ποσοτικές μέθοδοι (κινητού μέσου, γραμμική προβολή, εκθετική εξομάλυνση) με τις οποίες είναι εξοικειωμένοι αρκετά οι χρήστες έχουν το πλεονέκτημα ότι απαιτούν μικρή ή μηδενική παρέμβαση ανθρώπου, καθώς οι υπολογισμοί και οι εκτιμήσεις εκτελούνται από υπολογιστές. Έχουν όμως το μειονέκτημα ότι δεν είναι τόσο ευέλικτες όσον αφορά την αντιμετώπιση εποχικών διακυμάνσεων. Παρόλα αυτά, αποδεικνύονται αρκετά αξιόπιστες σε μεσοπρόθεσμες και μακροπρόθεσμες προβλέψεις. Η μέθοδος της παλινδρόμησης είναι αρκετά αξιόπιστη για όλους τους τύπους προβλέψεων και προσφέρει το πλεονέκτημα της «εξήγησης» για τη συμπεριφορά του μεγέθους που μελετάται. Απαιτεί όμως την απασχόληση ειδικευμένου προσωπικού με εμπειρία στη μοντελοποίηση και στον στατιστικό έλεγχο.

Συγκεκριμένα για την πρόβλεψη της ζήτησης, ένα μέγεθος καίριας σημασίας στον προγραμματισμό και έλεγχο αποθεμάτων, οι απλές μέθοδοι θα μπορούσαν να εφαρμοστούν σε προϊόντα μικρής αξίας, κίνησης ή σημασίας όπως αποθέματα κατηγορίας Β ή C της ανάλυσης ABC (βλέπε παράγραφο 3.3 Ανάλυση ABC).

Για τα πολύ σημαντικά προϊόντα (κατηγορίας Α), υπάρχουν πολύπλοκες μέθοδοι, όπως αυτή που παρουσιάζουμε στην εργασία, που προσφέρουν μεγαλύτερη ακρίβεια, αποτελεσματικότητα και φερεγγυότητα και οι οποίες, αν εφαρμοστούν σωστά από εξειδικευμένο προσωπικό, είναι σε θέση να προσφέρουν ένα σημαντικό συγκριτικό πλεονέκτημα στην επιχείρηση. Είναι προφανές ότι τέτοιες μέθοδοι απαιτούν ανθρώπινη παρέμβαση και ενδεχομένως να είναι οι περισσότερο χρονοβόρες. Πιθανώς αυτοί είναι και οι λόγοι για τους οποίους δε χρησιμοποιούνται αρκετά. Στη βιβλιογραφία προτείνεται να γίνεται η χρήση τους μόνο σε καίριας σημασίας μεγέθη (Nahmias, 2004).

3 ΑΠΟΘΕΜΑΤΑ

Όταν αναφερόμαστε στα αποθέματα μιας επιχείρησης που κατασκευάζει και διακινεί προϊόντα, μπορούμε να τα κατηγοριοποιήσουμε ανάλογα με το στάδιο της αλυσίδας κατασκευής στην οποία βρίσκονται. Αυτός δεν είναι ο μόνος τρόπος αλλά είναι φυσικό να υιοθετήσουμε αυτήν την κατηγοριοποίηση όταν αναφερόμαστε σε βιομηχανικές εφαρμογές.

Η διατήρηση αποθέματος γίνεται μόνο όταν η ωφέλεια που παρουσιάζεται είναι μεγαλύτερη από το κόστος που δαπανάται για την εξασφάλισή τους. Το μέγεθος του αποθέματος σε μια ορισμένη θέση πρέπει να είναι τόσο, ώστε να μην παρατηρείται ούτε έλλειψη αλλά ούτε και υπερεπάρκεια.

Τόσο η έλλειψη, όσο και η υπερεπάρκεια αποθεμάτων έχουν οικονομικές συνέπειες. Συγκεκριμένα, η έλλειψη προκαλεί απώλεια της ωφέλειας που μας εξασφαλίζουν τα αποθέματα, η δε υπερεπάρκεια προκαλεί δέσμευση μεγαλύτερων κεφαλαίων. Συμπεραίνουμε λοιπόν, ότι πρέπει να διατηρούμε τέτοιο μέγεθος αποθεμάτων που ελαχιστοποιεί τις δύο οικονομικές απώλειες που προκαλούν η έλλειψη και η υπερεπάρκεια τους.

Ο Προγραμματισμός και Έλεγχος Αποθεμάτων προσπαθεί από τη μαθηματική διερεύνηση της χρονικής και ποσοτικής διακίνησης του αποθέματος, να διαμορφώσει κατάλληλους κανόνες για την τήρηση των αποθεμάτων σε μια επιχείρηση. Η διατήρηση του αποθέματος καθορίζεται από την ποσότητα που παραγγέλλεται και από το χρόνο στον οποίο δίνεται αυτή η παραγγελία

Κατηγοριοποίηση αποθεμάτων

- Πρώτες ύλες
- Εξαρτήματα
- Ημιτελή προϊόντα
- Έτοιμα προϊόντα

Λόγοι για την τήρηση αποθεμάτων

- Οικονομίες κλίμακας. Είναι φθηνότερο να αγοράσει ή να παρασκευάσει κανείς προϊόντα σε μεγάλες ποσότητες απ' ό,τι σε μικρές.
- Αβεβαιότητα όσον αφορά τη ζήτηση, το χρόνο παράδοσης, τη διαθεσιμότητα του προϊόντος στο μέλλον, την ενδεχόμενη αύξηση της τιμής του, κ.α. είναι όλοι λόγοι υπέρ της τήρησης αποθεμάτων.
- Εξομάλυνση πιθανής έκρηξης της ζήτησης στο μέλλον
- Κόστος ελέγχου.

Χαρακτηριστικά των συστημάτων αποθεμάτων

Μπορούμε να κατηγοριοποιήσουμε τα διάφορα συστήματα ελέγχου αποθεμάτων που χρησιμοποιούνται συνήθως από επιχειρήσεις, αναγνωρίζοντας τα βασικά χαρακτηριστικά τους.

- *Χαρακτηρισμός της ζήτησης.* Η αρχική υπόθεση που κάνει κανείς για τη φύση της ζήτησης, συχνά αποδεικνύεται ότι είναι και η πιο σημαντική στον καθορισμό της πολυπλοκότητας του μοντέλου ελέγχου που επιλέγεται.
 - i) *Σταθερή ή μεταβαλλόμενη ζήτηση.* Τα απλούστερα συστήματα ελέγχου αποθεμάτων αντιμετωπίζουν τη ζήτηση ως σταθερή. Το μοντέλο EOQ (Economic Order Quantity) και οι εφαρμογές του βασίζονται σε αυτήν την υπόθεση. Η διαχείριση της ζήτησης ως μεταβαλλόμενου μεγέθους ανακύπτει σε πληθώρα εφαρμογών και συχνά οδηγεί σε πολυπλοκότερα συστήματα.
 - ii) *Τυχαία ζήτηση.* Η ζήτηση θεωρείται τυχαία με σταθερή μέση τιμή. Σχεδόν το σύνολο των στοχαστικών μοντέλων υποθέτουν ότι η μέση ζήτηση είναι σταθερή. Αυτά τα μοντέλα είναι συνήθως και τα πιο ρεαλιστικά αλλά ταυτόχρονα πιο πολύπλοκα από τα ντετερμινιστικά.
- *Χρόνος αναπλήρωσης (lead time).* Ο χρόνος μεταξύ της τοποθέτησης μιας παραγγελίας και της ολοκλήρωσής της.
- *Χρόνος επιθεώρησης.* Οι χρονικές στιγμές κατά τις οποίες ελέγχεται το επίπεδο του αποθέματος.
- *Αντιμετώπιση της υπερβολικής ζήτησης.* Σε περίπτωση που το απόθεμα δεν επαρκεί για την κάλυψη της ζήτησης, αυτή η ζήτηση μπορεί να ικανοποιείται σε μεταγενέστερο χρόνο (backorder) ή να χάνεται (απώλεια πωλήσεων).

Σχετικά κόστη

Επειδή μας ενδιαφέρει η βελτιστοποίηση του συστήματος ελέγχου αποθεμάτων, πρέπει να αποφασίσουμε ως προς ποιο μέγεθος θα γίνει αυτή η βελτιστοποίηση. Σχεδόν όλα τα μοντέλα ελέγχου αποθεμάτων θεωρούν την ελαχιστοποίηση του συνολικού κόστους διαχείρισης ως τη βέλτιστη λύση. Μια εναλλακτική προσέγγιση είναι η προσπάθεια εύρεσης μιας λύσης που μεγιστοποιεί το αναμενόμενο ολικό κέρδος. Οι δυο προσεγγίσεις στα περισσότερα μοντέλα μπορεί να θεωρηθούν ισοδύναμες, αφού σχεδόν όλα διαχειρίζονται περίπου τις ίδιες κατηγορίες κόστους.

- *Κόστος διατήρησης.* Περιλαμβάνει το κόστος αποθήκευσης, το κόστος ενδεχόμενης καταστροφής ή λήξης του αποθέματος, το κόστος ασφάλισης, το κόστος δέσμευσης κεφαλαίου και την πιθανή απώλεια εσόδων από εναλλακτική επένδυση του δεσμευμένου στα αποθέματα κεφαλαίου. Το τελευταίο είναι συχνά και το σημαντικότερο όταν υπολογίζουμε το κόστος διατήρησης. Το κεφάλαιο που δεν είναι «αποθηκευμένο» σε αποθέματα, παρόλο που τα αποθέματα είναι ζωτικής σημασίας, θα μπορούσε να επενδυθεί εσωτερικά σε άλλους τομείς της επιχείρησης, ή εξωτερικά. Ο υπολογισμός του κόστους δέσμευσης κεφαλαίου είναι δύσκολος και έχει να κάνει με παραμέτρους όπως τα επιτόκια, τις επενδυτικές συνθήκες, κ.α.. Είναι επίσης εμφανές, ότι το επενδυμένο κεφάλαιο σε αποθέματα, επιβαρύνεται και από τις υπόλοιπες κατηγορίες κόστους που αναφέρθηκαν, όπως ασφάλιση, κίνδυνος καταστροφής, κτλ.
- *Κόστος παραγγελίας.* Ενώ το κόστος διατήρησης είναι ανάλογο της ποσότητας αποθέματος που είναι διαθέσιμο, το κόστος παραγγελίας εξαρτάται από την ποσότητα αποθέματος που παραγγέλλεται ή παράγεται. Αποτελείται κυρίως από δύο συνιστώσες. Μια σταθερή συνιστώσα K , που σχετίζεται με το κόστος που έχει η τοποθέτηση μιας παραγγελίας και είναι ανεξάρτητο από το μέγεθός της, και μια μεταβλητή συνιστώσα c που σχετίζεται με το κόστος ανά μονάδα προϊόντος που παραγγέλλεται. Το K ονομάζεται και *setup cost*, ενώ το c ονομάζεται *proportional order cost*. Ορίζουμε $C(x)$ το κόστος παραγγελίας x μονάδων προϊόντος. Είναι:
$$C(x) = \begin{cases} 0, & x = 0 \\ K + cx, & x > 0 \end{cases}$$
- *Κόστος ποινής (penalty costs).* Αυτά επιβαρύνουν την επιχείρηση όταν το απόθεμα δεν υπερκαλύπτει τη ζήτηση. Σε αυτήν την περίπτωση παρατηρούνται κόστη που σχετίζονται με την απώλεια πωλήσεων και την απώλεια της καλής πίστης των πελατών (good will).

3.1 Έλεγχος αποθεμάτων με γνωστή (καθορισμένη) ζήτηση

Το βασικό μοντέλο EOQ (Economic Order Quantity)

Είναι η βάση όλων των μοντέλων αναπλήρωσης αποθεμάτων και αναπτύχθηκε περί το 1915. Περιγράφει το βασικό συμβιβασμό μεταξύ του κόστους τήρησης αποθεμάτων και του κόστους που συνεπάγεται η παραγγελία αποθεμάτων.

Υποθέτουμε:

- Ο ρυθμός της ζήτησης είναι σταθερός και ίσος με λ μονάδες προϊόντος ανά μονάδα χρόνου.
- Η έλλειψη σε προϊόντα δεν επιτρέπεται.
- Ο χρόνος αναπλήρωσης είναι μηδέν.
- Το κόστος περιλαμβάνει: Το αρχικό σταθερό κόστος (setup cost) K για κάθε παραγγελία, το αναλογικό κόστος παραγγελίας (proportional order cost) c ανά μονάδα προϊόντος που παραγγέλλεται και τέλος, το κόστος διατήρησης h ανά μονάδα προϊόντος και ανά μονάδα χρόνου.

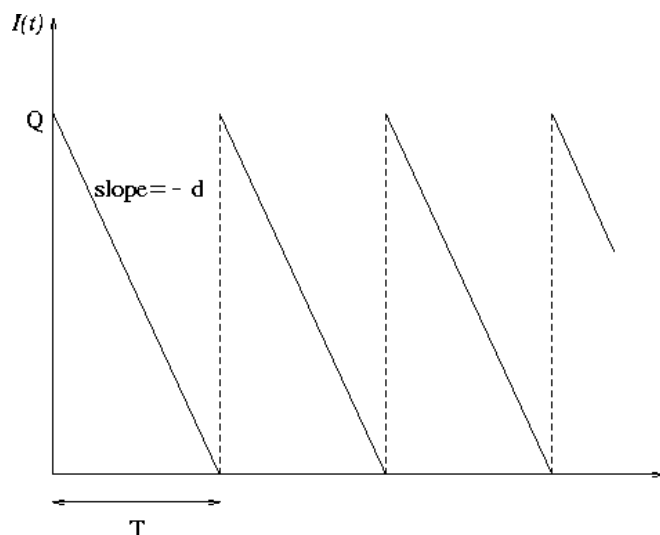
Χωρίς βλάβη της γενικότητας υποθέτουμε ότι τη χρονική στιγμή 0 έχουμε στη διάθεσή μας 0 μονάδες προϊόντος. Συνεπώς τοποθετούμε μία παραγγελία τη χρονική στιγμή 0, έστω για Q μονάδες προϊόντος. Τότε το απόθεμα αναπληρώνεται άμεσα και γίνεται Q μονάδες για $t=0$ (Εικόνα 1. Επίπεδα αποθέματος για το μοντέλο ΕΟQ.)

Κατά την επόμενη χρονική περίοδο τοποθετούμε ξανά μία παραγγελία. Τότε το απόθεμα είναι είτε θετικό, είτε μηδέν. Είναι προφανές ότι αν περιμένουμε το επίπεδο του αποθέματος να φτάσει πάλι στο μηδέν πριν παραγγείλουμε ξανά, τότε ελαχιστοποιούμε το κόστος διατήρησης ανά μονάδα προϊόντος, επειδή στη συγκεκριμένη περίπτωση ο χρόνος αναπλήρωσης είναι μηδέν. Σκοπός είναι να επιλέξουμε το κατάλληλο Q ώστε να ελαχιστοποιήσουμε το μέσο κόστος ανά μονάδα χρόνου.

Εκφράζουμε το μέσο κόστος ανά μονάδα χρόνου ως συνάρτηση του Q . Σε κάθε κύκλο, το σταθερό συν το αναλογικό κόστος παραγγελίας είναι:

$$C(Q) = K + cQ$$

Αν $\lambda=Q/T$, T η διάρκεια του κύκλου, τότε $T=Q/\lambda$. Συμπεραίνουμε ότι το μέσο κόστος για τη διάρκεια του κύκλου T , έστω $G(Q)$ είναι:



Εικόνα 1. Επίπεδα αποθέματος για το μοντέλο ΕΟQ.

$$G(Q) = \frac{K + cQ}{T} + \frac{hQ}{2} = \frac{K + cQ}{\frac{Q}{\lambda}} + \frac{hQ}{2} = \frac{K\lambda}{Q} + \lambda c + \frac{hQ}{2}$$

Οι τρεις συνιστώσες του $G(Q)$ είναι το αρχικό σταθερό κόστος (setup cost), το κόστος αγοράς και το κόστος διατήρησης κατά τη διάρκεια ενός κύκλου. Προσπαθούμε να ελαχιστοποιήσουμε το $G(Q)$ για να βρούμε το βέλτιστο Q .

$$G'(Q) = -\frac{K\lambda}{Q^2} + \frac{h}{2}$$

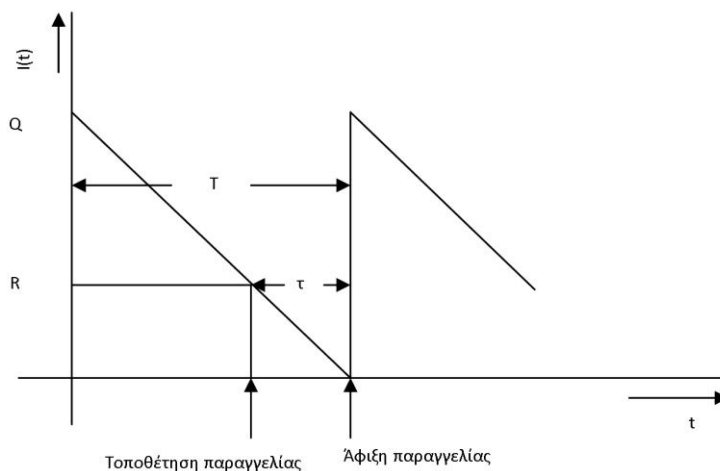
$$G''(Q) = \frac{2K\lambda}{Q^3} > 0, \text{ για } Q > 0$$

Ακόμα είναι $G'(0)=-\infty$ και $G(\infty)=h/2$. Η βέλτιστη ποσότητα Q προκύπτει για $G'(Q)=0$. Συνεπώς $Q^2 = \frac{2K\lambda}{h} \Rightarrow Q^* = \sqrt{2K\lambda/h}$

Το Q^* είναι γνωστό ως η *οικονομική ποσότητα παραγγελίας* (EOQ).

Αυτό το αποτέλεσμα γενικεύεται ώστε να συμπεριλαμβάνει κάποιο θετικό χρόνο αναπλήρωσης, ως εξής.

Αν $\tau > 0$ είναι ο χρόνος αναπλήρωσης, δηλαδή ο χρόνος που μεσολαβεί από την τοποθέτηση της παραγγελίας μέχρι την παράδοση των αποθεμάτων, τότε ορίζουμε $R=\lambda\tau$, το σημείο αναπαραγγελίας, δηλαδή το επίπεδο του αποθέματος που απομένει τη χρονική στιγμή κατά την οποία πρέπει να τοποθετηθεί μια νέα παραγγελία (Εικόνα 2).



Εικόνα 2. Σημείο αναπαραγγελίας

3.2 Έλεγχος αποθεμάτων με άγνωστη (στοχαστική) ζήτηση

Στην προηγούμενη ενότητα θεωρήσαμε ότι η ζήτηση ήταν σταθερή και ίση με λ μονάδες προϊόντος ανά μονάδα χρόνου. Στη συνέχεια βρήκαμε την οικονομική ποσότητα παραγγελίας, σύμφωνα με τις ασθενέστερες δυνατές προϋποθέσεις. Στην πραγματικότητα, η ζήτηση είναι σπάνια γνωστή εκ των προτέρων. Συνήθως η μελλοντική ζήτηση είναι άγνωστη.

Επειδή η διαχείριση αποθεμάτων επηρεάζεται σε σημαντικό βαθμό από την αβεβαιότητα και τη μεταβλητότητα της ζήτησης, η χρησιμότητα των ντετερμινιστικών πρακτικών φαντάζει ισχνή. Η συνεισφορά τους όμως είναι αρκετά σημαντική, κυρίως όσον αφορά τη διαισθητική αντίληψη που αποκτά κανείς μελετώντας τέτοιες πρακτικές. Επίσης, υπάρχουν περιπτώσεις όπου η ζήτηση μπορεί να προσεγγιστεί ικανοποιητικά από ντετερμινιστικά μοντέλα, ειδικά όταν ο βαθμός μεταβλητότητας είναι σχετικά μικρός.

Αν D είναι η ζήτηση για κάποιο προϊόν κατά τη διάρκεια της χρονικής περιόδου T , μπορούμε να θεωρήσουμε τη ζήτηση αυτή ως το άθροισμα δύο συνιστωσών:

$$D = D_{Det} + D_{Ran},$$

δηλαδή ως σύνθεση της ντετερμινιστικής και τυχαίας συνιστώσας.

Οι περιπτώσεις στις οποίες μπορεί να θεωρηθεί η ζήτηση D ως ντετερμινιστική, παρόλο που η συνιστώσα D_{Rand} δεν είναι μηδέν, είναι:

- Όταν η μεταβλητότητα της τυχαίας συνιστώσας είναι μικρή σε σχέση με το επίπεδο τιμών του D .
- Όταν η στοχαστική συμπεριφορά της τυχαίας συνιστώσας είναι ιδιαίτερα πολύπλοκη για να μοντελοποιηθεί.

Παρόλα αυτά, για πολλά προϊόντα η στοχαστική συμπεριφορά της ζήτησης είναι αρκετά σημαντική και δεν μπορεί να αγνοηθεί. Εφόσον η αναμενόμενη ζήτηση ανά μονάδα χρόνου είναι σχετικά σταθερή και η δομή του προβλήματος που προκύπτει δεν είναι υπερβολικά πολύπλοκη, είναι προτιμότερο να μελετηθεί η ζήτηση συμπεριλαμβάνοντας και την τυχαία συνιστώσα. Θα δούμε στη συνέχεια τρόπους με τους οποίους μπορούμε να προβλέψουμε τη ζήτηση και σε πιο πολύπλοκες περιπτώσεις, εκτελώντας μετασχηματισμούς στο σύνολο των παρατηρήσεων (ιστορικά στοιχεία) που διαθέτουμε, ώστε να προκύψει μια πιο ομαλή χρονοσειρά (στάσιμη).

Το μοντέλο του εφημεριδοπώλη

Ας υποθέσουμε ότι προσπαθούμε να προβλέψουμε την ποσότητα ενός προϊόντος (π.χ. μιας συγκεκριμένης εβδομαδιαίας εφημερίδας) που πρέπει να παραγγείλουμε, ώστε να ικανοποιήσουμε την εβδομαδιαία ζήτηση. Θεωρούμε ότι σε περίπτωση που δεν πωληθεί μία εφημερίδα, μπορεί να επιστραφεί στον προμηθευτή με μία τιμή μικρότερη από την τιμή αγοράς. Σκοπός είναι να καλύψουμε τη ζήτηση χωρίς να διατηρούμε υπερβολικό αποθέμα το οποίο θα πρέπει να επιστρέψουμε, μιας και μετά το πέρας της εβδομάδας δε θα μπορεί να διατεθεί (ημερομηνία λήξης).

Από προηγούμενες εβδομάδες έχει παρατηρηθεί ότι η ζήτηση έχει μέση τιμή μ και τυπική απόκλιση σ . Μία προφανής επιλογή θα μπορούσε να είναι η προμήθεια ποσότητας που να καλύπτει τη μέση ζήτηση. Αυτή όμως η προσέγγιση δεν είναι σωστή διότι αν για παράδειγμα η τιμή αγοράς είναι $x=25$, η τιμή πώλησης είναι $y=75$ και η τιμή επιστροφής στον προμηθευτή είναι $z=10$, τότε για κάθε απούλητη εφημερίδα το αποτέλεσμα είναι $x-z=15$. Από την άλλη, κάθε φορά που παρουσιάζεται αδυναμία πώλησης μιας εφημερίδας, το αποτέλεσμα είναι $y-x=50$. Φαίνεται πως η ποινή για την έλλειψη είναι μεγαλύτερη από την ποινή της υπερβολικής τήρησης αποθέματος.

Ορίζουμε:

- c_0 = κόστος ανά μονάδα αποθέματος που απομένει στο τέλος της περιόδου
- c_u = κόστος ανά μονάδα ανικανοποίητης ζήτησης.

Επιπλέον, θεωρούμε ότι η ζήτηση D είναι συνεχής, μη αρνητική τυχαία μεταβλητή με σππ $f(x)$ και αθροιστική κατανομή $F(x)$. Ορίζουμε τη συνάρτηση κόστους και ελαχιστοποιούμε ως προς το αναμενόμενο κόστος. Έτσι υπολογίζουμε τη βέλτιστη ποσότητα παραγγελίας Q στην αρχή της περιόδου.

