



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

ΤΟΜΕΑΣ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ & ΕΠΙΧΕΙΡΗΣΙΑΚΗΣ ΕΡΕΥΝΑΣ

**Μέθοδοι Πρόγνωσης Ζήτησης : Εφαρμογή σε Παραγωγική
Επιχείρηση του Κλάδου των Ποτών - Τροφίμων**

Θεοδοσίου Ανδρέας

Επιβλέπων: Νικόλαος Α. Παναγιώτου

Αθήνα, 2015



Περίληψη

Η παρούσα διπλωματική εργασία εκπονήθηκε στα πλαίσια του προγράμματος σπουδών της σχολής Μηχανολόγων Μηχανικών του ΕΜΠ. Το θέμα της εργασίας εντάσσεται στα ερευνητικά ενδιαφέροντα του Τομέα Βιομηχανικής Διοίκησης και Επιχειρησιακής Έρευνας και στο επιστημονικό πεδίο της κατεύθυνσης του Μηχανικού Παραγωγής.

Ο αποτελεσματικός και αποδοτικός προγραμματισμός των λειτουργικών διαδικασιών σε μια επιχείρηση, εξαρτάται σε μεγάλο βαθμό από τις προβλέψεις της ζήτησης. Αρκεί να αναλογιστεί κανείς ότι οι κρίσιμες αποφάσεις που σχετίζονται με τον έλεγχο των αποθεμάτων, όπως οι ποσότητες και ο χρόνος έκδοσης των εντολών παραγγελίας ή παραγωγής, τα αποθέματα ασφάλειας και άλλα, βασίζονται σε προβλέψεις της ζήτησης, η ακρίβεια των οποίων επηρεάζει τη συνολική απόδοση των λειτουργιών μιας επιχείρησης.

Συνειδητοποιώντας την κρισιμότητα του ρόλου των προβλέψεων για τις επιχειρήσεις που δραστηριοποιούνται στο γεμάτο αβεβαιότητα πλαίσιο της σημερινής αγοράς, μέσω της παρούσας διπλωματικής εργασίας επιχειρείται η παρουσίαση και ανάλυση των βασικών μεθόδων πρόβλεψης που χρησιμοποιούνται ως εργαλείο για την υποβοήθηση των στελεχών και συμβούλων, στις διαδικασίες λήψης κρίσιμων αποφάσεων.

Ειδικότερα, παρουσιάζονται οι ποσοτικές μέθοδοι πρόβλεψης που βασίζονται στην ανάλυση χρονοσειρών, ενώ αναφορά γίνεται και στις μεθόδους πρόβλεψης μέσω των μοντέλων ARIMA. Τα μοντέλα ARIMA αναπτύσσονται μέσα από τη μεθοδολογία Box – Jenkins, η οποία επίσης αποτελεί αντικείμενο μελέτης. Η εργασία ολοκληρώνεται με τη μελέτη περίπτωσης εταιρείας που δραστηριοποιείται στο χώρο της παραγωγικής βιομηχανίας, για τις ανάγκες της οποίας αναλύθηκε η δοθείσα χρονοσειρά των πραγματικών πωλήσεων ενός προϊόντος της και αναπτύχθηκε σχετικό μοντέλο πρόβλεψης ζήτησης.



Πίνακας περιεχομένων

1. Εισαγωγή.....	5
1.1 Η σημασία των προβλέψεων	5
1.2 Η πρόβλεψη ζήτησης στην εφοδιαστική αλυσίδα	6
1.3 Χρονικός ορίζοντας προβλέψεων.....	8
1.4 Μέθοδοι προβλέψεων.....	8
1.4.1 Ποιοτικές μέθοδοι πρόβλεψης.....	9
1.4.2 Ποσοτικές μέθοδοι πρόβλεψης.....	10
1.5 Χαρακτηριστικά των προβλέψεων	11
1.6 Διαδικασία πρόβλεψης	12
2. Ανάλυση Χρονοσειρών και Μέθοδοι Εξομάλυνσης	15
2.1 Στοιχεία ανάλυσης χρονοσειρών	15
2.2 Αποσύνθεση χρονοσειράς	19
2.3 Ποσοτικές μέθοδοι με εξομάλυνση	20
2.3.1 Η μέθοδος του κινητού μέσου όρου	21
2.3.2 Απλή εκθετική εξομάλυνση	23
2.3.3 Εκθετική εξομάλυνση με γραμμική τάση (Μέθοδος Holt)	31
2.3.4 Εκθετική εξομάλυνση με τάση και εποχικότητα (Μέθοδος Winter)	32
2.4 Ανάλυση Χρονοσειρών στις επιχειρήσεις.....	33
3. Σφάλματα πρόβλεψης	34
3.1 Ορισμοί και τύποι σφαλμάτων	34
3.2 Ερμηνεία των σφαλμάτων	36
3.3 Παρακολούθηση πρόβλεψης	37
4. Ανάλυση ARIMA	41
4.1 Αυτοπαλίνδρομα μοντέλα και ανάλυση χρονοσειρών	41
4.2 Βασικά στοιχεία της ανάλυσης ARIMA	41
4.2.1 Συντελεστής αυτοσυσχέτισης.....	41
4.2.2 Κανονική και Εποχική διαφόριση χρονοσειρών	48
4.2.3 Συντελεστής μερικής αυτοσυσχέτισης.....	53
4.3 Μοντέλα ARIMA	53
4.3.1 Αυτοπαλίνδρομα μοντέλα AR(p)	53



4.3.2 Μοντέλα κινητού μέσου όρου MA(q).....	54
4.3.3 Αυτοπαλίνδρομα μοντέλα κινητού μέσου όρου ARMA(p,q)	55
4.3.4 Ολοκληρωμένα αυτοπαλίνδρομα μοντέλα κινητού μέσου όρου	56
4.3.5 Ολοκληρωμένα αυτοπαλίνδρομα μοντέλα κινητού μέσου όρου με εποχικότητα.....	56
4.4 Μεθοδολογία Box-Jenkins	57
4.5 Θεωρητικές κατανομές συντελεστών αυτοσυσχέτισης και μερικής αυτοσυσχέτισης.....	62
4.6 Πρόσθετα κριτήρια επιλογής μοτέλου ARIMA	63
5. Μελέτη Περίπτωσης	65
5.1 Μοντέλο πρόβλεψης ARIMA	65
5.2 Μοντέλο πρόβλεψης Winters	79
5.3 Μοντέλο πρόβλεψης Holts	83
5.4 Μοντέλα πρόβλεψης για διμηνιαίες και τριμηνιαίες πωλήσεις.....	85
5.4.1 Μοντέλο πρόβλεψης διμηνιαίων πωλήσεων ARIMA.....	85
5.4.2 Μοντέλο πρόβλεψης τριμηνιαίων πωλήσεων Winters.....	91
5.5 Σύγκριση προβλέψεων ARIMA και Winters	95
6. Συμπεράσματα	98
7. Παραρτήματα.....	100
8. Βιβλιογραφία	110



"Είναι πολύ καλύτερο το να προβλέπει κανείς χωρίς βεβαιότητα από το να μην προβλέπει καθόλου"
(Henri Poincare, *The Foundations of Science*)

1. Εισαγωγή

1.1 Η σημασία των προβλέψεων

Η γνώση των μελλοντικών τιμών διαφόρων μεγεθών που έχουν άμεση σχέση με το περιβάλλον μιας επιχείρησης και τα αποτελέσματα από την καθημερινή λειτουργία της (όπως π.χ. έσοδα και όγκος πωλήσεων, ύψος της ζήτησης των παραγόμενων προϊόντων ή των νέων προϊόντων που πρόκειται να παραχθούν, τεχνολογικές εξελίξεις κ.α.), είναι απολύτως απαραίτητα, μια και μόνο σε αυτή τη γνώση μπορούμε να στηρίξουμε τη λήψη βέλτιστων επιχειρηματικών αποφάσεων, για τη σχεδίαση και τον προγραμματισμό του συνόλου των διαδικασιών της. Επομένως κάλλιστα θα μπορούσε να πει κανείς ότι η αποδοτική λειτουργία μιας επιχείρησης, εξαρτάται κατά πολύ από την ικανότητα της διοίκησης και των στελεχών της, να αποκτούν έγκυρες και αξιόπιστες προβλέψεις για τις μελλοντικές τιμές των διαφόρων μεταβλητών που επηρεάζουν την καθημερινή λειτουργία της.

Σύμφωνα με τον Bernstein (1996) οι προβλέψεις είναι ένα από τα σημαντικότερα δεδομένα εισόδου, βάσει των οποίων τα στελέχη καλούνται να υποστηρίξουν τις διαδικασίες λήψης επιχειρηματικών αποφάσεων. Πρακτικά, κάθε επιχειρηματική απόφαση που αφορά τη διοίκηση λειτουργιών, βασίζεται σε μια πρόβλεψη. Η συσσώρευση αποθεμάτων σχετίζεται με την πρόβλεψη για την αναμενόμενη ζήτηση, το τμήμα παραγωγής πρέπει να προγραμματίσει τις ανάγκες σε δυναμικό και πρώτες ύλες για τον επόμενο μήνα, το οικονομικό τμήμα πρέπει να διευθετήσει τις χρηματοοικονομικές ανάγκες της επιχείρησης, το τμήμα ανθρωπίνου δυναμικού πρέπει να αποφασίσει εάν θα αυξήσει ή να μειώσει το εργατικό δυναμικό κ.α.

Όπως αντιλαμβάνεται λοιπόν κανείς, σε ένα επιχειρησιακό περιβάλλον, οι λειτουργίες για τις οποίες απαιτείται η εκ προοιμίου λήψη μιας συγκεκριμένης απόφασης, ποσοτικής ή/και ποιοτικής, ώστε να προγραμματιστούν με αποδοτικό και οικονομικό τρόπο, εκτείνονται σε όλα τα οργανωτικά και λειτουργικά του επίπεδα. Υπό το πρίσμα αυτό, η ανάγκη για τη διενέργεια προβλέψεων οι οποίες σε συνδυασμό με τις γνώμες των έμπειρων στελεχών και συμβούλων, θα χρησιμοποιηθούν ως δεδομένα υποστήριξης των διαδικασιών λήψης κρίσιμων αποφάσεων για τη λειτουργία μιας επιχείρησης, είναι κάτι παραπάνω από επιτακτική.



1.2 Η πρόβλεψη ζήτησης στην εφοδιαστική αλυσίδα

Σε κάθε επιχείρηση, φτάνει από την αγορά μια ζήτηση για τα προϊόντα ή τις υπηρεσίες που παράγει. Από την εξωτερική αυτή ζήτηση προκύπτουν οι ανάγκες για την παραγωγή προϊόντων ή υπηρεσιών καθώς και οι ανάγκες για τους πόρους που πρέπει να εξασφαλίσει η επιχείρηση όπως : πρώτες ύλες, διαθέσιμη παραγωγική δυναμικότητα, εργατικό δυναμικό κ.α.. Ισχύει ότι η πρόβλεψη ζήτησης είναι το κυριότερο δεδομένο εισόδου για την κατάρτιση του γενικού πλάνου παραγωγής μιας επιχείρησης.

Είναι αποδεδειγμένο ότι ένα αποτελεσματικό πλάνο παραγωγής, είναι η βάση για την αποδοτική διοίκηση των λειτουργιών μιας επιχείρησης. Στόχος είναι μέσω της κατάρτισης του πλάνου παραγωγής, να επιτυγχάνεται ικανοποίηση των αναγκών σε προϊόντα, διαθεσιμότητα και χρόνους παράδοσης. Ακριβείς προβλέψεις, συνεπάγονται αυτόματα και πιο αντιπροσωπευτικό πλάνο παραγωγής που με τη σειρά του θα έχει ως αποτέλεσμα :

- Αυξημένη αξιοποίηση του δυναμικού της επιχείρησης
- Αποτελεσματικότερη οργάνωση παραγωγής και προμηθειών
- Μείωση του κόστους αποθεμάτων
- Αυξημένη ικανοποίηση των πελατών – ενίσχυση αξιοπιστίας της επιχείρησης

Η διαδικασία που ακολουθείται για την πρόβλεψη ζήτησης, είναι μια γέφυρα επικοινωνίας μεταξύ της αγοράς και της επιχείρησης. Μια γέφυρα μέσω της οποίας η επιχείρηση αποκτά ποσοτικές και ποιοτικές πληροφορίες για τις μελλοντικές, πιθανές καταναλωτικές τάσεις της αγοράς. Οι πληροφορίες αυτές είναι απαραίτητες προκειμένου κάθε επιχείρηση να είναι σε θέση να συνδυάσει το βέλτιστο βαθμό εξυπηρέτησης πελατών με το χαμηλότερο δυνατό λειτουργικό κόστος. Δικαιολογημένα επομένως οι προβλέψεις θεωρούνται ως απαραίτητο εργαλείο υποβοήθησης στη λήψη αποφάσεων σχετικά με τον προγραμματισμό των κρίσιμων λειτουργιών της επιχείρησης.

Στην παραγωγική βιομηχανία, οι επιχειρησιακές λειτουργίες για τις οποίες είναι απαραίτητες οι προβλέψεις είναι οι εξής :

Παραγωγή : προγραμματισμός παραγωγής, έλεγχος αποθεμάτων, συγκεντρωτικός προγραμματισμός, προμήθεια πρώτων υλών, προγραμματισμός δυναμικότητας

Μάρκετινγκ : προωθήσεις προϊόντων, εισαγωγή νέων προϊόντων

Ανθρώπινο δυναμικό : προγραμματισμός εργατικού δυναμικού, πόρων

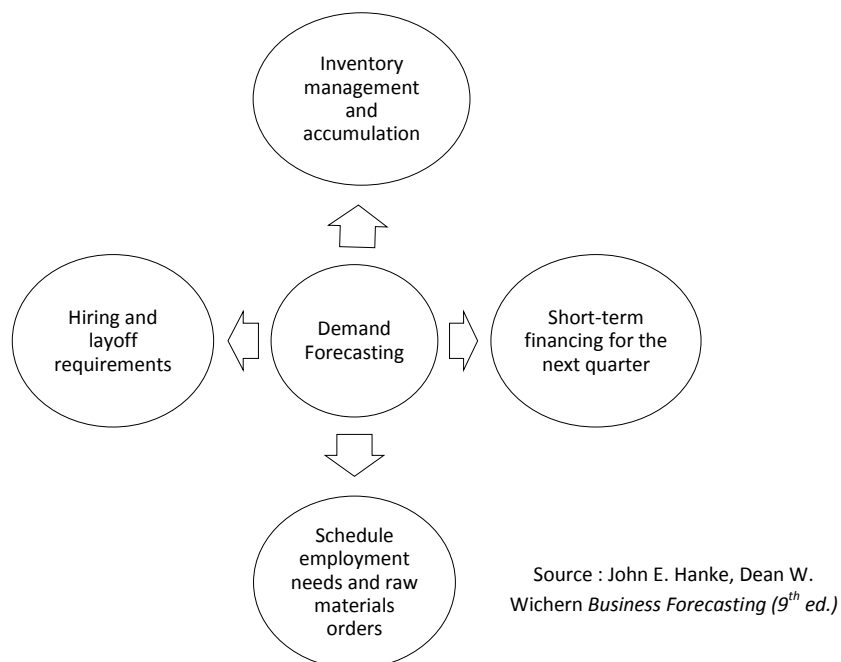
Χρηματοοικονομικά : επενδύσεις εργοστασίων / εξοπλισμού



Αντιλαμβάνεται λοιπόν κανείς ότι οι κατά το δυνατό πιο ακριβείς προβλέψεις, αποδεικνύεται να είναι ζωτικής σημασίας για την αποδοτική λειτουργία των παραγωγικών επιχειρήσεων στο σύγχρονο ανταγωνιστικό περιβάλλον της αγοράς που δραστηριοποιούνται. Ιδιαίτερα, η πρόβλεψη ζήτησης είναι ο πιο κρίσιμος παράγοντας στον τομέα διαχείρισης αποθεμάτων των επιχειρήσεων. Αναλογιζόμενοι τις οικονομικές συγκυρίες της εποχής, κυρίως όσον αφορά το ιδιάζον και πολύκροτο κεφάλαιο της ρευστότητας των επιχειρήσεων, η δέσμευση μεγάλου κεφαλαίου σε απόθεμα λόγω της απουσίας ενός πλάνου παραγωγής και προμηθειών το οποίο είναι άρρηκτα συνδεδεμένο με την πρόβλεψη ζήτησης του προϊόντος/υπηρεσίας στις μελλοντικές χρονικές περιόδους, δυσχεραίνει σε μεγάλο βαθμό την οικονομική ευελιξία της.

Ενδεικτικά αξίζει να αναφερθεί ότι μια μελέτη στην Hewlett-Packard αποκάλυψε ότι το 40% των αποθεμάτων της επιχείρησης ήταν αποθέματα κύκλου και διερχόμενα αποθέματα, ενώ το 60% οφειλόταν στην αβεβαιότητα. Από αυτό το 60%, το 2% οφείλονταν σε διακυμάνσεις της απόδοσης των προμηθευτών, 2% σε διακυμάνσεις της απόδοσης του παραγωγού και 96% στην αβεβαιότητα της ζήτησης.

Το παρακάτω σχήμα είναι αντιπροσωπευτικό για να αντιληφθεί κανείς τη σημασία που έχει συγκεκριμένα η πρόβλεψη ζήτησης για τις διαδικασίες που αφορούν το λειτουργικό επίπεδο μιας παραγωγικής επιχείρησης.



Σχήμα 1.1 : Πρόγνωση ζήτησης και επιχειρησιακές λειτουργίες



1.3 Χρονικός Ορίζοντας Προβλέψεων

Για τον προγραμματισμό όλων των κρίσιμων επιχειρησιακών λειτουργιών, υπάρχει η ανάγκη για λήψη βραχυπρόθεσμων, μεσοπρόθεσμων ή μακροπρόθεσμων αποφάσεων. Σύμφωνα λοιπόν με το είδος της απόφασης που πρέπει να ληφθεί και το χρονικό ορίζοντα στον οποίο εκτείνονται, υπάρχουν τα ακόλουθα επίπεδα προβλέψεων :

Λειτουργικό επίπεδο : οι βραχυπρόθεσμες προβλέψεις (ημέρα, εβδομάδα), είναι κρίσιμες για τον καθημερινό προγραμματισμό σε θέματα παραγωγής, πρώτων υλών, διαχείρισης αποθεμάτων, τοποθέτησης παραγγελιών και αξιοποίησης του εργατικού δυναμικού.

Τακτικό επίπεδο : οι μεσοπρόθεσμες προβλέψεις (εβδομάδες ή μήνες), χρησιμοποιούνται κατά κύριο λόγο για τον καθορισμό του επιπέδου των αποθεμάτων και για τον προγραμματισμό των απαιτήσεων της παραγωγής σε πόρους.

Στρατηγικό επίπεδο : οι μακροπρόθεσμες προβλέψεις (έτος ή έτη) αποτελούν εργαλείο για λήψη αποφάσεων σε θέματα χωροταξίας παραγωγής και αποθηκών, αύξηση δυναμικότητας, αναζήτηση εναλλακτικών προμηθευτών και ανανέωση μηχανολογικού εξοπλισμού.

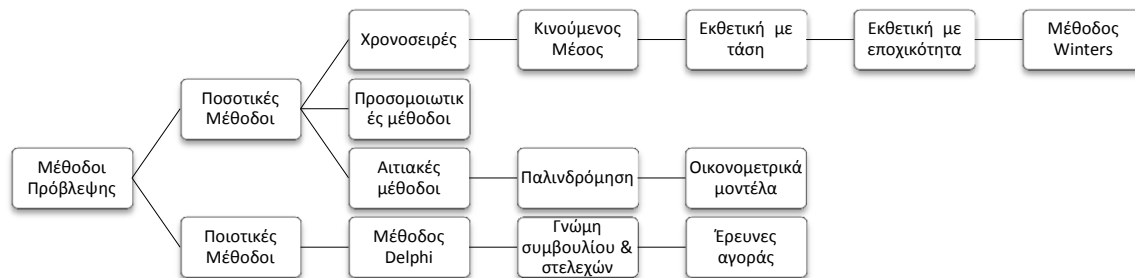
1.4 Μέθοδοι προβλέψεων

Αρκετές μέθοδοι έχουν αναπτυχθεί για τη διενέργεια προβλέψεων στο επιχειρηματικό περιβάλλον. Η επιλογή της καταλληλότερης κάθε φορά μεθόδου, εξαρτάται από διάφορους παράγοντες, όπως το είδος των αποφάσεων που θα ληφθούν με βάση τις προβλέψεις, την περίοδο και το χρονικό ορίζοντα πρόβλεψης, την επιζητούμενη ακρίβεια καθώς και τα διαθέσιμα στοιχεία. Οι δύο βασικές κατηγορίες των χρησιμοποιούμενων μεθόδων πρόβλεψης είναι οι ποιοτικές και οι ποσοτικές (βλ. **Σχήμα 1.1**). Οι ποιοτικές μέθοδοι είναι υποκειμενικές και βασίζονται κυρίως σε γνώμες ειδικών και στελεχών μιας επιχείρησης. Στην κατηγορία των ποσοτικών μεθόδων, κυριαρχούν οι μέθοδοι ανάλυσης των χρονοσειρών.

Τα τελευταία χρόνια, παρασυρόμενοι από τη χαμηλή σε κόστος και υψηλή σε επίπεδο υπολογιστική ισχύ που κυκλοφορεί στο επιχειρησιακό περιβάλλον, αναπτύσσονται νέες μέθοδοι που αποσκοπούν στην επίτευξη όλο και μεγαλύτερης ακρίβειας προβλέψεων, όπως για παράδειγμα τα νευρωνικά δίκτυα. Παρά ταύτα, στις επιχειρήσεις, ακόμη χρησιμοποιούνται κατά κόρον οι ‘συμβατικές’ μέθοδοι προβλέψεων (ανάλυση χρονοσειρών).



Στο παρακάτω σχήμα απεικονίζονται οι βασικές κατηγορίες μεθόδων πρόβλεψης.



Σχήμα 1.2 : Κατηγορίες Μεθόδων Πρόβλεψης

1.4.1 Ποιοτικές Μέθοδοι Πρόβλεψης

Μια ποιοτική πρόβλεψη βασίζεται στην κρίση του ανθρώπινου παράγοντα. Οι πιο χαρακτηριστικές μέθοδοι που χρησιμοποιούνται είναι η συμπλήρωση και ανάλυση ερωτηματολογίων (έρευνα αγοράς), οι εκτιμήσεις του ανθρώπινου δυναμικού και των συμβούλων μιας επιχείρησης, καθώς και η μέθοδος Delphi. Οι ποιοτικές μέθοδοι, χρησιμοποιούνται όταν δεν υπάρχουν καθόλου ή υπάρχουν περιορισμένα ιστορικά στοιχεία. Εφαρμόζονται για προβλέψεις μελλοντικών εξελίξεων στην τεχνολογία, για αγορές αγαθών, πρώτων υλών, κ.λπ. και χαρακτηρίζονται από μεγάλο χρονικό ορίζοντα πρόβλεψης στηρίζοντας στρατηγικές αποφάσεις της διοίκησης.

Γνώμη Συμβουλίου και Στελεχών

Η μέθοδος αυτή βασίζεται στο γεγονός ότι μια πολυπληθής ομάδα ανθρώπων από διαφορετικές θέσεις μπορεί να κάνει μια πιο αξιόπιστη πρόβλεψη απ' ό,τι ένας μεμονωμένος ή λίγοι άνθρωποι. Έτσι, διοργανώνονται ανοιχτές συναντήσεις με ελεύθερη ανταλλαγή απόψεων μεταξύ ανθρώπων από όλο το φάσμα των θέσεων ενός οργανισμού. Ένα μειονέκτημα της μεθόδου αυτής είναι ότι η άποψη των υφιστάμενων «σκεδάζεται» ή υποβαθμίζεται από την άποψη των ανώτερων στην ιεραρχία. Αυτό το μειονέκτημα προσπαθεί να διορθώσει η μέθοδος Delphi.

Μέθοδος Delphi

Η μέθοδος αυτή είναι μια τεχνική πρόβλεψης, στόχος της οποίας είναι η προσέγγιση μιας συμφωνίας μεταξύ μιας ομάδας ειδικών, διατηρώντας την ανωνυμία τους. Η ιδέα πίσω από αυτήν είναι ότι ενώ οι



ειδικοί δεν θα συμφωνήσουν σε όλα τα ζητήματα, εντούτοις σε ότι συμφωνήσουν αυτά κατά πάσα πιθανότητα θα συμβούν.

Σύμφωνα με τη μέθοδο αυτή:

1. Πραγματοποιείται επιλογή των συμμετεχόντων, συνήθως ειδικών από διαφορετικές θέσεις ή επιστημονικά υπόβαθρα,
2. Μέσω ενός ερωτηματολογίου συλλέγονται οι απόψεις όλων (χωρίς ο ένας να δει ή να γνωρίζει τους υπόλοιπους συμμετέχοντες),
3. Οι απαντήσεις όλων ταξινομούνται και επανατροφοδοτούνται (feedback) στους συμμετέχοντες μαζί με ένα καινούργιο ερωτηματολόγιο,
4. Το βήμα 3 επαναλαμβάνεται όσες φορές κρίνεται απαραίτητο προκειμένου να επιτευχθεί μια συμφωνία μεταξύ των συμμετεχόντων. Συνήθως 3 ή 4 επαναλήψεις είναι αρκετές.

Λόγω της ανωνυμίας του καθενός και της ίδιας βαρύτητας όλων των απόψεων, με τη μέθοδο Delphi αποφεύγεται το μειονέκτημα της πρώτης μεθόδου. Από την άλλη, η μέθοδος είναι σχετικά χρονοβόρα.

Έρευνα αγοράς

Αποτελεί μια προσέγγιση που χρησιμοποιεί ερωτηματολόγια και συνεντεύξεις για τον καθορισμό των αναγκών, των προτιμήσεων, των επιλογών κτλ, μιας ομάδας στόχου (π.χ. των καταναλωτών). Η μέθοδος χρησιμοποιείται ευρέως για την βελτίωση και την δημιουργία καινούργιων προϊόντων. Σημαντικό στοιχείο για την επιτυχία της μεθόδου είναι ο σχεδιασμός των ερωτηματολογίων (ή των συνεντεύξεων).

1.4.2 Ποσοτικές Μέθοδοι Πρόβλεψης

Οι ποσοτικές μέθοδοι μπορούν να διακριθούν σε αυτές που βασίζονται σε **μοντέλα χρονοσειρών** (time series models), σε αυτές που βασίζονται σε **αιτιακά μοντέλα** (causal models) και στα **προσομοιωτικά μοντέλα**. Τα πρώτα προϋποθέτουν ότι η απαραίτητη πληροφορία για την πρόβλεψη περιέχεται στη χρονοσειρά των στοιχείων. Χρονοσειρά είναι μια σειρά παρατηρήσεων που λαμβάνονται σε ίσες διαδοχικές χρονικές περιόδους, μέσα σε ένα καθορισμένο χρονικό ορίζοντα. Μέσω της ανάλυσης χρονοσειράς, υπονοείται σιωπηρά ότι μπορεί να γίνει πρόβλεψη με βάση τα μοτίβα (patterns) διακύμανσης των διαθέσιμων δεδομένων από τα ιστορικά στοιχεία. Έτσι, η ανάλυση αυτή επικεντρώνεται στην αναζήτηση προτύπων τάσης, κυκλικότητας, περιοδικότητας κτλ στα δεδομένα, προκειμένου να επιλεγεί το κατάλληλο μοντέλο πρόβλεψης.

Τα αιτιακά μοντέλα χρησιμοποιούν μια αρκετά διαφορετική προσέγγιση για την δημιουργία πρόβλεψης.



Συγκεκριμένα, θεωρούν ότι η μεταβλητή για την οποία θέλουμε να κάνουμε πρόβλεψη είναι εξαρτημένη με κάποιο τρόπο από μία ή περισσότερες παραμέτρους. Η δυσκολία έγκειται στην εύρεση της μαθηματικής σχέσης με την οποία επηρεάζεται η ζητούμενη μεταβλητή από τις παραμέτρους αυτές.

1.5 Χαρακτηριστικά των προβλέψεων

Η ζήτηση που φτάνει στην επιχείρηση από την αγορά για το εκάστοτε προϊόν σε κάθε χρονική περίοδο και η οποία χρησιμοποιείται ως δεδομένο εισόδου στις ποσοτικές μεθόδους πρόβλεψης, αποτελείται από το συστηματικό και το τυχαίο μέρος. Το συστηματικό μέρος είναι ακριβώς εκείνο το οποίο επιχειρούν οι χρησιμοποιούμενες μέθοδοι χρονοσειρών να προβλέψουν. Το τυχαίο μέρος, αποτελεί αναπόσπαστο κομμάτι των προβλέψεων και είναι η κυριότερη πηγή των σφαλμάτων. Εξ ορισμού οι προβλέψεις εμπεριέχουν τον παράγοντα του σφάλματος. Για τον λόγο αυτό κάθε μεθοδολογία που χρησιμοποιείται, έχει ως έξοδο μια πρόβλεψη και ένα μέτρο του αναμενόμενου σφάλματος, το οποίο όπως θα δούμε στη συνέχεια, θα πρέπει να βρίσκεται εντός αποδεκτών ορίων.

Το σφάλμα πρόβλεψης που πηγάζει μέσα από το τυχαίο μέρος της ζήτησης, ενισχύεται ακόμη περισσότερο στις περιπτώσεις όπου πραγματοποιούνται μακροπρόθεσμες προβλέψεις. Ακόμη και διαισθητικά, εύκολα μπορεί κανείς να κατανοήσει ότι όσο μεγαλύτερος είναι ο χρονικός ορίζοντας πρόβλεψης, τόσο λιγότερο ακριβείς είναι οι προβλέψεις. Αυτό συμβαίνει κυρίως λόγω της αδυναμίας να συμπεριληφθούν στην πρόβλεψη οι πιθανοί παράγοντες που θα μεταβάλλουν τη ζήτηση στο μέλλον.

Πέραν όμως από τον παράγοντα χρόνο, στις προβλέψεις υπεισέρχεται και το πλήθος κωδικών των προϊόντων για τα οποία πραγματοποιείται πρόβλεψη. Είναι αποδεδειγμένο ότι προβλέψεις που πραγματοποιούνται για μια ομάδα προϊόντων (αθροιστικές προβλέψεις), είναι ακριβέστερες από τις προβλέψεις για ένα μεμονωμένο προϊόν, λόγω της πιθανής αντιστάθμισης θετικών με αρνητικών αποκλίσεων στην τελική ζήτηση.

Έχοντας λάβει υπόψη σφάλματα, χρονικό ορίζοντα και πλήθος κωδικών, η εφαρμογή κατάλληλης μεθόδου πρόβλεψης, σωστά παραμετροποιημένης, είναι δυνατό να δώσει σχετικά καλές προβλέψεις. Η πρόβλεψη όμως που θα προκύψει, θα πρέπει να χρησιμοποιείται παράλληλα με την αξιοποίηση πρόσθετων πληροφοριών προερχόμενων από τα διάφορα τμήματα μιας επιχείρησης, όπως λόγου χάριν μια πληροφορία που αφορά τη μελλοντική ζήτηση (π.χ. διαφημιστική εκστρατεία) η οποία ενδεχομένως να μην λήφθηκε υπόψη για την εξαγωγή της πρόβλεψης.



Η ακρίβεια της κάθε μεθόδου, μπορεί να ελέγχεται μέσα από τον υπολογισμό και την παρακολούθηση των σφαλμάτων πρόβλεψης. Σε κάθε περίπτωση όμως, τα σφάλματα είναι αναπόφευκτα και ο στόχος δεν είναι να εξαλειφθούν, αυτό είναι εκ των πραγμάτων αδύνατο, αλλά να κυμαίνονται εντός αποδεκτών ορίων. Είναι κοινά αποδεκτό ότι το ιδιαίτερα ασταθές περιβάλλον της αγοράς στο οποίο δραστηριοποιούνται σήμερα οι επιχειρήσεις, επιβάλλει την ύπαρξη ενός στιβαρού συστήματος παραγωγικού σχεδιασμού για καλύτερη απόκριση στα αναμενόμενα σφάλματα και μεγαλύτερη ευελιξία στις μεταβολές της ζήτησης, έτσι ώστε να είναι σε θέση να ανταποκρίνονται στις απαιτήσεις των πελατών τους.

1.6 Διαδικασία Πρόβλεψης

Η εξαγωγή προβλέψεων είναι μια ιδιαίτερα σημαντική διαδικασία που απαιτεί προσεκτική μελέτη και σχεδιασμό. Σύμφωνα με τη διεθνή βιβλιογραφία, έχουν καθοριστεί πέντε βασικά βήματα που πρέπει να ακολουθηθούν για τη σωστή εξαγωγή και αξιολόγηση προβλέψεων :

Καθορισμός του προβλήματος : είναι το σημαντικότερο βήμα στην διαδικασία της πρόβλεψης. Προϋποθέτει βαθιά κατανόηση του τρόπου με τον οποίο θα χρησιμοποιηθούν οι προβλέψεις καθώς και των αναγκών του ανθρώπου που τις χρειάζεται. Καθορίζονται βασικές παράμετροι της πρόβλεψης, όπως η προς εξέταση μεταβλητή, ο χρονικός ορίζοντας πρόβλεψης, η μορφή παρουσίασης των αποτελεσμάτων της πρόβλεψης κ.α.

Συλλογή δεδομένων : αποτελεί μια χρονοβόρα διαδικασία τόσο για την ορθή συλλογή όσο και για την αποθήκευση και συντήρηση των δεδομένων. Οι απαραίτητες πληροφορίες, εκτός από αριθμητικά δεδομένα, μπορεί να είναι κρίσεις και απόψεις ειδικών και εμπειρογνομόνων. Τα τελευταία χρόνια, λόγω και της εξέλιξης των πληροφοριακών συστημάτων διοίκησης στις επιχειρήσεις, όλες οι απαραίτητες πληροφορίες τηρούνται σε βάσεις δεδομένων και είναι άμεσα προσβάσιμες για ανάλυση και επεξεργασία.

Προκαταρκτική ανάλυση και διερεύνηση : στο βήμα αυτό γίνεται προσπάθεια αναγνώρισης και απομόνωσης των συνιστωσών που συνθέτουν κάθε χρονοσειρά, όπως η εποχικότητα, η τάση, ο κύκλος, το οριζόντιο στοιχείο και ο θόρυβος. Προς την κατεύθυνση αυτή, είναι πολλαπλά χρήσιμη η γραφική παράσταση των δεδομένων και η στατιστική επεξεργασία τους. Επίσης, πραγματοποιείται και μια



πρώτη αναπροσαρμογή των δεδομένων ώστε να προκύψει μια εξομαλυμένη χρονοσειρά επί της οποίας θα επιχειρηθούν οι προβλέψεις.

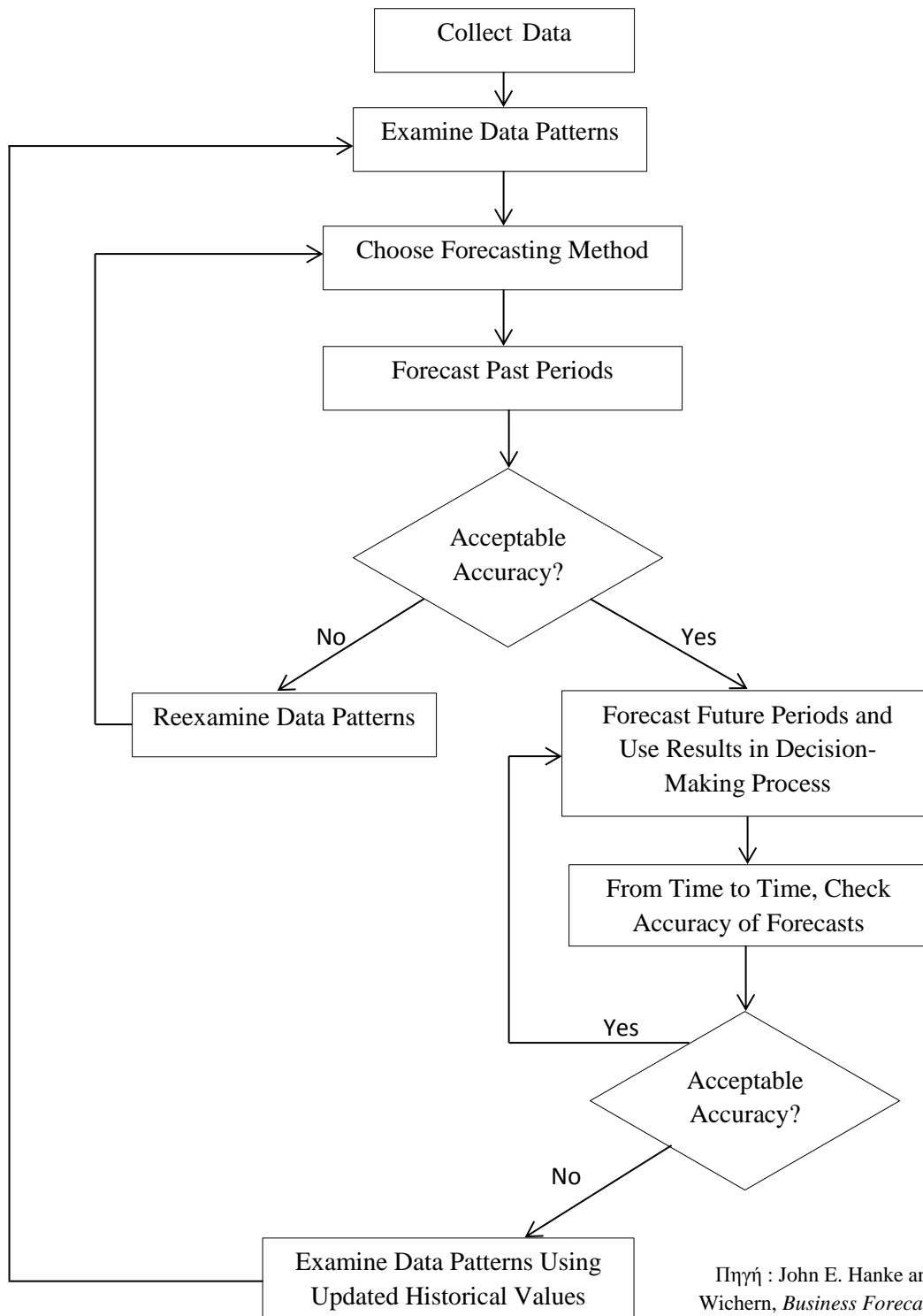
Επιλογή μεθόδων πρόβλεψης : Για να αποφανθεί κατά πόσο μια μέθοδος πρόβλεψης είναι η κατάλληλη, αξιολογούνται διεξοδικά τα μοντέλα πρόβλεψης που έχουμε στη διάθεσή μας, όπως π.χ., μοντέλα κινητών μέσων, εκθετικής εξομάλυνσης, απλής ή πολλαπλής παλινδρόμησης ARIMA κ.α. Βάση κριτηρίων, ανάλογα με το σφάλμα, επιλέγεται κάθε φορά το κατάλληλο μοντέλο πρόβλεψης. Η αποτελεσματική επιλογή του κατάλληλου μοντέλου στηρίζεται, αμφοτέρωθεν, στα μαθηματικά εργαλεία που υπάρχουν διαθέσιμα αλλά και στην εμπειρία και τις αναλυτικές ικανότητες του μελετητή που διενεργεί τις προβλέψεις

Χρήση και αξιολόγηση των μοντέλων πρόβλεψης : Μετά την επιλογή του κατάλληλου μοντέλου πρόβλεψης και την εκτίμηση των παραμέτρων του, αυτό είναι έτοιμο να χρησιμοποιηθεί συστηματικά στο μέλλον και να έχουμε στη διάθεσή μας προβλέψεις μελλοντικών τιμών της μεταβλητής που μελετάται. Από τη στιγμή όμως που θα έχουμε στη διάθεσή μας και τις πραγματικές τιμές της υπό πρόβλεψη μεταβλητής, είμαστε πλέον σε θέση να αξιολογήσουμε –εκ του αποτελέσματος- την ακρίβεια ή όχι των προβλέψεων. Η αξιολόγηση πραγματοποιείται μέσω εξειδικευμένων στατιστικών δεικτών συστηματικά με την πάροδο του χρόνου. Η παρακολούθηση των σφαλμάτων αυτών βοηθά στην έγκαιρη αναπροσαρμογή των παραμέτρων του μοντέλου ή και ακόμα, ολόκληρης της διαδικασίας πρόβλεψης.

Όπως εύκολα αντιλαμβάνεται κανείς, η ευχέρεια (ή δυσχέρεια) στην εφαρμογή των διαδικασιών που απαιτούνται για την πραγματοποίηση προβλέψεων, εξαρτάται σε μεγάλο βαθμό από τη χρησιμοποιούμενη μέθοδο αλλά και από τη σχετική κατάρτιση των στελεχών της επιχείρησης. Το κόστος μιας πρόβλεψης σε καμία περίπτωση δεν μπορεί να θεωρηθεί αμελητέο, αφού εμπεριέχει δαπάνες σε πόρους, όπως οι ανθρωποώρες και πιθανά κάποιο εξειδικευμένο λογισμικό που απαιτείται για τη διενέργεια των προβλέψεων. Για το λόγο αυτό οι επιχειρήσεις προτού προχωρήσουν στην υιοθέτηση ενός συστήματος πρόβλεψης, πρέπει να διεξάγουν μια ανάλυση τύπου κόστους οφέλους. Στόχος της ανάλυσης είναι να μελετηθεί κατά πόσο το κόστος υιοθέτησης μιας περίπλοκης και κοστοβόρας μεθόδου πρόβλεψης, αντισταθμίζεται με το όφελος που θα προκύψει από τις, πιθανά, πιο ακριβείς προβλέψεις που θα παράγει.



Στο ακόλουθο διάγραμμα, απεικονίζονται συνοπτικά τα βήματα που εφαρμόζονται κατά τη διαδικασία εφαρμογής ενός μοντέλου πρόβλεψης.



Πηγή : John E. Hanke and Dean W. Wichern, *Business Forecasting* (9th ed.)



2. Ανάλυση Χρονοσειρών και Μέθοδοι Εξομάλυνσης

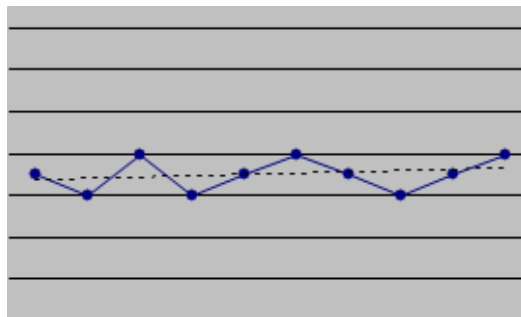
2.1 Στοιχεία ανάλυσης χρονοσειρών

Η ανάλυση χρονοσειρών είναι μια από τις πλέον βασικές λειτουργίες της διοίκησης των επιχειρήσεων, ενώ οι προβλέψεις που στηρίζονται στην ανάλυση χρονοσειρών αποτελούν απαραίτητη πηγή πληροφόρησης, η οποία υποστηρίζει τη λήψη επιχειρηματικών αποφάσεων. Κάθε επιχείρηση ή οργανισμός λαμβάνει αποφάσεις με βάση προβλέψεις οικονομικών μεγεθών που προέρχονται από την επεξεργασία των διαθέσιμων δεδομένων και την εφαρμογή κατάλληλων μεθόδων πρόβλεψης – προεκβολής των ιστορικών δεδομένων στο μέλλον.

