

# **ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ ΓΕΝΙΚΟ ΤΜΗΜΑ**



## **ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ ΕΦΑΡΜΟΣΜΕΝΕΣ ΕΠΙΣΤΗΜΕΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑ**

**ΔΙΠΛΩΜΑΤΙΚΗ ΔΙΑΤΡΙΒΗ ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ  
ΚΑΤΕΥΘΥΝΣΗ : «ΕΦΑΡΜΟΣΜΕΝΑ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΑ ΜΑΘΗΜΑΤΙΚΑ»**

## **ΕΦΑΡΜΟΓΗ ΤΗΣ ΜΕΘΟΔΟΥ ΧΡΟΝΟΣΕΙΡΩΝ ΣΤΗΝ ΠΡΟΒΛΕΨΗ ΠΑΡΑΓΩΓΗΣ ΠΕΤΡΕΛΑΙΟΥ**

**ΠΑΠΑΔΟΠΟΥΛΟΥ ΓΕΩΡΓΙΑ**

Επιβλέπων: Επίκουρος Καθηγητής **Δάρας Τρύφων**

**ΧΑΝΙΑ, 2015**

# Περιεχόμενα

ΕΙΣΑΓΩΓΗ    iii

## I    ΑΝΑΛΥΣΗ ΧΡΟΝΟΣΕΙΡΩΝ    1

### 1    ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΧΡΟΝΟΣΕΙΡΩΝ    3

- 1.1    Η ΕΝΝΟΙΑ ΤΗΣ ΧΡΟΝΟΣΕΙΡΑΣ    3
- 1.2    ΣΤΑΤΙΣΤΙΚΑ ΜΕΓΕΘΗ ΧΡΟΝΟΣΕΙΡΩΝ    5
- 1.3    ΣΤΑΣΙΜΟΤΗΤΑ ΚΑΙ ΑΥΤΟΣΥΣΧΕΤΗΣΗ    6

### 2    ΜΕΘΟΔΟΙ ΠΡΟΒΛΕΨΗΣ    9

- 2.1    ΚΡΙΤΗΡΙΑ ΑΞΙΟΛΟΓΗΣΗΣ ΤΩΝ ΜΕΘΟΔΩΝ ΠΡΟΒΛΕΨΗΣ    9
- 2.2    ΜΕΘΟΔΟΣ ΕΞΟΜΑΛΥΝΣΗΣ    11
  - 2.2.1    ΑΠΛΟΣ ΚΙΝΗΤΟΣ ΜΕΣΟΣ    11
  - 2.2.2    ΑΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ    11
  - 2.2.3    ΔΙΠΛΟΣ ΚΙΝΗΤΟΣ ΜΕΣΟΣ    11
  - 2.2.4    ΜΕΘΟΔΟΣ BROWN    11
  - 2.2.5    ΜΕΘΟΔΟΣ HOLT    11
  - 2.2.6    ΜΕΘΟΔΟΣ WINTERS    11
- 2.3    ΜΕΘΟΔΟΣ ΔΙΑΣΠΑΣΗΣ ΧΡΟΝΟΣΕΙΡΩΝ    11
  - 2.3.1    ΑΝΑΛΥΣΗ ΠΟΧΙΚΟΤΗΤΑΣ    11
  - 2.3.2    ΑΝΑΛΥΣΗ ΜΑΚΡΟΧΡΟΝΙΑΣ ΤΑΣΗΣ    11
  - 2.3.3    ΑΝΑΛΥΣΗ ΚΥΚΛΙΚΟΤΗΤΑΣ ΚΑΙ ΜΗ ΚΑΝΟΝΙΚΟΤΗΤΑΣ    11

## II    ΕΦΑΡΜΟΓΗ ΤΩΝ ΜΕΘΟΔΩΝ ΓΙΑ ΤΗΝ ΠΡΟΒΛΕΨΗ ΠΑ- ΡΑΓΩΓΗΣ ΠΕΤΡΕΛΑΙΟΥ    13

### 3    Όνομα τρίτου κεφαλαίου    15

- 3.1    Όνομα πρώτης ενότητας    15
  - 3.1.1    Όνομα πρώτης υποενότητας    15
- 3.2    Όνομα δεύτερης ενότητας    15



# ΕΙΣΑΓΩΓΗ

Εδώ η Γεωργία θα γράψει την εισαγωγή της. Στην πορεία θα συζητήσουμε την δομή της εισαγωγής [3]. Πρέπει η εισαγωγή να έχει παρόμοιο θεματικό σκελετό με την [2, 1] περίληψη και τα συμπεράσματα στο τέλος της εργασίας.

