ΣΥΓΧΡΟΝΟΙ ΜΕΘΟΔΟΙ ΑΝΑΛΥΣΕΩΣ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ

Τοῦ **κ. Ὄθωνος Παπαδήμα (M. Sc)**
Ἀνωτάτη Βιομηχανικὴ Σχολὴ Πειραιῶς.

ΕΙΣΑΓΩΓΗ

Ἡ περιγραφὴ τῆς διαχρονικῆς ἐξελίξεως καὶ ἡ πρόβλεψις τῆς μελλοντικῆς συμπεριφορᾶς μιᾶς στοχαστικῆς μεταβλητῆς ἀποτελεῖ ἕνα ἀπὸ τὰ σπουδαιότερα ἀντικείμενα μελέτης τῆς οἰκονομετρίας, τῆς ἐπιχειρησιακῆς ἔρευνας τοῦ Marketing, τῆς Δημογραφίας τῆς Μετεωρολογίας καὶ πολλῶν ἄλλων ἐπιστημῶν.
Ὡς πλέον ἐνδεδειγμένες μέθοδοι γιὰ τὴν μελέτη τῶν διαχρονικῶν φαινομένων θεωροῦνται σήμερα οἱ στατιστικὲς μέθοδοι. Ὁ κλάδος τῆς Στατιστικῆς ὁ ὁποῖος μελετᾶ τὰ φαινόμενα αὐτὰ ἀποδίδεται μὲ τὸν ὅρο:
Ἀνάλυση χρονολογικῶν σειρῶν (Time series Analysis). Χρονολογικὴ σειρὰ καλεῖται μία συλλογὴ παρατηρήσεων ποὺ γίνονται σὲ διαδοχικὲς χρονικὲς στιγμές.
Πολλὲς χρονολογικὲς σειρὲς συναντοῦμε π.χ. εἰς τὴν οἰκονομία ὅπως οἱ τιμὲς τῶν μετοχῶν σὲ διαδοχικὲς ἡμέρες, ἡ ἀξία τῶν ἐξαγωγῶν μιᾶς χώρας σὲ διαδοχικοὺς μῆνες τὰ κέρδη μιᾶς ἐπιχειρήσεως σὲ διαδοχικὰ χρόνια κ.λ.π. Κατὰ τὴν τελευταία δεκαπενταετία ἀναπτύχθησαν πολλὲς στατιστικὲς μέθοδοι ἀναλύσεως τῶν χρονολογικῶν σειρῶν καὶ ὁ σύγχρονος ἐρευνητὴς ἔχει σήμερα στὴν διαθεσή του ὅλα ἐκεῖνα τὰ ἐφόδια ὥστε νὰ μελετήση μὲ ἐπιτυχία μιὰ χρονολογικὴ σειρὰ ποὺ θὰ τοῦ παρουσιασθεῖ στὴν πράξη. Ἐκτὸς ὅμως ἀπὸ αὐτό, ἐπειδὴ ἡ θεωρία ἀναλύσεως τῶν χρονολογικῶν σειρῶν ἔχει ἀναπτυχθεῖ πρόσφατα ἀποτελεῖ καὶ σήμερα πόλο ἕλξεως πολλῶν ἐρευνητῶν.

2. Η «ΦΙΛΟΣΟΦΙΑ» ΤΩΝ ΜΕΘΟΔΩΝ ΑΝΑΛΥΣΕΩΣ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ

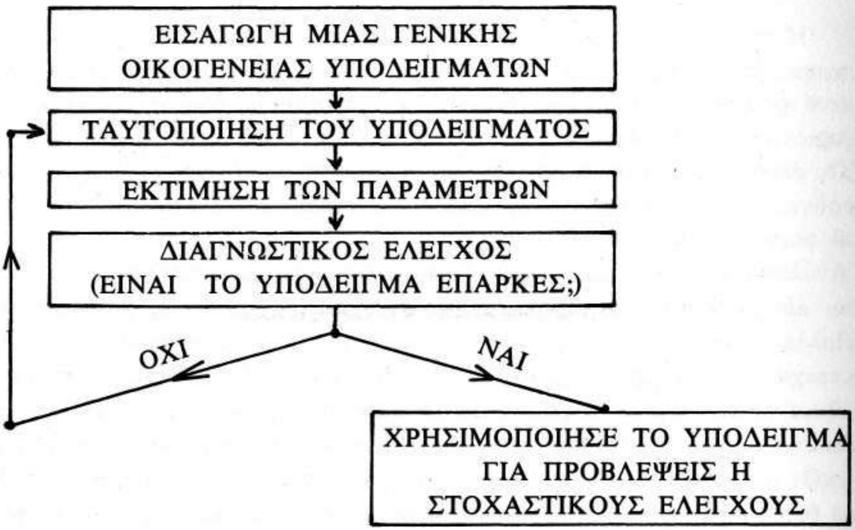
Ὅλες ἀναξαιρέτως οἱ μέθοδοι ἀναλύσεως χρονολογικῶν σειρῶν ἀναζητοῦν ὑποδείγματα ποὺ θὰ περιγράφουν κατὰ τὸν καλύτερο δυνατὸ τρόπο τὰ ὑπάρχοντα ἐμπειρικὰ δεδομένα.

Δηλαδὴ ὑποδείγματα ποὺ θὰ διαθέτουν τὴν μεγαλύτερη δυνατὴ ἁπλότητα καὶ συγχρόνως τὸν ἐλάχιστο ἀριθμὸ παραμέτρων, χωρὶς ὅμως νὰ βλάπτεται ἡ εὐελιξία τους. Ἡ κατασκευὴ τῶν ὑποδειγμάτων αὐτῶν εἶναι ἐνδιαφέρουσα διότι:

α. Αὐτὰ δύνανται νὰ μᾶς ἀποκαλύψουν τὴν ὑπάρχουσα νομοτέλεια ἡ ὁποία πιθανὸν ὑφίσταται μεταξὺ τῶν διαδοχικῶν παρατηρήσεων μιᾶς μεταβλητῆς.
β. Αὐτὰ δύνανται νὰ χρησιμοποιηθοῦν γιὰ νὰ μᾶς δώσουν «ἄριστες» προβλέψεις γιὰ τὴν μελλοντικὴ ἐξέλιξη τῆς ὑπὸ ἔρευνα μεταβλητῆς.
γ. Ἡ πρόβλεψη τῆς μελλοντικῆς συμπεριφορᾶς τῆς διαδικασίας παρέχει στὴν διοίκηση τὴν εὐχέρεια νὰ ἐπέμβη ἔγκαιρα, ἂν χρειαστεῖ καὶ νὰ ἐλέγξει αὐτὴν πρὶν ἐξελιχθεῖ δυσμενῶς (optimal control policy).

3. Η ΜΕΘΟΔΟΛΟΓΙΑ ΑΝΑΛΥΣΕΩΣ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ

Ἡ μεθοδολογία ἀναλύσεως τῶν χρονολογικῶν σειρῶν γνωστὴ ὡς μεθοδολογία τῶν Box-Jenkins περιγράφεται ἀπό τό παρακάτω διάγραμμα.



4. ΟΙΚΟΓΕΝΕΙΕΣ ΥΠΟΔΕΙΓΜΑΤΩΝ

Ἡ εὐρύτερη οἰκογένεια στοχαστικῶν ὑποδειγμάτων γνωστὴ στὴν διεθνὴ βιβλιογραφία ὡς ARIMA MODELS ἔχει μελετηθεῖ ἀπὸ τοὺς σύγχρονους ἐρευνητὲς BOX-JENKINS. Ἡ δύναμη τῶν ὑποδειγμάτων αὐτῶν ἔγκειται στὸ γεγονὸς ὅτι αὐτὰ μποροῦν νὰ περιγράψουν μιὰ ὁποιαδήποτε χρονολογικὴ σειρά, νὰ δώσουν ἱκανοποιητικὲς προβλέψεις καὶ συνεπῶς νὰ χρησιμεύσουν στὴ διαδικασία λήψεως ἀποφάσεων καὶ στὸν στοχαστικὸ ἔλεγχο τῆς διαδικασίας. Τὸ κυριότερο ἴσως μειονέκτημα τῶν ὑποδειγμάτων BOX-JENKINS εἶναι ὅτι αὐτὰ εἶναι μάλλον σύνθετα καὶ ἀπαιτοῦν μεγάλη ἐμπειρία ἀπὸ τὸ μέρος τοῦ ἐρευνητοῦ καθὼς ἐπίσης καὶ πολλοὺς ἀριθμητικοὺς ὑπολογισμούς. Ἡ μελέτη τῆς οἰκογένειας ARIMA θὰ ἀποτελέσει τὸ κύριο ἀντικείμενο αὐτῆς τῆς δημοσιεύσεως.