Ορίζουμε $G(Q, D)$ το συνολικό κόστος που επιβαρύνεται η επιχείρηση στο τέλος της περιόδου όταν παραγγέλνονται Q μονάδες προϊόντος και όταν η ζήτηση είναι D . Τότε

$$\max\{Q - D, 0\} = \begin{cases} Q - D, & Q \geq D \\ 0, & Q \leq D \end{cases}$$

είναι η ποσότητα που απομένει στο τέλος της περιόδου. Αντίστοιχα, $\max\{D - Q, 0\}$ δηλώνει την περίπτωση που η ζήτηση ήταν μεγαλύτερη από το διαθέσιμο απόθεμα κατά το τέλος της περιόδου. Έτσι,

$$G(Q, D) = c_0 \max\{0, Q - D\} + c_u \max\{0, D - Q\}.$$

Συνεπάγεται ότι η αναμενόμενη τιμή της συνάρτησης κόστους είναι

$$\begin{aligned} G(Q) &= E[G(Q, D)] = \\ &= c_0 \int_0^\infty \max(0, Q - x) f(x) dx + c_u \int_0^\infty \max(0, x - Q) f(x) dx = c_0 \int_0^Q (Q - \\ & x) f(x) dx + c_u \int_Q^\infty (x - Q) f(x) dx. \end{aligned}$$

$$\frac{dG(Q)}{dQ} = c_0 \int_0^Q 1 f(x) dx + c_u \int_Q^\infty (-1) f(x) dx = c_0 F(Q) - c_u (1 - F(Q)).$$

$$\frac{d^2 G(Q)}{dQ^2} = (c_0 + c_u) f(Q) \geq 0 \text{ για κάθε } Q \geq 0.$$

Συνεπώς $G(Q)$ είναι κυρτή. Παρατηρούμε ότι:

$$\left. \frac{dG(Q)}{dQ} \right|_{(0)} = c_0 F(0) - c_u (1 - F(0)) = -c_u < 0, F(0) = 0.$$

Η βέλτιστη λύση προκύπτει για $G'(Q) = 0$. Δηλαδή,

$$G'(Q) = (c_0 + c_u) F(Q^*) - c_u = 0 \Rightarrow F(Q^*) = \frac{c_u}{c_0 + c_u}.$$

Η παραπάνω, έχει λύση για κάθε συνεχή κατανομή αφού είναι $0 \leq \frac{c_u}{c_0 + c_u} \leq 1$.

Όταν υπάρχει αρχικό απόθεμα

Υποθέσαμε ότι το αρχικό απόθεμα είναι ίσο με μηδέν. Στη γενική περίπτωση όπου το αρχικό απόθεμα είναι $u > 0$, η βέλτιστη ποσότητα αναπαραγγελίας διαμορφώνεται σε

$$Q^{**} = \begin{cases} Q^* - u, & u < Q^* \\ 0, & u \geq Q^* \end{cases}$$

3.3 Συστήματα πολλαπλών προϊόντων – Ανάλυση ABC

Ένα θέμα που προκύπτει κατά τη μελέτη ενός συστήματος ελέγχου αποθεμάτων είναι αυτό της διαχείρισης κόστους-ωφέλειας από την εφαρμογή αυτού του συστήματος. Μία επιχείρηση που διακινεί πολλά προϊόντα αναγκάζεται να εστιάσει στα πιο προσοδοφόρα. Κάποιες κατηγορίες προϊόντων δικαιολογούν τη δαπάνη πόρων (κόστος διαχείρισης) για την ακριβέστερη μελέτη τους, ενώ για κάποιες άλλες κατηγορίες αυτό το κόστος δεν έχει νόημα. Για παράδειγμα, η επιβάρυνση 200 ευρώ ετησίως για την παρακολούθηση ενός προϊόντος που αποφέρει καθαρό κέρδος 100 ευρώ δεν είναι οικονομικά συμφέρουσα.

Για το λόγο αυτό είναι αναγκαίο να ξεχωρίσουμε τα προϊόντα σε κερδοφόρα, λιγότερο κερδοφόρα και ενδεχομένως σε μη κερδοφόρα. Ακολουθώντας το νόμο Pareto και

εφαρμόζοντάς τον στη διαχείριση αποθεμάτων, προκύπτει πως συχνά, ένα μεγάλο μέρος των πωλήσεων δημιουργείται από ένα σχετικά μικρό ποσοστό του πλήθους των προϊόντων.

Χονδρικά, το 20% των προϊόντων (κατηγορία A) αποφέρει των 80% των πωλήσεων, το επόμενο 30% των προϊόντων αποφέρει το 15% των πωλήσεων και τέλος το 50% των προϊόντων (κατηγορία C) αποφέρει το υπόλοιπο 5% των πωλήσεων.

Σύμφωνα με αυτήν την κατηγοριοποίηση, διανέμονται και οι ανάλογοι πόροι στην παρακολούθηση των προϊόντων. Είναι λοιπόν λογικό, για την κατηγορία A να εφαρμόζονται ακριβέστερες μέθοδοι παρακολούθησης και πρόβλεψης, οι παραγγελίες για αναπλήρωση να εκτελούνται μεμονωμένα για κάθε προϊόν αν αυτό είναι αναγκαίο, ώστε η τελική απόδοση του συστήματος παρακολούθησης αποθεμάτων να είναι η καλύτερη δυνατή. Αντίστοιχα, για τη B κατηγορία η παρακολούθηση μπορεί να γίνει λιγότερο αυστηρή, καθώς και τα μοντέλα πρόβλεψης που εφαρμόζονται μπορούν να απλοποιηθούν. Τέλος, για την τελευταία κατηγορία C οι παραγγελίες μπορεί να τοποθετούνται μαζικά και σε μικρότερη συχνότητα, ώστε να μειωθεί το κόστος διαχείρισης για αυτήν την κατηγορία, απαιτώντας τη λιγότερη δυνατή ενασχόλησή μας. Ειδικά για ακριβά προϊόντα με ελάχιστες διακινούμενες ποσότητες, είναι προτιμότερο να μην τηρούνται καθόλου αποθέματα, αντίθετα η ενδεχόμενη ζήτηση να αντιμετωπίζεται με επί τόπου παραγγελίες.

3.4 Σύνδεση προβλέψεων με τον έλεγχο αποθεμάτων

Στο σημείο αυτό πρέπει να τονιστεί η μία σημαντική παράμετρος που αφορά τη χρήση των ιστορικών στοιχείων στις προβλέψεις της ζήτησης. Μέχρι στιγμής, έχουμε αναφερθεί στις καταγραμμένες πωλήσεις ως ζήτηση. Αυτό όμως δεν είναι πάντα σωστό. Συνήθως, οι παρατηρήσεις που έχουμε στη διάθεσή μας αντιπροσωπεύουν τις πωλήσεις και όχι τη ζήτηση. Υπάρχει δηλαδή περίπτωση, για κάποιες χρονικές περιόδους, το διαθέσιμο απόθεμα να μηδενίστηκε δημιουργώντας κολοβές παρατηρήσεις.

Παρατηρούμε λοιπόν, ότι υπάρχουν περιπτώσεις που τα ίδια τα δεδομένα μπορούν να δημιουργήσουν προβλήματα στις προβλέψεις που καλούμαστε να κάνουμε, αν προηγουμένως δεν έχουμε μεριμνήσει να αντιμετωπίσουμε τέτοιου είδους καταστάσεις (αντιμετώπιση outliers, inliers). Υπάρχουν μέθοδοι συμπλήρωσης κολοβών παρατηρήσεων, που στην πράξη βοηθούν στο σχηματισμό καλύτερων προβλέψεων.

Επιπρόσθετα, αρκετές φορές οι παρατηρήσεις παρουσιάζουν διακυμάνσεις ανάλογα με την εποχή (εποχικότητα) και εξαρτήσεις μεταξύ τους. Για την αντιμετώπιση αυτών των φαινομένων απαιτείται η χρήση διαφορετικών μεθόδων τις οποίες περιγράφουμε στα επόμενα κεφάλαια.

Στην παρούσα εργασία μελετάμε τη ζήτηση για διάφορα φάρμακα με ιδιαίτερο ενδιαφέρον, καθώς εντάσσονται στην κατηγορία των ταχέως κινούμενων προϊόντων στο χώρο της υγείας. Σε αυτά τα προϊόντα εφαρμόζουμε μοντέλα πρόβλεψης για τη ζήτηση με χρήση χρονοσειρών. Συγκεκριμένα, εφαρμόζουμε μοντέλα πρόβλεψης χρονοσειρών, ARMA, ARIMA και SARIMA βασιζόμενοι σε ιστορικά στοιχεία πωλήσεων και στη συνέχεια επιλέγουμε το κατάλληλο μοντέλο για κάθε ένα προϊόν. Τέλος, προσπαθούμε να

καταλήξουμε στη δημιουργία αξιόπιστων προβλέψεων για τη ζήτηση της αμέσως επόμενης χρονικής περιόδου, όπου αυτό είναι δυνατό. Για το σκοπό αυτό, παραθέτουμε μια εισαγωγή στην Ανάλυση Χρονοσειρών.

4 ΧΡΟΝΟΣΕΙΡΕΣ

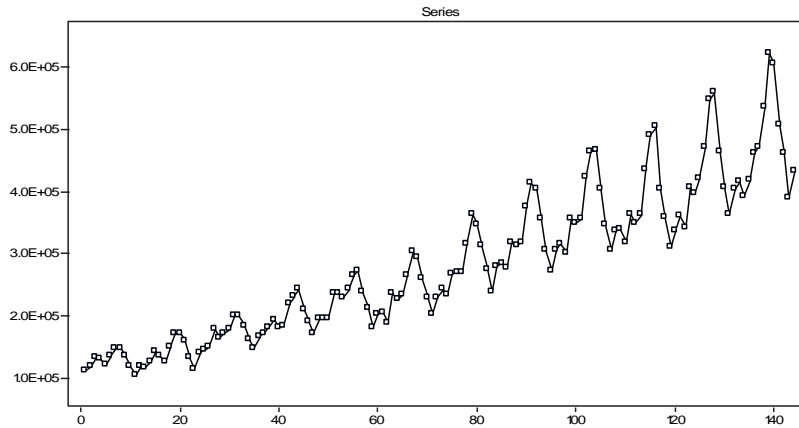
Με τον όρο χρονοσειρά εννοούμε ένα σύνολο καταστάσεων $\{X_t: t = 0, 1, 2, \dots\}$ κάθε μία από τις οποίες καταγράφεται σε συγκεκριμένη χρονική στιγμή t . Το σύνολο των δυνατών καταστάσεων ονομάζεται *χώρος καταστάσεων* και συμβολίζεται με $S \subseteq \mathbb{R}$, ενώ το σύνολο τιμών του t ονομάζεται *παραμετρικός χώρος* και συμβολίζεται με T . Ανάλογα με τη μορφή του χώρου καταστάσεων και του παραμετρικού χώρου, οι χρονοσειρές ταξινομούνται στις ακόλουθες κατηγορίες:

- Διακριτές καταστάσεις σε διακριτό χρόνο
- Διακριτές καταστάσεις σε συνεχή χρόνο
- Συνεχείς καταστάσεις σε διακριτό χρόνο
- Συνεχείς καταστάσεις σε συνεχή χρόνο

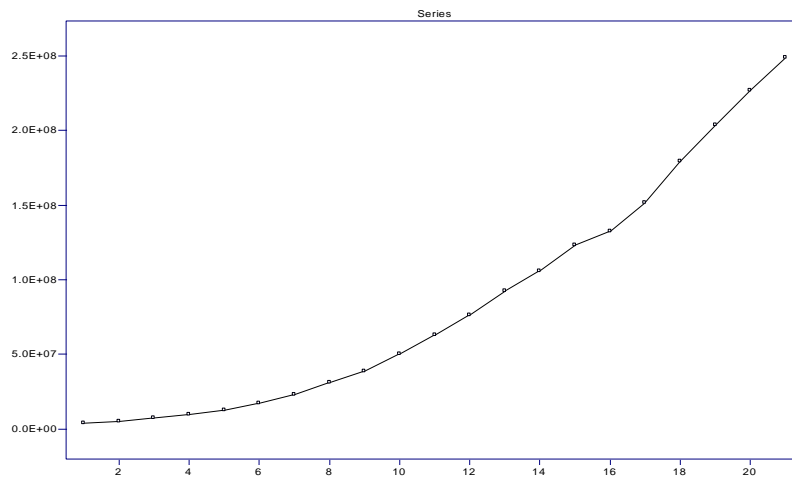
Συνήθως η μέθοδος συλλογής των δεδομένων υπαγορεύει και τη μορφή του παραμετρικού χώρου, καθιστώντας έτσι τον παραμετρικό χώρο π.χ. διακριτό. Για παράδειγμα, οι αφίξεις πελατών σε ένα κατάστημα ανά ώρα, η θερμοκρασία σε κάποιο συγκεκριμένο χώρο ανά λεπτό, κ.α.

Παρακάτω φαίνονται μερικά παραδείγματα χρονοσειρών που χρησιμοποιούνται εκτενώς στη βιβλιογραφία. Συγκεκριμένα στο Σχήμα 1 φαίνεται το πλήθος των διεθνών αεροπορικών επιβατών από τον Ιανουάριο του 1949 ως το Δεκέμβριο του 1960. Στο Σχήμα 2 φαίνεται ο πληθυσμός των Ηνωμένων Πολιτειών της Αμερικής από το 1790 ως το 1990, όπως καταγράφεται στις δεκαετίες απογραφές. Τέλος, στο Σχήμα 3 βλέπουμε τις γεννήσεις στην Ελλάδα κατά την περίοδο 1955-2010, όπως καταγράφονται από την Ελληνική Στατιστική Υπηρεσία.

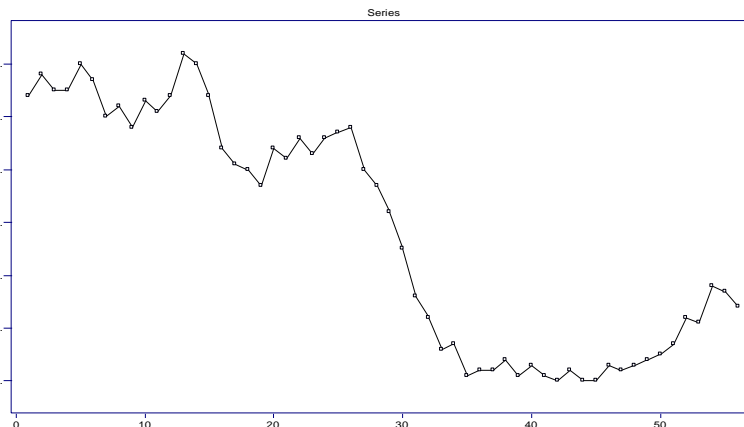
Παρατηρούμε ότι αυτές οι χρονοσειρές διαφέρουν αρκετά ως προς τη δομή τους. Το πλήθος των αεροπορικών επιβατών εμφανίζει αυξητική τάση, και παρουσιάζει έντονες αυξήσεις κατά τους καλοκαιρινούς μήνες, σταδιακή πτώση κατά τους φθινοπωρινούς και μικρή άνοδο κατά τους χειμερινούς. Από την άλλη, ο πληθυσμός παρουσιάζει αυξητική τάση, ίσως εκθετική, χωρίς εμφανή περιοδική συμπεριφορά, γεγονός που συμφωνεί με την διαίσθησή μας.



Σχήμα 1. Πλήθος αεροπορικών επιβατών ανά μήνα από το 1949 έως το 1960 Πηγή: (Box & Jenkins, 1970)



Σχήμα 2. Πληθυσμός ΗΠΑ ανά 10ετία από το 1790 έως το 1990.



Σχήμα 3. Γεννήσεις στην Ελλάδα την περίοδο 1955-2010 ($\times 10^3$) (Πηγή: Ελλ. Στατιστική Υπηρεσία)

Κύριος στόχος της ανάλυσης χρονοσειρών είναι να κατασκευάσουμε μαθηματικά μοντέλα που θα μπορούσαν να περιγράψουν τις παρατηρήσεις. Αναζητούμε, δηλαδή, κάποιο θεωρητικό στατιστικό μοντέλο ώστε η χρονοσειρά που έχουμε στα χέρια μας να μπορεί να θεωρηθεί ως μία τροχιά ή πραγματοποίηση αυτού. Αφού επιλέξουμε μία οικογένεια μοντέλων που είναι ικανοποιητική στο να περιγράψει τα αρχικά δεδομένα, μπορούμε να προχωρήσουμε στην εκτίμηση των παραμέτρων για το μοντέλο αυτό, να ελέγξουμε την προσαρμογή (goodness of fit) του συγκεκριμένου μοντέλου και πιθανώς να

χρησιμοποιήσουμε το προσαρμοσμένο μοντέλο ώστε να αποκτήσουμε πολύτιμη γνώση για το μηχανισμό που δημιούργησε τη χρονοσειρά. Στη συνέχεια, μας ενδιαφέρει να προβλέψουμε μελλοντικές τιμές x_{n+k} : $k = 1, \dots, m$ χρησιμοποιώντας τις παρατηρήσεις x_1, x_2, \dots, x_n .

Φαίνεται λοιπόν ότι ένας σημαντικός παράγοντας στην ανάλυση μιας χρονοσειράς είναι η επιλογή ενός κατάλληλου μοντέλου πιθανοτήτων για τα δεδομένα. Είναι φυσικό να υποθέσουμε ότι κάθε παρατήρηση x_t , $t \in T$ είναι μία πραγματοποίηση μίας συγκεκριμένης τυχαίας μεταβλητής X_t . Για να αποφύγουμε τη σύγχυση μεταξύ του μοντέλου χρονοσειράς, που είναι μία πιθανοθεωρητικά δομημένη οικογένεια τυχαίων μεταβλητών, και των παρατηρήσεων που θεωρείται πως είναι μία πραγματοποίηση αυτού, θα αναφερόμαστε τόσο στο μοντέλο, όσο και στις παρατηρήσεις με τον όρο *χρονοσειρά*.

4.1 Μερικά απλά μοντέλα χρονοσειρών

Θόρυβος iid

Ένα απλό μοντέλο για μία χρονοσειρά είναι αυτό που οι παρατηρήσεις είναι ανεξάρτητες τυχαίες μεταβλητές με την ίδια κατανομή (independent identically distributed ή αλλιώς iid) και μέση τιμή 0. Επειδή σε αυτό το μοντέλο δεν υπάρχει εξάρτηση μεταξύ των παρατηρήσεων, μπορούμε να πούμε πως για κάθε $h \geq 1$ και για κάθε x, x_1, \dots, x_n ,

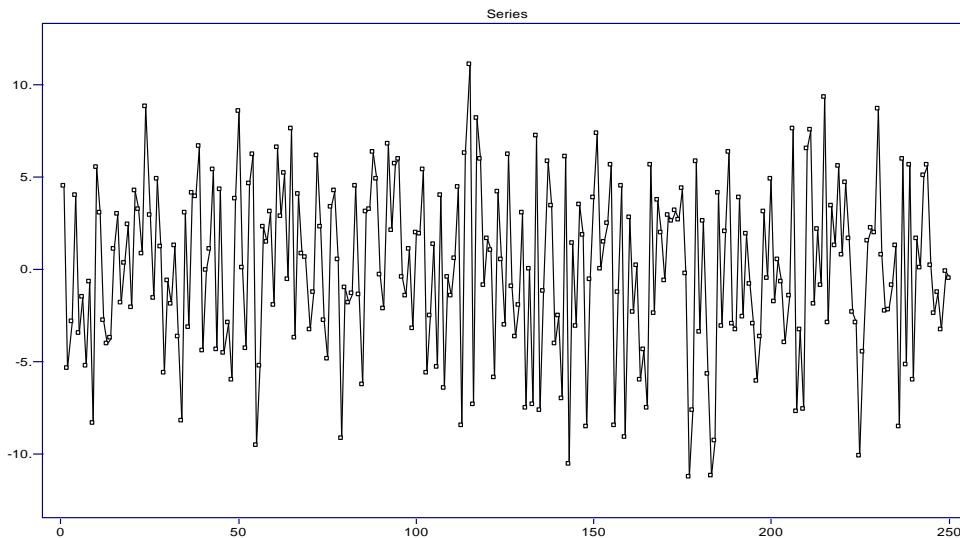
$$P[X_{n+h} \leq x | X_1 = x_1, \dots, X_n = x_n] = P[X_{n+h} \leq x]$$

υποδεικνύοντας πως η γνώση των X_1, \dots, X_n δεν συνεισφέρει στην πρόβλεψη του X_{n+h} .

Λευκός θόρυβος (White noise)

Είναι μία οικογένεια από ασυσχέτιστες τυχαίες μεταβλητές X_t , $t \in T$ με μέση τιμή 0 και διασπορά σ^2 .

$$w_t \sim wn(0, \sigma_w^2)$$



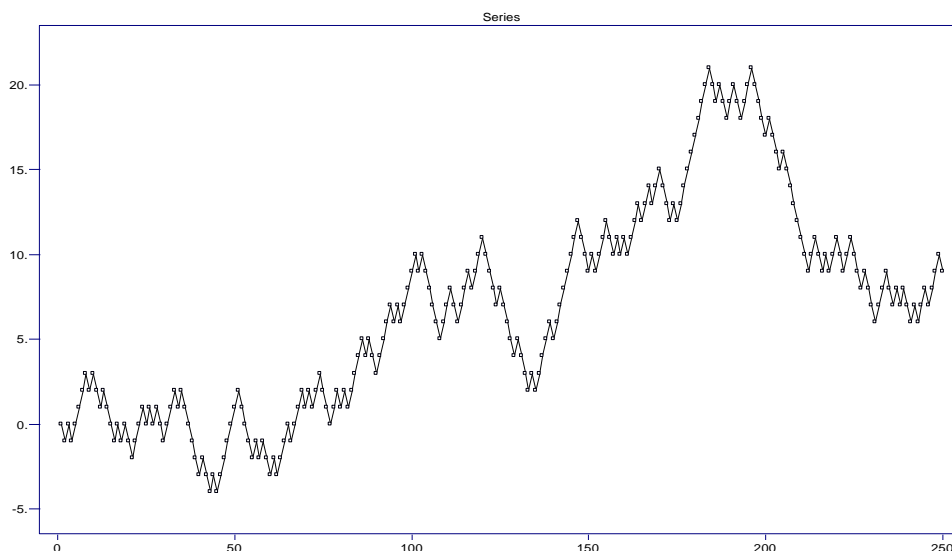
Σχήμα 4 Διαδικασία $WN(0, \sigma^2)$, με white noise variance 19

Τυχαίος περίπατος (Random Walk)

Ο τυχαίος περίπατος $\{S_t: t = 0, 1, 2, \dots\}$ (που ξεκινάει από το μηδέν) προκύπτει αθροίζοντας ανεξάρτητες και ισόνομες τυχαίες μεταβλητές. Έτσι, ένας τυχαίος περίπατος προκύπτει ορίζοντας $S_0 = 0$ και

$$S_n = X_1 + X_2 + \dots + X_n, \quad \text{για } n = 1, 2, \dots,$$

Όπου $\{X_n, n \in \mathbb{N}\}$ είναι ανεξάρτητες και ισόνομες τ.μ. (iid).



4.2 Στασιμότητα

Σημαντικό ρόλο στην Ανάλυση Χρονοσειρών διαδραματίζουν διαδικασίες των οποίων κάποια βασικά χαρακτηριστικά δε μεταβάλλονται με το χρόνο. Προκειμένου να προβλέψουμε μελλοντικές τιμές για μία χρονοσειρά, είναι απαραίτητο να υποθέσουμε ότι κάποιες ιδιότητες ή δομές στις παρατηρήσεις θα συνεχίσουν να υπάρχουν και στο μέλλον. Για το λόγο αυτό, εισάγουμε την έννοια της στασιμότητας, αφού αναφέρουμε κάποιους βασικούς ορισμούς.

Έστω $\{X_t, t \in T\}$ χρονοσειρά με $E(X_t^2) < \infty$. Τότε,

Η **συνάρτηση μέσης τιμής** είναι

$$\mu_X(t) = E(X_t).$$

Η **συνάρτηση συνδιακύμανσης** της $\{X_t\}$ είναι

$$\gamma_X(r, s) = \text{Cov}(X_r, X_s) = E[(X_r - \mu_X(r))(X_s - \mu_X(s))]$$

για όλους τους ακέραιους r, s .

Η χρονοσειρά $\{X_t: t \in T\}$ ονομάζεται **(ασθενώς) στάσιμη** αν

- i) $\mu_X(t) = \mu$ για κάθε $t \in T$.
- ii) $\gamma_X(t+h, t) = \gamma(h)$, για κάθε $h \in T$.

Η **συνάρτηση αυτοσυνδιακύμανσης** (auto covariance function ή ACVF) μιας στάσιμης χρονοσειράς $\{X_t, t \in T\}$ στην υστέρηση h είναι

$$\gamma_X(h) = \text{Cov}(X_{t+h}, X_t).$$

Η **συνάρτηση αυτοσυσχέτισης** (autocorrelation function ή ACF) μιας στάσιμης χρονοσειράς στην υστέρηση h είναι

$$\rho(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \text{Cor}(X_{t+h}, X_t), \quad h \in T.$$

Σημειώνεται πως $|\rho(h)| \leq 1$.

