

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ  
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΕΠΙΣΤΗΜΗΣ



Μέτρα Εξάρτησης και Υποφαινομενικές Συμπεριφορές για  
Γραμμικές Σχέσεις που Δημιουργούνται από Μη-  
Γραμμικής Μορφής Χρονοσειρές

Εμμανουήλ Δεληγιαννάκης

Διδακτορική Διατριβή

Υποβληθείσα στο Τμήμα Οικονομικής Επιστήμης του Πανεπιστημίου Πειραιώς

Πειραιάς, Μάιος 2024

*Αυτή η σελίδα αφήνεται σκόπιμα κενή.*

UNIVERSITY OF PIRAEUS  
DEPARTMENT OF ECONOMICS



Measures of Dependence and Spurious Behaviors for  
Linear Relationships Generated from Non-Linearly  
Structured Time Series

Emmanouil Deligiannakis

PhD Thesis

Submitted to the Department of Economics of the University of Piraeus

Piraeus, May 2024

*Αυτή η σελίδα αφήνεται σκόπιμα κενή.*

## Ευχαριστίες

Η παρούσα διδακτορική διατριβή αποτελεί το αποκορύφωμα μιας μακράς και απαιτητικής ακαδημαϊκής πορείας στον τομέα της Οικονομετρίας, η οποία δεν θα ήταν δυνατή χωρίς την καθοδήγηση, την υποστήριξη και την πίστη πολλών ανθρώπων.

Κατ' αρχάς, θα ήθελα να εκφράσω τη βαθύτατη ευγνωμοσύνη μου προς τον επιβλέποντα καθηγητή μου, κ. Χρήστο Αγιακλόγλου, για την αδιάκοπη υποστήριξη, την εμπιστοσύνη και την πολύτιμη καθοδήγησή του. Το επιστημονικό έργο και η ευρύτητα των ερευνητικών ενδιαφερόντων του αποτελούν πηγή έμπνευσης για εμένα, καθώς η αφοσίωση και η εμπειρία του στον τομέα της Οικονομετρίας ήταν οδηγός για την επιστημονική μου ανάπτυξη και έρευνα.

Ένα ιδιαίτερο ευχαριστώ οφείλω επίσης στα μέλη της τριμελούς επιτροπής της διδακτορικής διατριβής μου, τον καθηγητή κ. Σωτήριο Καρκαλάκο και την καθηγήτρια κ. Μαρία Ψυλλάκη για τις πολύτιμες παρατηρήσεις και την κριτική τους που βελτίωσαν σημαντικά το επίπεδο της εργασίας μου. Επίσης, θέλω να ευχαριστήσω τον καθηγητή κ. Anil Bera για την υποστήριξη και τη συνεργασία. Σημαντικό μέρος αυτής της ερευνητικής προσπάθειας βασίζεται στο σπουδαίο ερευνητικό του έργο.

Στη συνέχεια θέλω να εκφράσω την ευγνωμοσύνη στην οικογένειά μου, στη σύζυγό μου Ξένια και στα παιδιά μου, Δημήτρη και Εύα, καθώς και στους γονείς μου, Δημήτρη και Γεωργία, για την αδιάλειπτη αγάπη, υπομονή και υποστήριξή τους. Η σταθερότητα και η πίστη τους στις δυνατότητές μου ήταν ο πυλώνας πάνω στον οποίο στήριξα τις προσπάθειές μου.

Τέλος, θέλω να ευχαριστήσω όλους όσους με υποστήριξαν με οποιονδήποτε τρόπο κατά τη διάρκεια αυτής της πορείας. Η κάθε συμβουλή, η κάθε πράξη ενθάρρυνσης, αποτελούν μέρος αυτού που κατάφερα να επιτύχω.

*Εμμανουήλ Δεληγιαννάκης*

*Αυτή η σελίδα αφήνεται σκόπιμα κενή.*

# Μέτρα Εξάρτησης και Υποφαινομενικές Συμπεριφορές για Γραμμικές Σχέσεις που Δημιουργούνται από Μη-Γραμμικής Μορφής Χρονοσειρές

**Σημαντικοί Όροι:** χρονοσειρές, συντελεστής συσχέτισης, μέτρα εξάρτησης, *Copulas*,  $AR(1)$  –  $ARCH(1)$ , υποφαινομενικές σχέσεις

## Περίληψη

Ο προσδιορισμός της σχέσης μεταξύ δύο μεταβλητών είναι μία από τις πιο σημαντικές πτυχές κάθε ποσοτικής ανάλυσης. Στόχος αυτής της προσπάθειας είναι να διερευνηθεί σε αρχικό στάδιο εάν υπάρχει και σε ποιο βαθμό εξάρτηση μεταξύ των μεταβλητών, έτσι ώστε η ανάλυση να επεκταθεί στον ποσοτικό καθορισμό της σχέσης τους. Η τεχνική του συντελεστή συσχέτισης που εφαρμόζεται κατά κανόνα στην ποσοτική ανάλυση έχει αποδειχθεί ότι αρκετά συχνά δεν είναι ικανή να εντοπίσει ικανοποιητικά την ύπαρξη της σχέσης των μεταβλητών, δεδομένου ότι η τιμή του συντελεστή συσχέτισης εξαρτάται από τη φύση των δεδομένων, εμφανίζοντας μάλιστα ακόμα και εικονικές σχέσεις για την περίπτωση ανεξάρτητων χρονοσειρών. Για το λόγο αυτό, χρησιμοποιούνται για τη μέτρηση της εξάρτησης και άλλες πιο πολύπλοκες τεχνικές, όπως οι *Copula*, προκειμένου να συλλάβουν καλύτερα την έννοια της εξάρτησης μεταξύ των μεταβλητών ακόμα και για δεδομένου που παρουσιάζουν μεγάλη μεταβλητότητα και ποικιλομορφία.

Σε αυτή τη διδακτορική διατριβή διερευνάται η συμπεριφορά διαφόρων μέτρων εξάρτησης, συμπεριλαμβανομένου και του συντελεστή συσχέτισης, για γραμμικές σχέσεις που παράγονται από μη γραμμικές χρονοσειρές που έχουν έντονη μεταβλητότητα. Πιο

συγκεκριμένα, για την εφαρμογή των προσομοιώσεων χρησιμοποιούνται τα υποδείγματα  $AR(1) - ARCH(1)$  των Bera, Higgins και Lee (1992 και 1996), τα οποία παρουσιάζουν μεγάλη μεταβλητότητα στις τιμές τους, δεδομένου ότι η ετεροσκεδαστικότητα προέρχεται από τις τιμές της μεταβλητής και όχι από τις τιμές του τυχαίου σφάλματος. Μάλιστα, όπως αναφέρουν οι συγγραφείς τα υποδείγματα αυτά μπορούν να παρουσιαστούν και ως μη γραμμικά υποδείγματα με μεταβαλλόμενους συντελεστές.

Στο πρώτο κεφάλαιο αναφέρονται συνοπτικά τα βασικά στοιχεία που χρησιμοποιούνται στην ανάλυση τυχαίων μεταβλητών και περιγράφουν τον τρόπο συμπεριφοράς τους. Στη συνέχεια, αναλύεται η έννοια της από κοινού κατανομής δύο τυχαίων μεταβλητών με σκοπό την παρουσίαση του συντελεστή γραμμικής συσχέτισης, ενώ γίνεται ακόμα και μία σύντομη αναφορά στο πρόβλημα των υποφαινομενικών συσχετίσεων που υπάρχει από τη χρήση στάσιμων και μη στάσιμων χρονοσειρών.

Στο δεύτερο κεφάλαιο παρουσιάζονται οι βασικές έννοιες των συναρτήσεων *Copula* μαζί με το θεώρημα του Sklar, το οποίο θεμελιώνει την ύπαρξη αυτών των συναρτήσεων και ορίζει τα άνω και κάτω όρια τους. Κατόπιν, περιγράφονται οι πιο σημαντικές συναρτήσεις *Copula*, οι οποίες χρησιμοποιούνται ευρέως και οι οποίες θα εφαρμοστούν σε αυτή τη διδακτορική διατριβή για τη διερεύνηση της συμπεριφοράς τους σε περίπλοκες χρονοσειρές, ενώ στο τέλος γίνεται μία σύντομη αναφορά στον τρόπο εκτίμησης των παραμέτρων τους.

Στο τρίτο κεφάλαιο αναλύονται τα χαρακτηριστικά και οι ιδιότητες των στάσιμων χρονοσειρών και παρουσιάζονται τα βασικά υποδείγματα της μεθοδολογίας των Box και Jenkins. Κατόπιν, γίνεται μία σύντομη αναφορά στις μη στάσιμες χρονοσειρές και περιγράφονται οι στατιστικοί έλεγχοι που εφαρμόζονται για τη διαπίστωση της ύπαρξης μοναδιαίας αυτοσυσχετιζόμενης ρίζας, σημειώνοντας ότι οι χρονοσειρές αυτές αποτυπώνουν συμπεριφορές φαινομένων με έντονη μεταβλητότητα. Στα πλαίσια αυτής της συμπεριφοράς παρουσιάστηκαν και τα υποδείγματα υπό συνθήκης



ετεροσκεδαστικότητας τα οποία και αυτά απεικονίζουν τη μεταβλητότητα αλλά με ένα διαφορετικό τρόπο, όπου η διακύμανση των τιμών της χρονοσειράς μεταβάλλεται διαχρονικά με συγκεκριμένο υπόδειγμα. Τέλος, αναφορά γίνεται και στα υποδείγματα υπό συνθήκης ετεροσκεδαστικότητας στις τιμές της μεταβλητής που παρουσιάστηκαν από τους Bera, Higgins και Lee (1992 και 1996) ως μία άλλη μορφή υποδειγμάτων που μπορεί να περιλάβει ακόμα μεγαλύτερη μεταβλητότητα στις τιμές μιας χρονοσειράς.

Στο τέταρτο κεφάλαιο διερευνάται η συμπεριφορά του συντελεστή συσχέτισης, καθώς και ορισμένων τεχνικών *Copula* για ανεξάρτητες, για γραμμικά εξαρτημένες χρονοσειρές με ετεροσκεδαστικότητα, όπου η συμπεριφορά ορισμένων από των οποίων ομοιάζει και με μη γραμμική, χρησιμοποιώντας ανάλυση Monte Carlo. Επιπρόσθετα, διερευνώνται και οι υποφαινομενικές συσχετίσεις που εμφανίζονται από τέτοια μορφής υποδείγματα, χρησιμοποιώντας το στατιστικό έλεγχο για τη μηδενική υπόθεση ότι δεν υπάρχει γραμμική συσχέτιση με βάση το συντελεστή συσχέτισης. Τέλος, εξετάζονται τόσο οι συμπεριφορές των μέτρων εξάρτησης όσο και της δύναμης τους στατιστικού ελέγχου της μηδενικής συσχέτισης για γραμμικές εξαρτήσεις δύο μεταβλητών που κατασκευάζονται.

Στο πέμπτο κεφάλαιο διερευνάται μέσω Monte Carlo ανάλυσης η συμπεριφορά του φαινομένου των υποφαινομενικών παλινδρομήσεων για δύο ανεξάρτητες στάσιμες αυτοπαλίνδρομου πρώτου βαθμού AR(1) χρονοσειρές αλλά και δύο ανεξάρτητες μη στάσιμες τυχαίου περπατήματος χρονοσειρές, των οποίων τα σφάλματα δεν έχουν σταθερή διακύμανση, αλλά η διαχρονικά μεταβαλλόμενη διακύμανσή τους κατασκευάζεται από ένα αυτοπαλίνδρομο υπό όρους ετεροσκεδαστικότητας υπόδειγμα πρώτου βαθμού, δηλαδή από ένα ARCH(1) υπόδειγμα, όπου η μεταβλητότητα στη διακύμανση προέρχεται είτε από τις τιμές των σφαλμάτων είτε από τις τιμές της ίδιας της χρονοσειράς, με την τελευταία μορφή να σχηματίζει μια οικογένεια χρονοσειρών που μπορούν να θεωρηθούν και ως μη γραμμικές διεργασίες. Επίσης, διερευνάται η ύπαρξη αυτοσυσχέτισης στις τιμές των σφαλμάτων της παλινδρόμησης, καθώς και ο τρόπος αντιμετώπισης και των δύο προαναφερθέντων προβλημάτων.

Στο έκτο και τελευταίο κεφάλαιο εξετάζονται οι σχέσεις εξάρτησης μεταξύ διαφόρων μεταβλητών που συνδέονται με την αγορά ενέργειας και ειδικότερα με την τιμή του πετρελαίου για συγκεκριμένη χρονική περίοδο χρησιμοποιώντας όλες τις τεχνικές *Copula* που εφαρμόστηκαν στην διατριβή αυτή. Αρχικά εξετάζονται οι σχέσεις των μεταβλητών στις τιμές τους και κατόπιν στις τιμές των πρώτες διαφορών τους, δεδομένου ότι οι χρονοσειρές βρέθηκαν να είναι μη στάσιμες πρώτου βαθμού. Τέλος, η ανάλυση της εξάρτησής τους επεκτάθηκε στις εκτιμηθείσες τιμές αλλά και στις τιμές των καταλοίπων που προέκυψαν από την εκτίμηση του καταλληλότερου ARIMA υποδείγματος, διαπιστώνοντας σημαντικές αλλαγές στις σχέσεις τους.

# Measures of Dependence and Spurious Behaviors for Linear Relationships Generated from Non-Linearly Structured Time Series

**Keywords:** time series, correlation coefficient, measures of dependence, *Copulas*, AR(1) – ARCH(1), spurious relationships

## Abstract

Determining the relationship between two variables is one of the most crucial aspects of any quantitative analysis. The goal of this effort is to explore if there is, and to what extent, a dependency between the two variables, so that the analysis can expand to quantitatively determine their relationship. The technique of correlation coefficient commonly used in quantitative analysis has often been proven to be insufficient in adequately identifying the existence of variable relationships, given that the value of the correlation coefficient depends on the nature of the data, showing sometimes spurious relationships in the case of independent time series. For this reason, more complex techniques, such as Copulas, are used to measure dependency in a better way, capturing the concept of dependency between variables even for data with high variability and diversity.

This doctoral thesis investigates the behavior of various dependency measures, including the correlation coefficient, for linear relationships generated by non-linear time series with significant variability. Specifically, the AR(1) - ARCH(1) models introduced by Bera, Higgins, and Lee (1992 and 1996) are used for the simulation application, which exhibit high variability in their values, as heteroskedasticity stems from the variable values rather than from the random error values. Moreover, as the authors' state, these models can also be presented as non-linear models with time varying coefficients.

The first chapter briefly outlines the basic elements used in the analysis of random variables and describes their behavior. The concept of joint distribution of two random variables is then analyzed with the aim of presenting the linear correlation coefficient, while there is even a brief mention of the problem of spurious correlations that arises from the use of stationary and non-stationary time series.

In the second chapter, the basic concepts of Copula functions are presented along with Sklar's theorem, which establishes the existence of these functions and defines their upper and lower bounds. Subsequently, the most important and widely used Copula functions are described which they will be applied in this doctoral thesis to explore their behavior in complex time series, with a brief reference at the end on how to estimate their parameters.

In the third chapter, the characteristics and properties of stationary time series are introduced along with the basic concepts of the Box and Jenkins methodology. Then, non-stationary time series are described along with their corresponding statistical tests used to detect the presence of unit autoregressive root, noting that these time series reflect behaviors of phenomena with significant variability. Within this context, the conditional heteroskedasticity models are presented, which depict variability in a different way, where the fluctuation of time series values changes over time with a specific pattern. Finally, the Bera, Higgins, and Lee (1992 and 1996) models are presented as another form of models that may include even greater variability in time series values.

In the fourth chapter, the behavior of the correlation coefficient and various Copula techniques for independent, linearly dependent time series with heteroskedasticity is explored, where the behavior of some of them resembles non-linear behavior, using Monte Carlo analysis. Additionally, the correlation behavior that arise from such models is investigated based on the statistical tests for the null hypothesis that there is no linear

correlation along with the power of the test to understand better the spurious correlation phenomenon as well as the behaviors of the dependency measures used in this research.

In the fifth chapter, a Monte Carlo analysis is applied to explore the phenomenon of spurious regression for two independent stationary first-order autoregressive AR(1) time series, as well as two independent non-stationary random walk time series, with errors that do not have constant variance. Indeed, the errors were constructed by a time-varying variance following an autoregressive conditional heteroskedasticity first-order model, namely an ARCH(1) model, where the variability of the variance arises either from error values or from the time series values themselves, with the latter forming a family of time series that can also be considered as non-linear processes. The error structures of this regression analysis is also investigated to detect serially correlated errors along with an effort to show how to address both issues.

In the sixth and final chapter, the dependency relationships between various variables related to the energy market is examined using all the Copula techniques applied in this thesis. Specifically, the effort is consecrated on the price of oil for several other macro and finance variables for a specific period. Initially, the relationships of all variables is examined in their levels and then in their first differences, given that these time series were found to be first-order non-stationary processes. Finally, the analysis of their dependency is extended to the fitted values as well as the residuals resulting from the estimation of the most suitable ARIMA model, noting significant changes in their relationships.

*Αυτή η σελίδα αφήνεται σκόπιμα κενή.*

# Περιεχόμενα

## ΚΕΦΑΛΑΙΟ 1      ΑΝΑΛΥΣΗ ΔΙΜΕΤΑΒΛΗΤΩΝ ΣΧΕΣΕΩΝ

1.1	ΕΙΣΑΓΩΓΗ	1
1.2	ΤΥΧΑΙΑ ΜΕΤΑΒΛΗΤΗ	2
1.3	ΡΟΠΕΣ ΚΑΤΑΝΟΜΗΣ	6
1.4	ΑΠΟ-ΚΟΙΝΟΥ ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ	8
1.5	ΣΥΣΧΕΤΙΣΗ	12
1.6	ΥΠΟΦΑΙΝΟΜΕΝΙΚΗ ΣΥΣΧΕΤΙΣΗ	16
1.7	ΑΝΑΚΕΦΑΛΑΙΩΣΗ	19

## ΚΕΦΑΛΑΙΟ 2      ΣΥΝΑΡΤΗΣΕΙΣ COPULA

2.1	ΕΙΣΑΓΩΓΗ	21
2.2	ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ COPULA	23
2.3	ΟΙΚΟΓΕΝΕΙΕΣ COPULA	28
2.4	ΕΚΤΙΜΗΣΗ ΠΑΡΑΜΕΤΡΩΝ	38
2.5	ΑΝΑΚΕΦΑΛΑΙΩΣΗ	41

## ΚΕΦΑΛΑΙΟ 3      ΓΡΑΜΜΙΚΑ ΚΑΙ ΜΗ ΓΡΑΜΜΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ ΧΡΟΝΟΣΕΙΡΩΝ

3.1	ΕΙΣΑΓΩΓΗ	43
3.2	ΣΤΑΣΙΜΕΣ ΧΡΟΝΟΣΕΙΡΕΣ	45
3.3	ΥΠΟΔΕΙΓΜΑΤΑ ΚΙΝΗΤΟΥ ΜΕΣΟΥ	52
3.4	ΑΥΤΟΠΑΛΙΝΔΡΟΜΑ ΥΠΟΔΕΙΓΜΑΤΑ	56
3.5	ΜΕΙΚΤΑ ΥΠΟΔΕΙΓΜΑΤΑ	63
3.6	ΜΗ ΣΤΑΣΙΜΕΣ ΧΡΟΝΟΣΕΙΡΕΣ	66

3.7	ΑΥΤΟΠΑΛΙΝΔΡΟΜΑ ΕΤΕΡΟΣΚΕΔΑΣΤΙΚΟΤΗΤΑΣ	ΥΠΟΔΕΙΓΜΑΤΑ	ΥΠΟ	ΣΥΝΘΗΚΗΣ	73
3.8	ΑΥΤΟΠΑΛΙΝΔΡΟΜΑ ΕΤΕΡΟΣΚΕΔΑΣΤΙΚΟΤΗΤΑΣ	ΥΠΟΔΕΙΓΜΑΤΑ	ΥΠΟ	ΣΥΝΘΗΚΗΣ ΥΠΟ ΟΡΟΥΣ ΤΥΧΑΙΟΥ ΣΥΝΤΕΛΕΣΤΗ ΚΑΙ ΕΠΑΥΞΗΜΕΝΗΣ ΕΤΕΡΟΣΚΕΔΑΣΤΙΚΟΤΗΤΑΣ	79
3.9	ΑΝΑΚΕΦΑΛΑΙΩΣΗ				83

#### **ΚΕΦΑΛΑΙΟ 4 ΔΙΜΕΤΑΒΛΗΤΕΣ ΓΡΑΜΜΙΚΕΣ ΣΧΕΣΕΙΣ ΓΙΑ ΜΗ ΓΡΑΜΜΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ**

4.1	ΕΙΣΑΓΩΓΗ	85
4.2	ΟΙ ΜΕΤΑΒΛΗΤΕΣ ΤΗΣ ΠΡΟΣΟΜΕΙΩΣΗΣ	88
4.3	ΟΙ ΤΕΧΝΙΚΕΣ ΤΗΣ ΠΡΟΣΟΜΕΙΩΣΗΣ	95
4.4	ΑΝΕΞΑΡΤΗΣΙΑ	98
4.5	ΥΠΟΦΑΙΝΟΜΕΝΙΚΕΣ ΣΥΣΧΕΤΙΣΕΙΣ	102
4.6	ΓΡΑΜΜΙΚΗ ΕΞΑΡΤΗΣΗ	112
4.7	ΔΥΝΑΜΗ ΕΛΕΓΧΟΥ	132
4.8	ΑΝΑΚΕΦΑΛΑΙΩΣΗ	137

#### **ΚΕΦΑΛΑΙΟ 5 ΥΠΟΦΑΙΝΟΜΕΝΙΚΕΣ ΠΑΛΙΝΔΡΟΜΗΣΕΙΣ ΓΙΑ ΥΠΟΔΕΙΓΜΑΤΑ ΜΕ ΥΠΟ ΣΥΝΘΗΚΗ ΕΤΕΡΟΣΚΕΔΑΣΤΗΚΟΤΗΤΑ ΠΡΩΤΟΥ ΒΑΘΜΟΥ**

5.1	ΕΙΣΑΓΩΓΗ	141
5.2	ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ ΥΠΟΔΕΙΓΜΑΤΩΝ ΤΗΣ ΠΡΟΣΟΜΕΙΩΣΗΣ	143
5.3	ΠΑΡΟΥΣΙΑΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΠΡΟΣΟΜΕΙΩΣΗΣ	147
5.4	ΑΥΤΟΣΥΣΧΕΤΙΣΗ	151
5.5	ΑΝΑΚΕΦΑΛΑΙΩΣΗ	157



**ΚΕΦΑΛΑΙΟ 6            ΕΜΠΕΙΡΙΚΗ ΕΦΑΡΜΟΓΗ ΣΕ ΜΕΤΡΑ ΕΞΑΡΤΗΣΗΣ ΓΙΑ  
ΜΕΤΑΒΛΗΤΕΣ ΠΟΥ ΣΥΝΔΕΟΝΤΑΙ ΜΕ ΤΗΝ ΑΓΟΡΑ  
ΕΝΕΡΓΕΙΑΣ**

6.1	ΕΙΣΑΓΩΓΗ	159
6.2	ΠΑΡΟΥΣΙΑΣΗ ΜΕΤΑΒΛΗΤΩΝ	162
6.3	ΜΕΤΡΗΣΗ ΕΞΑΡΤΗΣΗΣ	170
6.4	ΑΝΑΚΕΦΑΛΑΙΩΣΗ	185

<b>ΠΑΡΑΡΤΗΜΑ</b>	<b>189</b>
------------------	------------

<b>ΒΙΒΛΙΟΓΡΑΦΙΑ</b>	<b>195</b>
---------------------	------------

*Αυτή η σελίδα αφήνεται σκόπιμα κενή.*

## Πίνακες

Πίνακας 4.1 Τιμές μέσων όρων για μέτρα εξάρτησης από δύο ανεξάρτητες στάσιμες χρονοσειρές AR(1) – ARCH(1) στα σφάλματα βασισμένες σε 10000 επαναλήψεις 100

Πίνακας 4.2 Τιμές μέσων όρων για μέτρα εξάρτησης από δύο ανεξάρτητες στάσιμες χρονοσειρές AR(1) – ARCH(1) στη μεταβλητή βασισμένες σε 10000 επαναλήψεις 101

Πίνακας 4.3 Ποσοστά απόρριψης της υπόθεσης της μηδενικής συσχέτισης για 10% και 5% ονομαστικό επίπεδο σημαντικότητας για δύο ανεξάρτητες στάσιμες χρονοσειρές AR(1) – ARCH(1) στα σφάλματα μαζί με τις τιμές του μέσου όρου και της τυπικής απόκλισης της στατιστικής  $t$  που χρησιμοποιήθηκε βασισμένες σε 10000 επαναλήψεις 109

Πίνακας 4.4 Ποσοστά απόρριψης της υπόθεσης της μηδενικής συσχέτισης για 10% και 5% ονομαστικό επίπεδο σημαντικότητας για δύο ανεξάρτητες στάσιμες χρονοσειρές AR(1) – ARCH(1) στη μεταβλητή μαζί με τις τιμές του μέσου όρου και της τυπικής απόκλισης της στατιστικής  $t$  που χρησιμοποιήθηκε βασισμένες σε 10000 επαναλήψεις 110

Πίνακας 4.5 Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές για  $\rho = 0.2$  βασισμένες σε 10000 επαναλήψεις 116

Πίνακας 4.6 Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές για  $\rho = 0.5$  βασισμένες σε 10000 επαναλήψεις 117

Πίνακας 4.7 Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές για  $\rho = 0.8$  βασισμένες σε 10000 επαναλήψεις 118

Πίνακας 4.8 Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές για  $\rho = 0.9$  βασισμένες σε 10000 επαναλήψεις 119

Πίνακας 4.9 Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για  $\rho = 0.2$  βασισμένες σε 10000 επαναλήψεις 123

Πίνακας 4.10 Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για  $\rho = 0.5$  βασισμένες σε 10000 επαναλήψεις

124

Πίνακας 4.11 Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για  $\rho = 0.8$  βασισμένες σε 10000 επαναλήψεις

125

Πίνακας 4.12 Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για  $\rho = 0.9$  βασισμένες σε 10000 επαναλήψεις

126

Πίνακας 4.13 Ποσοστά απόρριψης της μηδενικής υπόθεσης για μηδενική συσχέτιση σε ονομαστικό επίπεδο 5% ( $|t| > 1.96$ ) για δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές για  $\rho = 0.2, 0.5, 0.8$  and  $0.9$  βασισμένες σε 10000 επαναλήψεις

135

Πίνακας 4.14 Ποσοστά απόρριψης της μηδενικής υπόθεσης για μηδενική συσχέτιση σε ονομαστικό επίπεδο 5% ( $|t| > 1.96$ ) για δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για  $\rho = 0.2, 0.5, 0.8$  and  $0.9$  βασισμένες σε 10000 επαναλήψεις

136

Πίνακας 5.1 Ποσοστά απόρριψης της μηδενικής υπόθεσης ότι  $\beta = 0$  σε επίπεδο σημαντικότητας 5% για δύο ανεξάρτητες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές βασισμένα σε 10000 επαναλήψεις

148

Πίνακας 5.2 Ποσοστά απόρριψης της μηδενικής υπόθεσης ότι  $\beta = 0$  σε επίπεδο σημαντικότητας 5% για δύο ανεξάρτητες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές βασισμένα σε 10000 επαναλήψεις

149

Πίνακας 5.3 Τιμές μέσων όρων της στατιστικής των Durbin-Watson για δύο ανεξάρτητες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές βασισμένες σε 10000 επαναλήψεις

152

Πίνακας 5.4 Τιμές μέσων όρων της στατιστικής των Durbin-Watson για δύο ανεξάρτητες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές βασισμένες σε 10000 επαναλήψεις

153

Πίνακας 5.5 Ποσοστά απόρριψης της μηδενικής υπόθεσης ότι  $\beta = 0$  σε επίπεδο σημαντικότητας 5% για δύο ανεξάρτητες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές βασισμένα σε 10000 επαναλήψεις χρησιμοποιώντας τη μέθοδο των Cochran-Orcutt

155

Πίνακας 5.6 Ποσοστά απόρριψης της μηδενικής υπόθεσης ότι $\beta = 0$ σε επίπεδο σημαντικότητας 5% για δύο ανεξάρτητες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές βασισμένα σε 10000 επαναλήψεις χρησιμοποιώντας τη μέθοδο των Cochrane-Orcutt	156
Πίνακας 6.1 Περιγραφικά στατιστικά των μεταβλητών	169
Πίνακας 6.2 Τιμές μέτρων εξάρτησης σε τιμές των μεταβλητών σε σχέση με το Πετρέλαιο	171
Πίνακας 6.3 Αποτελέσματα εφαρμογής του ADF test με τάση	174
Πίνακας 6.4 Τιμές μέτρων εξάρτησης σε πρώτες διαφορές των τιμών των μεταβλητών σε σχέση με το Πετρέλαιο	174
Πίνακας 6.5 Παρουσίαση των πρώτων τριών υποδειγμάτων που επιλέχθηκαν για κάθε μεταβλητή με βάση τα πληροφοριακά κριτήρια AIC και BIC	178
Πίνακας 6.6 Εκτιμήσεις του αυτοπαλίνδρομου όρου του υποδείγματος ARIMA(1, 1, 0) για τις τιμές όλων των μεταβλητών	179
Πίνακας 6.7 Τιμές μέτρων εξάρτησης σε εκτιμηθείσες τιμές μεταβλητών από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα σε σχέση με το Πετρέλαιο	180
Πίνακας 6.8 Τιμές μέτρων εξάρτησης σε τιμές καταλοίπων των μεταβλητών από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα σε σχέση με το Πετρέλαιο	181
Πίνακας 6.9 Τιμές της στατιστικής LM για τον έλεγχο ετροσκεδαστικότητας ενός ARCH(1) υποδείγματος	184
Πίνακας A.1: Τιμές μέτρων εξάρτησης σε τιμές όλων των άλλων Μεταβλητών	190
Πίνακας A.2: Τιμές μέτρων εξάρτησης σε πρώτες διαφορές των τιμών όλων των άλλων μεταβλητών	191
Πίνακας A.3: Τιμές μέτρων εξάρτησης σε εκτιμηθείσες τιμές μεταβλητών από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα	192
Πίνακας A.4: Τιμές μέτρων εξάρτησης σε τιμές καταλοίπων των μεταβλητών από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα	193

*Αυτή η σελίδα αφήνεται σκόπιμα κενή.*

## Διαγράμματα

Διάγραμμα 1.1: Συναρτήσεις Κατανομής Πιθανοτήτων και Αθροιστική	4
Διάγραμμα 1.2: Συναρτήσεις Πυκνότητας Πιθανοτήτων και Αθροιστική	5
Διάγραμμα 3.1 Διαδικασία Λευκού Θορύβου	47
Διάγραμμα 3.2 Διαδικασία MA(1)	54
Διάγραμμα 3.3 Διαδικασία MA(2)	56
Διάγραμμα 3.4 Διαδικασία AR(1)	59
Διάγραμμα 3.5 Διαδικασία AR(2)	61
Διάγραμμα 3.6 Διαδικασία ARMA(1, 1)	65
Διάγραμμα 3.7 Διαδικασία Τυχαίου Περιπάτου	69
Διάγραμμα 3.8 Εξέλιξη Τιμών-Αποδόσεων Δείκτη S&P 500	74
Διάγραμμα 3.9 Εφαρμογή Υποδείγματος ARCH(1)	77
Διάγραμμα 3.10 Εφαρμογή Υποδείγματος GARCH(1, 1)	79
Διάγραμμα 4.1 Διαγραμματική απεικόνιση ενός AR(1) υποδείγματος για $\phi = 0.2$ και $0.8$	89
Διάγραμμα 4.2 Διαγραμματική απεικόνιση ενός AR(1) – ARCH(1) στα σφάλματα υποδείγματος για $\phi = 0.2$ και για $a_1 = 0.2$ και $0.8$	91
Διάγραμμα 4.3 Διαγραμματική απεικόνιση ενός AR(1) – ARCH(1) στα σφάλματα υποδείγματος για $\phi = 0.8$ και για $a_1 = 0.2$ και $0.8$	92
Διάγραμμα 4.4 Διαγραμματική απεικόνιση ενός AR(1) – ARCH(1) στη μεταβλητή υποδείγματος για $\phi = 0.2$ και για $a_1 = 0.2$ και $0.8$	93
Διάγραμμα 4.5 Διαγραμματική απεικόνιση ενός AR(1) – ARCH(1) στη μεταβλητή υποδείγματος για $\phi = 0.8$ και για $a_1 = 0.2$ και $0.8$	94
Διάγραμμα 4.6 Κατανομή συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές AR(1) – ARCH(1) στα σφάλματα για $\phi = 0.2$ και για $a_1 = 0.0, 0.2, 0.5$ και $0.8$ βασισμένες σε 10000 επαναλήψεις και για μεγέθη δείγματος 100	104
Διάγραμμα 4.7 Κατανομή συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές AR(1) – ARCH(1) στα σφάλματα για $\phi = 0.8$ και για $a_1 = 0.0, 0.2, 0.5$ και $0.8$ βασισμένες σε 10000 επαναλήψεις και για μεγέθη δείγματος 100	105

Διάγραμμα 4.8 Κατανομή συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές AR(1) – ARCH(1) στη μεταβλητή για $\phi = 0.2$ και για $a_1 = 0.0, 0.2, 0.5$ και $0.8$ βασισμένες σε 10000 επαναλήψεις και για μεγέθη δείγματος 100	105
Διάγραμμα 4.9 Κατανομή συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές AR(1) – ARCH(1) στη μεταβλητή για $\phi = 0.5$ και για $a_1 = 0.0, 0.2, 0.5$ και $0.8$ βασισμένες σε 10000 επαναλήψεις και για μεγέθη δείγματος 100	106
Διάγραμμα 4.10 Κατανομή συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές AR(1) – ARCH(1) στη μεταβλητή για $\phi = 0.8$ και για $a_1 = 0.0, 0.2, 0.5$ και $0.8$ βασισμένες σε 10000 επαναλήψεις και για μεγέθη δείγματος 100	106
Διάγραμμα 4.11 Διαγραμματική απεικόνιση δύο γραμμικά εξαρτημένων AR(1) – ARCH(1) στα σφάλματα χρονοσειρών για $\rho = 0.2$ , $\phi = 0.2$ και για $a_1 = 0.0, 0.2, 0.5$ και $0.8$ και για μεγέθη δείγματος 100 παρατηρήσεων	129
Διάγραμμα 4.12 Διαγραμματική απεικόνιση δύο γραμμικά εξαρτημένων AR(1) – ARCH(1) στα σφάλματα χρονοσειρών για $\rho = 0.8$ , $\phi = 0.8$ και για $a_1 = 0.0, 0.2, 0.5$ και $0.8$ και για μεγέθη δείγματος 100 παρατηρήσεων	129
Διάγραμμα 4.13 Διαγραμματική απεικόνιση δύο γραμμικά εξαρτημένων AR(1) – ARCH(1) στη μεταβλητή χρονοσειρών για $\rho = 0.2$ , $\phi = 0.2$ και για $a_1 = 0.0, 0.2, 0.5$ και $0.8$ και για μεγέθη δείγματος 100 παρατηρήσεων	130
Διάγραμμα 4.14 Διαγραμματική απεικόνιση δύο γραμμικά εξαρτημένων AR(1) – ARCH(1) στη μεταβλητή χρονοσειρών για $\rho = 0.8$ , $\phi = 0.8$ και για $a_1 = 0.0, 0.2, 0.5$ και $0.8$ και για μεγέθη δείγματος 100 παρατηρήσεων	130
Διάγραμμα 4.15 Διαγραμματική απεικόνιση δύο γραμμικά εξαρτημένων AR(1) – ARCH(1) στη μεταβλητή χρονοσειρών για $\rho = 0.9$ , $\phi = 0.9$ και για $a_1 = 0.9$ και για μεγέθη δείγματος 50, 100, 500 και 1000 παρατηρήσεων	132
Διάγραμμα 6.1 Διαχρονική Εξέλιξη Τιμής Νομισμάτων USD και YEN	163
Διάγραμμα 6.2 Διαχρονική Εξέλιξη Χρηματιστηριακών Δεικτών S&P 500, Nikkei 225 και VIX	165
Διάγραμμα 6.3 Διαχρονική Εξέλιξη Τιμής Ελαφρών Αγαθών Wheat, Rice και Cotton	166
Διάγραμμα 6.4 Διαχρονική Εξέλιξη Τιμής Ενεργειακού Δείκτη και Δείκτη Αγαθών	167
Διάγραμμα 6.5 Διαχρονική Εξέλιξη Τιμής Βαρέων Αγαθών Crude Oil, Natural Gas και Gold	168





# ΚΕΦΑΛΑΙΟ 1

## ΑΝΑΛΥΣΗ ΔΙΜΕΤΑΒΛΗΤΩΝ ΣΧΕΣΕΩΝ

### 1.1 Εισαγωγή

Η διαδικασία λήψεων αποφάσεων στηρίζεται σε δύο βασικές τεχνικές: α) στην ποιοτική ανάλυση (qualitative analysis) και β) στην ποσοτική ανάλυση (quantitative analysis), με την πρώτη τεχνική να αναφέρεται σε θέματα Διοίκησης (management) που δεν συνδέονται με εξέταση δεδομένων. Αντίθετα, η ποσοτική ανάλυση ασχολείται με τη μελέτη αριθμητικών δεδομένων που περιγράφουν ένα φαινόμενο, με σκοπό να αναγνωρίζει την εξέλιξη ή τη συμπεριφορά του προκειμένου να σχηματίσει προβλέψεις για την μελλοντική του πορεία, έτσι ώστε να παρθούν σωστές αποφάσεις.

Κάθε φαινόμενο προσδιορίζεται από μία τυχαία μεταβλητή (random variable), η οποία εκφράζει στατιστικά τον τρόπο συμπεριφοράς του. Χρησιμοποιούνται τυχαίες μεταβλητές για να περιγράψουν φαινόμενα διότι η συμπεριφορά τους δεν είναι προσδιοριστική (deterministic), και άρα προβλέψιμη με βεβαιότητα, αλλά στοχαστική (stochastic) και εξαρτάται από πολλούς άλλους μετρήσιμους ή μη παράγοντες. Για αυτό άλλωστε η συμπεριφορά των τυχαίων μεταβλητών καθορίζεται από κάποια συνάρτηση πιθανοτήτων που δηλώνει τη δυνατότητα εμφάνισης των τιμών που λαμβάνει η τυχαία μεταβλητή. Για το λόγο αυτό, η ποσοτική ανάλυση των φαινομένων ενέχει την έννοια της αβεβαιότητας για αυτό και η προσέγγισή τους είναι πάντα στατιστική, προσπαθώντας να περιορίσει την αβεβαιότητα για τη λήψη αποφάσεων.

Πολλές φορές όμως η συμπεριφορά διαφόρων φαινομένων συνδέεται άμεσα με μία ή περισσότερες άλλες μετρήσιμες μεταβλητές, όπου η ανάλυση που ακολουθείται στην περίπτωση αυτή για τη λήψη των αποφάσεων είναι διαφορετική. Μάλιστα στις περιπτώσεις όπου το φαινόμενο συνδέεται μόνο με μία άλλη μεταβλητή (bivariate case), τότε το ενδιαφέρον της ανάλυσης επικεντρώνεται στη σχέση αυτών των δύο μεταβλητών.



Η αλήθεια είναι ότι υπάρχουν διάφοροι τρόποι με τους οποίους μπορεί να εξετασθεί η σχέση ανάμεσα σε δύο τυχαίες μεταβλητές. Οι περισσότερες όμως προσεγγίσεις ξεκινάνε από την ανάλυση συσχέτισης (correlation analysis), η οποία μέσω του συντελεστή συσχέτισης (correlation coefficient) μελετά το βαθμό, την ύπαρξη αλλά και την κατεύθυνση της γραμμικής σχέσης των δύο μεταβλητών. Αναμφισβήτητα, η ανάλυση συσχέτισης δεν προσδιορίζει την ποσοτική σχέση των μεταβλητών, όπως η οικονομετρική ανάλυση (econometric analysis), αλλά δίνει την πρώτη ένδειξη στον αναλυτή για το εάν υπάρχει σχέση μεταξύ των μεταβλητών, έτσι ώστε να συνεχίσει τη διερεύνηση του φαινομένου.

Στο κεφάλαιο αυτό αρχικά αναλύεται η έννοια της τυχαίας μεταβλητής, καθώς και ο τρόπος συμπεριφοράς της όπως αυτός δίνεται από τις ροπές (moments) της. Στη συνέχεια, περιγράφεται η έννοια της από κοινού κατανομής δύο τυχαίων μεταβλητών, καθώς και η συνδιακύμανση (covariance). Κατόπιν, παρουσιάζονται ο συντελεστής γραμμικής συσχέτισης, καθώς και οι δύο συντελεστές συσχέτισης κατάταξης, ενώ τέλος γίνεται μία σύντομη αναφορά στο πρόβλημα των υποφαινομενικών συσχετίσεων.

## 1.2 Τυχαία Μεταβλητή

Η τυχαία μεταβλητή είναι μια πραγματική συνάρτηση που ορίζεται σε ένα δειγματικό χώρο. Για κάθε τυχαία μεταβλητή  $X$  και κάθε σύνολο  $C$  πραγματικών αριθμών, μπορεί να υπολογιστεί η πιθανότητα ότι η  $X$  παίρνει την τιμή της στο σύνολο  $C$ . Η συλλογή όλων αυτών των πιθανοτήτων είναι η κατανομή της τυχαίας μεταβλητή  $X$  και υπάρχουν δύο κατηγορίες τυχαίων μεταβλητών, οι διακριτές ή ασυνεχείς (discrete) και οι συνεχείς (continuous).

### *Διακριτή Τυχαία Μεταβλητή*

Η διακριτή ή ασυνεχής τυχαία μεταβλητή μπορεί να λάβει συγκεκριμένες ή ακέραιες τιμές από ένα πεπερασμένο πλήθος τιμών. Για παράδειγμα, σε κάποιο πείραμα ένα νόμισμα ρίχνεται 5 φορές. Ως εκ τούτου, ο δειγματικός χώρος  $\mathcal{S}$ , που είναι το σύνολο όλων των δυνατών αποτελεσμάτων του πειράματος, είναι οι  $2^5$  διαφορετικές



ακολουθίες από 5 γράμματα ή από κορώνες που είναι πιθανές. Εάν η πραγματική συνάρτηση  $X$  οριστεί στο δειγματικό χώρο  $\mathcal{S}$ , έτσι ώστε να μετράει τον αριθμό των κορωνών, τότε το αποτέλεσμα πέντε ρίψεων είναι η ακολουθία  $\{K, K, \Gamma, \Gamma, K\}$  και  $X(s) = 3$ . Για κάθε πιθανή ακολουθία που αποτελείται από 5 γράμματα ή κορώνες, η τιμή  $X(s)$  ισούται με τον αριθμό των κορωνών στην ακολουθία. Οι πιθανές τιμές για τη συνάρτηση  $X$  είναι  $0, 1, \dots, 5$ . Επιπλέον, εάν το νόμισμα είναι δίκαιο, τότε κάθε ακολουθία από τις 32 πιθανές έχει την ίδια πιθανότητα εμφάνισης και ίση με  $1/32$ .

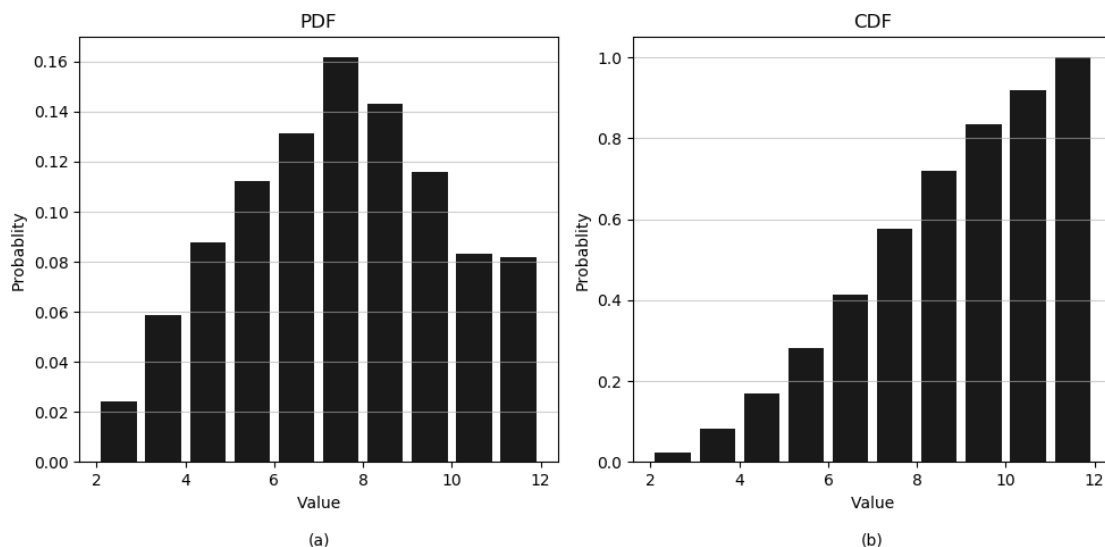
Οι τιμές των πιθανοτήτων που λαμβάνει μία τυχαία μεταβλητή μπορούν να εκφραστούν μέσω μίας συνάρτησης, έτσι ώστε η τιμή της συνάρτησης να ισούται με την πιθανότητα εμφάνισης της τιμής αυτής. Για μία διακριτή μεταβλητή, η συνάρτηση αυτή ονομάζεται συνάρτηση κατανομής πιθανοτήτων (probability distribution function ή PDF) και ορίζεται ως:

$$f(x) = P(X = x) \quad (1.1)$$

όπου πρέπει να ισχύει  $0 \leq f(x) \leq 1$  και  $\sum_x f(x) = 1$ . Επομένως, αν ο σκοπός της ανάλυσης είναι να υπολογιστεί η πιθανότητα ότι η τιμή μίας τυχαίας μεταβλητής είναι μικρότερη ή ίση με κάποιο πραγματικό αριθμό  $x$ , τότε χρησιμοποιείται η αθροιστική συνάρτηση κατανομής (cumulative distribution function ή CDF) που ορίζεται ως:

$$F(x_0) = P(X \leq x_0) = \sum_{x \leq x_0} f(x) \quad (1.2)$$

όπου  $F(-\infty) = 0$  και  $F(\infty) = 1$ . Επιπρόσθετα, εφαρμόζοντας δύο αριθμούς  $a$  και  $b$  με  $a \leq b$ , τότε πρέπει να ισχύει  $F(a) \leq F(b)$ .



**Διάγραμμα 1.1: Συναρτήσεις Κατανομής Πιθανοτήτων και Αθροιστική**

Για παράδειγμα, από τη ρίψη 2 τίμιων ζαριών προκύπτει ένα δειγματικός χώρος με 36 διαφορετικά ενδεχόμενα τα οποία μπορούν να χρησιμοποιηθούν για να σχηματιστεί μία τυχαία μεταβλητή βασιζόμενη στο άθροισμα των αριθμών που εμφανίζονται από τη ρίψη των δύο ζαριών. Κατά συνέπεια, από αυτή τη διαδικασία προκύπτει μία τυχαία μεταβλητή  $X$  που λαμβάνει ακέραιες τιμές από 2 έως 12, ενώ η συνάρτηση πιθανοτήτων προσδιορίζεται με βάση τη δυνατότητα εμφάνισης των τιμών της  $X$ . Εφαρμόζοντας αυτό το φαινόμενο 10,000 φορές προέκυψαν τα ακόλουθα αποτελέσματα που δίνονται στο διάγραμμα 1.1 σε μορφή συνάρτησης κατανομής πιθανοτήτων και αθροιστικής συνάρτησης κατανομής πιθανοτήτων, όπου κατά προσέγγιση χρησιμοποιούνται οι όροι αυτοί αντί των όρων σχετικής συχνότητας και σχετικής αθροιστικής συχνότητας.

### **Συνεχής Τυχαία Μεταβλητή**

Η συνεχής τυχαία μεταβλητή μπορεί να λάβει οποιαδήποτε πραγματική τιμή από το σύνολο των πραγματικών τιμών της. Η μεταβλητή αυτή, όπως και η διακριτή τυχαία μεταβλητή, ορίζεται από το πεδίο ορισμού της, δηλαδή τις τιμές που μπορεί να λάβει η τυχαία μεταβλητή και τις πιθανότητές της. Στη περίπτωση της συνεχούς τυχαίας μεταβλητής η συνάρτηση πιθανοτήτων ονομάζεται συνάρτηση πυκνότητας πιθανοτήτων (probability density function ή PDF) και συμβολίζεται ως  $f(x)$ , όπου η συνάρτηση αυτή θα πρέπει να ικανοποιεί τις ακόλουθες προϋποθέσεις,  $f(x) \geq 0$  για



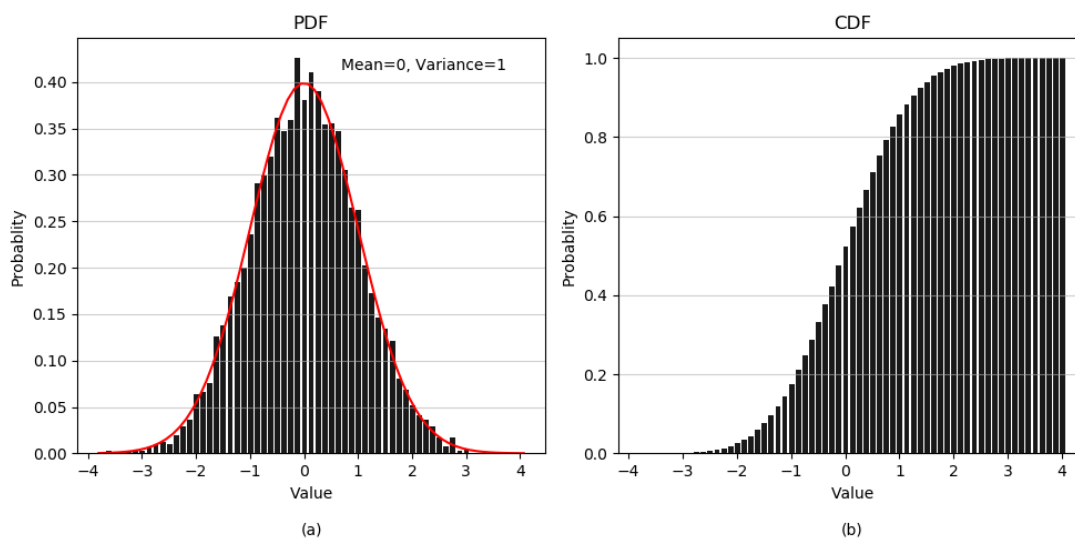
$-\infty \leq x \leq \infty$  και  $\int_{-\infty}^{\infty} f(x) dx = 1$ . Επίσης, η αθροιστική συνάρτηση πυκνότητας πιθανοτήτων ορίζεται ως:

$$F(x_0) = P(X \leq x_0) = \int_{-\infty}^{x_0} f(x) dx \quad (1.3)$$

όπου πρέπει να ισχύουν οι συνθήκες  $F(-\infty) = 0$  και  $F(\infty) = 1$ , ενώ η συνάρτηση μπορεί να εφαρμοστεί για να υπολογιστεί η ακόλουθη πιθανότητα:

$$P(a \leq X \leq b) = F(b) - F(a) \quad (1.4)$$

με τη συνάρτηση πυκνότητας πιθανοτήτων να προκύπτει με την ακόλουθη σχέση  $f(x) = dF(x)/dx$ .



**Διάγραμμα 1.2: Συναρτήσεις Πυκνότητας Πιθανοτήτων και Αθροιστική**

Στο Διάγραμμα 1.2 παρουσιάζονται η συνάρτηση πυκνότητας πιθανοτήτων και η αθροιστική συνάρτηση πυκνότητας πιθανοτήτων, η διαγραμματική απεικόνιση των οποίων προέκυψε από τυχαία πειραματική διαδικασία 10,000 δοκιμών, όπου οι αριθμοί επιλέχθηκαν με τυχαίο τρόπο από μία κανονική κατανομή με μέσο όρο 0 και διακύμανση 1. Από το διάγραμμα της συνάρτησης πυκνότητας πιθανοτήτων, το οποίο κατασκευάζεται από τις σχετικές συχνότητες παρόλο που χρησιμοποιείται ο όρος



πιθανότητες, παρατηρείται ότι η κατανομή είναι συμμετρική ως προς το μηδέν και από το διάγραμμα των αθροιστικών πιθανοτήτων παρατηρείται η σύγκλισή της στη μονάδα.

### 1.3 Ροπές Κατανομής

Η συμπεριφορά μίας τυχαίας μεταβλητής δίνεται από το πεδίο ορισμού της, δηλαδή τις τιμές που μπορεί να λάβει, και τις πιθανότητες που καθορίζουν την εμφάνιση αυτών των τιμών. Για να γίνει όμως κατανοητός ο τρόπος συμπεριφοράς της μεταβλητής, απαιτείται ένα συνοπτικός τρόπος που να προσδιορίζεται από ένα μικρό αριθμό παραμέτρων ικανό να περιγράψει την συμπεριφορά της μεταβλητής. Οι παράμετροι αυτοί ανήκουν σε μία ευρεία οικογένεια παραμέτρων που ονομάζονται ροπές (moments). Η ροπή  $r$  τάξης (βλέπε Αγιακλόγλου & Μπένος, 2014), για  $r \geq 2$ , μιας τυχαίας μεταβλητής  $X$  συμβολίζεται με  $\mu_r$  ορίζεται ως εξής:

$$\mu_r = E[(X - \mu)^r]$$

όπου  $\mu$  συμβολίζει τον μέσο όρο της τυχαίας μεταβλητής και  $r$  είναι θετικός ακέραιος αριθμός. Σημειώνεται ότι αν ο μέσος όρος της τυχαίας μεταβλητής είναι μηδέν τότε όλες οι ροπές προσδιορίζονται από την σχέση  $E(X^r)$  και ονομάζονται ροπές ως προς την αρχή μηδέν.

#### *Αναμενόμενη Τιμή*

Η ροπή πρώτης τάξης ονομάζεται αναμενόμενη τιμή (expected value) και εκφράζει το μέσον όρο (mean) της τυχαίας μεταβλητής  $X$  και ορίζεται ως:

$$\mu = E(X) = \sum_x xP(X = x) \quad (1.5)$$

αν η μεταβλητή  $X$  είναι διακριτή και

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x) dx \quad (1.6)$$

αν η μεταβλητή  $X$  είναι συνεχής.



### Διακύμανση

Η ροπή δεύτερης τάξης ονομάζεται διακύμανση (variance) και περιγράφει τον μέσο όρο των αποκλίσεων των τιμών της  $X$  από τον μέσο όρων της στο τετράγωνο, ενώ ορίζεται ως:

$$\sigma^2 = Var(X) = E[(X - \mu)^2] = \sum_x (X - \mu)^2 P(X = x) \quad (1.7)$$

αν η μεταβλητή  $X$  είναι διακριτή και

$$\sigma^2 = Var(X) = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (X - \mu)^2 f(x) dx \quad (1.8)$$

αν η μεταβλητή  $X$  είναι συνεχής. Η θετική τετραγωνική ρίζα της διακύμανσης ορίζει τη τυπική απόκλιση της μεταβλητής  $X$  (standard deviation) που εκφράζει διασπορά στις ίδιες μονάδες μέτρησης με αυτές της μεταβλητής  $X$ .

### Ασυμμετρία

Η ροπή τρίτης τάξης εκφράζει την ασυμμετρία (skewness) και ορίζεται ως:

$$\mu_3 = E[(X - \mu)^3] \quad (1.9)$$

ενώ ο βαθμός ασυμμετρίας προσδιορίζεται από τον συντελεστή ασυμμετρίας  $\beta_1$  που ορίζεται ως εξής:

$$\beta_1 = \frac{E[(X - \mu)^3]}{\sigma^3}$$

Σημειώνεται ότι ο συντελεστή ασυμμετρίας  $\beta_1$  δεν έχει μονάδες μέτρησης και εκφράζει καλύτερα την ασυμμετρία για κάθε μεταβλητή. Ειδικότερα, όταν η τιμή του είναι μηδέν, τότε η κατανομή είναι συμμετρική, ενώ όταν η τιμή του είναι θετική/αρνητική, η κατανομή είναι ασύμμετρη προς τα δεξιά/αριστερά (sskewed to the right/skewed to the left).



## Κύρτωση

Η ροπή τέταρτης τάξης εκφράζει την κύρτωση (kurtosis) της κατανομής και ορίζεται ως:

$$\mu_4 = E[(X - \mu)^4] \quad (1.10)$$

ενώ ο βαθμός κύρτωσης προσδιορίζεται από τον συντελεστή κύρτωσης  $\beta_2$  που δίνεται από την ακόλουθη σχέση:

$$\beta_2 = \frac{E[(X - \mu)^4]}{\sigma^4}$$

Σημειώνεται ότι ο συντελεστή κύρτωσης  $\beta_2$  δεν έχει μονάδες μέτρησης και η τιμή του δηλώνει τη συγκέντρωση των τιμών της μεταβλητής  $X$  γύρω από το μέσο όρο της. Εάν ο συντελεστής  $\beta_2 = 3$ , τότε η κατανομή ονομάζεται μεσόκυρτη (mesokurtic) ενώ αν ο συντελεστής  $\beta_2 > 3$  /  $\beta_2 < 3$  τότε η κατανομή ονομάζεται λεπτόκυρτη (leptokurtic) / πλατύκυρτη (platykurtic).

## 1.4 Από-κοινού Τυχαίες Μεταβλητές

Υπάρχουν καταστάσεις όπου ο αναλυτής ενδιαφέρεται να μελετήσει ταυτόχρονα τη συμπεριφορά δύο τυχαίων μεταβλητών, μία κατάσταση που εκφράζεται κατά κανόνα από την από-κοινού κατανομή (joint distribution) των τυχαίων μεταβλητών. Η από-κοινού κατανομή δύο τυχαίων μεταβλητών  $X$  και  $Y$  εκφράζει την πιθανότητα να συμβεί ένα ζεύγος ενδεχομένων των μεταβλητών.

Στη περίπτωση που οι τυχαίες μεταβλητές  $X$  και  $Y$  είναι διακριτές και η κατανομή τους ονομάζεται από-κοινού συνάρτηση κατανομής πιθανοτήτων (joint probability distribution function) και ορίζεται ως:

$$f(x, y) = P(X = x \cap Y = y) \quad (1.11)$$

όπου το σύμβολο  $\cap$  δηλώνει τη τομή, δηλαδή την από κοινού εμφάνιση δύο ενδεχομένων και πρέπει να ισχύει  $f(x, y) \geq 0$  και  $\sum_x \sum_y f(x, y) = 1$ . Επίσης, για δύο





διακριτές μεταβλητές, η από-κοινού αθροιστική συνάρτηση κατανομής (joint cumulative distribution function) ορίζεται ως:

$$F(x_0, y_0) = P(X \leq x_0 \cap Y \leq y_0) = \sum_{x \leq x_0} \sum_{y \leq y_0} f(x, y) \quad (1.12)$$

για κάθε τιμή των μεταβλητών  $X$  και  $Y$  του πεδίου ορισμού τους.

Αντίθετα όταν οι τυχαίες μεταβλητές  $X$  και  $Y$  είναι συνεχείς η συνάρτηση αυτή ονομάζεται από-κοινού συνάρτηση πυκνότητας πιθανοτήτων (joint probability density function) και ορίζεται ως:

$$P(X, Y) \in A = \iint_A f(x, y) dx dy \quad (1.13)$$

όπου πρέπει να ισχύει  $f(x, y) \geq 0$  και  $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$ . Ομοίως, για δύο συνεχείς μεταβλητές, η από-κοινού αθροιστική συνάρτηση κατανομής ορίζεται ως

$$F(x_0, y_0) = P(X \leq x_0 \cap Y \leq y_0) = \int_{-\infty}^{y_0} \int_{-\infty}^{x_0} f(x, y) dx dy \quad (1.14)$$

Επιπλέον, εάν έχουν οριστεί η από-κοινού συνάρτηση πυκνότητας πιθανοτήτων και η από-κοινού αθροιστική συνάρτηση κατανομής, τότε ισχύει ότι  $f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$ .

### **Οριακή Κατανομή**

Η κατανομή μιας τυχαίας μεταβλητής  $X$  που υπολογίζεται από την από-κοινού κατανομή δύο τυχαίων μεταβλητών  $X$  και  $Y$  ονομάζεται οριακή κατανομή (marginal distribution) της μεταβλητής  $X$  και για την περίπτωση των διακριτών μεταβλητών ονομάζεται οριακή συνάρτηση κατανομής πιθανοτήτων (marginal probability distribution function) και υπολογίζεται ως

$$f_x(x) = \sum_y f(x, y) \quad (1.15)$$

Κατά ανάλογο τρόπο υπολογίζεται η οριακή συνάρτηση κατανομής πιθανοτήτων για τη μεταβλητή  $Y$ .

Αντίστοιχα, αν οι μεταβλητές  $X$  και  $Y$  είναι συνεχείς, τότε η κατανομή αυτή ονομάζεται οριακή συνάρτηση πυκνότητας πιθανοτήτων (marginal probability density function) για τη μεταβλητή  $X$  και ορίζεται ως.



$$f_x(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad (1.16)$$

ενώ κατά ανάλογο τρόπο υπολογίζεται η οριακή συνάρτηση πυκνότητας πιθανοτήτων για τη μεταβλητή  $Y$ .

### ***Υπό Συνθήκη Κατανομή***

Η υπό συνθήκη κατανομή (conditional distribution) μιας τυχαίας μεταβλητής  $X$  υπολογίζεται από τη συλλογή των πιθανοτήτων που καθορίζονται από τη τυχαία μεταβλητή  $X$  σε ενδεχόμενα που εξαρτώνται από κάποια άλλη,  $Y$ , τυχαία μεταβλητή. Η συνάρτηση πιθανοτήτων που περιγράφει αυτή τη συμπεριφορά, τόσο για διακριτές όσο και για συνέχεις τυχαίες μεταβλητές, ονομάζεται υπό συνθήκη συνάρτηση πιθανοτήτων (conditional probability function) και ορίζεται για μία τυχαία μεταβλητή  $X$  ως:

$$f(x|y) = \frac{f(x, y)}{f_y(y)} \quad (1.17)$$

ενώ κατά ανάλογο τρόπο προσδιορίζεται η υπό συνθήκη συνάρτηση πιθανοτήτων της  $Y$  ως προς τις τιμές της  $X$ . Σημειώνεται ότι η υπό συνθήκη συνάρτηση πιθανοτήτων προκύπτει ως ο λόγος της από-κοινού συνάρτησης κατανομής πιθανοτήτων των μεταβλητών προς την οριακή συνάρτηση κατανομής πιθανοτήτων.

### ***Ανεξάρτητες Τυχαίες Μεταβλητές***

Δύο τυχαίες μεταβλητές  $X$  και  $Y$  είναι ανεξάρτητες όταν και μόνο όταν η από-κοινού συνάρτηση κατανομής πιθανοτήτων τους ισούται με το γινόμενο των οριακών συναρτήσεων πιθανοτήτων τους, δηλαδή όταν ισχύει η ακόλουθη σχέση:

$$f(x, y) = f_x(x) \cdot f_y(y) \quad (1.18)$$

για κάθε ζεύγος τιμών των μεταβλητών  $X$  και  $Y$ . Κατά συνέπεια, εάν οι τυχαίες μεταβλητές  $X$  και  $Y$  είναι ανεξάρτητες, τότε η υπό συνθήκη συνάρτηση πιθανοτήτων τους ισούται την οριακή συνάρτηση κατανομής πιθανοτήτων τους, δηλαδή ισχύει ότι  $f(x|y) = f_x(x)$  και  $f(y|x) = f_y(y)$ .



### Συνδιακύμανση

Η συνδιακύμανση δύο τυχαίων μεταβλητών  $X$  και  $Y$  είναι ένα μέτρο έκφρασης της κοινής μεταβλητότητάς τους και ορίζεται ως:

$$Cov(X, Y) = E[(X - \mu_x)(Y - \mu_y)] \quad (1.19)$$

όπου  $\mu_x = E(X)$  και  $\mu_y = E(Y)$ . Σημειώνεται ότι η συνδιακύμανση είναι μια γενικευμένη μορφή της διακύμανσης ή ότι η διακύμανση είναι μία ειδική περίπτωση της συνδιακύμανσης για  $X = Y$  και αν εξαιρέσει κανείς το μέγεθος και τις μονάδες μέτρησης της συνδιακύμανσης αυτό που μετράει είναι το πρόσημό της που καθορίζει τη κατεύθυνση της συμμεταβλητότητας των δύο μεταβλητών. Πιο συγκεκριμένα, η συνδιακύμανση είναι η μέση τιμή του γινομένου των αποκλίσεων των μεταβλητών  $X$  και  $Y$  από τους μέσους όρους τους αντίστοιχα. Στη περίπτωση που η τιμή της συνδιακύμανσης είναι θετικός αριθμός, οι μεταβλητές  $X$  και  $Y$  συμμεταβάλλονται θετικά μεταξύ τους, δηλαδή όταν οι τιμές της μεταβλητής  $X$  αυξάνονται, τότε αυξάνονται οι τιμές της μεταβλητής  $Y$ , ενώ όταν η τιμή της συνδιακύμανσης είναι αρνητικός αριθμός, οι μεταβλητές  $X$  και  $Y$  συμμεταβάλλονται αρνητικά μεταξύ τους, δηλαδή όταν οι τιμές της μεταβλητής  $X$  μειώνονται, τότε αυξάνονται οι τιμές της μεταβλητής  $Y$ . Αξίζει να σημειωθεί ότι όταν οι μεταβλητές  $X$  και  $Y$  είναι ανεξάρτητες μεταξύ τους τότε η συνδιακύμανση τους είναι ίση με το μηδέν. Η σχέση της συνδιακύμανσης με την ανεξαρτησία των δύο τυχαίων μεταβλητών γίνεται κατανοητή διευρύνοντας την εξίσωση (2.19), έτσι ώστε να προκύψει η ακόλουθη σχέση:

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

δεδομένου όταν αν οι μεταβλητές  $X$  και  $Y$  είναι ανεξάρτητες, τότε η αναμενόμενη τιμή του γινομένου τους είναι ίση με το γινόμενο των μέσων όρων τους, οπότε η συνδιακύμανση είναι ίση με το μηδέν.



## 1.5 Συσχέτιση

Η συνδιακύμανση δύο τυχαίων μεταβλητών  $X$  και  $Y$  παρέχει ένα αριθμητικό αποτέλεσμα, όπου το πρόσημό της καθορίζει την κατεύθυνση συμμεταβολής των δύο μεταβλητών, ενώ το μέγεθος της είναι σχετικά δύσκολο να ερμηνευτεί, λόγω των μονάδων μέτρησης των μεταβλητών που εμπεριέχονται στον αριθμό αυτόν. Για το λόγο αυτό χρησιμοποιούνται διάφορα άλλα μέτρα έκφρασης προσδιορισμού της σχέσης δύο μεταβλητών, τα κυριότερα από τα οποία παρουσιάζονται στη συνέχεια.

### *Συντελεστής Γραμμικής Συσχέτισης*

Ο συντελεστής συσχέτισης (correlation coefficient), που αναπτύχθηκε από τον Pearson (1895), ορίζεται ως ο λόγος της συνδιακύμανσης των δύο μεταβλητών προς το γινόμενο των τετραγωνικών ριζών των διακυμάνσεών τους, δηλαδή ως εξής:

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} \quad (1.20)$$

όπου από τη σχέση αυτή προκύπτουν οι ακόλουθες τρεις χρήσιμες παρατηρήσεις. Πρώτον, ο συντελεστής συσχέτισης δεν έχει μονάδες μέτρησης και κατά συνέπεια μπορεί να χρησιμοποιηθεί για σύγκριση μεταξύ διαφόρων ζευγών μεταβλητών. Δεύτερον, ο συντελεστής συσχέτισης αποδεικνύεται ότι λαμβάνει τιμές από -1 έως 1 και τρίτον, ο συντελεστής συσχέτισης, όπως και η συνδιακύμανση, έχει πρόσημο το οποίο καθορίζεται από το πρόσημο της συνδιακύμανσης και το οποίο δηλώνει τη κατεύθυνση της σχέσης των δύο μεταβλητών. Πιο συγκεκριμένα, ο συντελεστής συσχέτισης εκφράζει τον βαθμό, την ύπαρξη και την κατεύθυνση γραμμικής σχέσης μεταξύ δύο μεταβλητών.

Στη βιβλιογραφία χρησιμοποιούνται εναλλακτικά και οι ακόλουθοι συμβολισμοί για το συντελεστή συσχέτισης. Ειδικότερα, ο πληθυσμιακός συντελεστής συσχέτισης (population correlation coefficient) συμβολίζεται με το ελληνικό μικρό -γράμμα  $\rho$  και υπολογίζεται ως εξής:



$$\rho = \text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} = \frac{\sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum_{i=1}^N (x_i - \mu_x)^2} \sqrt{\sum_{i=1}^N (y_i - \mu_y)^2}}$$

όπου  $\mu_x$  και  $\mu_y$  συμβολίζουν αντίστοιχα τους πληθυσμιακούς μέσους όρους των  $X$  και  $Y$ , ενώ ο δειγματικός συντελεστής συσχέτισης (sample correlation coefficient) συμβολίζεται με το λατινικό μικρό γράμμα  $r$  και υπολογίζεται ως εξής:

$$r = \text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

όπου  $\bar{x}$  και  $\bar{y}$  συμβολίζουν αντίστοιχα τους δειγματικούς μέσους όρους των μεταβλητών  $X$  και  $Y$ . Σημειώνεται ότι ο συντελεστής συσχέτισης δεν επηρεάζεται από μονάδες μέτρησης των μεταβλητών αλλά ούτε και από την παρουσία του μέσου όρου.

Ο δειγματικός συντελεστής συσχέτισης έχει τα ίδια χαρακτηριστικά και ιδιότητες με τον πληθυσμιακό συντελεστή συσχέτισης και δηλώνει το βαθμό της γραμμικής σχέσης των δύο μεταβλητών στο δείγμα. Ωστόσο, αρκετά συχνά και ειδικότερα όταν η τιμή του σε απόλυτο μέγεθος είναι σχετικά μικρή δημιουργείται αίσθημα αμφισβήτησης για την ύπαρξη γραμμικής σχέσης μεταξύ των δύο μεταβλητών. Για το λόγο αυτό, εφαρμόζεται στατιστικός έλεγχος που στηρίζεται στη μηδενική υπόθεση ότι δεν υπάρχει γραμμική συσχέτιση μεταξύ των δύο μεταβλητών, δηλαδή στην υπόθεση:

$$H_0: \rho = 0$$

και πραγματοποιείται με βάση την ακόλουθη στατιστική:

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

όπου  $n$  ο αριθμός των παρατηρήσεων των μεταβλητών και η στατιστική  $t$  ακολουθεί την  $t$  κατανομή με  $(n - 2)$  βαθμούς ελευθερίας, πάντα με την προϋπόθεση της ύπαρξης της κανονικής κατανομής.

### **Συντελεστές Συσχέτισης Κατάταξης**

Ένας άλλος τρόπος μέτρησης της σχέσης δύο μεταβλητών που δεν βασίζεται όμως στις τιμές τους αλλά στις κατατάξεις (ranks) των τιμών τους είναι η συσχέτιση κατάταξης



(rank correlation) που χρησιμοποιείται αρκετά συχνά όταν οι παρατηρήσεις των μεταβλητών εμφανίζουν έντονα έκτροπες/ακραίες τιμές. Η εφαρμογή αυτών των μετρήσεων αποσκοπεί στην αποφυγή των αδικαιολόγητων επιδράσεων από ακραίες παρατηρήσεις στον υπολογισμό της συσχέτισης των μεταβλητών που μπορεί ενδεχομένως να προέρχονται ακόμα και από την απουσία κανονικής κατανομής. Μεταξύ των διάφορων συντελεστών συσχέτισης κατάταξης (rank correlation coefficients) που υπάρχουν, εξέχον θέση στη βιβλιογραφία έχουν ο συντελεστής συσχέτισης κατάταξης που προτάθηκε από τον Spearman (1904) και ο συντελεστής συσχέτισης κατάταξης που προτάθηκε από τον Kendall (1938). Μάλιστα, σύμφωνα με τους Kendall και Gibbons (1990), οι δειγματικοί συντελεστές συσχέτισης κατάταξης έχουν καλύτερη συμπεριφορά στο προσδιορισμό της συσχέτισης των μεταβλητών σε σχέση με τον δειγματικό συντελεστή συσχέτισης Pearson στη περίπτωση όπου ο αριθμός παρατηρήσεων του δείγματος είναι μικρός και στην περίπτωση όπου η απόκοινού κατανομή των μεταβλητών διαφέρει από την κανονική κατανομή.

### ***Συντελεστής Spearman***

Ο συντελεστής συσχέτισης κατάταξης Spearman (Spearman rank correlation coefficient) για δύο τυχαίες μεταβλητές  $X$  και  $Y$  υπολογίζεται με την ακόλουθη διαδικασία. Αρχικά, οι τιμές των παρατηρήσεων των μεταβλητών καταγράφονται σε σειρά κατάταξης, συνήθως σε αύξουσα σειρά, όπου οι ταξινομημένες σειρές έχουν τιμές ίσες με την κατάταξή τους. Στη συνέχεια, ο δειγματικός συντελεστής συσχέτισης κατάταξης Spearman,  $r_s$ , υπολογίζεται ως εξής (Newbold, 2013):

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

όπου  $d_i$  δηλώνει τη διαφορά στην κατάταξη για κάθε παρατήρηση  $i$  μεταξύ των μεταβλητών  $X$  και  $Y$  και  $n$  εκφράζει τον αριθμό των παρατηρήσεων.

Εναλλακτικά, ο συντελεστής συσχέτισης κατάταξης Spearman των μεταβλητών  $X$  και  $Y$ , όπως αναφέρουν οι Casella και Berger (2001), υπολογίζεται και ως:

$$r_s = \frac{\text{Cov}(r_X, r_Y)}{\sqrt{\text{Var}(r_X)}\sqrt{\text{Var}(r_Y)}} \quad (1.21)$$



όπου  $r_X$  και  $r_Y$  είναι ταξινομημένες σειρές των μεταβλητών  $X$  και  $Y$  αντίστοιχα. Κατά συνέπεια, ο συντελεστής συσχέτισης κατάταξης Spearman είναι η εφαρμογή του συντελεστή συσχέτισης Pearson στις ταξινομημένες σειρές των υπό εξέταση μεταβλητών και για αυτό η τιμή του δεν αναμένεται να διαφέρει σημαντικά εκτός κι αν υπάρχουν πολλές ακραίες τιμές. Σημειώνεται ότι στατιστική αναφορά για τη διερεύνηση της ύπαρξης της γραμμικής σχέσης μπορεί να πραγματοποιηθεί και με το συντελεστή συσχέτισης κατάταξης Spearman κάτω όμως από τελείως διαφορετικές καταστάσεις.

### **Συντελεστής Kendall's $\tau$**

Ο συντελεστής συσχέτισης κατάταξης Kendall's  $\tau$  (Kendall's  $\tau$  rank correlation coefficient) για δύο τυχαίες μεταβλητές  $X$  και  $Y$  υπολογίζεται με την ακόλουθη διαδικασία. Αρχικά δημιουργούνται οι ταξινομημένες σειρές  $r_X$  και  $r_Y$  των μεταβλητών  $X$  και  $Y$  αντίστοιχα ακολουθώντας τα ίδια βήματα, όπως και με τον συντελεστή συσχέτισης κατάταξης Spearman. Στη συνέχεια, για κάθε συνδυασμό από ζεύγη  $(r_{X_i}, r_{Y_i})$  και  $(r_{X_j}, r_{Y_j})$  όπου  $i < j$  υπολογίζονται όλα τα αρμονικά ζεύγη (concordant pairs) και όλα τα δυσαρμονικά ζεύγη (discordant pairs). Αρμονικά ζεύγη είναι αυτά για τα οποία ισχύει η σχέση  $(r_{X_i}, r_{Y_i}) \cdot (r_{X_j}, r_{Y_j}) > 0$  για κάθε  $i$  και  $j$  και δυσαρμονικά ζεύγη είναι αυτά για τα οποία ισχύει η σχέση  $(r_{X_i}, r_{Y_i}) \cdot (r_{X_j}, r_{Y_j}) \leq 0$  για κάθε  $i$  και  $j$ . Αν επιπλέον οριστεί ο συνολικός αριθμός των αρμονικών ζευγών ως  $P_n = \sum_{i,j=1}^n I[(r_{X_i}, r_{Y_i})(r_{X_j}, r_{Y_j})]$  και ο συνολικός αριθμός των δυσαρμονικών ζευγών ως  $Q_n = \sum_{i,j=1}^n I[-(r_{X_i}, r_{Y_i})(r_{X_j}, r_{Y_j})]$ , τότε ο δειγματικός συντελεστής συσχέτισης κατάταξης Kendall's  $\tau$  δίνεται από τη σχέση (Casella & Berger, 2001) ως:

$$\tau = \frac{2}{n(n-1)}(P_n - Q_n) \quad (1.22)$$

όπου  $I(A_{i,j}) = 1$  αν  $i = j$  και  $I(A_{i,j}) = 0$  αν  $i \neq j$ .



### Παρατηρήσεις

Στην περίπτωση όπου οι δύο τυχαίες μεταβλητές  $X$  και  $Y$  προέρχονται από κανονική κατανομή, τότε αποδεικνύεται ότι η αναμενόμενη τιμή του δειγματικού συντελεστή συσχέτισης κατάταξης Kendall's  $\tau$  δίνεται από την ακόλουθη σχέση (Coles, 2001):

$$E(\tau) = \frac{2}{\pi} \sin^{-1} \rho \quad (1.23)$$

όπου  $\rho$  είναι ο πληθυσμιακός συντελεστής συσχέτισης του Pearson. Επιπλέον, αποδεικνύεται ότι η αναμενόμενη τιμή του δειγματικού συντελεστή συσχέτισης κατάταξης Spearman δίνεται από την ακόλουθη σχέση (Coles, 2001):

$$E(r_s) = \rho_s + \frac{3(\tau - \rho_s)}{n+1} \quad (1.24)$$

όπου  $\rho_s$  είναι ο πληθυσμιακός συντελεστής συσχέτισης κατάταξης του Spearman. Από τις ανωτέρω σχέσεις διαπιστώνεται ότι και οι δύο δειγματικοί συντελεστές συσχέτισης κατάταξης δεν είναι αμερόληπτοι εκτιμητές, αλλά έχουν σημάδια μεροληπτικής συμπεριφοράς. Τέλος, αν η από-κοινού κατανομή των μεταβλητών  $X$  και  $Y$  είναι κανονική, τότε αποδεικνύεται ότι ισχύει η σχέση (Coles, 2001):

$$\rho = 2 \sin\left(\frac{\pi \rho_s}{6}\right) \quad (1.25)$$

που συνδέει τον πληθυσμιακό συντελεστή συσχέτισης  $\rho$  του Pearson με τον πληθυσμιακό συντελεστή συσχέτισης κατάταξης  $\rho_s$  του Spearman.