Το αντικείμενο είναι λοιπόν η ανάλυση χρονοσειρών, δηλαδή η χρήση μεθόδων που θα μας επιτρέψουν να διερευνήσουμε το μηχανισμό (στοχαστική διαδικασία ή δυναμικό σύστημα) που παράγει τη χρονοσειρά, να εκτιμήσουμε χαρακτηριστικά του, να αναπτύξουμε μοντέλο για να το περιγράψουμε και να κάνουμε προβλέψεις της εξέλιξης του, δηλαδή τις επόμενες τιμές στη χρονοσειρά.

Γ. Παπαδοπούλου, Χανιά 2015.



**Μέρος Ι**

**ΑΝΑΛΥΣΗ ΧΡΟΝΟΣΕΙΡΩΝ**



# Κεφάλαιο 1

## ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΧΡΟΝΟΣΕΙΡΩΝ

### 1.1 Η ΕΝΝΟΙΑ ΤΗΣ ΧΡΟΝΟΣΕΙΡΑΣ

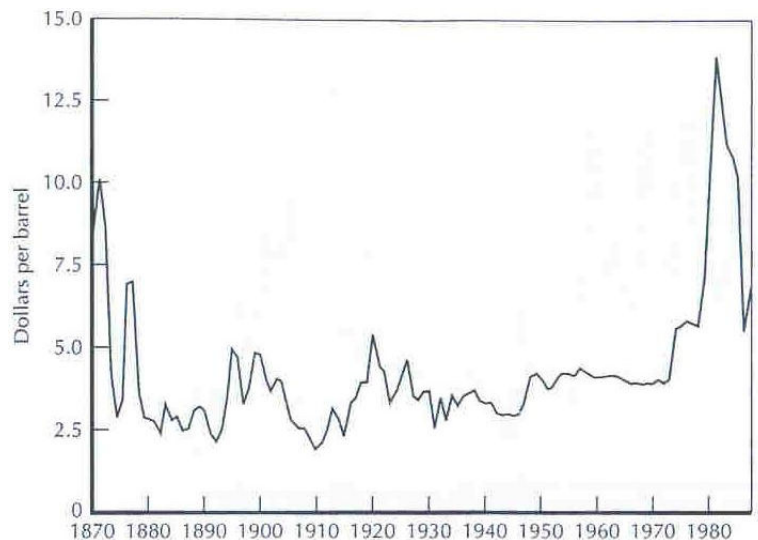
Η χρονοσειρά μπορεί να ορισθεί ως μια συλλογή διαδοχικών χρονικών παρατηρήσεων της τιμής κάποιου μεγέθους. Πρόκειται ουσιαστικά για μια στοχαστική διαδικασία, μιας και η εξέλιξη των τιμών του μεγέθους επηρεάζεται από τυχαίους παράγοντες, ενώ η τιμή κάθε χρονικής στιγμής συνιστά και μια ξεχωριστή τυχαία μεταβλητή. Με τον όρο χρονοσειρά, δηλαδή εννοούμε συνήθως μια ακολουθία  $X_t : t = 0, 1, 2, \dots$ , όπου κάθε  $X_t$  εκφράζει την κατά την χρονική στιγμή  $t$  κατάσταση ενός συστήματος το οποίο εξελίσσεται στο χρόνο κατά τυχαίο εν γένει τρόπο (stochastic system). Παραδείγματα τέτοιων χρονοσειρών είναι:

- (i) Οι ημερήσιες, αεροπορικές και οδικές, αφίξεις τουριστών στην χώρα μας  $X_t$  με  $t = 1, 2, \dots$
- (ii) Ο αριθμός  $X_t$  πελατών μέσα σε ένα πολυκατάστημα κατά τη χρονική στιγμή  $t$  με  $t \in [0, ]$ .
- (iii) Ο συνολικός αριθμός τροχαίων ατυχημάτων  $X_t$  κατά μήκος μιας οδικής αρτηρίας στο χρονικό διάστημα  $[0, t]$  με  $t \geq 0$ .
- (iv) Η ημερήσια κατανάλωση ηλεκτρικού ρεύματος καθώς και η ημερήσια κατανάλωση ύδατος,  $X_t$  και  $Y_t$  αντίστοιχα, σε μια μεγάλη γεωγραφική περιοχή της χώρας με  $t = 1, 2, \dots$
- (v) Οι οικονομικές χρονοσειρές, όπως το ετήσιο ακαθάριστο εθνικό προϊόν και ετήσιο ισοζύγιο εξωτερικών συναλλαγών  $X_t$  και  $Y_t$  αντίστοιχα, με  $t = 1, 2, \dots$
- (vi) Οι μετεωρολογικές χρονοσειρές, όπως η θερμοκρασία περιβάλλοντος και ατμοσφαιρική πίεση,  $X_t$  και  $Y_t$  αντίστοιχα, σε συγκεκριμένη γεωγραφική περιοχή με γεωγραφικές συντεταγμένες  $(l, a, h)$  κατά την χρονική στιγμή  $t$ . Εδώ η χρησιμοποιούμενη παράμετρος  $t$  είναι περισσότερο σύνθετη και συγκεκριμένα  $t = (l, a, h, t)$ .



Όπως διαπιστώνει κανείς από τα παραπάνω παραδείγματα, οι χρονοσειρές μπορούν να αφορούν *διακριτά* μεγέθη  $X_t$  σε *διακριτό* χρόνο  $t$ , περίπτωση (i), *διακριτά* μεγέθη  $X_t$  σε *συνεχή* χρόνο  $t$ , περιπτώσεις (ii) και (iii), *συνεχή* μεγέθη  $X_t$  σε *διακριτό* χρόνο  $t$ , περιπτώσεις (iv) και (v) και *συνεχή* μεγέθη  $X_t$  σε *συνεχή* χρόνο  $t$ , περίπτωση (vi). Το πρόβλημα είναι η “πρόβλεψη” μελλοντικών τιμών της χρονοσειράς με βάση τις μέχρι σήμερα τιμές της ίδιας χρονοσειράς, περιπτώσεις (i)–(iii), είτε ακόμα και σε συνδυασμό με τις μέχρι σήμερα τιμές μιας άλλης χρονοσειράς η οποία εξελίσσεται παράλληλα με την πρώτη και επιδρά πάνω σ’ αυτή, περιπτώσεις (iv)–(vi), οπότε μιλάμε για *πολυμεταβλητές χρονοσειρές*. Το σύνολο των δυνατών καταστάσεων ονομάζεται *χώρος καταστάσεων* και συμβολίζεται με  $S$ , ένα (μονοδιάστατο) υποσύνολο του  $R$ , ή γενικότερα ένα πολυδιάστατο υποσύνολο του  $R^d$ , ενώ το σύνολο τιμών του  $t$  ονομάζεται *παραμετρικός χώρος*, συμβολίζεται με  $T$  και μπορεί επίσης να είναι υποσύνολο του  $R^k$ , όταν χρειάζεται ένα πολυδιάστατο  $t$  για να καθορίσουμε πέραν του χρόνου  $t$  και γεωγραφικές π.χ. συντεταγμένες  $(l, a, h)$  σε χωρο-χρονοσειρές (spatial time series), βλ. παράδειγμα (vi) παραπάνω. Σημειώνεται οι όροι *διακριτά* και *συνεχή* μεγέθη είναι σε αντιστοιχία με τους όρους *διακριτές* και *συνεχείς* τυχαιές μεταβλητές.

Ένα παράδειγμα μίας χρονοσειράς απεικονίζεται στο ακόλουθο σχήμα 1.1, όπου παρουσιάζεται το γράφημα της ετήσιας τιμής του αργού πετρελαίου σε αμερικανικά δολάρια ανά βαρέλι. Ο χρονικός ορίζοντας είναι 117 χρόνια, από το έτος 1870 έως το έτος 1987. Το γράφημα ελήφθη από το βιβλίο των R. Pindyck και D. Rubinfeld (1998).