Οι ορισμοί αυτοί αφορούν μοντέλα χρονοσειρών. Στην πράξη ξεκινάμε από ένα σύνολο παρατηρήσεων και αναζητούμε ένα πιθανό μοντέλο που να «εξηγεί» τη συμπεριφορά των παρατηρήσεων. Προκειμένου να ποσοτικοποιήσουμε το βαθμό εξάρτησης μεταξύ των δεδομένων χρησιμοποιούμε κυρίως τη δειγματική συνάρτηση αυτοσυνδιακύμανσης.

Αν x_1, x_2, \dots, x_n είναι οι παρατηρήσεις μιας χρονοσειράς, τότε:

Ο **δειγματικός μέσος** των x_1, x_2, \dots, x_n είναι

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t.$$

Η **δειγματική συνάρτηση αυτοσυνδιακύμανσης** είναι

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}), \quad -n < h < n.$$

Η **δειγματική συνάρτηση αυτοσυσχέτισης** είναι

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}, \quad -n < h < n.$$

Η χρονοσειρά $\{X_t: t \in T\}$ ονομάζεται «**αυστηρά στάσιμη**» όταν

$$(X_1, \dots, X_n) \sim (X_{1+h}, \dots, X_{n+h}), \quad \forall n \in N$$

δηλαδή οι (X_1, \dots, X_n) και $(X_{1+h}, \dots, X_{n+h})$ έχουν τις ίδιες κατανομές για όλα τα $h \in T$.

Συνεπώς οι κατανομές πεπερασμένης διάστασης αυστηρώς στάσιμων χρονοσειρών παραμένουν αναλλοίωτες σε χρονικές μεταθέσεις.

4.3 Τάση και εποχικότητα

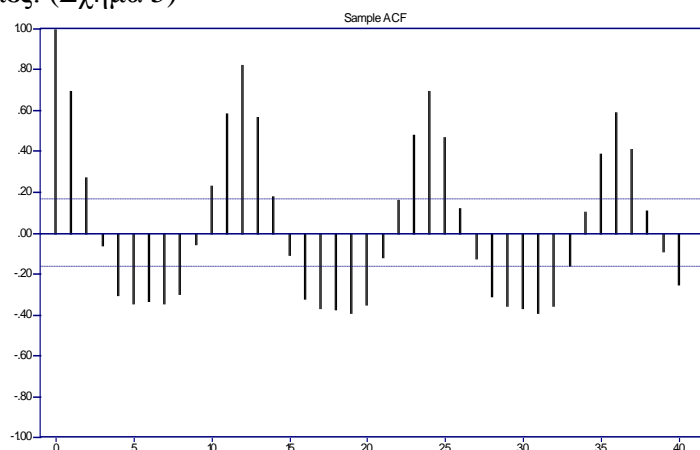
Πριν προσπαθήσουμε να αναζητήσουμε ένα κατάλληλο μοντέλο για τα δεδομένα μας, πρέπει να αναφερθούμε στη σημασία της στασιμότητας. Οι περισσότερες μέθοδοι

μοντελοποίησης χρονοσειρών εφαρμόζονται σε χρονοσειρές που έχουν (ασθενώς) στάσιμη συμπεριφορά.

Πολλές φορές κατά την επεξεργασία δεδομένων παρατηρείται το φαινόμενο η μέση τιμή τους να παρουσιάζει μία αυξητική ή φθίνουσα τάση. Επίσης τα δεδομένα μπορεί να παρουσιάζουν κάποια περιοδική συμπεριφορά με αυξητικές φάσεις και φθίνουσες φάσεις σε διαδοχικά χρονικά διαστήματα ίσης απόστασης μεταξύ τους. Τέτοια παραδείγματα έχουμε σε μετεωρολογικά ή οικονομικά στοιχεία που επηρεάζονται από την εποχή, την ηλιοφάνεια κτλ.

Ο έλεγχος για την ύπαρξη εποχικότητας στα δεδομένα γίνεται ως εξής:

- Παρατηρούμε το γράφημα των παρατηρήσεων στο χρόνο και ελέγχουμε οπτικά για την ύπαρξη συστηματικών (περιοδικών) «παλμών», είτε αύξησης, είτε μείωσης. Από το γράφημα επίσης μπορούμε συχνά να αναγνωρίσουμε και την περίοδο d . Για παράδειγμα, στην περίπτωση που έχουμε δεδομένα που εκφράζουν μηνιαίες πωλήσεις, ενδέχεται η περίοδος να είναι ετήσια όπως στο Σχήμα 1 των αεροπορικών επιβατών, εξαμηνιαία ή άλλη.
- Σχεδιάζουμε τις ACF των παρατηρήσεων για υστέρηση 1,2,... (correlogram). Αναζητούμε έντονες αυτοσυσχετίσεις που πιθανώς επαναλαμβάνονται κάθε d χρονικές στιγμές. (Σχήμα 5)



Σχήμα 5 ACF των αεροπορικών επιβατών (χωρίς τάση)

Πρέπει λοιπόν, με κατάλληλους μετασχηματισμούς και μεθόδους να εξαλείψουμε αυτά τα χαρακτηριστικά ώστε να προκύψει στο τέλος μία στάσιμη χρονοσειρά κατάλληλη για επεξεργασία με τις μεθόδους που θα περιγράψουμε στη συνέχεια. Αυτό σημαίνει πως από την αρχική χρονοσειρά, ύστερα από αυτούς τους μετασχηματισμούς και την εξάλειψη της εποχικότητας που τυχόν υπάρχει, θα προκύψει μία άλλη χρονοσειρά προς επεξεργασία. Αυτήν την τελική χρονοσειρά προσπαθούμε να μοντελοποιήσουμε, και αφού βρούμε ένα ικανοποιητικό μοντέλο για αυτήν, αντιστρέφουμε τον αρχικό μετασχηματισμό και έτσι προκύπτει ένα μοντέλο για την αρχική χρονοσειρά.

Για το σκοπό αυτό μπορούμε να αναπαραστήσουμε μία χρονοσειρά ως εξής:

$$X_t = m_t + s_t + Y_t, \quad t \in T.$$

- m_t είναι μία συνιστώσα που εκφράζει την τάση,
- s_t είναι μία περιοδική συνάρτηση με περίοδο d που εκφράζει την εποχικότητα,
- Y_t είναι μία στάσιμη χρονοσειρά και αποτελεί το θόρυβο.

Με τους κατάλληλους μετασχηματισμούς και την αφαίρεση της τάσης και της εποχικότητας προκύπτει η $\{Y_t: t \in T\}$. Δηλαδή παίρνουμε την

$$Y_t = X_t - m_t - s_t, \quad t \in T$$

Απάλειψη των διακυμάνσεων

Αρχικά ελέγχουμε το γράφημα της χρονοσειράς για να εξακριβώσουμε αν η μεταβλητότητα των δεδομένων παρουσιάζει αυξητική ή φθίνουσα τάση κατά την εξέλιξη στο χρόνο. Σε αυτήν την περίπτωση ίσως είναι χρήσιμο να λογαριθμήσουμε τα αρχικά δεδομένα, ή να πάρουμε την τετραγωνική ρίζα τους για να εξαλείψουμε το φαινόμενο.

Κάνοντας χρήση ενός μετασχηματισμού, μπορούμε να εξομαλύνουμε τις διακυμάνσεις στα δεδομένα, αν αυτό κριθεί απαραίτητο. Ένας τέτοιος μετασχηματισμός είναι ο μετασχηματισμός f_λ των Box-Cox.

$$f_\lambda(U_t) = \begin{cases} \lambda^{-1}(U_t^\lambda - 1), & U_t \geq 0, \lambda > 0 \\ \ln U_t, & U_t > 0, \lambda = 0 \end{cases}$$

Βλέπουμε πως για $\lambda=0$ ο μετασχηματισμός είναι ίδιος με το λογάριθμο, ενώ για τιμές διάφορες του μηδενός, τα δεδομένα εξομαλύνονται πιο ομαλά. Συνήθως επιλέγουμε $0 \leq \lambda \leq 1,5$.

4.4 Εκτίμηση της τάσης

Στο γενικό μοντέλο αναπαράστασης μιας χρονοσειράς χωρίς την εποχική συνιστώσα

$$X_t = m_t + Y_t, \quad t \in N$$

η τάση m_t είναι συνήθως μία συνάρτηση του χρόνου με χαμηλές μεταβολές. Για παράδειγμα μπορεί να είναι γραμμική ή χαμηλού βαθμού πολυώνυμο. Ακόμα, θεωρούμε $E(Y_t) = 0$. Αν $E(Y_t) \neq 0$ τότε αντικαθιστούμε τη m_t με $m_t - E(Y_t)$ και την Y_t με $Y_t - E(Y_t)$. Για να απαλείψουμε την τάση αυτή χρησιμοποιούμε τις εξής μεθόδους.

Παλινδρόμηση ως προς το χρόνο

Χρησιμοποιούμε τη μέθοδο ελαχίστων τετραγώνων για να βρούμε μία $m_t = a + bt$ ελαχιστοποιώντας το άθροισμα $\sum_{t=1}^n (x_t - a - bt)^2$. Παίρνουμε την $\bar{m}_t = \bar{a} + \bar{b}t$.

Ανάλογα αποτελέσματα προκύπτουν και για πολυωνυμικές τάσεις.

Εξομάλυνση κινητού μέσου

Σε αυτήν την περίπτωση μπορούμε να εκτιμήσουμε την τάση m_t χρησιμοποιώντας τη μέθοδο εξομάλυνσης κινητού μέσου. Για $1 \leq q \ll n$ εξομαλύνουμε τη $X_t = m_t + Y_t, t = 1, 2, \dots, n$ από την

$$W_t = \frac{1}{2q+1} \sum_{t=-q}^q X_{t-j} = \frac{1}{2q+1} \sum_{t=-q}^q m_{t-j} + \frac{1}{2q+1} \sum_{t=-q}^q Y_{t-j} \approx m_t, q+1 \leq t \leq n-q$$

Θεωρώντας ότι η m_t είναι περίπου γραμμική στο διάστημα $[t-q, t+q]$. Προκύπτουν οι εκτιμήσεις

$$\widehat{m}_t = \frac{1}{2q+1} \sum_{j=-q}^q X_{t-j}, q+1 \leq t \leq n-q.$$

Μέθοδος των διαφορών

Προκειμένου να εξαλείψουμε την τάση m_t που υπάρχει σε μία χρονοσειρά παρατηρήσεων $\{x_t\}$ μπορούμε να πάρουμε τις διαφορές διαδοχικών όρων της χρονοσειράς ως εξής:

$$v_t = x_t - x_{t-1}, \quad t = 1, 2, \dots, n$$

Έτσι, όταν η τάση είναι γραμμική, η νέα χρονοσειρά v_t έχει μηδενική τάση. Στην περίπτωση που η τάση είναι πολυωνυμική, η διαδικασία των διαφορών επαναλαμβάνεται τόσες φορές, όσο η τάξη του πολυωνύμου.

Ορίζουμε τους τελεστές:

Οπισθοδρομικός τελεστής: $Bx_t = x_{t-1}$

Τελεστής διαφοράς: $\nabla x_t = x_t - x_{t-1} = (1 - B)x_t$

Αν εφαρμοστεί ο τελεστής διαφοράς k φορές τότε προκύπτει

$$\nabla^k x_t = (I - B)^k x_t = \sum_{l=0}^k \binom{k}{l} (-1)^l x_{t-l}.$$

Η σταδιακή εφαρμογή του τελεστή ∇ οδηγεί στην εξάλειψη της τάσης m_t . Κατά την εφαρμογή του τελεστή είναι αναγκαίος ο έλεγχος της γραφικής παράστασης της χρονοσειράς που προκύπτει ώστε να επιβεβαιώσουμε πως έχουμε επιτύχει τη στασιμότητα. Αν χρειαστεί, εφαρμόζουμε τον τελεστή, μία ή δύο φορές ακόμα.

Σε αυτό το σημείο είναι χρήσιμο να τονίσουμε τη διαφορά μεταξύ ντετερμινιστικών και στοχαστικών τάσεων. Ντετερμινιστική τάση θεωρείται η συνιστώσα m_t που μπορεί να προσεγγιστεί μέσω κάποιας συνάρτησης του χρόνου, όπως πολυώνυμο, εκθετική

συνάρτηση, κ.α.. Στοχαστική τάση θεωρείται οποιαδήποτε τάση m_t δεν μπορεί να προσδιοριστεί με τη βοήθεια κάποιας απλής, σχετικά, συνάρτησης, και συνεπώς θα πρέπει να θεωρηθεί ως μία τυχαία, συνήθως αργή μεταβολή. Στην πρώτη περίπτωση προσπαθούμε να εκτιμήσουμε τη συνιστώσα της τάσης με παλινδρόμηση ή με διαδοχικές διαφορές τάξης 1, ενώ στη δεύτερη περίπτωση απαλείφουμε τη συνιστώσα μέσω της διαδικασίας κινητού μέσου, ή πάλι με διαφορές. Η μέθοδος των διαφορών είναι απλούστερη καθώς δεν απαιτεί τον υπολογισμό παραμέτρων.

4.5 Εκτίμηση της τάσης και της εποχικότητας

Αν στο γενικό μοντέλο $X_t = m_t + s_t + Y_t$, $t \in R$ είναι προφανής η ύπαρξη εποχικότητας, δηλαδή υπάρχει ένα $1 < d < n$ με $s_t = s_{t+d}$, τότε κατά την εκτίμηση της τάσης m_t πρέπει να λάβουμε υπόψη μας και την εποχική συνιστώσα s_t . Η εκτίμηση των δύο συνιστωσών τάσης και εποχικότητας γίνεται είτε με επαναληπτικές εκτιμήσεις, είτε με μία διαδικασία ταυτόχρονου προσδιορισμού και των δύο.

Εποχικοί συντελεστές

Ας πάρουμε για παράδειγμα μία χρονοσειρά $\{x_t\}$, $t = 1, 2, \dots, n$ όπου οι μετρήσεις έχουν γίνει σε μηνιαία βάση και ότι το αρχικό γράφημα της χρονοσειράς παρουσιάζει έντονη ετήσια περιοδικότητα, δηλαδή $d=12$ και ότι η συνιστώσα τάσης m_t παρουσιάζει χαμηλό βαθμό μεταβολής. Τότε η εποχική συνιστώσα s_t μπορεί να εκτιμηθεί ως εξής:

Εκτιμούμε την m_k , για το k έτος από την

$$\bar{m}_k = \frac{1}{d} \sum_{l=1}^d x_{k,l}, \quad k = 1, \dots, K \text{ όπου } K \text{ το πλήθος των ετών της χρονοσειράς.}$$

Τώρα υπολογίζουμε τις εποχικές συνιστώσες s_l , $l = 1, \dots, d$ ως τις μέσες αποκλίσεις από τη μέση τιμή του έτους. Δηλαδή,

$$\bar{s}_l = \frac{1}{K} \sum_{k=1}^K (x_{k,l} - \bar{m}_k) \text{ με } l = 1, \dots, d \text{ και } \sum_{l=1}^d \bar{s}_l = 0.$$

Έτσι προκύπτει η νέα χρονοσειρά χωρίς τάση και περιοδικότητα

$$\bar{Y}_t = x_{k,l} - \bar{m}_k - \bar{s}_l, \quad k = 1, \dots, K, \quad l = 1, \dots, 12.$$

Προσαρμογή περιοδικής συνάρτησης

Αν η περίοδος της εποχικής συνιστώσας είναι d , μπορούμε να την προσεγγίσουμε με κάποια γνωστή παραμετρική περιοδική συνάρτηση $s_t = f(t)$, όπως π.χ. μια ημιτονοειδή συνάρτηση. Συχνά όμως η μορφή της περιοδικής συνάρτησης δεν είναι συγκεκριμένη. Έτσι ένας απλός τρόπος εκτίμησης της s_t είναι από τους μέσους όρους των στοιχείων της περιοδικής συνάρτησης s_i , $i = 1, \dots, d$. Αν n είναι το πλήθος των παρατηρήσεων, τότε

$k = \frac{n}{d}$ είναι ο αριθμός των περιόδων στη χρονοσειρά $\{x_1, x_2, \dots, x_n\}$. Συνεπώς κάθε στοιχείο της περιοδικής συνάρτησης εκτιμάται από την εξής τιμή:

$$\bar{s}_t = \frac{1}{k} \sum_{j=1}^k y_{i+jd}$$

Μέθοδος διαφορών με υστέρηση d

Εφαρμόζουμε στην αρχική χρονοσειρά τη διαφορά $x_t - x_{t-d}$, δηλαδή τον τελεστή

$\nabla_d = 1 - B^d$ και προκύπτει η χωρίς εποχική συνιστώσα χρονοσειρά

$$\nabla_d x_t = m_t - m_{t-d} + y_t - y_{t-d},$$

στην οποία ύστερα εφαρμόζουμε τον τελεστή διαφορών τάξης k , για κατάλληλο k , ώστε να απαλειφθεί και η συνιστώσα τάσης m_t .

Προφανώς δεν επιθυμούμε μεγάλα διαστήματα προβλέψεων και γι' αυτόν το λόγο πρέπει να προσέχουμε τις διαφορές. Πρέπει να παίρνουμε όσο το δυνατό λιγότερες διαφορές για να μετασχηματίσουμε τα δεδομένα.

4.6 Αυτοπαλινδρομική Διαδικασία Κινητού Μέσου ARMA(p,q)

Μία χρονοσειρά $\{X_t, t = 0, \pm 1, \pm 2, \dots\}$ λέγεται **ARMA(p,q)** αν είναι στάσιμη και για κάθε t ,

$$X_t - \varphi_1 X_{t-1} - \varphi_2 X_{t-2} - \dots - \varphi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_q Z_{t-q},$$

όπου $Z_t \sim WN(0, \sigma^2)$, δηλαδή τα Z_t είναι τ.μ. λευκού θορύβου και τα πολυώνυμα $(1 - \varphi_1 z - \dots - \varphi_p z^p)$ και $(1 + \theta_1 z + \dots + \theta_q z^q)$ δεν έχουν κοινές ρίζες. Αν έχουν κοινές ρίζες, τα πολυώνυμα απλοποιούνται και ελαττώνονται οι αντίστοιχοι βαθμοί τους.

Μία διαδικασία ARMA(p,q) μπορεί να γραφεί ως εξής:

$$\varphi_p(B)X_t = \theta_q(B)Z_t,$$

όπου $\theta(z)$, $\varphi(z)$, $z \in \mathbb{C}$ πολυώνυμα βαθμού p , q αντίστοιχα και B ο οπισθοδρομικός τελεστής.

Αιτιατότητα και Αντιστρεψιμότητα

Μία διαδικασία ARMA(p,q) $\{X_t\}$ ονομάζεται **αιτιατή (causal)** συνάρτηση του $\{Z_t\}$ αν

υπάρχουν σταθερές $\{\psi_i, i \in \mathbb{Z}^+\}$ τέτοια ώστε $\sum_{j=0}^{\infty} |\psi_j| < \infty$ και

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \text{ για κάθε } t.$$

Η αιτιατότητα είναι ισοδύναμη με τη συνθήκη

$$\varphi(z) = 1 - \varphi_1 z - \dots - \varphi_p z^p \neq 0 \text{ για κάθε } |z| \leq 1.$$

Αποδεικνύεται ότι μία διαδικασία ARMA(1,1) είναι αιτιατή αν και μόνο αν $|\varphi_1| < 1$.

Μία διαδικασία ARMA(p,q) ονομάζεται **αντιστρέψιμη (invertible)** αν υπάρχουν σταθερές

$$\{\pi_j, j \in Z^+\} \text{ τέτοιες ώστε } \sum_{j=0}^{\infty} |\pi_j| < \infty \text{ και } Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \forall t.$$

Η αντιστρεψιμότητα είναι ισοδύναμη με τη συνθήκη

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q \neq 0 \text{ για κάθε } |z| \leq 1.$$

Η διαδικασία AR(p)

Όταν $\theta(z)=1$ και $\varphi(B)X_t = Z_t$, με $\varphi(z) = 1 - \varphi_1 z - \dots - \varphi_p z^p$ τότε λέμε ότι η διαδικασία είναι τύπου AR(p) ή *αυτοπαλινδρομική διαδικασία τάξης p*.

Η διαδικασία MA(q)

Όταν $\varphi(z)=1$ και $X_t = \theta(B)Z_t$, με $\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q$, τότε λέμε ότι η διαδικασία είναι τύπου MA(q) ή *διαδικασία κινητού μέσου τάξης q*.

4.7 Κριτήρια επιλογής μοντέλου

Κατά τη διαδικασία επιλογής ενός θεωρητικού μοντέλου ARMA(p,q) που να περιγράφει τα δεδομένα μας, είναι δυνατόν να προκύψουν μοντέλα διαφορετικής τάξης. Δεν επιθυμούμε τη χρησιμοποίηση μοντέλων πολύ υψηλής τάξης γιατί αυτά οδηγούν σε overfitting. Επιπλέον, όσο μεγαλύτερη η τάξη του μοντέλου, τόσες περισσότερες παραμέτρους έχουμε να υπολογίσουμε και κατά συνέπεια περισσότερες πηγές σφαλμάτων. Άρα είναι λογικό να αναζητήσουμε κάποια κριτήρια τα οποία να «προτιμούν» μικρότερα σε τάξη μοντέλα, έναντι πιο πολύπλοκων. Επιπλέον, τα κριτήρια αυτά πρέπει να δίνουν τον ελάχιστο δυνατό βαθμό μεταβλητότητας στο λευκό θόρυβο.

Το κριτήριο AIC

Το κριτήριο πληροφορίας Akaike (AIC) είναι το συνηθέστερο κριτήριο που χρησιμοποιείται εκτενώς από τα στατιστικά προγράμματα H/Y. Μεταξύ υποψήφιων μοντέλων επιλέγεται αυτό με το ελάχιστο AIC. Το κριτήριο αποθαρρύνει τη χρήση μοντέλων υψηλής τάξης.

$$AIC = -2 \ln(L) + 2(p + q + 1),$$

όπου L είναι η μέγιστη τιμή της πιθανοφάνειας του επιλεγμένου μοντέλου.

Το κριτήριο AICC (AIC corrected)

Το κριτήριο αυτό είναι παρόμοιο με το AIC αλλά με μεγαλύτερη «ποινή» για μεγάλα σε τάξη μοντέλα.

$$AICC = AIC + \frac{2(p + q + 1)n}{n - p - q - 2}$$

Αναζητούμε το μοντέλο με την ελάχιστη τιμή για το AICC.

Το κριτήριο BIC (Schwarz)

Ένα άλλο κριτήριο επιλογής μοντέλου, είναι το BIC (ή αλλιώς SBC, ή SBIC), το οποίο είναι παρόμοιο με το AIC και ορίζεται ως:

$$BIC = -2 \ln(L) + (p + q + 1) \ln(n),$$

όπου L είναι η μέγιστη τιμή της πιθανοφάνειας του επιλεγμένου μοντέλου. Όπως και στις προηγούμενες περιπτώσεις αναζητούμε το μοντέλο με την ελάχιστη τιμή για το BIC.

4.8 Έλεγχος των υπολοίπων

Για να αποδεχθούμε ότι ένα μοντέλο περιγράφει ικανοποιητικά τα δεδομένα, είναι αναγκαίο να εξετάσουμε τη συμπεριφορά των υπολοίπων (residuals). Αν η χρονοσειρά που εξετάζουμε μπορεί να μοντελοποιηθεί μέσω κάποιας διαδικασίας ARMA(p,q), τότε τα υπόλοιπα πρέπει να είναι ασυσχέτιστα, ανεξάρτητα, και κατανομημένα ανάλογα με την ακολουθία $\{Z_t, t = 0, 1, \dots, n\}$ που δημιουργεί τη διαδικασία ARMA(p,q) που εφαρμόστηκε.

Παρακάτω εξετάζουμε μερικούς τρόπους ελέγχου των υπολοίπων.

Η δειγματική ACF των υπολοίπων.

Για αρκετά μεγάλο n οι δειγματικές αυτοσυσχετίσεις μιας ακολουθίας iid Y_1, \dots, Y_n με πεπερασμένη διασπορά, είναι και αυτές iid κατά προσέγγιση με κατανομή $N(0, 1/n)$. Συνεπώς, αν y_1, \dots, y_n είναι μια πραγματοποίηση ή τροχιά μιας τέτοιας ακολουθίας iid τότε οι δειγματικές αυτοσυσχετίσεις θα πρέπει να περικλείονται μέσα στα όρια $\pm 1,96/\sqrt{n}$. Έτσι μπορούμε να εξετάσουμε το γράφημα αυτής της δειγματικής συνάρτησης αυτοσυσχετίσης και να ελέγξουμε οπτικά αν κάποια σημεία ξεφεύγουν από τα όρια αυτά.