Το σύνολο των δεδομένων, τα οποία συλλέγονται διαχρονικά και εκφράζουν την εξέλιξη των τιμών μιας μεταβλητής κατά τη διάρκεια ίσων διαδοχικών χρονικών περιόδων, ονομάζεται χρονοσειρά (ή χρονολογική σειρά, time series). Ειδικότερα, η χρονοσειρά αποτελείται από ένα σύνολο παρατηρήσεων μιας μεταβλητής, οι τιμές της οποίας είναι ιεραρχημένες με βάση τη χρονική περίοδο στην οποία αναφέρονται, π.χ. έτος, τρίμηνο, μήνας κ.α. Παραδείγματα χρονοσειρών είναι οι μηνιαίες πωλήσεις ενός προϊόντος μιας επιχείρησης κατά την τελευταία τετραετία, το ετήσιο ακαθάριστο εθνικό προϊόν μιας χώρας κ.α.

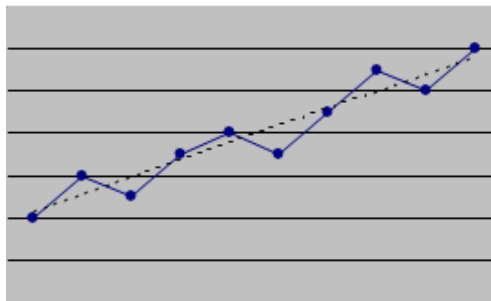
Με μαθηματικούς όρους, μία χρονοσειρά ορίζεται από τις τιμές Y_1, Y_2, \dots, Y_n κάποιας μεταβλητής Y , η οποία για παράδειγμα μπορεί να αντιπροσωπεύει τις πωλήσεις ενός προϊόντος, κατά τις χρονικές στιγμές t_1, t_2, \dots, t_n . Επομένως το Y είναι μια συνάρτηση του t , και συμβολίζεται ως $Y = F(t)$. Με λίγα λόγια η γραφική παράσταση της συνάρτησης $Y = F(t)$ παρουσιάζει την εξέλιξη της προς μελέτη μεταβλητής Y στο χρόνο t . Κάθε χρονοσειρά, αποτελείται από τα εξής χαρακτηριστικά στοιχεία :

Επίπεδο ή οριζόντιο στοιχείο (Level) : η χρονοσειρά εμφανίζει μια μέση τιμή και μια διακύμανση γύρω από αυτή.

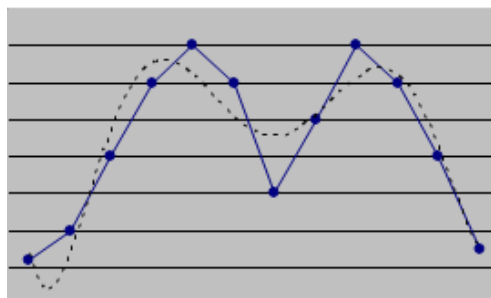




Τάση (Trend) : οι μακροπρόθεσμες κινήσεις ή τάση, αναφέρονται στη γενική κατεύθυνση που φαίνεται ότι ακολουθεί η γραφική απεικόνιση των παρατηρούμενων τιμών της χρονοσειράς, σε μία μεγάλη διάρκεια χρόνου. Σε πολλές περιπτώσεις οι τιμές των παρατηρήσεων ορισμένων χρονοσειρών τείνουν να αυξάνονται ή να μειώνονται με αρκετά σταθερό ρυθμό για μεγάλα χρονικά διαστήματα. Η συμπεριφορά αυτή εκφράζεται από την τάση που φανερώνει τη μακροχρόνια εξέλιξη της χρονοσειράς, η οποία μπορεί να είναι ανοδική ή καθοδική. Η τάση οφείλεται συνήθως σε πληθυσμιακές ή τεχνολογικές αλλαγές και σε οικονομικούς παράγοντες, όπως π.χ. στον πληθωρισμό, στην αύξηση της παραγωγικότητας κ.α.

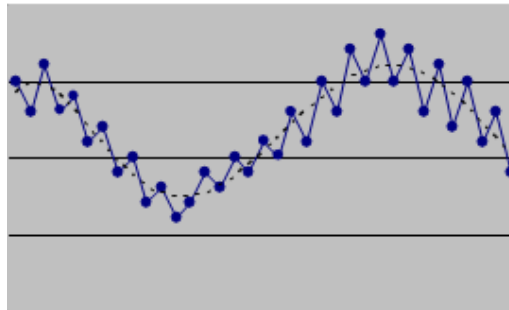


Εποχικότητα (seasonality) : Οι εποχικές κινήσεις ή εποχικότητα αναφέρονται στην ταυτόσημη ή σχεδόν ταυτόσημη εξέλιξη που έχει μία χρονοσειρά κατά τη διάρκεια κάποιων συγκεκριμένων μηνών ή τριμήνων διαδοχικών ετών. Η εποχικότητα οφείλεται σε επαναλαμβανόμενα γεγονότα. Τα δεδομένα ορισμένων χρονοσειρών αναφέρονται σε χρονικές περιόδους μικρότερες του έτους, όπως μήνες ή τρίμηνα, με αποτέλεσμα να παρατηρούνται εποχικές διακυμάνσεις, οι οποίες εμφανίζονται κατά τη διάρκεια του έτους και επαναλαμβάνονται με την ίδια ή περίπου την ίδια μορφή από έτος σε έτος. Γενικά, το φαινόμενο της εποχικότητας οφείλεται κυρίως σε μεταβολές του καιρού, σε πολιτικές της διοίκησης αναφορικά με περιόδους εκπτώσεων, καθώς και σε άλλους παράγοντες όπως θρησκευτικούς, κοινωνικούς κ.α. Οι εποχικές διακυμάνσεις, επειδή παρουσιάζονται με συστηματικό τρόπο, συνήθως, μπορούν εύκολα να αναλυθούν και να προσδιοριστούν και κατά συνέπεια να χρησιμοποιηθούν για την πρόβλεψη των μελλοντικών τιμών της χρονοσειράς.





Κύκλος (Cycle) : οι κυκλικές κινήσεις ή κυκλικότητα αναφέρονται σε μακροπρόθεσμες ταλαντώσεις γύρω από τη γραμμική ή καμπύλη τάσης. Η κυκλικότητα εμφανίζεται ακανόνιστα με κυματοειδή μορφή και διαρκεί για χρονικό διάστημα πολύ μεγαλύτερο του έτους. Η συμπεριφορά αυτή των τιμών των χρονοσειρών αποδίδεται κυρίως στους οικονομικούς κύκλους, οι οποίοι οφείλονται σε μεταβαλλόμενες οικονομικές, τεχνολογικές και άλλες συνθήκες. Επειδή όμως οι οικονομικοί κύκλοι δεν εμφανίζονται με την ίδια περιοδικότητα ή και την ίδια μορφή, το στοιχείο της κυκλικότητας σε αντίθεση με την τάση και την εποχικότητα, δεν θεωρείται ότι συμβάλλει άμεσα στη δημιουργία προβλέψεων. Ωστόσο, η κυκλικότητα μπορεί να χρησιμοποιηθεί για να προσδιοριστεί η μέχρι τώρα εξέλιξη των τιμών της χρονοσειράς.



Τυχαιότητα (Irregular – Random) : Οι ακανόνιστες ή τυχαίες κινήσεις αναφέρονται στις σποραδικές, ακανόνιστες κινήσεις μιας χρονοσειράς λόγω τυχαίων παραγόντων και γεγονότων. Οι τυχαίες κινήσεις επηρεάζουν τις τιμές των χρονοσειρών κατά ένα τυχαίο και μη συστηματικό τρόπο, ο οποίος δεν μπορεί να προσδιοριστεί. Η συνιστώσα αυτή λοιπόν δεν είναι δυνατόν να χρησιμοποιηθεί στη διαμόρφωση των μελλοντικών τιμών των χρονοσειρών. Οι τυχαίες κινήσεις οφείλονται σε όλους εκείνους τους τυχαίους και απρόσμενους παράγοντες που επηρεάζουν τις τιμές των χρονοσειρών και οι οποίοι δεν προσδιορίζονται από την τάση, την εποχικότητα και την κυκλικότητα. Οι παράγοντες αυτοί μπορεί να είναι πόλεμοι, σεισμοί, απρόσμενες καιρικές μεταβολές, απεργίες, διαδόσεις για συγκεκριμένο προϊόν, αιφνίδιες μεταβολές στις προτιμήσεις των καταναλωτών, απρόσμενες αλλαγές στη νομοθεσία κ.α.

Τα χαρακτηριστικά αυτά των χρονοσειρών, ο αναλυτής καλείται να τα εντοπίσει και να τα μεταφράσει από απλές αριθμητικές τιμές, σε πληροφορίες χρήσιμου επιχειρηματικού περιεχομένου για τη διεξαγωγή προβλέψεων. Η διαδικασία της αποσύνθεσης μιας χρονοσειράς είναι κρίσιμη για την επιτυχή επιλογή και εφαρμογή της κατάλληλης, κάθε φορά, μεθόδου πρόβλεψης.

Για την ανάλυση των χρονοσειρών χρησιμοποιούμε τους ακόλουθους συμβολισμούς:



- Y_t = Πραγματική τιμή της χρονοσειράς
- T_t = Τάση
- S_t = Εποχικότητα
- C_t = Κυκλικότητα
- I_t = Τυχαίες κινήσεις όπου $t = 1, 2, \dots, n$

Η εξέταση των στοιχείων αυτών γίνεται σύμφωνα με κάποιο μαθηματικό μοντέλο που φανερώνει τον τρόπο με τον οποίο οι παρατηρήσεις της χρονοσειράς προσδιορίζονται από τις συνιστώσες της. Τα χρησιμοποιούμενα μοντέλα στην ανάλυση των χρονοσειρών είναι, το προσθετικό (additive model) και το πολλαπλασιαστικό (multiplicative model).

Στο προσθετικό μοντέλο οι πραγματικές τιμές της χρονοσειράς για κάθε περίοδο θεωρούνται ως το άθροισμα των τεσσάρων προαναφερθέντων συνιστωσών της και δημιουργούνται σύμφωνα με την ακόλουθη σχέση :

$$Y_t = T_t + S_t + C_t + I_t$$

Αντίθετα στο πολλαπλασιαστικό μοντέλο οι πραγματικές τιμές της χρονοσειράς προσδιορίζονται από το γινόμενο των τεσσάρων συνιστωσών της, δηλαδή ως ακολούθως:

$$Y_t = T_t \cdot S_t \cdot C_t \cdot I_t$$

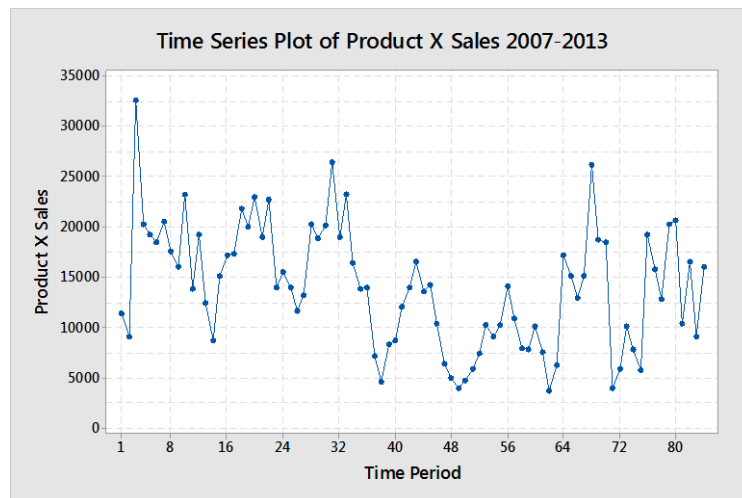
Από τα δύο παραπάνω μοντέλα το προσθετικό μοντέλο χρησιμοποιείται λιγότερο συχνά στην πράξη, επειδή είναι δύσκολο στην ανάλυση του, για υπολογιστικούς κυρίως λόγους. Επίσης βασίζεται στην υπόθεση ότι οι συνιστώσες της χρονοσειράς είναι ανεξάρτητες μεταξύ τους, που σημαίνει για παράδειγμα, ότι η τάση δεν επηρεάζει την εποχικότητα στον υπολογισμό των τιμών της χρονοσειράς. Η παραδοχή αυτή μπορεί να είναι σωστή κυρίως για φυσικά φαινόμενα, αλλά σπάνια ισχύει σε επιχειρησιακές και οικονομικές εφαρμογές, στις οποίες συνήθως η τάση, σε μεγάλο βαθμό, επηρεάζει μεταξύ άλλων και τις εποχικές διακυμάνσεις.

Σε μία συγκεκριμένη χρονοσειρά είναι δυνατόν να μην συνυπάρχουν και οι τέσσερις συνιστώσες αλλά μόνο κάποιες από αυτές. Η ανάλυση, συνίσταται ακριβώς στον εντοπισμό και την περιγραφή (εν γένει



με μαθηματικό τρόπο) των συνιστωσών εκείνων που ενυπάρχουν και επηρεάζουν τη διαχρονική εξέλιξη της προς μελέτη χρονοσειράς.

Στο ακόλουθο διάγραμμα παρατίθεται η χρονοσειρά που απεικονίζει τη ζήτηση ανά μήνα (πωλήσεις) ενός προϊόντος, κατά το χρονικό διάστημα 2007-2013.



Σχήμα 2.1 : Χρονοσειρά Πωλήσεων

Για τη χρονοσειρά πωλήσεων του παραπάνω σχήματος μπορούμε να παρατηρήσουμε τα εξής σημαντικά. Είναι εμφανής η παρουσία **προτύπου εποχικότητας** χρονικής διάρκειας 12 περιόδων (μηνών), αφού το μοτίβο διακύμανσης επαναλαμβάνεται μετά την πάροδο αυτού του χρόνου. Επίσης, η χρονοσειρά δεν φαίνεται να κυμαίνεται γύρω από μια σταθερή, μέση τιμή, γεγονός που υποδηλώνει ότι δεν είναι **στάσιμη** (nonstationary time series). Τα πρώτα χρόνια (χρονικές περιόδοι 1-36), παρουσιάζεται ένα **πρότυπο τάσης** μείωσης των πωλήσεων, το οποίο αντιστρέφεται τα τελευταία δυο χρόνια που οι πωλήσεις παρουσιάζουν αυξητική τάση.

2.2 Αποσύνθεση Χρονοσειράς

Οι πλείστες μέθοδοι πρόβλεψης μέσω ανάλυσης χρονοσειρών, απαιτούν την ύπαρξη καταγεγραμμένων δεδομένων για τις τιμές της προς εξέτασης μεταβλητής τα προηγούμενα χρόνια. Τα δεδομένα αυτά ωστόσο, σπάνια αποτελούν την προς μελέτη χρονοσειρά χωρίς να υποστούν κάποια επεξεργασία, η οποία θα οδηγήσει σε μια πιο εξομαλυμένη χρονοσειρά κατάλληλη για πρόβλεψη. Η επεξεργασία αυτή μπορεί να αφορά διαδικασίες όπως η εξαίρεση ορισμένων δεδομένων (outliers), η διαχείριση μηδενικών τιμών, η άθροιση των δεδομένων στην επιθυμητή συχνότητα, η εύρεση και διαχείριση ασυνήθιστων τιμών (irregular data), η από-εποχικοποίηση της χρονοσειράς κ.α. Οι παραπάνω διαδικασίες



πραγματοποιούνται με αυτοματοποιημένες τεχνικές και μεθοδολογίες που υπάρχουν διαθέσιμες στη διεθνή βιβλιογραφία. Στο κομμάτι της αποσύνθεσης μιας χρονοσειράς, σημαντικό ρόλο διαδραματίζει η σχετική πείρα και η κριτική σκέψη του αναλυτή, παράλληλα με τα επιπρόσθετα ποιοτικά δεδομένα που ενδέχεται να υπάρχουν για τις διακυμάνσεις της χρονοσειράς στο χρονικό ορίζοντα που μελετάται.

Γενικά, η αποσύνθεση χρονοσειρών (time series decomposition) στηρίζεται στην υπόθεση ότι οι τιμές μιας χρονοσειράς σχηματίζονται από τις προαναφερθείσες βασικές συνιστώσες: επίπεδο, τάση, εποχικότητα, κύκλος και τυχαιότητα. Για τη δημιουργία των προβλέψεων μέσω ανάλυσης χρονοσειρών, η χρονοσειρά διασπάται στις ανωτέρω συνιστώσες και προσδιορίζεται μέσω κατάλληλων συντελεστών, η επιρροή που θα ασκεί η κάθε συνιστώσα στη διαμόρφωση των τιμών της μεταβλητής στις μελλοντικές χρονικές περιόδους.

Η ικανότητα του αναλυτή να εντοπίσει τα στοιχεία εκείνα των συνιστωσών που επηρεάζουν τη χρονική εξέλιξη της χρονοσειράς και να τα μοντελοποιήσει με μαθηματικό τρόπο ώστε να επιτευχθεί η αποτελεσματική προεκβολή τους στο μέλλον, είναι ζωτικής σημασίας για μια επιτυχή πρόβλεψη.

2.3 Ποσοτικές μέθοδοι με εξομάλυνση

Με τις μεθόδους εξομάλυνσης προσδιορίζουμε τις μελλοντικές τιμές μιας μεταβλητής στηριζόμενοι αποκλειστικά στα διαθέσιμα ιστορικά στοιχεία, ανεξαρτήτως της σχέσης που μπορεί να έχει η μεταβλητή αυτή με άλλη ή άλλες μεταβλητές. Κύριο χαρακτηριστικό της εφαρμογής αυτών των μεθόδων είναι ότι μπορούμε χωρίς μεγάλο υπολογιστικό βαθμό δυσκολίας, να διαμορφώσουμε εύκολα και σχετικά γρήγορα προβλέψεις για μια μεταβλητή, που προέρχονται από τη διαχρονική μελέτη του τρόπου δημιουργίας των τιμών της συγκεκριμένης μεταβλητής.

Οι μέθοδοι εξομάλυνσης εφαρμόζονται για μεταβλητές, τα δεδομένα των οποίων προέρχονται αποκλειστικά από χρονοσειρές, δηλαδή από παρατηρήσεις που έχουν καταγραφεί κατά τη διάρκεια ίσων διαδοχικών χρονικών περιόδων. Ο αριθμός των διαθέσιμων παρατηρήσεων της μεταβλητής δεν αποτελεί κατά κανόνα ανασταλτικό παράγοντα για την εφαρμογή τους. Ορισμένες από αυτές τις μεθόδους μπορούν να χρησιμοποιηθούν ακόμα και για δείγματα παρατηρήσεων μικρού μεγέθους.

Στόχος των χρησιμοποιούμενων μεθόδων εξομάλυνσης είναι η αναγνώριση του τρόπου δημιουργίας των παρατηρήσεων της χρονοσειράς (pattern), δηλαδή πιθανά, ενός επαναλαμβανόμενου μοτίβου διακύμανσης της ζήτησης, ώστε να αποκτηθούν οι καλύτερες δυνατές προβλέψεις μέσω της



προεκβολής τους στο μέλλον. Η σύγκριση εναλλακτικών μεθόδων πρόβλεψης επιτυγχάνεται με τη βοήθεια των κριτηρίων αξιολόγησης, υπολογίζοντας και αξιολογώντας τα σφάλματα πρόβλεψης.

Ορολογία και συμβολισμοί

Για την ανάλυση των μεθόδων πρόβλεψης ζήτησης, θα χρησιμοποιηθεί η σχετική ορολογία και συμβολισμοί σύμφωνα με τη διεθνή βιβλιογραφία :

L_t : εκτίμηση του επιπέδου κατά την περίοδο t

T_t : εκτίμηση της τάσης κατά την περίοδο t

S_t : εκτίμηση του παράγοντα εποχικότητας κατά την περίοδο t

F_t : πρόβλεψη ζήτησης για την περίοδο t

D_t : πραγματική ζήτηση κατά την περίοδο t

e_t : σφάλμα πρόβλεψης κατά την περίοδο t

2.3.1 Η μέθοδος του κινητού μέσου όρου (Moving Average)

Η μέθοδος του κινητού μέσου όρου είναι μια πολύ απλή μέθοδος προβλέψεων η οποία χρησιμοποιεί ως πρόβλεψη τον αριθμητικό μέσο όρο των N πιο πρόσφατων παρατηρήσεων της χρονοσειράς που εξετάζουμε. Ο λόγος για τον οποίο χρησιμοποιούνται οι πιο πρόσφατες παρατηρήσεις είναι διότι σε σχέση με τις πιο απομακρυσμένες, χρονικά, παρατηρήσεις μιας δεδομένης χρονοσειράς, θεωρούνται περισσότερο αντιπροσωπευτικές για τη δημιουργία προβλέψεων. Ο μέσος αυτός αποκαλείται «κινούμενος» γιατί κάθε φορά που γίνεται διαθέσιμη μια νέα παρατήρηση, μπορεί να υπολογιστεί και να χρησιμοποιηθεί ως πρόβλεψη ένας νέος αριθμητικός μέσος.

Η μέθοδος του κινητού μέσου όρου, μας επιτρέπει να εξάγουμε προβλέψεις, χρησιμοποιώντας ιστορικά στοιχεία του παρελθόντος, μόνο για την αμέσως επόμενη χρονική περίοδο. Η προβλεπόμενη τιμή F_{t+1} της μεταβλητής για την περίοδο $t+1$ ισούται με τη μέση τιμή των τιμών $D_t, D_{t-1}, \dots, D_{t-N+1}$ της μεταβλητής κατά τις N αμέσως προηγούμενες περιόδους, είναι δηλαδή ίση με το επίπεδο (οριζόντιο στοιχείο) των προηγούμενων t περιόδων και δίνεται από την εξίσωση:

$$F_{t+1} = \frac{D_t + D_{t-1} + \dots + D_{t-N+1}}{N} = \frac{1}{N} \sum_{i=t-N+1}^t D_i = L_t$$

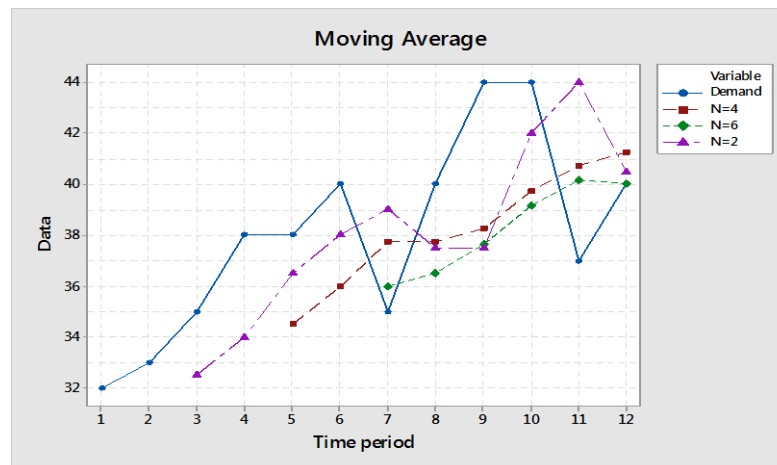


Η μέθοδος του κινούμενου μέσου όρου είναι κατάλληλη για πρόβλεψη μεταβλητών που δεν εμφανίζουν τάση ή εποχικότητα, άρα είναι στάσιμες. Για την εφαρμογή της μεθόδου επιλέγεται αρχικά ο αριθμός των περιόδων N , για τις οποίες θα υπολογιστεί ο κινούμενος μέσος όρος. Όσο μεγαλύτερο είναι το N τόσο μεγαλύτερη είναι η εξομάλυνση της τιμής της μεταβλητής. Συνήθως, για να προσδιοριστεί η τιμή του N για μια συγκεκριμένη χρονοσειρά υπολογίζεται ο κινούμενος μέσος για διάφορες τιμές του N και επιλέγεται εκείνη η τιμή που ελαχιστοποιεί την τιμή ενός κριτηρίου αξιολόγησης των μεθόδων πρόβλεψης (κριτήρια MSE ή MAPE, βλ. **Κεφάλαιο 3**).

Στον ακόλουθο πίνακα, δίνονται οι πωλήσεις ενός προϊόντος κατά τους τελευταίους 12 μήνες. Για κινητό μέσο όρο $N = 2, 4$ και 6 περιόδων αντίστοιχα, παρατηρούμε τις εκτιμημένες τιμές για την πραγματική ζήτηση και τα προκύπτοντα διαγράμματα.

Κινητός Μέσος Όρος (Moving Average)				
Περίοδος $\rightarrow (t)$	Ζήτηση $\rightarrow (D_t)$	$N = 2 \rightarrow (F_t)$	$N = 4 \rightarrow (F_t)$	$N = 6 \rightarrow (F_t)$
1	32			
2	33			
3	35	32.5		
4	38	34		
5	38	36.5	34.5	
6	40	38	36	
7	35	39	37.75	36
8	40	37.5	37.75	36.5
9	44	37.5	38.25	37.7
10	44	42	39.75	39.2
11	37	44	40.75	40.2
12	40	40.5	41.25	40

Πίνακας 2.1 : Κινητός Μέσος Όρος



Σχήμα 2.1 : Μέθοδος Κινητού Μέσου Όρου



Όπως αναμενόταν, παρατηρούμε ότι όσο μεγαλύτερο είναι το N , τόσο μεγαλύτερη είναι η εξομάλυνση της τιμής της προς εξέταση μεταβλητής, τόσο δηλαδή είναι μικρότερη η επίδραση των τελευταίων χρονικά παρατηρούμενων τιμών.

2.3.2 Απλή εκθετική εξομάλυνση (Simple exponential smoothing)

Η μέθοδος της απλής εκθετικής εξομάλυνσης αποτελεί μια εξέλιξη της μεθόδου του κινούμενου μέσου. Το κυριότερο μειονέκτημα της μεθόδου του κινούμενου μέσου N -περιόδων είναι ότι για την εξαγωγή προβλέψεων δίνεται ίση βαρύτητα σε κάθε παρατήρηση, ανεξάρτητα από το πόσο κοντά ή μακριά βρίσκεται σε σχέση με την προς πρόβλεψη περίοδο. Το μειονέκτημα αυτό εξαλείφεται με τη μέθοδο της απλής εκθετικής εξομάλυνσης, σύμφωνα με την οποία οι προβλέψεις δημιουργούνται με βάση καθορισμένο σταθμικό μέσο όρο, έτσι ώστε να δίνεται διαφορετική βαρύτητα σε κάθε παρατήρηση. Πιο συγκεκριμένα, με τη μέθοδο αυτή ανάλογα με την τιμή που λαμβάνει ο συντελεστής εξομάλυνσης, δίνεται μεγαλύτερη ή μικρότερη βαρύτητα στις πιο πρόσφατες παρατηρήσεις, σε σχέση με τις χρονικά, πιο απομακρυσμένες.

Το μοντέλο χρησιμοποιεί την πρόβλεψη καθώς και την αντίστοιχη πραγματική τιμή της μεταβλητής για την τρέχουσα περίοδο για να προβλέψει την τιμή της μεταβλητής κατά τις επόμενες περιόδους.

Η πρόβλεψη της τιμής της μεταβλητής για την επόμενη χρονική περίοδο δίνεται από τη σχέση :

$$F_{t+1} = \alpha D_t + (1 - \alpha)F_t \quad \text{όπου } 0 < \alpha < 1$$

Η επιλογή του συντελεστή εξομάλυνσης καθορίζεται κατά κύριο λόγο από δύο αλληλοεξαρτώμενους παράγοντες. Πρώτος παράγοντας είναι το ποσοστό τυχαιότητας που εμπερικλείεται στις τιμές της χρονοσειράς. Όσο πιο πολλές ακανόνιστες, τυχαίες και μη επαναλαμβανόμενες διακυμάνσεις παρατηρούνται στη χρονοσειρά, τόσο πιο μικρή πρέπει να είναι η τιμή του συντελεστή εξομάλυνσης ώστε να μειωθεί η επίδραση των τιμών αυτών στο μοντέλο πρόβλεψης που θα χρησιμοποιηθεί. Ο δεύτερος παράγοντας είναι η σταθερότητα του μέσου όρου της χρονοσειράς στη χρονική διάρκεια του ορίζοντα που εκτείνονται τα δεδομένα (π.χ. σταθερότητα μέσου όρου ανά έτος). Κι αυτό διότι όσο μεγαλύτερη είναι η μεταβολή του μέσου όρου της χρονοσειράς, τόσο μεγαλύτερος πρέπει να είναι ο συντελεστής εξομάλυνσης ώστε να ακολουθεί το επίπεδο, και αντιστρόφως. Γενικά, ισχύει ότι υψηλός συντελεστής εξομάλυνσης, προκαλεί ταχύτερη μείωση της επίδρασης των παλαιότερων, χρονικά,

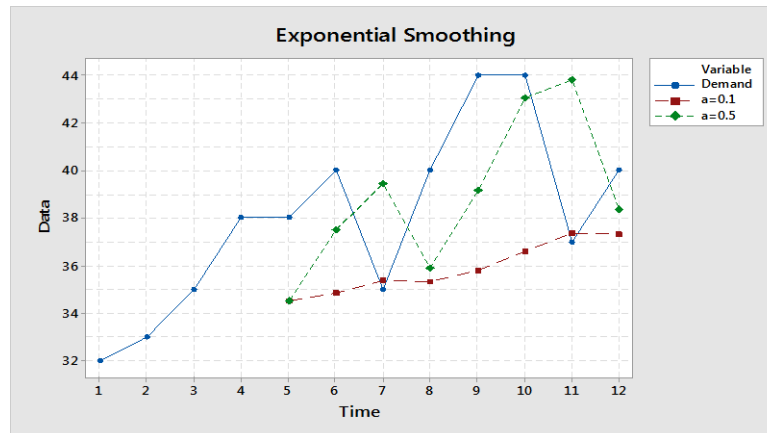


παρατηρήσεων και ταχύτερη προσαρμογή στις αλλαγές της ζήτησης, ενώ όσο πιο μικρές είναι οι τιμές που λαμβάνει ο συντελεστής τόσο περισσότερο εξομαλύνονται οι παρατηρήσεις της χρονοσειράς. Για τιμές συντελεστή εξομάλυνσης κοντά στο μηδέν το επίπεδο μεταβάλλεται ελάχιστα. Για $\alpha = 0$ κάθε πρόβλεψη είναι ίση με το αρχικό επίπεδο. Στην αντίθετη περίπτωση, όσο η τιμή που λαμβάνει ο συντελεστής είναι πιο κοντά στη μονάδα, τόσο πιο ενεργό ρόλο παίζει το σφάλμα στην πρόβλεψη. Για $\alpha = 1$ το μοντέλο ταυτίζεται με την απλοϊκή μέθοδο Naïve. Απαιτείται έλεγχος (προσομοίωση) με ιστορικά στοιχεία για τον προσδιορισμό της τιμής του συντελεστή εξομάλυνσης α , έχοντας ως στόχο την ικανοποίηση ενός κριτηρίου αξιολόγησης των σφαλμάτων.

Στον ακόλουθο πίνακα, δίνονται οι πωλήσεις ενός προϊόντος κατά τους τελευταίους 12 μήνες. Για συντελεστή εξομάλυνσης $\alpha = 0.1$ και $\alpha = 0.5$ αντίστοιχα, παρατηρούμε τις εκτιμημένες τιμές για την πραγματική ζήτηση και τα προκύπτοντα διαγράμματα. Για την αρχικοποίηση* της μεθόδου χρησιμοποιήθηκε η μέση τιμή των τεσσάρων πρώτων παρατηρήσεων της χρονοσειράς.

Απλή Εκθετική Εξομάλυνση (Single Exponential Smoothing)			
Περίοδος $\rightarrow (t)$	Ζήτηση $\rightarrow (D_t)$	$\alpha = 0.1 \rightarrow (F_t)$	$\alpha = 0.5 \rightarrow (F_t)$
1	32		
2	33		
3	35		
4	38		
5	38	34.5 *	34.5 *
6	40	34.85	37.5
7	35	35.37	39.46
8	40	35.33	35.89
9	44	35.8	39.18
10	44	36.62	43.04
11	37	37.35	43.81
12	40	37.32	38.36

Πίνακας 2.2 : Απλή Εκθετική Εξομάλυνση



Σχήμα 2.2 : Μέθοδος Απλής Εκθετικής Εξομάλυνσης



Παρατηρούμε ότι για μεγαλύτερες τιμές του συντελεστή α , η εξομάλυνση της χρονοσειράς είναι εντονότερη. Η επίδραση των παλαιότερων, χρονικά, παρατηρήσεων της χρονοσειράς στις πρόσφατες εκτιμημένες τιμές μειώνεται, με αποτέλεσμα να προσαρμόζεται ταχύτερα στις μεταβολές που παρουσιάζει η διακύμανση της χρονοσειράς.

Εκτίμηση των παραμέτρων τάσης, επιπέδου και εποχικότητας

Στη διεθνή βιβλιογραφία, υπάρχουν διάφοροι τρόποι υπολογισμού των παραμέτρων επιπέδου, τάσης και εποχικότητας για την αρχικοποίηση των μεθόδων πρόβλεψης. Εν προκειμένω, στην παρούσα εργασία, θα χρησιμοποιηθεί η μεθοδολογία που προτείνεται από τους Chopra S. and Meindl P. στο σύγγραμμά τους «Supply chain management : Strategy, Planning and Operations»

Σύμφωνα με τη μέθοδο αυτή το πρώτο βήμα για την εκτίμηση των παραμέτρων μιας χρονοσειράς η οποία παρουσιάζει πρότυπα τάσης και εποχικότητας, αφορά την από-εποχικοποίηση των δεδομένων (deseasonalized data). Η χρονική διάρκεια της περιοδικότητας (p), είναι ο αριθμός των χρονικών περιόδων μετά την πάροδο των οποίων επαναλαμβάνεται ο εποχικός κύκλος. Για παράδειγμα, εάν παρουσιάζεται εποχικό πρότυπο σε διαθέσιμα μηνιαία ιστορικά στοιχεία, η χρονική διάρκεια της περιοδικότητας είναι $p = 12$, ενώ για τριμηνιαία στοιχεία προκύπτει ότι $p = 4$ κ.ο.κ. Η χρονική διάρκεια της περιοδικότητας, γίνεται πιο εύκολα κατανοητή μέσα από την παρατήρηση της διακύμανσης των τιμών της χρονοσειράς στο χρονικό ορίζοντα που μελετάται.

Ο υπολογισμός της από-εποχικοποιημένης ζήτησης πραγματοποιείται μέσω των παρακάτω εξισώσεων :

$$\bar{D}_t = \frac{\left[D_{t-\left(\frac{p}{2}\right)} + D_{t+\left(\frac{p}{2}\right)} + \sum_{i=t+1-\left(\frac{p}{2}\right)}^{t-1+\left(\frac{p}{2}\right)} 2D_i \right]}{2p} \quad \text{όταν } p - \text{άρτιος}$$

$$\bar{D}_t = \sum_{i=t-\left(\frac{p}{2}\right)}^{t-1+\left(\frac{p}{2}\right)} \frac{D_i}{2p} \quad \text{όταν } p - \text{περιττός}$$



Στη συνέχεια, πραγματοποιείται γραφική παράσταση της από-εποχικοποιημένης ζήτησης της χρονοσειράς \bar{D}_t σε συνάρτηση με το χρόνο t . Μεταξύ της από-εποχικοποιημένης ζήτησης και του χρόνου, θα εμφανίζεται μια γραμμική σχέση η οποία είναι της μορφής $\bar{D}_t = L + Tt$. Με γραμμική παλινδρόμηση (βλ. **Παράρτημα ΣΤ**) στις τιμές της ζήτησης που προκύπτουν με την αφαίρεση της εποχικότητας, εκτιμώνται οι τιμές των L_0 και T_0 .

Στο σημείο αυτό, έχουμε εκτιμήσει τις τιμές των παραμέτρων του επιπέδου και της τάσης. Για την εκτίμηση των παραμέτρων της εποχικότητας, απαιτείται ο υπολογισμός της νέας ζήτησης για κάθε χρονική περίοδο, μέσω της σχέσης που προέκυψε από τη γραμμική παλινδρόμηση. Οι τιμές της νέας ζήτησης χωρίς εποχικότητα για κάθε χρονική περίοδο υπολογίζονται μέσω της σχέσης :

$$\bar{D}_t' = L + Tt$$

Ο συντελεστής εποχικότητας, για κάθε περίοδο t της προς μελέτη χρονοσειράς, δίνεται από τη σχέση μεταξύ πραγματικής και από-εποχικοποιημένης ζήτησης :

$$\bar{S}_t = \frac{D_t}{\bar{D}_t'}$$

Για χρονικές περιόδους που βρίσκονται σε αντίστοιχα σημεία της περιοδικότητας, ο συντελεστής εποχικότητας θα λαμβάνει σχεδόν την ίδια τιμή. Για παράδειγμα, σε μια χρονοσειρά με χρονική διάρκεια περιοδικότητας ίση με $p = 4$ περιόδους, ο συντελεστής των αντίστοιχων περιόδων 1, 5, 9 θα είναι κατά προσέγγιση ίσος λόγω της επανάληψης του εποχικού κύκλου. Ως αρχικές τιμές των συντελεστών εποχικότητας, λαμβάνονται οι μέσοι όροι των συντελεστών των αντίστοιχων εποχικών χρονικών περιόδων όπως υπολογίστηκαν, μέσω της σχέσης :

$$S_i = \frac{(\sum_{j=0}^{r-1} \bar{S}_{jp+i})}{r} \quad \text{όπου } r = \frac{N}{p}$$

Με τον υπολογισμό και των αρχικών τιμών των παραμέτρων εποχικότητας, είναι διαθέσιμα όλα τα απαιτούμενα δεδομένα για την εφαρμογή μεθόδων πρόβλεψης οι οποίες χρησιμοποιούνται σε χρονοσειρές που εμφανίζουν πρότυπα τάσης και εποχικότητας στη διακύμανση τους



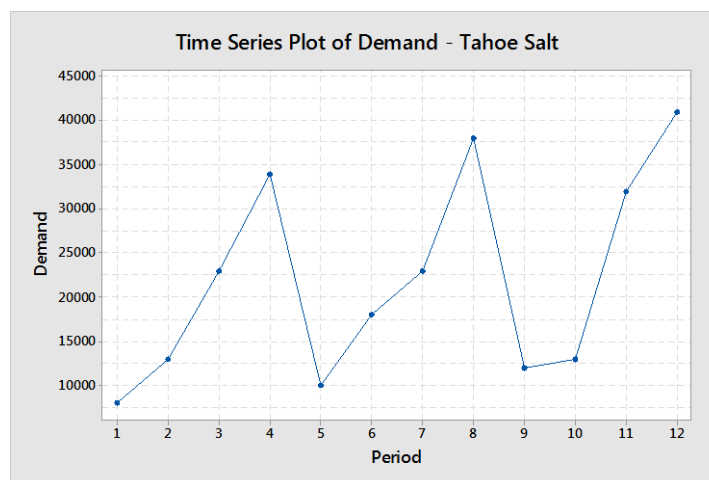
Παράδειγμα

Στον πίνακα που ακολουθεί, παρουσιάζονται οι πωλήσεις ανά τρίμηνο για την Tahoe Salt και η χρονοσειρά που προκύπτει από τα δεδομένα αυτά. Εφαρμόζοντας τη μεθοδολογία που περιγράφηκε προηγουμένως, των Chopra S. and Meindl P, θα υπολογιστούν οι παράμετροι επιπέδου, εποχικότητας και τάσης της χρονοσειράς, για την περαιτέρω εφαρμογή της κατάλληλης μεθόδου πρόβλεψης.

Έτος	Τρίμηνο	Περίοδος (t)	Ζήτηση (D_t)
1	2	1	8000
1	3	2	13000
1	4	3	23000
2	1	4	34000
2	2	5	10000
2	3	6	18000
2	4	7	23000
3	1	8	38000
3	2	9	12000
3	3	10	13000
3	4	11	32000
4	1	12	41000

Πίνακας 2.3 : Δεδομένα Πωλήσεων της Tahoe Salt

(Πηγή : Chopra S. and Meindl P., Supply chain management : Strategy, Planning and Operations, New Jersey, Pearson, Practice Hall, 2nd ed., 2004)



Σχήμα 2.3 : Χρονοσειρά Πωλήσεων της Tahoe Salt

Παρατηρώντας την παραπάνω χρονοσειρά για τις πωλήσεις της Tahoe Salt, συμπεραίνουμε τα εξής :



- Υπάρχει εποχικό πρότυπο στη ζήτηση του προϊόντος.
- Η χρονοσειρά παρουσιάζει πρότυπο τάσης, αφού η ζήτηση αυξάνεται χρόνο με το χρόνο.
- Η χαμηλότερη τιμή της ζήτησης παρατηρείται στο δεύτερο τρίμηνο κάθε έτους, ενώ η υψηλότερη τιμή στο τέταρτο τρίμηνο.
- Υποδηλώνεται περιοδικότητα στη ζήτηση, χρονικής διάρκειας τεσσάρων περιόδων.

Εφόσον λοιπόν η χρονοσειρά για τις πωλήσεις της Tahoe Salt παρουσιάζει στοιχεία τάσης και εποχικότητας, για την εξαγωγή μελλοντικών προβλέψεων, πρέπει να χρησιμοποιηθεί μοντέλο πρόβλεψης το οποίο να λαμβάνει υπόψη τους παράγοντες αυτούς. Για τα διαθέσιμα μοντέλα πρόβλεψης, απαιτείται αρχικοποίηση των εμπλεκόμενων παραμέτρων, δηλαδή αρχική εκτίμηση των παραμέτρων επιπέδου, τάσης και εποχικότητας.

Όπως επισημάνθηκε, στη χρονοσειρά, παρατηρείται περιοδικότητα της ζήτησης, χρονικής διάρκειας τεσσάρων περιόδων. Άρα για $p = 4$, άρτιος θα χρησιμοποιηθεί η παρακάτω εξίσωση, για την από-εποχικοποίηση της χρονοσειράς.

$$\bar{D}_t = \frac{\left[D_{t-\left(\frac{p}{2}\right)} + D_{t+\left(\frac{p}{2}\right)} + \sum_{i=t+1-\left(\frac{p}{2}\right)}^{t-1+\left(\frac{p}{2}\right)} 2D_i \right]}{2p} \quad \text{όταν } p - \text{άρτιος}$$

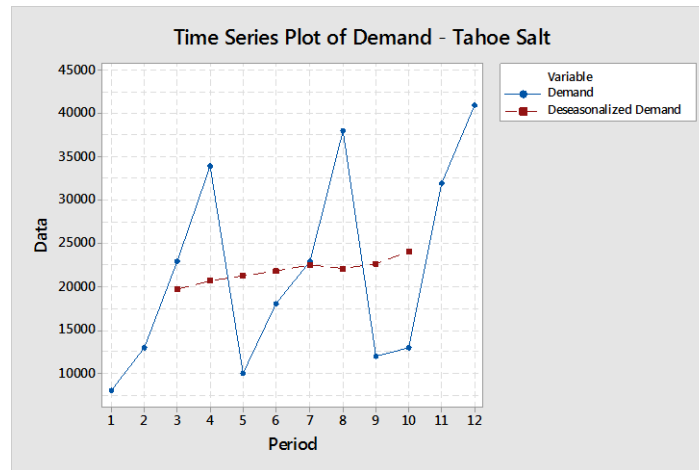
Τα αποτελέσματα από την αφαίρεση της εποχικότητας παρουσιάζονται στον ακόλουθο πίνακα.