## 1.6 Υποφαινομενική Συσχέτιση

Σε πολλές οικονομετρικές εφαρμογές η ανάλυση μεταξύ δύο μεταβλητών ξεκινάει από την ανάλυση της συσχέτισης ως μία προσπάθεια για τον προσδιορισμό της ύπαρξης, του βαθμού αλλά και της κατεύθυνσης της γραμμικής σχέσης μεταξύ των μεταβλητών. Σε αυτήν την προσπάθεια υπολογίζεται ο συντελεστής συσχέτισης του δείγματος και στη συνέχεια εφαρμόζεται ο στατιστικός έλεγχος προκυμμένου να διερευνηθεί στατιστικά αν υφίσταται η σχέση. Ωστόσο, υπάρχουν περιπτώσεις όπου ο αναλυτής





μπορεί να οδηγηθεί σε λανθασμένα συμπεράσματα για τον βαθμό και την ύπαρξη της γραμμικής σχέσης μεταξύ των δύο μεταβλητών χωρίς να υπάρχει καμία απολύτως σχέση μεταξύ τους. Το φαινόμενο αυτό είναι γνωστό ως το φαινόμενο υποφαινομενικής συσχέτισης (spurious correlation phenomenon) που παρουσιάστηκε αρχικά από τον Yule (1926), αν και έγινε γνωστό στη βιβλιογραφία ως το φαινόμενο υποφαινομενικής παλινδρόμησης (spurious regression phenomenon) από την πρωτοποριακή εργασία των Granger and Newbold (1974), οι οποίοι χρησιμοποιώντας Monte Carlo ανάλυση έδειξαν με προσομοιώσεις ότι η παλινδρόμηση δύο ανεξάρτητων διαδικασιών τυχαίων περιπάτων (random walk processes) χωρίς μέσο όρο δημιουργεί λανθασμένα/ψευδή αποτελέσματα περίπου στο 76% για σειρές εκατό παρατηρήσεων. Το φαινόμενο αυτό αναλύθηκε μαθηματικά από τον Phillips (1986) και επεκτάθηκε σε δύο ανεξάρτητες στάσιμες αυτοπαλινδρομου τύπου πρώτου βαθμού σειρές, δηλαδή σε AR(1) διαδικασίες, από τους Granger et al (2001), οι οποίοι διαπίστωσαν ότι ο βαθμός των υποφαινομενικών αποτελεσμάτων δεν εξαρτάται από το μέγεθος του δείγματος αλλά από την τιμή της παραμέτρου του AR(1) υποδείγματος.

Το φαινόμενο της υποφαινομενικής συσχέτισης διερευνήθηκε για δυο ανεξάρτητες στάσιμες αυτοπαλινδρομου τύπου πρώτου βαθμού χρονοσειρές από τους, Agiakloglou και Tsimpanos (2012) και επιβεβαιώθηκαν τα αποτελέσματα της παλινδρόμησης, ότι δηλαδή η μηδενική υπόθεση της μη ύπαρξης συσχέτισης απορρίπτεται κατά περίπου 53%, 36%, 13% και 6% για τιμές της παραμέτρου 0,9, 0,8, 0,5 και 0,2 αντίστοιχα. Ωστόσο, αυτό που πρέπει να σημειωθεί είναι ότι το θέμα αυτό είχε ξεκινήσει από τον Fisher (1915), ο οποίος μελέτησε τις ιδιότητες του δειγματικού συντελεστή συσχέτισης  $r$ , διαπιστώνοντας μάλιστα ότι δεν είναι αμερόληπτος εκτιμητής του πληθυσμιακού συντελεστή συσχέτισης  $\rho$ , αφού η αναμενόμενη τιμή του δίνεται από την ακόλουθη σχέση:

$$E(r) = \rho - \frac{\rho(1-\rho^2)}{2n} \quad (1.26)$$

για κανονικούς πληθυσμούς, όπως επίσης παρουσιάζεται και στους Kenny και Keeping (1951) και Sawkins (1944). Η αλήθεια είναι ότι η μεροληψία δεν μπορεί να έχει μεγάλο μέγεθος δεδομένου ότι ο συντελεστή συσχέτισης λαμβάνει τιμές από -1 έως +1. Ωστόσο, εάν κάποιος ενδιαφέρεται για την ακρίβεια του στατιστικού ελέγχου της μηδενικής υπόθεσης ότι δεν υπάρχει συσχέτιση, τότε η μικρή αυτή σε μέγεθος



μεροληψία, και ειδικότερα για μικρές τιμές του δειγματικού συντελεστή συσχέτισης, μπορεί να επηρεάσει τη διακύμανση του συντελεστή συσχέτισης και κατά συνέπεια το αποτέλεσμα του στατιστικού ελέγχου.

Η αλήθεια είναι ότι η αρχική παρουσίαση του φαινομένου της υποφαινομενικής συσχέτισης από τον Yule (1926) δεν βασίστηκε στον στατιστικό έλεγχο της μηδενικής συσχέτισης, αλλά στηρίχθηκε στην κατανομή των συχνοτήτων του συντελεστή συσχέτισης προκειμένου να διαπιστωθεί αυτή η παράξενη συμπεριφορά που την είχε αποκαλέσει παραίσθηση (illusion) διότι ενδεχομένως δεν μπορούσε να την εξηγήσει. Μάλιστα, στο βιβλίο των Banerjee et al. (1993) εμφανίζονται διαγράμματα των συχνοτήτων του συντελεστή συσχέτισης, σύμφωνα με την τοποθέτηση του Yule, όπου η διαγραμματική απεικόνιση δύο ανεξάρτητων μη στάσιμων χρονοσειρών  $I(2)$  έχει σχήμα U, με τις ακραίες τιμές -1 και +1 να έχουν τη μεγαλύτερη δυνατότητα εμφάνισης, ενώ για δύο ανεξάρτητες μη στάσιμες χρονοσειρές  $I(1)$ , η διαγραμματική απεικόνιση ομοιάζει με ημί-ελλειπτική (semi-ellipse) κατανομή. Για την περίπτωση των δύο ανεξάρτητων λευκών θορύβων χρονοσειρών (white noise processes) η κατανομή των συχνοτήτων τους είναι συμμετρική στο μηδέν και ομοιάζει σαν την κανονική κατανομή.

Τέλος, ο Yule (1926) ανέφερε ότι αυτή η παράξενη συμπεριφορά οφείλονταν στην τυπική απόκλιση του συντελεστή συσχέτισης, κάτι το οποίο διερευνήθηκε από τους Agiakloglou και Tsimpanos (2012) για δύο ανεξάρτητες στάσιμες αυτοπαλίνδρομου τύπου πρώτου βαθμού χρονοσειρές. Οι Agiakloglou και Tsimpanos (2012) έδειξαν ότι για αυτές τις σειρές το φαινόμενο της υποφαινομενικής συσχέτισης μπορεί να εξαλειφθεί τελείως αν στην στατιστική  $t$  για τον στατιστικό έλεγχο της μη ύπαρξης συσχέτισης χρησιμοποιηθεί η διακύμανση του συντελεστή συσχέτισης των δύο ανεξάρτητων  $AR(1)$  χρονοσειρών. Επιπρόσθετα, οι Agiakloglou και Agiropoulos (2016) εξέτασαν τη συμπεριφορά του ελέγχου για τιμές των παραμέτρων που προήλθαν από εκτίμηση και επιπρόσθετα διερεύνησαν τη δύναμη του ελέγχου, διαπιστώνοντας ότι τα αποτελέσματα δεν ήταν ενθαρρυντικά.



## 1.7 Ανακεφαλαίωση

Στο κεφάλαιο αυτό περιεγράφηκαν συνοπτικά τα βασικά στοιχεία που χρησιμοποιούνται στην ανάλυση δύο μεταβλητών. Αρχικά αναλύθηκε η έννοια της τυχαίας μεταβλητής, καθώς και ο συνοπτικός τρόπος συμπεριφοράς της, όπως αυτός προσδιορίζεται μέσα από τις ροπές της. Στη συνέχεια, παρουσιάστηκε η από-κοινού συμπεριφορά δύο μεταβλητών τόσο από την πλευρά των κατανομών τους όσο και από την πλευρά των σχέσεων τους, όπως αυτή εκφράζεται αρχικά από τη συνδιακύμανση τους. Κατόπιν, αναφέρθηκαν οι διάφορες μορφές συντελεστών συσχέτισης που χρησιμοποιούνται για το προσδιορισμό των γραμμικών σχέσεων μεταξύ δύο μεταβλητών μαζί με ορισμένες από τις ιδιότητές τους. Τέλος, σχολιάστηκε το πρόβλημα των υποφαιμενικών συσχετίσεων που εμφανίζεται στην ανάλυση συσχέτισης με χρονοσειρές.



*Αυτή η σελίδα αφήνεται σκόπιμα κενή.*



## ΚΕΦΑΛΑΙΟ 2

### ΣΥΝΑΡΤΗΣΕΙΣ COPULA

#### 2.1 Εισαγωγή

Η ανάλυση συσχέτισης αποτελεί σημείο εκκίνησης πολλών εμπειρικών μελετών για τη διερεύνηση της εξάρτησης μεταξύ δύο τυχαίων μεταβλητών  $X$  και  $Y$ , χωρίς όμως να είναι σε θέση να αντιληφθεί σχέσεις άλλης μορφής πλην της γραμμικής. Επιπρόσθετα, όπως αναφέρθηκε και στο προηγούμενο κεφάλαιο, η ανάλυση συσχέτισης, η οποία βασίζεται στη ύπαρξη κανονικότητας, δεν μπορεί να απεικονίσει σωστά σχέσεις μεταβλητών που εμφανίζουν έντονη μεταβλητότητα με μεγάλη ασυμμετρία αλλά και κύρτωση, όπως χαρακτηριστικά παρατηρείται σε χρηματοοικονομικά δεδομένα. Επίσης, δεν θα πρέπει να αγνοηθεί και η εμφάνιση υποφαινομενικών συσχετίσεων (spurious correlations) για δεδομένα που προέρχονται από χρονοσειρές, καθώς και ο τρόπος με τον οποίο αντιλαμβάνεται ο συντελεστής συσχέτισης την από κοινού κατανομή των δύο μεταβλητών.

Για το λόγο αυτό, έχουν δημιουργηθεί νέες τεχνικές μέτρησης των εξαρτήσεων μεταξύ μεταβλητών, που ονομάζονται *Copula*, έτσι ώστε να μπορούν να αποτυπωθούν καλύτερα πολύπλοκες σχέσεις εξάρτησης, που δεν ανταποκρίνονται στις προδιαγραφές υπολογισμού του συντελεστή συσχέτισης. Οι τεχνικές *Copula* είναι αρκετά περίπλοκες συναρτήσεις που προσπαθούν να προσδιορίσουν την εξάρτηση των μεταβλητών συνδέοντας την από κοινού κατανομή τους με τις οριακές κατανομές πιθανοτήτων τους. Άλλωστε, η λέξη “copula” προέρχεται από τη λατινική γλώσσα και δηλώνει τη σύνδεση ή καλύτερα τη σύζευξη, με το αποτέλεσμα αυτής της προσπάθειας να αποτυπώνεται σε ένα μοναδικό αριθμό που περιγράφει τον τρόπο με τον οποίο τα ζεύγη των μεταβλητών ταιριάζουν μεταξύ τους σε ένα πιο περίπλοκο σύστημα.

Πιο συγκεκριμένα, η εξέταση ολόκληρης της από-κοινού κατανομής των δύο μεταβλητών  $X$  και  $Y$  θα καθιστούσε δυνατή την αντιμετώπιση όλων των ειδών



εξαρτήσεων που υπάρχουν μεταξύ τους, προσεγγίζοντας καλύτερα τη σχέση των μεταβλητών αυτών, μία σχέση που εξαρτάται από πολλούς παράγοντες ακόμα και από τον τρόπο αλλά και από την χρονική περίοδο που συλλέγονται οι παρατηρήσεις τους. Μάλιστα, πολύ συχνά, και ειδικότερα στις χρηματοοικονομικές αναλύσεις, συλλέγονται πολύ περισσότερα δεδομένα από όσα απαιτούνται για τη διερεύνηση της σχέσης των δύο μεταβλητών, προσφέροντας παραδόξως σχετικά μικρότερη ροή πληροφοριών αναφορικά με την αλληλεξάρτησή τους, κάτι που πολλές φορές δημιουργεί σύγχυση ως προς τον καθορισμό της σχέσης τους. Για το λόγο αυτό, η μοντελοποίηση της από-κοινού κατανομής των δύο μεταβλητών, με βάση τα διαθέσιμα εμπειρικά δεδομένα τους, αποτελεί ένα πολύ σημαντικό και δύσκολο κομμάτι του καθορισμού της εξάρτησής τους.

Αναμφισβήτητα, μία στατιστική προσέγγιση που θα περιλάμβανε τόσο την εκτίμηση των οριακών κατανομών των μεταβλητών, χρησιμοποιώντας όλα τα διαθέσιμα δεδομένα αλλά και τις προηγούμενες πληροφορίες, όσο και την εκτίμηση της εξάρτησής τους, χρησιμοποιώντας τα περιορισμένα ζεύγη παρατηρήσεων, θα αποτελούσε την ιδανική τεχνική για τον καθορισμό της σχέσης τους. Άλλωστε, στην πρώτη περίπτωση υπάρχει άφθονη πληροφορία και ικανοποιητικό μέγεθος δείγματος, ενώ στη δεύτερη περίπτωση το μέγεθος δείγματος είναι κατά κανόνα μικρό. Το “πάντρεμα” (couple) αυτών των δύο διακριτών διαδικασιών γίνεται με τις συναρτήσεις *Copula* (*Copula functions*) που ονομάζονται και συναρτήσεις εξάρτησης (*dependence functions*).

Στο κεφάλαιο αυτό αρχικά θα παρουσιαστούν οι βασικές έννοιες των συναρτήσεων *Copula* μαζί με το θεώρημα του Sklar. Κατόπιν, θα αναφερθούν οι λειτουργίες των συναρτήσεων αυτών και θα καθοριστούν τα όριά τους, ενώ επίσης θα αναλυθούν και ορισμένες από τις πιο σημαντικές ιδιότητές τους. Τέλος, θα περιγραφούν οι βασικότερες οικογένειες *Copulas* μεταξύ δύο τυχαίων μεταβλητών, καθώς και οι τρόποι εκτίμησης των παραμέτρων τους.



## 2.2 Βασικές Έννοιες *Copula*

Οι συναρτήσεις *Copula* (*Copula functions*) διερευνούν τις σχέσεις των μεταβλητών βασιζόμενες στην από-κοινού κατανομή τους και παρουσιάζονται με τον εξής τρόπο για δύο μεταβλητές  $X$  και  $Y$ . Αρχικά, είναι γνωστό ότι κάθε τυχαία μεταβλητή  $X$  με συνεχή αθροιστική κατανομή  $F_X$  θα έχει ομοιόμορφη κατανομή (uniform distribution), δηλαδή  $U = F_X(X)$ , έτσι ώστε  $P(U \leq u)$  για  $0 \leq u \leq 1$ , σύμφωνα με το θεώρημα *Fisher Transform*.<sup>1</sup> Ομοίως, ομοιόμορφη κατανομή,  $V = F_Y(Y)$ , θα έχει και μια άλλη τυχαία μεταβλητή  $Y$  με συνεχή αθροιστική κατανομή  $F_Y$ . Επιπρόσθετα, αν ορίζονται οι αντίστροφες συναρτήσεις  $F^{-1}(\cdot)$  για τις  $F_X$  και  $F_Y$ , τότε η από-κοινού κατανομή  $F(X, Y)$  των μεταβλητών  $X$  και  $Y$  μπορεί να γραφεί ως:

$$F(X, Y) = F(F_X^{-1}(U), F_Y^{-1}(V)) \equiv C(U, V) \quad (2.1)$$

ενώ εάν αναδιατυπωθεί η από-κοινού κατανομή ως προς τις μεταβλητές  $U$  και  $V$  που ακολουθούν ομοιόμορφη κατανομή, τότε η σχέση (2.1) μπορεί να εκφραστεί και ως:

$$F(x, y) = F(F_X^{-1}(u), F_Y^{-1}(v)) \equiv C(u, v) \quad (2.2)$$

όπου  $u = F_X(x)$  και  $v = F_Y(y)$  και η  $C$  δηλώνει την συνάρτηση που συνδέει τις ομοιόμορφες μεταβλητές με την από-κοινού κατανομή τους και είναι η συνάρτηση *Copula* η οποία απλά ονομάζεται και *Copula*.<sup>2</sup>

Σημειώνεται επίσης ότι θα μπορούσε να ακολουθηθεί η ίδια διαδικασία και σε μία πολυδιάστατη κατανομή (multivariate distribution)  $m$  βαθμού, έτσι ώστε κάθε από-κοινού κατανομή να μπορεί να γραφτεί ως συνάρτηση  $m$  ομοιόμορφων κατανομών. Στη περίπτωση αυτή οι συναρτήσεις *Copula* είναι παραμετρικά καθορισμένες από-κοινού κατανομές που δημιουργούνται από δεδομένες οριακές κατανομές.

Στην συνέχεια θα παρουσιαστεί ο ορισμός της συνάρτησης *Copula* και θα γίνουν αναφορές σε κάποιες βασικές έννοιες της συμπεριφοράς της. Πιο συγκεκριμένα, θα περιγραφεί το θεώρημα του Sklar που ορίζει τις συναρτήσεις *Copula* μαζί με τον

<sup>1</sup> Το θεώρημα Fisher Transform επίσης ονομάζεται Probability Integral Transformation of Continuous Random Variables, όπως αρχικά περιγράφεται στην εργασία του Fisher (1924).

<sup>2</sup> Στην ελληνική βιβλιογραφία χρησιμοποιείται και ο όρος Σύζευξη.



καθορισμό των ορίων της συμπεριφοράς τους και θα παρουσιαστούν οι βασικές ιδιότητες αυτών των συναρτήσεων.

### 2.2.1 Θεώρημα του Sklar

Το θεώρημα του Sklar (1959) δηλώνει ότι μία *Copula*  $m$  βαθμού, ή  $m - Copula$ , είναι μία συνάρτηση  $C$  από τη μονάδα  $m - κύβου$  ( $m - cube$ )  $[0, 1]^m$  έως το διάστημα μονάδας  $[0, 1]$  που πληροί τις ακόλουθες προϋποθέσεις:

- i.  $C(1, \dots, 1, \dots, a_n, 1, \dots, 1) = a_n$  για κάθε  $n \leq m$  και όλα τα  $a_n$  να ανήκουν στο διάστημα  $[0, 1]$ ,
- ii.  $C(a_1, \dots, a_m) = 0$  αν  $a_n = 0$  για κάθε  $n \leq m$  και
- iii.  $C$  είναι αύξουσα  $m$  βαθμού.

Η πρώτη προϋπόθεση αναφέρει πως όταν είναι γνωστές οι τιμές των παρατηρήσεων των  $m - 1$  μεταβλητών και η καθεμία έχει οριακή πιθανότητα ίση με τη μονάδα, τότε η από-κοινού πιθανότητα των  $m$  αποτελεσμάτων είναι ίδια με την πιθανότητα του υπόλοιπου αβέβαιου αποτελέσματος. Η δεύτερη προϋπόθεση αναφέρει ότι η από-κοινού πιθανότητα όλων των αποτελεσμάτων είναι μηδέν εάν η οριακή πιθανότητα οποιουδήποτε αποτελέσματος είναι μηδέν, μία ιδιότητα η οποία αναφέρεται στη βιβλιογραφία ως *grounded property of Copula* (Nelsen, 1999). Τέλος, η τρίτη προϋπόθεση αναφέρει ότι ο όγκος κάθε  $m - Copula$  είναι μη-αρνητικός, δηλαδή κάθε συνάρτηση  $C$  θα πρέπει να είναι μονότονη.

Η μετάβαση στη περίπτωση διμερών κατανομών, δηλαδή  $m = 2$ , είναι μια σχετικά απλή διαδικασία. Σύμφωνα με τον Nelsen (1999), μία *Copula* δευτέρου βαθμού είναι μία συνάρτηση με πεδίο ορισμού  $[0, 1]^2$  και σύνολο τιμών  $[0, 1]$ , η οποία ικανοποιεί τις ακόλουθες απαραίτητες προϋποθέσεις:

- i. Ομοιόμορφα περιθώρια (uniform margins):  $C(u, 1) = u$  και  $C(1, v) = v$ .
- ii. Γειωμένη στο μηδέν (grounded at zero):  $C(0, v) = C(u, 0) = 0$ .
- iii. Αύξουσα  $2^{\text{ος}}$  βαθμού, δηλαδή να ισχύει:  

$$C(u_1, v_1) - C(u_1, v_2) - C(u_2, v_1) + C(u_2, v_2) \geq 0$$
για  $u_1 > u_2$  και  $v_1 > v_2$ .





Από τον ορισμό αυτόν είναι προφανές ότι εάν  $u$  και  $v$  είναι οι αθροιστικές συναρτήσεις κατανομών των μεταβλητών  $X$  και  $Y$  αντίστοιχα, δηλαδή  $u = F_X(x)$  και  $v = F_Y(y)$ , τότε οποιαδήποτε *Copula*  $C(u, v) = C(F_X(x), F_Y(y))$  είναι μία έγκυρη συνάρτηση διμερής κατανομής. Επίσης, από τον ορισμό αυτόν προκύπτει ότι ισχύει και το αντίστροφο, δηλαδή ότι κάθε από-κοινού αθροιστική κατανομή σχηματίζει μία *Copula*.

Τα παραπάνω συμπεράσματα δηλώνουν ότι οι *Copulas* μπορούν να χρησιμοποιηθούν για να εκφράσουν μία πολυδιάστατη κατανομή ως προς τις οριακές κατανομές της, δηλαδή μπορούν να εξετάσουν από-κοινού κατανομές διαφόρων μεταβλητών μόνο όταν οι οριακές κατανομές τους είναι γνωστές. Έτσι, η *Copula* που σχετίζεται με μία  $m$  – *διάστατη* συνάρτηση  $F$  είναι μία συνάρτηση  $C: [0, 1]^m \rightarrow [0, 1]$  που ικανοποιεί τη συνθήκη:

$$F(y_1, \dots, y_m) = C(F_1(y_1), \dots, F_m(y_m); \theta), \quad (2.3)$$

όπου  $\theta$  είναι η παράμετρος της *Copula* που μετρά την εξάρτηση μεταξύ των οριακών κατανομών και ονομάζεται παράμετρος εξάρτησης (*dependence parameter*). Η παράμετρος  $\theta$  αν και μπορεί να είναι ένα διάνυσμα από παραμέτρους, στη περίπτωση των διμερών κατανομών συνήθως ορίζεται ως ένας αριθμός που δηλώνει το βαθμό της εξάρτησης. Επιπρόσθετα, εάν οι οριακές κατανομές  $F_1(y_1), \dots, F_m(y_m)$  των μεταβλητών είναι συνεχείς, τότε η αντίστοιχη *Copula* είναι μοναδική (*unique*), ή εναλλακτικά, η συνάρτηση  $C$  είναι μοναδικά προσδιορισμένη στο διάστημα  $Range(F_1) \times \dots \times Range(F_m)$ . Ωστόσο, στη περίπτωση όπου όλες οι οριακές κατανομές  $F_1(y_1), \dots, F_m(y_m)$  των μεταβλητών δεν είναι συνεχείς, τότε σύμφωνα με τους Schweizer και Sklar (1983) η αντίστοιχη *Copula* που προκύπτει δεν είναι μοναδικά ορισμένη.

Η σχέση (2.3) δηλώνει ότι η από-κοινού κατανομή εκφράζεται από τις αντίστοιχες οριακές κατανομές και από μία συνάρτηση *Copula* που τις συνδέει μεταξύ τους, χωρίς να καθορίζει από πού προέρχονται αυτές, κάτι το οποίο αποτελεί σημαντικό πλεονέκτημα των τεχνικών *Copula*, δεδομένου ότι οι οριακές κατανομές μπορεί να προέρχονται από διαφορετικές οικογένειες κατανομών. Το γεγονός αυτό επιτρέπει στους αναλυτές να εξετάζουν τις οριακές κατανομές και την εξάρτηση ως δύο διακριτά, αλλά σχετικά, προβλήματα, ενώ στις εμπειρικές εφαρμογές η παράμετρος εξάρτησης



αποτελεί το βασικότερο σημείο αναφοράς για τις συναρτήσεις *Copula* που έρχεται από την εκτίμησή τους.

### 2.2.2 Όρια συναρτήσεων *Copula*

Για τον προσδιορισμό των συναρτήσεων *Copula*, ως πολυδιάστατες συναρτήσεις κατανομής πιθανοτήτων, απαιτείται ο καθορισμός των ορίων τους, ο οποίος επιτυγχάνεται χρησιμοποιώντας τα άνω και κάτω όρια (upper and lower bounds) Fréchet-Hoeffding, από όπου προκύπτουν αντίστοιχα η μέγιστη *Copula* (maximum *Copula*) και η ελάχιστη *Copula* (minimum *Copula*). Στην πραγματικότητα, τα δύο αυτά όρια αντιπροσωπεύουν τις ανώτερες και τις χαμηλότερες επιφάνειες των συναρτήσεων *Copula*, οι οποίες απεικονίζονται διαγραμματικά ωςάν υπέρ-επιφάνειες (hyper-surfaces) που επεκτείνονται στο  $m$  – κύμβο, τοποθετημένες έτσι ώστε όλες οι άλλες επιφάνειες *Copula* να συμπίεζονται μεταξύ αυτών των δύο.

Από την εφαρμογή των ορίων Fréchet-Hoeffding σε μία συνάρτηση  $m$  – *Copula* προκύπτουν, όπως αναφέρεται στον Joe (1997), οι ακόλουθες σχέσεις:

$$W = C(y_1, \dots, y_m) \geq \max[\sum_{j=1}^m F_j - m + 1, 0] \quad (2.4)$$

$$M = C(y_1, \dots, y_m) \leq \min[F_1, \dots, F_m] \quad (2.5)$$

για  $m \geq 2$ , όπου για  $m = 2$  το ανώτερο όριο είναι μία αθροιστική κατανομή πιθανοτήτων, από την οποία σχηματίζεται μία συνάρτηση *Copula*, ενώ το κατώτερο όριο είναι πάντα μία αθροιστική κατανομή πιθανοτήτων. Αντίθετα, για  $m > 2$  η συνάρτηση  $W$  είναι μία αθροιστική κατανομή πιθανοτήτων υπό ορισμένες προϋποθέσεις που περιγράφονται αναλυτικά στον Joe (1997) και σε περίπτωση που το κατώτερο όριο σχηματίζει μία συνάρτηση *Copula*, τότε τα όρια Fréchet-Hoeffding γράφονται ως:

$$C_L(y_1, \dots, y_m) \leq C(y_1, \dots, y_m) \leq C_U(y_1, \dots, y_m), \quad (2.6)$$

όπου  $C_L(y_1, \dots, y_m)$  είναι το κατώτερο όριο και  $C_U(y_1, \dots, y_m)$  είναι το ανώτερο όριο.



Σημειώνεται ότι ο υπολογισμός των ορίων Fréchet-Hoeffding είναι σημαντικός για την επιλογή της κατάλληλης *Copula*, δεδομένου ότι με βάση τα όρια αυτά θα πρέπει η συνάρτηση να ανταποκρίνεται στο δειγματικό χώρο μεταξύ του ανωτέρου και κατωτέρου ορίου, κάτι το οποίο εξασφαλίζεται από τη παράμετρο  $\theta$ , η οποία όσο πλησιάζει το ανώτερο (κατώτερο) όριο του επιτρεπόμενου εύρους της, τόσο η *Copula* πλησιάζει το ανώτερο (κατώτερο) όριο Fréchet-Hoeffding. Ωστόσο, η παραμετρική μορφή της *Copula* μπορεί να επιβάλλει περιορισμούς έτσι ώστε να μην επιτευχθεί η πλήρης κάλυψη μεταξύ των ορίων της. Κατά συνέπεια, μία συγκεκριμένη *Copula* μπορεί να είναι η καλύτερη επιλογή για ένα συγκεκριμένο σύνολο δεδομένων από ότι σε ένα άλλο σύνολο δεδομένων.

### 2.2.3 Περαιτέρω Ιδιότητες της *Copula*

Για κάθε *Copula* δευτέρου βαθμού  $C$  (ή διμερής *Copula*) και για κάθε  $(u, v) \in [0,1] \times [0,1]$ , η συνάρτηση  $C$  ακολουθεί τα ανώτερα και κατώτερα όρια Fréchet-Hoeffding, έτσι ώστε:

$$C_{min}(u, v) \leq C(u, v) \leq C_{max}(u, v)$$

όπου  $C_{max}(u, v)$  εκφράζει τη μέγιστη *Copula* και  $C_{min}(u, v)$  εκφράζει την ελάχιστη *Copula*. Ωστόσο, για να γίνουν κατανοητές οι έννοιες της μέγιστης και της ελάχιστης *Copula* θα πρέπει να καθοριστεί ο όρος εξάρτησης στο πλαίσιο της *Copula*. Πιο συγκεκριμένα, η τέλεια θετική εξάρτηση προσδιορίζεται από τον όρο της κομονοτονικότητας (comonotonicity), που δηλώνει ότι για κάθε ζεύγη  $(u_1, v_1)$  και  $(u_2, v_2)$  που προέρχονται από ένα κομονοτονικό σύνολο θα πρέπει να ισχύει ότι  $\{u_1 \leq v_1, u_2 \leq v_2\}$  ή  $\{u_1 \geq v_1, u_2 \geq v_2\}$ . Αντίθετα, η τέλεια αρνητική εξάρτηση προσδιορίζεται από τον όρο αντιμονοτονικότητα (countermonotonicity) που δηλώνει ότι για κάθε ζεύγη  $(u_1, v_1)$  και  $(u_2, v_2)$  που προέρχονται από ένα αντικομονοτονικό σύνολο θα πρέπει να ισχύει ότι  $\{u_1 \leq v_1, u_2 \geq v_2\}$  ή  $\{u_1 \geq v_1, u_2 \leq v_2\}$ . Έτσι, στην περίπτωση της κομονοτονικότητας, δηλαδή της τέλει θετικής εξάρτησης, ισχύει ότι  $C(u, v) = C_{max}(u, v)$ , ενώ στη περίπτωση της αντιμονοτονικότητας, δηλαδή της τέλει αρνητικής εξάρτησης, ισχύει ότι  $C(u, v) = C_{min}(u, v)$ .



### **Μέγιστη Copula**

Εάν δύο ενδεχόμενα είναι τέλεια θετικά εξαρτημένα θα πρέπει η πιθανότητα να συμβαίνουν και τα δύο ενδεχόμενα να ισούται με την ελάχιστη από τις οριακές πιθανότητες τους. Κατά αυτό τον τρόπο, στη περίπτωση όπου  $X$  είναι μία συνεχής τυχαία μεταβλητή,  $Y = f(X)$  και  $f(\cdot)$  είναι μία αύξουσα διαφορίσιμη συνάρτηση, τότε η μέγιστη *Copula* (maximum Copula) ορίζεται ως:

$$M(u, v) = C_{max}(u, v) = \min(u, v) \quad (2.7)$$

όπου  $u = F_X(x)$  και  $v = F_Y(y)$  και η συνάρτηση  $M$  ικανοποιεί όλες τις απαραίτητες προϋποθέσεις από τον ορισμό της *Copula*. Ως εκ τούτου, η μέγιστη *Copula* μοντελοποιεί μία άμεση μονοτονική σχέση μεταξύ των  $X$  και  $Y$ , η οποία θα μπορούσε να είναι τόσο γραμμική όσο και μη γραμμική. Αυτό αποτελεί μία σημαντική διαφορά από το συντελεστή συσχέτισης Pearson, ο οποίος αποτυπώνει τη γραμμική εξάρτηση αλλά δεν εντοπίζει μη-γραμμικές σχέσεις μεταξύ των εξαρτημένων μεταβλητών.

### **Ελάχιστη Copula**

Αντίστοιχα, για την περίπτωση που δύο μεταβλητές είναι τέλεια αρνητικά εξαρτημένες υπάρχει μία συνάρτηση *Copula* που περιγράφει την εξάρτησή τους. Πιο συγκεκριμένα, εάν  $X$  είναι μία συνεχής τυχαία μεταβλητή,  $Y = f(X)$  και  $f(\cdot)$  είναι μία φθίνουσα διαφορίσιμη συνάρτηση, τότε η ελάχιστη *Copula* (minimum Copula) ορίζεται ως:

$$W(u, v) = C_{min}(u, v) = \max(u + v - 1, 0) \quad (2.8)$$

όπου  $u = F_X(x)$  και  $v = F_Y(y)$ . Η συνάρτηση  $W$  ικανοποιεί όλες τις απαραίτητες προϋποθέσεις από τον ορισμό της *Copula* και μοντελοποιεί την άμεση αρνητική σχέση μεταξύ των  $X$  και  $Y$ , η οποία θα μπορούσε να είναι τόσο γραμμική όσο και μη-γραμμική.



### *Copula Ανεξαρτησίας*

Στην περίπτωση όπου δύο τυχαίες μεταβλητές  $X$  και  $Y$  είναι ανεξάρτητες μεταξύ τους, τότε, ως γνωστόν, η από-κοινού συνάρτηση κατανομής πιθανοτήτων τους προσδιορίζεται από το γινόμενο των οριακών συναρτήσεων πιθανοτήτων τους και κατά συνέπεια το γινόμενο αυτό σχηματίζει τη *Copula* ανεξαρτησίας (product *Copula*) και ορίζεται ως

$$P(u, v) = C_{\perp}(u, v) = u \cdot v \quad (2.9)$$

όπου  $u = F_X(x)$  και  $v = F_Y(y)$  και η συνάρτηση  $P$  ικανοποιεί όλες τις απαραίτητες προϋποθέσεις από τον ορισμό της *Copula*.

Σημειώνεται ότι οι όροι μέγιστη και ελάχιστη *Copula* δηλώνουν ευθέως ότι η τιμή μιας συνάρτησης *Copula* αυξάνεται με την εξάρτηση μεταξύ των τυχαίων μεταβλητών και επομένως, η *Copula* ανεξαρτησίας βρίσκεται στη μέση μεταξύ των ανώτερων και κατώτερων ορίων, όπως αυτά περιγράφονται από τη μέγιστη και ελάχιστη *Copula*. Τέλος, η *Copula* ανεξαρτησίας λειτουργεί ως ένα σημαντικό σημείο αναφοράς, διότι αντιστοιχεί στην ανεξαρτησία τους.

Επιπρόσθετα, για κάθε  $v \in [0, 1]$ , η μερική παράγωγος  $\partial C / \partial u$  ορίζεται για σχεδόν όλα τα  $u$  και ισχύει ότι:

$$0 \leq \frac{\partial}{\partial u} C(u, v) \leq 1$$

και ομοίως, για κάθε  $u \in [0, 1]$  ισχύει ότι:

$$0 \leq \frac{\partial}{\partial v} C(u, v) \leq 1$$

Κατά συνέπεια, η συνάρτηση πυκνότητας *Copula* (*Copula density*) ορίζεται ως:

$$c(u, v) = \frac{\partial}{\partial u \partial v} C(u, v) = \frac{\partial}{\partial v \partial u} C(u, v) \quad (2.10)$$

εάν οι δευτερές μερικές παράγωγοι  $\frac{\partial}{\partial u \partial v} C(u, v)$  και  $\frac{\partial}{\partial v \partial u} C(u, v)$  υπάρχουν και είναι συνεχή στο διάστημα  $[0, 1]^2$ .



Επιπλέον, ένας κυρτός συνδυασμός δύο συναρτήσεων *Copula*  $C_1$  και  $C_2$  είναι επίσης μία *Copula*. Για παράδειγμα, η ακόλουθη συνάρτηση:

$$C(u, v) = \alpha C_{\max}(u, v) + (1 - \alpha) C_{\min}(u, v) \quad (2.11)$$

είναι *Copula*, όπου  $0 \leq \alpha \leq 1$ . Μάλιστα, από την εξίσωση (2.10) προκύπτει ότι οποιοσδήποτε κυρτός γραμμικός συνδυασμός που αποτελείται από *Copulas* είναι μία *Copula*, δηλαδή  $\sum_{i=1}^n \alpha_i C_i$  είναι *Copula* για  $\alpha_i > 0$  και  $\sum_{i=1}^n \alpha_i = 1$ , ενώ μία οικογένεια από *Copulas* που περιλαμβάνει  $C_{\max}$ ,  $C_{\min}$  και  $C_{\perp}$  λέγεται ότι είναι ολοκληρωμένη *Copula* (comprehensive *Copula*).

Τέλος, οι συναρτήσεις *Copula* έχουν μία αμετάβλητη ή αναλλοίωτη ιδιότητα (invariance property) με την οποία η εξάρτηση που εκτιμάται με κάποια *Copula* παραμένει αμετάβλητη σε σχέση με τις συνεχείς μεταβολές των οριακών κατανομών. Αυτό σημαίνει ότι μπορεί να χρησιμοποιηθεί η ίδια *Copula* για την από-κοινού κατανομή των τυχαίων μεταβλητών  $(Y_1, Y_2)$  όπως και για την από-κοινού κατανομή των τυχαίων μεταβλητών  $(\ln Y_1, \ln Y_2)$ . Οι ιδιότητες της κομονοτονικότητας και της αμεταβλητότητας συνδυαστικά καθιστούν τις *Copulas* ελκυστικά εργαλεία για εμπειρικές εφαρμογές.

## 2.3 Οικογένειες *Copula*

Στην ενότητα αυτήν θα παρουσιαστούν οι βασικότερες οικογένειες *Copulas* μεταξύ δύο τυχαίων μεταβλητών ή εναλλακτικά μεταξύ ενός ζεύγους-*Copula* (pair-*Copula*) και θα αναλυθούν οι δομές εξάρτησής τους. Πιο συγκεκριμένα θα παρουσιαστούν αναλυτικά οι Fréchet *Copulas*, οι Farlie–Gumbel–Morgenstern *Copulas*, οι Ελλειπτικές (Elliptical) *Copulas* και οι Αρχιμήδειες (Archimedean) *Copulas*.

Σημειώνεται ότι στη βιβλιογραφία έχει καθιερώσει ένας μεγάλος αριθμός συναρτήσεων *Copula*, όπου η καθεμία από αυτές καθορίζει μία διαφορετική δομή εξάρτησης στις



μεταβλητές. Άλλωστε, ως γνωστό, οι συναρτήσεις *Copula* διαχωρίζουν τις οριακές κατανομές από τις δομές εξάρτησης των υπό εξέταση μεταβλητών και επιλέγεται, για μια συγκεκριμένη εφαρμογή, η συνάρτηση εκείνη που αντιλαμβάνεται καλύτερα από τις διαθέσιμες παρατηρήσεις των μεταβλητών τα χαρακτηριστικά της εξάρτησής τους.

### 2.3.1 Fréchet *Copulas*

Μία απλή μορφή οικογένειας *Copula* δύο παραμέτρων είναι η Fréchet *Copula* που παρουσιάστηκε από τον Fréchet (1958) και προκύπτει ως ένας κυρτός συνδυασμός από μία *Copula* ανεξαρτησίας και από μία κομοτοτονική *Copula*. Η *Copula* αυτή ορίζεται ως:

$$C_{\alpha,\beta}(u, v) = \alpha C_{max}(u, v) + \beta C_{min}(u, v) + (1 - \alpha - \beta)C_{\perp}(u, v) \quad (2.12)$$

όπου  $\alpha + \beta \leq 1$ . Η βασική ιδέα για την κατασκευή αυτής της *Copula* προκύπτει από το σκεπτικό της δημιουργίας μίας οικογένειας από *Copulas* που μπορούν να καλύπτουν τις χαρακτηριστικές ιδιότητες της εξάρτησης, δηλαδή την κομοτοτονικότητα, τη αντιμονοτονικότητα και την ανεξαρτησία.

Μία υποκατηγορία αυτής της οικογενείας *Copula* που χρησιμοποιεί ωστόσο μία παράμετρο έχει προταθεί από τον Mardia (1970) και ορίζεται ως:

$$C_{\alpha}(u, v) = \frac{\alpha^2(1+\alpha)}{2} C_{max}(u, v) + \frac{\alpha^2(1-\alpha)}{2} C_{min}(u, v) + (1 - \alpha^2)C_{\perp}(u, v) \quad (2.13)$$

όπου  $-1 \leq \alpha \leq 1$ , σημειώνοντας ότι αυτή η οικογένεια συναρτήσεων δεν μπορεί να επεκταθεί εύκολα σε υψηλότερες διαστάσεις, δηλαδή σε περισσότερες από δ'υο μεταβλητές.



### 2.3.2 Farlie–Gumbel–Morgenstern *Copulas*

Η Farlie–Gumbel–Morgenstern (FGM) *Copula* προτάθηκε από τον Morgenstern (1956) και ορίζεται ως:

$$C_{\theta}(u, v) = uv(1 + \theta(1 - u)(1 - v)), \quad (2.14)$$

όπου  $(u, v) \in [0, 1]^2$  και  $\theta \in [-1, 1]$  που είναι και η παράμετρος εξάρτησης. Η FGM *Copula* μπορεί να θεωρηθεί ως μία παραλλαγή της *Copula* ανεξαρτησίας εάν η παράμετρος εξάρτησης  $\theta$  ισούται με το μηδέν. Σημειώνεται ότι η οικογένεια αυτή των *Copula* είναι ελκυστική λόγω της απλότητάς της, ωστόσο στη βιβλιογραφία αναφέρεται ότι η χρησιμοποίησή της δεν ενδείκνυται για εξαρτήσεις μεταξύ μεταβλητών των οποίων οι οριακές κατανομές (marginal distributions) έχουν ένταση μετρίου βαθμού.

### 2.3.3 Ελλειπτικές *Copulas*

Στην υποενότητα αυτή παρουσιάζονται δύο από τις πιο σημαντικές περιπτώσεις οικογενειών Ελλειπτικών (elliptical) *Copulas* που είναι η Gaussian (ή normal) *Copula* και η Student's *t Copula*.

#### *Gaussian Copula*

Εάν  $\Phi$  δηλώνει την αθροιστική συνάρτηση σε μία μονομερή τυπική κανονική κατανομή και  $\Psi$  δηλώνει την αθροιστική συνάρτηση σε μία διμερή τυπική κανονική κατανομή, τότε η Gaussian *Copula* ορίζεται ως:

$$C_{\rho}(u, v) = \Psi(\Phi^{-1}(u), \Phi^{-1}(v))$$

δηλαδή ως:

$$C_{\rho}(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \times \left\{ \frac{-(s^2 - 2\rho st + t^2)}{2(1-\rho^2)} \right\} ds dt \quad (2.15)$$





όπου  $s = \Phi^{-1}(u)$ ,  $t = \Phi^{-1}(v)$  και  $\rho$  είναι ο συντελεστής συσχέτισης των μεταβλητών, όπου  $\rho \in (-1, 1)$ .

Η *Copula* αυτή παρουσιάστηκε αρχικά από τον Lee (1983), ο οποίος με αυτήν την προσέγγιση εξέφρασε τη μη-γραμμική εξάρτηση μεταξύ δύο μεταβλητών μέσω της γραμμικής συσχέτισης των μετασχηματισμένων σε τυπική κανονική κατανομή μεταβλητών. Μάλιστα, η συνάρτηση πυκνότητας αυτή της *Copula* δίνεται από την ακόλουθη σχέση:

$$c_{\rho}(u, v) = \frac{1}{\sqrt{(1-\rho^2)}} \exp\left(\frac{\rho^2 s^2 + \rho^2 t^2 - 2\rho st}{2(1-\rho^2)}\right) \quad (2.16)$$

η οποία προκύπτει από τη σχέση (2.10).

### ***Student's t-Copula***

Αντίστοιχα, εάν  $T_v$  δηλώνει την αθροιστική συνάρτηση σε μία μονομερή Student's  $t$  κατανομή και  $T_{v,\rho}$  δηλώνει την αθροιστική συνάρτηση σε μία διμερή Student's  $t$  κατανομή με τους ίδιους βαθμούς ελευθερίας  $v$ , τότε η Student's  $t - Copula$  ορίζεται ως:

$$C_{v,\rho}^t(u, v) = T_{v,\rho}\left(T_v^{-1}(u), T_v^{-1}(v)\right)$$

δηλαδή ως:

$$C_{v,\rho}^t(u, v) = \int_{-\infty}^{T_v^{-1}(u)} \int_{-\infty}^{T_v^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \times \left\{ \frac{-(s^2 - 2\rho st + t^2)}{v(1-\rho^2)} \right\}^{-(v+2)/2} ds dt \quad (2.17)$$

όπου  $s = T_v^{-1}(u)$ ,  $t = T_v^{-1}(v)$ ,  $v > 2$  και  $\rho \in (-1, 1)$ .

Σε αντίθεση με τη Gaussian *Copula* που έχει μοναδική παράμετρο εξάρτησης το συντελεστή συσχέτισης  $\rho$ , η Student's  $t - Copula$  έχει δύο παραμέτρους εξάρτησης, το συντελεστή συσχέτισης  $\rho$  και τους βαθμούς ελευθερίας  $v$ , όπου η παράμετρος αυτή ελέγχει τη βαρύτητα των ουρών των κατανομών. Μάλιστα, η Student's  $t - Copula$  προσεγγίζει ασυμπτωτικά την Gaussian *Copula*, καθώς το  $v$  τείνει στο συν άπειρο, κατά



ανάλογο τρόπο όπως η κατανομή  $t$  προσεγγίζει ασυμπτωτικά την τυπική κανονική κατανομή.

Από τη σχέση (2.10) προκύπτει η συνάρτηση πυκνότητας της Student's  $t - Copula$  που δίνεται από την ακόλουθη σχέση:

$$C^t_{v,\rho}(u,v) = \frac{\Gamma\left(\frac{v+2}{2}\right)}{\sqrt{(1-\rho^2)}\Gamma^2\left(\frac{v+2}{2}\right)} \cdot \frac{\left(\left(1+\frac{s^2}{v}\right)\left(1+\frac{t^2}{v}\right)\right)^{\frac{v+1}{2}}}{\left(1+\frac{s^2+t^2-2\rho st}{v(1-\rho^2)}\right)^{\frac{v+2}{2}}}, \quad (2.18)$$

όπου  $\Gamma$  είναι η συνάρτηση Γάμμα. Σημειώνεται ότι οι ιδιότητες της  $t - Copula$  μελετήθηκαν από τους Fang (2002) και τους Demarta και McNeil (2005), ενώ οι Marshall *et al.* (2003), Jondeau and Rockinger (2006), Ning (2010), και Janus *et al.* (2014) έχουν δείξει ότι η εμπειρική εφαρμογή της  $t - Copula$  σε χρηματοοικονομικά δεδομένα είναι συχνά ανώτερη της Gaussian *Copula*.

### 2.3.4 Αρχιμήδειες *Copulas*

Μία άλλη δημοφιλής οικογένεια *Copula* είναι οι Αρχιμήδειες (Archimedean) *Copulas* οι οποίες βασίζονται στο μετασχηματισμό Laplace των μονοδιάστατων συναρτήσεων κατανομής πιθανοτήτων και επιτρέπουν την ύπαρξη μίας μεγάλης ποικιλίας διαφορετικών δομών εξάρτησης. Οι συναρτήσεις αυτές έχουν σχεδιαστεί για να περιγράφουν σχέσεις εξάρτησης που δεν περιορίζονται κατ' ανάγκη στη συμμετρία, όπως οι Ελλειπτικές *Copulas* που βασίζονται στην ακτινωτή συμμετρία (radial symmetry), δεδομένου ότι προέρχονται από κατανομές που είναι συμμετρικές, όπως τονίζεται από τους Embrechts *et al.* (2002). Επίσης, ένα ακόμα πλεονέκτημα αυτών των συναρτήσεων είναι ότι έχουν κλειστού τύπου λύσεις (closed forms solutions), ενώ εκτενή αναφορά στις Αρχιμήδειες *Copulas* γίνεται από τους Genest και MacKay (1986), Joe (1997) και Nelsen (1999).



Σύμφωνα με τον Nelsen (1999) οι Αρχιμήδειες *Copulas* ορίζονται στη γενική τους μορφή με τον ακόλουθο τρόπο. Έστω  $\phi$  μία συνεχής, αυστηρά φθίνουσα και κυρτή συνάρτηση από  $[0, 1]$  σε  $[0, \infty]$ , έτσι ώστε  $\phi(1) = 0$  και  $\phi(0) \leq \infty$ , από την οποία ορίζεται η ψευδοαντίστροφος (pseudoinverse):

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t) & 0 \leq t \leq \phi(0) \\ 0 & \phi(0) \leq t \leq \infty \end{cases}$$

όπου η ψευδοαντίστροφος  $\phi^{[-1]}$  είναι συνεχής και φθίνουσα στο διάστημα  $[0, \infty]$  και αυστηρά φθίνουσα στο διάστημα  $[0, \phi(0)]$ . Επιπρόσθετα,  $\phi^{[-1]}(\phi(u)) = u$  στο διάστημα  $[0, 1]$  και

$$\phi(\phi^{[-1]}(t)) = \begin{cases} t & 0 \leq t \leq \phi(0) \\ \phi(0) & \phi(0) \leq t \leq \infty \end{cases}$$

ενώ εάν  $\phi(0) = \infty$ , τότε  $\phi^{[-1]} = \phi^{-1}$ . Κατά συνέπεια, εάν η  $\phi$  είναι συνεχής και αυστηρά φθίνουσα συνάρτηση  $[0, 1] \rightarrow [0, \infty]$ , έτσι ώστε  $\phi(1) = 0$ , και  $\phi^{[-1]}$  είναι η ψευδοαντίστροφος της  $\phi$ , τότε η συνάρτηση  $C [0, 1]^2 \rightarrow [0, 1]$  που ορίζεται ως:

$$C(u, v) = \phi^{[-1]}(\phi(u) + \phi(v)) \quad (2.19)$$

είναι η συνάρτηση *Copula* εάν και μόνο εάν η συνάρτηση  $\phi$  είναι κυρτή, ενώ όλες οι *Copulas* που προκύπτουν από τη σχέση (2.19) ονομάζονται Αρχιμήδειες *Copulas*.<sup>3</sup>

Σημειώνεται ότι η συνάρτηση  $\phi$  ονομάζεται γεννήτρια (generator) της *Copula* και εάν  $\phi(0) = \infty$ , τότε η συνάρτηση  $\phi$  είναι αυστηρή γεννήτρια (strict generator), όπου στην περίπτωση αυτή ισχύει ότι  $\phi^{[-1]} = \phi^{-1}$  και η συνάρτηση  $C$  ονομάζεται αυστηρή (strict) Αρχιμήδεια *Copula*. Επιπρόσθετα, εάν η δεύτερη παράγωγος  $\phi''$  υφίσταται, τότε η συνάρτηση πυκνότητας της *Copula* μπορεί να εκφραστεί μέσω της γεννήτριας και των παραγώγων της ως εξής:

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v} = - \frac{\phi''(C(u, v)) \phi'(u) \phi'(v)}{(\phi'(C(u, v)))^3} \quad (2.20)$$

<sup>3</sup> Η ψευδοαντίστροφη συνάρτηση ή γενικευμένη αντίστροφη συνάρτηση είναι μία επέκταση της αντιστρόφου συνάρτησης, η οποία όμως δεν έχει όλες τις ιδιότητες της αντίστροφης συνάρτησης. Μάλιστα, εάν η αντίστροφη συνάρτηση ορίζεται, τότε η ψευδοαντίστροφη συνάρτηση είναι ίδια με την αντίστροφη συνάρτηση.



όπου  $\phi'$  δηλώνει την πρώτη παράγωγο. Στη συνέχεια περιγράφονται οι τρεις πιο δημοφιλείς *Copulas* αυτής της κατηγορίας, που είναι οι Clayton, Gumbel και Frank *Copula*.

### **Clayton Copula**

Εάν η συνάρτηση  $\phi(t) = (t^{-\theta} - 1)/\theta$ , όπου η παράμετρος  $\theta$  λαμβάνει τιμές από  $-1$  έως συν άπειρο χωρίς την τιμή μηδέν, δηλαδή  $\theta \in [-1, 0) \cup (0, \infty)$ , και η ψευδοαντίστροφος  $\phi^{[-1]}(s) = \max\{(1 + \theta s)^{-1/\theta}, 0\}$ , τότε η Clayton *Copula* με παράμετρο εξάρτησης  $\theta$  ορίζεται ως ακολούθως:

$$C_{\theta}(u, v) = \max\{(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, 0\} \quad (2.21)$$

ενώ για  $\theta > 0$  η *Copula* αυτή είναι αυστηρή Αρχιμήδεια *Copula* και η σχέση (2.21) γράφεται ως εξής:

$$C_{\theta}(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta} \quad (2.22)$$

Σημειώνεται ότι όσο η τιμή της παραμέτρου  $\theta$  πλησιάζει το μηδέν, τόσο οι οριακές κατανομές των μεταβλητών γίνονται ανεξάρτητες. Αντίθετα, όσο η τιμή της παραμέτρου  $\theta$  πλησιάζει το άπειρο, η Clayton *Copula* πλησιάζει το ανώτερο όριο Fréchet-Hoeffding, ενώ για καμία τιμή δε πλησιάζει το κατώτερο όριο Fréchet-Hoeffding. Αυτός δε είναι και ο λόγος για τον οποίο η Clayton *Copula* δεν μπορεί να εξηγήσει αρνητική εξάρτηση.

### **Gumbel Copula**

Αντίστοιχα, εάν  $\phi(t) = (-\ln t)^{\theta}$  και η ψευδοαντίστροφος συνάρτηση δίνεται ως  $\phi^{[-1]}(s) = e^{-s^{1/\theta}}$ , τότε η Gumbel *Copula* ορίζεται ως ακολούθως:

$$C_{\theta}(u, v) = \exp\left[-\{(-\log u)^{\theta} + (-\log v)^{\theta}\}^{1/\theta}\right] \quad (2.23)$$

όπου η παράμετρος  $\theta$  λαμβάνει τιμές μεγαλύτερες ή ίσες της μονάδας, δηλαδή  $\theta \geq 1$ , και ονομάζεται παράμετρος εξάρτησης (dependency parameter).



Οι ακραίες τιμές της παραμέτρου  $\theta$ , δηλαδή οι τιμές 1 και  $\infty$ , αντιστοιχούν στην ανεξαρτησία και στο ανώτερο όριο Fréchet-Hoeffding, αντίστοιχα. Ωστόσο, αυτή η *Copula* δεν πλησιάζει το κατώτερο όριο Fréchet-Hoeffding για οποιαδήποτε τιμή της παραμέτρου  $\theta$ . Σημειώνεται ότι η *Copula* αυτή, όπως και η Clayton *Copula*, δεν προσδιορίζει αρνητική εξάρτηση, αλλά σε αντίθεση με την Clayton, η Gumbel *Copula* προσδιορίζει καλύτερα την ισχυρή εξάρτηση στη δεξιά ουρά της κατανομής από ότι την ασθενή εξάρτηση στην αριστερή ουρά της.

### Frank *Copula*

Τέλος, εάν  $\phi(t) = -\ln(e^{-\theta t} - 1/e^{-\theta} - 1)$ , με  $\theta \neq 0$ , και η ψευδοαντίστροφος συνάρτηση δίνεται ως  $\phi^{[-1]}(s) = -1/\theta \ln[1 + e^{-s}(e^{-\theta} - 1)]$ , τότε η Frank *Copula* με παράμετρο εξάρτησης  $\theta$  ορίζεται ως:

$$C_{\theta}(u, v) = -\theta^{-1} \ln \left( 1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right) \quad (2.24)$$

όπου η παράμετρος  $\theta$  λαμβάνει οποιαδήποτε πραγματική τιμή από -1 έως 1. Οι ακραίες τιμές της παραμέτρου -1 και 1, καθώς και η τιμή μηδέν αντιστοιχούν στο κατώτερο και ανώτερο όριο Fréchet-Hoeffding, καθώς και στην ανεξαρτησία, αντίστοιχα.

Η Frank *Copula* είναι αρκετά δημοφιλής στις εμπειρικές εφαρμογές διότι, σε αντίθεση με τις άλλες Αρχιμήδειες *Copulas*, επιτρέπει προσδιορισμούς αρνητικών εξαρτήσεων μεταξύ των οριακών κατανομών των μεταβλητών. Επίσης, η *Copula* αυτή εξετάζει περιπτώσεις εξαρτήσεων όπου οι κατανομές τους είναι συμμετρικές, κάτι ανάλογο που συμβαίνει και στις συναρτήσεις Gaussian και Student-t, ενώ, πάλι σε αντίθεση με τις άλλες Αρχιμήδειες *Copulas*, και τα δύο όρια Fréchet-Hoeffding περιλαμβάνονται στο εύρος της επιτρεπόμενης εξάρτησης.



Κατά συνέπεια, η *Frank Copula* μπορεί να χρησιμοποιηθεί για να διερευνήσει εξαρτήσεις με ισχυρή θετική ή αρνητική εξάρτηση. Ωστόσο, όπως αναφέρεται στη βιβλιογραφία, η εξάρτηση στις ουρές τείνει να είναι σχετικά αδύναμη σε σύγκριση με την *Gaussian Copula*, ενώ η ισχυρότερη εξάρτηση επικεντρώνεται στη μέση της κατανομής, πράγμα που δηλώνει ότι η *Copula* αυτή είναι η πιο κατάλληλη για δεδομένα που εμφανίζουν ασθενή εξάρτηση στις ουρές και για το λόγο αυτό χρησιμοποιείται ευρέως σε εμπειρικές εφαρμογές, όπως σημειώνουν οι Meester και MacKay (1994).

## 2.4 Εκτίμηση παραμέτρων

Η προσδιορισμός των τιμών των παραμέτρων των *Copulas* πραγματοποιείται με διάφορες τεχνικές, οι οποίες βασίζονται στις ιδιότητες του θεωρήματος του Sklar (1959) για να προσδιοριστεί η από κοινού συνάρτηση πιθανοφάνειας. Πιο συγκεκριμένα, από το θεώρημα αυτό προκύπτει ότι για την ύπαρξη μίας *Copula*  $m$  βαθμού θα πρέπει να ισχύει η σχέση (2.3), δηλαδή η σχέση:

$$F(x_1, \dots, x_m) = C(F_1(x_1), \dots, F_m(x_m); \theta)$$

όπου η συνάρτηση  $F$  είναι η αθροιστική συνάρτηση κατανομής, η  $F_i$  είναι η οριακή κατανομή της  $x_i$  και η  $\theta$  είναι η παράμετρος της *Copula* με διάσταση  $m$ , από την οποίαν έπεται ότι η συνάρτηση πυκνότητας της *Copula* (copula density) που δίνεται από τη σχέση (2.10) ορίζεται ως:

$$c(u_1, \dots, u_m; \theta) = \frac{\partial^m c(u_1, \dots, u_m; \theta)}{\partial u_1 \dots \partial u_m} \quad (2.25)$$

όπου  $u_i = F_i(x_i)$ .

Επομένως, η από κοινού συνάρτηση πιθανοφάνειας (likelihood function)  $L$  για  $n$  παρατηρήσεις δίνεται από την ακόλουθη σχέση:

$$L(x_1, \dots, x_m) = \prod_{i=1}^n f(x_{1,i}, \dots, x_{m,i}) \quad (2.26)$$



όπου η συνάρτηση πυκνότητας για κάθε μία παρατήρηση  $(x_{1,i}, \dots, x_{m,i})$  ορίζεται ως:

$$f(x_{1,i}, \dots, x_{m,i}) = c\{F_1(x_{1,i}), \dots, F_m(x_{m,i}); \theta\} \prod_{j=1}^m f_j(x_{j,i}) \quad (2.27)$$

και έτσι η λογαριθμική συνάρτηση πιθανοφάνειας (loglikelihood function)  $l$  γράφεται ως εξής:

$$l(x_1, \dots, x_m) = \sum_{i=1}^n [\log c\{F_1(x_{1,i}), \dots, F_m(x_{m,i})\} + \sum_{j=1}^m \log f_j(x_{j,i})] \quad (2.28)$$

όπου το  $\log$  συμβολίζει το λογαριθμικό μετασχηματισμό και η σχέση αυτή χρησιμοποιείται για την εκτίμηση των παραμέτρων των συναρτήσεων *Copula*.

Πιο συγκεκριμένα, υπάρχουν τρεις βασικές μέθοδοι που χρησιμοποιούν τη λογαριθμική συνάρτηση πιθανοφάνειας για την εκτίμηση των παραμέτρων των συναρτήσεων *Copula* και οι οποίες είναι η πλήρης μέγιστη πιθανοφάνεια (full maximum likelihood), η συμπερασματολογία για όρια (inference for margins) και οι ημι-παραμετρικές προσεγγίσεις (semi-parametric approaches), τεχνικές που περιγράφονται από τους Kim et al (2007) και αναφέρονται περιληπτικά στη συνέχεια.

### ***Μέθοδος της Πλήρους Μέγιστης Πιθανοφάνειας***

Οι εκτιμήσεις των παραμέτρων από τη μέθοδο αυτή προκύπτουν από τη μεγιστοποίηση της λογαριθμικής συνάρτησης πιθανοφάνειας, δηλαδή από την ακόλουθη διαδικασία:

$$\max l(x_1, \dots, x_m)$$

όπου για να επιτευχθεί αυτό θα πρέπει να προσδιοριστούν οι οριακές κατανομές των  $X_1, \dots, X_m$ , έτσι ώστε να ληφθούν τα  $f_i$  και  $F_{x_i}$  για κάθε  $i = 1, \dots, m$ , μία παραμετρική προσέγγιση που είναι αρκετά υπολογιστικά εντατική.

### ***Μέθοδος Συμπερασματολογίας για Όρια***

Η μέθοδος συμπερασματολογίας για όρια, που είναι και αυτή παραμετρική μέθοδος, λειτουργεί σε αντίθεση με τη μέθοδο της πλήρους μέγιστης πιθανοφάνειας σε δύο



στάδια. Στο πρώτο στάδιο εκτιμώνται, μέσω της διαδικασίας της μέγιστης πιθανοφάνειας, οι παράμετροι των οριακών κατανομών και στη συνέχεια εφαρμόζεται η μέθοδος της μέγιστης πιθανοφάνειας για την εκτίμηση των παραμέτρων των *Copulas*, δεδομένων όμως των εκτιμηθέντων οριακών κατανομών.

### Ημι-παραμετρικές Προσεγγίσεις

Για τις δύο προαναφερθείσες παραμετρικές μεθόδους, δηλαδή για την πλήρη μέγιστη πιθανοφάνεια και για τη συμπερασματολογία για όρια, η οριακή κατανομή των μεταβλητών θα πρέπει να είναι καθορισμένη. Ωστόσο, για ορισμένες περιπτώσεις, όπου υπάρχουν δυσκολίες προσδιορισμού των οριακών κατανομών, χρησιμοποιείται μία πιο ευέλικτη μέθοδος, η οποία συνίσταται στη χρήση ενός μη παραμετρικού εκτιμητή για τις οριακές κατανομές και στη συνέχεια οι τιμές αυτές χρησιμοποιούνται για την εκτίμηση των παραμέτρων *Copula* με τη μέθοδο της μέγιστης πιθανοφάνειας.

Πιο συγκεκριμένα, ο εκτιμητής των οριακών κατανομών προσδιορίζεται από την ακόλουθη εμπειρική συνάρτηση κατανομής:

$$\tilde{F}_{nj}(x) = \frac{1}{n+1} \sum_{i=1}^n I(X_{ji} \leq x) \quad (2.29)$$

για κάθε  $j = 1, \dots, m$ , όπου στη συνέχεια γίνεται η αντικατάστασή του στη λογαριθμική συνάρτηση πιθανοφάνειας για να προκύψει η εκτίμηση των παραμέτρων *Copula* από τη μεγιστοποίηση της συνάρτησης αυτής. Σημειώνεται ότι ο εκτίμηση της παραμέτρου, όπως περιγράφεται από τους Chen και Fan (2006), δίνεται από την ακόλουθη σχέση:

$$\tilde{\theta} = \arg \max_{\theta \in \Theta} \frac{1}{n} \sum_{t=1}^n \log c\{\tilde{F}_1(x_1), \dots, \tilde{F}_m(x_m); \theta\} \quad (2.30)$$

όπου  $\arg \max$  είναι η τιμή που μεγιστοποιεί τη συνάρτηση (the argument of the maxima).





## 2.5 Ανακεφαλαίωση

Στο κεφάλαιο αυτό παρουσιάστηκαν οι τεχνικές *Copula* για τη μέτρηση εξαρτήσεων μεταξύ μεταβλητών, στις οποίες μπορούν να αποτυπωθούν καλύτερα σχέσεις εξάρτησης που δεν ανταποκρίνονται κατά κανόνα στις προδιαγραφές υπολογισμού του συντελεστή συσχέτισης, δηλαδή στις γραμμικές σχέσεις. Οι τεχνικές αυτές, όπως αναφέρθηκε, είναι αρκετά πολύπλοκες συναρτήσεις που προσπαθούν να προσδιορίσουν την εξάρτηση των μεταβλητών συνδέοντας την από κοινού κατανομή τους με τις οριακές κατανομές πιθανοτήτων τους.

Αρχικά, παρουσιάστηκαν οι βασικές έννοιες των συναρτήσεων *Copula*, καθώς το θεώρημα του Sklar, το οποίο θεμελιώνει την ύπαρξη αυτών των συναρτήσεων και ορίζει τα άνω και κάτω όρια τους. Κατόπιν, περιεγράφηκαν οι πιο σημαντικές συναρτήσεις *Copula*, οι οποίες χρησιμοποιούνται ευρέως και θα εφαρμοστούν στη συνέχεια για τη διερεύνηση της συμπεριφοράς τους σε περίπλοκες χρονικές σειρές. Ειδικότερα, παρουσιάστηκαν οι Fréchet *Copulas*, οι Farlie–Gumbel–Morgenstern *Copulas*, οι Ελλειπτικές (Elliptical) *Copulas* και οι Αρχιμήδειες (Archimedean) *Copulas*, ενώ στο τέλος γίνεται μία σύντομη αναφορά στον τρόπο εκτίμησης των παραμέτρων τους.



*Αυτή η σελίδα αφήνεται σκόπιμα κενή.*



## **ΚΕΦΑΛΑΙΟ 3**

### **ΓΡΑΜΜΙΚΑ ΚΑΙ ΜΗ ΓΡΑΜΜΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ**

### **ΧΡΟΝΟΣΕΙΡΩΝ**

#### **3.1 Εισαγωγή**

Η οικονομετρική ανάλυση βασίζεται στην ανάπτυξη στατιστικών μεθόδων για την εκτίμηση οικονομικών σχέσεων, καθώς και για τον έλεγχο οικονομικών θεωριών, μία ενέργεια που ξεκινά από τον καθορισμό της θεωρητικής μορφής του υποδείγματος, στο οποίο ορίζονται και οι ρόλοι των μεταβλητών που εμπλέκονται στην ανάλυση. Μάλιστα, το οικονομετρικό υπόδειγμα προσπαθεί να συμπεριλάβει όσες περισσότερες ανεξάρτητες μεταβλητές για να ερμηνεύσει καλύτερα τη συμπεριφορά της εξαρτημένης μεταβλητής, αφήνοντας όλες τις άλλες άγνωστες επιρροές στην παρουσία του τυχαίου μέρους, επιδιώκοντας με αυτόν τον τρόπο να προσδιορίσει τη ξεχωριστή επιρροή που έχει κάθε μία ανεξάρτητη μεταβλητή στην εξαρτημένη, γνωρίζοντας την αιτιώδη σχέση των μεταβλητών.

Ωστόσο, η επιτυχία μιας τέτοιας προσπάθειας βασίζεται πολλές φορές και στην ποιότητα αλλά και στην διαθεσιμότητα των δεδομένων. Οι παρατηρήσεις που χρησιμοποιούνται για κάθε μεταβλητή εκφράζουν τη συμπεριφορά της, όπως αυτή καταγράφεται μέσα από τη δική της εξέλιξη αλλά και μέσα από την επιρροή πολλών άλλων μετρήσιμων και μη παραγόντων, δίνοντας στοχαστική διάσταση στην ανάλυση της συμπεριφοράς του φαινομένου.

Τα δεδομένα μπορούν να εμφανιστούν με διάφορους τύπους, όπως είναι τα διαστρωματικά δεδομένα (cross-sectional data), τα χρονικά δεδομένα ή χρονοσειρές (time-series data), τα διαστρωματικά δεδομένα με χρονοσειρές (panel data) και τα χωροταξικά δεδομένα (spatial data), κάθε ένα από τα οποία έχει κατά κανόνα και



διαφορετική ποσοτική προσέγγιση. Ειδικότερα όμως σημαντικό ενδιαφέρον παρουσιάζει η οικονομετρική ανάλυση χρονοσειρών, δεδομένου ότι πολλά ιδιαίτερα θέματα μπορούν να διερευνηθούν αποκλειστικά για μεταβλητές οι παρατηρήσεις των οποίων προέρχονται από χρονοσειρές, όπως χαρακτηριστικά η αιτιότητα (causality), η συνολοκλήρωση (co-integration), αλλά ακόμα περισσότερο η εκτίμηση ενός οικονομετρικού γραμμικού υποδείγματος απαιτεί οι μεταβλητές που χρησιμοποιούνται να είναι στάσιμες για να είναι αξιόπιστα τα αποτελέσματά του. Για το λόγο αυτό, είναι απαραίτητο να διερευνώνται πρωτίστως τα χαρακτηριστικά των χρονοσειρών προτού γίνει η εκτίμηση του υποδείγματος. Επιπρόσθετα, η γραμμική παλινδρόμηση επιτρέπει, στην περίπτωση των χρονοσειρών, τη διερεύνηση της συμπεριφοράς της εξαρτημένης μεταβλητής από προηγούμενες τιμές των ανεξάρτητων μεταβλητών και πιθανώς και από τις δικές της εκτιμώντας γραμμικά υποδείγματα χρονικών υστερήσεων (distributed lagged models) από τα οποία προκύπτουν χρήσιμα συμπεράσματα αναφορικά με βραχυχρόνιες μεταβολές καθώς και τη αναμενόμενη μακροχρόνια μεταβολή των τιμών της εξαρτημένης μεταβλητής. Επιπρόσθετα, η χρήση μη γραμμικών υποδειγμάτων στην ανάλυση χρονοσειρών γίνεται όλο και πιο επιτακτική ως μία προσπάθεια καλύτερης απεικόνισης της συμπεριφοράς μεταβλητών με έντονη μεταβλητότητα, κάτι το οποίο δεν περιλαμβάνεται στην κλασική προσέγγιση των χρονοσειρών με τη μεθοδολογία των Box και Jenkins (1970).

Στο κεφάλαιο αυτό παρουσιάζονται αρχικά τα χαρακτηριστικά και οι ιδιότητες των στάσιμων χρονοσειρών και αναλύονται τα υποδείγματα κινητού μέσου, τα αυτοπαλίνδρομα υποδείγματα, καθώς και τα μικτά υποδείγματα που χρησιμοποιούνται στην ανάλυση χρονοσειρών. Κατόπιν, γίνεται μία σύντομη αναφορά στις μη στάσιμες χρονοσειρές και περιγράφονται οι στατιστικοί έλεγχοι που εφαρμόζονται για τη διαπίστωση της ύπαρξης μοναδιαίας αυτοσυσχετιζόμενης ρίζας, καθώς και οι εμπειρικές μορφές διερεύνησης του τρόπου εξέλιξης της χρονοσειράς. Στη συνέχεια, αναπτύσσεται το πρόβλημα της ετεροσκεδαστικότητας όπως αυτό εμφανίζεται στην ανάλυση χρονοσειρών με τα αυτοπαλίνδρομα υποδείγματα υπό συνθήκης ετεροσκεδαστικότητας, ενώ στο τέλος περιγράφεται συνοπτικά μία κατηγορία μη-γραμμικών υποδειγμάτων χρονοσειρών που παρουσιάστηκαν στις εργασίες των Bera, Higgins και Lee (1992 και 1996).



### 3.2 Στάσιμες χρονοσειρές

Η ανάλυση χρονοσειρών (time series analysis) αποσκοπεί στον προσδιορισμό του τρόπου εξέλιξης των τιμών μίας μεταβλητής, από μόνες τους χωρίς την άμεση επιρροή άλλων παραγόντων, έτσι ώστε να κατανοηθεί η συμπεριφορά τους και να σχηματιστούν προβλέψεις για τη μελλοντική πορεία τους. Αποκαλείται και μεθοδολογία των Box και Jenkins (1970) (Box and Jenkins methodology), διότι αυτοί πρώτοι ανέπτυξαν και παρουσίασαν ένα ολοκληρωμένο πρότυπο τρόπο στοχαστικής καταγραφής της συμπεριφοράς των τιμών μιας μεταβλητής που αποτυπώνεται μέσα από μία οικογένεια υποδειγμάτων που ονομάζονται Αυτοπαλίνδρομα Κινητού Μέσου υποδείγματα ή απλά ARMA (AutoRegressive Moving Average models), ενώ σημαντική συμβολή στην ανάλυση χρονοσειρών έχουν και οι εργασίες του Yule (1926 και 1927), καθώς και του Slutsky (1937).

Στόχος αυτής της προσπάθειας είναι να ερμηνεύσει την εξέλιξη των τιμών μιας χρονοσειράς κάτω από την προϋπόθεση ότι το παρελθόν καθορίζει το παρόν και προκαλεί το μέλλον. Έτσι, η συμπεριφορά των τιμών της χρονοσειράς θεωρείται ότι σχηματίζεται γραμμικά από τις χρονικές υστερήσεις των ιδίων τιμών της ή από τις τιμές των τυχαίου μέρους, ενώ μπορεί να συνυπάρχουν και τα δύο μαζί. Ωστόσο, αυτό που έχει μεγαλύτερη σημασία, προκειμένου να ανακαλυφθεί ο τρόπος εξέλιξης των τιμών της χρονοσειράς, είναι η μορφή του τρόπου εξέλιξης των τιμών της χρονοσειράς, η οποία απαιτείται να ακολουθεί μία στάσιμη διαδικασία, τα χαρακτηριστικά της οποίας αναλύονται στη συνέχεια.

#### **Στασιμότητα**

Η έννοια της στασιμότητας έχει διαδραματίσει σημαντικό ρόλο στη θεωρία των στοχαστικών διαδικασιών και της ανάλυσης χρονοσειρών. Κάθε υπόδειγμα που περιγράφει τη κατανομή πιθανοτήτων σε μια ακολουθία παρατηρήσεων ονομάζεται στοχαστική διαδικασία. Μία χρονοσειρά  $T$  διαδοχικών παρατηρήσεων  $y_1, y_2, \dots, y_T$  είναι ένα δείγμα από έναν άπειρο πληθυσμό τέτοιων δειγμάτων, τα οποία θα



μπορούσαν να έχουν δημιουργηθεί από μία στοχαστική διαδικασία, οπότε οι χρονοσειρές δύναται να θεωρηθούν ως το αποτέλεσμα μίας στοχαστικής διαδικασίας. Η στασιμότητα (stationarity) μιας χρονοσειράς σχετίζεται με τις στατιστικές της ιδιότητες στο χρόνο. Πιο συγκεκριμένα, οι στάσιμες χρονοσειρές προέρχονται από στοχαστικές διαδικασίες που παραμένουν σε στατιστική ισορροπία και έχουν σταθερή κατανομή πιθανοτήτων στην πάροδο του χρόνου. Οι μαθηματικές αρχές των στάσιμων στοχαστικών διαδικασιών αναπτύχθηκαν αρχικά από τους Kolmogorov (1933), Khintchine (1934) και Wold (1938), ενώ μετέπειτα παρουσιάστηκαν αναλυτικά στο βιβλίο των Box και Jenkins (1970).

Μία χρονοσειρά  $\{y_t\}$  είναι αυστηρώς στάσιμη (strictly stationary) αν η από-κοινού κατανομή (joint distribution)  $y_{t_1}, y_{t_2}, \dots, y_{t_k}$  είναι όμοια με την από-κοινού κατανομή  $y_{t_1+t}, y_{t_2+t}, \dots, y_{t_k+t}$  για κάθε  $t$ , όπου  $k$  είναι ένας αυθαίρετος θετικός ακέραιος αριθμός και η ακολουθία  $t_1, t_2, \dots, t_k$  είναι μία συλλογή από  $k$  θετικούς ακέραιους αριθμούς. Επομένως, η αυστηρή στασιμότητα απαιτεί ότι η από-κοινού κατανομή  $y_{t_1}, y_{t_2}, \dots, y_{t_k}$  να είναι αμετάβλητη στο χρόνο. Αυτή είναι μία πολύ δυνατή υπόθεση η οποία είναι δύσκολο να επαληθευτεί εμπειρικά και για αυτό το λόγο είναι συνήθως αρκετό να λαμβάνονται υπ' όψιν μόνο οι δύο πρώτες ροπές της χρονοσειράς. Συνεπώς, μία χρονοσειρά  $\{y_t\}$  είναι ασθενώς στάσιμη (weakly stationary) αν ο μέσος όρος και η διακύμανση της κατανομής  $y_t$  παραμένουν σταθερά στο χρόνο, ενώ η συνδιακύμανση μεταξύ των παρατηρήσεων  $y_t$  και  $y_{t-l}$  της χρονοσειράς παραμένουν αμετάβλητες στο χρόνο, όπου  $l$  είναι ένας αυθαίρετος ακέραιος αριθμός.

Έτσι, για μία ασθενώς στάσιμη χρονοσειρά, η συνδιακύμανση μεταξύ δύο παρατηρήσεων εξαρτάται μόνο από το απόσταση του χρονικού διαστήματος που διαχωρίζει δύο παρατηρήσεις στο χρόνο και όχι από τη μεταβολή του χρόνου. Για αυτό το λόγο, η χρονοσειρά αυτή αναφέρεται και ως χρονοσειρά στάσιμης συνδιακύμανσης (covariance stationary). Ο ορισμός αυτός, ωστόσο, είναι πολύ περιοριστικός για τις περισσότερες οικονομικές χρονοσειρές στις οποίες συναντάται τάση. Μία συναφή έννοια, που επιτρέπει την ανάλυση στοχαστικών διαδικασιών με προσδιοριστική τάση (deterministic trend) είναι η στάσιμη χρονοσειρά με τάση (trend stationary process).

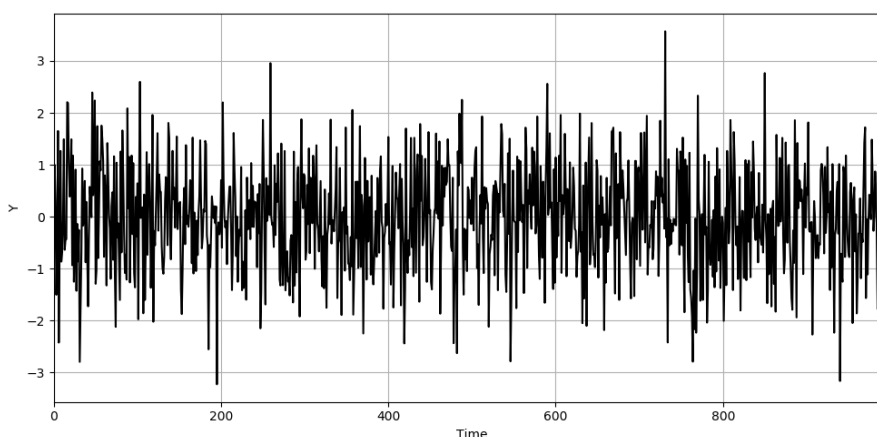


Συγκεκριμένα, μία χρονοσειρά  $\{y_t\}$  λογίζεται ως στάσιμη χρονοσειρά με τάση αν η κατανομή  $y_t = x_t - d_t$  είναι στάσιμης συνδιακύμανσης, όπου  $d_t$  είναι μία τέλεια προβλέψιμη συνιστώσα του  $y_t$ .

Τέλος, αξίζει να σημειωθεί ότι μία αυστηρώς στάσιμη χρονοσειρά, με πεπερασμένες ροπές δεύτερης τάξης είναι ασθενώς στάσιμη χρονοσειρά, αλλά μία ασθενώς στάσιμη χρονοσειρά δεν είναι απαραίτητα αυστηρώς στάσιμη. Ωστόσο, είναι δυνατόν για μια αυστηρώς στάσιμη χρονοσειρά να μην είναι ασθενώς στάσιμη αν αυτή δεν έχει ροπές δεύτερης τάξης. Επίσης, αν η χρονοσειρά ακολουθεί κανονική κατανομή, τότε η ασθενής στασιμότητα είναι ισοδύναμη της αυστηρής στασιμότητας.

### ***Λευκός Θόρυβος και Γραμμικά Υποδείγματα Χρονοσειρών***

Η απλούστερη μορφή μιας στάσιμης χρονοσειράς είναι ο λευκός θόρυβος (white noise). Μία χρονοσειρά  $\{y_t\}$  ονομάζεται λευκός θόρυβος αν προέρχεται από μία ακολουθία ανεξάρτητων και ομοιόμορφων κατανεμημένων τυχαίων μεταβλητών (independently and identically distributed ή iid) με πεπερασμένη μέση τιμή και σταθερή διακύμανση, ενώ εάν η χρονοσειρά  $\{y_t\}$  ακολουθεί κανονική κατανομή με μέσο μηδέν και διακύμανση  $\sigma^2$ , τότε η χρονοσειρά ονομάζεται γκαουσιανός λευκός θόρυβος (gaussian white noise).