Το τεστ portmanteau

Παρόμοια διαδικασία ελέγχου πραγματοποιείται εξετάζοντας την ποσότητα

$$Q = n \sum_{j=1}^h \hat{\rho}^2(j).$$

Αν το Q είναι αρκετά μεγάλο, τότε οι αυτοσυσχετίσεις είναι πολύ μεγάλες ώστε να θεωρηθούν iid. Έτσι απορρίπτουμε την υπόθεση iid στο επίπεδο σημαντικότητας α αν $Q > \chi_{1-\alpha}^2(h)$, όπου $\chi_{1-\alpha}^2(h)$ το 1- α ποσοστιαίο σημείο της χ -τετράγωνο κατανομής με h βαθμούς ελευθερίας.

Turning point test

Αν y_1, \dots, y_n είναι οι παρατηρήσεις, τότε ορίζουμε ως «σημείο αλλαγής» (turning point) το χρόνο i, $1 < i < n$

αν $y_{i-1} < y_i$ και $y_i > y_{i+1}$ ή αν $y_{i-1} > y_i$ και $y_i < y_{i+1}$.

Αν T είναι το πλήθος των σημείων αλλαγής μιας ακολουθίας iid μήκους n , η αναμενόμενη τιμή του T είναι $E(T)=2(n-2)/3$ και $\sigma^2(T)=(16n-29)/90$. Ελέγχοντας τη διαφορά $T-E(T)$ κρίνουμε αν η διαδικασία μεταβάλλεται γρηγορότερα από μία διαδικασία iid. Μπορούμε να πούμε ότι $T \sim N(\mu_T, \sigma_T^2)$ και συνεπώς απορρίπτουμε την υπόθεση iid στο επίπεδο α αν $\frac{|T-\mu_T|}{\sigma(T)} > \Phi_{1-\frac{\alpha}{2}}$.

Έλεγχος πρόσημου της διαφοράς

Ορίζουμε S το πλήθος των χρονικών στιγμών i κατά τις οποίες παρατηρείται

$$y_i > y_{i-1} \text{ ή αλλιώς } y_i - y_{i-1} > 0.$$

Για διαδικασίες iid αναμένουμε μέση τιμή $E(S) = \frac{1}{2}(n-1)$ $\sigma^2(S) = \frac{n+1}{12}$. Αν η διαφορά $S-E(S)$ είναι μεγάλη κατά απόλυτη τιμή τότε ενδέχεται να υπάρχει κάποια αύξουσα ή φθίνουσα τάση στα υπόλοιπα και μπορούμε να απορρίψουμε την αρχική υπόθεση ότι τα υπόλοιπα είναι τυχαία. Επειδή είναι $S \sim N(\mu_S, \sigma_S^2)$ απορρίπτουμε την υπόθεση της ύπαρξης τάσης αν $\frac{|S-\mu_S|}{\sigma_S} > \Phi_{1-\frac{\alpha}{2}}$.

Rank test

Ορίζουμε P το πλήθος των ζευγών (i,j) για τα οποία

$$y_j > y_i \text{ και } j > i, i = 1, \dots, n-1$$

Είναι $E(P) = \frac{1}{4}n(n-1)$ και συνεπώς μία μεγάλη θετική (αρνητική) τιμή για τη διαφορά $P-E(P)$ οδηγεί στο συμπέρασμα πως ενδέχεται να υπάρχει αυξητική (αρνητική) τάση στα υπόλοιπα.

Προσαρμογή αυτοπαλινδρομικού μοντέλου

Αν κατά την προσπάθεια προσαρμογής μοντέλου στα υπόλοιπα με το ελάχιστο AICC καταλήξει σε κάποιο μοντέλο μηδενικής τάξης, τότε συμπεραίνουμε ότι τα δεδομένα είναι white noise.

Έλεγχος κανονικότητας και ανεξαρτησίας υπολοίπων

Σε αυτό το σημείο ελέγχουμε γραφικά ορισμένες ιδιότητες που είναι αναμενόμενες σε δεδομένα που ακολουθούν την κανονική κατανομή. Εξετάζουμε τη χρονοσειρά των υπολοίπων, το ιστόγραμμα, το qqplot, κ.α.

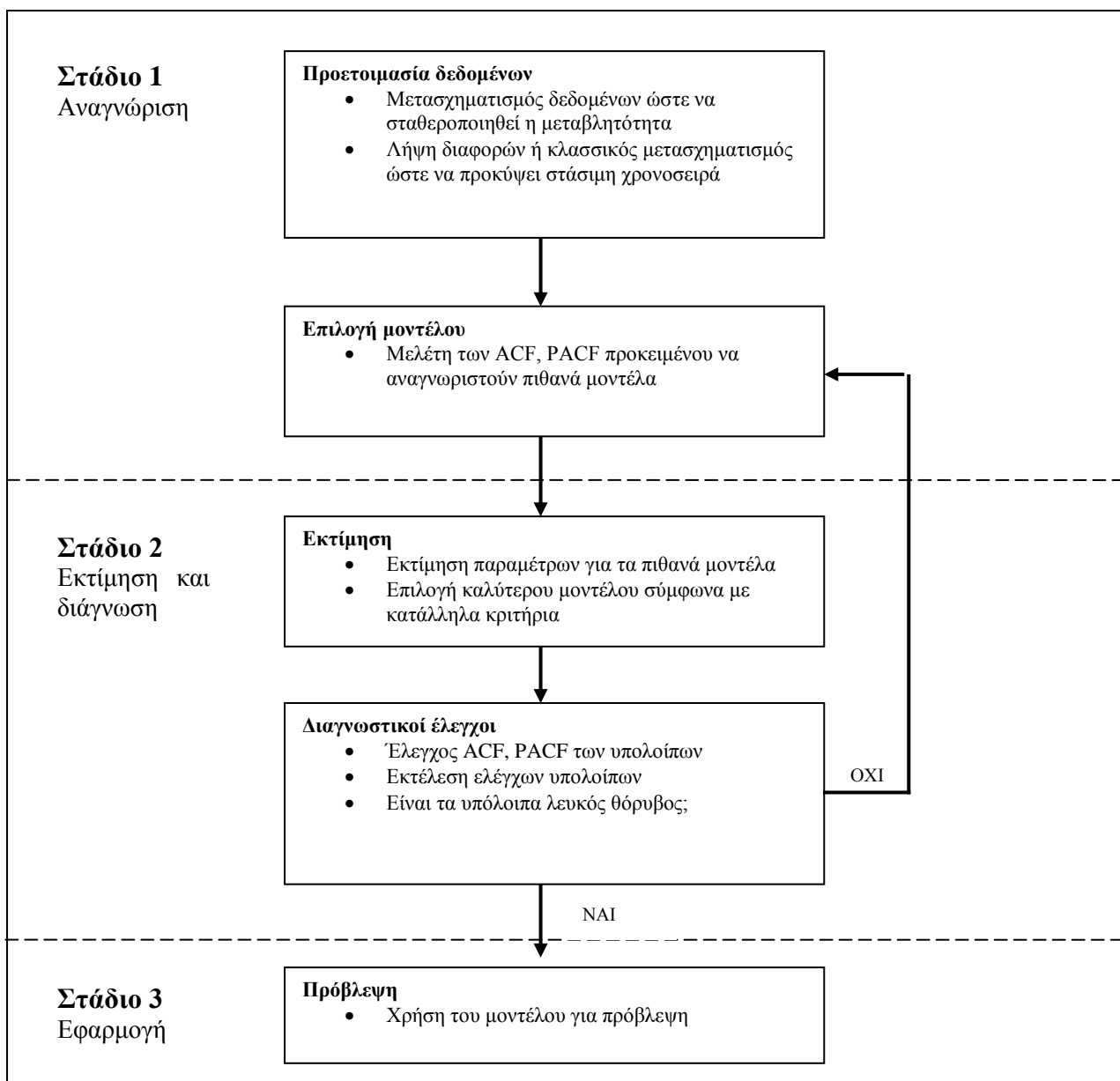
Συνήθως εκτελούμε όλους τους ελέγχους. Υπάρχει περίπτωση ένας ή περισσότεροι έλεγχοι να μας οδηγήσουν στην απόρριψη της υπόθεσης iid. Σε αυτήν την περίπτωση εξετάζουμε διαφορετικά μοντέλα ή ξεκινάμε ξανά τη διαδικασία εκτελώντας διαφορετικούς αρχικούς μετασχηματισμούς και προσπαθούμε να αφαιρέσουμε όσο το δυνατό περισσότερη «πληροφορία» από τα δεδομένα μας που οφείλονται στην ύπαρξη τάσης και εποχικότητας.

4.9 Επιλογή μοντέλου

Συνοψίζοντας μπορούμε να αναφέρουμε επιγραμματικά τα βήματα που ακολουθούμε για την επιλογή του κατάλληλου μοντέλου για τα δεδομένα προκειμένου να προχωρήσουμε στην πραγματοποίηση προβλέψεων.

- Εξετάζουμε το γράφημα της χρονοσειράς για την ύπαρξη μεταβλητότητας στην τιμή της διασποράς. Αν υπάρχει, τότε μπορούμε να μετασχηματίσουμε τα δεδομένα με το μετασχηματισμό Box-Cox.
- Εξετάζουμε γραφικά τις ACF, PACF για την ύπαρξη περιοδικότητας, αυξητικής ή φθίνουσας τάσης.
- Αν υπάρχουν, τότε εκτελούμε μετασχηματισμούς ή παίρνουμε τις διαφορές υστέρησης d για την εποχικότητα, τάξης k για την τάση.
- Εξετάζουμε γραφικά τις ACF, PACF των νέων δεδομένων για στασιμότητα
- Προσαρμόζουμε ένα μοντέλο AR(p), MA(q), ή ARMA(p,q) στα δεδομένα. Διαλέγουμε το καλύτερο χρησιμοποιώντας τα κριτήρια **Akaike** (AICC), **Schwartz** (BIC). Αναζητούμε μοντέλα χαμηλής τάξης για να αποφύγουμε πολύπλοκους υπολογισμούς και το over-fitting.
- Εξετάζουμε τα υπόλοιπα (residuals) ως προς την κανονικότητα με μεθόδους όπως normal probability plot, ιστόγραμμα κτλ και εκτελώντας τους ελέγχους υπολοίπων που αναφέρθηκαν παραπάνω.
- Αν το μοντέλο που επιλέξαμε δεν περνάει με επιτυχία τους ελέγχους τυχαιότητας τότε επιλέγουμε ένα άλλο μοντέλο και επαναλαμβάνουμε τους ελέγχους.
- Όταν σιγουρευτούμε ότι το μοντέλο περιγράφει ικανοποιητικά τα δεδομένα, τότε μπορούμε να προβούμε σε προβλέψεις ανασκευάζοντας το μοντέλο, συμπεριλαμβάνοντας τους αρχικούς μετασχηματισμούς.

Σχηματικά, μπορούμε να συνοψίσουμε τα παραπάνω ως εξής:



4.10 Προβλέψεις για στάσιμες χρονοσειρές

Πλήθος τεχνικών και αλγορίθμων έχουν αναπτυχθεί για να δημιουργήσουν προβλέψεις σε παρατηρήσεις που περιγράφονται από μαθηματικά μοντέλα ARMA. Στην πράξη σπάνια υπάρχουν χρονοσειρές που παράγονται από απλούς μαθηματικούς μηχανισμούς όπως οι διαδικασίες ARMA. Για την επιλογή της κατάλληλης μεθόδου απαιτείται συνήθως η σύγκριση μεταξύ διαφορετικών αλγορίθμων και των αποτελεσμάτων που παράγουν, τόσο αναφορικά με τα σφάλματα προσαρμογής όσο και με τα σφάλματα πρόβλεψης.

Αξιολόγηση προβλέψεων

Στην εισαγωγή είδαμε τους ορισμούς των MAD, MSE και MAPE. Οι προβλέψεις που δημιουργούμε επιθυμούμε να έχουν όσο το δυνατό μικρότερη απόκλιση από τις πραγματικές παρατηρήσεις, όταν γίνουν διαθέσιμες. Επίσης, αναζητούμε αμερόληπτες προβλέψεις, δηλαδή η αναμενόμενη τιμή του σφάλματος πρόβλεψης πρέπει να είναι

μηδέν. Μόνο τότε μπορούμε να πούμε ότι το μοντέλο πρόβλεψης είναι ολοκληρωμένο και ικανοποιητικό. Ένας εκτιμητής της μεροληψίας των προβλέψεων είναι το *tracking signal* του Trigg.

Αν e_t το παρατηρούμενο σφάλμα τη χρονική στιγμή t και $|e_t|$ η απόλυτη τιμή του, τότε ορίζουμε τα εξής:

$$E_t = \beta e_t + (1 - \beta)E_{t-1},$$

$$M_t = \beta |e_t| + (1 - \beta)M_{t-1},$$

δηλαδή το εξομαλυμένο σφάλμα και απόλυτο σφάλμα κατά τη χρονική στιγμή t για $0 < \beta < 1$. Το tracking signal είναι ο λόγος:

$$T_t = \left| \frac{E_t}{M_t} \right|$$

Αν οι προβλέψεις είναι αμερόληπτες τότε το εξομαλυμένο σφάλμα E_t θα είναι μικρό συγκριτικά με το εξομαλυμένο απόλυτο σφάλμα M_t . Συνεπώς, μεγάλη τιμή του tracking signal υποδεικνύει μεροληπτικές προβλέψεις, κάτι που με τη σειρά του υποδεικνύει ότι το επιλεγμένο μοντέλο πρόβλεψης είναι ακατάλληλο. Η τιμή του T_t που σηματοδοτεί σημαντική μεροληψία εξαρτάται από το συντελεστή εξομάλυνσης β . Για παράδειγμα, μία τιμή του $T_t > 0,51$ σηματοδοτεί την ύπαρξη μη τυχαίων σφαλμάτων για $\beta = 0,1$.

U-Statistic του Theil

Τα σχετικά σφάλματα δίνουν την ίδια βαρύτητα σε όλα τα σφάλματα, σε αντίθεση με το MSE που υψώνει τα σφάλματα στο τετράγωνο και δίνει περαιτέρω έμφαση στα μεγάλα σφάλματα. Ένα μέτρο που δίνει έμφαση στα μεγάλα σφάλματα και ταυτόχρονα προσφέρει μια σχετική σύγκριση μεταξύ του μοντέλου που επιλέξαμε και του «απλουστευτικού μοντέλου» (naïve model) είναι αυτό του Theil. Το απλουστευτικό μοντέλο πρόβλεψης είναι το απλούστερο δυνατό μοντέλο και χρησιμοποιεί την πιο πρόσφατη παρατήρηση ως πρόβλεψη για την αμέσως επόμενη.

Αν \hat{X}_t είναι η πρόβλεψη για την X_t τότε,

$$U = \sqrt{\frac{\sum_{t=1}^{n-1} (FPE_{t+1} - APE_{t+1})^2}{\sum_{t=1}^{n-1} (APE_{t+1})^2}}$$

Όπου,

$$FPE_{t+1} = \frac{\hat{X}_{t+1} - X_t}{X_t}, \text{ (σχετική μεταβολή πρόβλεψης) και}$$

$$APE_{t+1} = \frac{X_{t+1} - X_t}{X_t}, \text{ (πραγματική σχετική μεταβολή).}$$

Απλοποιώντας την παραπάνω προκύπτει:

$$U = \sqrt{\frac{\sum_{t=1}^{n-1} \left(\frac{\hat{X}_{t+1} - X_t - X_{t+1} + X_t}{X_t} \right)^2}{\sum_{t=1}^{n-1} \left(\frac{X_{t+1} - X_t}{X_t} \right)^2}} = \sqrt{\frac{\sum_{t=1}^{n-1} \left(\frac{\hat{X}_{t+1} - X_{t+1}}{X_t} \right)^2}{\sum_{t=1}^{n-1} \left(\frac{X_{t+1} - X_t}{X_t} \right)^2}}$$

Ο αριθμητής της τελευταίας σχέσης είναι παρόμοιος με το MAPE, ενώ ο παρονομαστής προσομοιάζει το MAPE του απλουστευτικού μοντέλου διότι στο μοντέλο αυτό η πρόβλεψη είναι η τελευταία παρατήρηση.

Για το U-statistic ισχύουν τα παρακάτω:

- $U=1$: η απλουστευτική μέθοδος είναι το ίδιο καλή με τη μέθοδο που αξιολογείται.
- $U>1$: η απλουστευτική μέθοδος είναι καλύτερη. Δεν υπάρχει λόγος να χρησιμοποιηθεί η μέθοδος που αξιολογείται.
- $U<1$: η μέθοδος που αξιολογείται είναι καλύτερη από την απλουστευτική. Όσο μικρότερη είναι η τιμή του U τόσο καλύτερο από την απλουστευτική μέθοδο είναι το μοντέλο που αξιολογείται.

Επιλογή απλού ή σύνθετου μοντέλου πρόβλεψης

Σύνθετα μοντέλα εννοούμε αυτά που απαιτούν την εκτίμηση πολλών παραμέτρων, ή που χρησιμοποιούν πολύπλοκες μαθηματικές έννοιες. Έχουν γίνει έρευνες για τη διαχρονική απόδοση πολύπλοκων και απλών μοντέλων. Υπήρχαν φορές που κάποιο απλό μοντέλο αποδείχθηκε ακριβέστερο από ένα πολύπλοκο, ειδικότερα όταν αυτά αφορούσαν μακροπρόθεσμες προβλέψεις (Armstrong, 1984). Γενικά, η επιλογή ενός μοντέλου πρόβλεψης είναι μια επίπονη διαδικασία κατά την οποία όλες οι παράμετροι πρέπει να λαμβάνονται υπόψη.

Υπάρχουν ενδείξεις ότι ο απλός αριθμητικός μέσος από προβλέψεις που προκύπτουν από διαφορετικά μοντέλα, είναι ακριβέστερος από ότι ένα μεμονωμένο απλό ή σύνθετο μοντέλο (Makridakis & Winkler, 1983). Αυτό ίσως έχει να κάνει με το γεγονός ότι ένα μεμονωμένο μοντέλο δεν είναι σε θέση να αντικατοπτρίσει ολόκληρη την πληροφορία που βρίσκεται ενσωματωμένη στα δεδομένα, ενώ διαφορετικά μοντέλα φέρνουν στην «επιφάνεια» διαφορετικά κομμάτια αυτής της πληροφορίας.

5 ΕΦΑΡΜΟΓΗ

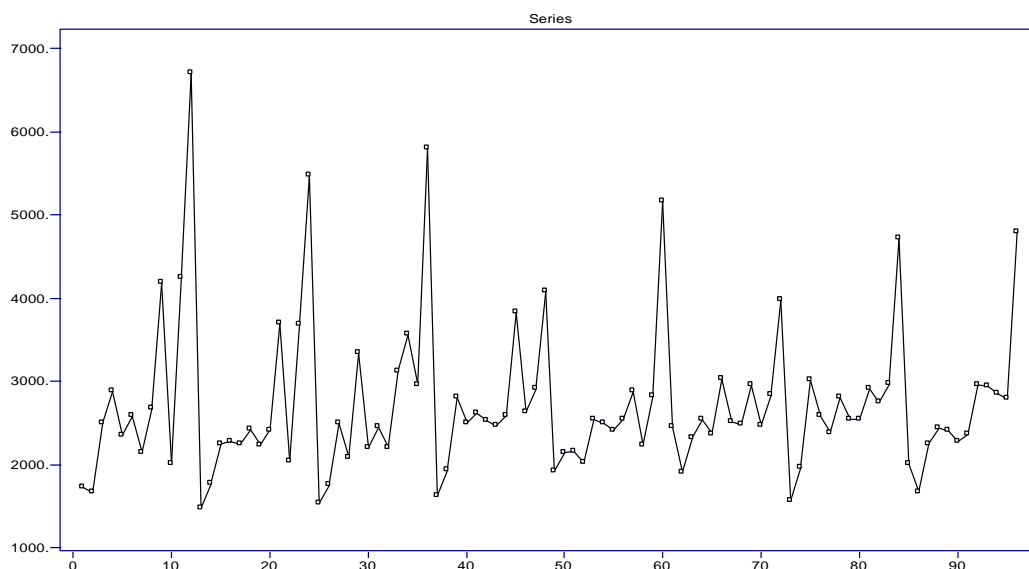
Στην παρούσα εργασία θα μελετήσουμε δεδομένα από πωλήσεις φαρμάκων στην περιοχή της Κοζάνης κατά την περίοδο 2004-2011. Από τη μελέτη αυτή προσπαθούμε να εξαγάγουμε συμπεράσματα για τη φύση της αγοράς του φαρμάκου, τον τρόπο διάθεσής του στα φαρμακεία και την εκτίμηση της μελλοντικής ζήτησης για τον έλεγχο των αποθεμάτων.

Έχουμε συμπεριλάβει διαφορετικές κατηγορίες φαρμάκων και εξετάζουμε αν αυτά παρουσιάζουν εποχικότητα λόγω της φύσης του φαρμάκου και της ασθένειας την οποία θεραπεύουν, αλλά και εξαιτίας άλλων φαινομένων τα οποία και θα μελετήσουμε. Επίσης θα δούμε τη διαχρονική συμπεριφορά των σκευασμάτων αυτών και, ειδικότερα, την όποια ανοδική ή καθοδική τάση παρουσιάζουν στις πωλήσεις.

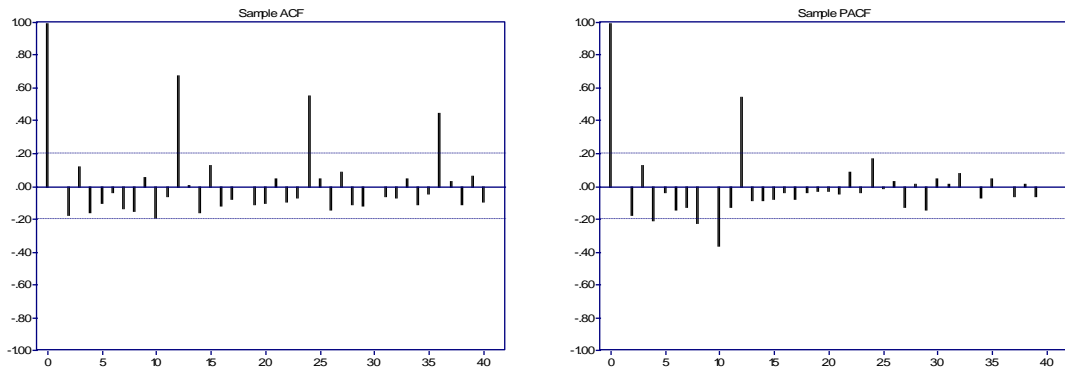
Επειδή τα φάρμακα είναι διατιμημένα προϊόντα από το κράτος και υπόκεινται σε ελέγχους τόσο ως προς την τιμή τους, όσο και ως προς τον ανταγωνισμό, δε θα εξετάσουμε το ρόλο που διαδραματίζει η τιμή στη ζήτησή τους, αλλά θα ασχοληθούμε μόνο με τις ποσότητες που διακινούνται από τη φαρμακαποθήκη προς τα φαρμακεία. Σε ορισμένες περιπτώσεις παρατηρείται εποχική συμπεριφορά, που είναι εντονότερη το Δεκέμβριο κάθε έτους και η οποία έχει να κάνει περισσότερο με οικονομικά κίνητρα των φαρμακοποιών, παρά με την έξαρση ασθενειών την περίοδο εκείνη.

5.1 Zantac

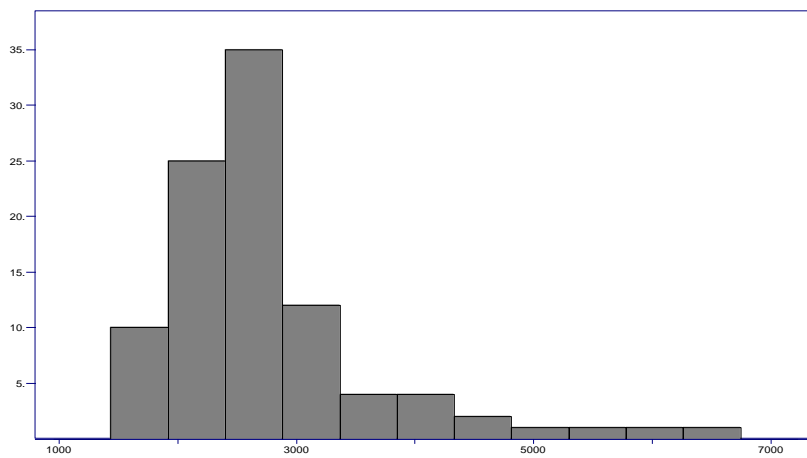
Πωλήσεις φαρμάκου Zantac στην πόλη της Κοζάνης από Ιαν 2004 έως Δεκ 2011 (96 μήνες).



Σχήμα 6. Πωλήσεις σε τεμάχια από Ιαν 2004 έως Δεκ 2011.