Περίοδος (t)	Ζήτηση (D_t)	Ζήτηση χωρίς εποχικότητα (\bar{D}_t)
1	8000	
2	13000	
3	23000	19750
4	34000	20625
5	10000	21250
6	18000	21750
7	23000	22500
8	38000	22125
9	12000	22625
10	13000	24125
11	32000	
12	41000	

Πίνακας 2.4 : Ζήτηση χωρίς εποχικότητα



Με την αφαίρεση της εποχικότητας από τη ζήτηση, παρατηρείται μια γραμμική συσχέτιση μεταξύ της νέας ζήτησης και του χρόνου. Η γραμμική σχέση είναι της μορφής $\bar{D}_t = L + Tt$. Με εφαρμογή γραμμικής παλινδρόμησης στα νέα δεδομένα, προκύπτει μια αρχική εκτίμηση των παραμέτρων του επιπέδου L και της τάσης T .

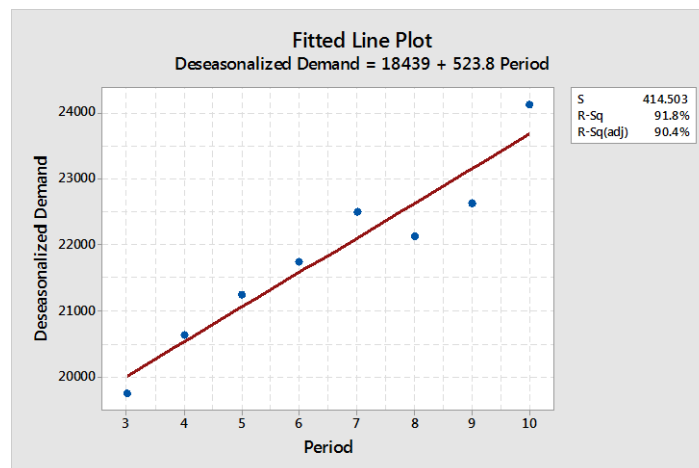


Σχήμα 2.4 : Από-εποχικοποιημένη Ζήτηση

Η εξίσωση που προκύπτει από την εφαρμογή γραμμικής παλινδρόμησης στην από-εποχικοποιημένη ζήτηση, είναι $\bar{D}_t' = 18439 + 523.8t$ γεγονός που σημαίνει ότι η αρχική εκτίμηση της παραμέτρου του επιπέδου είναι $L = 18439$ και της τάσης $T = 523.8$.

Οι τιμές της νέας ζήτησης χωρίς εποχικότητα για κάθε χρονική περίοδο υπολογίζονται από τη σχέση :

$$\bar{D}_t' = L + Tt$$



Σχήμα 2.5 : Γραμμική παλινδρόμηση



Έχοντας εκτιμήσει τις αρχικές τιμές των παραμέτρων επιπέδου και τάσης, μένει να λάβουμε αρχική εκτίμηση και για τις παραμέτρους εποχικότητας. Οι συντελεστές εποχικότητας, για κάθε περίοδο t της προς μελέτη χρονοσειράς, υπολογίζονται με δεδομένες τις τιμές πραγματικής και από-εποχικοποιημένης ζήτησης ως ακολούθως:

Περίοδος (t)	Ζήτηση (D_t)	Νέα Ζήτηση χωρίς εποχικότητα (D_t')	Συντ. εποχικότητας ($\bar{S}_t = \frac{D_t}{D_t'}$)
1	8000	18963	0.42
2	13000	19487	0.67
3	23000	20011	1.15
4	34000	20535	1.66
5	10000	21059	0.47
6	18000	21583	0.83
7	23000	22107	1.04
8	38000	22631	1.68
9	12000	23155	0.52
10	13000	23679	0.55
11	32000	24203	1.32
12	41000	24727	1.66

Πίνακας 2.5 : Υπολογισμός συντελεστών εποχικότητας

Στη χρονοσειρά, εφόσον η χρονική διάρκεια περιοδικότητας είναι ίση με $p = 4$ περιόδους, ο συντελεστής των αντίστοιχων περιόδων 1, 5, 9 θα είναι κατά προσέγγιση ίσος λόγω της επανάληψης του εποχικού κύκλου. Ως αρχικές τιμές των συντελεστών εποχικότητας, λαμβάνονται οι μέσοι όροι των συντελεστών των αντίστοιχων εποχικών χρονικών περιόδων, μέσω της σχέσης (όπου $N=12$) :

$$S_i = \frac{(\sum_{j=0}^{r-1} \bar{S}_{jp+i})}{r} \quad \text{όπου } r = \frac{N}{p}$$

$$S_1 = \frac{\bar{S}_1 + \bar{S}_5 + \bar{S}_9}{3} = \frac{0.42 + 0.47 + 0.52}{3} = 0.47$$

$$S_2 = \frac{\bar{S}_2 + \bar{S}_6 + \bar{S}_{10}}{3} = \frac{0.67 + 0.83 + 0.55}{3} = 0.68$$

$$S_3 = \frac{\bar{S}_3 + \bar{S}_7 + \bar{S}_{11}}{3} = \frac{1.15 + 1.04 + 1.32}{3} = 1.17$$

$$S_4 = \frac{\bar{S}_4 + \bar{S}_8 + \bar{S}_{12}}{3} = \frac{1.66 + 1.68 + 1.66}{3} = 1.67$$



Με την αρχική εκτίμηση και των συντελεστών εποχικότητας, είμαστε σε θέση να εφαρμόσουμε τις μεθόδους πρόβλεψης που εμπλέκουν στοιχεία επιπέδου, τάσης και εποχικότητας όπως για παράδειγμα τις μεθόδους Holts και Winters που παρουσιάζονται στις επόμενες παραγράφους.

2.3.3 Εκθετική εξομάλυνση με γραμμική τάση (Μέθοδος Holt)

Αν σε μια χρονοσειρά παρατηρείται ότι εκτός από το οριζόντιο στοιχείο, υπάρχει και γραμμική τάση τότε η μέθοδος της απλής εκθετικής εξομάλυνσης θα υποεκτιμά συστηματικά την προς πρόβλεψη μεταβλητή (π.χ. ζήτηση) όταν η τάση είναι αυξητική, ή αντίθετα θα υπερεκτιμά εν γένει, αν η τάση είναι φθίνουσα. Το μοντέλο της διπλής εκθετικής εξομάλυνσης (ή μοντέλο Holt) είναι κατάλληλο για χρονοσειρές όπου το συστηματικό στοιχείο της ζήτησης χαρακτηρίζεται από οριζόντιο στοιχείο και γραμμική τάση. Επομένως για την πρόβλεψη μιας μελλοντικής χρονικής περιόδου είναι απαραίτητες δύο εκτιμήσεις, μία για το οριζόντιο στοιχείο και μία για την τάση.

Η εκτίμηση του οριζόντιου στοιχείου και της τάσης για κάθε χρονική περίοδο που μελετάται η χρονοσειρά γίνεται μέσω των σχέσεων :

$$L_{t+1} = \alpha D_{t+1} + (1 - \alpha)(L_t + T_t)$$

$$T_{t+1} = \beta(L_{t+1} - L_t) + (1 - \beta)T_t$$

όπου α, β οι συντελεστές εξομάλυνσης για το επίπεδο και την τάση αντίστοιχα. Ισχύει : $0 \leq \alpha, \beta \leq 1$

Η επιλογή των συντελεστών εξομάλυνσης είναι ιδιαίτερος σημαντική. Οι βέλτιστες τιμές των παραμέτρων α και β για μια συγκεκριμένη χρονοσειρά προκύπτουν από την ελαχιστοποίηση της τιμής του μέσου τετραγωνικού σφάλματος MSE ή κάποιου άλλου κριτηρίου, εφαρμόζοντας τη μέθοδο για όλους τους δυνατούς συνδυασμούς των τιμών α και β .

Για να εφαρμοστεί η μέθοδος αυτή, απαιτείται μια αρχική εκτίμηση των τιμών των παραμέτρων του επιπέδου L_0 και της τάσης T_0 . Για την αρχική εκτίμηση των παραμέτρων υπάρχουν διάφοροι τρόποι διαθέσιμοι στη βιβλιογραφία. Ένας από αυτούς, σύμφωνα με τους John E.Hanke και Dean W.Wichern, είναι η τιμή L_0 να λαμβάνει την τιμή της πρώτης παρατήρησης της χρονοσειράς και η τιμή T_0 να ισούται με το μηδέν. Με δεδομένες τις αρχικές εκτιμήσεις του επιπέδου και της τάσης, μπορεί να



εφαρμοστεί η μέθοδος πρόβλεψης, υπολογίζοντας διαδοχικά τις τιμές των παραμέτρων επιπέδου και τάσης για κάθε χρονική περίοδο πρόβλεψης. Για την εξαγωγή προβλέψεων ισχύουν οι σχέσεις :

$$F_{t+1} = L_t + T_t \quad \text{και} \quad F_{t+n} = L_t + nT_t$$

2.3.4 Εκθετική εξομάλυνση με τάση και εποχικότητα (Μέθοδος Winters)

Σε πολλές χρονοσειρές, οι τιμές της προς εξέταση μεταβλητής που αναφέρονται σε χρονικές περιόδους μικρότερες του έτους, όπως για παράδειγμα μήνες ή τρίμηνα, είναι δυνατόν να παρουσιάζουν εποχικές διακυμάνσεις, οι οποίες επαναλαμβάνονται κάθε έτος με την ίδια περίπου μορφή. Η εποχικότητα στις παρατηρήσεις των χρονοσειρών είναι ένα φαινόμενο που εμφανίζεται συχνά κατά τη διερεύνηση οικονομικών φαινομένων όπως η πρόβλεψη ζήτησης. Σε αυτή την περίπτωση, για την εξαγωγή αποτελεσματικών προβλέψεων θα πρέπει να χρησιμοποιούνται μέθοδοι, οι οποίες να λαμβάνουν υπόψη τον παράγοντα της εποχικότητας. Μια τέτοια μέθοδος είναι αυτή της εκθετικής εξομάλυνσης με προσαρμογή στην τάση και στην εποχικότητα (exponential smoothing adjusted for trend and seasonality), γνωστή ως μέθοδος Winters, που αποτελεί επέκταση της μεθόδου Holt.

Η μέθοδος Winters έχει τρεις παραμέτρους, α , β , γ οι οποίες χρησιμοποιούνται για την εξομάλυνση του επιπέδου, της τάσης και της εποχικότητας της χρονοσειράς αντίστοιχα. Αυτό σημαίνει ότι χρησιμοποιείται μια επιπλέον εξίσωση σε σχέση με τη μέθοδο Holt η οποία στοχεύει στην εξομάλυνση της εποχικότητας. Αν η περιοδικότητα που παρατηρείται στις τιμές της χρονοσειράς είναι ίση με p χρονικές περιόδους, για την εφαρμογή της μεθόδου απαιτούνται αρχικές εκτιμήσεις για το οριζόντιο στοιχείο L_0 , για την τάση T_0 και για τους δείκτες εποχικότητας S_1, S_2, \dots, S_p . Η μεθοδολογία εκτίμησης των απαιτούμενων παραμέτρων της μεθόδου, είναι αυτή που περιγράφηκε σε προηγούμενη παράγραφο.

Οι εξισώσεις για την εξομάλυνση των στοιχείων του επιπέδου, της τάσης και της εποχικότητας είναι :

$$L_{t+1} = \alpha \left(\frac{D_{t+1}}{S_{t+1}} \right) + (1 - \alpha)(L_t + T_t)$$

$$T_{t+1} = \beta(L_{t+1} - L_t) + (1 - \beta)T_t$$

$$S_{t+p+1} = \gamma \left(\frac{D_{t+1}}{L_{t+1}} \right) + (1 - \gamma)S_{t+1}$$



Οι βέλτιστες τιμές των παραμέτρων α , β και γ (ισχύει $0 < \alpha, \beta, \gamma < 1$) προκύπτουν από την ελαχιστοποίηση του κριτηρίου MSE ή κάποιου άλλου κριτηρίου, εφαρμόζοντας τη μέθοδο αυτή για όλους τους δυνατούς συνδυασμούς των τιμών των παραμέτρων στα δεδομένα της χρονοσειράς. Συνήθως χρησιμοποιούνται στατιστικά πακέτα τα οποία διαθέτουν ενσωματωμένους επαναληπτικούς αλγορίθμους για την εξεύρεση των βέλτιστων τιμών για τις παραμέτρους του μοντέλου πρόβλεψης. Τελικά, η πρόβλεψη F_{t+1} για την επόμενη χρονική περίοδο προκύπτει συνδυάζοντας τα παραπάνω υπολογισθέντα στοιχεία, σύμφωνα με τη σχέση :

$$F_{t+1} = (L_t + T_t)S_{t+1}$$

2.4 Ανάλυση Χρονοσειρών στις επιχειρήσεις

Οι προβλέψεις, είναι μια από τις σημαντικότερες πηγές δεδομένων εισόδου για τα στελέχη τα οποία καλούνται να λάβουν αποφάσεις σχετικά με τη λειτουργία της επιχείρησης. Στις πλείστες όμως των περιπτώσεων, οι προβλέψεις παραβλέπονται ή δεν πραγματοποιούνται καθόλου και αντικαθίστανται από τις απόψεις των στελεχών οι οποίες στηρίζονται κατά κύριο λόγο στην εμπειρία και τη διαίσθησή τους. Αυτό συμβαίνει είτε λόγω έλλειψης εξειδικευμένων γνώσεων σχετικά με τις τεχνικές προβλέψεων, είτε λόγω μη διαθεσιμότητας των πόρων που απαιτούνται για τη διενέργειά αυτών.

Οι μέθοδοι εξομάλυνσης και κινούμενων μέσων όρων στην ανάλυση χρονοσειρών, είναι μια λύση στην οποία καταλήγουν συχνά πολλές επιχειρήσεις για την εξαγωγή βραχυπρόθεσμων προβλέψεων. Ο λόγος που χρησιμοποιούνται ευρέως αυτές οι μέθοδοι είναι διότι διακρίνονται για τη σχετική τους απλότητα, το χαμηλό κόστος εκτέλεσης και διαχείρισης τους, αλλά και κυρίως, διότι είναι ευκολότερες στην κατανόηση και εφαρμογή τους από τα εμπλεκόμενα στελέχη.

Αυτό όμως που πρέπει να λαμβάνεται ιδιαίτερα υπόψη είναι το γεγονός ότι όλες οι διαθέσιμες ποσοτικές μέθοδοι προβλέψεων, στηρίζονται αποκλειστικά σε ιστορικά δεδομένα, με αποτέλεσμα να μην συνυπολογίζονται πιθανοί παράγοντες που θα μεταβάλουν σημαντικά τη ζήτηση στο μέλλον. Για τη διενέργεια πιο αντιπροσωπευτικών προβλέψεων, απαιτείται επομένως ο συνδυασμός της ποσοτικής προσέγγισης μιας πρόβλεψης μέσω της ανάλυσης χρονοσειρών, μαζί με την ποιοτική προσέγγιση, τις γνώμες δηλαδή των έμπειρων στελεχών της επιχείρησης.



3. Σφάλματα πρόβλεψης

3.1 Ορισμοί και τύποι σφαλμάτων

Κατά την εφαρμογή οποιουδήποτε μοντέλου πρόβλεψης, ως δεδομένα εισόδου, χρησιμοποιούνται ιστορικά στοιχεία τα οποία αποτελούνται από συστηματικό και τυχαίο μέρος. Το τυχαίο αυτό μέρος σχετίζεται άμεσα με το σφάλμα πρόβλεψης το οποίο πρέπει να υπολογίζεται και να παρακολουθείται ούτως ώστε να ελέγχεται η καταλληλότητα της μεθόδου πρόβλεψης.

Ως αποτελεσματικότητα μιας μεθόδου πρόβλεψης ορίζουμε τη σχέση μεταξύ των αποτελεσμάτων (προβλέψεων) που παράγει η μέθοδος σε σχέση με τις πραγματικά εκδηλωμένες τιμές της προς εξέταση μεταβλητής. Η αποτελεσματικότητα αξιολογείται με μια σειρά δεικτών οι οποίοι μετρούν τις αποκλίσεις (σφάλματα) της πρόβλεψης και αποτελούν το μέτρο καταλληλότητας της εκάστοτε μεθόδου πρόβλεψης που χρησιμοποιείται.

Ορίζουμε ως e_t το σφάλμα της πρόβλεψης, δηλαδή τη διαφορά ανάμεσα στην πραγματική τιμή D_t και την πρόβλεψη F_t της μεταβλητής για την περίοδο t δηλαδή:

$$e_t = D_t - F_t$$

Στη συνέχεια παρουσιάζονται οι βασικοί δείκτες μέτρησης των σφαλμάτων, οι οποίοι χρησιμοποιούνται ως κριτήρια καταλληλότητας των μεθόδων πρόβλεψης (συνολικός πλήθος χρονικών περιόδων, N) :

- i. Μέσο Σφάλμα ή Μέση Απόκλιση (Mean Deviation - MD)

$$MD = \frac{1}{N} \sum_{t=1}^N e_t = \frac{1}{N} \sum_{t=1}^N (D_t - F_t)$$

- ii. Μέση Απόλυτη Απόκλιση (Mean Absolute Deviation – MAD)

$$MAD = \frac{1}{N} \sum_{t=1}^N |e_t| = \frac{1}{N} \sum_{t=1}^N |D_t - F_t|$$



- iii. Μέσο τετραγωνικό σφάλμα (Mean squared error – MSE)

$$MSE = \frac{1}{N} \sum_{t=1}^N e_t^2$$

- iv. Τετραγωνική ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Squared Error – RMSE)

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N e_t^2}$$

- v. Τυπική απόκλιση σφαλμάτων (Standard deviation of error, σ)

$$TAS = \sqrt{\frac{\sum_{t=1}^N e_t^2}{N - 1}}$$

- vi. Ποσοστιαίο σφάλμα (Percentage Error – PE)

$$PE = \frac{e_t}{D_t} 100$$

- vii. Μέσο ποσοστιαίο σφάλμα (Mean Percentage Error – MPE)

$$MPE = \frac{1}{N} \sum_{t=1}^N \frac{e_t}{D_t} 100$$

- viii. Μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error – MAPE)

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{e_t}{D_t} \right| 100$$



3.2 Ερμηνεία των σφαλμάτων

Η εγκυρότητα, η αξιοπιστία και άρα η καταλληλότητα μιας μεθόδου πρόβλεψης, εξαρτάται από τον βαθμό που αυτή προσαρμόζεται στα διαθέσιμα δεδομένα και από την ακρίβεια που επιτυγχάνει στην πρόβλεψη των πραγματικών τιμών της προς εξέταση μεταβλητής. Σε κάθε περίπτωση, για να κριθεί ο βαθμός καταλληλότητας του μοντέλου χρησιμοποιούνται τα παρακάτω σφάλματα – κριτήρια αξιολόγησης :

i. Μέση Απόλυτη Απόκλιση (Mean Absolute Deviation – MAD)

Το κριτήριο MAD εκφράζει τη μέση τιμή των απόλυτων αποκλίσεων των προβλεπόμενων τιμών της χρονοσειράς από τις αντίστοιχες πραγματικές και έχει τα ακόλουθα χαρακτηριστικά. Η μονάδα μέτρησης της μέσης απόλυτης απόκλισης είναι ίδια με τη μονάδα μέτρησης των τιμών της χρονοσειράς, με αποτέλεσμα την ευκολότερη ερμηνεία και κατανόησή του. Επιπλέον, στον υπολογισμό της απόκλισης λαμβάνονται υπ' όψιν μόνο οι απόλυτες τιμές των σφαλμάτων και όχι οι πραγματικές τιμές τους. Αυτό σημαίνει ότι το κριτήριο MAD είναι ανεξάρτητο από θετικές ή αρνητικές τιμές του σφάλματος, δηλαδή είναι ανεξάρτητο από το αν οι τιμές των προβλέψεων είναι μικρότερες (υποεκτίμηση) ή μεγαλύτερες (υπερεκτίμηση) των πραγματικών τιμών. Το MAD βασίζεται στην υπόθεση ότι το κόστος που δημιουργείται από το σφάλμα της πρόβλεψης, σχετίζεται γραμμικά με το μέγεθος του σφάλματος.

ii. Μέσο τετραγωνικό σφάλμα (Mean squared error – MSE) / Τετραγωνική ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Squared Error – RMSE)

Το MSE είναι η μέση τιμή των τετραγώνων των αποκλίσεων των προβλεπόμενων τιμών της χρονοσειράς από τις αντίστοιχες πραγματικές. Η μονάδα μέτρησης του MSE είναι εκφρασμένη στη μονάδα μέτρησης των τιμών των παρατηρήσεων της χρονοσειράς, υψωμένη στο τετράγωνο. Για το λόγο αυτό, μερικές φορές χρησιμοποιούμε τη θετική τιμή της τετραγωνικής του ρίζας, που ονομάζεται τετραγωνική ρίζα μέσου σφάλματος τετραγώνου RMSE (Root Mean Squared Error). Το RMSE εκφράζεται στην ίδια μονάδα μέτρησης με εκείνη των τιμών της χρονοσειράς.

Η ύπαρξη προβλέψεων που απέχουν πολύ από τις αντίστοιχες πραγματικές τιμές γίνεται πολύ περισσότερο αισθητή με το κριτήριο MSE από ότι με το κριτήριο MAD, επειδή οι τιμές των σφαλμάτων της πρόβλεψης υψώνονται στο τετράγωνο. Συνεπώς το κριτήριο MSE είναι στατιστικά πιο



αξιόπιστο από το κριτήριο MAD και χρησιμοποιείται συχνότερα για την επιλογή της ‘κατάλληλης’ μεθόδου πρόβλεψης.

iii. Μέσο ποσοστιαίο σφάλμα (Mean Percentage Error – MPE)

Το μέσο ποσοστιαίο σφάλμα χρησιμοποιείται όταν ο στόχος είναι να εξεταστεί κατά πόσο η μέθοδος πρόβλεψης είναι μεροληπτική, δηλαδή αν οι προβλεπόμενες τιμές είναι συστηματικά μεγαλύτερες ή μικρότερες από τις αντίστοιχες πραγματικές. Όσο πιο κοντά στο μηδέν είναι η τιμή του MPE, τόσο πιο αμερόληπτη και αντιπροσωπευτική είναι η μέθοδος πρόβλεψης που χρησιμοποιήθηκε. Αντίθετα, μεγάλες απόλυτες τιμές του MPE φανερώνουν θετική ή αρνητική, μεροληψία της μεθόδου ως προς τις πραγματικά εκδηλωμένες τιμές.

iv. Μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error – MAPE)

Το μέσο απόλυτο ποσοστιαίο σφάλμα εξετάζει τη συμπεριφορά της απόλυτης τιμής του σφάλματος της πρόβλεψης σε σχέση με την πραγματική τιμή της χρονοσειράς. Γενικά, το MAPE ορίζεται ως το άθροισμα των απόλυτων τιμών των σφαλμάτων της πρόβλεψης προς τις αντίστοιχες πραγματικές τιμές της χρονοσειράς, διαιρούμενο με τον αριθμό των χρονικών περιόδων N , στις οποίες έγιναν προβλέψεις. Το κριτήριο αυτό είναι απαλλαγμένο από μονάδες μέτρησης και χρησιμοποιείται, εκτός από κριτήριο καταλληλότητας των μεθόδων πρόβλεψης, για τη σύγκριση της ακρίβειας μιας ή περισσότερων μεθόδων προβλέψεων, για περισσότερες από μια χρονοσειρές.

3.3 Παρακολούθηση πρόβλεψης (Monitoring forecasting systems)

Η διαδικασία της πρόβλεψης, χαρακτηρίζεται από μια μεταβλητότητα η οποία δημιουργεί την ανάγκη για συνεχή υπολογισμό, παρακολούθηση και ερμηνεία των αποτελεσμάτων που εξάγει το μοντέλο, είτε αυτά είναι προβλέψεις, είτε σφάλματα των προβλέψεων. Η μεταβλητότητα αυτή, πηγάζει μέσα από την ενδεχόμενη εμφάνιση νέων, αστάθμητων παραγόντων που θα επηρεάσουν την τελική τιμή της εξεταζόμενης μεταβλητής, προκαλώντας μη αποδεκτές αποκλίσεις μεταξύ των πραγματικά εκδηλωμένων τιμών και των αντίστοιχων προβλέψεων του χρησιμοποιούμενου μοντέλου. Η έγκαιρη αναγνώριση της εμφάνισης των παραγόντων που προκάλεσαν μη αποδεκτές αποκλίσεις είναι κρίσιμη για τη διαδικασία της πρόβλεψης και για το λόγο αυτό αναπτύχθηκαν συστήματα παρακολούθησης των προβλέψεων, τα βασικά εργαλεία των οποίων παρουσιάζονται στη συνέχεια.



Για να προσδιορίσουμε αν μια μέθοδος πρόβλεψης συστηματικά υπερεκτιμά ή υποεκτιμά τη ζήτηση, πρέπει να χρησιμοποιηθεί το άθροισμα των σφαλμάτων που παράγονται, για να υπολογιστεί η *αμεροληψία* της μεθόδου. Η αμεροληψία υπολογίζεται από τον παρακάτω τύπο και πρέπει να κυμαίνεται γύρω από το 0 εφόσον τα λάθη είναι τελείως τυχαία και δεν έχουν τάση (υπερεκτίμησης ή υποεκτίμησης).

$$Bias = \sum_{t=1}^N e_t$$

Με δεδομένες τις τιμές αμεροληψίας της μεθόδου, μπορεί να υπολογισθεί το σήμα ανίχνευσης (Tracking Signal – TS). Όταν το *TS* βρίσκεται έξω από το εύρος των τιμών (-6,+6) αυτό αποτελεί ένδειξη ότι η πρόβλεψη έχει τάση είτε υποεκτίμησης (όταν $TS \leq -6$) είτε υπερεκτίμησης (όταν $TS \geq +6$), οπότε απαιτείται να ερευνηθεί η αιτία της απόκλισης και τελικά, η καταλληλότητα της μεθόδου.

$$TS = \frac{Bias}{MAD}$$

Το σήμα ανίχνευσης αν και είναι βασικό εργαλείο που χρησιμοποιείται για την παρακολούθηση πρόβλεψης, παρουσιάζει κάποια μειονεκτήματα. Συγκεκριμένα, το σήμα ανίχνευσης υπολογίζεται με βάση τα *σωρευτικά σφάλματα* που παρατηρήθηκαν στη χρονική διάρκεια που μελετάται. Αυτό έχει ως αποτέλεσμα τυχόν μεγάλες τιμές θετικών σφαλμάτων να αντισταθμίζονται από τυχόν μεγάλες τιμές αρνητικών σφαλμάτων, λειτουργώντας με τον τρόπο αυτό παραπλανητικά στην κρίση του αναλυτή, όσον αφορά τον έλεγχο της επάρκειας του μοντέλου πρόβλεψης. Χρησιμοποιώντας μόνο το σήμα ανίχνευσης σαν εργαλείο παρακολούθησης, εν γένει, δεν υπάρχει η δυνατότητα παρακολούθησης των σφαλμάτων, διακριτά, ανά χρονική περίοδο.

Για να αντιμετωπιστεί το πρόβλημα αυτό, τα τελευταία χρόνια χρησιμοποιείται παράλληλα με το σήμα ανίχνευσης, το διάγραμμα ελέγχου (control chart) σφαλμάτων. Στο διάγραμμα αυτό, προϋπολογίζονται τα άνω και κάτω όρια εντός των οποίων τα σφάλματα θεωρούνται αποδεκτά. Το πλεονέκτημα έναντι της χρήσης μεμονωμένα του σήματος ανίχνευσης για την παρακολούθηση των σφαλμάτων, εντοπίζεται στο γεγονός ότι η γραφική απεικόνιση της χρονικής εξέλιξης των σφαλμάτων, καθιστά πολύ πιο εύκολη την αναγνώριση τυχόν μοτίβων διακύμανσης ή προτύπων τάσης που ενδέχεται να παρουσιαστούν κατά τη χρονική διάρκεια παρακολούθησης των προβλέψεων.



Τα διαγράμματα ελέγχου, στηρίζονται στη θεώρηση ότι αν τα σφάλματα είναι τυχαία, θα κατανέμονται σύμφωνα με την κανονική κατανομή, με μέση τιμή ίση με το μηδέν και τυπική απόκλιση $s = \sqrt{MSE}$. Ισχύει ότι σε μια κανονική κατανομή, περίπου το 95% των παρατηρούμενων τιμών, αναμένεται να βρίσκονται εντός των ορίων των $0 \pm 2s$ (δηλαδή, 0 ± 2 τυπικές αποκλίσεις).

Το πρώτο βήμα για την κατασκευή ενός διαγράμματος ελέγχου, είναι ο υπολογισμός του μέσου τετραγωνικού σφάλματος (MSE). Η τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος, χρησιμοποιείται ως εκτίμηση της τυπικής απόκλισης της κατανομής των σφαλμάτων. Δηλαδή :

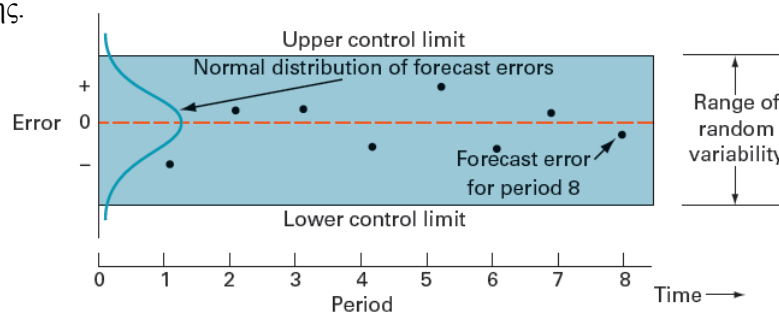
$$s = \sqrt{MSE}$$

Τα όρια, εντός των οποίων πρέπει να κατανέμονται τα σφάλματα πρόβλεψης, εφόσον είναι τυχαία και επομένως ακολουθούν την κανονική κατανομή, θα είναι :

Άνω όριο ελέγχου (Upper Control Limit - UCL) : $0 + 1.96\sqrt{MSE}$

Κάτω όριο ελέγχου (Lower Control Limit - LCL) : $0 - 1.96\sqrt{MSE}$

Στη συνέχεια παρατίθενται η μορφή ενός βασικού διαγράμματος ελέγχου καθώς τα άνω και κάτω όρια ελέγχου, εντός των οποίων πρέπει να βρίσκονται τα υπολογιζόμενα, σε κάθε χρονική περίοδο, σφάλματα πρόβλεψης.



Σχήμα 3.1 : Διάγραμμα Ελέγχου Σφαλμάτων

Η παρακολούθηση του χρησιμοποιούμενου συστήματος πρόβλεψης, είναι το ίδιο σημαντική με την υλοποίηση και χρήση του. Σε καμία περίπτωση ένα, προσωρινά επαρκές μοντέλο πρόβλεψης, δεν σημαίνει ότι θα παραμείνει επαρκές στο πέρασμα του χρόνου. Απαιτείται επομένως συνεχής παρακολούθηση των προβλέψεων και εκ των υστέρων των σφαλμάτων τους, για την έγκαιρη αναγνώριση της ανεπάρκειας ενός μοντέλου και άρα την αναγνώριση της ανάγκης για επανεκτίμηση των παραμέτρων ή ακόμη και τη χρησιμοποίηση διαφορετικού μοντέλου.



Παράδειγμα

Στον πίνακα που ακολουθεί, καταγράφονται ο ημερήσιος αριθμός των επισκεπτών που καταφθάνουν στο συνεργείο επισκευής Gary's Chevron, η αντίστοιχη πρόβλεψη που διενεργήθηκε με βάση τη χρησιμοποιούμενη μέθοδο και τα σφάλματα που προέκυψαν.

Time (t)	Customers (D_t)	Forecast (F_t)	Error (e_t)	$ e_t $	$(e_t)^2$	$ e_t /D_t$	e_t/D_t
1	58	-	-	-	-	-	-
2	54	58	-4	4	16	0.074	-0.074
3	60	54	6	6	36	0.100	0.100
4	55	60	-5	5	25	0.091	-0.091
5	62	55	7	7	49	0.113	0.113
6	62	62	0	0	0	0.000	0.000
7	65	62	3	3	9	0.046	0.046
8	63	65	-2	2	4	0.032	-0.032
9	70	63	7	7	49	0.100	0.100
Totals			12	34	188	0.556	0.162

Πίνακας 3.1 : Υπολογισμός Σφαλμάτων Πρόβλεψης

(Πηγή : John E. Hanke and Dean W. Wichern, *Business Forecasting* (9th ed.))

$$MAD = \frac{1}{N} \sum_{t=1}^N |e_t| = \frac{1}{N} \sum_{t=1}^N |D_t - F_t| = \frac{34}{8} = 4.3$$

$$MSE = \frac{1}{N} \sum_{t=1}^N e_t^2 = \frac{188}{8} = 23.5$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N e_t^2} = \sqrt{23.5} = 4.8$$

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{e_t}{D_t} \right| 100 = \frac{0.556}{8} 100 = 0.0695 = 6.95\%$$

$$MPE = \frac{1}{N} \sum_{t=1}^N \frac{e_t}{D_t} 100 = \frac{0.162}{8} 100 = 0.0203 = 2.03\%$$

Το MAD υποδηλώνει ότι κάθε πρόβλεψη, αποκλίνει από την πραγματική ζήτηση κατά 4.3 πελάτες. Το MSE της τάξεως του 23.5 ή (RMSE 4.8) και το MAPE της τάξεως του 6.95% θα συγκριθούν με τις αντίστοιχες τιμές οποιασδήποτε άλλης τεχνικής πρόβλεψης ενδεχομένως να μελετηθεί, για την εξεύρεση της καταλληλότερης τεχνικής που τελικά θα υιοθετηθεί.

Τέλος, η μικρή τιμή του MPE το 2.03%, υποδηλώνει ότι η τεχνική πρόβλεψης που χρησιμοποιείται είναι αμερόληπτη. Από τη στιγμή που η τιμή του MSE είναι κοντά στο μηδέν, σημαίνει δεν υπάρχει τάση υπό-εκτίμησης ή υπέρ-εκτίμησης των προβλέψεων.



4. Ανάλυση ARIMA (Autoregressive Integrated Moving Average)

4.1 Αυτοπαλίνδρομα μοντέλα και ανάλυση χρονοσειρών

Τα ολοκληρωμένα αυτοπαλίνδρομα μοντέλα κινητού μέσου όρου (autoregressive integrated moving average) είναι στοχαστικά μαθηματικά μοντέλα τα οποία χρησιμοποιούνται για την περιγραφή της διαχρονικής εξέλιξης κάποιου φυσικού μεγέθους. Τα στοχαστικά μοντέλα περιλαμβάνουν τον τυχαίο παράγοντα (τυχαίο σφάλμα), τις τιμές του προς εξέταση μεγέθους οι οποίες εμφανίζονται σε προηγούμενες χρονικές στιγμές και ίσως μερικούς πρόσθετους στοχαστικούς παράγοντες. Το μοντέλο που προκύπτει είναι ένας γραμμικός συνδυασμός των παραπάνω ποσοτήτων.

Οι *Box-Jenkins* πρότειναν μια οικογένεια θεωρητικών αλγεβρικών μοντέλων, από τα οποία κάθε φορά επιλέγεται το πιο κατάλληλο για την πρόβλεψη μιας χρονοσειράς. Η διαδικασία επιλογής του κατάλληλου μοντέλου περιγράφεται από τη σχετική μεθοδολογία των *Box-Jenkins*, η οποία θα παρουσιαστεί εκτενώς σε επόμενη παράγραφο. Αυτό που αξίζει να σημειωθεί, είναι ότι τα μοντέλα *ARIMA* διαφέρουν από τα παραδοσιακά μοντέλα στο γεγονός ότι οι προβλέψεις στηρίζονται και στις προηγούμενες χρονικά τιμές της προς εξέταση μεταβλητής καθώς και στα σφάλματα πρόβλεψης της χρονοσειράς. Επιπρόσθετα, η χρησιμοποίηση των μοντέλων *ARIMA* με στόχο την πρόβλεψη μιας χρονοσειράς, δεν απαιτεί την ύπαρξη κάποιου επαναλαμβανόμενου μοτίβου διακύμανσης των τιμών, από τις παρατηρήσεις των προηγούμενων χρόνων.

4.2 Βασικά στοιχεία της ανάλυσης ARIMA

4.2.1 Συντελεστής αυτοσυσχέτισης (autocorrelation function - ACF)

Ο υπολογισμός και η γραφική αναπαράσταση των συντελεστών αυτοσυσχέτισης μιας χρονοσειράς, για διάφορες χρονικές καθυστερήσεις, είναι βασική προϋπόθεση για να μπορέσουμε να αποφανθούμε κατά πόσο :

- τα δεδομένα είναι τυχαία
- η χρονοσειρά είναι σταθερή ή εμφανίζει κάποιο πρότυπο τάσης
- εάν υπάρχει το στοιχείο της εποχικότητας και ποιο το πρότυπο που ακολουθεί



Όταν μια μεταβλητή μετρείται κατά τη διάρκεια του χρόνου, όπως για παράδειγμα οι πωλήσεις ενός προϊόντος, οι παρατηρήσεις που καταγράφονται συχνά συσχετίζονται μεταξύ τους. Η συσχέτιση αυτή, υπολογίζεται μέσω του συντελεστή αυτοσυσχέτισης. Συγκεκριμένα, ο συντελεστής αυτοσυσχέτισης είναι ένας στατιστικός δείκτης ο οποίος χρησιμοποιείται στην ανάλυση χρονοσειρών για τον καθορισμό της τυχαιότητας ή μη των δεδομένων, για τον εντοπισμό στοιχείων τάσης ή εποχικότητας αλλά και για να αποφανθεί κατά πόσο η χρονοσειρά μπορεί να θεωρηθεί στάσιμη ή όχι.

Ο συντελεστής αυτοσυσχέτισης r_k δείχνει την συσχέτιση της χρονοσειράς με τον εαυτό της για παρατηρήσεις που υστερούν χρονικά μεταξύ τους κατά k περιόδους. Για παράδειγμα, ο συντελεστής αυτοσυσχέτισης r_1 φανερώνει τον τρόπο με τον οποίο σχετίζονται διαδοχικές παρατηρήσεις της χρονοσειράς, ο συντελεστής αυτοσυσχέτισης r_2 φανερώνει τον τρόπο με τον οποίο σχετίζονται παρατηρήσεις της χρονοσειράς που απέχουν δύο χρονικές περιόδους κ.ο.κ. Η γραφική παράσταση των συντελεστών αυτοσυσχέτισης για $1, 2, \dots, k$ χρονικές καθυστερήσεις πραγματοποιείται με τα αποτελέσματα που εξάγει η συνάρτηση αυτοσυσχέτισης, η οποία ορίζεται ως εξής :

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_1^n (Y_t - \bar{Y})^2}$$

r_k = συντελεστής αυτοσυσχέτισης για χρονική καθυστέρηση k χρονικών περιόδων

\bar{Y} = η μέση τιμή των παρατηρήσεων της χρονοσειράς

Y_t = η παρατήρηση (τιμή της μεταβλητής) τη χρονική περίοδο t

Y_{t-k} = η παρατήρηση κατά τη χρονική περίοδο $t-k$

Η συνάρτηση αυτοσυσχέτισης (autocorrelation function) ή correlogram είναι το γράφημα των αυτοσυσχετίσεων για διάφορες χρονικές καθυστερήσεις των παρατηρήσεων της χρονοσειράς. Η αποτελεσματική μελέτη των χρονοσειρών μέσω της ανάλυσης συσχετίσεων είναι το καθοριστικότερο βήμα για τον προσδιορισμό του κατάλληλου μοντέλου πρόβλεψης ARIMA.

Αν τα δεδομένα μιας χρονοσειράς είναι *τυχαία*, τότε οι συσχετίσεις μεταξύ των παρατηρήσεων Y_t, Y_{t-k} , για κάθε χρονική καθυστέρηση k , είναι περίπου μηδέν, δηλαδή οι τιμές της χρονοσειράς δεν σχετίζονται μεταξύ τους. Σύμφωνα με τους Quenouille et al. (1949), η συνάρτηση αυτοσυσχέτισης τυχαίων δεδομένων, προσεγγίζεται από την καμπύλη κανονικής κατανομής με μέση τιμή μηδέν και



τυπική απόκλιση $1/\sqrt{n}$. Επομένως, σε μία τυχαία χρονοσειρά το 95% των συντελεστών αυτοσυσχέτισης θα βρίσκονται στο διάστημα που ορίζεται από τις τιμές $\pm Z_{0.025}/\sqrt{n}$ όπου n είναι ο αριθμός των παρατηρήσεων της χρονοσειράς και $Z_{0.025}$ η τιμή που προκύπτει από την πίνακα κανονικής κατανομής για επίπεδο εμπιστοσύνης 95%. Στατιστικά σημαντικός (δηλαδή διάφορος του μηδενός) συντελεστής αυτοσυσχέτισης, θεωρείται εκείνος όπου υπερβαίνει τα όρια που ορίζουν οι τιμές $\pm 1,96/\sqrt{n}$.

Ένα παράδειγμα τυχαίας χρονοσειράς, είναι το θεμελιώδες μοντέλο λευκού θορύβου, το οποίο χρησιμοποιείται ευρέως στην ανάλυση χρονοσειρών και περιγράφεται από τη σχέση : $Y_t = c + e_t$ όπου Y_t είναι η παρατήρηση κατά τη χρονική στιγμή t , c είναι μία σταθερή τιμή και e_t είναι η τιμή του τυχαίου σφάλματος της συγκεκριμένης παρατήρησης.

Στην περίπτωση που μια χρονοσειρά παρουσιάζει **πρότυπο τάσης**, οι παρατηρήσεις της παρουσιάζουν υψηλή αυτοσυσχέτιση. Οι συντελεστές αυτοσυσχέτισης είναι σημαντικά διάφοροι του μηδενός για τις πρώτες χρονικές καθυστερήσεις και σταδιακά πέφτουν στο μηδέν καθώς ο αριθμός των χρονικών καθυστερήσεων αυξάνεται. Ο συντελεστής αυτοσυσχέτισης για χρονική καθυστέρηση (lag 1), συχνά θα είναι πολύ μεγάλος, περίπου κοντά στη μονάδα.

Αν η χρονοσειρά παρουσιάζει **εποχική διακύμανση**, θα εμφανίζεται σημαντικά διάφορος του μηδενός συντελεστής αυτοσυσχέτισης για τις εποχικές χρονικές καθυστερήσεις ή τις πολλαπλάσιες αυτών. Για παράδειγμα, υπονοείται εποχική διακύμανση για μηνιαία δεδομένα, εάν υπάρχει στατιστικά σημαντικός συντελεστής αυτοσυσχέτισης για χρονική υστέρηση 12 περιόδων. Αυτό μεταφράζεται στο ότι υπάρχει συσχέτιση του μηνός Ιανουαρίου με τον αντίστοιχο Ιανουάριο κάθε επόμενου έτους, του μηνός Φεβρουαρίου με τον αντίστοιχο Φεβρουάριο κάθε επόμενου έτους κ.ο.κ. Εάν τα δεδομένα είναι τριμηνιαία, υποδηλώνεται εποχική διακύμανση αν παρουσιαστεί στατιστικά σημαντικός συντελεστής αυτοσυσχέτισης για χρονική καθυστέρηση 4 περιόδων.