**Σχήμα 1.1: Γράφημα χρονοσειράς τιμής αργού πετρελαίου σε δολάρια ανά βαρέλι.**

Η συστηματική μελέτη μιας χρονοσειράς ξεκινάει με την επισκόπηση του γραφήματός της στο πεδίο του χρόνου, από το οποίο μπορούν αρχικά να ανιχνευθούν τρία βασικά ποιοτικά χαρακτηριστικά της: Η τάση, η εποχικότητα και οι ακραίες παρατηρήσεις.

- Η **τάση** (trend) γενικά θα μπορούσε να οριστεί ως η μακροπρόθεσμη μεταβολή του μέσου επιπέδου των τιμών μιας χρονοσειράς. Έτσι, μπορεί η τάση των τιμών

να είναι αυξητική, πτωτική ή σταθερή σε ένα συγκεκριμένο χρονικό διάστημα, ενώ μπορεί και να έχει τη μορφή κάποιας συνάρτησης στο εν λόγω διάστημα. Να σημειωθεί ότι η έννοια “μακροπρόθεσμη μεταβολή” εξαρτάται από την εκάστοτε εφαρμογή που εξετάζεται.

- Η **εποχικότητα** (seasonal) μπορεί να οριστεί σαν μια περιοδική διακύμανση που έχει σταθερό μήκος. Η εν λόγω διακύμανση τις περισσότερες φορές διακρίνεται εύκολα και μπορεί να ερμηνευθεί στα πλαίσια του υπό μελέτη φαινομένου. Φερ’ ειπείν, αν κανείς επιθυμούσε να αναλύσει τη χρονοσειρά των τιμών των καυσίμων σε βάθος χρόνων, είναι λογικό να περιμένει να παρατηρήσει μια σχετική άνοδο κατά τους χειμερινούς μήνες κάθε έτους.
- Οι **ακραίες παρατηρήσεις** (outliers) είναι οι απομονωμένες παρατηρήσεις που εμφανίζονται στο γράφημα κάποιας χρονοσειράς ως απότομες αλλαγές στο πρότυπο συμπεριφοράς της. Τα outliers μελετήθηκαν αρχικά κατά κύριο λόγο από τον A. J. Fox (1972), ο οποίος μάλιστα εισήγαγε δύο τύπους. Ο τύπος I αφορά περιπτώσεις όπου η ύπαρξη μιας ακραίας τιμής δεν έχει καμία επίδραση στις ακόλουθες παρατηρήσεις. Αντιθέτως, ο τύπος II αφορά περιπτώσεις όπου υπάρχει επίδραση στη μετέπειτα συμπεριφορά των τιμών της χρονοσειράς, παράγοντας μια σειρά λιγότερο ή περισσότερο ακραίων παρατηρήσεων ή αλλάζοντας εξ’ ολοκλήρου τα χαρακτηριστικά της.

## 1.2 ΣΤΑΤΙΣΤΙΚΑ ΜΕΓΕΘΗ ΧΡΟΝΟΣΕΙΡΩΝ

Ως γνωστόν, τα βασικότερα χαρακτηριστικά μιας τυχαίας μεταβλητής  $X$  είναι η μέση τιμή  $\mu = E[X]$ , η διασπορά  $\sigma^2 = V[X]$  και όταν έχουμε να κάνουμε με ζεύγη τυχαίων μεταβλητών, η μικτή ροπή 2ας τάξης, δηλαδή η συνδιακύμανση  $\sigma_{xy} = \text{Cov}(X, Y)$ . Κατ’ επέκταση, τα βασικότερα χαρακτηριστικά μιας χρονοσειράς  $\{X_t : t \in T\}$  είναι:

- Η συνάρτηση *μέσης τιμής*
- Η συνάρτηση *διασποράς*
- Η συνάρτηση *αυτοσυνδιακύμανσης*