**Διάγραμμα 3.1**

**Διαδικασία Λευκού Θορύβου**



Στο Διάγραμμα 3.1 παρουσιάζεται μία προσομοιωμένη ακολουθία 1000 παρατηρήσεων μίας μεταβλητής που προέκυψαν από κανονική κατανομή με μέσο μηδέν και μοναδιαία διακύμανση με ανεξάρτητη συμπεριφορά. Το διάγραμμα αυτό είναι χαρακτηριστικό και αποτελεί σημείο αναφοράς στην ανάλυση χρονοσειρών διότι προσφέρει μία πολύ απλή απεικόνιση των τιμών μιας χρονοσειράς, με τη μορφή της “αθόρυβης” προόδου, δηλώνοντας ότι η εξέλιξη των τιμών της βρίσκεται γύρω από την τιμή του μέσου όρου της χρονοσειράς που είναι το μηδέν.

### ***Γραμμικά Υποδείγματα Χρονοσειρών***

Μία χρονοσειρά  $\{y_t\}$  θεωρείται ότι είναι γραμμική αν μπορεί να γραφεί ως:

$$y_t = \mu + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \quad (3.1)$$

όπου  $\mu$  είναι η μέση τιμή της  $y_t$ ,  $\psi_0 = 1$  και  $\{\varepsilon_t\}$  είναι μία ακολουθία iid τυχαίων μεταβλητών με μέση τιμή μηδέν και με μία καλώς προσδιορισμένη κατανομή, όπως για παράδειγμα η διαδικασία του λευκού θορύβου. Η μεταβλητή  $\varepsilon_t$  εκφράζει τις νέες πληροφορίες της χρονοσειράς στο χρόνο  $t$  και συχνά αναφέρεται ως καινοτομία (innovation) ή αιφνίδια ταραχή (shock).

Σημαντικές αναφορές στη δημιουργία γραμμικών στοχαστικών διαδικασιών αναφέρονται στους Yule (1927), Slutsky (1937), Wold (1938), Barlett (1946), Grenander και Rosenblatt (1957) μεταξύ άλλων. Ειδικότερα, η εργασία του Wold (1938), που θεωρείται ως ένα θεμελιώδες αποτέλεσμα στην ανάπτυξη στάσιμων στοχαστικών διαδικασιών, περιγράφει τη διαπίστωση πως οποιαδήποτε στοχαστική διαδικασία με μη-μηδενική μέση τιμή μπορεί να αναπαρασταθεί σύμφωνα με την σχέση (3.1) και με την υπόθεση ότι  $\sum_{j=0}^{\infty} |\psi_j|^2 < \infty$ . Μάλιστα, για να αντιπροσωπεύει η σχέση (3.1) μια έγκυρη στάσιμη στοχαστική διαδικασία, είναι απαραίτητο οι συντελεστές  $\psi_j$  να είναι απολύτως αθροιστικοί (absolutely summable), δηλαδή να ισχύει ότι  $\sum_{j=0}^{\infty} |\psi_j| < \infty$ , ενώ οι τυχαίες μεταβλητές  $\varepsilon_t$  που έχουν κοινή διακύμανση θα μπορούσαν να μην είναι ανεξάρτητες μεταξύ τους.





### Αυτοσυνδιακύμανση και Αυτοσυσχέτιση

Η συνδιακύμανση  $\gamma_l = Cov(y_t, y_{t-l})$  μίας στάσιμης χρονοσειράς  $y_t$  ονομάζεται αυτοσυνδιακύμανση χρονικής υστέρησης  $l$  (lag- $l$  autocovariance) και έχει δύο σημαντικές ιδιότητες: α)  $\gamma_0 = Var(y_t)$  και β)  $\gamma_{-l} = \gamma_l$ , όπου στην πρώτη ιδιότητα ορίζεται η διακύμανση της χρονοσειράς, ενώ η δεύτερη προκύπτει από τη στασιμότητα, δεδομένου ότι ισχύει:

$$Cov(y_{t+l}, y_t) = Cov(y_t, y_{t-l})$$

για κάθε τιμή του  $l$ .

Επίσης, η συνάρτηση αυτοσυσχέτισης χρονικής υστέρησης  $l$  (lag- $l$  autocorrelation function ή ACF) για στάσιμες χρονοσειρές ορίζεται ως:

$$\rho_l = \frac{Cov(y_t, y_{t-l})}{\sqrt{Var(y_t) Var(y_{t-l})}} = \frac{Cov(y_t, y_{t-l})}{Var(y_t)} = \frac{\gamma_l}{\gamma_0} \quad (3.2)$$

όπου εξ 'ορισμού ισχύει ότι  $\rho_0 = 1$ ,  $\rho_l = \rho_{-l}$  και  $-1 \leq \rho_l \leq 1$ , σημειώνοντας ότι για μία χρονοσειρά  $y_t$  στάσιμης συνδιακύμανσης δεν συσχετίζεται σειριακά μόνο όταν  $\rho_l = 0$  για κάθε  $l > 0$ , όπως στη χρονοσειρά του λευκού θορύβου.

Συνδέοντας τις έννοιες της αυτοσυνδιακύμανσης και της αυτοσυσχέτισης με τις γραμμικές χρονοσειρές, μπορούν να προσδιοριστούν η μέση τιμή και η διακύμανση μίας ασθενώς στάσιμης χρονοσειράς  $y_t$ , χρησιμοποιώντας απλά την υπόθεση ανεξαρτησίας της ακολουθίας  $\{\varepsilon_t\}$  ως εξής:

$$E(y_t) = \mu, \quad Var(y_t) = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \psi_i^2 \quad (3.3)$$

όπου  $\sigma_\varepsilon^2$  είναι η διακύμανση της μεταβλητής  $\varepsilon_t$ . Επίσης, επειδή θα πρέπει η διακύμανση να είναι πεπερασμένος αριθμός, δηλαδή  $Var(y_t) < \infty$ , η ακολουθία  $\{\psi_i^2\}$  πρέπει να συγκλίνει έτσι ώστε  $\psi_i^2 \rightarrow 0$  καθώς  $i \rightarrow \infty$ . Μάλιστα, ως συνέπεια αυτής της κατάστασης θα πρέπει σε στάσιμες χρονοσειρές η επίδραση της αιφνίδιας ταραχής  $\varepsilon_{t-i}$  στην παρατήρηση  $y_t$  να εξαφανίζεται καθώς το  $i$  αυξάνει.



Η αυτοσυνδιακύμανση χρονικής υστέρησης  $l$  υπολογίζεται ως εξής:

$$\begin{aligned}
 \gamma_l &= \text{Cov}(y_t, y_{t-l}) \\
 &= E \left[ \left( \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \right) \left( \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-l-j} \right) \right] \\
 &= \sum_{j=0}^{\infty} \psi_{j+l} \psi_j E(\varepsilon_{t-l-j}) \\
 &= \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+l}
 \end{aligned} \tag{3.4}$$

όπου τα βάρη  $\psi$  συνδέονται με τις αυτοσυσχετίσεις της χρονοσειράς  $y_t$  με την παρακάτω συνάρτηση:

$$\rho_l = \frac{\gamma_l}{\gamma_0} = \frac{\sum_{i=0}^{\infty} \psi_i \psi_{i+l}}{1 + \sum_{i=0}^{\infty} \psi_i^2} \tag{3.5}$$

για κάθε  $l > 0$  και  $\psi_0 = 1$ . Σε στάσιμες χρονοσειρές οι τιμές των βαρών  $\psi$  τείνουν στο μηδέν, δηλαδή  $\psi_i \rightarrow 0$  καθώς  $i \rightarrow \infty$ , και ως εκ τούτου η αυτοσυσχέτιση  $\rho_l$  συγκλίνει στο μηδέν καθώς το  $l$  αυξάνει.

Επιπρόσθετα, ο δειγματικός συντελεστής αυτοσυσχέτισης χρονικής υστέρησης  $l$  (lag- $l$  sample autocorrelation)  $\hat{\rho}_l$ , ο οποίος συμβολίζεται και με  $r$ , ορίζεται ως:

$$\hat{\rho}_l = \frac{\sum_{t=l}^T (y_t - \bar{y})(y_{t-l} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} \tag{3.6}$$

για κάθε  $0 \leq l \leq T-1$ , όπου  $\bar{y}$  είναι η δειγματική τιμή του μέσου όρου της χρονοσειράς. Αν  $\{y_t\}$  είναι μία iid ακολουθία ικανοποιώντας τη συνθήκη  $E(y_t^2) < \infty$ , τότε ο δειγματικός συντελεστής συσχέτισης  $\hat{\rho}_l$  έχει ασυμπτωτικά κανονική κατανομή με μέση τιμή μηδέν και διακύμανση  $1/T$  για κάθε σταθερό θετικό ακέραιο αριθμό. Πιο συγκεκριμένα, εάν μία χρονοσειρά  $y_t$  είναι σταθερής διακύμανσης ικανοποιώντας τη συνθήκη  $y_t = \mu + \sum_{i=0}^q \psi_i \varepsilon_{t-i}$ , όπου  $\psi_0 = 1$  και  $\{\varepsilon_t\}$  είναι μία ακολουθία iid τυχαίων μεταβλητών με μέση τιμή μηδέν, τότε ο  $\hat{\rho}_l$  έχει ασυμπτωτικά κανονική κατανομή με μέση τιμή μηδέν και διακύμανση  $(1 + 2 \sum_{i=1}^q \rho_i^2)/T$  για  $l > q$ . Το αποτέλεσμα αυτό είναι γνωστό ως *Bartlett's formula* και περιγράφεται αναλυτικά από τους Bartlett (1946), Box και Jenkins (1976) και Fuller (1976). Μάλιστα, έχοντας ορίσει τη μορφή



του δειγματικού συντελεστή αυτοσυσχέτισης και γνωρίζοντας τις ασυμπτωματικές ιδιότητές του, το επόμενο βήμα είναι να ελεγχθεί η στατιστική σημαντικότητά του, όπου για έναν συγκεκριμένο θετικό ακέραιο αριθμό  $l$ , ο έλεγχος βασίζεται στις ακόλουθες υποθέσεις:

$$H_0: \rho_l = 0 \text{ και } H_1: \rho_l \neq 0$$

και εφαρμόζεται χρησιμοποιώντας την ακόλουθη στατιστική:

$$t_{\text{statistic}} = \frac{\hat{\rho}_l}{\sqrt{(1+2 \sum_{i=1}^{l-1} \hat{\rho}_i^2)/T}} \quad (3.7)$$

η οποία ακολουθεί ασυμπτωτικά την τυπική κανονική κατανομή εάν η χρονοσειρά  $y_t$  ακολουθεί την κανονική κατανομή και ικανοποιείται η υπόθεση  $\rho_j = 0$  για  $j > l$ . Κατά συνέπεια, η απόφαση για την αποδοχή ή όχι της μηδενικής υπόθεσης γίνεται με κριτική τιμή  $z_{\alpha/2}$  που προέρχεται από την τυπική κανονική κατανομή (Tsay, 2010).

Επίσης, οι Box και Pierce (1970), σε μία προσπάθεια να διερευνήσουν αν οι παρατηρήσεις μιας χρονοσειράς ακολουθούν το υπόδειγμα του λευκού θορύβου, κατασκεύασαν έναν στατιστικό έλεγχο για την ταυτόχρονη στατιστική σημαντικότητα των αυτοσυσχετίσεων σε διάφορες χρονικές υστερήσεις βασισμένο στην μηδενική υπόθεση  $H_0: \rho_1 = \dots = \rho_m = 0$  έναντι της εναλλακτικής υπόθεσης ότι τουλάχιστον μία τιμή δεν είναι μηδέν, δηλαδή  $H_1: \rho_l \neq 0$  για κάποιο  $l \in \{1, \dots, m\}$ , όπου  $m$  ένας θετικός ακέραιος αριθμός. Η έλεγχος αυτός εφαρμόζεται με βάση την ακόλουθη στατιστική  $Q$ :

$$Q(m) = T \sum_{l=1}^m r_l^2 \quad (3.8)$$

η οποία προσεγγίζει ασυμπτωτικά την κατανομή  $X^2$  με  $m$  βαθμούς ελευθερίας και ονομάζεται έλεγχος Portmanteau (Portmanteau test). Στη συνέχεια, οι Ljung και Box (1978) τροποποίησαν τη στατιστική  $Q$  προκειμένου να αποκτήσουν καλύτερη συμπεριφορά και διαπίστωσαν ότι η νέα στατιστική:

$$\tilde{Q}(m) = T(T+2) \sum_{l=1}^m \frac{r_l^2}{T-l} \quad (3.9)$$



ακολουθεί πλησιέστερα την  $X^2$  κατανομή με  $m$  βαθμούς ελευθερίας, όπου η επιλογή της τιμής του  $m$  δεν έχει καθοριστεί και μπορεί να επηρεάσει την απόδοση της στατιστικής  $Q(m)$ . Μάλιστα, αρκετά συχνά χρησιμοποιούνται διάφορες τιμές του  $m$ , προκειμένου να διαπιστωθεί η ευαισθησία του στατιστικού ελέγχου αν και μελέτες με προσομοιωμένα δεδομένα προτείνουν η τιμή του  $m$  να επιλέγεται από τη λογαριθμική τιμή του μεγέθους του δείγματος, δηλαδή  $m = \ln(T)$ , όπου ο έλεγχος έχει την καλύτερη απόδοση.

### 3.3 Υποδείγματα Κινητού Μέσου

Η σχέση (3.1) παρέχει μια γενικευμένη αναπαράσταση για οποιαδήποτε στάσιμη χρονοσειρά και μπορεί να χρησιμοποιηθεί για την περιγραφή υποδειγμάτων κινητού μέσου (moving average ή MA) πεπερασμένου βαθμού, που ικανοποιούν την ιδιότητα της αντιστρεψιμότητας (invertibility) και φυσικά εκτός αυτών που δημιουργούνται με μηδενικούς συντελεστές  $\psi_i$ .<sup>4</sup> Ειδικότερα, στα υποδείγματα κινητού μέσου, κατά σύμβαση ο συντελεστής  $\psi_0$  ισούται με τη μονάδα και οι υπόλοιποι μη-μηδενικοί συντελεστές αναπαρίστανται με το γράμμα  $\theta$ , έχοντας αρνητικό πρόσημο. Οπότε, μία στοχαστική διαδικασία κινητού μέσου με βαθμό  $q$  που συμβολίζεται ως  $MA(q)$  και έχει τη μορφή:

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (3.10)$$

όπου  $\mu$  είναι ο μέσος όρος της χρονοσειράς και η ακολουθία  $\{\varepsilon_t\}$  είναι λευκός θόρυβος. Από τη σχέση αυτή προκύπτει ότι η μέση τιμή και η διακύμανση μίας  $MA(q)$  διαδικασίας ισούνται με:

$$E(y_t) = E(\mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}) = \mu \quad (3.11)$$

$$Var(y_t) = Var(\mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}) = \sigma^2(1 + \theta_1^2 + \dots + \theta_q^2) \quad (3.12)$$

<sup>4</sup> Η αντιστρεψιμότητα εξασφαλίζει την μετατροπή ενός κινητού μέσου όρου υποδείγματος σε ένα αυτοπαλίνδρομο υπόδειγμα απείρων όρων.



όπου  $\sigma^2$  είναι η διακύμανση του λευκού θορύβου (Tsay, 2010).

Ομοίως, η αυτοσυνδιακύμανση χρονικής υστέρησης  $k$ , όπου για απλούστευση  $k$  ισούται με τη χρονική υστέρηση  $-l$ , μπορεί να υπολογισθεί ως:

$$\begin{aligned} \gamma_k &= Cov(y_t, y_{t+k}) \\ &= E[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q})(\varepsilon_{t+k} - \theta_1 \varepsilon_{t+k-1} - \dots - \theta_q \varepsilon_{t+k-q})] \\ &= \begin{cases} \sigma^2(-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q), & k = 1, 2, \dots, q \\ 0, & k > q \end{cases} \end{aligned} \quad (3.13)$$

ενώ από τις σχέσεις (3.12) και (3.13) προκύπτει πως η συνάρτηση αυτοσυσχέτισης χρονικής υστέρησης  $k$  μίας  $MA(q)$  διαδικασίας είναι:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2}, & k = 1, 2, \dots, q \\ 0, & k > q \end{cases}. \quad (3.14)$$

σημειώνοντας ότι η συνάρτηση αυτοσυσχέτισης είναι ιδιαίτερα χρήσιμη για τον προσδιορισμό του βαθμού  $q$  μίας  $MA(q)$  διαδικασίας, δεδομένου ότι μετά τη χρονική υστέρηση του βαθμού του υποδείγματος η τιμή του συντελεστή αυτοσυσχέτισης μηδενίζεται. Σημειώνεται ότι η τιμή του δειγματικού συντελεστή αυτοσυσχέτισης δεν θα είναι απαραίτητα ίση με μηδέν μετά τη χρονική υστέρηση  $q$ , όμως αναμένεται να έχει πάρα πολύ μικρή τιμή. Πιο συγκεκριμένα, για ένα σύνολο  $N$  παρατηρήσεων, η τιμή του δειγματικού συντελεστή αυτοσυσχέτισης ελέγχεται να βρίσκεται μεταξύ των ορίων  $\pm 2/\sqrt{N}$ , όπου  $1/\sqrt{N}$  είναι η κατά προσέγγιση τιμή για την τυπική απόκλιση του συντελεστή για οποιαδήποτε χρονική υστέρηση υπό την υπόθεση  $\rho_k = 0$  για κάθε  $k$ .

### ***Υπόδειγμα Κινητού Μέσου πρώτου βαθμού***

Η απλούστερη μορφή ενός υποδείγματος κινητού μέσου δημιουργείται από την σχέση (3.10) για  $q = 1$ , ονομάζεται υπόδειγμα κινητού μέσου όρου πρώτου βαθμού (moving average process of order one), συμβολίζεται ως  $MA(1)$  και έχει τη μορφή:

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (3.15)$$



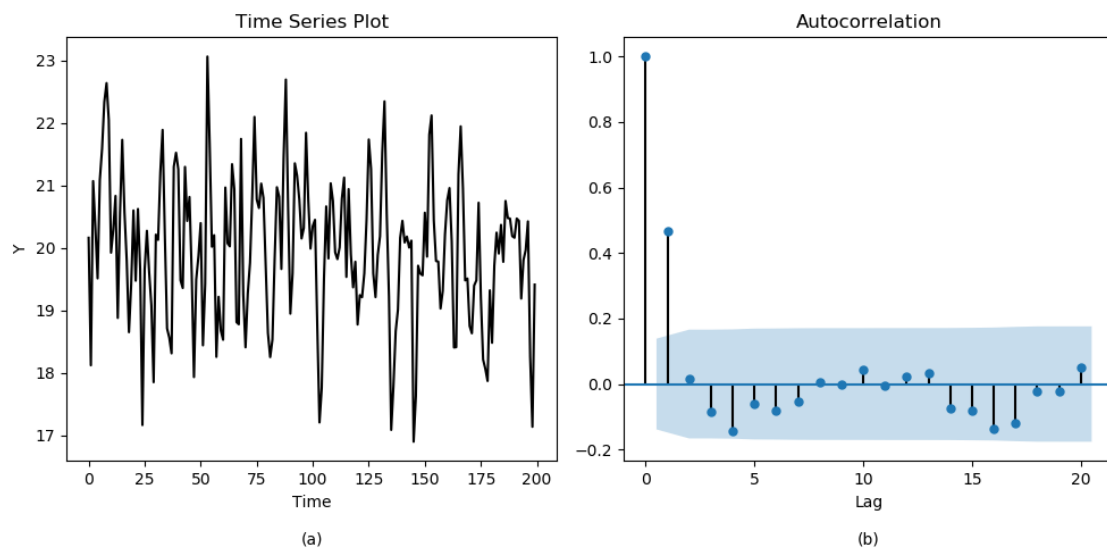
όπου  $\mu$  ο μέσος όρος της χρονοσειράς και η παράμετρος  $\theta_1$  λαμβάνει σε απόλυτο μέγεθος τιμές μικρότερες της μονάδας, ενώ στο υπόδειγμα αυτό η συνάρτηση αυτοσυνδιακύμανσης ορίζεται ως:

$$\begin{aligned}\gamma_0 &= \sigma^2(1 + \theta_1^2) \\ \gamma_1 &= -\theta_1\sigma^2 \\ \gamma_k &= 0, \quad k > 1\end{aligned}\tag{3.16}$$

και ομοίως, η συνάρτηση αυτοσυσχέτισης δίνεται ως:

$$\begin{aligned}\rho_1 &= \frac{-\theta_1}{1+\theta_1^2} \\ \rho_k &= 0, \quad k > 1\end{aligned}\tag{3.17}$$

όπου οι τιμές της μηδενίζονται από το  $\rho_2$  και μετά.



**Διάγραμμα 3.2**

### **Διαδικασία MA(1)**

Στο Διάγραμμα 3.2 απεικονίζεται μία διαδικασία ενός υποδείγματος κινητού μέσου πρώτου βαθμού από προσομοιωμένα δεδομένα έχουν παραχθεί από το ακόλουθο υπόδειγμα:



$$y_t = 20 + \varepsilon_t + 0.65\varepsilon_{t-1}$$

Επίσης παρουσιάζεται και το διάγραμμα των δειγματικών συντελεστών αυτοσυσχετίσεων, γνωστό ως κορελόγραμμα (correlogram, στο οποίο οι τιμές τους είναι σχεδόν μηδέν μετά την τιμή του συντελεστή αυτοσυσχέτισης πρώτου βαθμού.

### **Υπόδειγμα Κινητού Μέσου δευτέρου βαθμού**

Το υπόδειγμα κινητού μέσου δευτέρου βαθμού (moving average process of order two), συμβολίζεται ως MA(2) δημιουργείται από την σχέση (3.10) για  $q = 2$  και έχει την ακόλουθη μορφή:

$$y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} \quad (3.18)$$

όπου οι τιμές των παραμέτρων  $\theta_1$  και  $\theta_2$  ικανοποιούν την αντιστρεψιμότητα.

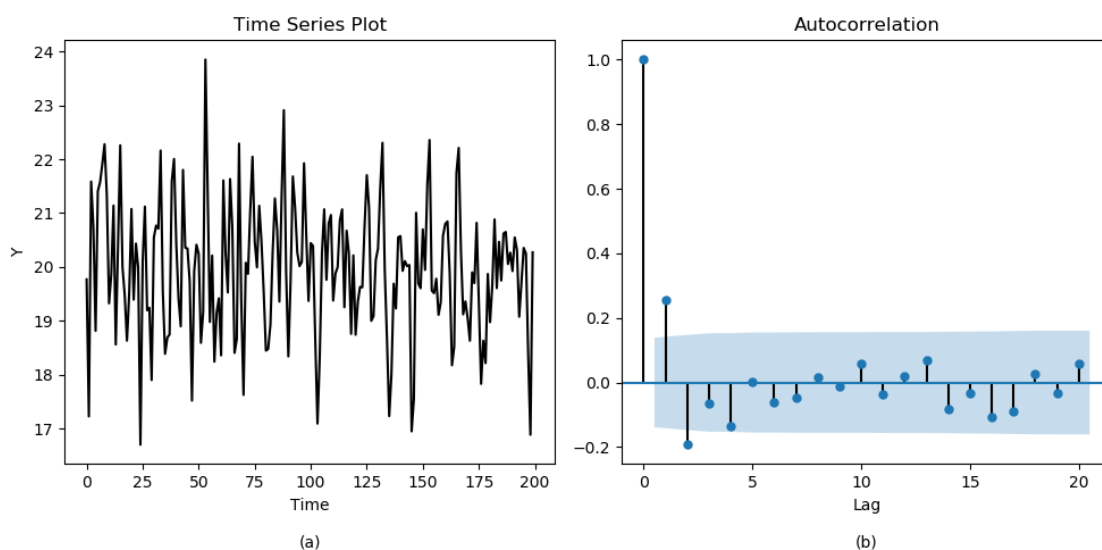
Η συνάρτηση αυτοσυνδιακύμανσης δίνεται από τις ακόλουθες σχέσεις:

$$\begin{aligned} \gamma_0 &= \sigma^2(1 + \theta_1^2 + \theta_2^2) \\ \gamma_1 &= \sigma^2(-\theta_1 + \theta_1\theta_2) \\ \gamma_2 &= \sigma^2(-\theta_2) \\ \gamma_k &= 0, \quad k > 1 \end{aligned} \quad (3.19)$$

και ομοίως η συνάρτηση αυτοσυσχέτισης ορίζεται ως:

$$\begin{aligned} \rho_1 &= \frac{-\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2} \\ \rho_2 &= \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} \\ \rho_k &= 0, \quad k > 1 \end{aligned} \quad (3.20)$$

όπου και εδώ οι τιμές της μηδενίζονται από το  $\rho_3$  και μετά.



**Διάγραμμα 3.3**

### **Διαδικασία MA(2)**

Στο Διάγραμμα 3.3 απεικονίζεται μία διαδικασία ενός υποδείγματος κινητού μέσου δευτέρου βαθμού από προσομοιωμένα δεδομένα έχουν παραχθεί από το ακόλουθο υπόδειγμα:

$$y_t = 20 + \varepsilon_t + 0.65\varepsilon_{t-1} - 0.40\varepsilon_{t-2}$$

ενώ επίσης παρουσιάζεται και το διάγραμμα των δειγματικών συντελεστών αυτοσυσχετίσεων, στο οποίο οι τιμές τους είναι σχεδόν μηδέν μετά την τιμή του συντελεστή αυτοσυσχετίσης δευτέρου βαθμού.

### **3.4 Αυτοπαλίνδρομα Υποδείγματα**

Μία άλλη κατηγορία υποδειγμάτων, εξίσου σημαντική με αυτή των υποδειγμάτων κινητού μέσου, που επίσης προέρχεται από τη γενικευμένη αναπαράσταση των γραμμικών χρονοσειρών, είναι τα αυτοπαλίνδρομα υποδείγματα (auto-regressive ή AR) τα οποία στη γενική τους μορφή δίνονται από την ακόλουθη σχέση:





$$y_t = \delta + \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t \quad (3.21)$$

όπου η ακολουθία  $\{\varepsilon_t\}$  είναι λευκός θόρυβος. Η χρονοσειρά  $y_t$  που δίνεται από τη σχέση (3.21) εκφράζει ένα αυτοπαλίνδρομο υπόδειγμα βαθμού  $p$ , που συμβολίζεται ως  $AR(p)$ , και είναι στάσιμη διαδικασία αν οι ρίζες του πολυώνυμου:

$$m^p - \varphi_1 m^{p-1} - \varphi_2 m^{p-2} - \dots - \varphi_p = 0$$

είναι μικρότερες της μονάδας σε απόλυτη τιμή. Έτσι, αν ισχύει αυτή η συνθήκη, τότε η χρονοσειρά  $y_t$  έχει μια άπειρη απολύτως αθροιστική αναπαράσταση κινητού μέσου:

$$y_t = \mu + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} = \mu + \Psi(B)\varepsilon_t, \quad (3.22)$$

όπου  $\Psi(B) = \Phi(B)^{-1}$ ,  $\Phi(B)$  το χαρακτηριστικό πολυώνυμο της αυτοπαλίνδρομης διαδικασίας και ισχύει ότι οι συντελεστές  $\psi_i$  να είναι απολύτως αθροιστικοί, δηλαδή  $\sum_{i=0}^{\infty} |\psi_i| < \infty$ , ενώ ο τελεστής (operator)  $B$  ορίζει τη χρονική υστέρηση της μεταβλητής που εκπροσωπεί, δηλαδή  $B^j = X_{t-j}$ .

Επίσης, από τη σχέση (3.23) προκύπτει ότι η μέση τιμή της χρονοσειράς ισούται με:

$$E(y_t) = \frac{\delta}{1 - \varphi_1 - \varphi_2 - \dots - \varphi_p} \quad (3.23)$$

και η αυτοσυνδιακύμανση με:

$$\begin{aligned} \gamma_k &= Cov(y_t, y_{t-k}) \\ &= Cov(\delta + \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t) \\ &= \sum_{i=1}^p \varphi_i Cov(y_{t-i}, y_{t-k}) + Cov(\varepsilon_t, y_{t-k}) \\ &= \sum_{i=1}^p \varphi_i \gamma(k-i) + \begin{cases} \sigma^2, & k = 0 \\ 0, & k > 0 \end{cases} \end{aligned} \quad (3.24)$$

ενώ για  $k = 0$  η διακύμανση είναι ίση με:

$$\gamma_0 = Var(y_t) = \sum_{i=1}^p \varphi_i \gamma(i) + \sigma^2 = \sigma^2 \quad (3.25)$$

Τέλος, από τη διαίρεση της εξίσωσης (3.24) με  $\gamma_0$  για όλες τις τιμές  $k > 0$  προκύπτει η συνάρτηση αυτοσυσχέτισης:



$$\rho_k = \sum_{i=1}^p \varphi_i \rho(k-i) \quad (3.26)$$

οι τιμές της οποίας φθίνουν εκθετικά καθώς αυξάνεται η τιμή του  $k$ .

### Αυτοπαλίνδρομο Υπόδειγμα πρώτου βαθμού

Το αυτοπαλίνδρομο υπόδειγμα πρώτου βαθμού  $AR(1)$  έχει τη ακόλουθη μορφή:

$$y_t = \delta + \varphi_1 y_{t-1} + \varepsilon_t \quad (3.27)$$

και μπορεί να θεωρηθεί ως μία παλινδρόμηση του  $y_t$  στο  $y_{t-1}$ , ενώ αν ισχύει η υπόθεση ότι  $-1 < \varphi < 1$ , τότε η χρονοσειρά  $y_t$  είναι στάσιμη. Η σχέση (3.27) μπορεί να γραφτεί στη πολυωνυμική μορφή ως:

$$(1 - \varphi_1 B)y_t = \delta + \varepsilon_t$$

όπου  $\Phi(B) = (1 - \varphi_1 B)$  είναι το πολυώνυμο του αυτοπαλίνδρου μέρους.

Έτσι, ο μέσος όρος της χρονοσειράς είναι:

$$E(y_t) = \frac{\delta}{1-\varphi} \quad (3.28)$$

και η συνάρτηση αυτοσυνδιακύμανσης δίνεται από τις ακόλουθες σχέσεις:

$$\gamma_k = \sigma^2 \varphi^k \frac{1}{1-\varphi^2} \quad (3.29)$$

για όλες τις θετικές ακέραιες τιμές του  $k$ , ενώ για  $k = 0$  προκύπτει η διακύμανση της χρονοσειράς:

$$\gamma_0 = \sigma^2 \frac{1}{1-\varphi^2}. \quad (3.30)$$

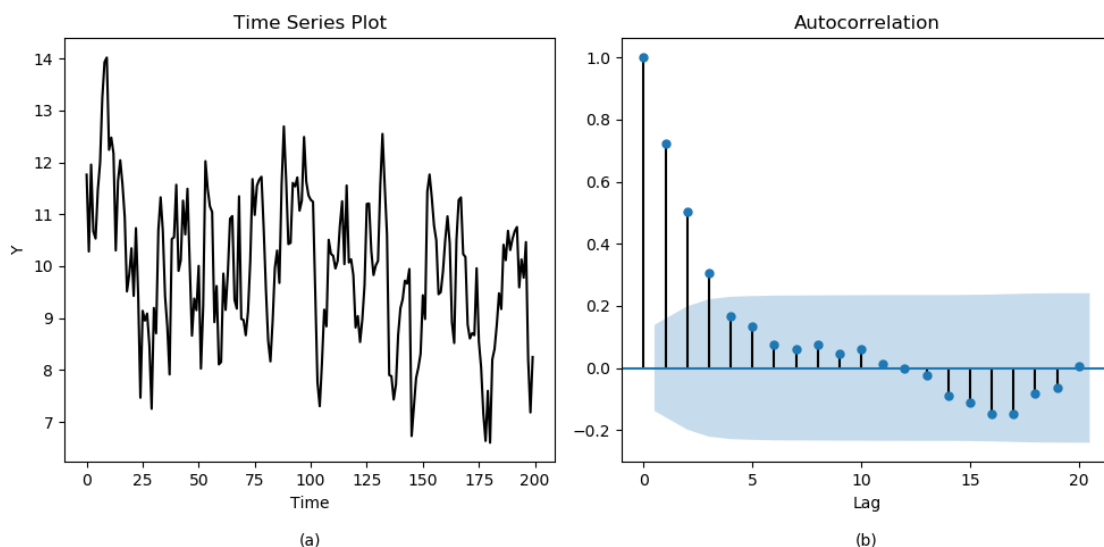
η οποία ορίζεται για τιμές της παραμέτρου  $\varphi$  μικρότερες σε απόλυτο μέγεθος της μονάδας, αναγκαία συνθήκη για να χαρακτηριστεί η χρονοσειρά ως στάσιμη.

Κατά ανάλογο τρόπο προκύπτει και η συνάρτηση αυτοσυσχέτισης ως εξής:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \varphi^k. \quad (3.31)$$



οι τιμές της οποίας φθίνουν εκθετικά αν όλες οι τιμές της είναι θετικές ή με μορφή φθίνουσας ταλάντωσης αν οι τιμές της είναι θετικές και αρνητικές.



**Διάγραμμα 3.4**

### Διαδικασία AR(1)

Στο Διάγραμμα 3.4 απεικονίζεται μία διαδικασία ενός αυτοπαλίνδρομου υποδείγματος πρώτου βαθμού από προσομοιωμένα δεδομένα έχουν παραχθεί από το ακόλουθο υπόδειγμα:

$$y_t = 10 + 0.75y_{t-1} + \varepsilon_t.$$

ενώ επίσης παρουσιάζεται και το διάγραμμα των δειγματικών συντελεστών αυτοσυσχετίσεων, στο οποίο οι τιμές τους φθίνουν εκθετικά.

### Αυτοπαλίνδρομο Υπόδειγμα δευτέρου βαθμού

Το αυτοπαλίνδρομο υπόδειγμα δευτέρου βαθμού AR(2) έχει τη ακόλουθη μορφή:

$$y_t = \delta + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varepsilon_t \quad (3.32)$$

και μπορεί να γραφτεί στην πολωνυμική μορφή ως εξής:

$$(1 - \varphi_1 B - \varphi_2 B^2)y_t = \delta + \varepsilon_t \quad (3.33)$$



όπου  $\Phi(B) = (1 - \varphi_1 B - \varphi_2 B^2)$  είναι το πολυώνυμο του αυτοπαλίνδρομου μέρους και για να είναι στάσιμη η χρονοσειρά  $y_t$  θα πρέπει να ικανοποιούνται οι κάτωθι συνθήκες για τις τιμές των παραμέτρων:

$$\varphi_1 + \varphi_2 < 1$$

$$\varphi_2 - \varphi_1 < 1$$

$$|\varphi_2| < 1$$

Έτσι, ο μέσος όρος της χρονοσειράς είναι:

$$E(y_t) = \frac{\delta}{1 - \varphi_1 - \varphi_2} \quad (3.34)$$

και η συνάρτηση αυτοσυνδιακύμανσης δίνεται από τις ακόλουθες σχέσεις:

$$\gamma_k = Cov(y_t, y_{t-k}) = \varphi_1 \gamma_{k-1} + \varphi_2 \gamma_{k-2} \quad (3.35)$$

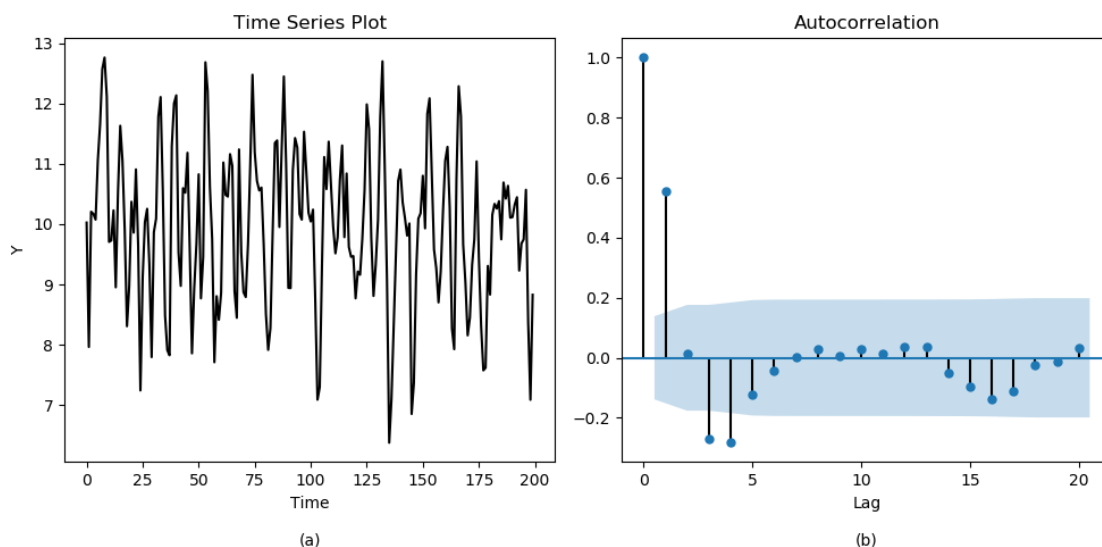
για όλες τις θετικές ακέραιες τιμές του  $k$ , ενώ για  $k = 0$  προκύπτει η διακύμανση της χρονοσειράς:

$$\gamma_0 = \varphi_1 \gamma_1 + \varphi_2 \gamma_2 + \sigma^2, \quad (3.36)$$

Από τη διαίρεση της σχέσης (3.35) με  $\gamma_0$  προκύπτει η συνάρτηση αυτοσυσχέτισης:

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} \quad (3.37)$$

οι τιμές της οποίας φθίνουν τείνοντας προς το μηδέν.



**Διάγραμμα 3.5**

### **Διαδικασία AR(2)**

Στο Διάγραμμα 3.5 απεικονίζεται μία διαδικασία ενός αυτοπαλίνδρομου υποδείγματος δευτέρου βαθμού από προσομοιωμένα δεδομένα έχουν παραχθεί από το ακόλουθο υπόδειγμα:

$$y_t = 10 + 0.75y_{t-1} - 0.40y_{t-2} + \varepsilon_t$$

ενώ επίσης παρουσιάζεται και το διάγραμμα των δειγματικών συντελεστών αυτοσυσχετίσεων, στο οποίο οι τιμές τους φθίνουν με μορφή φθίνουσας ταλάντωσης.

### **Συντελεστής μερικής αυτοσυσχέτισης**

Η δειγματικός συντελεστής αυτοσυσχέτισης είναι ένα ιδιαίτερα χρήσιμο εργαλείο στην ανάλυση υποδειγμάτων κινητού μέσου και στον καθορισμό του βαθμού  $q$  της διαδικασίας, δεδομένου ότι μετά το βαθμό του πολωνύμου όλες οι τιμές του μηδενίζονται απότομα. Αντίθετα, η συμπεριφορά του σε αυτοπαλίνδρομα υποδείγματα είναι διαφορετική και όπως αναφέρθηκε οι τιμές του μειώνονται εκθετικά τείνοντας προς το μηδέν και κατά συνέπεια αυτό το μέτρο ενδεχομένως να μην αρκεί για τον καθορισμό του βαθμού  $p$  της διαδικασίας.



Για το λόγο αυτό χρησιμοποιείται επιπρόσθετα, στην ανάλυση χρονοσειρών και ο δειγματικός συντελεστής μερικής αυτοσυσχέτισης (partial autocorrelation coefficient) που προκύπτει από τη συνάρτηση μερικής αυτοσυσχέτισης (partial autocorrelation function ή PACF). Η συνάρτηση μερικής αυτοσυσχέτισης για τις παρατηρήσεις  $y_t$  και  $y_{t-k}$  είναι απλά η αυτοσυσχέτιση τους θεωρώντας πως όλες οι άλλες οι ενδιάμεσες παρατηρήσεις  $y_{t-1}, y_{t-2}, \dots, y_{t-k+1}$  παραμένουν σταθερές και οι τιμές τους μπορούν να υπολογιστούν και από τις εξισώσεις των Yule και Walker. Πιο συγκεκριμένα, για μία στάσιμη χρονοσειρά  $y_t$  και για οποιαδήποτε σταθερή τιμή  $k$ , η συνάρτηση αυτοσυσχέτισης προκύπτει από τις εξισώσεις Yule-Walker ως:

$$\rho_j = \sum_{i=1}^k \varphi_{ik} \rho(j-i) \quad (3.38)$$

η οποία για κάθε τιμή του  $k$  γράφεται ως:

$$\begin{aligned} \rho_1 &= \varphi_{1k} + \varphi_{2k}\rho_1 + \dots + \varphi_{kk}\rho_{k-1} \\ \rho_2 &= \varphi_{1k}\rho_1 + \varphi_{2k} + \dots + \varphi_{kk}\rho_{k-2} \\ &\vdots \\ \rho_k &= \varphi_{1k}\rho_{k-1} + \varphi_{2k}\rho_{k-2} + \dots + \varphi_{kk} \end{aligned}$$

ενώ ο τελευταίος συντελεστής  $\varphi_{kk}$  προσδιορίζει την τιμή του συντελεστή μερικής αυτοσυσχέτισης.

Έτσι, σε αντίθεση με τη συμπεριφορά του συντελεστή αυτοσυσχέτισης, ο συντελεστής μερικής αυτοσυσχέτισης για κάθε διαδικασία  $AR(p)$  με  $k > p$ , λαμβάνει τιμές διάφορες του μηδενός μέχρι το βαθμό της διαδικασίας και μετά όλες οι τιμές τους μηδενίζονται απότομα. Για τις διαδικασίες κινητού μέσου όρου οι τιμές του συντελεστή μερικής συσχέτισης φθίνουν είτε εκθετικά είτε με μορφή φθίνουσας ταλάντωσης. Κατά συνέπεια, ο συντελεστής μερικής αυτοσυσχέτισης μπορεί να χρησιμοποιηθεί για τον προσδιορισμό τόσο της μορφής όσο και του βαθμού μιας διαδικασίας.



### 3.5 Μικτά Υποδείγματα

Η ανάλυση χρονοσειρών δεν επικεντρώνεται μόνο στα υποδείγματα κινητού μέσου ή στα αυτοπαλίνδρομα υποδείγματα, αλλά επεκτείνεται και στα μικτά υποδείγματα (mixed models) τα οποία αποτελούν μία γενικευμένη μορφή υποδειγμάτων στην οποία εμπεριέχονται και τα δύο είδη αυτών των υποδειγμάτων και για αυτό αποκαλούνται και ως αυτοπαλίνδρομα υποδείγματα κινητού μέσου (auto-regressive moving-average processes), γνωστά ως  $ARMA(p, q)$  υποδείγματα. Η βασική λογική αυτών των υποδειγμάτων είναι να συνδυαστούν τα χαρακτηριστικά των αυτοπαλίνδρομων υποδειγμάτων,  $AR$ , και των υποδειγμάτων κινητού μέσου,  $MA$ , έτσι ώστε να προσδιορίζεται ακόμα καλύτερα και με μεγαλύτερη ακρίβεια ο τρόπος εξέλιξης των τιμών μιας χρονοσειράς. Για το λόγο αυτό, τα καθαρά υποδείγματα κινητού μέσου ή τα καθαρά αυτοπαλίνδρομα υποδείγματα αποτελούν ειδικές περιπτώσεις των γενικευμένων  $ARMA$  υποδειγμάτων.

Η γενική μορφή ενός αυτοπαλίνδρομου υποδείγματος κινητού μέσου με βαθμούς  $(p, q)$ , που συμβολίζεται ως  $ARMA(p, q)$ , δίνεται από την ακόλουθη σχέση:

$$\begin{aligned} y_t &= \delta + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots \\ &\quad - \theta_q \varepsilon_{t-q} \\ &= \delta + \sum_{i=1}^p \varphi_i y_{t-i} + \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i}, \end{aligned} \quad (3.39)$$

όπου η ακολουθία  $\{\varepsilon_t\}$  είναι λευκός θόρυβος. Η χρονοσειρά  $y_t$  είναι στάσιμη και αντιστρέψιμη αν οι ρίζες του πολυώνυμου  $m^p - \varphi_1 m^{p-1} - \varphi_2 m^{p-2} - \dots - \varphi_p = 0$  είναι μικρότερες της μονάδας σε απόλυτη τιμή και αντίστοιχα αν οι ρίζες του πολυώνυμου  $m^q - \theta_1 m^{q-1} - \theta_2 m^{q-2} - \dots - \theta_q = 0$  είναι μικρότερες της μονάδας σε απόλυτη τιμή. Τα χαρακτηριστικά αυτά προσδίδουν καλές ιδιότητες στη χρονοσειρά τόσο στην αναγνώριση του τρόπου συμπεριφοράς της όσο και στον σχηματισμό προβλέψεων.

Η απλούστερη μορφή ενός αυτοπαλίνδρομου υποδείγματος κινητού μέσου δημιουργείται για  $p = 1$  και  $q = 1$  και ονομάζεται  $ARMA(1, 1)$  υπόδειγμα έχοντας την ακόλουθη μορφή:



$$y_t = \delta + \varphi_1 y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (3.40)$$

όπου η ακολουθία  $\{\varepsilon_t\}$  είναι λευκός θόρυβος. Από το υπόδειγμα αυτό προκύπτει ότι η μέση τιμή και η διακύμανση της εν λόγω χρονοσειράς είναι αντίστοιχα:

$$E(y_t) = \frac{\delta}{1-\varphi} \quad (3.41)$$

$$\gamma_0 = Var(y_t) = \sigma^2 \frac{1+2\varphi\theta+\theta^2}{1-\varphi^2}. \quad (3.42)$$

ενώ οι συναρτήσεις αυτοσυνδιακύμανσης ορίζονται ως:

$$\gamma_1 = Cov(y_t, y_{t-1}) = \sigma^2 \frac{(\varphi+\theta)(1+\varphi\theta)}{1-\varphi^2} \quad (3.43)$$

$$\gamma_k = \varphi \gamma_{k-1}, \quad k \geq 2 \quad (3.44)$$

από τις οποίες προκύπτει η συνάρτηση αυτοσυσχέτισης:

$$\rho_k = \varphi \rho_{k-1} = \rho_1 \varphi^{k-1} \quad (3.45)$$

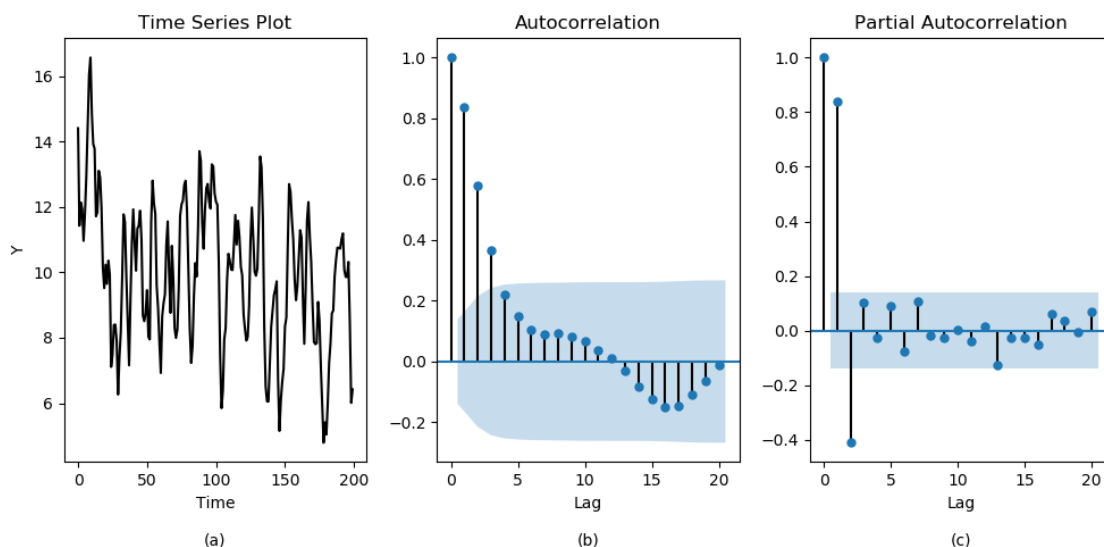
για  $k \geq 2$ , οι τιμές της οποίας τείνουν στο μηδέν καθώς αυξάνεται η τιμή του  $k$ , από την αρχική τιμή  $\rho_1$  που είναι ίση με:

$$\rho_1 = \frac{(1-\varphi\theta)(\varphi-\theta)}{1-2\varphi\theta+\theta^2} \quad (3.46)$$

και εξαρτάται από τις τιμές των παραμέτρων  $\varphi$  και  $\theta$ .

Σημειώνεται ότι κατά ανάλογο αλλά με περισσότερο πολύπλοκο τρόπο υπολογίζονται και οι αυτοσυσχετίσεις μικτών υποδειγμάτων μεγαλύτερου βαθμού. Η γενική εικόνα της συμπεριφοράς τους βασίζεται στη αρχή ότι οι συντελεστές αυτοσυσχέτισης ακολουθούν το αυτοπαλίνδρομο υπόδειγμα μόλις περάσουν το βαθμό του πολυωνύμου του κινητού μέσου, ενώ οι τιμές των παραμέτρων του κινητού μέσου παίζουν ρόλο μόνο στον υπολογισμό των αρχικών τιμών των συντελεστών αυτοσυσχέτισης μέχρι το βαθμό του πολυωνύμου τους.





**Διάγραμμα 3.6**

**Διαδικασία ARMA(1, 1)**

Στο Διάγραμμα 3.6 απεικονίζεται μία διαδικασία ενός αυτοπαλίνδρομου υποδείγματος κινητού μέσου  $ARMA(1, 1)$ , όπου τα προσομοιωμένα δεδομένα έχουν παραχθεί από το ακόλουθο υπόδειγμα:

$$y_t = 10 + 0.75y_{t-1} + \varepsilon_t - 0.65\varepsilon_{t-1}$$

για τιμές δηλαδή 0.75 και 0.65 του αυτοπαλίνδρομου και του κινητού μέσου όρου αντίστοιχα. Επίσης, στο διάγραμμα αυτό απεικονίζονται τα κορελογράμματα των δειγματικών συντελεστών αυτοσυσχέτισης, καθώς και των δειγματικών συντελεστών μερικής αυτοσυσχέτισης.

Σημειώνεται ότι η εκτίμηση  $ARMA$  υποδειγμάτων στην ανάλυση χρονοσειρών είναι μία ιδιαίτερη διαδικασία, σε αντίθεση με την εκτίμηση γραμμικών υποδειγμάτων στην ανάλυση παλινδρόμησης. Στη βιβλιογραφία υπάρχουν τρεις μέθοδοι εκτίμησης: α) η υπό-συνθήκη μέθοδος των ελαχίστων τετραγώνων CLS (Conditional Least Squares method), η οποία προσδιορίζει τους εκτιμητές με βάση την ελαχιστοποίηση του αθροίσματος των τετραγώνων των σφαλμάτων, β) η μέθοδος της μέγιστης πιθανοφάνειας MLE (Maximum Likelihood Estimation method), η οποία προσδιορίζει τους εκτιμητές με βάση την μεγιστοποίηση τη λογαριθμικής συνάρτησης



πιθανοφάνειας και γ) η χωρίς συνθήκη μέθοδος των ελαχίστων τετραγώνων ULS (Unconditional Least Squares method), η οποία προσδιορίζει τους εκτιμητές με βάση την ελαχιστοποίηση του αθροίσματος των τετραγώνων των σφαλμάτων πηγαίνοντας προς τα οπίσω χρονικά τις τιμές τους, για αυτό και αποκαλείται μέθοδος Backcasting, με αποτέλεσμα να δημιουργούνται διαφορετικές εκτιμήσεις των παραμέτρων για το ίδιο υπόδειγμα, όπως έδειξαν στην εργασία τους οι Newbold, Agiakloglou και Mills (1994). Από τις τρεις προαναφερθείσες τεχνικές εκτίμηση οι Newbold και Ansley (1980) έδειξαν ότι η καλύτερη είναι η μέθοδος της μεγίστης πιθανοφάνειας και αυτή πρότειναν να χρησιμοποιείται.

### 3.6 Μη στάσιμες χρονοσειρές

Μία άλλη κατηγορία χρονοσειρών είναι οι μη-στάσιμες διαδικασίες (non-stationary processes), οι οποίες προκύπτουν από τις στάσιμες διαδικασίες όταν παραβιάζεται μία από τις βασικές τους υποθέσεις. Έτσι όταν η τιμή του μέσου όρου ή της διακύμανσης της χρονοσειράς δεν παραμένουν σταθερές διαχρονικά ή ακόμα όταν η αυτοσυνδιακύμανση των τιμών της χρονοσειράς δεν καθορίζεται από την απόστασή τους, τότε εμφανίζονται διαδικασίες που δεν έχουν καλές στατιστικές ιδιότητες για να αναλυθούν και να προσδιοριστεί ο τρόπος εξέλιξης των τιμών τους. Ειδικότερα, οι μη-στάσιμες χρονοσειρές προέρχονται από διαδικασίες που παρουσιάζουν έντονη μεταβλητότητα και δεν έχουν ομοιογενή συμπεριφορά με την πάροδο του χρόνου. Για το λόγο αυτό απαιτείται ο μετασχηματισμός της χρονοσειράς σε στάσιμη διαδικασία, κάτι το οποίο συνήθως γίνεται με τη χρήση διαφορών στις τιμές της χρονοσειράς, έτσι ώστε να αποκτηθούν οι καλές στατιστικές ιδιότητες των στάσιμων χρονοσειρών και να αναλυθεί η χρονοσειρά σε πρώτες ή άλλες διαφορές.

Συγκεκριμένα, μία χρονοσειρά  $y_t$  είναι σε διαφορές στάσιμη (difference stationary) αν δεν είναι στάσιμη στις αρχικές τιμές της, αλλά μπορεί να μετατραπεί σε στάσιμη με τη χρήση πρώτων ή άλλων διαφορών. Οι στάσιμες σε διαφορές χρονοσειρές ονομάζονται



διαφοροποιημένες ή ολοκληρωμένες χρονοσειρές (integrated processes), ενώ ο αριθμός των φορών που χρειάζεται να εφαρμοστούν οι πρώτες διαφορές στις τιμές της χρονοσειράς, ώστε να μετατραπεί σε στάσιμη, ονομάζεται βαθμός ολοκλήρωσης και συμβολίζεται με  $d$ . Σημειώνεται ότι η προσπάθεια αυτή αποσκοπεί στην εμπειρική εξάλειψη της μη σταθερότητας της τιμής του μέσου όρου αλλά και της διακύμανσης, ενώ θα μπορούσε να εφαρμοζόταν, ειδικά για τη διακύμανση και λογαριθμικός μετασχηματισμός των τιμών της χρονοσειράς προκειμένου να μικρύνει τις μεταβολές των τιμών τη χρονοσειράς, πάντα με την προϋπόθεση ότι οι τιμές της λαμβάνουν μόνο θετικές τιμές.

Επιπρόσθετα, εάν η ολοκληρωμένη χρονοσειρά  $y_t$ , με βαθμό ολοκλήρωσης  $d$ , δημιουργεί μία στάσιμη  $ARMA(p, q)$  διαδικασία, τότε μπορεί να περιγραφεί από ένα ολοκληρωμένο αυτοπαλίνδρομο υπόδειγμα κινητού μέσου (auto-regressive integrated moving-average model) που συμβολίζεται ως  $ARIMA(p, d, q)$  και ορίζεται ως:

$$\Phi(B)(1 - B)^d y_t = \delta + \Theta(B)\varepsilon_t, \quad (3.47)$$

όπου  $\Phi(B)$  και  $\Theta(B)$  είναι τα πολυώνυμα του αυτοπαλίνδρου και κινητού μέσου αντίστοιχα, ενώ για  $d = 1$ , τότε  $(1 - B)y_t = y_t - y_{t-1}$ , δηλαδή σχηματίστηκε χρονοσειρά σε πρώτες διαφορές.

Το απλούστερο υπόδειγμα σε πρώτες διαφορές στάσιμης στοχαστικής διαδικασίας είναι ο τυχαίος περίπατος (random walk) που συμβολίζεται ως  $ARIMA(0,1,0)$  και γράφεται ως:

$$(1 - B)y_t = \delta + \varepsilon_t \quad (3.48)$$

όπου  $\delta$  είναι ο σταθερός όρος ή η παράμετρος περιπλάνησης (drift) και  $\{\varepsilon_t\}$  είναι ο λευκός θόρυβος. Το υπόδειγμα αυτό για μηδενικό βαθμό ολοκλήρωσης εξελίσσεται σύμφωνα με τη σχέση:

$$y_t = \delta + y_{t-1} + \varepsilon_t$$

ενώ λύνοντας ως προς  $y_t$  προκύπτει η ακόλουθη σχέση:



$$y_t = y_0 + \delta t + \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t$$

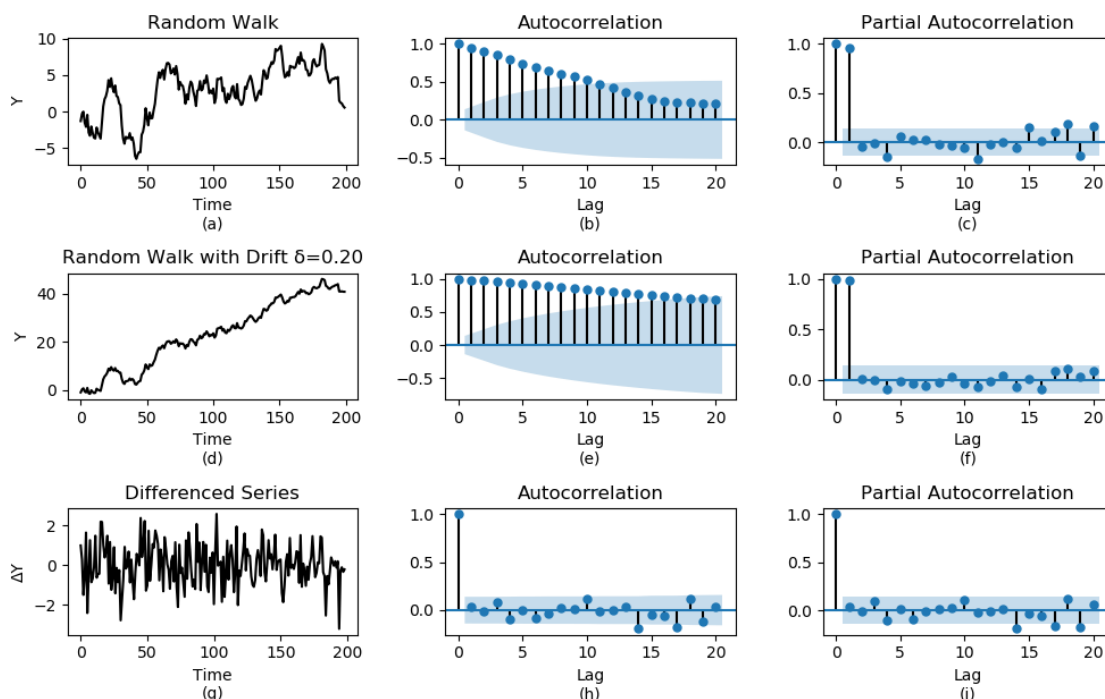
όπου  $y_0$  είναι η αρχική τιμή της χρονοσειράς και κατά συνέπεια η υπό-συνθήκη αναμενόμενη τιμή και διακύμανση της χρονοσειράς  $y_t$  είναι αντίστοιχα:

$$E(y_t|y_0) = y_0 + \delta t$$

$$Var(y_t|y_0) = \sigma^2 t$$

από όπου είναι προφανές ότι ο τυχαίος περίπατος είναι υπόδειγμα μη-στάσιμης διαδικασίας δεδομένου ότι τόσο ο μέσος όρος όσο και η διακύμανσή του δεν είναι σταθερά, αλλά οι τιμές τους μεταβάλλονται στο χρόνο.

Αξίζεις πάντως να αναφερθεί ότι το υπόδειγμα αυτό είναι σημείο αναφοράς για όλες τις μορφές μη στάσιμων χρονοσειρών, όπου κύριο χαρακτηριστικό τους αποτελεί το γεγονός ότι οι τιμές των συντελεστών αυτοσυσχέτισης δεν τείνουν στο μηδέν αλλά παραμένουν πολύ κοντά στην περιοχή της μονάδας και σβήνουν πολύ αργά. Μάλιστα, το υπόδειγμα της σχέσης (3.48) ονομάζεται τυχαίος περίπατος χωρίς περιπλάνηση (random walk without drift), εάν η παράμετρος περιπλάνησης (drift)  $\delta$  λάβει την τιμή μηδέν.



**Διάγραμμα 3.7**

### **Διαδικασία Τυχαίου Περιπάτου**

Τέλος, στο Διάγραμμα 3.7 παρουσιάζονται ενδεικτικά διαγραμματικές απεικονίσεις των υποδειγμάτων: α) τυχαίος περίπατος χωρίς περιπλάνηση, β) τυχαίος περίπατος με περιπλάνηση ίση με 0.20 και γ) υπόδειγμα σε πρώτες διαφορές, που έχουν παραχθεί από προσομοιωμένα δεδομένα μαζί με τα κορελογράμματα των δειγματικών συντελεστών αυτοσυσχετίσεων και των μερικών αυτοσυσχετίσεων. Από το διάγραμμα αυτό παρατηρείται ότι οι τιμές των δειγματικών συντελεστών αυτοσυσχετίσης για τα πρώτα δύο μη-στάσιμα υποδείγματα δεν τείνουν προς το μηδέν αλλά παραμένουν υψηλές και μάλιστα κοντά στη μονάδα, ενώ το κορελόγραμμα των δειγματικών συντελεστών αυτοσυσχετίσης για το στάσιμο υπόδειγμα σε πρώτες διαφορές ομοιάζει με αυτό του λευκού θορύβου.

Η διάκριση μεταξύ στάσιμων και μη-στάσιμων χρονοσειρών βασίζεται στην ύπαρξη μοναδιαίας αυτοσυσχετιζόμενης ρίζας (unit autoregressive root). Για παράδειγμα, ένα αυτοπαλίνδρομο υπόδειγμα πρώτου βαθμού είναι στάσιμη διαδικασία αν η παράμετρός του λαμβάνει τιμές μικρότερες σε απόλυτο μέγεθος της μονάδας, ενώ αν η παράμετρός



του λάβει την τιμή ένα, τότε η διαδικασία αυτή γίνεται μία μη-στάσιμη διαδικασία τυχαίου περιπάτου. Έτσι, όταν μία χρονοσειρά έχει μία μοναδιαία ρίζα, τότε ονομάζεται ολοκληρωμένη διαδικασία πρώτου βαθμού και συμβολίζεται ως  $I(1)$ , ενώ αν η χρονοσειρά είναι στάσιμη τότε είναι μία ολοκληρωμένη διαδικασία μηδενικού βαθμού  $I(0)$ . Σε περίπτωση όπου μία διαδικασία έχει περισσότερες από μία μοναδιαίες ρίζες, τότε η διαδικασία αυτή ονομάζεται ολοκληρωμένη διαδικασία  $d$  βαθμού και συμβολίζεται ως  $I(d)$ , ενώ για την μετατροπή της σε στάσιμη διαδικασία θα χρειαστούν να ληφθούν  $d$  φορές πρώτες διαφορές στις τιμές της, όπου  $d$  θετικός ακέραιος αριθμός.

Θα πρέπει ωστόσο να σημειωθεί ότι η διάκριση μεταξύ στάσιμων και μη-στάσιμων χρονοσειρών δεν είναι εύκολη διαδικασία, όπως επίσης, σε γενικότερο πλαίσιο, εύκολη δεν είναι και η εύρεση του καταλληλότερου υποδείγματος. Σύμφωνα με τη μεθοδολογία των Box και Jenkins (1970) η ανάλυση των χρονοσειρών ξεκινάει με το στάδιο της αναγνώρισης (identification stage), σύμφωνα με το οποίο ο ερευνητής καλείται να αναγνωρίσει και να επιλέξει το καταλληλότερο υπόδειγμα με βάση τα κορελογράμματα των δειγματικών απλών και μερικών συντελεστών αυτοσυσχέτισης στηριζόμενος στην αρχή της παρρησίας (the principal of parsimony), ότι δηλαδή ένα μικρό σε αριθμό παραμέτρων υπόδειγμα είναι πάντα προτιμότερο από ένα μεγαλύτερο υπόδειγμα. Η διαδικασία αυτή είναι αρκετά ιδιαίτερη διότι απαιτεί καλή γνώση και εμπειρία από τον αναλυτή για υποδείγματα χρονοσειρών. Για το λόγο αυτό, οι Newbold, Agiakloglou και Mills (1993) πρότειναν μία τεχνική στην οποία η επιλογή του καταλληλότερου υποδείγματος προέρχεται από την ελαχιστοποίηση της τιμής κάποιου πληροφοριακού κριτηρίου, όπως το Akaike Information Criterion (AIC) που προτάθηκε από τον Akaike (1974):

$$AIC = \ln(\hat{\sigma}^2) + \frac{2m}{T}$$

και το Bayesian Information Criterion (BIC) που προτάθηκε από τον Schwarz (1978):

$$BIC = \ln(\hat{\sigma}^2) + \frac{\ln(T) m}{T}$$



όπου  $m$  ο αριθμός των παραμέτρων που χρησιμοποιούνται για την εκτίμηση του υποδείγματος. Έτσι, απαιτείται εκτίμηση διαφόρων υποδειγμάτων για να επιλεγεί το καταλληλότερο, μία διαδικασία όπου και αυτή θέλει προσοχή διότι τα κριτήρια δεν επιλέγουν πάντα αυτό το σωστό υπόδειγμα. Σημειώνεται ότι το AIC έχει την τάση να επιλέγει μεγάλο σε αριθμό παραμέτρων υπόδειγμα, ενώ το BIC επιλέγει μικρό σε αριθμό παραμέτρων υπόδειγμα.

Ιδιαίτερη όμως αναφορά στη βιβλιογραφία γίνεται για τον έλεγχο της μοναδιαίας ρίζας (unit root test) ως μία προσπάθεια να αποσαφηνιστεί αν η χρονοσειρά είναι στάσιμη ή όχι. Ο έλεγχος αυτός παρουσιάστηκε από τους Dickey και Fuller (1979) και για αυτό ονομάστηκε και έλεγχος των Dickey και Fuller (Dickey and Fuller test) στηριζόμενος στην εκτίμηση με τη μέθοδο των ελαχίστων τετραγώνων σε ένα από τα ακόλουθα υποδείγματα:

$$y_t = \beta y_{t-1} + \varepsilon_t$$

$$y_t = \alpha + \beta y_{t-1} + \varepsilon_t$$

$$y_t = \alpha + \beta y_{t-1} + \gamma t + \varepsilon_t$$

δηλαδή με ή χωρίς σταθερό όρο ή και με τάση, όπου  $\{\varepsilon_t\}$  είναι το τυχαίο σφάλμα, προκειμένου να διερευνηθεί η μηδενική υπόθεση  $H_0: \beta = 1$ , η οποία υποστηρίζει την ύπαρξη μοναδιαίας ρίζας, έναντι της εναλλακτικής υπόθεσης  $H_1: \beta < 1$  που την απορρίπτει. Ο έλεγχος εφαρμόζεται με δύο στατιστικές  $\hat{t}$  και  $\hat{\rho}$  που ανέπτυξαν οι Dickey και Fuller:

$$\hat{t} = (\hat{\beta} - 1)/se(\hat{\beta}) \quad \text{και} \quad \hat{\rho} = T(\hat{\beta} - 1)$$

όπου κριτικές τιμές για την υλοποίηση του ελέγχου σε κάθε μία από τις παραπάνω μορφές υποδείγματος υπάρχουν στο βιβλίο του Fuller (1976 και 1996). Στην πραγματικότητα όμως ο έλεγχος αυτός, παρόλο που αποκαλείται έλεγχος μοναδιαίας ρίζας, είναι έλεγχος ενός υποδείγματος τυχαίου περιπάτου με ή χωρίς περιπλάνηση ή και με τάση, όπου υπάρχει μοναδιαία ρίζα, έναντι ενός αυτοπαλίνδρομου υποδείγματος πρώτου βαθμού, όπου δεν υπάρχει μοναδιαία ρίζα.



Κατόπιν, οι Dickey και Fuller (1981) ανέπτυξαν τον επαυξημένο έλεγχο των Dickey και Fuller (Augmented Dickey and Fuller test), γνωστό και ως ADF test, προκειμένου να διερευνήσουν αν υπάρχει μοναδιαία ρίζα σε μεγαλύτερο βαθμό αυτοπαλίνδρομου υποδείγματος. Πιο συγκεκριμένα, ο έλεγχος βασίζεται στην υπόθεση ότι υπάρχει μοναδιαία ρίζα σε μία διαδικασία  $ARIMA(p, 1, 0)$ , όπου ο βαθμός  $p$  είναι γνωστός, έναντι της εναλλακτικής ότι δεν υπάρχει, δηλαδή σε μια διαδικασία της μορφής ενός υποδείγματος  $ARIMA(p+1, 0, 0)$  και στηρίζεται στην εκτίμηση του ακόλουθου υποδείγματος:

$$y_t = \alpha + \beta y_{t-1} + \sum_{i=1}^k \varphi_i \Delta y_{t-i} + \varepsilon_t \quad (3.49)$$

για να διερευνηθεί η μηδενική υπόθεση ότι  $\beta = 1$  έναντι της εναλλακτικής υπόθεσης ότι  $\beta < 1$  χρησιμοποιώντας μία από τις δύο προαναφερθείσες στατιστικές  $\hat{t}$  και  $\hat{\rho}$  που ανέπτυξαν οι Dickey και Fuller, λαμβάνοντας κριτικές τιμές από το βιβλίο του Fuller (1976 και 1996), όπου  $k = p$ . Έτσι, αν η μηδενική υπόθεση γίνει δεκτή τότε το υπόδειγμα της σχέσης (3.49) εκφράζει μία διαδικασία  $ARIMA(p, 1, 0)$ .

Ωστόσο και αυτός ο στατιστικός έλεγχος δεν αναφερόταν στη γενική μορφή ελέγχου μοναδιαίας ρίζας, αλλά σε αυτοπαλίνδρομο υπόδειγμα σε πρώτες διαφορές γνωστού βαθμού. Για το λόγο αυτό, οι Said και Dickey (1984) ανέπτυξαν τον έλεγχο μοναδιαίας ρίζας για κάθε διαδικασία  $ARIMA(p, 1, q)$ , ο οποίος όπως απέδειξαν εφαρμόζεται με βάση το υπόδειγμα της σχέσης (3.49) και αποκαλείται και αυτός ADF test με τη μόνη διαφορά ότι μόνο η στατιστική  $\hat{t}$  μπορεί να χρησιμοποιηθεί για την εφαρμογή του ελέγχου διότι η συμπεριφορά της στατιστικής  $\hat{\rho}$  δεν ήταν εύκολο να προσδιοριστεί αφού εξαρτώνταν από τις άγνωστες τιμές των παραμέτρων του κινητού μέσου. Επιπρόσθετα, έρευνες έδειξαν ότι ο έλεγχος αυτός έχει ευαισθησία στην επιλογή του αριθμού των χρονικών υστερήσεων με τους Agiakloglou και Newbold (1992) να προτείνουν το Akaike Information Criterion (AIC) για την επιλογή του  $k$ , επειδή έχει καλύτερες ασυμπτωτικές ιδιότητες, ενώ οι Ng και Perron (1995) να προτείνουν τεχνικές εκτίμησης από μία πολύ μικρή τιμή σε μία πολύ μεγάλη τιμή του  $k$  ή και το αντίστροφο (specific to general or general to specific) με τη δεύτερη να έχει καλύτερη συμπεριφορά, κοντά στην επιλογή του  $k$  με το AIC. Σημειώνεται ότι όλες οι προσομοιώσεις έγιναν για ένα υπόδειγμα  $ARIMA(0, 1, 1)$ , ενώ οι Agiakloglou και



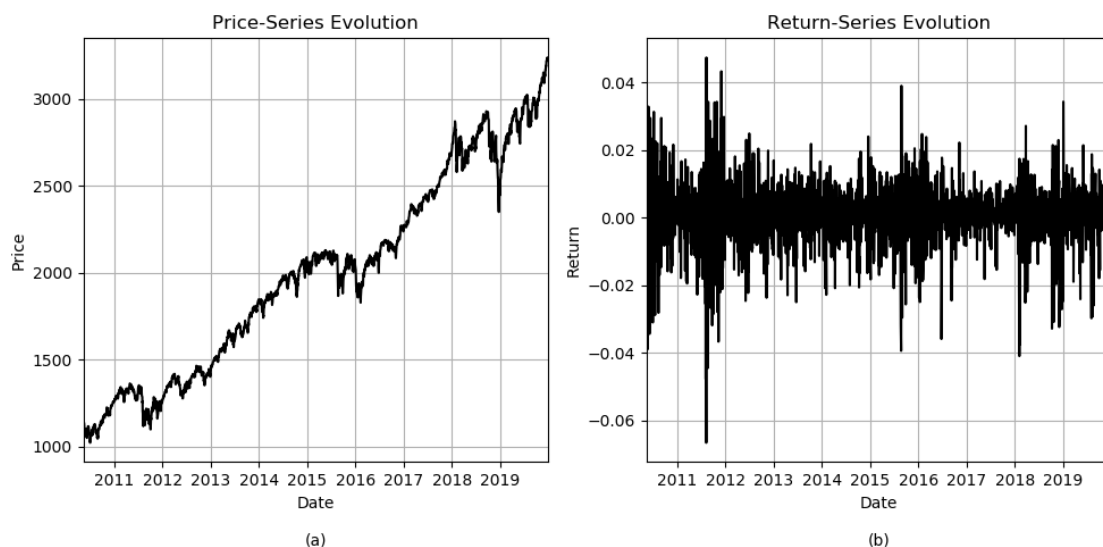


Newbold (1992) έδειξαν ότι η δύναμη αυτού του ελέγχου δεν είναι ικανοποιητική ακόμα και για το απλό αυτοπαλίνδρομο υπόδειγμα πρώτου βαθμού.

Αξίζει τέλος να αναφερθεί ότι υπάρχει και ο έλεγχος των Phillips και Perron (1988) οι οποίοι πρότειναν μια ημι-παραμετρική μέθοδο για τον έλεγχο της ύπαρξης μοναδιαίας ρίζας σε μία χρονοσειρά  $y_t$ . Σε αντίθεση μάλιστα με τον έλεγχο  $ADF$ , ο έλεγχος των Phillips και Perron (PP) προσπαθεί να διορθώσει το πρόβλημα της σειριακής συσχέτισης των καταλοίπων χρησιμοποιώντας μη-παραμετρικές εκτιμήσεις της μακροπρόθεσμης διακύμανσης. Με αυτό το τρόπο αντιμετωπίζουν άμεσα τη σειριακή συσχέτιση τροποποιώντας τις στατιστικές του ελέγχου. Ωστόσο, οι Leybourne και Newbold (1999) έδειξαν ότι αυτός ο έλεγχος έχει περισσότερα προβλήματα από ότι ο  $ADF$  ακόμα και όταν οι πραγματικές τιμές των διακυμάνσεων χρησιμοποιηθούν αντί των δειγματικών τους εκτιμήσεων.

### **3.7 Αυτοπαλίνδρομα Υποδείγματα υπό Συνθήκης Ετεροσκεδαστικότητας**

Τα αυτοπαλίνδρομα κινητού μέσου υποδείγματα ( $ARMA$ ) είναι στοχαστικές διαδικασίες οι τιμές των οποίων είναι εκφρασμένες ως συνάρτηση των προηγούμενων χρονικών παρατηρήσεων τους, καθώς και των τιμών του τυχαίου σφάλματος. Η βασικότερη παραδοχή που απαιτείται για να εφαρμοστούν τα υποδείγματα αυτά σε κάποιο σύνολο δεδομένων είναι ότι το τυχαίο σφάλμα ακολουθεί τη διαδικασία του λευκού θορύβου, δηλαδή έχει, μεταξύ άλλων, σταθερή διακύμανση. Ωστόσο, πολύ συχνά, η υπόθεση αυτή φαίνεται ασυμβίβαστη με την ύπαρξη ετεροσκεδαστικότητας που παρατηρείται σε συμπεριφορές οικονομικών αλλά κυρίως χρηματοοικονομικών δεδομένων.



**Διάγραμμα 3.8**

**Εξέλιξη Τιμών-Αποδόσεων Δείκτη S&P 500**

Στο Διάγραμμα 3.8 (α και β) απεικονίζονται η ημερήσια διαχρονική εξέλιξη των τιμών του δείκτη S&P 500 και η ημερήσια διαχρονική εξέλιξη των αποδόσεων του από τις 01/01/2010 έως τις 31/12/2019.<sup>5</sup> Όπως παρατηρείται από το διάγραμμα αυτό οι τιμές του δείκτη έχουν μία ανοδική πορεία για την περίοδο αυτή, ενώ οι αποδόσεις του δείκτη μεταβάλλονται γύρω από μία μέση τιμή, παρατηρώντας ότι η μεταβλητότητα των αποδόσεων δεν είναι σταθερή, αφού παρουσιάζονται συστάδες μεταβλητότητας (volatility clustering), κυρίως ύστερα από απότομες διορθώσεις του δείκτη. Το γεγονός αυτό, το οποίο μάλιστα χαρακτηρίζει τη συμπεριφορά των αποδόσεων των μετοχών γενικότερα, δηλώνει ότι η διακύμανση των αποδόσεων δεν θα πρέπει να είναι σταθερή διαχρονικά.

Για τη μελέτη αλλά και την κατανόηση τέτοιων συμπεριφορών ο Engle (1982) ανέπτυξε υποδείγματα που επιτρέπουν διακυμάνσεις στην μεταβλητότητα και κατά συνέπεια η διακύμανση δεν είναι σταθερή αλλά μεταβάλλεται διαχρονικά, γνωστά ως αυτοπαλίνδρομα υπό συνθήκης ετεροσκεδαστικότητας υποδείγματα (autoregressive conditional heteroscedastic models), που συμβολίζονται ως ARCH. Στη συνέχεια, ο

<sup>5</sup> Τα δεδομένα αντλήθηκαν από <https://finance.yahoo.com>.



Bollerslev (1986) παρουσίασε τα γενικευμένα αυτοπαλίνδρομα υπό συνθήκης ετεροσκεδαστικότητας υποδείγματα (generalized autoregressive conditional heteroscedastic models), που συμβολίζονται ως GARCH, και τα οποία θεωρούνται ως προέκταση των υποδειγμάτων ARCH όπου αντιλαμβάνονται με ένα πιο περίπλοκο τρόπο τη μεταβλητότητα της διακύμανσης μιας χρονοσειράς. Σε γενικές γραμμές τα υποδείγματα αυτά μπορούν να περιγράψουν ικανοποιητικά χρονοσειρές που παρουσιάζουν χαρακτηριστικά συστάδων μεταβλητότητας και έχουν βαριές ουρές (heavy tails), ενώ στη βιβλιογραφία υπάρχουν και άλλες πολλές επεκτάσεις αυτών των υποδειγμάτων.

### **Υπόδειγμα ARCH**

Το αυτοπαλίνδρομο υπό συνθήκης ετεροσκεδαστικότητας υπόδειγμα θεωρεί ότι η χωρίς συνθήκη διακύμανση (unconditional variance) των τυχαίων σφαλμάτων είναι σταθερή, ενώ η υπό συνθήκη διακύμανση (conditional variance) δεν είναι σταθερή και μεταβάλλεται σε συνάρτηση με τα παρελθοντικές τιμές του τυχαίου σφάλματος. Έτσι, αν  $\sigma_t^2 = Var[\varepsilon_t|F_{t-1}]$  είναι η υπό συνθήκη διακύμανση της διαδικασίας  $\varepsilon_t$  και  $F_{t-1}$  είναι η διαθέσιμη πληροφορία τη χρονική στιγμή  $t - 1$ , τότε η γενική μορφή ενός αυτοπαλίνδρομου υποδείγματος υπό συνθήκης ετεροσκεδαστικότητας με βαθμό  $s$ , που συμβολίζεται ως ARCH( $s$ ), είναι:

$$\varepsilon_t = \sigma_t e_t, \quad (3.50)$$

όπου  $\{e_t\}$  είναι μία ακολουθία iid τυχαίων μεταβλητών με μέση τιμή μηδέν και διακύμανση ίση με τη μονάδα, ενώ η υπό συνθήκη διακύμανση περιγράφεται από την εξίσωση:

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + \dots + a_s \varepsilon_{t-s}^2 \quad (3.51)$$

με τους περιορισμούς  $a_0 > 0$ ,  $a_i \geq 0$  για  $i = 1, 2, \dots, s - 1$  και  $a_s > 0$  ώστε να εξασφαλίζεται το θετικό πρόσημο της υπό συνθήκη διακύμανσης και επιπλέον, πρέπει να ισχύει ότι  $\sum_{i=1}^s a_i < 1$  για να είναι η διαδικασία  $\varepsilon_t$  στάσιμη συνδιακύμανσης με πεπερασμένη μόνιμη διακύμανση  $\sigma_\varepsilon^2$ . Από τη δομή του υποδείγματος, φαίνεται ότι τα τετράγωνα των παρελθοντικών σφαλμάτων (αιφνίδιες ταραχές ή shocks ή innovations) υπονοούν την ύπαρξη μεγάλης υπό συνθήκη διακύμανσης με αποτέλεσμα το τωρινό



σφάλμα να έχει μεγάλο συντελεστή. Ως εκ τούτου, στα πλαίσια ενός υποδείγματος *ARCH*, μεγάλες αιφνίδιες ταραχές είθισται να ακολουθούνται από νέες μεγάλες αιφνίδιες ταραχές.