Για τη μελέτη χρονοσειρών μέσω της ανάλυσης αυτοσυσχετίσεων, χρησιμοποιούνται στατιστικά πακέτα (βλ. Excel, Minitab, SPSS) που έχουν ενσωματωμένη τη μεθοδολογία υπολογισμού των συντελεστών και σαν έξοδο, παρουσιάζουν τα προκύπτοντα διαγράμματα αυτοσυσχέτισης για κάθε χρονική καθυστέρηση. Υπενθυμίζεται όμως ότι η επιτυχής αναγνώριση του μοντέλου ARIMA είναι



συνδεδεμένη με την πλήρη κατανόηση και κατάλληλη ερμηνεία της μορφής των συναρτήσεων αυτοσυσχέτισης (correlogram) της χρονοσειράς που μελετάται.

Στη συνέχεια, για σκοπούς πληρότητας της μελέτης, παρατίθεται ένα απλό παράδειγμα υπολογισμού των συντελεστών αυτών, καθώς και ένα παράδειγμα ερμηνείας μιας συνάρτησης αυτοσυσχέτισης.

Παράδειγμα 1 : Υπολογισμός συντελεστών αυτοσυσχέτισης

Για την καλύτερη κατανόηση του τρόπου υπολογισμού των συντελεστών αυτοσυσχέτισης για διάφορες χρονικές καθυστερήσεις, ας ρίξουμε μια ματιά στα δεδομένα του παρακάτω πίνακα. Τα δεδομένα απεικονίζουν τις μηνιαίες πωλήσεις (Y_t) του προϊόντος X, μιας εταιρείας, κατά τη διάρκεια του προηγούμενου έτους. Στις επόμενες στήλες, απεικονίζονται οι πωλήσεις του προϊόντος αλλά για χρονικές καθυστερήσεις μίας και δυο περιόδων αντίστοιχα. Στόχος είναι ο υπολογισμός των συντελεστών αυτοσυσχέτισης για καθυστερήσεις μίας και δύο χρονικών περιόδων.

Time (t)	Month	Original Data Y_t	Y Lagged One	Y Lagged Two
			Period Y_{t-1}	Periods Y_{t-2}
1	January	123	-	-
2	February	130	123	-
3	March	125	130	123
4	April	138	125	130
5	May	145	138	125
6	June	142	145	138
7	July	141	142	145
8	August	146	141	142
9	September	147	146	141
10	October	157	147	146
11	November	150	157	147
12	December	160	150	157

Πίνακας 4.1 : Δεδομένα πωλήσεων και χρονικές καθυστερήσεις

Πηγή : John E.Hanke and Dean W.Wichern, *Business Forecasting 9th ed.*)



Για τη διευκόλυνση των υπολογισμών, χρησιμοποιείται ο παρακάτω πίνακας. Ο συντελεστής αυτοσυσχέτισης για χρονική καθυστέρηση μιας περιόδου (lag 1) ή ισοδύναμα, η αυτοσυσχέτιση μεταξύ των παρατηρήσεων της χρονοσειράς Y_t, Y_{t-1} υπολογίζεται ως εξής :

$$r_1 = \frac{\sum_{t=1}^n (Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$$

Αρχικά, υπολογίζεται η μέση τιμή των παρατηρήσεων της χρονοσειράς, που ως γνωστό, δίνεται από τη σχέση (n ο αριθμός των παρατηρήσεων της χρονοσειράς) :

$$\bar{Y} = \frac{\sum Y_t}{n} = \frac{1704}{12} = 142$$

Στον πίνακα που ακολουθεί υπολογίζονται διακριτά, σε κάθε στήλη, οι ποσότητες που υπεισέρχονται στη σχέση υπολογισμού του συντελεστή αυτοσυσχέτισης για χρονική καθυστέρηση μιας περιόδου.

Time (t)	Y_t	Y_{t-1}	$Y_t - \bar{Y}$	$Y_{t-1} - \bar{Y}$	$(Y_t - \bar{Y})^2$	$(Y_t - \bar{Y}) \cdot (Y_{t-1} - \bar{Y})$
1	123	-	-19	-	361	-
2	130	123	-12	-19	144	228
3	125	130	-17	-12	289	204
4	138	125	-4	-17	16	68
5	145	138	3	-4	9	-12
6	142	145	0	3	0	0
7	141	142	-1	0	1	0
8	146	141	4	-1	16	-4
9	147	146	5	4	25	20
10	157	147	15	5	225	75
11	150	157	8	15	64	120
12	160	150	18	8	324	144
Total	1704				1474	843

Πίνακας 4.2 : Υπολογισμός συντελεστών αυτοσυσχέτισης



Έχοντας προϋπολογίσει όλες τις απαιτούμενες ποσότητες, προκύπτει η τιμή του συντελεστή αυτοσυσχέτισης :

$$r_1 = \frac{\sum_{t=1+1}^n (Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_1^n (Y_t - \bar{Y})^2} = \frac{843}{1474} = 0.572$$

Ομοίως, ο συντελεστής αυτοσυσχέτισης 2^{ης} τάξης, ή ισοδύναμα η αυτοσυσχέτιση μεταξύ των τιμών Y_t , Y_{t-2} της χρονοσειράς, υπολογίζεται ως :

Time (t)	Y_t	Y_{t-2}	$Y_t - \bar{Y}$	$Y_{t-2} - \bar{Y}$	$(Y_t - \bar{Y})^2$	$(Y_t - \bar{Y}) \cdot (Y_{t-2} - \bar{Y})$
1	123	-	-19	-	361	-
2	130	-	-12	-	144	-
3	125	123	-17	-19	289	323
4	138	130	-4	-12	16	48
5	145	125	3	-17	9	-51
6	142	138	0	-4	0	0
7	141	145	-1	3	1	-3
8	146	142	4	0	16	0
9	147	141	5	-1	25	-5
10	157	146	15	4	225	60
11	150	147	8	5	64	40
12	160	157	18	15	324	270
Total	1704				1474	685

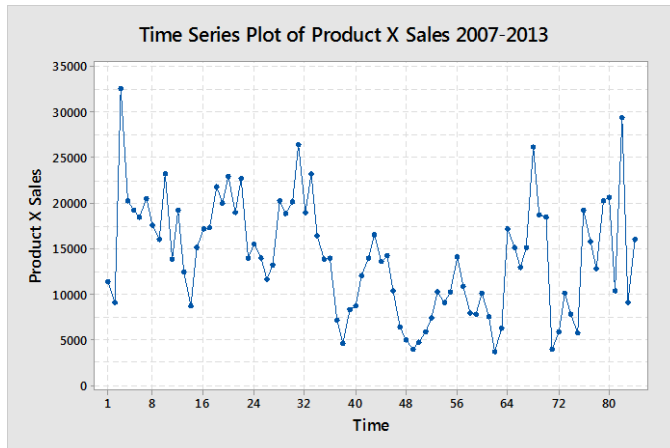
Πίνακας 4.3 : Υπολογισμός συντελεστή αυτοσυσχέτισης 2^{ης} τάξης

$$r_2 = \frac{\sum_{t=2+1}^n (Y_t - \bar{Y})(Y_{t-2} - \bar{Y})}{\sum_1^n (Y_t - \bar{Y})^2} = \frac{682}{1474} = 0.463$$

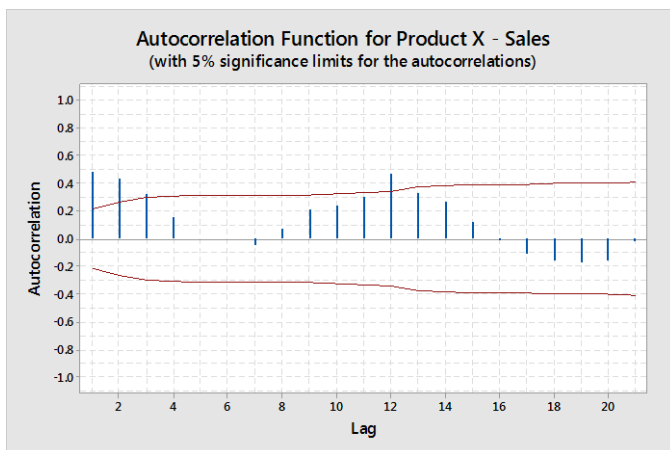
Στη μελέτη χρονοσειρών μέσω της ανάλυσης αυτοσυσχετίσεων, ο υπολογισμός των συντελεστών είναι το πρώτο βήμα που εφαρμόζεται και έχει ως στόχο την εξαγωγή του διαγράμματος της συνάρτησης αυτοσυσχέτισης. Αυτό όμως που έχει σημασία σε επόμενη φάση, είναι η κατάλληλη ερμηνεία των διαγραμμάτων που προκύπτουν, αφού μέσα σε αυτά, εμπερικλείονται κρίσιμες πληροφορίες σχετικά με τους παράγοντες που επηρεάζουν τη διακύμανση της χρονοσειράς. Ενδεικτικό, είναι το παράδειγμα που ακολουθεί.

Παράδειγμα 2 : Ερμηνεία συνάρτησης αυτοσυσχέτισης

Στο διάγραμμα που ακολουθεί απεικονίζεται η συνάρτηση αυτοσυσχέτισης (correlogram) για τις μηνιαίες πωλήσεις του προϊόντος X της εταιρείας για τις ανάγκες της οποίας πραγματοποιήθηκε η ανάλυση της χρονοσειράς. (βλ. **Παράρτημα Α**). Οι τιμές των συντελεστών αυτοσυσχέτισης έχουν υπολογιστεί σύμφωνα με τη μεθοδολογία που περιγράφηκε στο προηγούμενο παράδειγμα.



Σχήμα 4.1 : Χρονοσειρά Πωλήσεων



Σχήμα 4.2 : Συνάρτηση Αυτοσυσχέτισης

Αρχικά, παρατηρούμε ότι τα δεδομένα της χρονοσειράς μπορούν να χαρακτηριστούν *τυχαία*, αφού το 95% των συντελεστών αυτοσυσχέτισης βρίσκονται εντός των ορίων

$$\pm \frac{1.96}{\sqrt{n}} = \pm \frac{1.96}{\sqrt{84}} = \pm 0.214.$$

Από τη συνάρτηση αυτοσυσχέτισης προκύπτει επίσης ότι οι πωλήσεις του προϊόντος παρουσιάζουν πρότυπο *τάσης*, αφού οι συντελεστές αυτοσυσχέτισης είναι σημαντικά διάφοροι του μηδενός για τις πρώτες χρονικές καθυστερήσεις και σταδιακά πέφτουν στο μηδέν.

Επιπρόσθετα, αναλογιζόμενοι το γεγονός ότι τα δεδομένα είναι σε μηνιαία βάση, παρατηρούμε ότι υπάρχει *εποχική διακύμανση* των πωλήσεων του προϊόντος, με διάρκεια περιοδικότητας 12 χρονικές περιόδους. Η παρατήρηση αυτή τεκμηριώνεται στατιστικά από την ύπαρξη σημαντικού συντελεστή αυτοσυσχέτισης για χρονική καθυστέρηση 12 περιόδων, γεγονός που υποδηλώνει υψηλή συσχέτιση μεταξύ των αντίστοιχων πωλήσεων για τους ίδιους μήνες του κάθε έτους. Παρατηρούμε επίσης ότι η χρονοσειρά δεν είναι *στάσιμη*, αφού οι παρατηρήσεις δεν κυμαίνονται γύρω από μια σταθερή τιμή.



Αντιλαμβάνεται λοιπόν κανείς ότι η μελέτη της χρονοσειράς μέσω της ανάλυσης αυτοσυσχετίσεων, είναι απαραίτητη για να μπορέσει ο αναλυτής να εντοπίσει τους σημαντικούς παράγοντες που συμβάλλουν στη δημιουργία του μοτίβου διακύμανσης της προς εξέταση μεταβλητής. Αφού το επιτύχει αυτό, όπως θα δούμε στη συνέχεια, διαμορφώνοντας κατάλληλα τη χρονοσειρά μπορεί να προχωρήσει στην επιλογή του κατάλληλου μοντέλου πρόβλεψης ARIMA.

4.2.2 Κανονική και Εποχική Διαφόριση Χρονοσειρών

Η εξάλειψη των προτύπων τάσης ή/και εποχικής διακύμανσης που τυχόν να παρατηρηθούν στα δεδομένα της χρονοσειράς επιτυγχάνεται με τη χρήση της **μεθόδου κανονικής διαφόρισης**. Οι σειρές διαφορών πρώτης τάξης προκύπτουν από τις διαφορές των διαδοχικών παρατηρήσεων της αρχικής χρονοσειράς και ορίζονται ως εξής :

$$\Delta Y_t = Y_t - Y_{t-1}$$

Ομοίως, οι σειρές διαφορών δεύτερης τάξης προκύπτουν από τις διαφορές των ανά δύο χρονικών περιόδων, διαδοχικών παρατηρήσεων της αρχικής χρονοσειράς και ορίζονται ως :

$$\Delta^2 Y_t = \Delta(\Delta Y_t) = \Delta(Y_t - Y_{t-1}) = Y_t - 2Y_{t-1} + Y_{t-2}$$

Στην περίπτωση τώρα που μία χρονοσειρά εμφανίζει κάποιο εποχικό πρότυπο, τότε πρέπει το πρότυπο αυτό να εξαιρεθεί για την περαιτέρω ανάλυση της χρονοσειράς. Η εξάλειψη του εποχικού προτύπου πραγματοποιείται με τη χρήση της **μεθόδου των εποχιακών διαφορών**.

Ορίζουμε εποχιακή διαφορά, σε χρονοσειρά με διάρκεια περιοδικότητας 12 μήνες, τη διαφορά μιας παρατήρησης της χρονοσειράς από την αντίστοιχη παρατήρηση του προηγούμενου έτους. Σε ορισμένες περιπτώσεις, η εποχική διαφόριση των δεδομένων ακολουθείται από κανονική διαφόριση και η τελική χρονοσειρά που προκύπτει ορίζεται ως εξής :

$$\Delta_{12} Y_t = Y_t - Y_{t-12} = \text{η χρονοσειρά μετά την πρώτη εποχική διαφόριση } s = 12$$

$\Delta\Delta_{12} Y_t = \Delta(\Delta_{12} Y_t) = \Delta(Y_t - Y_{t-12}) = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13} = \text{η χρονοσειρά που προκύπτει μετά από εποχική διαφόριση η οποία ακολουθείται από κανονική διαφόριση των δεδομένων.}$



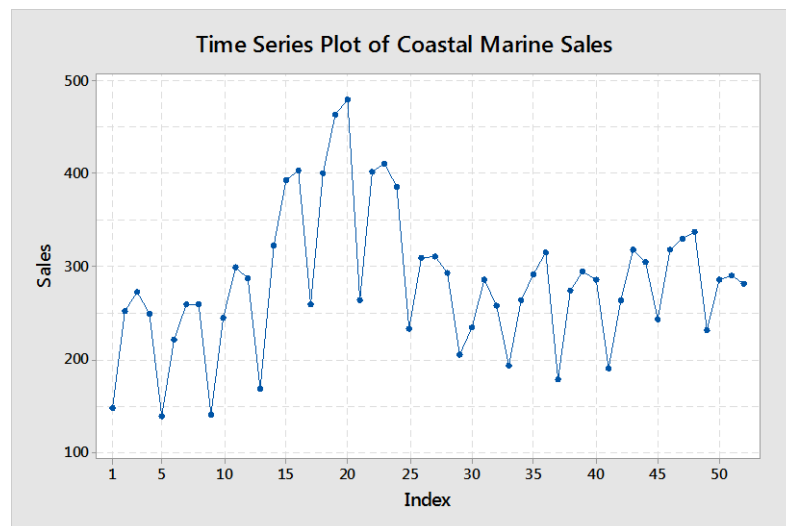
Παράδειγμα διαφορίσης δεδομένων της Χρονοσειράς

Στον πίνακα που ακολουθεί, απεικονίζονται τα δεδομένα πωλήσεων ανά τρίμηνο, της εταιρείας Coastal Marine. Σε πρώτο βήμα πραγματοποιείται γραφική παράσταση της χρονοσειράς και υπολογίζεται η συνάρτηση αυτοσυσχέτισης.

Fiscal Year	December 31	March 31	June 30	September 30
1994	147.6	251.8	273.1	249.1
1995	139.3	221.2	260.2	259.5
1996	140.5	245.5	298.8	287.0
1997	168.8	322.6	393.5	404.3
1998	259.7	401.1	464.4	479.7
1999	264.4	402.6	411.3	385.9
2000	232.7	309.2	310.7	293.0
2001	205.1	234.4	285.4	258.7
2002	193.2	263.7	292.5	315.2
2003	178.3	274.5	295.4	286.4
2004	190.8	263.5	318.8	305.5
2005	242.6	318.8	329.6	338.2
2006	232.1	285.6	291.0	281.4

Πίνακας 4.4 : Δεδομένα πωλήσεων της Coastal Marine

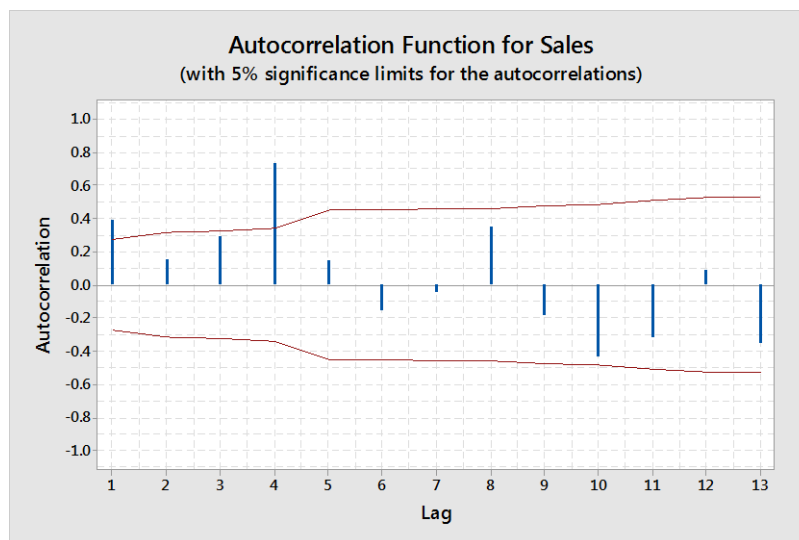
Πηγή : John E.Hanke and Dean W.Wichern, *Business Forecasting* (9th ed.)



Σχήμα 4.3 : Χρονοσειρά Πωλήσεων της Coastal Marine



Όπως εύκολα γίνεται αντιληπτό από την παρατήρηση της διακύμανσης της χρονοσειράς, υπάρχει έντονο το στοιχείο της εποχικότητας για τα δεδομένα πωλήσεων της εταιρείας Coastal Marine. Τα δεδομένα είναι τριμηνιαία γεγονός που σημαίνει ότι η διάρκεια της εποχικότητας είναι 4 χρονικές περίοδοι. Το πρότυπο εποχικότητας, υποδηλώνεται ξεκάθαρα και μέσω της συνάρτησης αυτοσυσχέτισης (βλ. correlogram), αφού σημειώνεται στατιστικά σημαντικός συντελεστής αυτοσυσχέτισης για χρονική καθυστέρηση 4 περιόδων. Αυτό σημαίνει υψηλή συσχέτιση μεταξύ των πωλήσεων των αντίστοιχων τριμήνων κάθε έτους της προς μελέτη χρονοσειράς.



Σχήμα 4.4 : Συνάρτηση αυτοσυσχέτισης της Coastal Marine

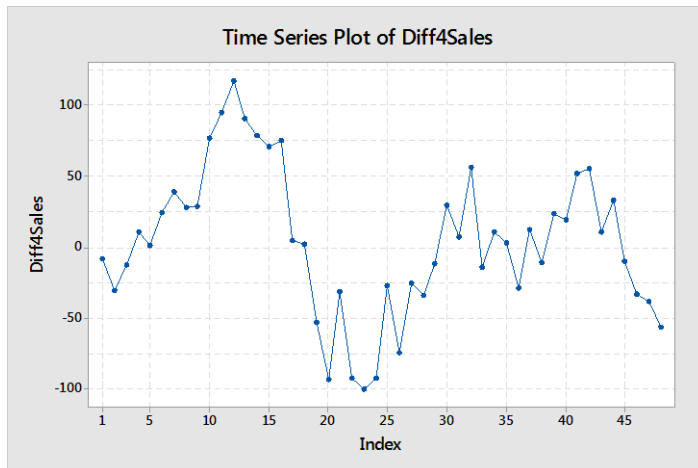
Η εξάλειψη της εποχικής διακύμανσης που τυχόν να παρατηρηθεί στα δεδομένα της χρονοσειράς επιτυγχάνεται με τη χρήση της μεθόδου της διαφοράς, τα αποτελέσματα της οποίας παρατίθενται στον παρακάτω πίνακα. Ισχύει : $\Delta_4 Y_t = Y_t - Y_{t-4}$ = η χρονοσειρά μετά την πρώτη εποχική διαφορά

Time (t)	Y_t	$\Delta_4 Y_t = Y_t - Y_{t-4}$
1	147.6	-
2	251.8	-
3	273.1	-
4	249.1	-
5	139.3	-8.3
6	221.2	-30.6
7	260.2	-12.9
8	259.5	10.4
9	140.5	1.2
...
52	281.4	-56.8

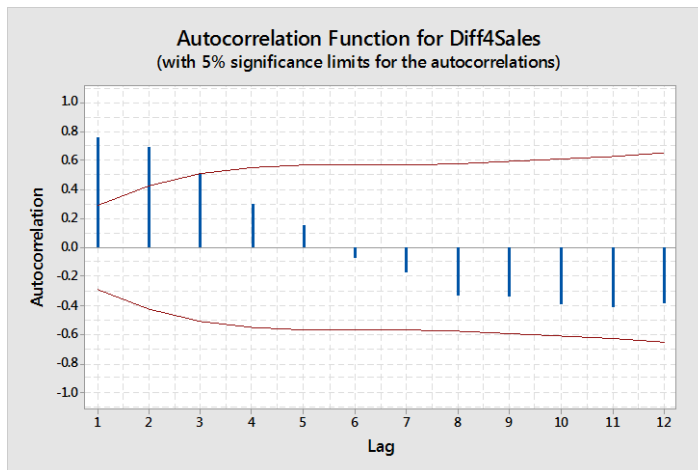
Πίνακας 4.5 : Εποχική διαφορά των δεδομένων της χρονοσειράς



Μετά την εποχική διαφορίση των δεδομένων, παριστάνουμε γραφικά τη νέα χρονοσειρά και υπολογίζουμε εκ νέου τη συνάρτηση αυτοσυσχέτισης.



Σχήμα 4.5 : Χρονοσειρά μετά την κανονική διαφορίση



Σχήμα 4.6 : Συνάρτηση αυτοσυσχέτισης διαφορισμένης χρονοσειράς

Η συμπεριφορά αυτή της συνάρτησης αυτοσυσχέτισης, υποδηλώνει ότι η χρονοσειρά δεν είναι σταθερή και παρουσιάζει κάποιο πρότυπο τάσης. Το πρότυπο τάσης, όπως και η εποχική διακύμανση, πρέπει να εξαλειφθεί από τη χρονοσειρά και αυτό επιτυγχάνεται μέσω της κανονικής διαφορίσης. Η κανονική διαφορίση, θα πραγματοποιηθεί στα ήδη εποχικά διαφορισμένα δεδομένα και τα αποτελέσματα – η νέα χρονοσειρά – που προκύπτει, φαίνονται στον παρακάτω πίνακα. Με την απαλοιφή των προτύπων τάσης και εποχικότητας από τα δεδομένα, οδηγούμαστε σε μια σταθερή χρονοσειρά στην οποία πρέπει να αντιστοιχίσουμε το κατάλληλο μοντέλο πρόβλεψης ARIMA.

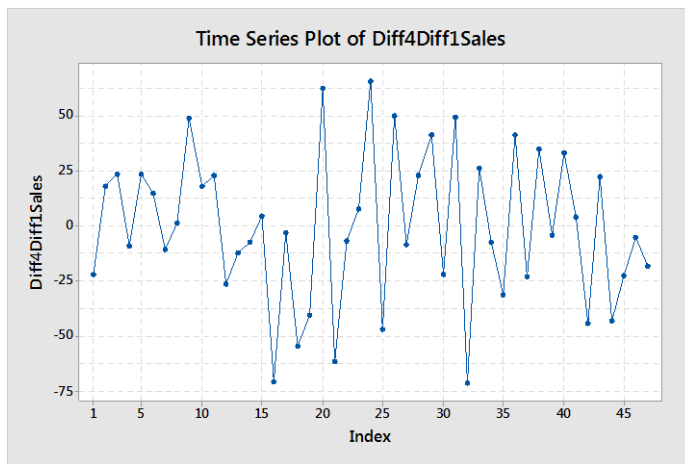
Από τα διπλανά διαγράμματα της εποχικά διαφορισμένης χρονοσειράς και της αντίστοιχης συνάρτησης αυτοσυσχέτισης μπορούμε να παρατηρήσουμε τα εξής. Το πρότυπο της εποχικότητας έχει εξαλειφθεί από τη χρονοσειρά. Η παρατήρηση αυτή, τεκμηριώνεται και στατιστικά από το γεγονός ότι ο συντελεστής αυτοσυσχέτισης για την εποχική χρονική καθυστέρηση των 4 περιόδων, δεν είναι πλέον σημαντικός.

Παρόλα αυτά όμως πρέπει να προσέξουμε ότι από το διάγραμμα αυτοσυσχετίσεων, προκύπτει ακόμη μια πληροφορία σχετικά με τη χρονοσειρά που μελετάμε. Οι τιμές του συντελεστή για τις πρώτες χρονικές καθυστερήσεις είναι στατιστικά σημαντικές και σταδιακά πέφτουν στο μηδέν.

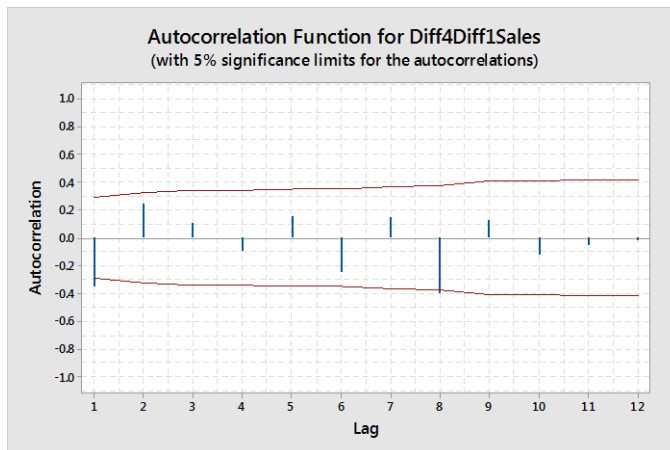


Time (t)	Y_t	$\Delta_4 Y_t = Y_t - Y_{t-4}$	$\Delta\Delta_4 Y_t = Y_t - Y_{t-1} - Y_{t-4} + Y_{t-5}$
1	147.6	-	-
2	251.8	-	-
3	273.1	-	-
4	249.1	-	-
5	139.3	-8.3	-
6	221.2	-30.6	-22.3
7	260.2	-12.9	17.7
8	259.5	10.4	23.3
9	140.5	1.2	-9.2
10	245.5	24.3	23.1
11	298.8	38.6	14.3
12	287	27.5	-11.1
...
52	281.4	-56.8	-18.2

Πίνακας 4.6 : Εποχική και Κανονική διαφόριση των δεδομένων



Σχήμα 4.7 : Χρονοσειρά μετά την κανονική και εποχική διαφόριση των δεδομένων



Σχήμα 4.8 : Συνάρτηση αυτοσυσχέτισης νέας χρονοσειράς

Μέσα από τη γραφική παράσταση της χρονοσειράς μετά και από την κανονική διαφόριση των δεδομένων, μπορούμε να συμπεράνουμε τα εξής σημαντικά. Αρχικά, παρατηρούμε ότι έχουν εξαλειφθεί τα πρότυπα τάσης και εποχικότητας που εμφανίζονταν στη χρονοσειρά και πλέον οι τιμές κυμαίνονται γύρω από μία σταθερή τιμή, εν προκειμένω, το μηδέν. Η γραφική αυτή παράσταση είναι αντιπροσωπευτική μιας σταθερής χρονοσειράς, για την οποία μένει πλέον να εντοπιστεί το κατάλληλο μοντέλο ARIMA, συγκρίνοντας την προκύπτουσα συνάρτηση αυτοσυσχέτισης με τις αντίστοιχες θεωρητικές.



4.2.3 Συντελεστής μερικής αυτοσυσχέτισης (partial autocorrelation function - PACF)

Οι συντελεστές μερικής αυτοσυσχέτισης μετρούν το βαθμό της σχέσης μεταξύ των Y_t και Y_{t-k} όταν οι επιδράσεις όλων των άλλων χρονικών υστερήσεων $1, 2, \dots, k-1$, έχουν αφαιρεθεί. Οι συντελεστές αυτοί και συγκεκριμένα τα διαγράμματα που παράγουν, χρησιμοποιούνται για την αναγνώριση του καταλληλότερου αυτοπαλίνδρομου μοντέλου κινητού μέσου, μέσα από τη σύγκρισή τους με τα θεωρητικά διαγράμματα μερικών αυτοσυσχετίσεων των μοντέλων ARIMA (Βλ. **Παράγραφο 4.5**). Ισχύει ότι η τάξη της διαδικασίας (p) ισούται με τον αριθμό των στατιστικά σημαντικών μερικών αυτοσυσχετίσεων όπως προκύπτει από τον υπολογισμό της σχετικής συνάρτησης.

Ο συντελεστής μερικής αυτοσυσχέτισης τάξης k συμβολίζεται με α_k και μπορεί να υπολογισθεί εφαρμόζοντας τη γνωστή μέθοδο πολλαπλής γραμμικής παλινδρόμησης με εξαρτημένη μεταβλητή την Y_t και ανεξάρτητες μεταβλητές $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k}$ ως εξής :

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \dots + \alpha_k Y_{t-k}$$

Οι συντελεστές μερικής αυτοσυσχέτισης αναλύονται εκτενώς από τους Newbold και Bos (1994) και από τους Box, Jenkins και Reinsel (1994) στα σχετικά συγγράμματά τους. Για τις ανάγκες της παρούσας εργασίας οι συναρτήσεις μερικής αυτοσυσχέτισης θα υπολογισθούν και θα χρησιμοποιηθούν για την αναγνώριση του κατάλληλου μοντέλου πρόβλεψης ARIMA.

4.3 Μοντέλα ARIMA

4.3.1 Αυτοπαλίνδρομα μοντέλα AR(p)

Ένα αυτοπαλίνδρομο (AutoRegressive) μοντέλο AR(p), p -τάξη αυτοπαλινδρόμησης ορίζεται ως :

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + e_t$$

Y_t : μεταβλητή απόκρισης της χρονοσειράς τη χρονική στιγμή t

$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$: μεταβλητές απόκρισης τις χρονικές στιγμές $t-1, t-2, \dots, t-p$

$\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_p$: οι παράμετροι του μοντέλου που πρέπει να εκτιμηθούν, όπου $-1 \leq \varphi_p \leq 1$

e_t : η τιμή του τυχαίου σφάλματος της μεταβλητής απόκρισης τη χρονική στιγμή t



Τα αυτοπαλινδρόμα μοντέλα είναι ιδανικά για στάσιμες χρονοσειρές. Εάν η προς μελέτη χρονοσειρά παρουσιάζει πρότυπο τάσης, το πρότυπο αυτό πρέπει να εξαλειφθεί για να μπορέσει να εφαρμοστεί το μοντέλο της αυτοπαλινδρόμησης. Για τη συνάρτηση αυτοσυσχέτισης που θα προκύψει μετά την αφαίρεση της τάσης από τα δεδομένα, (μέσω της μεθόδου κανονικής διαφόρισης που έχει περιγραφεί σε προηγούμενη παράγραφο) θα αναζητηθεί το αντίστοιχο θεωρητικό μοντέλο AR(p) που προσεγγίζει τη μορφή της. Γενικά χρησιμοποιούμε ένα AR(p) μοντέλο πρόβλεψης όταν οι συντελεστές αυτοσυσχέτισης φθίνουν εκθετικά στο μηδέν και ταυτόχρονα υπάρχουν p στατιστικά σημαντικοί συντελεστές μερικής αυτοσυσχέτισης.

Τα μοντέλα AR(p) παράγουν προβλέψεις Y_t ως γραμμική συνάρτηση των p παρελθοντικών πραγματικών τιμών απόκρισης της χρονοσειράς.

4.3.2 Μοντέλα κινητού μέσου όρου MA(q)

Ένα μοντέλο κινητού μέσου όρου (Moving Average) q-τάξης ορίζεται ως εξής :

$$Y_t = \mu + e_t - \omega_1 e_{t-1} - \omega_2 e_{t-2} - \dots - \omega_q e_{t-q}$$

Y_t : μεταβλητή απόκρισης της χρονοσειράς τη χρονική στιγμή t

μ : η σταθερή παράμετρος του μέσου επιπέδου της χρονοσειράς

$\omega_1, \omega_2, \dots, \omega_q$: οι παράμετροι του μοντέλου που πρέπει να εκτιμηθούν, όπου $-1 \leq \omega_p \leq 1$

$e_{t-1}, e_{t-2}, \dots, e_{t-q}$: οι τιμές των τυχαίων σφαλμάτων της μεταβλητής απόκρισης κατά τις προηγούμενες $t - 1, t - 2, \dots, t - q$ χρονικές στιγμές οι οποίες λαμβάνονται υπόψη για τον υπολογισμό της τιμής απόκρισης κατά το χρόνο t

Γενικά χρησιμοποιούμε ένα MA(q) μοντέλο πρόβλεψης όταν οι συντελεστές μερικής αυτοσυσχέτισης φθίνουν εκθετικά στο μηδέν και ταυτόχρονα υπάρχουν q στατιστικά σημαντικοί συντελεστές αυτοσυσχέτισης. Όπως και στην περίπτωση των μοντέλων αυτοπαλινδρόμησης, εάν η προς μελέτη χρονοσειρά παρουσιάζει πρότυπο τάσης, το πρότυπο αυτό πρέπει να εξαλειφθεί για να μπορέσει να εφαρμοστεί το μοντέλο κινητού μέσου όρου.



Στο σημείο αυτό πρέπει να γίνει σαφές ότι ο όρος *κινητός μέσος*, δεν πρέπει να συγχέεται με τα μοντέλα εξομάλυνσης κινητού μέσου που έχουν παρουσιαστεί σε προηγούμενη παράγραφο. Εδώ, ο όρος *κινητός μέσος*, αναφέρεται στο γεγονός ότι η απόκλιση της παρατηρούμενης τιμής από το μέσο επίπεδο της χρονοσειράς, είναι γραμμικός συνδυασμός των τρέχων και παρελθοντικών σφαλμάτων. Καθώς ο χρόνος προχωρά, τα σφάλματα που περιλαμβάνονται στο γραμμικό συνδυασμό κινούνται μπροστά, με αποτέλεσμα να διατηρείται σταθερός ο αριθμός των σφαλμάτων που συνυπολογίζονται στους υπολογισμούς των προβλέψεων. Ισχύει δηλαδή ότι :

$$Y_t - \mu = e_t - \omega_1 e_{t-1} - \omega_2 e_{t-2} - \dots - \omega_q e_{t-q}$$

$$Y_{t+1} - \mu = e_{t+1} - \omega_1 e_t - \omega_2 e_{t-1} - \dots - \omega_q e_{t-q+1}$$

Τα υποδείγματα *κινητού μέσου όρου MA(q)*, παράγουν προβλέψεις Y_t βασισμένες σε ένα γραμμικό συνδυασμό των παρελθοντικών σφαλμάτων πρόβλεψης.

4.3.3 Αυτοπαλίνδρομα μοντέλα κινητού μέσου όρου ARMA(p, q)

Ένα πιο γενικό μοντέλο πρόβλεψης είναι ο συνδυασμός ενός AR(p) και ενός MA(q) μοντέλου, το οποίο ονομάζεται αυτοπαλίνδρομο μοντέλο κινητού μέσου όρου, (**Auto**Regressive **M**oving **A**verage) συμβολίζεται ως ARMA(p, q) και ορίζεται ως εξής :

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + e_t - \omega_1 e_{t-1} - \omega_2 e_{t-2} - \dots - \omega_q e_{t-q}$$

Τα μοντέλα ARMA(p, q) μπορούν να περιγράψουν πληθώρα διαφορετικών μοτίβων διακύμανσης για στάσιμες χρονοσειρές. Όπως και σε κάθε μοντέλο πρόβλεψης στην ανάλυση ARIMA, αν η χρονοσειρά δεν είναι στάσιμη, πρέπει να εξαλειφθεί το πρότυπο τάσης για να εφαρμοστεί το μοντέλο. Για τον προσδιορισμό της τάξης αυτοπαλινδρόμησης (p) και κινητού μέσου όρου (q), απαιτείται παράλληλη προσεκτική μελέτη των συναρτήσεων αυτοσυσχέτισης και μερικής αυτοσυσχέτισης της χρονοσειράς, ενώ καθοριστικό ρόλο στη λήψη της τελικής απόφασης, διαδραματίζει και η σχετική εμπειρία του αναλυτή στη χρήση των μοντέλων πρόβλεψης ARIMA.

Οι προβλέψεις που παράγονται από τα μοντέλα αυτά, εξαρτώνται και από τις τρέχουσες και παρελθοντικές πραγματικές τιμές της μεταβλητής αλλά και από τις τρέχουσες και παρελθοντικές τιμές των σφαλμάτων πρόβλεψης.



Είναι σημαντικό να παρατηρήσουμε ότι καθώς αυξάνεται η τάξη των μοντέλων (p, q) αλλά και οι παράγοντες που εμπεριέχονται στους υπολογισμούς (παρελθούσες τιμές και σφάλματα), αυξάνεται ανάλογα και ο αριθμός των προς εκτίμηση παραμέτρων και άρα η πολυπλοκότητα του μοντέλου. Στις επιχειρήσεις, επικρατούσα είναι η **αρχή της απλότητας (principle of parsimony)**, σύμφωνα με την οποία προτιμάται ένα απλό, ευέλικτο και εύχρηστο μοντέλο πρόβλεψης παρά ένα πολύπλοκο. Κρίνεται λοιπόν σκόπιμο ο αναλυτής να καταλήξει στο απλούστερο μοντέλο πρόβλεψης, το οποίο στα πλαίσια των αναγκών της επιχείρησης, να εξάγει αποδεκτά αποτελέσματα όντας απλό στη χρήση του και κατανοητό στα στελέχη που θα το διαχειρίζονται.

4.3.4 Ολοκληρωμένα αυτοπαλίνδρομα μοντέλα κινητού μέσου $ARIMA(p,d,q)$

Σύμφωνα με το όσα παρουσιάστηκαν και στα προηγούμενα μοντέλα $ARIMA$, όταν μία χρονοσειρά είναι μη σταθερή, τότε πριν εφαρμοστεί ένα μοντέλο $ARMA(p, q)$ πρέπει να επιτευχθεί σταθερότητα με τη χρήση της μεθόδου διαφορίσης (κανονική διαφορίση για πρότυπα τάσης). Ο όρος I (integrated) στα μοντέλα $ARIMA$, αναφέρεται στην κανονική διαφορίση των δεδομένων της χρονοσειράς.

Συνεπώς σε μια σταθερή χρονοσειρά, εφαρμόζουμε ένα μοντέλο πρόβλεψης $ARIMA(p, d, q)$ όπου:

- p είναι η τάξη της αυτοπαλινδρόμησης
- d η τάξη της διαφορίσης για την επίτευξη σταθερότητας και
- q η τάξη του κινητού μέσου όρου

4.3.5 Ολοκληρωμένα αυτοπαλίνδρομα μοντέλα κινητού μέσου με εποχικότητα $ARIMA(p, d, q) \times (P, D, Q)_s$

Όταν μία χρονοσειρά παρουσιάζει εποχική διακύμανση η οποία επαναλαμβάνεται μετά από κάθε s χρονικές περιόδους ($s=12$ για μηνιαία δεδομένα και $s=4$ για τριμηνιαία δεδομένα), μπορεί να χρησιμοποιηθεί ένας συνδυασμός ενός εποχικού και ενός μη-εποχικού μοντέλου $ARIMA$.

Ένας πολύ χρήσιμος συμβολισμός που περιγράφει τις τάξεις των διαφόρων συνιστωσών στο πολλαπλασιαστικό μοντέλο είναι ο $(p, d, q) \times (P, D, Q)_s$ όπου :

- ♦ p, d, q αναφέρονται στο μη εποχικό μέρος της χρονοσειράς
- ♦ P, D, Q αναφέρονται στο εποχικό μέρος της χρονοσειράς.



Τα εποχικά μοντέλα πρόβλεψης $(P, D, Q)_s$ προσδιορίζονται για χρονοσειρές που παρουσιάζουν εποχικότητα στις παρατηρούμενες τιμές, με τρόπο παρόμοιο της αρχικής διαδικασίας ARIMA. Ωστόσο, αυτό που πρέπει να γίνει κατανοητό είναι ότι οι όροι κανονικής αυτοπαλινδρόμησης και κινητού μέσου όρου (p, q) του μοντέλου, χρησιμοποιούνται λαμβάνοντας υπόψη τις αυτοσυσχετίσεις που παρατηρούνται για τις πρώτες χρονικές καθυστερήσεις, ενώ οι όροι της εποχικής αυτοπαλινδρόμησης και κινητού μέσου όρου (P, Q) χρησιμοποιούνται με γνώμονα τις αυτοσυσχετίσεις που παρατηρούνται στις εποχικές χρονικές καθυστερήσεις. Για την αφαίρεση της εποχικότητας από τα δεδομένα, χρησιμοποιείται η μέθοδος της εποχικής διαφορίσης, η τάξη της οποίας προσδιορίζει και την τιμή (D) του τελικού μοντέλου.

Σε ένα πολλαπλασιαστικό υπόδειγμα ARIMA οι κανονικές και οι εποχικές συνιστώσες αυτοπαλινδρόμησης και οι κανονικές και εποχικές συνιστώσες κινητών μέσων όρων, πολλαπλασιάζονται μαζί στο γενικό μοντέλο. Συχνά, στον τελικό προσδιορισμό του μοντέλου οι περισσότερες από τις τιμές p, d, q ή P, D, Q είναι 0 ή 1 και τα μοντέλα που λαμβάνονται είναι σχετικά απλά, ακολουθώντας έτσι τη βασική αρχή της απλότητας των μοντέλων.