5. ΤΑΥΤΟΠΟΙΗΣΗ ΤΟΥ ΥΠΟΔΕΙΓΜΑΤΟΣ

Ἂς συμβολίσουμε μὲ X_t τὶς N τὸ πλῆθος παρατηρήσεις σὲ ἰσαπέχουσες χρονικὲς στιγμὲς t=1,2,...,N. Ὁ Wold (1938) ἀπέδειξε ὅτι κάθε στοχαστική, ἀσυνεχής, χρονολογικὴ σειρά* X_t μπορεῖ νὰ ἐξομαλυνθεῖ ἀπὸ τὴν παρακάτω οἰκογένεια ὑποδειγμάτων καλουμένων ὑποδειγμάτων ARMA.

$$X_t= \varphi_1 X_{t-1}+\varphi_2 X_{t-2}+...+\varphi_p X_{t-p}+u_t - \partial_1 u_{t-1}-\partial_2 u_{t-2}.... - \partial_q u_{t-q} \quad (1)$$

ὅπου: $\Phi' = (\varphi_1, \varphi_2,...,\varphi_p)$, $\Theta' = (\partial_1,\partial_2,...,\partial_q)$

διανύσματα παραμέτρων καί:
$V' = (u_t, u_{t-1},..., u_{t-q})$

ἕνα διάνυσμα ὁμοσκεδαστικῶν τυχαίων μεταβλητῶν μὲ μέσο μηδὲν καὶ ἀσυσχετίστων ἀνὰ δύο.

Δηλαδή:

$$E(u_t) = 0$$

$$Cov(u_t, u_s) = \begin{cases} \sigma_u^2<\infty, & s=t \\ 0, & s\neq t \end{cases}$$

Τὶς ἰδέες τοῦ Wold ἐκμεταλεύτηκαν πλήρως μετὰ ἀπὸ δύο καὶ πλέον δεκαετίες οἱ ἐρευνητὲς Box-Jenkins καὶ ὑπέδειξαν πρῶτοι μεθόδους:

α. Ταυτοποιήσεως μιᾶς συγκεκριμένης χρονολογικῆς σειρᾶς.
β. Μεθόδους ἐκτιμήσεως παραμέτρων.
γ. Στατιστικοὺς ἐλέγχους γιὰ τὴν ἐπάρκεια τοῦ ἐκτιμηθέντος ὑποδείγματος.

Μὲ τὸν ὅρο ταυτοποίηση (indetification) ἐννοοῦμε τὴν εὕρεση τῶν φυσικῶν ἀριθμῶν* p, q στὴν (1). Οἱ p, q ἀποφασίζονται ἀπὸ δύο στατιστικὲς (δειγματικὲς) συναρτήσεις, οἱ ὁποῖες ὁρίζονται στὸ δεύτερο μέρος τῆς δημοσιεύσεως μας, δηλαδὴ ἀπὸ τὴν συνάρτηση αὐτοσυσχετίσεως (autocorrelation fuction) καὶ ἀπὸ τὴν συνάρτηση μερικῆς αὐτοσυσχετίσεως (partial autocorrelation function) ἔτσι ὥστε οἱ p, q νὰ εἶναι ὅσον τὸ δυνατὸν «μικρότεροι» χωρὶς ὅμως νὰ βλάπτεται ἡ ἐπάρκεια τοῦ ὑποδείγματος γιὰ τὴν ἐξομάλυνση μιᾶς συγκεκριμένης χρονολογικῆς σειρᾶς. Ἐὰν q=0 δηλαδὴ ἐὰν τὸ διάνυσμα Θ τῶν παραμέτρων εἶναι τὸ μηδενικὸ διάνυσμα ἡ οἰκογένεια (1) λαμβαίνει τὴν μορφή.

$$X_t= \varphi_1 X_{t-1}+\varphi_2 X_{t-2}+...+\varphi_p X_{t-p}+u_t \quad (2)$$

---

* Σέ ὅτι ἀκολουθεῖ θὰ ὑποθέτουμε ὅτι ἡ σειρὰ X_t εἶναι ἀπαλλαγμένη ἐποχικῶν κυμάνσεων.
* Οἱ p, q μποροῦν νὰ εἶναι καὶ ἴσοι πρὸς τὸ μηδέν.

καὶ καλεῖται διαδικάσία αὐτοπαλινδρομήσεως τάξεως p (autovegressive process) ἢ ὑπόδειγμα $AR_{(p)}$

Ἐὰν P=0 τότε ἡ (1) λαβαίνει τὴν μορφὴ:

$$X_t = u_t - \partial_1 u_{t-1} - ... - \partial_q u_{t-q} \tag{3}$$

καὶ καλεῖτται διαδικασία κινητῶν μέσων τάξεως· q (Mooving average process) ἢ ὑπόδειγμα $MA_{(q)}$. Τέλος τὸ πλῆρες ὑπόδειγμα (1) θὰ καλεῖται ὑπόδειγμα ARMA (p,q) ὡς σύνθεσις τῶν ὑποδειγμάτων (2), (3). Γιὰ περισσότερες λεπτομέρειες ἐπὶ τῆς ἀκολουθητέας διαδικασίας ταυτοποιήσεως τοῦ ὑποδείγματος, ἡ ὁποία ἀποτελεῖ τὸ σπουδαιότερο καὶ δυσκολότερο μέρος τῆς ἀναλύσεως μιᾶς χρονολογικῆς σειρᾶς θὰ ἐπανέλθουμε στὸ δεύτερο μέρος τῆς δημοσιεύσεως.

6. ΕΚΤΙΜΗΣΙΣ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ

Ἀφοῦ ἀποφασισθοῦν οἱ ἀκέραιοι p,q τὸ ἀμέσως ἑπόμενο βῆμα εἶναι νὰ ἐκτιμηθοῦν οἱ παράμετροι $\varphi_1, \varphi_2, ..., \varphi_p, \partial_1\ \partial_2, ..., \partial_q$ ἐπὶ τῇ βάσει βεβαίως τῶν πειραματικῶν δεδομένων.

Ἡ διαδικασία ἐκτιμήσεως τῶν παραμέτρων συνίσταται στὴν εὕρεση τῶν τιμῶν ἐκείνων τῶν παραμέτρων οἱ ὁποῖες ἐλαχιστοποιοῦν τὸ ἄθροισμα $\sum_{t=1}^{N} u_t^2$ γνωστὴ ὡς μέθοδος ἐλαχίστων τετραγώνων. Δυστυχῶς ὅμως οἱ «κανονικὲς ἐξισώσεις» εἶναι κατὰ κανόνα μὴ γραμμικὲς ὡς πρὸς τὶς παραμέτρους καὶ συνήθως εἶναι ἀδύνατο νὰ λυθοῦν καὶ νὰ δώσουν ἐκτιμήσεις τῶν παραμέτρων. Γι' αὐτὸ ὁ ἐρευνητὴς καλεῖται νὰ «μαντέψει» ἀρχικὲς ἐκτιμήσεις τῶν παραμέτρων μέσα στὰ ἐπιτρεπόμενα ὅρια τοῦ παραμετρικοῦ χώρου καὶ νὰ ὑπολογίσει τὸ ἄθροισμα $\sum_{t=1}^{N} u_t^2$.
Στὴν συνέχεια νὰ ἐπαναλάβει ἴδια διαδικασία γιὰ ἄλλες τιμὲς τῶν παραμέτρων μέχρις ὅτου ἐπιτύχει τὴν ἐλαχιστοποίηση τοῦ ἐν λόγῳ ἀθροίσματος τετραγώνων. Ἡ ἀνωτέρω ἀριθμητικὴ τεχνικὴ τοῦ ὑπολογισμοῦ τῶν παραμέτρων εἶναι γνωστὴ ὡς μέθοδος μὴ γραμμικῆς ἐκτιμήσεως (non linear estimation) καὶ βεβαίως εἶναι δυνατὸν νὰ ἐφαρμοσθεῖ μόνον ὅταν ὁ ἐρευνητὴς ἔχει στὴ διάθεσή του ἕνα ταχὺ ἠλεκτρονικὸ ὑπολογιστή.