Το απλούστερο υπόδειγμα *ARCH* είναι αυτό του πρώτου βαθμού, που συμβολίζεται ως *ARCH*(1) και περιγράφεται από την εξίσωση:

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 \quad (3.52)$$

όπου ισχύει ότι  $a_0 > 0$  και  $0 < a_1 < 1$ . Ο μέσος όρος της διαδικασίας  $\varepsilon_t$  είναι μηδενικός, ενώ η χωρίς συνθήκη διακύμανσή της ισούται με:

$$Var(\varepsilon_t) = \sigma_\varepsilon^2 = \frac{1}{1-a_1} \quad (3.53)$$

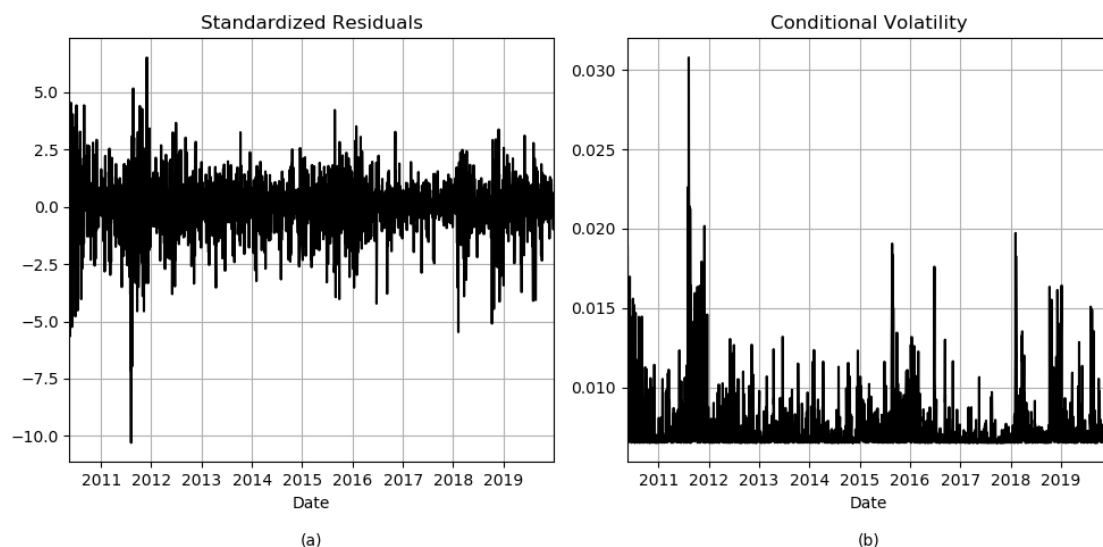
όπου, αν γίνει αντικατάσταση της σχέσης (3.53) στην εξίσωση (3.52), τότε προκύπτει η εξίσωση:

$$\sigma_t^2 = \sigma_\varepsilon^2 + a_1(\varepsilon_{t-1}^2 - \sigma_\varepsilon^2) \quad (3.54)$$

για την οποία η υπό συνθήκη διακύμανση θα είναι μεγαλύτερη της χωρίς συνθήκης διακύμανσης μόνο όταν  $\varepsilon_{t-1}^2 > \sigma_\varepsilon^2$ . Επιπρόσθετα, αν οριστεί ότι  $u_t = \varepsilon_t^2 - \sigma_t^2$ , τότε  $\varepsilon_t^2 = \sigma_t^2 + u_t$  και η τυχαία μεταβλητή  $u_t$  έχει μηδενικό μέσο όρο και οι τιμές της είναι σειριακά ασυσχέτιστες. Επομένως, κάνοντας αντικατάσταση στην εξίσωση (3.52), το υπόδειγμα *ARCH*(1) μπορεί να γραφεί ως:

$$\varepsilon_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + u_t \quad (3.55)$$

όπου η σχέση αυτή αποκαλύπτει πως τα τετράγωνα των σφαλμάτων είναι μία *AR*(1) διαδικασία με ασυσχέτιστες αιφνίδιες διαταραχές  $u_t$ .



**Διάγραμμα 3.9**

### **Εφαρμογή Υποδείγματος ARCH(1)**

Στο Διάγραμμα 3.9 (a και b) απεικονίζονται τα κανονικοποιημένα κατάλοιπα (standardized residuals) και αυτά που προκύπτουν από την εφαρμογή ενός  $ARCH(1)$  υποδείγματος για τα δεδομένα των ημερήσιων αποδόσεων του δείκτη  $S\&P\ 500$  που παρουσιάστηκαν στο Διάγραμμα 3.8. Από το διάγραμμα αυτό διαπιστώνεται πως η εφαρμογή ενός  $ARCH(1)$  υποδείγματος στις ημερήσιες αποδόσεις του δείκτη  $S\&P\ 500$  δεν φαίνεται να ερμηνεύει ικανοποιητικά την ετεροσκεδαστικότητα που παρατηρείται στις αποδόσεις του δείκτη, δεδομένου ότι τα κανονικοποιημένα κατάλοιπα εξακολουθούν να παρουσιάζουν συστάδες μεταβλητότητας.

### **Υπόδειγμα GARCH**

Το κυριότερο μειονέκτημα των  $ARCH$  υποδειγμάτων είναι ότι αρκετά συχνά απαιτείται μεγάλος αριθμός χρονικών υστερήσεων ώστε να περιγραφεί ικανοποιητικά η μεταβλητότητα των τιμών της υπό εξέταση χρονοσειράς, όπως ενδεχομένως στα δεδομένα του Διαγράμματος 3.9. Το πρόβλημα αυτό αντιμετωπίζεται από μία άλλη οικογένεια υποδειγμάτων που ονομάζονται γενικευμένα αυτοπαλίνδρομα υπό συνθήκης ετεροσκεδαστικότητας υποδείγματα με βαθμούς  $(s, r)$  που συμβολίζονται ως  $GARCH(s, r)$  και δίνονται από την ακόλουθη σχέση:



$$\sigma_t^2 = a_0 + \sum_{i=1}^s a_i \varepsilon_{t-i}^2 + \sum_{j=1}^r \beta_j \sigma_{t-j}^2 \quad (3.56)$$

όπου  $\{\varepsilon_t\}$  είναι μία ακολουθία iid τυχαίων μεταβλητών με μέση τιμή μηδέν και διακύμανση ίση με τη μονάδα, ενώ πρέπει να ισχύουν οι ακόλουθες προϋποθέσεις ότι  $a_0 > 0$ ,  $a_i \geq 0$  για  $i = 1, 2, \dots, s-1$ ,  $a_s > 0$ ,  $\beta_j \geq 0$  για  $j = 1, 2, \dots, r-1$  και  $\beta_r > 0$  για να εξασφαλίζεται το θετικό πρόσημο της υπό συνθήκη διακύμανσης. Επιπρόσθετα, θα πρέπει να ισχύει και ο περιορισμός  $\sum_{i=1}^m (a_i + \beta_i) < 1$ , για τιμές των παραμέτρων  $m = \max(s, r)$  με  $a_i = 0$  για  $i > s$  και με  $\beta_j = 0$  για  $j > r$ , έτσι ώστε να εξασφαλίζει ότι η μόνιμη διακύμανση ορίζεται και είναι πεπερασμένη. Σημειώνεται ότι η σχέση (3.56) απλοποιείται σε υπόδειγμα  $ARCH(s)$  αν  $r = 0$ , ενώ οι παράμετροι  $a_i$  και  $\beta_j$  ονομάζονται παράμετροι  $ARCH$  και  $GARCH$  αντίστοιχα.

Το απλούστερο υπόδειγμα  $GARCH$  είναι για βαθμούς  $(s, r) = (1, 1)$  και συμβολίζεται ως  $GARCH(1, 1)$ . Η υπό συνθήκη διακύμανση για το υπόδειγμα αυτό ορίζεται ως:

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \quad (3.57)$$

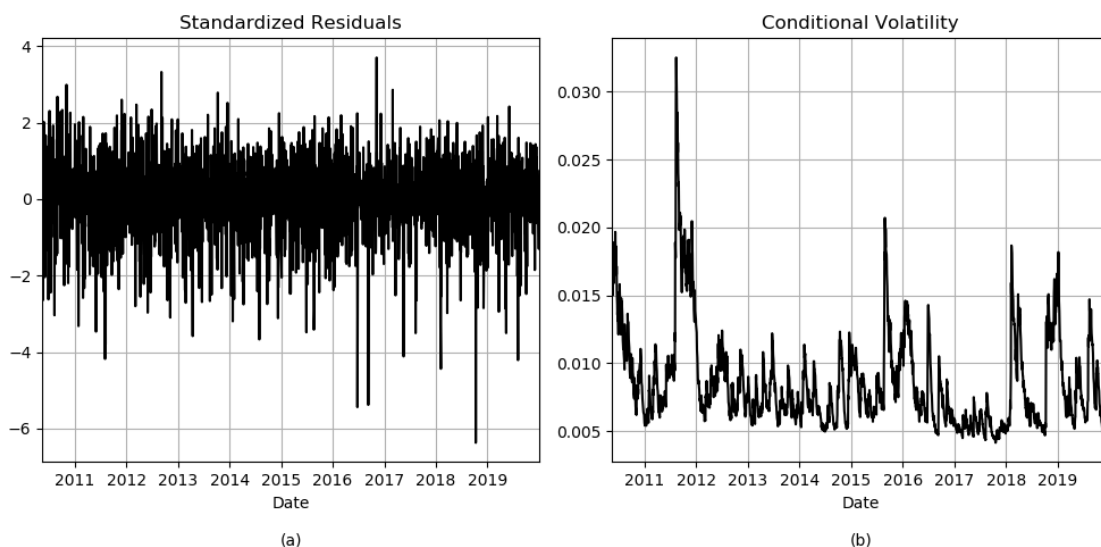
με την προϋπόθεση ότι οι σταθερές  $a_1$  και  $\beta_1$  είναι θετικοί αριθμοί και ότι θα πρέπει να ισχύει ότι  $a_1 + \beta_1 < 1$ , έτσι ώστε η ασθενής στασιμότητα της διαδικασίας  $\varepsilon_t$ , όπου η μόνιμη (unconditional) διακύμανσή της δίνεται από την ακόλουθη σχέση:

$$Var(\varepsilon_t) = \sigma_\varepsilon^2 = \frac{a_0}{1 - a_1 - \beta_1} \quad (3.58)$$

Τέλος, αν οριστεί  $u_t = \varepsilon_t^2 - \sigma_t^2$  και επομένως  $\varepsilon_t^2 = \sigma_t^2 + u_t$ , με τη μεταβλητή  $u_t$  να έχει μηδενικό μέσο όρο και να είναι σειριακά ασυσχέτιστες οι τιμές της, τότε η εξίσωση (3.57) μπορεί να λυθεί ως προς  $\varepsilon_t^2$ , έχοντας την εξής αναπαράσταση:

$$\varepsilon_t^2 = a_0 + a_1 (a_1 + \beta_1) \varepsilon_{t-1}^2 + u_t - \beta_1 u_{t-1}. \quad (3.59)$$

και κατά συνέπεια η διαδικασία των τετραγώνων των σφαλμάτων έχει τη μορφή μίας  $ARMA(1,1)$  χρονοσειράς υπό ετεροσκεδαστικά σφάλματα. Στην ειδική περίπτωση όπου  $\beta_1 = 0$ , τότε η διαδικασία  $\varepsilon_t^2$  καταλήγει σε μία χρονοσειρά  $AR(1)$ , δηλαδή σε αυτοπαλίνδρομη πρώτου βαθμού διαδικασία στις δεύτερες ροπές.



**Διάγραμμα 3.10**

### **Εφαρμογή Υποδείγματος $GARCH(1, 1)$**

Στο Διάγραμμα 3.10 (a και b), κατ' ανάλογο τρόπο με το Διάγραμμα 3.9, παρουσιάζονται τα κανονικοποιημένα κατάλοιπα (standardized residuals) και αυτά που προκύπτουν από την εφαρμογή ενός  $GARCH(1, 1)$  υποδείγματος για τα δεδομένα των ημερήσιων αποδόσεων του δείκτη  $S\&P\ 500$  που παρουσιάστηκαν στο Διάγραμμα 3.8. Από το διάγραμμα αυτό παρατηρείται πως η εφαρμογή ενός  $GARCH(1,1)$  υποδείγματος στις ημερήσιες αποδόσεις του δείκτη  $S\&P\ 500$  φαίνεται να έχει καλύτερη εφαρμογή στα δεδομένα, σε σχέση το υπόδειγμα  $ARCH(1)$  που εμφανίζεται στο Διάγραμμα 3.9, δεδομένου ότι τα κατάλοιπα έχουν μικρότερη μεταβλητότητα.

## **3.8 Αυτοπαλίνδρομα Υποδείγματα υπό Συνθήκης Ετεροσκεδαστικότητας υπό όρους Τυχαίου Συντελεστή και Επαυξημένης Ετεροσκεδαστικότητας**

Τα αυτοπαλίνδρομα υποδείγματα υπό συνθήκης ετεροσκεδαστικότητας προσφέρουν ένα βολικό πλαίσιο για την ταυτόχρονη ανάλυση αυτοσυσχέτισης και



ετεροσκεδαστικότητα. Μετά από το υπόδειγμα αυτοπαλίνδρομης ετεροσκεδαστικότητας υπό όρους *ARCH* που εισήχθη από τον Engle (1982) και τις γενικεύσεις που ακολούθησαν από τους Engle and Bollerslev (1986), Bera, Bubnys, and Park (1988) και Chou και Kroner (1992), τα υποδείγματα αυτά έγιναν ιδιαίτερα δημοφιλή στην ανάλυση χρηματοοικονομικών μεταβλητών, δεδομένου ότι οι μεταβλητές αυτές εμφανίζουν έντονη μεταβλητότητα. Σε εμπειρικές αναλύσεις που πραγματοποιήθηκαν σε χρονοσειρές με ετεροσκεδαστικότητα, διαπιστώθηκε ότι η παρουσία σειριακής συσχέτισης πρέπει να επανεκτιμηθεί επειδή οι συνήθεις δοκιμές για αυτοσυσχέτιση δεν είναι έγκυρες παρουσία *ARCH*.

Επιπρόσθετα, ο Diebold (1986) έδειξε ότι η παρουσία του υποδείματος *ARCH* ακυρώνει την τυπική θεωρία ασυμπτωτικής κατανομής των αυτοσυσχετίσεων του δείγματος και ως εκ τούτου των στατιστικών τεστ Box-Pierce και Box-Ljung για σειριακή συσχέτιση. Επιπλέον, όπως υποδεικνύεται από τους Engle, Hendry και Trumble (1985), η παρουσία αυτοσυσχέτισης μπορεί εύκολα να εκληφθεί εσφαλμένα ως *ARCH* όταν, στην πραγματικότητα, δεν υπάρχει *ARCH*. Μάλιστα, η συμπεριφορά αυτή έχει παρατηρηθεί από τον Agiakloglou (2009) στα πλαίσια της διερεύνησης των υποφαινομενικών σχέσεων στάσιμων αυτοπαλίνδρομων πρώτου βαθμού χρονοσειρών όπου η εμφάνιση της αυτοσυσχέτισης συνδέονταν με την εμφάνιση της ετεροσκεδαστικότητας με μορφή ενός *ARCH*(1) υποδείματος.

Οι θεωρητικές και εμπειρικές συνέπειες της αλληλεπίδρασης μεταξύ *ARCH* και αυτοσυσχέτισης θα μπορούσε να ερμηνευθούν με διάφορους τρόπους, όπως η μη κανονικότητα του όρου διαταραχής (Engle 1982) και η μη γραμμικότητα της διαδικασίας (Higgins και Bera 1990). Οι Bera και Lee (1991) καθιέρωσαν μια στενή σχέση μεταξύ αυτοσυσχέτισης και *ARCH* μέσω της ετερογένειας των παραμέτρων εφαρμόζοντας τη δοκιμή μήτρας πληροφοριών (Information Matrix - IM) του White (1982) στο τυπικό μοντέλο γραμμικής παλινδρόμησης και διαπίστωσαν ότι ο LM έλεγχος για ετεροσκεδαστικότητα με τον πολλαπλασιαστή Lagrange (LM test) του Engle (1982) για *ARCH* υποδείματος είναι μια ειδική περίπτωση μιας συνιστώσας του τεστ IM. Δεδομένης της ερμηνείας του Chesher (1984) του ελέγχου IM ως στατιστικός





έλεγχος για τη μεταβολή των παραμέτρων, μπορεί να ειπωθεί ότι η παρουσία του *ARCH* είναι ισοδύναμη με τυχαία διακύμανση στους συντελεστές αυτοσυσχέτισης (Tsay 1987). Αυτό οδήγησε τους Bera, Higgins και Lee (1992) να εξετάσουν το μοντέλο παλινδρόμησης με παραμέτρους τυχαίας αυτοσυσχέτισης.

Έτσι, στο πλαίσιο της ανάλυσης της αυτοσυσχέτισης ταυτόχρονα με *ARCH*, οι Bera et al. (1992 και 1996) καθορίσαν ένα γενικό πλαίσιο που περιλαμβάνει υποδείγματα τα οποία συναντώνται στην πράξη. Για το λόγο αυτό, εξετάζεται ένα υπόδειγμα στατικής γραμμικής παλινδρόμησης με τυχαίες αυτοπαλίνδρομες (AR) διαταραχές τυχαίου συντελεστή, όπως αυτό δίνεται στη συνέχεια:

$$y_t = \beta' X_t + \varepsilon_t \quad (3.60)$$

$$\varepsilon_t = \sum_{j=1}^p \varphi_{jt} \varepsilon_{t-j} + u_t = \sum_{j=1}^p (\varphi_j - \eta_{jt}) \varepsilon_{t-j} + u_t$$

όπου  $y$  είναι η  $t$  παρατήρηση στην εξαρτημένη μεταβλητή  $X$  είναι  $k \times 1$  διάνυσμα που αντιπροσωπεύει την  $t$  παρατήρηση και το  $\beta$  είναι ένα  $k \times 1$  διάνυσμα αγνώστων σταθερών συντελεστών. Η διαταραχή  $\varepsilon$  ακολουθεί μια στοχαστική αυτοπαλίνδρομη διαδικασία (AR) τάξης  $p$  με σταθερούς όρους  $\varphi_j$  και στοχαστικούς όρους  $\eta_{jt}$  για όλα τα  $j$ . Αν τα  $X$  είναι στοχαστικά, τότε θεωρείται ότι είναι ανεξάρτητα από τα  $\varepsilon_s$ , για όλα τα  $t$  και  $s$ . Το υπόδειγμα της σχέσης (3.60) περιλαμβάνει ένα ευρύ φάσμα υποδειγμάτων ως ειδικές περιπτώσεις, τα οποία προσπαθούν να μελετήσουν διαχρονικά μεταβαλλόμενες σχέσεις μεταξύ μεταβλητών ή καλύτερα σχέσεις με έντονη μεταβλητότητα.

Έστω  $\Psi_{t-1}$  το σύνολο πληροφοριών που είναι διαθέσιμο τη χρονική στιγμή  $t$ . Η σειρά των παρελθοντικών διαταραχών παλινδρόμησης  $\{\varepsilon_{t-1} | t = 1, 1, \dots\}$  θεωρείται ότι ανήκει στο  $\Psi_{t-1}$ . Τότε ο υπό όρους μέσος όρος  $\mu$  και η διακύμανση της διαταραχής  $\varepsilon_t$ , σύμφωνα με τις υποθέσεις μπορούν να αναπαρασταθούν ως ακολούθως. Αρχικά ο μέσος όρος ορίζεται ως:

$$\mu = E(\varepsilon_t | \Psi_{t-1}) = E \left[ \sum_{j=1}^p \varphi_{jt} \varepsilon_{t-j} + u_t | \Psi_{t-1} \right] = \sum_{j=1}^p \varphi_j \varepsilon_{t-j} = \varphi' \varepsilon_{t-1} \quad (3.61)$$



όπου  $\varphi = (\varphi_1, \dots, \varphi_p)'$ ,  $\varepsilon_{t-1} = (\varepsilon_{t-1}, \dots, \varepsilon_{t-p})'$  και αφήνοντας  $\eta_t = (\eta_{1t}, \dots, \eta_{pt})'$ , τότε προκύπτει ότι η διακύμανση είναι ίση με:

$$\begin{aligned} \eta_t &= \text{var}(\varepsilon_t | \Psi_{t-1}) \\ &= E \left[ \left( \sum_{j=1}^p \varphi_{jt} \varepsilon_{t-j} + u_t \right)^2 | \Psi_{t-1} \right] \\ &= E [\varepsilon_{t-1}' + \eta_t' \eta_t \varepsilon_{t-1} + 2u_t \varepsilon_{t-1}' \eta_t + u_t^2 | \Psi_{t-1}] \\ &= \varepsilon_{t-1}' \Sigma \varepsilon_{t-1} + \sigma^2 \end{aligned} \quad (3.62)$$

Μάλιστα, εάν  $\varphi \neq 0$ , η διαδικασία σφάλματος συσχετίζεται σειριακά και αν  $\Sigma \neq 0$ , τότε η διαδικασία σφάλματος έχει *ARCH*. Η διαγώνιος  $\Sigma$  καθορίζει τη γραμμική διαδικασία *ARCH* που προτάθηκε από τον Engle (1982), όμως η μη- διαγώνιος  $\Sigma$  καθορίζει μια διεργασία *ARCH* με πρόσθετους όρους διανυσματικού γινομένου μεταξύ προηγούμενων σφαλμάτων. Το τελευταίο υπόδειγμα ονομάζεται επαυξημένο αυτοπαλίνδρομο υπόδειγμα υπό συνθήκης ετεροσκεδαστικότητας (Augmented Autoregressive Conditional Heteroscedasticity ή *AARCH*). Σε αντίθεση με το *ARCH*, το *AARCH* δεν είναι συμμετρικό με την έννοια ότι η υπό συνθήκη διακύμανση  $\eta_t$  εξαρτάται από τα πρόσημα των μεμονωμένων με καθυστέρηση σφαλμάτων  $\varepsilon_t$ . Επομένως, το *AARCH* εμφανίζει το αποτέλεσμα μόχλευσης του μη-συμμετρικού υποδείγματος *ARCH* που προτείνεται από τον Nelson (1991).

Στο σημείο αυτό θα πρέπει να σημειωθεί ότι υπάρχουν και άλλοι τρόποι εισαγωγής της από κοινού. αυτοσυσχέτισης μαζί με *ARCH*. Μία τέτοια εναλλακτική προσέγγιση στη διατύπωση του τυχαίου συντελεστή είναι να εξεταστεί μια διαταραχή *AR* με τις καινοτομίες *ARCH*, όπως παρουσίασαν οι Bera, Lee, and Higgins (1990) ένα γενικευμένο υπόδειγμα *AARCH* έτσι ώστε η υπό όρους διακύμανση του  $\varepsilon_t$  να είναι επίσης συνάρτηση των καθυστερημένων  $\eta_t$ .



### 3.9 Ανακεφαλαίωση

Στο κεφάλαιο αυτό παρουσιάστηκαν συνοπτικά τα χαρακτηριστικά και οι ιδιότητες στάσιμων χρονοσειρών στη συνέχεια περιγράφηκε η μεθοδολογία των Box και Jenkins (1970). Κατόπιν, αναλύθηκε η έννοια των μη στάσιμων χρονοσειρών, μαζί με όλους τους στατιστικούς ελέγχους που εφαρμόζονται για τη διαπίστωση της ύπαρξης μοναδιαίας αυτοσυσχετιζόμενης ρίζας. Οι χρονοσειρές αυτές αποτυπώνουν συμπεριφορές φαινομένων με έντονη μεταβλητότητα και απαιτούν προσεκτική αντιμετώπιση διότι η ανάλυσή τους δεν είναι εύκολη. Στα πλαίσια αυτής της συμπεριφοράς παρουσιάστηκαν και τα υποδείγματα υπό συνθήκης ετεροσκεδαστικότητας τα οποία και αυτά απεικονίζουν τη μεταβλητότητα αλλά με ένα συγκεκριμένο τρόπο όπου η διακύμανση της χρονοσειράς μεταβάλλεται διαχρονικά. Τέλος, αναφορά έγινε και στα υποδείγματα υπό συνθήκης ετεροσκεδαστικότητας υπό όρους τυχαίου συντελεστή και επανυξημένης ετεροσκεδαστικότητας τα οποία παρουσιάστηκαν από τους Bera, Higgins και Lee (1992 και 1996) ως μία άλλη μορφή υποδειγμάτων που μπορεί να περιλάβει ακόμα μεγαλύτερη μεταβλητότητα στις τιμές μιας χρονοσειράς.



*Αυτή η σελίδα αφήνεται σκόπιμα κενή.*



## ΚΕΦΑΛΑΙΟ 4

### ΔΙΜΕΤΑΒΛΗΤΕΣ ΓΡΑΜΜΙΚΕΣ ΣΧΕΣΕΙΣ ΓΙΑ ΜΗ ΓΡΑΜΜΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ

#### 4.1 Εισαγωγή

Ο προσδιορισμός της σχέσης μεταξύ δύο μεταβλητών είναι αρκετά συχνά ένα δύσκολο ερευνητικό θέμα για πολλές οικονομικές και χρηματοοικονομικές εφαρμογές, κυρίως όταν οι παρατηρήσεις των μεταβλητών προέρχονται από χρονοσειρές, λόγω της φύσεως των δεδομένων. Η δυσκολία γίνεται ακόμα μεγαλύτερη όταν οι αναλυτές στηρίζουν την ανάλυση τους στην σιωπηρή παραδοχή ότι η σχέση των δύο μεταβλητών παραμένει διαχρονικά σταθερή, ενώ αρκετοί παράγοντες μπορούν να την αλλάξουν ανά πάσα στιγμή, ή επίσης ότι ο μηχανισμός δημιουργίας των χρονοσειρών προέρχεται από γραμμικής μορφής υποδείγματα, κινδυνεύοντας να οδηγηθούν σε λανθασμένα συμπεράσματα.

Και όμως κάθε μελέτη για τη συμμεταβολή δύο μεταβλητών ξεκινάει από την ανάλυση συσχέτισης, γνωρίζοντας από την αρχή ότι θα υπάρχουν ορισμένες ενστάσεις σχετικά με την αξιοπιστία των αποτελεσμάτων. Άλλωστε, όπως έχει αναφερθεί, ο Fisher (1915), ο οποίος μελέτησε τις ιδιότητες του συντελεστή συσχέτισης του δείγματος,  $r$ , δύο μεταβλητών διαπίστωσε ότι δεν είναι αμερόληπτος εκτιμητής του πληθυσμιακού συντελεστή συσχέτισης τους  $\rho$ , όπως επίσης παρουσιάζεται και στους Kenny και Keeping (1951), αλλά και στον Kendall (1954). Βέβαια, η μεροληψία του δειγματικού συντελεστή συσχέτισης, αυτή καθαυτή, δεν είναι ένα σημαντικό θέμα, δεδομένου ότι το μέγεθός του είναι πολύ μικρό, αφού λαμβάνει τιμές από -1 έως +1, και έτσι η μεροληψία ως μέγεθος θα είναι πολύ μικρή. Αυτό όμως που είναι σημαντικό, αναφορικά με τη μεροληψία του δειγματικού συντελεστή συσχέτισης, είναι ότι μπορεί να επηρεάσει το αποτέλεσμα του στατιστικού ελέγχου για τον έλεγχο της μηδενικής



υπόθεσης ότι δεν υπάρχει συσχέτιση, δημιουργώντας λανθασμένες ενδείξεις στον αναλυτή για τη σχέση των μεταβλητών.

Εσφαλμένες ενδείξεις μπορεί να αποκτηθούν και μόνο από το γεγονός ότι στην ανάλυση συσχέτισης χρησιμοποιούνται δεδομένα από χρονοσειρές. Πιο συγκεκριμένα, ο αναλυτής μπορεί να οδηγηθεί στο συμπέρασμα ότι δύο μεταβλητές εμφανίζουν να έχουν γραμμική σχέση μεταξύ τους, ενώ είναι ανεξάρτητες, δηλαδή δεν έχουν κοινή συμπεριφορά, δεδομένου ότι στο στατιστικό έλεγχο για την ύπαρξη της συσχέτισής τους η μηδενική υπόθεση απορρίφθηκε. Το φαινόμενο αυτό παρουσιάστηκε αρχικά από τον Yule (1926), ονομάζεται υποφαινομενική συσχέτιση (*spurious correlation*), και εντάσσεται στη γενικότερη έννοια των λανθασμένων συμπερασμάτων (*miss-specification issues*) που προκύπτουν από τη φύση των δεδομένων.

Ειδικότερα, ο Yule (1926) παρατήρησε ότι ο συντελεστής συσχέτισης δύο μεταβλητών δεν έχει κανένα νόημα αν προέρχεται από μία κατανομή συχνοτήτων που έχει σχήμα  $U$ , αποτέλεσμα το οποίο εμφανίζεται και στο βιβλίο των Banerjee et al. (1993). Μάλιστα, στο βιβλίο αυτό υπάρχουν διαγράμματα, τα οποία προέρχονται από προσομοιώσεις χρησιμοποιώντας ανάλυση Monte Carlo και τα οποία δείχνουν ότι η κατανομή συχνοτήτων του συντελεστή συσχέτισης δύο ανεξάρτητων μη στάσιμων  $I(2)$  χρονοσειρών έχει σχήμα  $U$  με τις τιμές  $-1$  και  $+1$  να εμφανίζονται πολύ έντονα, ενώ η κατανομή συχνοτήτων του συντελεστή συσχέτισης δύο ανεξάρτητων μη στάσιμων  $I(1)$  χρονοσειρών έχει ημι-ελλειπτικό σχήμα (*semi-ellipse*). Αντίθετα από αυτές τις συμπεριφορές, η κατανομή συχνοτήτων του συντελεστή συσχέτισης δύο ανεξάρτητων λευκού θορύβου χρονοσειρών (*white noise processes*) είναι συμμετρική στο μηδέν και συμπεριφέρεται ως κανονική κατανομή. Ωστόσο, συμμετρία ως προς το μηδέν στην κατανομή συχνοτήτων εμφανίζεται και σε δύο ανεξάρτητες στάσιμες αυτοπαλίνδρομου πρώτου βαθμού χρονοσειρές, δηλαδή σε  $AR(1)$  χρονοσειρές, όπως έδειξαν οι Agiakloglou and Tsimpanos (2012), αλλά η κατανομή τους δεν ομοιάζει με κανονική κατανομή, και έτσι δημιουργούνται προβλήματα υποφαινομενικών σχέσεων μέσω της εφαρμογής του στατιστικού ελέγχου για την διερεύνηση της ύπαρξης της



συσχέτισης, ενώ οι Agiakloglou και Agiropoulos (2016) έδειξαν ότι και η δύναμη του στατιστικού ελέγχου έχει πρόβλημα για αυτές τις χρονοσειρές.

Τα αποτελέσματα του Yule (1926), που έδειχναν ότι υπάρχει γραμμική σχέση σε χρονοσειρές που δεν είχαν τίποτε κοινό μεταξύ τους, αποτέλεσαν σημείο αναφοράς για την εργασία των Granger and Newbold (1974) οι οποίοι μελέτησαν το φαινόμενο αυτό από την πλευρά της παλινδρόμησης και ονομάστηκε υποφαινομενική παλινδρόμηση (spurious regression). Πιο συγκεκριμένα, οι Granger and Newbold (1974) έδειξαν ότι η παλινδρόμηση δύο ανεξάρτητων μη στάσιμων τυχαίου περπατήματος (random walk) χωρίς μέσο όρο χρονοσειρών θα οδηγήσει τον αναλυτή σε εσφαλμένα συμπεράσματα, κάτι το οποίο διερευνήθηκε αργότερα και για δύο ανεξάρτητες στάσιμες AR(1) χρονοσειρές από τους Granger et al. (2001), σημειώνοντας ότι η έρευνα των υποφαινομενικών σχέσεων έχει βασιστεί σε γραμμικές σχέσεις προερχόμενες από γραμμικά παραγόμενες διαδικασίες, ανεξάρτητα από το αν η γραμμικότητα μπορεί να μην είναι η καταλληλότερη προσέγγιση μιας υπάρχουσας σχέσης ούτε ο καταλληλότερος τρόπος για να περιγραφεί η εξέλιξη μιας ασταθούς διαδικασίας. Άλλωστε, είναι πολύ δύσκολο να διατυπωθούν συμπεριφορές μη γραμμικών σχέσεων αλλά και συνθηκών που να είναι γενικά αποδεκτές, δεδομένου ότι η μη γραμμικότητα μπορεί να εμφανιστεί με πάρα πολλούς τρόπους.

Κατά ανάλογο τρόπο και ο συντελεστής συσχέτισης αναφέρεται σε γραμμικές σχέσεις χωρίς να είναι σε θέση να προσδιορίσει άλλες μορφές σχέσεων μεταξύ δύο μεταβλητών. Μάλιστα, όπως αναφέρεται στους Embrechts et al. (1999), ο συντελεστής συσχέτισης δεν είναι επαρκές μέτρο μέτρησης εξάρτησης, εκτός εάν οι σειρές είναι από κοινού Gaussian, κάτι που δεν εμφανίζεται συχνά στις χρηματοοικονομικές μελέτες. Για το λόγο αυτό, έχουν δημιουργηθεί άλλες τεχνικές μέτρησης της σχέσης εξάρτησης μεταξύ δύο μεταβλητών που να περιλαμβάνουν περισσότερο πολύπλοκες σχέσεις, όπως για παράδειγμα οι τεχνικές *Copula*, όπου είναι μια πολυμεταβλητή συνάρτηση κατανομής που μπορεί να περιγράψει πλήρως την εξάρτηση μεταξύ δύο (ή περισσότερων) μεταβλητών, ενώ ταυτόχρονα είναι ανεξάρτητη από τις οριακές κατανομές τους.



Στο κεφάλαιο αυτό διερευνάται η συμπεριφορά του συντελεστή συσχέτισης, καθώς και ορισμένων τεχνικών *Copula* για ανεξάρτητες, καθώς και γραμμικά εξαρτημένες, χρονοσειρές με ετεροσκεδαστικότητα, ορισμένες από τις οποίες συμπεριφέρονται ως μη γραμμικές. Για το λόγο αυτό παρουσιάζονται αρχικά οι μεταβλητές και οι τεχνικές εξάρτησης που θα χρησιμοποιηθούν στην προσομοίωση, δηλαδή για την ανάλυση Monte Carlo. Στη συνέχεια, εμφανίζονται τα αποτελέσματα της συμπεριφοράς των τεχνικών εξάρτησης των ερευνητικών προσπαθειών για ανεξάρτητες μεταβλητές, καθώς και για υποφαινομενικές συσχετίσεις χρησιμοποιώντας το στατιστικό έλεγχο για τη μηδενική υπόθεση ότι δεν υπάρχει γραμμική συσχέτιση με βάση το συντελεστή συσχέτισης. Κατόπιν, παρουσιάζονται τα αποτελέσματα για γραμμικές εξαρτήσεις δύο μεταβλητών με όλες τις τεχνικές εξάρτησης, καθώς και η δύναμη του στατιστικού ελέγχου της μηδενικής συσχέτισης.

## 4.2 Οι μεταβλητές της προσομοίωσης

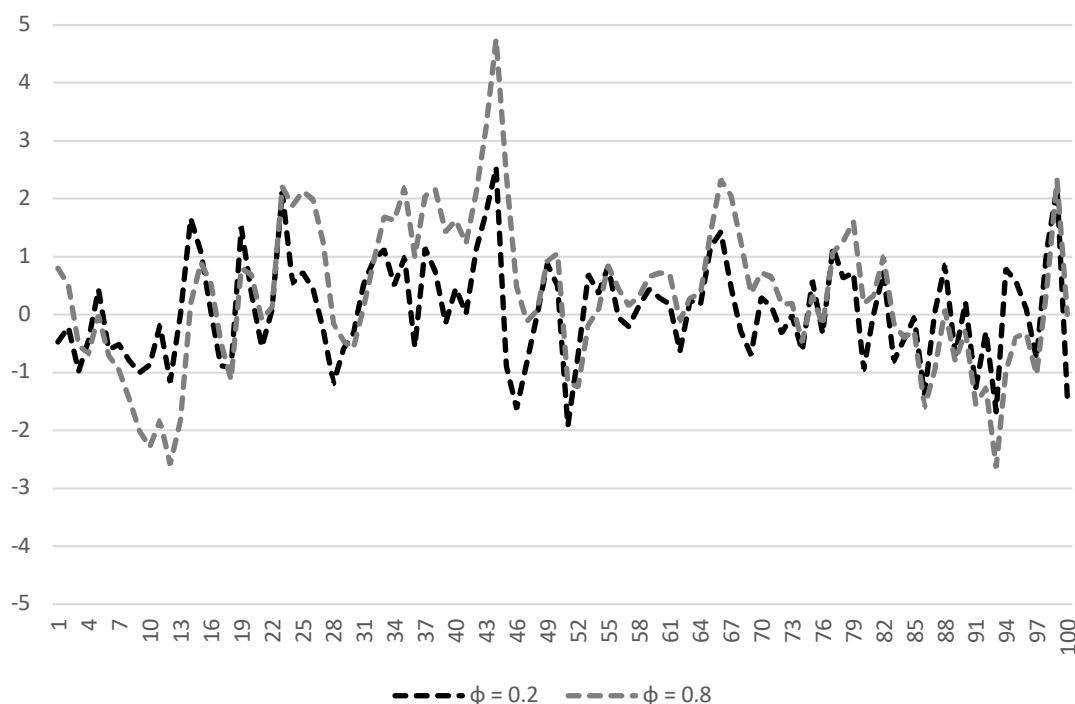
Η ανάλυση Monte Carlo (Monte Carlo Analysis – MCA) που θα εφαρμοστεί θα βασιστεί σε στάσιμες διαδικασίες χρονοσειρών αυτοπαλίνδρομου πρώτου βαθμού (Autoregressive process of order one – AR(1)), δηλαδή σε:

$$X_t = \varphi X_{t-1} + \varepsilon_t \quad (4.1)$$

όπου οι τιμές του αυτοπαλίνδρομου όρου  $\varphi$  είναι μικρότερες της μονάδας για να είναι στάσιμη η χρονοσειρά και το τυχαίο σφάλμα ακολουθεί τη διαδικασία του λευκού θορύβου. Σημειώνεται ότι για  $\varphi = 1$  η διαδικασία είναι μη στάσιμη χρονοσειρά, δηλαδή χρονοσειρά τυχαίου περπατήματος χωρίς μέσο όρο (random walk process without drift), ενώ για  $\varphi = 0$  η διαδικασία είναι λευκός θόρυβος (white noise process). Σημειώνεται ότι το υπόδειγμα της σχέσης (4.1) είναι γραμμικό υπόδειγμα με μέσο όρο μηδέν και σταθερή παράμετρο που προέρχεται από μία μεγάλη οικογένεια υποδειγμάτων, γνωστά ως αυτοπαλίνδρομα κινητού μέσου όρου υποδείγματα (AutoRegressive Moving Average processes – ARMA( $p$ ,  $q$ )), τα οποία



παρουσιάστηκαν από τους Box και Jenkins (1976), και έχει σταθερό μέσο όρο και σταθερή διακύμανση διαχρονικά, δύο βασικές προϋποθέσεις για τον καθορισμό του τρόπου εξέλιξης της χρονοσειράς. Στο Διάγραμμα 4.1 απεικονίζονται ενδεικτικά η διαχρονική συμπεριφορά ενός AR(1) υποδείγματος με μέσο όρο μηδέν για τιμές της παραμέτρου του 0.2 και 0.8.



**Διάγραμμα 4.1**

**Διαγραμματική απεικόνιση ενός AR(1) υποδείγματος για  $\phi = 0.2$  και 0.8**

Οι δύο προαναφερθείσες προϋποθέσεις των σταθερών τιμών του μέσου όρου και της διακύμανσης της χρονοσειράς αποτελούν βασικό σημείο αναφοράς για τη στασιμότητα αλλά και για την ανάλυση της χρονοσειράς τόσο ως προς την εκτίμηση του υποδείγματος όσο και ως προς την πρόβλεψη των μελλοντικών τιμών της. Ωστόσο, πολύ συχνά σε εμπειρικές μελέτες κυρίως για χρηματοοικονομικά δεδομένα έχει διαπιστωθεί ότι οι υποθέσεις αυτές δεν υφίστανται με αποτέλεσμα να δημιουργούνται πολύ σοβαρά προβλήματα για την διερεύνηση της συμπεριφοράς ενός φαινομένου. Πολύ συχνά η αντιμετώπιση αυτού του προβλήματος στηρίζεται στη μετατροπή της



χρονοσειράς σε πρώτες διαφορές ελπίζοντας ότι η νέα σειρά θα είναι στάσιμη, μία τεχνική που έχει πολύ καλά αποτελέσματα στην περίπτωση όπου το πρόβλημα επικεντρώνεται στο μέσο όρο. Αντίθετα, αν το πρόβλημα είναι στη διακύμανση ή ακόμα και στα δύο, δηλαδή στο μέσο όρο και στη διακύμανση, συνήθως η απόκτηση της στασιμότητας επιτυγχάνεται κατά κανόνα με λογαριθμικές διαφορές των τιμών της χρονοσειράς.

Η αλήθεια είναι ότι το πρόβλημα του μέσου όρου εμφανίζεται στην πράξη λιγότερο σημαντικό και η στασιμότητα επιτυγχάνεται σχετικά εύκολα. Το πρόβλημα όμως της διακύμανσης είναι πιο σημαντικό, διότι δεν επιτυγχάνεται εύκολα η στασιμότητα και εμφανίζεται πολύ πιο συχνά στην ανάλυση δεδομένου ότι τα περισσότερα χρηματοοικονομικά δεδομένα έχουν έντονη μεταβλητότητα. Για το λόγο αυτό έχουν αναπτυχθεί διάφορα υποδείγματα στα οποία καθορίζεται ο τρόπος συμπεριφοράς της μεταβλητότητας και περιγράφουν την εξέλιξη της μεταβαλλόμενης διαχρονικά διακύμανσης. Η πιο απλή μορφή καθορισμού της μεταβαλλόμενης διακύμανσης είναι να θεωρηθεί ότι προέρχεται από ένα αυτοπαλίνδρομο υπό συνθήκης ετεροσκεδαστικότητας πρώτου βαθμού (AutoRegressive Conditionally Heteroscedastic model of order one – ARCH(1)), υπόδειγμα που παρουσιάστηκε από τον Engle (1982) και ορίζεται ως:

$$\varepsilon_t = u_t \sqrt{a_0 + a_1 \varepsilon_{t-1}^2} \quad (4.2)$$

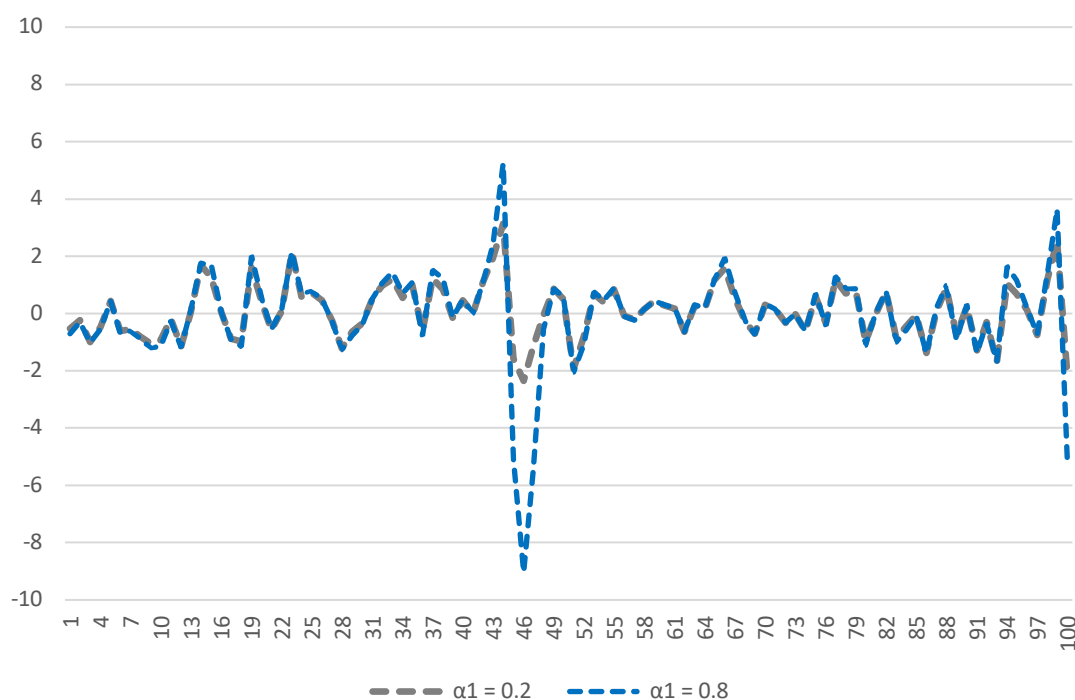
όπου  $a_0 > 0$ ,  $0 < a_1 < 1$  και το τυχαίο σφάλμα  $u_t$  ακολουθεί κανονική κατανομή με μέσο όρο μηδέν και διακύμανση ένα, δηλαδή  $u_t \sim iid N(0, 1)$ . Από τη σχέση (4.1) προκύπτει ότι η διακύμανση του σφάλματος ορίζεται από τις προηγούμενες τιμές της υστερούμενες κατά μία περίοδο.

Ο συνδυασμός των σχέσεων (4.1) και (4.2) προσδιορίζει μία άλλη οικογένεια υποδειγμάτων στα οποία λαμβάνεται υπ' όψιν η μεταβαλλόμενη διακύμανση, γνωστά ως αυτοπαλίνδρομα πρώτου βαθμού υποδείγματα με ετεροσκεδαστικότητα τύπου ARCH(1) στα σφάλματα (AR(1) – ARCH(1) in errors models). Τα υποδείγματα αυτά παραμένουν γραμμικά με σταθερό μέσο όρο και με σταθερή τιμή για την παράμετρο



τους  $\varphi$ , αλλά με μη σταθερή διακύμανση εξαρτωμένη στο παρελθόν (non-constant variance conditional on past), ενώ η χωρίς εξάρτηση διακύμανσή τους είναι σταθερή (constant unconditional variance). Στα Διαγράμματα 4.2 και 4.3 παρουσιάζονται ενδεικτικά οι διαχρονικές εξελίξεις ενός  $AR(1) - ARCH(1)$  στα σφάλματα με μέσο όρο μηδέν για τιμές της παραμέτρου τους  $\varphi$  ίσες με 0.2 και 0.8 αντίστοιχα και για τιμές της παραμέτρου  $\alpha_1$  ίσες με 0.2 και 0.8, σημειώνοντας ότι όταν η τιμή του  $\alpha_1$  είναι ίση με μηδέν, τότε το υπόδειγμα αυτό ταυτίζεται με ένα απλό στάσιμο αυτοπαλίνδρομο πρώτου βαθμού υπόδειγμα,  $AR(1)$ , όπως χαρακτηριστικά απεικονίζεται στο Διάγραμμα 4.1.

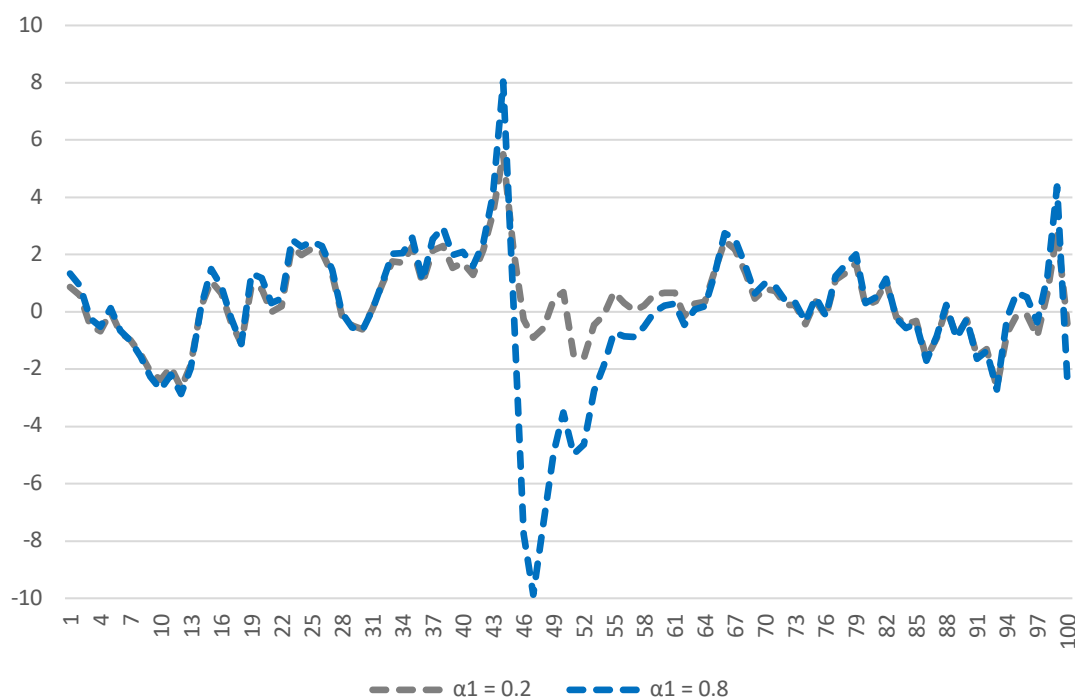
Σημειώνεται ότι η ύπαρξη ετεροσκεδαστικότητας που προέρχεται από τις τιμές του τυχαίου σφάλματος αυξάνει τη μεταβλητότητα των τιμών της χρονοσειράς, όπως εμφανίζεται στα Διαγράμματα 4.2 και 4.3. Μάλιστα, η μεταβλητότητα αυξάνεται όσο αυξάνεται η τιμή της παραμέτρου του αυτοπαλίνδρομου όρου, καθώς και η τιμή του της παραμέτρου  $\alpha_1$ .



**Διάγραμμα 4.2**



**Διαγραμματική απεικόνιση ενός AR(1) – ARCH(1) στα σφάλματα υποδείγματος  
για  $\phi = 0.2$  και για  $\alpha_1 = 0.2$  και  $0.8$**



**Διάγραμμα 4.3**

**Διαγραμματική απεικόνιση ενός AR(1) – ARCH(1) στα σφάλματα υποδείγματος  
για  $\phi = 0.8$  και για  $\alpha_1 = 0.2$  και  $0.8$**

Τα πράγματα διαφοροποιούνται δραστικά ως προς τον καθορισμό της συμπεριφοράς της διακύμανσης όταν η μεταβλητότητα προέρχεται από προηγούμενες τιμές της μεταβλητής και όχι από προηγούμενες τιμές του σφάλματος. Στην περίπτωση αυτή το υπόδειγμα ARCH(1) δίνεται από την ακόλουθη σχέση:

$$\varepsilon_t = u_t \sqrt{a_0 + a_1 X_{t-1}^2} \quad (4.3)$$

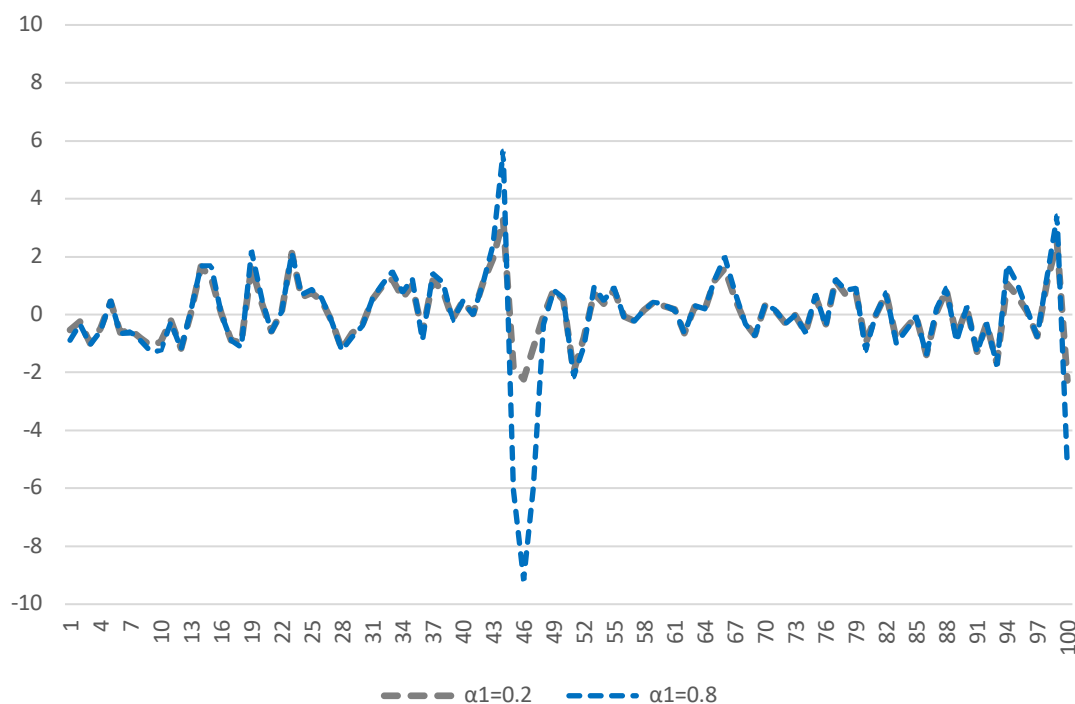
ενώ ο συνδυασμός των σχέσεων (4.1) και (4.3) προσδιορίζει μία νέα οικογένεια υποδειγμάτων, τα οποία παρουσίασαν οι Bera, Higgins και Lee το 1992 και το 1996, γνωστά ως αυτοπαλίνδρομα πρώτου βαθμού υποδείγματα με ετεροσκεδαστικότητα



τύπου ARCH(1) στη μεταβλητή (AR(1) – ARCH(1) in variable models) και στα οποία η υπό συνθήκη κατανομή της μεταβλητής  $X_t$  δίνεται από την ακόλουθη σχέση:

$$X_t|X_{t-1} \sim N(\varphi X_{t-1}, a_0 + a_1 X_{t-1}^2) \quad (4.4)$$

Στα Διαγράμματα 4.4 και 4.5 παρουσιάζονται ενδεικτικά οι διαχρονικές εξελίξεις ενός AR(1) – ARCH(1) στη μεταβλητή για τιμές της παραμέτρου τους  $\varphi$  ίσες με 0.2 και 0.8 αντίστοιχα και για τιμές της παραμέτρου  $a_1$  ίσες με 0.2 και 0.8, σημειώνοντας ότι όταν η τιμή του  $a_1$  είναι ίση με μηδέν, τότε το υπόδειγμα αυτό και σε αυτή την περίπτωση ταυτίζεται με ένα απλό στάσιμο αυτοπαλίνδρομο πρώτου βαθμού υπόδειγμα, AR(1), όπως χαρακτηριστικά απεικονίζεται στο Διάγραμμα 4.1.



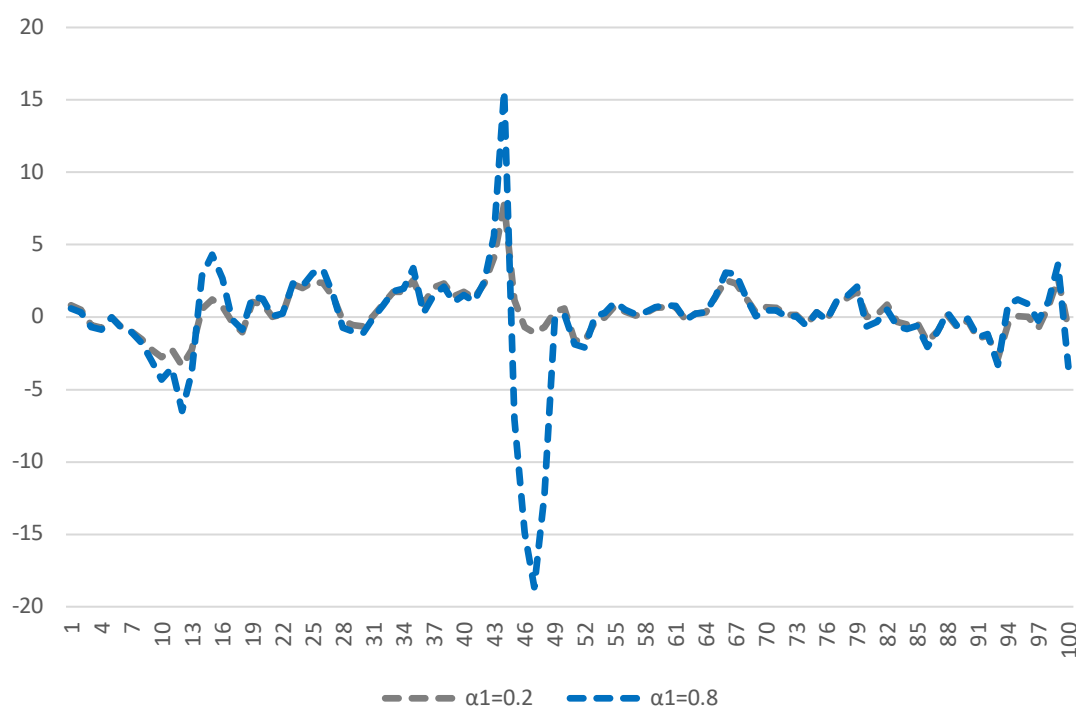
**Διάγραμμα 4.4**

**Διαγραμματική απεικόνιση ενός AR(1) – ARCH(1) στη μεταβλητή υποδείγματος για  $\varphi = 0.2$  και για  $a_1 = 0.2$  και 0.8**

Όπως χαρακτηριστικά απεικονίζεται στα Διαγράμματα 4.4 και 4.5 η μεταβλητότητα αυξήθηκε ακόμα περισσότερο σε σχέση με την προηγούμενη περίπτωση. Μάλιστα,



στο Διάγραμμα 4.5 έχει αλλάξει και η κλίμακα του κάθετου άξονα για να συμπεριλάβει τις τιμές της χρονοσειράς, όπου και σε αυτήν την περίπτωση η μεταβλητότητα των τιμών της χρονοσειράς αυξάνεται όσο αυξάνονται οι τιμές των παραμέτρων  $\phi$  και  $a_1$ . Σημειώνεται ότι όλες οι χρονοσειρές που απεικονίζονται στα Διαγράμματα 4.1 έως 4.5 προέρχονται από τις ίδιες τιμές του τυχαίου σφάλματος για να είναι εύκολα συγκρίσιμες.



**Διάγραμμα 4.5**

**Διαγραμματική απεικόνιση ενός AR(1) – ARCH(1) στη μεταβλητή υποδείγματος για  $\phi = 0.8$  και για  $a_1 = 0.2$  και  $0.8$**

Ωστόσο, το πιο συναρπαστικό σημείο της διαφοράς των δύο αυτών οικογενειών υποδειγμάτων, δηλαδή του AR(1) – ARCH(1) στα σφάλματα και του αντίστοιχου AR(1) – ARCH(1) στη μεταβλητή, δεν είναι ο τρόπος σχηματισμού της μεταβαλλόμενης διακύμανσης, και κατ' επέκταση ότι η δεύτερη οικογένεια υποδειγμάτων εμφανίζει μεγαλύτερη μεταβλητότητα, αλλά το γεγονός ότι η πρώτη οικογένεια παραμένει στη μορφή γραμμικών υποδειγμάτων, ενώ η δεύτερη οικογένεια



υποδειγμάτων προέρχεται από μη γραμμικά υποδείγματα, όπως απέδειξαν οι Bera et al. (1992 και 1996). Πιο συγκεκριμένα, τα υποδείγματα  $AR(1) - ARCH(1)$  στη μεταβλητή μπορούν εξίσου να εκφραστούν ως οικογένεια μη γραμμικών υποδειγμάτων τα οποία παρουσιάζονται με διαχρονικά μεταβαλλόμενο συντελεστή (time varying coefficient) σε γραμμική δομή υποδείγματος, δηλαδή ως:

$$X_t = \varphi_t X_{t-1} + w_t \quad (4.5)$$

όπου  $\varphi_t \sim iid(\varphi, a_1)$  and  $w_t \sim iid N(0, a_0)$ . Στην περίπτωση αυτή, η υπό συνθήκη κατανομή της μεταβλητής  $X_t$  που δημιουργείται από ένα αυτοπαλίνδρομο υπόδειγμα πρώτου βαθμού με τυχαίο συντελεστή (random coefficient) είναι ίδια με αυτή που δίνεται στη σχέση (4.4), ενώ η χωρίς συνθήκη διακύμανση (unconditional variance) της μεταβλητής  $X_t$  δίνεται από την ακόλουθη σχέση:

$$Var(X_t) = \frac{a_0}{1 - \varphi^2 - a_1}$$

και η οποία εξαρτάται από τις τιμές των παραμέτρων  $\varphi$  και  $a_1$ . Κατά συνέπεια, για μεγάλες τιμές και των δύο παραμέτρων η τυχαιότητα (randomness) του αυτοπαλίνδρου όρου (autoregressive parameter) αυξάνει την χωρίς συνθήκη μεταβλητότητα της μεταβλητής, αποσβένοντας κατά μία έννοια το βαθμό της γραμμικής δομής του υποδείγματος, όπως σχολίασαν οι Bera et al. (1992 και 1996).

### 4.3 Οι τεχνικές της προσομοίωσης

Η ανάλυση με μία μεταβλητή επικεντρώνεται στο να προσδιοριστούν λίγοι αριθμοί, τέσσερεις το πολύ, που να περιγράφουν με λιτό τρόπο την εξέλιξή της. Οι αριθμοί αυτοί εκφράζουν τις ροπές (moments) της μεταβλητής και δίνουν στον αναλυτή μία γρήγορη και συνοπτική εικόνα της συμπεριφοράς της. Αντίθετα, η ανάλυση με δύο μεταβλητές επικεντρώνεται στον να προσδιοριστεί με έναν αριθμό η σχέση εξάρτησής τους, ξεκινώντας αρχικά με την προϋπόθεση ότι οι μεταβλητές αυτές συνδέονται μεταξύ τους.



Η αλήθεια είναι ότι κάθε ποσοτική ανάλυση με δύο μεταβλητές αρχίζει με την ανάλυση συσχέτισης (correlation analysis), δηλαδή με το συντελεστή συσχέτισης, ένα μέτρο έκφρασης της ύπαρξης, του βαθμού, αλλά και της κατεύθυνσης της γραμμικής σχέσης των δύο μεταβλητών, αν και ως πρώτη προσπάθεια προσδιορισμού της συμμεταβολής τους θα μπορούσε να θεωρηθεί η συνδιακύμανση (covariance). Ωστόσο, επειδή η συνδιακύμανση έχει σημαντικά θέματα ερμηνείας που ξεκινάνε από το μέγεθός της και καταλήγουν στις μονάδες μέτρησής της, για αυτό αναμφισβήτητα ο συντελεστής συσχέτισης, που λαμβάνει τιμές από -1 έως +1 και δεν έχει μονάδες μέτρησης, είναι καλύτερο μέτρο μέτρησης της εξάρτησής των δύο μεταβλητών και φυσικά η τιμή του είναι και συγκρίσιμη.

Ως εκ τούτου, η ανάλυση συσχέτισης έχει ευρείες εφαρμογές σε πολλές μελέτες ως μία πρώτη προσπάθεια ανίχνευσης της ύπαρξης μιας γραμμικής σχέσης μεταξύ δύο μεταβλητών, αν και υπάρχουν και άλλοι τρόποι με τους οποίους μπορεί να προσδιοριστεί η εξάρτηση σε ένα ζεύγος μεταβλητών, η ανάγκη των οποίων προέκυψε από τις αδυναμίες του συντελεστή συσχέτισης. Για παράδειγμα, η υπόθεση της γραμμικότητας είναι ένα πολύ σημαντικό θέμα στην ανάλυση της συσχέτισης, η οποία όμως μπορεί να παραβιάζεται, κυρίως σε περιπτώσεις όπου τα φαινόμενα παρουσιάζουν μεγάλη μεταβλητότητα. Έτσι, ακόμα και αν θεωρηθεί ότι η γραμμικότητα μπορεί να εκφράσει κατά προσέγγιση μη γραμμικές σχέσεις, πολύ συχνά η εφαρμογή του συντελεστή συσχέτισης μπορεί να οδηγήσει τον ερευνητή σε λανθασμένα συμπεράσματα αν οι μεταβλητές συνδέονται μη γραμμικά μεταξύ τους ή ακόμα κι αν σχηματίζονται από μη γραμμικές δομές. Λανθασμένα συμπεράσματα μπορεί επίσης να προέλθουν με τη χρήση του συντελεστή συσχέτισης και από τη φύση των χρονοσειρών, όπως έχει αναφερθεί η περίπτωση των υποφαινομενικών συσχετίσεων, και όχι λόγω της μη γραμμικότητας. Επιπρόσθετα, η παρουσία έντονων ακραίων συμπεριφορών αλλά και η απουσία της κανονικότητας μπορούν εύκολα να παραπλανήσουν τον ερευνητή, καθιστώντας το συντελεστή συσχέτισης μη ικανό μέτρο έκφρασης της εξάρτησης των μεταβλητών.





Εναλλακτικά, μπορεί κανείς να εξετάσει τη σχέση δύο μεταβλητών με πιο περίπλοκους τύπους εξαρτήσεων που χρησιμοποιούνται αρκετά σε εξαιρετικά ασταθή δεδομένα, όπως είναι οι τεχνικές *Copula*, και μπορούν να ενσωματώσουν καλύτερα άλλους τύπους σχέσεων μεταξύ δύο μεταβλητών. Σημειώνεται ότι οι τεχνικές *Copula*, όπως έχουν παρουσιαστεί με λεπτομέρεια σε προηγούμενο κεφάλαιο, εκφράζουν μία πολυμεταβλητή συνάρτηση κατανομής της οποίας οι οριακές κατανομές είναι ομοιόμορφες στον κύβο μονάδας  $[0, 1]^m$ , ενώ για τη διμεταβλητή περίπτωση είναι ένα 2-διάστατο, όπου το *Copula* είναι η κοινή κατανομή δύο τυχαίων μεταβλητών, καθεμία από τις οποίες είναι οριακά ομοιόμορφα κατανεμημένη ως  $U(0, 1)$ .

Για την ανάλυση Monte Carlo θα χρησιμοποιηθούν οι τεχνικές *Copula* που εφαρμόζονται περισσότερο συχνά στην πράξη για χρηματοοικονομικές μελέτες και είναι κατά κανόνα οι Ελλειπτικές και οι Αρχιμήδειες (Elliptical και Archimedean) *Copulas*. Οι Ελλειπτικές *Copulas*, οι οποίες είναι απλώς οι συζεύξεις των ελλειπτικών κατανομών, συνδέονται με μια τετραγωνική μορφή συσχέτισης μεταξύ των οριακών κατανομών τους, ενώ η δομή εξάρτησής τους είναι απλώς ο γραμμικός συντελεστής συσχέτισης, όπως οι Gaussian και οι Student-t *Copulas*, που έχουν παρουσιαστεί σε προηγούμενο κεφάλαιο. Σημειώνεται ότι το πλεονέκτημα της Student-t *Copula*, έναντι της Gaussian *Copula*, είναι ότι χρησιμοποιεί την παράμετρο  $n$  για να συλλάβει τη συμμετρική εξάρτηση ακραίου (ή ουράς), ενώ και οι δύο *Copulas* αναφέρονται στην ίδια δομή εξάρτησης, όπως τεκμηριώνεται από τους Embrechts et al. (2003), Jondeau and Rockinger (2006), Ning (2010) και Janus et al. (2014). Αντίθετα, οι Αρχιμήδειες *Copulas* βασίζονται στους μετασχηματισμούς Laplace των συναρτήσεων μονομεταβλητής κατανομής, όπως παρουσιάζονται από τους Genest και MacKay (1986), Joe (1997) και Nelsen (1999), και θα χρησιμοποιηθούν οι τρεις πιο δημοφιλείς μιας παραμέτρου *Copulas* που είναι οι Clayton, Gumbel και Frank. Οι Αρχιμήδειες *Copulas* έχουν σχεδιαστεί για να ξεπερνούν το κύριο μειονέκτημα των Ελλειπτικών *Copulas*, που επικεντρώνονται στο γεγονός ότι προέρχονται από ορισμένες οικογένειες πολυμεταβλητών κατανομών που περιορίζονται να έχουν ακτινική συμμετρία, όπως τονίζεται από τους Embrechts et al. (2002). Τέλος και για να υπάρχει σύγκριση σε γραμμικότητα θα χρησιμοποιηθούν μαζί με το συντελεστή συσχέτισης και οι



συντελεστές συσχέτισης  $\rho$  του Spearman και  $\tau$  του Kendall, δύο μέτρα μέτρησης της εξάρτησης που βασίζονται στην κατάταξη.

#### 4.4 Ανεξαρτησία

Η συμπεριφορά των προαναφερθέντων μέτρων εξάρτησης μεταξύ δύο τυχαίων μεταβλητών  $X$  και  $Y$  που δημιουργούνται από υποδείγματα  $AR(1) - ARCH(1)$  στα σφάλματα και από υποδείγματα  $AR(1) - ARCH(1)$  στη μεταβλητή που δίνονται αντίστοιχα από τις σχέσεις (4.1) με (4.2) και (4.1) με (4.3), με τα δεύτερα να θεωρούνται και ως μη γραμμικά υποδείγματα, διερευνάται χρησιμοποιώντας ανάλυση Monte Carlo. Αναμφισβήτητα, η μη γραμμικότητα μπορεί να παρουσιαστεί με διάφορους τρόπους, πολλούς από τους οποίους δεν είναι και εύκολη η προσέγγισή τους. Για αυτό το λόγο, στην έρευνα αυτή η μη γραμμικότητα παρουσιάζεται με τα  $AR(1) - ARCH(1)$  στη μεταβλητή υποδείγματα, που εμφανίστηκαν από τους Bera et al. (1992 και 1996), καθώς αυτά τα υποδείγματα μπορούν εξίσου να θεωρηθούν και ως μη γραμμικά υποδείγματα που δημιουργούνται από μια διαχρονικά μεταβαλλόμενη αυτοπαλίνδρομη παράμετρο σε μια διάταξη γραμμικού μοντέλου. Σημειώνεται ότι και οι δύο οικογένειες υποδειγμάτων εμφανίζουν πρόβλημα ετεροσκεδαστικότητας στις τιμές του τυχαίου μέρους, απλά η πηγή προέλευσης που προκαλεί τη μεταβολή στη διακύμανση του τυχαίου σφάλματος είναι διαφορετική. Στην πρώτη, η πηγή προέλευσης είναι στις τιμές του τυχαίου μέρους, ενώ στη δεύτερη είναι στις τιμές της μεταβλητής, όπου ειδικότερα σε αυτήν περίπτωση το υπόδειγμα μπορεί να παρουσιαστεί και ως μη γραμμικό.

Έτσι, για το σκοπό της έρευνας δημιουργούνται δύο ανεξάρτητες χρονοσειρές  $X_t$  και  $Y_t$  με βάση τα υποδείγματα  $AR(1) - ARCH(1)$  στα σφάλματα και τα υποδείγματα  $AR(1) - ARCH(1)$  στη μεταβλητή, που δίνονται αντίστοιχα από τις σχέσεις (4.1) με (4.2) και (4.1) με (4.3), για τιμές των παραμέτρων  $\varphi = 0.0, 0.2, 0.5, 0.8$  και  $0.9$ ,  $a_0 = 1$ ,  $a_1 = 0.0, 0.2, 0.5, 0.8$  και  $0.9$ , για μεγέθη δείγματος 100 και 500 παρατηρήσεων και με



το τυχαίο σφάλμα να ακολουθεί την τυπική κανονική κατανομή.<sup>6</sup> Σημειώνεται ότι οι σειρές που προέκυψαν από αυτή τη διαδικασία παραγωγής δεδομένων (Data Generating Process – DGP) περιέχουν τόσο γραμμικές όσο και μη γραμμικές δομές υποδειγμάτων στην μορφή της ετεροσκεδαστικότητας, προκειμένου να απεικονιστεί η μεταβλητότητα ενός φαινομένου, καθώς εάν  $a_1 = 0$ , η μορφή αναπαραγωγής της χρονοσειράς είναι απλώς μια στάσιμη αυτοπαλίνδρομη διαδικασία AR(1). Οι δύο ανεξάρτητες διαδικασίες παραγωγής χρησιμοποιούνται για τον υπολογισμό του συντελεστή συσχέτισης του δείγματος  $r$ , το  $\rho$  του Spearman, το  $\tau$  του Kendall, το *Gaussian Copula*, το *Student-t Copula*, το *Clayton Copula*, το *Gumbel Copula* και το *Frank Copula*, ως μια προσπάθεια αντίληψης της συμπεριφοράς γραμμικών ή άλλων πιο σύνθετων τύπων εξάρτησης για χρονοσειρές με έντονη μεταβλητότητα.

---

<sup>6</sup> Η ανάλυση Monte Carlo εφαρμόστηκε στην Python με το 25% των αρχικών τιμών των χρονοσειρών που δημιουργήθηκαν να διαγράφεται, έτσι ώστε οι εναπομείναντες παρατηρήσεις των χρονοσειρών να εκφράζουν καλύτερα την συμπεριφορά τους, ενώ οι δεκαδικοί αριθμοί συμβολίζονται με τελεία αντί με κόμμα.



### Πίνακας 4.1

Τιμές μέσων όρων για μέτρα εξάρτησης από δύο ανεξάρτητες στάσιμες  
χρονοσειρές AR(1) – ARCH(1) στα σφάλματα βασισμένες σε 10000 επαναλήψεις

$\phi$	$a_1$	Correlation Coefficient $r$	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
0.0	0.0	0.000	0.000	0.000	0.000	0.000	0.005	0.499	1.051
		0.000	0.000	0.000	0.000	0.000	0.001	0.216	1.017
	0.2	0.000	0.000	0.000	0.000	0.000	0.005	0.500	1.051
		0.000	0.000	0.000	0.000	0.000	0.001	0.216	1.017
	0.5	0.000	0.000	0.000	0.000	0.000	0.005	0.501	1.051
		0.000	0.000	0.000	0.000	0.001	0.001	0.216	1.017
0.2	0.8	0.000	0.000	0.000	0.000	0.000	0.005	0.499	1.051
		0.000	0.000	0.000	0.000	0.000	0.001	0.216	1.017
	0.9	0.000	0.000	0.000	0.000	0.000	0.005	0.497	1.051
		0.000	0.000	0.000	0.000	0.000	0.001	0.215	1.017
	0.0	0.001	0.000	0.000	0.001	0.001	0.005	0.511	1.052
		0.000	0.000	0.000	0.000	0.000	0.001	0.221	1.017
0.5	0.2	0.001	0.000	0.000	0.001	0.000	0.005	0.508	1.052
		0.000	0.000	0.000	0.000	0.000	0.001	0.221	1.017
	0.5	0.001	0.000	0.000	0.000	0.000	0.005	0.511	1.052
		0.000	0.000	0.000	0.000	0.000	0.001	0.222	1.017
	0.8	0.000	0.001	0.000	0.000	0.001	0.005	0.510	1.052
		0.000	0.000	0.000	0.000	0.000	0.001	0.220	1.017
0.8	0.9	0.000	0.001	0.000	0.001	0.001	0.005	0.511	1.052
		0.000	0.000	0.000	0.000	0.000	0.001	0.220	1.017
	0.0	0.001	0.001	0.001	0.002	0.001	0.011	0.626	1.065
		0.000	0.000	0.000	0.000	0.000	0.001	0.272	1.022
	0.2	0.001	0.001	0.001	0.001	0.001	0.010	0.622	1.065
		0.000	0.000	0.000	0.000	0.000	0.001	0.270	1.022
0.9	0.5	0.001	0.001	0.001	0.001	0.001	0.009	0.621	1.065
		0.000	0.000	0.000	0.000	0.000	0.001	0.267	1.022
	0.8	0.001	0.001	0.001	0.001	0.001	0.009	0.616	1.064
		0.000	0.000	0.000	0.000	-0.001	0.001	0.267	1.021
	0.9	0.001	0.001	0.001	0.001	0.001	0.009	0.615	1.064
		0.000	0.000	0.000	0.000	0.000	0.001	0.266	1.021
0.0	0.0	0.002	0.002	0.001	0.003	0.002	0.030	1.004	1.112
		-0.001	-0.001	0.000	-0.001	-0.001	0.004	0.443	1.038
	0.2	0.002	0.002	0.001	0.003	0.002	0.030	1.014	1.113
		-0.001	-0.001	-0.001	-0.001	-0.001	0.004	0.446	1.039
	0.5	0.002	0.002	0.002	0.003	0.003	0.031	1.030	1.114
		-0.001	-0.001	-0.001	-0.001	-0.001	0.004	0.450	1.039
0.2	0.8	0.002	0.003	0.002	0.002	0.003	0.031	1.039	1.116
		-0.001	-0.001	-0.001	-0.001	-0.001	0.004	0.454	1.039
	0.9	0.002	0.003	0.002	0.002	0.003	0.031	1.046	1.116
		-0.001	-0.001	-0.001	-0.001	-0.001	0.004	0.455	1.040
	0.0	0.003	0.002	0.002	0.003	0.003	0.054	1.399	1.164
		-0.002	-0.002	-0.001	-0.002	-0.002	0.009	0.643	1.059
0.5	0.2	0.003	0.003	0.002	0.003	0.003	0.055	1.411	1.165
		-0.002	-0.002	-0.001	-0.002	-0.002	0.009	0.640	1.059
	0.5	0.003	0.003	0.002	0.003	0.003	0.057	1.432	1.167
		-0.002	-0.002	-0.001	-0.002	-0.002	0.009	0.645	1.059
	0.8	0.002	0.004	0.003	0.004	0.004	0.059	1.469	1.172
		-0.002	-0.002	-0.001	-0.002	-0.002	0.010	0.669	1.062
0.8	0.9	0.002	0.004	0.003	0.004	0.004	0.059	1.491	1.175
		-0.002	-0.002	-0.001	-0.002	-0.002	0.010	0.675	1.063

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



**Πίνακας 4.2**

**Τιμές μέσων όρων για μέτρα εξάρτησης από δύο ανεξάρτητες στάσιμες  
χρονοσειρές AR(1) – ARCH(1) στη μεταβλητή βασισμένες σε 10000 επαναλήψεις**

$\phi$	$a_1$	Correlation Coefficient $r$	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
0.0	0.0	0.000	0.000	0.000	0.000	0.000	0.005	0.499	1.051
		0.000	0.000	0.000	0.000	0.000	0.001	0.216	1.017
	0.2	0.000	0.000	0.000	0.000	0.000	0.005	0.500	1.051
		0.000	0.000	0.000	0.000	0.000	0.001	0.216	1.017
	0.5	0.000	0.000	0.000	0.000	0.000	0.005	0.501	1.051
		0.000	0.000	0.000	0.000	0.001	0.001	0.216	1.017
0.2	0.8	0.000	0.000	0.000	0.000	0.000	0.005	0.499	1.051
		0.000	0.000	0.000	0.000	0.000	0.001	0.216	1.017
	0.9	0.000	0.000	0.000	0.000	0.000	0.005	0.497	1.051
		0.000	0.000	0.000	0.000	0.000	0.001	0.215	1.017
	0.0	0.001	0.000	0.000	0.001	0.001	0.005	0.511	1.052
		0.000	0.000	0.000	0.000	0.000	0.001	0.221	1.017
0.5	0.2	0.001	0.000	0.000	0.001	0.000	0.005	0.508	1.052
		0.000	0.000	0.000	0.000	0.000	0.001	0.222	1.017
	0.5	0.001	0.000	0.000	0.001	0.001	0.005	0.509	1.052
		0.000	0.000	0.000	0.000	0.000	0.001	0.221	1.017
	0.8	0.000	0.001	0.000	0.001	0.001	0.005	0.511	1.052
		0.000	0.000	0.000	0.000	0.000	0.001	0.220	1.017
0.8	0.9	0.000	0.001	0.001	0.001	0.001	0.005	0.509	1.052
		0.000	0.000	0.000	0.000	0.000	0.001	0.220	1.017
	0.0	0.001	0.001	0.001	0.002	0.001	0.011	0.626	1.065
		0.000	0.000	0.000	0.000	0.000	0.001	0.272	1.022
	0.2	0.001	0.001	0.001	0.001	0.001	0.009	0.609	1.064
		0.000	0.000	0.000	0.000	0.000	0.001	0.265	1.021
0.9	0.5	0.001	0.001	0.001	0.002	0.002	0.008	0.590	1.062
		0.000	0.000	0.000	0.000	0.000	0.001	0.254	1.021
	0.8	0.001	0.001	0.001	0.001	0.001	0.007	0.575	1.060
		0.000	0.000	0.000	0.000	0.000	0.001	0.248	1.020
	0.9	0.001	0.001	0.001	0.002	0.001	0.006	0.572	1.060
		0.000	0.000	0.000	0.000	0.000	0.001	0.245	1.020
0.0	0.0	0.002	0.002	0.001	0.003	0.002	0.030	1.004	1.112
		-0.001	-0.001	0.000	-0.001	-0.001	0.004	0.443	1.038
	0.2	0.002	0.003	0.002	0.003	0.003	0.024	0.916	1.102
		-0.001	-0.001	-0.001	-0.001	-0.001	0.002	0.402	1.035
	0.5	0.003	0.003	0.002	0.003	0.003	0.018	0.805	1.087
		-0.001	-0.001	-0.001	-0.001	-0.001	0.000	0.352	1.030
0.2	0.8	0.003	0.002	0.002	0.003	0.003	0.013	0.735	1.078
		0.000	-0.001	-0.001	-0.001	-0.001	-0.001	0.318	1.027
	0.9	0.003	0.002	0.001	0.002	0.002	0.011	0.713	1.075
		0.000	-0.001	-0.001	-0.001	-0.001	-0.001	0.310	1.026
	0.0	0.003	0.002	0.002	0.003	0.003	0.054	1.399	1.164
		-0.002	-0.002	-0.001	-0.002	-0.002	0.009	0.643	1.059
0.5	0.2	0.002	0.003	0.002	0.003	0.003	0.039	1.193	1.136
		-0.002	-0.002	-0.001	-0.002	-0.002	0.003	0.522	1.047
	0.5	0.003	0.003	0.002	0.004	0.003	0.023	0.932	1.103
		-0.001	-0.001	-0.001	-0.001	-0.001	0.000	0.408	1.035
	0.8	0.002	0.003	0.002	0.003	0.003	0.016	0.820	1.088
		0.000	-0.001	-0.001	-0.001	-0.002	-0.001	0.355	1.031
0.8	0.9	0.002	0.003	0.002	0.003	0.003	0.014	0.791	1.084
		0.000	-0.001	-0.001	-0.001	-0.001	-0.001	0.346	1.030

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.