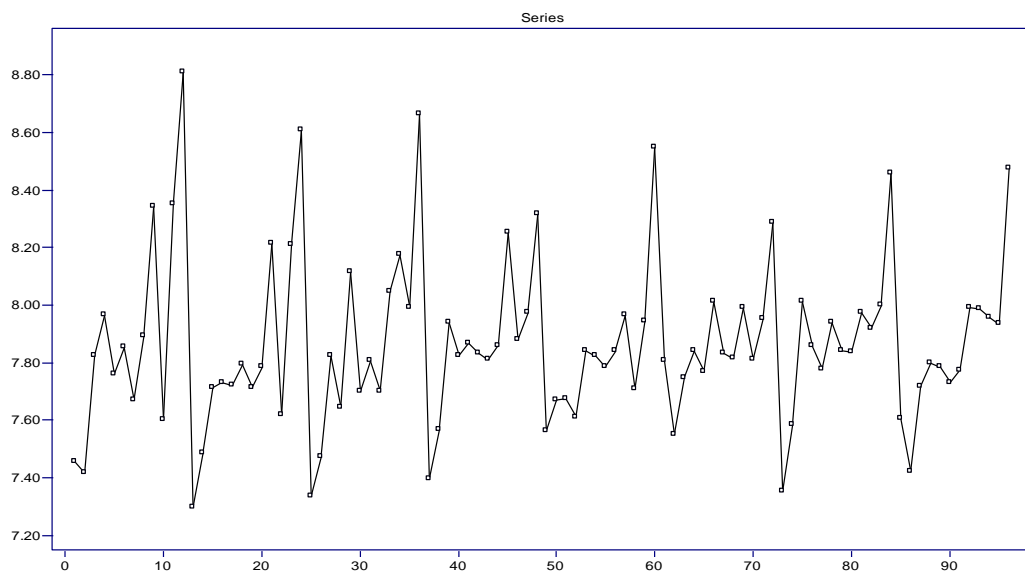


Σχήμα 7. ACF, PACF της χρονοσειράς

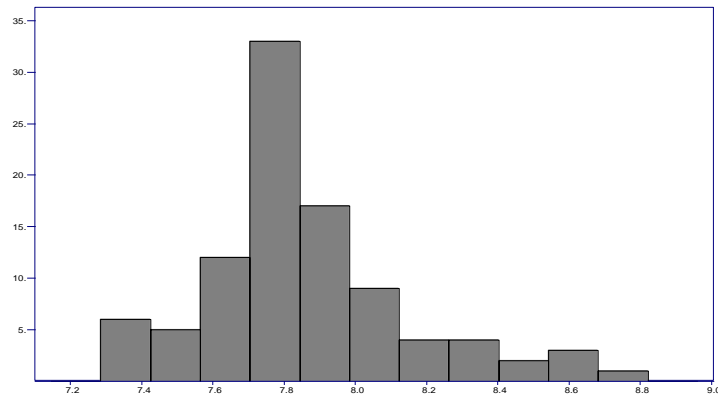


Προκειμένου να αντιμετωπίσουμε τη λοξότητα που παρουσιάζουν οι παρατηρήσεις εκτελούμε τον μετασχηματισμό Box-Cox με παράμετρο $\lambda=0$, δηλαδή λογαριθμίζουμε τις παρατηρήσεις.

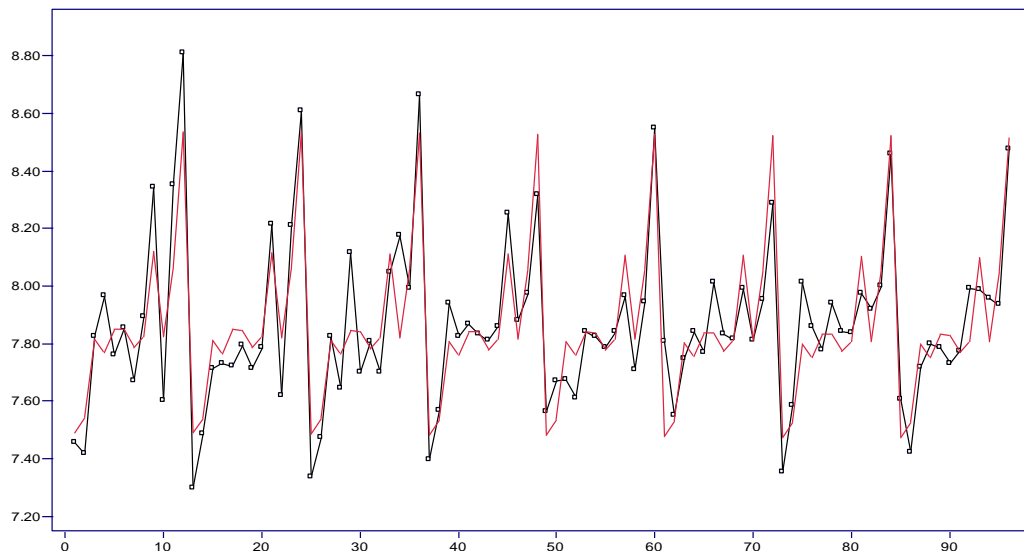
Προκύπτει η νέα μετασχηματισμένη χρονοσειρά που φαίνεται στο Σχήμα 8.



Σχήμα 8 Χρονοσειρά πωλήσεων μετά το μετασχηματισμό Box-Cox ($\lambda=0$).



Παρατηρούμε την ύπαρξη έντονης εποχικότητας με περίοδο 12 μηνών. Πρέπει συνεπώς είτε να πάρουμε διαφορές με υστέρηση 12 μηνών είτε να εφαρμόσουμε κάποια άλλη μέθοδο εκτίμησης της τάσης και της εποχικότητας. Εδώ θα αφαιρέσουμε την εποχική συνιστώσα και τη γραμμική τάση με τη μέθοδο των εποχικών συντελεστών. Η νέα χρονοσειρά φαίνεται στο Σχήμα 9.



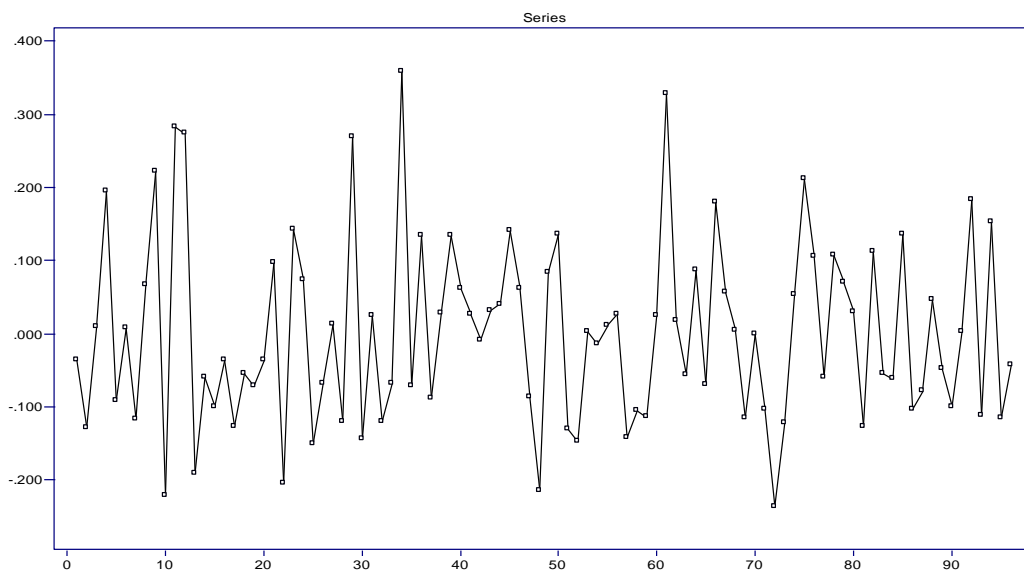
Σχήμα 9. Προσαρμογή της γραμμικής τάσης και της εποχικής συνιστώσας στα δεδομένα

Με διακεκομμένη γραμμή φαίνεται η σύνθεση της γραμμικής συνάρτησης της τάσης με τις αντίστοιχες αποκλίσεις από το μέσο σύμφωνα με τους εποχικούς συντελεστές για κάθε μήνα.

Linear Fit:	
$X(t) = -.00022187 * t + 7.8754$	
Seasonal fit of period = 12	
Seasonal components:	
1)	-.38401
2)	-.32847
3)	-.058246
4)	-.10538
5)	-.023160
6)	-.024926
7)	-.086393
8)	-.048286
9)	.24579
10)	-.048395
11)	.19708
12)	.66439

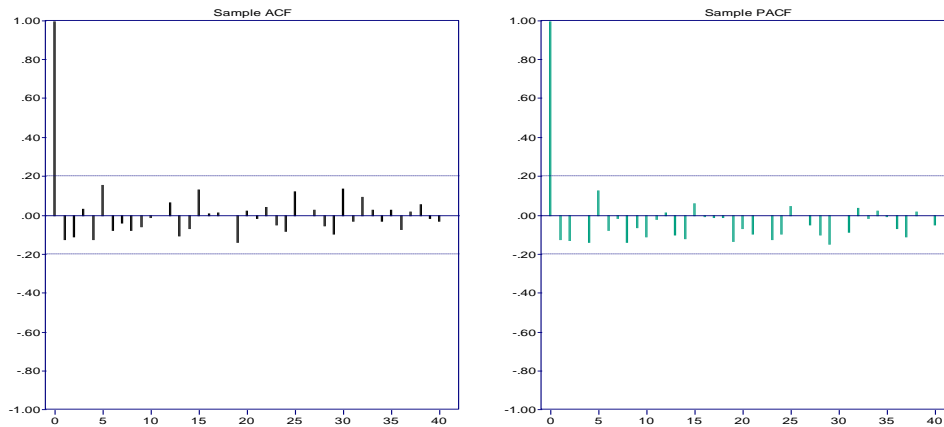
Πίνακας 2. Εποχικοί συντελεστές της εποχικής συνιστώσας των δεδομένων

Η περαιτέρω ανάλυση θα γίνει στη χρονοσειρά που προκύπτει αφού αφαιρέσουμε την προσαρμοσμένη εποχική και γραμμική συνιστώσα από τις παρατηρήσεις. Η χρονοσειρά αυτή φαίνεται στο Σχήμα 10.



Σχήμα 10. Η χρονοσειρά μετά την απαλοιφή της γραμμικής και εποχικής συνιστώσας

Σε αυτό το σημείο μπορούμε να ελέγξουμε τις ACF, PACF του δείγματος, που απεικονίζονται στο Σχήμα 11.

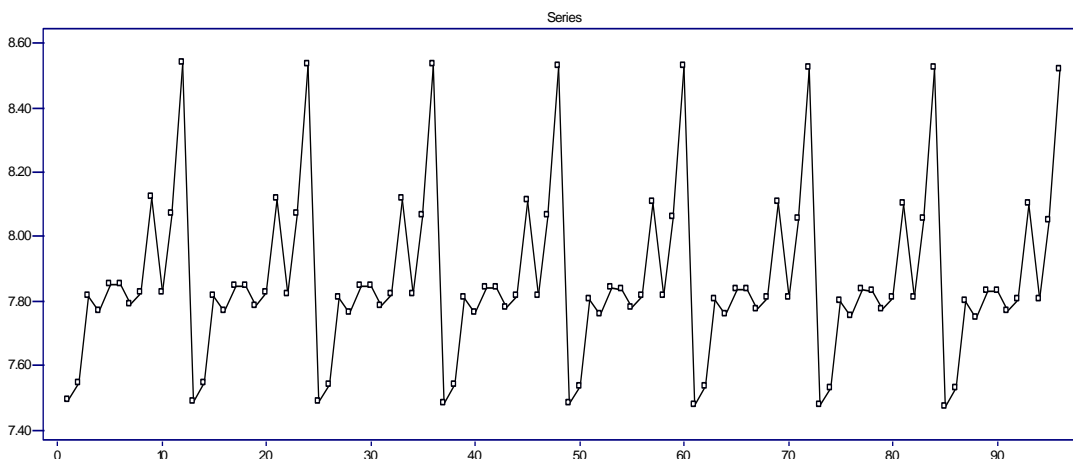


Σχήμα 11. ACF, PACF της χρονοσειράς χωρίς την εποχική και γραμμική συνιστώσα

Δεν παρατηρούμε κάποια σημαντική εξάρτηση (έξω από τα όρια 95%) σε κάποια υστέρηση, οπότε μπορούμε να ελέγξουμε την υπόθεση ότι τα υπόλοιπα είναι λευκός θόρυβος και αν ο έλεγχος αποτύχει να προσαρμόσουμε ένα μοντέλο μικρής τάξης προκειμένου να περιγράψουμε τα υπόλοιπα.

Εδώ οι έλεγχοι (), το ιστόγραμμα και το qq-plot των υπολοίπων μας οδηγούν στο συμπέρασμα ότι τα υπόλοιπα είναι λευκός θόρυβος, και συνεπώς δεν απαιτείται περαιτέρω μοντελοποίηση για τα δεδομένα.

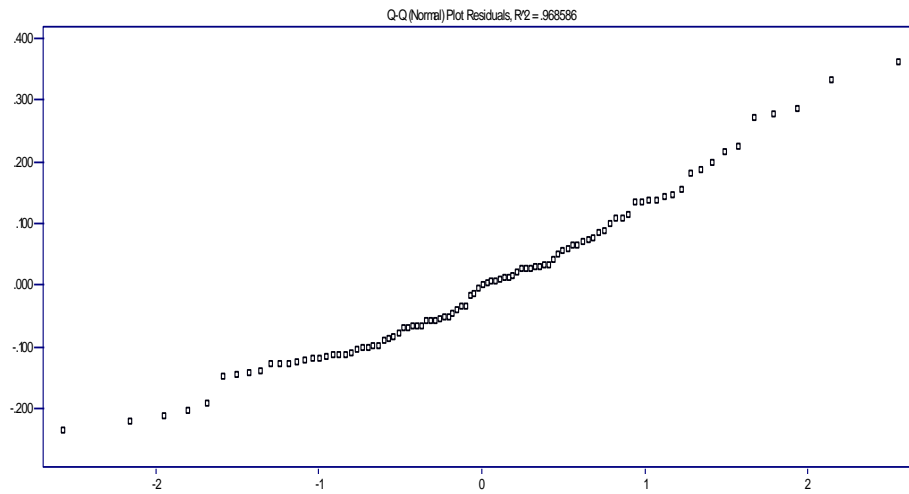
Η απαλοιφή της εποχικότητας με τη μέθοδο των εποχικών συντελεστών αποδείχθηκε επαρκής ως μοντέλο πρόβλεψης για τη ζήτηση του φαρμάκου. Παρακάτω φαίνονται οι έλεγχοι υπολοίπων.



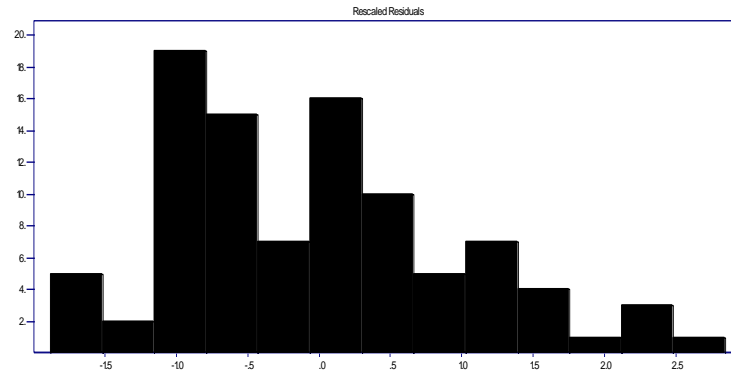
Σχήμα 12. Γράφημα χρονοσειράς του μοντέλου (96 μήνες).

Ljung - Box statistic = 16.774 Chi-Square (20), p-value = .66757
McLeod - Li statistic = 12.156 Chi-Square (21), p-value = .93530
Turning points = 63.000~AN(62.667,sd = 4.0920), p-value = .93508
Diff sign points = 47.000~AN(47.500,sd = 2.8431), p-value = .86040
Rank test statistic = .23510E+04~AN(.22800E+04,sd = .15797E+03), p-value = .65310
Jarque-Bera test statistic (for normality) = 5.9098 Chi-Square (2), p-value = .05208
Order of Min AICC YW Model for Residuals = 0

Πίνακας 3. Έλεγχοι τυχαιότητας των υπολοίπων



Σχήμα 13. Normal qq plot των υπολοίπων



Σχήμα 14. Ιστόγραμμα των υπολοίπων

Η χρονοσειρά του Σχήματος 12 προκύπτει από τη σχέση:

$$\hat{X}_t = -0,00022187t + 7,8754 + s_i^d$$

S_i^d οι εποχικοί συντελεστές με $i=1,2,\dots,12$ και $d=1,2,\dots,8$. $t=1,2,\dots,96$.

Κανένας έλεγχος δε μας οδηγεί στο να απορρίψουμε το μοντέλο που επιλέξαμε, οπότε είναι ασφαλές στο να προβούμε σε πρόβλεψη για τους μήνες 1 έως 12 για το έτος 2012, αφού αντιστρέψουμε τον αρχικό μετασχηματισμό Box-Cox.

Το μοντέλο προβλέπει για το 2012 τις τιμές που φαίνονται στον πίνακα 4, μαζί με ένα 95% διάστημα εμπιστοσύνης για την πρόβλεψη.

Step	Prediction	Lower	Upper
1	.17538E+04	.13707E+04	.22440E+04
2	.18535E+04	.14487E+04	.23715E+04
3	.24282E+04	.18978E+04	.31068E+04
4	.23157E+04	.18099E+04	.29629E+04
5	.25136E+04	.19645E+04	.32161E+04
6	.25086E+04	.19606E+04	.32096E+04
7	.23584E+04	.18432E+04	.30175E+04
8	.24495E+04	.19145E+04	.31341E+04
9	.32862E+04	.25684E+04	.42046E+04
10	.24482E+04	.19135E+04	.31324E+04
11	.31285E+04	.24451E+04	.40028E+04
12	.49909E+04	.39008E+04	.63858E+04

Πίνακας 4. Προβλέψεις με τα 95% διαστήματα εμπιστοσύνης τους

Συγκρίνοντας τις προβλέψεις για το 2012 με τα πραγματικά στοιχεία μπορούμε να αξιολογήσουμε την εγκυρότητά τους και, χρησιμοποιώντας το U-Statistic για την πρώτη πρόβλεψη (πωλήσεις Ιαν 2012), να αποκτήσουμε μια ένδειξη για το αν η μέθοδος πρόβλεψης που χρησιμοποιήσαμε είναι καλύτερη από το απλουστευτικό μοντέλο.

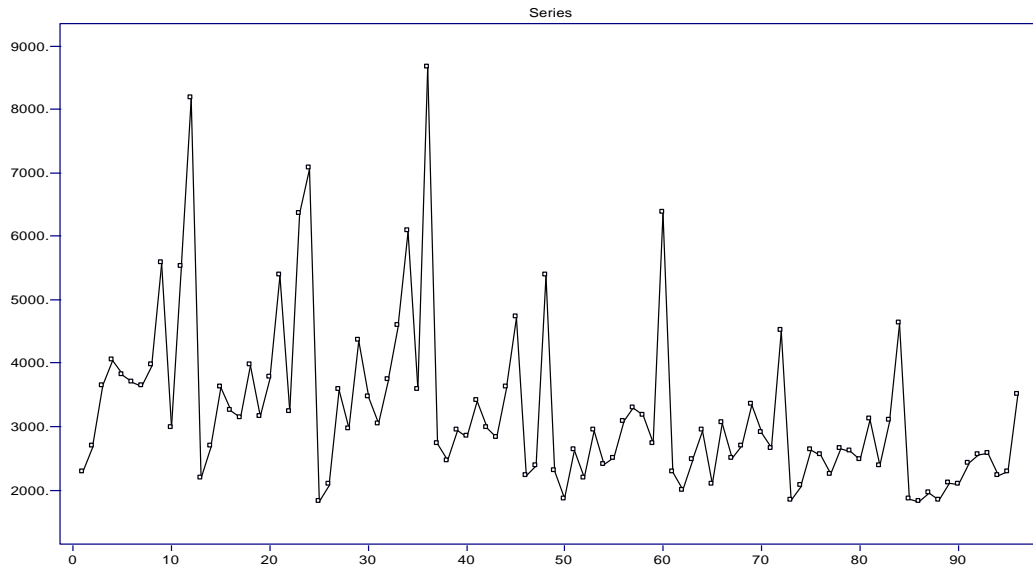
2012	X_t	\hat{X}_t
1	2108	1753,90
2	2392	1853,61
3	2858	2428,30
4	2727	2315,84
5	2895	2513,70
6	2945	2508,68
7	2514	2358,46
8	2546	2449,61
9	2311	3286,35
10	2697	2448,33
11	2629	3128,65
12	3186	4991,16

Πίνακας 5. Σύγκριση προβλέψεων με τις πραγματικές πωλήσεις για το έτος 2012.

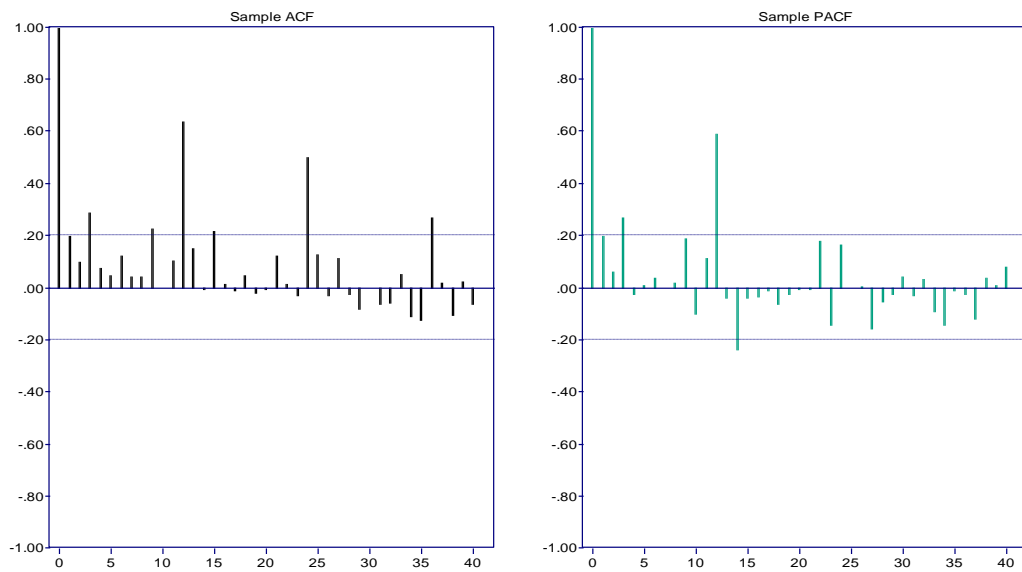
U statistic = 0,13 < 1 και το για το tracking signal είναι $T=0,39 < 0,51$ για $\beta=0,1$.

5.2 Mesulid

Μηνιαίες πωλήσεις Mesulid την περίοδο 2004-2011 (96 μήνες).



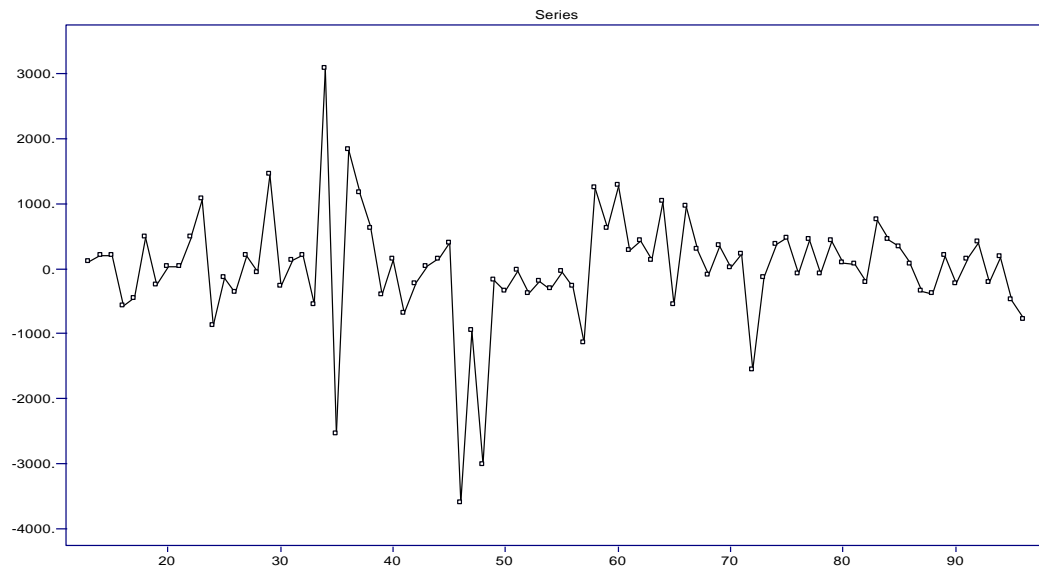
Σχήμα 15. Πωλήσεις σε τεμάχια από Ιαν 2004 έως Δεκ 2011.