Τέλος ὑποθέτοντες ὅτι τά σφάλματα u_t κατανέμονται κανονικά, εἶναι δυνατόν νά εὑρεθοῦν ἐκτιμήτριες μεγίστης πιθανοφανείας τῶν παραμέτρων, διαστήματα ἐμπιστοσύνης ἢ νά ἐλεγχθοῦν στατιστικές ὑποθέσεις ἀναφερόμενες στίς ἐν λόγῳ παραμέτρους.

7. ΔΙΑΓΝΩΣΤΙΚΟΣ ΕΛΕΓΧΟΣ.

Μετὰ τὴν ταυτοποίηση καὶ ἐκτίμηση τῶν παραμέτρων τοῦ ὑποδείγματος ὁ ἐρευνητὴς θὰ πρέπει νὰ προχωρήσει στὸν ἔλεγχο τῆς ἐπάρκειας τοῦ ὑποδείγματος (diagnosting checking). Δηλαδὴ τὰ ἐξακριβώσει κατὰ πόσο τὸ ὑπόδειγμα περιγράφει ἱκανοποιητικὰ τὰ ἐμπειρικὰ δεδομένα. Γι' αὐτὸ ὁ ἐρευνητὴς θὰ πρέπει νὰ ἐξετάσει τὶς προκύπτουσες διαφορὲς τῶν παρατηρηθεισῶν τιμῶν ἀπὸ τὶς τιμὲς ποὺ προβλέπει τὸ ἐκτιμηθὲν ὑπόδειγμα ([3] σελὶς 73).

Ἄλλος ἔλεγχος τῆς ἐπάρκειας τοῦ ὑποδείγματος στηρίζεται στοὺς συντελεστὲς αὐτοσυσχετίσεως τοῦ δείγματος γνωστὸς ὡς Portmanteau lack –of – fit test ([1] σελὶς 290). Ἐὰν ὁ διαγνωστικὸς ἔλεγχος δώσει ἀρνητικὰ ἀποτελέσματα ὁ ἐρευνητής, θὰ πρέπει νὰ ἀνακαλύψει τὴν «διεύθυνση» κατὰ τὴν ὁποία τὸ ὑπόδειγμα εἶναι ἐπαρκὲς καὶ νὰ ἐπαναλάβει τὴν διαδικασία ταυτοποιήσεως.

8. ΠΡΟΒΛΕΨΕΙΣ

Ο σπουδαιότερος ἴσως λόγος ποὺ μᾶς ὠθεῖ στὴν ἀνάλυση μιᾶς χρονολογικῆς σειρᾶς, εἶναι ἡ πρόβλεψη τῆς μελλοντικῆς ἐξελίξεως αὐτῆς. Ἐνῶ ὅμως σέ κάθε μὴ στοχαστικὴ διαδικασία (deterministic) ἡ μελλοντικὴ ἐξέλιξη αὐτῆς ἐξαρτᾶται ἀποκλειστικὰ καὶ μόνο ἀπὸ τὴν συμπεριφορά της στὸ παρελθὸν καὶ συνεπῶς εἶναι δυνατὸν νὰ προβλεφθεῖ χωρὶς σφάλματα, αὐτὸ δὲν εἶναι δυνατὸν νὰ ἐπιτευχθεῖ σὲ μία στοχαστικὴ διαδικασία, διότι ἡ ἐξέλιξη της ἐξαρτᾶται κατὰ ἕνα μέρος καὶ ἀπὸ ἄλλους «τυχαίους» καὶ συνεπῶς ἀπρόβλεπτους παράγοντες.

Ἐὰν δηλαδὴ κατὰ τὴν χρονικὴ στιγμὴ t ἔχουμε στὴ διάθεση μας τὰ δεδομένα $X_1, X_2 .... X_t$ καὶ ἐπιθυμοῦμε νὰ προβλέψουμε τὴν ἑπόμενη τιμὴ X_{t+1} τῆς σειρᾶς θά πρέπει νὰ λάβουμε ὑπ' ὄψη ὅτι ἡ τιμὴ αὐτὴ εἶναι μία τυχαία μεταβλητὴ καὶ συνεπῶς ἀκολουθεῖ κάποιον νόμο πιθανότητος. Μία «λογική» ἐκτίμησις $\hat{X}_{t+1}$ τῆς X_{t+1} καλούμενη πρόβλεψις ἐλαχίστου μέσου τετραγωνικοῦ σφάλματος (mean squave error forecast, [1] σελὶς 128) ὁρίζεται κατωτέρω:

$$\hat{X}_{t+1} = E(X_{t+1} / X_t, X_{t-1} .....) \tag{4}$$

δηλαδὴ ὁρίζεται ὡς ἡ ἀναμενομένη τιμὴ τῆς μεταβλητῆς X_{t+1} δοθείσης τῆς ἱστορίας τῆς διαδικασίας στὸ παρελθόν.

Ὁ νόμος πιθανότητος τῆς μεταβλητῆς X_{t+1} δὲν εἶναι ἀπαραίτητο νὰ εἶναι γνωστός, ἀρκεῖ μόνο νὰ ἔχουμε στὴ διάθεση μας τὸ ὑπόδειγμα τὸ ὁποῖο νὰ περιγράφει τὰ δεδομένα μέχρις τῆς χρονικῆς στιγμῆς t. Ὁ ἐρευνητὴς ὅμως θὰ πρέπει προηγουμένως νὰ ἔχει πεισθεῖ ὅτι τὸ ἐκτιμηθὲν ὑπόδειγμα εἶναι τὸ πλέον ἐνδεδειγμένο γιὰ νὰ ἐξομαλύνει τὰ δεδομένα του καὶ ἐπὶ πλέον νὰ δεχθεῖ ὅτι ἡ μελλοντικὴ ἐξέλιξη τῆς διαδικασίας θὰ εἶναι τῆς ἴδιας μορφῆς ὡς καὶ στὸ παρελθόν. Ἔτσι π.χ. ἐὰν τὰ δεδομένα ἀντιστοιχοῦν στὸν ὄγκο τῆς παραγωγῆς ἑνὸς προϊόντος μέχρις τῆς χρονικῆς στιγμῆς t ἡ πρόβλεψις $\hat{X}_{t+1}$ τῆς παραγωγῆς κατὰ τὸ χρονικὸ διάστημα (t, t+1) δὲν θὰ εἶναι ρεαλιστικὴ ἐὰν μέσα στὸ διάστημα αὐτὸ συμβεῖ ἀπεργία, θεομηνία, πόλεμος, κλπ. Τέλος ὑπὸ τὴν προϋπόθεση τῆς κανονικότητος ὁ ἐρευνητὴς μπορεῖ νὰ ὑπολογίσει διαστήματα ἐμπιστοσύνης γιὰ τὶς προβλέψεις $\hat{X}_{t+1}$, $\hat{X}_{t+2}$,.... ἢ μὲ ἄλλα λόγια νὰ κατασκευάσει μία ζώνη ἐμπιστοσύνης ἐντὸς τῆς ὁποίας θὰ ἀναμένει νὰ ἐξελιχθεῖ ἡ διαδικασία μὲ ὁποιαδήποτε ἐπιθυμητὴ πιθανότητα.