Στους Πίνακες 4.1 και 4.2 παρουσιάζονται οι τιμές των μέσων όρων όλων των μέτρων εξάρτησης που χρησιμοποιούνται στη μελέτη αυτή για δύο ανεξάρτητες χρονοσειρές που κατασκευάστηκαν από την προαναφερθείσα διαδικασία παραγωγής DGP για



10000 επαναλήψεις.<sup>7</sup> Όπως φαίνεται από αυτούς πίνακες αυτούς, οι τιμές των μέσων όρων όλων των μέτρων εξάρτησης είναι πολύ κοντά στις αναμενόμενες τιμές τους που δηλώνουν ανεξαρτησία, ανεξάρτητα από το μέγεθος του δείγματος, τις τιμές και των δύο παραμέτρων,  $\varphi$  και  $a_1$ , καθώς και τη μορφή της ετεροσκεδαστικότητας αλλά και το είδος των χρονοσειρών που μπορεί να προκύπτουν, δηλαδή γραμμικές ή μη, σημειώνοντας ωστόσο ότι οι Αρχιμήδειες *Copulas* φαίνεται να επηρεάζονται ελαφρώς από τα χαρακτηριστικά των χρονοσειρών. Πιο συγκεκριμένα, οι τιμές των μέσων όρων των Clayton, Gumbel και Frank *Copulas* έχουν την τάση να τείνουν προς τις τιμές της ανεξαρτησίας τους, καθώς η τιμή της παραμέτρου  $a_1$  αυξάνεται, για δεδομένες τιμές των  $\varphi$  και  $n$ , καθώς το μέγεθος του δείγματος αυξάνεται, για δεδομένες τιμές των  $\varphi$  και  $a_1$  και καθώς η τιμή της παραμέτρου του αυτοπαλίνδρομου όρου μειώνεται, για δεδομένες τιμές των  $n$  και  $a_1$ . Επίσης, για τις *Copulas* αυτές εμφανίζεται μία ελαφρώς μεγαλύτερη τιμή για τις χρονοσειρές που δημιουργούνται από ετεροσκεδαστικότητα στα σφάλματα από ότι στις τιμές της μεταβλητής. Ενδεχομένως, οι Αρχιμήδειες αυτές *Copulas* να αντιλαμβάνονται την έννοια της ανεξαρτησίας ελαφρώς διαφορετικά από όλα τα άλλα μέτρα εξάρτησης για μεταβλητές που δημιουργούνται από πιο περίπλοκες δομές.

#### 4.5 Υποφαινομενικές συσχετίσεις

Το θέμα του καθορισμού της ανεξαρτησίας μεταξύ δύο μεταβλητών είναι αρκετά περίπλοκο και μπορεί να κρύβει λανθασμένες ενδείξεις, δεδομένου ότι η χρήση του συντελεστή συσχέτισης σε δεδομένα μεταβλητών που προέρχονται από χρονοσειρές μπορεί να εκφράζει υποφαινομενικές σχέσεις, όπως αναφέρθηκε σε προηγούμενο κεφάλαιο. Πραγματικά, οι τιμές των μέσων όρων του συντελεστή συσχέτισης που δίνονται από τους Πίνακες 4.1 και 4.2 είναι όλες μηδέν, για όλες τις περιπτώσεις των χρονοσειρών που αναπαράχθηκαν, δηλώνοντας ότι οι μεταβλητές αυτές είναι ανεξάρτητες μεταξύ τους. Ωστόσο, ο Yule (1926), που εισήγαγε το φαινόμενο των

---

<sup>7</sup> Η ανάλυση Monte Carlo εφαρμόστηκε και για χρονοσειρές με 50 και 1000 παρατηρήσεις χωρίς να αλλάζουν καθόλου τα αποτελέσματα.



υποφαινομενικών συσχετίσεων για δύο ανεξάρτητες χρονοσειρές μέσω της ανάλυσης συσχέτισης, παρατήρησε ότι ο συντελεστής συσχέτισης μπορεί να παράγει ψευδή αποτελέσματα, κάτι που μπορεί να ανιχνευθεί από την κατανομή συχνότητας του συντελεστή συσχέτισης του δείγματος. Μάλιστα, έχοντας ως σημείο αναφοράς το υπόδειγμα του λευκού θορύβου, δήλωσε ότι όταν η κατανομή συχνότητας του συντελεστή συσχέτισης του δείγματος είναι συμμετρική στο μηδέν, τότε δεν υπάρχουν ενδείξεις υποφαινομενικών συσχετίσεων. Την τοποθέτηση αυτή διερεύνησαν οι Agiakloglou και Tsimpanos (2012) όπου και διαπίστωσαν ότι για δύο ανεξάρτητες στάσιμες χρονοσειρές  $AR(1)$  η κατανομή συχνότητας του συντελεστή συσχέτισης του δείγματος είναι συμμετρική γύρω από το μηδέν, όπως αυτό προκύπτει από διαδικασίες λευκού θορύβου, εύρημα σύμφωνα με το οποίο δεν θα έπρεπε να υπήρχαν ενδείξεις υποφαινομενικών σχέσεων.

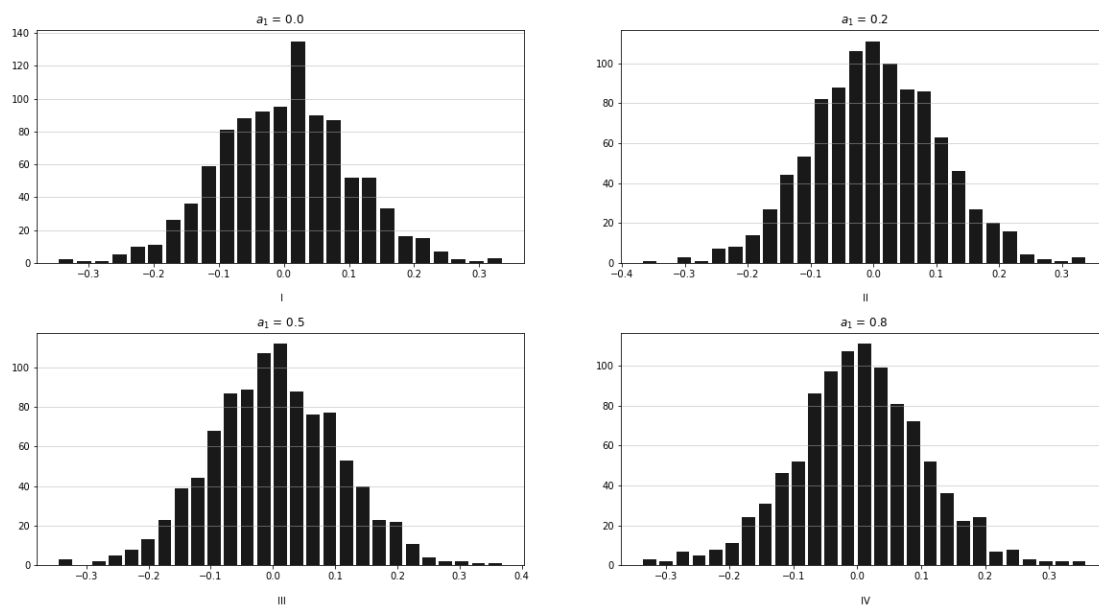
Κατά ανάλογο τρόπο οι διαγραμματικές απεικονίσεις της κατανομής συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές  $AR(1) - ARCH(1)$  στα σφάλματα, όπως ενδεικτικά παρουσιάζονται στα Διαγράμματα 4.6 και 4.7 για τιμές του  $\varphi = 0.2$  και  $0.8$  αντίστοιχα, και για τιμές του  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$ , βασισμένα σε 10000 επαναλήψεις και για μεγέθη δείγματος 100 παρατηρήσεων, δείχνουν να είναι συμμετρικές στο μηδέν και πολύ όμοιες μεταξύ τους ανεξάρτητα από την τιμή της παραμέτρου  $a_1$ .<sup>8</sup> Η παρατήρηση αυτή δηλώνει ότι η παρουσία ετεροσκεδαστικότητας που προέρχεται από τις τιμές του τυχαίου σφάλματος δεν φαίνεται να επηρεάζει την ήδη γνωστή συμπεριφορά της κατανομής του συντελεστή συσχέτισης για δύο ανεξάρτητες στάσιμες χρονοσειρές  $AR(1)$ , συμπεριφορά η οποία άλλωστε απεικονίζεται και στα διαγράμματα αυτά, δεδομένου ότι όταν η τιμή της παραμέτρου  $a_1$  είναι ίση με το μηδέν τα υποδείγματα έχουν σταθερή διακύμανση, δηλαδή δεν έχουν ετεροσκεδαστικότητα, και είναι απλά στάσιμες αυτοπαλίνδρομου πρώτου βαθμού χρονοσειρές, δηλαδή στάσιμες  $AR(1)$ . Κατά συνέπεια η μορφή αυτή της ετεροσκεδαστικότητας, όπως εμφανίζεται σε αυτά τα υποδείγματα, δεν θα πρέπει να

---

<sup>8</sup> Παρόμοια διαγραμματική απεικόνιση εμφανίζεται και για δύο ανεξάρτητες χρονοσειρές  $AR(1) - ARCH(1)$  στα σφάλματα του  $\varphi = 0.0, 0.5$  και  $0.9$  αντίστοιχα, και για τιμές του  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$  βασισμένες σε 10000 επαναλήψεις και για μέγεθος δείγματος 100 παρατηρήσεων.



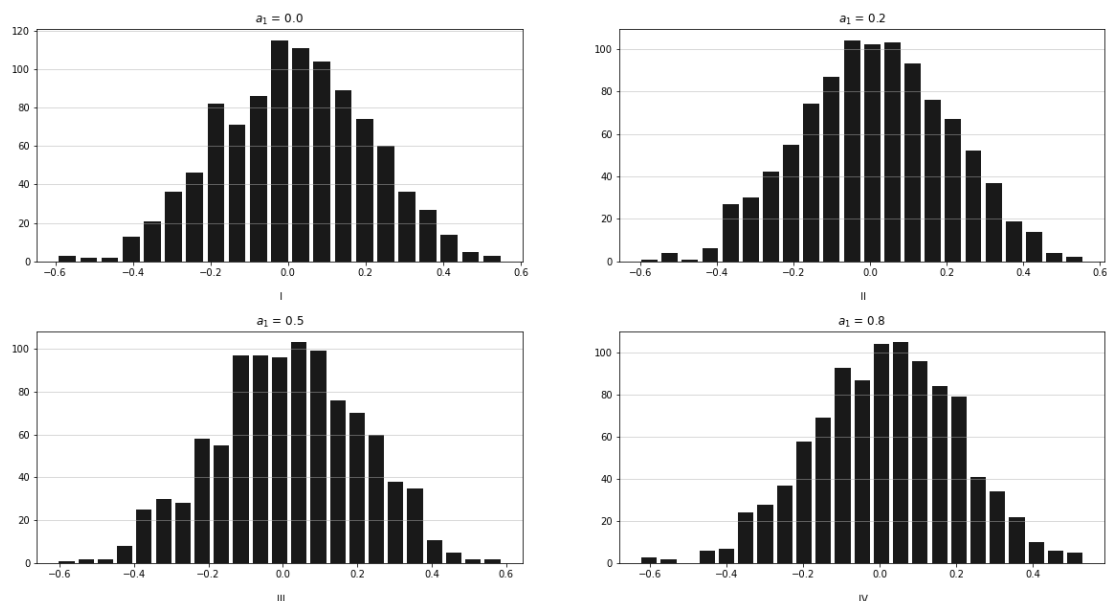
αναμένεται να αλλάξει τον τρόπο συμπεριφοράς του φαινομένου των υποφαινομενικών συσχετίσεων.



**Διάγραμμα 4.6**

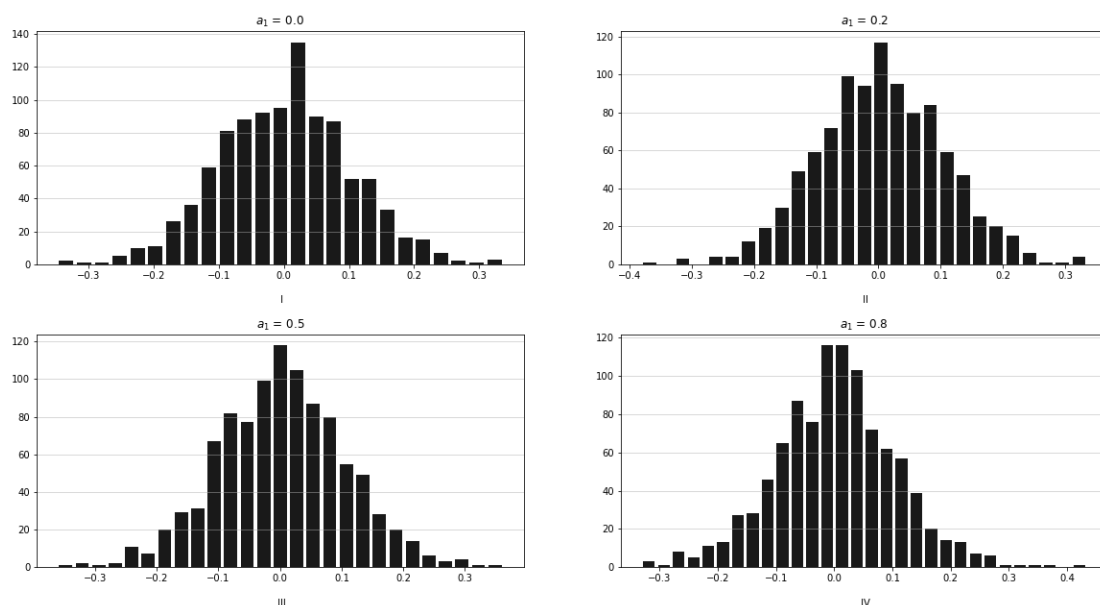
**Κατανομή συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές AR(1) – ARCH(1) στα σφάλματα για  $\phi = 0.2$  και για  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$  βασισμένες σε 10000 επαναλήψεις και για μεγέθη δείγματος 100**





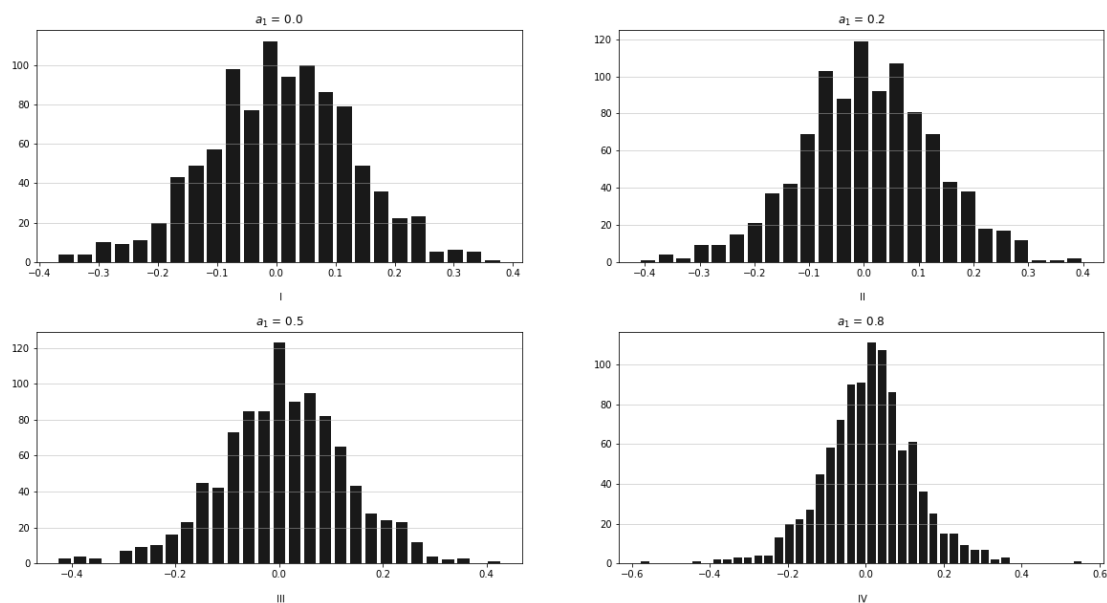
Διάγραμμα 4.7

Κατανομή συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές AR(1) – ARCH(1) στα σφάλματα για  $\phi = 0.8$  και για  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$  βασισμένες σε 10000 επαναλήψεις και για μεγέθη δείγματος 100



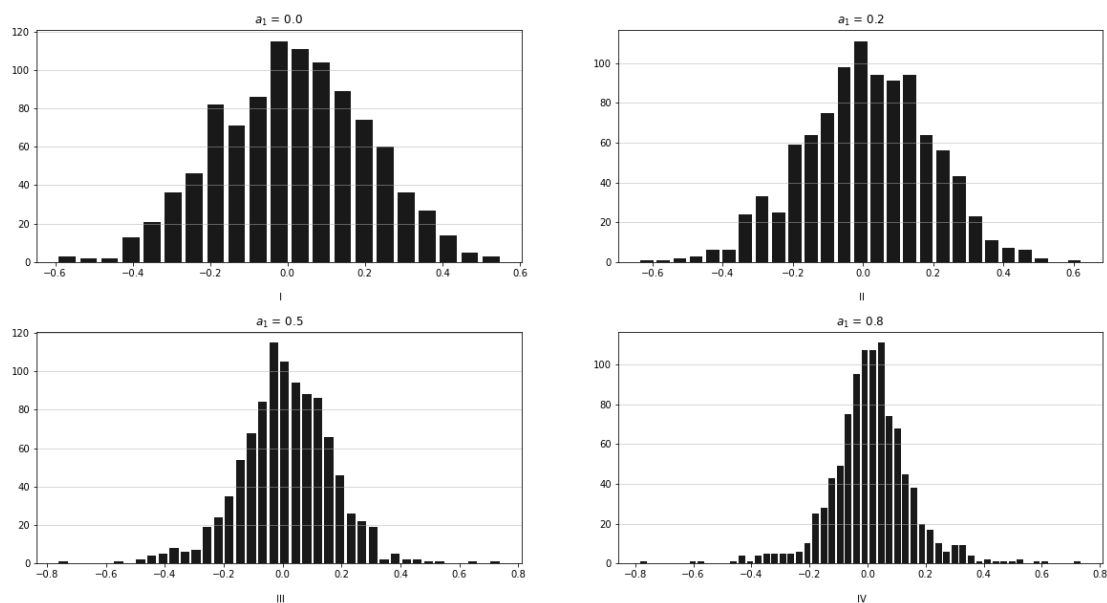
Διάγραμμα 4.8

Κατανομή συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές AR(1) – ARCH(1) στη μεταβλητή για  $\phi = 0.2$  και για  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$  βασισμένες σε 10000 επαναλήψεις και για μεγέθη δείγματος 100



Διάγραμμα 4.9

Κατανομή συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές AR(1) – ARCH(1) στη μεταβλητή για  $\varphi = 0.5$  και για  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$  βασισμένες σε 10000 επαναλήψεις και για μεγέθη δείγματος 100



Διάγραμμα 4.10

Κατανομή συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές AR(1) – ARCH(1) στη μεταβλητή για  $\varphi = 0.8$  και για  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$  βασισμένες σε 10000 επαναλήψεις και για μεγέθη δείγματος 100



Αντίθετα, στα Διαγράμματα 4.8 έως 4.10 απεικονίζεται η κατανομή συχνότητας του συντελεστή συσχέτισης για δύο ανεξάρτητες χρονοσειρές  $AR(1) - ARCH(1)$  στη μεταβλητή για τιμές του  $\varphi = 0.2, 0.5$  και  $0.8$ , αντίστοιχα, και για τιμές του  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$ , βασισμένα σε 10000 επαναλήψεις και για μεγέθη δείγματος 100 παρατηρήσεων. Από τα διαγράμματα αυτά είναι εμφανές ότι η συμπεριφορά της κατανομής συχνοτήτων του συντελεστή συσχέτισης των δύο μεταβλητών δεν είναι όμοια με αυτή των δύο ανεξάρτητων στάσιμων  $AR(1)$  χρονοσειρών, όπως στη προηγούμενη περίπτωση των ανεξάρτητων χρονοσειρών  $AR(1) - ARCH(1)$  στα σφάλματα. Η συμμετρία της γραφικής παράστασης της κατανομής συχνότητας του συντελεστή συσχέτισης του δείγματος γύρω από το μηδέν παραμένει ακόμη και για αυτές τις χρονοσειρές όπου η ετεροσκεδαστικότητα έρχεται από τις τιμές της μεταβλητής. Ωστόσο, στην περίπτωση αυτή το σχήμα των κατανομών γίνεται διαφορετικό και πιο απότομο με τις παρατηρήσεις να συγκεντρώνονται πιο έντονα στο κέντρο της κατανομής, δηλαδή στο μηδέν, καθώς η τιμή της παραμέτρου  $a_1$  αυξάνει για κάθε τιμή της παραμέτρου  $\varphi$ . Η διαφορετική αυτή διαγραμματική ένδειξη για αυτά τα υποδείγματα, στα οποία η ετεροσκεδαστικότητα προέρχεται από τις τιμές της μεταβλητής και τα οποία ομοιάζουν με μη γραμμικές ανελίξεις, πολύ πιθανότατα να διαφοροποιεί τη συμπεριφορά του φαινομένου της υποφαινομενικής συσχέτισης.

Η συμπεριφορά όμως του φαινομένου των υποφαινομενικών συσχετίσεων εξετάζεται με τον κλασικό στατιστικό έλεγχο της υπόθεσης της μηδενικής συσχέτισης, από τον οποίον, όπως έχει αναφερθεί σε προηγούμενο κεφάλαιο, παράγονται παραπλανητικά αποτελέσματα για δύο ανεξάρτητες στάσιμες αλλά και μη στάσιμες χρονοσειρές. Στην πραγματικότητα, το ποσοστό των απορρίψεων της μηδενικής υπόθεσης για δύο ανεξάρτητες στάσιμες χρονοσειρές  $AR(1)$  αυξάνεται καθώς αυξάνεται η τιμή του αυτοπαλίνδρουμου όρου, ανεξάρτητα από το μέγεθος του δείγματος, φτάνοντας σε επίπεδο 52% για  $\varphi = 0.9$  σε ονομαστικό 5% επίπεδο σημαντικότητας, εύρημα που σχετίζεται με την κατανομή της σχετικής  $t$  - στατιστικής, δεδομένου ότι η τυπική απόκλιση της αποκλίνει από το 1 όσο αυξάνεται η τιμή του  $\varphi$ , ενώ ο μέσος όρος της παραμένει στο μηδέν, όπως δείχνουν στην εργασία τους οι Αγιακλόγλου και Τσιμπάνος (2012) για δύο ανεξάρτητες στάσιμες  $AR(1)$  χρονοσειρές.



Αρχικά, ο έλεγχος για τη διερεύνηση της μηδενικής υπόθεσης ότι δεν υπάρχει γραμμική σχέση μεταξύ των μεταβλητών, δηλαδή  $H_0: \rho = 0$ , έναντι της εναλλακτικής υπόθεσης ότι υπάρχει, δηλαδή  $H_1: \rho \neq 0$ , εφαρμόστηκε σε δύο ανεξάρτητες χρονοσειρές που δημιουργήθηκαν από υποδείγματα  $AR(1) - ARCH(1)$  στα σφάλματα με βάση την υφιστάμενη δομή παραγωγής των χρονοσειρών DGP, όπου μετά από τον προσδιορισμό της τιμής του δειγματικού συντελεστή συσχέτισής τους υπολογίστηκε η ακόλουθη στατιστική:

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

η οποία ακολουθεί την  $t$  κατανομή με  $(n-2)$  βαθμούς ελευθερίας και κατά συνέπεια η μηδενική υπόθεση απορρίπτεται όταν η τιμή της είναι μεγαλύτερη της αντίστοιχης κριτικής τιμής.<sup>9</sup>

---

<sup>9</sup> Για την εφαρμογή του στατιστικού ελέγχου χρησιμοποιήθηκαν κριτικές τιμές από την τυπική κανονική κατανομή, ήτοι 1,645 και 1,96 για 10% και 5% επίπεδο σημαντικότητας αντίστοιχα.



### Πίνακας 4.3

Ποσοστά απόρριψης της υπόθεσης της μηδενικής συσχέτισης για 10% και 5% ονομαστικό επίπεδο σημαντικότητας για δύο ανεξάρτητες στάσιμες χρονοσειρές AR(1) – ARCH(1) στα σφάλματα μαζί με τις τιμές του μέσου όρου και της τυπικής απόκλισης της στατιστικής t που χρησιμοποιήθηκε βασισμένες σε 10000 επαναλήψεις

$\varphi$	$a_1$	Mean values	Standard deviation	$\alpha = 0.10$	$\alpha = 0.05$
0.0	0.0	0.004	1.005	10.21	5.12
		0.008	1.006	10.25	5.17
	0.2	0.004	1.004	10.36	5.21
		0.008	1.005	10.34	5.32
	0.5	0.004	1.006	10.11	5.18
		0.009	1.000	9.80	5.22
	0.8	0.001	1.011	10.21	5.38
		0.010	0.992	9.31	5.02
	0.9	0.000	1.014	10.00	5.79
		0.010	0.996	9.12	5.23
0.2	0.0	0.005	1.042	11.23	5.98
		0.003	1.048	11.71	6.32
	0.2	0.006	1.042	11.32	5.99
		0.002	1.047	11.77	6.21
	0.5	0.005	1.044	11.18	6.21
		0.004	1.040	11.22	5.93
	0.8	0.004	1.048	11.11	6.40
		0.005	1.032	10.55	5.96
	0.9	0.003	1.050	11.22	6.60
		0.004	1.034	9.97	5.95
0.5	0.0	0.011	1.290	19.78	12.77
		-0.006	1.300	20.65	13.17
	0.2	0.011	1.289	19.59	12.41
		-0.008	1.298	20.41	13.35
	0.5	0.011	1.286	19.18	12.16
		-0.008	1.291	20.08	12.79
	0.8	0.010	1.279	18.80	11.79
		-0.008	1.276	18.30	11.55
	0.9	0.009	1.278	18.56	11.76
		-0.008	1.273	17.39	11.14
0.8	0.0	0.025	2.138	43.06	34.99
		-0.017	2.144	43.67	35.49
	0.2	0.024	2.139	43.03	35.03
		-0.023	2.142	43.42	35.07
	0.5	0.022	2.130	42.97	34.79
		-0.028	2.134	43.15	35.12
	0.8	0.018	2.105	41.63	33.29
		-0.029	2.112	41.65	33.29
	0.9	0.017	2.096	41.04	32.80
		-0.029	2.100	40.48	32.12
0.9	0.0	0.027	3.043	57.33	50.54
		-0.043	3.108	58.81	51.83
	0.2	0.026	3.050	57.13	50.54
		-0.049	3.108	58.32	51.15
	0.5	0.025	3.048	57.33	50.46
		-0.053	3.101	58.27	51.58
	0.8	0.022	3.028	56.70	49.75
		-0.048	3.075	57.75	50.42
	0.9	0.021	3.021	56.83	49.16
		-0.046	3.058	56.86	49.49

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



#### Πίνακας 4.4

Ποσοστά απόρριψης της υπόθεσης της μηδενικής συσχέτισης για 10% και 5% ονομαστικό επίπεδο σημαντικότητας για δύο ανεξάρτητες στάσιμες χρονοσειρές AR(1) – ARCH(1) στη μεταβλητή μαζί με τις τιμές του μέσου όρου και της τυπικής απόκλισης της στατιστικής  $t$  που χρησιμοποιήθηκε βασισμένες σε 10000 επαναλήψεις

$\phi$	$a_1$	Mean values	Standard deviations	$\alpha = 0.10$	$\alpha = 0.05$
0.00	0.00	0.004	1.005	10.21	5.12
		0.008	1.006	10.25	5.17
	0.20	0.004	1.004	10.36	5.21
		0.008	1.005	10.34	5.32
	0.50	0.004	1.006	10.11	5.18
		0.009	1.000	9.80	5.22
0.20	0.80	0.001	1.011	10.21	5.38
		0.010	0.992	9.31	5.02
	0.90	0.000	1.014	10.00	5.79
		0.010	0.996	9.12	5.23
	0.00	0.005	1.042	11.23	5.98
		0.003	1.048	11.71	6.32
0.50	0.20	0.006	1.041	11.35	5.97
		0.003	1.045	11.86	6.27
	0.50	0.005	1.044	10.98	6.17
		0.005	1.036	11.09	6.01
	0.80	0.005	1.045	10.94	6.15
		0.006	1.022	10.01	5.77
0.80	0.90	0.005	1.045	10.69	6.16
		0.005	1.021	9.73	5.56
	0.00	0.011	1.290	19.78	12.77
		-0.006	1.300	20.65	13.17
	0.20	0.010	1.277	19.26	12.10
		-0.007	1.290	20.15	13.08
0.90	0.50	0.009	1.242	17.53	10.88
		-0.006	1.252	17.78	11.02
	0.80	0.012	1.204	15.06	9.50
		-0.002	1.216	13.71	8.58
	0.90	0.015	1.193	14.03	9.08
		0.001	1.210	12.30	7.91
0.90	0.00	0.025	2.138	43.06	34.99
		-0.017	2.144	43.67	35.49
	0.20	0.021	1.972	38.03	30.40
		-0.033	2.026	39.24	31.23
	0.50	0.025	1.690	26.83	19.85
		-0.027	1.682	24.04	17.36
0.90	0.80	0.026	1.502	18.53	13.11
		-0.006	1.456	13.02	9.72
	0.90	0.026	1.453	16.73	11.98
		-0.004	1.417	11.30	8.65
	0.00	0.027	3.043	57.33	50.54
		-0.043	3.108	58.81	51.83
0.90	0.20	0.018	2.449	46.67	38.85
		-0.058	2.591	46.18	38.83
	0.50	0.028	1.894	29.29	22.84
		-0.033	1.867	22.37	17.02
	0.80	0.019	1.637	19.07	14.24
		0.001	1.548	12.25	9.41
0.90	0.90	0.022	1.579	17.19	12.88
		-0.001	1.501	10.58	8.39

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



Τα ποσοστά απόρριψης της μηδενικής υπόθεσης για όλους τους συνδυασμούς των παραμέτρων  $\varphi$  και  $a_1$ , καθώς και οι τιμές των μέσων όρων αλλά και της τυπικής απόκλισης της στατιστικής  $t$  για χρονοσειρές των 100 και 500 παρατηρήσεων δίνονται στον Πίνακα 4.3 βασισμένα σε 10000 επαναλήψεις. Από τον πίνακα αυτόν είναι φανερό ότι δεν παρατηρείται καμία σημαντική μεταβολή της συμπεριφοράς του φαινομένου σε σχέση με τη συμπεριφορά που έχει για δύο ανεξάρτητες στάσιμες AR(1) χρονοσειρές που έχουν σταθερή διακύμανση, δεδομένου ότι τα ποσοστά απόρριψης της μηδενικής υπόθεσης παρέμειναν σχεδόν αναλλοίωτα. Αυτό σημαίνει ότι η ύπαρξη ετεροσκεδαστικότητας με τη μορφή ενός υποδείγματος ARCH(1) στα σφάλματα δεν άλλαξε καθόλου τη συμπεριφορά του φαινομένου, κάτι που επίσης επιβεβαιώνεται και από τις τιμές των μέσων όρων αλλά και της τυπικής απόκλισης της στατιστικής  $t$ . Πραγματικά, ο μέσο όρος των τιμών της στατιστικής είναι κοντά στο μηδέν, αλλά η τιμή της τυπικής του απόκλισης αυξάνει όσο αυξάνει η τιμή του αυτοπαλίνδρομου όρου, με αποτέλεσμα η κατανομή της στατιστικής να αποκλίνει από την τυπική κανονική κατανομή και έτσι να αυξάνονται τα ποσοστά απόρριψης της μηδενικής υπόθεσης.

Στη συνέχεια, το φαινόμενο των υποφαινομενικών συσχετίσεων, μέσω του στατιστικού ελέγχου της μηδενικής συσχέτισης, διερευνήθηκε για δύο ανεξάρτητες χρονοσειρές που δημιουργήθηκαν από υποδείγματα AR(1) – ARCH(1) στη μεταβλητή, με βάση την υφιστάμενη δομή παραγωγής των χρονοσειρών DGP. Όπως παρατηρείται από τον Πίνακα 4.4 η συμπεριφορά του φαινομένου άλλαξε δραστικά, δεδομένου ότι το ποσοστό απόρριψης της μηδενικής υπόθεσης μειώνεται καθώς η τιμή της παραμέτρου  $a_1$  αυξάνεται για κάθε δεδομένη τιμή του  $\varphi$  φτάνοντας ασυμπτωτικά τα ονομαστικά επίπεδα απόρριψης του στατιστικού ελέγχου. Για παράδειγμα, για  $\varphi = 0.9$  η μηδενική υπόθεση απορρίπτεται 51.8% και 8.4% για μέγεθος δείγματος 500 παρατηρήσεων και για τιμές  $a_1 = 0.0$  και 0.9 αντίστοιχα για ονομαστικό επίπεδο 5%. Τα αποτελέσματα αυτής της συμπεριφοράς επιβεβαιώνονται και από τις τιμές του μέσου όρου και της τυπικής απόκλισης της στατιστικής  $t$ , από όπου παρατηρείται ότι καθώς αυξάνονται οι τιμές των παραμέτρων των χρονοσειρών η κατανομή της συγκλίνει στην τυπική κανονική κατανομή με μέσο όρο μηδέν και τυπική απόκλιση ίση με τη μονάδα. Κατά συνέπεια, η μορφή αυτή της ετεροσκεδαστικότητας, η οποία όμως συνδέεται με μη



γραμμικά υπόδειγμα, περιορίζει έως και εξαφανίζει την εμφάνιση υποφαινομενικών συσχετίσεων κυρίως όσο αυξάνεται η τιμή της παραμέτρου  $a_1$ , όπου για τα δεδομένα αυτής της προσέγγισης η παράμετρος αυτή εκφράζει το βαθμό της μη γραμμικότητας. Επομένως, θα μπορούσε να ειπωθεί ότι η υποφαινομενική συσχέτιση είναι φαινόμενο που αφορά γραμμικά υποδείγματα.

#### 4.6 Γραμμική εξάρτηση

Τα αποτελέσματα της προσομοίωσης που διεξήχθησαν για δύο ανεξάρτητες μεταβλητές έδειξαν ότι η συμπεριφορά όλων των μέτρων εξάρτησης που χρησιμοποιήθηκαν ήταν πάρα πολύ ικανοποιητική και ότι η ετεροσκεδαστικότητα, με όποια μορφή κι αν εμφανίστηκε, δεν επηρέασε τα μέτρα αυτά στο να αναγνωρίσουν σωστά την ανεξαρτησία των μεταβλητών. Στη συνέχεια, το ενδιαφέρον της έρευνας επικεντρώνεται στη διερεύνηση της συμπεριφοράς αυτών των μέτρων εξάρτησης στο να αναγνωρίζουν σχέσεις εξάρτησης μεταξύ δύο μεταβλητών. Παρόλο που η δημιουργία μη γραμμικών αλλά και άλλων πολύπλοκων σχέσεων εξάρτησης δεν είναι εύκολη υπόθεση, ωστόσο, για τα δεδομένα των χρονοσειρών της προσομοίωσης που χρησιμοποιούνται στην μελέτη αυτή, υπάρχει η δυνατότητα να σχηματιστούν γραμμικές σχέσεις εξάρτησης μεταξύ δύο μεταβλητών ακόμα και για τις περιπτώσεις των μη γραμμικών χρονοσειρών.

Ακολουθώντας τη μεθοδολογία που παρουσιάστηκε από τους Agiakloglou και Agiropoulos (2016) για την προσομοίωση γραμμικά συσχετιζόμενων δύο στάσιμων αυτοπαλίνδρομου πρώτου βαθμού AR(1) χρονοσειρών δημιουργούνται γραμμικά εξαρτημένες σχέσεις με την ακόλουθη διαδικασία. Έστω οι δύο χρονοσειρές:

$$x_t = \varphi_1 x_{t-1} + v_t \text{ και } y_t = \varphi_2 y_{t-1} + w_t$$

όπου και οι δύο παράμετροι λαμβάνουν τιμές σε απόλυτο μέγεθος μικρότερες της μονάδας και έχουν γραμμική σχέση μεταξύ τους που ορίζεται από το συντελεστή συσχέτισής τους  $\rho = \text{Corr}(x_t, y_t)$ . Οι δύο αυτές χρονοσειρές μπορούν να γραφτούν





από κοινού ως ένα διανυσματικό αυτοπαλίνδρομο υπόδειγμα VAR (Vector Autoregressive Model) πρώτου βαθμού, δηλαδή ως:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \varphi_1 & 0 \\ 0 & \varphi_2 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} v_t \\ w_t \end{bmatrix} \quad (4.6)$$

ή ως της μορφής:

$$Z_t = FZ_{t-1} + \eta_t \quad (4.7)$$

όπου  $Z_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}$ ,  $F = \begin{bmatrix} \varphi_1 & 0 \\ 0 & \varphi_2 \end{bmatrix}$  και  $\eta_t = \begin{bmatrix} v_t \\ w_t \end{bmatrix}$ . Επίσης, με βάση το διαχωρισμό κατά Cholesky (Cholesky Decomposition), η σχέση (4.7) μπορεί να γραφτεί και ως εξής:

$$\Sigma = F\Sigma'F + Q \quad (4.8)$$

όπου  $\Sigma$  και  $Q$  είναι οι μήτρες συνδιακυμάνσεων των μεταβλητών  $Z_t$  και  $\eta_t$  αντίστοιχα που ορίζονται ως:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \quad \text{και} \quad Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

ενώ η λύση της σχέσης (4.8) δίνεται από τη σχέση:

$$\text{vec}\Sigma = [I - F \otimes F']^{-1} \text{vec}Q \quad (4.9)$$

όπου  $\text{vec}$  είναι η διανυσματοποίηση,  $\otimes$  είναι το γινόμενο *Kronecker* και  $I$  είναι η μοναδιαία μήτρα, η οποία μάλιστα γράφεται και ως εξής:

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{12} \\ \sigma_{21} \\ \sigma_{22} \end{bmatrix} = \left( \begin{bmatrix} 1 - \varphi_1^2 & 0 & 0 & 0 \\ 0 & 1 - \varphi_1\varphi_2 & 0 & 0 \\ 0 & 0 & 1 - \varphi_1\varphi_2 & 0 \\ 0 & 0 & 0 & 1 - \varphi_2^2 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 \\ q_{12} \\ q_{21} \\ 1 \end{bmatrix} \quad (4.10)$$



δεδομένου ότι πρόκειται για δύο μεταβλητές, ενώ θα πρέπει να σημειωθεί ότι οι όροι  $\sigma_{11}$  και  $\sigma_{22}$  δηλώνουν τις διακυμάνσεις των μεταβλητών  $x_t$  και  $y_t$ , αντίστοιχα, που προκύπτουν από ένα  $AR(1)$  υπόδειγμα, ενώ οι όροι  $\sigma_{12}$  και  $\sigma_{21}$  εκφράζουν τη συνδιακύμανση των μεταβλητών και είναι ίσα μεταξύ τους, ενώ ανάλογα ισχύει και για τους όρους  $q_{12}$  και  $q_{21}$  που εκφράζουν τη συνδιακύμανση των σφαλμάτων με τη διακύμανσή τους να λαμβάνεται ίση με μονάδα, δηλαδή  $q_{11} = q_{22} = 1$ .

Λύνοντας τη σχέση (4.10) προκύπτει η ακόλουθη σχέση:

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{21} \\ \sigma_{12} \\ \sigma_{22} \end{bmatrix} = \begin{pmatrix} \frac{1}{1-\varphi_1^2} & 0 & 0 & 0 \\ 0 & \frac{1}{1-\varphi_1\varphi_2} & 0 & 0 \\ 0 & 0 & \frac{1}{1-\varphi_1\varphi_2} & 0 \\ 0 & 0 & 0 & \frac{1}{1-\varphi_2^2} \end{pmatrix} \begin{bmatrix} 1 \\ q_{21} \\ q_{12} \\ 1 \end{bmatrix}$$

σύμφωνα με την οποία υπολογίζονται όλες οι τιμές των  $\sigma$  και κατ' επέκταση ο συντελεστής συσχέτισης  $\rho$  των δύο μεταβλητών ως εξής:

$$\rho = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}} = \frac{q_{12}\sqrt{(1-\varphi_1^2)}\sqrt{(1-\varphi_2^2)}}{1-\varphi_1\varphi_2} \quad (4.11)$$

Κατά συνέπεια, για να δημιουργηθούν δύο στάσιμες  $AR(1)$  χρονοσειρές με επιθυμητή συσχέτιση  $\rho$ , πρέπει πρώτα να παραχθούν τα τυχαία σφάλματα  $u_t$  και  $v_t$  με μοναδιαία διακύμανση και συνδιακύμανση:

$$q_{12} = \rho \frac{1-\varphi_1\varphi_2}{\sqrt{(1-\varphi_1^2)}\sqrt{(1-\varphi_2^2)}} \quad (4.12)$$

η οποία μάλιστα εκφράζει και τη συσχέτιση των σφαλμάτων δεδομένου ότι έχει θεωρηθεί ότι οι διακυμάνσεις τους είναι ίσες με τη μονάδα. Εναλλακτικά, ο συντελεστής συσχέτισης των τυχαίων σφαλμάτων θα υπολογιστεί λαμβάνοντας υπ'



όψιν και τις τιμές των διακυμάνσεων των σφαλμάτων, δηλαδή τις τιμές των  $q_{11}$  και  $q_{22}$ .

Η διαδικασία αυτή παραγωγής συσχετισμένων γραμμικά χρονοσειρών εφαρμόστηκε για τα υποδείγματα  $AR(1) - ARCH(1)$  στα σφάλματα, όπου αρχικά κατασκευάστηκαν τα σφάλματα με βάση τη σχέση (4.2) και κατόπιν με βάση την προκαθορισμένη τιμή του συντελεστή συσχέτισης  $\rho$  σχηματίστηκαν οι χρονοσειρές χρησιμοποιώντας την ίδια αναπαραγωγή χρονοσειρών DGP για τιμές συσχέτισης 0.2, 0.5, 0.8 και 0.9. Στη συνέχεια, υπολογίστηκαν πρώτα οι τιμές όλων των μέτρων εξάρτησης που έχουν χρησιμοποιηθεί στην εργασία αυτή από τις παραχθείσες δύο γραμμικά εξαρτημένες χρονοσειρές και κατόπιν οι τιμές των μέσων όρων τους βασισμένες σε 10000 επαναλήψεις. Τα αποτελέσματα αυτής της προσπάθειας δίνονται στους Πίνακες 4.5 έως 4.8.

Αρχικά, θα πρέπει να σημειωθεί, όπως φαίνεται από τους πίνακες αυτούς, ότι η παρουσία αυτής της μορφής ετεροσκεδαστικότητας δεν επιφέρει σημαντικές αλλαγές στην τιμή του μέσου όρου του δειγματικού συντελεστή συσχέτισης. Η αλήθεια είναι ότι η τιμή του συντελεστή συσχέτισης μειώνεται όσο αυξάνεται η τιμή των παραμέτρων  $\phi$  και  $a_1$  ή κάθε μία από αυτές τις δύο παραμέτρους ξεχωριστά θεωρώντας ότι η τιμή της άλλης παραμέτρου παραμένει σταθερή. Ωστόσο, η γενική εικόνα της γραμμικής συσχέτισης μεταξύ των δύο μεταβλητών προσδιορίζεται αρκετά καλά και οι τιμές των μέσων όρων των δειγματικών συντελεστών συσχέτισης δεν αλλοιώνουν την πραγματικότητα, δηλαδή την πληθυσμιακή τιμή του συντελεστή συσχέτισης που χρησιμοποιείται για να δημιουργηθούν οι δύο χρονοσειρές.

Για παράδειγμα, για  $\rho = 0.9$ ,  $\phi = 0.9$  και  $a_1 = 0.9$  η τιμή του μέσου όρου του δειγματικού συντελεστή συσχέτισης για δύο χρονοσειρές με 500 παρατηρήσεις που προκύπτει από 10000 επαναλήψεις είναι 0.790, μία τιμή πολύ κοντά στην πληθυσμιακή τιμή 0.9. Άλλωστε, όπως έχει αναφερθεί, ο δειγματικός συντελεστής συσχέτισης δεν είναι αμερόληπτος εκτιμητής του πληθυσμιακού συντελεστή συσχέτισης και κατά συνέπεια είναι κατά κάποιο τρόπο αποδεκτό να παρουσιάζει κάποιο βαθμό μεροληψίας.



### Πίνακας 4.5

Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές για  $\rho = 0.2$  βασισμένες σε 10000 επαναλήψεις

$\phi$	$a_1$	Correlation Coefficient $r$	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
0.0	0.0	0.200	0.190	0.129	0.213	0.211	0.249	1.244	1.147
		0.200	0.191	0.128	0.204	0.203	0.217	1.176	1.123
	0.2	0.197	0.189	0.128	0.210	0.209	0.247	1.240	1.146
		0.197	0.190	0.128	0.201	0.201	0.214	1.169	1.122
	0.5	0.183	0.185	0.125	0.202	0.202	0.237	1.219	1.141
		0.180	0.186	0.125	0.192	0.193	0.204	1.142	1.117
	0.8	0.158	0.179	0.121	0.191	0.193	0.223	1.184	1.135
		0.144	0.179	0.121	0.180	0.182	0.191	1.104	1.110
	0.9	0.148	0.177	0.120	0.188	0.189	0.218	1.171	1.132
		0.128	0.177	0.119	0.176	0.179	0.186	1.089	1.107
0.2	0.0	0.199	0.190	0.128	0.212	0.210	0.250	1.249	1.148
		0.200	0.191	0.128	0.204	0.204	0.217	1.176	1.123
	0.2	0.197	0.189	0.128	0.210	0.208	0.246	1.243	1.146
		0.197	0.190	0.128	0.201	0.201	0.214	1.168	1.122
	0.5	0.183	0.184	0.125	0.202	0.201	0.236	1.221	1.142
		0.180	0.185	0.125	0.192	0.193	0.204	1.141	1.117
	0.8	0.158	0.178	0.121	0.191	0.192	0.222	1.187	1.135
		0.144	0.179	0.120	0.180	0.182	0.191	1.102	1.109
	0.9	0.148	0.176	0.119	0.187	0.189	0.217	1.176	1.133
		0.128	0.177	0.119	0.176	0.179	0.186	1.087	1.107
0.5	0.0	0.198	0.189	0.128	0.211	0.209	0.251	1.311	1.157
		0.200	0.191	0.129	0.204	0.203	0.217	1.179	1.124
	0.2	0.196	0.187	0.127	0.209	0.207	0.248	1.300	1.155
		0.197	0.190	0.127	0.201	0.201	0.215	1.168	1.123
	0.5	0.182	0.182	0.124	0.200	0.199	0.237	1.270	1.149
		0.180	0.184	0.124	0.192	0.193	0.204	1.133	1.117
	0.8	0.158	0.174	0.118	0.188	0.189	0.222	1.227	1.142
		0.144	0.176	0.118	0.179	0.181	0.190	1.082	1.108
	0.9	0.148	0.171	0.116	0.184	0.184	0.216	1.211	1.139
		0.129	0.173	0.116	0.175	0.176	0.184	1.063	1.105
0.8	0.0	0.195	0.186	0.127	0.205	0.203	0.263	1.605	1.199
		0.200	0.191	0.129	0.203	0.203	0.220	1.209	1.129
	0.2	0.193	0.185	0.126	0.203	0.201	0.261	1.606	1.199
		0.197	0.189	0.127	0.200	0.200	0.217	1.197	1.127
	0.5	0.181	0.177	0.121	0.194	0.192	0.249	1.583	1.195
		0.180	0.181	0.122	0.189	0.189	0.204	1.151	1.121
	0.8	0.158	0.165	0.113	0.178	0.177	0.230	1.544	1.188
		0.145	0.167	0.113	0.172	0.173	0.184	1.079	1.110
	0.9	0.148	0.160	0.110	0.172	0.171	0.223	1.533	1.186
		0.130	0.162	0.109	0.165	0.166	0.176	1.051	1.106
0.9	0.0	0.192	0.183	0.127	0.199	0.197	0.285	1.990	1.256
		0.199	0.190	0.128	0.201	0.200	0.224	1.306	1.143
	0.2	0.190	0.182	0.126	0.197	0.195	0.283	1.985	1.255
		0.196	0.188	0.127	0.198	0.198	0.221	1.296	1.141
	0.5	0.178	0.173	0.120	0.187	0.185	0.271	1.964	1.251
		0.179	0.177	0.120	0.185	0.185	0.206	1.254	1.135
	0.8	0.155	0.156	0.109	0.168	0.166	0.250	1.933	1.245
		0.145	0.160	0.108	0.164	0.165	0.182	1.185	1.124
	0.9	0.146	0.150	0.105	0.161	0.159	0.242	1.927	1.243
		0.130	0.153	0.103	0.155	0.157	0.172	1.159	1.120

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



**Πίνακας 4.6**

**Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές για  $\rho = 0.5$  βασισμένες σε 10000 επαναλήψεις**

$\phi$	$a_1$	Correlation Coefficient $r$	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
0.0	0.0	0.499	0.479	0.334	0.517	0.515	0.764	3.362	1.476
		0.500	0.482	0.333	0.505	0.505	0.690	3.309	1.439
	0.2	0.493	0.476	0.332	0.513	0.511	0.756	3.344	1.471
		0.494	0.479	0.332	0.500	0.500	0.682	3.291	1.435
	0.5	0.464	0.467	0.326	0.496	0.497	0.728	3.276	1.455
		0.458	0.470	0.325	0.481	0.484	0.654	3.222	1.417
	0.8	0.409	0.453	0.316	0.474	0.479	0.691	3.172	1.431
		0.378	0.455	0.315	0.456	0.463	0.616	3.118	1.394
	0.9	0.386	0.448	0.312	0.466	0.472	0.678	3.132	1.422
		0.342	0.450	0.311	0.447	0.456	0.603	3.078	1.385
0.2	0.0	0.498	0.478	0.333	0.516	0.514	0.764	3.359	1.476
		0.500	0.482	0.334	0.505	0.505	0.690	3.310	1.440
	0.2	0.493	0.476	0.332	0.512	0.511	0.756	3.341	1.471
		0.494	0.479	0.332	0.500	0.500	0.683	3.290	1.435
	0.5	0.464	0.466	0.325	0.496	0.497	0.728	3.269	1.454
		0.458	0.469	0.325	0.481	0.484	0.654	3.217	1.417
	0.8	0.409	0.452	0.315	0.474	0.478	0.690	3.163	1.431
		0.378	0.455	0.315	0.456	0.463	0.616	3.111	1.393
	0.9	0.387	0.446	0.311	0.465	0.471	0.676	3.121	1.421
		0.342	0.449	0.311	0.447	0.455	0.602	3.070	1.384
0.5	0.0	0.497	0.477	0.333	0.515	0.512	0.767	3.370	1.479
		0.500	0.482	0.334	0.505	0.504	0.692	3.314	1.441
	0.2	0.492	0.474	0.331	0.511	0.508	0.760	3.344	1.474
		0.494	0.479	0.331	0.500	0.500	0.684	3.287	1.435
	0.5	0.463	0.462	0.322	0.494	0.493	0.729	3.247	1.454
		0.458	0.466	0.322	0.481	0.483	0.653	3.190	1.415
	0.8	0.410	0.444	0.310	0.469	0.471	0.685	3.106	1.425
		0.379	0.448	0.309	0.454	0.458	0.609	3.050	1.387
	0.9	0.387	0.437	0.305	0.460	0.463	0.668	3.053	1.415
		0.343	0.440	0.304	0.443	0.449	0.593	2.996	1.376
0.8	0.0	0.492	0.472	0.333	0.506	0.503	0.786	3.446	1.497
		0.499	0.481	0.334	0.503	0.503	0.696	3.331	1.444
	0.2	0.487	0.469	0.330	0.502	0.499	0.778	3.421	1.492
		0.493	0.477	0.331	0.498	0.498	0.687	3.293	1.438
	0.5	0.460	0.451	0.318	0.482	0.480	0.739	3.285	1.468
		0.458	0.458	0.317	0.475	0.476	0.646	3.141	1.410
	0.8	0.408	0.422	0.297	0.448	0.448	0.678	3.070	1.429
		0.381	0.428	0.296	0.437	0.440	0.583	2.905	1.368
	0.9	0.386	0.410	0.289	0.435	0.435	0.654	2.985	1.414
		0.346	0.416	0.287	0.422	0.427	0.558	2.812	1.352
0.9	0.0	0.485	0.466	0.333	0.495	0.493	0.816	3.647	1.535
		0.497	0.479	0.334	0.500	0.499	0.703	3.350	1.449
	0.2	0.480	0.462	0.330	0.491	0.488	0.808	3.621	1.531
		0.491	0.474	0.330	0.494	0.493	0.692	3.305	1.441
	0.5	0.454	0.442	0.315	0.469	0.466	0.763	3.486	1.506
		0.456	0.451	0.313	0.467	0.467	0.643	3.117	1.408
	0.8	0.403	0.405	0.290	0.428	0.427	0.690	3.262	1.466
		0.380	0.412	0.286	0.420	0.424	0.564	2.810	1.356
	0.9	0.381	0.390	0.279	0.412	0.410	0.662	3.176	1.451
		0.346	0.396	0.274	0.401	0.406	0.533	2.689	1.336

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



### Πίνακας 4.7

Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές για  $\rho = 0.8$  βασισμένες σε 10000 επαναλήψεις

$\phi$	$a_1$	Correlation Coefficient $r$	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
0.0	0.0	0.799	0.780	0.590	0.805	0.804	1.926	7.626	2.330
		0.800	0.785	0.590	0.802	0.801	1.807	7.566	2.268
	0.2	0.794	0.778	0.588	0.802	0.801	1.906	7.593	2.318
		0.795	0.782	0.588	0.798	0.798	1.788	7.532	2.255
	0.5	0.769	0.768	0.580	0.787	0.789	1.833	7.451	2.269
		0.764	0.772	0.580	0.781	0.784	1.711	7.390	2.204
	0.8	0.714	0.752	0.568	0.765	0.774	1.727	7.215	2.196
		0.681	0.756	0.567	0.756	0.765	1.602	7.158	2.129
	0.9	0.689	0.746	0.562	0.757	0.768	1.688	7.121	2.169
		0.640	0.750	0.562	0.746	0.758	1.561	7.064	2.101
0.2	0.0	0.799	0.780	0.590	0.805	0.804	1.927	7.629	2.331
		0.800	0.785	0.590	0.802	0.802	1.808	7.568	2.268
	0.2	0.794	0.778	0.588	0.802	0.801	1.907	7.591	2.319
		0.795	0.782	0.588	0.798	0.798	1.789	7.529	2.255
	0.5	0.769	0.767	0.580	0.787	0.789	1.833	7.435	2.268
		0.764	0.772	0.580	0.781	0.784	1.711	7.378	2.203
	0.8	0.714	0.751	0.567	0.765	0.773	1.726	7.193	2.194
		0.681	0.756	0.567	0.756	0.765	1.601	7.136	2.127
	0.9	0.690	0.745	0.561	0.757	0.768	1.686	7.096	2.166
		0.641	0.749	0.561	0.746	0.758	1.559	7.039	2.098
0.5	0.0	0.798	0.779	0.590	0.804	0.803	1.936	7.663	2.337
		0.800	0.785	0.590	0.801	0.801	1.811	7.576	2.270
	0.2	0.793	0.776	0.588	0.800	0.800	1.917	7.605	2.322
		0.795	0.782	0.588	0.798	0.798	1.790	7.518	2.255
	0.5	0.768	0.764	0.577	0.786	0.787	1.834	7.384	2.264
		0.764	0.770	0.577	0.781	0.783	1.706	7.303	2.196
	0.8	0.714	0.744	0.560	0.761	0.768	1.713	7.050	2.177
		0.682	0.749	0.560	0.753	0.760	1.581	6.975	2.108
	0.9	0.690	0.735	0.553	0.751	0.760	1.668	6.921	2.144
		0.642	0.740	0.553	0.742	0.751	1.533	6.845	2.074
0.8	0.0	0.794	0.775	0.590	0.797	0.796	1.981	7.816	2.366
		0.799	0.784	0.591	0.800	0.800	1.822	7.617	2.277
	0.2	0.789	0.771	0.587	0.794	0.793	1.958	7.742	2.349
		0.794	0.780	0.587	0.796	0.796	1.795	7.532	2.257
	0.5	0.765	0.754	0.572	0.776	0.776	1.858	7.411	2.275
		0.764	0.762	0.571	0.775	0.777	1.683	7.181	2.174
	0.8	0.712	0.722	0.545	0.742	0.746	1.695	6.868	2.154
		0.684	0.729	0.543	0.737	0.743	1.509	6.631	2.048
	0.9	0.688	0.708	0.533	0.728	0.734	1.632	6.654	2.108
		0.644	0.715	0.531	0.720	0.729	1.443	6.416	2.001
0.9	0.0	0.788	0.768	0.590	0.789	0.788	2.055	8.066	2.412
		0.798	0.782	0.590	0.797	0.796	1.837	7.666	2.284
	0.2	0.784	0.764	0.586	0.785	0.784	2.033	7.983	2.395
		0.793	0.778	0.586	0.792	0.792	1.806	7.566	2.262
	0.5	0.760	0.744	0.569	0.765	0.764	1.916	7.579	2.307
		0.762	0.756	0.566	0.768	0.769	1.671	7.128	2.162
	0.8	0.706	0.703	0.535	0.723	0.725	1.725	6.899	2.164
		0.683	0.713	0.530	0.720	0.727	1.458	6.413	2.006
	0.9	0.681	0.685	0.521	0.705	0.708	1.654	6.644	2.112
		0.644	0.694	0.515	0.699	0.709	1.378	6.134	1.948

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



### Πίνακας 4.8

Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές για  $\rho = 0.9$  βασισμένες σε 10000 επαναλήψεις

$\phi$	$a_1$	Correlation Coefficient $r$	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
0.0	0.0	0.899	0.886	0.713	0.900	0.899	3.149	11.716	3.252
		0.900	0.890	0.713	0.900	0.900	3.016	11.705	3.184
	0.2	0.897	0.885	0.711	0.898	0.898	3.117	11.668	3.233
		0.897	0.889	0.711	0.898	0.898	2.985	11.656	3.162
	0.5	0.880	0.878	0.704	0.888	0.890	2.995	11.451	3.148
		0.877	0.882	0.704	0.887	0.889	2.857	11.439	3.075
	0.8	0.843	0.865	0.693	0.873	0.881	2.816	11.074	3.021
		0.819	0.869	0.692	0.869	0.876	2.665	11.063	2.941
	0.9	0.824	0.860	0.688	0.866	0.877	2.746	10.913	2.971
		0.788	0.864	0.687	0.862	0.871	2.592	10.906	2.889
0.2	0.0	0.899	0.886	0.713	0.900	0.899	3.153	11.725	3.254
		0.900	0.890	0.713	0.900	0.900	3.017	11.707	3.185
	0.2	0.896	0.884	0.711	0.898	0.897	3.122	11.666	3.234
		0.897	0.889	0.711	0.898	0.898	2.986	11.651	3.162
	0.5	0.880	0.877	0.704	0.888	0.890	2.994	11.425	3.146
		0.877	0.882	0.704	0.887	0.889	2.857	11.418	3.073
	0.8	0.843	0.865	0.692	0.872	0.880	2.812	11.034	3.017
		0.819	0.869	0.692	0.869	0.876	2.663	11.027	2.937
	0.9	0.824	0.860	0.687	0.866	0.877	2.744	10.874	2.968
		0.788	0.864	0.687	0.862	0.870	2.589	10.864	2.885
0.5	0.0	0.899	0.886	0.713	0.899	0.899	3.168	11.793	3.263
		0.900	0.890	0.713	0.900	0.900	3.021	11.721	3.188
	0.2	0.896	0.884	0.711	0.897	0.897	3.135	11.705	3.240
		0.897	0.889	0.711	0.898	0.898	2.987	11.631	3.162
	0.5	0.880	0.875	0.701	0.887	0.889	2.997	11.354	3.142
		0.877	0.880	0.702	0.887	0.888	2.847	11.292	3.061
	0.8	0.842	0.860	0.686	0.870	0.877	2.790	10.809	2.990
		0.820	0.864	0.686	0.867	0.873	2.629	10.757	2.905
	0.9	0.825	0.853	0.679	0.862	0.872	2.711	10.587	2.931
		0.789	0.857	0.679	0.858	0.866	2.545	10.538	2.844
0.8	0.0	0.896	0.882	0.712	0.895	0.895	3.234	12.047	3.295
		0.899	0.890	0.713	0.899	0.899	3.036	11.785	3.196
	0.2	0.894	0.880	0.710	0.893	0.892	3.198	11.942	3.270
		0.896	0.887	0.710	0.896	0.896	2.993	11.652	3.163
	0.5	0.878	0.868	0.697	0.881	0.882	3.030	11.406	3.146
		0.877	0.876	0.696	0.883	0.884	2.805	11.098	3.023
	0.8	0.840	0.844	0.672	0.856	0.863	2.759	10.532	2.941
		0.821	0.851	0.671	0.856	0.861	2.509	10.217	2.805
	0.9	0.822	0.832	0.662	0.845	0.855	2.658	10.184	2.862
		0.790	0.839	0.660	0.843	0.852	2.396	9.869	2.722
0.9	0.0	0.893	0.878	0.712	0.889	0.889	3.351	12.429	3.340
		0.899	0.888	0.713	0.897	0.897	3.057	11.860	3.203
	0.2	0.890	0.875	0.710	0.887	0.887	3.313	12.326	3.317
		0.896	0.886	0.709	0.894	0.894	3.006	11.694	3.164
	0.5	0.874	0.861	0.695	0.873	0.873	3.123	11.651	3.171
		0.876	0.871	0.693	0.879	0.879	2.781	11.010	2.998
	0.8	0.836	0.829	0.664	0.843	0.847	2.815	10.613	2.940
		0.820	0.839	0.660	0.844	0.850	2.420	9.874	2.731
	0.9	0.817	0.815	0.651	0.828	0.834	2.697	10.210	2.855
		0.790	0.824	0.646	0.827	0.838	2.286	9.434	2.631

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



Ανάλογη συμπεριφορά παρατηρείται και για τα άλλα δύο μέτρα εξάρτησης του Spearman's  $\rho$  και του Kendall's  $\tau$ , με το πρώτο να προσδιορίζει καλύτερα τη γραμμική συσχέτιση από το συντελεστή συσχέτισης, ενώ το δεύτερο όχι καλύτερα. Για παράδειγμα, για  $\rho = 0.9$ ,  $\varphi = 0.9$  και  $a_1 = 0.9$  η τιμή του μέσου όρου του δειγματικού συντελεστή Spearman's  $\rho$  για δύο χρονοσειρές με 500 παρατηρήσεις που προκύπτει από 10000 επαναλήψεις είναι 0.824, ενώ αντίστοιχα η τιμή του μέσου όρου του δειγματικού συντελεστή του Kendall's  $\tau$  είναι ίση με 0.646.

Η συμπεριφορά των Ελλειπτικών *Copulas*, δηλαδή της Gaussian και της Student-*t*, είναι παρόμοια μεταξύ τους και παρόμοια με τη συμπεριφορά του δειγματικού συντελεστή του Spearman's  $\rho$ , ενδεχομένως και λίγο καλύτερη δεδομένου ότι οι τιμές τους είναι λίγο μεγαλύτερες από τις αντίστοιχες του Spearman's  $\rho$  και πιο κοντά στις πληθυσμιακές τιμές του συντελεστή συσχέτισης με τις οποίες δημιουργούνται οι εξαρτημένες γραμμικά χρονοσειρές. Για παράδειγμα, για  $\rho = 0.9$ ,  $\varphi = 0.9$  και  $a_1 = 0.9$  οι τιμές των μέσων όρων των δειγματικών συντελεστών των Gaussian και Student-*t copulas* για δύο χρονοσειρές με 500 παρατηρήσεις που προκύπτουν από 10000 επαναλήψεις είναι αντίστοιχα ίσες με 0.827 και 0.838.

Τα πράγματα όμως ως προς την ερμηνεία των τιμών των μέσων όρων των Αρχιμήδειων *Copulas*, δηλαδή των Clayton, Gumbel και Frank, είναι αρκετά περίπλοκη δεδομένου ότι για τους αριθμούς αυτούς η ερμηνεία τους δεν είναι πλήρως κατανοητή. Εκείνο όμως που πρέπει να σημειωθεί είναι ότι οι τιμές των μέσων όρων τους αυξάνονται καθώς αυξάνονται οι τιμές όλων των παραμέτρων της προσομοίωσης, δηλώνοντας πιο έντονη εξάρτηση μεταξύ των μεταβλητών. Ειδικότερα, οι τιμές των μέσων όρων τους αυξάνεται καθώς αυξάνεται η τιμή του πληθυσμιακού συντελεστή συσχέτισης που χρησιμοποιείται για τη δημιουργία των χρονοσειρών για δεδομένες τιμές των παραμέτρων  $\varphi$  και  $a_1$ . Για παράδειγμα, για  $\varphi = 0.9$  και  $a_1 = 0.9$  οι τιμές των μέσων όρων των δειγματικών συντελεστών των Clayton, Gumbel και Frank είναι αντίστοιχα ίσες με 0.172, 1.159 και 1.120 για δύο εξαρτημένες γραμμικά χρονοσειρές με 500 παρατηρήσεις που προκύπτουν από 10000 επαναλήψεις για  $\rho = 0.2$ , ενώ για  $\rho = 0.9$  οι τιμές αυτές είναι ίσες με 2.286, 9.434 και 2.631.





Η διαδικασία παραγωγής γραμμικά συσχετισμένων χρονοσειρών επεκτάθηκε και στα υποδείγματα  $AR(1) - ARCH(1)$  στη μεταβλητή, όπου αρχικά κατασκευάστηκαν τα σφάλματα με βάση τη σχέση (4.3) και κατόπιν με βάση την προκαθορισμένη τιμή του συντελεστή συσχέτισης  $\rho$  σχηματίστηκαν οι χρονοσειρές χρησιμοποιώντας την ίδια αναπαραγωγική διαδικασία (DGP) των χρονοσειρών για τιμές συσχέτισης 0.2, 0.5, 0.8 και 0.9. Σημειώνεται μάλιστα ότι ενδεχομένως για την περίπτωση αυτή η παραγωγική διαδικασία που εφαρμόζεται για το σχηματισμό των εξαρτημένων μεταβλητών να μεγαλώνει λίγο περισσότερο τη μεταβλητότητα των χρονοσειρών δεδομένου ότι το υπόδειγμα χρησιμοποιήθηκε σε δύο στάδια τόσο για την παραγωγή των σφαλμάτων όσο και για την παραγωγή της γραμμικής εξάρτησης, κάτι το οποίο μπορεί να συμβάλλει ακόμα καλύτερα στην αποτύπωση γραμμικών σχέσεων σε μη γραμμικά υποδείγματα. Στη συνέχεια, όπως και στην προηγούμενη περίπτωση, υπολογίστηκαν αρχικά από τις κατασκευασμένες γραμμικά εξαρτημένες χρονοσειρές οι τιμές όλων των μέτρων εξάρτησης που χρησιμοποιούνται στην εργασία αυτή και κατόπιν οι τιμές των μέσων όρων τους, βασισμένες σε 10000 επαναλήψεις, οι τιμές των οποίων δίνονται στους Πίνακες 4.9 έως 4.12.

Όπως φαίνεται από τους πίνακες αυτούς, ο συντελεστής συσχέτισης του δείγματος δύο χρονοσειρών χάνει τη συνέπεια του στο να προσδιορίσει σωστά το βαθμό της γραμμικής εξάρτησης με τον οποίο κατασκευάστηκαν οι χρονοσειρές, καθώς αυξάνονται οι τιμές των παραμέτρων  $\phi$  και  $a_1$  για κάθε τιμή του πληθυσμιακού συντελεστή συσχέτισης. Για παράδειγμα, για  $\rho = 0.9$ ,  $\phi = 0.9$  και  $a_1 = 0.9$  η τιμή του μέσου όρου του δειγματικού συντελεστή συσχέτισης για δύο χρονοσειρές με 500 παρατηρήσεις που προκύπτει από 10000 επαναλήψεις είναι 0.476, αρκετά πολύ μικρότερη από την πληθυσμιακή τιμή 0.9. Επίσης, η τιμή του δειγματικού συντελεστή συσχέτισης μειώνεται για κάθε τιμή του  $\rho$  και  $\phi$ , καθώς αυξάνεται η τιμή της παραμέτρου  $a_1$ , δηλαδή η τιμή της παραμέτρου της μη γραμμικότητας. Για παράδειγμα, για  $\rho = 0.8$  και  $\phi = 0.8$ , οι τιμές των μέσων όρων του συντελεστή συσχέτισης του δείγματος για σειρές 500 παρατηρήσεων βασισμένες σε 10000 επαναλήψεις είναι 0.799, 0.596 και 0.375 για τιμές του  $a_1 = 0.0, 0.5$  και  $0.9$  αντίστοιχα. Επιπλέον, η τιμή του συντελεστή συσχέτισης δείγματος μειώνεται, καθώς η τιμή της



παραμέτρου του αυτοπαλίνδρομου όρου αυξάνεται για δεδομένη τιμή της παραμέτρου της μη γραμμικότητας. Για παράδειγμα, για  $\rho = 0.8$  και για  $a_1 = 0.5$ , οι τιμές των μέσων όρων του συντελεστή συσχέτισης του δείγματος για σειρές 500 παρατηρήσεων βασισμένες σε 10000 επαναλήψεις είναι 0.760, 0.728, 0.596 και 0.505 για τιμές του  $\varphi = 0.2, 0.5, 0.8$  και  $0.9$  αντίστοιχα.

Η συμπεριφορά του δειγματικού συντελεστή συσχέτισης για τις γραμμικά εξαρτημένες  $AR(1) - ARCH(1)$  στη μεταβλητή χρονοσειρές είναι σε γενικές γραμμές παρόμοια με αυτή των γραμμικά εξαρτημένων  $AR(1) - ARCH(1)$  στα σφάλματα χρονοσειρών. Απλά στην περίπτωση αυτή και λόγω της ιδιομορφίας των χρονοσειρών η απόκλιση των εκτιμήσεων του δειγματικού συντελεστή από την πραγματική πληθυσμιακή τιμή με την οποία δημιουργούνται οι γραμμικά εξαρτημένες μεταβλητές είναι μεγαλύτερη για κάθε τιμή όλων των παραμέτρων. Για παράδειγμα, για  $\rho = 0.9, \varphi = 0.9$  και  $a_1 = 0.9$  η τιμή του μέσου όρου του δειγματικού συντελεστή συσχέτισης για δύο γραμμικά εξαρτώμενες  $AR(1) - ARCH(1)$  στη μεταβλητή χρονοσειρές με 500 παρατηρήσεις που προκύπτει από 10000 επαναλήψεις είναι 0.476, σε αντίθεση από την τιμή 0.790 που προκύπτει από τις  $AR(1) - ARCH(1)$  στα σφάλματα χρονοσειρές. Αυτό σημαίνει ότι η μεροληψία του συντελεστή συσχέτισης γίνεται ακόμα μεγαλύτερη υπό την επήρεια της μεγάλης μεταβλητότητας αυτών των μη γραμμικών χρονοσειρών.

Ωστόσο, το πιο σημαντικό στοιχείο αυτών των προσομοιώσεων για γραμμικά εξαρτημένες  $AR(1) - ARCH(1)$  στη μεταβλητή χρονοσειρές είναι ότι η τιμή του δειγματικού συντελεστή συσχέτισης μειώνεται σημαντικά καθώς αυξάνεται το μέγεθος των παρατηρήσεων των χρονοσειρών, για δεδομένες τιμές όλων των παραμέτρων της προσομοίωσης, φτάνοντας σε εξαιρετικά μεγάλα επίπεδα μεροληψίας. Για παράδειγμα, για  $\rho = 0.9, \varphi = 0.9$  και  $a_1 = 0.9$  η τιμή του μέσου όρου του δειγματικού συντελεστή συσχέτισης για δύο γραμμικά εξαρτημένες  $AR(1) - ARCH(1)$  στη μεταβλητή χρονοσειρές με 10000 παρατηρήσεις που προκύπτει από 10000 επαναλήψεις είναι 0.247, μία τιμή πάρα πολύ μικρή σε σχέση με την πραγματική τιμή της δημιουργίας της γραμμικής εξάρτησης των μεταβλητών. Στην περίπτωση αυτή ο βαθμός μεροληψίας του συντελεστή συσχέτισης ξεπερνά κάθε ανεκτό όριο απόκλισης



από τη θεωρητική πραγματική τιμή που χρησιμοποιήθηκε για τη δημιουργία των χρονοσειρών.