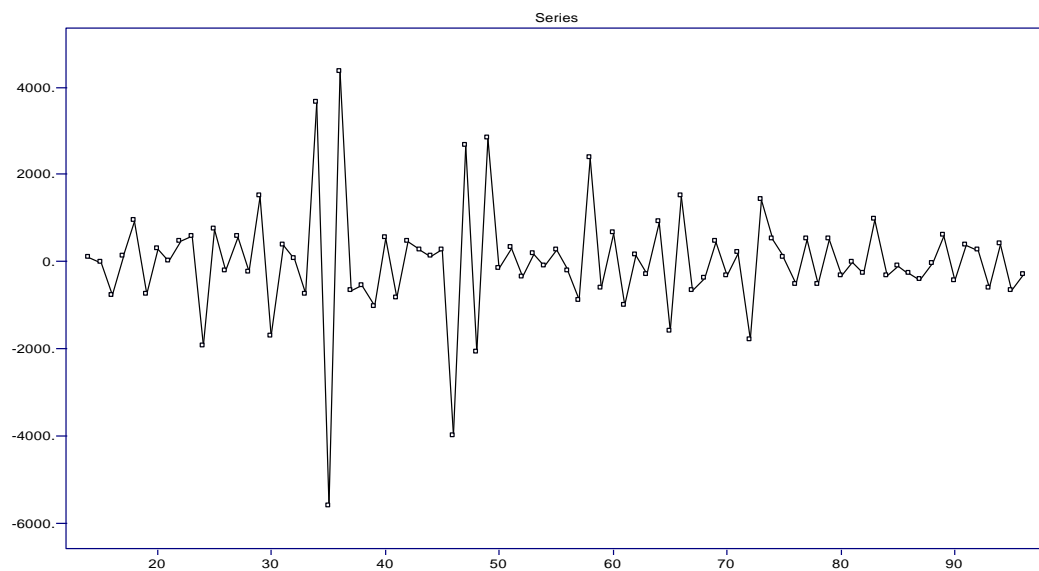
Σχήμα 16. Δειγματικές αυτοσυσχετίσεις και μερικές αυτοσυσχετίσεις των αρχικών παρατηρήσεων

Τα δεδομένα χρειάζονται μετασχηματισμό προκειμένου να μετατραπούν σε στάσιμη χρονοσειρά. Δοκιμάζουμε διαφορές με υστέρηση 12, για την απαλοιφή της εποχικότητας. Η νέα χρονοσειρά φαίνεται στο Σχήμα 17.

Υπάρχει ακόμα μία ασθενής γραμμική τάση, οπότε παίρνουμε επιπλέον μία διαφορά τάξης 1, όπως φαίνεται στο Σχήμα 18. Ύστερα υπολογίζουμε πάλι τις δειγματικές ACF, PACF της νέας σειράς (Σχήμα 19).

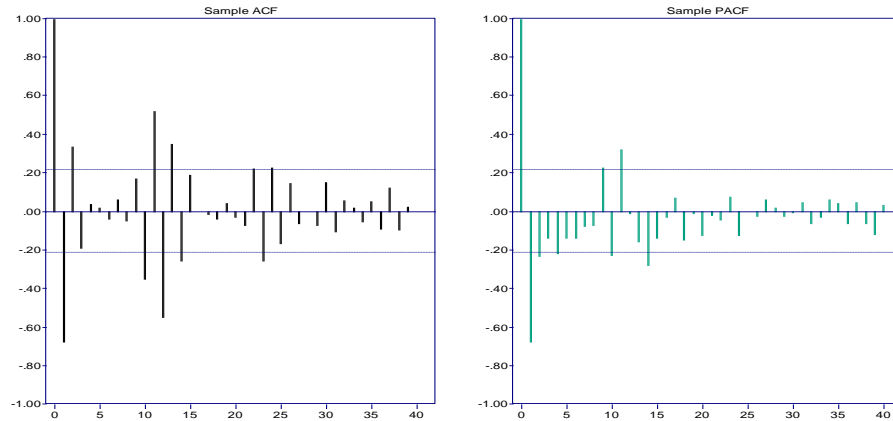


Σχήμα 17. Η διαφοροποιημένη χρονοσειρά με υστέρηση 12.



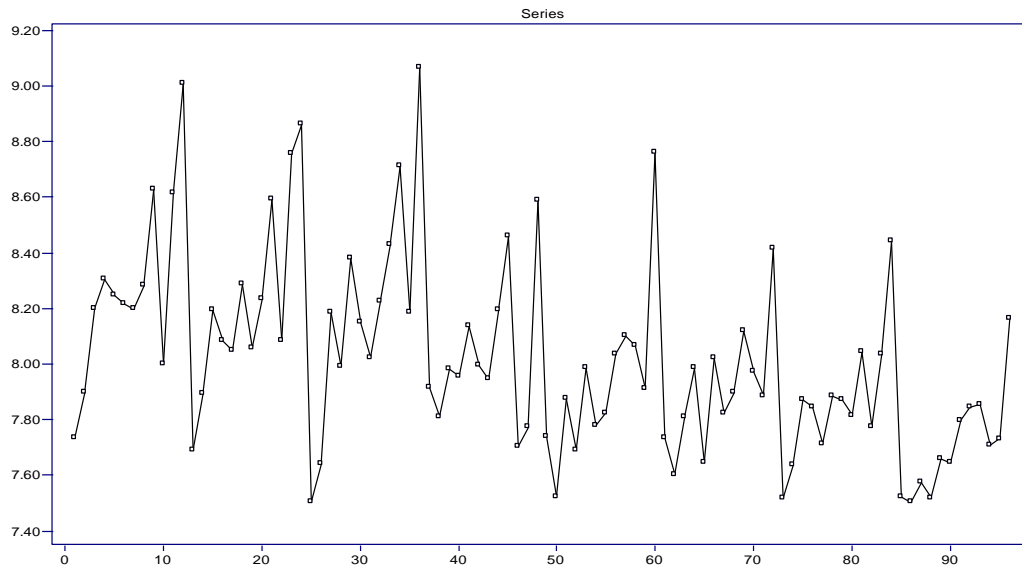
Σχήμα 18. Παρατηρήσεις μετά τις διαφορές 12 και 1

Παρατηρούμε ότι αυτός ο μετασχηματισμός δεν παράγει μία στάσιμη χρονοσειρά. Συγκεκριμένα, η διασπορά ανάμεσα στις χρονικές στιγμές 30 και 50 δεν παραμένει σταθερή. Φαίνεται επίσης σημαντική εξάρτηση σε υστέρηση αρκετά μεγάλη ($h=14$ και $h=23$), πράγμα που θα οδηγούσε σε υπερβολικά μεγάλης τάξης μοντέλο. Είμαστε αναγκασμένοι να δοκιμάσουμε έναν άλλο μετασχηματισμό προκειμένου να καταλήξουμε σε μία στάσιμη χρονοσειρά.

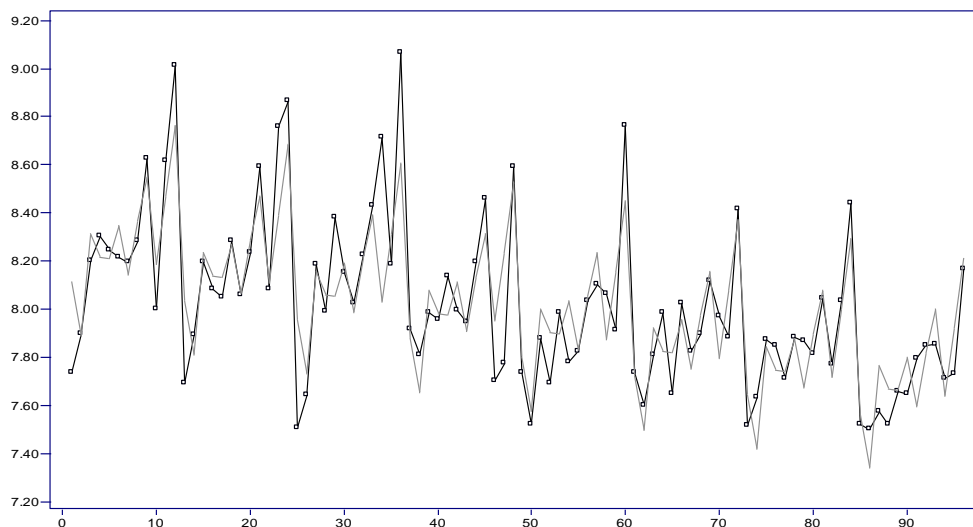


Σχήμα 19. ACF, PACF της διαφοροποιημένης χρονοσειράς (12,1)

Έτσι στα αρχικά δεδομένα εκτελούμε το μετασχηματισμό Box-Cox με $\lambda=0$, ύστερα προσαρμόζουμε μία γραμμική τάση και αναζητούμε μία περιοδική συνάρτηση που να απαλείφει την εποχική συνιστώσα (Σχήμα 21).



Σχήμα 20. Χρονοσειρά μετά το μετασχηματισμό Box-Cox $\lambda=0$

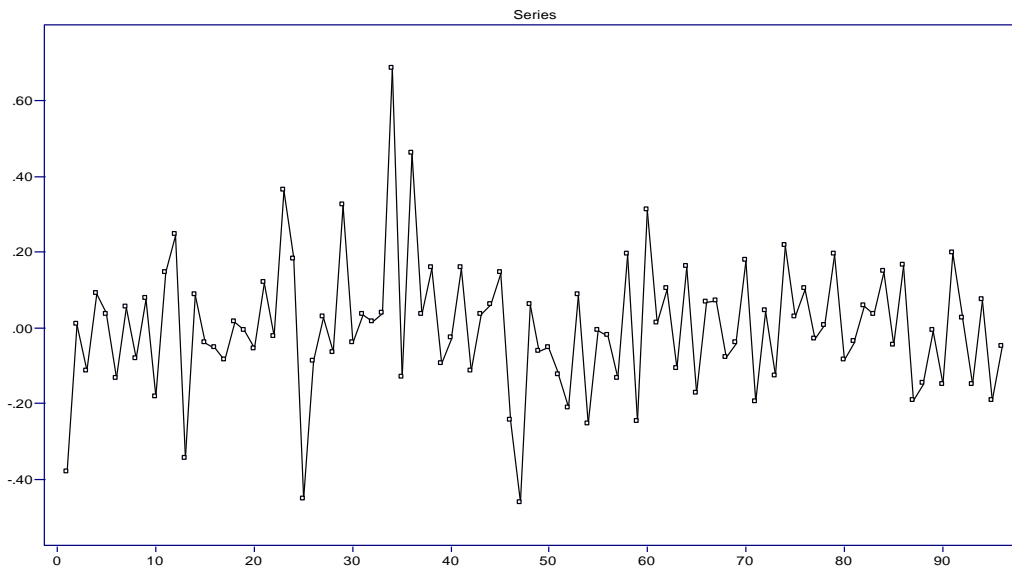


Σχήμα 21. Προσαρμογή της γραμμικής, και περιοδικής συνιστώσας με Harmonic Regression

Polynomial Fit:		
$X(t) = -.0065247 * t + .31645$		
Harmonic regression with 4 Fourier frequencies.		
frequency	alpha	beta
0	8.0236	.00000
8	3.1862	-7.3329
16	3.4582	-4.4820
24	7.7132	-.72413
32	9.7024	-2.1349

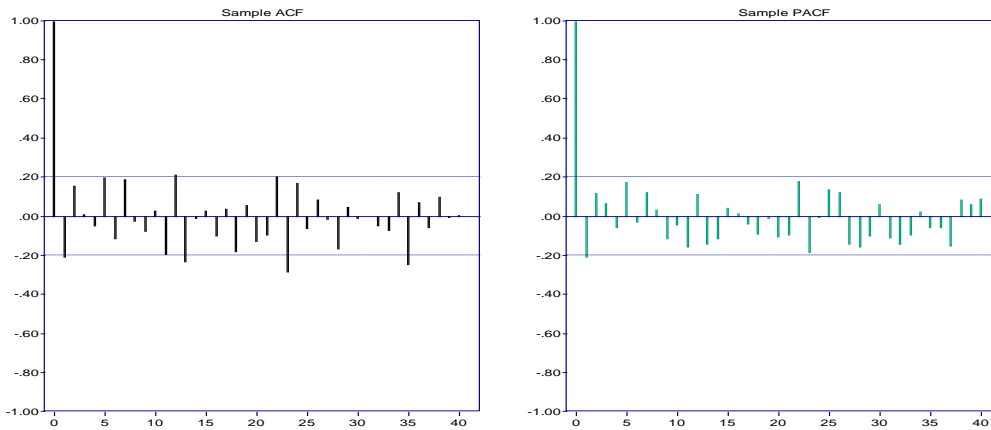
Πίνακας 6. Συντελεστές Harmonic Regression

Προκύπτει έτσι η μετασχηματισμένη χρονοσειρά του Σχήματος 22.



Σχήμα 22. Η χρονοσειρά μετά την απαλοιφή της εποχικής και γραμμικής συνιστώσας

Στο Σχήμα 23 φαίνονται οι δειγματικές αυτοσυσχετίσεις και μερικές αυτοσυσχετίσεις της νέας χρονοσειράς. Παρατηρούμε ότι αυτή η χρονοσειρά μπορεί να θεωρηθεί στάσιμη. Ακόμα, από τις οριακά σημαντικές συσχετίσεις που φαίνονται στην υστέρηση $h=1, 5, 12$ παίρνουμε μία ένδειξη για την τάξη του μοντέλου που θα προσαρμόσουμε. Προφανώς προτιμούμε ένα μικρότερης τάξης μοντέλο, π.χ. ARMA(1,1), AR(1), MA(1) ή ARMA(2,2). Παρακάτω παραθέτουμε έναν πίνακα σύγκρισης των διαφορετικών μοντέλων προσαρμογής με τα αντίστοιχα AIC, BIC και log-likelihood (Πίνακας 7).



Σχήμα 23. ACF, PACF της χρονοσειράς μετά την απαλοιφή της εποχικής και γραμμικής συνιστώσας

	AICC	BIC	-2Log(Like)
AR(1)	-63.084742	-64.689191	-67.214880
AR(2)	-62.307804	-64.366196	-68.568674
ARMA(1,1)	-61.640528	-63.925284	-67.901397
ARMA(2,2)	-68.752130	-72.459479	-79.418797
ARMA(1,2)	-60.463248	-63.631636	-68.902809
ARMA(2,1)	-60.787345	-63.814871	-69.226906

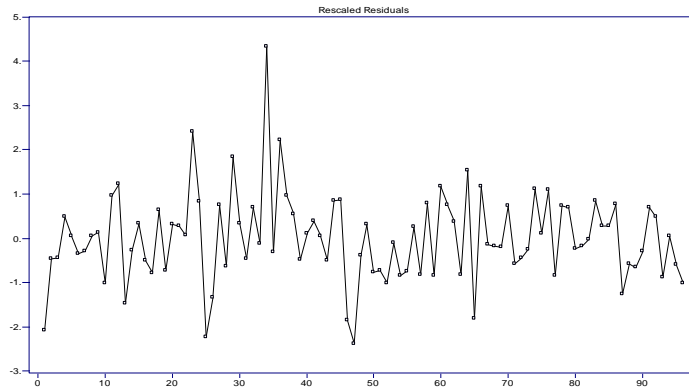
Πίνακας 7. Συγκριτικός πίνακας μοντέλων προσαρμογής

Ύστερα από σύγκριση των κριτηρίων AICC και BIC είναι φανερό ότι το καλύτερο από αυτά τα μοντέλα είναι το ARMA(2,2) επειδή σε αυτό παρουσιάζονται οι μικρότερες τιμές για τα κριτήρια. Ύστερα από τον υπολογισμό των παραμέτρων προκύπτει το μοντέλο

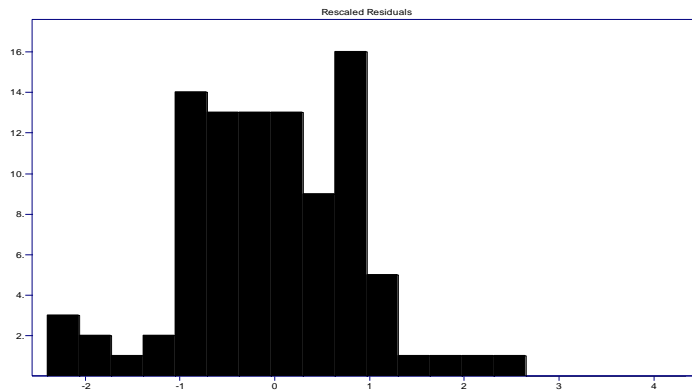
$$X(t) = -1.724 X(t-1) - .9991 X(t-2) + Z(t) + 1.703 Z(t-1) + .9746 Z(t-2)$$

$$\text{WN Variance} = .024551$$

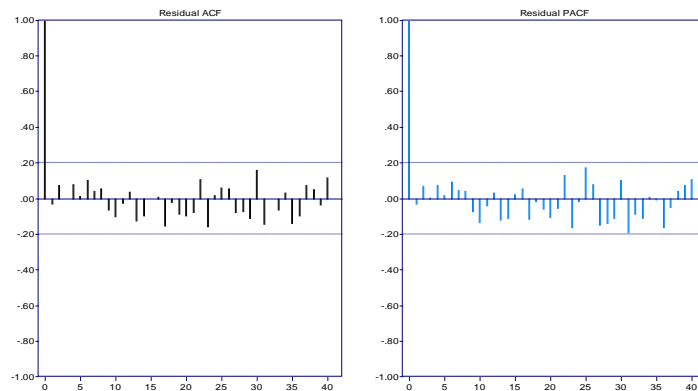
Από το έλεγχο προσαρμογής των υπολοίπων (Σχήμα 24) παρατηρούμε ότι ενδεχομένως τα υπόλοιπα να μην είναι τυχαία. Από τις δειγματικές ACF, PACF συμπεραίνουμε ότι δεν υπάρχουν σημαντικές συσχετίσεις στα υπόλοιπα, ενώ από το ιστόγραμμα, μόνο 4 σημεία (<5%) βρίσκονται εκτός ορίων. Τέλος, τα tests of randomness (εκτός από το Jarque-Bera test statistic) συναινούν στην αποδοχή ότι τα υπόλοιπα είναι αρκετά τυχαία, ώστε να αποδεχτούμε ότι το μοντέλο ARMA(2,2) που υπολογίσαμε περιγράφει ικανοποιητικά τη στάσιμη, μετασχηματισμένη χρονοσειρά του Σχήματος 22.



Σχήμα 24. Time series plot των residuals



Σχήμα 25. Ιστόγραμμα υπολοίπων

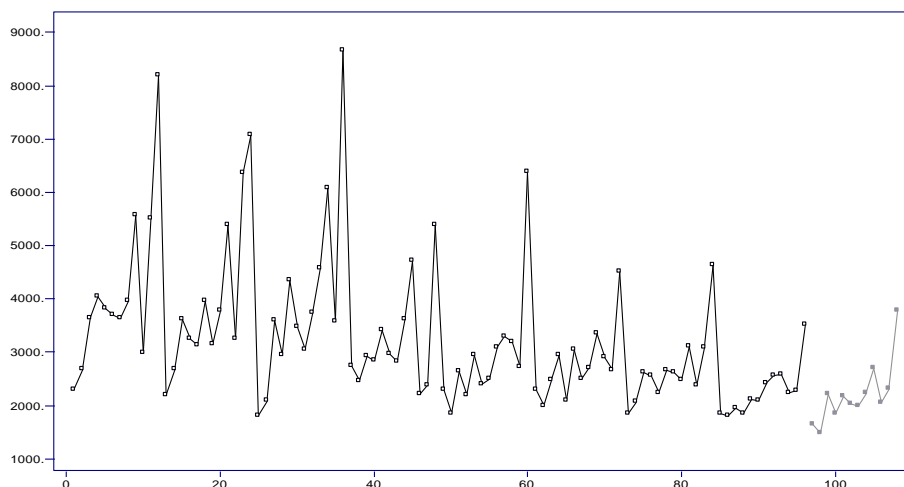


Σχήμα 26. Δειγματικές ACF/PACF των υπολοίπων

Ljung - Box statistic = 19.548 Chi-Square (20), p-value = .48650
McLeod - Li statistic = 25.016 Chi-Square (24), p-value = .40492
Turning points = 61.000~AN(62.667,sd = 4.0920), p-value = .68379
Diff sign points = 47.000~AN(47.500,sd = 2.8431), p-value = .86040
Rank test statistic = .22620E+04~AN(.22800E+04,sd = .15797E+03), p-value = .90928
Jarque-Bera test statistic (for normality) = 41.358 Chi-Square (2), p-value = .00000
Order of Min AICC YW Model for Residuals = 0

Πίνακας 8. Tests of randomness των υπολοίπων

Μπορούμε λοιπόν να προβούμε σε προβλέψεις για τους 12 μήνες του έτους 2012, αντιστρέφοντας τους μετασχηματισμούς, αφού συνθέσουμε το μοντέλο ARMA(2,2) που υπολογίσαμε. Οι προβλέψεις φαίνονται στο Σχήμα 27 και στον Πίνακα 9.



Σχήμα 27. Πρόβλεψη 12 μηνών για την αρχική χρονοσειρά

Approximate 95 Percent Prediction Bounds			
Step	Prediction	Lower	Upper
1	.16463E+04	.12103E+04	.22395E+04
2	.14781E+04	.10865E+04	.20109E+04
3	.22191E+04	.16311E+04	.30190E+04
4	.18433E+04	.13549E+04	.25077E+04
5	.21790E+04	.16017E+04	.29644E+04
6	.20336E+04	.14947E+04	.27668E+04
7	.19920E+04	.14640E+04	.27105E+04
8	.22283E+04	.16375E+04	.30323E+04
9	.26974E+04	.19822E+04	.36708E+04
10	.20598E+04	.15137E+04	.28030E+04
11	.23074E+04	.16956E+04	.31400E+04
12	.37876E+04	.27831E+04	.51547E+04

Πίνακας 9. Προβλέψεις 12 μηνών για την αρχική χρονοσειρά, με τα αντίστοιχα διαστήματα εμπιστοσύνης. Συγκρίνουμε τις προβλέψεις με τα πραγματικά στοιχεία πωλήσεων για το 2012.

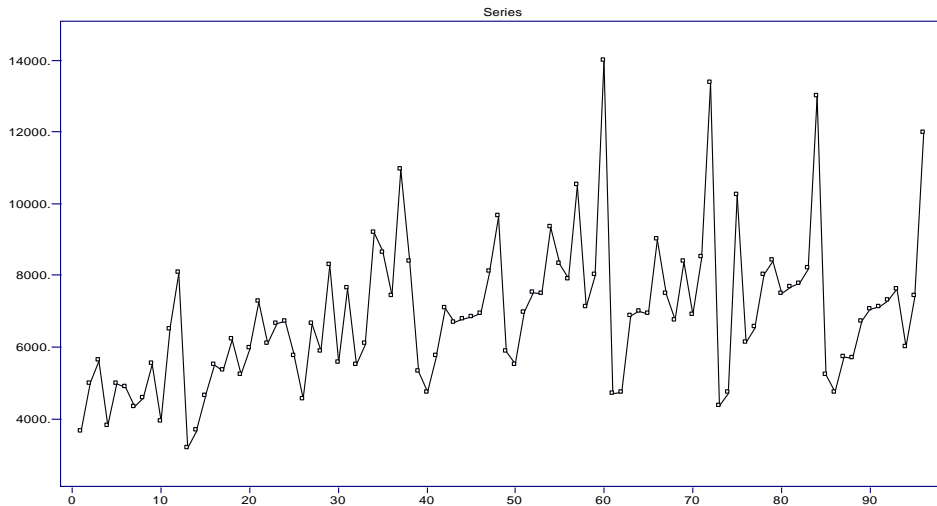
2012	X_t	F_t		
1	1796	1646,3	ME	-77,18
2	2091	1478,1	MAE	152,48
3	2344	2219,1	MAPE	6,93%
4	2391	1843,3	MSE	28457,23
5	2554	2179,0	U stat	0,058
6	2520	2033,6		
7	2321	1992,0		
8	2433	2228,3		
9	1992	2697,4		
10	2220	2059,8		
11	1814	2307,4		
12	2417	3783,6		

Πίνακας 10. Σύγκριση προβλέψεων με τις πραγματικές πωλήσεις για το έτος 2012.