ΜΕΡΟΣ ΔΕΥΤΕΡΟ

9. ΟΡΙΣΜΟΣ ΜΙΑΣ ΧΡΟΝΟΛΟΓΙΚΗΣ ΣΕΙΡΑΣ

Μιὰ συλλογὴ τυχαίων μεταβλητῶν $\{X_t\}_{t\in T}$ ὅπου $T = \{1,2,...\}$ ἕνα σύνολο διαδοχικῶν χρονικῶν στιγμῶν καλεῖται στοχαστικὴ διαδικασία στὸν διακριτὸ χρόνο (stochastic process in discrete time). Μία χρονολογικὴ σειρὰ $X_1,X_2,...$ θὰ θεωρεῖται σὰν δεῖγμα (μεγέθους μιᾶς μονάδας) ἀπὸ ἕναν πληθυσμὸ χρονολογικῶν σειρῶν οἱ ὁποῖες δύνανται νὰ παραχθοῦν ἀπὸ τὴν στοχαστικὴ διαδικασία $\{X_t\}_{t\in T}$. Τὸ σύνολο αὐτὸ τῶν χρονολογικῶν σειρῶν θὰ καλεῖται δειγματικὸς χῶρος τῆς διαδικασίας. Ἑπομένως μία παρατηρηθεῖσα χρονολογικὴ σειρὰ καλουμένη καὶ δειγματικὴ πραγματοποίηση τῆς διαδικασίας μπορεῖ νὰ θεωρηθεῖ σὰν ἕνα στοιχεῖο q τοῦ δειγματικοῦ χώρου:.

$Q = \Omega\times \Omega\times\Omega\times...$ ὅπου Ω ὁ δειγματικὸς χῶρος τῶν τυχαίων μεταβλητῶν $\{X_t\}_{t\in T}$ μὲ στοιχεῖα ω. Δηλαδή: $q' = (X_1, X_2,....)$, ὅπου πλέον τὰ στοιχεῖα τοῦ διανύσματος q δὲν θὰ εἶναι τυχαῖες μεταβλητές ἀλλὰ οἱ συγκεκριμένες παρατηρήσεις.

Δοθείσης μιᾶς χρονικῆς στιγμῆς $t_o\in T$ ἡ παρατηρηθεῖσα τιμὴ X_{t_o} τῆς χρονολογικῆς σειρᾶς θὰ δύναται νὰ θεωρηθεῖ ὡς τιμὴ μιᾶς τυχαίας μεταβλητῆς $X_{t_o}(\omega)$, $\omega\in\Omega$. Κατ' αὐτὴ τὴν ἔννοια μποροῦμε νὰ ὁμιλοῦμε π.χ. γιὰ διαστήματα ἐμπιστοσύνης ἢ ἐλέγχους ὑποθέσεων:

10 ΑΛΛΟΙ ΟΡΙΣΜΟΙ

Ένα μεγάλο μέρος τῆς στατιστικῆς μεθοδολογίας ἀναφέρεται σὲ καταστάσεις ὅπου οἱ παρατηρήσεις ὑποτίθεται ὅτι κατανέμονται ἀνεξάρτητα μεταξύ τους. Ἐν τούτοις πολλὰ δεδομένα π.χ. στὴν οἰκονομία στὶς φυσικὲς ἐπιστῆμες στὸ Marketing κ.λ.π. λαμβάνονται ὑπὸ τὴν μορφὴν χρονολογικῶν σειρῶν οἱ δὲ παρατηρήσεις εἶναι ἐξηρτημένες μεταξύ τους καὶ ἐπὶ πλέον ἡ ἐξάρτηση αὐτὴ παρουσιάζει ἰδιαίτερο ἐνδιαφέρον. Ἔτσι π.χ. ἡ σημερινὴ τιμὴ μιᾶς μετοχῆς θὰ ἐξαρτᾶται ἐν μέρει ἀπὸ τὴν χθεσινή, προχθεσινή. κλπ. τιμή της ἢ τὸ εἰσόδημα μιᾶς οἰκογένειας γιὰ τὸ τρέχον ἔτος θὰ σχετίζεται κατὰ κανόνα πρὸς τὰ εἰσοδήματα τῶν προηγουμένων ἐτῶν.
Ἡ συσχέτιση μεταξὺ διαδοχικῶν τιμῶν μιᾶς μεταβλητῆς καλεῖται αὐτοσυσχέτιση

(autocorrelation). Δοθέντων τῶν N παρατηρήσεων X_1, X_2,..., X_N μιᾶς διακριτῆς χρονολογικῆς σειρᾶς δυνάμεθα νὰ σχηματίσουμε N-1 ζεύγη παρατηρήσεων:

$$(X_1,X_2),\ (X_2,\ X_3),\ \ldots,\ (X_{N-1},X_N)$$

Θεωροῦντες τὴν πρώτη παρατήρηση σὲ κάθε ζεῦγος (X_t, X_{t+1}), t= 1,2,..., N−1 σὰν μιὰ μεταβλητὴ καὶ τὴν δεύτερη παρατήρηση σὰν ἄλλη μεταβλητή, ἡ (δειγματικὴ) συνδιακύμανση μεταξὺ τῶν μεταβλητῶν αὐτῶν, καλουμένη αὐτοσυνδιακύμανση μὲ ὑστέρηση μία χρονικὴ μονάδα (autocovariance at lag one) δίδεται ἀπὸ τὴν σχέση:

$$c_1 = \mathrm{Cov}(X_t,\ X_{t+1}),\ t\ =\ 1,2,\ldots,\ N\text{-}1 \tag{5}$$

Ὁ (δειγματικὸς) συντελεστὴς συσχετίσεως μεταξὺ τῶν μεταβλητῶν X_t, X_{t+1} καλούμενος συντελεστὴς αὐτοσυσχετίσεως μὲ ὑστέρηση μία χρονικὴ μονάδα δίδεται ἀπὸ τὴν παρακάτω σχέση.

$$r_1 = \frac{\mathrm{Cov}(X_t.X_{t+1})}{\sqrt{\mathrm{Var}X_t}\ \sqrt{\mathrm{Var}X_{t+1}}}\quad t= 1,2,\ldots,N-1 \tag{6}$$

Ἀναλόγως ὁρίζεται ὁ συντελεστὴς αὐτοσυνδιακυμάνσεως C_h ἢ αὐτοσυσχετίσεως r_h μὲ ὑστέρηση h χρονικῶν μονάδων. Ἐὰν τὸ πλῆθος τῶν παρατηρήσεων εἶναι ἀρκετὰ μεγάλο (N>50) τότε οἱ σχέσεις (5) καὶ (6) προσεγγίζονται ὑπὸ τῶν:

$$c_1 = \frac{1}{N} \sum_{t=1}^{N-1} (X_t - \bar{X})\,(X_{t+1} - \bar{X}) \tag{7}$$

$$r_1 \doteq \frac{\sum_{t=1}^{N-1} (X_t - \bar{X})\,(X_{t+1} - \bar{X})}{\sum_{t=1}^{N} (X_t - \bar{X})^2} \tag{8}$$

ὅπου

$$\bar{X} = \frac{1}{N} \sum_{t=1}^{N} X_t$$

Γενικώτερα οἱ (δειγματικοι) συντελεστὲς αὐτοσυνδιακυμάνσεως καὶ αὐτοσυσχετίσεως θὰ ὑπολογίζονται ἐκ τῶν (9) καὶ (10).

$$c_h = \frac{1}{N} \sum_{t=1}^{N-h} (X_t - \bar{X})(X_{t+h} - \bar{X}) \tag{9}$$

$$r_h = \frac{\sum_{t=1}^{N-h} (X_t - \bar{X})(X_{t+h} - \bar{X})}{\sum_{t=1}^{N-h} (X_t - \bar{X})^2} \tag{10}$$

Καὶ θὰ καλοῦνται δειγματικὲς συναρτήσεις αὐτοσυνδιακυμάνσεως καὶ αὐτοσυσχετίσεως τῆς σειρᾶς X_t.
Σημειωτέον ὅτι:

$$c_h = c_{-h},\ r_h = r_{-h},\ r_o = 1$$

Τέλος μία ἐξ' ἴσου σημαντικὴ συνάρτηση εἶναι ἡ δειγματικὴ συνάρτηση μερικῆς αὐτοσυσχετίσεως (Partial autocorrelation function) ἡ ὁποία ὁρίζεται* ὡς ἑξῆς:

$$\hat{\varphi}_1 = r_4, \quad \hat{\varphi}_2 = \frac{\begin{vmatrix} r_1 & r_1 \\ r_1 & r_2 \end{vmatrix}}{\begin{vmatrix} 1 & r_1 \\ r_1 & 1 \end{vmatrix}} \ \ldots\ldots,$$

$$\hat{\varphi}_k = \frac{\begin{vmatrix} 1 & r_1 & r_2 & \ldots & r_1 \\ r_1 & 1 & r_1 & \ldots & r_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{k-1} & r_{k-2} & r_{k-3} & \ldots & r_k \end{vmatrix}}{\begin{vmatrix} 1 & r_1 & r_2 & \ldots & r_{\kappa-1} \\ r_1 & 1 & r_1 & \ldots & r_{k-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{\kappa-1} & r_{\kappa-2} & r_{\kappa-3} & \ldots & 1 \end{vmatrix}} \tag{11}$$

* Ἀποδεικνύεται ὅτι οἱ συντελεστές $\hat{\varphi}_k$ ταυτίζονται μέ τίς ἐκτιμήτριες ἐλαχίστων τετραγώνων τῶν παραμέτρων φ_k τοῦ ὑποδείγματος (2).

