**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

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 I

# 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 $t \geq 0$.

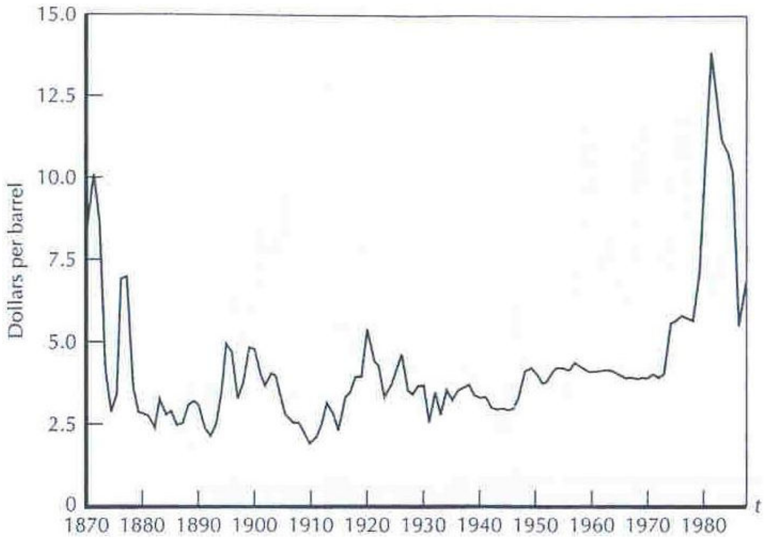
(iv) Η ημερήσια κατανάλωση ηλεκτρικού ρεύματος καθώς και η ημερήσια κατανάλωση ύδατος, $X_t$ και $Y_t$ αντίστοιχα, σε μια μεγάλη γεωγραφική περιοχή της χώρας με $t = 1, 2, \dots$

(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 τις 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* τυχαίες 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 ορισθεί 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 ορισθεί 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 $\sigma_{xy} = \mathrm{Cov}(X, Y)$. 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**:

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

(iii) **Η αυτοσυνδιακύμανση**:

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

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

## 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 $\{X_t : 0\leqslant t\leqslant s\}$, 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 $\{X_t : 0\leqslant t\leqslant s\}$ 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 $\forall n\in N,\ t_i\in T$ με $i=1,\ldots,n$ και $h\in T$ ισχύει η παρακάτω σχέση ισοδυναμίας

$$\left(X_{t_1},\ldots,X_{t_n}\right)\sim\left(X_{t_1+h},\ldots,X_{t_n+h}\right)$$

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: $\mu=E\left[X_t\right],\quad t\in T$
- $\gamma(h)=\mathrm{Cov}(X_t,X_{t+h}),\quad t,h\in T$

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

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

# Κεφάλαιο 2

# ΜΕΘΟΔΟΙ ΠΡΟΒΛΕΨΗΣ

## Μέρος ΙΙ

# ΕΦΑΡΜΟΓΗ 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.