(i) **Η μέση τιμή** ή αναμενόμενη τιμή:

$$\mu(t) = E[X_t], \quad t \in T$$

Στο σημείο αυτό θα πρέπει να αναφέρουμε ότι η μέση τιμή  $\mu_t$  σχετίζεται άμεσα με την έννοια της *τάσης* της χρονοσειράς, εφόσον εκφράζεται ως συνάρτηση της χρονικής στιγμής  $t$  της παρατήρησης  $X_t$ . Συγκεκριμένα, αν μια χρονοσειρά παρουσιάζει αυξητική ή πτωτική τάση αντιστοίχως σε ένα χρονικό διάστημα, αυτό θα αποτυπώνεται και στη μέση τιμή της  $\mu_t$  ως συνάρτηση του χρόνου. Ακόμη, η μη ύπαρξη τάσης σε ένα χρονικό διάστημα αποτυπώνεται στη σταθερή αναμενόμενη τιμή  $\mu_t$  στο εν λόγω διάστημα.

(ii) **Η διασπορά:**

$$\sigma^2(t) = V[X_t] = E[(X_t - \mu_t)^2], \quad t \in T$$

## (iii) Η αυτοσυνδιακύμανση:

$$\gamma(t, h) = \text{Cov}(X_t, X_{t+h}) = E[(X_t - \mu_t)(X_{t+h} - \mu_{t+h})], \quad t, h \in T$$

Είναι προφανές ότι ισχύει:  $\sigma^2(t) = \gamma(t, 0) = \text{Cov}(X_t, X_t)$ ,  $t, h \in T$

Είναι γνωστό επίσης ότι όταν οι διασπορές δύο τυχαίων μεταβλητών  $X$  και  $Y$ , είναι πεπερασμένες, τότε τόσο οι μέσες τιμές αυτών όσο και η συνδιακύμανση αυτών είναι πεπερασμένες ποσότητες. Τούτο διότι αφενός πρέπει να έχουμε  $E[|X|^2] < \infty$ ,  $E[|Y|^2] < \infty$ , αφού διαφορετικά δεν ορίζονται οι διασπορές, και αφετέρου, με εφαρμογή της ανισότητας Cauchy-Schwarz

$$E[|X||Y|] \leq \sqrt{E[X^2] E[Y^2]},$$

προκύπτει:

$$|E[X]| \leq E[|X|] \leq \sqrt{E[X^2]} < \infty, \quad |E[Y]| \leq E[|Y|] \leq \sqrt{E[Y^2]} < \infty,$$

και

$$|\text{Cov}(X, Y)| \leq \sqrt{V[X] V[Y]} < \infty.$$

### 1.3 ΣΤΑΣΙΜΟΤΗΤΑ ΚΑΙ ΑΥΤΟΣΥΣΧΕΤΗΣΗ

Είναι φανερό ότι έχοντας παρατηρήσει μια χρονοσειρά  $\{X_t : t \in T\}$  από κάποια χρονική στιγμή  $t = 0$ , έστω μέχρι και την παρούσα χρονική στιγμή  $t = s$ , αν δηλαδή γνωρίζουμε την τροχιά αυτής  $\{X_t : 0 \leq t \leq s\}$ , και θέλουμε να προβλέψουμε μελλοντικές τιμές αυτής  $X_{s+h}$ , με  $h > 0$ , θα πρέπει να βασιστούμε στις μέχρι τώρα γνωστές τιμές της και στην εξάρτηση που ενδέχεται να υπάρχει μεταξύ  $X_{s+h}$  και των τιμών  $\{X_t : 0 \leq t \leq s\}$  της χρονοσειράς στο παρελθόν. Τούτο βέβαια με την προϋπόθεση ότι όλα τα πιθανοθεωρητικά χαρακτηριστικά μιας χρονοσειράς, ή τουλάχιστον τα βασικότερα εξ αυτών, παραμένουν αναλλοίωτα στον χρόνο.