Ἂς σημειωθεῖ ὅτι ἡ ὁρίζουσα τοῦ ἀριθμητοῦ διαφέρει ἀπὸ τὴν ὁρίζουσα τοῦ παρανομαστοῦ κατὰ τὴν τελευταία μόνο στήλη ἡ ὁποία ἀποτελεῖται ἀπὸ τὰ στοιχεῖα

$r_1, r_2, \ldots, r_\kappa$.

Ἐπίσης ἂς σημειωθῆ ὅτι οἱ συναρτήσεις (9), (10), (11) θὰ νοοῦνται ὡς ἐκτιμήτριες τῶν πληθυσμιακῶν συναρτήσεων γ_h, ρ_h, φ_κ τῆς στοχαστικῆς διαδικασίας $\{X_t\}_{t\in T}$.

11. ΕΥΣΤΑΘΕΙΣ ΔΙΑΔΙΚΑΣΙΕΣ

Μία σημαντικὴ κλάση στοχαστικῶν διαδικασιῶν εἶναι οἱ λεγόμενες εὐσταθεῖς (stationary) διαδικασίες. Δειγματικὲς πραγματοποιήσεις εὐσταθῶν διαδικασιῶν παράγουν κατ' ἐπέκταση εὐσταθεῖς χρονολογικὲς σειρές. Ὑποδείγματα δυνάμενα νὰ ἐξομαλύνουν εὐσταθεῖς χρονολογικὲς σειρὲς καλοῦνται κατ' ἀναλογία εὐσταθῆ ὑποδείγματα. Οἱ προκύπτουσες ὅμως στὴν πράξη χρονολογικὲς σειρὲς εἶναι κατὰ κανόνα «ἀσταθεῖς». Ἐν τούτοις ἡ ἀνάλυση τῶν χρονολογικῶν σειρῶν στηρίζεται στὰ εὐσταθῆ ὑποδείγματα, διότι μία ὁποιαδήποτε χρονολογικὴ σειρὰ εἶναι δυνατὸν νὰ μετασχηματισθεῖ σὲ εὐσταθῆ ([1] σελὶς 185).

Μία στοχαστικὴ διαδικασία θὰ λέμε ὅτι εἶναι εὐσταθὴς δευτέρας τάξεως (second order stationary), ἐὰν ἔχει σταθερὸ μέσο καὶ ἡ συνάρτηση αὐτοσυνδιακυμάνσεως γ_h αὐτῆς ἐξαρτᾶται μόνο ἀπὸ τὴν ὑστέρηση h (καὶ ὄχι ἀπὸ τὸ t) δηλαδὴ ἐὰν $t\in T$ μία χρονικὴ στιγμὴ τότε:

$$E\left(X_t(\omega)\right) = \mu,\ \omega\in\Omega \tag{12}$$

$$\text{καὶ } Cov(X_t, X_{t+h}) = E\,(X_t - \mu)\,(X_{t+h} - \mu) = \gamma_h \tag{13}$$

Συνεπῶς μιὰ εὐσταθὴς χρονολογικὴ σειρὰ μεταξὺ τῶν ἄλλων παραμένει διαχρονικῶς σὲ κατάσταση ἰσορροπίας γύρω ἀπὸ τὸν αὐτὸ μέσο καὶ ἔχει σταθερὰ διακύμανση διότι ἐκ τῆς (13) θέτοντες h=0 λαβαίνουμε:

$$Cov(X_t, X_t) = Var^1(X_t) = \gamma_o = \text{σταθερά.}$$

Μὲ ἄλλα λόγια σὲ μία εὐσταθῆ χρονολογικὴ σειρὰ δὲν παρατηρεῖται μακροχρόνια τάση ἢ μεταβολὴ στὴν διακύμανση ἢ ἄλλες συστηματικὲς μεταβολὲς ὅπως π.χ. ἐποχικὲς κυμάνσεις καθὼς καὶ ἄλλες περιοδικὲς κινήσεις.

12. ΤΑΥΤΟΠΟΙΗΣΗ ΤΟΥ ΥΠΟΔΕΙΓΜΑΤΟΣ

Ἀς ὑποθέσουμε ὅτι ὁ ἐρευνητὴς ἔχει στὴ διάθεση του τὰ δεδομένα X_t, $t = 1,2,\ldots, N$ μιᾶς εὐσταθοῦς χρονολογικῆς σειρᾶς καὶ ἐπιθυμεῖ νὰ προσδιορίσει τοὺς φυσικοὺς p,q (βλέπε ἐξίσωση (1).

Ὅπως εἴπαμε στὴν 1.2 δύο ἀλληλοσυγκρουόμενα κριτήρια πρέπει νὰ ληφθοῦν ὑπ' ὄψιν. Δηλαδὴ τὸ κριτήριο τῆς ἁπλότητας τοῦ ὑποδείγματος καὶ τὸ κριτήριο τῆς ἐπάρκειας αὐτοῦ. Ἁπλότης στὸ ὑπόδειγμα σημαίνει ὅτι οἱ φυσικοὶ p,q θὰ πρέπει νὰ εἶναι ὅσον τὸ δυνατὸν μικρότεροι ἐνῶ ἐπάρκεια σημαίνει ὅτι οἱ p,q θὰ πρέπει νὰ εἶναι «ἀρκετά» μεγάλοι οὕτως ὥστε τὸ ὑπόδειγμα νὰ περιγράφει ἱκανοποιητικὰ τὰ δεδομένα. Πάντως στὶς περισσότερες τῶν ἐφαρμογῶν οἱ p,q δὲν ὑπερβαίνουν συνήθως τὸ τρία.

Ἡ διαδικασία ταυτοποιήσεως τοῦ ὑποδείγματος ἔχει ὡς ἑξῆς:

Κατ' ἀρχὴν ὑπολογίζονται ἀπὸ τὴν (10) οἱ ἐκτιμήσεις r_h τῶν συντελεστῶν αὐτοσυσχετίσεως ρ_h (συνήθως οἱ πρῶτοι 15 ἢ 20 συντελεστές).

Καθὼς ἐπίσης ἐκ τῆς (11) ἴσος ἀριθμὸς ἐκτιμήσεων $\hat{\varphi}_k$ τῶν συντελεστῶν μερικῆς αὐτοσυσχετίσεως φ_k. Εἶναι γνωστό ([1] σελὶς 174) ὅτι γιὰ ἀρκετὰ μεγάλα δείγματα οἱ ἐκτιμήσεις r_h καὶ $\hat{\varphi}_\kappa$ τείνουν νὰ μιμηθοῦν τὶς πληθυσμιακὲς συναρτήσεις ρ_h, φ_k ὑποκείμενες βεβαίως σὲ σφάλματα δειγματοληψίας [1], σελὶς 177). Ἔτσι λοιπὸν ὁ ἐρευνητὴς θὰ πρέπει νὰ συγκρίνει τὶς παρατηρηθείσες τιμὲς τῶν ἀντιστοίχων παραμέτρων μέ τιμές παραμέτρων συγκεκριμένων ὑποδειγμάτων. Εἶναι γνωστὸ ὅτι σὲ μία, εὐσταθῆ διαδικασία αὐτοπαλινδρομήσεως τάξεως P ἤτοι A.R. (p), οἱ θεωρητικοὶ συντελεστὲς μερικῆς αὐτοσυσχετίσεως εἶναι μηδέν, ἐάν $k>p$ δηλαδὴ $\varphi_k = 0$, $\kappa = p+1, p+2,\ldots$, οἱ δὲ συντελεστὲς αὐτοσυσχετίσεως φθίνουν ἀπολύτως (εἴτε ἐκθετικῶς εἴτε ὡς μία ἀποσβενυμένη ταλάντωση) μετὰ τῆς ὑστερήσεως (Βλέπε σχῆμα 1)

Σὲ μία διαδικασία κινητῶν μέσων MA(q) οἱ ρόλοι τῶν ἀνωτέρω δύο συναρτήσεων ρ_h φ_κ ἀντιστρέφονται.