**Πίνακας 4.9**

**Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για  $\rho = 0.2$  βασισμένες σε 10000 επαναλήψεις**

$\Phi$	$a_1$	Correlation Coefficient $r$	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
0.0	0.0	0.200	0.190	0.129	0.213	0.211	0.249	1.244	1.147
		0.200	0.191	0.128	0.204	0.203	0.217	1.176	1.123
	0.2	0.197	0.189	0.128	0.210	0.209	0.247	1.240	1.146
		0.197	0.190	0.128	0.201	0.201	0.214	1.169	1.122
	0.5	0.183	0.185	0.125	0.202	0.202	0.237	1.219	1.141
		0.180	0.186	0.125	0.192	0.193	0.204	1.142	1.117
	0.8	0.158	0.179	0.121	0.191	0.193	0.223	1.184	1.135
		0.144	0.179	0.121	0.180	0.182	0.191	1.104	1.110
	0.9	0.148	0.177	0.120	0.188	0.189	0.218	1.171	1.132
		0.128	0.177	0.119	0.176	0.179	0.186	1.089	1.107
0.2	0.0	0.199	0.190	0.128	0.212	0.210	0.250	1.249	1.148
		0.200	0.191	0.128	0.204	0.204	0.217	1.176	1.123
	0.2	0.196	0.189	0.128	0.210	0.208	0.246	1.243	1.146
		0.197	0.190	0.127	0.201	0.201	0.214	1.168	1.122
	0.5	0.181	0.184	0.125	0.201	0.201	0.235	1.220	1.142
		0.178	0.185	0.124	0.191	0.192	0.203	1.139	1.116
	0.8	0.155	0.178	0.120	0.190	0.191	0.221	1.184	1.135
		0.139	0.179	0.120	0.180	0.182	0.190	1.099	1.109
	0.9	0.144	0.175	0.119	0.186	0.188	0.216	1.172	1.132
		0.123	0.176	0.118	0.175	0.178	0.185	1.084	1.106
0.5	0.0	0.198	0.189	0.128	0.211	0.209	0.251	1.311	1.157
		0.200	0.191	0.129	0.204	0.203	0.217	1.179	1.124
	0.2	0.193	0.186	0.126	0.207	0.205	0.246	1.288	1.153
		0.194	0.189	0.127	0.200	0.200	0.213	1.164	1.122
	0.5	0.170	0.180	0.122	0.196	0.196	0.230	1.240	1.145
		0.161	0.182	0.122	0.187	0.188	0.199	1.118	1.114
	0.8	0.138	0.172	0.117	0.183	0.184	0.214	1.192	1.136
		0.110	0.173	0.116	0.173	0.175	0.183	1.065	1.104
	0.9	0.127	0.169	0.115	0.179	0.181	0.208	1.176	1.133
		0.094	0.170	0.114	0.169	0.171	0.178	1.046	1.101
0.8	0.0	0.195	0.186	0.127	0.205	0.203	0.263	1.605	1.199
		0.200	0.191	0.129	0.203	0.203	0.220	1.209	1.129
	0.2	0.180	0.182	0.124	0.199	0.197	0.251	1.512	1.185
		0.175	0.186	0.125	0.195	0.195	0.210	1.162	1.122
	0.5	0.139	0.168	0.114	0.181	0.181	0.221	1.349	1.160
		0.106	0.170	0.115	0.174	0.175	0.184	1.061	1.107
	0.8	0.102	0.156	0.106	0.165	0.167	0.197	1.238	1.142
		0.058	0.157	0.106	0.155	0.158	0.162	0.979	1.093
	0.9	0.092	0.152	0.104	0.160	0.162	0.190	1.211	1.138
		0.047	0.153	0.103	0.150	0.153	0.155	0.953	1.089
0.9	0.0	0.192	0.183	0.127	0.199	0.197	0.285	1.990	1.256
		0.199	0.190	0.128	0.201	0.200	0.224	1.306	1.143
	0.2	0.166	0.176	0.121	0.191	0.190	0.258	1.719	1.215
		0.149	0.181	0.122	0.190	0.190	0.207	1.189	1.126
	0.5	0.120	0.159	0.109	0.171	0.171	0.216	1.436	1.172
		0.077	0.161	0.109	0.163	0.165	0.172	1.029	1.102
	0.8	0.086	0.146	0.100	0.154	0.157	0.187	1.276	1.148
		0.040	0.148	0.099	0.144	0.147	0.149	0.934	1.088
	0.9	0.078	0.142	0.098	0.149	0.153	0.180	1.236	1.142
		0.033	0.143	0.097	0.138	0.142	0.142	0.906	1.084

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



Πίνακας 4.10

Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για  $\rho = 0.5$  βασισμένες σε 10000 επαναλήψεις

$\phi$	$a_1$	Correlation Coefficient $r$	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
0.0	0.0	0.499	0.479	0.334	0.517	0.515	0.764	3.362	1.476
		0.500	0.482	0.333	0.505	0.505	0.690	3.309	1.439
	0.2	0.493	0.476	0.332	0.513	0.511	0.756	3.344	1.471
		0.494	0.479	0.332	0.500	0.500	0.682	3.291	1.435
	0.5	0.464	0.467	0.326	0.496	0.497	0.728	3.276	1.455
		0.458	0.470	0.325	0.481	0.484	0.654	3.222	1.417
	0.8	0.409	0.453	0.316	0.474	0.479	0.691	3.172	1.431
		0.378	0.455	0.315	0.456	0.463	0.616	3.118	1.394
	0.9	0.386	0.448	0.312	0.466	0.472	0.678	3.132	1.422
		0.342	0.450	0.311	0.447	0.456	0.603	3.078	1.385
0.2	0.0	0.498	0.478	0.333	0.516	0.514	0.764	3.359	1.476
		0.500	0.482	0.334	0.505	0.505	0.690	3.310	1.440
	0.2	0.492	0.475	0.332	0.512	0.510	0.755	3.339	1.471
		0.493	0.479	0.331	0.499	0.500	0.682	3.288	1.434
	0.5	0.461	0.465	0.324	0.495	0.495	0.726	3.263	1.453
		0.454	0.468	0.324	0.480	0.482	0.653	3.212	1.416
	0.8	0.403	0.451	0.314	0.472	0.476	0.687	3.155	1.429
		0.367	0.453	0.314	0.454	0.461	0.613	3.102	1.391
	0.9	0.379	0.445	0.311	0.464	0.470	0.673	3.113	1.420
		0.330	0.448	0.310	0.445	0.454	0.599	3.060	1.382
0.5	0.0	0.497	0.477	0.333	0.515	0.512	0.767	3.370	1.479
		0.500	0.482	0.334	0.505	0.504	0.692	3.314	1.441
	0.2	0.487	0.472	0.330	0.508	0.506	0.755	3.327	1.470
		0.488	0.477	0.330	0.497	0.497	0.680	3.274	1.433
	0.5	0.438	0.456	0.319	0.485	0.486	0.713	3.202	1.444
		0.419	0.461	0.319	0.471	0.474	0.639	3.152	1.406
	0.8	0.365	0.437	0.305	0.457	0.463	0.664	3.055	1.413
		0.303	0.441	0.305	0.440	0.448	0.590	3.004	1.375
	0.9	0.339	0.431	0.301	0.448	0.455	0.648	3.004	1.403
		0.263	0.434	0.300	0.430	0.439	0.574	2.952	1.365
0.8	0.0	0.492	0.472	0.333	0.506	0.503	0.786	3.446	1.497
		0.499	0.481	0.334	0.503	0.503	0.696	3.331	1.444
	0.2	0.460	0.462	0.324	0.494	0.492	0.756	3.324	1.474
		0.448	0.469	0.325	0.488	0.488	0.668	3.225	1.425
	0.5	0.367	0.430	0.301	0.456	0.457	0.672	3.036	1.419
		0.296	0.436	0.301	0.442	0.446	0.592	2.958	1.375
	0.8	0.281	0.402	0.281	0.419	0.426	0.604	2.808	1.375
		0.174	0.405	0.280	0.399	0.410	0.523	2.731	1.331
	0.9	0.256	0.393	0.276	0.408	0.418	0.584	2.740	1.362
		0.145	0.396	0.273	0.387	0.399	0.503	2.661	1.318
0.9	0.0	0.485	0.466	0.333	0.495	0.493	0.816	3.647	1.535
		0.497	0.479	0.334	0.500	0.499	0.703	3.350	1.449
	0.2	0.429	0.450	0.318	0.479	0.476	0.755	3.356	1.482
		0.393	0.460	0.319	0.476	0.477	0.655	3.165	1.416
	0.5	0.323	0.410	0.288	0.433	0.436	0.644	2.930	1.403
		0.225	0.414	0.286	0.417	0.423	0.553	2.799	1.349
	0.8	0.242	0.379	0.267	0.395	0.405	0.567	2.665	1.353
		0.127	0.381	0.263	0.372	0.384	0.481	2.551	1.303
	0.9	0.221	0.370	0.260	0.383	0.396	0.546	2.596	1.340
		0.107	0.372	0.256	0.360	0.373	0.461	2.481	1.290

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



**Πίνακας 4.11**

**Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για  $\rho = 0.8$  βασισμένες σε 10000 επαναλήψεις**

$\phi$	$a_1$	Correlation Coefficient $r$	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
0.0	0.0	0.799	0.780	0.590	0.805	0.804	1.926	7.626	2.330
		0.800	0.785	0.590	0.802	0.801	1.807	7.566	2.268
	0.2	0.794	0.778	0.588	0.802	0.801	1.906	7.593	2.318
		0.795	0.782	0.588	0.798	0.798	1.788	7.532	2.255
	0.5	0.769	0.768	0.580	0.787	0.789	1.833	7.451	2.269
		0.764	0.772	0.580	0.781	0.784	1.711	7.390	2.204
0.2	0.8	0.714	0.752	0.568	0.765	0.774	1.727	7.215	2.196
		0.681	0.756	0.567	0.756	0.765	1.602	7.158	2.129
	0.9	0.689	0.746	0.562	0.757	0.768	1.688	7.121	2.169
		0.640	0.750	0.562	0.746	0.758	1.561	7.064	2.101
	0.0	0.799	0.780	0.590	0.805	0.804	1.927	7.629	2.331
		0.800	0.785	0.590	0.802	0.802	1.808	7.568	2.268
0.5	0.2	0.794	0.777	0.588	0.801	0.800	1.906	7.587	2.318
		0.795	0.782	0.588	0.797	0.797	1.787	7.526	2.254
	0.5	0.766	0.767	0.579	0.786	0.788	1.827	7.424	2.264
		0.760	0.771	0.579	0.780	0.783	1.707	7.366	2.199
	0.8	0.708	0.750	0.566	0.763	0.772	1.717	7.175	2.188
		0.670	0.754	0.565	0.753	0.763	1.592	7.116	2.121
0.8	0.9	0.682	0.743	0.560	0.754	0.766	1.677	7.078	2.160
		0.628	0.747	0.560	0.743	0.756	1.549	7.016	2.092
	0.0	0.798	0.779	0.590	0.804	0.803	1.936	7.663	2.337
		0.800	0.785	0.590	0.801	0.801	1.811	7.576	2.270
	0.2	0.789	0.775	0.586	0.798	0.798	1.904	7.572	2.314
		0.790	0.780	0.587	0.795	0.796	1.781	7.493	2.248
0.9	0.5	0.745	0.758	0.572	0.778	0.780	1.794	7.289	2.236
		0.728	0.764	0.572	0.772	0.775	1.670	7.222	2.170
	0.8	0.668	0.736	0.554	0.749	0.760	1.657	6.939	2.140
		0.599	0.741	0.554	0.738	0.750	1.527	6.876	2.071
	0.9	0.637	0.728	0.547	0.739	0.753	1.610	6.813	2.107
		0.547	0.732	0.547	0.726	0.741	1.478	6.749	2.036
0.0	0.0	0.794	0.775	0.590	0.797	0.796	1.981	7.816	2.366
		0.799	0.784	0.591	0.800	0.800	1.822	7.617	2.277
	0.2	0.766	0.765	0.580	0.787	0.786	1.903	7.545	2.305
		0.757	0.773	0.580	0.787	0.787	1.748	7.381	2.222
	0.5	0.674	0.730	0.550	0.749	0.756	1.685	6.888	2.151
		0.596	0.737	0.550	0.741	0.749	1.537	6.766	2.071
0.2	0.8	0.565	0.693	0.520	0.706	0.725	1.492	6.313	2.014
		0.424	0.698	0.518	0.690	0.710	1.336	6.202	1.931
	0.9	0.530	0.682	0.511	0.692	0.716	1.437	6.143	1.975
		0.375	0.685	0.508	0.673	0.698	1.277	6.028	1.890
	0.0	0.788	0.768	0.590	0.789	0.788	2.055	8.066	2.412
		0.798	0.782	0.590	0.797	0.796	1.837	7.666	2.284
0.5	0.2	0.738	0.752	0.573	0.774	0.773	1.895	7.505	2.293
		0.705	0.764	0.573	0.776	0.777	1.710	7.244	2.191
	0.5	0.621	0.706	0.532	0.725	0.735	1.604	6.574	2.086
		0.505	0.712	0.529	0.713	0.726	1.427	6.385	1.990
	0.8	0.509	0.664	0.498	0.677	0.701	1.394	5.923	1.938
		0.344	0.666	0.493	0.655	0.683	1.217	5.762	1.843
0.8	0.9	0.476	0.652	0.488	0.662	0.691	1.338	5.748	1.899
		0.302	0.652	0.482	0.637	0.670	1.158	5.579	1.802

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



Πίνακας 4.12

Τιμές μέσων όρων για όλα τα μέτρα εξάρτησης που προήλθαν από δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για  $\rho = 0.9$  βασισμένες σε 10000 επαναλήψεις

$\phi$	$a_1$	Correlation Coefficient $r$	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
0.0	0.0	0.899	0.886	0.713	0.900	0.899	3.149	11.716	3.252
		0.900	0.890	0.713	0.900	0.900	3.016	11.705	3.184
	0.2	0.897	0.885	0.711	0.898	0.898	3.117	11.668	3.233
		0.897	0.889	0.711	0.898	0.898	2.985	11.656	3.162
	0.5	0.880	0.878	0.704	0.888	0.890	2.995	11.451	3.148
		0.877	0.882	0.704	0.887	0.889	2.857	11.439	3.075
	0.8	0.843	0.865	0.693	0.873	0.881	2.816	11.074	3.021
		0.819	0.869	0.692	0.869	0.876	2.665	11.063	2.941
	0.9	0.824	0.860	0.688	0.866	0.877	2.746	10.913	2.971
		0.788	0.864	0.687	0.862	0.871	2.592	10.906	2.889
0.2	0.0	0.899	0.886	0.713	0.900	0.899	3.153	11.725	3.254
		0.900	0.890	0.713	0.900	0.900	3.017	11.707	3.185
	0.2	0.896	0.884	0.711	0.897	0.897	3.119	11.661	3.232
		0.897	0.889	0.711	0.898	0.898	2.984	11.647	3.161
	0.5	0.879	0.877	0.703	0.887	0.890	2.987	11.408	3.139
		0.875	0.881	0.703	0.886	0.888	2.849	11.400	3.067
	0.8	0.838	0.864	0.691	0.871	0.880	2.798	11.001	3.006
		0.811	0.868	0.691	0.867	0.874	2.647	10.993	2.926
	0.9	0.819	0.858	0.686	0.864	0.876	2.726	10.836	2.955
		0.779	0.862	0.685	0.859	0.869	2.570	10.825	2.872
0.5	0.0	0.899	0.886	0.713	0.899	0.899	3.168	11.793	3.263
		0.900	0.890	0.713	0.900	0.900	3.021	11.721	3.188
	0.2	0.894	0.883	0.710	0.896	0.896	3.116	11.647	3.225
		0.894	0.888	0.710	0.896	0.896	2.973	11.593	3.151
	0.5	0.865	0.871	0.697	0.882	0.885	2.932	11.205	3.094
		0.853	0.876	0.698	0.881	0.883	2.786	11.168	3.016
	0.8	0.809	0.853	0.680	0.861	0.872	2.695	10.624	2.922
		0.757	0.858	0.680	0.856	0.865	2.533	10.593	2.836
	0.9	0.784	0.846	0.674	0.852	0.867	2.609	10.414	2.863
		0.714	0.850	0.673	0.846	0.858	2.442	10.376	2.772
0.8	0.0	0.896	0.882	0.712	0.895	0.895	3.234	12.047	3.295
		0.899	0.890	0.713	0.899	0.899	3.036	11.785	3.196
	0.2	0.879	0.876	0.704	0.888	0.888	3.107	11.622	3.198
		0.873	0.883	0.705	0.891	0.891	2.915	11.418	3.105
	0.5	0.815	0.850	0.677	0.861	0.869	2.749	10.573	2.944
		0.758	0.855	0.677	0.858	0.865	2.556	10.429	2.843
	0.8	0.726	0.818	0.648	0.826	0.848	2.411	9.595	2.698
		0.605	0.822	0.646	0.816	0.838	2.196	9.481	2.585
	0.9	0.694	0.808	0.638	0.814	0.841	2.312	9.309	2.628
		0.556	0.810	0.636	0.801	0.829	2.086	9.181	2.507
0.9	0.0	0.893	0.878	0.712	0.889	0.889	3.351	12.429	3.340
		0.899	0.888	0.713	0.897	0.897	3.057	11.860	3.203
	0.2	0.861	0.867	0.698	0.879	0.880	3.092	11.537	3.161
		0.839	0.876	0.698	0.883	0.884	2.849	11.198	3.049
	0.5	0.773	0.830	0.661	0.842	0.855	2.603	10.043	2.822
		0.682	0.835	0.658	0.836	0.850	2.364	9.812	2.699
	0.8	0.673	0.793	0.626	0.801	0.829	2.239	8.969	2.561
		0.522	0.795	0.621	0.786	0.819	1.985	8.769	2.428
	0.9	0.641	0.781	0.615	0.788	0.821	2.142	8.674	2.491
		0.476	0.782	0.610	0.769	0.810	1.879	8.458	2.351

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



Το εύρημα αυτό της μεγάλης ασυμπτωτικής μεροληψίας του συντελεστή συσχέτισης επισημαίνει ότι ο συντελεστής συσχέτισης του δείγματος μπορεί να δώσει ενδείξεις χαμηλής ή ακόμη και καθόλου γραμμικής συσχέτισης για γραμμικά εξαρτημένες από μη γραμμικά παραγόμενες χρονοσειρές, καταλογίζοντας ψεύτικες συμπεριφορές όχι όμως με τη μορφή των υποφαινομενικών συσχετίσεων. Στην περίπτωση αυτή υπάρχει γραμμική εξάρτηση αλλά ο συντελεστής συσχέτισης δηλώνει ότι δεν υπάρχει. Ως εκ τούτου, σοβαρά σφάλματα μέτρησης της εξάρτησης δύο μεταβλητών μπορεί να προκύψουν όταν ο συντελεστής συσχέτισης χρησιμοποιείται για τον προσδιορισμό του επιπέδου εξάρτησης για χρονοσειρές υψηλής συχνότητας ή υψηλής μεταβλητότητας, οδηγώντας τον αναλυτή σε εσφαλμένα συμπεράσματα.

Σε αντίθεση με το συντελεστή συσχέτισης, οι τιμές των μέσων όρων για τα μέτρα εξάρτησης του Spearman's  $\rho$  και του Kendall's  $\tau$  παρέμειναν σταθερές ή σχετικά σταθερές γύρω από την προκαθορισμένη τιμή του  $\rho$  που χρησιμοποιήθηκε για τη δημιουργία της εξάρτησης, με μικρές αλλαγές που βασίζονται στην παράμετρο της μη γραμμικότητας, χωρίς να επηρεάζονται από την παρουσία της αυτοπαλίνδρομης παραμέτρου. ή το μέγεθος του δείγματος. Αναμφισβήτητα όμως ο συντελεστής του Spearman's  $\rho$  προσδιορίζει πολύ καλύτερα τη γραμμική σχέση των μεταβλητών από ότι ο συντελεστής του Kendall's  $\tau$ , δεδομένου ότι όπως φαίνεται από τους Πίνακες 4.9 έως 4.12 οι τιμές των μέσων όρων του Spearman's  $\rho$  είναι κοντά στο παραγόμενο επίπεδο εξάρτησης, ενώ οι τιμές των μέσων όρων του Kendall's  $\tau$  είναι συνήθως μικρότερες. Για παράδειγμα, για  $\rho = 0.9$ ,  $\varphi = 0.9$  και  $a_1 = 0.9$  η τιμή του μέσου όρου του συντελεστή του Spearman's  $\rho$  που προκύπτει από 10000 επαναλήψεις γραμμικά εξαρτώμενων χρονοσειρών 500 παρατηρήσεων είναι ίση με 0.782, ενώ αντίστοιχα η τιμή του μέσου όρου του συντελεστή του Kendall's  $\tau$  είναι ίση με 0.610.

Παρόμοια συμπεριφορά, αν όχι πανομοιότυπη, με το μέτρο εξάρτησης του Spearman's  $\rho$  βρίσκεται και στις Ελλειπτικές *Copulas*, Gaussian και Student-t, όπου οι τιμές των μέσων όρων τους είναι πολύ κοντά μεταξύ τους και κοντά στην τιμή του πληθυσμιακού συντελεστή συσχέτισης που χρησιμοποιείται για τη δημιουργία της γραμμικής εξάρτησης των μεταβλητών. Κατά συνέπεια, τα μέτρα αυτά αποτυπώνουν καλύτερα



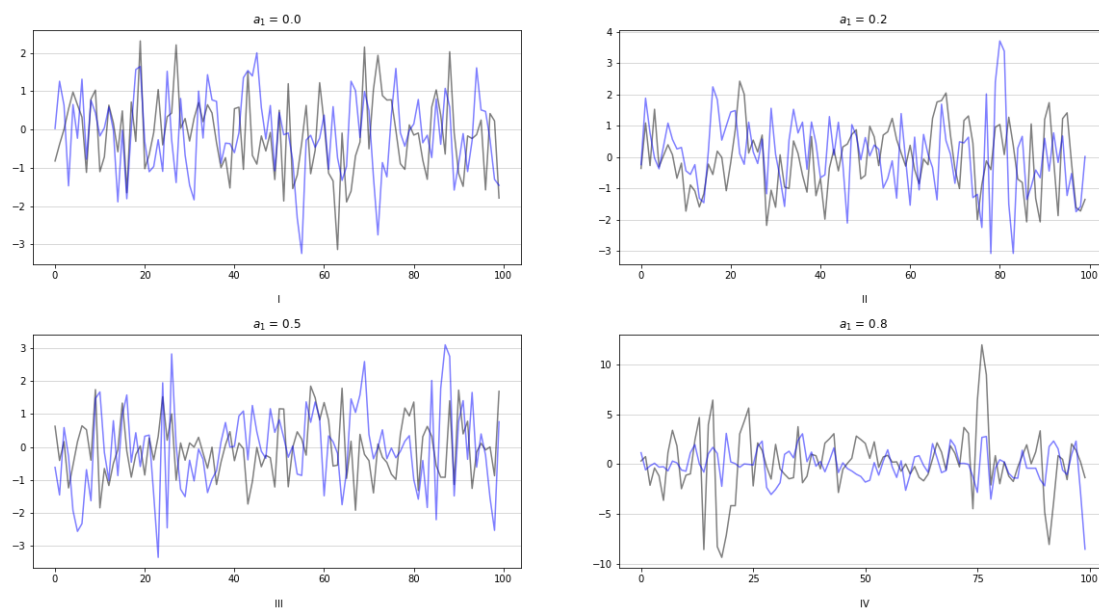
από τον συντελεστή συσχέτισης την έννοια της εξάρτησης μεταξύ των δύο μεταβλητών χωρίς οι τιμές τους να επηρεάζονται σημαντικά από την τιμή της παραμέτρου της μη γραμμικότητας, το μέγεθος του δείγματος ή/και την τιμή της αυτοπαλίνδρομης παραμέτρου.

Κατά ανάλογο τρόπο οι Αρχιμήδειες *Copulas*, δηλαδή οι Clayton, Gumbel και Frank, αντιλαμβάνονται επίσης την γραμμική εξάρτηση με επιτυχία, αν και είναι δύσκολο να ερμηνευτούν οι τιμές τους, δεδομένου ότι οι τιμές των μέσων όρων τους αυξάνονται καθώς αυξάνεται το παραγόμενο επίπεδο εξάρτησης για κάθε τιμή των παραμέτρων  $\varphi$  και  $a_1$  αλλά και του δείγματος των παρατηρήσεων  $n$ , συλλαμβάνοντας ωστόσο με διαφορετικό τρόπο τα παραγόμενα επίπεδα της γραμμικής εξάρτησης. Για παράδειγμα, οι τιμές των μέσων όρων των Clayton, Frank και Gumbel που προήλθαν από 10000 επαναλήψεις είναι 0.235, 1.220 και 1.142 αντίστοιχα για  $\rho = 0.2$  και 1.827, 7.424 και 2.424  $\rho = 0.8$  για  $\varphi = 0.2$ ,  $a_1 = 0.5$  και χρονοσειρές 100 παρατηρήσεων.

Επίσης, η παράμετρος της μη γραμμικότητας φαίνεται να μην παίζει σημαντικό ρόλο στην αλλαγή της γενικής εικόνας του προσδιορισμού της εξάρτησης που προκύπτει από τις οι Αρχιμήδειες *Copulas*, αν και οι τιμές των μέσων όρων τους μειώνονται ελαφρώς όσο αυξάνεται η τιμή του  $a_1$ , διατηρώντας όλες τις άλλες παραμέτρους αμετάβλητες. Ωστόσο, σε αντίθεση με τις Ελλειπτικές *Copulas* η παρουσία της αυτοπαλίνδρομης παραμέτρου αλλάζει, αλλά όχι σημαντικά, τις τιμές των μέσων όρων των Αρχιμήδειων *Copulas*, παράγοντας διαφορετικές συμπεριφορές ανάλογα με τον τύπο της κάθε μίας και το μέγεθος του συντελεστή συσχέτισης.

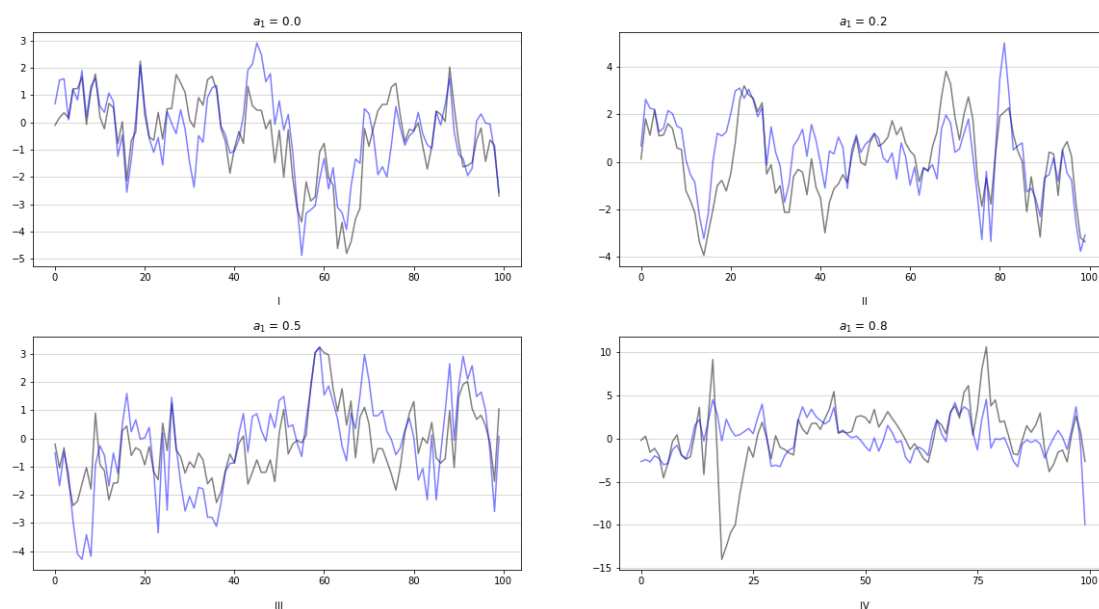
Επιπρόσθετα, οι τιμές των μέσων όρων τους επηρεάζονται από το μέγεθος των παρατηρήσεων των χρονοσειρών, σε αντίθεση με τις Ελλειπτικές *Copulas*. Πραγματικά, οι τιμές των μέσων όρων των Αρχιμήδειων *Copulas* μειώνονται, αλλά με διαφορετικό ρυθμό, καθώς αυξάνεται το μέγεθος του δείγματος, διατηρώντας όλες τις άλλες παραμέτρους σταθερές, με τη συμμετρική *Copula* Frank να έχει μικρότερες μεταβολές από τις μη συμμετρικές *Copulas* Clayton και Gumbel.





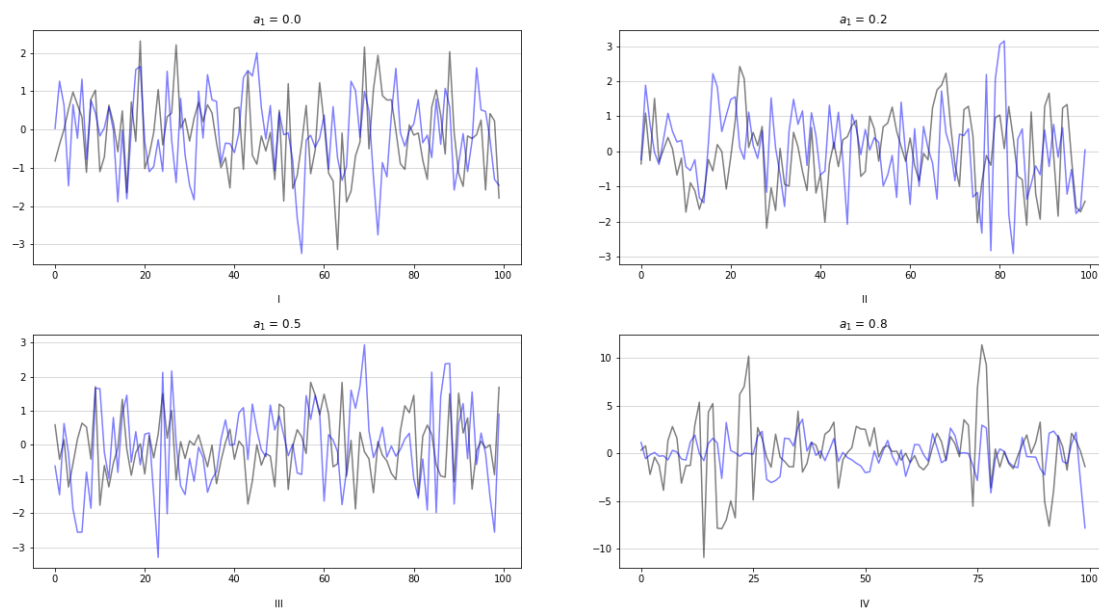
**Διάγραμμα 4.11**

**Διαγραμματική απεικόνιση δύο γραμμικά εξαρτημένων AR(1) – ARCH(1) στα σφάλματα χρονοσειρών για  $\rho = 0.2$ ,  $\phi = 0.2$  και για  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$  και για μεγέθη δείγματος 100 παρατηρήσεων**



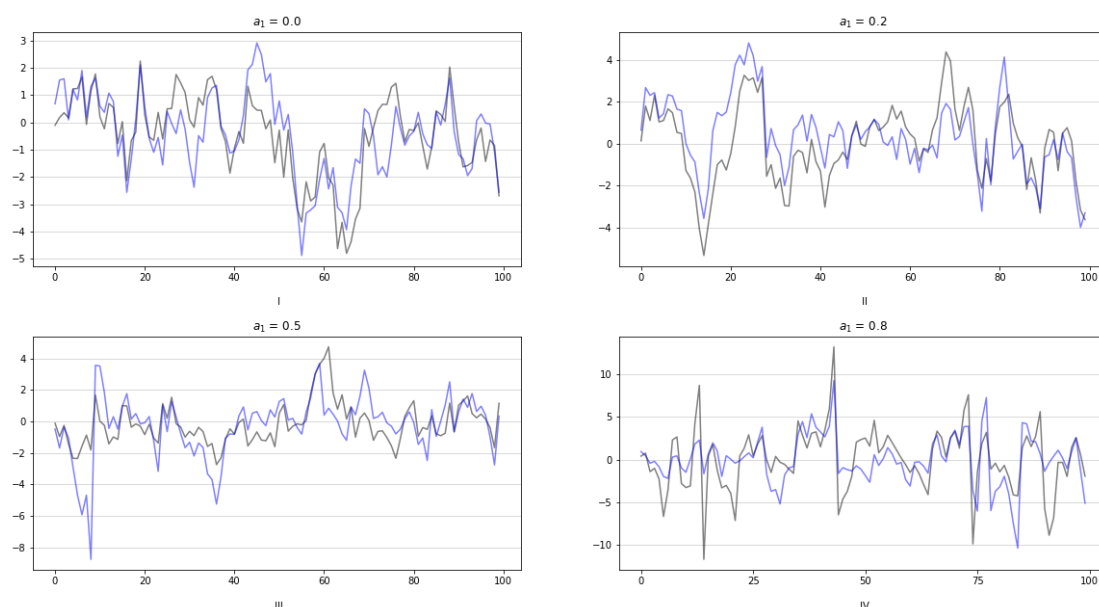
**Διάγραμμα 4.12**

**Διαγραμματική απεικόνιση δύο γραμμικά εξαρτημένων AR(1) – ARCH(1) στα σφάλματα χρονοσειρών για  $\rho = 0.8$ ,  $\phi = 0.8$  και για  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$  και για μεγέθη δείγματος 100 παρατηρήσεων**



**Διάγραμμα 4.13**

**Διαγραμματική απεικόνιση δύο γραμμικά εξαρτημένων AR(1) – ARCH(1) στη μεταβλητή χρονοσειρών για  $\rho = 0.2$ ,  $\phi = 0.2$  και για  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$  και για μεγέθη δείγματος 100 παρατηρήσεων**

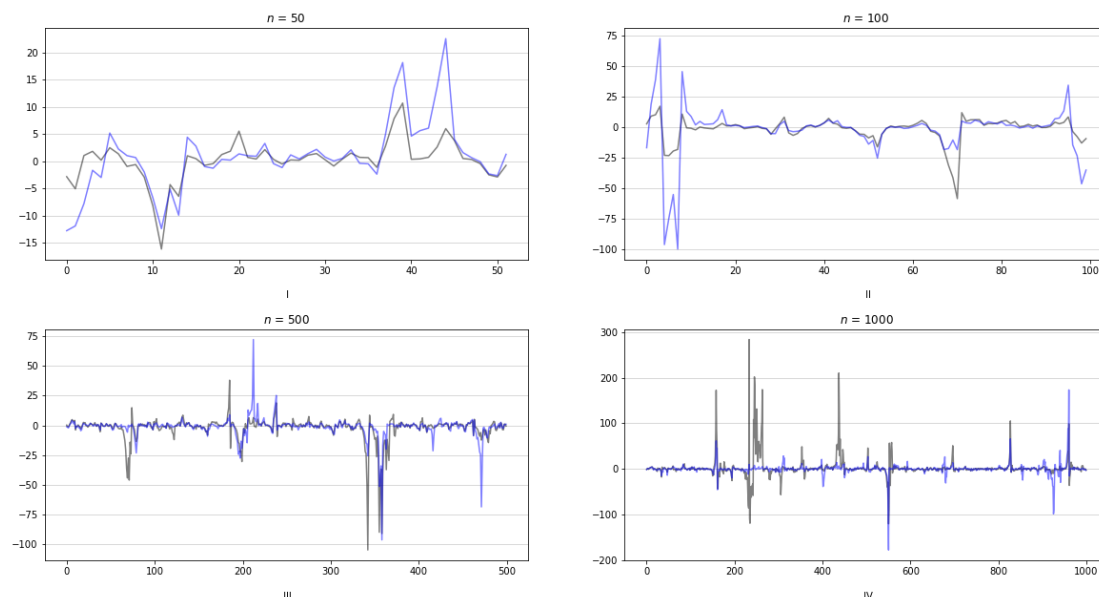


**Διάγραμμα 4.14**

**Διαγραμματική απεικόνιση δύο γραμμικά εξαρτημένων AR(1) – ARCH(1) στη μεταβλητή χρονοσειρών για  $\rho = 0.8$ ,  $\phi = 0.8$  και για  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$  και για μεγέθη δείγματος 100 παρατηρήσεων**



Τέλος, στα Διάγραμμα 4.11 και 4.12 απεικονίζονται ενδεικτικά δύο γραμμικά εξαρτημένες  $AR(1) - ARCH(1)$  στα σφάλματα χρονοσειρές με 100 παρατηρήσεις και για τιμές των παραμέτρων  $\rho = 0.2$  &  $\varphi = 0.2$  και  $\rho = 0.8$  &  $\varphi = 0.8$  για  $a_1 = 0.0, 0.2, 0.5$  και  $0.8$  αντίστοιχα, ενώ ανάλογα στα Διαγράμματα 4.13 και 4.14 απεικονίζουν δύο γραμμικά εξαρτημένες  $AR(1) - ARCH(1)$  στη μεταβλητή χρονοσειρές με τα ίδια χαρακτηριστικά. Από τα διαγράμματα αυτά παρατηρείται ότι για μικρές τιμές του συντελεστή συσχέτισης οι χρονοσειρές βαδίζουν πολύ έντονα μεταξύ τους, ενώ για μεγάλες τιμές του υπάρχει μια σχετική απόκλιση από την κοινή συμπεριφορά τους. Το γεγονός μάλιστα ότι τα διαγράμματα αυτά χρησιμοποιούν χρονοσειρές 100 παρατηρήσεων δυσκολεύει στο να απεικονιστεί καλά η διαφορά στη συσχέτισή τους μεταξύ αυτών των διαφορετικών τύπων χρονοσειρών ακόμα και για τη τιμή  $0.8$  του συντελεστή συσχέτισης. Άλλωστε, η τιμή του μέσου όρου του συντελεστή συσχέτισης για δύο γραμμικά εξαρτημένες  $AR(1) - ARCH(1)$  στα σφάλματα χρονοσειρές με 100 παρατηρήσεις που παρήχθησαν με τιμές παραμέτρων  $\rho = 0.8, \varphi = 0.8$  και  $a_1 = 0.8$  για 10000 επαναλήψεις είναι  $0.712$ , όπως εμφανίζεται στον Πίνακα 4.7, ενώ για δύο γραμμικά εξαρτημένες  $AR(1) - ARCH(1)$  στη μεταβλητή χρονοσειρές με τα ίδια χαρακτηριστικά είναι  $0.565$ , όπως εμφανίζεται στον Πίνακα 4.11.



**Διάγραμμα 4.15**

**Διαγραμματική απεικόνιση δύο γραμμικά εξαρτημένων AR(1) – ARCH(1) στη μεταβλητή χρονοσειρών για  $\rho = 0.9$ ,  $\varphi = 0.9$  και για  $a_1 = 0.9$  και για μεγέθη δείγματος 50, 100, 500 και 1000 παρατηρήσεων**

Για το λόγο αυτό, στο Διάγραμμα 4.15 απεικονίζονται δύο γραμμικά εξαρτημένες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για δείγματα 50, 100, 500 και 1000 παρατηρήσεων με τιμές παραμέτρων  $\rho = 0.9$ ,  $\varphi = 0.9$  και  $a_1 = 0.9$  από όπου διαπιστώνεται ότι για αυτές τις μη γραμμικές χρονοσειρές η γραμμικότητα συμπτωτικά εξαφανίζεται. Σημειώνεται ότι σε αυτές τις περιπτώσεις η γραμμική εξάρτηση λαμβάνει την πιο μεγάλη τιμή που χρησιμοποιήθηκε για τον πληθυσμιακό συντελεστή συσχέτισης στην ανάλυση προσομοίωσης που εφαρμόστηκε.

## **4.7 Δύναμη ελέγχου**

Οι γραμμικά εξαρτημένες χρονοσειρές που δημιουργήθηκαν στην προηγούμενη παράγραφο χρησιμοποιήθηκαν στη συνέχεια για τη διερεύνηση του δύναμης του στατιστικού ελέγχου της μηδενικής συσχέτισης με τον ίδιο ακριβώς τρόπο, όπως εφαρμόστηκε και για τη διερεύνηση της δύναμής του για ανεξάρτητες χρονοσειρές, και



τα ποσοστά απόρριψης της μηδενικής υπόθεσης για ονομαστικό επίπεδο 5% και για τιμές του  $\rho = 0.2, 0.5, 0.8$  και  $0.9$  μαζί με όλους τους συνδυασμούς των παραμέτρων  $\varphi$  και  $a_1$  βασισμένα σε 10000 επαναλήψεις δίνονται στους Πίνακες 4.13 και 4.14 για υποδείγματα  $AR(1) - ARCH(1)$  στα σφάλματα και για υποδείγματα  $AR(1) - ARCH(1)$  στη μεταβλητή αντίστοιχα. Αρχικά, θα πρέπει να σημειωθεί ότι τα αποτελέσματα της δύναμης του ελέγχου δεν είναι καθόλου ικανοποιητικά και για  $a_1 = 0$ , δηλαδή για στάσιμες αυτοπαλίνδρομες πρώτου βαθμού χρονοσειρές ή καλύτερα για στάσιμες  $AR(1)$  χρονοσειρές με ομοσκεδαστικότητα, είναι όμοια με αυτά που παρουσιάζονται στις εργασίες των Agiakloglou και Agiropoulos (2016) και στην εργασία των Zimmerman *et. al.* (2003) για κανονικούς πληθυσμούς.

Πραγματικά, για να είχε πολύ καλή δύναμη ο στατιστικός έλεγχος θα έπρεπε όλοι οι αριθμοί των Πινάκων 4.13 και 4.14 να ήταν 100.00, δηλαδή η μηδενική υπόθεση της μη ύπαρξης γραμμικής σχέσης να απορρίπτονταν με βεβαιότητα, δηλαδή 100%, για κάθε τιμή όλων των παραμέτρων. Η αλήθεια όμως είναι ότι και οι δύο πίνακες δείχνουν κάτι άλλο εκτός από αυτό, κάτι που σε γενικές γραμμές θα έπρεπε να ήταν αναμενόμενους δεδομένου ότι ανακαλύφθηκε μεροληψία στην τιμή του δειγματικού συντελεστή συσχέτισης και για τα δύο είδη χρονοσειρών, ειδικότερα όμως για τα υποδείγματα  $AR(1) - ARCH(1)$  στη μεταβλητή, πέρα από το κλασσικό πρόβλημα της μεροληψίας που είχε προσδιορίσει ο Fisher (1915), το μέγεθος της οποίας εξαρτιόταν από διάφορες συνθήκες ακόμα και από το μέγεθος του δείγματος.

Στην πραγματικότητα, η δύναμη του ελέγχου αυξάνεται καθώς το μέγεθος της γραμμικής εξάρτησης αυξάνεται για δεδομένες τιμές των παραμέτρων  $\varphi$ ,  $a_1$  και του μεγέθους του δείγματος των παρατηρήσεων  $n$  των χρονοσειρών και για τα δύο είδη υποδειγμάτων. Ωστόσο, η παρουσία της μη γραμμικής δομής στη δεύτερη μορφή υποδειγμάτων, δηλαδή στα υποδείγματα  $AR(1) - ARCH(1)$  στη μεταβλητή, έχει πολύ μεγαλύτερη αρνητική επιρροή στη συμπεριφορά της ισχύς του στατιστικού ελέγχου από ότι η πρώτη μορφή υποδειγμάτων, δηλαδή τα υποδείγματα  $AR(1) - ARCH(1)$  στα σφάλματα, δεδομένου ότι, όπως αναφέρθηκε, η μεροληψία της εκτίμησης του συντελεστή συσχέτισης ήταν πολύ μεγαλύτερη σε αυτά τα υποδείγματα. Για παράδειγμα και για τις τιμές των παραμέτρων,  $\varphi = 0.5$ ,  $a_1 = 0.8$  και  $n = 500$ , η μηδενική



υπόθεση απορρίπτεται, σε ονομαστικό επίπεδο σημαντικότητας 5%, βασισμένη σε 10000 επαναλήψεις 83.06% για  $\rho = 0.2$  και 100.00% για  $\rho = 0.9$  για τα υποδείγματα AR(1) – ARCH(1) στα σφάλματα, όπως φαίνεται στον Πίνακα 4.13, ενώ για τα υποδείγματα AR(1) – ARCH(1) στη μεταβλητή τα ανάλογα ποσοστά απόρριψης της μηδενικής υπόθεσης είναι 63.19% και 100.00%, όπως φαίνεται από τον Πίνακα 4.14.

Επίσης, ο έλεγχος χάνει την ισχύ του καθώς αυξάνεται η τιμή της παραμέτρου  $a_1$  διατηρώντας όλες τις άλλες τιμές των παραμέτρων σταθερές καθώς και το μέγεθος δείγματος, όπου και σε αυτή την περίπτωση μεγαλύτερες απώλειες παρουσιάζονται στη δεύτερη μορφή υποδειγμάτων. Για παράδειγμα και για τιμές των παραμέτρων  $\rho = 0.5$ ,  $\phi = 0.8$  και  $n = 500$ , η μηδενική υπόθεση απορρίπτεται, σε ονομαστικό επίπεδο σημαντικότητας 5%, βασισμένη σε 10000 επαναλήψεις 100% για  $a_1 = 0.2$  και 98.9% για  $a_1 = 0.9$  για τα υποδείγματα AR(1) – ARCH(1) στα σφάλματα, όπως φαίνεται στον Πίνακα 4.13, ενώ για τα υποδείγματα AR(1) – ARCH(1) στη μεταβλητή τα ανάλογα ποσοστά απόρριψης της μηδενικής υπόθεσης είναι 99.97% και 62.37%, όπως φαίνεται από τον Πίνακα 4.14.



**Πίνακας 4.13**

**Ποσοστά απόρριψης της μηδενικής υπόθεσης για μηδενική συσχέτιση σε ονομαστικό επίπεδο 5% ( $|t| > 1.96$ ) για δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές για  $\rho = 0.2, 0.5, 0.8$  and  $0.9$  βασισμένες σε 10000 επαναλήψεις**

$\phi$	$\alpha_1$	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$	$\rho = 0.9$
0.0	0.0	52.32	99.99	100.00	100.00
		99.51	100.00	100.00	100.00
	0.2	51.60	99.99	100.00	100.00
		99.42	100.00	100.00	100.00
	0.5	45.48	99.91	100.00	100.00
		98.18	100.00	100.00	100.00
0.2	0.8	34.85	98.15	100.00	100.00
		87.94	99.95	100.00	100.00
	0.9	30.83	96.12	99.97	100.00
		79.11	99.78	100.00	100.00
	0.0	52.88	99.99	100.00	100.00
		99.35	100.00	100.00	100.00
0.5	0.2	51.70	100.00	100.00	100.00
		99.20	100.00	100.00	100.00
	0.5	45.95	99.83	100.00	100.00
		97.78	100.00	100.00	100.00
	0.8	35.07	97.80	100.00	100.00
		86.91	99.94	100.00	100.00
0.8	0.9	31.28	95.76	99.96	100.00
		78.57	99.78	100.00	100.00
	0.0	52.29	99.67	100.00	100.00
		97.74	100.00	100.00	100.00
	0.2	51.49	99.63	100.00	100.00
		97.44	100.00	100.00	100.00
0.9	0.5	47.01	99.18	100.00	100.00
		94.91	100.00	100.00	100.00
	0.8	38.21	95.97	100.00	100.00
		83.06	99.92	100.00	100.00
	0.9	34.68	93.39	99.96	100.00
		75.02	99.69	100.00	100.00
0.9	0.0	54.81	95.86	100.00	100.00
		88.96	100.00	100.00	100.00
	0.2	53.74	95.64	100.00	100.00
		88.19	100.00	100.00	100.00
	0.5	51.17	93.49	100.00	100.00
		84.17	100.00	100.00	100.00
0.9	0.8	46.75	87.62	99.82	99.98
		73.86	99.66	99.99	100.00
	0.9	44.78	84.52	99.60	99.97
		68.28	98.90	99.96	100.00
	0.0	60.99	89.70	99.93	100.00
		81.56	99.99	100.00	100.00
0.9	0.2	60.98	89.38	99.96	100.00
		80.89	99.99	100.00	100.00
	0.5	59.42	87.00	99.84	100.00
		78.04	99.86	100.00	100.00
	0.8	56.76	81.58	98.88	99.89
		71.00	98.48	99.96	99.99
0.9	0.9	55.41	79.32	98.14	99.76
		67.55	96.67	99.88	99.99

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.



**Πίνακας 4.14**

**Ποσοστά απόρριψης της μηδενικής υπόθεσης για μηδενική συσχέτιση σε ονομαστικό επίπεδο 5% ( $|t| > 1.96$ ) για δύο γραμμικά εξαρτημένες στάσιμες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για  $\rho = 0.2, 0.5, 0.8$  and  $0.9$  βασισμένες σε 10000 επαναλήψεις**

$\phi$	$a_1$	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$	$\rho = 0.9$
0.00	0.00	52.32	99.99	100.00	100.00
		99.51	100.00	100.00	100.00
	0.20	51.60	99.99	100.00	100.00
		99.42	100.00	100.00	100.00
	0.50	45.48	99.91	100.00	100.00
		98.18	100.00	100.00	100.00
	0.80	34.85	98.15	100.00	100.00
		87.94	99.95	100.00	100.00
	0.90	30.83	96.12	99.97	100.00
		79.11	99.78	100.00	100.00
0.20	0.00	52.88	99.99	100.00	100.00
		99.35	100.00	100.00	100.00
	0.20	51.57	99.96	100.00	100.00
		99.19	100.00	100.00	100.00
	0.50	45.20	99.78	100.00	100.00
		97.41	100.00	100.00	100.00
	0.80	34.10	97.25	99.98	100.00
		85.12	99.91	100.00	100.00
	0.90	30.08	94.76	99.96	100.00
		75.10	99.66	100.00	100.00
0.50	0.00	52.29	99.67	100.00	100.00
		97.74	100.00	100.00	100.00
	0.20	50.59	99.63	100.00	100.00
		97.11	100.00	100.00	100.00
	0.50	42.33	98.45	100.00	100.00
		90.39	100.00	100.00	100.00
	0.80	29.88	90.95	99.88	99.98
		63.19	98.66	100.00	100.00
	0.90	26.20	85.70	99.69	99.96
		50.14	96.31	99.92	100.00
0.80	0.00	54.81	95.86	100.00	100.00
		88.96	100.00	100.00	100.00
	0.20	49.83	94.88	100.00	100.00
		84.06	99.97	100.00	100.00
	0.50	35.50	85.63	99.71	99.99
		55.80	96.63	99.94	100.00
	0.80	23.10	66.42	97.32	99.53
		24.43	73.84	97.56	99.68
	0.90	20.20	59.38	95.27	99.18
		18.63	62.37	94.94	99.06
0.90	0.00	60.99	89.70	99.93	100.00
		81.56	99.99	100.00	100.00
	0.20	51.61	87.87	99.91	99.99
		71.23	99.40	100.00	100.00
	0.50	32.81	74.52	98.48	99.81
		37.12	84.69	99.06	99.88
	0.80	21.24	54.57	92.18	98.43
		17.07	52.01	90.39	98.12
	0.90	18.68	48.72	88.69	97.23
		13.23	42.19	84.67	96.21

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα μεγέθους 100 και 500 παρατηρήσεων.





Επιπρόσθετα, η δύναμη του ελέγχου μειώνεται όσο αυξάνεται η τιμή της παραμέτρου του αυτοπαλίνδρομου όρου, διατηρώντας όλες τις άλλες τιμές των παραμέτρων σταθερές. Για παράδειγμα και για τιμές των παραμέτρων  $\rho = 0.5$ ,  $a_1 = 0.9$  και  $n = 500$ , η μηδενική υπόθεση απορρίπτεται, σε ονομαστικό επίπεδο σημαντικότητας 5%, βασισμένη σε 10000 επαναλήψεις 99.78% για  $\phi = 0.2$  και 96.67% για  $\phi = 0.9$  για τα υποδείγματα AR(1) – ARCH(1) στα σφάλματα, όπως φαίνεται στον Πίνακα 4.13, ενώ για τα υποδείγματα AR(1) – ARCH(1) στη μεταβλητή τα ανάλογα ποσοστά απόρριψης της μηδενικής υπόθεσης είναι 99.91% και 42.19%, όπως φαίνεται από τον Πίνακα 4.14.

Τέλος, η δύναμη του ελέγχου αυξάνεται καθώς αυξάνεται το μέγεθος του δείγματος των παρατηρήσεων των χρονοσειρών για δεδομένες τιμές των άλλων παραμέτρων. Για παράδειγμα και για τιμές των παραμέτρων  $\rho = 0.5$ ,  $\phi = 0.9$  και  $a_1 = 0.5$ , η μηδενική υπόθεση απορρίπτεται, σε ονομαστικό επίπεδο σημαντικότητας 5%, βασισμένη σε 10000 επαναλήψεις 87.00% για  $n = 100$  και 99.86% για  $n = 500$  για τα υποδείγματα AR(1) – ARCH(1) στα σφάλματα, όπως φαίνεται στον Πίνακα 4.13, ενώ για τα υποδείγματα AR(1) – ARCH(1) στη μεταβλητή τα ανάλογα ποσοστά απόρριψης της μηδενικής υπόθεσης είναι 74.52% και 84.69%, όπως φαίνεται από τον Πίνακα 4.14.

Εν κατακλείδι και παρατηρώντας τους Πίνακες 4.13 και 4.14 παρατηρείται ότι η απώλεια της δύναμης του στατιστικού ελέγχου της μηδενικής συσχέτισης έχει μία τριγωνική συμπεριφορά, δηλαδή μειώνεται καθώς κινούμαστε από την επάνω αριστερή γωνία διαγώνια στην κάτω δεξιά γωνία.

## 4.8 Ανακεφαλαίωση

Η ανάλυση συσχέτισης έχει πολλές εφαρμογές σε πολλές επιστήμες ως μια προσπάθεια να διερευνηθεί η παρουσία μιας γραμμικής σχέσης μεταξύ δύο μεταβλητών, αν και υπάρχουν διάφοροι άλλοι τρόποι με τους οποίους μπορεί να συσχετιστεί ένα ζεύγος



τυχαίων μεταβλητών, αλλά συχνά έχει νόημα να θεωρείται η γραμμικότητα ως προσέγγιση σε οποιοδήποτε άλλο είδος σχέσης. Για να αποτυπωθεί ο βαθμός, η ύπαρξη και η κατεύθυνση μιας γραμμικής σχέσης μεταξύ δύο μεταβλητών, υπολογίζεται ο συντελεστής συσχέτισης που ορίζεται ως η συνδιακύμανση των δύο μεταβλητών διαιρούμενη με το γινόμενο των τυπικών αποκλίσεων τους, υποθέτοντας, μεταξύ άλλων, ότι αυτά τα στοιχεία παραμένουν σταθερά με την πάροδο του χρόνου για δεδομένα χρονοσειρών. Επιπλέον, ο συντελεστής συσχέτισης που προκύπτει για δύο χρονοσειρές σχετίζεται επίσης με την λανθασμένη συμπεριφορά, δίνοντας στον αναλυτή παραπλανητικές πληροφορίες σχετικά με την ύπαρξη της σχέσης των δύο μεταβλητών, ανεξάρτητα από το εάν οι χρονοσειρές είναι στάσιμες ή όχι. Ωστόσο, σε περιπτώσεις όπου παραβιάζονται ορισμένες από τις βασικές παραδοχές που αφορούν τον υπολογισμό του συντελεστή συσχέτισης, εφαρμόζονται άλλα μέτρα εξάρτησης, όπως *Copulas*, για να απεικονίσουν καλύτερα την εξάρτηση πιο περίπλοκων μορφών σχέσεων, συμπεριλαμβανομένων των περιπτώσεων όπου οι χρονοσειρές δημιουργούνται από μη γραμμικά υποδείγματα δεδομένα υψηλής συχνότητας ή υψηλής μεταβλητότητας.

Έτσι, στο κεφάλαιο αυτό μελετήθηκε η συμπεριφορά διαφόρων μέτρων εξάρτησης για δύο χρονοσειρές που δημιουργούνται από συνθήκες ετεροσκεδαστικότητας και εμφανίζονται είτε ως υποδείγματα  $AR(1) - ARCH(1)$  στα σφάλματα είτε ως υποδείγματα  $AR(1) - ARCH(1)$  στη μεταβλητή που παρουσιάστηκαν από τους Bera et al. (1992 και 1996) ως μια πιθανή περίπτωση μη γραμμικών χρονοσειρών, η οποία καταγράφει τη γραμμικότητα ως ειδική περίπτωση. Πράγματι, τα υποδείγματα αυτά μπορούν να φανούν ότι προέρχονται από στάσιμες χρονοσειρές αυτοπαλίνδρομου πρώτου βαθμού  $AR(1)$  όπου η διακύμανση του σφάλματος προκύπτει από ένα σχήμα  $ARCH(1)$  για να περιλαμβάνει δεδομένα υψηλής συχνότητας και υψηλής μεταβλητότητας, ενώ το ετεροσκεδαστικό μέρος καθορίζεται από το τιμές της ίδιας της μεταβλητής. Επιπλέον, έχοντας κατά νου τον ισχυρισμό του Granger (2008) ότι οποιοδήποτε μη γραμμικό υπόδειγμα μπορεί να προσεγγιστεί με ένα γραμμικό υπόδειγμα μεταβαλλόμενης διαχρονικά παραμέτρου, αυτές οι χρονοσειρές περιέχουν μια μορφή μη γραμμικής συμπεριφοράς που μπορεί να θεωρηθεί ότι η παράμετρος του  $AR(1)$  υποδείγματος μεταβάλλεται διαχρονικά. Αναμφισβήτητα, είναι πολύ δύσκολο



να διατυπωθούν και να αντιμετωπιστούν μη γραμμικά υποδείγματα, αλλά η χρήση αυτών των  $AR(1) - ARCH(1)$  στη μεταβλητή υποδειγμάτων διευκολύνει την παρουσία μιας μη γραμμικής δομής με πολύ αποδεκτό τρόπο που μπορεί να θεωρηθεί ως σημείο αναφοράς για άλλες μη γραμμικές δομές.

Χρησιμοποιώντας Monte Carlo ανάλυση διαπιστώθηκε ότι για την περίπτωση δύο ανεξάρτητων χρονοσειρών όλα τα μέτρα εξάρτησης λαμβάνουν τιμές κοντά στις αναμενόμενες τιμές τους που υποδεικνύουν ανεξαρτησία, και για τα δύο είδη υποδειγμάτων, με τις Αρχιμήδειες *Copulas* να δείχνουν μικρά σημάδια ευαισθησίας στη μη γραμμική δομή παραγωγής των χρονοσειρών. Ωστόσο, το πιο ενδιαφέρον εύρημα για αυτές τις χρονοσειρές προέρχεται από το συντελεστή συσχέτισης, όπου η λανθασμένη συμπεριφορά του (υποφαινομενική συσχέτιση) περιορίστηκε ως και αφαιρέθηκε για τα υποδείγματα  $AR(1) - ARCH(1)$  στη μεταβλητή μέσω του στατιστικού ελέγχου της μηδενικής υπόθεσης για μηδενική συσχέτιση, αν και οι τιμές των μέσων όρων του παρέμειναν στο μηδέν. Πιο συγκεκριμένα, το ποσοστό απόρριψης της μηδενικής υπόθεσης του μηδενικού συσχετισμού μειώνονταν καθώς η τιμή της παραμέτρου της μη γραμμικότητας αυξανόταν, φτάνοντας ασυμπτωτικά στα ονομαστικά επίπεδα της εφαρμογής του ελέγχου και, ως εκ τούτου, εξαλείφοντας την παρουσία υποφαινομενικών συμπεριφορών. Ίσως, το φαινόμενο αυτό να αναφέρεται μόνο σε γραμμικά παραγόμενες χρονοσειρές. Αντίθετα, η συμπεριφορά του φαινομένου της υποφαινομενικής συσχέτισης παρέμεινε αμετάβλητη για τα υποδείγματα  $AR(1) - ARCH(1)$  στα σφάλματα, δηλώνοντας ότι αυτή η μορφή της ετεροσκεδαστικότητας δεν αλλοίωσε καθόλου τη συμπεριφορά του φαινομένου η οποία ήταν ολόιδια με αυτή για στάσιμες  $AR(1)$  χρονοσειρές με ομοσκεδαστικότητα.

Για την περίπτωση δύο γραμμικά εξαρτημένων χρονοσειρών τα ευρήματα είναι ακόμη πιο συναρπαστικά. Η μεροληψία του συντελεστή συσχέτισης αυξάνεται τρομερά καθώς αυξάνεται η παράμετρος της μη γραμμικότητας φτάνοντας ασυμπτωτικά σε παραπλανητικά αποτελέσματα, όπως ενδείξεις ανεξάρτητων χρονοσειρών για εξαρτημένες χρονοσειρές ή με άλλα λόγια υποφαινομενικές συμπεριφορές από διαφορετική οπτική γωνία. Η μεροληψία όμως του συντελεστή συσχέτισης



εμφανίζεται και για τα υποδείγματα  $AR(1) - ARCH(1)$  στα σφάλματα αλλά με πολύ μικρότερο μέγεθος. Τέλος, σοβαρά προβλήματα στη δύναμη του στατιστικού ελέγχου της μηδενικής υπόθεσης της μηδενικής συσχέτισης εμφανίστηκαν και για τα δύο είδη υποδειγμάτων με τα  $AR(1) - ARCH(1)$  στη μεταβλητή να έχουν μεγαλύτερες απώλειες στη ισχύ του ελέγχου.

Σε αντίθεση με τον συντελεστή συσχέτισης, οι τιμές των μέσων όρων των μέτρων εξάρτησης του Spearman's  $\rho$  και του Kendall's  $\tau$  μαζί με τις Ελλειπτικές *Copulas* παρέμειναν σταθερές γύρω από την προκαθορισμένη τιμή του συντελεστή συσχέτισης που χρησιμοποιήθηκε για τη δημιουργία του επιπέδου γραμμικής εξάρτησης και δεν επηρεάζονται από τη δομή της παραγόμενης σειράς ή από οποιαδήποτε άλλο παράγοντα. Σημειώνεται ότι οι τιμές των μέσων όρων του Spearman's  $\rho$  και των Ελλειπτικών *Copulas*, δηλαδή του Gaussian και του Student-t, ήταν πολύ κοντά στις παραγόμενες τιμές του συντελεστή συσχέτισης, γεγονός που δηλώνει ότι αυτά τα μέτρα εξάρτησης ήταν σε θέση να αντιληφθούν τη γραμμική εξάρτηση για μη γραμμικές χρονοσειρές πολύ καλύτερα από τον συντελεστή συσχέτισης του δείγματος. Τα ίδια συμπεράσματα ισχύουν και για υποδείγματα  $AR(1) - ARCH(1)$  στα σφάλματα.

Παρομοίως, οι Αρχιμήδειες *Copulas* ήταν σε θέση να αναγνωρίσουν διαφορετικά επίπεδα εξάρτησης που δεν επηρεάζονταν από την παράμετρο της μη γραμμικότητας ή από οποιονδήποτε άλλο παράγοντα, εκτός από το μέγεθος της παραγόμενης τιμής του συντελεστή συσχέτισης, αν και είναι δύσκολο να ερμηνευτούν οι τιμές που λαμβάνουν. Κατά συνέπεια, όλα τα προαναφερθέντα μέτρα εξάρτησης μπόρεσαν να αποτυπώσουν τη γραμμική εξάρτηση και για τα δύο διαφορετικά είδη υποδειγμάτων καλύτερα από τον συντελεστή συσχέτισης, χωρίς να επηρεάζονται σημαντικά από τις τιμές των παραμέτρων που χρησιμοποιήθηκαν και το μέγεθος του δείγματος και κυρίως για τα μη γραμμικά υποδείγματα, δηλαδή για τα υποδείγματα  $AR(1) - ARCH(1)$  στη μεταβλητή.



## **ΚΕΦΑΛΑΙΟ 5**

### **ΥΠΟΦΑΙΝΟΜΕΝΙΚΕΣ ΠΑΛΙΝΔΡΟΜΗΣΕΙΣ ΓΙΑ ΥΠΟΔΕΙΓΜΑΤΑ ΜΕ ΥΠΟ ΣΥΝΘΗΚΗ ΕΤΕΡΟΣΚΕΔΑΣΤΗΚΟΤΗΤΑ ΠΡΩΤΟΥ ΒΑΘΜΟΥ**

#### **5.1 Εισαγωγή**

Το πρόβλημα της αυτοσυσχέτισης (autocorrelation) των σφαλμάτων εμφανίζεται πολύ συχνά στην ανάλυση της παλινδρόμησης όταν κατά τη διερεύνηση της συμπεριφοράς ενός φαινομένου, χρησιμοποιούνται μεταβλητές, τα δεδομένα των οποίων προέρχονται από χρονοσειρές, οδηγώντας τον αναλυτή σε λανθασμένα συμπεράσματα εάν η ύπαρξή του δεν αντιμετωπιστεί σωστά. Ωστόσο, τον ενδιαφέρον του προβλήματος αυτού επικεντρώνεται στον τρόπο δημιουργίας του και ειδικότερα αν είναι αληθινό σύμπτωμα αυτοσυσχέτισης των σφαλμάτων ή ψεύτικο, δηλαδή να εμφανίζεται ως ένδειξη κάποιου άλλου προβλήματος, όπως ενδεικτικά μπορεί να προέλθει από υπό συνθήκη ετεροσκεδαστικότητα (autoregressive conditional heteroscedasticity) στις τιμές των σφαλμάτων αλλά και από λανθασμένο σχηματισμό του υποδείγματος (model misspecification).

Πραγματικά, η παρουσία της αυτοσυσχέτισης στις τιμές των σφαλμάτων, ως ψευδής ένδειξη, συνδέεται σε μεγάλο βαθμό με το πρόβλημα των υποφαινομενικών παλινδρομήσεων (spurious regression), ένα φαινόμενο που παρουσιάστηκε αρχικά στη βιβλιογραφία από την πρωτοποριακή εργασία των Granger και Newbold (1974) και τεκμηριώθηκε μαθηματικά από τον Phillips (1986), αν και εισήχθη το φαινόμενο αυτό από τον Yule (1926) ως φαινόμενο υποφαινομενικών συσχετίσεων. Πιο συγκεκριμένα, οι Granger και Newbold (1974) χρησιμοποιώντας Monte Carlo ανάλυση έδειξαν ότι η



παλινδρόμηση δύο ανεξάρτητων χρονοσειρών τυχαίου περπατήματος χωρίς μέσο όρο (random walk processes without drift) θα δημιουργήσει ισχυρές ενδείξεις γραμμικής σχέσης αυτών των μεταβλητών, δεδομένου ότι η μηδενική υπόθεση για τον έλεγχο της στατιστικής σημαντικότητας του συντελεστή της ανεξάρτητης μεταβλητής θα απορριφθεί κατά 76% για σειρές 100 παρατηρήσεων σε ονομαστικό επίπεδο 5%, ενώ για δύο ανεξάρτητες χρονοσειρές τυχαίου περπατήματος με μέσο όρο το ποσοστό απόρριψης της εν λόγω μηδενικής υπόθεσης έφτανε στο 100% όπως παρουσιάστηκε από τον Entorf (1997). Η εργασία των Granger και Newbold (1974) επεκτάθηκε από τους Granger et al. (2001) σε δύο ανεξάρτητες στάσιμες αυτοπαλίνδρομες πρώτου βαθμού χρονοσειρές AR(1) (Autoregressive processes of order one) ανακαλύπτοντας ποσοστά απόρριψης της μηδενικής υπόθεσης για το έλεγχο της στατιστικής σημαντικότητας του συντελεστή της ανεξάρτητης μεταβλητής μικρότερα από αυτά των δύο ανεξάρτητων χρονοσειρών τυχαίου περπατήματος χωρίς μέσο όρο, που καθορίζονταν όμως από την τιμή της παραμέτρου του αυτοπαλίνδρομου όρου ανεξάρτητα από το μέγεθος δείγματος, π.χ. τα ποσοστά απόρριψης της μηδενικής για τιμές της παραμέτρου του αυτοπαλίνδρομου όρου 0.2, 0.5, 0.8 και 0.9 είναι αντίστοιχα 6%, 13%, 36% και 52% για ονομαστικό επίπεδο σημαντικότητας 5%.

Επιπρόσθετα, οι Granger και Newbold (1974) επεσήμαναν ότι μαζί με τις μεγάλες τιμές της στατιστικής  $t$  που απέρριπταν πολύ συχνά τον έλεγχο της μηδενικής υπόθεσης, χαμηλές τιμές της στατιστικής των Durbin-Watson θα εμφανιστούν στην ανάλυση παλινδρόμησης, ένα εύρημα που διερευνήθηκε από τον Marmol (1995) και ο οποίος αποκάλυψε ότι η στατιστική των Durbin-Watson συγκλίνει κατά πιθανότητα στο μηδέν. Ενδείξεις αυτοσυσχέτισης αναφέρονται επίσης στην εργασία των Newbold και Davies (1978), για ανεξάρτητες μη στάσιμες χρονοσειρές κινητού μέσου όρου (non-stationary moving average processes), στον Agiakloglou (2009), για ανεξάρτητες στάσιμες χρονοσειρές AR(1) και στους Agiakloglou et al. (2015), για ανεξάρτητες στάσιμες χωρικές αυτοπαλίνδρομες χρονοσειρές πρώτου βαθμού SAR(1) (stationary spatial autoregressive processes of order one). Θα πρέπει ωστόσο να σημειωθεί ότι η εμφάνιση της αυτοσυσχέτισης στην ανάλυση παλινδρόμησης βοηθά σημαντικά τον αναλυτή να εντοπίσει πρόβλημα στην παλινδρόμηση και να αμφισβητήσει την



αξιοπιστία των αποτελεσμάτων του, κάτι που στην ανάλυση συσχέτισης ο αναλυτής δεν έχει ενδείξεις αλλά και στοιχεία για ψευδή σχέση.

Στο κεφάλαιο αυτό διερευνάται μέσω ανάλυσης Monte Carlo η συμπεριφορά του φαινομένου των υποφαινομενικών παλινδρομήσεων (spurious regressions) για δύο ανεξάρτητες στάσιμες αυτοπαλίνδρομου πρώτου βαθμού AR(1) χρονοσειρές αλλά και δύο ανεξάρτητες μη στάσιμες τυχαίου περπατήματος χρονοσειρές, των οποίων τα σφάλματα δεν έχουν σταθερή διακύμανση, αλλά η διαχρονικά μεταβαλλόμενη διακύμανσή τους κατασκευάζεται από ένα αυτοπαλίνδρομο υπό όρους ετεροσκεδαστικότητας υπόδειγμα πρώτου βαθμού, δηλαδή από ένα ARCH( 1) υπόδειγμα, όπου η μεταβλητότητα στη διακύμανση προέρχεται είτε από τις τιμές των σφαλμάτων είτε από τις τιμές της ίδιας της χρονοσειράς, με την τελευταία μορφή να σχηματίζει μια οικογένεια χρονοσειρών που μπορούν να θεωρηθούν ως μη γραμμικές διεργασίες. Αναμφισβήτητα, η μη γραμμικότητα μπορεί να εμφανιστεί με διάφορους τρόπους, αλλά στην εργασία αυτή επιλέγεται να παρουσιαστεί μέσω της αυτοπαλίνδρομης υπό όρους ετεροσκεδαστικότητας δεδομένου ότι αυτά τα υποδείγματα έχουν ήδη χρησιμοποιηθεί για τη διερεύνηση των υποφαινομενικών συσχετίσεων στο προηγούμενο κεφάλαιο και διαθέτουν και καλές γραμμικές ιδιότητες. Σημειώνεται μάλιστα ότι η μελέτη του φαινομένου αυτού για μη γραμμικά υποδείγματα δεν έχει παρουσιαστεί στην βιβλιογραφία αν εξαιρέσει κανείς την εργασία των Lee et al. (2005) στην οποία διερεύνησαν την υποφαινομενικές μη γραμμικές σχέσεις μεταξύ δύο ανεξάρτητων τυχαίων περιπάτου χρονοσειρών. Επίσης, διερευνάται η ύπαρξη αυτοσυσχέτισης στις τιμές των σφαλμάτων, καθώς και ο τρόπος αντιμετώπισης και των δύο προβλημάτων.

## 5.2 Παρουσίαση των υποδειγμάτων της προσομοίωσης

Το αυτοπαλίνδρομο υπόδειγμα πρώτου βαθμού AR(1), το οποίο δίνεται από την ακόλουθη σχέση:

$$X_t = \varphi X_{t-1} + \varepsilon_t \quad (5.1)$$





θεωρώντας ότι ο μέσος όρος της χρονοσειράς είναι μηδέν και ότι το τυχαίο σφάλμα ακολουθεί τη διαδικασία του λευκού θορύβου (white noise process), χρησιμοποιείται ως σημείο αναφοράς σε όλες τις προσπάθειες διερεύνησης υποφαινομενικών παλινδρομήσεων μέσω ανάλυσης Monte Carlo διότι είναι απλό, έχει καλές ειδικότητες και μπορεί να περιλάβει στάσιμες και μη στάσιμες διαδικασίες, όπως ακόμα και τη διαδικασία του λευκού θορύβου, δίνοντας τη δυνατότητα να εξαχθούν πολύ χρήσιμα συμπεράσματα για τη συμπεριφορά του φαινομένου. Πιο συγκεκριμένα, το υπόδειγμα αυτό, το οποίο προέρχεται από μία μεγάλη οικογένεια υποδειγμάτων, γνωστά ως αυτοπαλίνδρομα κινητού μέσου όρου υποδείγματα ARMA(p, q) (AutoRegressive Moving Average processes), τα οποία παρουσιάστηκαν από τους Box και Jenkins (1976), είναι γραμμικό με σταθερή τιμή της παραμέτρου  $\varphi$  του αυτοπαλίνδρομου όρου και διακρίνεται α) σε στάσιμη χρονοσειρά, όταν η τιμή του  $\varphi$  είναι σε απόλυτο μέγεθος μικρότερη της μονάδας, με τον μέσο όρο του και τη διακύμανσή του να έχουν σταθερές τιμές, β) σε μη στάσιμη χρονοσειρά και ειδικότερα σε υπόδειγμα τυχαίου περπατήματος χωρίς μέσο όρο (random walk process without drift), όταν η παράμετρος  $\varphi$  λάβει την τιμή ένα, και γ) σε χρονοσειρά λευκού θορύβου, όταν η παράμετρος  $\varphi$  λάβει την τιμή μηδέν.

Αναμφισβήτητα, η στασιμότητα είναι ένα πάρα πολύ σημαντικό στοιχείο για την ανάλυση των χρονοσειρών, διότι εξασφαλίζει πολύ καλές ιδιότητες και επιτρέπει την εκτίμηση της χρονοσειράς, καθώς και το σχηματισμό των προβλέψεων της. Ωστόσο, πολύ συχνά η ιδιότητα αυτή δεν παρατηρείται σε πολλές οικονομικές αλλά και χρηματοοικονομικές μεταβλητές, είτε διότι παραβιάζεται η υπόθεση του σταθερού μέσου όρου είτε διότι παραβιάζεται η υπόθεση της σταθερής διακύμανσης, δημιουργώντας δυσκολίες στην ανάλυσή τους, δεδομένου ότι οι μεταβλητές αυτές εμφανίζονται ως μη στάσιμες διαδικασίες. Για το λόγο αυτό, απαιτείται να εφαρμοστούν κυρίως πρακτικοί τρόποι αντιμετώπισης του προβλήματος αυτού, έτσι ώστε η χρονοσειρά να μετατραπεί σε στάσιμη διαδικασία για να μπορεί να αναλυθεί. Ενδεικτικά και για την περίπτωση όπου το πρόβλημα της μη στασιμότητας επικεντρώνεται στο μέσο όρο της χρονοσειράς, τότε μια μη στάσιμη χρονοσειρά που δεν έχει σταθερό μέσο όρο μετατρέπεται τυπικά σε στάσιμη λαμβάνοντας πρώτες διαφορές. Αντίθετα, αν το πρόβλημα της μη στασιμότητας προέρχεται από μη σταθερή





διακύμανση, τότε ένας πιθανός μετασχηματισμός των τιμών της χρονοσειράς λαμβάνοντας λογαρίθμους, μπορεί να συμβάλλει στην απόκτηση της στασιμότητας, αν και σε γενικές γραμμές δεν είναι εύκολη η αντιμετώπιση αυτού του προβλήματος.

Η πιο απλή μορφή μίας όχι σταθερής αλλά διαχρονικά μεταβαλλόμενης διακύμανσης δίνεται από ένα Αυτοπαλίνδρομο Υπό Συνθήκης Ετεροσκεδαστικότητας υπόδειγμα πρώτου βαθμού (AutoRegressive Conditionally Heteroscedastic model of order one), γνωστό ως υπόδειγμα ARCH(1), που παρουσιάστηκε από τον Engle (1982) και ορίζεται ως εξής:

$$\varepsilon_t = u_t \sqrt{a_0 + a_1 \varepsilon_{t-1}^2} \quad (5.2)$$

όπου  $a_0 > 0$ ,  $0 < a_1 < 1$  και το τυχαίο σφάλμα ακολουθεί την τυπική κανονική κατανομή με μέσο όρο μηδέν και διακύμανση 1, δηλαδή  $u_t \sim iid N(0, 1)$ . Το υπόδειγμα της σχέσης (5.2) δηλώνει ότι οι τιμές της διακύμανσης καθορίζονται από τα τετράγωνα των τιμών του τυχαίου σφάλματος της προηγούμενης περιόδου, ενώ ο συνδυασμός των σχέσεων (5.1) και (5.2) σχηματίζει μία οικογένεια υποδειγμάτων, γνωστά ως AR(1) – ARCH(1) στα σφάλματα υποδείγματα, όπου από εδώ και στο εξής θα αποκαλούνται υποδείγματα Σειράς I. τα οποία είναι γραμμικά με σταθερή παράμετρο και μέσο όρο, έχοντας όμως διακύμανση μη σταθερή που εξαρτάται από τις προηγούμενες τιμές του σφάλματος αλλά που είναι σταθερή χωρίς εξάρτηση (constant unconditional variance). Σημειώνεται ότι διαγραμματικές απεικονίσεις των υποδειγμάτων αυτών έχουν παρουσιαστεί στο προηγούμενο κεφάλαιο και ότι όταν η τιμή της παραμέτρου  $a_1$  λάβει την τιμή μηδέν, τότε η οικογένεια αυτή των υποδειγμάτων, δηλαδή η Σειρά I, γίνεται μία στάσιμη αυτοπαλίνδρομου πρώτου βαθμού AR(1) χρονοσειρά.

Οι προδιαγραφές των προαναφερθέντων υποδειγμάτων αλλάζουν όταν η μεταβλητότητα των τιμών της χρονοσειράς γίνεται πιο έντονη και κυρίως όταν αυτή προέρχεται από τις ίδιες τις τιμές της. Μάλιστα, στην περίπτωση όπου η διαχρονικά μεταβαλλόμενη διακύμανση καθορίζεται από ένα υπόδειγμα ARCH(1) της μορφής:

$$\varepsilon_t = u_t \sqrt{a_0 + a_1 X_{t-1}^2} \quad (5.3)$$

όπου  $a_0 > 0$ ,  $0 < a_1 < 1$  και το τυχαίο σφάλμα ακολουθεί την τυπική κανονική κατανομή με μέσο όρο μηδέν και διακύμανση 1, δηλαδή  $u_t \sim iid N(0, 1)$ , τότε οι τιμές της



διακύμανσης καθορίζονται από τα τετράγωνα των τιμών της μεταβλητής της προηγούμενης περιόδου και όχι από τα τετράγωνα των τιμών του τυχαίου σφάλματος. Ο συνδυασμός των σχέσεων (5.1) και (5.3) σχηματίζει μία οικογένεια υποδειγμάτων, γνωστά ως AR(1) – ARCH(1) στη μεταβλητή υποδείγματα, τα οποία παρουσιάστηκαν από τους Bera et al. (1992 και 1996), όπου από εδώ και στο εξής θα αποκαλούνται υποδείγματα Σειράς II, και ομοιάζουν με τα υποδείγματα της Σειράς I. Μάλιστα, στο προηγούμενο κεφάλαιο έχουν παρουσιαστεί διαγραμματικές απεικονίσεις των υποδειγμάτων αυτών, τα οποία έχουν την ακόλουθη υπό συνθήκη κατανομή:

$$X_t|X_{t-1} \sim N(\varphi X_{t-1}, a_0 + a_1 X_{t-1}^2) \quad (5.4)$$

ενώ στην πιο απλή μορφή τους, δηλαδή όταν η τιμή της παραμέτρου  $a_1$  λάβει την τιμή μηδέν, γίνονται στάσιμα αυτοπαλίνδρομα πρώτου βαθμού υποδείγματα, όπως και τα υποδείγματα των Σειρών I.

Η διαφορά μεταξύ αυτών των δύο οικογενειών υποδειγμάτων δεν επικεντρώνεται στον τρόπο εμφάνισης της μεταβλητότητας ή καλύτερα στο τρόπο σχηματισμού των τιμών της διακύμανσης, αλλά στο γεγονός ότι η δεύτερη οικογένεια υποδειγμάτων μπορεί να εκφραστεί και ως μια μη γραμμική διαδικασία. Ακολουθώντας τον ισχυρισμό του Granger (2008) ότι οποιοδήποτε μη γραμμικό μοντέλο μπορεί κατά προσέγγιση να παρουσιαστεί με ένα γραμμικό μοντέλο μεταβαλλόμενης διαχρονικά παραμέτρου (time varying process), η Σειρά II μπορεί να γραφτεί ως:

$$X_t = \varphi_t X_{t-1} + w_t \quad (5.5)$$

όπου  $\varphi_t \sim iid(\varphi, a_1)$  και  $w_t \sim iid N(0, a_0)$ . Στην περίπτωση αυτή, το υπόδειγμα της σχέσης (5.5) μπορεί εξίσου να θεωρηθεί ως Σειρά II, δηλαδή ως AR(1) – ARCH(1) στη μεταβλητή, αφού η υπό συνθήκη κατανομή της μεταβλητή  $X_t$ , ως ένα αυτοπαλίνδρομο υπόδειγμα πρώτης τάξης με ένα τυχαίο συντελεστή (random coefficient), είναι ακριβώς η ίδια με αυτή που προκύπτει από τη Σειρά II και δίνεται από τη σχέση (5.4), ενώ η άνευ όρων διακύμανση (unconditional variance) της μεταβλητής  $X_t$  είναι σταθερή και ορίζεται ως:

$$Var(X_t) = \frac{a_0}{1 - \varphi^2 - a_1}$$

και η οποία υπολογίζεται με τον ακόλουθο τρόπο:



$$\begin{aligned}\text{Var}(X_t) &= E[\text{Var}(X_t|X_{t-1})] + \text{Var}[E(X_t|X_{t-1})] = E[a_0 + a_1 X_{t-1}^2] + \text{Var}[\varphi X_{t-1}] \\ &= a_1 \text{Var}(X_t) + \varphi^2 \text{Var}(X_t)\end{aligned}$$

και εξαρτάται από τις τιμές των παραμέτρων  $\varphi$  and  $a_1$ . Κατά συνέπεια, για μεγάλες τιμές και των δύο παραμέτρων η τυχειότητα της αυτοπαλίνδρομης παραμέτρου αυξάνει την άνευ όρων τη μεταβλητότητα της μεταβλητής, μειώνοντας ουσιαστικά τον βαθμό της γραμμικής δομής της χρονοσειράς, όπως αναλύθηκε από τους Bera et al. (1992 και 1996), σε αντίθεση με τη Σειρά I, όπου η άνευ όρων διακύμανση αυτής της σειράς δεν περιλαμβάνει την τιμή της μη γραμμικής παραμέτρου  $a_1$  στον παρονομαστή της.

### 5.3 Παρουσίαση αποτελεσμάτων προσομοίωσης

Για τη διερεύνηση της συμπεριφοράς του φαινομένου των υποφαινομενικών παλινδρομήσεων από διαδικασίες με ετεροσκεδαστικότητα στα σφάλματα τύπου ARCH(1) εφαρμόστηκε ανάλυση Monte Carlo εις την οποίαν εκτιμήθηκε, για δύο ανεξάρτητες χρονοσειρές  $Y_t$  και  $X_t$ , για  $t = 1, 2, \dots, T$ , το ακόλουθο απλό γραμμικό υπόδειγμα:

$$Y_t = \alpha + \beta X_t + \zeta_t \quad (5.6)$$

όπου το τυχαίο σφάλμα ικανοποιεί όλες τις απαραίτητες προϋποθέσεις για να εκτιμηθεί το υπόδειγμα αυτό με τη μέθοδο των ελαχίστων τετραγώνων, δηλαδή  $\zeta_t \sim iid N(0, \sigma^2)$ , και οι μεταβλητές  $Y_t$  και  $X_t$  κατασκευάζονται με δύο διαφορετικές διαδικασίες, δηλαδή ως Σειρές I, AR(1) – ARCH(1) στα σφάλματα, και ως Σειρές II, AR(1) – ARCH(1) στη μεταβλητή, για ακόλουθες τιμές των παραμέτρων, of  $\varphi = 0.0, 0.2, 0.5, 0.8, 0.9$  και  $1.0$ ,  $a_0 = 1.0$ ,  $a_1 = 0.0, 0.2, 0.5, 0.8$  και  $0.9$ , καθώς και για δείγματα μεγέθους 50, 100, 500 και 1000 παρατηρήσεων, ενώ η προσομοίωση πραγματοποιήθηκε στην Python.