Παράδειγμα

Θεωρούμε ένα μηνιαίο εποχικό μοντέλο, στο οποίο τα δεδομένα της χρονοσειράς έχουν υποστεί εποχική διαφορίση και εφαρμόζεται κανονική και εποχική αυτοπαλινδρόμηση πρώτης τάξης :

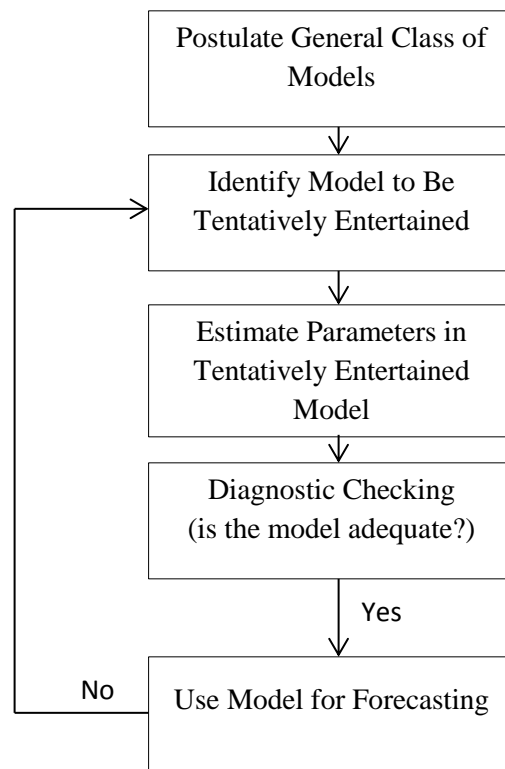
$$(p, d, q)x(P, D, Q)_{12} = (1, 0, 0)X(1, 1, 0)_{12}$$

το οποίο σε όρους του Y_t ορίζεται ως :

$$Y_t = \varphi Y_{t-1} + (1 + \varphi_{12})Y_{t-12} - \varphi_1(1 + \varphi_{12})Y_{t-13} - \varphi_{12}Y_{t-24} + \varphi_1\varphi_2Y_{t-25} + e_t$$

4.4 Μεθοδολογία Box – Jenkins

- i. Ταυτοποίηση μοντέλου (Model Identification)
- ii. Εκτίμηση των παραμέτρων του μοντέλου (Model Estimation)
- iii. Διαγνωστικός έλεγχος του μοντέλου (Model Checking)
- iv. Πρόβλεψη χρονοσειράς με χρήση του μοντέλου (Forecasting with the Model)



Source : Box, G.E.P.,
Jenkins, G.M., and G.C.
Reinsel, *Time Series
Analysis : Forecasting and
Control (3rd ed.)*

Βήμα 1^ο : Ταυτοποίηση του μοντέλου

Το πρώτο βήμα της διαδικασίας ταυτοποίησης του μοντέλου, είναι η γραφική παράσταση της αρχικής χρονοσειράς καθώς και των διαγραμμάτων αυτοσυσχέτισης και μερικής αυτοσυσχέτισης. Στόχος είναι μέσα από την ερμηνεία των διαγραμμάτων, να διαφανεί κατά πόσο η χρονοσειρά μπορεί να θεωρηθεί στάσιμη, τυχαία ή κατά πόσο εμφανίζει κάποιο εποχικό πρότυπο ή πρότυπο τάσης. Αρχικά, με τον υπολογισμό και τη γραφική παράσταση των συντελεστών αυτοσυσχέτισης, εξετάζουμε εάν οι πραγματικές τιμές της χρονοσειράς τείνουν στο μηδέν μετά τη δεύτερη ή τρίτη χρονική υστέρηση. Αν πράγματι συμβαίνει αυτό, η χρονοσειρά θεωρείται σταθερή στην αρχική της μορφή και μπορεί να αναζητηθεί το θεωρητικό μοντέλο ARIMA. Εάν η χρονοσειρά δεν είναι σταθερή τότε διαφορίζεται μέχρι να αποκτήσει σταθερότητα και στη συνέχεια παράγονται τα νέα διαγράμματα αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων.

Για την προς μελέτη μεταβλητή εξετάζουμε τη συνάρτηση αυτοσυσχέτισης με στόχο την αναγνώριση τυχόν εποχικού προτύπου. Αν διαπιστωθεί ότι η χρονοσειρά ακολουθεί εποχικό πρότυπο, εφαρμόζεται η μέθοδος της εποχικής διαφορίσεως για την εξάλειψη του προτύπου. Ενδέχεται, σε μια χρονοσειρά να



χρειαστεί εποχική διαφόριση των δεδομένων για την εξάλειψη του εποχιακού προτύπου, η οποία να ακολουθείται από κανονική διαφόριση για την εξάλειψη του προτύπου τάσης.

Η ταυτοποίηση του μοντέλου είναι δυνατό να επιτευχθεί συγκρίνοντας τα διαγράμματα αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων των ακατέργαστων ή διαφορισμένων δεδομένων με τα θεωρητικά (βλ. **Παράγραφο 4.5**). Κρίσιμη, στο σημείο αυτό, αποδεικνύεται η εμπειρία, η γνώση και η κριτική σκέψη του αναλυτή για την επιλογή του κατάλληλου μοντέλου ARIMA. Το αποτέλεσμα του πρώτου βήματος της μεθοδολογίας, είναι η δημιουργία ενός πρώτου δοκιμαστικού μοντέλου ARIMA.

Βήμα 2^ο : Εκτίμηση των παραμέτρων του μοντέλου

Η εκτίμηση του μοντέλου περιλαμβάνει την εκτίμηση των παραμέτρων $\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_p$ ή/και $\omega_1, \omega_2, \dots, \omega_q$ που ορίζουν το εκάστοτε μοντέλο ARIMA. Κριτήριο για τον τελικό προσδιορισμό των τιμών των παραμέτρων, είναι η ελαχιστοποίηση του αθροίσματος των τετραγώνων των σφαλμάτων.

Για την εκτίμηση των παραμέτρων συνήθως χρησιμοποιούνται μη γραμμικοί επαναληπτικοί αλγόριθμοι ελαχίστων τετραγώνων, οι οποίοι είναι ενσωματωμένοι στα ειδικά υπολογιστικά εργαλεία ανάλυσης χρονοσειρών. Η διαδικασία επαναλαμβάνεται μέχρι το σημείο όπου οι τελικές τιμές των εκτιμήσεων των παραμέτρων είναι σημαντικά διαφορετικές από τις αρχικές τιμές της διαδικασίας εκτίμησης. Συνήθως οι εκτιμήσεις συγκλίνουν σε μία άριστη τιμή για τις παραμέτρους μετά από σχετικά μικρό αριθμό επαναλήψεων.

Το υπολειπόμενο μέσο τετραγωνικό σφάλμα (residual mean square error) ορίζεται ως :

$$s^2 = \frac{\sum_{t=1}^n e_t^2}{n - r} = \frac{\sum_{t=1}^n (Y_t - \bar{Y}_t)^2}{n - r}$$

$e_t = Y_t - \bar{Y}_t$: το υπολειπόμενο σφάλμα κατά τη χρονική περίοδο t

n : ο αριθμός των σφαλμάτων

r : ο συνολικός αριθμός των εκτιμώμενων παραμέτρων του μοντέλου

Το υπολειπόμενο μέσο τετραγωνικό σφάλμα, είναι χρήσιμο για τη σύγκριση των υποψηφίων μοντέλων πρόβλεψης ARIMA και χρησιμοποιείται επίσης για τον υπολογισμό των ορίων (άνω και κάτω) των σφαλμάτων πρόβλεψης στο διάγραμμα ελέγχου.



Βήμα 3^ο : Διαγνωστικός έλεγχος του μοντέλου

Ο διαγνωστικός έλεγχος πραγματοποιείται για τον έλεγχο της καταλληλότητας και της επάρκειας του μοντέλου. Για να ελέγξουμε την καταλληλότητα του μοντέλου κατασκευάζουμε το διάγραμμα αυτοσυσχετίσεων των σφαλμάτων πρόβλεψης της χρονοσειράς (residuals). Εάν οι συντελεστές αυτοσυσχετίσης δεν είναι στατιστικά σημαντικοί, τότε το υπόδειγμα κρίνεται επαρκές. Τα άνω και κάτω όρια, εντός των οποίων τα σφάλματα κρίνονται στατιστικά σημαντικά ορίζονται από τη σχέση $\pm \frac{2}{\sqrt{n}}$, όπου n είναι ο αριθμός των παρατηρήσεων της χρονοσειράς.

Ένας άλλος τρόπος που χρησιμοποιείται για να ελέγξουμε την επάρκεια του χρησιμοποιούμενου μοντέλου πρόβλεψης, είναι να εξετάσουμε διαγνωστικά τις αυτοσυσχετίσεις των υπολοίπων με το κριτήριο ελέγχου (Ljung – Box Q statistic) :

$$Q = n(n + 2) \sum_{k=1}^m \frac{r_k^2(e)}{n - k}$$

το οποίο κατά προσέγγιση, κατανέμεται σύμφωνα με την κατανομή χ^2 με $m - r$ βαθμούς ελευθερίας. Στην παραπάνω εξίσωση :

- $r_k(e)$ ο συντελεστής αυτοσυσχετίσης του υπολοίπου για χρονική καθυστέρηση k
- m ο αριθμός των χρονικών καθυστερήσεων που θα ελεγχθούν
- k η χρονική καθυστέρηση
- r ο αριθμός των παραμέτρων που εκτιμώνται για το υπό μελέτη μοντέλο

Σύμφωνα με τη θεωρία της στατιστικής (hypothesis testing) ισχύουν τα εξής :

$$H_0 = \text{το μοντέλο είναι επαρκές}$$

$$H_A = \text{το μοντέλο είναι ανεπαρκές}$$

Εάν η υπολογισθείσα τιμή του Q είναι μεγαλύτερη από την τιμή $\chi^2(a, m - r)$, τότε απορρίπτεται η μηδενική υπόθεση και το μοντέλο θεωρείται ανεπαρκές. Η τιμή $\chi^2(a, m - r)$, προκύπτει από τον πίνακα τιμών της κατανομής για $m - r$ βαθμούς ελευθερίας.



Εναλλακτικά, εάν η υπολογιζόμενη τιμή p -value σε σχέση με την υπολογισθείσα τιμή του Q είναι μικρή, (π.χ. p – value < 0.05) τότε δεν μπορούμε να δεχθούμε τη μηδενική υπόθεση και καταλήγουμε στο συμπέρασμα ότι το μοντέλο θεωρείται στατιστικά ανεπαρκές.

Σε μια τέτοια περίπτωση, ο αναλυτής θα πρέπει επαναλάβει τη μεθοδολογία Box-Jenkins μέχρι να βρεθεί ένα ικανοποιητικό μοντέλο. Αν δύο ή περισσότερα μοντέλα θεωρούνται επαρκή και εξάγουν περίπου ίδια αποτελέσματα, μολονότι κανένα μοντέλο δεν εμφανίζει ακριβή προσαρμογή στα πραγματικά δεδομένα, κατά την επιλογή πρέπει να επικρατήσει η ιδιότητα της απλότητας.

Βήμα 4^ο : Πρόβλεψη με χρήση του μοντέλου

Μετά την ταυτοποίηση, την εκτίμηση των παραμέτρων και τον έλεγχο επάρκειας του επιλεχθέντος μοντέλου, μπορούν να γίνουν προβλέψεις μιας ή περισσότερων χρονικών περιόδων στο μέλλον. Πρέπει να σημειωθεί ότι καθώς η περίοδος πρόβλεψης αυξάνεται, αυξάνονται επίσης οι πιθανότητες να εμφανιστούν σημαντικά αποκλίσεις των προβλέψεων, με τις εκ των υστέρων, πραγματικά εκδηλωμένες τιμές της μεταβλητής που εξετάζεται..

Με την απόκτηση νέων παρατηρήσεων για μια χρονοσειρά, το χρησιμοποιούμενο μοντέλο πρόβλεψης πρέπει να επανεξετάζεται και να ελέγχεται ως προς την ακρίβεια. Ο έλεγχος, επιτυγχάνεται με τον υπολογισμό, παρακολούθηση και ερμηνεία των σφαλμάτων πρόβλεψης που προκύπτουν. Εάν η χρονοσειρά φαίνεται ότι μεταβάλλεται κατά τη διάρκεια του χρόνου είτε όσον αφορά την τάση, είτε την εποχικότητα, οι παράμετροι του μοντέλου πρέπει να επαναυπολογιστούν ή και σε ορισμένες περιπτώσεις, θα πρέπει να αναπτυχθεί ένα εντελώς νέο μοντέλο. Σε μια τέτοια περίπτωση, θα πρέπει να επιστρέψουμε στο πρώτο στάδιο της διαδικασίας Box- Jenkins.

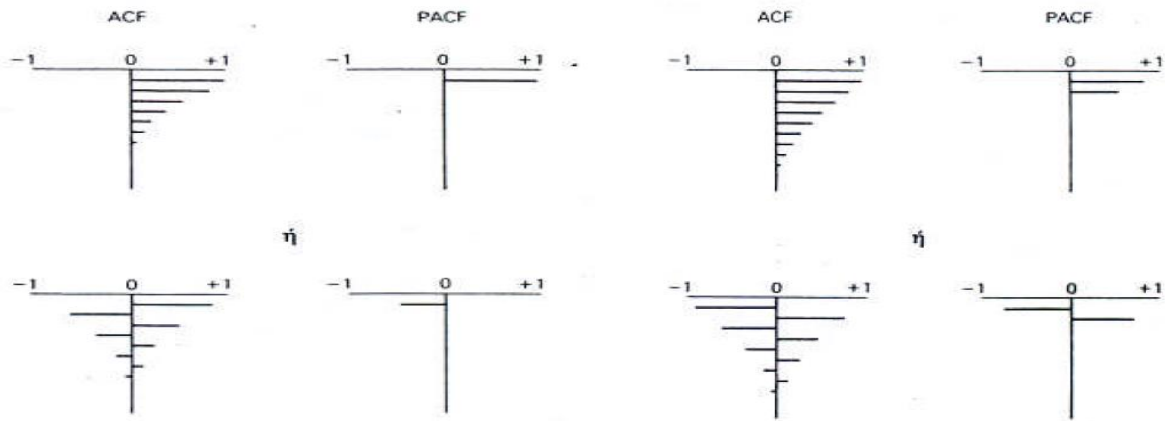
Σχόλιο

Η χρησιμοποίηση μοντέλων ARIMA για τη διενέργεια προβλέψεων, είναι μια διαδικασία για την οποία οι αναλυτές πρέπει να επενδύσουν, συγκριτικά με τις απλές μεθόδους εξομάλυνσης, πολύ περισσότερο χρόνο και υπολογιστική ισχύ. Αναμφισβήτητα, η ακρίβεια των προβλέψεων, εφόσον καταλήξει κανείς στο κατάλληλο μοντέλο ARIMA, θα είναι επίσης συγκριτικά πολύ μεγαλύτερη. Εναπόκειται στις ανάγκες της κάθε επιχείρησης και στην κατάρτιση του προσωπικού, για το ποια μέθοδος θα υιοθετηθεί, αντισταθμίζοντας τα οφέλη - κόστη που προκύπτουν.



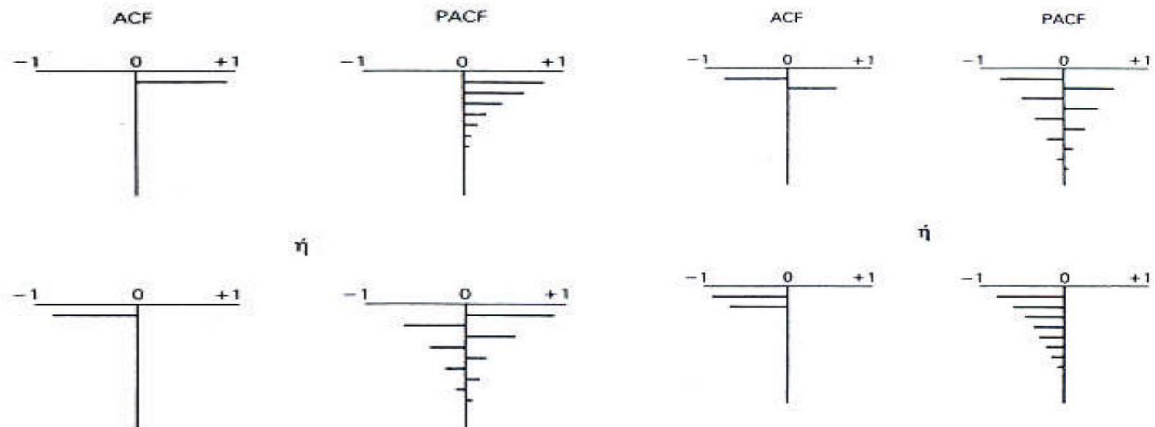
4.5 Θεωρητικές κατανομές συντελεστών αυτοσυσχέτισης και μερικής αυτοσυσχέτισης

Η ταυτοποίηση του μοντέλου πρόβλεψης που είναι κατάλληλο για την εξεταζόμενη μεταβλητή, είναι δυνατό να επιτευχθεί συγκρίνοντας τις κατανομές των συντελεστών αυτοσυσχέτισης και μερικής αυτοσυσχέτισης των παρατηρήσεων της χρονοσειράς, με τις αντίστοιχες θεωρητικές. Εάν η αρχική χρονοσειρά παρουσιάζει πρότυπα τάσης ή/και εποχικότητας, πρέπει να αφαιρεθούν προκειμένου να επιτευχθεί στάσιμη χρονοσειρά. Σε κάθε περίπτωση, οι συντελεστές αυτοσυσχέτισης στάσιμης χρονοσειράς θα μας οδηγήσουν στην επιλογή του κατάλληλου μοντέλου για πρόβλεψη, αφού κατά πάσα πιθανότητα η κατανομή τους, θα προσεγγίζει μία από τις παρακάτω θεωρητικές κατανομές. Στη συνέχεια απεικονίζονται οι θεωρητικές κατανομές των συντελεστών που αντιστοιχούν σε βασικά μοντέλα ARIMA 1^{ης} και 2^{ης} τάξης.



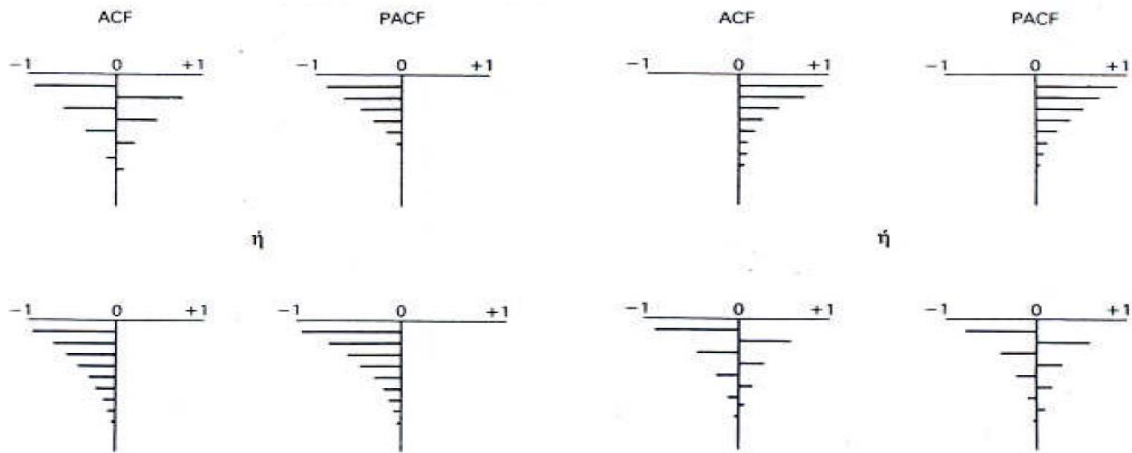
Μοντέλο AR(1)

Μοντέλο AR(2)



Μοντέλο MA(1)

Μοντέλο MA(2)



Μοντέλο ARMA(1,1)

4.6 Πρόσθετα κριτήρια επιλογής μοντέλου ARIMA

Τα μοντέλα ARIMA, όπως περιγράφεται στη μεθοδολογία των Box-Jenkins, ταυτοποιούνται μέσα από την παρατήρηση του μοτίβου διακύμανσης της χρονοσειράς και την αντιστοίχιση των κατανομών αυτοσυσχέτισης και μερικής αυτοσυσχέτισης με τις γνωστές, θεωρητικές κατανομές των προς επιλογή μοντέλων. Εντούτοις, σε περίπτωση που δύο ή περισσότερα μοντέλα φαίνεται να είναι κατάλληλα και επαρκή για την αναπαράσταση και πρόβλεψη της χρονοσειράς, η διαδικασία αυτή παρουσιάζει στοιχειά υποκειμενικότητας, αφού η τελική επιλογή εναπόκειται σε μεγάλο βαθμό στην εμπειρία και στις αναλυτικές ικανότητες του μελετητή.

Στην ανάλυση και πρόβλεψη των χρονοσειρών, ισχύουν δυο βασικοί κανόνες :

- ♦ Αν δυο μοντέλα, που φαίνεται να είναι επαρκή, περιέχουν τον ίδιο αριθμό παραμέτρων προς εκτίμηση, επιλέγεται αυτό που παρουσιάζει το μικρότερο μέσο τετραγωνικό σφάλμα.
- ♦ Στην περίπτωση που δύο μοντέλα ARIMA είναι επαρκή, αλλά έχουν διαφορετικό αριθμό παραμέτρων προς εκτίμηση, επιλέγεται το απλούστερο μοντέλο, δηλαδή εκείνο με τις λιγότερες παραμέτρους προς εκτίμηση.

Παρόλα αυτά όμως, ιδιαίτερο ενδιαφέρον παρουσιάζει η περίπτωση που το μοντέλο με τις περισσότερες προς εκτίμηση παραμέτρους, εμφανίζει κατά πολύ μικρότερο μέσο τετραγωνικό σφάλμα. Μια προσέγγιση για τη σύγκριση των υποψήφιων μοντέλων ARIMA, η οποία λαμβάνει υπόψη αμφότερα και τον αριθμό των προς εκτίμηση παραμέτρων αλλά και το μέσο τετραγωνικό σφάλμα, πραγματοποιείται με τη χρήση των κριτηρίων επιλογής *AIC* – Akaike (1974) και *BIC* – Schwarz (1978).



Σύμφωνα με το κριτήριο AIC , επιλέγεται ως καταλληλότερο, το μοντέλο που ελαχιστοποιεί την τιμή που προκύπτει από την παρακάτω σχέση :

$$AIC = \ln \hat{\sigma}^2 + \frac{2}{n}r$$

$\ln \hat{\sigma}^2$: ο φυσικός λογάριθμος του τετραγώνου του αθροίσματος των σφαλμάτων διαιρεμένου με το συνολικό αριθμό των παρατηρήσεων της χρονοσειράς

n : ο αριθμός των παρατηρήσεων

r : ο συνολικός αριθμός των παραμέτρων προς εκτίμηση του υποψηφίου μοντέλου $ARIMA$, συμπεριλαμβανομένου του σταθερού όρου εφόσον θα συνυπολογιστεί.

Το Bayesian Information Criterion, BIC , που αναπτύχθηκε από τον Schwartz, θεωρεί ως καταλληλότερο μοντέλο $ARIMA$, εκείνο το οποίο ελαχιστοποιεί την τιμή που προκύπτει από την παρακάτω εξίσωση :

$$BIC = \ln \hat{\sigma}^2 + \frac{\ln n}{n}r$$

Ο δεύτερος όρος στις εξισώσεις των κριτηρίων AIC , BIC είναι ουσιαστικά εκείνος που «τιμωρεί» την επιπρόσθετο προς εκτίμηση παράμετρο του υποψηφίου μοντέλου $ARIMA$. Το κριτήριο BIC τιμωρεί σε μεγαλύτερο βαθμό τον αριθμό των παραμέτρων από ότι το κριτήριο AIC . Το γεγονός αυτό μεταφράζεται στο ότι η χρήση του κριτηρίου BIC σε αντιπαράβολή με τη χρήση του μοντέλου AIC για την επιλογή του καταλληλότερου μοντέλου, θα έχει ως αποτέλεσμα το μοντέλο που θα επιλεγεί, εν γένει, να έχει μικρότερο αριθμό προς εκτίμηση παραμέτρων.

Τα κριτήρια AIC και BIC , ενδέχεται ορισμένες φορές να παρουσιάζουν τα ίδια αποτελέσματα ως προς την επιλογή του καταλληλότερου μοντέλου $ARIMA$. Αυτό που είναι κρίσιμο να αναφερθεί, είναι ότι τα κριτήρια πρέπει να αντιμετωπίζονται και να χρησιμοποιούνται ως πρόσθετα εργαλεία υποβοήθησης στη λήψη αποφάσεων για την επιλογή των μοντέλων πρόβλεψης. Σε καμία περίπτωση δεν πρέπει να αντικαθιστούν την προσεκτική και εις βάθος μελέτη της χρονοσειράς και συγκεκριμένα των κατανομών των συντελεστών αυτοσυσχέτισης και μερικής αυτοσυσχέτισης, αφού στη μελέτη αυτή εναπόκειται σε μεγάλο βαθμό η επιτυχής επιλογή του μοντέλου.



5. Μελέτη Περίπτωσης

5.1 Μοντέλο Πρόβλεψης ARIMA

Στα πλαίσια της διπλωματικής εργασίας πραγματοποιήθηκε μελέτη και ανάπτυξη μοντέλου πρόβλεψης ARIMA για τη ζήτηση του προϊόντος παραγωγικής επιχείρησης που δραστηριοποιείται στον κλάδο των ποτών. Η ανάγκη της επιχείρησης για επίτευξη υψηλότερης ακρίβειας προβλέψεων πηγάζει μέσα από την αλληλένδετη ανάγκη για πιο αποτελεσματική διαχείριση των αποθεμάτων της. Το γεγονός ότι μέχρι σήμερα χρησιμοποιούνταν απλές μέθοδοι εξομάλυνσης με περιορισμένη δυνατότητα επίτευξης υψηλής ακρίβειας προβλέψεων, είχε ως αποτέλεσμα να τηρούνται υψηλά επίπεδα αποθεμάτων για την κάλυψη των μεγάλων αποκλίσεων που προέκυπταν μεταξύ των προβλέψεων και της πραγματικής ζήτησης.

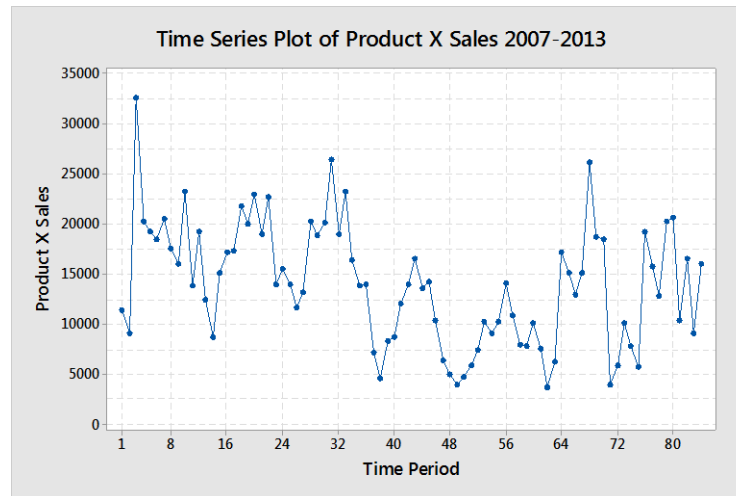
Τα δεδομένα που δόθηκαν αφορούν την πραγματική ζήτηση (πωλήσεις) ανά μήνα για το συγκεκριμένο προϊόν και εκτείνονται χρονικά από το 2007 έως το 2014. Στη συνέχεια παρουσιάζεται η μεθοδολογία ανάπτυξης του μοντέλου. Στο τέλος, αξιολογούνται τα αποτελέσματα που προκύπτουν συγκρίνοντας τις προβλέψεις που παρήγαγαν τα μοντέλα, με την πραγματική ζήτηση του προϊόντος για το έτος 2014.

Βήμα 1^ο : Ταυτοποίηση του μοντέλου

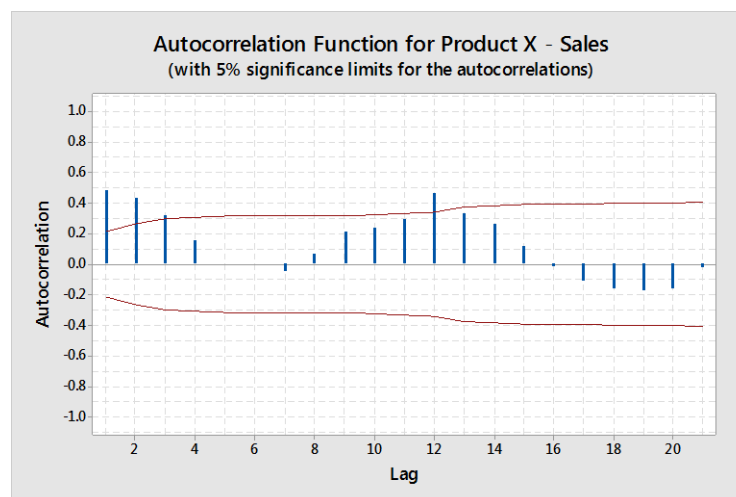
Το πρώτο βήμα της διαδικασίας ταυτοποίησης του μοντέλου είναι η γραφική παράσταση των δεδομένων της αρχικής χρονοσειράς (βλ. **Παράρτημα Α**) και ο υπολογισμός των συναρτήσεων αυτοσυσχέτισης και μερικής αυτοσυσχέτισης. Στόχος, είναι μέσα από τη μελέτη των διαγραμμάτων να διαφανεί κατά πόσο η χρονοσειρά πωλήσεων στην αρχική της μορφή παρουσιάζει πρότυπα τάσης ή/και εποχικότητας, είναι σταθερή ή όχι και τα δεδομένα της είναι τυχαία ή μη.

Υπενθυμίζεται ότι στην περίπτωση που εμφανίζεται πρότυπο τάσης στη χρονοσειρά, το πρότυπο αυτό πρέπει να εξαλειφθεί μέσω της μεθόδου κανονικής διαφόρισης, αφού προϋπόθεση εφαρμογής των μοντέλων πρόβλεψης ARIMA είναι η σταθερή διακύμανση της χρονοσειράς. Ομοίως, με την εποχική διαφόριση των δεδομένων, εξαλείφεται τυχόν εποχικό πρότυπο που εμφανίζεται στη χρονοσειρά.

Στο διάγραμμα που ακολουθεί απεικονίζεται η χρονοσειρά με τις πωλήσεις του προϊόντος από το 2007 έως το 2013 και η αντίστοιχη συνάρτηση αυτοσυσχέτισης που προκύπτει για τις συσχετίσεις των χρησιμοποιούμενων δεδομένων. Οι πωλήσεις για το έτος 2014, θα χρησιμοποιηθούν για την αξιολόγηση των προβλέψεων του μοντέλου ARIMA που θα αναπτυχθεί.



Σχήμα 5.1.1 : Χρονοσειρά Πωλήσεων



Σχήμα 5.1.2 : Συνάρτηση αυτοσυσχέτισης της χρονοσειράς

Παρατηρώντας τα διαγράμματα της χρονοσειράς και της συνάρτησης αυτοσυσχέτισης (οι τιμές των συντελεστών υπολογίστηκαν με τη σχετική μεθοδολογία που παρουσιάστηκε σε προηγούμενη παράγραφο) καταλήγουμε στα εξής σημαντικά συμπεράσματα :

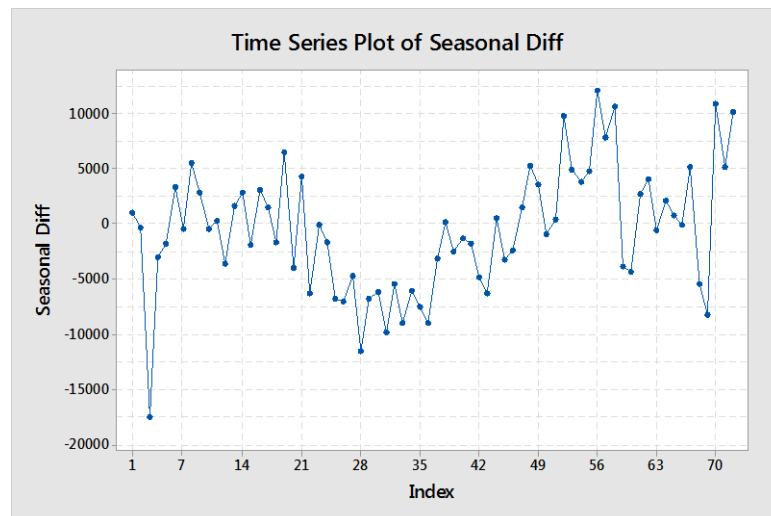
- Η χρονοσειρά παρουσιάζει εποχικότητα, αφού επαναλαμβάνει το μοτίβο διακύμανσής της μετά από κάθε χρονική παρέλευση 12 περιόδων (μηνών)
- Η ύπαρξη του προτύπου εποχικότητας, τεκμηριώνεται και μέσω της συνάρτησης αυτοσυσχέτισης από το γεγονός ότι σημειώνεται στατιστικά σημαντικός συντελεστής αυτοσυσχέτισης στην εποχική χρονική καθυστέρηση των 12 περιόδων.



- Από τη συνάρτηση αυτοσυσχέτισης υποδηλώνεται εν μέρει και η ενδεχόμενη ύπαρξη προτύπου τάσης στα δεδομένα, αφού οι συντελεστές αυτοσυσχέτισης είναι στατιστικά σημαντικοί για τις πρώτες χρονικές καθυστερήσεις και σταδιακά πέφτουν στο μηδέν.
- Τα όρια, πέρα από τα οποία οι συντελεστές αυτοσυσχέτισης θεωρούνται στατιστικά σημαντικοί (διάφοροι του μηδενός), προσδιορίζονται από : $\pm \frac{2}{\sqrt{n}} = \pm \frac{2}{\sqrt{84}} = \pm 0.218$

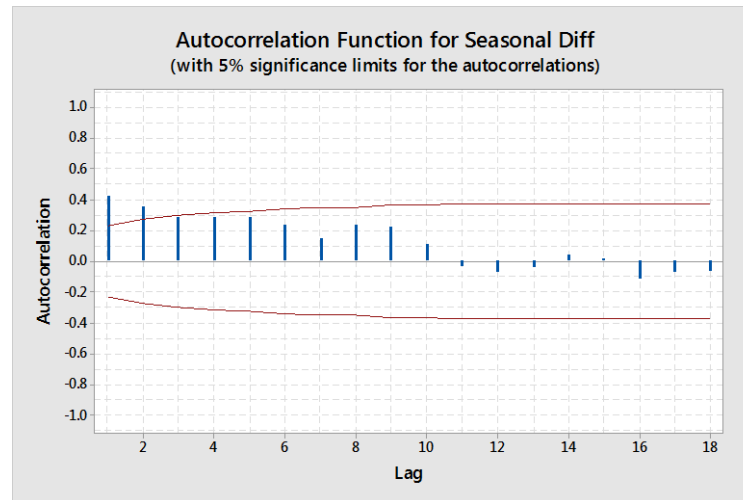
Εντοπίζοντας επομένως την ύπαρξη προτύπων τάσης και εποχικότητας στα δεδομένα, ο αναλυτής, αποκτά μια πρώτη άποψη για τους παράγοντες που διαμορφώνουν τη χρονική εξέλιξη των παρατηρήσεων της χρονοσειράς. Για την περαιτέρω ανάλυση της χρονοσειράς ώστε να βρεθεί το κατάλληλο μοντέλο ARIMA, πρέπει να αφαιρεθούν τα πρότυπα τάσης και εποχικότητας από τα δεδομένα. Η αφαίρεση των εμφανιζόμενων προτύπων με σκοπό τη δημιουργία σταθερής χρονοσειράς πραγματοποιείται με τις μεθόδους κανονικής και εποχικής διαφορίσης.

Τα αποτελέσματα μετά την **εποχική διαφορίση** των δεδομένων παρουσιάζονται στο **Παράρτημα Α**. Στη συνέχεια παρίστανται γραφικά η νέα χρονοσειρά και η συνάρτηση αυτοσυσχέτισης.



Σχήμα 5.1.3 : Χρονοσειρά μετά την εποχική διαφορίση

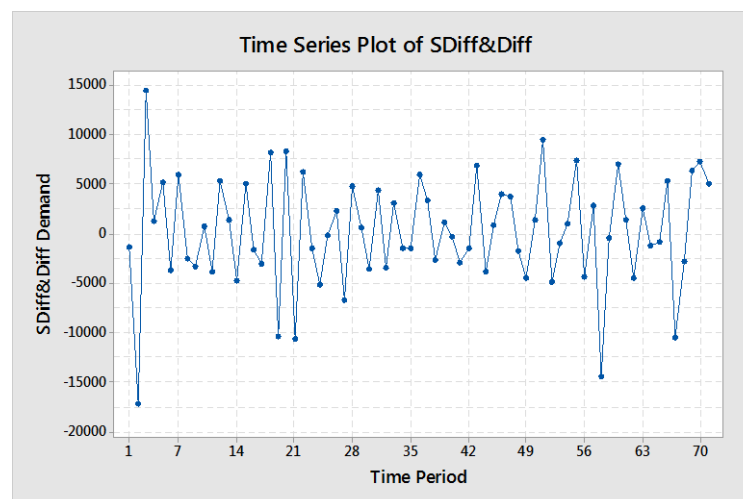
Στη νέα χρονοσειρά που προέκυψε παρατηρούμε ότι έχει εξαλειφθεί το πρότυπο εποχικότητας. Η παρατήρηση αυτή, τεκμηριώνεται και στατιστικά μέσω της συνάρτησης αυτοσυσχέτισης, αφού ο συντελεστής αυτοσυσχέτισης για τη χρονική καθυστέρηση των 12 περιόδων δεν είναι στατιστικά διάφορος του μηδενός.



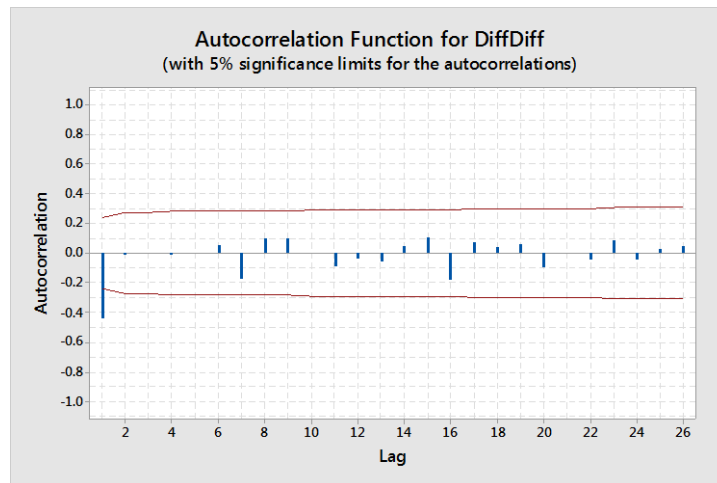
Σχήμα 5.1.4 : Συνάρτηση αυτοσυσχέτισης μετά την εποχική διαφύριση

Παρόλα αυτά όμως, η συνάρτηση αυτοσυσχέτισης, παρουσιάζει στατιστικά σημαντικούς συντελεστές στις πρώτες χρονικές καθυστερήσεις και σταδιακά οι συντελεστές αυτοί πλησιάζουν το μηδέν. Η συμπεριφορά αυτή, παραπέμπει σε μια μη σταθερή χρονοσειρά, η οποία παρουσιάζει πρότυπο τάσης. Όπως λοιπόν και στην περίπτωση του εποχικού προτύπου, το πρότυπο τάσης εξαλείφεται μέσω της διαφύρισης (κανονικής) των δεδομένων.

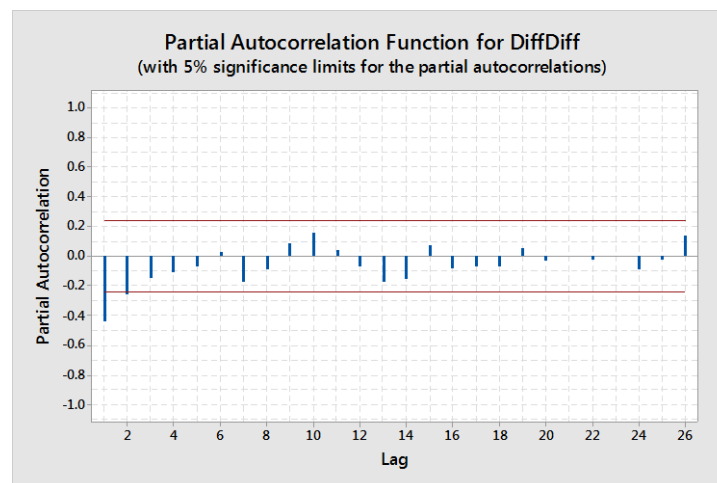
Να σημειωθεί, ότι τα δεδομένα που χρησιμοποιούνται για τον υπολογισμό των αυτοσυσχετίσεων, είναι επίσης αυτά που προκύπτουν μετά τη διαφύριση των δεδομένων της αρχικής χρονοσειράς μέχρι να επιτευχθεί η επιθυμητή σταθερότητα. Η νέα χρονοσειρά που προκύπτει και οι αντίστοιχες συναρτήσεις αυτοσυσχέτισης παρουσιάζονται στη συνέχεια.



Σχήμα 5.1.5 : Χρονοσειρά μετά την εποχική και κανονική διαφύριση των δεδομένων



Σχήμα 5.1.6 : Συνάρτηση αυτοσυσχέτισης μετά την εποχική και κανονική διαφορίση



Σχήμα 5.1.7 : Συνάρτηση μερικής αυτοσυσχέτισης μετά την εποχική και κανονική διαφορίση

Η χρονοσειρά και οι αντίστοιχες συναρτήσεις αυτοσυσχέτισης, αναφέρονται στα δεδομένα που προέκυψαν μετά την εφαρμογή εποχικής διαφορίσης η οποία ακολουθήθηκε από κανονική διαφορίση, για την εξάλειψη των προτύπων εποχικότητας και τάσης αντίστοιχα. Συγκεκριμένα, οι τιμές της νέας χρονοσειράς, παρουσιάζονται στο **Παράρτημα Α** και υπολογίζονται από την ακόλουθη σχέση :

$$\Delta\Delta_{12}Y_t = \Delta(\Delta_{12}Y_t) = \Delta(Y_t - Y_{t-12}) = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$$

Στο σημείο αυτό, μέσω των μεθόδων διαφορίσης (εποχικής και κανονικής) των δεδομένων έχουμε επιτύχει την ύπαρξη σταθερότητας στη διακύμανση της χρονοσειράς. Η προκύπτουσα γραφική παράσταση παρατηρούμε ότι κυμαίνεται γύρω από μια σταθερή, μέση τιμή, η οποία εν προκειμένω



είναι η μηδενική*. Το γεγονός ότι η συνάρτηση αυτοσυσχέτισης δεν παρουσιάζει κάποιο συγκεκριμένο μοτίβο διακύμανσης (σημειώνεται μόνο μία σημαντική αυτοσυσχέτιση για χρονική καθυστέρηση μιας περιόδου) που να υποδηλώνει τάση ή εποχικότητα των δεδομένων, επαληθεύει την παρατήρηση αυτή.