Δηλαδή:

$$\rho_h = 0, \qquad h = q+1, \qquad q+2,\ldots$$

Καὶ ἡ συνάρτηση φ_κ φθίνει ἀπολύτως μετὰ τοῦ κ (βλέπε σχῆμα 2).

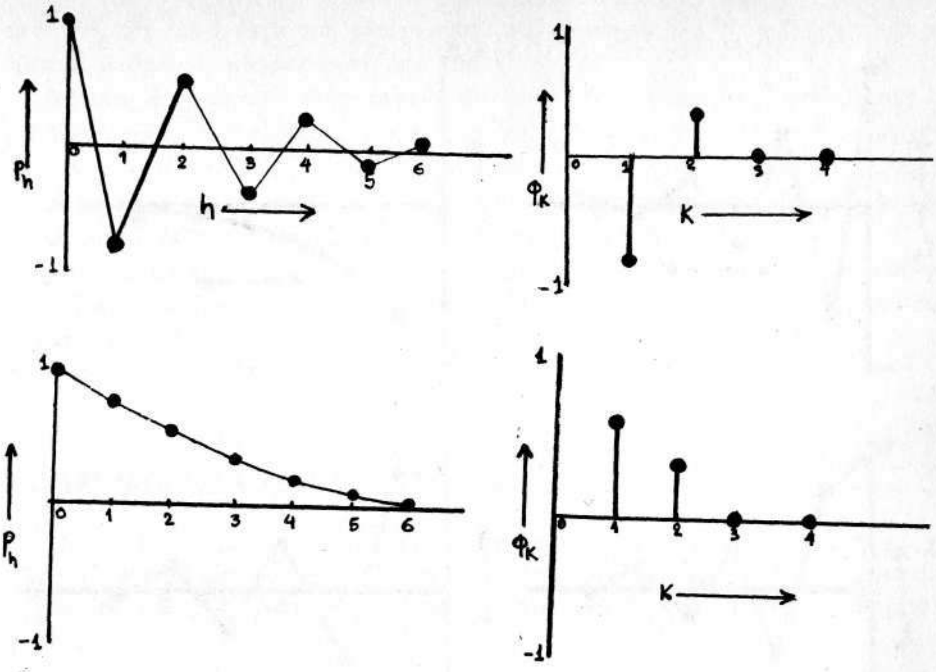
Τέλος οἱ θεωρητικοὶ συντελεστὲς αὐτοσυσχετίσεως καὶ μερικῆς αὐτοσυσχετίσεως μιᾶς μικτῆς διαδικασίας ARMA (p,q) φθίνουν ἀμφότεροι ἀπολύτως μετὰ τῶν h καὶ k ἀντιστοίχως (Βλέπε σχῆμα 3).

13. ΔΕΙΓΜΑΤΙΚΕΣ ΙΔΙΟΤΗΤΕΣ ΤΩΝ ΕΚΤΙΜΗΤΡΙΩΝ r_h, $\hat{\varphi}_k$.

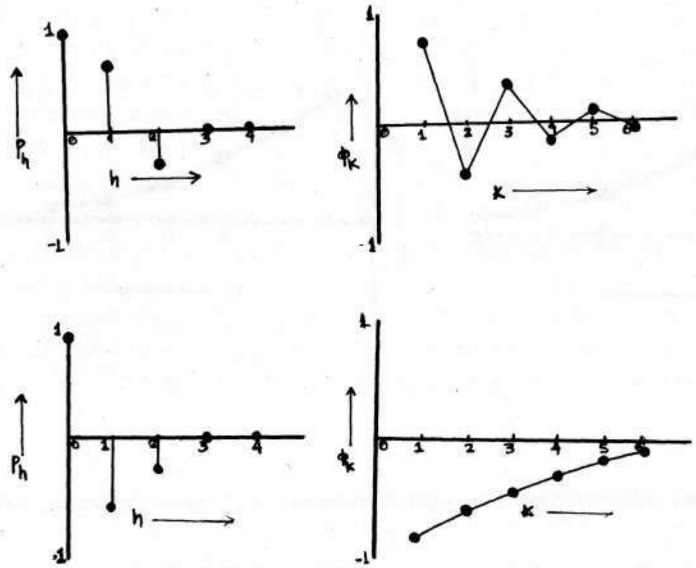
Ἐπειδὴ οἱ θεωρητικοὶ συντελεστὲς αὐτοσυσχετίσεως καὶ μερικῆς αὐτοσυσχετίσεως εἶναι ἄγνωστοι καὶ βεβαίως οἱ ἐκτιμήσεις r_h, $\hat{\varphi}_k$ θὰ διαφέρουν ἀπὸ αὐτούς, θὰ πρέπει νὰ λάβουμε ὑπ' ὄψη ὅτι ([1] κεφάλαιο 48)

$$E\,(r_h) \simeq -\,1/N$$

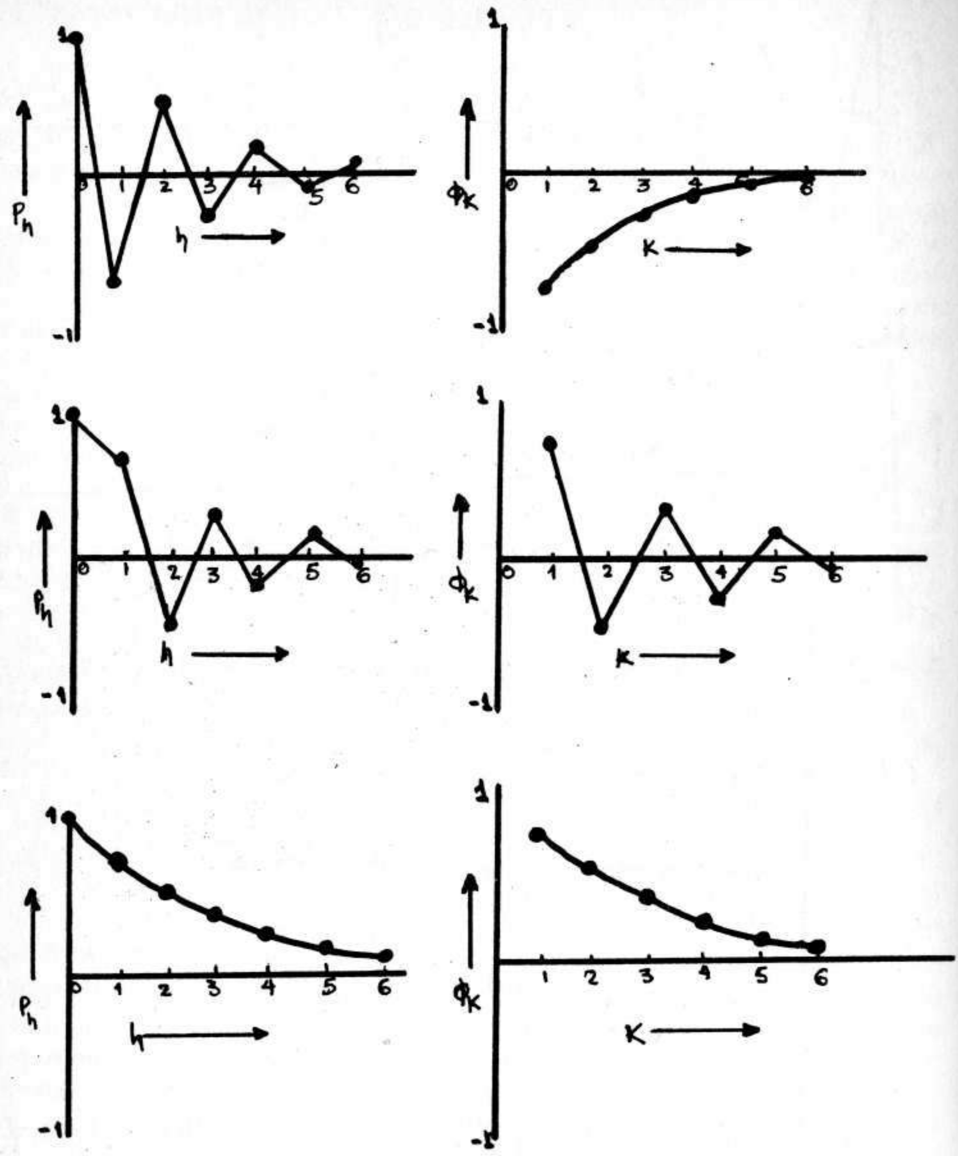
$$\mathrm{Var}\,(r_h) = 1/N$$



Σχῆμα 1: Συναρτήσεις αὐτοσυσχετίσεως καὶ μερικῆς αὐτοσυσχετίσεως δύο διαδικασιῶν AR(2).