Όταν όλα τα πιθανοθεωρητικά χαρακτηριστικά μιας χρονοσειράς παραμένουν αναλλοίωτα στο χρόνο τότε μιλάμε για *αυστηρή στασιμότητα*. Συγκεκριμένα έχουμε:

#### (i) Αυστηρή Στασιμότητα

Η χρονοσειρά  $\{X_t : t \in T\}$  ονομάζεται *αυστηρά στάσιμη* όταν  $\forall n \in \mathbb{N}$ ,  $t_i \in T$  με  $i = 1, \dots, n$  και  $h \in T$  ισχύει η παρακάτω σχέση ισοδυναμίας

$$(X_{t_1}, \dots, X_{t_n}) \sim (X_{t_1+h}, \dots, X_{t_n+h})$$

Εδώ το σύμβολο “ $\sim$ ” διαβάζεται “κατανέμεται όπως”, ή “ισοκατανέμεται με”.

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

#### (ii) Ασθενής Στασιμότητα

Η χρονοσειρά  $\{X_t : t \in T\}$  ονομάζεται *ασθενώς στάσιμη*, ή υπό ευρεία έννοια, *στάσιμη* όταν οι ροπές πρώτης και δεύτερης τάξης είναι σταθερές στο χρόνο, δηλαδή:

- Η μέση τιμή είναι σταθερή:  $\mu = E[X_t]$ ,  $t \in T$
- $\gamma(h) = \text{Cov}(X_t, X_{t+h})$ ,  $t, h \in T$

Η συνάρτηση *αυτοσυσχέτισης* μίας χρονοσειράς ορίζεται από τη σχέση:

$$\rho(h) = \text{Corr}(X_0, X_h) = \frac{\text{Cov}(X_0, X_h)}{\sqrt{V[X_0] V[X_h]}} = \frac{\gamma(h)}{\gamma(0)}, \quad h \in T$$

Προφανώς η συνάρτηση αυτοσυσχέτισης έχει όλες τις ιδιότητες της συνάρτησης αυτοσυνδιακύμανσης με την επιπρόσθετη ιδιότητα  $|\rho(h)| \leq 1$ , ως συνέπεια της ανισότητας Cauchy-Schwarz. Συνεπώς, όλες οι δυνατές τιμές της αυτοσυσχέτισης βρίσκονται εντός του διαστήματος  $[-1, 1]$ .



## Κεφάλαιο 2

# ΜΕΘΟΔΟΙ ΠΡΟΒΛΕΨΗΣ

### 2.1 ΚΡΙΤΗΡΙΑ ΑΞΙΟΛΟΓΗΣΗΣ ΤΩΝ ΜΕΘΟΔΩΝ ΠΡΟΒΛΕΨΗΣ

Η ανάλυση χρονοσειρών (time series analysis) ασχολείται αποκλειστικά με τη διερεύνηση της διαχρονικής συμπεριφοράς των τιμών μιας μεταβλητής, οι παρατηρήσεις της οποίας προέρχονται από χρονοσειρά. Η πρόβλεψη των μελλοντικών τιμών της μεταβλητής σύμφωνα με την ανάλυση χρονοσειρών μπορεί να προέλθει από διάφορες κατηγορίες μεθόδων προβλέψης, όπως η μέθοδος Εξομάλυνσης και η Διάσπαση των χρονοσειρών, με τις οποίες θα ασχοληθούμε στην παρούσα εργασία.

Για την επιλογή της κατάλληλης μεθόδου χρησιμοποιούνται τα κριτήρια αξιολόγησης των μεθόδων προβλέψεων. Τα κριτήρια αυτά βασίζονται στις τιμές των αποκλίσεων των προβλεπόμενων τιμών από τις αντίστοιχες πραγματικές τιμές της χρονοσειράς.

Για μία μεταβλητή  $Y$ , η απόκλιση της προβλεπόμενης τιμής της  $Y_t$  από την αντίστοιχη πραγματική τιμή της  $Y_t$  για την περίοδο  $t$ , όπου  $t = 1, 2, 3, \dots, n$ , ονομάζεται σφάλμα της πρόβλεψης (forecast error), συμβολίζεται με  $e_t$  και ορίζεται ως:  $e_t = Y_t - \hat{Y}_t$

Η παραπάνω σχέση εκφράζει για κάθε περίοδο  $t$  τη διαφορά μεταξύ της πραγματικής τιμής  $Y_t$  και της αντίστοιχης προβλεπόμενης τιμής  $\hat{Y}_t$  που προήλθε από τη μέθοδο πρόβλεψης που χρησιμοποιήθηκε.