Στον Πίνακα 5.1 παρουσιάζονται τα ποσοστά απόρριψης της μηδενικής υπόθεσης ότι ο συντελεστής  $\beta$  του υποδείγματος (5.6) είναι μηδέν για ονομαστικό επίπεδο 5% και για δύο ανεξάρτητες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές βασιζόμενα σε



10000 επαναλήψεις. Όπως παρατηρείται από τον πίνακα αυτόν, η παρουσία αυτής της δομής σφαλμάτων δεν άλλαξε καθόλου τη συμπεριφορά του φαινομένου όπως αυτή προκύπτει από δύο ανεξάρτητες στάσιμες AR(1) διαδικασίες ή από δύο ανεξάρτητες μη στάσιμες τυχαίου περπατήματος χωρίς μέσο όρο διαδικασίες με ομοσκεδαστικά σφάλματα. Αυτό σημαίνει ότι η ετεροσκεδαστικότητα αυτής της μορφής δεν επηρέασε καθόλου τη διακύμανση του συντελεστή  $\beta$  του υποδείγματος και έτσι η συμπεριφορά του ελέγχου δεν άλλαξε καθόλου δίνοντας τα ίδια ευρήματα.

### Πίνακας 5.1

**Ποσοστά απόρριψης της μηδενικής υπόθεσης ότι  $\beta = 0$  σε επίπεδο σημαντικότητας 5% για δύο ανεξάρτητες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές βασισμένα σε 10000 επαναλήψεις**

$a_1$	$\varphi = 0.0$	$\varphi = 0.2$	$\varphi = 0.5$	$\varphi = 0.8$	$\varphi = 0.9$	$\varphi = 1.0$
<b>0.0</b>	0.054	0.062	0.127	0.335	0.472	0.678
	0.051	0.060	0.128	0.350	0.505	0.763
	0.052	0.063	0.132	0.355	0.518	0.895
	0.051	0.060	0.128	0.355	0.524	0.930
<b>0.2</b>	0.053	0.061	0.125	0.336	0.474	0.684
	0.052	0.060	0.124	0.350	0.505	0.770
	0.053	0.062	0.134	0.351	0.512	0.898
	0.053	0.061	0.127	0.353	0.526	0.928
<b>0.5</b>	0.054	0.062	0.124	0.331	0.476	0.679
	0.052	0.062	0.122	0.348	0.505	0.772
	0.052	0.059	0.128	0.351	0.516	0.898
	0.054	0.062	0.128	0.361	0.527	0.926
<b>0.8</b>	0.058	0.066	0.121	0.322	0.470	0.675
	0.054	0.064	0.118	0.333	0.498	0.771
	0.050	0.060	0.116	0.333	0.504	0.900
	0.051	0.061	0.121	0.338	0.507	0.928
<b>0.9</b>	0.058	0.067	0.119	0.316	0.466	0.673
	0.058	0.066	0.118	0.328	0.492	0.772
	0.052	0.060	0.111	0.321	0.495	0.895
	0.052	0.056	0.114	0.326	0.495	0.929

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα 50, 100, 500 και 1000 παρατηρήσεων.



Πραγματικά, το ποσοστό απόρριψης της μηδενικής υπόθεσης για  $\beta = 0$  για στάσιμες AR(1) διαδικασίες καθορίζεται μόνο από την τιμή του αυτοπαλίνδρομου όρου και όχι από το μέγεθος του δείγματος και αυξάνεται καθώς η τιμή του  $\varphi$  αυξάνεται, φτάνοντας για  $\varphi = 0.9$  στο 0.524 για ονομαστικό επίπεδο σημαντικότητας 5%. Αντίθετα, για μη στάσιμες χρονοσειρές και συγκεκριμένα για χρονοσειρές τυχαίου περιπάτου χωρίς μέσο όρο η μηδενική υπόθεση απορρίπτεται συχνότερα καθώς το μέγεθος του δείγματος αυξάνεται, φτάνοντας ασυμπτωτικά στο επίπεδο απόρριψης 100%. Ως εκ τούτου, η ετεροσκεδαστικότητα σε αυτή τη μορφή δεν έχει καμία επίδραση στη γνωστή συμπεριφορά του φαινομένου της υποφαινομενικής παλινδρόμησης.

**Πίνακας 5.2**

**Ποσοστά απόρριψης της μηδενικής υπόθεσης ότι  $\beta = 0$  σε επίπεδο σημαντικότητας 5% για δύο ανεξάρτητες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές βασισμένα σε 10000 επαναλήψεις**

$a_1$	$\varphi = 0.0$	$\varphi = 0.2$	$\varphi = 0.5$	$\varphi = 0.8$	$\varphi = 0.9$	$\varphi = 1.0$
<b>0.0</b>	0.054	0.062	0.127	0.335	0.472	0.678
	0.051	0.060	0.128	0.350	0.505	0.763
	0.052	0.063	0.132	0.355	0.518	0.895
	0.051	0.060	0.128	0.355	0.524	0.930
<b>0.2</b>	0.053	0.061	0.123	0.288	0.378	0.471
	0.052	0.060	0.121	0.304	0.389	0.485
	0.053	0.063	0.131	0.312	0.388	0.392
	0.053	0.061	0.127	0.326	0.383	0.335
<b>0.5</b>	0.054	0.064	0.111	0.203	0.235	0.265
	0.052	0.062	0.109	0.199	0.228	0.238
	0.052	0.060	0.110	0.174	0.170	0.140
	0.054	0.061	0.114	0.168	0.147	0.112
<b>0.8</b>	0.058	0.064	0.098	0.146	0.161	0.178
	0.054	0.062	0.095	0.131	0.142	0.153
	0.050	0.058	0.086	0.097	0.094	0.088
	0.051	0.055	0.081	0.081	0.073	0.061
<b>0.9</b>	0.058	0.064	0.093	0.136	0.145	0.160
	0.058	0.062	0.091	0.120	0.129	0.139
	0.052	0.056	0.079	0.087	0.084	0.078
	0.052	0.054	0.071	0.069	0.064	0.053

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα 50, 100, 500 και 1000 παρατηρήσεων.



Η συμπεριφορά του φαινομένου ωστόσο αλλάζει από τη στιγμή που η ετεροσκεδαστικότητα προέρχεται από τις τιμές της μεταβλητής και όχι από τις τιμές των σφαλμάτων. Όπως φαίνεται από τον Πίνακα 5.2, ο οποίος παρουσιάζει τα ποσοστά απόρριψης της μηδενικής υπόθεσης ότι ο συντελεστής  $\beta$  του υποδείγματος (5.6) είναι μηδέν σε επίπεδο σημαντικότητας 5% για δύο ανεξάρτητες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές για 10000 επαναλήψεις, το ποσοστό απόρριψης της μηδενικής υπόθεσης μειώνεται καθώς η τιμή της παραμέτρου της μη γραμμικότητας αυξάνεται για δεδομένες τιμές της παραμέτρου του αυτοπαλίνδρομου όρου και του μεγέθους του δείγματος, ανεξάρτητα από το εάν οι διαδικασίες είναι στάσιμες ή μη.

Πραγματικά, το φαινόμενο της υποφαινομενικής παλινδρόμησης περιορίζεται σταδιακά, και για τα δύο είδη χρονοσειρών, δηλαδή στάσιμες και μη στάσιμες, έως ότου εξαλειφθεί τελείως ασυμπτωτικά, καθώς το μέγεθος του δείγματος αυξάνεται μαζί με τις τιμές των παραμέτρων  $\phi$  και  $a_1$ , σε αντίθεση με την προηγούμενη περίπτωση ετεροσκεδαστικότητας, όπου ο αναλυτής θα είχε λάβει, σχεδόν με βεβαιότητα, παραπλανητικά αποτελέσματα, για τη σχέση δύο ανεξάρτητων μη στάσιμων χρονοσειρών, δεδομένου ότι το ποσοστό απόρριψης της μηδενικής υπόθεσης φτάνει στο 100%. Πιο συγκεκριμένα, για την περίπτωση δύο ανεξάρτητων τυχαίου περπατήματος διαδικασιών χωρίς μέσο όρο, δηλαδή για  $\phi = 1$ , το ποσοστό απόρριψης της μηδενικής υπόθεσης για χρονοσειρές 1000 παρατηρήσεων και για  $a_1 = 0.9$  γίνεται 0.053, επιτυγχάνοντας στην ουσία το ονομαστικό επίπεδο 5% , όπως φαίνεται στον Πίνακα 5.2, ενώ ήταν 0.929 και 0.930 υπό την παρουσία του ARCH(1) στα σφάλματα ή υπό την απουσία ετεροσκεδαστικότητας αντίστοιχα, όπως παρουσιάζονται στον Πίνακα 5.1.

Ως εκ τούτου, η παρουσία αυτής της μορφής ετεροσκεδαστικότητας, δηλαδή υποδείγματος ARCH(1) στη μεταβλητή, που εκφράζει μεγάλη μεταβλητότητα μειώνει το βαθμό εμφάνισης λανθασμένων ενδείξεων, ένα εύρημα το οποίο αναμφισβήτητα θα επηρεάζει τη συμπεριφορά των ελέγχων μοναδιαίας ρίζας καθώς και συνολοκλήρωσης. Επίσης, οι λανθασμένες ενδείξεις εξαφανίζονται ασυμπτωτικά, καθώς αυξάνεται η τιμή



της παραμέτρου της μη γραμμικότητας, όταν οι χρονοσειρές είναι λιγότερο συνεπείς και σταθερές (less persistent), δηλαδή όταν η τιμή του αυτοπαλίνδρομου όρου είναι διαφορετική από τη μονάδα. Για παράδειγμα, τα ποσοστά απόρριψης της μηδενικής υπόθεσης για  $\rho = 0.9$  σε ονομαστικό επίπεδο 5% μειώνονται από 0.524 σε 0.064 για σειρές 1000 παρατηρήσεων και για τιμές  $a_1$  ίσες με 0 και 0.9 αντίστοιχα, σημειώνοντας ότι τα διαφορετικά μεγέθη δείγματος έχουν πολύ μικρή επίδραση στα ποσοστά απόρριψης της μηδενικής υπόθεσης για στάσιμες χρονοσειρές, ειδικά για μικρές και μεσαίες τιμές του  $a_1$ , αλλά ελαφρώς μεγαλύτερο αποτέλεσμα για μεγάλες τιμές της παραμέτρου της μη γραμμικότητας.

## 5.4 Αυτοσυσχέτιση

Τα αποτελέσματα που ανακαλύφθηκαν για τη συμπεριφορά του φαινομένου υποφαινομενικών παλινδρομήσεων για χρονοσειρές με ετεροσκεδαστικότητα και ειδικότερα για αυτές τις χρονοσειρές που εμφανίζουν μεγάλη μεταβλητότητα διερευνήθηκαν περαιτέρω ως προς την αξιοπιστία των εκτιμήσεών τους. Πιο συγκεκριμένα, διερευνήθηκε η ύπαρξη αυτοσυσχέτισης στις τιμές των καταλοίπων της εκτίμησης του απλού γραμμικού υποδείγματος, μία ένδειξη που έχει αναφερθεί στη βιβλιογραφία ότι συνδέεται με το φαινόμενο αυτό, αρχικά από τους Granger και Newbold (1974) για υποδείγματα τυχαίου περπατήματος χωρίς μέσω όρο, όπου ο Marmol (1995) απέδειξε ότι η τιμή της στατιστικής των Durbin-Watson ασυμπτωτικά τείνει στο μηδέν, και στη συνέχεια από τους Agiakloglou και Agiropoulos (2021) για στάσιμα αυτοπαλίνδρομα πρώτου βαθμού υποδείγματα AR(1), οι οποίοι μάλιστα γενίκευσαν τα αποτελέσματα του Marmol (1995) και έδειξαν ότι η τιμή της στατιστικής των Durbin-Watson,  $d$ , τείνει ασυμπτωτικά στη σχέση  $d = 2(1 - \rho)$  και έτσι όταν  $\rho = 1$ , τότε  $d = 1$ .<sup>10</sup> Σημειώνεται δε ότι τα αποτελέσματα αυτά θα μπορούσαν να είχαν επιτευχθεί και με την ανάλυση της συσχέτισης κάτι που παρουσιάστηκε στο

---

<sup>10</sup> Αναφέρεται ότι η παράμετρος  $\rho$  εκφράζει τον συντελεστή του αυτοπαλίνδρομου όρου AR(1) των τιμών του τυχαίου μέρους του υποδείγματος που συνδέεται με την παράμετρο του αυτοπαλίνδρομου υποδείγματος AR(1) των μεταβλητών.



προηγούμενο κεφάλαιο. Ωστόσο, στην περίπτωση αυτή ο αναλυτής δεν θα είχε καμία ένδειξη για λανθασμένες ενδείξεις, δεδομένου ότι η ύπαρξη αυτοσυσχέτισης στα αποτελέσματα της παλινδρόμησης συνιστά στον αναλυτή να διερευνήσει περαιτέρω την αξιοπιστία των αποτελεσμάτων του.

**Πίνακας 5.3**

**Τιμές μέσων όρων της στατιστικής των Durbin-Watson για δύο ανεξάρτητες AR(1) – ARCH(1) στα σφάλματα χρονοσειρές βασισμένες σε 10000 επαναλήψεις**

$a_1$	$\phi = 0.0$	$\phi = 0.2$	$\phi = 0.5$	$\phi = 0.8$	$\phi = 0.9$	$\phi = 1.0$
<b>0.0</b>	1.997	1.642	1.112	0.595	0.437	0.321
	2.000	1.623	1.059	0.500	0.319	0.173
	2.000	1.605	1.012	0.419	0.222	0.036
	1.999	1.602	1.005	0.409	0.211	0.018
<b>0.2</b>	1.996	1.645	1.120	0.602	0.442	0.326
	1.999	1.625	1.064	0.505	0.323	0.175
	2.001	1.606	1.014	0.420	0.223	0.037
	2.000	1.602	1.006	0.410	0.211	0.018
<b>0.5</b>	1.995	1.651	1.133	0.615	0.453	0.336
	1.998	1.630	1.076	0.516	0.331	0.181
	2.003	1.610	1.020	0.426	0.227	0.037
	2.000	1.604	1.010	0.413	0.213	0.018
<b>0.8</b>	1.992	1.657	1.150	0.634	0.471	0.354
	1.995	1.637	1.095	0.537	0.349	0.195
	2.005	1.621	1.040	0.445	0.240	0.041
	2.001	1.612	1.025	0.428	0.224	0.020
<b>0.9</b>	1.991	1.660	1.157	0.642	0.479	0.362
	1.994	1.640	1.102	0.546	0.356	0.201
	2.006	1.626	1.050	0.455	0.248	0.043
	2.002	1.617	1.035	0.438	0.231	0.021

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα 50, 100, 500 και 1000 παρατηρήσεων.

Για το σκοπό αυτό, οι τιμές των μέσων όρων της στατιστικής Durbin-Watson (DW) υπολογίζονται από τις τιμές των καταλοίπων που προέκυψαν από την εκτίμηση του υποδείγματος (5.6) και για τις δύο υπό εξέταση Σειρές, με βάση την ίδια διαδικασία προσομοίωσης, και τα αποτελέσματα αυτής της προσπάθειας παρουσιάζονται





αντίστοιχα στους Πίνακες 5.3 και 5.4. Όπως φαίνεται από τον Πίνακα 5.3, οι τιμές των μέσων όρων της στατιστικής DW για τη Σειρά I, δηλαδή για τα υποδείγματα που δημιουργούνται από AR(1) – ARCH(1) στα σφάλματα, είναι πανομοιότυπες με αυτές που λαμβάνονται για στάσιμες AR(1) διαδικασίες και για μη στάσιμες, δηλαδή για διαδικασίες τυχαίου περπατήματος χωρίς μέσο όρο, υποδεικνύοντας ότι αυτή η μορφή ετεροσκεδαστικότητας παρήγαγε στοιχεία αυτοσυσχέτισης ίδια με αυτήν των υποδειγμάτων σταθερής διακύμανσης. Επομένως, η συνολική συμπεριφορά του φαινομένου της υποφαινομενικής παλινδρόμησης δεν άλλαξε καθόλου με την ύπαρξη ετεροσκεδαστικότητας

**Πίνακας 5.4**

**Τιμές μέσων όρων της στατιστικής των Durbin-Watson για δύο ανεξάρτητες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές βασισμένες σε 10000 επαναλήψεις**

$a_1$	$\varphi = 0.0$	$\varphi = 0.2$	$\varphi = 0.5$	$\varphi = 0.8$	$\varphi = 0.9$	$\varphi = 1.0$
<b>0.0</b>	1.997	1.642	1.112	0.595	0.437	0.321
	2.000	1.623	1.059	0.500	0.319	0.173
	2.000	1.605	1.012	0.419	0.222	0.036
	1.999	1.602	1.005	0.409	0.211	0.018
<b>0.2</b>	1.996	1.649	1.137	0.666	0.546	0.476
	1.999	1.628	1.077	0.565	0.431	0.354
	2.001	1.607	1.017	0.455	0.311	0.244
	2.000	1.602	1.008	0.436	0.290	0.231
<b>0.5</b>	1.995	1.668	1.198	0.816	0.731	0.675
	1.998	1.643	1.133	0.728	0.644	0.593
	2.003	1.615	1.055	0.624	0.551	0.520
	2.000	1.607	1.036	0.605	0.539	0.515
<b>0.8</b>	1.992	1.694	1.278	0.961	0.892	0.843
	1.995	1.669	1.218	0.897	0.832	0.786
	2.005	1.641	1.149	0.828	0.773	0.742
	2.001	1.629	1.127	0.818	0.770	0.738
<b>0.9</b>	1.991	1.704	1.306	1.006	0.939	0.892
	1.994	1.680	1.250	0.949	0.885	0.840
	2.006	1.655	1.189	0.892	0.838	0.802
	2.002	1.643	1.169	0.884	0.836	0.801

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα 50, 100, 500 και 1000 παρατηρήσεων.



Ωστόσο, ενδείξεις αυτοσυσχέτισης λαμβάνονται επίσης και από την εκτίμηση του υποδείγματος (5.6) χρησιμοποιώντας δεδομένα από τα υποδείγματα της Σειράς II, δηλαδή από τα υποδείγματα που δημιουργούνται από  $AR(1) - ARCH(1)$  στη μεταβλητή, όπως παρουσιάζονται στον Πίνακα 5.4, αλλά όχι με το ίδιο μοτίβο με αυτό της Σειράς I. Σε αυτήν την περίπτωση, οι τιμές των μέσων όρων της στατιστικής των Durbin-Watson είναι μεγαλύτερες από αυτές που αναφέρονται στον Πίνακα 5.3 για όλες τις τιμές των παραμέτρων  $\phi$  και  $a_1$ , αλλά όχι αρκετά μεγάλες ώστε να υποδεικνύουν ανεξαρτησία στις τιμές των σφαλμάτων του υποδείγματος της παλινδρόμησης. Για παράδειγμα, για τη Σειρά II η τιμή του μέσου όρου της στατιστικής των Durbin-Watson είναι 0.836 για τιμές  $\phi = 0.9$  και  $a_1 = 0.9$  χρησιμοποιώντας δείγματα χρονοσειρών 1000 παρατηρήσεων σε αντίθεση με τη Σειρά I που είναι 0.231. Στην ουσία, η μη γραμμικότητα κάνει μια προσπάθεια να αφαιρέσει την αυτοσυσχέτιση στην ανάλυση παλινδρόμησης, αλλά όχι τόσο επιτυχημένη όσο στο να διορθώσει τις λανθασμένες ενδείξεις ψευδούς συμπεριφοράς.



### Πίνακας 5.5

Ποσοστά απόρριψης της μηδενικής υπόθεσης ότι  $\beta = 0$  σε επίπεδο  
σημαντικότητας 5% για δύο ανεξάρτητες AR(1) – ARCH(1) στα σφάλματα  
χρονοσειρές βασισμένα σε 10000 επαναλήψεις χρησιμοποιώντας τη μέθοδο των  
Cochrane-Orcutt

$a_1$	$\varphi = 0.0$	$\varphi = 0.2$	$\varphi = 0.5$	$\varphi = 0.8$	$\varphi = 0.9$	$\varphi = 1.0$
<b>0.0</b>	0.066	0.066	0.067	0.069	0.075	0.101
	0.057	0.057	0.058	0.060	0.062	0.074
	0.054	0.054	0.053	0.054	0.054	0.058
	0.052	0.052	0.052	0.052	0.052	0.053
<b>0.2</b>	0.063	0.063	0.064	0.070	0.074	0.103
	0.057	0.057	0.056	0.060	0.064	0.077
	0.054	0.054	0.056	0.054	0.055	0.061
	0.052	0.052	0.053	0.053	0.053	0.054
<b>0.5</b>	0.064	0.065	0.066	0.069	0.078	0.108
	0.057	0.057	0.058	0.060	0.064	0.082
	0.054	0.054	0.054	0.053	0.053	0.059
	0.054	0.054	0.054	0.054	0.054	0.055
<b>0.8</b>	0.068	0.068	0.068	0.076	0.084	0.117
	0.059	0.058	0.059	0.063	0.068	0.092
	0.052	0.052	0.053	0.053	0.053	0.062
	0.052	0.052	0.052	0.054	0.053	0.056
<b>0.9</b>	0.069	0.070	0.070	0.076	0.087	0.121
	0.062	0.061	0.062	0.066	0.070	0.097
	0.053	0.053	0.053	0.054	0.055	0.062
	0.050	0.051	0.052	0.053	0.053	0.056

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα 50, 100, 500 και 1000 παρατηρήσεων.

**Πίνακας 5.6**

**Ποσοστά απόρριψης της μηδενικής υπόθεσης ότι  $\beta = 0$  σε επίπεδο σημαντικότητας 5% για δύο ανεξάρτητες AR(1) – ARCH(1) στη μεταβλητή χρονοσειρές βασισμένα σε 10000 επαναλήψεις χρησιμοποιώντας τη μέθοδο των Cochrane-Orcutt**

$a_1$	$\varphi = 0.0$	$\varphi = 0.2$	$\varphi = 0.5$	$\varphi = 0.8$	$\varphi = 0.9$	$\varphi = 1.0$
<b>0.0</b>	0.066	0.066	0.067	0.069	0.075	0.101
	0.057	0.057	0.058	0.060	0.062	0.074
	0.054	0.054	0.053	0.054	0.054	0.058
	0.052	0.052	0.052	0.052	0.052	0.053
<b>0.2</b>	0.063	0.064	0.065	0.072	0.076	0.085
	0.057	0.058	0.058	0.062	0.062	0.068
	0.054	0.052	0.053	0.051	0.053	0.051
	0.052	0.051	0.052	0.048	0.049	0.044
<b>0.5</b>	0.064	0.065	0.070	0.074	0.078	0.084
	0.057	0.058	0.059	0.064	0.068	0.070
	0.054	0.054	0.050	0.046	0.047	0.045
	0.054	0.053	0.048	0.043	0.043	0.035
<b>0.8</b>	0.068	0.068	0.072	0.073	0.076	0.081
	0.059	0.059	0.063	0.062	0.064	0.067
	0.052	0.055	0.051	0.046	0.044	0.044
	0.052	0.050	0.046	0.041	0.037	0.033
<b>0.9</b>	0.069	0.070	0.073	0.073	0.077	0.080
	0.062	0.061	0.063	0.062	0.063	0.067
	0.053	0.053	0.051	0.045	0.047	0.046
	0.050	0.050	0.046	0.038	0.036	0.032

Σημείωση: Οι αριθμοί σε κάθε κελί αντιπροσωπεύουν δείγματα 50, 100, 500 και 1000 παρατηρήσεων.

Ως εκ τούτου, το επόμενο βήμα είναι η διερεύνηση της συμπεριφοράς του φαινομένου της υποφαινομενικής παλινδρόμησης υπό την παρουσία σειριακών συσχετισμένων σφαλμάτων. Για το σκοπό αυτό, το υπόδειγμα (5.6) εκτιμάται χρησιμοποιώντας τη μέθοδο των Cochrane-Orcutt (1949) για την αντιμετώπιση της αυτοσυσχέτισης, όπως έχει προτείνει ο McCallum (2010), τόσο για τις Σειρές Ι όσο και για τις Σειρές ΙΙ και τα ποσοστά απόρριψης της μηδενικής υπόθεσης ότι ο συντελεστής  $\beta$  του υποδείγματος (5.6) είναι μηδέν σε ονομαστικό επίπεδο σημαντικότητας 5% παρουσιάζονται αντίστοιχα στους Πίνακες 5.5 και 5.6, ένα θέμα το οποίο έχει επίσης αντιμετωπιστεί



από τον Kolen (2011), χρησιμοποιώντας τη Γενικευμένη Μέθοδο των ελαχίστων τετραγώνων GLS (Generalized Least Squares Method), και από τον Agiakloglou (2013), χρησιμοποιώντας εμπειρικές πρακτικές, δηλαδή εκτιμώντας το αρχικό υπόδειγμα με την παρουσία χρονικής υστέρησης της εξαρτημένης μεταβλητής.

Όπως φαίνεται και από τους δύο πίνακες αυτούς, η διαδικασία των Cochrane-Orcutt συμβάλλει καθοριστικά στην επίλυση της λανθασμένης συμπεριφοράς, καθώς τα εμπειρικά ποσοστά απόρριψης της μηδενικής υπόθεσης είναι όχι μόνο σημαντικά μικρότερα από αυτά που αναφέρονται στους Πίνακες 5.1 και 5.2, δηλαδή και για τα δύο είδη σειρών, αλλά τείνουν και ασυμπτωτικά. στο ονομαστικό επίπεδο 5% για όλες τις τιμές και των δύο παραμέτρων, ενώ η μορφή της ετεροσκεδαστικότητας δεν έχει καμία επίδραση στην περίπτωση αυτή.

## 5.5 Ανακεφαλαίωση

Ψευδείς σχέσεις εμφανίζονται στην ανάλυση της παλινδρόμησης όταν παραβιάζεται μία από τις βασικές υποθέσεις της μεθόδου των ελαχίστων τετραγώνων ή από λανθασμένο σχηματισμό ενός υποδείγματος. Κατά έναν όμως πολύ παράξενο τρόπο, ψευδείς σχέσεις εμφανίζονται και στην ανάλυση της παλινδρόμησης χρονοσειρών για μεταβλητές που δεν έχουν κανένα κοινό μεταξύ τους. Πραγματικά, το μοναδικό αυτό φαινόμενο που ονομάζεται υποφαινομενικές παλινδρομήσεις εμφανίζεται στην παλινδρόμηση όταν ανεξάρτητες στάσιμες ή μη στάσιμες διαδικασίες οδηγούν τον αναλυτή σε ψευδή συμπεράσματα αναφορικά με τη σχέση τους.

Από την πρωτοποριακή εργασία των Granger και Newbold (1974) μέχρι πρόσφατα, έχουν γίνει αρκετές προσπάθειες για την καλύτερη κατανόηση της συμπεριφοράς αυτού του φαινομένου, αλλά όχι σε ποιο βαθμό η ετεροσκεδαστικότητα επηρεάζει αυτή τη μοναδική συμπεριφορά. Χρησιμοποιώντας ανάλυση Monte Carlo στο κεφάλαιο



αυτό διερευνήθηκε η επίδραση της αυτοπαλίνδρομης υπό συνθήκης ετεροσκεδαστικότητας (ARCH) στην υποφαινομενική παλινδρόμηση για στάσιμες και μη στάσιμες διαδικασίες και διαπιστώθηκε ότι μόνο το ARCH(1) στη μεταβλητή υπόδειγμα μπορεί να εξουδετερώσει το μεγαλύτερο μέρος της παρατυπίας, αν και αυτός ο τύπος ετεροσκεδαστικότητας μπορεί να θεωρηθεί ως ένα μη γραμμικό υπόδειγμα ή ως ένα υπόδειγμα με διαχρονικά μεταβαλλόμενου συντελεστή, ένα εύρημα που θα έχει επιπτώσεις στην ανάλυση της μοναδιαίας ρίζας και της συνολοκλήρωσης. Αντίθετα, η παρουσία ετεροσκεδαστικότητας με τη μορφή ενός υποδείγματος ARCH(1) στα σφάλματα δεν έχει καμία επίδραση στη συμπεριφορά του φαινομένου, απλά και μόνο επειδή στην περίπτωση αυτή η παράμετρος της μη γραμμικότητας δεν επηρεάζει τη διακύμανση του συντελεστή  $\beta$ .

Τέλος, και οι δύο τύποι ετεροσκεδαστικότητας αποκάλυψαν στοιχεία αυτοσυσχέτισης στις τιμές των καταλοίπων του απλού γραμμικού υποδείγματος, μία συμπεριφορά που έχει επισημανθεί στη βιβλιογραφία, ενώ η εφαρμογή της διαδικασίας των Cochrane-Orcutt ήταν αρκετά ικανή να εξαλείψει και τα δύο προβλήματα, αυτό δηλαδή που αφορούσε το λανθασμένο συμπέρασμα και αυτό που αφορούσε το πρόβλημα της αυτοσυσχέτισης στα σφάλματα. Ενδεχομένως, να μπορεί να ειπωθεί ότι αυτή η παράξενη συμπεριφορά είναι πιθανότατα ένα φαινόμενο για γραμμικά παραγόμενες χρονοσειρές, κάτι που θα μπορούσε να μελετηθεί και για άλλες μορφές μη γραμμικών υποδειγμάτων που παρουσιάστηκαν στους Granger και Terasvirta (1993).



## ΚΕΦΑΛΑΙΟ 6

# ΕΜΠΕΙΡΙΚΗ ΕΦΑΡΜΟΓΗ ΣΕ ΜΕΤΡΑ ΕΞΑΡΤΗΣΗΣ ΓΙΑ ΜΕΤΑΒΛΗΤΕΣ ΠΟΥ ΣΥΝΔΕΟΝΤΑΙ ΜΕ ΤΗΝ ΑΓΟΡΑ ΕΝΕΡΓΕΙΑΣ

### 6.1 Εισαγωγή

Σε πολλές οικονομικές εφαρμογές είναι απαραίτητο να μοντελοποιηθεί τόσο η σειριακή εξάρτηση όσο και η εξάρτηση μεταξύ διαφορετικών χρονοσειρών. Αυτό μπορεί να γίνει άμεσα προτείνοντας ένα πλήρες παραμετρικό υπόδειγμα ή εξετάζοντας τη σειριακή εξάρτηση για κάθε σειρά και στη συνέχεια επιτυγχάνοντας την αλληλεξάρτηση μεταξύ των σειρών μέσω συζεύξεων, που είναι συναρτήσεις κοινής κατανομής τυχαίων μεταβλητών με ομοιόμορφα περιθώρια. Επιπρόσθετα, θα πρέπει να εξεταστεί και η συμπεριφορά των καταλοίπων του υποδείγματος, δεδομένου ότι οι καινοτομίες των τιμών τους δεν είναι εκ των προτέρων γνωστές, ενώ η διερεύνηση της συμπεριφοράς των τιμών τους δίνει άλλη διάσταση στην σχέση των μεταβλητών. Η έρευνα με τις τιμές των καταλοίπων περιπλέκει τη διαδικασία εξαγωγής συμπερασμάτων, καθώς η υπό συνθήκη κατανομή των στατιστικών μέτρων εξάρτησης εξαρτάται γενικά από άγνωστες παραμέτρους (βλέπε, για παράδειγμα, Bai 1994 και Ghoudi και Rémillard 2004). Ειδικά για την αγορά ενέργειας υπάρχουν ενδείξεις ότι κοινές κινήσεις των τιμών διαφόρων σημαντικών μεταβλητών μπορούν να επεξηγηθούν πιθανώς και από τον ταυτόχρονο αντίκτυπο της μεταβολής οικονομικών και πολιτικών δραστηριοτήτων σε διαφορετικές γεωγραφικές περιοχές.

Παραδοσιακά, σε ένα πρώτο βήμα στατιστικής ανάλυσης, εφαρμόζονται διάφορα υποδείγματα χρονοσειρών, μεμονωμένων εξισώσεων, για να προσδιοριστεί η αυτοσυσχέτιση των τιμών της χρονοσειράς και η ομαδοποίηση της μεταβλητότητας



στις οριακές κατανομές των τιμών της. Μάλιστα, όταν η μεταβλητότητα των τιμών του υποδείγματος είναι μεγάλη, τότε η προσπάθεια επικεντρώνεται στην μοντελοποίηση των οριακών κατανομών των τιμών του σφάλματος, όπου σύμφωνα με την βασική παραδοχή, οι τιμές τους θεωρείται ότι κατανέμονται ομοιόμορφα και ανεξάρτητα μεταξύ τους και ακολουθούν κανονική κατανομή (identically, independently and normally distributed). Ωστόσο, αυτή η παραδοχή αρκετά συχνά παραβιάζεται στην πράξη, όπου τα δεδομένα για διάφορους λόγους δεν ακολουθούν αυτήν την τυπική συμπεριφορά. Για παράδειγμα, ο Alexander (2005) μελέτησε την εξάρτηση μεταξύ των τιμών των συμβολαίων μελλοντικής εκπλήρωσης σε αργό πετρέλαιο και φυσικό αέριο και κατέληξε στο συμπέρασμα ότι όταν οι τιμές αυτών των εμπορευμάτων προσαρμόζονται σε υποδείγματα χρονοσειρών, η κατανομή των τιμών των σφαλμάτων τους δεν ήταν συμμετρική και για αυτό το λόγο χρησιμοποιήθηκαν πιο πολύπλοκες τεχνικές μέτρησης εξάρτησης.

Έτσι, η προσέγγιση της χρήσης τεχνικών *Copula* για τη μοντελοποίηση των σχέσεων σε μεταβλητές της αγοράς ενέργειας εφαρμόστηκε πολύ συχνά σε εμπειρικές αναλύσεις. Για παράδειγμα, οι Gregoire et al. (2008) ανέλυσαν τις τιμές του αργού πετρελαίου και του φυσικού αερίου με βάση τα συμβόλαια μελλοντικής εκπλήρωσης ενός μηνός που διαπραγματεύονται στο Χρηματιστήριο Εμπορευμάτων της Νέας Υόρκης (NYMEX) από την 1η Ιουλίου 2003 έως τις 19 Ιουλίου 2006 χρησιμοποιώντας τις τεχνικές αυτές. Πιο συγκεκριμένα, οι συγγραφείς αυτοί μοντελοποίησαν τις ημερήσιες αποδόσεις μεμονωμένα ως χρονοσειρές, επιλέγοντας στη συνέχεια το καταλληλότερο υπόδειγμα τύπου GARCH για τις τιμές των καταλοίπων και στη συνέχεια επέλεξαν την καλύτερη σύζευξη για να απεικονίσει την υπάρχουσα εξάρτηση μεταξύ των τιμών των σφαλμάτων. Επίσης, ο Reboredo (2011) χρησιμοποίησε ένα υπόδειγμα GARCH με βάση τη τεχνική *Copula* για να μελετήσει τη δομή εξάρτησης μεταξύ των τιμών αναφοράς του αργού πετρελαίου σε 12 διεθνείς αγορές αργού πετρελαίου όπου διαπιστώθηκε ότι σε περιόδους πίεσης της αγοράς, οι τιμές του αργού πετρελαίου σε κάθε αγορά τείνουν να παρουσιάζουν κινήσεις με την ίδια ένταση, τονίζοντας ότι το αργό πετρέλαιο και το φυσικό αέριο παρουσιάζουν μη γραμμική εξάρτηση.





Επιπρόσθετα, οι Wen et al. (2012) ερευνήσαν τις πιθανές επιδράσεις μεταξύ τιμών πετρελαίου και κινεζικών χρηματιστηρίων χρησιμοποιώντας μια χρονικά μεταβαλλόμενη προσέγγιση *Copula* και έδειξαν ότι οι οικονομικά ασταθείς περίοδοι, όπως η περίοδος της αποτυχίας της Lehman Brothers, χαρακτηρίζονται από αύξηση της εξάρτησης μεταξύ των μεταβλητών. Επίσης, οι Aloui et al. (2013) ανέλυσαν τις ημερήσιες τιμές κλεισίματος για το δείκτη αργού πετρελαίου Brent με τους δείκτες MSCI (Morgan Stanley Capital International) των έξι μεταβατικών οικονομιών της Κεντρικής Ανατολικής Ευρώπης (Βουλγαρία, Τσεχία, Ουγγαρία, Πολωνία, Ρουμανία και Σλοβενία) κατά την περίοδο από 1η Δεκεμβρίου 2005 έως και τις 20 Αυγούστου 2012 και διαπίστωσαν ότι η δομή εξάρτησης μεταξύ των αποδόσεων των τιμών του πετρελαίου και των χρηματιστηριακών αποδόσεων περιγράφεται καλύτερα με διαχρονικά μεταβαλλόμενη τεχνική *Copula*, όπως και οι Cherubini et al. (2012).

Για τη μέτρηση της εξάρτησης ακραίων γεγονότων στις αγορές ενέργειας οι Jaschke et al. (2012) διαπίστωσαν ότι η εξέταση των δεικτών εξάρτησης ουράς μπορεί να οδηγήσει εσφαλμένα σε στοιχεία ασύμμετρης σχέσης μεταξύ των αποδόσεων και πρότειναν τη χρήση τοπικής μέτρησης με βάση την προσέγγιση *tail-Copula* (βλέπε επίσης Schmidt and Stadtmuller, 2006). Η έρευνα τους επικεντρώθηκε στην εξάρτηση μεταξύ των τιμών του αργού πετρελαίου και του φυσικού αερίου με βάση τα συμβόλαια μελλοντικής εκπλήρωσης ενός μήνα που διαπραγματεύονται στη NYMEX μεταξύ 2 Ιουλίου 2007 και 2 Ιουλίου 2010. Τέλος, οι Beckmann et al. (2015) ανέλυσαν τις τιμές του πετρελαίου και τις συναλλαγματικές ισοτιμίες έναντι του δολαρίου όπου και οι δύο χρονοσειρές παρουσίασαν μεγάλες διακυμάνσεις τις τελευταίες δεκαετίες και διαπίστωσαν ότι η προσέγγιση *Copula* έδινε καλύτερες μετρήσεις της σχέσης εξάρτησης των μεταβλητών αυτών. Μάλιστα, λαμβάνοντας ως διαχωριστικό σημείο τη συντριβή της Lehman Brothers το 2008, κατέληξαν στο συμπέρασμα ότι για ημερήσια δεδομένα για δύο πενταετείς περιόδους μεταξύ Σεπτεμβρίου 2003 και 2013 η αυξανόμενη εξάρτηση της ουράς δείχνει ότι τα ακραία γεγονότα είναι πιο πιθανό να συμβούν ταυτόχρονα και για τις δύο σειρές.



Κατά συνέπεια, στο κεφάλαιο αυτό γίνεται μία προσπάθεια να μελετηθούν εμπειρικά οι σχέσεις εξάρτησης μεταξύ διαφόρων μεταβλητών που συνδέονται με την αγορά ενέργειας χρησιμοποιώντας τις τεχνικές *Copula* που προσδιορίζουν καλύτερα τη σχέση εξάρτησης των μεταβλητών σε διάφορα στάδια μελέτης των τιμών των χρονοσειρών. Αρχικά γίνεται μία σύντομη παρουσίαση των μεταβλητών που θα χρησιμοποιηθούν στην ανάλυση και διερευνώνται οι σχέσεις εξάρτησης των μεταβλητών στις τιμές τους. Στη συνέχεια, οι σχέσεις αυτές εξετάστηκαν σε τιμές πρώτων διαφορών, έχοντας δηλαδή μετατρέψει σε στάσιμες διαδικασίες. Κατόπιν, η ανάλυση επεκτάθηκε στις τιμές των καταλοίπων που προέκυψαν από την εκτίμηση του καταλληλότερου επιλεγθέν ARIMA υποδείγματος, όπως, και μετά από τη απομόνωση της ετεροσκεδαστικότητας.

## 6.2 Παρουσίαση μεταβλητών

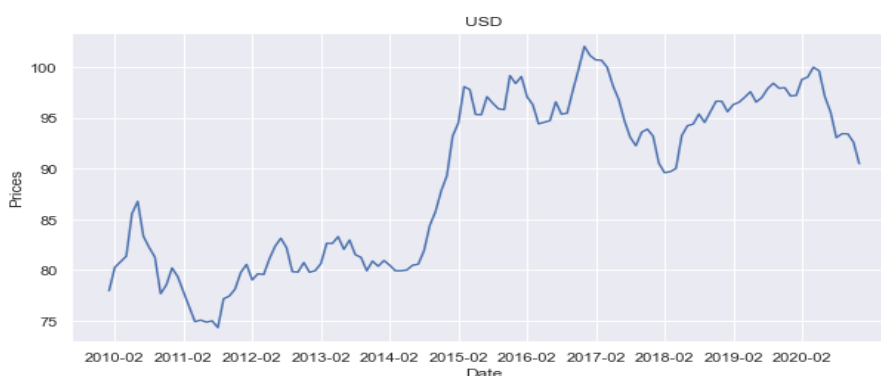
Η αγορά ενέργειας (Energy Market) είναι μία αγορά, με την ευρύτερη έννοια του όρου, που απευθύνεται σε προϊόντα ηλεκτρικής ενέργειας από την παραγωγή μέχρι την κατανάλωση και συνδέεται με χρηματοοικονομικές, καθώς και οικονομικές μεταβλητές, δεδομένου ότι η κατανάλωση ενέργειας αποτελεί βασικό συστατικό οικονομικής ανάπτυξης. Για το λόγο αυτό, όπως αναφέρθηκε και στην προηγούμενη παράγραφο, πολλές μελέτες προσπάθησαν να συνδέσουν τη συμπεριφορά βασικών ενεργειακών αγαθών με άλλες μεταβλητές. Έτσι, το φυσικό αέριο, ο άνθρακας και κυρίως το πετρέλαιο αποτέλεσαν σημεία αναφοράς για τη διερεύνηση της συμπεριφοράς της τιμής τους σε σχέση με τιμές άλλων βασικών και πολύτιμων αγαθών, καθώς και με τιμές χρηματιστηριακών δεικτών.

Για την έρευνα αυτή χρησιμοποιούνται μεταβλητές από πέντε κατηγορίες προκειμένου να διερευνηθεί η συμπεριφορά των διαφόρων μέτρων εξάρτησης σε σχέση με την τιμή του πετρελαίου έναντι όλων των άλλων μεταβλητών και όχι να προσδιοριστούν ποσοτικά οι σχέσεις εξάρτησης μεταξύ των μεταβλητών, αναφέροντας ότι τα δεδομένα

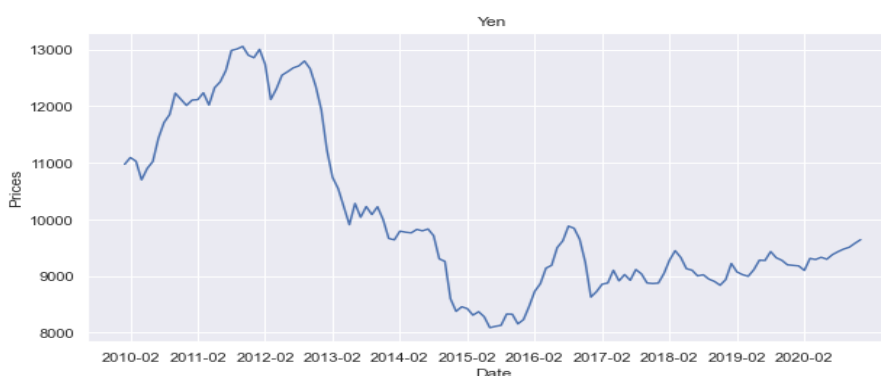


αντλήθηκαν από το Yahoo Finance και αποτελούνται από μηνιαίες παρατηρήσεις για τη χρονική περίοδο από 01/01/2010 έως 31/12/2020, έχοντας συνολικά 132 παρατηρήσεις.

Ειδικότερα, για την κατηγορία Νομίσματα (Currencies) χρησιμοποιούνται: (a) η τιμή του δείκτη που μετράει την αξία του δολαρίου Αμερικής USD (ICE U.S. Dollar Index futures contract) σχέση με ένα καλάθι από άλλα διεθνή νομίσματα και (b) η τιμή του δείκτη του γιέν Ιαπωνίας Yen (Dow Jones FXCM Yen) που μετράει την αγοραστική δύναμη του γιέν σε σχέση με οκτώ άλλα νομίσματα, η διαγραμματική απεικόνιση των οποίων δίνεται στο Διάγραμμα 6.1 (a) και (b) αντίστοιχα. Από το διάγραμμα αυτό παρατηρείται μία αντίστροφη διαχρονική πορεία. Το δολάριο Αμερικής δείχνει να κερδίζει συναλλαγματική αξία, ενώ το γιέν Ιαπωνίας δείχνει να χάνει την αξία του.



(a)



(b)

**Διάγραμμα 6.1**

**Διαχρονική Εξέλιξη Τιμής Νομισμάτων USD και YEN**



Για την κατηγορία Χρηματιστηριακοί Δείκτες (Stock Market Indices) χρησιμοποιούνται: (a) ο δείκτης S&P 500 (S&P 500 Index), (b) ο Nikkei 225 (Nikkei 225 Index) από το Χρηματιστήριο της Νέας Υόρκης και (c) ο δείκτης VIX (CBOE Volatility Index) από το Χρηματιστήριο παραγώγων του Σικάγου. Ο S&P 500 και ο Nikkei 225 είναι δύο από τους πιο σημαντικούς δείκτες στον κόσμο, ενώ ο δείκτης μεταβλητότητας (Volatility Index – VIX) δημιουργήθηκε το 1993 από τον Robert Whaley με σκοπό να απεικονίσει τη μεταβλητότητα καθώς και το ρίσκο που υφίσταται στις χρηματιστηριακές αγορές και ο υπολογισμός του πραγματοποιείται από το Χρηματιστήριο παράγωγων του Σικάγο. Από το Διάγραμμα 6.2 (a), (b) και (c) στο οποίο παρουσιάζονται η διαχρονική εξέλιξη των τιμών όλων των ανωτέρω χρηματιστηριακών δεικτών αντίστοιχα παρατηρείται μία σχετικά ομοιόμορφη εξέλιξη των τιμών των χρηματιστηριακών δεικτών εκτός από τη εξέλιξη του δείκτη VIX, ο οποίος φαίνεται να έχει μία δική του ξεχωριστή πορεία.

Για την κατηγορία Ελαφρά Αγαθά (Soft Commodities) χρησιμοποιούνται οι τιμές των μελλοντικών συμβολαίων που αναφέρονται σε ένα μήνα μπροστά για τα αγαθά: (a) Σιτάρι – Wheat (Chicago SRW Wheat Futures), (b) Ρύζι – Rice (Rough Rice Futures) και (c) Βαμβάκι – Cotton (Cotton ICE Futures), η διαγραμματική παρουσίαση των τιμών των οποίων δίνεται αντίστοιχα στο Διάγραμμα 6.3 (a), (b) και (c). Από το διάγραμμα αυτό παρατηρείται ότι η τιμή του Σιταριού – Wheat έχει μία πτώση από το αρχικό υψηλό επίπεδο που είχε στην αρχή της περιόδου μελέτης, το Ρύζι – Rice έχει μεταβλητότητα αλλά κρατάει το μέσο όρο του και το Βαμβάκι – Cotton έχει στο μεγαλύτερο μέρος της περιόδου μελέτης χαμηλή μεταβλητότητα γύρω από μία μέση τιμή με εξαίρεση την αρχική του περίοδο όπου εκεί έχει μία μεγάλη αύξηση.

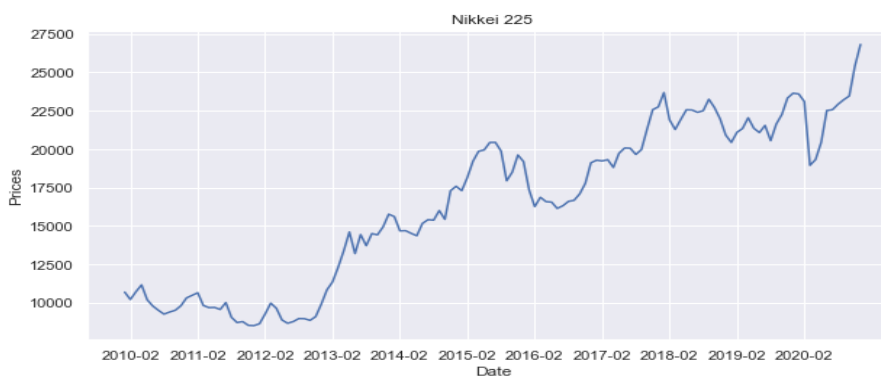
Για την κατηγορία Βασικοί Δείκτες (Basic Indices) χρησιμοποιούνται οι τιμές των δεικτών: (a) Ενέργειας – Energy Index (Euronext Rogers International Energy Index) και (b) Αγαθών – Commodity Index (Euronext Rogers International Commodity Index), η διαγραμματική απεικόνιση των οποίων εμφανίζεται στο Διάγραμμα 6.4 (a) και (b) αντίστοιχα, από όπου παρατηρείται μία σχεδόν ολοίδια και ομοιόμορφη



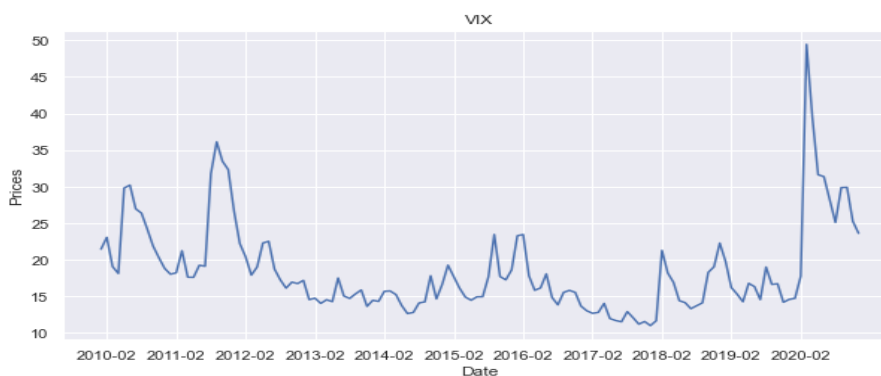
συμπεριφορά στους δείκτες αυτούς, όπου σε κάποιο χρονικό σημείο, κοντά στη μέση της περιόδου μελέτης, εμφανίζεται μία σημαντική και απότομη πτώση στις τιμές τους.



(a)



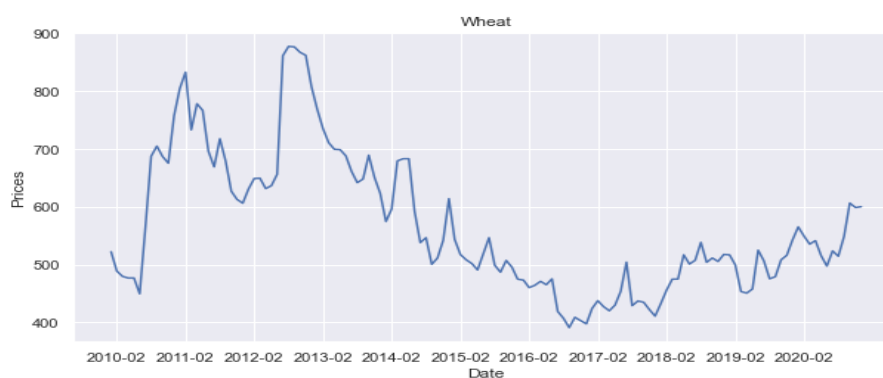
(b)



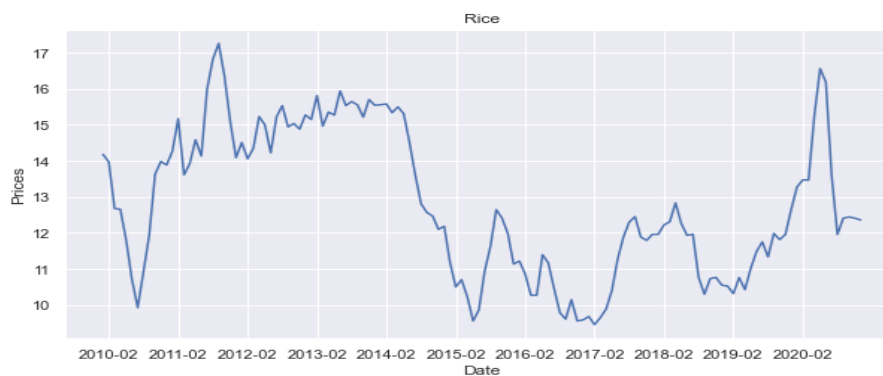
(c)

**Διάγραμμα 6.2**

**Διαχρονική Εξέλιξη Χρηματιστηριακών Δεικτών S&P 500, Nikkei 225 και VIX**



(a)



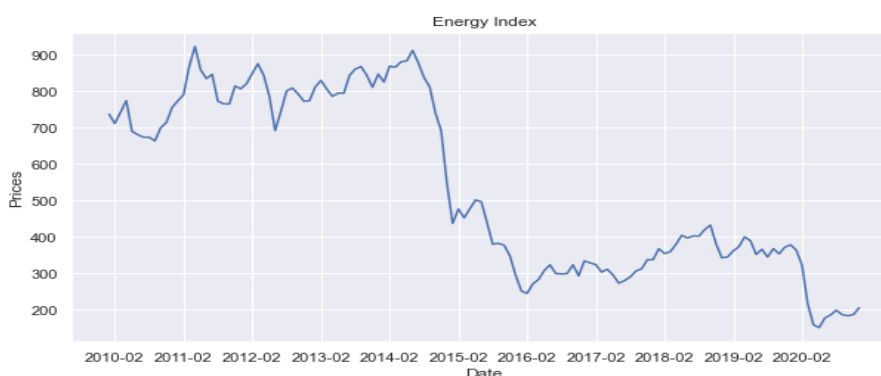
(b)



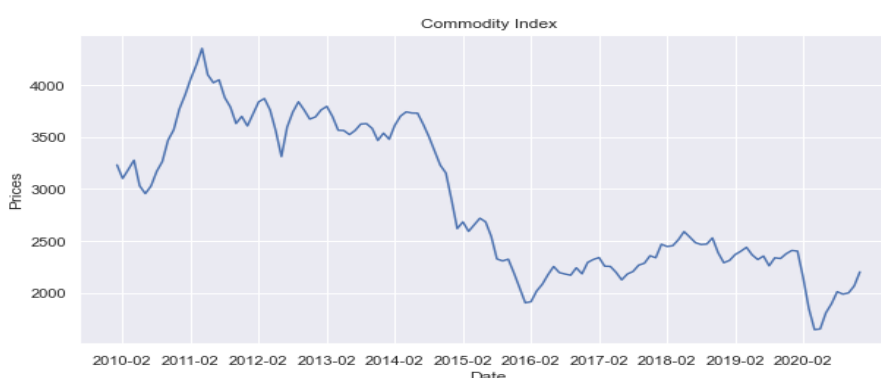
(c)

**Διάγραμμα 6.3**

**Διαχρονική Εξέλιξη Τιμής Ελαφρών Αγαθών Wheat, Rice και Cotton**



(a)

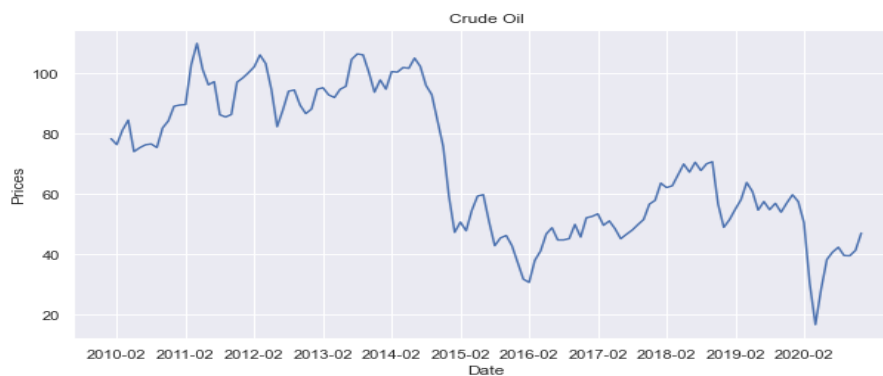


(b)

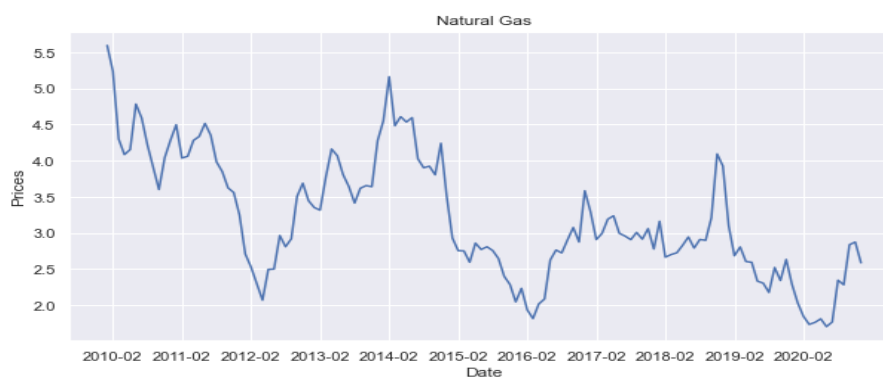
#### Διάγραμμα 6.4

##### Διαχρονική Εξέλιξη Τιμής Ενεργειακού Δείκτη και Δείκτη Αγαθών

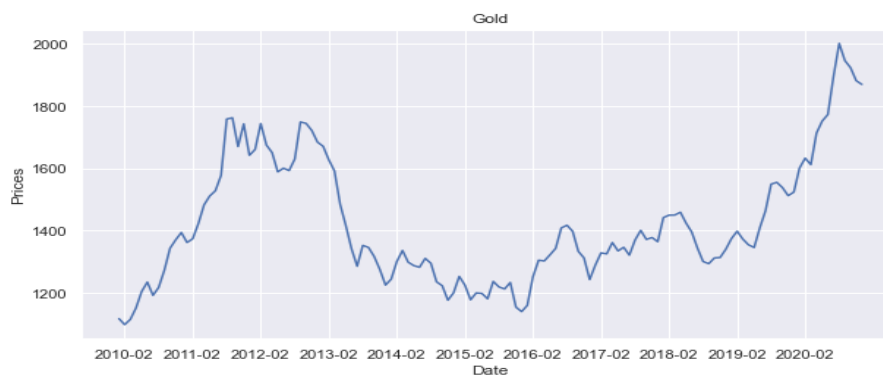
Για την τελευταία κατηγορία Βαρέα Αγαθά (Hard Commodities) χρησιμοποιούνται οι τιμές των μελλοντικών συμβολαίων που αναφέρονται σε ένα μήνα μπροστά για τα αγαθά: (a) Πετρέλαιο – Crude Oil (Crude Oil WTI Futures), (b) Φυσικό Αέριο – Natural Gas (Natural Gas - NYMEX (Henry Hub) Natural Gas Futures) και (c ) Χρυσός – Gold (Gold COMEX Futures), η διαγραμματική απεικόνιση των οποίων εμφανίζεται αντίστοιχα στο Διάγραμμα 6.5 (a), (b) και (c). Από τα διαγράμματα αυτά παρατηρείται μία σχεδόν παρόμοια πτωτική συμπεριφορά των τιμών του πετρελαίου και του φυσικού αερίου, με το πρώτο να έχει μία πιο απότομη πτώση, ενώ αντίθετα οι τιμές του χρυσού να έχουν μία ανοδική πορεία.



(a)



(b)



(c)

### Διάγραμμα 6.5

**Διαχρονική Εξέλιξη Τιμής Βαρέων Αγαθών Crude Oil, Natural Gas και Gold**





Πίνακας 6.1

Περιγραφικά στατιστικά των μεταβλητών

Variable	# of Obs	Mean	Stand Dev	Min	Max	Skewness	Kurtosis
<i>USD</i>	132	89.00	8.24	74.35	102.01	-0.18	1.44
<i>Yen</i>	132	10,033.36	1,455.81	8,090.98	13,058.33	0.81	2.30
<i>S&amp;P 500</i>	132	2,073.49	676.41	1,072.52	3,692.35	0.30	2.08
<i>Nikkei 225</i>	132	16,434.27	5,143.10	8,501.36	26,820.00	-0.17	1.69
<i>VIX</i>	132	18.58	6.23	11.02	49.43	1.89	7.67
<i>Crude Oil</i>	132	69.46	23.13	16.70	110.04	0.07	1.73
<i>Natural Gas</i>	132	3.20	0.85	1.70	5.60	0.39	2.50
<i>Gold</i>	132	1,421.62	202.18	1,098.01	2,001.59	0.81	2.93
<i>Wheat</i>	132	566.48	119.63	390.98	876.92	0.81	2.88
<i>Rice</i>	132	12.81	2.05	9.45	17.26	0.17	1.84
<i>Cotton</i>	132	80.70	25.34	53.75	203.33	2.80	11.84
<i>Energy Index</i>	132	541.63	242.63	151.61	923.78	0.11	1.38
<i>Commodity Index</i>	132	2,876.11	709.98	1,646.52	4,350.52	0.24	1.59

Τέλος, στον Πίνακα 6.1 παρουσιάζονται τα βασικά περιγραφικά στατιστικά μέτρα όλων των μεταβλητών που χρησιμοποιούνται στην ανάλυση. Από τον πίνακα αυτόν παρατηρείται ότι μόνο δύο μεταβλητές αντιμετωπίζουν πρόβλημα με τη κατανομή τους δεδομένου ότι οι τιμές των δειγματικών συντελεστών ασυμμετρίας και κύρτωσης για το δείκτη VIX και για το αγαθό Βαμβάκι δεν φαίνεται να υποστηρίζουν κανονικότητα, κάτι που διαπιστώνεται και από τα αντίστοιχα διαγράμματά τους, αφού ο δείκτης στην αρχή της περιόδου και το Βαμβάκι στο τέλος της περιόδου μελέτης εμφανίζουν απότομες ακραίες μεταβολές. Αντίθετα, όλες οι άλλες μεταβλητές υποστηρίζουν χαρακτηριστικά συμπεριφοράς κανονικής κατανομής.



### 6.3 Μέτρηση Εξάρτησης

Οι μεταβλητές που παρουσιάστηκαν στην προηγούμενη ενότητα χρησιμοποιούνται στη συνέχεια για να διερευνηθεί η σχέση εξάρτησής τους με όλα τα μέτρα εξάρτησης που αναλύθηκαν σε προηγούμενα κεφάλαια αλλά και σε διάφορες μορφές και πτυχές των τιμών τους. Αρχικά, θα εξεταστούν οι σχέσεις των μεταβλητών στις αρχικές τιμές τους και στη συνέχεια σε πρώτες διαφορές, αφού εξεταστεί πρωτύτερα η στασιμότητα των τιμών τους. Κατόπιν, η εξέταση των σχέσεων τους θα επεκταθεί σε τιμές καταλοίπων που θα προκύψουν από την εκτίμηση του καταλληλότερου υποδείγματος χρονοσειρών που θα προσδιοριστεί και θα ληφθεί υπ' όψιν και η περίπτωση ετεροσκεδαστικότητας στις τιμές τους.

Σημειώνεται ότι η ανάλυση που θα ακολουθηθεί θα έχει ως σημείο αναφοράς το πετρέλαιο, ένα πολύ σημαντικό αγαθό για την αγορά ενέργειας αλλά και για την οικονομική λειτουργία των χωρών, δεδομένου ότι αποτελεί βασική πρώτη ύλη στην πετροχημική βιομηχανία. Επιπρόσθετα, στο Παράρτημα Α θα δίνονται και όλες οι άλλες εκτιμήσεις των σχέσεων εξάρτησης με όλες τις άλλες μεταβλητές, έτσι ώστε να εξασφαλίζεται μεγαλύτερος αριθμός εκτιμήσεων των μέτρων εξάρτησης για την καλύτερη κατανόησή της συμπεριφοράς τους.

#### *Μέτρηση εξάρτησης σε τιμές των μεταβλητών*

Στον Πίνακα 6.2 δίνονται οι τιμές του συντελεστή συσχέτισης και όλων των άλλων μέτρων εξάρτησης που παρουσιάστηκαν σε προηγούμενα κεφάλαια για τις αρχικές τιμές των μεταβλητών σε σχέση με τις τιμές του πετρελαίου, όπου όλες οι εκτιμήσεις πραγματοποιήθηκαν στην Python. Όπως παρατηρείται από τον πίνακα αυτόν, ο συντελεστής συσχέτισης, που εκφράζει βαθμό, ύπαρξη και κατεύθυνση γραμμικής σχέσης μεταξύ δύο μεταβλητών, παρέχει μία πρώτη εικόνα των σχέσεων των μεταβλητών αυτών.



Πιο συγκεκριμένα, το πετρέλαιο με τα νομίσματα συνδέεται πολύ έντονα και αρνητικά και ειδικότερα με το δολάριο -0.90, ενδεχομένως δεδομένου ότι όλες οι συναλλαγές του γίνονται σε αυτό το νόμισμα, και λιγότερο έντονα αλλά θετικά με το γιέν 0.69. Επίσης, το πετρέλαιο συνδέεται αρνητικά με όλους του χρηματιστηριακούς δείκτες και σε μεγάλο βαθμό με τον S&P 500 και Nikkei 225 (-0.71), ενώ με πολύ μικρό βαθμό με τον VIX (-0.16). Θετική και έντονη σχέση παρατηρείται του πετρελαίου με τα ελαφρά αγαθά σε βαθμό κατά μέσο όρο 0.65, ενώ πολύ έντονη θετική σχέση εμφανίζεται με τον ενεργειακό δείκτη και με το δείκτη αγαθών (0.96). Τέλος, το πετρέλαιο συνδέεται θετικά και έντονα με το φυσικό αέριο (0.66), ενώ με το χρυσό διακρίνεται μία πάρα πολύ ασθενής θετική σχέση (0.03). Για τη σημαντικότητα όλων των ανωτέρων ευρημάτων εφαρμόστηκε ο στατιστικός έλεγχος για τη μηδενική συσχέτιση, που παρουσιάστηκε σε προηγούμενο κεφάλαιο, και βρέθηκε ότι όλες οι σχέσεις είναι στατιστικά σημαντικές, εκτός από τη σχέση του πετρελαίου με το δείκτη VIX και το χρυσό, όπως διαπιστώνεται από τις τιμές της στατιστικής  $t$  που δίνονται στον Πίνακα 6.2, κάτι που άλλωστε για αυτές τις δύο μεταβλητές ήταν αναμενόμενο αφού η τιμή του συντελεστή συσχέτισης είναι πολύ μικρή.

## Πίνακας 6.2

### Τιμές μέτρων εξάρτησης σε τιμές των μεταβλητών σε σχέση με το Πετρέλαιο

Crude Oil with	Correlation Coefficient $r$	$t$ -statistic	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
<i>USD</i>	-0.90	-22.97	-0.81	-0.59	-0.74	-0.74	0.00	-7.46	1.00
<i>Yen</i>	0.69	10.78	0.67	0.43	0.60	0.60	0.81	4.81	1.59
<i>S&amp;P 500</i>	-0.67	-10.30	-0.64	-0.34	-0.56	-0.56	0.00	-4.24	1.00
<i>Nikkei 225</i>	-0.71	-11.61	-0.64	-0.36	-0.57	-0.57	0.00	-4.28	1.00
<i>VIX</i>	-0.16	-1.80	-0.09	-0.06	-0.16	-0.12	0.00	-0.53	1.00
<i>Natural Gas</i>	0.66	10.00	0.64	0.46	0.64	0.66	1.45	5.18	1.71
<i>Gold</i>	0.03	0.32	0.06	0.05	0.00	0.08	0.03	0.40	1.01
<i>Wheat</i>	0.69	10.93	0.66	0.43	0.59	0.59	0.72	4.67	1.57
<i>Rice</i>	0.65	9.86	0.62	0.43	0.52	0.59	0.66	4.93	1.62
<i>Cotton</i>	0.56	7.77	0.75	0.55	0.76	0.76	1.84	6.50	1.95
<i>Energy Index</i>	0.96	41.46	0.96	0.85	0.95	0.95	4.03	21.87	5.02
<i>Commodity Index</i>	0.96	37.83	0.94	0.80	0.91	0.92	5.84	16.30	3.14



Τη συμπεριφορά του συντελεστή συσχέτισης, ως προς τον καθορισμό και την κατεύθυνση των σχέσεων, την επιβεβαιώνουν οι τιμές των συντελεστών του Spearman και του Kendall, καθώς και οι τιμές των *Copulas* Gaussian και Student-*t*, οι οποίες όπως έχει αναφερθεί συνδέονται με την τιμή του συντελεστή συσχέτισης. Αντίθετα, από τις εναπομείναντες *Copulas*, η Frank είναι αυτή που ξεχωρίζει σε σχέση με τις άλλες δύο Αρχιμήδειες *Copulas*, τη Clayton και τη Gumbel, διότι δίνει πολύ καλή εικόνα για το τι συμβαίνει στη σχέση του πετρελαίου με τις άλλες μεταβλητές. Καταρχάς, η *Copula* Frank έχει πρόσημο που προσδιορίζει την κατεύθυνση των σχέσεων και επαληθεύει τα ευρήματα του συντελεστή συσχέτισης. Κατά δεύτερον, έχει μέγεθος που παραθέτει ξεκάθαρα τη σειρά της έντασης των σχέσεων, δηλώνοντας ότι η πιο έντονη σχέση του πετρελαίου είναι με τον ενεργειακό δείκτη (θετική 21.87), μετά με το δείκτη αγαθών (θετική 16.30) και μετά με το δολάριο (αρνητική -7.46), ενώ η χαμηλότερη σε ένταση σχέση που αγγίζει το επίπεδο ανεξαρτησίας είναι με το χρυσό (θετική 0.40) και μετά με το δείκτη VIX (αρνητική -0.53).

Αντίθετα, η Clayton και η Gumbel Αρχιμήδειες *Copulas* δείχνουν να έχουν μία σχεδόν παρόμοια συμπεριφορά ως προς τον προσδιορισμό των σχέσεων των μεταβλητών αυτών με το πετρέλαιο χωρίς να μπορούν ωστόσο να προσδιορίζουν την κατεύθυνση της σχέσης των μεταβλητών, με τη Clayton να επιλέγει πρώτη τη σχέση του πετρελαίου με τον δείκτη αγαθών (5.84) και δεύτερη τη σχέση του με τον ενεργειακό δείκτη (4.03), ενώ η Gumbel επιλέγει πρώτη τη σχέση του με το ενεργειακό δείκτη (5.02) και δεύτερη τη σχέση του με το δείκτη αγαθών (3.14). Αξιοσημείωτο πάντως είναι το γεγονός ότι αυτές οι δύο *Copulas* θεωρούν ότι το πετρέλαιο δεν συνδέεται καθόλου με το δολάριο και με του χρηματιστηριακούς δείκτες, αφού οι τιμές τους είναι μηδέν και ένα για την Clayton και την Gumbel αντίστοιχα, φανερώνοντας σημάδια ανεξαρτησίας ακόμα και με το χρυσό.

Τέλος, να αναφερθεί ότι στο Παράρτημα Α δίνονται στον Πίνακα Α.1 όλες οι τιμές των μέτρων εξάρτησης όλων των άλλων διμεταβλητών σχέσεων χωρίς της συμμετοχής του



πετρελαίου. Από τον πίνακα αυτόν μπορούν να εξαχθούν παρόμοια και ανάλογα συμπεράσματα για την ένταση και την κατεύθυνση των σχέσεων.

### ***Μέτρηση εξάρτησης σε τιμές πρώτων διαφορών των μεταβλητών***

Στη συνέχεια, οι μεταβλητές της μελέτης εξετάζονται ως προς τη στασιμότητά τους προκειμένου να διαπιστωθεί εάν οι τιμές τους ακολουθούν στάσιμη διαδικασία. Για το λόγο αυτό, εφαρμόστηκε ο επαυξημένος έλεγχος μοναδιαίας αυτοσυσχετιζόμενης ρίζας των Dickey-Fuller, γνωστός ως ADF test, που παρουσιάστηκε σε προηγούμενο κεφάλαιο, και τα αποτελέσματα αυτού του ελέγχου δίνονται στον Πίνακα 6.3. Πιο συγκεκριμένα, στο πίνακα αυτόν δίνονται οι τιμές της στατιστικής  $t$ , καθώς και οι δειγματικές τιμές της πιθανότητας απόρριψης της μηδενικής υπόθεσης ( $p$ -values) του ελέγχου που προέκυψαν από την εκτίμηση του ADF υποδείγματος με τάση μαζί με τον αριθμό των χρονικών υστερήσεων που χρησιμοποιήθηκαν για την εκτίμηση του υποδείγματος, η επιλογή των οποίων βασίστηκε στην ελαχιστοποίηση της τιμής του πληροφοριακού κριτηρίου AIC, όπως πρότειναν οι Agiakloglou και Newbold (1992). Τα αποτελέσματα αυτού του ελέγχου δηλώνουν ότι όλες οι σειρές των τιμών των μεταβλητών είναι μη στάσιμες, δεδομένου ότι η μηδενική υπόθεση της ύπαρξης μοναδιαίας αυτοσυσχετιζόμενης ρίζας γίνεται δεκτή σε επίπεδο σημαντικότητας 5% για όλες τις μεταβλητές εκτός από το δείκτη VIX που γίνεται δεκτή σε επίπεδο 1%.<sup>11</sup>

---

<sup>11</sup> Οι κριτικές τιμές που χρησιμοποιήθηκαν για τον έλεγχο αυτόν είναι -3.45 και -4.05 για επίπεδο σημαντικότητας 5% και 1% αντίστοιχα και δίνονται στο βιβλίο του Fuller (1996). Επίσης, ο έλεγχος αυτός εφαρμόστηκε και χωρίς τάση στις τιμές των μεταβλητών και τα αποτελέσματα ως προς τη στασιμότητα των χρονοσειρών παρέμειναν τα ίδια.



**Πίνακας 6.3**

**Αποτελέσματα εφαρμογής του ADF test με τάση**

Metric	USD	Yen	S&P 500	Nikkei 225	VIX	Crude Oil	Natural Gas	Gold	Wheat	Rice	Cotton	Energy Index	Com/dity Index
<i>ADF Test-Stat</i>	-2.33	-1.72	-3.37	-3.08	-3.88	-2.50	-3.37	-1.45	-2.38	-2.82	-3.32	-2.39	-2.84
<i>ADF p-Value</i>	0.416	0.741	0.055	0.112	0.013	0.326	0.055	0.847	0.393	0.191	0.063	0.385	0.183
<i>Lag Order</i>	4	2	1	1	1	2	1	5	1	1	1	1	1

Κατά συνέπεια, η ανάλυση των σχέσεων των μεταβλητών, έχοντας διαπιστώσει ότι οι τιμές τους προέρχονται από μη στάσιμες διαδικασίες, εφαρμόστηκε στις πρώτες διαφορές των τιμών τους, μία μετατροπή, η οποία εξασφάλισε τη στασιμότητα των τιμών τους, η οποία επιβεβαιώθηκε και από την εφαρμογή του ADF ελέγχου στις τιμές των πρώτων διαφορών των μεταβλητών. Έτσι, όλα τα προαναφερθέντα μέτρα μέτρησης της εξάρτησης εφαρμόστηκαν στις τιμές των πρώτων διαφορών των μεταβλητών σε σχέση με το πετρέλαιο και δίνονται στον Πίνακα 6.4.

**Πίνακας 6.4**

**Τιμές μέτρων εξάρτησης σε πρώτες διαφορές των τιμών των μεταβλητών σε σχέση με το Πετρέλαιο**

Crude Oil	Correlation Coefficient <i>r</i>	<i>t</i> -statistic	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- <i>t</i>	Clayton	Frank	Gumbel
<i>USD</i>	-0.34	-4.11	-0.28	-0.18	-0.34	-0.34	0.00	-1.76	1.00
<i>Yen</i>	-0.12	-1.35	-0.15	-0.10	-0.14	-0.15	0.00	-0.92	1.00
<i>S&amp;P 500</i>	0.51	6.67	0.44	0.31	0.48	0.48	0.82	3.19	1.44
<i>Nikkei 225</i>	0.41	5.05	0.38	0.27	0.40	0.42	0.60	2.66	1.37
<i>VIX</i>	-0.47	-5.97	-0.45	-0.32	-0.45	-0.49	0.00	-3.25	1.00
<i>Natural Gas</i>	0.05	0.56	0.07	0.04	0.07	0.07	0.07	0.39	1.04
<i>Gold</i>	-0.02	-0.25	0.05	0.04	0.03	0.03	0.00	0.33	1.04
<i>Wheat</i>	-0.05	-0.53	-0.06	-0.05	-0.10	-0.08	0.00	-0.40	1.00
<i>Rice</i>	-0.06	-0.72	-0.03	-0.02	-0.06	-0.03	0.00	-0.18	1.01
<i>Cotton</i>	0.30	3.53	0.33	0.23	0.36	0.36	0.49	2.08	1.25
<i>Energy Index</i>	0.93	28.28	0.93	0.79	0.92	0.94	5.24	16.92	74.27
<i>Commodity Index</i>	0.83	16.74	0.81	0.63	0.81	0.82	2.81	8.78	2.25



Από το πίνακα αυτόν παρατηρείται καταρχάς ότι αρκετές σχέσεις των μεταβλητών που προέκυψαν από το συντελεστή συσχέτισης άλλαξαν σε σχέση με την πρώτη ποσοτική προσέγγιση στις τιμές των μεταβλητών, είτε ως προς την κατεύθυνση είτε ως προς την ένταση και τη σημαντικότητα της σχέσης τους, ενώ ορισμένες και πολύ λίγες παρέμειναν όπως ήταν. Έτσι, η σχέση του πετρελαίου με το γιέν που ήταν θετική και σημαντική (0.69) έγινε αρνητική και μη σημαντική (-0.12), ενώ αντίθετα η σχέση του με το δολάριο παρέμεινε αρνητική και σημαντική αλλά πολύ λιγότερο έντονη, δηλαδή από -0.90 έγινε -0.34. Ομοίως, η σχέση του πετρελαίου με τους χρηματιστηριακούς δείκτες S&P 500 και Nikkei 225 παρέμεινε στατιστικά σημαντική αλλά με διαφορετικό πρόσημο και μικρότερο μέγεθος, δηλαδή από -0.67 και -0.71 σε 0.51 και 0.41 αντίστοιχα, ενώ η σχέση του πετρελαίου με το δείκτη VIX παρέμεινε αρνητική αλλά έγινε πολύ πιο έντονη και στατιστικά σημαντική, δηλαδή από -0.16 έγινε -0.47. Τα πράγματα με το φυσικό αέριο επίσης άλλαξαν σημαντικά αφού σε πρώτες διαφορές στις τιμές τους φαίνεται να μην υπάρχει σχέση με το πετρέλαιο, δεδομένου ότι η αρκετά έντονη και σημαντικά θετική τους σχέση έγινε πολύ ασθενής και μη σημαντική παραμένοντας όμως θετική, δηλαδή από 0.66 έγινε 0.05, ενώ με το χρυσό ότι ίσχυε ισχύει, δηλαδή η μεταξύ τους σχέση που δεν υπήρχε εξακολουθεί να μην υφίσταται και σε πρώτες διαφορές των τιμών τους και απλά έγινε αρνητική από θετική με μέγεθος πάρα πολύ μικρό, δηλαδή από 0.03 σε -0.02. Αλλαγές στο πρόσημο, μέγεθος και κατά συνέπεια στη σημαντικότητα παρατηρούνται στα ελαφρά αγαθά σιτάρι και ρύζι όπου οι θετικά έντονες και στατιστικά σημαντικές σχέσεις τους με το πετρέλαιο γίνονται στις πρώτες διαφορές των τιμών τους αρνητικές πολύ ασθενείς και ασήμαντες, δηλαδή από 0.69 και 0.65 γίνονται -0.05 και -0.06 αντίστοιχα, ενώ με το βαμβάκι το πετρέλαιο διατήρησε τα χαρακτηριστικά της αρχικής τους σχέσης που ήταν θετική και σημαντική αλλά έχασε σε μέγεθος και έγινε πιο ασθενής, δηλαδή από 0.56 έγινε 0.30.