Σχῆμα 2: Συναρτήσεις αὐτοσυσχετίσεως καὶ μερικῆς αὐτοσυσχετίσεως δύο διαδικασιῶν M.A(2).



Σχῆμα 3: Συναρτήσεις αὐτοσυσχετίσεως καὶ μερικῆς αὐτοσυσχετίσεως τριῶν διαδικασιῶν ARMA(1,1).

Ὑποθέτοντες τώρα ὅτι $\rho_h = 0$ ἡ πυκνότης τοῦ $\hat{\rho}_h$ (γιὰ N>50) εἶναι κατὰ προσέγγιση ἡ κανονική. Συνεπῶς ἔχοντες ὑπολογίσει ἀπὸ τὰ δεδομένα τὴν συνάρτηση, (r_h, h = 1,2..., 20) ἡ ὑπόθεση $\rho_h = 0$ ἔναντι τῆς $\rho_h \neq 0$ θὰ γίνεται δεκτὴ σὲ ἐπίπεδο σημαντικότητος 5% ἐὰν ἡ ἐκτίμηση r_h εὑρίσκεται στὸ διάστημα $-1/N \pm 2/\sqrt{N}$ τὸ ὁποῖο δύναται ἀκόμη νὰ προσεγγισθεῖ ὑπὸ τοῦ $\pm 2/\sqrt{N}$.

Βεβαίως θὰ πρέπει νὰ ἐνθυμούμεθα ὅτι ἔστω ἐὰν οἱ θεωρητικοὶ συντελεστὲς εἶναι ὅλοι μηδέν, θὰ ἀναμένουμε μία στὶς εἴκοσι πρῶτες ἐκτιμήσεις r_h νὰ διαφέρει σημαντικὰ ἀπὸ τὸ μηδέν, ἐὰν βεβαίως ἐργαζόμεθα σὲ ἐπίπεδο σημαντικότητος 5% Ἡ παρατήρηση αὐτὴ διευκολύνει στὴν ἑρμηνεία τοῦ διαγράμματος τῶν δειγματικῶν συντελεστῶν αὐτοσυσχετίσεως μιᾶς χρονολογικῆς σειρᾶς.

Ἴδια συμπεράσματα ἰσχύουν γιὰ τὸν ἐκτιμητὴ $\hat{\varphi}_k$, δηλαδὴ ἡ ὑπόθεση $\varphi_k = o$ θὰ γίνεται δεκτὴ σὲ ἐπίπεδο σημαντικότητος 5% ἐὰν ἡ ἐκτιμήτρια $\hat{\varphi}_k$ εὑρίσκεται στὸ διάστημα $\pm 2/\sqrt{N}$.

14. ΑΡΙΘΜΗΤΙΚΑ ΠΑΡΑΔΕΙΓΜΑΤΑ

Εἰς τὴν (11) εἴπαμε ὅτι μία ὁποιαδήποτε χρονολογικὴ σειρὰ ἀπηλαγμένη ἐποχικῶν κυμάνσεων) εἶναι δυνατὸ νὰ μετασχηματισθεῖ σὲ εὐσταθή. Οἱ Box καὶ Jenkins ([1] σελὶς 11) ἰσχυρίζονται ὅτι ἂν ἡ χρονολογικὴ σειρὰ X_t δὲν εἶναι εὐσταθὴς τότε ἡ σειρά:

$Z_t = \Delta X_t = X_t - X_{t-1}$, $t = 1,2,..., N$ τῶν διαδοχικῶν διαφορῶν τῆς X_t θὰ εἶναι εὐσταθής. Ἐὰν ὄχι, τότε ἡ σειρὰ τῶν δευτέρων διαφορῶν $W_t = \Delta Z_t = \Delta^2 X_t$ θὰ εἶναι εὐσταθὴς κ.ο.κ. Συνεπῶς ὁ τελεστὴς Δ^d μετασχηματίζει μιὰ σειρὰ σὲ εὐσταθῆ ἐὰν τὸ d εἶναι «ἀρκετά» μεγάλο. Στὴν πράξη τὸ d δὲν ὑπερβαίνει τὸν φυσικὸ τρία. Ἐπομένως ἡ τριάδα (p, d, q) παρέχει τὴν πλήρη ταυτοποίηση μιᾶς χρονολογικῆς σειρᾶς. Δοθείσης μιᾶς χρονολογικῆς σειρᾶς X_t τῆς ὁποίας οἱ d τάξεως διαφορὲς συνιστοῦν εὐσταθῆ χρονολογικὴ σειρά, τότε αὐτὴ θὰ δύναται νὰ ἐξομαλυνθεῖ ἀπὸ τὴν λεγόμενη οἰκογένεια, ARIMA (p, d, q) ἡ ὁποία προφανῶς εἶναι εὐρύτερη ἀπὸ τὴν οἰκογένεια (1). Θὰ πρέπει νὰ σημειωθεῖ ὅτι ἡ συνάρτηση αὐτοσυσχετίσεως r_h μιᾶς μὴ εὐσταθοῦς χρονολογικῆς σειρᾶς ἔχει τὴν ἰδιότητα νὰ μὴ φθίνει ἐκθετικὰ μετὰ τῆς ὑστερήσεως ἀλλὰ πολὺ «ἀργά» γεγονὸς τὸ ὁποῖο μᾶς ὁδηγεῖ νὰ μετασχηματίσουμε τὰ δεδομένα διὰ μέσω τοῦ τελεστοῦ Δ^d.

Οἱ κατωτέρω δύο πίνακες ἀναφέρονται στὶς συναρτήσεις αὐτοσυσχετίσεως καὶ μερικῆς αὐτοσυσχετίσεως δύο χρονολογικῶν σειρῶν. Γιὰ τὴν σειρὰ Β περιέχονται ἐπίσης οἱ ἴδιες συναρτήσεις γιὰ τὶς πρῶτες διαφορὲς ΔX_t αὐτῆς.

ΣΕΙΡΑ Α (100 παρατηρήσεις)

Πίνακας 1: συντελεστῶν αὐτοσυσχετίσεως r_h, $h = 1,2,..., 20$

ΥΣΤΕΡΗΣΕΙΣ	ΣΥΝΤΕΛΕΣΤΕΣ ΑΥΤΟΣΥΣΧΕΤΙΣΕΩΣ									
1 – 10	– 0,60	0,40	0,22	0,15	–0,10	0,04	0,07	0,01	–0,01	0,03
11 – 20	– 0,09	0,08	0,01	0,00	–0,01	–0,03	–0,07	0,03	–0,09	0,01

ΣΥΝΤΕΛΕΣΤΕΣ ΜΕΡΙΚΗΣ ΑΥΤΟΣΥΣΧΕΤΙΣΕΩΣ

1 – 10	0,50	–0,22	0,07	0,05	0,13	–0,14	0,03	–0,01	0,00	0,10
11 – 20	0,13	–0,01	0,02	0,04	0,07	0,13	0,18	0,00	0,02	–0,01

ΣΕΙΡΑ Β (225 παρατηρήσεις)