Με δεδομένη την ύπαρξη σταθερότητας στη χρονοσειρά, μπορεί πλέον να αναζητηθεί ένα πρώτο μοντέλο πρόβλεψης ARIMA. Όπως περιγράφηκε στη σχετική παράγραφο της μεθοδολογίας των Box-Jenkins, η αναγνώριση του μοντέλου είναι δυνατό να επιτευχθεί συγκρίνοντας τις πραγματικές συναρτήσεις αυτοσυσχέτισης, με τις αντίστοιχες θεωρητικές των μοντέλων ARIMA που υπάρχουν στη βιβλιογραφία.

Στη χρονοσειρά που μελετάμε, η συνάρτηση αυτοσυσχέτισης παρουσιάζει ένα σημαντικό συντελεστή στο lag 1, ενώ η συνάρτηση μερικής αυτοσυσχέτισης παρατηρούμε ότι ξεκινώντας από μεγάλες αρνητικές τιμές των συντελεστών στις πρώτες χρονικές καθυστερήσεις, πλησιάζει σταδιακά το μηδέν. Η διακύμανση αυτή των συναρτήσεων αυτοσυσχέτισης, παραπέμπει, σε ένα μοντέλο κινητού μέσου όρου 1^{ης} τάξης. Επομένως, το πρώτο δοκιμαστικό μοντέλο στο οποίο καταλήγει η ανάλυση της χρονοσειράς, είναι το εποχικό μοντέλο :

$$ARIMA(p, d, q)(P, D, Q)_{12} = ARIMA(0, 1, 1)(0, 1, 0)_{12}$$

αφού έχει εφαρμοστεί μια κανονική διαφόριση ($d = 1$), μια εποχική διαφόριση ($D = 1$) και ένας κινητός μέσο όρος 1^{ης} τάξης ($q = 1$). Έχοντας καταλήξει στο δοκιμαστικό μοντέλο, μπορούμε πλέον να προχωρήσουμε στο επόμενο βήμα της μεθοδολογίας Box-Jenkins το οποίο αφορά την εκτίμηση των παραμέτρων που υπεισέρχονται στο μοντέλο πρόβλεψης ARIMA.

*Λόγω του ότι η μέση τιμή γύρω από την οποία σημειώνεται η διακύμανση της χρονοσειράς είναι η μηδενική, στο τελικό μοντέλο ARIMA που θα επιλεγεί δεν συμπεριλαμβάνεται σταθερός όρος.

Βήμα 2^ο : Εκτίμηση των παραμέτρων

Το δεύτερο βήμα της μεθοδολογίας Box-Jenkins, περιλαμβάνει την εκτίμηση των παραμέτρων $\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_p$ ή/και $\omega_1, \omega_2, \dots, \omega_q$ που ορίζουν το επιλεχθέν δοκιμαστικό μοντέλο πρόβλεψης ARIMA. Κριτήριο για την τελική εκτίμηση των παραμέτρων, είναι η ελαχιστοποίηση του αθροίσματος των τετραγώνων των σφαλμάτων μεταξύ της πραγματικά εκδηλωμένης τιμής και της αντίστοιχης εκτίμησης της (fitted value), όπως υπολογίζεται από το μοντέλο. Για την εκτίμηση των παραμέτρων



χρησιμοποιούνται αυτοματοποιημένοι μη γραμμικοί επαναληπτικοί αλγόριθμοι ελαχιστοποίησης των σφαλμάτων, οι οποίοι συνήθως συγκλίνουν σε μία βέλτιστη τιμή για τις παραμέτρους μετά από σχετικά μικρό αριθμό επαναλήψεων.

Το μοντέλο πρόβλεψης στο οποίο έχουμε καταλήξει είναι το $ARIMA(0,1,1)(0,1,0)_{12}$. Είναι σημαντικό, ο αναλυτής/στέλεχος της επιχείρησης που χρησιμοποιεί την ανάλυση ARIMA, να κατανοήσει την εξίσωση μέσω της οποίας πραγματοποιούνται οι προβλέψεις, για το μοντέλο που επιλέχθηκε. Για το σκοπό αυτό, στη συνέχεια επαναλαμβάνονται οι βασικές εξισώσεις που περιγράφουν τα μοντέλα ARIMA καθώς και οι εξισώσεις διαφορίσης των δεδομένων.

$$AR(p) \rightarrow Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + e_t$$

$$*MA(q) \rightarrow Y_t = \mu + e_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \dots - \omega_q \varepsilon_{t-q}$$

$$\Delta_{12} Y_t = Y_t - Y_{t-12} = \text{η χρονοσειρά μετά την πρώτη εποχική διαφορίση } s = 12$$

$$\Delta Y_t = Y_t - Y_{t-1} = \text{η χρονοσειρά μετά την πρώτη κανονική διαφορίση}$$

$$** \Delta \Delta_{12} Y_t = \Delta(\Delta_{12} Y_t) = \Delta(Y_t - Y_{t-12}) = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13} = \text{η χρονοσειρά που προκύπτει μετά από εποχική διαφορίση η οποία ακολουθείται από κανονική διαφορίση των δεδομένων.}$$

από * και ** προκύπτει : $Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13} = \mu + e_t - \omega_1 \varepsilon_{t-1}$ όπου $\mu = 0$ λόγω του ότι η σταθερή χρονοσειρά κυμαίνεται γύρω από τη μηδενική μέση τιμή. Λύνοντας ως προς Y_t :

$$Y_t = e_t - \omega_1 \varepsilon_{t-1} + Y_{t-1} + Y_{t-12} - Y_{t-13}$$

Για το μοντέλο $ARIMA(0,1,1)(0,1,0)_{12}$ που επιλέχθηκε, απαιτείται ο υπολογισμός μιας παραμέτρου, της ω_1 . Η τιμή της παραμέτρου υπολογίζεται μέσω επαναληπτικού αλγορίθμου, σύμφωνα με τον οποίο επιλέγεται η βέλτιστη τιμή της παραμέτρου μεταξύ $0 \leq \omega_1 \leq 1$ με κριτήριο την ελαχιστοποίηση του αθροίσματος των τετραγώνων των υπολοίπων (residuals).

Το στατιστικό πακέτο Minitab, που χρησιμοποιείται ευρέως για την ανάλυση χρονοσειρών, εξάγει τη σχετική έκθεση αποτελεσμάτων που παρατίθεται στη συνέχεια συμπεριλαμβανομένης της τελικής εκτίμησης της παραμέτρου του μοντέλου πρόβλεψης ARIMA. Στη σχετική αυτή έκθεση υπολογίζονται επίσης τα προκύπτοντα σφάλματα και οι στατιστικές παράμετροι που απαιτούνται για το διαγνωστικό έλεγχο του μοντέλου όπως θα δούμε στη συνέχεια.



ARIMA (0, 1, 1) (0, 1, 0)

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
MA 1	0.6736	0.0920	7.32	0.000

Differencing: 1 regular, 1 seasonal of order 12
Number of observations: Original series 84, after differencing 71
Residuals: SS = 1653583881 (backforecasts excluded)
MS = 23622627 DF = 70

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	11.7	20.9	27.3	36.4
DF	11	23	35	47
P-Value	0.388	0.585	0.819	0.869

Η τελική εκτίμηση της τιμής της παραμέτρου, όπως προέκυψε από τον επαναληπτικό αλγόριθμο υπολογισμού για την ελαχιστοποίηση του αθροίσματος των τετραγώνων των σφαλμάτων, είναι $\omega_1 = 0.6736$. Παρατηρούμε επίσης στη σχετική έκθεση αποτελεσμάτων, ότι στα δεδομένα της αρχικής χρονοσειράς, εφαρμόστηκε μία κανονική διαφόριση (1 regular) και μια εποχική με χρονική διάρκεια 12 περιόδων (1 seasonal of order 12) ακριβώς όπως προνοεί το δοκιμαστικό μοντέλο ARIMA που επιλέχθηκε. Η αρχική χρονοσειρά αποτελείται από 84 παρατηρήσεις (πωλήσεις, ανά μήνα, από το 2007-2013) και μετά την εφαρμογή των διαφορίσεων ώστε να επιτευχθεί σταθερότητα, η τελική χρονοσειρά αποτελείται από 71 παρατηρήσεις.

Η τελική εξίσωση υπολογισμού των προβλέψεων για το δοκιμαστικό μοντέλο $ARIMA(0,1,1)(0,1,0)_{12}$ μετά και από την εκτίμηση της παραμέτρου είναι επομένως η ακόλουθη:

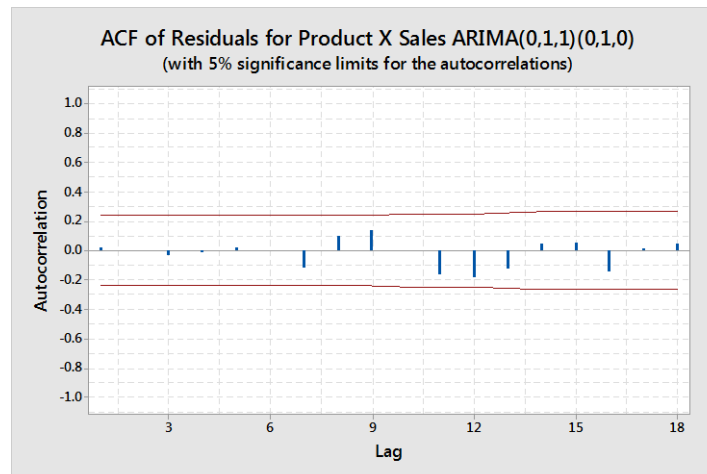
$$Y_t = -0.6736\varepsilon_{t-1} + Y_{t-1} + Y_{t-12} - Y_{t-13}$$

Το δοκιμαστικό μοντέλο ARIMA που αναπτύχθηκε, προτού χρησιμοποιηθεί για τη διενέργεια προβλέψεων, πρέπει να ελεγχθεί στατιστικά για την επάρκεια και την καταλληλότητα του. Η διαδικασία του στατιστικού ελέγχου της επάρκειας του υπολογισθέντος μοντέλου πρόβλεψης, πραγματοποιείται στο επόμενο βήμα της μεθοδολογίας των Box-Jenkins.

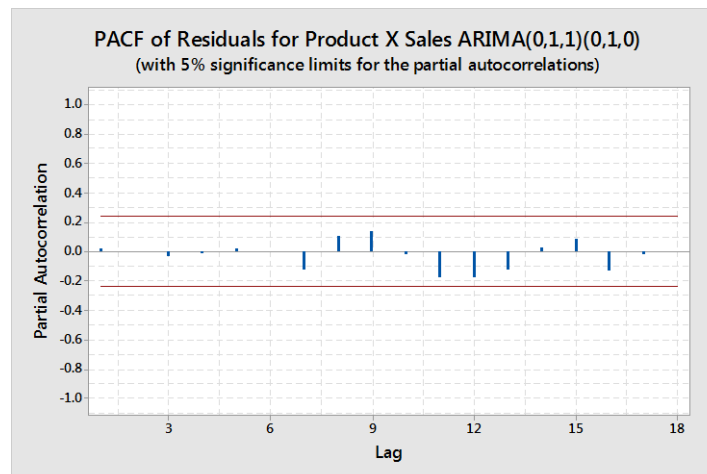


Βήμα 3^ο : Διαγνωστικός Έλεγχος του μοντέλου

Ο διαγνωστικός έλεγχος πραγματοποιείται για τον έλεγχο της καταλληλότητας και της επάρκειας του επιλεγμένου μοντέλου πρόβλεψης. Για να ελέγξουμε την καταλληλότητα του μοντέλου, κατασκευάζεται το διάγραμμα αυτοσυσχετίσεων των σφαλμάτων πρόβλεψης (residuals) της χρονοσειράς. Εάν οι συντελεστές αυτοσυσχέτισης για τα υπολογισθέντα σφάλματα, δεν είναι στατιστικά σημαντικοί, δηλαδή βρίσκονται μέσα στα προκαθορισμένα όρια και δεν παρουσιάζουν κάποιο επαναλαμβανόμενο μοτίβο διακύμανσης, τότε το χρησιμοποιούμενο μοντέλο πρόβλεψης θεωρείται επαρκές. Τα άνω και κάτω όρια, εντός των οποίων τα σφάλματα κρίνονται στατιστικά σημαντικά ορίζονται από τη σχέση $\pm \frac{2}{\sqrt{n}}$ όπου n είναι ο αριθμός των παρατηρήσεων της χρονοσειράς.



Σχήμα 5.1.8 : Συνάρτηση αυτοσυσχέτισης των υπολοίπων



Σχήμα 5.1.9 : Συνάρτηση μερικής αυτοσυσχέτισης των υπολοίπων



Όπως προκύπτει από τις παραπάνω συναρτήσεις αυτοσυσχέτισης των σφαλμάτων πρόβλεψης, οι υπολογιζόμενοι συντελεστές βρίσκονται καθαρά εντός των προκαθορισμένων ορίων, δεν σημειώνεται κάποιο μοτίβο διακύμανσης, και επομένως το μοντέλο μπορεί να θεωρηθεί επαρκές.

Ένας άλλος τρόπος που χρησιμοποιείται για τον έλεγχο της επάρκειας του μοντέλου πρόβλεψης ARIMA, είναι να εξετάσουμε διαγνωστικά τις αυτοσυσχετίσεις των υπολοίπων, χρησιμοποιώντας το στατιστικό έλεγχο X^2 . Το κριτήριο ελέγχου είναι (Ljung – Box Q statistic) :

$$Q = n(n + 2) \sum_{k=1}^m \frac{r_k^2(e)}{n - k}$$

το οποίο κατανέμεται κατά προσέγγιση, σύμφωνα με την κατανομή X^2 με $m - r$ βαθμούς ελευθερίας. Σύμφωνα με τη θεωρία της στατιστικής (hypothesis testing) ισχύουν τα εξής :

$$H_0 = \text{το μοντέλο είναι επαρκές}$$

$$H_A = \text{το μοντέλο είναι ανεπαρκές}$$

Εάν η υπολογισθείσα τιμή του Q είναι μεγαλύτερη από την $X^2(a, m - r)$ (η τιμή προκύπτει από τον πίνακα τιμών της σχετικής κατανομής για $m - r$ βαθμούς ελευθερίας), τότε το μοντέλο θεωρείται ανεπαρκές. Εναλλακτικά, εάν η υπολογιζόμενη τιμή **p-value** σε σχέση με την υπολογισθείσα τιμή του Q statistic είναι μικρή, (π.χ. $p - value < 0.05$) τότε δεν μπορούμε να δεχθούμε τη μηδενική υπόθεση και καταλήγουμε στο συμπέρασμα ότι το μοντέλο θεωρείται στατιστικά ανεπαρκές.

Για τις ανάγκες του διαγνωστικού ελέγχου του μοντέλου πρόβλεψης, στη σχετική έκθεση αποτελεσμάτων του Minitab, παρατίθενται οι απαραίτητοι υπολογισμοί των εμπλεκόμενων ποσοτήτων. Συγκεκριμένα, πραγματοποιείται ο έλεγχος Ljung – Box Q statistic, για τα γκρουπ χρονικών καθυστερήσεων $m = 12,24,36,48$. Οι υπολογισμοί είναι :

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	11.7	20.9	27.3	36.4
DF	11	23	35	47
P-Value	0.388	0.585	0.819	0.869

Όπως και στην περίπτωση των συναρτήσεων αυτοσυσχέτισης των υπολοίπων, το μοντέλο φαίνεται να είναι στατιστικά επαρκές, λόγω των υψηλών τιμών p-value που προέκυψαν για τα $m = 12,24,26,48$.



Η τιμή του Q για $m = 12$ όπως υπολογίστηκε, είναι $Q = 11.7$. Ο αριθμός των βαθμών ελευθερίας της κατανομής X^2 είναι $m - r = 12 - 1 = 11$ (DF). Επομένως, από τον πίνακα τιμών της κατανομής X^2 για 11 βαθμούς ελευθερίας, σε επίπεδο σημαντικότητας 0.05, λαμβάνεται η ακόλουθη τιμή: $X^2(0.025, 11) = 19.6751$. Ισχύει δηλαδή ότι $Q = 11.7 < X^2(0.025, 11) = 19.6751$, επομένως μπορούμε να δεχθούμε τη μηδενική υπόθεση στο επίπεδο σημαντικότητας 0.05 και να καταλήξουμε και πάλι στο συμπέρασμα ότι το μοντέλο είναι στατιστικά επαρκές. Ομοίως, πραγματοποιούνται οι αντίστοιχοι έλεγχοι επάρκειας του μοντέλου για $m = 24, 36, 38$ ούτως ώστε να ελεγχθεί πλήρως το δοκιμαστικό μοντέλο ARIMA.

Βήμα 4^ο : Πρόβλεψη με χρήση του μοντέλου

Μετά την ταυτοποίηση και τον έλεγχο επάρκειας του επιλεγθέντος μοντέλου, μπορούν να πραγματοποιηθούν προβλέψεις για την τιμή που θα λάβει η προς εξέταση μεταβλητή σε μια ή περισσότερες χρονικές περιόδους στο μέλλον. Πρέπει να σημειωθεί ότι καθώς η περίοδος πρόβλεψης αυξάνεται, ενισχύονται οι πιθανότητες να εμφανιστούν μεγαλύτερα και πιθανά, μη αποδεκτά σφάλματα πρόβλεψης. Επομένως, με την απόκτηση νέων παρατηρήσεων για μια χρονοσειρά, το χρησιμοποιούμενο μοντέλο πρόβλεψης πρέπει να επανεξετάζεται και να ελέγχεται ως προς την ακρίβεια του. Εάν η χρονοσειρά φαίνεται ότι αλλάζει διαχρονικά είτε όσον αφορά την τάση, είτε την εποχικότητα, οι παράμετροι του μοντέλου πρέπει να επαναυπολογιστούν ή και σε ορισμένες περιπτώσεις, θα πρέπει να αναπτυχθεί ένα εντελώς νέο μοντέλο.

Οι προβλέψεις για το επόμενο έτος (2014) ανά μήνα, μαζί με τα άνω και κάτω όρια τιμών (95% limits) πραγματοποιούνται μέσω της εξίσωσης του μοντέλου :

$$Y_t = -0.6736\varepsilon_{t-1} + Y_{t-1} + Y_{t-12} - Y_{t-13}$$

Υπενθυμίζουμε ότι η αρχική χρονοσειρά, αποτελείτο από 84 παρατηρήσεις (πωλήσεις ανά μήνα από 2007 έως 2013), με την τελευταία παρατήρηση να αντιπροσωπεύει τις πωλήσεις του προϊόντος X για το μήνα Δεκέμβριο του 2013. Θέτοντας χρονικό ορίζοντα πρόβλεψης τις 12 περιόδους, υπολογίζονται οι τιμές πρόβλεψης ανά μήνα για το έτος 2014.

Τα αποτελέσματα που προκύπτουν από τον υπολογισμό των προβλέψεων και των άνω και κάτω ορίων σε επίπεδο εμπιστοσύνης 95% σύμφωνα με το μοντέλο $ARIMA(0,1,1)(0,1,0)_{12}$ είναι τα εξής:



Forecasts from period 84

Period	Forecast	95% Limits		Actual
		Lower	Upper	
85	13457.4	3929.3	22985.6	
86	11045.5	1022.6	21068.3	
87	8978.5	1515.7	19472.7	
88	22468.5	11523.2	33413.8	
89	19082.4	7703.9	30461.0	
90	16077.5	4281.6	27873.4	
91	23496.4	11297.5	35695.4	
92	24006.5	11417.3	36595.6	
93	13689.5	721.9	26657.0	
94	19801.9	6466.7	33137.2	
95	12424.5	1268.6	26117.6	
96	19311.5	5269.7	33353.3	

Επαλήθευση τιμών πρόβλεψης

Η εξίσωση υπολογισμού των τιμών για τις προβλέψεις των επόμενων χρονικών περιόδων, είναι :

$$Y_t = -0.6736\varepsilon_{t-1} + Y_{t-1} + Y_{t-12} - Y_{t-13}$$

Άρα ισχύει :

$$Y_{85} = -0.6736\varepsilon_{84} + Y_{84} + Y_{73} - Y_{72}$$

Οι τιμές των ε_{84} , Y_{84} , Y_{73} και Y_{72} δίνονται στο σχετικό πίνακα του **Παράρτηματος Α**. Αντικαθιστώντας στην παραπάνω εξίσωση προκύπτει:

$$Y_{85} = 0.6736 \times 10236.9 + 16019.1 + 10165.0 - 5831.0 = 13457.4$$

ομοίως

$$Y_{86} = -0.6736\varepsilon_{85} + Y_{85} + Y_{74} - Y_{73}$$

και αντικαθιστώντας

$$Y_{86} = -0.6736 \times 0 + 13457.4 + 7753.032 - 10165.01 = 11045.5$$

Οι τιμές πρόβλεψης για τις χρονικές περιόδους 85 και 86 (αντιστοιχούν στον Ιανουάριο και Φεβρουάριο του 2014) επαληθεύονται με τις αντίστοιχη τιμές πρόβλεψης στην έκθεση αποτελεσμάτων

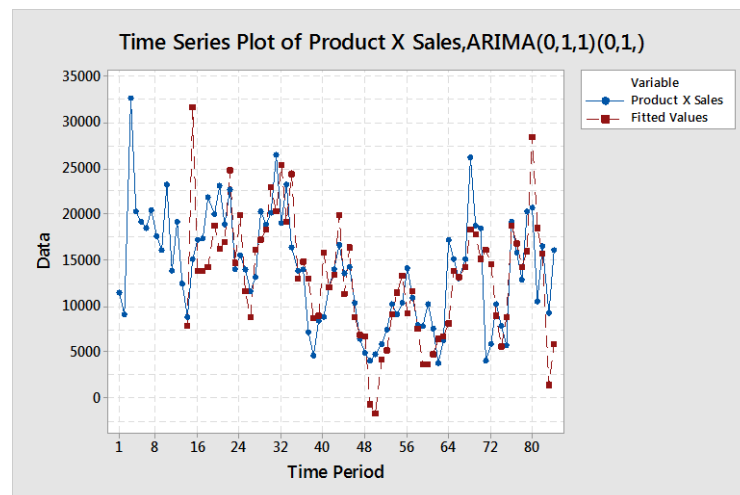


του Minitab. Είναι πολύ σημαντικό για τον αναλυτή να κατανοήσει το μηχανισμό υπολογισμού προβλέψεων για το εκάστοτε μοντέλο ARIMA που χρησιμοποιεί, ούτως ώστε να είναι σε θέση να ερμηνεύσει ποιοτικά τα αποτελέσματα που προκύπτουν αλλά και να αναγνωρίσει ευκολότερα τυχόν πηγές μεγάλων αποκλίσεων και σφαλμάτων που θα προκύψουν, λόγω ανεπάρκειας του μοντέλου.

Το μοντέλο πρόβλεψης ARIMA που επιλέχθηκε, είναι δοκιμαστικό. Από στατιστικής πλευράς, έχει ελεγχτεί και θεωρείται επαρκές, όπως προνοεί η μεθοδολογία των Box-Jenkins. Πέρα όμως από το στατιστικό έλεγχο επάρκειας του μοντέλου, πρέπει κάθε φορά που υπάρχει διαθέσιμη η πραγματικά εκδηλωμένη τιμή της εξεταζόμενης μεταβλητής, να συγκρίνεται με την τιμή πρόβλεψης του μοντέλου για την αντίστοιχη χρονική περίοδο και να υπολογίζεται το σφάλμα πρόβλεψης. Η παρακολούθηση του συστήματος πρόβλεψης είναι κρίσιμη διαδικασία που πρέπει να υλοποιείται για την έγκαιρη αναγνώριση της ανάγκης για επανεκτίμηση των παραμέτρων του μοντέλου.

Σε ορισμένες περιπτώσεις η εμφάνιση παραγόντων που θα επηρεάσουν σε μεγάλο βαθμό τη διακύμανση της ζήτησης, ενδέχεται να οδηγήσει στην ανάγκη για αναζήτηση ενός εντελώς καινούργιου μοντέλου πρόβλεψης ARIMA. Εάν συμβεί αυτό, απαιτείται να εφαρμοστεί από την αρχή η μεθοδολογία των Box-Jenkins, αφού πρώτα αναβαθμιστούν τα χρησιμοποιούμενα ιστορικά στοιχεία πωλήσεων αντικαθιστώντας τις παλαιότερες τιμές με τις πιο πρόσφατες και πιο αντιπροσωπευτικές.

Στο πιο κάτω διάγραμμα απεικονίζονται οι εκτιμώμενες τιμές (Fitted Values) του μοντέλου σε σχέση με τη χρονοσειρά των πραγματικών πωλήσεων (Product X Sales).



Σχήμα 5.1.10 : Χρονοσειρές της πραγματικής ζήτησης και των εκτιμημένων τιμών το μοντέλου ARIMA



Όπως έχει αποδειχθεί και στο βήμα του διαγνωστικού ελέγχου, παρατηρούμε ότι οι εκτιμώμενες τιμές από το μοντέλο ARIMA προσεγγίζουν σε ικανοποιητικό βαθμό το μοτίβο διακύμανσης των πραγματικών τιμών της χρονοσειράς. Οι τιμές των κριτηρίων αξιολόγησης μετά τον υπολογισμό των σφαλμάτων, είναι $MAPE = 0.36$ και $MPE = -0.04$

Από στατιστικής πλευράς, το χρησιμοποιούμενο μοντέλο θα μπορέσει να προβλέψει με αποδεκτή ακρίβεια τις τιμές της προς εξέταση μεταβλητής στο μέλλον, νοουμένου ότι θα διατηρηθούν (ή ισοδύναμα, δεν θα μεταβληθούν σε ακραίο βαθμό) οι παράγοντες που διαμορφώνουν το μοτίβο διακύμανσης της χρονοσειράς. Και λέμε από στατιστικής πλευράς για το λόγο ότι οποιοδήποτε μοντέλο πρόβλεψης υιοθετείται από μια επιχείρηση, αδυνατεί να συνυπολογίσει όλους τους αστάθμητους παράγοντες που πιθανά να εμφανιστούν και να επηρεάσουν την προβλεπόμενη ζήτηση.

Οι προβλέψεις που παράγονται από τα μοντέλα πρόβλεψης μέσω της ανάλυσης χρονοσειρών, στηρίζονται αποκλειστικά στα ιστορικά στοιχεία πωλήσεων των προηγούμενων ετών και αποσκοπούν στην αναγνώριση των παραγόντων εκείνων που επηρέασαν τη διακύμανση της ζήτησης. Υποδηλώνεται επομένως σιωπηρά ότι οι προβλέψεις που διενεργούνται, είναι υπό την προϋπόθεση ότι οι ίδιοι παράγοντες που επηρέασαν τη ζήτηση στα προηγούμενα έτη, θα εξακολουθήσουν να επιδρούν στη ζήτηση και κατά τη χρονική περίοδο των προβλέψεων.

Αυτός είναι ο κυριότερος λόγος που οι προβλέψεις πρέπει να εμπλουτίζονται με τις γνώμες εμπειρών στελεχών οι οποίοι έχουν ανεπτυγμένη αίσθηση του περιβάλλοντος ανταγωνισμού και η δική τους κρίση είναι ιδιαίτερος σημαντική για την «εξομάλυνση» του συνεχώς μεταβαλλόμενου περιβάλλοντος της αγοράς που δραστηριοποιούνται σήμερα οι επιχειρήσεις.

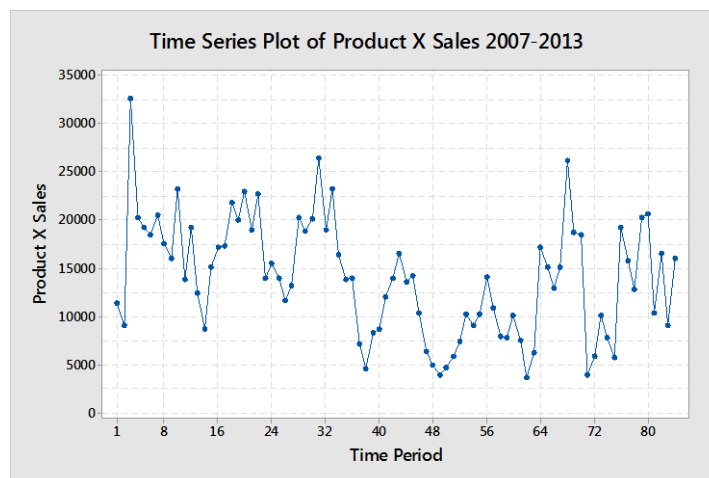


5.2 Μοντέλο Πρόβλεψης Winters

Στην παράγραφο που προηγήθηκε και αφορούσε τη μελέτη της χρονοσειράς μέσω της ανάλυσης αυτοσυσχετίσεων, είχαμε καταλήξει στο συμπέρασμα ότι στη διαμόρφωση του μοτίβου διακύμανσης των πωλήσεων του προϊόντος, εκτός από την τυχαιότητα, υπεισέρχονται οι παράγοντες της τάσης και της εποχικότητας. Καθίσταται λοιπόν αναγκαίο, η μέθοδος πρόβλεψης που θα χρησιμοποιηθεί, εξ ορισμού να λαμβάνει υπόψη αυτούς τους παράγοντες. Από την κατηγορία των ποσοτικών μεθόδων εξομάλυνσης, ιδανική θεωρείται η μέθοδος της εκθετικής εξομάλυνσης με προσαρμογή στην τάση και στην εποχικότητα (exponential smoothing adjusted for trend and seasonality), γνωστή ως μέθοδος Winters. Να υπενθυμίσουμε ότι με τις ποσοτικές μεθόδους εξομάλυνσης, επιχειρούνται προβλέψεις για μια μεταβλητή μέσω της προεκβολής στο μέλλον, του μοτίβου διακύμανσης της χρονοσειράς.

Η μέθοδος Winters έχει τρεις συντελεστές, α , β και γ , οι οποίοι χρησιμοποιούνται για την εξομάλυνση του επιπέδου, της τάσης και της εποχικότητας αντίστοιχα. Οι βέλτιστες τιμές των συντελεστών εξομάλυνσης εκτιμώνται εφαρμόζοντας επαναληπτική διαδικασία δοκιμών για όλους τους δυνατούς συνδυασμούς των τιμών των συντελεστών, με κριτήριο την ελαχιστοποίηση του MSE. Όπως περιγράφεται και στο σχετικό εδάφιο, το πρώτο βήμα εφαρμογής της μεθόδου Winters είναι η εκτίμηση των τιμών των παραμέτρων επιπέδου (L), τάσης (T) και εποχικότητας (S) ούτως ώστε να αποτελέσουν τις αρχικές τιμές εισόδου στις εξισώσεις υπολογισμού των τιμών των προβλέψεων.

Η μέθοδος Winters θα εφαρμοστεί, όπως και στην ανάλυση ARIMA που προηγήθηκε, για τα δεδομένα πωλήσεων του προϊόντος X. Η χρονοσειρά που προκύπτει από τα δεδομένα πωλήσεων, ανά μήνα, από το 2007 έως το 2013, είναι η εξής :



Σχήμα 5.2.1 : Χρονοσειρά πωλήσεων

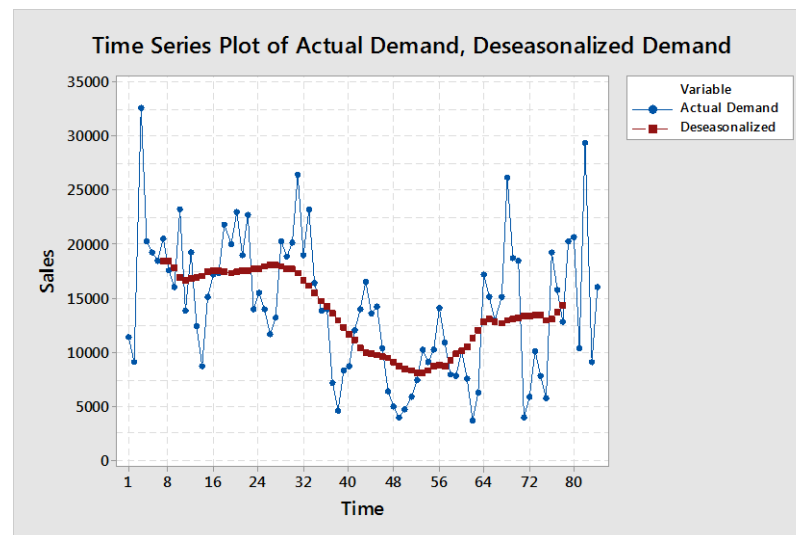


Παρατηρώντας την παραπάνω χρονοσειρά, επιβεβαιώνονται οι ισχυρισμοί για την ύπαρξη προτύπων τάσης και εποχικότητας στα δεδομένα πωλήσεων. Υπενθυμίζουμε ότι, η χρονική διάρκεια της περιοδικότητας (p) είναι ο αριθμός των χρονικών περιόδων, μετά την πάροδο των οποίων επαναλαμβάνεται ο εποχικός κύκλος (επαναλαμβανόμενη διακύμανση). Για παράδειγμα, σε διαθέσιμα μηνιαία ιστορικά στοιχεία στα οποία υπάρχει εποχικότητα, η χρονική διάρκεια της περιοδικότητας είναι $p = 12$, ενώ για τριμηνιαία ιστορικά στοιχεία, όπως εν προκειμένω, προκύπτει ότι $p = 4$.

Το πρώτο βήμα που απαιτείται για την εκτίμηση των παραμέτρων του μοντέλου Winters της χρονοσειράς, αφορά την από-εποχικοποίηση των δεδομένων (deseasonalizing data). Ο υπολογισμός της από-εποχικοποιημένης ζήτησης πραγματοποιείται μέσω της εξίσωσης ($p = 12$, άρτιος) :

$$\bar{D}_t = \frac{\left[D_{t-\left(\frac{p}{2}\right)} + D_{t+\left(\frac{p}{2}\right)} + \sum_{i=t+1-\left(\frac{p}{2}\right)}^{t-1+\left(\frac{p}{2}\right)} 2D_i \right]}{2p}$$

Τα αποτελέσματα από την αφαίρεση της εποχικότητας από τη ζήτηση παρουσιάζονται στο σχετικό πίνακα στο **Παράρτημα Β**. Η γραφική παράσταση των από-εποχικοποιημένων δεδομένων σε σύγκριση με τα αρχικά δεδομένα της χρονοσειράς είναι η ακόλουθη :



Σχήμα 5.2.2 : Χρονοσειρά πωλήσεων και από-εποχικοποιημένη ζήτηση

Με γραμμική παλινδρόμηση* (βλ. **Παράρτημα ΣΤ**) στην καμπύλη δεδομένων της από-εποχικοποιημένης ζήτησης, εκτιμώνται οι τιμές των L_0 και T_0 , αφού ισχύει η σχέση : $\bar{D}_t = L + Tt$

*Για την εφαρμογή γραμμικής παλινδρόμησης στα δεδομένα χρησιμοποιήθηκε η ειδική εντολή (Regression) από το Analysis Tool Pak του Microsoft Excel.



Η εξίσωση που προκύπτει από τη γραμμική παλινδρόμηση είναι :

$$\bar{D}_t = 18318 - 109.82t$$

επομένως λαμβάνουμε $L_0 = 18318.0$ και $T_0 = -109.82$

Οι τιμές της νέας ζήτησης χωρίς εποχικότητα για κάθε χρονική περίοδο δίνονται από τη σχέση :

$$\bar{D}_t' = L + Tt$$

και ο συντελεστής εποχικότητας, για κάθε περίοδο t της προς μελέτη χρονοσειράς, δίνεται από τη σχέση μεταξύ πραγματικής και από-εποχικοποιημένης ζήτησης :

$$\bar{S}_t = \frac{D_t}{\bar{D}_t'}$$

Τέλος, ως αρχικές τιμές των συντελεστών εποχικότητας, λαμβάνονται οι μέσοι όροι των συντελεστών των αντίστοιχων εποχικών χρονικών περιόδων μέσω της σχέσης :

$$S_i = \frac{(\sum_{j=0}^{r-1} \bar{S}_{jp+i})}{r} \quad \text{όπου } r = \frac{N}{p}$$

Στο μοντέλο Winters χρησιμοποιούνται τρεις εξισώσεις για την εξομάλυνση των στοιχείων του επιπέδου, της τάσης και της εποχικότητας για κάθε χρονική περίοδο αντίστοιχα :

$$L_{t+1} = \alpha \left(\frac{D_{t+1}}{S_{t+1}} \right) + (1 - \alpha)(L_t + T_t)$$

$$T_{t+1} = \beta(L_{t+1} - L_t) + (1 - \beta)T_t$$

$$S_{t+p+1} = \gamma \left(\frac{D_{t+1}}{L_{t+1}} \right) + (1 - \gamma)S_{t+1}$$

όπου α, β, γ , οι συντελεστές εξομάλυνσης επιπέδου, τάσης και εποχικότητας αντίστοιχα για τους οποίους ισχύει $0 \leq \alpha, \beta, \gamma \leq 1$. Για την εύρεση των βέλτιστων τιμών των συντελεστών, χρησιμοποιήθηκε το πρόσθετο εργαλείο Solver του Excel, θέτοντας ως κριτήριο (αντικειμενική συνάρτηση) την ελαχιστοποίηση του MSE*.

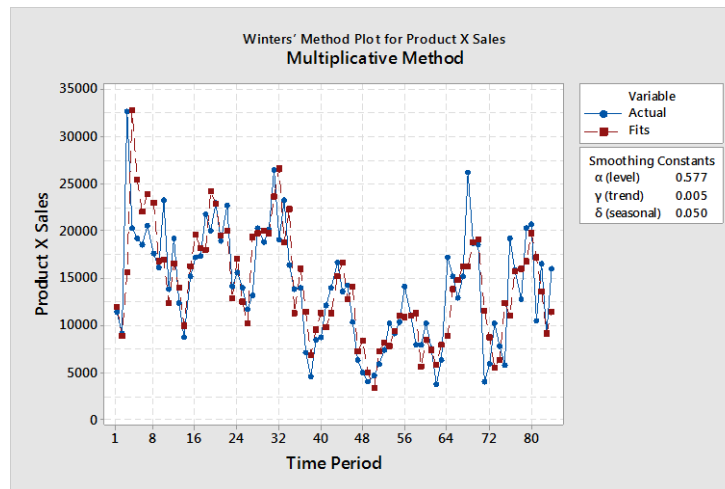
Τελικά, η πρόβλεψη F_{t+1} για την επόμενη χρονική περίοδο γίνεται συνδυάζοντας τα παραπάνω υπολογισθέντα στοιχεία, σύμφωνα με τη σχέση :

$$F_{t+1} = (L_t + T_t)S_{t+1}$$

*Όλα τα αποτελέσματα που προκύπτουν από τους παραπάνω υπολογισμούς, παρουσιάζονται στο **Παράρτημα Β**.

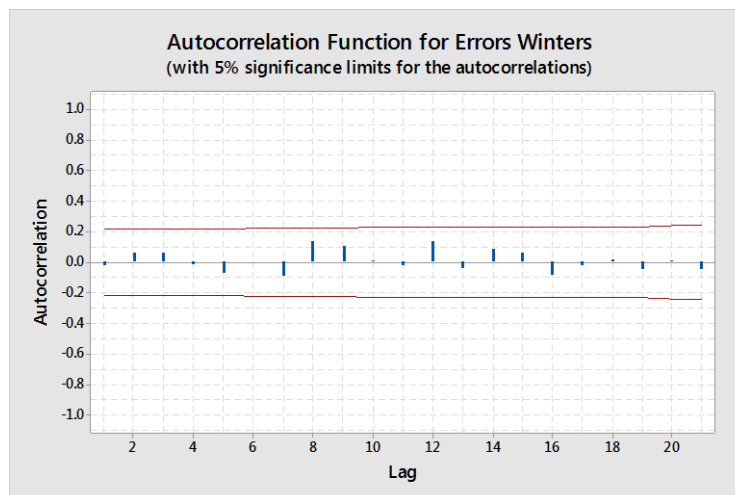


Η γραφική παράσταση που ακολουθεί, απεικονίζει τις πραγματικές τιμές της χρονοσειράς (Actual) σε σχέση με τις εκτιμημένες (Fits) από το μοντέλο Winters για τις συγκεκριμένες τιμές των συντελεστών εξομάλυνσης. Οι εκτιμημένες τιμές προσεγγίζουν σε ικανοποιητικό βαθμό τις πραγματικές, συμπεράσμα που τεκμηριώνεται και στατιστικά από τις τιμές των MAPE και MPE.



Σχήμα 5.2.3 : Χρονοσειρά πωλήσεων και εκτιμημένες τιμές του μοντέλου Winters

Εναλλακτικά, μπορούμε να ελέγξουμε την επάρκεια του χρησιμοποιούμενο μοντέλου Winters μέσω της συνάρτησης αυτοσυσχέτισης των υπολοίπων. Όπως και στην ανάλυση ARIMA, αφού οι συντελεστές αυτοσυσχέτισης για των υπολοίπων (residuals) δεν είναι στατιστικά σημαντικοί (όρια : $\pm \frac{2}{\sqrt{n}} = \pm \frac{2}{\sqrt{84}} = \pm 0.218$) και δεν παρουσιάζουν συγκεκριμένο μοτίβο διακύμανσης, το μοντέλο μπορεί να θεωρηθεί στατιστικά επαρκές. Ενδεικτικό είναι το διάγραμμα που ακολουθεί στο οποίο αποδεικνύεται η επάρκεια του μοντέλου μέσα από τον υπολογισμό της συνάρτησης αυτοσυσχέτισης.



Σχήμα 5.2.4 : Συνάρτηση αυτοσυσχέτισης των υπολοίπων



5.3 Μοντέλο Πρόβλεψης Holts

Ένα ακόμη μοντέλο πρόβλεψης που θα χρησιμοποιηθεί και ανήκει στην οικογένεια των ποσοτικών μεθόδων εξομάλυνσης, είναι το μοντέλο του Holt. Το μοντέλο της διπλής εκθετικής εξομάλυνσης (ή μοντέλο Holt) χρησιμοποιείται όταν στη χρονοσειρά της εξεταζόμενης μεταβλητής παρατηρείται ότι εκτός από το οριζόντιο στοιχείο, υπάρχει και γραμμική τάση. Στο σημείο αυτό, να αναφέρουμε ότι ο παράγοντας εποχικότητα δεν συνυπολογίζεται στο μοντέλο του Holt. Νοούμενου ότι έχει αποδειχθεί στατιστικά η ύπαρξη προτύπου εποχικότητας παράλληλα με το πρότυπο τάσης στα δεδομένα της χρονοσειράς, αναμένουμε ότι οι εκτιμώμενες τιμές και οι προβλέψεις του μοντέλου θα είναι μικρότερης ακριβείας από τις αντίστοιχες τιμές των προηγούμενων μοντέλων ARIMA και Winters.

Η εκτίμηση του οριζόντιου στοιχείου και της τάσης της χρονοσειράς γίνεται μέσω των σχέσεων :

$$\begin{aligned}L_{t+1} &= \alpha D_{t+1} + (1 - \alpha)(L_t + T_t) \\T_{t+1} &= \beta(L_{t+1} - L_t) + (1 - \beta)T_t\end{aligned}$$

όπου α, β οι συντελεστές εξομάλυνσης για το επίπεδο και την τάση. Ισχύει ότι : $0 \leq \alpha, \beta \leq 1$

Οι βέλτιστες τιμές των παραμέτρων α και β για μια συγκεκριμένη χρονοσειρά προκύπτουν μετά από δοκιμές, με κριτήριο την ελαχιστοποίηση της τιμής του μέσου τετραγωνικού σφάλματος MSE. Όπως και στο μοντέλο Winters οι βέλτιστες τιμές των συντελεστών εξομάλυνσης υπολογίστηκαν με χρήση του πρόσθετου εργαλείου Solver.