Αυτό όμως το οποίο είναι ίσως το πιο αξιοσημείωτο εύρημα αυτής της προσπάθειας και πρέπει έντονα να καταγραφεί είναι ότι το πετρέλαιο διατηρεί την αρχική του έντονη θετική και σημαντική σχέση με τον ενεργειακό δείκτη και με το δείκτη αγαθών που ήταν 0.96 και η οποία δεν άλλαξε σχεδόν καθόλου και έγινε αντίστοιχα 0.93 και 0.83. Έτσι, οι σχέσεις του πετρελαίου με αυτά τα δύο αγαθά παραμένουν ακλόνητες και στις



πρώτες διαφορές των τιμών τους, κάτι που δεν συμβαίνει με όλα τα άλλα αγαθά αφού οι σχέσεις τους άλλαξαν σημαντικά.

Τις αλλαγές στη συμπεριφορά του συντελεστή συσχέτισης των μεταβλητών με το πετρέλαιο στις πρώτες διαφορές των τιμών τους τις ακολουθούν πιστά και οι συντελεστές του Spearman και του Kendall, καθώς και οι Gaussian και Student-t *Copulas*, οι οποίες όπως έχει αναφερθεί συνδέονται με την τιμή του συντελεστή συσχέτισης. Παρόμοια, συμπεριφορά ως προς τις αλλαγές στην ένταση της σχέσης των μεταβλητών σε πρώτες διαφορές εμφανίζεται σε όλες τις άλλες Αρχιμήδειες *Copulas* με την Frank να δηλώνει ακόμα και την αλλαγή στην κατεύθυνση της σχέσης σε ορισμένες από αυτές. Ωστόσο, και οι τρεις αυτές *Copulas* δηλώνουν ταυτόχρονα ότι η πιο έντονη σχέση εξάρτησης του πετρελαίου είναι με τον ενεργειακό δείκτη και μετά με το δείκτη αγαθών, κάτι που επιβεβαιώνεται και από όλα τα άλλα μέτρα εξάρτησης. Αξίζει πάντως να σημειωθεί ότι οι Clayton και Gumbel *Copulas* δηλώνουν ότι περισσότερες σχέσεις του πετρελαίου με τις άλλες μεταβλητές δεν υφίστανται όταν η σχέση εξάρτησής τους μετρηθεί στις πρώτες διαφορές των τιμών τους από ότι στις ίδιες τις τιμές τους, δεδομένου ότι στον Πίνακα 6.4 εμφανίζονται περισσότερες τιμές κοντά ή στην τιμή τους που δηλώνει ανεξαρτησία που είναι το μηδέν και το ένα αντίστοιχα, από ότι στον Πίνακα 6.2.

Τέλος, στον Πίνακα A.2 του Παραρτήματος Α δίνονται όλες οι τιμές των μέτρων εξάρτησης όλων των άλλων διμεταβλητών σχέσεων σε τιμές πρώτων διαφορών των μεταβλητών. Από τον πίνακα αυτόν διαπιστώνονται παρόμοια συμπεράσματα για τις σημαντικές αλλαγές που προκύπτουν στη διερεύνηση των σχέσεων των μεταβλητών σε πρώτες διαφορές των τιμών τους από ότι στις αρχικές τιμές τους τόσο ως προς την ένταση όσο και ως προς την κατεύθυνση της σχέσης τους.





### ***Μέτρηση εξάρτησης σε τιμές των μεταβλητών και καταλοίπων που προκύπτουν από την εκτίμηση επιλεγμένου ARIMA υποδείγματος***

Στη συνέχεια η διερεύνηση των σχέσεων των μεταβλητών επεκτείνεται στις τιμές των μεταβλητών καθώς και των καταλοίπων που θα προκύψουν από την εκτίμηση του καταλληλότερου ARIMA υποδείγματος, μία προσπάθεια που αποσκοπεί να διερευνηθεί η συμπεριφορά των σχέσεων των μεταβλητών με βάση του τρόπου εξέλιξης των τιμών τους αλλά και με βάση όλων των άλλων παραγόντων που επηρεάζουν κατά έμμεσο τρόπο την εξέλιξη των τιμών τους και κατά επέκταση τις σχέσεις τους που εμφανίζονται στις τιμές των καταλοίπων. Για το λόγο αυτό, εφαρμόστηκε η μεθοδολογία των Box και Jenkins (1976) στις πρώτες διαφορές των τιμών των μεταβλητών, δεδομένου ότι οι τιμές τους βρέθηκαν να ακολουθούν μη στάσιμη διαδικασία, και εκτιμήθηκαν με τη μέθοδο μεγίστης πιθανοφάνειας (MLE)  $ARIMA(p, 1, q)$  υποδείγματα για όλους τους συνδυασμούς  $p + q \leq 5$  και στον Πίνακα 6.5 παρουσιάζονται τα τρία πρώτα υποδείγματα που επιλέχθηκαν με σειρά επιλογής με βάση την ελαχιστοποίηση των τιμών των πληροφοριακών κριτηρίων AIC και BIC.



## Πίνακας 6.5

Παρουσίαση των πρώτων τριών υποδειγμάτων που επιλέχθηκαν για κάθε μεταβλητή με βάση τα πληροφοριακά κριτήρια AIC και BIC

Variable	AIC			BIC		
	1st	2nd	3rd	1st	2nd	3rd
USD	(1,1,0)	(0,1,1)	(2,1,0)	(1,1,0)	(0,1,1)	(2,1,0)
Yen	(2,1,0)	(1,1,0)	(0,1,2)	(2,1,0)	(1,1,0)	(0,1,2)
S&P 500	(0,1,1)	(1,1,0)	(2,1,2)	(0,1,1)	(1,1,0)	(2,1,0)
Nikkei 225	(0,1,1)	(1,1,0)	(2,1,0)	(0,1,1)	(1,1,0)	(2,1,0)
VIX	(1,1,1)	(1,1,2)	(2,1,0)	(1,1,1)	(0,1,2)	(1,1,2)
Crude Oil	(0,1,1)	(2,1,0)	(3,1,0)	(0,1,1)	(1,1,0)	(2,1,0)
Natural Gas	(2,1,2)	(3,1,2)	(3,1,0)	(1,1,0)	(0,1,1)	(2,1,2)
Gold	(0,1,4)	(1,1,0)	(5,1,0)	(0,1,1)	(1,1,0)	(1,1,1)
Wheat	(0,1,1)	(2,1,0)	(1,1,0)	(0,1,1)	(1,1,0)	(2,1,0)
Rice	(0,1,1)	(1,1,0)	(2,1,0)	(0,1,1)	(1,1,0)	(2,1,0)
Cotton	(1,1,0)	(5,1,0)	(2,1,0)	(1,1,0)	(2,1,0)	(1,1,1)
Energy Index	(1,1,0)	(0,1,1)	(0,1,2)	(1,1,0)	(0,1,1)	(0,1,2)
Com/dity Index	(1,1,0)	(0,1,2)	(2,1,0)	(1,1,0)	(0,1,1)	(0,1,2)

Από τον πίνακα αυτόν παρατηρείται ότι το υπόδειγμα  $ARIMA(1, 1, 0)$  είναι το καταλληλότερο (best fitted model) για όλες τις μεταβλητές και επιλέγεται. Καταρχάς, το υπόδειγμα αυτό επιλέγεται ως πρώτο σε προτίμηση με βάση το κριτήριο BIC σε πέντε μεταβλητές και δεύτερο στις υπόλοιπες επτά μεταβλητές, με πρώτο το υπόδειγμα  $ARIMA(0, 1, 1)$  που είναι πολύ κοντά στο υπόδειγμα  $ARIMA(1, 1, 0)$  και με πολύ μικρή διαφορά στην τιμή του κριτηρίου. Το υπόδειγμα  $ARIMA(1, 1, 0)$  επιλέγεται και ως πέμπτο με το κριτήριο BIC για το δείκτη VIX, παρόλο που δεν εμφανίζεται στον πίνακα αυτόν, ενώ η παρουσία του ως το καταλληλότερο υπόδειγμα για όλες τις τιμές των μεταβλητών υποστηρίζεται και από το κριτήριο AIC, αν και το κριτήριο αυτό έχει την τάση να επιλέγει μεγάλα σε αριθμό παραμέτρων υποδείγματα. Κατά δεύτερον, η επιλογή του υποδείγματος αυτού είναι προτιμητέα όχι διότι είναι ένα απλό υπόδειγμα, αλλά διότι συμβαδίζει και με το θεωρητικό μέρος της ανάλυσης που εφαρμόστηκε για τη μελέτη της συμπεριφοράς των μέτρων εξάρτησης σε προηγούμενα κεφάλαια.



Κατά συνέπεια, όλες οι τιμές των μεταβλητών εκτιμήθηκαν σε πρώτες διαφορές με ένα αυτοπαλίνδρομο πρώτου βαθμού υπόδειγμα και οι τιμές του συντελεστή του αυτοπαλίνδρομου όρου δίνονται στον Πίνακα 6.6. Από τον πίνακα αυτόν παρατηρείται ότι οι τιμές του συντελεστή κατά κανόνα είναι μικρές και κυμαίνονται από 0.09 στο φυσικό αέριο μέχρι 0.37 στο δείκτη αγαθών, δεδομένου ότι πολύ συχνά αυτές οι χρονοσειρές ομοιάζουν περισσότερο με διαδικασία τυχαίου περπατήματος, όπως το φυσικό αέριο και ο δείκτης VIX. Όλες οι τιμές του συντελεστή είναι θετικές, με εξαίρεση την τιμή του δείκτη VIX, όπου το υπόδειγμα αυτό δεν επιλέγεται μέσα στα πρώτα τρία καλύτερα και από τα δύο κριτήρια, όπως φαίνεται και από τον Πίνακα 6.5, ενώ οι τιμές όλων είναι στατιστικά σημαντικές κατά κανόνα σε επίπεδο 5% με δύο εξαιρέσεις σε επίπεδο σημαντικότητας 10%, ενώ οι τιμές του φυσικού αερίου και του δείκτη VIX δεν είναι στατιστικά σημαντικές.

**Πίνακας 6.6**

**Εκτιμήσεις του αυτοπαλίνδρομου όρου του υποδείγματος ARIMA(1, 1, 0) για τις τιμές όλων των μεταβλητών**

Metric	USD	Yen	S&P 500	Nikkei 225	VIX	Crude Oil	Natural Gas	Gold	Wheat	Rice	Cotton	Energy Index	Com/dity Index
<i>AR(1) Coefficient</i>	0.28	0.31	0.11	0.12	-0.04	0.30	0.09	0.23	0.14	0.25	0.43	0.31	0.37
<i>t-Statistic</i>	2.78	4.82	1.95	1.59	-0.78	4.55	1.04	2.91	1.78	3.26	6.72	5.25	5.20

Έτσι, οι εκτιμηθείσες τιμές των μεταβλητών από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα χρησιμοποιήθηκαν για να προσδιοριστούν οι τιμές των μέτρων εξάρτησής των που δίνονται στον Πίνακα 6.7. Από τον πίνακα αυτόν παρατηρείται ότι σε γενικές γραμμές οι εκτιμήσεις όλων των μέτρων εξάρτησης ομοιάζουν κατά πολύ με αυτές που δίνονται στον Πίνακα 6.2, δηλαδή σε αυτές που προέκυψαν με βάση τις αρχές τιμές των μεταβλητών σε σχέση με το πετρέλαιο, σημειώνοντας μάλιστα ότι αυτές οι τιμές των μεταβλητών εκφράζουν την καλύτερη δυνατή δειγματική συμπεριφορά τους για τη χρονική περίοδο μελέτης. Πραγματικά, οι τιμές του συντελεστή συσχέτισης ομοιάζουν σε πολύ μεγάλο βαθμό τόσο ως προς την κατεύθυνση όσο και ως προς το



μέγεθος της σχέσης όλων των μεταβλητών σε σχέση με το πετρέλαιο, καθώς και ως προς τη στατιστική σημαντικότητά τους. Από όλες τις σχέσεις του πετρελαίου ξεχωρίζει η σχέση του με το δολάριο η οποία μειώνεται σε απόλυτο μέγεθος από -0.96 σε -0.46, οι σχέσεις του με τον ενεργειακό δείκτη και το δείκτη αγαθών που παραμένουν πολύ ισχυρές σε επίπεδο 0.96 και 0.95 αντίστοιχα, καθώς και οι σχέσεις του με το δείκτη VIX και το χρυσό που εξακολουθούν να μην υφίστανται. Επομένως, το πετρέλαιο συνδέεται πολύ έντονα με τον ενεργειακό δείκτη και το δείκτη αγαθών και όχι τόσο με το δολάριο, όπως στα αρχικά δεδομένα.

### Πίνακας 6.7

Τιμές μέτρων εξάρτησης σε εκτιμηθείσες τιμές μεταβλητών από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα σε σχέση με το Πετρέλαιο

Crude Oil	Correlation Coefficient $r$	t-statistic	Spearman $\rho$	Kendall $\tau$	Gaussian	Student-t	Clayton	Frank	Gumbel
USD	-0.46	-5.84	-0.76	-0.55	-0.66	-0.72	0.00	-6.96	1.00
Yen	0.69	10.85	0.68	0.45	0.63	0.63	1.00	5.03	1.63
S&P 500	-0.55	-7.54	-0.60	-0.32	-0.50	-0.52	0.00	-4.01	1.00
Nikkei 225	-0.59	-8.25	-0.61	-0.35	-0.50	-0.53	0.00	-4.12	1.00
VIX	-0.09	-1.01	-0.05	-0.04	-0.09	-0.07	0.02	-0.34	1.00
Natural Gas	0.67	10.40	0.63	0.45	0.64	0.66	1.49	5.10	1.70
Gold	0.17	1.91	0.10	0.07	0.09	0.17	0.20	0.69	1.06
Wheat	0.71	11.37	0.67	0.44	0.62	0.62	0.85	4.74	1.59
Rice	0.66	10.03	0.62	0.43	0.54	0.60	0.81	4.92	1.62
Cotton	0.58	8.10	0.72	0.53	0.73	0.73	1.74	6.13	1.87
Energy Index	0.96	40.60	0.96	0.85	0.96	0.96	4.75	22.59	5.30
Commodity Index	0.95	36.68	0.94	0.80	0.92	0.92	6.13	16.65	3.24

Τη συμπεριφορά του συντελεστή συσχέτισης των μεταβλητών με το πετρέλαιο στις εκτιμηθείσες τιμές των μεταβλητών από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα την ακολουθούν πιστά, και μάλιστα με μικρότερες μεταβολές στα μεγέθη των τιμών τους από ότι ο συντελεστής συσχέτισης, και οι συντελεστές του Spearman και του Kendall, καθώς και οι Gaussian και Student-t *Copulas*, οι οποίες όπως έχει αναφερθεί συνδέονται με την τιμή του συντελεστή συσχέτισης. Όλα όμως τα ανωτέρω μέτρα εξάρτησης δηλώνουν κατηγορηματικά, όπως και ο συντελεστής συσχέτισης, ότι η



ισχυρότερη σχέση του πετρελαίου είναι με τον ενεργειακό δείκτη. Το συμπέρασμα αυτό το επιβεβαιώνουν και οι Frank και Gumbel *Copulas*, ενώ η Clayton, όπως και στις αρχικές τιμές των μεταβλητών θεωρεί ότι η σχέση του πετρελαίου με το δείκτη αγαθών είναι η πιο έντονη. Αντίθετα, οι Clayton και Gumbel *Copulas* εξακολουθούν να θεωρούν αρκετές σχέσεις του πετρελαίου με άλλες μεταβλητές ότι δεν υφίστανται.

Τέλος στον Πίνακα Α.3 του Παραρτήματος Α δίνονται οι τιμές των μέτρων εξάρτησης όλων των άλλων διμεταβλητών σχέσεων που προέκυψαν από τις εκτιμηθείσες τιμές των μεταβλητών από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα από όπου παρατηρείται και εκεί ότι αυτές οι σχέσεις ομοιάζουν αρκετά με εκείνες που δίνονται στον Πίνακα Α.1 του Παραρτήματος Α για τις τιμές των μεταβλητών.

### Πίνακας 6.8

**Τιμές μέτρων εξάρτησης σε τιμές καταλοίπων των μεταβλητών από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα σε σχέση με το Πετρέλαιο**

Crude Oil	Correlation Coefficient <i>r</i>	<i>t</i> -statistic	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- <i>t</i>	Clayton	Frank	Gumbel
USD	0.74	12.54	-0.24	-0.16	-0.23	-0.25	0.00	-1.49	1.01
Yen	0.75	12.88	-0.16	-0.10	-0.14	-0.17	0.00	-0.99	1.02
S&P 500	0.79	14.50	0.42	0.29	0.47	0.45	0.76	3.01	1.43
Nikkei 225	0.76	13.18	0.39	0.28	0.43	0.43	0.61	2.80	1.42
VIX	0.06	0.71	-0.37	-0.26	-0.37	-0.41	0.00	-2.64	1.00
Natural Gas	0.69	10.87	0.10	0.07	0.13	0.12	0.07	0.60	1.11
Gold	0.70	11.17	0.03	0.02	0.04	0.03	0.00	0.17	1.06
Wheat	0.58	8.06	-0.07	-0.05	-0.05	-0.08	0.00	-0.47	1.03
Rice	0.69	10.86	0.02	0.01	0.04	0.02	0.00	0.12	1.08
Cotton	0.62	8.96	0.18	0.12	0.24	0.23	0.25	1.05	1.15
Energy Index	0.96	40.23	0.96	0.82	0.95	0.96	5.47	20.13	4.66
Commodity Index	0.91	25.69	0.84	0.64	0.85	0.85	2.46	9.12	2.51

Η ανάλυση των σχέσεων εξάρτησης των μεταβλητών επεκτείνεται και στις τιμές των καταλοίπων που ανέκυψαν από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα. Στόχος αυτής της προσπάθειας είναι να διερευνηθούν οι έμμεσες σχέσεις εξάρτησης των μεταβλητών που προκύπτουν από το τυχαίο μέρος και οφείλονται σε όλους τους άλλους παράγοντες



που επηρεάζουν την εξέλιξη των τιμών τους. Από τον Πίνακα 6.8 παρατηρείται ότι πολλά πράγματα έχουν αλλάξει ριζικά. Οι σχέσεις των τιμών των καταλοίπων του πετρελαίου με τις τιμές των καταλοίπων όλων των άλλων μεταβλητών εμφανίζονται θετικές και κατά κανόνα ισχυρές, αφού όλες είναι μεγαλύτερες από την τιμή 0.58, με εντονότερη πάλι τη σχέση του πετρελαίου με τον ενεργειακό δείκτη, κατά πρώτον, και κατά δεύτερον με το δείκτη αγαθών, όπου εδώ η έντονη σχέση τους παραμένει και στο τυχαίο μέρος τους. Εξαίρεση αποτελεί η σχέση των τιμών των καταλοίπων του πετρελαίου με τις τιμές των καταλοίπων του δείκτη VIX όπου παραμένει πολύ μικρή, 0.06, δηλώνοντας ότι δεν υπάρχει σχέση μεταξύ τους ούτε και στο τυχαίο μέρος τους, ενώ για πρώτη φορά εμφανίζεται να υπάρχει σχέση με το χρυσό, έστω και με έμμεσο τρόπο.

Για όλες τις τιμές των μέτρων εξάρτησης που συνδέονται με το συντελεστή συσχέτισης, και ειδικότερα για τις τιμές των συντελεστών του Spearman και του Kendall, καθώς και για τις τιμές των Gaussian και Student-t *Copulas*, παρατηρούνται διαφορετικές ενδείξεις για το μέγεθος αλλά και για την κατεύθυνση των σχέσεων τους. Πιο συγκεκριμένα, το μέγεθος όλων των ανωτέρω μέτρων εξάρτησης σε τιμές καταλοίπων είναι πολύ μικρότερο από αυτό του συντελεστή συσχέτισης σε απόλυτο μέγεθος με εξαίρεση της σχέσης των τιμών των καταλοίπων του πετρελαίου με τις τιμές των καταλοίπων του ενεργειακού δείκτη και του δείκτη αγαθών, όπου παραμένει η σχέση πολύ ισχυρή και θετική με τον ενεργειακό δείκτη να κατέχει τα πρωτεία. Επιπλέον παρατηρείται αλλαγή στο πρόσημο των σχέσεων για τέσσερις μεταβλητές που είναι το δολάριο, το γιέν, ο δείκτης VIX και το σιτάρι, μία κατεύθυνση σχέσεων που επαληθεύεται και από την τιμή της Frank.

Τα ευρήματα των σχέσεων των τιμών των καταλοίπων του πετρελαίου με τις τιμές των καταλοίπων όλων των άλλων μεταβλητών είναι πιο ξεκάθαρα στα μέτρα εξάρτησης που δίνονται από τις Αρχιμήδειες *Copulas*, οι οποίες κατά κανόνα δηλώνουν ανεξαρτησία σε όλες σχεδόν τις σχέσεις αυτές, κυρίως από τις Clayton και Gumbel *Copulas* και λιγότερο από την Frank. Πάντως και οι τρεις αυτές *Copulas* θεωρούν πολύ έντονη τη σχέση του τυχαίου μέρους του πετρελαίου με το τυχαίο μέρος του



ενεργειακού δείκτη και δεύτερη την αντίστοιχη με τον δείκτη αγαθών, υπογραμμίζοντας για άλλη μία φορά τη μοναδική και πολύ έντονη σχέση τους, ακόμα και στη συμπεριφορά του τυχαίου μέρους τους, ενώ τη σχέση του με το φυσικό αέριο δεν την υπολογίζουν σοβαρά ούτε και στο τυχαίο μέρος τους.

Τέλος, ανάλογα συμπεράσματα ανακύπτουν και για όλες τις άλλες σχέσεις των τιμών των καταλοίπων όλων των άλλων μεταβλητών που ανακύπτουν από την εκτίμηση του επιλεχθέντος  $ARIMA(1, 1, 0)$  υποδείγματος και δίνονται στο Πίνακα Α.4 του Παραρτήματος Α.

### ***Μέτρηση Εξάρτησης σε τιμές των καταλοίπων από υπόδειγμα $ARCH(1)$ των μεταβλητών***

Η διερεύνηση των σχέσεων των μεταβλητών που αφορά το τυχαίο μέρος τους συνεχίζεται και στις τιμές των καταλοίπων με τη διερεύνηση της ύπαρξης ετεροσκεδαστικότητας. Ως γνωστόν, οι μεταβλητές αυτές παρουσιάζουν μία σχετικά μεγάλη ευαισθησία σε κάθε αιφνίδια κατάσταση ή νέα πληροφόρηση που θα προκύψει και η οποία εκδηλώνεται με έντονη μεταβλητότητα στις τιμές τους, μεταβάλλοντας τη σταθερότητα της διακύμανσής τους, κάτι το οποίο έχει ήδη διαπιστωθεί στις τιμές τους, δεδομένου ότι οι σειρές αυτές δεν είναι στάσιμες. Για το λόγο αυτό, οι τιμές των καταλοίπων που προέκυψαν από την εκτίμηση του  $ARIMA(1, 1, 0)$  υποδείγματος χρησιμοποιούνται για να προσδιοριστεί όχι μόνο η ύπαρξη αλλά και ο τρόπος δημιουργίας της μεταβαλλόμενης διαχρονικά διακύμανσής τους.

Κατά συνέπεια, εκτιμήθηκαν με τη μέθοδο μεγίστης πιθανοφάνειας υποδείγματα  $GARCH(h, m)$  για όλους του συνδυασμούς  $h + m \leq 3$ , από όπου τα πληροφορικά κριτήρια επέλεξαν ως το καλύτερο υπόδειγμα το  $ARCH(1)$ , συμπεριφορά του οποίου εμφανίζεται αρκετά συχνά σε μηνιαία δεδομένα. Ωστόσο, οι εκτιμήσεις του συντελεστή του υποδείγματος δεν ήταν στατιστικά σημαντικές και για το λόγο αυτό εφαρμόστηκε και ο στατιστικός έλεγχος  $LM$  (Lagrange Multiplier test).



Πιο συγκεκριμένα, οι τιμές των αρχικών καταλοίπων  $\hat{\varepsilon}_t$  από το ARIMA(1, 1, 0) υπόδειγμα υψώθηκαν στο τετράγωνο και εκτιμήθηκε το ακόλουθο υπόδειγμα:

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + u_t \quad (6.1)$$

με τις τιμές του τυχαίου σφάλματος  $u$  να ικανοποιούν όλες τις βασικές υποθέσεις, προκειμένου να ελεγχθεί η μηδενική υπόθεση ότι η τιμή του συντελεστή  $\alpha_1$  είναι μηδέν έναντι της εναλλακτικής ότι δεν είναι μηδέν, δηλαδή:

$$H_0: \alpha_1 = 0$$

$$H_1: \alpha_1 \neq 0$$

Ο έλεγχος εφαρμόζεται με τη στατιστική  $LM$ , η τιμή της οποίας υπολογίζεται ως εξής:

$$LM = n R^2$$

όπου  $n$  ο αριθμός των παρατηρήσεων που χρησιμοποιήθηκαν για την εκτίμηση του υποδείγματος (6.1) και  $R^2$  ο συντελεστής προσδιορισμού αυτού του υποδείγματος, και η στατιστική  $LM$  ακολουθεί τη  $\chi^2$  κατανομή με ένα βαθμό ελευθερίας. Έτσι, αν

$$LM < \chi_{1,\alpha}^2$$

τότε η μηδενική υπόθεση γίνεται δεκτή και κατά συνέπεια δεν υπάρχει πρόβλημα ετεροσκεδαστικότητας.

## Πίνακας 6.9

**Τιμές της στατιστικής  $LM$  για τον έλεγχο ετεροσκεδαστικότητας ενός ARCH(1) υποδείγματος**

Metric	USD	Yen	S&P 500	Nikkei 225	VIX	Crude Oil	Natural Gas	Gold	Wheat	Rice	Cotton	Energy Index	Com/dity Index
<i>LM-Test</i>	1.28	0.34	0.00	0.05	0.29	0.19	0.05	0.23	0.00	0.21	0.04	0.02	0.33
<i>p-Value</i>	0.258	0.558	0.962	0.816	0.592	0.665	0.823	0.632	0.981	0.645	0.839	0.888	0.565

Στον Πίνακα 6.9 δίνονται οι τιμές της στατιστικής  $LM$ , καθώς και οι δειγματικές τιμές της πιθανότητας απόρριψης της μηδενικής υπόθεσης (p-values) από την εφαρμογή του





ελέγχου, σημειώνοντας επιπρόσθετα ότι η κριτική τιμή της  $X^2$  κατανομής για ένα βαθμό ελευθερίας και επίπεδο σημαντικότητας 5% είναι ίση με 3.84. Από τον πίνακα αυτόν επιβεβαιώνονται οι προαναφερθέντες παρατηρήσεις ότι δεν υπάρχουν ενδείξεις ετεροσκεδαστικότητας, αφού η μηδενική υπόθεση γίνεται δεκτή για όλες τις σειρές των τιμών των μεταβλητών.

Κατά συνέπεια, το εκτιμηθέν ARIMA(1, 1, 0) υπόδειγμα που επιλέχθηκε ως το καταλληλότερο υπόδειγμα για όλες σχεδόν τις τιμές των μεταβλητών δεν παρουσιάζει ενδείξεις ετεροσκεδαστικότητας στις τιμές των καταλοίπων του. Μάλιστα, εφαρμόστηκε επιπρόσθετα και ο έλεγχος των Ljung και Box (1978) στις τιμές των καταλοίπων προκειμένου να διερευνηθεί εάν αυτές προέρχονται από ένα πληθυσμό που έχει τα χαρακτηριστικά του λευκού θορύβου (white noise), ένα από τα οποία είναι και η σταθερή διακύμανση, και έτσι το υπόδειγμα είναι σωστά ορισμένο. Από την εφαρμογή του ελέγχου διαπιστώθηκε ότι σε όλες τις μεταβλητές η μηδενική υπόθεση έγινε δεκτή δηλώνοντας ότι το υπόδειγμα είναι σωστά ορισμένο, με εξαίρεση το δείκτη VIX, όπου, όπως έχει αναφερθεί, θα πρέπει οι τιμές του να ακολουθούν διαδικασία τυχαίου περπατήματος. Κατά συνέπεια, έστω και με έμμεσο τρόπο επαληθεύεται ότι δεν υπάρχουν ενδείξεις ετεροσκεδαστικότητας στις τιμές των καταλοίπων όλων των μεταβλητών.

## 6.4 Ανακεφαλαίωση

Στο κεφάλαιο αυτό έγινε μία προσπάθεια εμπειρικής διερεύνησης της συμπεριφοράς όλων των μέτρων εξάρτησης που παρουσιάστηκαν και μελετήθηκαν σε θεωρητικό επίπεδο σε προηγούμενα κεφάλαια χρησιμοποιώντας μηνιαία δεδομένα από μεταβλητές μεγάλου μακροοικονομικού ενδιαφέροντος και οι οποίες συνδέονται με την αγορά ενέργειας έχοντας ως σημείο αναφοράς την τιμή του πετρελαιο. Οι σχέσεις των μεταβλητών διερευνήθηκαν σε διάφορα στάδια προκειμένου να διαπιστωθεί η συνέπεια των εκτιμήσεών τους. Έτσι, οι σχέσεις των μεταβλητών μελετήθηκαν αρχικά



στις αρχικές τιμές τους, κατόπιν στις τιμές των πρώτων διαφορών τους, έχοντας διαπιστώσει ότι οι τιμές τους προέρχονται από μη στάσιμες διαδικασίες, και στη συνέχεια στις εκτιμηθείσες τιμές τους που προέκυψαν από την εκτίμηση του καταλληλότερου ARIMA υποδείγματος, όπως επίσης και στις τιμές των καταλοίπων αυτού του υποδείγματος. Σημειώνεται ότι η έρευνα της συμπεριφοράς των μέτρων εξάρτησης δεν συνεχίστηκε δεδομένου ότι δεν βρέθηκαν στοιχεία μεταβαλλόμενης διακύμανσης.

Από την εφαρμογή όλων αυτών των διαφορετικών μέτρων εξάρτησης σε πραγματικά δεδομένα διαφόρων μεταβλητών και διαφόρων καταστάσεων προέκυψαν αρκετά χρήσιμα συμπεράσματα, ορισμένα από τα οποία είχαν επισημανθεί στη θεωρητική ανάλυση της συμπεριφοράς τους. Ο συντελεστής συσχέτισης είχε μία ιδιαίτερη συμπεριφορά το μέγεθος και η κατεύθυνση του οποίου άλλαζε για όλες σχεδόν τις σχέσεις μεταβλητών με το πετρέλαιο ανάλογα με την μορφή των δεδομένων και έτσι ισχυρές σχέσεις γινόντουσαν ασθενείς και κάποιες άλλαζαν την κατεύθυνσή τους, δίνοντας συμπεριφορές που δύσκολα μπορούν να ερμηνευτούν. Ενδεχομένως να έλεγε κανείς ότι μπορεί να υπήρχαν σχέσεις υποφαινόμενικών συσχετίσεων (spurious correlations). Εξαίρεση όλων αυτών αποτέλεσε η σχέση της τιμής του πετρελαίου με την τιμή του ενεργειακού δείκτη και του δείκτη αγαθών, η οποία παρέμεινε με συνέπεια έντονα θετική σε όλες τις μορφές δεδομένων που εξετάστηκε, ακόμα και στις τιμές των καταλοίπων.

Για όλες τις τιμές των μέτρων εξάρτησης που συνδέονται με το συντελεστή συσχέτισης, και ειδικότερα για τις τιμές των συντελεστών του Spearman και του Kendall, καθώς και για τις τιμές των Gaussian και Student-t *Copulas*, παρατηρήθηκε ότι η συμπεριφοράς τους ήταν πολύ κοντά στη συμπεριφορά του συντελεστή συσχέτισης με εξαίρεση την τελευταία περίπτωση όπου μελετήθηκαν οι σχέσεις τους σε τιμές καταλοίπων. Σε αυτή την περίπτωση εμφανίστηκαν σχέσεις με διαφορετικό μέγεθος και αντίθετο πρόσημο κάτι που δεν υπήρχε σε όλες τις άλλες. Ωστόσο και αυτά τα μέτρα μέτρησης της εξάρτησης αποφάνθηκαν ότι η σχέση του πετρελαίου με



τον ενεργειακό δείκτη είναι η εντονότερη και κατόπιν είναι η σχέση του με το δείκτη αγαθών.

Τα αποτελέσματα όμως ήταν αρκετά διαφορετικά ως προς τη μέτρηση των σχέσεων όλων των μεταβλητών με το πετρέλαιο με τη χρήση των Αρχιμήδειων *Copulas*, όπου οι δύο από τις τρεις, δηλαδή η Clayton και η Gumbel *Copulas*, έβλεπαν ανεξαρτησία σχέσεων του πετρελαίου με τις περισσότερες μεταβλητές, εκτός από τον ενεργειακό δείκτη και το δείκτη αγαθών και για αυτό και αναφέρθηκε η περίπτωση υποφαινομενικών συσχετίσεων. Έτσι, ενώ η σχέση της τιμής του πετρελαίου με το χρυσό για παράδειγμα εμφανίζονταν πολύ έντονη σε απόλυτο μέγεθος, κάτι που επιβεβαιωνόταν και από άλλα μέτρα μέτρησης της εξάρτησης, αυτές οι δύο προαναφερθέντες *Copulas* έβλεπαν ανεξαρτησία. Αντίθετα, η Frank *Copula* είχε μία έντονη κάπως ξεχωριστή παρουσία από τις άλλες δύο Αρχιμήδειες *Copulas* τόσο ως προς το μέγεθος όσο και ως προς την κατεύθυνση των σχέσεων των μεταβλητών, αλλά και αυτή η *Copula* από την αρχή παρατήρησε την έντονη σχέση της τιμής του πετρελαίου με την τιμή τους ενεργειακού δείκτη κατά πρώτον και κατά δεύτερον με το δείκτη αγαθών.

Η αλήθεια είναι ότι ο αναλυτής πολύ πιθανόν να είχε οδηγηθεί σε εσφαλμένα συμπεράσματα αν δεν είχαν προσδιοριστεί οι τιμές όλων αυτών των μέτρων μέτρησης της εξάρτησης για τη σχέση του πετρελαίου με όλες τις άλλες μεταβλητές, οι οποίες δεν επιλέχθηκαν κατά τυχαίο τρόπο αλλά χρησιμοποιούνται ευρέως στη διεθνή βιβλιογραφία.



*Αυτή η σελίδα αφήνεται σκόπιμα κενή.*



## **ΠΑΡΑΡΤΗΜΑ Α**



Πίνακας Α.1: Τιμές μέτρων εξάρτησης σε τιμές όλων των άλλων Μεταβλητών

Variables Pair	Correlation Coefficient $r$	$z$ -statistic	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
<i>USD with Yen</i>	-0.82	-16.38	-0.79	-0.58	-0.75	-0.75	0.00	-7.31	1.00
<i>USD with S&amp;P 500</i>	0.77	13.61	0.74	0.52	0.65	0.65	1.31	6.15	1.69
<i>USD with Nikkei 225</i>	0.83	16.95	0.73	0.49	0.64	0.64	1.29	5.74	1.61
<i>USD with VIX</i>	-0.11	-1.26	-0.22	-0.15	-0.24	-0.24	0.00	-1.33	1.00
<i>USD with Natural Gas</i>	-0.68	-10.65	-0.63	-0.44	-0.60	-0.61	0.00	-4.98	1.00
<i>USD with Gold</i>	-0.14	-1.63	-0.18	-0.12	-0.19	-0.19	0.00	-1.12	1.00
<i>USD with Wheat</i>	-0.76	-13.29	-0.71	-0.50	-0.69	-0.69	0.00	-5.75	1.00
<i>USD with Rice</i>	-0.74	-12.59	-0.69	-0.49	-0.67	-0.70	0.00	-6.18	1.00
<i>USD with Cotton</i>	-0.61	-8.86	-0.71	-0.52	-0.72	-0.72	0.00	-6.01	1.00
<i>USD with Energy Index</i>	-0.92	-26.38	-0.79	-0.57	-0.73	-0.73	0.00	-7.10	1.00
<i>USD with Commodity Index</i>	-0.94	-31.03	-0.85	-0.65	-0.82	-0.82	0.00	-9.43	1.00
<i>Yen with S&amp;P 500</i>	-0.71	-11.44	-0.67	-0.43	-0.57	-0.57	0.00	-4.97	1.00
<i>Yen with Nikkei 225</i>	-0.86	-19.10	-0.77	-0.58	-0.71	-0.73	0.00	-7.14	1.00
<i>Yen with VIX</i>	0.33	3.97	0.39	0.27	0.41	0.41	0.39	2.44	1.34
<i>Yen with Natural Gas</i>	0.39	4.87	0.42	0.26	0.38	0.38	0.40	2.54	1.27
<i>Yen with Gold</i>	0.42	5.31	0.47	0.36	0.50	0.51	0.83	3.30	1.45
<i>Yen with Wheat</i>	0.73	12.34	0.67	0.47	0.61	0.61	0.70	5.16	1.71
<i>Yen with Rice</i>	0.62	9.06	0.67	0.47	0.65	0.65	0.95	5.27	1.73
<i>Yen with Cotton</i>	0.59	8.33	0.63	0.43	0.62	0.62	0.81	4.53	1.64
<i>Yen with Energy Index</i>	0.70	11.06	0.60	0.33	0.50	0.50	0.52	3.85	1.45
<i>Yen with Commodity Index</i>	0.77	13.73	0.68	0.44	0.61	0.61	0.65	5.15	1.73
<i>S&amp;P 500 with Nikkei 225</i>	0.95	36.07	0.95	0.80	0.90	0.91	2.06	17.42	76.23
<i>S&amp;P 500 with VIX</i>	-0.10	-1.16	-0.29	-0.25	-0.27	-0.32	0.00	-2.04	1.00
<i>S&amp;P 500 with Natural Gas</i>	-0.60	-8.60	-0.62	-0.43	-0.63	-0.63	0.00	-4.68	1.00
<i>S&amp;P 500 with Gold</i>	0.22	2.63	0.19	0.15	0.34	0.08	0.33	1.30	1.28
<i>S&amp;P 500 with Wheat</i>	-0.55	-7.42	-0.53	-0.33	-0.43	-0.43	0.00	-3.47	1.00
<i>S&amp;P 500 with Rice</i>	-0.43	-5.38	-0.44	-0.26	-0.34	-0.35	0.00	-2.78	1.00
<i>S&amp;P 500 with Cotton</i>	-0.47	-6.04	-0.55	-0.34	-0.54	-0.54	0.00	-3.68	1.00
<i>S&amp;P 500 with Energy Index</i>	-0.80	-15.06	-0.70	-0.41	-0.64	-0.64	0.00	-5.24	1.00
<i>S&amp;P 500 with Commodity Index</i>	-0.78	-14.29	-0.74	-0.49	-0.67	-0.67	0.00	-5.91	1.00
<i>Nikkei 225 with VIX</i>	-0.21	-2.48	-0.30	-0.23	-0.33	-0.33	0.00	-2.06	1.00
<i>Nikkei 225 with Natural Gas</i>	-0.54	-7.38	-0.52	-0.34	-0.46	-0.46	0.00	-3.47	1.00
<i>Nikkei 225 with Gold</i>	-0.03	-0.31	0.00	0.00	0.05	-0.01	0.00	0.01	1.10
<i>Nikkei 225 with Wheat</i>	-0.66	-10.13	-0.58	-0.37	-0.52	-0.52	0.00	-3.99	1.00
<i>Nikkei 225 with Rice</i>	-0.55	-7.43	-0.50	-0.30	-0.43	-0.43	0.00	-3.12	1.00
<i>Nikkei 225 with Cotton</i>	-0.50	-6.58	-0.51	-0.32	-0.48	-0.48	0.00	-3.28	1.00
<i>Nikkei 225 with Energy Index</i>	-0.82	-16.07	-0.69	-0.42	-0.62	-0.62	0.00	-4.86	1.00
<i>Nikkei 225 with Commodity Index</i>	-0.82	-16.45	-0.72	-0.48	-0.66	-0.66	0.00	-5.60	1.00
<i>VIX with Natural Gas</i>	-0.10	-1.16	-0.10	-0.08	-0.15	-0.13	0.00	-0.62	1.00
<i>VIX with Gold</i>	0.43	5.41	0.25	0.17	0.28	0.28	0.00	1.62	1.27
<i>VIX with Wheat</i>	0.13	1.45	0.28	0.18	0.32	0.32	0.55	1.70	1.13
<i>VIX with Rice</i>	0.26	3.12	0.23	0.15	0.31	0.31	0.31	1.42	1.21
<i>VIX with Cotton</i>	0.03	0.34	0.02	0.00	-0.03	-0.01	0.00	0.09	1.03
<i>VIX with Energy Index</i>	-0.07	-0.80	-0.04	-0.03	-0.12	-0.05	0.00	-0.29	1.00
<i>VIX with Commodity Index</i>	-0.06	-0.66	0.02	0.02	-0.03	0.00	0.00	0.14	1.00
<i>Natural Gas with Gold</i>	-0.34	-4.07	-0.30	-0.21	-0.37	-0.33	0.00	-2.11	1.00
<i>Natural Gas with Wheat</i>	0.37	4.57	0.38	0.23	0.32	0.32	0.26	2.23	1.25
<i>Natural Gas with Rice</i>	0.32	3.83	0.31	0.19	0.24	0.27	0.13	1.91	1.22
<i>Natural Gas with Cotton</i>	0.50	6.55	0.63	0.47	0.65	0.66	1.51	5.25	1.68
<i>Natural Gas with Energy Index</i>	0.70	11.32	0.64	0.46	0.67	0.67	1.30	5.18	1.77
<i>Natural Gas with Commodity Index</i>	0.68	10.46	0.65	0.46	0.66	0.67	1.60	5.25	1.70
<i>Gold with Wheat</i>	0.37	4.60	0.34	0.21	0.33	0.34	0.24	1.88	1.22
<i>Gold with Rice</i>	0.43	5.45	0.41	0.28	0.39	0.39	0.43	2.57	1.32
<i>Gold with Cotton</i>	0.06	0.69	0.15	0.10	0.12	0.14	0.19	0.96	1.07
<i>Gold with Energy Index</i>	-0.06	-0.67	-0.07	-0.05	-0.17	-0.11	0.00	-0.46	1.00
<i>Gold with Commodity Index</i>	0.06	0.65	0.07	0.06	0.02	0.09	0.00	0.47	1.08
<i>Wheat with Rice</i>	0.73	12.12	0.73	0.53	0.71	0.71	1.51	6.28	1.79
<i>Wheat with Cotton</i>	0.53	7.11	0.49	0.33	0.46	0.46	0.37	3.01	1.44
<i>Wheat with Energy Index</i>	0.72	11.72	0.69	0.47	0.61	0.61	0.84	5.31	1.63
<i>Wheat with Commodity Index</i>	0.79	14.86	0.75	0.54	0.70	0.70	0.86	6.53	1.97
<i>Rice with Cotton</i>	0.35	4.29	0.43	0.28	0.37	0.39	0.34	2.70	1.33
<i>Rice with Energy Index</i>	0.66	10.05	0.61	0.41	0.49	0.58	0.63	4.75	1.61
<i>Rice with Commodity Index</i>	0.68	10.57	0.61	0.41	0.50	0.58	0.64	4.77	1.61
<i>Cotton with Energy Index</i>	0.53	7.09	0.65	0.46	0.68	0.68	1.15	4.89	1.73
<i>Cotton with Commodity Index</i>	0.64	9.58	0.73	0.55	0.78	0.78	1.55	6.34	2.31
<i>Energy Index with Commodity Index</i>	0.98	52.26	0.95	0.83	0.93	0.93	6.56	19.71	3.37



**Πίνακας Α.2: Τιμές μέτρων εξάρτησης σε πρώτες διαφορές των τιμών όλων των άλλων μεταβλητών**

Variables Pair	Correlation Coefficient $r$	$t$ -statistic	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
USD with Yen	-0.36	-4.41	-0.39	-0.28	-0.41	-0.42	0.00	-2.74	1.00
USD with S&P 500	-0.15	-1.72	-0.13	-0.08	-0.20	-0.19	0.00	-0.80	1.00
USD with Nikkei 225	0.06	0.70	0.09	0.06	0.06	0.09	0.03	0.59	1.07
USD with VIX	0.10	1.12	0.07	0.04	0.11	0.11	0.07	0.41	1.07
USD with Natural Gas	0.09	1.08	0.16	0.10	0.13	0.14	0.14	0.99	1.08
USD with Gold	-0.31	-3.65	-0.31	-0.22	-0.34	-0.35	0.00	-2.13	1.00
USD with Wheat	-0.14	-1.60	-0.16	-0.11	-0.15	-0.17	0.00	-1.02	1.00
USD with Rice	-0.12	-1.33	-0.16	-0.11	-0.14	-0.18	0.01	-1.04	1.00
USD with Cotton	-0.12	-1.43	-0.15	-0.10	-0.15	-0.15	0.00	-0.94	1.00
USD with Energy Index	-0.35	-4.21	-0.27	-0.18	-0.33	-0.33	0.00	-1.77	1.00
USD with Commodity Index	-0.47	-6.04	-0.43	-0.30	-0.48	-0.48	0.00	-3.07	1.00
Yen with S&P 500	-0.25	-2.88	-0.27	-0.18	-0.30	-0.29	0.00	-1.58	1.00
Yen with Nikkei 225	-0.56	-7.71	-0.63	-0.45	-0.63	-0.63	0.00	-4.89	1.00
Yen with VIX	0.24	2.86	0.23	0.16	0.25	0.25	0.17	1.44	1.20
Yen with Natural Gas	-0.17	-1.98	-0.18	-0.12	-0.20	-0.20	0.00	-1.12	1.00
Yen with Gold	0.41	5.15	0.42	0.30	0.42	0.45	0.71	3.05	1.37
Yen with Wheat	0.12	1.40	0.07	0.05	0.10	0.10	0.06	0.45	1.07
Yen with Rice	0.08	0.95	0.08	0.05	0.11	0.11	0.00	0.46	1.08
Yen with Cotton	-0.08	-0.95	-0.01	-0.01	-0.04	-0.03	0.00	-0.06	1.01
Yen with Energy Index	-0.15	-1.78	-0.19	-0.13	-0.20	-0.21	0.00	-1.23	1.00
Yen with Commodity Index	-0.05	-0.54	-0.03	-0.02	-0.03	-0.03	0.00	-0.19	1.02
S&P 500 with Nikkei 225	0.79	14.56	0.69	0.52	0.74	0.74	1.87	5.90	1.93
S&P 500 with VIX	-0.85	-18.07	-0.71	-0.53	-0.74	-0.74	0.00	-6.35	1.00
S&P 500 with Natural Gas	0.02	0.18	-0.03	-0.02	-0.02	-0.02	0.00	-0.19	1.00
S&P 500 with Gold	0.03	0.29	0.02	0.01	0.03	0.03	0.00	0.10	1.06
S&P 500 with Wheat	0.03	0.38	0.05	0.03	0.04	0.04	0.05	0.30	1.00
S&P 500 with Rice	-0.06	-0.70	-0.08	-0.06	-0.08	-0.08	0.00	-0.50	1.00
S&P 500 with Cotton	0.20	2.32	0.29	0.19	0.31	0.30	0.39	1.70	1.16
S&P 500 with Energy Index	0.44	5.51	0.42	0.29	0.43	0.43	0.77	2.82	1.30
S&P 500 with Commodity Index	0.46	5.85	0.46	0.31	0.48	0.48	0.81	3.04	1.36
Nikkei 225 with VIX	-0.68	-10.62	-0.53	-0.38	-0.57	-0.57	0.00	-3.95	1.00
Nikkei 225 with Natural Gas	0.10	1.09	0.10	0.06	0.11	0.11	0.10	0.57	1.07
Nikkei 225 with Gold	-0.24	-2.82	-0.30	-0.21	-0.31	-0.32	0.00	-1.94	1.00
Nikkei 225 with Wheat	-0.02	-0.25	-0.05	-0.03	-0.03	-0.03	0.03	-0.25	1.00
Nikkei 225 with Rice	-0.05	-0.53	-0.06	-0.03	-0.05	-0.05	0.00	-0.32	1.00
Nikkei 225 with Cotton	0.13	1.50	0.21	0.13	0.21	0.21	0.29	1.22	1.09
Nikkei 225 with Energy Index	0.35	4.27	0.36	0.26	0.36	0.39	0.56	2.50	1.30
Nikkei 225 with Commodity Index	0.35	4.20	0.31	0.20	0.33	0.34	0.54	1.92	1.20
VIX with Natural Gas	-0.03	-0.33	0.03	0.02	0.03	0.03	0.08	0.19	1.00
VIX with Gold	0.03	0.35	-0.01	0.00	-0.01	0.00	0.02	-0.06	1.03
VIX with Wheat	-0.08	-0.88	-0.10	-0.07	-0.11	-0.12	0.00	-0.60	1.00
VIX with Rice	0.02	0.21	0.09	0.06	0.07	0.09	0.15	0.58	1.04
VIX with Cotton	-0.16	-1.90	-0.17	-0.12	-0.19	-0.19	0.00	-1.07	1.00
VIX with Energy Index	-0.43	-5.37	-0.41	-0.28	-0.41	-0.44	0.00	-2.84	1.00
VIX with Commodity Index	-0.43	-5.33	-0.45	-0.32	-0.46	-0.48	0.00	-3.26	1.00
Natural Gas with Gold	-0.13	-1.54	-0.14	-0.09	-0.15	-0.15	0.00	-0.87	1.00
Natural Gas with Wheat	-0.01	-0.11	-0.02	-0.02	-0.04	-0.03	0.00	-0.13	1.00
Natural Gas with Rice	-0.04	-0.48	-0.04	-0.03	-0.04	-0.05	0.04	-0.23	1.00
Natural Gas with Cotton	0.00	-0.02	0.00	0.00	0.00	0.00	0.01	0.03	1.02
Natural Gas with Energy Index	0.16	1.84	0.14	0.10	0.17	0.16	0.19	0.85	1.12
Natural Gas with Commodity Index	0.09	1.02	0.08	0.05	0.10	0.09	0.09	0.50	1.09
Gold with Wheat	0.06	0.66	0.06	0.05	0.07	0.07	0.02	0.38	1.03
Gold with Rice	0.01	0.13	0.04	0.03	0.03	0.04	0.00	0.24	1.06
Gold with Cotton	-0.01	-0.12	0.04	0.03	0.04	0.04	0.06	0.25	1.01
Gold with Energy Index	-0.02	-0.26	0.06	0.04	0.04	0.04	0.00	0.36	1.04
Gold with Commodity Index	0.15	1.70	0.22	0.15	0.21	0.22	0.17	1.40	1.17
Wheat with Rice	0.21	2.44	0.20	0.14	0.22	0.23	0.34	1.36	1.15
Wheat with Cotton	0.04	0.48	0.02	0.01	0.03	0.01	0.04	0.12	1.06
Wheat with Energy Index	0.01	0.12	0.01	0.01	-0.01	0.01	0.02	0.04	1.03
Wheat with Commodity Index	0.29	3.44	0.24	0.16	0.28	0.27	0.26	1.54	1.23
Rice with Cotton	-0.06	-0.67	0.07	0.05	0.03	0.11	0.03	0.48	1.07
Rice with Energy Index	-0.03	-0.30	-0.01	-0.01	-0.04	-0.02	0.02	-0.07	1.00
Rice with Commodity Index	0.05	0.61	0.06	0.04	0.08	0.08	0.05	0.40	1.08
Cotton with Energy Index	0.30	3.59	0.33	0.23	0.36	0.36	0.50	2.15	1.27
Cotton with Commodity Index	0.41	5.17	0.47	0.33	0.50	0.51	0.67	3.34	1.51
Energy Index with Commodity Index	0.89	21.67	0.88	0.71	0.88	0.89	4.10	11.93	2.75



**Πίνακας Α.3: Τιμές μέτρων εξάρτησης σε εκτιμηθείσες τιμές μεταβλητών από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα**

Variables Pair	Correlation Coefficient $r$	$t$ -statistic	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
USD with Yen	-0.16	-1.86	-0.75	-0.55	-0.67	-0.72	0.00	-6.79	1.00
USD with S&P 500	0.72	11.93	0.75	0.52	0.67	0.68	1.49	6.27	1.73
USD with Nikkei 225	0.77	13.78	0.73	0.50	0.66	0.66	1.47	5.90	1.64
USD with VIX	0.09	1.09	-0.19	-0.13	-0.16	-0.19	0.00	-1.17	1.00
USD with Natural Gas	-0.25	-2.98	-0.59	-0.42	-0.52	-0.59	0.00	-4.65	1.00
USD with Gold	0.26	3.12	-0.16	-0.11	-0.14	-0.17	0.00	-1.00	1.00
USD with Wheat	-0.25	-2.96	-0.67	-0.47	-0.61	-0.66	0.00	-5.39	1.00
USD with Rice	-0.15	-1.68	-0.64	-0.47	-0.57	-0.68	0.00	-5.71	1.00
USD with Cotton	-0.24	-2.85	-0.64	-0.47	-0.61	-0.67	0.00	-5.27	1.00
USD with Energy Index	-0.53	-7.04	-0.75	-0.55	-0.66	-0.71	0.00	-6.79	1.00
USD with Commodity Index	-0.42	-5.24	-0.80	-0.62	-0.73	-0.80	0.00	-8.64	1.00
Yen with S&P 500	-0.46	-5.89	-0.64	-0.42	-0.52	-0.57	0.00	-4.80	1.00
Yen with Nikkei 225	-0.58	-8.04	-0.74	-0.57	-0.64	-0.73	0.00	-6.86	1.00
Yen with VIX	0.42	5.26	0.42	0.30	0.46	0.46	0.54	2.70	1.39
Yen with Natural Gas	0.47	6.15	0.43	0.27	0.41	0.41	0.52	2.65	1.30
Yen with Gold	0.58	8.22	0.49	0.37	0.54	0.54	1.02	3.50	1.50
Yen with Wheat	0.77	13.75	0.66	0.47	0.63	0.63	0.85	5.19	1.73
Yen with Rice	0.70	11.27	0.67	0.46	0.66	0.66	1.10	5.24	1.76
Yen with Cotton	0.61	8.70	0.62	0.42	0.63	0.63	0.91	4.43	1.64
Yen with Energy Index	0.68	10.58	0.61	0.35	0.55	0.55	0.70	4.06	1.49
Yen with Commodity Index	0.79	14.57	0.69	0.45	0.65	0.65	0.84	5.35	1.78
S&P 500 with Nikkei 225	0.95	36.67	0.95	0.80	0.91	0.91	2.49	17.50	76.07
S&P 500 with VIX	-0.05	-0.57	-0.29	-0.24	-0.22	-0.33	0.00	-2.09	1.00
S&P 500 with Natural Gas	-0.47	-6.01	-0.57	-0.40	-0.55	-0.59	0.00	-4.29	1.00
S&P 500 with Gold	0.30	3.53	0.18	0.15	0.34	0.10	0.39	1.24	1.27
S&P 500 with Wheat	-0.40	-4.96	-0.50	-0.32	-0.38	-0.44	0.00	-3.33	1.00
S&P 500 with Rice	-0.24	-2.80	-0.41	-0.24	-0.28	-0.36	0.00	-2.63	1.00
S&P 500 with Cotton	-0.36	-4.44	-0.50	-0.31	-0.46	-0.49	0.00	-3.37	1.00
S&P 500 with Energy Index	-0.70	-11.13	-0.66	-0.39	-0.56	-0.60	0.00	-4.84	1.00
S&P 500 with Commodity Index	-0.62	-9.05	-0.69	-0.46	-0.59	-0.64	0.00	-5.50	1.00
Nikkei 225 with VIX	-0.15	-1.77	-0.31	-0.23	-0.29	-0.33	0.00	-2.12	1.00
Nikkei 225 with Natural Gas	-0.40	-5.04	-0.47	-0.32	-0.39	-0.43	0.00	-3.18	1.00
Nikkei 225 with Gold	0.09	1.03	-0.01	0.00	0.07	0.05	0.01	-0.05	1.10
Nikkei 225 with Wheat	-0.50	-6.54	-0.55	-0.35	-0.47	-0.51	0.00	-3.82	1.00
Nikkei 225 with Rice	-0.33	-4.01	-0.46	-0.28	-0.37	-0.41	0.00	-2.98	1.00
Nikkei 225 with Cotton	-0.38	-4.74	-0.46	-0.29	-0.39	-0.42	0.00	-2.98	1.00
Nikkei 225 with Energy Index	-0.71	-11.54	-0.64	-0.39	-0.54	-0.57	0.00	-4.54	1.00
Nikkei 225 with Commodity Index	-0.65	-9.82	-0.68	-0.45	-0.58	-0.62	0.00	-5.30	1.00
VIX with Natural Gas	-0.01	-0.14	-0.07	-0.06	-0.08	-0.09	0.00	-0.44	1.00
VIX with Gold	0.49	6.43	0.27	0.18	0.32	0.30	0.19	1.74	1.29
VIX with Wheat	0.21	2.46	0.30	0.20	0.35	0.35	0.66	1.80	1.16
VIX with Rice	0.34	4.16	0.25	0.17	0.35	0.34	0.41	1.58	1.24
VIX with Cotton	0.09	1.02	0.05	0.02	0.05	0.04	0.02	0.28	1.06
VIX with Energy Index	-0.02	-0.23	-0.01	-0.01	-0.05	-0.01	0.12	-0.05	1.00
VIX with Commodity Index	0.03	0.32	0.05	0.04	0.03	0.03	0.09	0.33	1.01
Natural Gas with Gold	-0.10	-1.18	-0.26	-0.19	-0.29	-0.24	0.00	-1.82	1.00
Natural Gas with Wheat	0.44	5.65	0.40	0.24	0.36	0.36	0.40	2.38	1.29
Natural Gas with Rice	0.40	5.03	0.31	0.20	0.28	0.30	0.30	1.98	1.23
Natural Gas with Cotton	0.53	7.05	0.62	0.46	0.64	0.66	1.53	5.19	1.68
Natural Gas with Energy Index	0.71	11.54	0.65	0.46	0.67	0.68	1.42	5.23	1.78
Natural Gas with Commodity Index	0.70	11.29	0.65	0.46	0.67	0.68	1.70	5.37	1.72
Gold with Wheat	0.49	6.41	0.35	0.22	0.37	0.37	0.36	1.98	1.25
Gold with Rice	0.57	7.95	0.43	0.29	0.44	0.44	0.57	2.79	1.38
Gold with Cotton	0.18	2.11	0.17	0.12	0.18	0.19	0.33	1.10	1.10
Gold with Energy Index	0.07	0.78	-0.03	-0.02	-0.07	-0.01	0.01	-0.17	1.00
Gold with Commodity Index	0.23	2.72	0.11	0.08	0.10	0.14	0.10	0.73	1.12
Wheat with Rice	0.77	13.61	0.74	0.53	0.72	0.72	1.67	6.42	1.82
Wheat with Cotton	0.56	7.70	0.48	0.33	0.48	0.48	0.47	3.00	1.46
Wheat with Energy Index	0.72	11.97	0.70	0.48	0.64	0.64	1.03	5.49	1.67
Wheat with Commodity Index	0.82	16.23	0.75	0.55	0.72	0.72	1.04	6.72	2.03
Rice with Cotton	0.41	5.06	0.42	0.27	0.40	0.41	0.45	2.63	1.33
Rice with Energy Index	0.65	9.80	0.60	0.42	0.52	0.60	0.80	4.78	1.61
Rice with Commodity Index	0.71	11.49	0.61	0.42	0.53	0.60	0.79	4.86	1.64
Cotton with Energy Index	0.54	7.33	0.64	0.45	0.67	0.67	1.18	4.77	1.71
Cotton with Commodity Index	0.66	10.10	0.71	0.53	0.77	0.76	1.50	6.13	2.20
Energy Index with Commodity Index	0.97	43.55	0.96	0.83	0.93	0.94	7.36	20.20	3.56





**Πίνακας Α.4: Τιμές μέτρων εξάρτησης σε τιμές καταλοίπων των μεταβλητών από το επιλεγθέν ARIMA(1, 1, 0) υπόδειγμα.**

Variables Pair	Correlation Coefficient $r$	$t$ -statistic	Spearman $\rho$	Kendall $\tau$	Gaussian	Student- $t$	Clayton	Frank	Gumbel
USD with Yen	0.95	33.13	-0.36	-0.27	-0.33	-0.39	0.00	-2.60	1.00
USD with S&P 500	0.73	12.13	-0.09	-0.06	-0.09	-0.10	0.00	-0.56	1.02
USD with Nikkei 225	0.75	12.82	0.13	0.09	0.12	0.13	0.05	0.82	1.13
USD with VIX	0.43	5.38	0.09	0.06	0.14	0.14	0.07	0.56	1.12
USD with Natural Gas	0.84	17.64	0.20	0.14	0.21	0.21	0.20	1.26	1.14
USD with Gold	0.86	19.06	-0.24	-0.18	-0.23	-0.29	0.00	-1.74	1.01
USD with Wheat	0.73	12.23	-0.06	-0.04	-0.06	-0.08	0.00	-0.41	1.03
USD with Rice	0.86	19.19	-0.04	-0.02	-0.01	-0.05	0.03	-0.23	1.03
USD with Cotton	0.66	10.12	-0.09	-0.05	-0.04	-0.06	0.00	-0.52	1.03
USD with Energy Index	0.84	17.87	-0.22	-0.15	-0.21	-0.23	0.00	-1.39	1.01
USD with Commodity Index	0.89	22.50	-0.36	-0.25	-0.37	-0.39	0.00	-2.50	1.00
Yen with S&P 500	0.70	11.31	-0.25	-0.17	-0.26	-0.26	0.00	-1.50	1.00
Yen with Nikkei 225	0.66	10.04	-0.59	-0.42	-0.56	-0.60	0.00	-4.47	1.00
Yen with VIX	0.46	5.95	0.29	0.21	0.34	0.34	0.27	1.86	1.28
Yen with Natural Gas	0.81	15.49	-0.16	-0.11	-0.15	-0.17	0.00	-0.99	1.01
Yen with Gold	0.91	25.41	0.41	0.30	0.42	0.45	0.62	2.96	1.39
Yen with Wheat	0.76	13.27	0.12	0.07	0.14	0.14	0.08	0.70	1.11
Yen with Rice	0.87	20.06	0.07	0.05	0.12	0.12	0.00	0.43	1.10
Yen with Cotton	0.67	10.18	0.02	0.01	0.04	0.02	0.00	0.09	1.06
Yen with Energy Index	0.84	17.99	-0.20	-0.14	-0.20	-0.22	0.00	-1.31	1.01
Yen with Commodity Index	0.91	25.09	-0.06	-0.04	-0.06	-0.07	0.00	-0.38	1.04
S&P 500 with Nikkei 225	0.91	24.98	0.70	0.53	0.76	0.74	1.87	6.19	2.03
S&P 500 with VIX	-0.20	-2.32	-0.69	-0.52	-0.70	-0.77	0.00	-6.31	1.00
S&P 500 with Natural Gas	0.65	9.73	0.00	-0.01	0.04	0.01	0.00	-0.03	1.06
S&P 500 with Gold	0.69	10.97	0.05	0.04	0.09	0.07	0.00	0.33	1.10
S&P 500 with Wheat	0.59	8.30	0.06	0.04	0.08	0.08	0.03	0.34	1.05
S&P 500 with Rice	0.66	9.97	-0.07	-0.05	-0.02	-0.06	0.00	-0.41	1.05
S&P 500 with Cotton	0.59	8.33	0.21	0.14	0.24	0.24	0.26	1.23	1.14
S&P 500 with Energy Index	0.80	15.21	0.42	0.29	0.46	0.46	0.75	2.83	1.36
S&P 500 with Commodity Index	0.80	15.33	0.45	0.31	0.49	0.49	0.71	3.02	1.41
Nikkei 225 with VIX	-0.11	-1.27	-0.53	-0.39	-0.55	-0.57	0.00	-4.10	1.00
Nikkei 225 with Natural Gas	0.67	10.41	0.13	0.09	0.17	0.16	0.15	0.76	1.12
Nikkei 225 with Gold	0.62	8.96	-0.26	-0.17	-0.23	-0.27	0.00	-1.64	1.00
Nikkei 225 with Wheat	0.57	7.88	-0.02	-0.02	0.02	0.02	0.03	-0.11	1.03
Nikkei 225 with Rice	0.66	9.96	-0.01	0.00	0.03	0.02	0.00	-0.06	1.04
Nikkei 225 with Cotton	0.56	7.78	0.13	0.09	0.17	0.17	0.20	0.81	1.09
Nikkei 225 with Energy Index	0.78	14.19	0.38	0.27	0.42	0.42	0.60	2.66	1.38
Nikkei 225 with Commodity Index	0.78	14.38	0.33	0.23	0.39	0.37	0.51	2.24	1.31
VIX with Natural Gas	0.34	4.10	0.05	0.03	0.08	0.06	0.07	0.31	1.06
VIX with Gold	0.38	4.73	0.02	0.01	0.04	0.03	0.06	0.11	1.07
VIX with Wheat	0.28	3.28	-0.05	-0.04	-0.05	-0.05	0.00	-0.32	1.03
VIX with Rice	0.36	4.46	0.08	0.06	0.08	0.09	0.11	0.55	1.08
VIX with Cotton	0.21	2.50	-0.08	-0.06	-0.08	-0.09	0.00	-0.54	1.02
VIX with Energy Index	0.18	2.14	-0.36	-0.25	-0.35	-0.39	0.00	-2.48	1.00
VIX with Commodity Index	0.26	3.07	-0.40	-0.28	-0.39	-0.43	0.00	-2.83	1.00
Natural Gas with Gold	0.72	11.92	-0.16	-0.10	-0.12	-0.15	0.00	-0.97	1.02
Natural Gas with Wheat	0.64	9.45	-0.01	-0.01	0.01	0.00	0.00	-0.05	1.06
Natural Gas with Rice	0.74	12.37	-0.01	0.00	0.02	0.00	0.04	-0.09	1.04
Natural Gas with Cotton	0.56	7.77	-0.04	-0.02	0.00	-0.04	0.00	-0.25	1.05
Natural Gas with Energy Index	0.79	14.90	0.18	0.12	0.23	0.21	0.22	1.12	1.19
Natural Gas with Commodity Index	0.81	15.82	0.10	0.07	0.14	0.13	0.09	0.69	1.14
Gold with Wheat	0.70	11.25	0.11	0.07	0.13	0.12	0.03	0.69	1.11
Gold with Rice	0.79	14.49	0.00	0.00	0.02	-0.01	0.00	-0.03	1.07
Gold with Cotton	0.64	9.39	0.14	0.09	0.18	0.17	0.18	0.84	1.11
Gold with Energy Index	0.79	14.59	0.06	0.04	0.06	0.06	0.00	0.39	1.09
Gold with Commodity Index	0.86	19.48	0.21	0.15	0.21	0.23	0.20	1.40	1.18
Wheat with Rice	0.74	12.61	0.19	0.13	0.25	0.22	0.32	1.24	1.20
Wheat with Cotton	0.54	7.25	0.02	0.02	0.04	0.03	0.00	0.12	1.07
Wheat with Energy Index	0.68	10.48	-0.02	-0.02	0.01	-0.01	0.00	-0.15	1.06
Wheat with Commodity Index	0.78	14.22	0.19	0.13	0.25	0.22	0.22	1.23	1.20
Rice with Cotton	0.58	8.16	0.06	0.05	0.08	0.07	0.00	0.42	1.11
Rice with Energy Index	0.78	14.34	0.05	0.03	0.07	0.06	0.05	0.31	1.08
Rice with Commodity Index	0.84	17.71	0.08	0.05	0.13	0.09	0.05	0.52	1.14
Cotton with Energy Index	0.67	10.26	0.18	0.12	0.23	0.22	0.25	1.08	1.16
Cotton with Commodity Index	0.71	11.65	0.29	0.20	0.36	0.32	0.43	1.86	1.29
Energy Index with Commodity Index	0.97	47.88	0.88	0.70	0.89	0.89	3.27	11.29	3.08



*Αυτή η σελίδα αφήνεται σκόπιμα κενή.*



## Βιβλιογραφία

Χρήστος Ν. Αγιακλόγλου και Θεοφάνης Ε. Μπένος, *Αρχές Οικονομετρικής Ανάλυσης*, 2014, Εκδόσεις Ε. Μπένου, Αθήνα.

Agiakloglou, C. (2009). Evidence of ARCH(1) errors in the context of spurious regressions, *Communications in Statistics – Simulations and Computation*, 38, 1803-1810.

Agiakloglou, C. (2013). Resolving spurious regressions and serially correlated errors, *Empirical Economics*, 45, 1361-1366.

Agiakloglou, C. and Agiropoulos, C. (2016). The Balance between Size and Power in testing for linear association for two stationary AR(1) processes, *Applied Economics Letters*, 23, 230-234.

Agiakloglou, C. and Agiropoulos, C. (2021). Dealing with serially correlated errors in the context of spurious regression for two independent stationary AR(1) processes, *Applied Economic Letters*, 29(7), 1-7.

Agiakloglou, C. and Tsimpanos, A. (2012). An alternative approach for testing for linear association for two independent stationary AR(1) processes, *Applied Economics*, 44, 4799-4803.

Agiakloglou, C., Tsimbos, C. and Tsimpanos, A. (2015). Is spurious behavior an issue for two independent stationary spatial autoregressive SAR(1) processes? *Applied Economics Letters*, 22, 1372-1377.

Alexander, C. (2005). Correlation in crude oil and natural gas markets. In V. Kaminski (Ed.), *Managing Energy Price Risk: The New Challenges and Solutions*, 573-604. Risk Books, London.

Aloui, R., Hammoudeh, S. and Nguyen, D. K. (2013). A time-varying copula approach to oil and stock market dependence: the case of transition economies, *Energy Economics*, 39, 208-221. <https://doi.org/10.1016/j.eneco.2013.04.012>



- Ansley, C. F. and Newbold, P. (1980). Finite sample properties of estimators for autoregressive moving average models, *Journal of Econometrics*, 13(2), 159–183. [https://doi.org/10.1016/0304-4076\(80\)90013-5](https://doi.org/10.1016/0304-4076(80)90013-5)
- Bai, J. (1994). Least Squares Estimation of a Shift in Linear Processes, *Journal of Time Series Analysis*, 15(5), 453–472.
- Banerjee, A., Dolado, J. J., Galbraith, J. W. and Hendry, D. F. (1993). Co-integration, error correction, and the econometric analysis of non-stationary data. In Oxford University Press eBooks. <https://doi.org/10.1093/0198288107.001.0001>
- Beckmann, J., Berger, T. and Czudaj, R. (2015). Oil price and FX-rates dependency, *Quantitative Finance*, 16(3), 477–488.
- Bera, A., Higgins, M. L. and Lee, S. (1992). Interaction between autocorrelation and conditional heteroscedasticity: A random-coefficient approach, *Journal of Business & Economic Statistics*, 10, 133–142.
- Bera, A., Higgins, M. L. and Lee, S. (1996). Random coefficient formulation of conditional heteroscedasticity and augmented ARCH models, *The Indian Journal of Statistics*, 58, 199–220.
- Box, G. E. P. and Jenkins, G. M. (1976). Time Series Analysis, Forecasting and Control, 2nd Edition, Holden Day, San Francisco.
- Casella, G. and Berger, R.L. (2001) Statistical Inference. 2nd Edition, Duxbury Press.
- Chen, X. and Fan, Y. (2006). Estimation of copula-based semiparametric time series models, *Journal of Econometrics*, 130(2), 307–335. <https://doi.org/10.1016/j.jeconom.2005.03.004>
- Cherubini, U., Gobbi, F., Mulinacci, S. and Romagnoli, S. (2012). Dynamic Copula Models in Finance. Chichester, UK: John Wiley & Sons.
- Cochrane, D. and Orcutt, G. H. (1949). Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms, *Journal of the American Statistical Association*, 44, 32–61.
- Coles, S. (2001). An Introduction to Statistical Modeling of Extreme Values. Springer Verlag, Berlin. <http://dx.doi.org/10.1007/978-1-4471-3675-0>



Demarta, S. and McNeil, A. J. (2007). The t Copula and Related Copulas, *International Statistical Review*, 73(1), 111–129. <https://doi.org/10.1111/j.1751-5823.2005.tb0024.x>

Embrechts, P., Hoing, A. and Juri, A. (2003). Using copulae to bound the value-at-risk for functions of dependent risks, *Finance Stochastics*, 7, 145–167.

Embrechts, P., McNeil, A. and Straumann, D. (2002). Correlation and Dependence in Risk Management: Properties and Pitfalls, Risk Management: Value at Risk and Beyond, edited by Dempster, M., 176-223, Cambridge University Press, Cambridge.

Embrechts, P., McNeil, A. J. and Straumann, D. (1999). Correlation: Pitfalls and Alternatives, *Risk*, 12, 69–71.

Engle, R. E. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50, 987-1008.

Entorf, H. (1997). Random walks with drifts: nonsense regression and spurious fixed-effect estimation, *Journal of Econometrics*, 80, 287–296.

Fang, H. B., Fang, K. T. and Kotz, S. (2002). The Meta-elliptical Distributions with Given Marginals, *Journal of Multivariate Analysis*, 82(1), 1–16. <https://doi.org/10.1006/jmva.2001.2017>

Fisher, R. A. (1915). Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population, *Biometrika*, 10(4), 507. <https://doi.org/10.2307/2331838>

Fréchet, M. (1958). Remarques au sujet de la note précédente, *CR Acad. Sci. Paris Sér. I Math*, 246, 2719-2720.

Genest, C. and MacKay, R. J. (1986). The joy of copulas: Bivariate distributions with uniform marginal, *The American Statistician*, 40, 280-283.

Genest, C. and Remillard, B. (2004). Test of independence and randomness based on the empirical copula process, *TEST*, 13(2), 335–369.

Ghoudi, K. and McDonald, D. J. (1994). A nonparametric test for homogeneity: applications to parameter estimation. In *Lecture notes-monograph series* (pp. 149–156).



Granger, C. W. J. (2008). Non-linear Models: Where do we go next – time varying parameter models? *Studies in Nonlinear Dynamics & Econometrics*, 12, 1-9.

Granger, C. W. J. and Newbold, P. (1974). Spurious regressions in econometrics, *Journal of Econometrics*, 2, 111-120.

Granger, C. W. J. and Terasvirta, T. (1993). Modeling nonlinear economic relationships, Oxford University Press, New York.

Granger, C. W. J., Hyung, N. and Jeon, Y. (2001). Spurious regressions with stationary series, *Applied Economics*, 33, 899-904.

Granger, C. and Newbold, P. (1974). Spurious regressions in econometrics, *Journal of Econometrics*, 2(2), 111–120. [https://doi.org/10.1016/0304-4076\(74\)90034-7](https://doi.org/10.1016/0304-4076(74)90034-7)

Gregoire, V., Genest, C. and Gendron, M. (2008). Using copulas to model price dependence in energy markets, *Energy Risk*, 8, 62-68.

Janus, P., Koopman, S.J. and Lucas, A. (2014). Long memory dynamics for multivariate dependence under heavy tails, *Journal of Empirical Finance*, 29, 187–206.

Joe, H. (1997) Multivariate Models and Dependence Concepts. Monographs in Statistics and Probability. Chapman and Hall, London. <http://dx.doi.org/10.1201/b13150>

Jondeau, E. and Rockinger, M. (2006). The Copula-GARCH model of conditional dependencies: an international stock market application, *Journal of International Money and Finance*, 25, 827–853.

Kendall, M. (1938) A New Measure of Rank Correlation, *Biometrika*, 30, 81-89. <https://doi.org/10.1093/biomet/30.1-2.81>

Kendall, M. and Gibbons, J. D. (1990) Rank Correlation Methods. 5th Edition, Edward Arnold, London.

Kendall, M. G. (1954). Exercises in theoretical statistics, Griffin, London.

Kenny, J. F. and Keeping, E. S. (1951). Mathematics of Statistics, part two, 2nd edition, New York, Van Nostrand.



- Kim, G., Silvapulle, M. J. and Silvapulle, P. (2007). Comparison of semiparametric and parametric methods for estimating copulas, *Computational Statistics & Data Analysis*, 51(6), 2836–2850. <https://doi.org/10.1016/j.csda.2006.10.009>
- Kolev, G. I. (2011). The “spurious regression problem” in the classical regression model framework, *Economics Bulletin*, 31, 925-937.
- Lee, L. F. (1983). Generalized Econometric Models with Selectivity, *Econometrica*, 51(2), 507. <https://doi.org/10.2307/1912003>
- Lee, Y. S., Kim, T. H. and Newbold, P. (2005). Spurious nonlinear regressions in econometrics, *Economics Letters*, 87, 301-306.
- Leybourne, S. J. and Newbold, P. (1999). On the Size Properties of Phillips-Perron Tests, *Journal of Time Series Analysis*, 20(1), 51–61. <https://doi.org/10.1111/1467-9892.00125>
- Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models, *Biometrika*, 65(2), 297–303. <https://doi.org/10.1093/biomet/65.2.297>
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications, *Biometrika*, 57(3), 519–530. <https://doi.org/10.1093/biomet/57.3.519>
- Marmol, F. (1995). Spurious regressions between I(d) processes, *Journal of time series analysis*, 16, 313-321.
- Mashal, R., Naldi, M. and Zeevi, A. (2003). On the dependence of equity and asset returns, *RISK*, 16(10), 83-88.
- McCallum, B. T. (2010). Is the spurious regression problem spurious? *Economics Letters*, 107, 321-323.
- Meester, S. G. and MacKay, J. (1994). A Parametric Model for Cluster Correlated Categorical Data, *Biometrics*, 50(4), 954. <https://doi.org/10.2307/2533435>
- Morgenstern, D. (1956). Einfache beispiele zweidimensionaler verteilungen, *Mitteilungsblatt für Mathematische Statistik*, 8, 234-235.
- Nelsen, R. (1999). An Introduction to Copulas, Springer, New York.



Newbold, P., Agiakloglou, C. and Miller, J. Long-term Inference Based on Short-term Forecasting Models with Paul Newbold and John Miller, 1993, In *Developments in Time Series Analysis*, In Honor of M.B. Priestley, edited by T.S. Rao, 9-25, Chapman & Hall: London.

Newbold, P., Agiakloglou, C. and Miller, J. (1994). Adventures with ARIMA software, *International Journal of Forecasting*, 10(4), 573-581.

Newbold, P., Carlson, W. L. and Thorne, B. (2013). *Statistics for Business and Economics* (8th.Ed), Pearson, London.

Newbold, P. and Davies, N. (1978). Error misspecification and spurious regressions, *International Economic Review*, 19, 513-519.

Ning, C. (2010). Dependence structure between the equity market and the foreign exchange market – A copula approach, *Journal of International Money and Finance*, 29, 743-759.

Pearson, K. (1895). Contributions to the mathematical theory of evolution.—II. Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London*, 186, 343–414. <https://doi.org/10.1098/rsta.1895.0010>

Pearson, K. (1895a). Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58, 347–352, 240–242. <https://doi.org/10.1098/rspl.1895.0041>

Phillips, P. C. B. (1986). Understanding spurious regressions in econometrics, *Journal of Econometrics*, 33, 311-340

Reboredo, J. C. (2011). How do crude oil prices co-move? A copula approach, *Energy Economics*, 33, 948-955. <https://doi.org/10.1016/j.eneco.2011.04.006>

Sawkins, D.T. (1944). Simple regression and correlation. *Journal and Proceedings of the Royal Society of New South Wales*, 77, 85-95.

Schmidt, R. and Stadtmuller, U. (2006). Nonparametric estimation of tail dependence, *Scandinavian Journal of Statistics. Theory and Applications*, 33, 307-335

Schweizer, B. and Sklar, A. (1983). *Probabilistic metric spaces*. Elsevier Science Publishing Co., Inc, Amsterdam.





Sklar, A. (1959). Fonctions de répartition à  $n$  dimensions et leurs marges, Publications de l'Institut de Statistique de L'Université de Paris, 8, 229–231.

Spearman, C. (1904). The Proof and Measurement of Association between Two Things, *American Journal of Psychology*, 15(1), 72. <https://doi.org/10.2307/1412159>

Wen, X., Wei, Y. and Huang, D. (2012). Measuring contagion between energy market and stock market during financial crisis: A copula approach, *Energy Economics*, 34(5), 1435–1446. <https://doi.org/10.1016/j.eneco.2012.06.021>

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series, *Journal of the Royal Statistical Society*, 89, 1-64.

Zimmerman, W., D., Zumbo, D., B. and Williams, H., R. (2003). Bias in Estimation and Hypothesis Testing of Correlation, *Psicológica*, 24, 133-158.