Επομένως, για να προσδιορίσουμε την αξιοπιστία μιας συγκεκριμένης μεθόδου πρόβλεψης, θα πρέπει να μελετήσουμε τη διαχρονική συμπεριφορά των τιμών των σφαλμάτων της πρόβλεψης. Αυτό γίνεται με την εφαρμογή διάφορων κριτηρίων, σύμφωνα με τα οποία αξιολογούμε τη χρησιμοποιούμενη μέθοδο πρόβλεψης. Κάθε ένα από τα κριτήρια αυτά ορίζεται από μία συγκεκριμένη συναρτησιακή σχέση των σφαλμάτων της πρόβλεψης και μπορεί να χρησιμοποιηθεί όχι μόνο για την αξιολόγηση μιας μεθόδου πρόβλεψης αλλά και για την επιλογή της “καλύτερης” μεταξύ δύο ή περισσότερων εναλλακτικών μεθόδων προβλέψεων. Τα κριτήρια αυτά είναι:

- **Μέση απόλυτη απόκλιση MAD (Mean Absolute Deviation)**

Η μέση απόλυτη απόκλιση ορίζεται ως το άθροισμα των απόλυτων τιμών του σφάλματος της πρόβλεψης διαιρούμενο με τον αριθμό των περιόδων  $n$ , στις οποίες έγιναν προβλέψεις, δηλαδή:

$$MAD = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t| = \frac{1}{n} \sum_{t=1}^n |e_t|$$

Το MAD εκφράζει τη μέση τιμή των απολύτων αποκλίσεων των προβλεπόμενων τιμών της χρονοσειράς από τις αντίστοιχες πραγματικές και έχει τα ακόλουθα χαρακτηριστικά. Πρώτον, η μονάδα μέτρησης του είναι η ίδια με εκείνη των τιμών της χρονοσειράς και έτσι είναι εύκολη η ερμηνεία του. Δεύτερον, στον υπολογισμό του λαμβάνονται υπ' όψιν μόνο οι απόλυτες τιμές των σφαλμάτων και όχι οι πραγματικές τιμές τους. Αυτό σημαίνει ότι το MAD είναι ανεξάρτητο από θετικές ή αρνητικές τιμές του σφάλματος, δηλαδή είναι ανεξάρτητο από το αν οι τιμές των προβλέψεων είναι μικρότερες ή μεγαλύτερες των πραγματικών τιμών. Και τέλος, το MAD βασίζεται στην υπόθεση ότι η αξιοπιστία του σφάλματος ή το κόστος που δημιουργείται από το σφάλμα της πρόβλεψης, σχετίζεται γραμμικά με το μέγεθος του σφάλματος.

- **Μέσο σφάλμα τετραγώνου MSE (Mean Squared Error)**

Το μέσο σφάλμα τετραγώνου ορίζεται ως το άθροισμα των τετραγώνων των σφαλμάτων διαιρούμενο με τον αριθμό των χρονικών περιόδων  $n$ , στις οποίες έγιναν οι προβλέψεις, δηλαδή:

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2 = \frac{1}{n} \sum_{t=1}^n e_t^2$$

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 υψωμένη στο τετράγωνο.

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

- **Ρίζα μέσου σφάλματος τετραγώνου RMSE (Root Mean Squared Error)**

Η τετραγωνική ρίζα μέσου σφάλματος τετραγώνου είναι η θετική τιμή της τετραγωνικής του ρίζας, δηλαδή είναι:

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

Το RMSE εκφράζεται στην ίδια μονάδα μέτρησης με εκείνη των τιμών της χρονοσειράς.