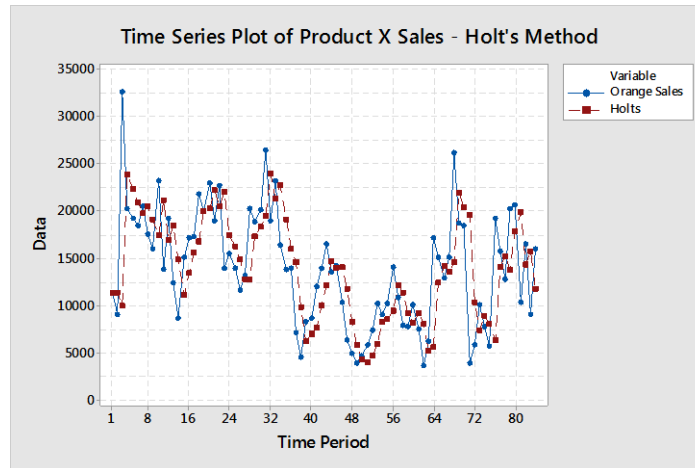
Για να εφαρμοστεί το μοντέλο του Holt, απαιτείται μια αρχική εκτίμηση των τιμών των παραμέτρων του επιπέδου L_0 και της τάσης T_0 . Η αρχική εκτίμηση της τιμής του επιπέδου λαμβάνεται ίση με την πρώτη παρατήρηση της χρονοσειράς και η αντίστοιχη αρχική εκτίμηση της τάσης ίση με το μηδέν.

$$L_0 = 11391.04 \text{ και } T_0 = 0$$

Με τις αρχικές εκτιμήσεις των παραμέτρων τάσης και επιπέδου, είναι διαθέσιμα όλα τα απαιτούμενα δεδομένα για την εξαγωγή προβλέψεων*. Υπολογίζονται διαδοχικά οι τιμές των L, T σε κάθε χρονική περίοδο και μέσω της ακόλουθης σχέσης προκύπτουν οι τιμές πρόβλεψης ως :

$$F_{t+n} = L_t + nT_t$$

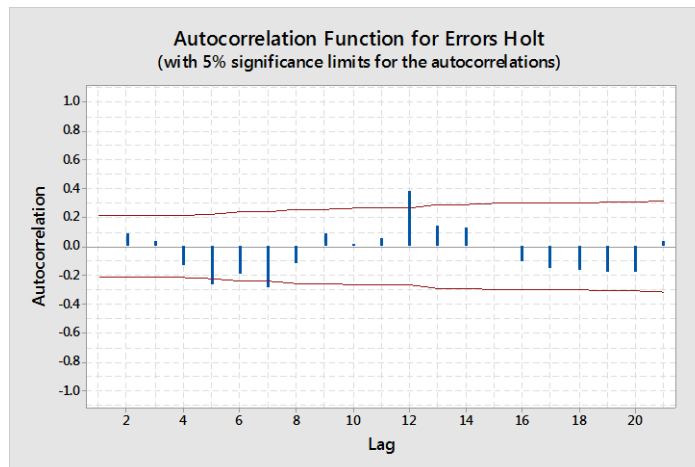
*Όλα τα αποτελέσματα που προκύπτουν από τους παραπάνω υπολογισμούς, παρουσιάζονται στο **Παράρτημα Γ**.



Σχήμα 5.3.1 : Χρονοσειρά πωλήσεων και εκτιμημένες τιμές του μοντέλου Holt's

Παρατηρώντας το παραπάνω διάγραμμα, αντιλαμβανόμαστε την αδυναμία του μοντέλου να προσεγγίσει ικανοποιητικά την πραγματική χρονοσειρά. Στατιστικά, το συμπέρασμα αυτό τεκμηριώνεται από τις μεγάλες τιμές των κριτηρίων αξιολόγησης των σφαλμάτων MAPE και MPE οι υπολογισθείσες τιμές των οποίων παρατίθενται στο **Παράρτημα Γ**.

Η ανεπάρκεια του χρησιμοποιούμενου μοντέλου πρόβλεψης, γίνεται εμφανής και μέσω της συνάρτησης αυτοσυσχέτισης των σφαλμάτων. Παρατηρούμε ότι υπάρχουν στατιστικά σημαντικοί συντελεστές αυτοσυσχέτισης των σφαλμάτων (errors), που ξεπερνούν δηλαδή τα όρια : $\pm \frac{2}{\sqrt{n}} = \pm \frac{2}{\sqrt{84}} = \pm 0.218$ για διάφορες χρονικές καθυστερήσεις. Αυτό συνεπάγεται ότι τα προκύπτοντα σφάλματα από τους υπολογισμούς του μοντέλου, δεν είναι τυχαία. Επίσης, το γεγονός ότι παρουσιάζεται στατιστικά σημαντικός συντελεστής αυτοσυσχέτισης στο lag 12 υποδηλώνει την ύπαρξη προτύπου εποχικότητας.



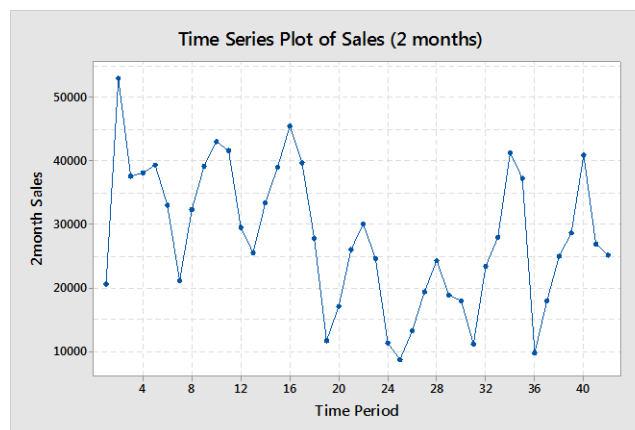
Σχήμα 5.3.2 : Συνάρτηση αυτοσυσχέτισης των σφαλμάτων

5.4 Μοντέλα Πρόβλεψης για διμηνιαίες και τριμηνιαίες πωλήσεις

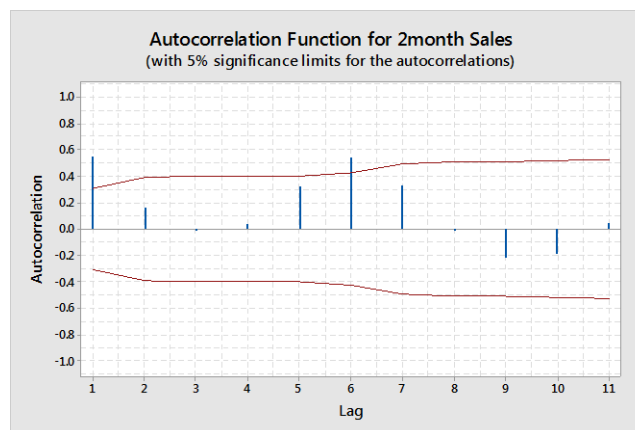
5.4.1 Μοντέλο Πρόβλεψης διμηνιαίων πωλήσεων ARIMA

Σε συνέχεια της μελέτης που προηγήθηκε για τη χρονοσειρά των μηνιαίων πωλήσεων του προϊόντος X, πραγματοποιήθηκε αναγωγή των πωλήσεων σε διμηνιαία βάση για την ανάπτυξη ενός μοντέλου το οποίο θα παράγει προβλέψεις για τη δίμηνη ζήτηση του προϊόντος στο μέλλον. Για την ανάπτυξη του μοντέλου, εφαρμόστηκε η διαδικασία των Box-Jenkins.

Ως γνωστό, το πρώτο βήμα της μεθοδολογίας των Box-Jenkins είναι η γραφική παράσταση των δεδομένων της αρχικής χρονοσειράς και ο υπολογισμός της συνάρτησης αυτοσυσχέτισης με στόχο την ταυτοποίηση του μοντέλου. Η μελέτη της χρονοσειράς μέσω της ανάλυσης των αυτοσυσχετίσεων για τις διαδοχικές παρατηρήσεις της, είναι καθοριστική για την εξαγωγή χρήσιμων συμπερασμάτων όπως :



Σχήμα 5.4.1 : Χρονοσειρά διμηνιαίων πωλήσεων



Σχήμα 5.4.2 : Συνάρτηση αυτοσυσχέτισης



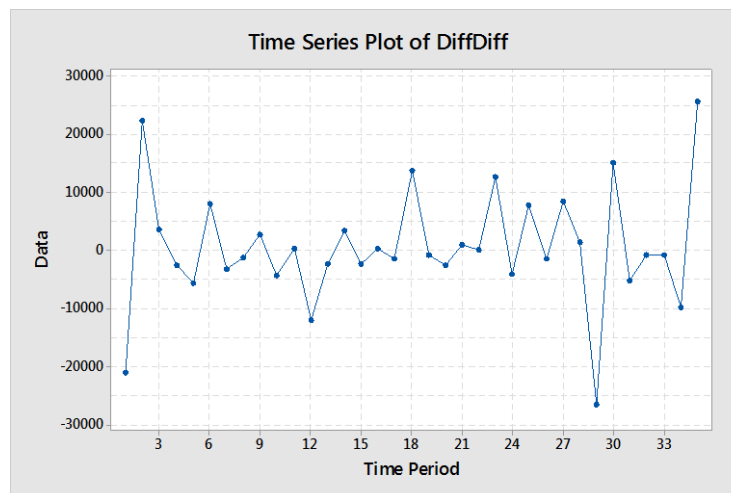
- Η χρονοσειρά παρουσιάζει εποχικότητα, αφού επαναλαμβάνει το μοτίβο διακύμανσης της μετά από κάθε πάροδο 6 χρονικών περιόδων.
- Η ύπαρξη του προτύπου εποχικότητας, τεκμηριώνεται και μέσω της συνάρτησης αυτοσυσχέτισης από το γεγονός ότι σημειώνεται στατιστικά σημαντικός συντελεστής αυτοσυσχέτισης στην εποχική χρονική καθυστέρηση των 6 περιόδων.
- Παρατηρώντας τη διακύμανση της χρονοσειράς, υποδηλώνεται πρότυπο τάσης αφού οι τιμές που εκδηλώνονται δεν κυμαίνονται γύρω από μια σταθερή τιμή.
- Τα όρια, πέρα από τα οποία οι συντελεστές αυτοσυσχέτισης θεωρούνται στατιστικά σημαντικοί (διάφοροι του μηδενός), προσδιορίζονται από : $\pm \frac{2}{\sqrt{n}} = \pm \frac{2}{\sqrt{42}} = \pm 0.308$

Για την περαιτέρω ανάλυση της χρονοσειράς ώστε να βρεθεί το κατάλληλο μοντέλο ARIMA, πρέπει να αφαιρεθούν τα πρότυπα εποχικότητας και τάσης από τα δεδομένα ώστε να καταλήξουμε σε μια σταθερή χρονοσειρά.. Η αφαίρεση των εμφανιζόμενων προτύπων με σκοπό τη δημιουργία σταθερής χρονοσειράς, πραγματοποιείται μέσω της κανονικής και εποχικής διαφορίσης των δεδομένων.

Συγκεκριμένα, οι τιμές της χρονοσειράς που προκύπτουν με την αφαίρεση των προτύπων τάσης και εποχικότητας υπολογίζονται από την ακόλουθη σχέση :

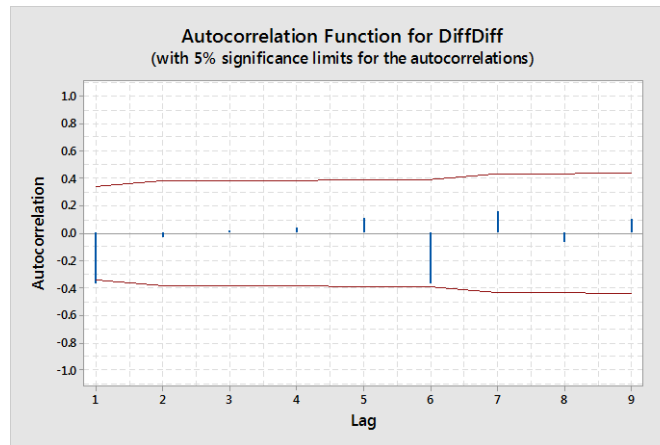
$$\Delta\Delta_6 Y_t = \Delta(\Delta_6 Y_t) = \Delta(Y_t - Y_{t-6}) = Y_t - Y_{t-1} - Y_{t-6} + Y_{t-7}$$

Η νέα χρονοσειρά και οι αντίστοιχες συναρτήσεις αυτοσυσχέτισης που υπολογίζονται για τα απαλλαγμένα από τάση και εποχικότητα δεδομένα, παρουσιάζονται στη συνέχεια.

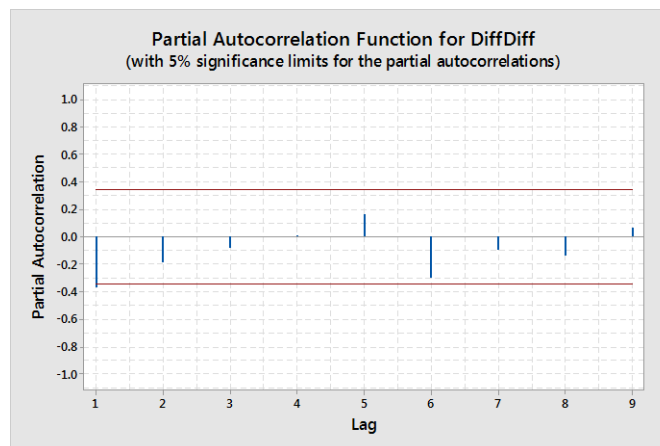


Σχήμα 5.4.3 : Χρονοσειρά μετά την εποχική και κανονική διαφορίση

*Όλα τα αποτελέσματα που προκύπτουν από τους παραπάνω υπολογισμούς, παρουσιάζονται στο Παράρτημα Α.



Σχήμα 5.4.4 : Συνάρτηση αυτοσυσχέτισης νέας χρονοσειράς



Σχήμα 5.4.5 : Συνάρτηση μερικής αυτοσυσχέτισης νέας χρονοσειράς

Η προκύπτουσα γραφική παράσταση της νέας χρονοσειράς παρατηρούμε ότι κυμαίνεται γύρω από μια σταθερή, μέση τιμή, η οποία εν προκειμένω είναι η μηδενική. Επιπρόσθετα, το γεγονός ότι η συνάρτηση αυτοσυσχέτισης δεν παρουσιάζει κάποιο συγκεκριμένο μοτίβο διακύμανσης (σημειώνεται μόνο μία σημαντική αυτοσυσχέτιση για χρονική καθυστέρηση μιας περιόδου) επαληθεύει την παρατήρηση ότι τα πρότυπα τάσης και εποχικότητας έχουν εξαλειφθεί από τη χρονοσειρά.

Με δεδομένη πλέον τη σταθερότητα της χρονοσειράς, έπεται να αναζητηθεί ένα πρώτο μοντέλο πρόβλεψης ARIMA. Όπως περιγράφηκε στη σχετική παράγραφο της μεθοδολογίας των Box-Jenkins, η αναγνώριση του μοντέλου είναι δυνατό να επιτευχθεί συγκρίνοντας τις πραγματικές συναρτήσεις αυτοσυσχέτισης, με τις αντίστοιχες θεωρητικές των μοντέλων ARIMA. Παρατηρούμε ότι η συνάρτηση αυτοσυσχέτισης παρουσιάζει ένα στατιστικά σημαντικό συντελεστή στο lag 1, ενώ η συνάρτηση μερικής αυτοσυσχέτισης παρατηρούμε ότι ξεκινώντας από μεγάλες αρνητικές τιμές των συντελεστών στις πρώτες χρονικές καθυστερήσεις, πλησιάζει σταδιακά το μηδέν.



Η διακύμανση αυτή των συναρτήσεων αυτοσυσχέτισης, παραπέμπει, σε ένα μοντέλο κινητού μέσου όρου 1^{ης} τάξης. Επομένως, το δοκιμαστικό μοντέλο στο οποίο καταλήγει η ανάλυση της χρονοσειράς, είναι το εποχικό μοντέλο :

$$ARIMA(p, d, q)(P, D, Q)_6 = ARIMA(0, 1, 1)(0, 1, 0)_6$$

αφού έχει εφαρμοστεί μια κανονική διαφόριση ($d = 1$), μια εποχική διαφόριση ($D = 1$) και ένας κινητός μέσο όρος 1^{ης} τάξης ($q = 1$), με διάρκεια εποχικότητας 6 χρονικές περιόδους.

Για το μοντέλο $ARIMA(0,1,1)(0,1,0)_6$ η εξίσωση υπολογισμού των τιμών πρόβλεψης της χρονοσειράς, είναι η εξής :

$$Y_t = e_t - \omega_1 e_{t-1} + Y_{t-1} + Y_{t-6} - Y_{t-7}$$

Για το μοντέλο $ARIMA(0,1,1)(0,1,0)_6$ που επιλέχθηκε, απαιτείται ο υπολογισμός μιας παραμέτρου, της ω_1 . Το στατιστικό πακέτο Minitab, μετά την ολοκλήρωση των υπολογισμών, εξάγει σχετική έκθεση αποτελεσμάτων (report) στην οποία συμπεριλαμβάνεται η τελική εκτίμηση της παραμέτρου του μοντέλου πρόβλεψης ARIMA.

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
MA 1	0.5705	0.1431	3.99	0.000

Differencing: 1 regular, 1 seasonal of order 6

Number of observations: Original series 42, after differencing 35

Residuals: SS = 2643150930 (backforecasts excluded)

MS = 77739733 DF = 34

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	12.8	21.7	*	*
DF	11	23	*	*
P-Value	0.305	0.538	*	*

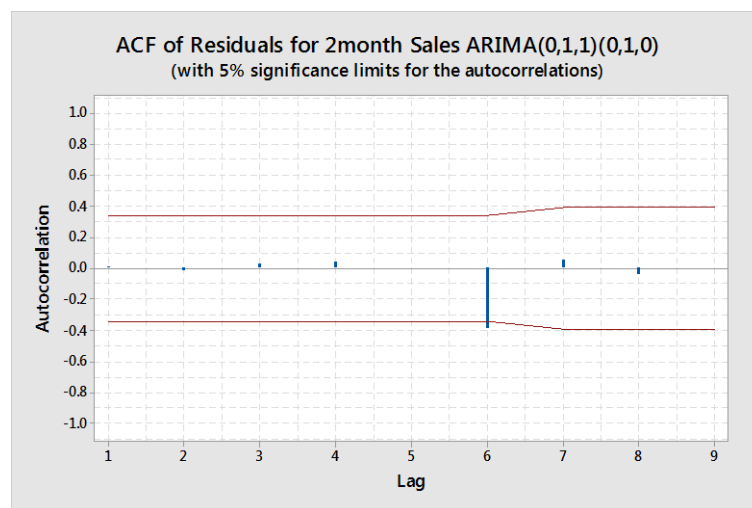
Η τελική εκτίμηση της τιμής της παραμέτρου, όπως προέκυψε από τον επαναληπτικό αλγόριθμο υπολογισμού για την ελαχιστοποίηση του αθροίσματος των τετραγώνων των σφαλμάτων, είναι $\omega_1 = 0.5705$.



Επομένως η τελική εξίσωση υπολογισμού του μοντέλου είναι :

$$Y_t = e_t - 0.5705\varepsilon_{t-1} + Y_{t-1} + Y_{t-6} - Y_{t-7}$$

Απομένει να ελέγξουμε την καταλληλότητα του δοκιμαστικού μοντέλου στο οποίο έχουμε καταλήξει. Για το διαγνωστικό έλεγχο, κατασκευάζεται το διάγραμμα αυτοσυσχετίσεων των σφαλμάτων πρόβλεψης (residuals) της χρονοσειράς και ελέγχουμε αν ικανοποιούνται οι συνθήκες επάρκειας του μοντέλου (έλεγχος για τυχαιότητα, ή ενδεχόμενη ύπαρξη προτύπων στα σφάλματα).



Σχήμα 5.4.6 : Συνάρτηση αυτοσυσχέτισης των υπολοίπων

Παρατηρούμε ότι το μοντέλο θεωρείται στατιστικά επαρκές αφού δεν υπάρχουν στατιστικά σημαντικοί συντελεστές αυτοσυσχέτισης των υπολοίπων (πλην της χρονικής καθυστέρησης των 6 περιόδων). Για τον έλεγχο της επάρκειας του μοντέλου, μπορεί να χρησιμοποιηθεί και το κριτήριο ελέγχου Ljung – Box Q statistic.

Μετά την ταυτοποίηση και τον έλεγχο επάρκειας του επιλεχθέντος μοντέλου, μπορούν να πραγματοποιηθούν προβλέψεις για την τιμή που θα λάβει η προς εξέταση μεταβλητή σε μια ή περισσότερες χρονικές περιόδους στο μέλλον. Οι προβλέψεις για το επόμενο έτος (2014) ανά δίμηνο, πραγματοποιούνται μέσω της εξίσωσης του μοντέλου :

$$Y_t = - 0.5705\varepsilon_{t-1} + Y_{t-1} + Y_{t-6} - Y_{t-7}$$

Υπενθυμίζουμε ότι η αρχική χρονοσειρά, αποτελείτο από 42 παρατηρήσεις, με την τελευταία παρατήρηση να αντιπροσωπεύει τις πωλήσεις του προϊόντος για το δίμηνο Νοέμβριο-Δεκέμβριο του

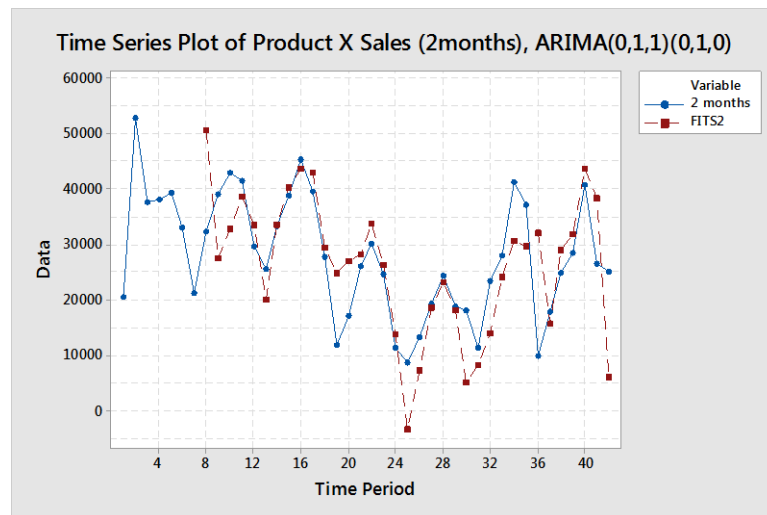


έτους 2013. Οι προβλέψεις για το επόμενο έτος (ανά δίμηνο για το 2014) όπως υπολογίστηκαν από το μοντέλο πρόβλεψης ARIMA μέσω της παραπάνω εξίσωσης, είναι οι εξής :

Forecasts from period 42

Period	Forecast	95% Limits		Actual
		Lower	Upper	
43	22291.4	5006.6	39576.3	
44	29235.5	10423.5	48047.4	
45	32948.4	12724.3	53172.5	
46	45291.4	23747.6	66835.3	
47	30829.4	8042.1	53616.7	
48	29524.5	5558.1	53490.8	

Στο πιο κάτω διάγραμμα απεικονίζονται οι εκτιμώμενες τιμές (FITS2) του μοντέλου σε σχέση με τη χρονοσειρά των πραγματικών πωλήσεων (2month Sales) ανά δίμηνο από το 2007 έως το 2013. Όπως έχει αποδειχθεί και μέσω του διαγνωστικού ελέγχου, παρατηρούμε ότι οι εκτιμώμενες τιμές από το μοντέλο ARIMA προσεγγίζουν σε ικανοποιητικό βαθμό το μοτίβο διακύμανσης των πραγματικών τιμών της χρονοσειράς.

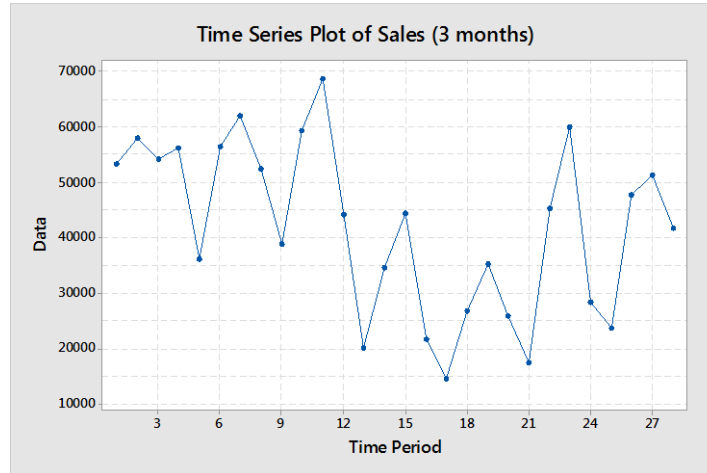


Σχήμα 5.4.7 : Χρονοσειρά πωλήσεων και εκτιμημένες τιμές από το μοντέλο ARIMA

Το χρησιμοποιούμενο μοντέλο ARIMA, παράγει εκτιμήσεις (προβλέψεις) οι οποίες προσεγγίζουν σε ικανοποιητικό βαθμό τις πραγματικά εκδηλωμένες τιμές της χρονοσειράς στις προηγούμενες περιόδους. Εκτός από τη γραφική παράσταση, το συμπέρασμα αυτό τεκμηριώνεται και στατιστικά από τις τιμές των $MAPE = 0.34$ και $MPE = -0.01$. Από στατιστικής πλευράς, το μοντέλο είναι σε θέση να προβλέψει με αποδεκτή ακρίβεια τις τιμές που θα λάβει η εξεταζόμενη μεταβλητή στο μέλλον.

5.4.2 Μοντέλο Πρόβλεψης τριμηνιαίων πωλήσεων Winters

Η μέθοδος Winters θα εφαρμοστεί για τα δεδομένα πωλήσεων του προϊόντος X, τα οποία έχουν αναχθεί σε περιόδους χρονικής διάρκειας 3 μηνών. Η χρονοσειρά πωλήσεων, ανά τρίμηνο από το 2007 έως το 2013, είναι η εξής :



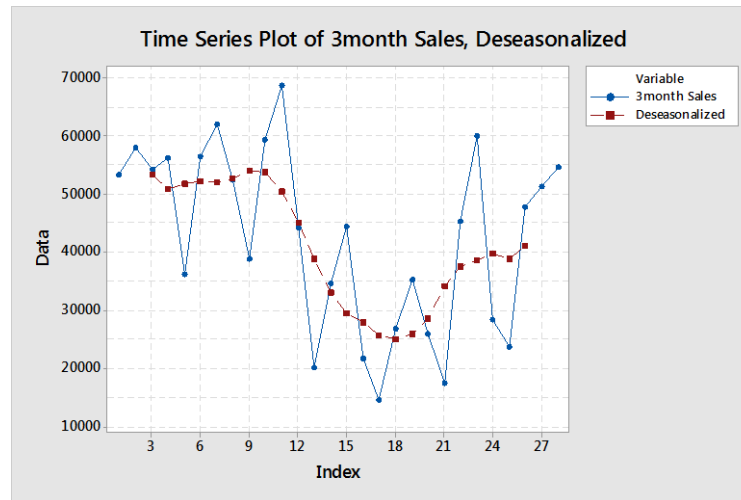
Σχήμα 5.4.8 : Χρονοσειρά τριμηνιαίων πωλήσεων

Παρατηρώντας την παραπάνω χρονοσειρά, επιβεβαιώνονται οι ισχυρισμοί για την ύπαρξη προτύπων τάσης και εποχικότητας στα δεδομένα πωλήσεων. Υπενθυμίζουμε ότι η χρονική διάρκεια της περιοδικότητας (p) είναι ο αριθμός των χρονικών περιόδων, μετά την πάροδο των οποίων επαναλαμβάνεται ο εποχικός κύκλος (επαναλαμβανόμενη διακύμανση). Για παράδειγμα, σε διαθέσιμα μηνιαία ιστορικά στοιχεία στα οποία υπάρχει εποχικότητα, η χρονική διάρκεια της περιοδικότητας είναι $p = 12$, ενώ για τριμηνιαία ιστορικά στοιχεία, όπως εν προκειμένω, προκύπτει ότι $p = 4$.

Το πρώτο βήμα που απαιτείται για την εκτίμηση των παραμέτρων του μοντέλου Winters της χρονοσειράς, αφορά την από-εποχικοποίηση των δεδομένων (deseasonalizing data). Υπενθυμίζουμε ότι ο υπολογισμός της από-εποχικοποιημένης ζήτησης πραγματοποιείται μέσω της ακόλουθης εξίσωσης ($p = 4$, άρτιος) :

$$\bar{D}_t = \frac{\left[D_{t-\left(\frac{p}{2}\right)} + D_{t+\left(\frac{p}{2}\right)} + \sum_{i=t+1-\left(\frac{p}{2}\right)}^{t-1+\left(\frac{p}{2}\right)} 2D_i \right]}{2p}$$

Τα αποτελέσματα από την αφαίρεση της εποχικότητας από τη ζήτηση παρουσιάζονται στο σχετικό πίνακα στο **Παράρτημα Δ**. Η γραφική παράσταση των από-εποχικοποιημένων δεδομένων σε σύγκριση με τα αρχικά δεδομένα της χρονοσειράς είναι η ακόλουθη :



Σχήμα 5.4.9 : Χρονοσειρά τριμηνιαίων πωλήσεων και από-εποχικοποιημένη ζήτηση

Με γραμμική παλινδρόμηση στην καμπύλη δεδομένων της από-εποχικοποιημένης ζήτησης, εκτιμώνται οι τιμές των L_0 και T_0 , αφού ισχύει η σχέση :

$$\bar{D}_t = L + Tt$$

Η εξίσωση που προκύπτει από την εφαρμογή γραμμικής παλινδρόμησης είναι :

$$\bar{D}_t = 55450 - 1015.3t$$

επομένως λαμβάνουμε $L_0 = 55450$ και $T_0 = -1015.3$.

Οι τιμές της νέας ζήτησης χωρίς εποχικότητα για κάθε χρονική περίοδο δίνονται από τη σχέση :

$$\bar{D}_t' = L + Tt$$

και ο συντελεστής εποχικότητας, για κάθε περίοδο t της προς μελέτη χρονοσειράς, δίνεται από τη σχέση μεταξύ πραγματικής και από-εποχικοποιημένης ζήτησης :

$$\bar{S}_t = \frac{D_t}{\bar{D}_t'}$$

Τέλος, ως αρχικές τιμές των συντελεστών εποχικότητας, λαμβάνονται οι μέσοι όροι των συντελεστών των αντίστοιχων εποχικών χρονικών περιόδων μέσω της σχέσης :

$$S_i = \frac{(\sum_{j=0}^{r-1} \bar{S}_{jp+i})}{r} \quad \text{όπου } r = \frac{N}{p}$$

Στο μοντέλο Winters χρησιμοποιούνται τρεις εξισώσεις για την εξομάλυνση των στοιχείων του επιπέδου, της τάσης και της εποχικότητας για κάθε χρονική περίοδο αντίστοιχα :



$$L_{t+1} = \alpha \left(\frac{D_{t+1}}{S_{t+1}} \right) + (1 - \alpha)(L_t + T_t)$$

$$T_{t+1} = \beta(L_{t+1} - L_t) + (1 - \beta)T_t$$

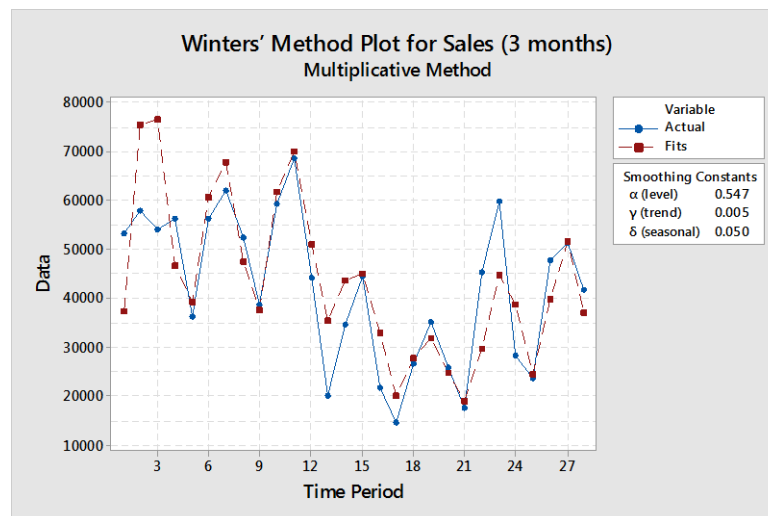
$$S_{t+p+1} = \gamma \left(\frac{D_{t+1}}{L_{t+1}} \right) + (1 - \gamma)S_{t+1}$$

όπου α , β , γ , οι συντελεστές εξομάλυνσης επιπέδου, τάσης και εποχικότητας αντίστοιχα για τους οποίους ισχύει $0 \leq \alpha, \beta, \gamma \leq 1$. Για την εύρεση των βέλτιστων τιμών των συντελεστών, χρησιμοποιήθηκε το πρόσθετο εργαλείο Solver του Excel, θέτοντας ως κριτήριο την ελαχιστοποίηση του MSE.

Τελικά, η πρόβλεψη F_{t+1} για την επόμενη χρονική περίοδο γίνεται συνδυάζοντας τα παραπάνω υπολογισθέντα στοιχεία, σύμφωνα με τη σχέση :

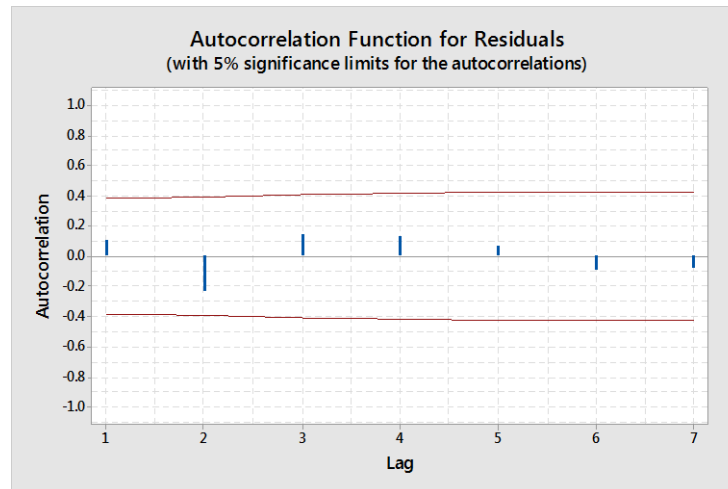
$$F_{t+1} = (L_t + T_t)S_{t+1}$$

Η γραφική παράσταση που ακολουθεί, απεικονίζει τις πραγματικές τιμές της χρονοσειράς (Actual) σε σχέση με τις εκτιμημένες (Fits) από το μοντέλο Winters για τις συγκεκριμένες τιμές των συντελεστών εξομάλυνσης. Οι εκτιμημένες τιμές προσεγγίζουν σε ικανοποιητικό βαθμό τις πραγματικές, συμπεράσμα που τεκμηριώνεται και στατιστικά από τις τιμές των MAPE και MPE.



Σχήμα 5.4.10 : Χρονοσειρά τριμηνιαίων πωλήσεων και εκτιμημένες τιμές από το μοντέλο Winters

Εναλλακτικά, μπορούμε να ελέγξουμε την επάρκεια του χρησιμοποιούμενου μοντέλου Winters μέσω της συνάρτησης αυτοσυσχέτισης των υπολοίπων.



Σχήμα 5.4.11 : Συνάρτηση αυτοσυσχέτισης των υπολοίπων

Εφόσον οι υπολογισθέντες συντελεστές αυτοσυσχέτισης για των υπολοίπων (residuals) δεν είναι στατιστικά σημαντικοί (όρια : $\pm \frac{2}{\sqrt{n}} = \pm \frac{2}{\sqrt{28}} = \pm 0.38$) και δεν παρουσιάζουν συγκεκριμένο μοτίβο διακύμανσης, το μοντέλο Winters μπορεί να θεωρηθεί στατιστικά επαρκές και κατάλληλο για προβλέψεις των μελλοντικών τιμών της προς εξέταση μεταβλητής.

Στο σημείο αυτό κρίνεται σκόπιμο να διασαφηνιστεί το εξής σημαντικό. Το γεγονός ότι ένα μοντέλο πρόβλεψης αποδεικνύεται ότι είναι στατιστικά επαρκές, δεν συνεπάγεται ότι είναι σε θέση να λάβει υπόψη όλους τους πιθανούς παράγοντες που πρόκειται να μεταβάλουν και να διαμορφώσουν τη ζήτηση στο χρονικό ορίζοντα των προβλέψεων. Αυτό που συνεπάγεται είναι ότι το μοντέλο αποδεικνύεται επαρκές για την προεκβολή στο μέλλον του μοτίβου διακύμανσης της χρονοσειράς των πωλήσεων, όπως αυτό διαμορφώθηκε στο παρελθόν από τους παράγοντες που επηρεάζουν την τιμή της εξεταζόμενης μεταβλητής, π.χ. τάση, εποχικότητα, επίπεδο.



5.5 Σύγκριση προβλέψεων ARIMA και Winters

Στη συνέχεια, με βάση τα μοντέλα στα οποία καταλήξαμε μέσω της ανάλυσης της χρονοσειράς πωλήσεων του προϊόντος για τα έτη 2007-2013, θα πραγματοποιηθεί σύγκριση των παραγόμενων προβλέψεων για τις πωλήσεις του έτους 2014 ανά μήνα, δίμηνο και τρίμηνο, με τις πραγματικές πωλήσεις όπως αυτές έχουν εκδηλωθεί. Η σύγκριση θα πραγματοποιηθεί μεταξύ των προβλέψεων που παρήγαγαν τα μοντέλα ARIMA και Winters (βλ. **Παράρτημα Ε**), αφού όπως αποδείχθηκε, το μοντέλο Holts είναι στατιστικά ανεπαρκές.

Επαναλαμβάνεται ότι οι προβλέψεις του μοντέλου Winters για τις μελλοντικές τιμές της εξεταζόμενης μεταβλητής, δίνονται από τη σχέση :

$$F_{t+1} = (L_t + T_t)S_{t+1}$$

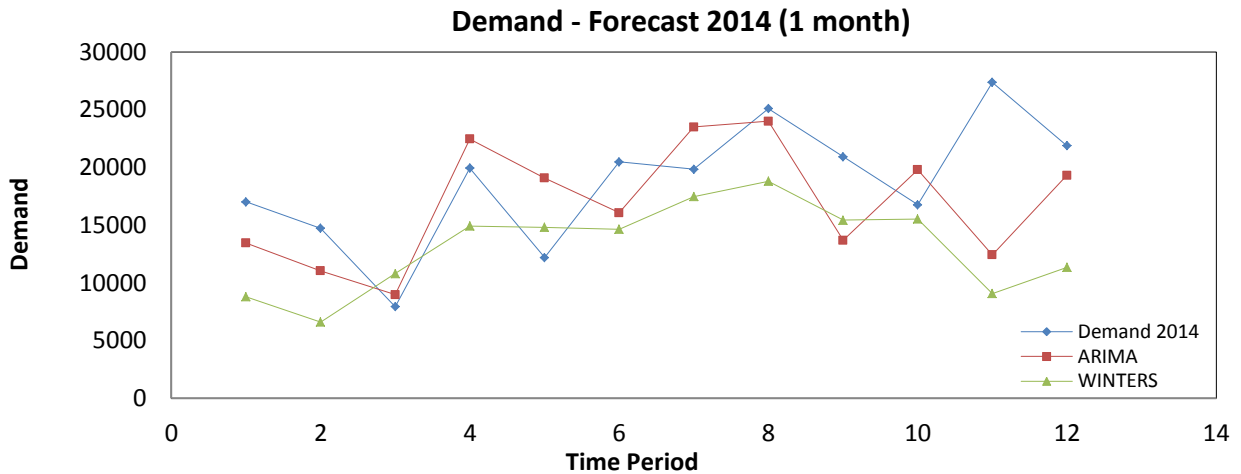
όπου L_t , T_t και S_{t+1} είναι οι εκτιμήσεις του επιπέδου, της τάσης και της εποχικότητας αντίστοιχα. Στη μέθοδο Winters, χρησιμοποιούνται ισάριθμοι αντίστοιχοι συντελεστές εξομάλυνσης α , β , γ οι τιμές των οποίων προσδιορίζονται με επαναληπτική διαδικασία, έχοντας ως στόχο την ελαχιστοποίηση του αθροίσματος των τετραγώνων των σφαλμάτων.

Στο μοντέλο πρόβλεψης ARIMA οι προβλέψεις που διενεργούνται εξαρτώνται και από τις τρέχουσες και παρελθοντικές πραγματικές τιμές της μεταβλητής αλλά και από τις τρέχουσες και παρελθοντικές τιμές των σφαλμάτων πρόβλεψης, με βαρύτητα που καθορίζεται από τους προς εκτίμηση συντελεστές. Στο χρησιμοποιούμενο μοντέλο $ARIMA(0,1,1)(0,1,0)_s$ οι προβλέψεις υπολογίζονται μέσω της σχέσης:

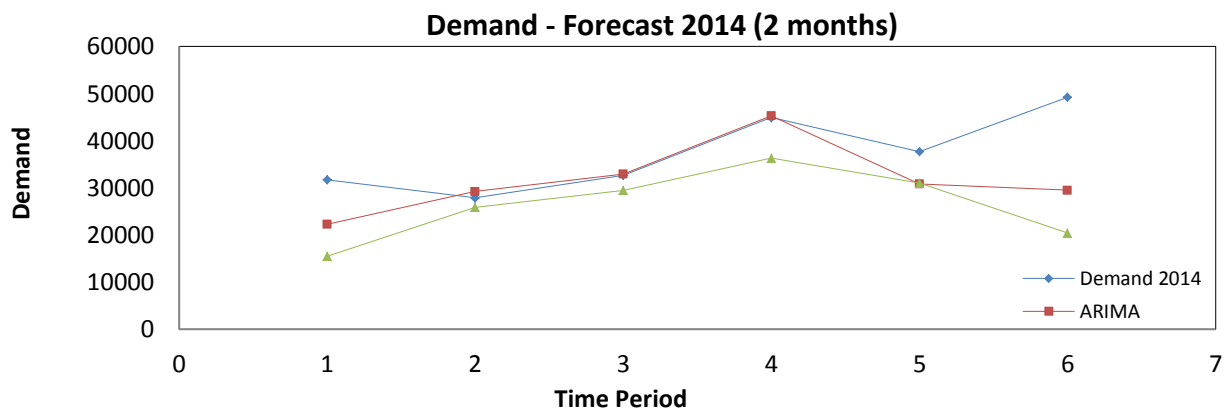
$$Y_t = e_t - \omega_1 e_{t-1} + Y_{t-1} + Y_{t-s} - Y_{t-s-1}$$

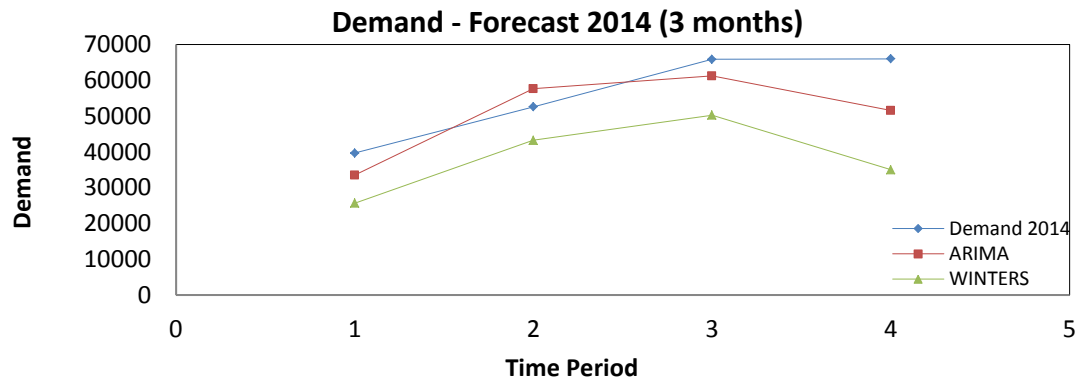
όπου s η χρονική διάρκεια της εποχικότητας ($s = 12$ για μηνιαία δεδομένα | $s = 4$ για τριμηνιαία δεδομένα πωλήσεων), Y_{t-k} η παρατήρηση της χρονοσειράς κατά τη χρονική περίοδο $t - k$, e_{t-k} το σφάλμα (διαφορά μεταξύ εκτιμημένης τιμής από το μοντέλο και πραγματικά εκδηλωμένης τιμής) κατά τη χρονική περίοδο $t - k$ και ω_1 η προς εκτίμηση παράμετρος.