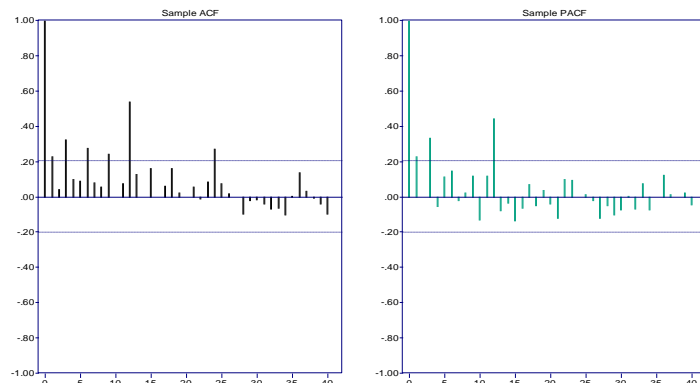
5.3 Glucophage

Μηνιαίες πωλήσεις Glucophage κατά την περίοδο 2004-2011 (96 μήνες).

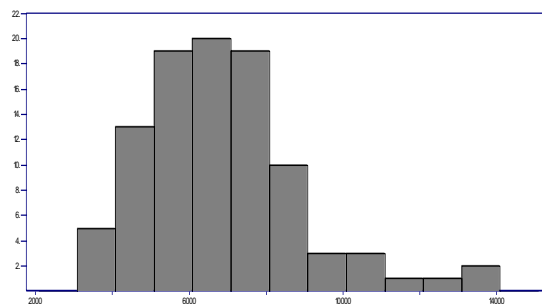
Από το γράφημα των πωλήσεων παρατηρούμε μία έντονη εποχικότητα, ειδικότερα κατά τα τελευταία 4 έτη. Ακόμα, φαίνεται μία αρχικά αυξανόμενη τάση μέχρι περίπου το 2007, η οποία αργότερα υποχωρεί και οι πωλήσεις μεταβάλλονται γύρω από μία σχεδόν σταθερή μέση τιμή. Από το Σχήμα 29 φαίνεται αυτή η ισχυρή συσχέτιση στην υστέρηση 12, όπως αναμενόταν. Προφανώς θα πρέπει να εκτελέσουμε κατάλληλους μετασχηματισμούς για να επιτύχουμε στασιμότητα.



Σχήμα 28. Πωλήσεις σε τεμάχια από Ιαν 2004 έως Δεκ 2011.

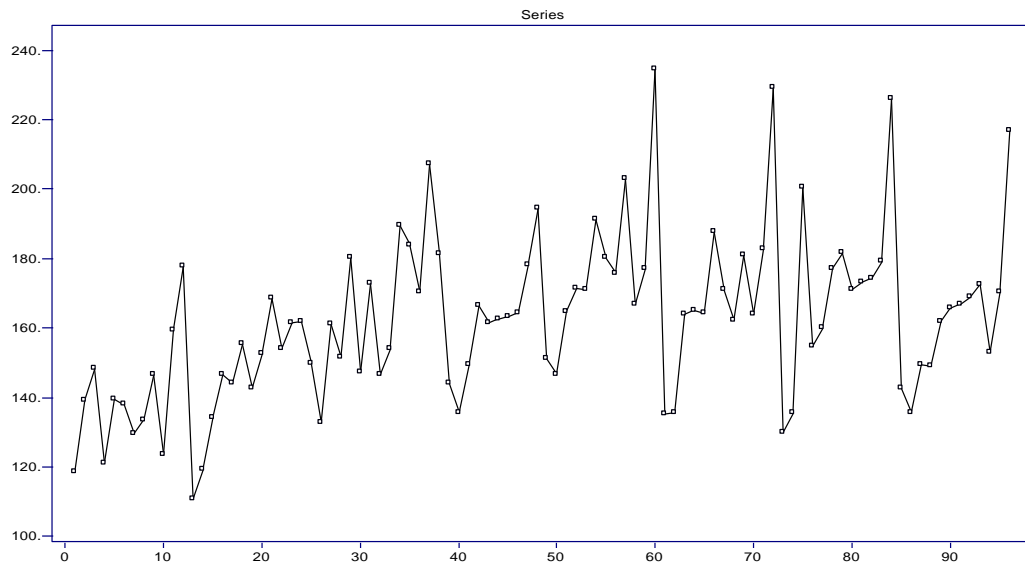


Σχήμα 29. ACF, PACF των παρατηρήσεων.



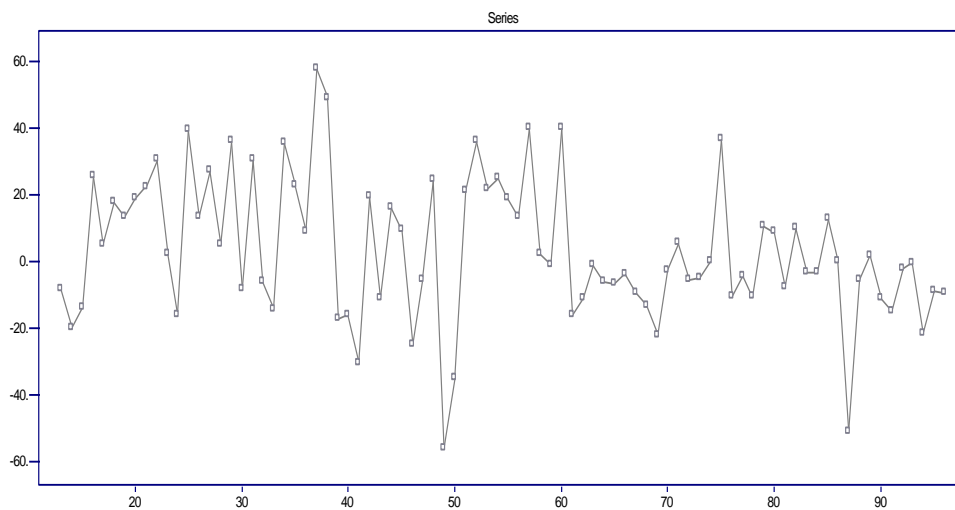
Σχήμα 30. Ιστόγραμμα χρονοσειράς πωλήσεων.

Αρχικά μετασχηματίζουμε με $\lambda=0.5$ στο μετασχηματισμό Box-Cox (που ισοδυναμεί με την τετραγωνική ρίζα) για να ελαττώσουμε τη λοξότητα (Σχήμα 31)

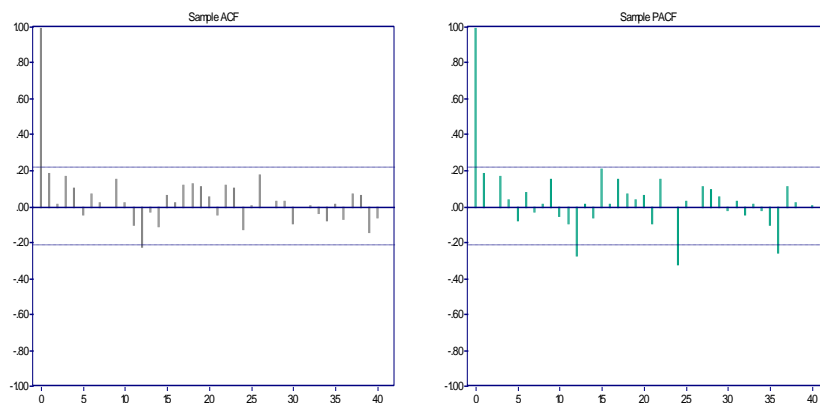


Σχήμα 31 Box-Cox με $\lambda=0,5$ (ισοδυναμεί με το μετασχηματισμό $\sqrt{x_t}$)

Ύστερα επιλέγουμε να πάρουμε διαφορές υστέρησης 12 για να απαλείψουμε την εποχική συνιστώσα. Παρατηρούμε πως η χρονοσειρά των διαφορών είναι στάσιμη (Σχήμα 32). Μελετάμε τις ACF, PACF της νέας χρονοσειράς για να πάρουμε μία ένδειξη για την τάξη και το είδος του μοντέλου που θα προσαρμόσουμε. Ο Πίνακας 11 συγκρίνει τα κριτήρια AIC, BIC για τα μοντέλα AR, MA, τάξης 1, 2, 3 και μικτά μοντέλα ARMA μέχρι τάξης 3.



Σχήμα 32. Χρονοσειρά διαφορών με υστέρηση $d=12$



Σχήμα 33. ACF, PACF των διαφοροποιημένων παρατηρήσεων.

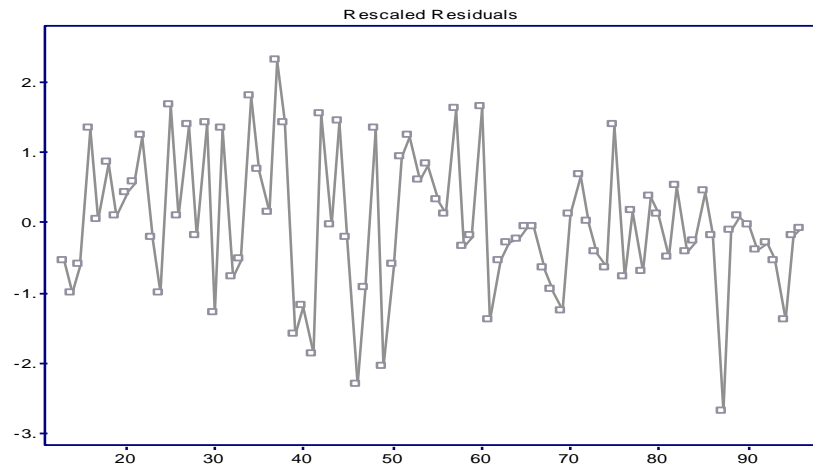
MODEL	AICc	BIC
AR(1)	.750585E+03	.748551E+03
AR(2)	.752722E+03	.749298E+03
AR(3)	.752132E+03	.748637E+03
MA(1)	.750412E+03	.748428E+03
MA(2)	.752494E+03	.749209E+03
MA(3)	.753302E+03	.749158E+03
ARMA(1,1)	.752412E+03	.749175E+03
ARMA(1,2)	.754601E+03	.749494E+03
ARMA(2,1)	.754573E+03	.749491E+03
ARMA(2,2)	.755690E+03	.749483E+03
ARMA(3,1)	.754289E+03	.749143E+03
ARMA(3,2)	.756365E+03	.749394E+03
ARMA(2,3)	.749983E+03	.744043E+03
ARMA(3,3)	.754701E+03	.746991E+03

Πίνακας 11. Σύγκριση μοντέλων προσαρμογής

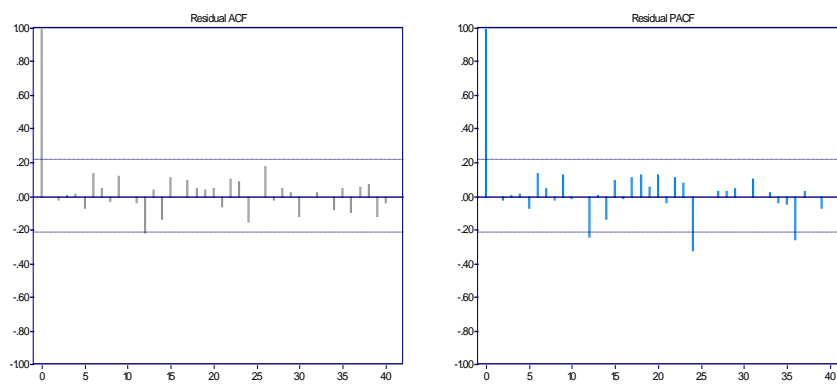
Από το συγκριτικό πίνακα επιλέγουμε το ARMA(2,3)

ARMA Model:
$X(t) = 0,8474X(t-1) - 0,6900X(t-2) + Z(t) - 0,7139Z(t-1) + 0,6620Z(t-2) + 0,3252 Z(t-3)$
WN Variance = 0,359004E+03

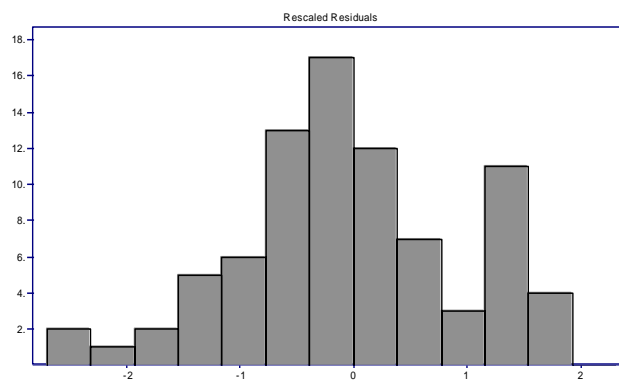
Προχωρούμε στον έλεγχο των υπολοίπων. Από τους παρακάτω πίνακες και τους ελέγχους του πίνακα 12 αποδεχόμαστε την υπόθεση ότι τα υπόλοιπα είναι λευκός θόρυβος, οπότε προχωρούμε στην πραγματοποίηση προβλέψεων για τους επόμενους 12 μήνες (Πίνακας 13).



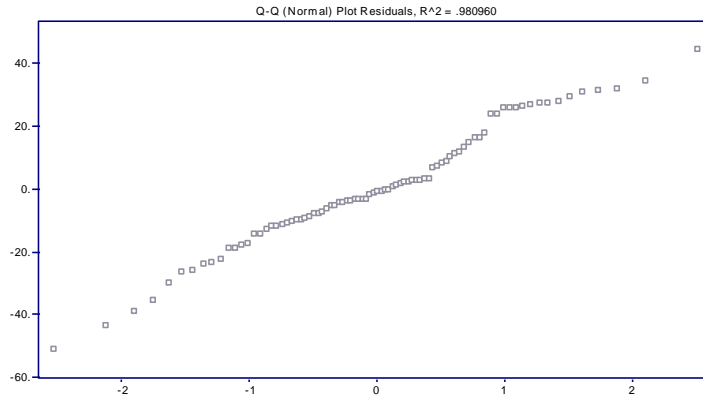
Σχήμα 34. Rescaled residuals



Σχήμα 35. ACF, PACF των υπολοίπων.



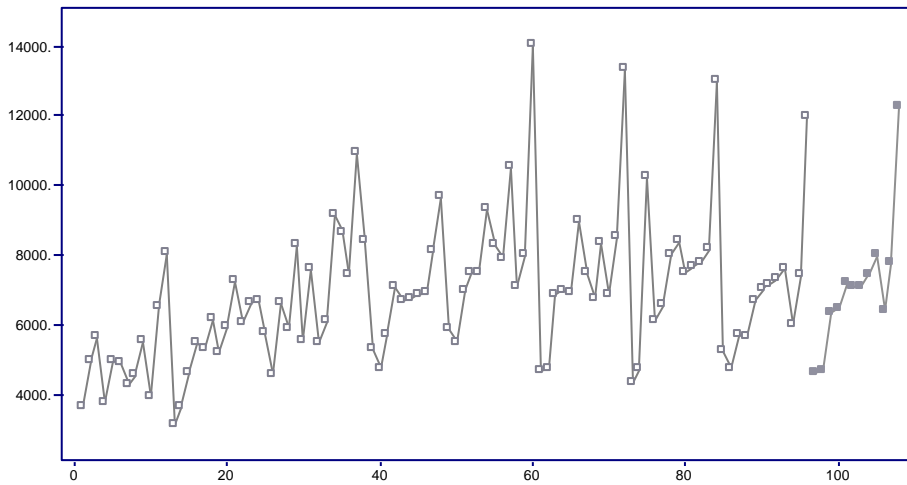
Σχήμα 36. Ιστόγραμμα residuals



Σχήμα 37. Normal qq plot of residuals.

Ljung - Box statistic = 21.596 Chi-Square (20), p-value = .36281
McLeod - Li statistic = 27.362 Chi-Square (25), p-value = .33803
Turning points = 53.000~AN(54.667,sd = 3.8224), p-value = .66282
Diff sign points = 42.000~AN(41.500,sd = 2.6615), p-value = .85098
Rank test statistic = .15180E+04~AN(.17430E+04,sd = .12943E+03), p-value = .08214
Jarque-Bera test statistic (for normality) = .089283 Chi-Square (2), p-value = .95634
Order of Min AICC YW Model for Residuals = 0

Πίνακας 12. Tests of randomness των υπολοίπων



Σχήμα 38. Προβλέψεις 12 μηνών

Step	Prediction	Lower	Upper
1	.46220E+04	.24421E+04	.74915E+04
2	.46978E+04	.24808E+04	.76167E+04
3	.63539E+04	.37103E+04	.97044E+04
4	.64638E+04	.36921E+04	.10007E+05
5	.71853E+04	.41969E+04	.10972E+05
6	.71108E+04	.41381E+04	.10883E+05
7	.70914E+04	.40905E+04	.10913E+05
8	.74698E+04	.43634E+04	.11406E+05
9	.80341E+04	.47967E+04	.12102E+05
10	.64279E+04	.35663E+04	.10126E+05
11	.77916E+04	.45933E+04	.11830E+05
12	.12268E+05	.81478E+04	.17229E+05

Πίνακας 13. Προβλέψεις 12 μηνών για το έτος 2012 με τα 95% διαστήματα εμπιστοσύνης τους.

Συγκρίνουμε τις προβλέψεις με τις πραγματικές πωλήσεις για το 2012.

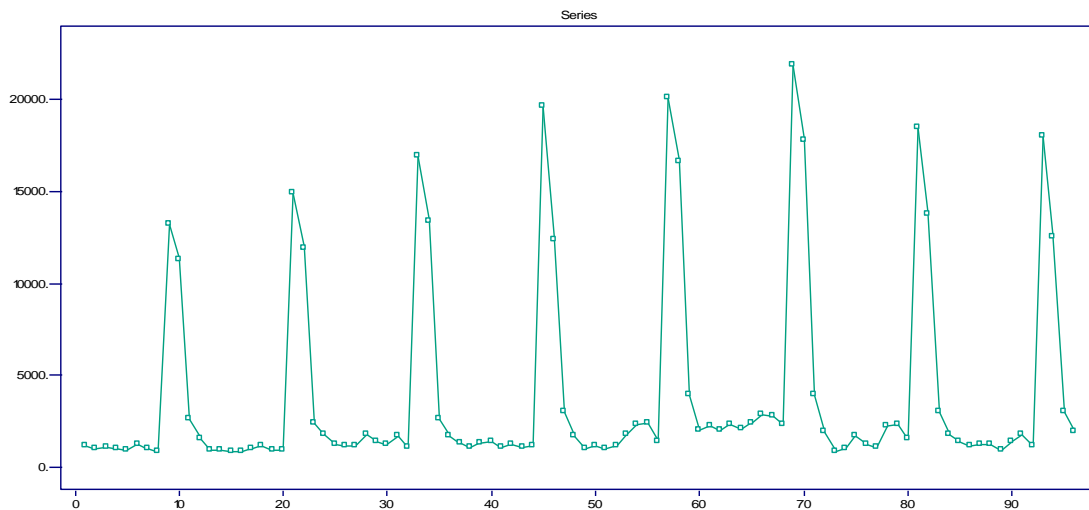
2012	X_t	F_t
1	4461	4622,00
2	5172	4697,80
3	6562	6353,90
4	6070	6463,80
5	6463	7185,30
6	7199	7110,80
7	6846	7091,40
8	7045	7469,80
9	7689	8034,10
10	5910	6427,90
11	7408	7791,60
12	10908	12268,00

ME	315,28
MAE	443,70
MAPE	5,98%
MSE	300283,90
U stat	0,17

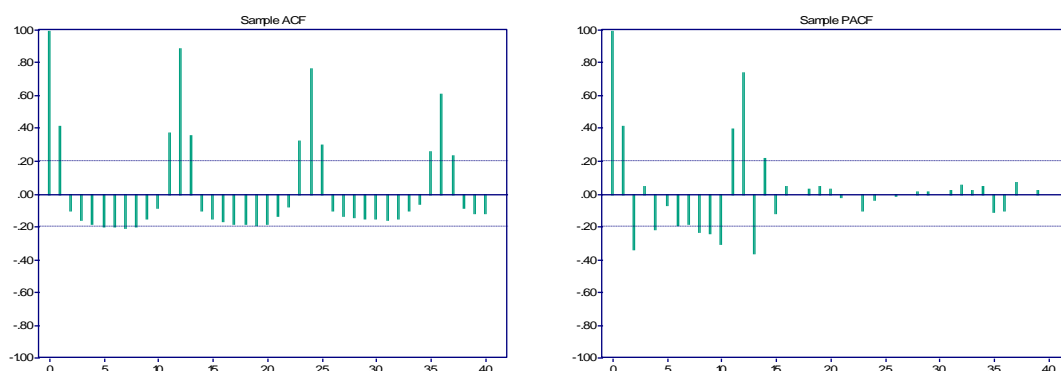
Πίνακας 14. Σύγκριση προβλέψεων με τις πραγματικές πωλήσεις για το έτος 2012.

5.4 Εμβόλια Ιών

Εδώ μελετάμε τις πωλήσεις εμβολίων από τον Ιανουάριο του 2004 έως το Δεκέμβριο του 2011 (96 μήνες). Οι παρατηρήσεις αφορούν ομαδοποιημένες πωλήσεις 32 σκευασμάτων σε τεμάχια.



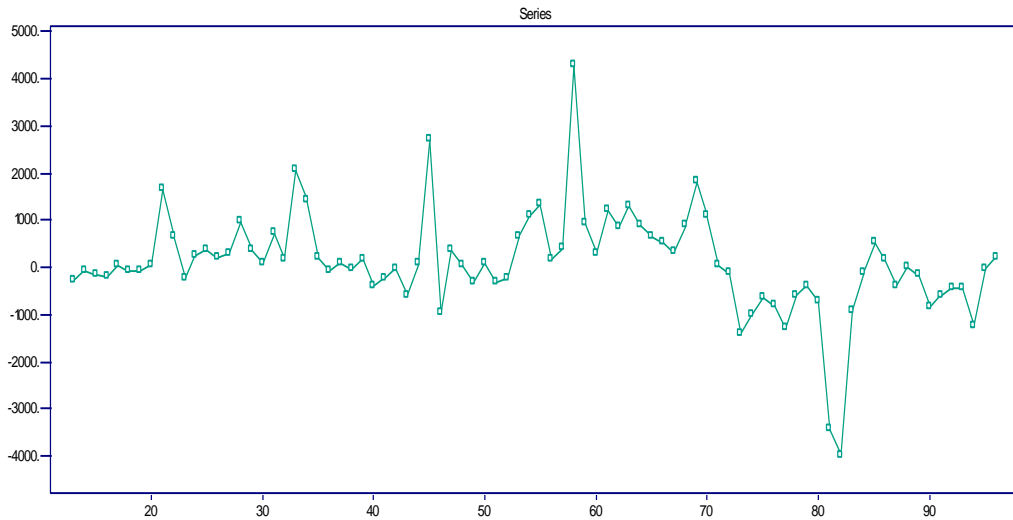
Σχήμα 39. Πωλήσεις σε τεμάχια 32 εμβολίων από Ιαν 2004 έως Δεκ 2011



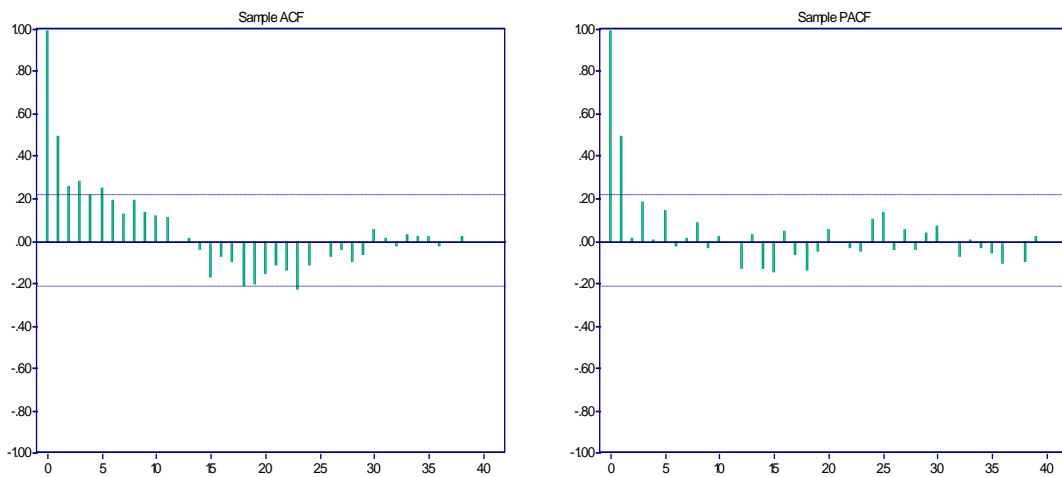
Σχήμα 40. ACF, PACF της χρονοσειράς πωλήσεων εμβολίων.