- **Μέσο απόλυτο ποσοστιαίο σφάλμα MAPE (Mean Absolute Percentage Error)**

Το μέσο απόλυτο ποσοστιαίο σφάλμα εξετάζει τη συμπεριφορά της απόλυτης τιμής του σφάλματος της πρόβλεψης σε σχέση με την πραγματική τιμή της χρονοσειράς. 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, δηλαδή:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Y_t - \hat{Y}_t|}{Y_t} = \frac{1}{n} \sum_{t=1}^n \frac{|e_t|}{Y_t}$$

Το κριτήριο αυτό είναι απαλλαγμένο από μονάδες μέτρησης και το χρησιμοποιούμε για να συγκρίνουμε την ακρίβεια μιας ή περισσότερων μεθόδων προβλέψεων και για περισσότερες από μια χρονοσειρές.

- **Μέσο ποσοστιαίο σφάλμα MPE (Mean Percentage Error)**

Το μέσο ποσοστιαίο σφάλμα το χρησιμοποιούμε όταν ενδιαφερόμαστε να προσδιορίσουμε αν η μέθοδος πρόβλεψης είναι μεροληπτική, δηλαδή αν οι προβλεπόμενες τιμές είναι συστηματικά μεγαλύτερες ή μικρότερες από τις αντίστοιχες πραγματικές.

$$MPE = \frac{1}{n} \sum_{t=1}^n \frac{Y_t - \hat{Y}_t}{Y_t} = \frac{1}{n} \sum_{t=1}^n \frac{e_t}{Y_t}$$

Χωρίς αμφιβολία, όσο πιο κοντά στο μηδέν είναι η τιμή του MPE, τόσο πιο αμερόληπτη και καλή είναι η μέθοδος πρόβλεψης που χρησιμοποιήθηκε. Αντιθέτως, μεγάλες απόλυτες τιμές του MPE φανερώνουν μεγάλη μεροληψία της μεθόδου.

## 2.2 ΜΕΘΟΔΟΣ ΕΞΟΜΑΛΥΝΣΗΣ

### 2.2.1 ΑΠΛΟΣ ΚΙΝΗΤΟΣ ΜΕΣΟΣ

### 2.2.2 ΑΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ

### 2.2.3 ΔΙΠΛΟΣ ΚΙΝΗΤΟΣ ΜΕΣΟΣ

### 2.2.4 ΜΕΘΟΔΟΣ BROWN

### 2.2.5 ΜΕΘΟΔΟΣ HOLT

### 2.2.6 ΜΕΘΟΔΟΣ WINTERS

## 2.3 ΜΕΘΟΔΟΣ ΔΙΑΣΠΑΣΗΣ ΧΡΟΝΟΣΕΙΡΩΝ

### 2.3.1 ΑΝΑΛΥΣΗ ΠΟΧΙΚΟΤΗΤΑΣ

### 2.3.2 ΑΝΑΛΥΣΗ ΜΑΚΡΟΧΡΟΝΙΑΣ ΤΑΣΗΣ

### 2.3.3 ΑΝΑΛΥΣΗ ΚΥΚΛΙΚΟΤΗΤΑΣ ΚΑΙ ΜΗ ΚΑΝΟΝΙΚΟΤΗΤΑΣ





**Μέρος II**

**ΕΦΑΡΜΟΓΗ ΤΩΝ  
ΜΕΘΟΔΩΝ ΓΙΑ ΤΗΝ  
ΠΡΟΒΛΕΨΗ ΠΑΡΑΓΩΓΗΣ  
ΠΕΤΡΕΛΑΙΟΥ**



## Κεφάλαιο 3

### Όνομα τρίτου κεφαλαίου

3.1 Όνομα πρώτης ενότητας

3.1.1 Όνομα πρώτης υποενότητας

3.2 Όνομα δεύτερης ενότητας



# Βιβλιογραφία

- [1] C. Li. Chen and A.C. Reynolds. Robust constrained optimization of short- and long-term net present value for closed-loop reservoir management. *SPE Journal*, 17(3):849–864, 2012.
- [2] Jan Dirk Jansen. *A Systems Description of Flow Through Porous Media*. SpringerBriefs in Earth Sciences, 2013.
- [3] Drosos Kourounis, Louis J. Durlofsky, Jan Dirk Jansen, and Khalid Aziz. Adjoint formulation and constraint handling for gradient-based optimization of compositional reservoir flow. *Computational Geosciences*, pages 1–21, 2014.