Σημειώνεται ότι η ανάπτυξη των μοντέλων για προβλέψεις που αφορούν τις πωλήσεις του προϊόντος ανά μήνα, δίμηνο ή τρίμηνο, πραγματοποιήθηκαν εφαρμόζοντας τις ίδιες μεθοδολογίες (Winters και Box-Jenkins) με μόνη διαφορά την αναγωγή των μηνιαίων πωλήσεων σε δύο ή τρεις μήνες αντίστοιχα. Τα αποτελέσματα της σύγκρισης των προβλέψεων για κάθε μοντέλο παρουσιάζονται στη συνέχεια.



Οι γραφικές παραστάσεις σύγκρισης των προβλέψεων για τα μοντέλα Winters και ARIMA, με τις πραγματικές πωλήσεις για το έτος 2014, επαληθεύουν την από στατιστικής πλευράς τεκμηρίωση της επάρκειας και των δυο χρησιμοποιούμενων μοντέλων. Οι προβλέψεις που προέκυψαν κρίνονται ιδιαίτερα ικανοποιητικές αν αναλογιστεί κανείς αφενός μεν ότι πραγματοποιούνται για μεμονωμένο προϊόν και όχι για ομάδα προϊόντων (προβλέψεις που πραγματοποιούνται για μια ομάδα προϊόντων, είναι ακριβέστερες από τις προβλέψεις για ένα μεμονωμένο προϊόν λόγω της πιθανής αντιστάθμισης θετικών με αρνητικών αποκλίσεων στην τελική ζήτηση) και αφετέρου δε ότι κατά το έτος 2014 λόγω διαφόρων παραγόντων (όπως π.χ. έντονες προωθητικές ενέργειες, κλείσιμο συμβολαίων με νέους πελάτες) σε ορισμένες χρονικές περιόδους, είχε παρατηρηθεί ασυνήθιστα αυξημένη ζήτηση σε σχέση με αντίστοιχες χρονικές περιόδους προηγούμενων ετών. Η παρατήρηση αυτή πρέπει οπωσδήποτε να αναλυθεί περεταίρω, ούτως ώστε να προκύψουν κρίσιμα ποιοτικά συμπεράσματα που να εξηγούν την έντονη μεταβολή της ζήτησης που σημειώθηκε κατά τους μήνες αυτούς. Η ανάλυση των αποκλίσεων κρίνεται αναγκαία για να διαπιστωθεί κατά πόσο το μοντέλο πρέπει να επανεκτιμηθεί και να ελεγχθεί ως προς την επάρκειά του.





Συγκρίνοντας τις προβλέψεις των μοντέλων Winters και ARIMA, είναι εμφανής η επίτευξη μεγαλύτερης ακρίβειας προβλέψεων με το μοντέλο ARIMA. Η μέθοδος Winters, απλούστερη από πλευράς κατανόησης και υπολογιστικής υλοποίησης, παράγει προβλέψεις οι οποίες πηγάζουν από την εξομάλυνση, μέσω των χρησιμοποιούμενων συντελεστών, του επιπέδου και των προτύπων τάσης και εποχικότητας που παρουσιάζει διαχρονικά η χρονοσειρά. Από την άλλη, στο μοντέλο ARIMA οι προβλέψεις στηρίζονται στις προηγούμενες παρατηρήσεις και σφάλματα πρόβλεψης (αυτοπαλινδρόμηση και κινητός μέσος). Αυτό έχει ως αποτέλεσμα οι προβλέψεις των μοντέλων ARIMA, να υπολογίζονται με τη χρήση πιο πρόσφατων δεδομένων, προσδίδοντας τους τη δυνατότητα να προσαρμόζονται με μεγαλύτερη ακρίβεια στις μεταβολές της ζήτησης.

Το πλεονέκτημα των μοντέλων ARIMA, εντοπίζεται επίσης στο γεγονός ότι για την ανάπτυξή τους και για την πραγματοποίηση προβλέψεων δεν χρησιμοποιούνται ανεξάρτητες μεταβλητές, παρά μόνο γίνεται χρήση της ποιοτικής και ποσοτικής πληροφορίας που εμπερικλείεται στη χρονοσειρά. Συγκεκριμένα, στην ανάλυση ARIMA μελετώνται εκτενώς οι συναρτήσεις αυτοσυσχέτισης και μερικής αυτοσυσχέτισης, με αποτέλεσμα να αποσαφηνίζεται με μεγαλύτερη ακρίβεια η σχέση που υπάρχει μεταξύ των παρατηρήσεων της χρονοσειράς (διαχωρισμός συστηματικού και τυχαίου μέρους της παρατήρησης). Επιπρόσθετα, στο τελικό στοχαστικό μοντέλο πρόβλεψης ARIMA που θα επιλεγεί, υπεισέρχονται μεταβλητές (προηγούμενες παρατηρήσεις και σφάλματα πρόβλεψης) που είναι πιο αντιπροσωπευτικές για τον υπολογισμό των προβλέψεων που αφορούν την τελική ζήτηση του προϊόντος κατά τις μελλοντικές χρονικές περιόδους.

Ο αναλυτής που πραγματοποιεί τις προβλέψεις, στην περίπτωση που οι αποκλίσεις είναι εκτός των αποδεκτών προκαθορισμένων ορίων, πρέπει να ερμηνεύσει ποιοτικά τις αποκλίσεις αναζητώντας εάν τα πιθανά αίτια των μεγάλων αποκλίσεων οφείλονται σε ανεπάρκεια του χρησιμοποιούμενου μοντέλου ή σε εξωγενείς αστάθμητους παράγοντες που επηρέασαν τη ζήτηση.



6. Συμπεράσματα

Ανάγκη : Μέσα στο σύγχρονο, συνεχώς μεταβαλλόμενο ευρύτερο πλαίσιο της αγοράς που δραστηριοποιούνται σήμερα οι επιχειρήσεις, η επίτευξη προβλέψεων που εκ των υστέρων θα μπορούσαν να χαρακτηριστούν ακριβείς, δυσχεραίνεται ακόμη περισσότερο. Η ανάγκη όμως για λήψη αποφάσεων σχετικά με τις κρίσιμες λειτουργίες μιας επιχείρησης σε βραχυπρόθεσμο, μεσοπρόθεσμο και μακροπρόθεσμο επίπεδο, εξακολουθεί να είναι επιτακτική. Αρκεί κανείς να αναλογιστεί τη σημασία που έχει για μια παραγωγική make-to-stock επιχείρηση, η λήψη απόφασης σχετικά με τις ποσότητες που θα πρέπει να παραχθούν στους επόμενους μήνες, ούτως ώστε να ικανοποιηθεί με τον πλέον αποδοτικό και οικονομικό (επίπεδα αποθεμάτων) τρόπο η αναμενόμενη ζήτηση.

Για τη λήψη αποφάσεων ζωτικής σημασίας για τη λειτουργία των επιχειρήσεων, εκτός από τη σχετική εμπειρία και κατάρτιση των στελεχών, απαιτείται να υπάρχουν διαθέσιμες ποσοτικές ή/και ποιοτικές πληροφορίες σχετικά με την πιθανή συμπεριφορά της εξεταζόμενης μεταβλητής στο μέλλον. Η μόνη πηγή για την εξαγωγή αυτών των πληροφοριών, είναι η ανάλυση και μελέτη της συμπεριφοράς της εξεταζόμενης μεταβλητής στο παρελθόν (π.χ. για πρόβλεψη ζήτησης, πηγή πληροφορίας είναι η ανάλυση και μελέτη της χρονοσειράς πωλήσεων στο παρελθόν). Όλες λοιπόν οι διαθέσιμες μέθοδοι πρόβλεψης, στοχεύουν ακριβώς στην προσπάθεια μεταβίβασης της ποιοτικής και ποσοτικής πληροφορίας που εμπεριέχεται στη χρονοσειρά της εξεταζόμενης μεταβλητής, από το παρελθόν, στο μέλλον, αποτελώντας με τον τρόπο αυτό ένα χρήσιμο εργαλείο για την υποστήριξη της λήψης επιχειρηματικών αποφάσεων.

Ακρίβεια : Η ακρίβεια των μεθόδων πρόβλεψης, είναι ένα σημαντικό κεφάλαιο για το οποίο υπάρχουν ποικίλες απόψεις και προσεγγίσεις. Σε δημοσίευση στο *Control Magazine, Institute of Operations Management*, αναφέρεται το εξής : "Συχνά λέγεται ότι υπάρχουν κατηγορίες προγνώσεων : οι τυχερές και οι λανθασμένες" . Είναι σημαντικό να αντιληφθεί ο εκάστοτε αναλυτής, ότι καμία μέθοδος πρόβλεψης δεν μπορεί να θεωρηθεί πανάκεια, αφού εν γένει όλες οι μέθοδοι στηρίζονται στην προσπάθεια να προβλεφθεί το μέλλον, μελετώντας το παρελθόν. Αυτό από μόνο του αποτελεί την κυριότερη πηγή σφαλμάτων, αφού σε καμιά μέθοδο δεν μπορεί να συνυπολογιστεί η επίδραση παραγόντων που θα εμφανιστούν στο χρονικό ορίζοντα της πρόβλεψης και θα διαμορφώσουν την τιμή της μεταβλητής.



Επομένως, η ακρίβεια μιας μεθόδου πρόβλεψης είναι στατιστικά υποκειμενική έννοια. Έχουν όμως αναπτυχθεί κριτήρια παρακολούθησης και μέτρησης της ακρίβειας των μεθόδων, στόχος των οποίων είναι να μελετάται κατά πόσον οι αποκλίσεις της χρησιμοποιούμενης μεθόδου πρόβλεψης από τις πραγματικά εκδηλωμένες τιμές της μεταβλητής, βρίσκονται εντός των αποδεκτών προκαθορισμένων ορίων. Με τα εργαλεία αυτά αξιολογείται στατιστικά η επάρκεια της μεθόδου πρόβλεψης και αναγνωρίζεται η ανάγκη για επανεκτίμηση των παραμέτρων του μοντέλου ή ακόμη και για αναζήτηση ενός διαφορετικού, πιο προηγμένου μοντέλου πρόβλεψης.

Επιλογή : Ο μόνος τρόπος με τον οποίο ο ανθρώπινος παράγοντας, εν προκειμένω ο εκάστοτε αναλυτής, μπορεί να επηρεάσει την ακρίβεια των προβλέψεων, είναι μέσω της επιλογής της καταλληλότερης για κάθε περίπτωση μεθόδου. Η καταλληλότερη μέθοδος πρόβλεψης είναι διαφορετική για κάθε χρονοσειρά και κατ' επέκταση για κάθε εξεταζόμενη μεταβλητή. Ο ερευνητής θα πρέπει να διαθέτει εμπειρία, οξυδέρκεια, διορατικότητα και παρατηρητικότητα, για να επιλέξει την καταλληλότερη κάθε φορά μέθοδο αξιοποιώντας έτσι με τον καλύτερο τρόπο τα διαθέσιμα στατιστικά εργαλεία για την πραγματοποίηση προβλέψεων. Ενδεικτικό είναι το παράδειγμα της μεθόδου Holts με την οποία επιχειρήσαμε να πραγματοποιήσουμε προβλέψεις σε μια χρονοσειρά η οποία παρουσίαζε έντονο το στοιχείο της εποχικότητας. Το αποτέλεσμα ήταν να αποδειχθεί η ανεπάρκεια της χρησιμοποιούμενης μεθόδου και η αναζήτηση εναλλακτικής, καταλληλότερης για πρόβλεψη σε χρονοσειρές με εποχικότητα, όπως οι μέθοδοι Winters και ARIMA. Οι αποφάσεις σχετικά με την επιλογή του μοντέλου πρόβλεψης, είναι καθοριστικής σημασίας για την επίτευξη 'υποκειμενικά' υψηλής ακρίβειας προβλέψεων.

Αξιολόγηση : "Πρόγνωση καλείται η τέχνη του να λες τι θα συμβεί και στη συνέχεια να εξηγείς γιατί δε συνέβη". Η διαδικασία των προβλέψεων απαιτεί συνεχή ενασχόληση και αξιολόγηση των αποτελεσμάτων. Δεν αρκεί απλά να αποδειχθεί στατιστικά η ανεπάρκεια ενός μοντέλου λόγω της ύπαρξης μεγάλων αποκλίσεων των προβλέψεων από τις πραγματικές τιμές που εκδηλώθηκαν. Απαιτείται ποιοτική αιτιολόγηση των παραγόντων που οδήγησαν στην ανεπάρκεια του μοντέλου ούτως ώστε να αποκρυπτογραφηθεί στο μέγιστο δυνατό βαθμό το πλαίσιο της αγοράς που δραστηριοποιείται η επιχείρηση. Το τελευταίο, είναι απαραίτητη προϋπόθεση για την αποδοτική λειτουργία των επιχειρήσεων. Κατ' επέκταση η υιοθέτηση ενός ολοκληρωμένου συστήματος πρόβλεψης καθίσταται αναγκαία επένδυση για κάθε επιχείρηση.



Παράρτημα Α

Δεδομένα πωλήσεων ανά μήνα του προϊόντος Χ											
Period	Year	Time	D_t	Period	Year	Time	D_t	Period	Year	Time	D_t
1	2007	1	11393.04	37	2010	1	7120.092	73	2013	1	10165.01
2		2	9112.04	38		2	4531.044	74		2	7753.032
3		3	32675.03	39		3	8361.048	75		3	5686.032
4		4	20263.03	40		4	8709.012	76		4	19176.05
5		5	19181.04	41		5	12009.07	77		5	15790.01
6		6	18449.03	42		6	13920.04	78		6	12785.03
7		7	20472.05	43		7	16571.05	79		7	20204
8		8	17557.04	44		8	13514.1	80		8	20714.03
9		9	16076.03	45		9	14240.04	81		9	10397.03
10		10	23196.01	46		10	10363.04	82		10	16509.5
11		11	13820.99	47		11	6326.04	83		11	9132.024
12		12	19180.03	48		12	4895.028	84		12	16019.09
13	2008	1	12358.04	49	2011	1	3970.032	85	2014	1	17014.07
14		2	8729.05	50		2	4720.032	86		2	14723.02
15		3	15099.06	51		3	5843.04	87		3	7932
16		4	17179.01	52		4	7340.04	88		4	19945.04
17		5	17348.09	53		5	10229	89		5	12199.03
18		6	21784.09	54		6	9105.048	90		6	20484.04
19		7	20019.98	55		7	10254.04	91		7	19849.03
20		8	23004.06	56		8	14041.06	92		8	25289.02
21		9	18937.06	57		9	10910.03	93		9	20925.02
22		10	22711.07	58		10	7897.032	94		10	16761.01
23		11	14022.01	59		11	7827.996	95		11	27368.03
24		12	15553.03	60		12	10160.04	96		12	21876.01
25	2009	1	13946.03	61	2012	1	7471.056				
26		2	11601.05	62		2	3705.012				
27		3	13144.04	63		3	6238.02				
28		4	20278.97	64		4	17143.08				
29		5	18813.04	65		5	15087				
30		6	20117.05	66		6	12915.04				
31		7	26450.10	67		7	15081.02				
32		8	19009.06	68		8	26179.03				
33		9	23283.11	69		9	18676.02				
34		10	16390.04	70		10	18500				
35		11	13867.06	71		11	3947.04				
36		12	13918.04	72		12	5831.028				



ARIMA(0,1,1) (0,1,0) – 1 month – Εποχική και Κανονική Διαφύριση			
Time (t)	Y_t	$\Delta_{12}Y_t = Y_t - Y_{t-12}$	$\Delta\Delta_{12}Y_t = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$
1	11393.04		
2	9112.044		
3	32675.028		
4	20263.032		
5	19181.04		
6	18449.028		
7	20472.048		
8	17557.044		
9	16076.028		
10	23196.012		
11	13820.988		
12	19180.032		
13	12358.044	965.004	
14	8729.052	-382.992	-1347.996
15	15099.06	-17575.968	-17192.976
16	17179.008	-3084.024	14491.944
...
84	16019.088	10188.06	5003.076

$$Y_t = -0.6736\varepsilon_{t-1} + Y_{t-1} + Y_{t-12} - Y_{t-13}$$

ARIMA(0,1,1) (0,1,0) – 1 month – Εκτιμημένες τιμές και Σφάλματα Πρόβλεψης			
Time (t)	Y_t	Fits	Errors
1	11393.04		
2	9112.044		
3	32675.028		
4	20263.032		
5	19181.04		
6	18449.028		
7	20472.048		
8	17557.044		
9	16076.028		
10	23196.012		
11	13820.988		
12	19180.032		
13	12358.044		
14	8729.052	7728.03204	1001.02
15	15099.06	31617.74181	-16518.7
16	17179.008	13814.16596	3364.8
17	17348.088	13830.43439	3517.7
18	21784.092	14246.55942	7537.5
19	20019.984	18729.77626	1290.2
...
84	16019.88	5782.22	10236.9



Δεδομένα πωλήσεων ανά δίμηνο του προϊόντος Χ							
Period	Year	Time	Y_t	Period	Year	Time	Y_t
1	2007	1	20505.1	22		4	30085.2
2		2	52938.1	23		5	24603.1
3		3	37630.1	24		6	11221.1
4		4	38029.1	25	2011	1	8690.1
5		5	39272.0	26		2	13183.1
6		6	33001.0	27		3	19334.1
7	2008	1	21087.1	28		4	24295.1
8		2	32278.1	29		5	18807.1
9		3	39132.2	30		6	17988.0
10		4	43024.0	31	2012	1	11176.1
11		5	41648.1	32		2	23381.1
12		6	29575.0	33		3	28002.0
13	2009	1	25547.1	34		4	41260.1
14		2	33423.0	35		5	37176.0
15		3	38930.1	36		6	9778.1
16		4	45459.2	37	2013	1	17918.0
17		5	39673.2	38		2	24862.1
18		6	27785.1	39		3	28575.0
19	2010	1	11651.1	40		4	40918.0
20		2	17070.1	41		5	26906.5
21		3	25929.1	42		6	25151.1

$$Y_t = -0.5705\varepsilon_{t-1} + Y_{t-1} + Y_{t-6} - Y_{t-7}$$

ARIMA(0,1,1) (0,1,0) – 2 months – Εποχική και Κανονική Διαφοράση

Time (t)	Y_t	$\Delta_6 Y_t = Y_t - Y_{t-6}$	$\Delta\Delta_6 Y_t = Y_t - Y_{t-1} - Y_{t-6} + Y_{t-7}$
1	20505.1		
2	52938.1		
3	37630.1		
4	38029.1		
5	39272.0		
6	33001.0		
7	21087.1	582.0	
8	32278.1	-20660.0	-21242.0
9	39132.2	1502.1	22162.1
10	43024.0	4995.0	3492.8
11	41648.1	2376.1	-2618.9
12	29575.0	-3426.0	-5802.1
13	25547.1	4460.0	7886.0
14	33423.0	1144.9	-3315.0
15	38930.1	-202.1	-1347.0
16	45459.2	2435.1	2637.2
...
84	25151.1	15373.0	25642.5



Παράρτημα Β

Winter's Method – Υπολογισμοί παραμέτρων

Period	Year	Time	D_t	\bar{D}_t	\bar{D}'_t	\bar{S}_t	S_t	L_t	T_t	F_t
0								18407.0	-112.8	
1	2007	1	11393.04		18294.2	0.62	0.67	17953.5	-114.5	12290.8
2		2	9112.04		18181.4	0.50	0.51	17864.7	-114.4	9061.0
3		3	32675.03		18068.6	1.81	0.82	23366.8	-86.3	14579.7
4		4	20263.03		17955.8	1.13	1.17	21762.0	-93.9	27229.5
5		5	19181.04		17843.0	1.07	1.15	20395.8	-100.2	24921.1
6		6	18449.03		17730.2	1.04	1.14	19249.1	-105.5	23125.8
7		7	20472.05	18488.2	17617.4	1.16	1.39	18011.3	-111.1	26657.0
8		8	17557.04	18512.4	17504.6	1.00	1.50	16313.7	-119.1	26912.7
9		9	16076.03	17764.1	17391.8	0.92	1.22	15421.7	-122.9	19778.8
10		10	23196.01	16903.3	17279.0	1.34	1.27	16046.6	-119.2	19464.1
11		11	13820.99	16698.4	17166.2	0.81	0.74	16620.1	-115.7	11806.8
12		12	19180.03	16761.0	17053.4	1.12	0.96	17408.8	-111.2	15786.7
13	2008	1	12358.04	16881.1	16940.6	0.73	0.67	17590.2	-109.7	11589.0
14		2	8729.05	17089.2	16827.8	0.52	0.51	17404.4	-110.1	8880.7
15		3	15099.06	17435.4	16715.0	0.90	0.85	17412.7	-109.5	14704.0
16		4	17179.01	17534.4	16602.2	1.03	1.16	16674.9	-112.7	20031.9
17		5	17348.09	17522.6	16489.4	1.05	1.14	16220.7	-114.4	18875.0
18		6	21784.09	17379.8	16376.6	1.33	1.13	16913.1	-110.4	18206.6
19		7	20019.98	17294.9	16263.8	1.23	1.38	16218.4	-113.3	23182.4
20		8	23004.06	17480.7	16151.0	1.42	1.48	15956.2	-114.0	23869.8
21		9	18937.06	17518.9	16038.2	1.18	1.21	15785.5	-114.3	19206.6
22		10	22711.07	17566.6	15925.4	1.43	1.28	16196.1	-111.7	20073.6
23		11	14022.01	17756.8	15812.6	0.89	0.75	16777.1	-108.2	11995.8
24		12	15553.03	17748.4	15699.8	0.99	0.96	16533.4	-108.9	16065.1
25	2009	1	13946.03	17946.9	15587.0	0.89	0.67	17531.1	-103.4	11030.8
26		2	11601.05	18048.3	15474.2	0.75	0.51	18810.0	-96.4	8848.3
27		3	13144.04	18063.0	15361.4	0.86	0.85	17880.1	-100.6	15926.6
28		4	20278.97	17980.7	15248.6	1.33	1.15	17737.1	-100.8	20470.0
29		5	18813.04	17710.8	15135.8	1.24	1.14	17361.6	-102.2	20037.3
30		6	20117.05	17636.3	15023.0	1.34	1.14	17364.9	-101.7	19646.1
31		7	26450.10	17283.7	14910.2	1.77	1.37	17775.5	-99.1	23692.3
32		8	19009.06	16704.7	14797.4	1.28	1.48	16444.2	-105.3	26162.9
33		9	23283.11	16210.8	14684.6	1.59	1.21	17072.1	-101.6	19798.4
34		10	16390.04	15529.5	14571.8	1.12	1.29	15890.7	-107.0	21840.8
35		11	13867.06	14763.9	14459.0	0.96	0.75	16471.6	-103.6	11842.5
36		12	13918.04	14222.2	14346.2	0.97	0.96	15881.2	-106.0	15756.2
...	
73	2013	1	10165.01	10172.6	1.00	0.67	11138.9	-106.8	6516.4	13510.9
74		2	7753.032	10059.8	0.77	0.51	12130.2	-101.3	5576.1	13496.6
75		3	5686.032	9947.0	0.57	0.83	10708.1	-107.9	9987.5	12924.0
76		4	19176.05	9834.2	1.95	1.16	12120.6	-100.3	12271.9	12496.1
77		5	15790.01	9721.4	1.62	1.16	12433.6	-98.2	13913.7	12629.2
78		6	12785.03	9608.6	1.33	1.15	12021.4	-99.8	14203.3	13269.7
79		7	20204	9495.8	2.13	1.38	12601.6	-96.4	16510.1	
80		8	20714.03	9383.0	2.21	1.50	12831.2	-94.7	18792.5	
81		9	10397.03	9270.2	1.12	1.24	11625.0	-100.3	15808.0	
82		10	16509.5	9157.4	1.80	1.26	11931.5	-98.3	14501.6	
83		11	9132.024	9044.6	1.01	0.74	11960.3	-97.6	8763.0	
84		12	16019.09	8931.8	1.79	0.93	13214.6	-90.9	11070.5	



Winters Method – Υπολογισμοί Σφαλμάτων Πρόβλεψης					
Period	Error	 Error 	Error²	 Error /Demand	Error/Demand
0					
1	-897.8	897.8	805983.73	0.08	-0.08
2	51.1	51.1	2609.01	0.01	0.01
3	18095.4	18095.4	327441778.84	0.55	0.55
4	-6966.5	6966.5	48532056.78	0.34	-0.34
5	-5740.1	5740.1	32948381.84	0.30	-0.30
6	-4676.8	4676.8	21872331.30	0.25	-0.25
7	-6185.0	6185.0	38253816.69	0.30	-0.30
8	-9355.6	9355.6	87527701.53	0.53	-0.53
...
84	4948.6	4948.6	24488604.53	0.31	0.31
Total	8654.9	248169.5	1355921424.8	20.1	-4.5

Winters Errors	
MAD	2954.399
MSE	16141922
RMSE	4017.701
MAPE	0.23905
MPE	-0.05313

$$L_{t+1} = \alpha \left(\frac{D_{t+1}}{S_{t+1}} \right) + (1 - \alpha)(L_t + T_t)$$

$$T_{t+1} = \beta(L_{t+1} - L_t) + (1 - \beta)T_t$$

$$S_{t+p+1} = \gamma \left(\frac{D_{t+1}}{L_{t+1}} \right) + (1 - \gamma)S_{t+1}$$

$$F_{t+1} = (L_t + T_t)S_{t+1}$$

Χρονική Διάρκεια Περιοδικότητας	4	
Συντελεστής α	0.577	Level
Συντελεστής β	0.005	Trend
Συντελεστής γ	0.05	Seasonality



Παράρτημα Γ

Holt's Method – Υπολογισμοί Παραμέτρων						
Period	Year	Time	D_t	\bar{L}_t	\bar{T}_t	\bar{F}_t
0				11393.04	0	
1	2007	1	11393.04	11393.04	0	11393.04
2		2	9112.04	10051.6	-67.1	11393.0
3		3	32675.03	23328.8	600.1	9984.5
4		4	20263.03	21773.0	492.3	23928.9
5		5	19181.04	20451.5	401.7	22265.4
6		6	18449.03	19439.3	331.0	20853.1
7		7	20472.05	20183.0	351.6	19770.2
8		8	17557.04	18783.5	264.0	20534.6
9		9	16076.03	17300.0	176.7	19047.5
10		10	23196.01	20840.2	344.8	17476.7
11		11	13820.99	16854.2	128.3	21185.0
12		12	19180.03	18274.9	192.9	16982.5
13	2008	1	12358.04	14874.7	13.3	18467.8
14		2	8729.05	11265.9	-167.8	14887.9
15		3	15099.06	13451.0	-50.2	11098.1
16		4	17179.01	15622.8	60.9	13400.9
17		5	17348.09	16662.5	109.8	15683.7
18		6	21784.09	19719.8	257.2	16772.4
19		7	20019.98	20002.3	258.5	19977.0
20		8	23004.06	21874.1	339.1	20260.8
21		9	18937.06	20286.5	242.8	22213.2
22		10	22711.07	21812.4	307.0	20529.3
23		11	14022.01	17357.3	68.9	22119.4
24		12	15553.03	16324.6	13.8	17426.2
25	2009	1	13946.03	14931.4	-56.6	16338.4
26		2	11601.05	12949.5	-152.8	14874.9
27		3	13144.04	13001.0	-142.6	12796.7
28		4	20278.97	17222.4	75.6	12858.4
29		5	18813.04	18189.0	120.1	17298.0
30		6	20117.05	19372.4	173.3	18309.1
31		7	26450.10	23606.2	376.3	19545.7
32		8	19009.06	21057.6	230.1	23982.5
33		9	23283.11	22461.2	288.8	21287.7
34		10	16390.04	19009.7	101.7	22750.0
35		11	13867.06	16027.2	-52.5	19111.4
36		12	13918.04	14765.2	-112.9	15974.7
...				
73	2013	1	10165.0	9041.6	-157.6	7437.6
74		2	7753.0	8218.9	-190.8	8884.0
75		3	5686.0	6650.7	-259.7	8028.1
76		4	19176.0	13909.9	116.3	6391.0
77		5	15790.0	15063.5	168.1	14026.2
78		6	12785.0	13792.8	96.2	15231.6
79		7	20204.0	17602.8	281.9	13888.9
80		8	20714.0	19548.6	365.1	17884.7
81		9	10397.0	14316.9	85.2	19913.7
82		10	16509.5	15641.5	147.2	14402.2
83		11	9132.0	11873.9	-48.5	15788.7
84		12	16019.1	14291.7	74.8	11825.4



Holt's Method – Υπολογισμοί Σφαλμάτων					
Period	Error	 Error 	Error^2	 Error /Demand	Error/Demand
0					
1	0	0	0	0	0
2	-2281.0	2281.0	5202942.8	0.3	-0.3
3	22690.5	22690.5	514859418.7	0.7	0.7
4	-3665.9	3665.9	13438891.8	0.2	-0.2
5	-3084.3	3084.3	9513070.5	0.2	-0.2
6	-2404.1	2404.1	5779686.7	0.1	-0.1
7	701.8	701.8	492541.4	0.0	0.0
8	-2977.5	2977.5	8865639.8	0.2	-0.2
9	-2971.5	2971.5	8829809.6	0.2	-0.2
10	5719.4	5719.4	32711058.0	0.2	0.2
11	-7364.1	7364.1	54229395.2	0.5	-0.5
12	2197.5	2197.5	4828933.7	0.1	0.1
13	-6109.8	6109.8	37329203.3	0.5	-0.5
...
84	4193.7	4193.7	17587300.4	0.3	0.3
Total	2542.94783	330728.3	2363622551	29.46937858	-9.908024803

Holts Errors			
MAD	3937.241128	MAPE	0.350826
MSE	28138363.7	MPE	-0.11795
RMSE	5304.560651		

$$L_{t+1} = \alpha D_{t+1} + (1 - \alpha)(L_t + T_t)$$

$$T_{t+1} = \beta(L_{t+1} - L_t) + (1 - \beta)T_t$$

$$F_{t+n} = L_t + nT_t$$

Συντελεστής α	0.5881	Level
Συντελεστής β	0.05	Trend



Παράρτημα Δ

Winter's Method – Υπολογισμοί παραμέτρων (3 months)

Period	Year	Time	D_t	\bar{D}_t	\bar{D}'_t	\bar{S}_t	S_t	L_t	T_t	F_t
0								55450	-1015.3	
1	2007	1	53180.1		54434.7	0.98	0.7	68169.8	-946.6	36395.3
2		2	57893.1		53419.4	1.08	1.2	57829.7	-993.6	77750.9
3		3	54105.1	53219.6	52404.1	1.03	1.4	47235.7	-1041.6	78269.4
4		4	56197.0	50897.6	51388.8	1.09	1.0	51852.8	-1013.3	45916.5
5	2008	1	36186.2	51681.9	50373.5	0.72	0.7	52390.5	-1005.5	34274.9
6		2	56311.2	52175.0	49358.2	1.14	1.1	50088.8	-1012.0	59032.7
7		3	61961.1	51999.3	48342.9	1.28	1.4	47051.6	-1022.2	67015.6
8		4	52286.1	52674.6	47327.6	1.10	1.0	49496.1	-1004.8	45959.6
9	2009	1	38691.1	53884.5	46312.3	0.84	0.7	53321.5	-980.7	32731.9
10		2	59209.1	53718.3	45297.0	1.31	1.1	51932.1	-982.7	60066.5
11		3	68742.3	50369.5	44281.7	1.55	1.4	50665.8	-984.1	69448.6
12		4	44175.1	44963.3	43266.4	1.02	1.0	46635.8	-999.4	49750.0
13	2010	1	20012.2	38839.8	42251.1	0.47	0.7	36827.9	-1043.4	30920.3
14		2	34638.1	32963.8	41235.8	0.84	1.1	32725.2	-1058.7	41053.0
15		3	44325.2	29455.0	40220.5	1.10	1.4	32136.6	-1056.3	43154.4
16		4	21584.1	27774.6	39205.2	0.55	1.0	25900.5	-1082.2	31038.8
17	2011	1	14533.1	25639.1	38189.9	0.38	0.7	23092.7	-1090.9	16648.8
18		2	26674.1	25036.7	37174.6	0.72	1.1	22734.6	-1087.2	25143.5
19		3	35205.1	25934.5	36159.3	0.97	1.4	23929.1	-1075.8	29518.4
20		4	25885.1	28603.5	35144.0	0.74	1.0	24649.3	-1066.8	22633.9
21	2012	1	17414.1	34003.7	34128.7	0.51	0.7	24926.8	-1060.1	15770.9
22		2	45145.1	37394.2	33113.4	1.36	1.1	32393.5	-1017.5	27311.0
23		3	59936.1	38467.1	32098.1	1.87	1.4	38163.3	-983.5	42953.3
24		4	28278.1	39566.6	31082.8	0.91	1.0	32412.4	-1007.4	36934.0
25	2013	1	23604.1	38814.7	30067.5	0.79	0.7	33490.6	-996.9	21049.2
26		2	47751.1	39409.9	29052.2	1.64	1.2	37300.5	-972.9	37588.1
27		3	51315.1		28036.9	1.83	1.4	36810.5	-970.5	50097.8
28		4	41660.5		27021.6	1.54	1.0	39316.7	-953.1	35386.3

Winter's Method – Υπολογισμοί Σφαλμάτων (3 months)

Period	Error	Error	Error ²	Error /Demand	Error/Demand
0					
1	16784.8	16784.8	281730875.8	0.32	0.32
2	-19857.8	19857.8	394331027.7	0.34	-0.34
3	-24164.3	24164.3	583913432.2	0.45	-0.45
4	10280.5	10280.5	105689148.7	0.18	0.18
5	1911.2	1911.2	3652769.5	0.05	0.05
6	-2721.5	2721.5	7406740.2	0.05	-0.05
...
28	6274.2	6274.2	39365603.2	0.15	0.15
Total	13084.86685	206057.1242	2645506939	5.155941694	-0.22501279

MAD	7359.183008	MAPE	0.184140775
MSE	94482390.69	MPE	-0.008036171



Παράρτημα Ε

Demand – Forecast (1 month)				
Year	Month	Demand	Forecast	
			ARIMA	WINTERS
2014	1	17014.07	13457.44	8801.133
	2	14723.02	11045.46	6609.785
	3	7932	8978.465	10803.96
	4	19945.04	22468.48	14918.29
	5	12199.03	19082.44	14814.79
	6	20484.04	16077.46	14640.99
	7	19849.03	23496.44	17472.41
	8	25089.02	24006.46	18808.07
	9	20925.02	13689.46	15437.95
	10	16761.01	19801.93	15530.47
	11	27368.03	12424.46	9080.371
	12	21876.01	19311.52	11346.88

Demand – Forecast (2 months)				
Year	Time Period	Demand	Forecast	
			ARIMA	WINTERS
2014	1	31737.084	22291.42	15514.65
	2	27877.044	29235.46	25880.95
	3	32683.068	32948.42	29483.25
	4	44938.052	45291.41	36310.93
	5	37686.032	30829.38	31010.65
	6	49244.04	29524.49	20411.45

Demand – Forecast (3 months)				
Year	Time Period	Demand	Forecast	
			ARIMA	WINTERS
2014	1	39669.084	33514.89405	25713.13
	2	52628.112	57661.90605	43275.78
	3	65863.072	61225.88205	50276.92
	4	66005.052	51571.32205	35054.88



ΠΑΡΑΡΤΗΜΑ ΣΤ

Γραμμική Παλινδρόμηση (Linear Regression)

Η ανάλυση απλής συσχέτισης (παλινδρόμηση) στηρίζεται στην υπόθεση ότι η σχέση μεταξύ μιας εξηγημένης μεταβλητής Y (π.χ. ζήτηση ενός προϊόντος) και μιας ανεξάρτητης X (π.χ. τιμή πώλησης) είναι γραμμική, είναι δηλαδή της μορφής:

$$Y = \alpha + \beta X$$

Η εύρεση της καταλληλότερης ευθείας, δηλαδή ο προσδιορισμός των τιμών των α και β , πραγματοποιείται εφαρμόζοντας τη μέθοδο των ελαχίστων τετραγώνων. Οι τιμές των α και β , προσδιορίζονται από τις σχέσεις :

$$\beta = \frac{n\Sigma XY - \Sigma X \Sigma Y}{n\Sigma X^2 - (\Sigma X)^2} = \frac{\Sigma(X - \bar{X})(Y - \bar{Y})}{\Sigma(X - \bar{X})^2}$$

$$\alpha = \frac{\Sigma Y}{n} - \frac{\beta \Sigma X}{n} = \bar{Y} - \beta \bar{X}$$

Η υπόθεση της γραμμικότητας μεταξύ Y και X και κατά συνέπεια η αξιοπιστία της πρόβλεψης, μετριέται με το συντελεστή προσδιορισμού R^2 (coefficient of determination) που ορίζεται από τη σχέση :

$$R^2 = 1 - \frac{\Sigma(Y - \hat{Y})^2}{(\Sigma(Y - \bar{Y}))^2}$$

Ο συντελεστής προσδιορισμού παίρνει τιμές μεταξύ 0 και 1 και αποτελεί μέτρο του πόσο καλά η προσδιορισθείσα σχέση $Y = \alpha + \beta X$ ταιριάζει με τα δεδομένα $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$. Όταν $R^2 = 1$ όλα τα δεδομένα ταιριάζουν ακριβώς με την ευθεία συσχέτισης κατά συνέπεια η προσδιορισθείσα σχέση $Y = \alpha + \beta X$ είναι καλό πρότυπο για μελλοντικές προβλέψεις. Αντίθετα, αν ο συντελεστής προσδιορισμού έχει τιμή κοντά στο 0, τότε η ευθεία συσχέτισης $Y = \alpha + \beta X$ είναι ακατάλληλη για να εξηγήσει τη σχέση που συνδέει τις δύο μεταβλητές Y και X .

Ομοίως, το στατιστικό μοντέλο πολλαπλής γραμμικής παλινδρόμησης (multiple linear regression), στο οποίο Y είναι η εξαρτημένη μεταβλητή και X_1, X_2, \dots, X_n οι ανεξάρτητες μεταβλητές, ορίζεται ως :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$



ΒΙΒΛΙΟΓΡΑΦΙΑ

Hanke, John E., and Arthur G. Reitsch. *Business Forecasting*. 9th ed. Boston: Allyn and Bacon, 1981.

Box, George E. P., and Gwilym M. Jenkins. *Time Series Analysis: Forecasting and Control*. Rev. ed. San Francisco: Holden-Day, 1976.

Makridakis, Spyros G. *Arma Models and the Box Jenkins Methodology*. Fontainebleau, France: INSEAD, 1995.

Chopra, Sunil, and Peter Meindl. *Supply Management: Strategy, Planning, and Operation*. 3rd ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2007.

Τατσιόπουλος Η., Μαρμαράς Ν. και Χατζηγιαννάκης Δ., *Βιομηχανική Διοίκηση και Κοστολόγηση*, Εκδόσεις Παπασωτηρίου, 2010.

Silver, Edward A., and D. F. Pyke. *Inventory Management and Production Planning and Scheduling*. 3rd ed. New York: Wiley, 1998.

Δημητριάδης, Σ.Γ και Μιχιώτης, Α.Ν., *Διοίκηση Παραγωγικών Συστημάτων : Βασικές Θεωρητικές Αρχές και Εφαρμογές στη Λήψη Επιχειρηματικών Αποφάσεων*, Εκδόσεις Κριτική, Αθήνα, 2007.

Ασημακόπουλος Β. και Πετρόπουλος Φ., *Επιχειρησιακές Προβλέψεις*, Εκδόσεις Συμμετρία, Αθήνα 2011.

Slack, Nigel, and Stuart Chambers. *Operations Management*. 6th ed. Harlow, England: Financial Times Prentice Hall, 2010.

Ιωάννου Γ., *Διοίκηση Παραγωγής & Υπηρεσιών*, Εκδόσεις Σταμούλη, 2005.

Jarrett, Jeffrey. *Business Forecasting Methods*. Oxford, UK: Blackwell, 1987.

"Time Series Forecasting by Using Seasonal Autoregressive Integrated Moving Average: Subset, Multiplicative or Additive Model." *Journal of Mathematics and Statistics* (2011): 20-27.

Lawrence, Kenneth D. *Advances in Business and Management Forecasting*. Bingley, U.K.: Emerald, 2013.

Βαϊδάνης Μ. *Αρχές Διοίκησης και Οργάνωσης Παραγωγής : Σημειώσεις Πρόβλεψης*, Αθήνα 2005



Estampe, Dominique. *Supply Chain Performance and Evaluation Models*. Hoboken: Wiley, 2014.

Jenkins, G. M. "Autoregressive-Integrated Moving Average (ARIMA) Models." *Encyclopedia of Statistical Sciences* (2004).

Brockwell, Peter J., and Richard A. Davis. *Time Series: Theory and Methods*. 2nd ed. New York: Springer-Verlag, 1991

Διαδικτυακές Πηγές

"Business Forecasting - Systems, Data, Model, and Methods." Web. June 2015
<<http://ecommerce.hostip.info/pages/457/Forecasting-Business.html>>.

"Fit an ARIMA Model - Minitab." Web. June 2015. <<http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/time-series/time-series-models/fit-an-arima-model/>>.

"Time Series Analysis for Business Forecasting - Ubalt.edu." Web. June 2015.
<<http://home.ubalt.edu/ntsbarsh/stat-data/Forecast.htm>>.

"Demand Forecasting in Managerial Economics." Web. June 2015.
<<http://www.mbaknol.com/managerial-economics/demand-forecasting-in-managerial-economics/>>.

"Forecasting & Strategy Unit - Home." Web. June 2015.
<<http://www.fsu.gr/en>>

"IBF - Demand Planning, Forecasting, Supply Chain Management" Web. June 2015.
<<https://ibf.org/>>.