Στο Σχήμα 39 και το Σχήμα 40 φαίνεται έντονα η εποχική συμπεριφορά των πωλήσεων. Συγκεκριμένα παρατηρείται ότι οι πωλήσεις παραμένουν σε χαμηλά επίπεδα καθ' όλη τη διάρκεια του έτους εκτός από τους μήνες Σεπτέμβριο, Οκτώβριο και Νοέμβριο όπου παρουσιάζουν μεγάλη αύξηση. Το γεγονός αυτό εξηγείται από την έξαρση των ιώσεων που παρατηρείται κατά τους χειμερινούς μήνες και για να αντιμετωπίσουν αυτήν την έξαρση οι φαρμακοποιοί προμηθεύονται νωρίτερα τα εμβόλια. Σε αυτήν την περίπτωση η άνοδος αυτή δε σχετίζεται με οικονομικά κίνητρα των φαρμακείων όπως στις προηγούμενες περιπτώσεις.

Για να απαλείψουμε την εποχικότητα παίρνουμε διαφορές υστέρησης 12 (Σχήμα 41) και μελετώντας τις δειγματικές αυτοσυσχετίσεις (Σχήμα 42) βλέπουμε σημαντικές εξάρτηση στις μερικές αυτοσυσχετίσεις για $h = 1$. Σε συνδυασμό με την συμπεριφορά των αυτοσυσχετίσεων (ACF) συμπεραίνουμε ότι πιθανώς πρόκειται για ένα μοντέλο AR(1).



Σχήμα 41. Χρονοσειρά διαφοροποιημένων παρατηρήσεων με υστέρηση $d=12$



Σχήμα 42. ACF, PACF των διαφοροποιημένων παρατηρήσεων

Πράγματι, το μοντέλο με τις μικρότερες τιμές, τόσο για το AICc, όσο και για το BIC είναι το AR(1) του πίνακα παρακάτω.

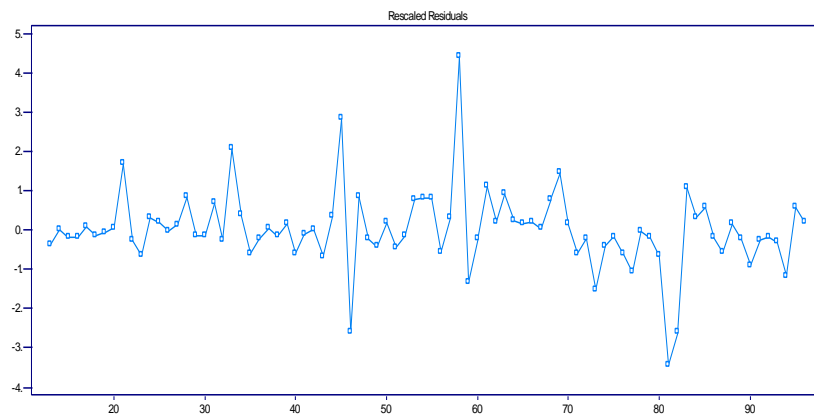
$$\text{ARMA Model: } X(t) = 0,4912 X(t-1) + Z(t)$$

$$\text{WN Variance} = 0,832757\text{E}+06$$

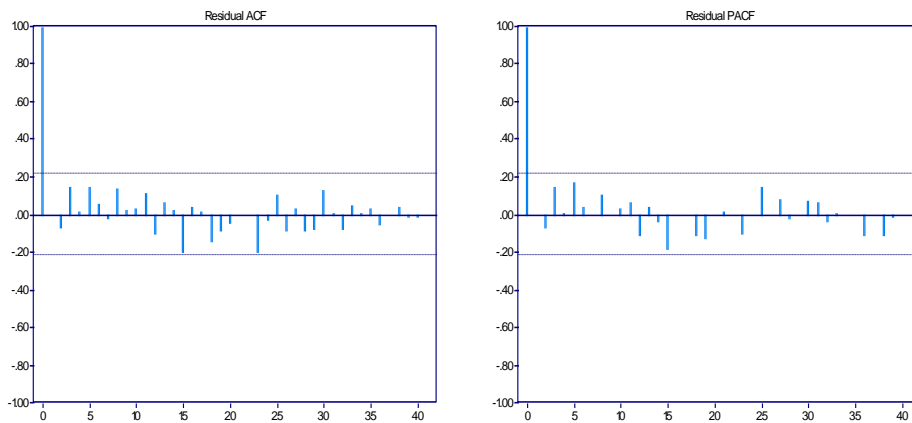
$$\text{AICC} = .138794\text{E}+04$$

$$\text{BIC} = .138782\text{E}+04$$

Σχεδιάζοντας το γράφημα της χρονοσειράς των υπολοίπων, καθώς και τις ACF, PACF των υπολοίπων, γίνεται φανερό πως η συμπεριφορά τους προσομοιάζει λευκό θόρυβο.

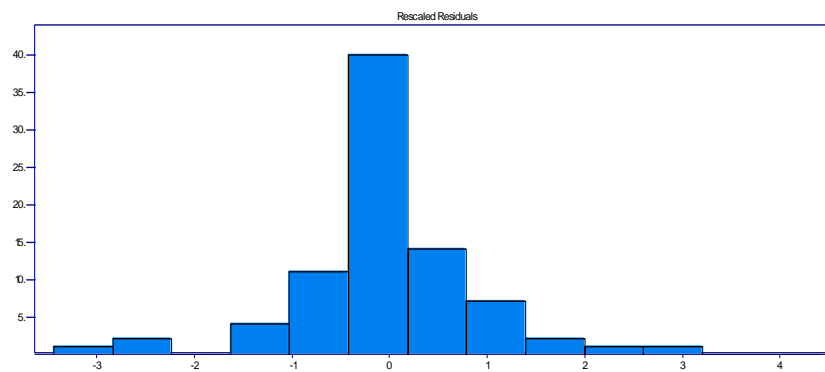


Σχήμα 43. Χρονοσειρά των υπολοίπων

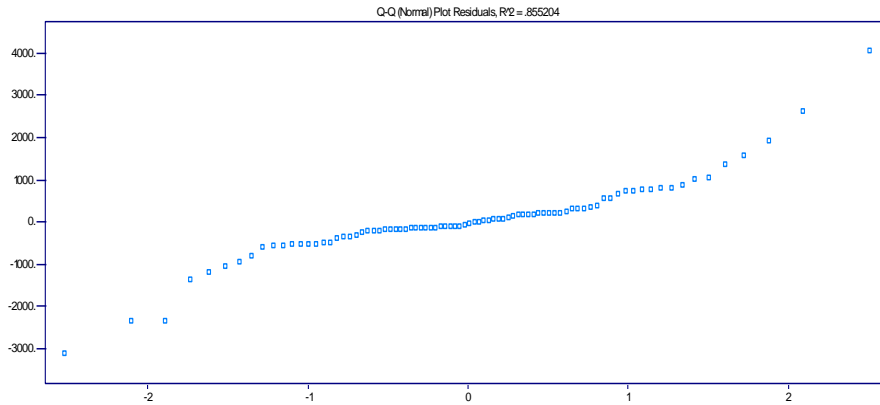


Σχήμα 44. ACF, PACF των υπολοίπων

Οι έλεγχοι των υπολοίπων που φαίνονται παρακάτω, μας οδηγούν να δεχθούμε πως τα υπόλοιπα είναι λευκός θόρυβος και πως το μοντέλο είναι κατάλληλο για τη δημιουργία προβλέψεων.



Σχήμα 45. Ιστόγραμμα των υπολοίπων

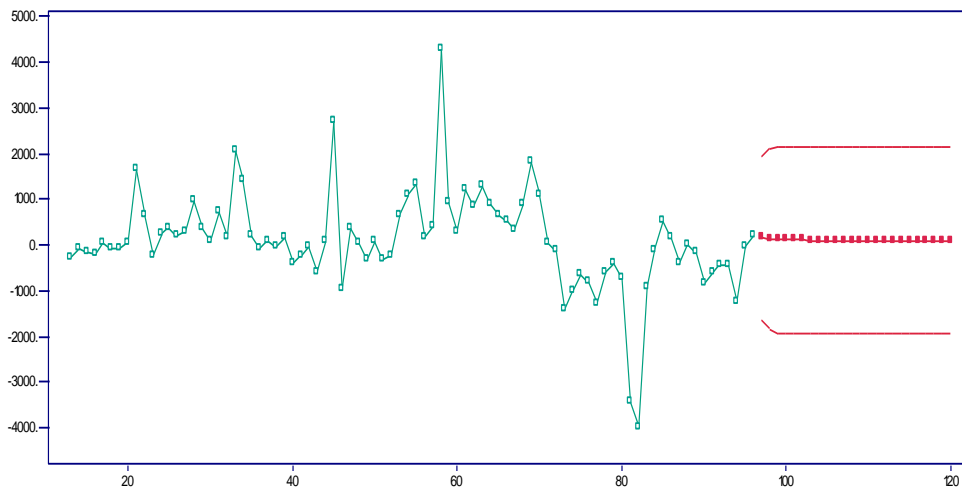


Σχήμα 46. Normal qq-plot των υπολοίπων

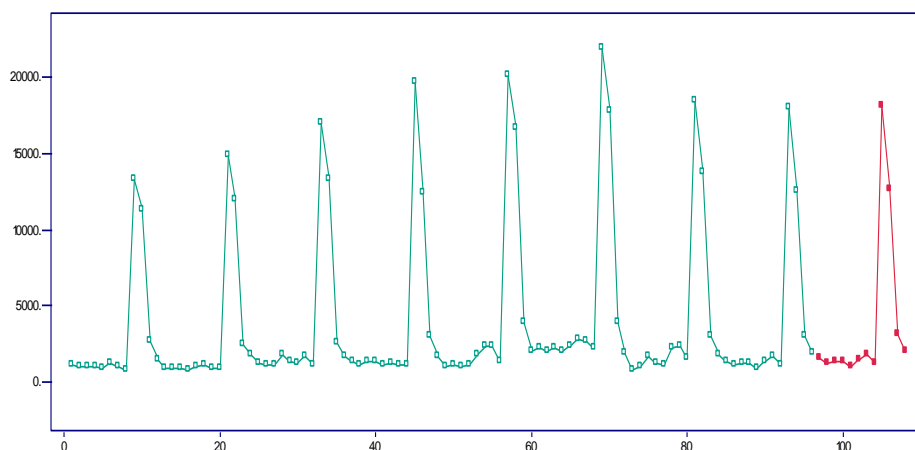
Ljung - Box statistic = 18.580 Chi-Square (20), p-value = .54923
McLeod - Li statistic = 21.325 Chi-Square (21), p-value = .43924
Turning points = 53.000~AN(54.667,sd = 3.8224), p-value = .66282
Diff sign points = 42.000~AN(41.500,sd = 2.6615), p-value = .85098
Rank test statistic = .15780E+04~AN(.17430E+04,sd = .12943E+03), p-value = .20237
Jarque-Bera test statistic (for normality) = .11625E+03 Chi-Square (2), p-value = .00000
Order of Min AICC YW Model for Residuals = 0

Πίνακας 15. Tests of randomness των υπολοίπων

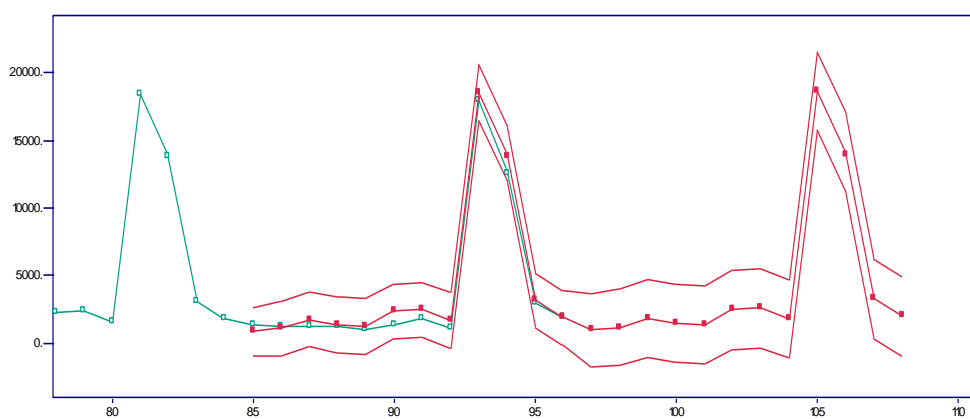
Στο Σχήμα 47 φαίνονται οι προβλέψεις για τη διαφοροποιημένη χρονοσειρά, ενώ στο Σχήμα 48 οι προβλέψεις για την αρχική.



Σχήμα 47. Προβλέψεις 24 μηνών για τη διαφοροποιημένη χρονοσειρά πωλήσεων



Σχήμα 48. Προβλέψεις 12 μηνών για την αρχική χρονοσειρά



Σχήμα 49. Προβλέψεις για το 2011, 2012 με 95% διαστήματα εμπιστοσύνης

Τέλος, στον παρακάτω πίνακα φαίνονται οι προβλέψεις με τα αντίστοιχα 95% διαστήματα εμπιστοσύνης τους.

Step	Prediction	sqrt(MSE)	Lower	Upper
1	.15383E+04	.91256E+03	-.25023E+03	.33269E+04
2	.13034E+04	.10167E+04	-.68926E+03	.32961E+04
3	.13842E+04	.10403E+04	-.65470E+03	.34231E+04
4	.13886E+04	.10459E+04	-.66122E+03	.34385E+04
5	.10789E+04	.10472E+04	-.97358E+03	.31314E+04
6	.14976E+04	.10475E+04	-.55555E+03	.35507E+04
7	.18569E+04	.10476E+04	-.19636E+03	.39102E+04
8	.12506E+04	.10476E+04	-.80271E+03	.33040E+04
9	.18126E+05	.10476E+04	.16073E+05	.20180E+05
10	.12626E+05	.10476E+04	.10573E+05	.14680E+05
11	.31023E+04	.10476E+04	.10490E+04	.51557E+04
12	.20733E+04	.10476E+04	19.98121	.41267E+04

Πίνακας 16. Προβλέψεις 12 μηνών για το έτος 2012 με 95% διάστημα εμπιστοσύνης

Συγκρίνοντας τις τιμές των προβλέψεων με τις πραγματικές πωλήσεις του 2012 μπορούμε να αξιολογήσουμε το μοντέλο εξετάζοντας τα σφάλματα που φαίνονται στον παρακάτω πίνακα.

2012	X_t	F_t	ME	-21,08
1	1484	1538,3	MAE	202,03
2	1311	1303,4	MAPE	4,84%
3	1523	1384,2	MSE	136271,70
4	1433	1388,6	U stat	0,0032
5	985	1078,9		
6	1512	1497,6		
7	1808	1856,9		
8	1256	1125,6		
9	19129	18126,0		
10	11875	12626,0		
11	3073	3102,3		
12	1965	2073,3		

Πίνακας 17. Σύγκριση προβλέψεων με τις πραγματικές πωλήσεις για το 2012

6 ΣΥΜΠΕΡΑΣΜΑΤΑ

Στόχος της παρούσας διπλωματικής εργασίας ήταν η χρήση μοντέλων ARMA, ARIMA και SARIMA για την πρόβλεψη της μελλοντικής ζήτησης στο χώρο των επιχειρήσεων, με εφαρμογή στη φαρμακευτική αγορά. Παραδοσιακά, τέτοιες τεχνικές πρόβλεψης δε χρησιμοποιούνται από τις επιχειρήσεις. Συνήθως οι μέθοδοι που χρησιμοποιούνται δεν επιλέγονται με βάση τη διαχρονική ποιότητα των προβλέψεων που παράγουν, ούτε με βάση την ποιότητα προσαρμογής των μοντέλων σε κάθε κατηγορία προϊόντων ξεχωριστά. Αντίθετα, επιλέγονται με κριτήρια όπως η ευκολία στη χρήση, ο βαθμός κατανόησης των μεθόδων και η ταχύτητα παραγωγής προβλέψεων.

Επίσης έγινε μία προσπάθεια να αναδειχθεί η σημαντικότητα των ποιοτικών προβλέψεων στον έλεγχο αποθεμάτων. Τα αποθέματα διαδραματίζουν σημαντικό ρόλο στη λειτουργία μιας επιχείρησης και συνεπώς η σωστή παρακολούθησή τους είναι αναγκαία για την επιβίωση της επιχείρησης. Αντίθετα, η υπερβολική, ή ανεπαρκής τήρηση αποθεμάτων έχει εξαιρετικά αρνητικές επιπτώσεις, ιδιαίτερα στο σημερινό ανταγωνιστικό οικονομικό κλίμα, όπου οι επιχειρήσεις όχι μόνο ανταγωνίζονται μεταξύ τους στο επίπεδο εξυπηρέτησης και ποιότητας των προσφερόμενων υπηρεσιών και προϊόντων, αλλά προσπαθούν να επιβιώσουν σε ένα διαρκώς μεταβαλλόμενο και αβέβαιο περιβάλλον.

Στην εργασία αυτή έγινε μια προσπάθεια να αναδειχθούν τα θετικά αποτελέσματα που προκύπτουν από την ενσωμάτωση προηγμένων μεθόδων πρόβλεψης για τη ζήτηση στη λειτουργία των επιχειρήσεων. Παράλληλα, έγινε μια εισαγωγή στις μεθόδους αυτές με εφαρμογές και παραδείγματα ενδιάμεσων σταδίων, μέχρι τη δημιουργία προβλέψεων και τη σύγκρισή τους με τις πραγματικές παρατηρήσεις, όταν αυτές έγιναν διαθέσιμες.

Όπως αναφέρθηκε στα πρώτα κεφάλαια, χρήση τέτοιων μεθόδων (ARMA, ARIMA) δεν είναι αναγκαία σε όλες τις κατηγορίες προϊόντων. Σύμφωνα με έρευνες (Makridakis, 1984) ενδεχομένως τέτοιες μέθοδοι να έχουν χειρότερα αποτελέσματα από τις απλούστερες δυνατές μεθόδους. Η χρήση τους, όμως σε καίριας σημασίας για την επιχείρηση προϊόντα είναι σε θέση να παράγουν θετικά αποτελέσματα.

Γενικά συνιστάται η χρήση συνδυασμών μεθόδων πρόβλεψης. Συγκεκριμένα για τα παραδείγματα του προηγούμενου κεφαλαίου, έγινε φανερό πως ενώ οι βραχυπρόθεσμες προβλέψεις, δηλαδή οι προβλέψεις που έγιναν για τους 2-3 πρώτους μήνες του επόμενου έτους ήταν κατά κανόνα ακριβείς, όσο μεγάλωνε ο ορίζοντας της πρόβλεψης, τόσο απέκλιναν τα διαστήματα εμπιστοσύνης και υπήρχαν περιπτώσεις όπου η πρόβλεψη ήταν ανεπιτυχής. Αυτό οφείλεται κυρίως στην οικονομική κρίση του 2012 που επηρέασε αρνητικά τις πωλήσεις της συγκεκριμένης επιχείρησης. Συνεπώς τα μοντέλα που επιλέχθηκαν με βάση το ιστορικό των πωλήσεων δεν ήταν σε θέση να προβλέψουν την πτώση των πωλήσεων.

Για το λόγο αυτό, η χρήση οποιουδήποτε μοντέλου πρόβλεψης πρέπει να εφαρμόζεται σε συνδυασμό με την υποκειμενική κρίση του χρήστη ή της διοίκησης προκειμένου να προσαρμοστούν στα πραγματικά δεδομένα της οικονομίας και της πορείας της εταιρίας.

7 ΠΑΡΑΡΤΗΜΑ

Παραθέτουμε τα αριθμητικά δεδομένα που χρησιμοποιήθηκαν στο Κεφάλαιο 5.

Έτος	Μήνας	Εμβόλια	Zantac	Glucophage	Mesulide
2004	1	1176	1759	3642	2285
2004	2	1024	1666	4975	2688
2004	3	1051	2507	5644	3638
2004	4	1025	2878	3792	4040
2004	5	945	2345	4995	3817
2004	6	1237	2582	4893	3698
2004	7	983	2144	4316	3631
2004	8	851	2678	4590	3965
2004	9	13252	4194	5525	5579
2004	10	11306	2004	3940	2985
2004	11	2653	4242	6503	5518
2004	12	1526	6706	8074	8186
2005	1	898	1476	3168	2190
2005	2	949	1780	3685	2684
2005	3	882	2241	4644	3621
2005	4	837	2278	5508	3244
2005	5	981	2255	5344	3136
2005	6	1170	2423	6209	3967
2005	7	925	2238	5223	3159
2005	8	915	2409	5971	3772
2005	9	14897	3694	7270	5382
2005	10	11944	2034	6086	3242
2005	11	2431	3678	6667	6359
2005	12	1769	5475	6705	7078
2006	1	1265	1534	5753	1817
2006	2	1155	1760	4549	2085
2006	3	1164	2503	6654	3590
2006	4	1804	2086	5890	2956
2006	5	1365	3345	8294	4358
2006	6	1267	2209	5564	3468
2006	7	1676	2456	7632	3049
2006	8	1094	2210	5520	3736
2006	9	16943	3124	6096	4581
2006	10	13347	3562	9178	6082
2006	11	2630	2960	8643	3584
2006	12	1705	5799	7431	8665
2007	1	1341	1627	10945	2739
2007	2	1105	1933	8402	2465
2007	3	1334	2813	5325	2933
2007	4	1405	2498	4721	2848
2007	5	1120	2616	5742	3412
2007	6	1218	2522	7090	2973
2007	7	1092	2468	6681	2821
2007	8	1198	2585	6777	3626
2007	9	19642	3836	6841	4723
2007	10	12364	2640	6927	2218

2007	11	2992	2909	8112	2379
2007	12	1731	4086	9665	5381
2008	1	1040	1927	5878	2294
2008	2	1193	2147	5518	1850
2008	3	1001	2156	6957	2635
2008	4	1164	2020	7524	2187
2008	5	1762	2548	7493	2937
2008	6	2322	2500	9345	2392
2008	7	2420	2410	8313	2496
2008	8	1383	2545	7903	3088
2008	9	20066	2885	10510	3300
2008	10	16653	2230	7120	3181
2008	11	3926	2825	8010	2725
2008	12	2032	5171	14009	6382
2009	1	2271	2456	4709	2287
2009	2	2034	1903	4735	2000
2009	3	2301	2315	6875	2470
2009	4	2066	2546	6983	2938
2009	5	2409	2366	6926	2096
2009	6	2859	3027	8998	3054
2009	7	2762	2519	7475	2495
2009	8	2293	2480	6747	2694
2009	9	21897	2955	8378	3355
2009	10	17756	2467	6897	2900
2009	11	3963	2846	8518	2655
2009	12	1902	3973	13370	4516
2010	1	854	1561	4359	1838
2010	2	1013	1966	4722	2070
2010	3	1669	3018	10244	2623
2010	4	1269	2587	6141	2559
2010	5	1111	2384	6566	2238
2010	6	2243	2807	8030	2655
2010	7	2349	2545	8432	2617
2010	8	1569	2538	7489	2475
2010	9	18454	2910	7672	3114
2010	10	13760	2755	7772	2374
2010	11	3038	2980	8197	3092
2010	12	1771	4724	13014	4636
2011	1	1388	2014	5238	1849
2011	2	1176	1676	4728	1815
2011	3	1268	2251	5732	1947
2011	4	1278	2435	5694	1843
2011	5	971	2405	6715	2117
2011	6	1391	2277	7047	2097
2011	7	1751	2372	7130	2427
2011	8	1145	2953	7308	2557
2011	9	18021	2948	7618	2572
2011	10	12521	2859	5996	2230
2011	11	2997	2799	7426	2275
2011	12	1968	4795	11986	3514

Βιβλιογραφία

- Armstrong, S. J. (1984, November). Forecasting by Extrapolation: Conclusions from 25 Years of Research. *Interfaces*, #14 , σσ. 52-66.
- Box, G. P., & Jenkins, G. (1970). *Time Series Analysis, Forecasting and Control*. Holden-Day.
- Brockwell, P. J., & Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. New York: Springer.
- Diebold, F. X. (2004). *Elements of Forecasting*. Thomson South-Western.
- Makridakis, S. (1984). *The Forecasting Accuracy of Major Time Series Methods*. John Wiley & Sons.
- Makridakis, S., & Winkler, R. L. (1983, September). Averages of Forecasts: Some Empirical Results. *Management Science*, #29 , pp. 987-996.
- Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting Methods and Applications*. New York: John Wiley & Sons.
- Mentzer, J. T., & Cox, J. E. (1984, 3 1). Familiarity, Application and Performance of Sales Forecasting Techniques. *Journal of Forecasting* , σσ. 27-36.
- Nahmias, S. (2004). *Production and Operations Analysis*. McGraw-Hill.
- Silver, E. A., Pyke, D. F., & Peterson, R. (1998). *Inventory Management and Production Planning and Scheduling*. Wiley.
- Yaffee, R. (2000). *Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS*. San Diego: Academic Press.
- Κοκολάκης, Γ. Ε. (2006). *Σημειώσεις Ανάλυσης Χρονοσειρών*. Ε.Μ.Π.