Πίνακας 2: συντελεστῶν αὐτοσυσχετίσεως r_h , $h = 1,2,..., 20$

ΣΕΙΡΑ	ΥΣΤΕΡΗΣΕΙΣ	ΣΥΝΤΕΛΕΣΤΕΣ ΑΥΤΟΣΥΣΧΕΤΙΣΕΩΣ									
X_t	1 – 10	0,55	0,50	0,40	0,32	0,34	0,40	0,30	0,30	0,25	0,32
	10 – 20	0,18	0,15	0,16	0,05	0,02	0,18	0,10	0,10	0,08	0,01
Δx_t	1 – 10	–0,40	0,03	–0,06	–0,02	–0,02	–0,11	0,03	0,01	0,03	0,02
	10 – 20	0,05	0,04	–0,01	0,00	0,10	–0,10	–0,19	0,02	–0,11	0,03

ΣΥΝΤΕΛΕΣΤΕΣ ΜΕΡΙΚΗΣ ΑΥΤΟΣΥΣΧΕΤΙΣΕΩΣ

X_t	1 – 10	0,60	0,20	0,08	0,06	0,08	0,12	0,15	–0,10	0,03	–0,07
	10 – 20	0,10	–0,07	–0,03	–0,06	0,09	0,12	–0,10	0,02	0,01	0,03
Δx_t	1 – 10	–0,68	–0,39	–0,28	–0,16	–0,10	–0,05	0,07	–0,02	–0,01	0,02
	10 – 20	0,02	–0,03	0,10	0,12	–0,02	0,03	0,10	–0,01	0,02	0,08

Ἂς δεχθοῦμε ὅτι οἱ συντελεστὲς τοῦ πίνακα Α προέρχονται ἀπὸ μία χρονολογικὴ σειρὰ 100 παρατηρήσεων. Εἶναι φανερὸ ὅτι οἱ συντελεστὲς αὐτοσυσχετίσεως φθίνουν «ταχέως» μετὰ τῆς ὑστερήσεως καὶ μάλιστα ὅλοι μετὰ τὸν τρίτο συντελεστὴ δύνανται νὰ θεωρηθοῦν πρακτικὰ ἴσοι πρὸς τὸ μηδέν, διότι βρίσκονται μέσα στὸ διάστημα $\pm 2/\sqrt{100} = \pm 0{,}20$. Ἐπίσης οἱ συντελεστὲς μερικῆς αὐτοσυσχετίσεως δύνανται νὰ θεωρηθοῦν μηδενικοὶ πλὴν τῶν δύο πρώτων ποὺ εἶναι ἀντιστοίχως $\hat{\varphi}_1 = 0{,}50$, $\hat{\varphi}_2 = -0{,}22$. Οἱ ἀνωτέρω παρατηρήσεις μᾶς ὁδηγοῦν στὸ νὰ δεχθοῦμε (βλέπε καὶ σχῆμα 1) ὅτι ἡ σειρὰ X_t προέρχεται ἀπὸ μία διαδικασία αὐτοπαλινδρομήσεως τάξεως δύο, δηλαδή:

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

ὅπου $\varphi_1 \varphi_2$ παράμετροι πρὸς ἐκτίμηση:

Ἂς δεχθοῦμε ἐπίσης ὅτι οἱ συντελεστὲς τοῦ πίνακα Β προέρχονται ἀπὸ μία χρονολογικὴ σειρὰ 225 παρατηρήσεων. Ἐπειδὴ οἱ συντελεστὲς αὐτοσυσχετίσεως τῆς ἀρχικῆς (μὴ μετασχηματισμένης) σειρᾶς X_t δὲν φθίνουν «ταχέως» μετὰ τῆς ὑστερήσεως ἀλλὰ πολὺ «ἀργὰ» ἡ σειρὰ δὲν μπορεῖ νὰ θεωρηθεῖ εὐσταθὴς Γι' αὐτὸ θὰ πρέπει νὰ ἐξετασθοῦν οἱ πρῶτες διαφορὲς Δx_t τῆς σειρᾶς μήπως αὐτὲς μποροῦν νὰ θεωρηθοῦν ὅτι προέρχονται ἀπὸ μία εὐσταθῆ χρονολογικὴ σειρά.

Πράγματι οἱ συντελεστὲς αὐτοσυσχετίσεως μὲ ὑστέρηση πάνω ἀπὸ μία χρονικὴ μονάδα, τῶν πρώτων διαφορῶν, μποροῦν νὰ θεωρηθοῦν μηδέν, διότι εὑρίσκονται ὅλοι πλὴν τοῦ δεκάτου ἑβδόμου μέσα στὸ διάστημα. $\pm 2/\sqrt{225} = \pm 0{,}13$. Ὁ δέκατος ἕβδομος συντελεστής εἶναι $r_{17} = -0{,}19$ γεγονὸς τὸ ὁποῖο μποροῦμε νὰ ἀποδόσουμε στὴν τύχη, δεδομένου ὅτι ἐργαζόμαστε σὲ ἐπίπεδο σημαντικότητας 5% καὶ ἐπομένως θὰ πρέπει νὰ περιμένουμε μία στὶς εἴκοσι πρῶτες ἐκτιμήσεις τῶν συντελεστῶν αὐτοσυσχετίσεως νὰ διαφέρει σημαντικὰ ἀπὸ τὸ μηδὲν χωρὶς αὐτὸ νὰ συμβαίνει ὑποχρεωτικὰ στὸν πληθυσμιακὸ συντελεστὴ αὐτοσυσχετίσεως .

Ἐξ' ἄλλου οἱ συντελεστὲς μερικῆς αὐτοσυσχετίσεως φθίνουν «ἐκθετικά» μετὰ τῆς ὑστερήσεως γεγονὸς τὸ ὁποῖο μᾶς ὁδηγεῖ στὸ νὰ δεχθοῦμε ὅτι ἡ μετασχηματισμένη σειρὰ $Z_t = \Delta x_t$ δύναται νὰ θεωρηθεῖ ὅτι προέρχεται (Βλέπε σχῆμα 2) ἀπὸ μία διαδικασία κινητῶν μέσων ΜΑ (1), δοθέντος ὅτι μόνο ὁ πρῶτος συντελεστὴς αὐτοσυσχετίσεως διαφέρει σημαντικὰ ἀπὸ τὸ μηδὲν ($r_1 = -0{,}40$). Ἑπομένως ἡ ἀρχικὴ σειρὰ δύναται νὰ προσεγγισθεῖ ἀπὸ τὸ ὑπόδειγμα ARIMA (0,1,1).

δηλαδή:

$$Z_t = \Delta X_t = u_t - \partial_1 u_{t-1}$$

ἢ

$$X_t - X_{t-1} = u_t - \partial_1 u_{t-1}$$

ὅπου ∂_1 παράμετρος πρὸς ἐκτίμηση.

Ἀναλόγως μποροῦμε νὰ ταυτοποιήσουμε ὁποιαδήποτε χρονολογικὴ σειρὰ ποὺ ἤθελε παρουσιασθεῖ στὴν πράξη.

Θὰ πρέπει ὅμως νὰ τονισθεῖ ὅτι γιὰ μία ἐπιτυχημένη ταυτοποίηση ἀπαιτεῖται ἀφ' ἑνὸς μὲν μεγάλη πεῖρα ἐκ μέρους τοῦ ἐρευνητοῦ ἐφ' ἑτέρου δὲ ὕπαρξη κατάλληλου ὑπολογιστοῦ, καὶ προγραμμάτων. Διότι τότε μόνο εἶναι δυνατὸ νὰ ἐφαρμοσθοῦν μὲ ἐπιτυχία οἱ μέθοδοι τῶν Box-Jenkins.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. BOX-JENKINS (1970): Time series analysis, Forecasting and control. San Fransisco: Holden-Day.
2. KENDALL (1973). Time Servies. Lon-Griffin.
3. CHATFIELD (1974). The analysis of time Series, theory and practice. Chapman and Hall.
4. ANDERSON (1971). The Statistical analysis of time Series. New York, Willey.
5. HANNAN (1960). Time Series Analysis, London, Methuen
6. NELSON (1973). Applied Time Series analysis for managerial forecasting. San Francisco. Holden Day.
7. Ο. ΠΑΠΑΔΗΜΑ (1978). Γραμμικὰ στοχαστικὰ ὑποδείγματα.